

# Bull Escala série EPC

**EPC et solutions HA**

**Guide d'installation**

AIX



# Bull Escala série EPC

**EPC et solutions HA**

**Guide d'installation**

AIX

---

**Logiciel**

Septembre 1999

**BULL ELECTRONICS ANGERS  
CEDOC  
34 rue du Nid de Pie – BP 428  
49004 ANGERS CEDEX 01  
FRANCE**

**86 F2 79HX 05**

The following copyright notice protects this book under the Copyright laws of the United States of America and other countries which prohibit such actions as, but not limited to, copying, distributing, modifying, and making derivative works.

Copyright © Bull S.A. 1992, 1999

Imprimé en France

Vos suggestions sur le fond et la forme de ce document seront les bienvenues.  
Une feuille destinée à recevoir vos remarques se trouve à la fin de ce document.

Pour commander d'autres exemplaires de ce manuel ou d'autres publications techniques Bull, veuillez utiliser le bon de commande également fourni en fin de manuel.

### **Marques déposées**

Toutes les marques déposées sont la propriété de leurs titulaires respectifs.

AIX<sup>®</sup> est une marque déposée d'IBM Corp. et est utilisée sous licence.

UNIX est une marque déposée licenciée exclusivement par Open Group.

### **A 2000**

Le produit présenté dans ce manuel est prêt pour l'an 2000.

*La loi du 11 mars 1957, complétée par la loi du 3 juillet 1985, interdit les copies ou reproductions destinées à une utilisation collective. Toute représentation ou reproduction intégrale ou partielle faite par quelque procédé que ce soit, sans consentement de l'auteur ou de ses ayants cause, est illicite et constitue une contrefaçon sanctionnée par les articles 425 et suivants du code pénal.*

*Ce document est fourni à titre d'information seulement. Il n'engage pas la responsabilité de Bull S.A. en cas de dommage résultant de son application. Des corrections ou modifications du contenu de ce document peuvent intervenir sans préavis ; des mises à jour ultérieures les signaleront éventuellement aux destinataires.*

---

# A propos de ce manuel

Ce manuel s'adresse aux administrateurs système et aux ingénieurs client, chargés de mettre en oeuvre un Powercluster ou une solution Escala HA Bull. Les produits sont abordés tant d'un point de vue matériel que logiciel.

Ce manuel décrit en outre les caractéristiques du logiciel BullCluster livré avec Powercluster et les solutions Escala HA.

Utilisez ce guide comme point de départ à la mise en oeuvre de Powercluster ou d'une solution HA avant de vous référer à d'autres documents.

## Terminologie

Dans cet ouvrage :

- Le terme générique de *grappe* couvre les grappes de type Powercluster et solution HA.
- Ethernet fait référence à Ethernet standard (10 Mbps) et à Fast Ethernet (10/100 Mbps).
- Reprise après sinistre signifie capacité de récupération de site à site dans un rayon limité, suite à une défaillance ayant entraîné une perte d'accès aux noeuds et aux dispositifs de stockage (sur un campus, par exemple).

## Télécontrôle

Une fonction de télésurveillance est proposée en option, à utiliser avec ISM (Integrated System Management). ISM n'est pas fourni dans ce module logiciel.

## Contenu

Ce manuel est composé des chapitres suivants :

- **Chapitre 1. Concepts et composants** présente les composants matériels et logiciels de Powercluster et des solutions HA.
- **Chapitre 2. Planification et installation** est le point de départ de la procédure générale de planification/installation/configuration.
- **Chapitre 3. Informations de planifications complémentaires** offre un complément au guide *HACMP Planning Guide*. Il contient des informations sur le câblage et les unités de disque partagées.
- **Chapitres 4 à 8** : décrivent en détail les procédures de configuration du matériel et du logiciel.
- **Chapitre 9. Gestion des ressources de grappe et des modules LVM partagés** décrit les modifications des ressources de grappe (modules LVM, serveurs d'application, groupes de ressources...) via les menus de BullCluster intégrés aux menus standard de HACMP. Ces menus sont intégrés aux menus HACMP standard.
- **Chapitre 10. Outils et techniques de gestion du système** présente les outils de gestion BullCluster (outil de diagnostic, gestionnaire de personnalisation d'événement, etc.) et décrit certains outils (timed, dsmit, iFOR/LS, ArrayGUIDe, Performance Toolbox...) particulièrement utiles pour la gestion de système en environnement de grappe.
- **Chapitre 11. Télécontrôle de grappe avec ISM** explique comment contrôler à distance une grappe à partir d'une station de gestion ISM. ISM est un produit Bull dédié au contrôle et à la gestion des systèmes et réseaux distribués, qui offre un environnement convivial avec une interface graphique que vous pouvez personnaliser.

- **Chapitre 12. Intégration d'un tiroir PCI** décrit, à partir d'une configuration classique, le processus d'intégration d'un tiroir PCI dans un rack Escala EPC.
- **Annexe A. Mise en oeuvre des solutions console** traite des différentes solutions de base possibles pour gérer les consoles des noeuds de grappe. Présente la console PowerConsole, la console système, l'affichage graphique, la console de grappe et les unités associées au concentrateur de console et donne le mode d'emploi de l'utilitaire BullCluster pwcons.
- **Annexe B. Implantation de la solution console de grappe** : installation du terminal X.

## Bibliographie

### Powercluster et solutions HA

Les documents suivants concernent à la fois Powercluster et les solutions HA :

- *Escala série EPC – Powercluster & solutions HA – Guide d'installation*, CEDOC 86 F2 79HX, le présent document.
- *Powercluster & solutions HA - Utilisation des scripts d'exemple HA*, CEDOC 86 F7 82HX, est destiné principalement au personnel technique Bull. Il donne des conseils sur l'utilisation des scripts de début et de fin proposés en exemple avec le module logiciel BullCluster (pour en savoir plus, reportez-vous à "Planification d'applications spécifiques", page 3-1).
- *Powercluster Outils de gestion de grappe - ClusterWatch*, CEDOC 86 F2 80HX, traite des techniques et de l'outil de gestion Powercluster-spécifique "ClusterWatch", fourni avec le module logiciel BullCluster.
- *Escala série EPC - Préparation site pour les systèmes rack*, CEDOC 86 F1 30PX, décrit la préparation des sites pour les systèmes Escala Racks et la gamme de produits Escala Powercluster (EPC).
- *Escala série EPC – EPC Guide de câblage*, CEDOC 86 F1 65JX, fournit des informations relatives au câblage et à la configuration des produits Escala EPC. Reportez-vous à *Autres publications*, page vi.

Le document suivant concerne exclusivement Powercluster :

- *Escala série EPC – PowerConsole & ClusterAssistant – Guide d'installation*, CEDOC 86 F2 81HX, livré avec le PowerConsole, matériel en option, et proposé uniquement avec le produit (et non avec les solutions HA). Le PowerConsole est une station de travail qui offre un point de contrôle unique pour gérer les grappes. ClusterAssistant est un ensemble d'applications exécutées sur le PowerConsole pour configurer, gérer et surveiller les grappes. ClusterAssistant est intégré à CDE (Common Desktop Environment), qui propose un environnement basé sur la trilogie fenêtres/icônes/souris.

### Bulletin de logiciel

Veillez consulter les *bulletins de logiciel* (SRB) livrés avec le logiciel : Solutions HA, CEDOC 86 F2 81WF et Powercluster, CEDOC 86 F2 79WF. Ils contiennent de précieuses informations sur les conditions préalables et les contraintes de l'environnement, la procédure d'installation du logiciel, ainsi que les nouveautés de dernière minute.

### Documentation HACMP pour AIX

La documentation HACMP (High Availability Cluster MultiProcessing) pour AIX regroupe un ensemble d'ouvrages, répertoriés ci-dessous.

#### Notes de version

Les *notes de version* (*Release Notes*) pour HACMP sont stockées dans le répertoire `/usr/lpp/cluster/doc` (accessible dès que HACMP est installé).

## **Documentation HACMP destinée aux administrateurs**

Les ouvrages suivants s'adressent aux administrateurs et à toute personne travaillant en environnement de grappe :

### **HACMP 4.2**

- *HACMP Concepts and Facilities*, CEDOC 86 A2 54KX
- *HACMP Planning Guide*, CEDOC 86 A2 55KX
- *HACMP Installation Guide*, CEDOC 86 A2 56KX
- *HACMP Administration Guide*, CEDOC 86 A2 57KX
- *HACMP Troubleshooting Guide*, CEDOC 86 A2 58KX
- *HANFS Installation and Administration*, CEDOC 86 A2 61KX
- *Enhanced Scalability Installation and Administration Guide*, CEDOC 86 A2 62KX
- *Event Management Programming Guide and Reference*, CEDOC 86 A2 63KX
- *Group Services Programming Guide and Reference*, CEDOC 86 A2 64KX
- *HACMP Master Index and Glossary*, CEDOC 86 A2 51HX
- *AIX - Guide d'administration : système d'exploitation et unités*, CEDOC 86 F2 53AP.

### **HACMP 4.3**

- *HACMP Concepts and Facilities*, CEDOC 86 A2 54KX
- *HACMP Planning Guide*, CEDOC 86 A2 55KX
- *HACMP Installation Guide*, CEDOC 86 A2 56KX
- *HACMP Administration Guide*, CEDOC 86 A2 57KX
- *HACMP Troubleshooting Guide*, CEDOC 86 A2 58KX
- *HANFS Installation and Administration*, CEDOC 86 A2 61KX
- *Enhanced Scalability Installation and Administration Guide*, CEDOC 86 A2 62KX
- *Event Management Programming Guide and Reference*, CEDOC 86 A2 63KX
- *Group Services Programming Guide and Reference*, CEDOC 86 A2 64KX
- *HACMP Master Index and Glossary*, CEDOC 86 A2 65KX
- *Bull System Management Guide Operating System and Devices for AIX*, CEDOC 86 A2 53AP.

## **Documentation HACMP destinée aux programmeurs**

Deux ouvrages supplémentaires sont dédiés à la programmation d'application HA en environnement de grappe HACMP :

### **HACMP 4.2**

- *HACMP Programming Client Applications*, CEDOC 86 A2 46HX  
Ce manuel décrit les interfaces de programmation d'application client (API), qui fournissent notamment les interfaces au Cluster Information Program (clinfo). Il répertorie le contenu de la base MIB HACMP (que vous pouvez consulter directement, celle-ci étant livrée avec le logiciel).
- *HACMP Locking Applications*, CEDOC 86 A2 50HX  
Ce manuel décrit l'interface de programmation (API) Cluster Lock Manager (CLM).

### **HACMP 4.3**

- *HACMP Programming Client Applications*, CEDOC 86 A2 46HX  
Ce manuel décrit les interfaces de programmation d'application client (API), qui fournissent notamment les interfaces au Cluster Information Program (clinfo). Il répertorie le contenu de la base MIB HACMP (que vous pouvez consulter directement, celle-ci étant livrée avec le logiciel).
- *HACMP Locking Applications*, CEDOC 86 A2 59KX  
Ce manuel décrit l'API Cluster Lock Manager (CLM).

### **Systemes de stockage sur disque**

Les manuels ci-après fournissent les informations nécessaires à l'installation, la gestion et la configuration des systèmes de stockage sur disque.

*Escala JBOD Storage Subsystem - Setup & Operator Guide*, CEDOC 86 A1 79GX.

*DAS 5700 Series Disk array Storage System, Installation & Service for Rackmount Models*, CEDOC 86 A1 43KX.

*DAS 5700 Series Disk array Storage System, Installation & Service for Deskside Models*, CEDOC 86 A1 44KX.

*DAS 5300 Disk array Storage System, Installation & Service for Rackmount Models*, CEDOC 86 A1 24KX.

*DAS 5300 Disk array Storage System, Installation & Service for Deskside Models*, CEDOC 86 A1 25KX.

*DAS 5000 Disk array Storage System, Installation & Service for Rackmount Models*, CEDOC 86 A1 45KX.

*DAS 5000 Disk array Storage System, Installation & Service for Deskside Models*, CEDOC 86 A1 46KX.

*DAS 3500 Disk array Storage System, Installation & Service for Rackmount Models*, CEDOC 86 A1 47JX.

*DAS 3500 Disk array Storage System, Installation & Service for Deskside Models*, CEDOC 86 A1 48JX.

*7133 SSA Disk Subsystems Service Guide*, CEDOC 86 A1 94GX.

*Adaptateurs PCI Fibre Channel - Guide d'installation & de configuration*, CEDOC 86 F1 95HX.

*Escala Disk Expansion Unit Service Guide*, CEDOC 86 A1 13PX.

*Planning a DAS Installation (Fibre Channel Environment)*, CEDOC 86 A1 94JX.

*Navisphere Setup and Operation*, CEDOC 86 A2 47KX.

*Configuring and Managing a DAS in an AIX Environment*, CEDOC 86 A2 20PN.

### **Autres publications**

La documentation AIX et les manuels livrés avec le matériel sont très utiles : conservez-les à portée de main.

*Cabling Guide*, CEDOC 86 A1 87AQ.

### **Commandes d'ouvrages**

Pour commander d'autres exemplaires de ce manuel, précisez la référence CEDOC 86 F2 79HX.

---

# Table des matières

<b>A propos de ce manuel</b> .....	<b>iii</b>
<b>Chapitre 1. Concepts et composants</b> .....	<b>1-1</b>
Concepts de grappe .....	1-1
Haute disponibilité .....	1-1
Reprise après sinistre .....	1-1
Hautes performances .....	1-2
Composants de Powercluster et solution HA .....	1-3
Vue synoptique des composants d'une grappe .....	1-3
Serveurs (noeuds de grappe) .....	1-4
HACMP .....	1-4
Module BullCluster .....	1-4
Outils généraux .....	1-5
Unités de disque .....	1-7
Stockage partagé sur configuration "HA native" .....	1-7
Sous-systèmes RAID DAS .....	1-8
Sous-systèmes disque SSA .....	1-10
Systèmes de stockage Symmetrix CDA d'EMC .....	1-11
Bibliothèques de bandes .....	1-11
<b>Chapitre 2. Planification et installation</b> .....	<b>2-1</b>
Repères .....	2-1
Étape suivante .....	2-2
<b>Chapitre 3. Informations de planifications complémentaires</b> .....	<b>3-1</b>
Planification d'applications spécifiques .....	3-1
Planification des ports série .....	3-2
Port série S2 .....	3-2
Ports série pour consoles système et connexion SP DAS (A) .....	3-2
Ports série pour connexions de réseau série (B) .....	3-2
Ports série pour UPS (C) .....	3-2
Planification des réseaux .....	3-3
Réseaux public, privé et série .....	3-3
Planification de la topologie de réseau FDDI privé .....	3-4
Planification du stockage partagé "HA native" pour les solutions Escala HA .....	3-6
Configuration "HA native" .....	3-6
Bus SCSI partagés et fonction miroir .....	3-6
Disques internes à utiliser .....	3-7
Cartes SCSI et câblage des bus .....	3-8
Planification des sous-systèmes DAS partagés .....	3-10
Configurations DAS possibles (à deux bus distincts, deux initiateurs/deux cartes) .....	3-10
Planification des unités de disque physiques (niveaux RAID, LUN) .....	3-13
Planification du câblage .....	3-13
Configurations DAS possibles avec Fibre Channel .....	3-17
Insertion d'emplacements de disques en grappe (DAE) .....	3-19
Planification du stockage partagé HA étendue pour solutions Escala HA .....	3-20
Câblage .....	3-20
Remarque spécifique des boucles FC-AL partagées .....	3-21
Planification de la configuration des unités de disque physiques .....	3-21

Planification des sous-systèmes SSA partagés .....	3-22
Cartes SSA .....	3-22
Topologie SSA .....	3-23
Exemple de topologie SSA .....	3-23
Planification de sous-systèmes EMC Symmetrix partagés .....	3-25
Configuration de base avec HACMP .....	3-26
Sous-systèmes de disque EMC en mode concurrent .....	3-26
<b>Chapitre 4. Installation du matériel de base .....</b>	<b>4-1</b>
Préalables .....	4-1
Installation physique des serveurs (noeuds) et des unités de disque .....	4-1
Installation du matériel pour la gestion de la console système .....	4-2
Raccordement des connexions réseau .....	4-2
Amorçage des noeuds et vérification du fonctionnement .....	4-4
Configuration TCP/IP des interfaces d'amorçage/de service .....	4-4
Prérequis : IPAT/adresses d'amorçage, de secours et de service .....	4-4
Conventions d'appellation .....	4-5
Procédure .....	4-6
Configuration des ID SCSI de carte .....	4-7
Préalables .....	4-7
Configuration des ID SCSI de carte .....	4-8
Réamorçage des noeuds .....	4-9
Étape suivante .....	4-9
<b>Chapitre 5. Vérification de l'installation logicielle .....</b>	<b>5-1</b>
Vérification de l'installation logicielle .....	5-1
Bulletin de logiciel .....	5-1
HACMP et BullCluster .....	5-1
<b>Chapitre 6. Finalisation de l'installation matérielle .....</b>	<b>6-1</b>
Arrêt et mise hors tension des noeuds .....	6-1
Interconnexion des noeuds via des réseaux série (RS-232) .....	6-1
Préalables .....	6-1
Installation physique des lignes série RS-232 .....	6-2
Préparation des cartes SCSI au câblage .....	6-2
Préparation des cartes SCSI Differential type 4-6, 4-B, 4-L ou WSA-1 .....	6-2
Préparation des cartes SCSI Differential type B4-6 .....	6-3
Préparation des cartes SCSI SE type B4-5 à terminaison simple .....	6-3
Préparation des cartes Fibre Channel type B4-7 .....	6-4
Préparation des cartes Fibre Channel type B4-8 .....	6-4
Préparation des cartes Fibre Channel type B4-A et B4-B .....	6-5
Préparation du concentrateur Fibre Channel .....	6-5
Installation des unités de disque partagées .....	6-6
Câblage des unités de disque et définition de leur ID .....	6-6
Configuration des sous-systèmes DAS .....	6-7
Configuration complémentaire du stockage EMC .....	6-8
Étape suivante .....	6-8
<b>Chapitre 7. Configuration de la topologie de la grappe .....</b>	<b>7-1</b>
Choix de la méthode de configuration .....	7-1
Les deux méthodes possibles .....	7-1
Critères de choix de la méthode .....	7-2
Étape suivante .....	7-2

Préalables à l'utilisation de l'outil "BullCluster Quick Initial Configuration" .....	7-3
Préalables .....	7-3
Récapitulatif .....	7-3
Étape préliminaire : mise à jour des fichiers /etc/hosts et /.rhosts .....	7-4
Édition du fichier /etc/hosts sur le noeud local .....	7-4
Édition du fichier /.rhosts sur chaque noeud .....	7-4
Définition d'un fichier de configuration .....	7-5
Accès à l'outil "Bullcluster Quick Initial Configuration" .....	7-5
Procédure de configuration rapide .....	7-6
Définition d'un fichier de configuration .....	7-6
Application du fichier de configuration .....	7-11
Processus de configuration initiale rapide .....	7-12
Opérations de configuration .....	7-12
Opérations de vérification .....	7-13
Annulation de la configuration initiale rapide .....	7-14
Vérification d'un fichier de configuration modifié manuellement .....	7-15
Utilisation de l'outil d'instantané BullCluster .....	7-15
Étape suivante .....	7-15
<b>Chapitre 8. Finalisation de l'installation de la grappe .....</b>	<b>8-1</b>
Remarque à l'attention des utilisateurs de PowerConsole .....	8-1
Configuration des serveurs d'application et ressources HACMP .....	8-1
Configuration des modules LVM partagés (groupes de volumes...) .....	8-2
Méthode rapide .....	8-2
Méthode manuelle .....	8-2
Synchronisation des horloges de la grappe .....	8-3
Implémentation d'alias d'adresses IP .....	8-3
Activation de la notification des événements de grappe via la messagerie électronique .....	8-3
Groupes de volumes partagés sur sous-systèmes DAS .....	8-5
Configuration deux initiateurs - deux serveurs .....	8-5
Configuration "à bus distincts" .....	8-6
Étape suivante .....	8-8
<b>Chapitre 9. Gestion des ressources de grappe et des modules LVM partagés .....</b>	<b>9-1</b>
Présentation du menu .....	9-1
Accès au menu Cluster Resources Modification .....	9-1
Remarque à l'attention des utilisateurs de PowerConsole .....	9-1
Options .....	9-2
Ajout, suppression et affichage d'un groupe de ressources .....	9-4
Ajout d'un groupe de ressources .....	9-4
Suppression d'un groupe de ressources .....	9-6
Affichage des ressources d'un groupe de ressources .....	9-7
Ajout et suppression d'un groupe de volumes .....	9-8
Ajout d'un groupe de volumes à un groupe de ressources .....	9-8
Suppression d'un groupe de volumes d'un groupe de ressources .....	9-10
Ajout et suppression d'un volume logique .....	9-11
Ajout d'un groupe logique à un groupe de ressources .....	9-11
Suppression d'un groupe logique d'un groupe de ressources .....	9-13
Ajout et suppression d'un système de fichiers .....	9-14
Ajout d'un système de fichiers à un groupe de ressources .....	9-14
Suppression d'un système de fichiers d'un groupe de ressources .....	9-16

Ajout et suppression d'un label IP de service .....	9-17
Ajout d'un label IP de service à un groupe de ressources .....	9-17
Suppression d'un label IP de service d'un groupe de ressources .....	9-18
Ajout et suppression d'un serveur d'application .....	9-19
Création et ajout d'un serveur d'application à un groupe de ressources .....	9-19
Suppression d'un serveur d'application d'un groupe de ressources .....	9-19
Diffusion générale de la définition d'un groupe de volumes .....	9-20
<b>Chapitre 10. Outils et techniques de gestion du système .....</b>	<b>10-1</b>
Outil de diagnostic BullCluster (bclerrdiag) .....	10-1
Objet .....	10-1
Description .....	10-2
Syntaxe .....	10-2
Fichiers .....	10-3
Utilitaire clusterview .....	10-3
Objet .....	10-3
Syntaxe .....	10-3
Description .....	10-3
Exemple de sortie .....	10-4
Fichiers .....	10-4
Voir aussi .....	10-4
Implémentation d'alias d'adresses IP .....	10-5
Objet .....	10-5
Procédure .....	10-5
Implémentation du transfert d'applications (Rolling Applications) .....	10-7
Objet .....	10-7
Procédure .....	10-9
Présentation du menu Rolling Applications .....	10-9
Définition d'un serveur de transfert d'applications .....	10-10
Modification/Affichage des caractéristiques d'un serveur de transfert d'applications .....	10-10
Retrait d'un serveur de transfert d'applications .....	10-11
Mise en ligne d'un serveur de transfert d'applications .....	10-12
Mise hors ligne d'un serveur de transfert d'applications .....	10-13
Déplacement d'un serveur de transfert d'applications .....	10-14
Restauration de l'emplacement par défaut d'un serveur de transfert d'applications .....	10-15
Afficher l'état du serveur de transfert d'applications .....	10-16
Implémentation de HA native (avec ou sans armoire d'extension disque) .....	10-17
Objet .....	10-17
Configuration .....	10-17
Procédure .....	10-18
Reprise sur incident matériel .....	10-19
Implémentation du stockage partagé HA étendue .....	10-24
Précautions .....	10-24
Configuration .....	10-24
Procédure .....	10-24
Reprise sur défaillance matérielle .....	10-25
Reprise sur défaillance de disques DAS .....	10-27
Personnalisation du traitement des événements de grappe .....	10-27
Objet .....	10-27

Synchronisation des horloges avec timed .....	10-27
timed en bref .....	10-27
Conseils pour implémenter timed dans votre grappe .....	10-28
Clés de licence iFOR/LS Nodelocked .....	10-28
Conseils relatifs à DSMIT .....	10-28
Gestion des sous-systèmes DAS avec ArrayGUIDe .....	10-29
Gestion des sous-systèmes DAS avec Navisphere .....	10-29
<b>Chapitre 11. Télécontrôle de grappe avec ISM .....</b>	<b>11-1</b>
Concepts .....	11-1
Qu'est-ce que ISM ? .....	11-1
ISM et le contrôle de grappe .....	11-1
Applications ISM requises .....	11-2
Récapitulatif .....	11-3
Configuration de l'environnement SNMP sur les noeuds de grappe .....	11-3
Définition des fichiers de configuration snmpd .....	11-3
Mise à jour des fichiers /etc/hosts et /etc/inittab .....	11-4
Vérification de l'installation des fichiers de définition MIB .....	11-4
Vérification des opérations SNMP .....	11-4
Installation des fichiers requis sur la station ISM Manager .....	11-5
Installation de l'ensemble de fichiers dédiés à ISM .....	11-5
Installation de la base MIB HACMP au sein d'ISM .....	11-5
Tracé d'une image ISM de base pour représenter une grappe .....	11-6
Recherche des réseaux et des unités associés à la grappe .....	11-6
Tracé d'une image comportant les unités découvertes .....	11-7
Animation des icônes .....	11-9
Étape suivante .....	11-9
Interprétation des informations de base .....	11-9
Couleur des icônes .....	11-9
Recherche du noeud contenant une adresse de service donnée .....	11-10
Exemple de scénario .....	11-11
Examen de la base MIB HACMP et ajout d'objets à l'image .....	11-12
Prérequis : vérification des opérations clsmuxpd .....	11-12
Examen de la base MIB HACMP .....	11-12
Tracé des objets de la base MIB HACMP .....	11-12
Choix des attributs pour l'animation d'icône .....	11-13
Remarques sur les objets MIB HACMP .....	11-14
Remarque sur le répertoire hacmp.ext .....	11-14
Objet cluster .....	11-14
Objet Node .....	11-15
Autres objets MIB HACMP .....	11-16
Traitement des interruptions .....	11-17
Prérequis .....	11-17
Remarques sur les interruptions HACMP .....	11-17
Affichage des interruptions reçues .....	11-18
Animation des images ISM en fonction des interruptions reçues .....	11-19
Contrôle des erreurs liées au disque .....	11-19
<b>Chapitre 12. Intégration d'un tiroir PCI .....</b>	<b>12-1</b>
Intégration d'un tiroir PCI .....	12-1
Exemple de configuration .....	12-1
Procédure d'intégration .....	12-1

<b>Annexe A. Mise en oeuvre des solutions console .....</b>	<b>A-1</b>
Présentation des solutions console .....	A-1
Solutions console pour solutions HA .....	A-2
Solution console traditionnelle .....	A-2
Solutions basées sur une carte multiport série .....	A-3
Solutions console pour Escala Powerclusters .....	A-5
Composants des solutions console pour Escala Powerclusters .....	A-5
Rôle des composants des solutions console .....	A-5
Exemples de configuration console pour PowerClusters .....	A-8
Guides pour le câblage du concentrateur .....	A-11
Configuration et utilisation du concentrateur de console .....	A-13
Remarques préliminaires .....	A-13
Connexions réseau du PowerConsole .....	A-13
Amorçage du concentrateur de console .....	A-14
Configuration des paramètres de base du concentrateur .....	A-15
Configuration des ports du concentrateur de console .....	A-17
Vérification des paramètres du concentrateur .....	A-17
Vérification de la connexion réseau du concentrateur .....	A-17
Vérification de l'état LISTEN sur les ports du concentrateur de console .....	A-18
Mise à jour du fichier /etc/hosts sur le PowerConsole .....	A-18
Accès aux ports console des noeuds via le concentrateur .....	A-18
Utilisation de Telnet .....	A-18
Configuration et exploitation de l'utilitaire pwcons .....	A-19
Présentation de pwcons .....	A-19
Définition des fichiers de configuration pwcons .....	A-20
Syntaxe de la commande pwcons .....	A-21
 <b>Annexe B. Implémentation de la solution console de grappe .....</b>	 <b>B-1</b>
Procédure d'installation du terminal X .....	B-1
 <b>Index .....</b>	 <b>X-1</b>

---

# Chapitre 1. Concepts et composants

Ce chapitre traite des points suivants :

- Concepts de grappe
- Composants de Powercluster et solution HA, page 1-3.
- Serveurs (noeuds de grappe), page 1-4.
- HACMP, page 1-4.
- Module BullCluster, page 1-4.
- Unités de disque, page 1-7.

---

## Concepts de grappe

**Remarque :** Pour en savoir plus sur les concepts de grappe, reportez-vous au manuel *HACMP Concepts and Facilities*.

La technologie des grappes se révèle particulièrement adaptée lorsqu'il s'agit de concevoir des plates-formes dédiées à des missions stratégiques. Elle permet d'atteindre deux objectifs principaux : haute disponibilité et hautes performances.

### Haute disponibilité

La structure en grappe couple plusieurs serveurs en réseau. Avec cette technique de couplage, chaque serveur fait office de serveur de secours auprès des autres serveurs. Ceci confère aux serveurs et aux applications une **haute disponibilité**. Ainsi, une application en cours sur un système défaillant ou en maintenance peut être transférée vers un autre serveur et rester disponible pour les utilisateurs, qui ne subissent alors qu'une **interruption minimale**.

Sur la plupart des configurations de grappe (telles que le "basculement mutuel"), il est inutile de maintenir inactifs les serveurs de secours : ils sont disponibles en exploitation normale.

#### Applications concernées

La fonction haute disponibilité (HA) est essentielle pour les applications stratégiques de l'entreprise. Toute application peut en bénéficier et être exploitée dans ce mode.

#### Terminologie

Les grappes HA (haute disponibilité) sont parfois qualifiées de grappes "quasi-tolérantes aux pannes". "HA" est l'abréviation de "high availability" (haute disponibilité) ou de "highly available" (hautement disponible).

### Reprise après sinistre

Il existe deux niveaux de reprise après sinistre :

- **Haute disponibilité étendue** (solution Bull). Dans ce cas, la grappe est partagée entre deux sites avec stockage commun de données. Ce partage du stockage limite la longueur du câblage requis entre les deux sites. La fonction HA locale est prise en charge si certaines précautions sont prises. Pour plus de détails, reportez-vous page 10-24.
- **Reprise géographique après sinistre**. Dans ce cas, les sites fonctionnent indépendamment et ne partagent pas le stockage. Des mécanismes assurent la réplication des données sur le site de secours et la haute disponibilité est limitée à la surveillance mutuelle et à la reprise du site.

## Hautes performances

Outre la **haute disponibilité**, les structures en grappes peuvent bénéficier des hautes performances et d'une grande évolutivité.

### Mode accès concurrent

Une grappe peut être configurée pour une exploitation en mode **d'accès concurrent**, dans lequel les différents serveurs ont accès simultanément aux mêmes données sur disques partagés. Ce type de configuration présente les avantages suivants :

- Performances et évolutivité exceptionnelles :
  - La charge d'une application peut être répartie sur plusieurs noeuds.
  - Pour adapter les performances d'une grappe à des besoins toujours croissants, il suffit d'ajouter des serveurs ou d'autres ressources à la configuration.
- Disponibilité accrue :
  - Les temps mort sont réduits du fait que le basculement de disque n'est pas requis quand un noeud de grappe vient à faillir.
  - Quand un noeud est en panne, les autres noeuds ont accès aux données. Les applications peuvent basculer instantanément sur un autre noeud.

Seules les applications compatibles avec l'accès concurrent peuvent bénéficier de la fonction haute performance, à savoir :

- Application OPS (Oracle Parallel Server), extension du SGBDR Oracle.
- Autres applications SGBDR et OLTP exécutables en accès concurrent.

Renseignez-vous auprès de votre représentant Bull pour connaître les applications SGBDR et OLTP exécutables en accès concurrent.

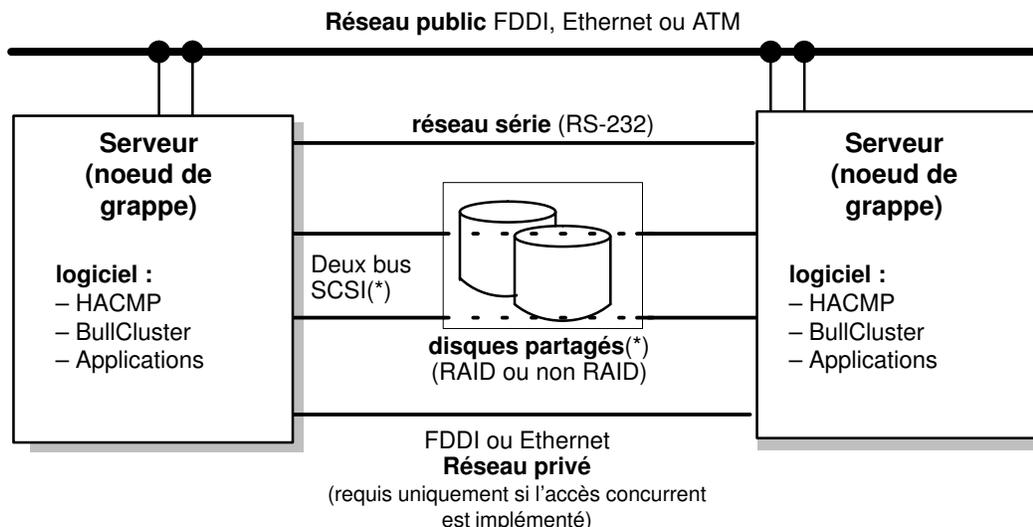
**Remarque :** Les expressions "accès concurrent", "multitraitement concurrent", "multitraitement en grappe" et "traitement CMP" font référence au même concept. Notez que, dans la documentation HACMP, le mode d'accès concurrent est parfois appelé "mode 3".

# Composants de Powercluster et solution HA

Powerclusters et les solutions HA sont des grappes résultant de la combinaison de composants matériels et logiciels. Ces deux produits, très puissants, offrent une grande disponibilité.

## Vue synoptique des composants d'une grappe

La figure ci-après illustre les principaux composants matériels d'une grappe de 2 noeuds.



(\*) Remarques :

- La plupart des configurations utilisent deux bus SCSI. Toutefois, des configurations économiques sont possibles avec un seul bus.
- Dans la configuration dite "Native HA", disponible sur les séries Escala E et T, des disques internes peuvent remplacer les disques externes pour le stockage partagé.
- Il est également possible d'utiliser, au lieu de bus et de disques partagés SCSI, des sous-systèmes disque SSA (Serial Storage Architecture) et des boucles de connexion.
- De même, à la place de bus et de disques partagés, il est possible d'utiliser des sous-systèmes disque Fibre Channel (DAS 3500), des boucles FC-AL et des cartes Fibre Channel.

### Logiciels

Une grappe comporte les logiciels suivants :

- HACMP pour AIX, coeur de la structure en grappe,
- BullCluster, module Bull intégrant divers utilitaires et outils de productivité,
- applications exécutées sur les serveurs, à doter de la haute disponibilité.

### Matériel

Une grappe est constituée du matériel de base suivant :

- serveurs (noeuds de la grappe),
- unités de disque partagé (RAID ou non – les unités prises en charge sont décrites plus loin dans ce chapitre),
- connexions matérielles : cartes SCSI (ou SSA), cartes Fibre Channel, cartes réseau (public et privé), connexions série RS-232 entre noeuds.

**Remarque** : Pour en savoir plus sur les réseaux (public, privé et série), reportez-vous à la page 3-3.

## Solutions console

Le produit Powercluster propose différentes solutions matérielles pour gérer la console. Le **PowerConsole**, notamment, est une station de travail qui offre un point de contrôle unique pour gérer les grappes. Il sert à la fois de point d'accès (via un concentrateur) aux consoles système des différents noeuds, et de puissante station de travail susceptible d'exécuter n'importe quel utilitaire d'administration.

## Duplications

Un certain nombre de duplications matérielles sont nécessaires en contexte HA pour limiter les défaillances sur point unique (SPOF).

Par exemple, les configurations illustrées précédemment intègrent deux noeuds, deux cartes réseau par noeud, ainsi que deux bus SCSI distincts pour l'accès aux disques partagés externes. Les données stockées sur disque doivent être redondantes, soit via la fonction miroir d'AIX (sur les disques non RAID), soit via l'utilisation de disques RAID (qui assurent eux-mêmes la duplication de leurs données).

---

## Serveurs (noeuds de grappe)

Les solutions HA sont des grappes intégrées aux séries Escala D, T, M et E.

Le produit Powercluster est constitué de serveurs Escala montés en armoires (racks). Trois modèles de serveur d'un seul noeud sont disponibles :

- EPC400, EPC430 et EPC440 avec un bus PCI,
- EPC800, avec un bus MCA,
- EPC1200, EPC1200A et EPC2400 avec un bus PCI, constitué de deux armoires (une armoire d'ordinateur avec un tiroir CPU et une armoire d'extension avec un tiroir d'E/S).

En outre, des configurations multi-noeud peuvent être configurées sur la base de ces modèles en ajoutant des noeuds.

---

## HACMP

HACMP (High Availability Cluster Multi-Processing) for AIX, coeur de la structure en grappes, fournit les mécanismes de contrôle de la grappe.

Le module HACMP principal, appelé "**sous-système HA**", est chargé de :

- détecter les modifications intervenues sur la grappe (notamment les défaillances matérielles et les arrêts du serveur pour maintenance);
- lancer et coordonner les interventions consécutives à une défaillance, pour reconfigurer l'environnement et basculer les activités sur d'autres ressources.

En outre, HACMP peut être complété du module **Concurrent Resource Manager** (CRM) qui gère les accès partagés concurrents sur disque. Ce module, requis pour la mise en oeuvre de l'accès concurrent, est proposé en option.

**Remarque :** Pour en savoir plus sur les fonctions HACMP, reportez-vous au manuel *HACMP Concepts and Facilities*.

---

## Module BullCluster

BullCluster est un module logiciel développé par Bull. Il regroupe divers utilitaires et outils de productivité, utiles dans un environnement de grappe.

Les caractéristiques de BullCluster sont répertoriées aux sections "Outils généraux" et "Outils spécifiques de Powercluster", ci-après.

## Outils généraux

### Outil de configuration rapide

BullCluster intègre un outil convivial simplifiant la configuration initiale d'une grappe, implémenté comme extension des menus SMIT de HACMP (SMIT, System Management Interface Tool, est l'outil standard AIX de gestion du système, piloté par menus).

L'outil de configuration rapide permet de configurer la topologie de la grappe via une jeu réduit de menus. Il offre en outre des options permettant de personnaliser le traitement d'événements HACMP.

### Menus simples pour la gestion des ressources de grappe

Un menu est ajouté aux menus HACMP d'origine pour faciliter :

- la création et la modification des ressources de grappe à l'usage de HACMP ;
- la configuration et la gestion des modules LVM partagés (groupes de volumes, volumes logiques, systèmes de fichiers).

### Exemple de scripts de basculement

BullCluster comporte des scripts de début et de fin, utiles pour installer les services Informix, Oracle, OPS, Tuxedo, HVX, X25 et Internet. Ces scripts, seulement fournis à titre de référence, doivent être adaptés aux besoins des utilisateurs (pour en savoir plus, reportez-vous à "Planification d'applications spécifiques", page 3-1).

### Utilitaires dédiés aux différentes configurations de stockage partagé

BullCluster fournit des scripts complémentaires à ceux de HACMP, pour prendre en charge les sous-systèmes DAS (Disk Array Storage) RAID, les sous-systèmes disque SSA (Serial Storage Architecture), les sous systèmes de disque miroir partagé permettant une reconstitution en cas de sinistre (Disaster Recovery) et la configuration dite "HA native" (qui comprend deux serveurs Escala partageant leurs disques internes, comme décrit page 3-6).

### Outil de diagnostic BullCluster

BullCluster propose un outil de diagnostic qui permet de générer des informations de diagnostic. Cet outil analyse l'état et la cohérence des ressources de grappe, et balaye les fichiers journaux HACMP et système à la recherche des événements majeurs survenue sur la grappe et des erreurs matérielles. Cet outil est fourni sous forme de commandes Unix.

### Utilitaires de télécontrôle de grappe (et OPS) via ISM

BullCluster fournit des utilitaires facilitant l'exploitation des stations de gestion ISM (Integrated System Management) pour :

- le télécontrôle de l'état de la grappe,
- le télécontrôle et l'optimisation d'OPS (Oracle Parallel Server).

BullCluster comprend également des utilitaires générant des interruptions SNMP lorsque des erreurs relatives aux disques partagés sont détectées. ISM peut contrôler ces interruptions.

**Remarque :** ISM est un produit Bull qui permet de contrôler et de gérer les systèmes et réseaux distribués. Il offre aux administrateurs une interface graphique conviviale et personnalisable qui leur donne la possibilité de visualiser et de contrôler les systèmes en réseau.

### ClusterWatch

ClusterWatch est un utilitaire basé sur le Web, qui contrôle les opérations de grappe. Il permet aux administrateurs système d'avoir instantanément accès aux données stratégiques dont ils ont besoin pour gérer les grappes et optimiser l'exploitation.

**Remarque :** Une version SMIT de ClusterWatch est également fournie.

### **Autres utilitaires**

BullCluster intègre également les utilitaires **pwcons** et **clusterview**, fournis sous forme de commandes Unix :

- **pwcons** est un utilitaire qui donne à l'opérateur le contrôle des consoles système des noeuds dépourvus de terminal.
- **clusterview** est un utilitaire qui rend compte de l'état des grappes. Il est comparable au programme **clstat** livré avec HACMP.

---

## Unités de disque

Cette section traite des disques partagés, c'est-à-dire des disques accessibles par les différents noeuds de la grappe.

### Remarques sur les disques internes

Les serveurs sont équipés d'unités de disque interne. Ces disques sont exclusivement dédiés au stockage des fichiers et des applications système (et du groupe de volume système **rootvg**). Aucune donnée destinée aux applications HA ne doit y être stockée. Autrement dit, les unités de disque interne ne peuvent pas être partagées sur une grappe (seule exception : la configuration "Native HA", décrite page 1-7).

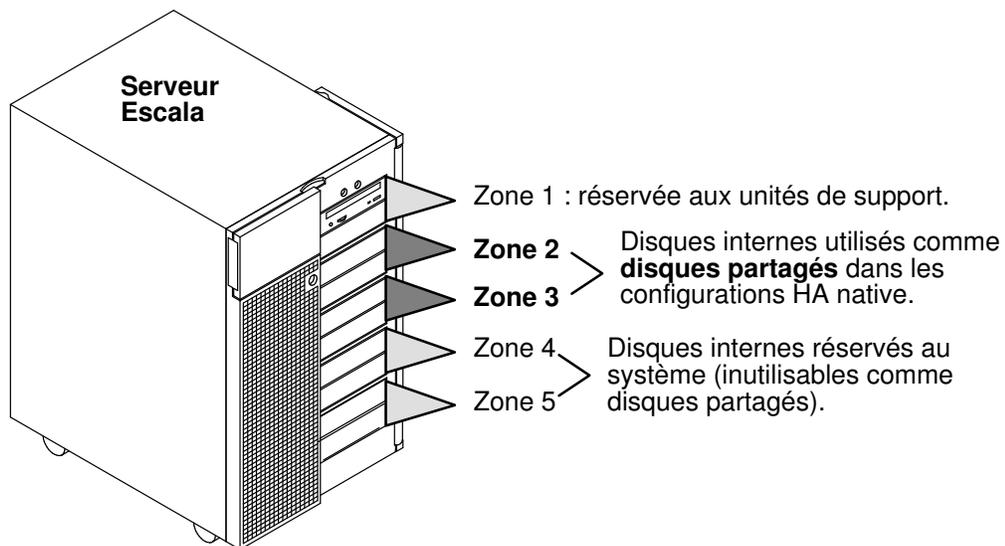
Notez que les disques système internes (et notamment le groupe de volume système **rootvg**) peuvent être mis en miroir pour accroître la disponibilité des noeuds. La fonction miroir est une fonction standard du système AIX.

## Stockage partagé sur configuration "HA native"

Pour les solutions HA basées sur serveur Escala E ou T, il est possible d'implémenter la configuration dite "HA native" (haute disponibilité native).

Une configuration HA native est constituée de deux serveurs Escala E ou T qui partagent leurs disques internes. Ces configurations offrent une alternative économique quoique puissante aux configurations qui utilisent des sous-système externes pour le stockage partagé.

Pour en savoir plus sur les configurations HA native, reportez-vous à la page 3-6.



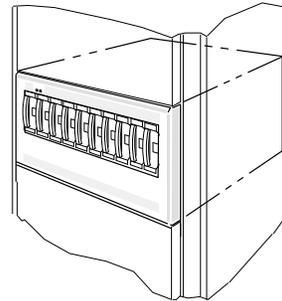
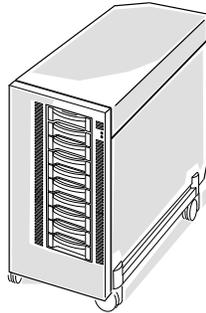
## Sous-systèmes RAID DAS

Les sous-systèmes de stockage DAS (Disk Array Storage) intègrent la fonction RAID permettant de récupérer des données de façon transparente en cas de défaillance de disque. Ils sont particulièrement adaptés à la haute disponibilité et aux grappes qui requièrent une capacité disque importante. Ils sont disponibles sur les Powerclusters et sur toutes les solutions HA basées sur serveurs Escala.

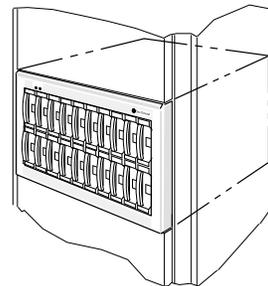
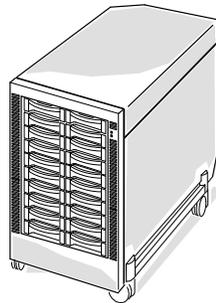
Versions verticales

Versions montées en armoire

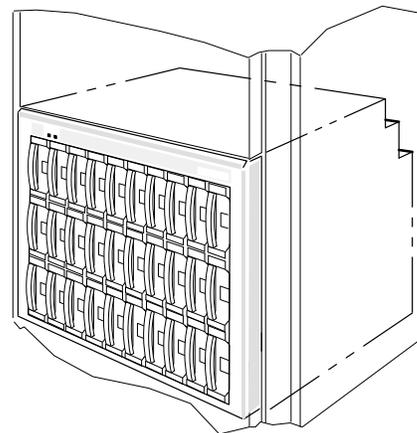
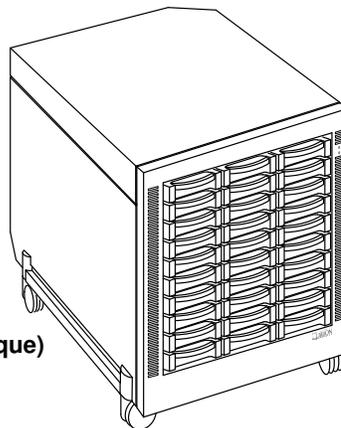
**DAS 1300**  
(jusqu'à 10 modules disque)



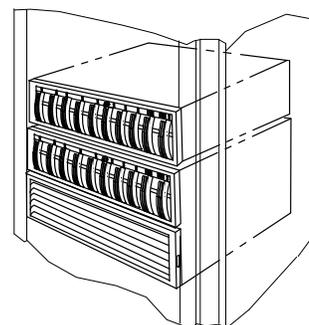
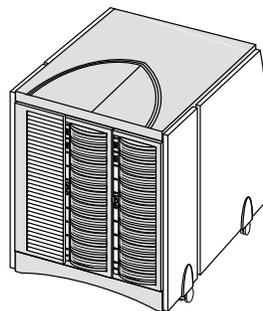
**DAS 2900**  
(jusqu'à 20 modules disque)



**DAS 3x00**  
(jusqu'à 30 modules disque)



**DAS 5700**  
(jusqu'à 120 modules disque)  
**DAS 5720, DAS5300**  
(jusqu'à 30 modules disque)



## **Principaux atouts des sous-systèmes DAS 1300/2900/3200**

Ces sous-systèmes sont prémunis contre les défaillances sur point unique (SPOF) du fait des configurations double bus SCSI et des processeurs de stockage et blocs d'alimentation redondants. Ils utilisent des connexions SCSI-2 DE Fast/Wide. Pour optimiser la haute disponibilité des données, les principaux composants (unités de disque, processeurs de stockage et blocs d'alimentation) sont remplaçables sans mise hors tension du système.

La structure modulaire des sous-systèmes DAS leur confère une très grande souplesse et un rapport performance/fiabilité optimal (disques ultra-rapides et écriture en mémoire cache). Les DAS 2900 et 3200 étant équipés de processeurs de stockage nouveaux et améliorés, offrent les performances optimales.

**Remarque :** Les processeurs de stockage (SP) sont des unités intégrées qui contrôlent les modules disque via un bus SCSI.

## **Atouts des sous-systèmes DAS 3500 (Fibre Channel)**

Ces sous-systèmes sont prémunis contre les défaillances sur point unique (SPOF) du fait des configurations double bus Fibre Channel et des processeurs de stockage et des blocs d'alimentation redondants. Pour optimiser la haute disponibilité des données, les principaux composants (unités de disque, processeurs de stockage et blocs d'alimentation) sont remplaçables sans mise hors tension du système.

Un câble Fibre Channel raccorde le système de stockage à une carte HBA (Host Bus Adapter) Fibre Channel au niveau du serveur. La circulation des données et des commandes entre le système de stockage et les serveurs se fait via un système FC-AL (Fibre Channel Arbitrated Loop).

## **Atouts des sous-systèmes DAS 5700/5720/5300 (Full Fibre Channel)**

Les systèmes de stockage sur disques en grappe Fibre Channel offrent une capacité de plusieurs téraoctets, des taux de transfert de données élevés, des configurations plus flexibles et des données hautement disponibles. Les sous-systèmes Full Fibre Channel font appel aux unités de disque à fibre optique (contrairement aux systèmes DAS 3500 qui utilisent des disques SCSI).

Optimisation de la disponibilité des données : tous les composants sont conçus pour être retirés et remplacés à chaud, sans interruption de fonctionnement.

## **Niveaux RAID**

Les sous-systèmes DAS prennent en charge les niveaux RAID 0, 0/1, 1, 3 et 5.

**Remarque :** Pour en savoir plus sur les configurations DAS, reportez-vous à la page 3-10.

## Sous-systèmes disque SSA

Les sous-systèmes disque SSA (Serial Storage Architecture) sont disponibles sur les Powerclusters et sur toutes les solutions HA basées sur serveurs Escala.

### Présentation de l'architecture SSA

L'architecture SSA (Serial Storage Architecture) est une architecture point-à-point pour les unités de stockage. SSA, devenu un standard de l'industrie, peut être une aux configurations de stockage basées sur SCSI.

Hautes performances :

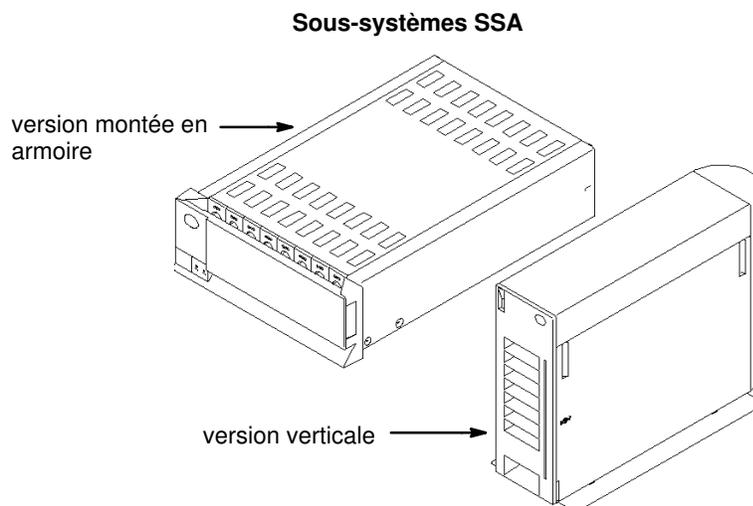
- SSA assure les communications en duplex intégral et offre une bande passante totale de 80 Mo/s (deux canaux de lecture et deux canaux d'écriture sont utilisés simultanément, chacun à 20 Mo/s).

Haute disponibilité inhérente au mécanisme SSA :

- Dans une grappe, les liaisons SSA sont implantées en boucle, de sorte que les voies d'accès aux données ne constituent pas des points uniques pour les défaillances (SPOF). Avec cette topologie en boucle, les noeuds ont accès aux sous-systèmes disque via deux voies distinctes d'accès aux données. Si l'une est indisponible, l'autre est automatiquement accessible. La haute disponibilité est ainsi optimisée et les connexions "à chaud" (sans mise hors tension) possibles.

### Sous-systèmes disque SSA compatibles

Les sous-systèmes disque SSA en version verticale et armoire. La version armoire est conçue pour s'intégrer aux serveurs Escala montés en armoire.



Voici les principales fonctions des sous-systèmes disque SSA :

- Un sous-système peut abriter de 4 à 16 unités de disque. Pour augmenter la capacité de stockage, il suffit d'interconnecter plusieurs systèmes : dans une boucle, vous pouvez connecter jusqu'à 48 disques et sur une carte SSA, jusqu'à 96.
- La plupart des composants des sous-systèmes disque SSA (unités de disque, alimentations, câbles SSA...) peuvent être ajoutés, retirés ou remplacés sans interrompre l'exploitation du sous-système ni des noeuds utilisés. Alimentation et ventilation redondantes sont disponibles en option.

### Remarque sur la disponibilité des données

Les sous-systèmes disque SSA pris en charge ne proposent pas les fonctions de disponibilité des données inhérentes à RAID. Aussi, pour implanter la redondance des données, l'administrateur peut implémenter la mise en miroir (avec les utilitaires standard d'AIX).

**Remarque :** Pour en savoir plus sur les configurations SSA, reportez-vous à la page 3-22.

## **Systèmes de stockage Symmetrix CDA d'EMC**

Les Powerclusters sont généralement équipés de sous-systèmes RAID DAS, ou de sous-systèmes disque SSA comme unités de disque partagées. Il peuvent également être, en option, équipés de systèmes de stockage Symmetrix CDA.

Les systèmes Symmetrix CDA (proposés par EMC) assure la protection des données via RAID, dont RAID-1, RAID-S (stratégie RAID développée par EMC) et la répartition dynamique des volumes disque.

N'étant pas traitée dans ce document, l'installation de ces systèmes incombe à des spécialistes. Pour plus de détails, contactez votre représentant Bull.

---

## **Bibliothèques de bandes**

Les unités de bandes hautes performances multichargeur de type Overland DLT4000/DLT7000 assurent les fonctionnalités de sauvegarde des données. Les unités de bandes, logées dans des unités montées en armoire, constituent les bibliothèques LibraryXpress LXB4000/LXB7000.

Une bibliothèque LXB peut être rattachée à un seul noeud. La bibliothèque ne peut être partagée pour la haute disponibilité que par deux noeuds de type similaire avec l'application HA NetBackup. Le serveur maître ou le serveur esclave peut être redondant en implémentant le mode de secours à chaud ou de reprise mutuelle.



---

## Chapitre 2. Planification et installation

Ce chapitre est une introduction à la procédure de planification, d'installation et de configuration d'une grappe.

La section Repères résume les différentes étapes de la procédure.

Une fois toutes les étapes de la planification menées à terme, passez au chapitre suivant pour procéder à l'installation. Exécutez les étapes dans l'ordre indiqué. Des renvois aux chapitres ou documents associés vous sont indiqués chaque fois que nécessaire.

---

### Repères

Cette section récapitule les étapes de la procédure de planification, d'installation et de configuration.

La procédure consiste à :

1. **Planifier l'installation de la grappe.**
2. **Installer la console et les composants matériels de base.**
  - Raccorder les consoles systèmes et installer les connexions réseau. C'est à un stade ultérieur que vous connectez les unités de disque partagées externes.
3. **Amorcer chaque noeud en vérifiant qu'il est opérationnel.**
4. **Configurer sur chaque noeud l'adresse d'amorçage TCP/IP pour la carte réseau de service et l'adaptateur de secours.**
  - Pour ce faire, lancez la commande `smit mktcpip`.
5. **Configurer les ID SCSI de la carte** (non applicable si vous exploitez des sous-systèmes disque SSA ou fibre).
  - Pour appliquer les modifications, cette étape est suivie du réamorçage des noeuds.
6. **Vérifier l'installation du logiciel.**
7. **Le cas échéant, installer le logiciel** (le logiciel est généralement préchargé en usine).
8. **Arrêter les noeuds et les mettre hors tension.**
9. **Interconnecter les noeuds via les réseaux série RS-232.**
10. **Préparer les cartes SCSI en vue du câblage** (non applicable si vous utilisez un sous-système de stockage disque SSA).

Si vous utilisez un sous-système de stockage disque SCSI, cette étape n'est requise que si vous ajoutez un serveur à une grappe préexistante ; elle est inutile si vous procédez à l'installation complète d'un nouvelle grappe déjà configurée en usine.
11. **Préparer les boucles Fibre Channel.**

Si vous utilisez des cartes Fibre Channel avec des sous-systèmes DAS (concentrateur FC-AL en option), établissez les différentes boucles FC-AL entre les sous-systèmes DAS et les cartes Fibre Channel, directement ou non, et au besoin avec des concentrateurs FC-AL.
12. **Installer les unités de disque partagées** (disques RAID ou non).
  - Câbler les unités de disque et, au besoin, définir leur ID SCSI.

**13. Démarrer les disques/noeuds et configurer le sous-système RAID DAS.**

- Créer les unités logiques et définir les LUN (Logical Unit Number = numéro d'unité logique) sur les sous-systèmes RAID DAS.
- C'est à un stade ultérieur que vous créez les groupes de volumes partagés.
- Si vous utilisez des concentrateurs FC-AL, mettez-les, ainsi que les sous-systèmes, sous tension. Le témoin (jaune) s'éteint quand le sous-système est prêt. Mettez ensuite les noeuds sous tension et configurez le sous-système.

**14. Déterminer la méthode de configuration de la grappe.**

- Vous avez le choix entre la méthode de configuration standard HACMP et la méthode de configuration initiale rapide (spécifique de BullCluster).

**15. Configurer la topologie de la grappe pour HACMP**

- soit avec l'outil de configuration rapide,
- soit via la méthode de configuration standard HACMP.

**16. Finaliser la configuration** tel que décrit au chapitre "Finalisation de l'installation de la grappe", page 8-1.

---

## Étape suivante

Une fois la première étape de planification menée à son terme, passez au chapitre "Installation du matériel de base", page 4-1, et suivez les instructions pour procéder à l'installation et à la configuration de la grappe.

---

## Chapitre 3. Informations de planifications complémentaires

Ce chapitre complète le manuel *HACMP Planning Guide*. Il apporte des précisions sur le câblage, les unités de disque partagées, les consoles et les applications spécifiques.

- Planification des applications
- Planification des ports série, page 3-2.
- Planification des réseaux, page 3-3.
- Planification du stockage partagé "HA native" pour les solutions Escala HA, page 3-6.
- Planification des sous-systèmes DAS partagés, page 3-10.
- Insertion d'emplacements de disques en grappe (DAE), page 3-19.
- Planification du stockage partagé HA étendue pour solutions Escala HA, page 3-20.
- Planification des sous-systèmes SSA partagés, page 3-22.
- Planification des sous-systèmes EMC Symmetrix partagés, page 3-25.

Pour en savoir plus sur les câbles, reportez-vous au document *Cabling Guide* livré avec le matériel.

---

### Planification d'applications spécifiques

Le module logiciel BullCluster comporte des scripts de début et de fin, utiles pour implémenter diverses applications sur la grappe comme des applications HA. Insistons sur le fait que ces scripts ne sont que des exemples (donc non pris en charge), qui doivent être modifiés selon vos besoins.

- En fonction de votre environnement et de ses contraintes, vous pouvez utiliser ces scripts comme base pour écrire des scripts adaptés. Vous pouvez également décider d'écrire vos propres scripts ex nihilo, sans tenir compte des scripts proposés.
- Dans l'un ou l'autre cas, n'oubliez pas qu'écrire des scripts de début et de fin fiables suppose une connaissance technique approfondie de HACMP et de l'application à implémenter. Nous vous conseillons de demander l'aide des services techniques Bull.

#### Exemple de scripts pour les principales applications

Les exemples de scripts pour les applications ci-après se trouvent dans le répertoire `/usr/sbin/bullcluster/install/ha_conf :`

- Oracle,
- OPS (Oracle Parallel Server),
- Informix,
- Tuxedo,
- Services Internet (Netscape Enterprise Server, Netscape Proxy Server, NetWall),
- HVX,
- X25.

Notez, en outre, que le sous-menu **Create and Add an Application Server to a Resource Group** du menu SMIT BullCluster **Cluster Resources Modification** permet de définir facilement pour HACMP un serveur d'application pour une quelconque des applications ci-dessus.

## Exemple de scripts pour diverses applications

Le module BullCluster est fourni avec d'autres exemples de scripts pour d'autres applications. Ces exemples et les fichiers README associés se trouvent dans le répertoire `/usr/sbin/bullcluster/install/samples`. Pour en savoir plus, consultez ces fichiers.

## Documentation pour l'utilisation des scripts exemple

Le document *Powercluster & solutions HA - Utilisation des scripts d'exemple HA*, principalement destiné au personnel technique Bull, donne des conseils sur l'utilisation des scripts de début et de fin proposés en exemple avec le module logiciel BullCluster.

---

## Planification des ports série

Pour prendre le matériel d'une grappe en charge, chaque noeud requiert plusieurs ports série RS-232. Il s'agit de déterminer le nombre total de ports série nécessaires.

### Port série S2

Tenez compte de ce qui suit au sujet du port série S2 des serveurs Escala.

### Serveur EPC 1200/1200A/2400

Le port série S2 du système Escala EPC 1200/1200A et EPC2400 est exclusivement dédié à la télémaintenance. Des cartes 8 ports (ou 16) doivent être utilisées pour toute connexion série nécessaire à la configuration HACMP.

### Autres serveurs ESCALA

EPC400/430/440 et ESCALA E ou T utilisent les ports COM2 (modem externe) ou COM3 (UPS) plutôt que le port 8-pour les connexions heartbeat et DAS.

## Ports série pour consoles système et connexion SP DAS (A)

Chaque noeud de grappe requiert :

- 1 port série (le port S1) pour la connexion à la console système, que le noeud soit ou non raccordé directement à une console.
- 1 port série supplémentaire (généralement le port série S3 ou COM3) par sous-système DAS, le cas échéant, pour la connexion du noeud avec un processeur de stockage (SP) DAS.

Soit au total pour chaque noeud :

- 1 port série pour une grappe dépourvue de sous-système DAS ;
- 2 ports série pour une grappe dotée d'un sous-système DAS ;
- 3 ports série pour une grappe dotée de deux sous-systèmes DAS en série.

## Ports série pour connexions de réseau série (B)

Par ailleurs, chaque noeud requiert des ports série supplémentaires pour la connexion aux réseaux en série :

- 1 port série pour une grappe de deux noeuds,
- 2 ports série pour une grappe de plus de deux noeuds.

## Ports série pour UPS (C)

Si vous utilisez des blocs d'alimentation UPS (Uninterruptible Power Supply) en option, chaque noeud requiert un port série supplémentaire.

---

## Planification des réseaux

Cette section complète les informations sur les réseaux fournies dans le manuel *HACMP Planning Guide*. Elle traite de quelques spécificités et fournit des conseils concernant les offres Powercluster et solutions HA.

### Réseaux public, privé et série

Dans le contexte HACMP :

- Un réseau *public* est un réseau qui relie plusieurs noeuds et donne aux clients accès aux noeuds de la grappe.
- Un réseau *privé* est un réseau de communication point à point entre les noeuds de la grappe. Un réseau privé est essentiellement utilisé pour le trafic verrouillé entre noeuds. Il est requis pour implémenter l'accès concurrent. Il est également requis pour l'architecture 3-tier.
- Un réseau *série* consiste en liaisons RS-232 qui relient les noeuds entre eux pour le transport du flux de détection de survie ("keep-alive") généré par HACMP.

### Réseau public

Le réseau public peut être implémenté via Ethernet, FDDI, Token-Ring ou ATM.

#### Ethernet

Pour la connexion au réseau public, les configurations de réseau Ethernet courantes requièrent généralement deux cartes Ethernet 10/100 Mbps ou de 1 Go par noeud (une carte de service et une de secours). Sur certaines configurations de grappe plus complexes, qui implémentent deux réseaux publics Ethernet à des fins de redondance, quatre cartes par noeud sont nécessaires.

#### FDDI

Sur les réseaux public faisant appel à FDDI, il est recommandé d'utiliser une configuration à deux anneaux plutôt qu'un seul pour augmenter la disponibilité.

Pour la connexion au réseau public, les configurations de réseau FDDI courantes requièrent généralement deux cartes à deux anneaux par noeud (une carte de service et une de secours). Sur certaines configurations de grappe plus complexes, qui implémentent deux réseaux publics FDDI à des fins de redondance, quatre cartes double anneau par noeud sont nécessaires.

#### ATM

L'implémentation du réseau public peut également passer par le mode ATM (Asynchronous Transmission Mode), qui fournit un débit de 155 mégabits par seconde.

Une configuration classique comporte un commutateur ATM et chaque noeud est équipé de deux cartes ATM (une carte de service et une de secours).

## Réseau privé

Les réseaux privés ne concernent que les grappes sur lesquelles est implémenté l'accès concurrent ou des noeuds avec architecture 3-tier, les uns jouant le rôle de serveurs d'applications et les autres faisant office de serveurs de bases de données.

Le réseau privé est implémenté via Ethernet ou FDDI double-anneau.

- Informations sur Ethernet :
  - interconnexion simple ou double avec un concentrateur standard (> 2 noeuds) ;
  - interconnexion simple ou double avec un commutateur haute performance (> 2 noeuds).

Pour prévoir un commutateur Fast Ethernet, adressez-vous à votre représentant.

- Pour en savoir plus sur FDDI, reportez-vous à "Planification de la topologie de réseau FDDI privé", ci-après.

**Concentrateurs FDDI et Ethernet :** pour le montage de ces concentrateurs dans des armoires, vous pouvez commander des kits spécifiques (un pour le concentrateur FDDI et un pour le concentrateur Ethernet).

## Planification de la topologie de réseau FDDI privé

**Remarque :** Cette section ne concerne que les grappes qui doivent être équipées de réseaux privés.

Pour l'interconnexion de noeuds de grappe, une installation type ne fait appel qu'à un seul réseau. Toutefois, pour une optimisation maximale de la disponibilité, deux réseaux privés peuvent intervenir. Ce type de configuration complexe n'est pas décrit ici.

Quand vous planifiez un réseau FDDI privé, rappelez-vous que du degré de disponibilité dépend de la topologie choisie. Quelques conseils sont donnés plus loin.

## Grappes de 2 ou 3 noeuds

Pour une grappe de 2 ou 3 noeuds, vous pouvez implémenter le réseau FDDI privé simplement via un anneau double, sans faire appel à d'autres unités FDDI.

En cas de bouclage d'anneau (si un anneau est rompu en raison d'un noeud hors tension ou d'une section FDDI défectueuse), aucune fragmentation n'a lieu : l'intercommunication des noeuds reste possible.

### Remarques :

1. Du point de point évolutivité, un réseau privé FDDI est construit avec un concentrateur (ou un commutateur, pour de meilleures performances).
2. Pour assurer une grande disponibilité, un réseau FDDI privé basé sur un seul anneau n'est pas approprié. Le FDDI à double anneau est formé de deux boucles indépendantes (un anneau principal et un secondaire faisant office d'anneau de secours) et offrant ainsi une plus grande disponibilité.
3. Comme expliqué plus haut, techniquement, une grappe de 3 noeuds ne requiert pas l'emploi d'unités FDDI particulières (un double anneau ordinaire étant suffisant). Toutefois, sachez que les modèles Powercluster 3 noeuds peuvent être livrés avec 2 concentrateurs FDDI ; vous pouvez ainsi mettre en oeuvre une configuration FDDI jumelée, comme expliqué à la section Grappes de 3 noeuds ou plus ci-après. Adopter ce type de configuration pour une grappe de 3 noeuds facilite les mises à niveau ultérieures (ajout de noeuds à la grappe).

## Grappes de 3 noeuds ou plus

Pour une grappe de 3 noeuds ou plus, implémenter le réseau privé simplement via un double anneau FDDI peut paraître inapproprié en matière de disponibilité. Sachez toutefois que :

- En cas de bouclage d'anneau (si un anneau est rompu en raison d'un noeud hors tension ou d'une section FDDI défectueuse), l'anneau continue à fonctionner normalement. L'intercommunication des noeuds reste possible.
- C'est l'éventualité d'un second bouclage d'anneau qui pose problème. Dans ce cas, le réseau FDDI se fragmente en deux anneaux indépendants. L'intercommunication des noeuds d'un fragment à l'autre n'est alors plus possible.

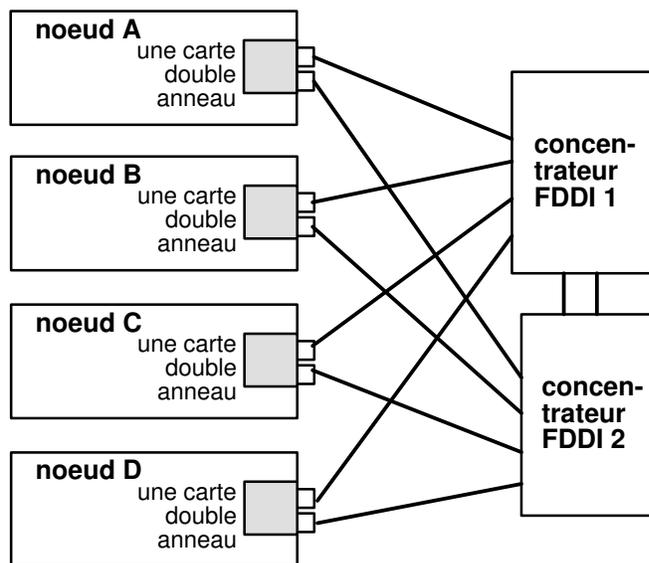
### Configuration FDDI jumelée

Pour empêcher la fragmentation du réseau FDDI privé, les noeuds doivent être connectés à un concentrateur FDDI. Par ce biais, la mise hors tension d'un noeud ou la défaillance d'une carte FDDI ne provoque pas de fragmentation du réseau.

Ceci peut s'avérer insuffisant si vous souhaitez une disponibilité optimale du réseau privé. Sachant que le concentrateur constitue un point de défaillance unique (SPOF), s'il vient à faillir, il en résulte une panne du réseau FDDI. Pour éliminer ce SPOF, la solution est de mettre en oeuvre une configuration jumelée. Ce type de configuration repose sur deux concentrateurs FDDI indépendants auxquels chaque noeud est connecté : ainsi, un accès alternatif est en place si l'accès principal vient à faillir.

Pour l'offre Powercluster, dans ce type de configuration, chaque noeud est équipé d'une carte FDDI double anneau reliée aux deux concentrateurs FDDI (modèle Linkbuilder de 3COM).

En voici illustré un exemple (grappe de 4 noeuds) :



**RESEAU PRIVE : Implémentation FDDI jumelée**

# Planification du stockage partagé "HA native" pour les solutions Escala HA

## Préalables

- Consultez le chapitre "Planning Shared Disk Devices" dans le document *HACMP Planning Guide*.

## Configuration "HA native"

Pour les solutions HA basées sur les serveurs Escala E ou T, il est possible d'implémenter la configuration dite "HA native".

## Bus SCSI partagés et fonction miroir

Deux types de configuration sont possibles, dans lesquelles sont utilisés :

- uniquement les disques internes ou
- les disques internes plus une armoire d'extension disque.

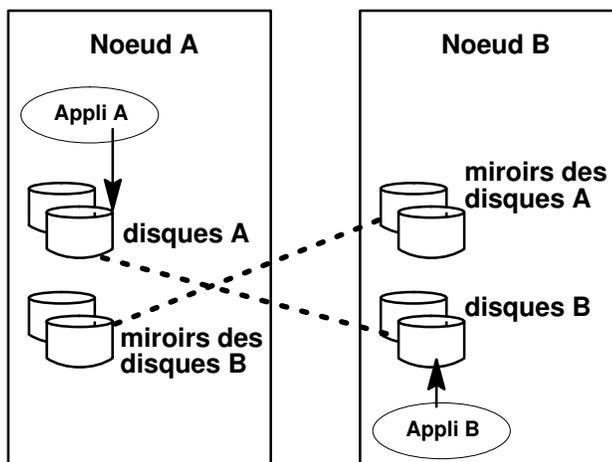
Dans une configuration HA native, deux disques SCSI partagés sont implémentés avec des disques non RAID.

Pour la haute disponibilité, vous devez implanter la fonction disque miroir, via les utilitaires standard d'AIX. Elle est mise en oeuvre par la fonction miroir standard du système AIX. Reportez-vous à "LVM Mirroring" dans le document *HACMP Planning Guide*.

## Scénario de fonction miroir avec une configuration HA native simple

La mise en miroir doit être implémentée sur les deux noeuds. Ce qui signifie que les disques internes d'un noeud (accessibles via un bus SCSI partagé) doivent être dupliqués sur les disques internes de l'autre noeud (accessibles via l'autre bus SCSI partagé).

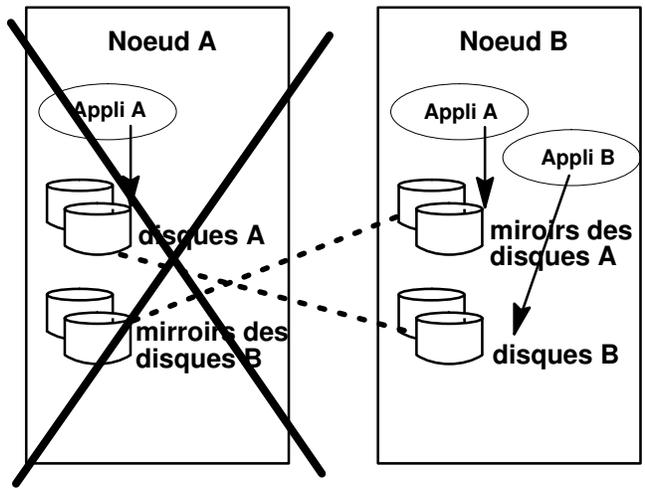
La figure suivante illustre une configuration permettant le basculement mutuel en cours d'exploitation normale :



Exploitation normale :

- Noeud A travaille avec l'application Appli A qui stocke ses données sur les disques A. Ces disques sont dupliqués sur les disques de noeud B. Les deux noeuds ont accès aux disques A et à leurs miroirs, ceux-ci étant sur des bus SCSI distincts.
- Sur noeud B a été défini un environnement similaire et symétrique. Une autre application, Appli B, utilise les disques B et ceux-ci sont dupliqués sur l'autre noeud.

Supposons une défaillance du noeud A. HACMP lance alors le processus de basculement : l'application appli A est démarrée sur le noeud B et peut ainsi travailler en utilisant les miroirs des disques A (voir illustration suivante).



Exploitation après basculement :

- Noeud A est en panne.
- Noeud B active les deux applications, Appli A et Appli B. Appli A utilise à présent les miroirs des disques A.

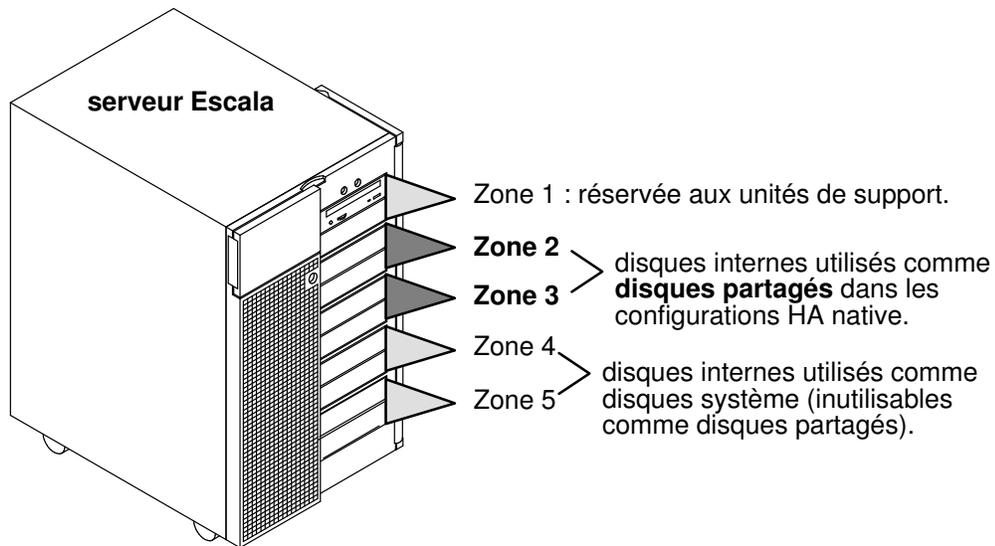
### Scénario de fonction miroir avec une armoire d'extension disque

Si vous utilisez une armoire d'extension disque, les disques internes de noeud A sont dupliqués sur les disques de l'armoire d'extension. De même, les disques internes de noeud B sont dupliqués sur les disques de l'armoire d'extension.

Pour en savoir plus, reportez-vous à Unité d'extension disque - Guide de maintenance et Tiroir d'extension PCI - Installation rapide.

### Disques internes à utiliser

Dans les configurations HA native, les disques utilisés pour le stockage partagé sont les disques internes résidant dans la zone unité 3, et éventuellement dans la zone unité 2, de l'unité de base du serveur.



Notez que :

- Les zones unités 3 et 2 utilisent le même bus SCSI. Chaque zone peut loger jusqu'à deux disques de 1,6 pouces ou trois disques de 1 pouce de haut.
- Pour la mise en miroir, la configuration doit être aussi symétrique que possible. Les zones unités partagées (zone 3 et éventuellement 2) d'un noeud doivent être occupées de la même façon que les zones correspondantes sur l'autre noeud : mêmes disques, même capacité, même nombre et même emplacement.

## Cartes SCSI et câblage des bus

Notez que :

- Pour implémenter les deux bus SCSI, chaque noeud doit être équipé de deux cartes SCSI Ultra SE. Ces cartes sont libellées "B4-5". (Pour les configurations HA native avec armoire d'extension disque, quatre cartes sont nécessaires).
- L'utilisation de câbles en "Y" accroît la disponibilité. Les opérations de maintenance se révèlent plus aisées si vous adoptez cette méthode : elle permet de débrancher le câble en Y sans rompre la continuité du bus SCSI.

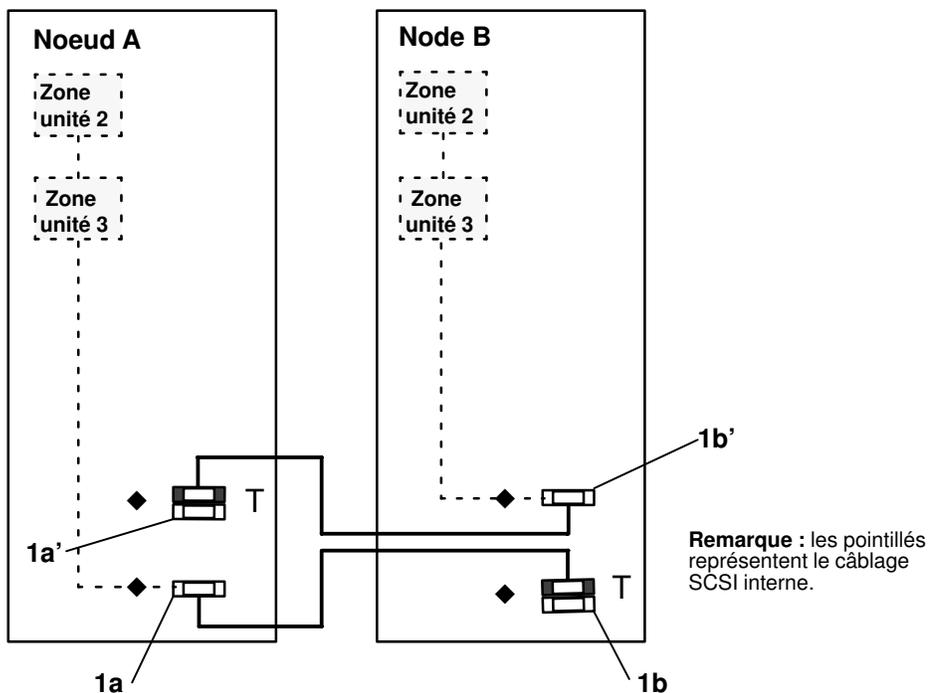
Deux bus SCSI partagés doivent être installés. Le schéma ci-après illustre le câblage du premier bus SCSI partagé (arrière des serveurs Escala).

### ATTENTION :

**Tous les serveurs livrés sont dotés de câbles internes et de connecteurs standard, qui doivent être adaptés aux configurations illustrées.**

### Configuration HA native simple

La configuration HA native simple est une solution d'extension disque HA économique, permettant de dupliquer et de partager les disques internes via deux chaînes SCSI.



**La longueur totale de câble ne doit pas excéder 1,5 m.**

### Référence des composants :

Câble SCSI interne, réf. : CBLG178  
(P/N: 78172894-001)

Adaptateurs SCSI (type B4-5) 

Câble SCSI + 2 fiches de bouclage (90982001-001)  
+ 1 terminaison de type pass-thru (91076001-001) 

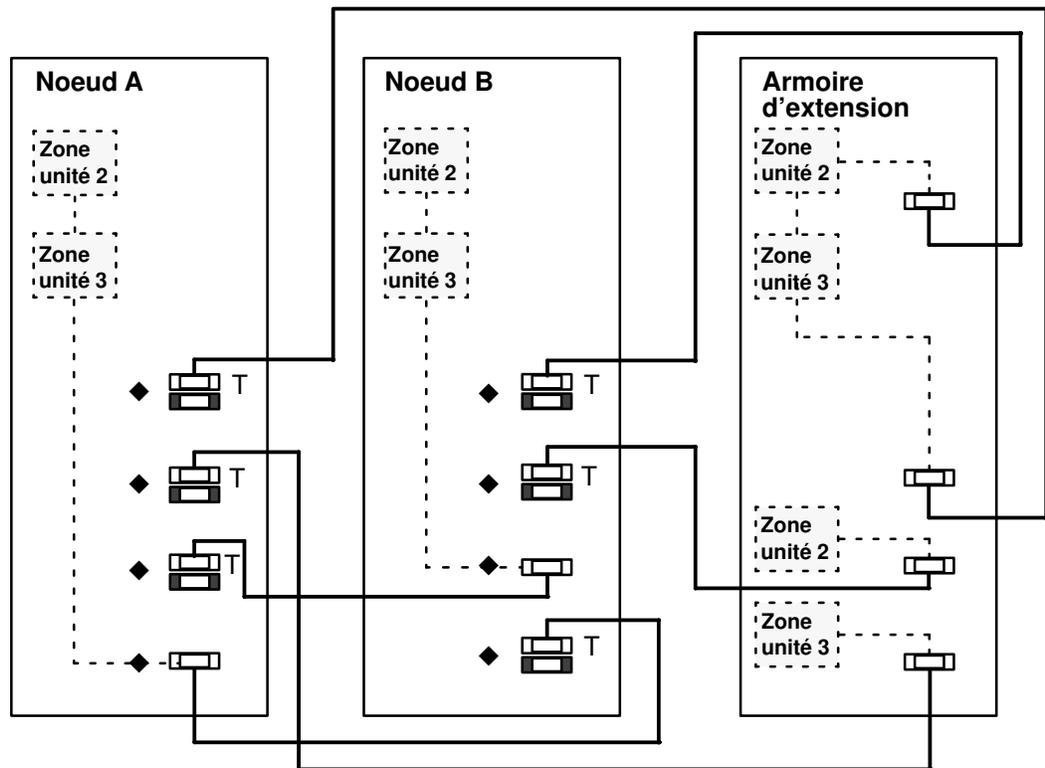
Ces références sont susceptibles de changer.

## Légende des exemples

<b>1a, 1a', 1b, 1b'</b>	<p>Chaque noeud est équipé de deux <b>cartes SCSI Ultra SE</b> (type B4-5), installées aux emplacements PCI.</p> <p>A l'intérieur de noeud A, un accès SCSI interne (ligne pointillée sur la figure) relie la carte <b>1a</b> aux disques de la zone 3 (et éventuellement de la zone 2), via la cloison interne (non illustrée). De même, à l'intérieur de noeud B, un accès SCSI interne relie la carte <b>1b'</b> aux disques de la zone 3 (et éventuellement de la zone 2), via la cloison interne.</p>
<b>T</b>	<p>Les terminaisons de type Pass-thru se trouvent sur une extrémité des câbles externes et sont connectées aux cartes SCSI <b>1a'</b> et <b>1b'</b> (auxquelles aucun disque interne n'est relié).</p>
<b>◆</b>	<p>Remarques sur les <b>fiches de bouclage</b> (représentées par le symbole <b>◆</b>) :</p> <ul style="list-style-type: none"> <li>• Les deux accès SCSI internes mentionnés ci-dessus doivent être équipés d'une fiche de bouclage pour neutraliser la terminaison SCSI au niveau de la carte. Cette fiche doit être branchée sur le connecteur intermédiaire du câble SCSI interne, qui relie les cartes SCSI (<b>1a</b> et <b>1b</b>) à la cloison interne.</li> <li>• Les deux cartes SCSI <b>1a'</b> et <b>1b</b> doivent également être équipées d'une fiche de bouclage pour neutraliser la terminaison SCSI au niveau de la carte. Cette fiche doit être branchée sur le connecteur d'extrémité de la carte SCSI.</li> <li>• Sur les nouveaux systèmes, les fiches de bouclage sont installées en usine, vous n'avez donc pas à vous en soucier. Si vous mettez des serveurs Escala à niveau, installez les fiches de bouclage livrées avec les cartes (dans le kit de mise à niveau).</li> </ul>

## Configuration HA native avec armoire d'extension disque partagée

Une configuration HA native avec armoire d'extension disque partagée constitue une solution économique pour partager des disque HA entre deux hôtes et dupliquer des disques internes et externes via deux chaînes SCSI.



---

## Planification des sous-systèmes DAS partagés

### Préalables

- Consultez le chapitre "Planning Shared Disk Devices" dans le document *HACMP Planning Guide*.
- Pour une vue d'ensemble des unités de disque prise en charge, reportez-vous également à "Unités de disque", page 1-7 de ce manuel.

## Configurations DAS possibles (à deux bus distincts, deux initiateurs/deux cartes)

### Présentation

Dans un environnement de grappe, plusieurs configurations sont possibles pour les sous-systèmes DAS. Elles diffèrent par le mode d'implémentation des bus SCSI.

Deux configurations (SCSI ou Fibre Channel) sont particulièrement intéressantes :

- configuration à deux bus distincts, comportant deux bus, mais un seul par noeud ;
- configuration à deux initiateurs/deux cartes, comprenant deux bus, partagés l'un et l'autre par les deux noeuds.

En réalité, la **configuration "deux initiateurs/deux cartes" est préférable car elle offre une plus grande disponibilité**. Ceci parce qu'elle implémente deux bus partagés distincts.

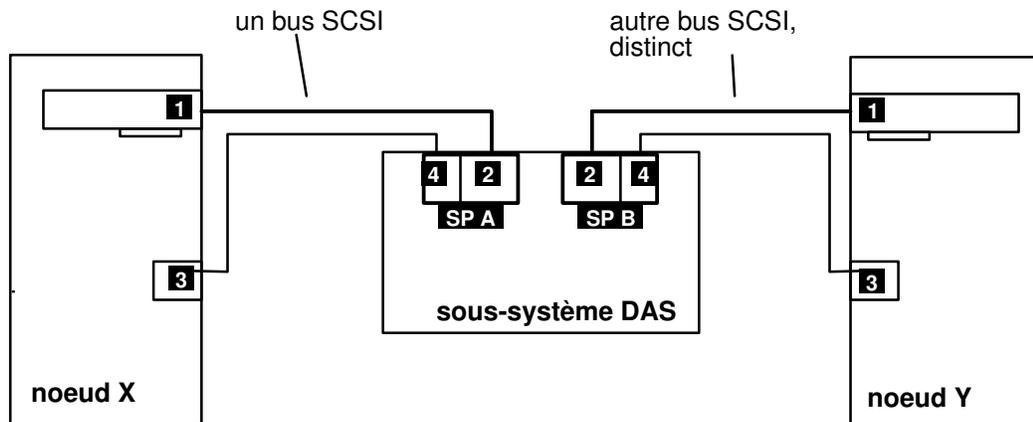
Les configurations proposées ne sont données qu'à titre d'exemple pour vous guider dans votre tâche. Voici quelques conseils pour vous aider à planifier l'installation.

Pour en savoir plus, reportez-vous à la documentation livrée avec le sous-système DAS, et tout particulièrement au chapitre "Understanding configurations" du manuel *Configuring and Managing a DAS and Using ATF* et au chapitre "About the disk-array storage system" du guide *Installing and Maintaining a DAS*.

**Remarque** : Les exemples illustrés mettent l'accent sur les configurations à deux bus distincts et à deux initiateurs/deux cartes les plus courantes. D'autres types de configuration peuvent être envisagés : reportez-vous à la documentation des sous-systèmes DAS.

## Configuration à deux bus distincts

La figure ci-dessous illustre une configuration à deux bus distincts.



SP A, SP B	Le sous-système DAS est équipé de deux processeurs de stockage (SP).
1, 2	Chaque noeud est équipé d'une carte SCSI-2 DE Fast/Wide (1) raccordée à la carte SCSI-2 (2) d'un processeur de stockage. Les deux bus SCSI distincts ne sont pas partagés par les noeuds (chaque bus commence à un noeud et se termine à un processeur de stockage).
3, 4	Chaque noeud est connecté à un processeur de stockage DAS via des ports série RS-232 (numéros 3 et 4 sur la figure). Cette ligne série permet l'accès aux fonctions de configuration et de gestion DAS depuis le noeud.

### Points forts

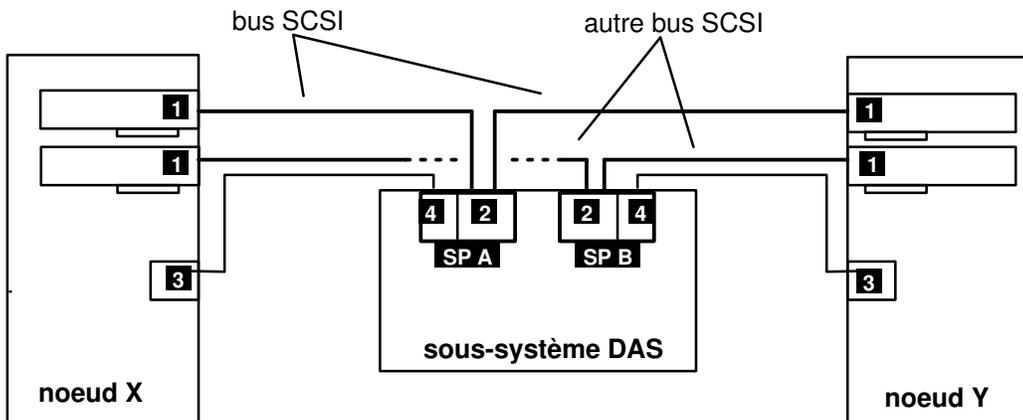
Chaque noeud est équipé d'une carte SCSI connectée par un bus SCSI à un processeur de stockage dans le sous-système DAS. Les deux noeuds ont accès à n'importe quelle unité de disque physique DAS, mais pas simultanément à la même unité.

Avec une configuration à deux bus distincts, si une carte SCSI ou un processeur de stockage est défaillant, le noeud concerné ne peut poursuivre, mais l'autre noeud reste opérationnel. Ainsi, un noeud peut prendre en charge les noeuds de l'autre, si celui-ci est défaillant : le noeud opérationnel, via le processeur de stockage accessible, peut prendre le contrôle de toutes les unités physiques de disques partagées.

Ce type de configuration est avantageux pour certains sites. Mais le plus souvent, c'est la configuration deux initiateurs/deux cartes qui est adoptée car elle offre une disponibilité maximale. Voir ci-après.

## Configuration deux initiateurs/deux cartes (SCSI)

La figure ci-dessous illustre une configuration deux initiateurs/deux cartes



SP A, SP B	Le sous-système DAS est équipé de deux processeurs de stockage (SP).
1, 2	Chaque noeud est équipé de deux cartes SCSI Corvette DE (1), chacune raccordée à la carte SCSI2 (2) d'un processeur de stockage. Les deux bus SCSI sont partagés par chaque noeud. Chaque bus commence à un noeud, passe par un processeur de stockage et se termine à un autre noeud.
3, 4	Chaque noeud est connecté à un processeur de stockage DAS via des ports série RS-232 (numéros 3 et 4 sur la figure). Cette ligne série permet l'accès aux fonctions de configuration et de gestion DAS depuis le noeud.

**Remarque :** Deux sous-systèmes DAS maximum peuvent être connectés en série pour fournir une capacité de stockage supérieure.

### Points forts

Chaque noeud est équipé de deux cartes SCSI, chacune connectée par un bus SCSI distinct du sous-système DAS. Comme dans la configuration à deux bus distincts, les deux noeuds ont accès à n'importe quelle unité de disque physique DAS, mais pas simultanément à la même unité.

Cette configuration confère une très grande disponibilité, dans la mesure où elle fournit deux routes indépendantes vers le sous-système DAS. Elle protège contre toute défaillance pouvant survenir sur le bus SCSI.

Si un noeud, une carte SCSI ou un processeur de stockage est défaillant, l'autre noeud peut prendre la relève au niveau de l'unité de disque du noeud défaillant. Le noeud opérationnel peut transférer le contrôle de toutes les unités physiques de disque partagées vers le SP accessible.

Cette configuration est recommandée pour la disponibilité maximale qu'elle confère au système. Pour sa mise en oeuvre, le logiciel ATF (Application-Transparent Failover) est nécessaire (voir ci-après).

### Logiciel ATF

Le logiciel ATF (Application-Transparent Failover) doit être commandé avec les sous-systèmes DAS. Il doit être installé sur les deux noeuds. Si un accès SCSI vient à faillir (carte, câble, connexion ou processeur de stockage), ATF en bascule le trafic sur le second accès SCSI. ATF opère automatiquement, de façon transparente et sans impact sur les applications et opérations en cours.

**Remarque :** ATF est un logiciel entièrement indépendant, qui n'est pas lié aux mécanismes de basculement HACMP. (Le logiciel ATF est un produit de CLARiiON Business Unit of Data General Corporation.)

## Planification des unités de disque physiques (niveaux RAID, LUN)

Les sous-systèmes DAS acceptent les niveaux RAID 0, 0/1, 1, 3 et 5. Il est possible de panacher des groupes de différents niveaux RAID dans un même sous-système DAS.

- Les niveaux 3 et 5 sont généralement les plus adaptés à la haute disponibilité.
- Il est du ressort de l'administrateur système de déterminer le niveau RAID adapté.

Par ailleurs, vous devez planifier la configuration des unités logiques (LUN). Une unité logique est une entité unique constituée d'un ou plusieurs modules disque (en environnement DAS, "LUN" et "unité de disque physique" sont synonymes). Vous pouvez créer jusqu'à 8 unités logiques.

Reportez-vous à la documentation des sous-systèmes DAS pour des informations détaillées sur la planification et pour les fiches de configuration.

## Planification du câblage

Vous devez prévoir les câbles requis pour l'installation des sous-systèmes DAS. Tous les éléments de câblage doivent être compatibles avec la norme **SCSI-2 Differential Fast/Wide**. Sur un bus SCSI, la longueur totale de câble ne doit pas dépasser 18 mètres (25 mètres étant le maximum théorique).

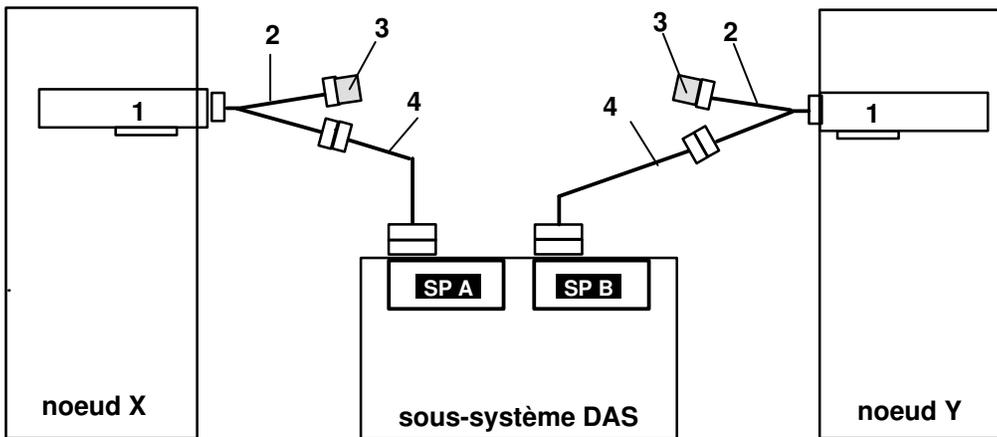
Les câbles en "Y" assurent une disponibilité accrue. Les opérations de maintenance se révèlent plus aisées si vous adoptez cette méthode : elle permet de débrancher le câble en Y sans rupture de continuité sur le bus SCSI.

Les figures ci-après représentent quatre configurations avec des commentaires axés sur le câblage (les câbles entre noeuds RS-232 et processeurs de stockage ne sont pas représentés) :

- câbles pour configuration à deux bus distincts,
- câbles pour configuration à deux initiateurs/deux cartes,
- câbles pour connexion en série de deux sous-systèmes DAS,
- autres câbles pour grappes de plus de deux noeuds.

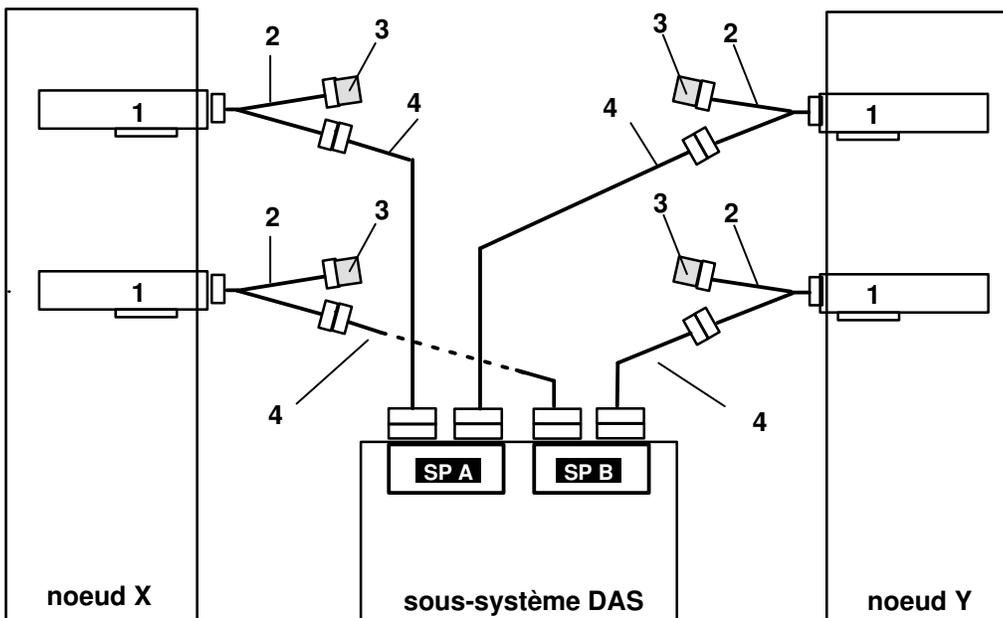
Ces informations sont données à titre indicatif. Pour des informations de référence (notamment sur la longueur maximale des câbles), reportez-vous à la documentation des sous-systèmes DAS.

## Câbles pour configuration à deux bus distincts



1	Une carte SCSI-2 DE Fast/Wide par noeud. Leur terminaison interne doit être retirée.
2	Câbles en Y (différentiels).
3	Terminaisons (différentielles). Raccordées à la branche courte des câbles en Y.
4	Câbles entre câble en Y et contrôleur DAS (différentiels).
*	Deux câbles RS-232 (non représentés sur la figure) sont nécessaires pour connecter chaque noeud à un processeur de stockage. Pour les ports série, reportez-vous à la page 3-2.

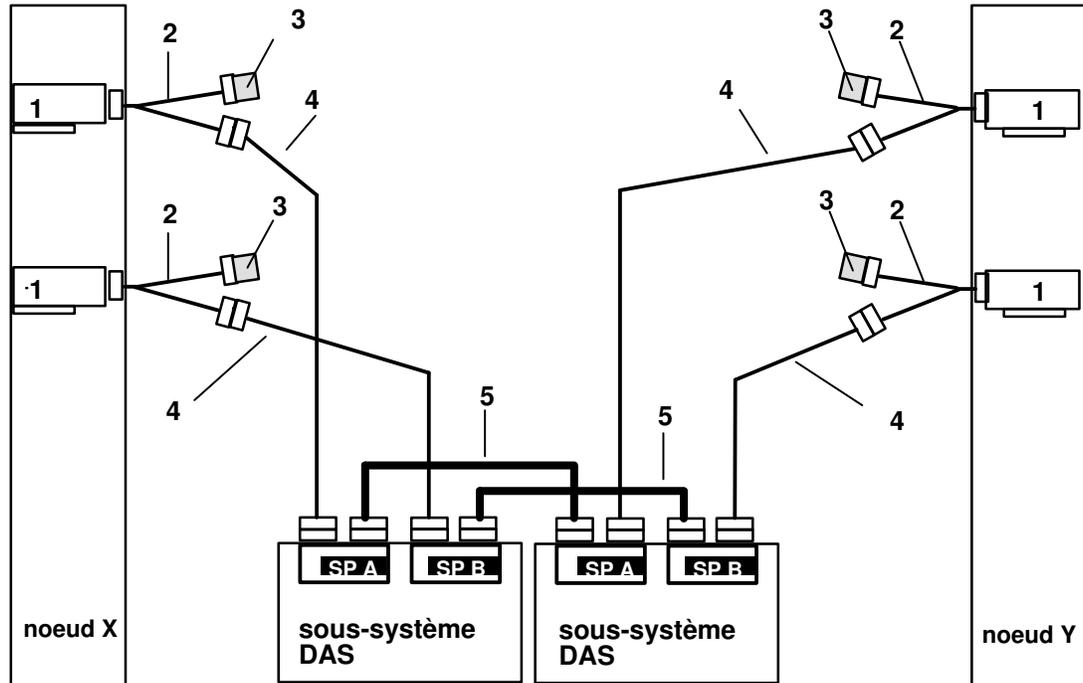
## Câbles pour configuration à deux initiateurs/deux cartes



1	Deux cartes SCSI-2 DE Fast/Wide par noeud pour fournir deux bus SCSI distincts partagés. Leur terminaison interne doit être retirée.
2	Câbles en Y (différentiels).
3	Terminaisons (différentielles). Raccordées à la branche courte des câbles en Y.
4	Câbles entre câble en Y et contrôleur DAS (différentiels).
*	Deux câbles RS-232 (non représentés sur la figure) sont nécessaires pour connecter chaque noeud à un processeur de stockage. Pour les ports série, reportez-vous à la page 3-2.

## Câbles pour connexion en série de deux sous-systèmes DAS

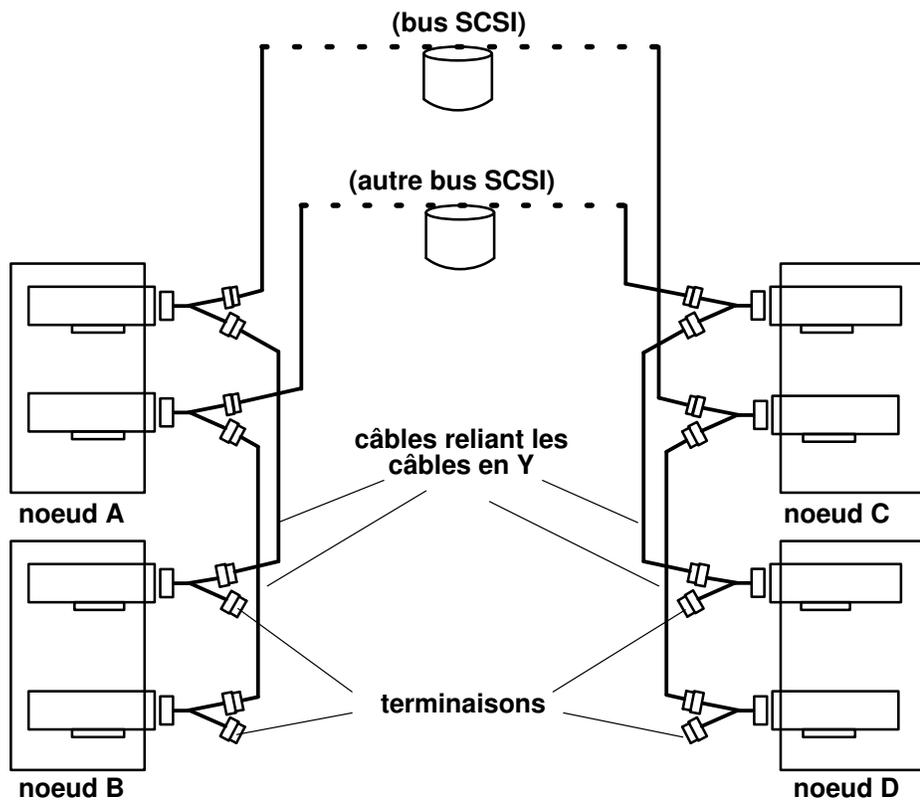
Deux sous-systèmes DAS maximum peuvent être connectés en série pour fournir une capacité de stockage supérieure. La figure ci-après illustre une configuration à deux initiateurs/deux cartes intégrant deux sous-systèmes DAS en série (pour en savoir plus, reportez-vous à la documentation des sous-systèmes DAS).



1	Deux cartes SCSI-2 DE Fast/Wide par noeud pour fournir deux bus SCSI distincts partagés. Leur terminaison interne doit être retirée.
2	Câbles en Y (différentiels).
3	Terminaisons (différentielles). Raccordées à la branche courte des câbles en Y.
4	Câbles entre câble en Y et contrôleur DAS (différentiels).
5	Câbles entre DAS (différentiels), représentés en gras sur la figure. Ils servent à relier en série deux sous-systèmes DAS.
*	Quatre câbles RS-232 (non représentés sur la figure) sont nécessaires. Chaque noeud est connecté à deux SP, une sur chaque sous-système DAS. Pour les ports série, reportez-vous à la page 3-2.

## Grappes de plus de deux noeuds

Si la grappe comporte plus de deux noeuds, les mêmes câbles sont utilisés. Toutefois, des câbles supplémentaires entre câbles en Y sont en outre requis pour relier les chemins SCSI entre deux noeuds mitoyens. Reportez-vous à la figure de la page 3-16, qui représente une grappe de 4 noeuds et met l'accent sur la nécessité de câbles entre câbles en Y et sur la position adéquate des terminaisons.

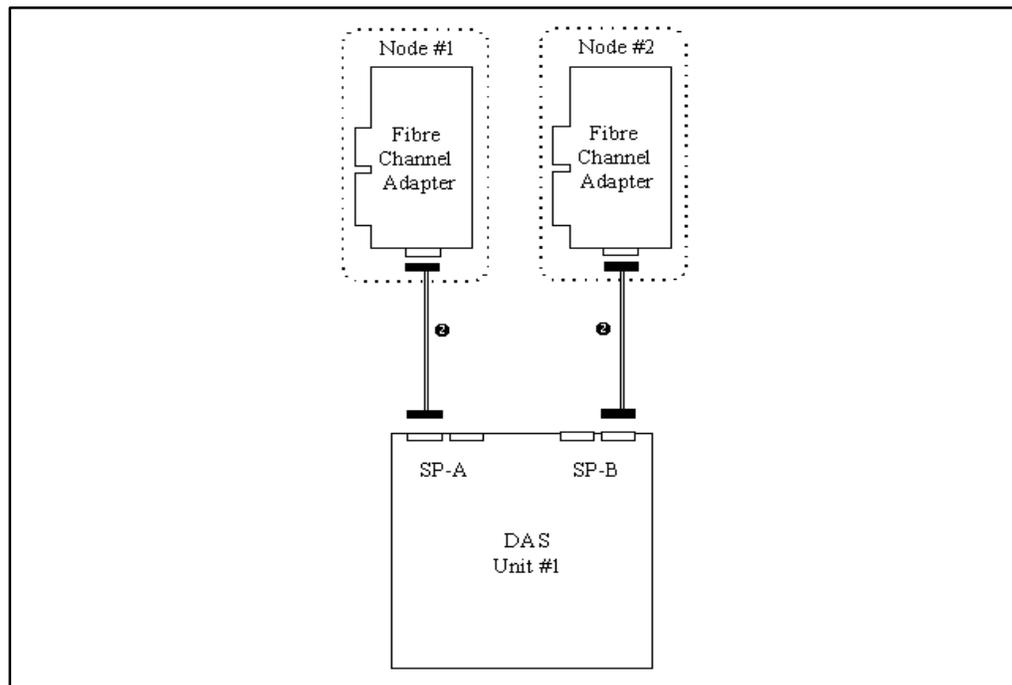


La figure ci-dessus illustre le câblage SCSI sur une grappe de 4 noeuds. Elle met l'accent sur la nécessité de câbles entre câbles en Y et sur la position adéquate des terminaisons. (Sur ces points, le schéma s'applique que les disques partagés soient des sous-systèmes DAS ou des disques SCSI simples, bien que les unités de disque ne soient pas représentées).

# Configurations DAS possibles avec Fibre Channel

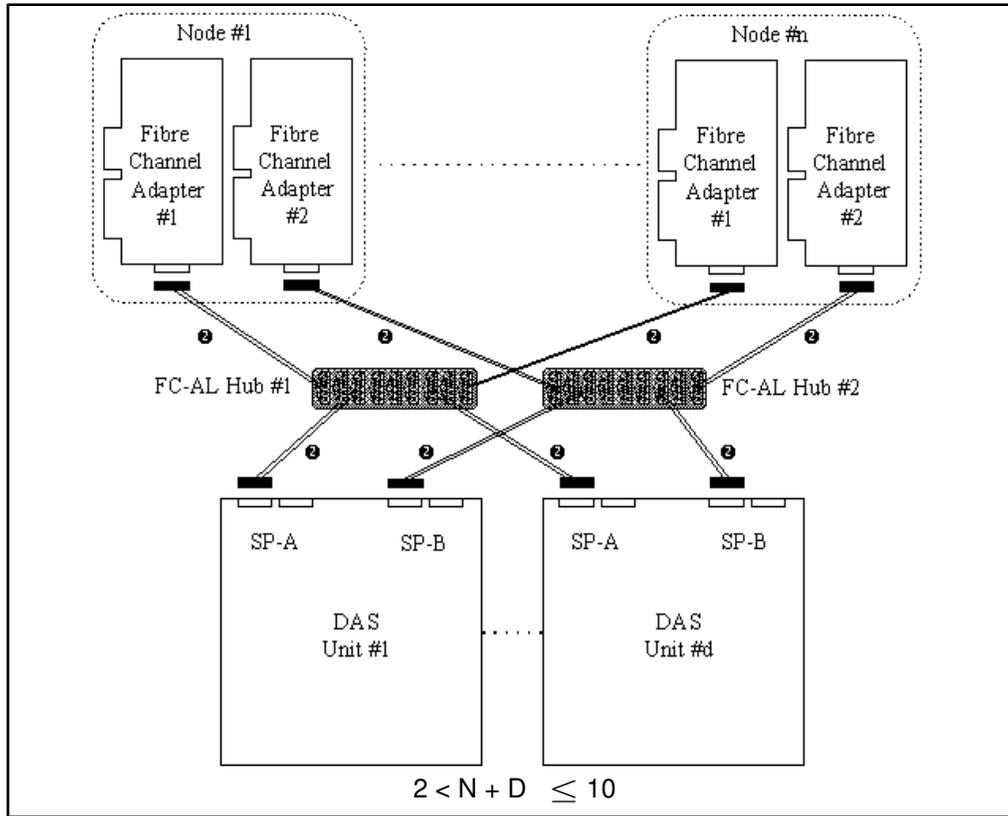
## Double boucle

Avec deux initiateurs (2 noeuds) utilisant 1 système DAS avec 2 (SP).



### Plus de deux noeuds

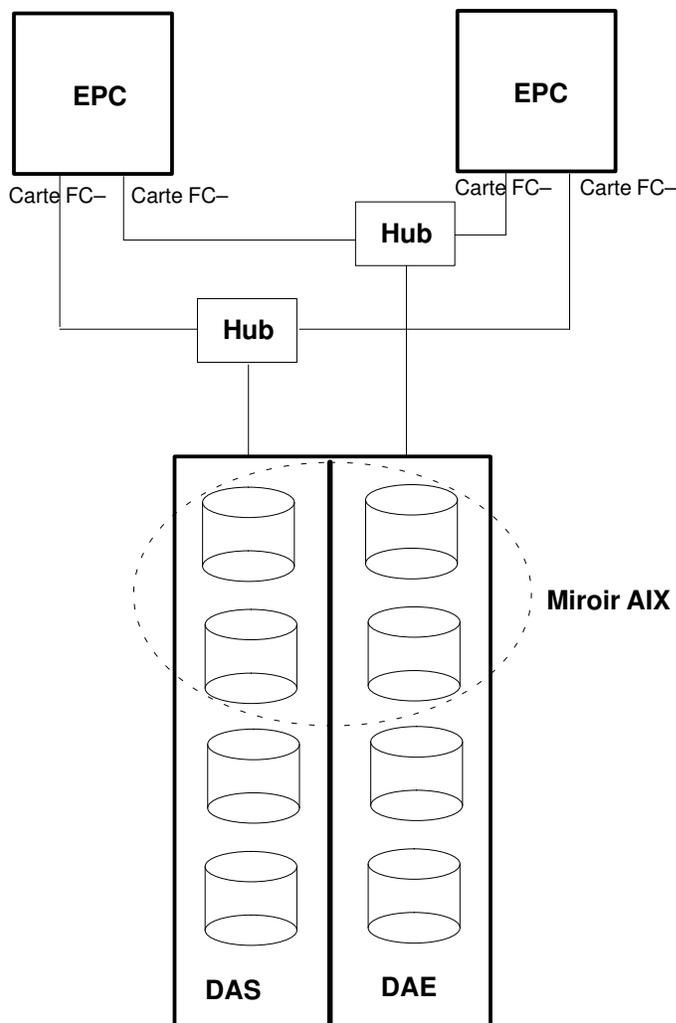
Avec doubles boucles, 2 concentrateurs, N noeuds, utilisant D DAS avec 2 SP.



## Insertion d'emplacements de disques en grappe (DAE)

Le système de stockage sur disques en grappe ultra-performant, haute capacité, utilisant la technique des boucles FC AL (Fibre Channel Arbitrated Loop) peut être exploité pour fournir des caractéristiques de haute disponibilité des données entre deux noeuds Escala Power Cluster.

La figure suivante illustre la configuration dans laquelle interviennent deux noeuds EPC :



Cette configuration permet d'assurer une grande disponibilité des données mises en miroir entre les deux systèmes DAE.

---

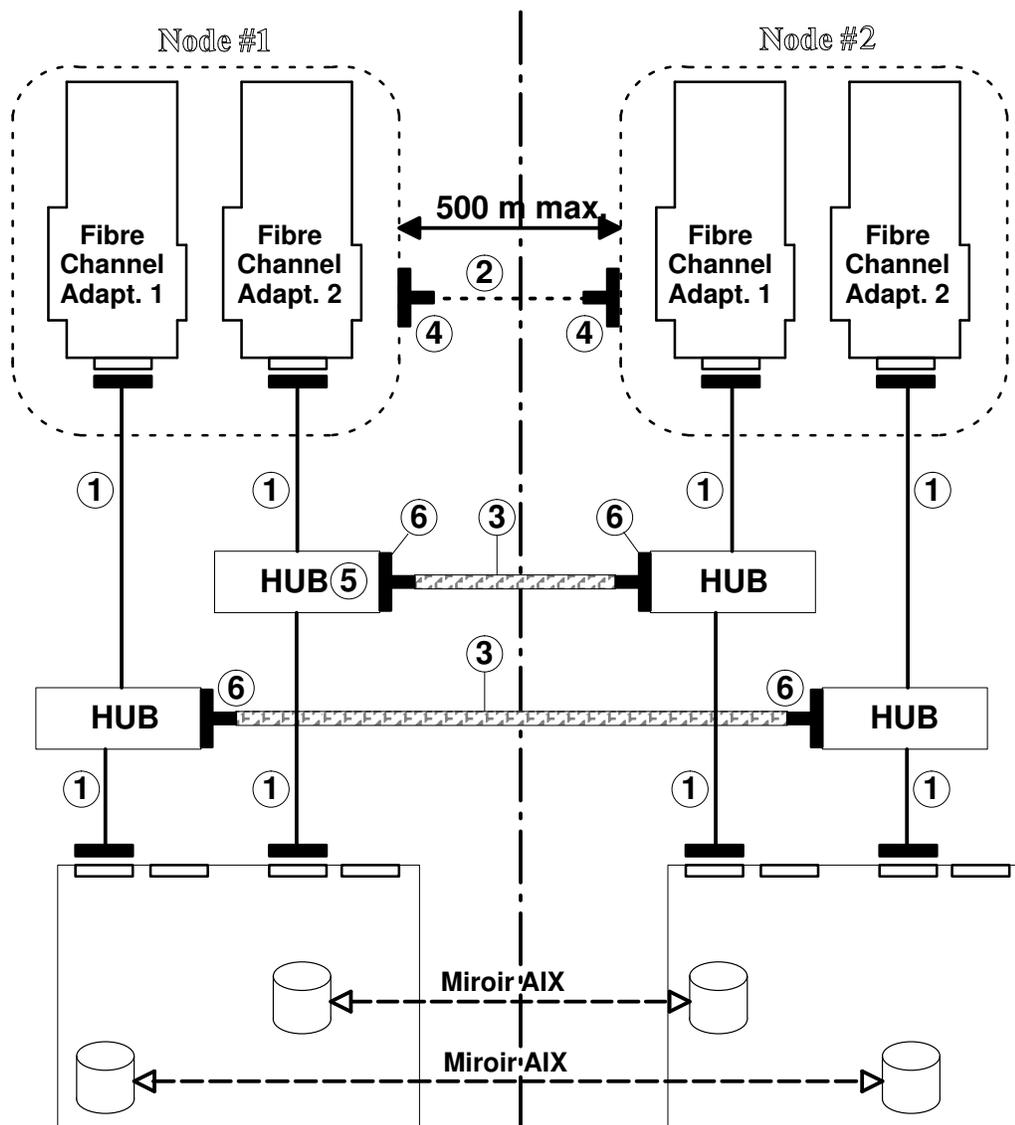
## Planification du stockage partagé HA étendue pour solutions Escala HA

La haute disponibilité étendue est la solution de reprise après incident proposée par Bull. Elle repose sur une technologie de sous-système à fibre permettant la connectivité de disques partagés pour une distance supérieure à 500 mètres. La fonction de duplication des données est assurée par la fonction miroir AIX.

Les fonctionnalités HA sont les mêmes que pour les clusters locaux, à condition de prendre un certain nombre de précautions (voir page 10-24).

### Câblage

Configurez les boucles FC–AL ainsi que la liaison série étendue (HACMP heartbeat de survie) comme indiqué dans la figure. Aucune interconnexion n'est disponible pour cette configuration.



1	Câble en cuivre
2	Câble RS232 étendu (keep-alive heartbeat)
3	Câble en fibre optique
4	Micro-modem pour RS232
5	Concentrateur FC-AL
6	Adaptateur d'interface MIA (Media Interface Adapter) (cuivre à fibre optique)

### Remarque spécifique des boucles FC-AL partagées

Définissez un ID de boucle différent sur chaque SP dans les deux systèmes DAS à fibre. Mettez sous tension les concentrateurs et les deux systèmes DAS, puis attendez que le voyant de service sur chaque DAS s'allume puis s'éteigne avant de mettre sous tension les serveurs Escala.

### Planification de la configuration des unités de disque physiques

Configurez les deux systèmes DAS comme décrit dans la documentation DAS : liez les LUN et rendez les disques accessibles aux systèmes.

**Remarque :** Les systèmes DAS à fibre permettent de créer jusqu'à 32 LUN et de les rendre accessibles à chaque SP.

---

## Planification des sous-systèmes SSA partagés

**Remarque** : Pour une présentation de SSA (Serial Storage Architecture), reportez-vous à la page 1-10 de ce manuel.

Reportez-vous aux documents de référence suivants :

- le chapitre "Planning Shared Disk Devices" du manuel *HACMP Planning Guide*,
- la documentation des cartes SSA et des sous-systèmes de disque SSA.

En complément, les informations et conseils suivants vous seront utiles.

### Cartes SSA

#### Bus MCA

Il existe deux modèles de carte SSA :

**Carte SSA 4 ports (type 4-D)** (MI = MSCG021)

Avec ce modèle (libellé "4-D"), il est impossible de connecter plus de deux cartes SSA dans une boucle SSA.

**Carte SSA étendue 4-ports (type 4-G)** (MI = MSCU-101)

Avec ce modèle (libellé "4-G"), jusqu'à 8 cartes SSA peuvent être connectées dans une boucle SSA. De ce fait :

- Si la grappe est formée de plus de deux noeuds ou si les noeuds sont équipés de 2 cartes SSA, ce modèle est indispensable.
- En outre, envisagez toujours d'utiliser des cartes étendues (même si la grappe ne comporte que 2 noeuds). Vous aurez ainsi plus de latitude pour des évolutions ultérieures de la grappe (ajout de noeud ou de carte SSA).

**Carte SSA étendue multi-initiateur / boucle RAID (4-M)** (MI = MSCU-038)

Deux cartes SSA au plus peuvent être connectées dans une boucle SSA. Dans les configurations RAID, cette carte fonctionne uniquement en mode mono-initiateur.

#### Bus PCI

Un modèle de carte SSA est disponible :

**Carte SSA étendue multi-initiateur / boucle RAID (4-N)** (MI = MSCG-039)

Deux cartes SSA au plus peuvent être connectées dans une boucle SSA. Dans les configurations RAID, cette carte fonctionne uniquement en mode mono-initiateur.

**Remarque** : Le panachage PCI/MCA sur une seule boucle SSA est possible entre cartes SSA PCI multi-initiateur (4-N) et SSA MCA 4 ports (type 4-M) exclusivement.

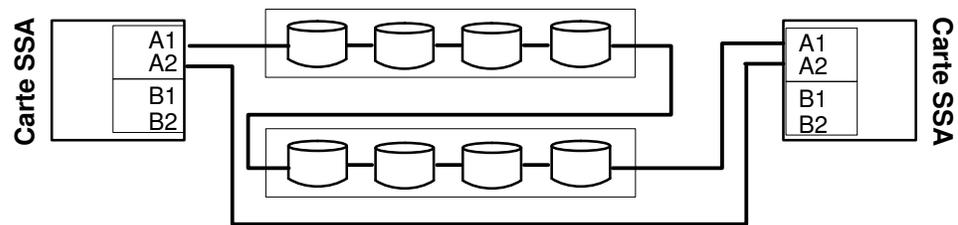
## Topologie SSA

Pour la planification du câblage, pensez aux points suivants :

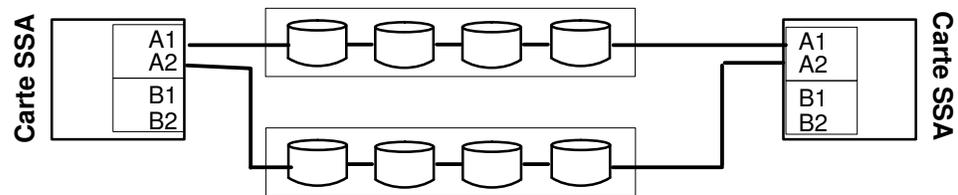
- La topologie en boucle est la plus adaptée à l'environnement de grappe. La topologie en chaîne point-à-point est inadaptée.
- Étant dotée de 4 ports, chaque carte SSA peut être connectée à une boucle ou à deux boucles distinctes.
- Pour la haute disponibilité, vous devez implanter la fonction disque miroir. Cette fonction doit être implantée via les utilitaires standard d'AIX.

Cette fonction, si vous décidez de câbler deux boucles SSA (ce qui augmente la disponibilité), doit être implantée sur les deux boucles, c'est-à-dire que les disques d'une boucle doivent être dupliqués par la fonction miroir sur les disques de l'autre boucle. En outre, si vous pouvez accéder à la copie miroir par le biais de deux cartes, la disponibilité est encore accrue.

- Pour les performances, évitez de connecter des cartes adjacentes dans un noeud. La figure ci-après illustre cette recommandation : dans la première topologie illustrée, deux cartes adjacentes figurent dans la boucle ; ceci n'est pas le cas de la topologie suivante qui est celle recommandée. Les performances sont légèrement supérieures, puisque le chemin d'accès entre les cartes et les disques est moindre.



La topologie ci-dessus est opérationnelle, mais réduit le débit. Celle ci-dessous offre des performances légèrement supérieures.



## Exemple de topologie SSA

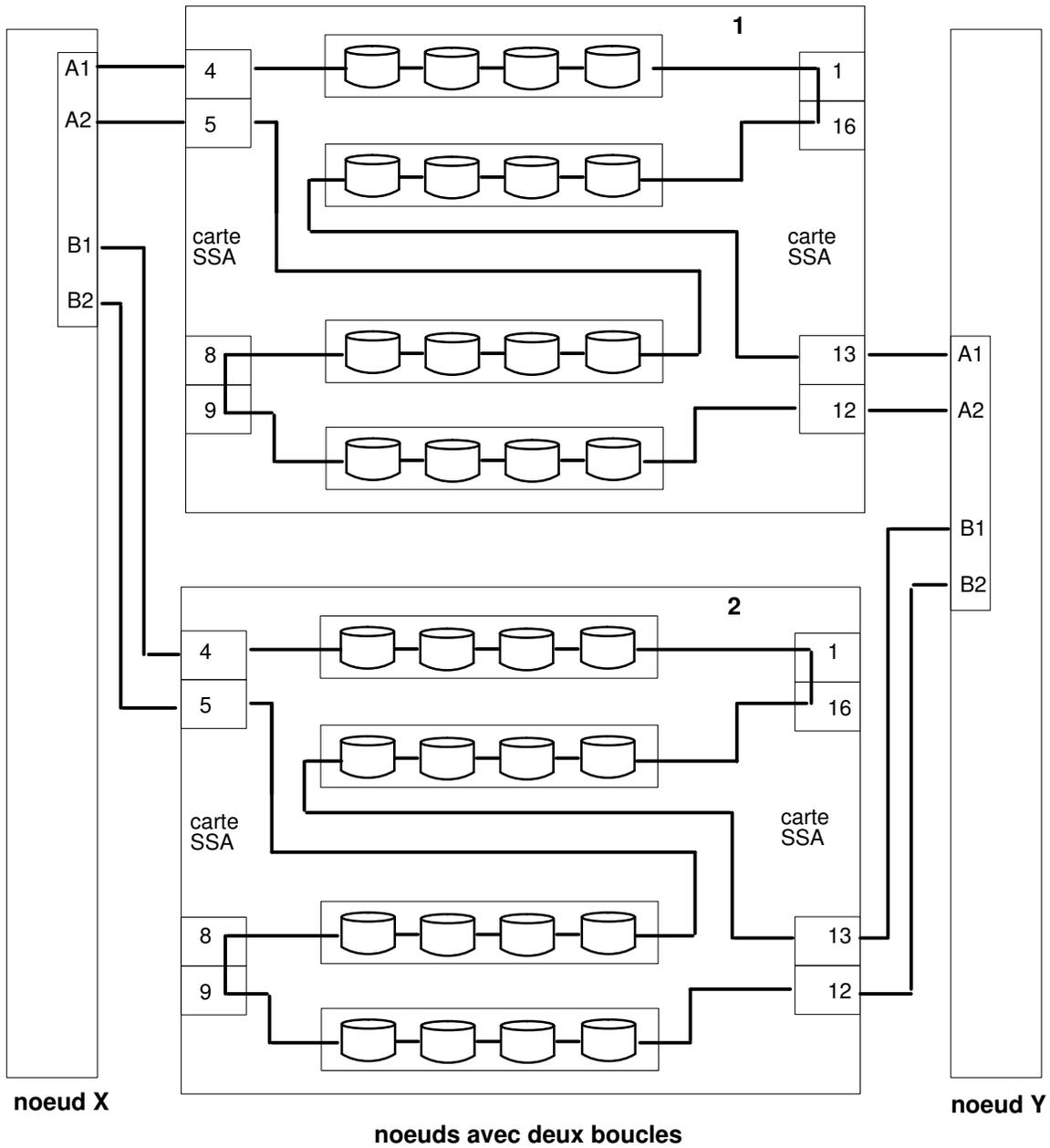
Comme indiqué plus haut, pour les configurations avec deux boucles SSA, la mise en miroir doivent être (autant que possible) implémentée sur les deux boucles. En pratique, toutefois, concevoir une topologie adaptée demande une planification précise et une sérieuse réflexion sur le sujet. A titre d'exemple, voici un cas pas très simple à résoudre.

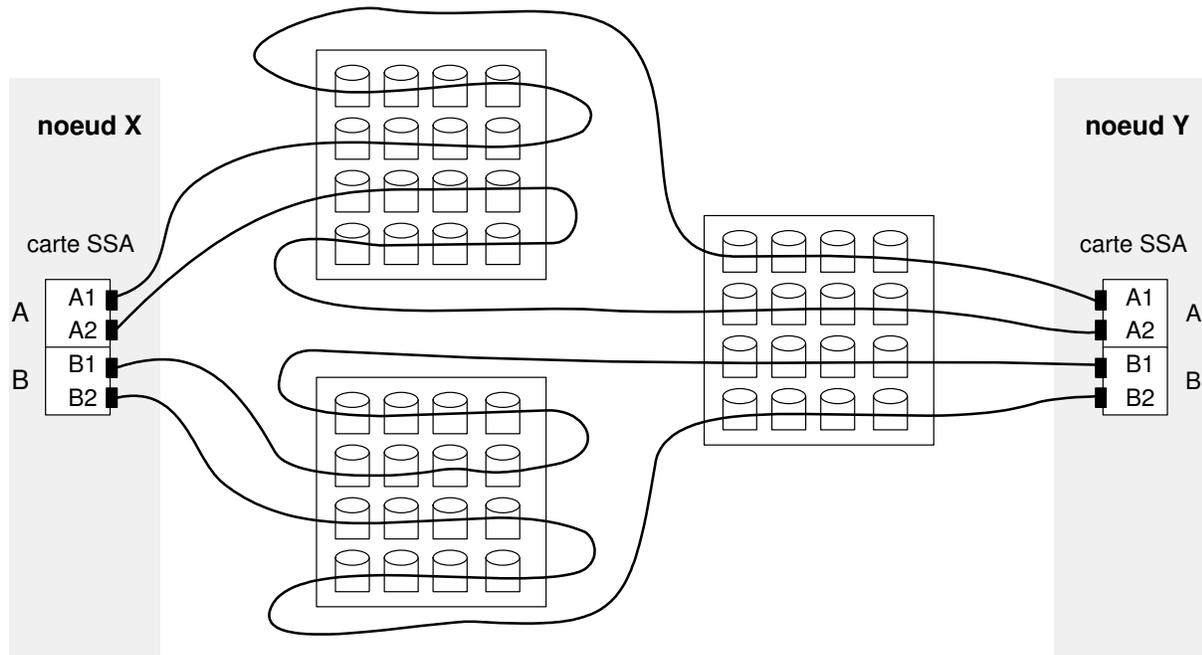
### Exemple de topologie avec une carte SSA par noeud

Supposons que vous souhaitez implémenter une mise en miroir inter-boucle et que vous disposiez de l'équipement suivant :

- la grappe, constituée de deux noeuds,
- une seule carte SSA par noeud,
- les sous-systèmes SSA peuvent recevoir jusqu'à 16 disques, mais ils peuvent être équipés de 32 disques au total à connecter dans une configuration à deux boucles (16 disques sur une boucle dupliqués sur 16 disques de l'autre boucle),
- 8 câbles de sous-système SSA.

Un exemple classique est illustré ci-après.





#### Variante de topologie avec deux cartes SSA par noeud

Si les noeuds sont chacun équipés de deux cartes SSA, la disponibilité obtenue est encore accrue et la mise en miroir peut être non seulement inter-boucle, mais aussi inter-carte (une boucle étant implémentée via une carte sur chaque noeud, l'autre boucle utilisant les deux autres cartes).

## Planification de sous-systèmes EMC Symmetrix partagés

EMC Symmetrix est un sous-système de disques offrant toutes les fonctions de stockage dans une seule armoire. Il est constitué des éléments suivants :

- un double bus interne,
- sélecteurs de canaux prenant en charge la gestion des liens hôtes et les fonctions de contrôle du stockage,
- cartes de mémoire cache,
- sélecteurs de disque prenant en charge les fonctions de stockage des données,
- bloc d'alimentation, sous-systèmes de batterie et modules de refroidissement,
- processeur de services intégré (ordinateur portable) permettant de télécharger la configuration Symmetrix vers les sélecteurs et d'exécuter des outils de diagnostic et de maintenance pour Symmetrix.

Les sélecteurs de canaux sont disponibles dans deux versions qui peuvent cohabiter sur un sous-système Symmetrix :

- sélecteurs SCSI Ultra Wide SCSI avec quatre ports hôtes
- sélecteurs Fibre Channel avec deux ports hôtes

Le sélecteur de disque fournit une interface entre les unités de cache et de disque, et il gère quatre bus SCSI Fast Wide dans les disques internes. Selon le modèle, quatre ou six disques sont connectés à chaque bus SCSI interne.

La connexion d'un port SCSI différentiel Fast Wide Symmetrix à un serveur AIX s'effectue via des cartes MCA ou PCI.

La connexion d'un port Fibre Channel ou d'un sélecteur Fibre Channel à un serveur AIX/PCI s'effectue via une carte Bull Fibre Channel. Le pilote Bull Fibre Channel AIX supporte le sous-système Symmetrix FC-AL. Le port Symmetrix Fibre Channel utilisant un connecteur à fibre optique, les connexions à Symmetrix doivent être assurées par un câble à fibre optique équipé d'un adaptateur MIA à l'autre extrémité de celui-ci. La partie cuivrée de l'adaptateur est connectée à un adaptateur Emulex.

Les fonctions de disque suivantes sont fournies par Symmetrix :

- RAID-1
- RAID-S, protection RAID de niveau 5 ultra performante
- Changement dynamique de disque

## Configuration de base avec HACMP

La configuration HA classique pour des sous-systèmes Symmetrix est une connexion point à point (appelée configuration en bus distincts dans le présent document) sur chaque serveur, accompagnée d'une configuration du sous-système Symmetrix de sorte que le volume de données soit disponible pour les deux serveurs à l'aide des deux ports hôtes distincts.

Pour le modèle de serveur parallèle Oracle, les volumes stockés doivent être partagés par tous les noeuds de la grappe.

La configuration CMP classique pour les sous-systèmes Symmetrix est une connexion point à point entre tous les serveurs et le sous-système Symmetrix, ainsi qu'une configuration de ce dernier permettant de faire en sorte que les volumes de données soient disponibles pour les deux serveurs via tous les ports hôtes distincts.

## Sous-systèmes de disque EMC en mode concurrent

Avec les sous-systèmes EMC Symmetrix, les disques peuvent être utilisés en mode concurrent si les commandes suivantes sont exécutées sur chaque disque concerné :

```
chdev -l hdisk<x> -a 'reserve_lock=no'
```

---

## Chapitre 4. Installation du matériel de base

Ce chapitre traite des points suivants :

- Installation physique des serveurs (noeuds) et des unités de disques
- Installation du matériel pour la gestion de la console système, page 4-2.
- Raccordement des connexions réseau, page 4-2.
- Amorçage des noeuds et vérification du fonctionnement, page 4-4.
- Configuration TCP/IP des interfaces d'amorçage/de service, page 4-4.
- Configuration des ID SCSI de carte (et réamorçage des noeuds concernés), page 4-7.

Suivez les instructions pas à pas avant de passer au chapitre suivant.

---

### Préalables

- Ce chapitre donne des conseils généraux. Pour les instructions autres que celles spécifiques aux configurations en grappe, reportez-vous à la documentation fournie avec le matériel.
- Avant de procéder à l'installation du matériel, nous vous conseillons de parcourir le manuel *HACMP Planning Guide* ainsi que le chapitre "Informations de planifications complémentaires", page 3-1 de ce manuel.

---

### Installation physique des serveurs (noeuds) et des unités de disque

#### Généralités

- Déballez les serveurs et les unités de disque externes (le cas échéant).
- Déterminez leur emplacement respectif sur le site en tenant compte des diverses contraintes : espace adapté à l'exploitation et la maintenance, longueur et cheminement des câbles, etc.
- Mettez les unités en place.

#### ATTENTION :

**Ne connectez pas maintenant l'unité de disque partagée externe aux noeuds !**

(Les ID SCSI et la configuration matérielle des cartes SCSI externes ne sont peut-être pas correctes.)

- Pour l'instant, ne mettez aucun noeud ni aucune unité sous tension.

#### Instructions spécifiques des modèles Powercluster

Si vous installez un Powercluster (c'est-à-dire une grappe basée sur les modèles en armoire Escala), installez dès à présent physiquement les unités dans l'armoire. Cependant, comme il est conseillé plus haut, ne connectez pas maintenant les sous-systèmes RAID DAS partagés aux noeuds.

Selon le modèle de Powercluster, vous devez installer une ou deux armoires. Si vous mettez en service une configuration qui inclut plusieurs armoires, gardez à l'esprit qu'elles sont totalement indépendantes du point de vue de l'alimentation : vous devez donc disposer d'une prise secteur distincte par armoire.

---

## Installation du matériel pour la gestion de la console système

Pour gérer les des différents noeuds de la grappe, il existe plusieurs solutions, sur la base des différents composants matériels.

- Le plus simple consiste à raccorder un terminal à chaque noeud. C'est ce que nous préconisons pour les solutions HA. Si vous optez pour cette solution, passez à la "Procédure avec consoles standard", ci-après.
- Si vous optez pour une autre solution (un seul terminal pour tous les noeud, ou une solution basée sur les terminaux X ou le PowerConsole), reportez-vous à l'annexe "Mise en oeuvre des solutions console", page A-1.

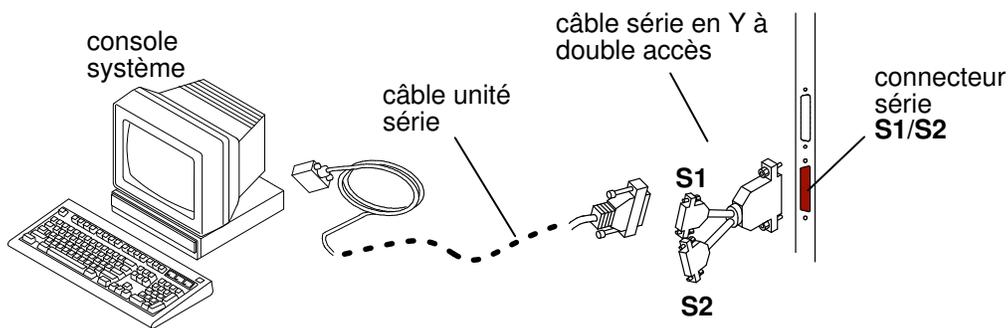
### Procédure avec consoles standard

Cette procédure s'applique aux configurations classiques qui utilisent une console système (une unité terminale) par noeud de grappe.

Connectez une console système à chaque noeud. Utilisez à cet effet le premier port série (généralement libellé "S1" ou "async L1").

### Remarques :

- Les modèles Escala mini-tour sont équipés d'un double port série. Dans ce cas, un seul connecteur, généralement libellé "S1/S2", fournit deux ports via un câble série en Y à double accès. Utilisez ce câble comme illustré ci-après.



---

## Raccordement des connexions réseau

### Raccordement des connexions au réseau public

Chaque noeud est équipé de deux cartes réseau par réseau connecté : une carte de service et une carte de secours. Chaque paire de cartes doit être connectée à chaque réseau, comme illustré plus loin. Voir illustration ci-dessous.

Connectez les carte au(x) réseau(x). Au besoin, reportez-vous à la documentation matérielle pour des précisions sur les connexions réseau (Ethernet, FDDI, ou ATM).

### Raccordement des connexions au réseau privé

Les réseaux série ne concernent que les grappes qui implémentent l'accès concurrent ou les architecture client/serveur.

Si vous installez un réseau privé (Ethernet ou FDDI), effectuez les connexions appropriées.

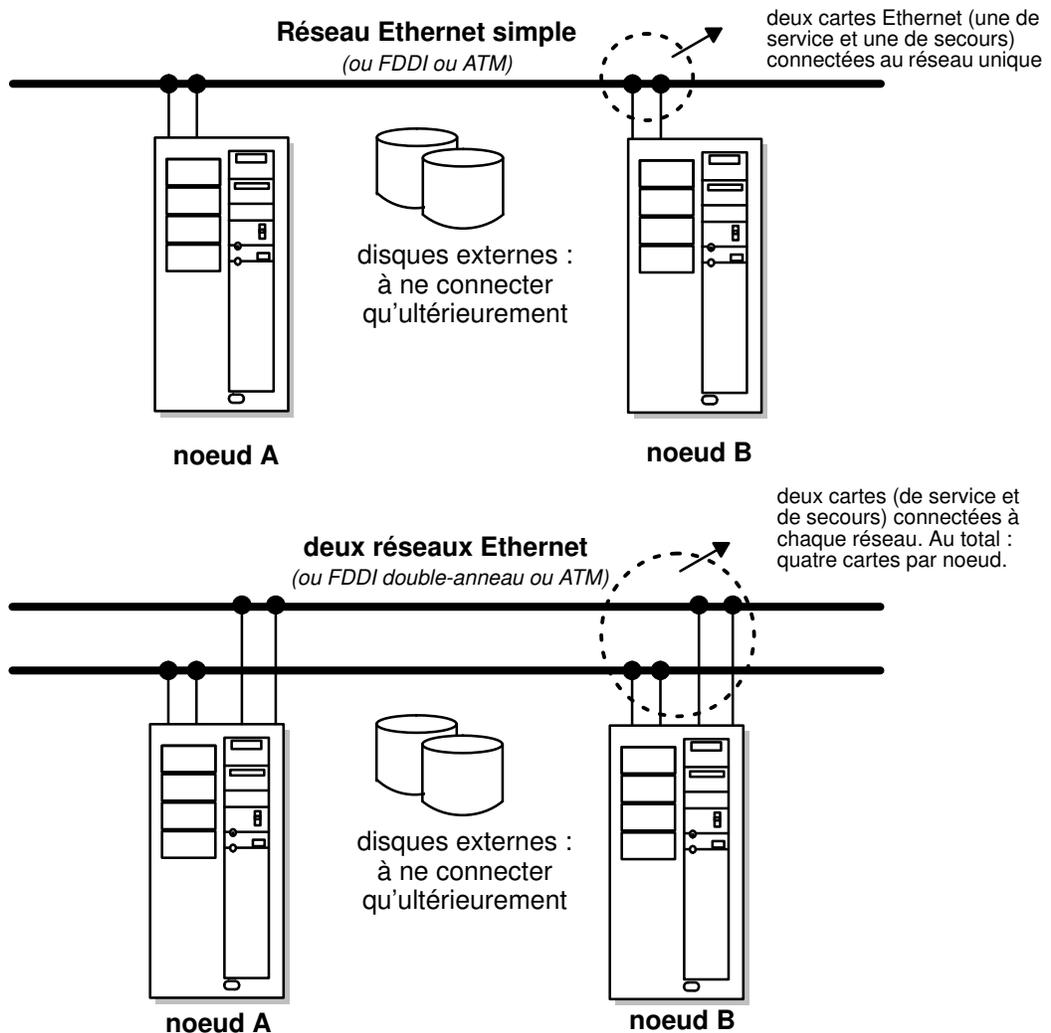
### Remarques :

- Reportez-vous à Planification des réseaux, page 3-3, pour les informations relatives aux réseaux publics et privés. Notez que :
  - En ce qui concerne le réseau public, les configurations de base de grappe n'utilisent qu'un seul réseau Ethernet : ce type de configuration est courant. Cependant, pour accroître la disponibilité, les noeuds d'une grappe peuvent être interconnectés par plusieurs réseaux (généralement, deux).
  - Si vous utilisez un réseau privé FDDI, vous pouvez décider d'implémenter une configuration jumelée.
- Pour en savoir plus sur la câblage des réseaux Ethernet et FDDI, reportez-vous également à la documentation matérielle.
- Pour implémenter un réseau privé basé sur Fibre Channel, prenez contact avec votre représentant.

### Votre configuration matérielle de base est désormais semblable à ceci :

Deux cas sont représentés ci-dessous. La première figure illustre une configuration classique : un réseau Ethernet avec deux cartes réseau sur chaque noeud. La seconde représente une configuration-HA (haute disponibilité) comprenant deux réseaux Ethernet distincts avec quatre cartes réseau sur chaque noeud.

**Remarque :** Ethernet a été choisi comme exemple pour illustrer les instructions (si vous exploitez FDDI ou ATM pour implémenter vos réseaux, adaptez les exemples en conséquence).



---

## Amorçage des noeuds et vérification du fonctionnement

Mettez chaque noeud sous tension et vérifiez que l'amorçage s'effectue normalement.

### Remarques :

- La configuration actuelle n'est pas encore organisée en grappe. L'opération d'amorçage à ce stade ne vise qu'à vérifier l'état du matériel avant de poursuivre l'installation.
- Si la grappe est équipée d'un nombre limité de consoles système, il est plus pratique d'amorcer en premier les noeuds auxquels une console est connectée. Vous pouvez ensuite lancer les commandes spéciales pour vous connecter aux ports sérieS1 (dédiés à la gestion des console système) des noeuds qui ne sont pas directement connectés à une console. Ces commandes spéciales sont décrites dans l'annexe "Mise en oeuvre des solutions console", page A-1.

---

## Configuration TCP/IP des interfaces d'amorçage/de service

**Remarque :** Les instructions ci-après concernent les cartes connectées au réseau public.

Cette étape consiste à exécuter la commande **mktcpip** sur chacun des noeuds pour définir le label IP, l'adresse IP et le masque de réseau de la carte réseau qui doit servir d'interface d'amorçage ou de service. Pour le moment, il est inutile d'intervenir sur la carte de secours, elle sera configurée ultérieurement avec l'outil de configuration rapide.

Cette étape suppose que vous êtes familiarisé avec HACMP et la mise en réseau. Et, plus particulièrement, que les concepts de cartes d'amorçage/de service/de secours, labels IP, adresses IP, masques de réseau, etc. ne vous sont pas étrangers. Pour mémoire, reportez-vous à "Prérequis" ci-après, sinon passez directement à "Conventions d'appellation".

### Prérequis : IPAT/adresses d'amorçage, de secours et de service

Pour plus de détails sur les sujets suivants, reportez-vous à la documentation HACMP.

### Basculement d'adresse IP

Généralement, vous souhaitez configurer la grappe pour le basculement d'adresse IP (parfois appelé IPAT). Le basculement d'adresse IP consiste à déplacer l'adresse de service IP du noeud défaillant vers un noeud restant. Lorsque cette fonction est active, les clients se reconnectent à la même adresse IP pour communiquer avec leur application (ce qui suppose que l'application concernée a également été reprise par le noeud restant).

Notez cependant que le basculement d'adresse-IP n'est pas une fonction obligatoire : vous pouvez l'activer ou non, en fonction de vos besoins propres. En cas de défaillance d'un noeud, et si vous n'avez pas activé le basculement, ses ressources sont reprises par un noeud restant, mais non son adresse de service IP. Les clients ne peuvent donc plus utiliser cette adresse IP pour communiquer avec leur application : ils doivent utiliser l'adresse IP de service du noeud qui a repris l'application en question.

## Adresses IP d'amorçage, de secours et de service

Les noeuds de grappe sont généralement équipés de deux cartes réseau : une carte de service et une carte de secours (il n'est pas question ici des configurations où chaque noeud est connecté à deux réseaux distincts, et équipé de quatre cartes). Voici un bref rappel concernant les adresses IP.

- Si vous décidez de ne pas activer le basculement d'adresse IP sur un noeud (si vous n'avez pas besoin que son adresse IP soit basculée sur un autre noeud), alors la carte de service du noeud prend toujours en compte l'adresse de service IP (c'est-à-dire que la notion d'adresse IP d'amorçage n'a pas de sens pour ce noeud).
- Si vous décidez d'activer le basculement d'adresse IP sur un noeud, la notion d'adresse IP d'amorçage a un sens. Les adresses d'amorçage et de service sont prises en charge par la même carte de réseau, appelée carte de service :
  - Au moment de l'amorçage, tant que les services de grappe ne sont pas lancés, la carte de service prend en compte l'adresse IP d'amorçage. Cette adresse n'est donc utilisée que temporairement.
  - Au lancement des services de grappe, l'adresse IP sur la carte de service est commutée de l'adresse d'amorçage sur l'adresse de service. A partir de ce moment, la carte de service prend en compte l'adresse de service IP. Cette adresse devient ainsi l'adresse "normale", utilisée par les clients pour communiquer avec les applications lorsque HACMP est actif.
- L'adresse de secours IP est prise en charge par une seconde carte de réseau, appelée carte de secours. Cette carte est utile lorsque la carte de service du noeud ou un autre noeud défaille. Dans ce cas, la carte de secours est reconfigurée pour prendre en charge l'adresse de service IP impliquée dans l'incident.

Notez que **la carte de secours est réservée aux opérations de basculement HACMP** : elle n'est pas conçue pour gérer les connexions des applications qui ne font pas partie des ressources de grappe (c'est-à-dire qui n'appartiennent pas à un groupe de ressources HACMP).

## Conventions d'appellation

Selon les règles d'appellation appliquées par l'outil *Bullcluster Quick Initial Configuration*, les chaînes `_boot` et `_stby` sont toujours utilisées comme suffixes standard des noms de cartes d'amorçage et de secours.

Supposons, par exemple, que vous prévoyiez d'implémenter une grappe à deux noeuds appelés `foo` et `bar`. L'outil *Bullcluster quick initial configuration* exécute une grande partie des tâches de configuration et il en résulte :

- Le noeud `foo` est configuré avec les trois interfaces :

<code>foo_boot</code>	( <i>amorçage</i> )
<code>foo</code>	( <i>service</i> )
<code>foo_stby</code>	( <i>secours</i> )

- De la même façon, le noeud `bar` est configuré avec les trois interfaces :

<code>bar_boot</code>	( <i>amorçage</i> )
<code>bar</code>	( <i>service</i> )
<code>bar_stby</code>	( <i>secours</i> )

Notez que vous pouvez utiliser l'outil *Bullcluster Quick Initial Configuration* que vous ayez planifié ou non l'activation du basculement d'adresse IP. Dans la négative, les noeuds sont configurés avec seulement deux interfaces (de service et de secours), `foo` et `foo_stby`, par exemple car les adresses d'amorçage IP n'ont pas de sens.

L'application des conventions d'appellation est :

- obligatoire si vous utilisez l'outil *Bullcluster quick initial configuration*,
- facultative sinon.

## Procédure

Pour chaque noeud, procédez comme suit :

1. En tant qu'utilisateur racine, entrez :

```
smit mktcpip
```

L'écran **Interfaces de réseau disponibles** s'affiche avec la liste des interfaces disponibles sur le noeud pour les communications réseau.

2. Sélectionnez la carte qui doit jouer le rôle de carte de service.

L'écran **Configuration minimale & lancement** s'affiche. Son contenu dépend du type de carte réseau sélectionnée à l'étape précédente (Ethernet ou FDDI).

3. Passez à la zone **NOM HÔTE**. Si vous souhaitez utiliser l'outil *Bullcluster Quick Initial Configuration* pour configurer la grappe, vous devez observer les conventions d'appellation exposées plus haut.

- Si vous envisagez d'activer le basculement d'adresse IP, entrez dans la zone **NOM HÔTE** le nom de l'interface d'amorçage, de la forme :

```
nom-noeud_boot
```

où :

- *nom-noeud* est le nom du noeud, c'est-à-dire le nom que vous devrez indiquer lors de la configuration de la grappe.
- La chaîne *\_boot* est, par convention, le suffixe à utiliser avec l'outil *Bullcluster quick initial configuration*.
- Si vous n'envisagez pas d'activer le basculement d'adresse IP, dans la zone **NOM HÔTE**, entrez le nom de l'interface de service, de la forme :

```
nom-noeud
```

4. Dans la zone **ADRESSE Internet**, spécifiez l'adresse IP à attribuer à l'interface (d'amorçage ou de service).
5. Le cas échéant, renseignez la zone **MASQUE de réseau** en fonction de votre configuration.
6. Renseignez la zone **Votre type de CÂBLE**.
7. Appuyez sur Entrée pour valider la configuration.
8. Veillez à appliquer cette procédure **sur chaque noeud** à configurer.

### Remarques :

- Lors de la validation, le nom spécifié dans la zone **NOM HÔTE** devient le nom hôte par défaut du noeud. Par exemple, si vous avez indiqué *foo\_boot*, le nom hôte du noeud est désormais *foo\_boot*.
- Par ailleurs, une entrée est automatiquement ajoutée à **/etc/hosts** pour le nom hôte spécifié.
- Rappelez-vous que le nom hôte (au sens courant d'AIX) et le nom de noeud sont deux notions différentes. Ils ne se confondent que dans HACMP.

## Remarques importantes sur les noms d'hôte

Lors de la configuration via l'outil *Bullcluster Quick Initial Configuration*, les autres interfaces sont configurées automatiquement. Si vous décidez d'activer le basculement d'adresse IP, l'outil exécute les tâches de configuration de façon à définir l'adresse IP de service IP et la carte de service. **La procédure de configuration rapide modifier également le nom hôte, pour qu'il corresponde à l'interface de service.** Par exemple, le nom-hôte `foo_boot`, devient `foo`. Il est en effet peu souhaitable qu'un nom hôte corresponde à l'interface d'amorçage.

Si vous décidez de ne pas activer le basculement d'adresse IP, la procédure de configuration rapide ne modifie pas le nom-hôte (par exemple, le nom initial `foo` reste `foo`).

---

## Configuration des ID SCSI de carte

**Remarque :** Cette section ne vous concerne que si vous installez une grappe où le stockage sur disque partagé est assuré par des unités de disque SCSI. Si vous utilisez des sous-systèmes disque SSA, passez au chapitre suivant.

### Préalables

#### Notion d'ID SCSI

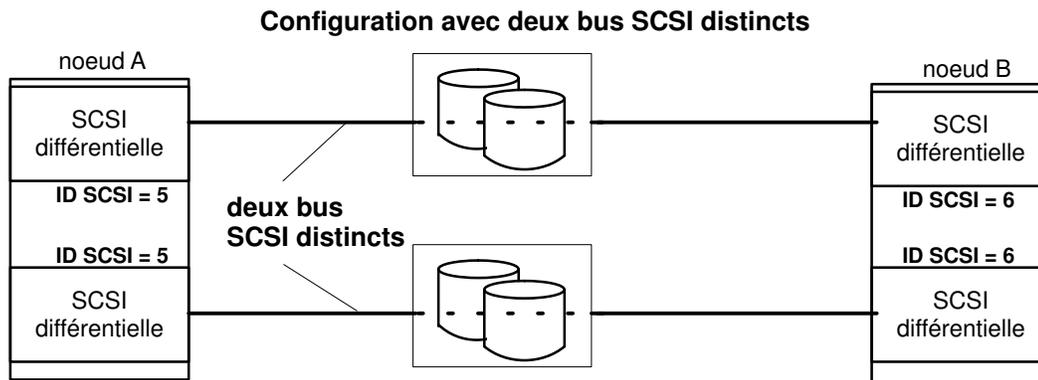
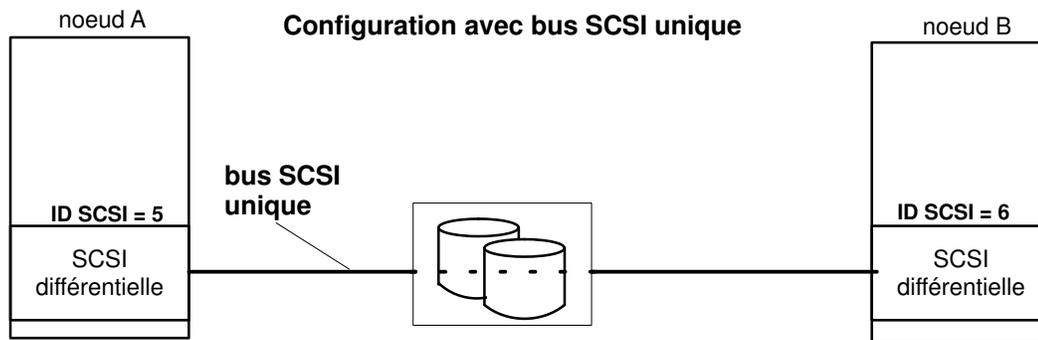
Dans une configuration en grappe, les noeuds partagent des unités de disque via des bus SCSI externes. Chaque bus SCSI partagé est pourvu d'une carte de noeud à chacune de ses extrémités. Ce type de configuration, avec deux serveurs raccordés au même bus SCSI, est aussi appelé "bus à deux têtes".

En ce qui concerne les ID SCSI :

- Chaque unité d'un bus SCSI doit disposer d'un ID SCSI unique.
- L'ID 7 ne doit pas être affecté aux cartes SCSI qui équipent ces noeuds (ceci parce que, si des diagnostics sont lancés, ils forcent l'ID SCSI de la carte à 7, provoquant alors un conflit d'adresse). Malheureusement, l'ID SCSI par défaut des cartes du noeud est 7.

Pour éviter qu'un ID SCSI ne soit utilisé deux fois sur un même bus et pour être sûr que l'ID 7 n'est utilisé sur aucun noeud, vous devez changer l'ID SCSI des cartes SCSI externes sur tous les noeuds.

**Remarque :** Certaines configurations de grappe ne prévoient qu'un seul bus SCSI partagé. Mais, dans un souci de haute disponibilité, les configurations classiques comportent deux bus SCSI partagés. La figure ci-dessous illustre ces deux cas.



Pour vérifier ou pour effectuer la configuration des ID SCSI, procédez comme suit.

## Configuration des ID SCSI de carte

**Remarque :** Si vous avez suivi la procédure d'installation expliquée dans ce guide, les unités de disque partagées externes ne sont toujours pas raccordées aux noeuds. Le câblage SCSI sera effectué plus tard, comme indiqué plus loin.

### Procédure de configuration

Pour configurer l'ID SCSI de la carte :

1. Sur un noeud, connectez-vous en tant qu'utilisateur racine et entrez :
 

```
smit chgscsi
```
2. La liste des cartes installées s'affiche. Sélectionnez la carte dont l'ID SCSI doit être modifiée.
  - a. Si la carte que vous configurez est de type "Corvette" (type "4-6"), choisissez-lui un nom approprié (commençant par le préfixe `ascsi` et libellée "Wide SCSI I/O Controller Adapter").
  - b. Si la carte que vous configurez est de type "WSA" (type "WSA1"), choisissez-lui un nom approprié (commençant par le préfixe `scsi` et libellée "Wide SCSI Adapter SCSI Driver").
  - c. Si vous configurez une carte SCSI PCI (cartes de type "B4-5" et "B4-6" utilisées sur les serveurs Escala), sélectionnez le nom approprié (commençant par le préfixe `scsi` et libellée "Wide/Fast-20 SCSI I/O Controller").

3. Un nouveau menu s'affiche. Examinez les valeurs indiquées dans la zone **External SCSI ID** (ou **Adapter card SCSI ID**, selon le type de carte). Si nécessaire, entrez l'ID SCSI ad hoc (cette étape n'est pas indispensable car il est possible que l'unité ait été configurée correctement à l'usine), sans oublier que :
  - L'ID 7 ne doit pas être utilisée pour les cartes de noeud.
  - Chaque unité de bus SCSI doit avoir un ID SCSI unique.
  - Il est conseillé de réserver les ID les plus élevés aux cartes de noeud et d'utiliser les ID les moins élevés pour les unités externes.
4. Spécifiez **yes** dans la zone **Apply change to database only** (conservez la valeur des autres zones).
5. Appuyez sur Entrée pour valider les modifications.
6. Si la grappe utilise deux bus SCSI, modifiez de la même façon, sur le même noeud, l'ID SCSI de la carte SCSI installée sur l'autre bus.
7. Répétez ces étapes sur les autres noeuds de grappe.

### Grappes composées de plus de deux noeuds

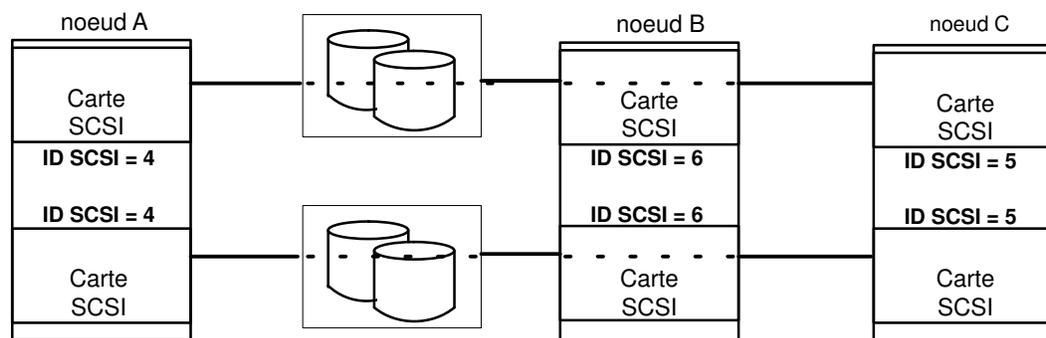
Si la grappe comprend plus de deux noeuds, les mêmes contraintes s'appliquent : les ID SCSI devant être uniques sur chaque bus SCSI, vous devez les modifier en conséquence sur plusieurs noeuds (voir figure ci-dessous).

Dans ce cas, répétez la procédure complète (**smit chgscsi**) pour les autres noeuds de grappe qui partagent des bus SCSI avec plusieurs noeuds. Les ID SCSI 6 et 5 étant déjà affectés (aux cartes de deux premiers noeuds), choisissez 4 pour le noeud suivant (optez toujours pour la valeur disponible la plus élevée).

### ATTENTION :

**Ne connectez pas maintenant les unités de disque externe partagées aux noeuds !**  
(Les ID SCSI et la configuration matérielle des cartes SCSI ne sont pas nécessairement définies correctement).

### Exemple de configuration avec trois noeuds et deux bus SCSI distincts



## Réamorçage des noeuds

Une fois les ID SCSI de carte modifiés sur un ou plusieurs noeuds, arrêtez le(s) noeud(s) et réamorcez-le(s) pour appliquer les nouveaux ID.

---

## Étape suivante

Passez au chapitre suivant pour vérifier l'installation logicielle.



---

## Chapitre 5. Vérification de l'installation logicielle

Ce chapitre traite des points suivants :

- Vérification de l'installation logicielle
- Bulletin de logiciel
- Vérification du préchargement de HACMP et BullCluster

Suivez les instructions avant de passer au chapitre suivant.

---

### Vérification de l'installation logicielle

Avant de poursuivre l'installation du matériel (câblage des réseaux série et des unités de disque), vérifiez que les noeuds sont équipés du logiciel approprié.

#### Remarques préliminaires

Pour exploiter une configuration en grappe, les noeuds de grappe requièrent les modules suivants :

- HACMP pour AIX,
- module BullCluster.

Ces deux modules, complémentaires, sont livrés sur le même support (CD-ROM).

### Bulletin de logiciel

Consultez le *Bulletin de logiciel* (SRB) livré avec le logiciel. Il fait état de toutes les conditions et contraintes associées à l'environnement, et fournit les informations de dernière minute.

Le bulletin indique également la procédure d'installation du logiciel. Ne l'appliquez pas à ce stade, le logiciel requis étant normalement préchargé en usine sur les systèmes livrés (voir ci-dessous).

### HACMP et BullCluster

Les modules HACMP et BullCluster sont normalement préchargés (c'est-à-dire pré-installés) sur les noeuds commandés pour constituer votre solution de grappe. Néanmoins, il est conseillé de vérifier que cette opération a bien été effectuée en usine.

Sur **chaque noeud** de la configuration, entrez :

```
lslpp -L '*cluster*'
```

La commande **lslpp** doit renvoyer :

- des ensembles de fichiers commençant par "cluster" : ils constituent le module HACMP.  
ET
- des ensembles de fichiers commençant par "bullcluster" : ils constituent le module BullCluster.

**Remarque :** L'ensemble de fichiers **cluster.clvm** correspondant à la partie HACMP qui gère l'accès concurrent. Cet ensemble de fichiers, fourni exclusivement sur le support dédié, n'est requis que dans les environnements avec accès concurrent. Il est superflu si vous installez une grappe haute-disponibilité qui n'implémente pas ce type d'accès ; dans ce cas, votre support ne contient pas l'ensemble de fichiers **cluster.clvm**.



---

## Chapitre 6. Finalisation de l'installation matérielle

Ce chapitre traite des points suivants :

- Arrêt et mise hors tension des noeuds
- Interconnexion des noeuds via des réseaux série (RS-232)
- Préparation des cartes SCSI au câblage, page 6-2.
- Installation des unités de disque partagées, page 6-6.

Suivez les instructions avant de passer au chapitre suivant.

---

### Arrêt et mise hors tension des noeuds

Avant de poursuivre l'installation des réseaux série et des disques partagés :

- Arrêtez tous les serveurs destinés à jouer le rôle de noeuds de grappe et mettez-les hors tension.
- Mettez toutes les unités raccordées hors tension.
- Débranchez tous les cordons d'alimentation (serveurs et unités). Ceci est vivement recommandé pour des raisons de sécurité.

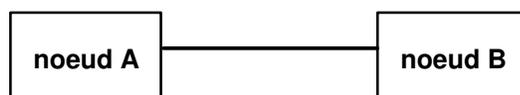
---

### Interconnexion des noeuds via des réseaux série (RS-232)

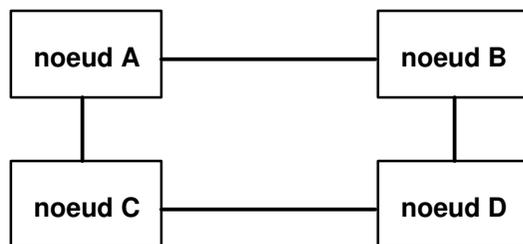
#### Préalables

Dans une grappe, les réseaux série sont des liaisons RS-232 qui relient les noeuds entre eux. Ces liaisons non TCP/IP transportent les flux de détection de survie ("keep-alive") générés par HACMP. En cas de défaillance du sous-système TCP/IP, les liaisons série continuent de prendre en charge ce flux de noeud à noeud.

Chaque noeud d'une grappe doit être connecté en série aux noeuds limitrophes. Les grappes limitées à deux noeuds ne requièrent qu'une seule liaison série entre les deux noeuds.



**Réseaux série pour grappe à 2 noeuds :**  
les deux noeuds sont interconnectés par une liaison RS-232.



**Réseaux série pour grappe à 4 noeuds :**  
les noeuds sont interconnectés par quatre liaisons RS-232 pour former un anneau logique.

#### Remarques :

- Pour en savoir plus sur le rôle des réseaux série, reportez-vous à "Planification des réseaux", page 3-3.
- Les réseaux série sont toujours de type RS-232 (le mode cible SCSI n'est pas pris en charge).

## Installation physique des lignes série RS-232

Installez physiquement les lignes série RS-232 entre les noeuds concernés.

Cette opération consiste à connecter chaque extrémité du câble RS-232 à un port série de chaque noeud. Notez que si vous décidez d'utiliser le logiciel RSF pour procéder à la télémaintenance via un modem, le port S2 doit être réservé aux opérations RSF (dans ce cas, n'utilisez pas le port S2 pour implémenter le réseau série). Pour en savoir plus sur les ports série, reportez-vous à la page 3-2.

---

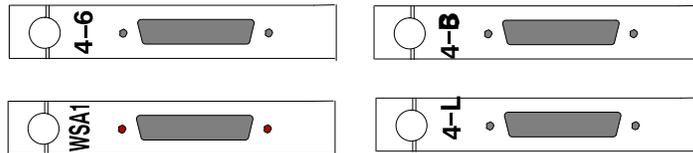
## Préparation des cartes SCSI au câblage

**Quand cette étape est-elle requise ?**

- Si vous installez un nouveau Powercluster ou une solution HA, cette étape est superflue car elle est exécutée en usine. Dans ce cas, passez directement à "Installation des unités de disque partagées", page 6-6.
- Cette étape n'est requise que si vous mettez à niveau des serveurs pour les intégrer à une grappe.

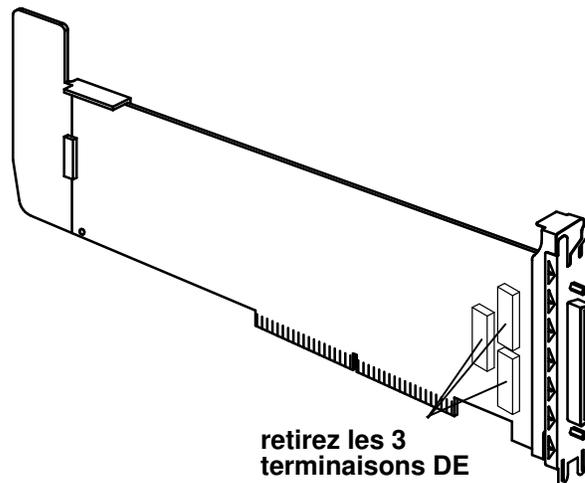
## Préparation des cartes SCSI Differential type 4-6, 4-B, 4-L ou WSA-1

Repérez les cartes SCSI-2 DE Fast/Wide installées à l'arrière des serveurs. Elles sont libellées "4-6", "4-B", "4-L" ou "WSA1" comme illustré ci-dessous.



A l'aide de la figure ci-dessous, repérez sur la carte les trois terminaisons à retirer. Si vous ne les voyez pas, cela signifie qu'elles ont été retirées en usine, et que vous pouvez réinstaller la carte immédiatement.

Retirez-les avec précaution. Mettez-les de côté pour un éventuel usage ultérieur.

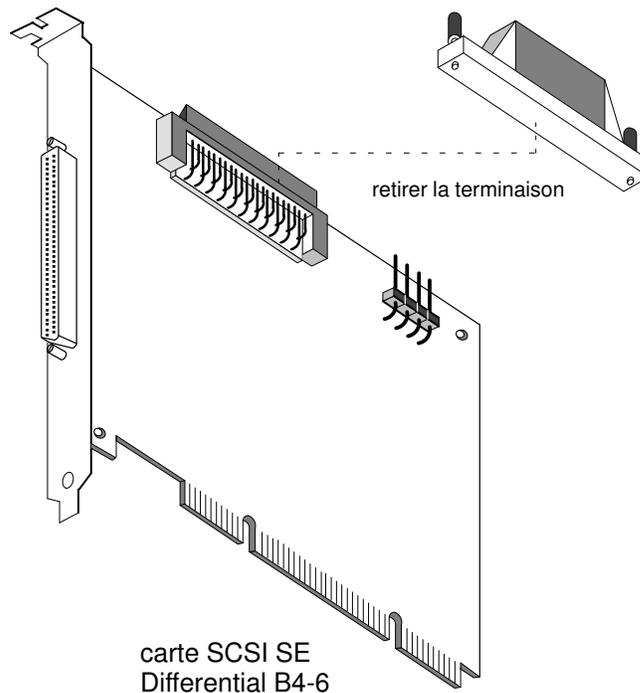


Une fois les terminaisons retirées sur toutes les cartes, procédez au câblage des unités de disque comme indiqué ci-après.

## Préparation des cartes SCSI Differential type B4-6

Les serveurs Escala EPC400 séries E et T peuvent être mis à niveau en ajoutant des cartes SCSI supplémentaires pour les configurer comme noeuds de grappe.

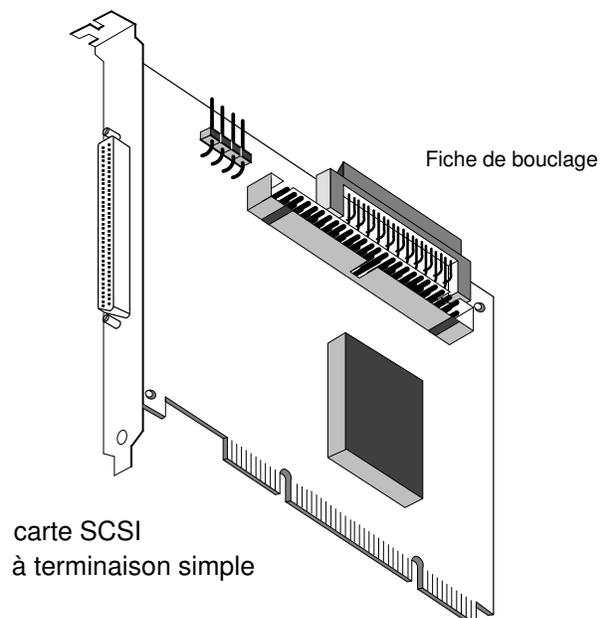
**Remarque :** Ces cartes sont également utilisées pour les noeuds en grappe Escala E200 et T400.



Pour en savoir plus sur les fiches de bouclage, reportez-vous à "Planification du stockage partagé "HA native" pour les solutions Escala HA", page 3-6.

## Préparation des cartes SCSI SE type B4-5 à terminaison simple

Les serveurs Escala série E et T peuvent être mis à niveau à l'aide de cartes SCSI SE supplémentaires afin de les configurer en tant que noeuds en grappe. **Remarque :** Pour installer ces cartes vous devez utiliser des fiches de bouclage.



Pour plus de détails sur ces fiches de bouclage, reportez-vous à la section "Planification du stockage partagé HA Native pour les solutions Escala HA", page 3-6.

## Préparation des cartes Fibre Channel type B4-7

Ces cartes sont utilisées dans les systèmes de stockage DAS 3500.

Pour la connexion à la couche physique (FC-0), un module GLM (Gigabaud Link Module) aux normes de l'industrie est utilisé. Ce module permet d'atteindre une vitesse de 1 063 Mbps et offre une connexion DB9 de type cuivre.

En outre, un adaptateur MIA (Module Interface Adapter) est utilisé pour fournir des doubles connecteurs SC fibre optique.

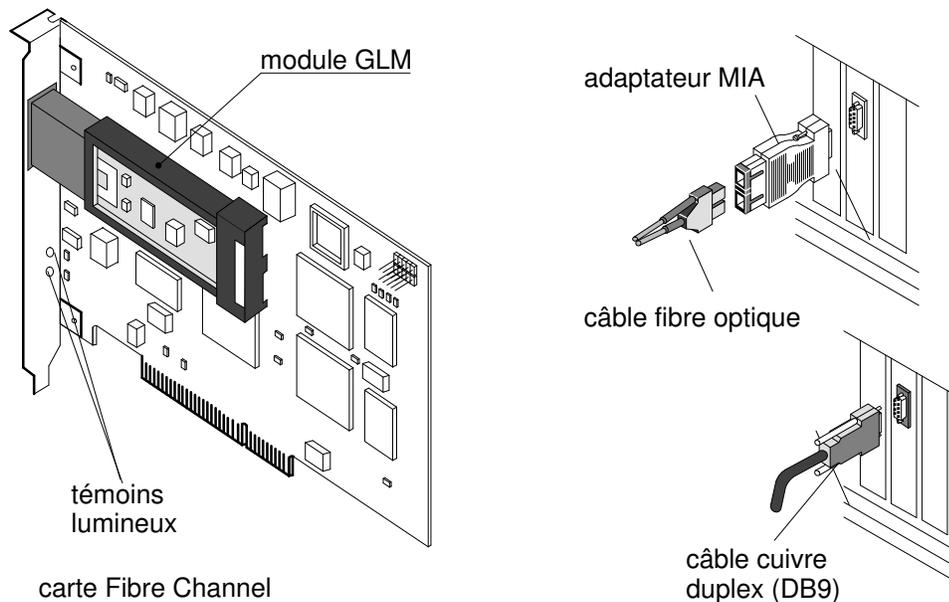


Figure 1. Carte PCI Fibre Channel (type B4-7)

## Préparation des cartes Fibre Channel type B4-8

Ces cartes sont utilisées sur des sous-systèmes Fibre DAS (DAS5700, DAS5720 et DAS5300).

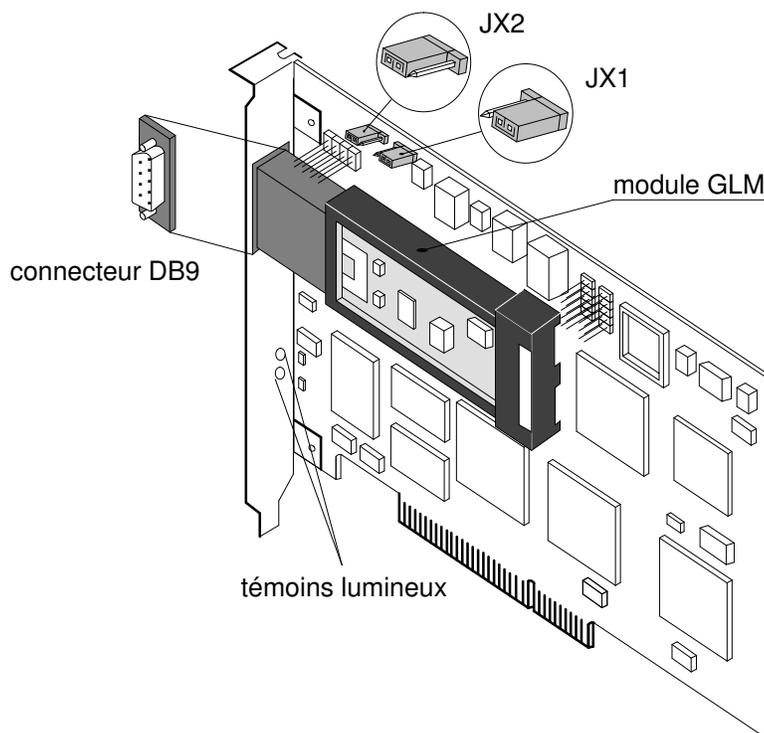


Figure 2. Carte PCI Enhanced Fibre Channel (type B4-8)

## Préparation des cartes Fibre Channel type B4-A et B4-B

Ces cartes s'utilisent sur des sous-systèmes Fibre DAS (DAS5700, DAS5720 et DAS5300).

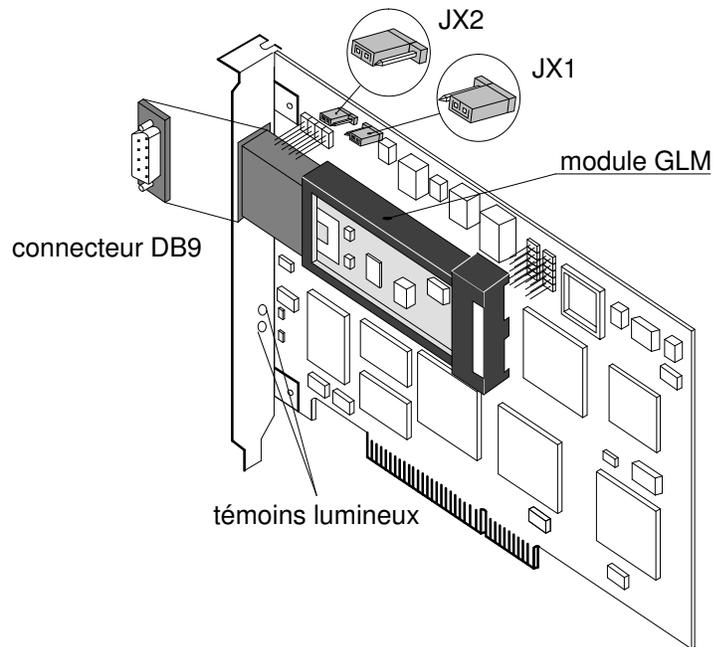


Figure 3. PCI 64 bits Carte Fibre Channel en cuivre (type B4-A)

Les cavaliers JX1 et JX2 doivent être installés comme illustré dans la figure 3.

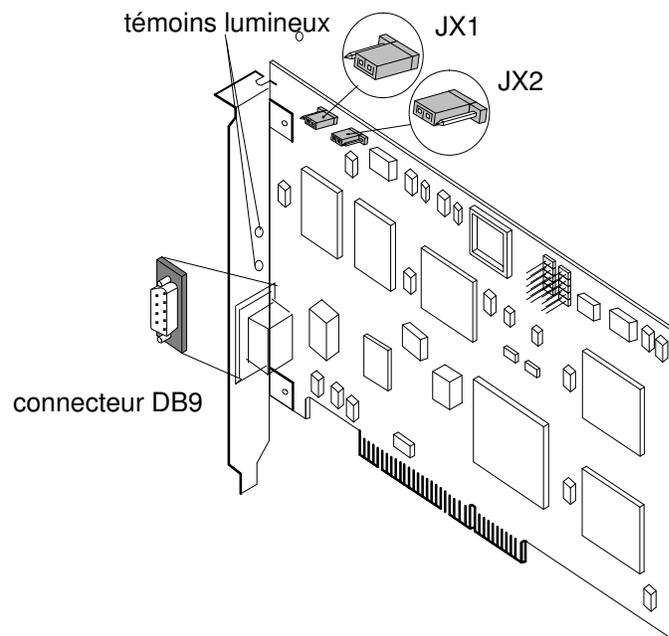


Figure 4. Carte optique Fibre channel PCI 64 bits (type B4-B)

Les cavaliers JX1 et JX2 doivent être installés comme illustré dans la figure 4.

## Préparation du concentrateur Fibre Channel

Le concentrateur (Hub) Fibre Channel doit être connecté aux noeuds avec des cartes Fibre Channel. Il doit aussi être raccordé aux unités DAS, page 3-18.

**Remarque :** Le concentrateur Fibre Channel n'a pas d'interrupteur d'alimentation. Il doit être le premier composant opérationnel, avant la mise sous tension des noeuds avec les adaptateurs FC-AL.

---

## Installation des unités de disque partagées

Il s'agit de câbler et d'installer les unités de disque partagées.

- Si vous installez des sous-systèmes disque SSA, reportez-vous à "Planification des sous-systèmes SSA partagés", page 3-22, et à la documentation de votre équipement SSA. Suivez la procédure décrite à la section "Installing Shared IBM 7133 SSA Disk Subsystems" dans le chapitre "Installing Shared Disk Devices" du manuel *HACMP Installation Guide* et passez ensuite au chapitre "Carrying Out Pre-Configuration Tasks".
- Si vous installez des unités SCSI ou Fibre Channel partagées, suivez la procédure ci-après. Pour un complément d'informations, reportez-vous à la documentation livrée avec les unités de disque et au chapitre Installing Shared Disk Devices du manuel *HACMP Installation Guide*.

### Câblage des unités de disque et définition de leur ID

**Avertissement :** Avant toute manipulation, vérifiez que les systèmes et unités sont hors tension. Débranchez tous les cordons d'alimentation, côtés secteur et unités. Observez les consignes de sécurité indiquées dans la documentation du matériel.

1. Connectez tous les câbles requis pour la configuration. Consultez au besoin le chapitre "Informations de planifications complémentaires", page 3-1, qui renseigne sur les câbles de disque (avec illustrations à l'appui). Pour en savoir plus, reportez-vous également à la documentation livrée avec vos unités de disque.
2. Définissez les ID SCSI de vos unités de disque. Chaque unité d'un bus SCSI doit être affecté d'un ID unique compris entre 0 et 7. Veillez à attribuer des ID disponibles (non assignés aux cartes SCSI des noeuds). (Pour en savoir plus sur les ID SCSI, reportez-vous à la page 4-7.)
3. Pour le système DAS Fibre Channel, affectez un ID de boucle à chaque SP. Reportez-vous à la documentation du système DAS.
4. Terminez l'installation matérielle comme indiqué dans la documentation qui accompagne vos unités de disque (cordons d'alimentation...). Si vous installez des sous-systèmes DAS, raccordez les processeurs de stockage DAS au noeud à l'aide de câbles RS-232.
5. Pour les boucles FC-AL (en option), mettez les concentrateurs sous tension.
6. Le cas échéant, mettez sous tension les unités de disque, puis mettez sous tension et amorcez les noeuds de grappe.

Lors de l'amorçage (ou à chaque émission de la commande **cfgmgr**), AIX configure toutes les unités connectées au système : un nom logique leur est attribué, sous la forme `hdiskX`, où `X` est un nombre entier identifiant le disque de façon unique.

7. C'est à un stade ultérieur que vous créez les groupes de volumes partagés. Sinon, vous devrez les créer manuellement, comme indiqué au chapitre "Finalisation de l'installation".

8. Vérifiez l'installation :
  - a. Si vous installez des sous-systèmes DAS, passez à la section ci-dessous.
  - b. Si vous installez d'autres unités de disque SCSI, utilisez la commande AIX

```
lsdev -Cc -H
```

pour vérifier que les disques sont à l'état "Available" (disponible).

Pour en savoir plus sur cette procédure, reportez-vous à la section "Installing Shared IBM SCSI-2 Differential Disks" dans le chapitre "Installing Shared Disk Devices" du manuel *HACMP Installation Guide* puis passez au chapitre "Carrying Out Pre-Configuration Tasks".

## Configuration des sous-systèmes DAS

Cette section récapitule les étapes de configuration des sous-systèmes DAS. A ce stade, l'installation matérielle est supposée achevée (câbles SCSI et RS-232, configuration des ID SCSI, etc.).

**Avertissement** : Les indications ne constituent que les grandes lignes de la procédure. Pour des informations détaillées, reportez-vous à la documentation du livrée avec le sous-système DAS (*Configuring and Managing a DAS* et *Installing and Maintaining a DAS*).

La configuration intervient à divers niveaux présentés ci-dessous.

### Configuration tty

Sur chaque noeud, configurez les ports série qui raccordent les processeurs de stockage aux sous-systèmes DAS. Utilisez la commande AIX **smit tty**, comme indiqué dans le chapitre "Installing your software" du manuel *Configuring and Managing a DAS*.

### Configuration des unités de disque physiques (LUN)

A présent, vous devez configurer les unités de disque physiques. Cette étape est obligatoire pour créer des groupes de volumes (que vous utilisiez ou non l'outil Bullcluster de configuration initiale rapide).

Suivez les instructions du chapitre "Setting physical disk units and storage-system caching" du manuel *Configuring and Managing a DAS and Using ATF*. Voici quelques conseils :

#### Accès au gestionnaire du DAS (DAS 2900, 3x00 et 3500)

Pour accéder au gestionnaire du DAS, lancez la commande `smit disk`, puis sélectionnez l'option **DiskArray Storage-System Manager**.

#### Accès au gestionnaire du DAS pour Fibre DAS

Pour accéder au gestionnaire du DAS, sélectionnez l'option **Navisphere**.

#### Installation d'une unité de disque physique (LUN et niveaux RAID)

Vous devez créer des LUN, autrement dit, regrouper des modules disque en unités logiques. Une unité logique est une entité unique constituée d'un ou plusieurs modules disque (en environnement DAS, "LUN" et "unité de disque physique" sont synonymes).

Pour définir des unités logiques, appelez le menu **Bind Physical Units**. Lisez attentivement les recommandations du manuel *Configuring and Managing a DAS and Using ATF*, et :

- **Ne créez pas plus de 8 LUN par sous-système DAS.**
- Les sous-systèmes DAS acceptent les niveaux RAID 0, 0/1, 1, 3 et 5. Il est du ressort de l'administrateur système de déterminer le niveau RAID le mieux adapté. Les niveaux 3 et 5 sont généralement les plus appropriés pour la haute disponibilité.

- Comme expliqué dans *Configuring and Managing a DAS and Using ATF* ("Installing the ATF Software"), si vous utilisez ATF, vérifiez que vous avez désactivé l'option **auto-assign** pour chaque LUN.
- La valeur par défaut de l'option **Maximum rebuild time** est de 4 heures. Cependant, si votre site requiert un temps de réponse rapide et que vous souhaitez limiter la dégradation des activités d'E/S normale, étendez le processus de reconstitution sur une période plus longue, 24 heures par exemple.
- Une fois les unités logiques définies, procédez au rafraîchissement de la base de données ODM AIX, comme indiqué ci-après.

### Groupes de volumes partagés

Ne créez pas de groupes de volumes partagés maintenant :

- Si vous utilisez l'outil *Bullcluster quick initial configuration* pour implémenter un environnement dédié à des applications spécifiques, les groupes de volumes partagés sont créés automatiquement.
- Sinon, il vous faudra les créer manuellement. Pour ce faire, reportez-vous au chapitre "Finalisation de l'installation".

### Rafraîchissement de la base de données ODM AIX

Une fois les unités logiques créées et configurées, il est conseillé de rafraîchir la base de données ODM AIX comme suit :

1. Demandez la liste des disques par la commande :  
`lsdev -C -c disk`
2. Relevez dans la liste, les disques appartenant aux sous-systèmes DAS. Du point de vue AIX, toute unité de disque physique (LUN) créée est présentée comme un disque distinct (`hdisk<n>`).
3. Pour chaque disque DAS de la liste, exécutez la commande :  
`rmdev -d -l hdiskn`  
où *n* est le numéro du disque concerné. Cette commande supprime la définition de l'unité dans la base de données ODM AIX.
4. Exécutez la commande `cfgmgr` pour demander à AIX de reconfigurer les unités.  
`cfgmgr`
5. Dès lors, vous êtes assuré de la cohérence de la base de données ODM.

**Il est conseillé d'exécuter cette procédure chaque fois que vous modifiez les liens des unités de disque DAS, pour garantir la cohérence la base de données.**

## Configuration complémentaire du stockage EMC

Une armoire EMC peut loger un grand nombre de disques et, en cas de panne de disque, le basculement de disques peut prendre du temps. C'est pourquoi, par défaut, le processus de basculement n'est pas lancé automatiquement.

Les menus `smit` permettent de définir le comportement d'une grappe à l'aide du stockage EMC dans une configuration en bus distincts (activation ou non du basculement après échec d'un câble ou d'une carte).

```
smit hacmp
  Bull Cluster Easy Configurator
    EMC Split-bus Behaviour Configuration
```

---

## Étape suivante

L'étape suivante consiste à configurer la topologie de la grappe pour HACMP. Passez au chapitre suivant pour effectuer les tâches de pré-configuration.

---

# Chapitre 7. Configuration de la topologie de la grappe

Ce chapitre explique comment configurer une grappe HACMP en définissant sa topologie. Il comporte les sections suivantes :

- Choix de la méthode de configuration.
- Préalables à l'utilisation de l'outil "BullCluster Quick Initial Configuration", page 7-3.
- Etape préliminaire : mise à jour des fichiers */etc/hosts* et */.rhosts*, page 7-4.
- Définition d'un fichier de configuration, page 7-5.
- Application du fichier de configuration, page 7-11.
- Processus de configuration initiale rapide, page 7-12.
- Annulation de la configuration initiale rapide, page 7-14.
- Vérification d'un fichier de configuration modifié manuellement, page 7-15.
- Utilisation de l'outil d'instantané BullCluster, page 7-15.
- Étape suivante.

---

## Choix de la méthode de configuration

### Les deux méthodes possibles

Il existe deux méthodes possibles pour définir la topologie d'une grappe HACMP : la méthode standard basée sur HACMP et la méthode rapide BullCluster.

**Remarque :** Pour en savoir plus sur le concept de "topologie" dans le cadre HACMP, reportez-vous à la documentation HACMP.

#### Méthode standard HACMP

Cette méthode consiste à faire appel à des menus SMIT d'AIX et de HACMP. Reportez-vous au manuel *HACMP Installation Guide*.

Cette méthode est applicable à toute configuration de grappe (simple ou complexe).

#### Méthode de configuration initiale rapide Bullcluster

Cette méthode consiste à utiliser l'outil *Bullcluster Quick Initial Configuration* intégré au logiciel BullCluster. Cet outil est implémenté comme une extension aux menus SMIT de HACMP (accessible via la commande **smit hacmp**).

Le but de l'outil *Bullcluster Quick Initial Configuration* est de faciliter la configuration des interfaces réseau et de la topologie de grappe HACMP (il évite de naviguer en permanence entre les menus AIX et SMIT HACMP).

L'outil *Bullcluster Quick Initial Configuration* ne convient toutefois pas aux configurations de grappe complexes. Si vous décidez d'utiliser l'outil *Bullcluster Quick Initial Configuration*, déterminez d'abord s'il peut gérer votre configuration de grappe : reportez-vous aux critères du tableau ci-après.

## Critères de choix de la méthode

Reportez-vous au tableau suivant pour savoir si l'outil *Bullcluster Quick Initial Configuration* convient à votre configuration de grappe.

<b>Vous ne pouvez utiliser la méthode de configuration rapide Bullcluster que si la grappe remplit toutes les conditions suivantes :</b>	<b>Vous ne pouvez PAS utiliser la méthode de configuration rapide Bullcluster que si la grappe remplit une des conditions suivantes :</b>
La grappe inclut <b>2, 3 ou 4 noeuds</b> .	La grappe inclut <b>plus de 4 noeuds</b> .
Les noeuds de grappe sont connectés à un <b>réseau public unique (FDDI ou Ethernet)</b> .	Les noeuds de grappe sont connectés à <b>plusieurs réseaux publics</b> Ethernet ou FDDI (configuration à deux réseaux), ou à un ou plusieurs réseaux publics <b>ATM</b> .

### Remarques :

- Si vous prévoyez d'implanter un réseau privé Fibre Channel, prenez contact avec votre représentant.
- Les notions de "réseau public" et "réseau privé" sont expliquées page 3-3.
- N'oubliez pas que vous pouvez toujours utiliser la méthode de configuration HACMP standard, même si la configuration désirée répond aux impératifs de la méthode rapide.

## Étape suivante

- Si l'outil *Bullcluster Quick Initial Configuration* convient à la configuration de grappe prévue, poursuivez avec la section suivante, "Préalables à l'utilisation de l'outil BullCluster Quick Initial Configuration".
- Sinon :
  - Utilisez la méthode de configuration standard HACMP, comme expliqué dans *HACMP Installation Guide*, pour configurer les interfaces de réseau et la topologie de la grappe.
  - Une fois les interfaces de réseau et la topologie de grappe configurées, reportez-vous au chapitre Finalisation de l'installation (page 8-1) avant de créer des groupes de volumes partagés, et de définir des serveurs d'application et des groupes de ressources.
  - Notez que même lorsque la topologie de la grappe est définie, vous pouvez utiliser les menus et les fonctions du logiciel BullCluster. Utilisez-les pour définir les ressources de grappe (groupe de ressources, serveurs d'application...) et les modules LVM partagés (groupes de volumes...), et pour préparer la grappe à exploiter des applications spécifiques.

De plus, si la grappe est équipée de sous-systèmes RAID DAS partagés ou de sous-systèmes de disque SSA, nous vous recommandons d'utiliser ces menus et ces fonctions : ils simplifient l'utilisation des groupes de volumes partagés, et offrent un risque d'erreur moindre que la méthode manuelle.

---

# Préalables à l'utilisation de l'outil "BullCluster Quick Initial Configuration"

## Préalables

Voici quelques rappels avant de commencer.

### Conventions d'appellation des interfaces de carte réseau

L'outil de configuration rapide applique des conventions d'appellation particulières pour les interfaces des cartes réseau. Il utilise toujours les chaînes `_boot` et `_stby` comme suffixes standard des labels des cartes d'amorçage et de secours, respectivement.

Ce point est exposé à la section "Conventions d'appellation", page 4-5.

### Autres contraintes

Avant de mettre en oeuvre l'outil de configuration initiale rapide, vous devez mettre à jour les fichiers `/etc/hosts` et `./rhosts`. Cette opération est explicitée plus bas.

### Terminologie : noeud local et noeud distant

L'outil de configuration rapide doit être utilisé à partir d'un des noeuds de la grappe. Dans ce contexte précis :

- le terme **noeud local** fait référence au noeud sur lequel l'outil sera appliqué pour appliquer un fichier de configuration,
- le terme **noeud distant** fait référence aux autres noeuds.

## Récapitulatif

**Important** : Avant de commencer, nous vous conseillons de passer en revue les différentes étapes de la procédure décrite dans ce chapitre. Accordez une attention particulière aux sections suivantes :

- "Définition d'un fichier de configuration", page 7-5
- "Processus de configuration initiale rapide", page 7-12

Une fois cette tâche accomplie, suivez les instructions ci-dessous.

---

## Étape préliminaire : mise à jour des fichiers `/etc/hosts` et `/.rhosts`

Lors de la validation de la configuration, l'outil de configuration initiale rapide doit être en mesure d'accéder au noeud distant pour exécuter les commandes de configuration sur ce noeud.

Cette condition n'est remplie que si vous mettez à jour les fichiers `/etc/hosts` et `/.rhosts` sur le noeud local et sur le noeud distant, comme indiqué ci-dessous.

### Édition du fichier `/etc/hosts` sur le noeud local

Editez le fichier `/etc/hosts` du noeud local.

- Ajoutez les entrées pour les interfaces d'amorçage des noeuds distants. Si vous n'envisagez pas d'implémenter le basculement d'adresse IP, la notion d'interface d'amorçage n'a pas de sens, dans la mesure où vous ajoutez des entrées pour les interfaces de service des noeuds distants. Notez que le basculement d'adresse IP est traité à "Configuration TCP/IP des interfaces d'amorçage/de service", page 4-4.
- Vérifiez également la présence de l'entrée suivante dans le fichier `/etc/hosts` :

```
127.0.0.1    loopback localhost
```

Par exemple, si vous envisagez d'exécuter l'outil de configuration rapide sur le noeud `foo` et que l'unique noeud distant soit `bar`, vous devez modifier le fichier `/etc/hosts` sur le noeud `foo`. Après avoir ajouté l'entrée qui convient, et si vous envisagez d'implémenter le basculement d'adresse IP, le fichier se présente comme suit :

```
127.0.0.1    loopback localhost
# existing entries:

# added entry for bar's boot interface
192.9.200.11 bar_boot
```

### Édition du fichier `/.rhosts` sur chaque noeud

L'outil de configuration rapide fait appel à des commandes shell distantes. Pour qu'il fonctionne correctement, éditez le fichier `/.rhosts` sur chaque noeud de la grappe. Ajoutez des entrées pour l'interface d'amorçage (ou pour l'interface de service, si vous n'implémentez pas le basculement d'adresse IP) pour tous les noeuds, avec des droits d'accès **racine**.

Par exemple, si vous installez une grappe comprenant les noeuds `foo` et `bar`, et que vous souhaitez implémenter le basculement d'adresse IP, le fichier `/.rhosts` doit comporter les entrées suivantes, sur les deux noeuds :

```
foo_boot root
bar_boot root
```

Avec une grappe identique, mais sans basculement d'adresse IP, les entrées seraient :

```
foo root
bar root
```

Pour des raisons de sécurité, n'insérez ces entrées que lorsqu'elles sont nécessaires (c'est-à-dire avant d'effectuer la configuration rapide ou standard via les menus HACMP) et supprimez-les ensuite (dès que la configuration est terminée).

# Définition d'un fichier de configuration

## Accès à l'outil "Bullcluster Quick Initial Configuration"

L'outil de configuration rapide est une extension des menus SMIT HACMP.

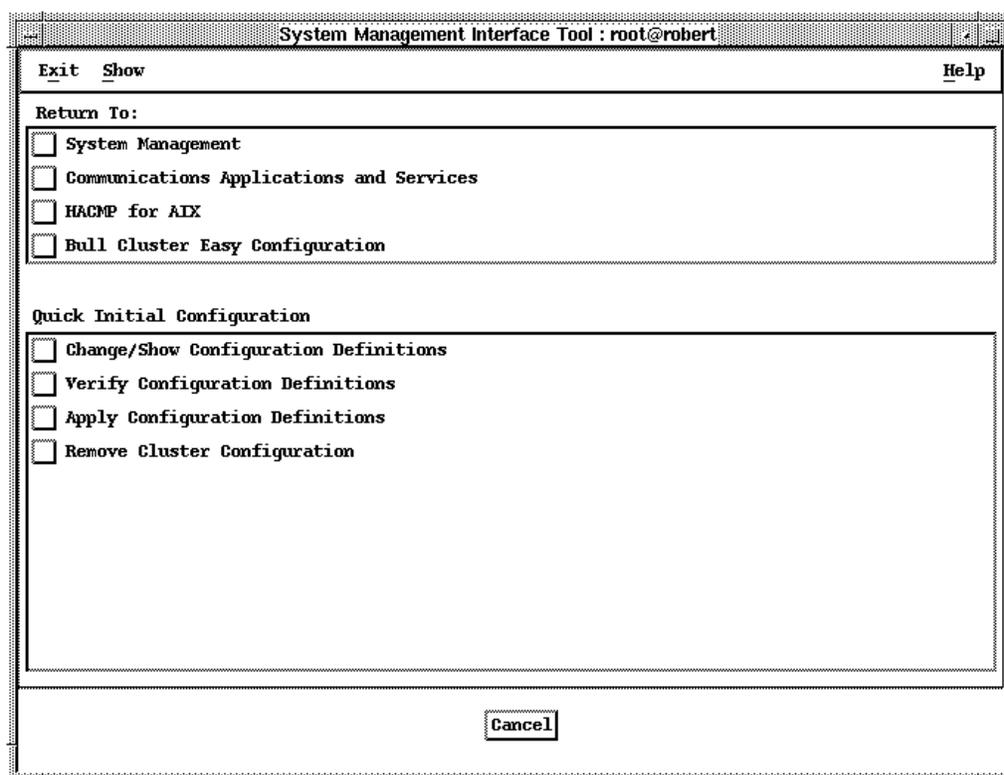
### Remarque à l'attention des utilisateurs d'un PowerConsole

Si vous êtes équipé d'un PowerConsole, vous pouvez tirer parti du module ClusterAssistant, qui affiche une icône pour lancer l'outil Bullcluster Quick Initial Configuration.

### Procédure générale

Si non (si vous n'êtes pas équipé d'un PowerConsole), procédez comme suit :

1. Connectez-vous en tant qu'utilisateur racine et entrez `smit hacmp` pour afficher le menu principal HACMP. Il s'agit du menu HACMP standard, à ceci près qu'il comporte l'option **Bull Cluster Easy Configuration**.
2. Sélectionnez **Bull Cluster Easy Configuration**, puis **Quick Initial Configuration**. Le menu principal de l'outil de configuration initiale rapide s'affiche, comme ci-dessous.



**Remarque :** Vous pouvez également accéder à l'outil de configuration rapide via le raccourci SMIT **quickconf**.

## Procédure de configuration rapide

Cette procédure se décompose en deux tâches principales :

- La première consiste à définir un fichier de configuration via les fonctions du menu **Change/Show Cluster Definitions**. Dans ce menu, vous spécifiez les informations concernant la configuration de base de la grappe, c'est-à-dire essentiellement la configuration réseau (interfaces IP, adresses) et la topologie de la grappe (nombre de noeuds, noms de la grappe et des noeuds).
- La seconde tâche consiste à appliquer le fichier de configuration, via l'option **Apply Configuration Definitions**. Une fois le fichier de configuration appliqué, les cartes réseau sont configurées en conséquence et la topologie de grappe HACMP est définie.

Ces tâches sont décrites ci-après.

## Définition d'un fichier de configuration

Définir un fichier de configuration consiste à renseigner les champs du menu **Change/Show Cluster Definition**. Procédez comme suit :

### Appel du menu Change/Show Cluster Definition :

1. Dans le menu **Quick Initial Configuration**, sélectionnez **Change/Show Configuration Definitions**.
2. Vous êtes invité à préciser le fichier de configuration que vous souhaitez configurer. Dans la mesure où vous n'avez pas encore créé de fichier de configuration, choisissez le fichier par défaut en entrant `template.cfg` (vous pouvez également retrouver son nom en appuyant sur F4 ou en cliquant sur le bouton List).
3. Un nouvel écran s'affiche, vous invitant à indiquer les noms de noeuds. Entrez, dans le champ **Node Names**, les noms des noeuds de la grappe.
  - Les noms ne doivent pas dépasser 31 caractères. Ils sont constitués de caractères alphabétiques et numériques ainsi que des traits de soulignement.
  - Séparez les noms par un espace.
  - Si vous souhaitez afficher ou modifier un fichier de configuration précédemment défini et sauvegardé (c'est-à-dire un fichier de configuration autres que le fichier par défaut **template.cfg**), appuyez sur F4 pour retrouver les noms de fichier définis dans ce fichier de configuration.
4. Une fois entrés les noms de noeuds, le menu **Change/Show Cluster Definition** s'affiche. Il s'agit d'un long menu (semblable à celui illustré page 7-7), composé de plusieurs sections :
  - En tête, se trouve une section dont les champs s'appliquent à l'intégralité de la grappe.
  - En dessous, apparaissent autant de sections qu'il y a de noeuds, chacune spécifique d'un noeud.

Change/Show Cluster Definition	
Input Configuration file	template.cfg
* Output Configuration file	<input type="text"/> List
* Cluster ID	<input type="text"/>
* Cluster Name	<input type="text"/>
PUBLIC NETWORK	
* Network Type	ether List
Cable Type (meaningless for fddi)	tp List ▲ ▼
* Network MASK (dotted decimal)	255.255.255.0
Default GATEWAY Address	<input type="text"/>
* Serial Network Type	rs232
PRIVATE NETWORK	
* Private Network Type	fddi List
Cable Type (meaningless for fddi)	tp List ▲ ▼

*Menu 1/2 : Partie supérieure du menu "Change/Show Cluster Definition", relative aux paramètres globaux de la grappe.*

Menu 2/2 : Partie inférieure du menu "Change/Show Cluster Definition", relative aux paramètres spécifiques de chaque noeud.

### Nom du fichier de configuration où enregistrer les paramètres :

1. Notez que les informations de **Input Configuration file** indiquent, pour mémoire, le fichier de configuration actuel. Comme expliqué plus haut, le fichier par défaut est **template.cfg**.
2. Indiquez, dans le champ **Output Configuration file**, le nom du fichier de configuration que vous créez. Les valeurs des paramètres y seront enregistrées. Choisissez un nom "parlant", et ajoutez-lui le suffixe ".cfg", comme dans l'exemple ci-dessous :

Output Configuration file                      *foobar.cfg*

Il est déconseillé de choisir le nom `template.cfg` : le fichier modèle par défaut serait écrasé.

**Remarque:** Si vous modifiez une configuration précédemment créée, spécifiez, dans le champ **Output Configuration file**, le nom du fichier à modifier. Dans ce cas, les fichiers d'entrée et de sortie sont identiques.

## Paramètres globaux (ID et nom de grappe, caractéristiques du réseau) :

1. Dans le champ **Cluster ID**, entrez un entier positif propre à votre site.
2. Dans le champ **Cluster Name**, entrez une chaîne de texte identifiant la grappe, par exemple `foobar`. Le nom de la grappe peut comporter des caractères alphanumériques et des traits de soulignement. Il ne doit pas dépasser 31 caractères.
3. Définissez les paramètres relatifs au **PUBLIC NETWORK** :

### Network Type

Sélectionnez `ether` (s'il s'agit d'un réseau Ethernet) ou `fdi`, selon le cas.

### Cable Type (sans signification pour fddi)

Si le réseau est de type FDDI, laissez ce champ vide. S'il est de type Ethernet, sélectionnez `tp`, `dix` ou `bnc`, selon le type du réseau Ethernet. Ces valeurs correspondent respectivement à Ethernet à paires torsadées, Ethernet épais et Ethernet fin.

### Network MASK (décimal avec point)

Spécifiez le masque de réseau adéquat, en notation décimale avec point. Par exemple, `255.255.255.0`. Cette valeur dépend du site. Pour en savoir plus sur les masques réseau, reportez-vous à la section "Defining a Network Mask", au chapitre 3, "Planning TCP/IP Networks", du manuel *HACMP Planning Guide*.

### Default GATEWAY Address

Ce champ est facultatif. Il n'est significatif que si le réseau est équipé d'une passerelle. Entrez l'adresse de la passerelle en notation décimale avec point (par exemple, `192.9.201.1`). Pour en savoir plus sur les passerelles, reportez-vous à la documentation AIX.

4. Le champ **Serial Network Type** n'est pas modifiable. Sa valeur prédéfinie `rs232` est indiquée pour mémoire. En fait, l'outil de configuration initiale rapide ne s'applique qu'aux configurations de grappe exploitant RS-232 pour le réseau série.
5. Définissez les paramètres relatifs au **PRIVATE NETWORK**. Renseignez les champs **Network Type** et, si vous exploitez Ethernet sur le réseau privé, **Cable Type**.

Ces trois derniers champs s'appliquent à l'intégralité de la grappe. Les autres sont spécifiques des noeuds.

## Paramètres spécifiques des noeuds :

Les autres champs sont relatifs aux paramètres spécifiques de chaque noeud. Notez qu'il y a autant de sections que de noeuds. Toutes les sections sont identiques. Au début de chaque section, un en-tête tel que `--- FIRST NODE ---` est affiché, qui indique le début de la section, et un champ **Node Name** (non éditable) affiche le nom du noeud concerné.

L'ordre d'affichage des noeuds dans le menu correspond à l'ordre dans lequel vous avez spécifié les noms des noeuds au premier appel du menu **Change/Show Cluster Definition**.

Renseignez les champs pour chaque noeud (les étapes qui suivent doivent être répétées pour chaque "section noeud").

1. Spécifiez les informations relatives à la carte **SERVICE ADAPTER** du noeud, en renseignant les champs comme suit :

### Service Interface

Entrez le nom d'une interface réseau, tel que `en0`, pour la carte de service du noeud.

**Service Address (décimal avec point)**

Entrez l'adresse IP de la carte de service. Utilisez la notation décimale avec point (par exemple, 192.9.201.100).

**Boot Address (décimal avec point)**

Entrez l'adresse d'amorçage IP. Utilisez la notation décimale avec point (par exemple, 192.9.201.120). Si vous ne prévoyez pas d'implémenter le basculement d'adresse IP, la notion d'adresse d'amorçage n'a pas de sens, laissez donc ce champ à blanc. Si vous spécifiez une adresse d'amorçage, faites-le sur le même sous-réseau que l'adresse de service.

**Alternate Hardware Address**

Champ facultatif. Il n'a de sens que si vous prévoyez d'implémenter la commutation d'adresse matérielle. Dans ce cas, entrez une seconde adresse matérielle pour la carte (obligatoirement de 12 chiffres, par exemple 02608c2fb512). Cette adresse doit être unique sur le réseau physique, de façon à ne pas entrer en conflit avec l'adresse physique d'une autre carte du réseau. Pour des conseils sur le choix d'une seconde adresse matérielle, reportez-vous au chapitre "Planning TCP/IP Networks" du manuel *HACMP Planning Guide*.

2. Spécifiez les informations relatives à la carte **STANDBY ADAPTER** du noeud, en renseignant les champs comme suit :

**Standby Interface**

Entrez le nom d'une interface réseau, tel que `en1`, pour la carte de secours du noeud.

**Standby Address (décimal avec point)**

Entrez l'adresse IP de la carte de secours. Utilisez la notation décimale avec point (par exemple, 192.9.202.100). La carte de secours ne doit pas se trouver sur le même sous-réseau que la carte de service. Pour en savoir plus sur les adresses IP des cartes, reportez-vous au manuel *HACMP Planning Guide*.

3. Spécifiez les informations relatives au réseau **SERIAL NETWORK**, en renseignant les champs comme suit. Notez l'absence de ces champs si la grappe que vous configurez comporte plus de deux noeuds. Dans ce cas, vous devez configurer manuellement le réseau série après exécution de la configuration initiale rapide. Notez également que le seul réseau série pris en charge est basé sur les liaisons série RS-232.

**Serial Adapter**

Spécifiez la carte série (par exemple `sa1`) utilisée pour le raccordement au réseau série. Appuyez sur F4 pour afficher la liste des cartes série disponibles.

**PORT Number**

Spécifiez le numéro du port correspondant (par exemple, `s2`). Appuyez sur F4 pour afficher la liste des ports série disponibles.

4. Spécifiez les informations relatives au réseau **PRIVATE NETWORK**, en renseignant les champs comme suit. Laissez ces champs à blanc si vous n'implémentez pas de réseau privé.

**Network Interface**

Entrez le nom d'une interface réseau, tel que `en1`, pour la carte réseau privé du noeud.

**IP Address (décimal avec point)**

Entrez l'adresse IP correspondante. Utilisez la notation décimale avec point (par exemple, 192.9.200.3). Dans la mesure où il s'agit d'un réseau privé, il doit se trouver sur un sous-réseau distinct de tous les autres réseaux à usage du réseau public.

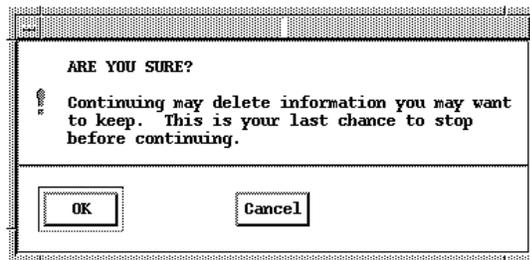
5. N'oubliez pas de répéter ces étapes pour chaque "section noeud" du menu.
6. Une fois tous les champs renseignés, validez l'écran. Un message s'affiche, et les paramètres sont sauvegardés dans le fichier **Output Configuration file** spécifié en tête du menu.
7. Appuyez plusieurs fois sur F3 pour revenir au menu **Quick Initial Configuration**.

---

## Application du fichier de configuration

Une fois défini le fichier de configuration, procédez comme suit pour l'appliquer :

1. Sur le noeud local, appelez le menu **Quick Initial Configuration** et sélectionnez **Apply Configuration Definitions**. Vous êtes invité à renseigner le champ **Configuration file to apply**.
2. Appuyez sur F4 pour afficher la liste des fichiers de configuration définis. Dans la liste, sélectionnez le fichier de configuration à appliquer. Une fois le champ **Configuration to apply** renseigné, appuyez sur Entrée.
3. Vous êtes ensuite invité à préciser si vous souhaitez l'affichage de messages détaillés pendant le processus de configuration rapide. Sélectionnez *Yes* ou *No*.
4. Lorsque vous validez l'écran, un message de confirmation s'affiche, illustré ci-dessous.



5. Appuyez sur Entrée pour lancer le processus de configuration initiale rapide.
6. Attendez la fin de l'opération, puis prenez connaissance des messages affichés pour vous assurer que la configuration a abouti.

Votre grappe est maintenant configurée. Avant de passer au chapitre suivant, qui traite des tâches de post-configuration, vous pouvez poursuivre ce chapitre pour comprendre ce qui se passe réellement lorsqu'un fichier de configuration est appliqué.

**Remarque :** Si, pour une raison quelconque, vous souhaitez annuler la configuration initiale rapide, reportez-vous à "Annulation de la configuration initiale rapide", page 7-14.

---

## Processus de configuration initiale rapide

Cette section présente brièvement les mécanismes appliqués par l'outil de configuration rapide lors du traitement du fichier de configuration via la fonction **Apply Configuration Definitions**.

Ces mécanismes se répartissent en deux catégories : opérations de configuration et opérations de vérification.

### Opérations de configuration

L'outil de configuration rapide définit les cartes réseau et configure la topologie de grappe HACMP.

### Cartes réseau, basculement d'adresse IP, noms des hôtes, fichiers /etc/hosts et /rhosts

L'outil de configuration rapide :

- définit le masque de réseau, le label IP et l'adresse IP des cartes de service et de secours des deux noeuds ;
- implémente le basculement d'adresse IP si vous avez indiqué les adresses d'amorçage dans le fichier de configuration et configure chaque carte de service dont l'adresse IP peut être basculée pour que l'amorçage soit exécuté à partir de l'adresse de la carte d'amorçage et non de celle de la carte de service.

Les adresses IP et masques de réseau sont établis en fonction des valeurs sauvegardées dans le fichier de configuration appliqué par l'outil de configuration rapide.

Concernant les noms hôte :

- Si vous avez choisi de **ne pas** activer le basculement d'adresse IP (c'est-à-dire si vous n'avez pas spécifié d'adresses d'amorçage), le processus de configuration initiale rapide ne change pas le nom hôte (par exemple, le nom initial `foo` reste `foo`).
- Si vous avez choisi d'activer le basculement d'adresse IP, il change le nom hôte. Par exemple, le nom hôte `foo_boot` devient `foo`. Ce changement est généralement bienvenu : il est rare que vous souhaitiez que le nom hôte soit le même que celui de l'interface d'amorçage.

#### Conventions d'appellation

Selon les conventions d'appellation appliquées par l'outil de configuration rapide, les chaînes `_boot` et `_stby` sont toujours utilisées comme suffixes standard des noms des interfaces d'amorçage et de secours, l'interface de service étant dépourvue de suffixe.

Soit par exemple une grappe composée des deux noeuds `foo` et `bar` :

- Le noeud `foo` est configuré avec les trois interfaces :

<code>foo_boot</code>	<i>(amorçage)</i>
<code>foo</code>	<i>(service)</i>
<code>foo_stby</code>	<i>(secours)</i>

- De même, le noeud `bar` est configuré avec les trois interfaces :

<code>bar_boot</code>	<i>(amorçage)</i>
<code>bar</code>	<i>(service)</i>
<code>bar_stby</code>	<i>(secours)</i>

Cela suppose que chaque serveur a été correctement configuré, avec un nom hôte suffixé par `_boot`. Pour en savoir plus sur ces conventions, reportez-vous à "Configuration TCP/IP des interfaces d'amorçage/de service", page 4-4.

## **/etc/hosts et /.rhosts**

L'outil de configuration rapide met à jour les fichiers **/etc/hosts** et **/.rhosts** sur les deux noeuds, en y ajoutant des entrées pour les adresses de service et de secours.

## **Permutation d'adresse matérielle**

Si vous avez renseigné les champs facultatifs **Hardware Address** de l'écran **Change/Show Cluster Definition**, l'outil de configuration rapide implémente la fonction de permutation matérielle. Les adresses matérielles spécifiées (une par noeud) sont utilisées comme adresses secondaires.

## **Réseau série RS-232**

Sur chaque noeud, l'outil de configuration rapide définit et configure l'unité tty connectée à la liaison RS-232 du réseau série. Cet outil se base sur les valeurs des champs **Serial Adapter** et **PORT Number** de l'écran **Change/Show Cluster Definition**.

Si l'unité tty est déjà configurée, l'outil de configuration rapide s'arrête. Vous devez alors supprimer cette unité tty, via la commande **rmdev** ou l'option **Remove tty** du menu SMIT **tty**, avant de relancer l'outil de configuration rapide.

Pour en savoir plus sur les liaisons série RS-232, reportez-vous à la documentation AIX et à la section "Configuring Networks" du manuel *HACMP Installation Guide*.

## **Configuration de la topologie de grappe HACMP**

L'outil de configuration rapide configure la topologie de grappe pour HACMP et crée les objets HACMP requis correspondants : ID de grappe, nom de grappe, noms de noeud, cartes (label, type, attribut de réseau...), etc.

La configuration terminée, l'outil de configuration rapide synchronise la définition de la topologie de la grappe HACMP sur tous les noeuds.

## **Opérations de vérification**

Avant d'entamer la phase de configuration, l'outil de configuration rapide effectue des vérifications en rapport avec les cartes concernées (Ethernet, SCSI et RS-232).

Après la phase de configuration, l'outil de configuration rapide procède à de nouvelles vérifications pour s'assurer que l'environnement de grappe a été correctement configuré.

## Annulation de la configuration initiale rapide

Si, après application de la configuration initiale rapide, vous souhaitez l'annuler, procédez comme suit :

1. A partir de du noeud local, appelez le menu **Quick Initial Configuration** et sélectionnez **Remove Cluster Configuration**.
2. Sur l'écran qui s'affiche, spécifiez le nom du fichier de configuration que vous venez d'appliquer (veillez à ne pas vous tromper !).
3. L'écran suivant s'affiche :

The screenshot shows a dialog box titled "Remove Cluster Configuration". It has three main input sections:

- \* Configuration to remove**: A text box containing "pat.cfg".
- Remove hacmp only or all definitions ?**: A list box containing "hacmp only", with a "List" button and up/down arrow buttons.
- Detailed Output**: A list box containing "no", with a "List" button and up/down arrow buttons.

At the bottom of the dialog, there are several buttons: "OK", "Command", "Reset", "Cancel", "?", and "Help".

4. Renseignez les champs :

### Remove hacmp only or all definitions ?

- Entrez **hacmp only** si vous souhaitez ne supprimer que les définitions HACMP (c'est-à-dire les objets liés à HACMP dans la base de données ODM).
- Entrez **all** si vous souhaitez supprimer les définitions HACMP et les groupes de volumes créés par l'outil de configuration rapide. Supprimer des groupes de volumes est une action destructive, ne spécifiez donc pas **all** à la légère.

### Detailed Output

Sélectionnez `yes` ou `no` selon que vous souhaitez des messages détaillés ou succincts.

5. Appuyez sur Entrée pour appliquer la commande avec vos paramètres.

---

## Vérification d'un fichier de configuration modifié manuellement

Le menu **Quick Initial Configuration** propose l'option **Verify Configuration Definitions**, qui permet de vérifier un fichier de configuration donné.

Cette fonction est superflue si vous avez fait appel aux menus SMIT de configuration initiale rapide (méthode normale). Elle est destinée aux personnels de maintenance sur site qui créent ou modifient un fichier de configuration avec un éditeur de texte.

---

## Utilisation de l'outil d'instantané BullCluster

L'outil d'instantané de BullCluster permet d'enregistrer et de restaurer la topologie d'une grappe ainsi que les classes ODM de la topologie BullCluster.

Menus SMIT permettant d'accéder aux nouvelles options d'instantané :

```
->Communications Applications and Services
    ->HACMP for AIX
        ->Bull Cluster Easy Configuration
            ->Bull Cluster Snapshots
                Add a Bull Cluster Snapshot
                Change/Show a Bull Cluster Snapshot
                Remove a Bull Cluster Snapshot
                Apply a Bull Cluster Snapshot
```

---

## Étape suivante

### Grappe équipée d'un PowerConsole

- Une fois définie la topologie de la grappe HACMP, vous pouvez enregistrer la grappe via ClusterAssistant (utilisez l'icône "Register Cluster" du dossier "Configuration" du groupe d'applications ClusterAssistant). Une fois enregistrée, la grappe peut être gérée à partir du PowerConsole par le biais des différents utilitaires ClusterAssistant.

### Étape suivante

- Passez au chapitre suivant pour finaliser l'installation de la grappe. Si la grappe est équipée d'un PowerConsole, reportez-vous également au manuel *PowerConsole & ClusterAssistant Setup Guide*.



---

## Chapitre 8. Finalisation de l'installation de la grappe

Ce chapitre traite des points suivants :

- Remarque à l'attention des utilisateurs de PowerConsole.
- Configuration des serveurs d'application et ressources HACMP.
- Configuration des modules LVM partagés (groupes de volumes...), page 8-2.
- Synchronisation des horloges de la grappe, page 8-3.
- Implémentation d'alias d'adresses IP, page 8-3.
- Activation de la notification des événements de grappe via la messagerie électronique, page 8-3.
- Groupes de volumes partagés sur sous-systèmes DAS, page 8-5.
- Etape suivante

---

### Remarque à l'attention des utilisateurs de PowerConsole

Si votre grappe est équipée d'un PowerConsole (solution console en option pour PowerCluster), notez que vous pouvez tirer parti du module ClusterAssistant pour effectuer les tâches de configuration décrites dans cette section. ClusterAssistant propose en effet nombre d'icônes qui permettent d'appeler rapidement les différents menus de configuration concernés par ces tâches.

---

### Configuration des serveurs d'application et ressources HACMP

Une fois configurée la topologie de la grappe HACMP, vous devez définir les serveurs d'application HACMP et les groupes de ressources souhaités, via BullCluster et/ou les utilitaires HACMP. Quelle que soit la méthode choisie, reportez-vous à la documentation HACMP pour des informations de fond sur les concepts relatifs aux serveurs d'application et aux groupes de ressources.

#### Utilitaires BullCluster

- Le logiciel BullCluster intègre le menu **Cluster Resources Modification**. Ce menu propose des fonctions simples à utiliser pour configurer groupes de ressources et serveurs d'application spécifiques. Par exemple, vous pouvez créer un groupe de ressources via l'option **Add a Resource Group**, puis créer et ajouter les ressources souhaitées à ce groupe. Pour en savoir plus, reportez-vous au chapitre "Gestion des ressources de grappe et des modules LVM partagés", page 9-1.
- Le module logiciel BullCluster propose également des exemples de scripts de début et de fin, utiles pour implémenter diverses applications de la grappe comme applications hautement disponible. Pour une brève présentation de ces scripts, reportez-vous à "Planification d'applications spécifiques", page 3-1. Pour des informations plus détaillées, reportez-vous au manuel *Powercluster & solutions HA - Utilisation des scripts d'exemple HA*, principalement destiné au personnel technique Bull.

#### Utilitaires HACMP

- Au lieu de passer par le menu BullCluster **Cluster Resources Modification** pour définir les groupes de ressource et les serveurs d'application spécifiques, vous pouvez également utiliser les menus et les utilitaires HACMP standard. Le manuel *HACMP Installation Guide* donne toutes les précisions utiles pour définir les serveurs d'application et les groupes de ressource via les menus HACMP standard.

---

## Configuration des modules LVM partagés (groupes de volumes...)

Si, via l'outil de configuration initiale rapide, vous avez appliqué un fichier de configuration personnalisé pour une application spécifique (X25, Informix, Oracle, OPS, Tuxedo ou HVX), les groupes de volumes partagés et les systèmes de fichiers sont déjà définis et prêts à l'emploi.

Sinon, vous devez créer les groupes de volumes partagés et les systèmes de fichiers souhaités, soit par la méthode rapide soit par la méthode manuelle (voir ci-dessous).

### Méthode rapide

Le logiciel BullCluster intègre le menu **Cluster Resources Modification**. Ce menu propose des fonctions simples à utiliser pour créer et définir des composants LVM partagés (groupes de volume, volumes logiques, systèmes de fichiers). La méthode rapide consiste à utiliser les options ci-après du menu **Cluster Resources Modification** pour créer les modules LVM partagés souhaités et les déclarer membres d'un groupe de ressources donné :

- **Add a Volume Group to a Resource Group**
- **Add a Logical Volume to a Resource Group**
- **Add a File System to a Resource Group.**

Pour en savoir plus, reportez-vous au chapitre "Gestion des ressources de grappe et des modules LVM partagés", page 9-1.

**Remarque :** Si vous exploitez des sous-systèmes DAS, vous devez, avant de créer des groupes de volume, avoir défini les LUN (unités de disque physique) souhaités et rafraîchi la base de données ODM AIX, comme expliqué à "Configuration des unités de disque physiques (LUN)", page 6-7.

### Méthode manuelle

La procédure manuelle de configuration des groupes de volumes diffère selon le type d'unités de disque partagées installées :

- Si vous utilisez des unités de disque SCSI non RAID SCSI ou des sous-système disque SSA :

Aucune remarque particulière ne s'applique : reportez-vous simplement à la documentation AIX et HACMP. Le chapitre "Defining Shared LVM Components" du manuel *HACMP Installation Guide* notamment, contient de précieuses informations.

- Si vous exploitez des sous-systèmes DAS :

Reportez-vous aux remarques de la section "Groupes de volumes partagés sur sous-systèmes DAS", page 8-5.

Pour en savoir plus sur la configuration manuelle des groupes de volumes et des systèmes de fichiers, reportez-vous à la documentation AIX et HACMP.

---

## Synchronisation des horloges de la grappe

Il est conseillé d'implémenter le démon **timed** dans votre grappe, pour maintenir la synchronisation des horloges sur tous les noeuds. Pour en savoir plus, reportez-vous à "Synchronisation des horloges avec timed", page 10-27.

---

## Implémentation d'alias d'adresses IP

Si le basculement d'adresse IP est activé dans votre configuration, vous pouvez implémenter la fonction d'alias d'adresse IP (facultatif). Elle fournit les moyens d'appliquer des alias aux adresses IP de service et d'amorçage. Pour en savoir plus, reportez-vous à la page 10-5.

---

## Activation de la notification des événements de grappe via la messagerie électronique

Vous pouvez personnaliser le traitement des événements de grappe pour qu'un message électronique soit adressé à l'administrateur dès qu'un événement se produit.

Pour implémenter cette fonction, vous pouvez appeler le menu HACMP standard **Change/Show Cluster Events** et spécifier la commande appropriée dans le champ **Notify Command**. Toutefois, ce menu ne s'applique qu'à un événement à la fois. Pour configurer la notification de plusieurs événements en une fois, appelez le menu **Send Mail for Cluster Events Notification** : menu BullCluster spécifique, il est une extension des menus HACMP standard.

### Procédure

Faites appel au menu **Send Mail for Cluster Events Notification** comme suit :

1. Sur le menu principal HACMP SMIT, sélectionnez successivement **Bull Cluster Easy Configuration**, et **Send Mail for Cluster Events Notification**.
2. La liste des événements s'affiche.
3. Sur la liste, sélectionnez un ou plusieurs événements via la touche F7. Appuyez sur Entrée. Un menu semblable au suivant s'affiche :

```
Send Mail for Cluster Events Notification
Type or select values in entry fields.
Press Enter AFTER making all desired changes.

* Node ID
* Event Name
* mailing List

[Entry Fields]
bruni
acquire_service_addr >
[boot]

F1=Help      F2=Refresh   F3=Cancel    F4=List
F5=Reset     F6=Command  F7=Edit     F8=Image
F9=Shell     F10=Exit    Enter=Do
```

4. Notez que le champ (non modifiable) **Event Name** contient les noms des événements que vous venez de sélectionner.
5. Dans le champ **Mailing List**, spécifiez les adresses électroniques auxquelles doivent être envoyées les notifications. Vous pouvez indiquer plusieurs adresses, séparées par des espaces.

```
root cladmin@foosite.com jack@barsite.com
```

6. Appuyez sur Entrée. Lorsque le message de confirmation, appuyez à nouveau sur Entrée.

### Conséquences du recours au menu "Send Mail for Cluster Events Notification"

La procédure ci-dessus modifie la commande de notification associée à chacun des événements que vous avez sélectionnés.

- Aussi, si vous appelez le menu HACMP **Change/Show Cluster Events** (accessible via le menu **Manage Node Environment**), vous pouvez constater que, pour ces événements, la valeur du champ **Notify Command** a changé. Ce champ contient la commande :

```
/usr/sbin/bullcluster/admin/mailtoroot
```

La commande définie via **Notify Command**, intégrée au logiciel BullCluster, est exécutée avant et après la commande de l'événement. Aussi mailtoroot envoie-t-il un message électronique avant et après l'exécution de la commande d'événement.

**Remarque :** Pour en savoir plus sur le traitement des événements de grappe, reportez-vous au chapitre "Maintaining Cluster Events Processing" du manuel *HACMP Administration Guide*.

La procédure ci-dessus modifie également le fichier `/usr/sbin/bullcluster/admin/mailtoroot.cfg`, en y intégrant les adresses électroniques spécifiées dans le champ **Mailing List** de **Send Mail for Cluster Events Notification**. Le script `mailtoroot` est basé sur `mailtoroot.cfg`, c'est-à-dire qu'il envoie des messages électroniques à toutes les adresses spécifiées dans ce fichier.

- Lorsque vous utilisez l'utilitaire **Send Mail for Cluster Events Notification**, si vous modifiez le champ **Mailing List**, la liste `mailtoroot.cfg` est mise à jour en conséquence. Les événements actuellement configurés pour utiliser `mailtoroot` comme commande de notification en sont affectés.

**Remarque :** Le script `mailtoroot` et les fichiers `mailtoroot.cfg` font partie du logiciel BullCluster et résident dans le répertoire `/usr/sbin/bullcluster/admin/utills`. Vous pouvez les consulter.

---

## Groupes de volumes partagés sur sous-systèmes DAS

Cette section traite des tâches à effectuer pour créer et configurer manuellement des groupes de volumes partagés sur des sous-systèmes DAS.

**Avertissement : La méthode manuelle décrite ci-après est plutôt complexe. Nous vous conseillons d'opter pour la méthode rapide, présentée page 8-2. Cette méthode permet en effet de simplifier le maniement des groupes de volumes partagés, avec un risque d'erreur moindre qu'avec la méthode manuelle.**

### Remarques préliminaires

- Nous supposons que vous avez défini les unités de disque physique (LUN) souhaitées et mis à jour la base de données ODM d'AIX, comme expliqué à "Configuration des unités de disque physiques (LUN)", page 6-7.
- Les tâches à effectuer diffèrent selon qu'il s'agit d'une configuration DAS" à bus distincts" ou "deux initiateurs - deux cartes".
- Ces deux types de configuration sont traitées à la section "Planification des sous-systèmes DAS partagés", page 3-10.

## Configuration deux initiateurs - deux serveurs

Pour créer et configurer manuellement un groupe de volumes partagé :

- Suivez simplement la procédure manuelle standard décrite à la section "Maintaining Shared Volume Groups" du manuel *HACMP Administration Guide*. Cette procédure est résumée ci-après.
- Notez que la configuration "deux initiateurs-deux cartes" requiert le logiciel ATF (Application-Transparent Failover). Pour en savoir plus, reportez-vous au SRB (*Bulletin de logiciel*) livré avec le logiciel de grappe.

## Procédure

La procédure explicitée dans la documentation HACMP comprend les étapes décrites ci-après. Pour en savoir plus, reportez-vous au chapitre 4 "Maintaining Shared Volume Groups", du manuel *HACMP Administration Guide*.

1. Le cas échéant, exécutez les tâches prérequis. Les volumes physiques partagés (**hdisks**) doivent être installés, configurés et disponibles (vous disposez de la commande `lsdev -Cc disk` pour vérifier leur état). En outre, HACMP doit être arrêté sur tous les noeuds.
2. Sur un noeud, créez le groupe de volume partagé souhaité via le raccourci **smit mkvg**. N'oubliez pas de donner au champ **Activate volume group AUTOMATICALLY at system restart?** la valeur **No**, et au champ **ACTIVATE volume group after it is created?** la valeur **Yes**.
3. Une fois le groupe de volumes en place, créez les volumes logiques et les systèmes de fichiers souhaités.
4. Démontez le(s) système(s) de fichiers nouvellement créé(s).
5. Mettez hors fonction le groupe de volumes nouvellement créé, via la commande `varyoffvg`.
6. Importez les informations du groupe de volumes sur le second noeud : connectez-vous sur ce second noeud et lancez la commande `importvg`.
7. Configurez le groupe de volumes importé de sorte qu'il reste inactif à l'initialisation : lancez le raccourci **smit mkvg** et donnez au champ **Activate volume group AUTOMATICALLY at system restart?** la valeur **No**.
8. Mettez le groupe de volumes hors fonction via la commande `varyoffvg`.

Notez que tous les groupes de volumes doivent être laissés hors fonction (état "varied off"). Ils seront mis en fonction par HACMP au démarrage des services de grappe.

## Configuration "à bus distincts"

Avec une configuration "à bus distincts", la configuration du groupe de volumes partagé, plus complexe, suppose l'emploi de la commande **trespass** (fournie avec le logiciel DAS).

Vous trouverez ci-après les principales étapes de cette opération. Pour des instructions plus détaillées, reportez-vous :

- Au manuel *Configuring and Managing a DAS* et notamment au chapitre "Making physical disk units available to AIX".
- A la documentation AIX et HACMP ; et notamment au chapitre 4 "Maintaining Shared Volume Groups", du manuel *HACMP Administration Guide*.

## Procédure

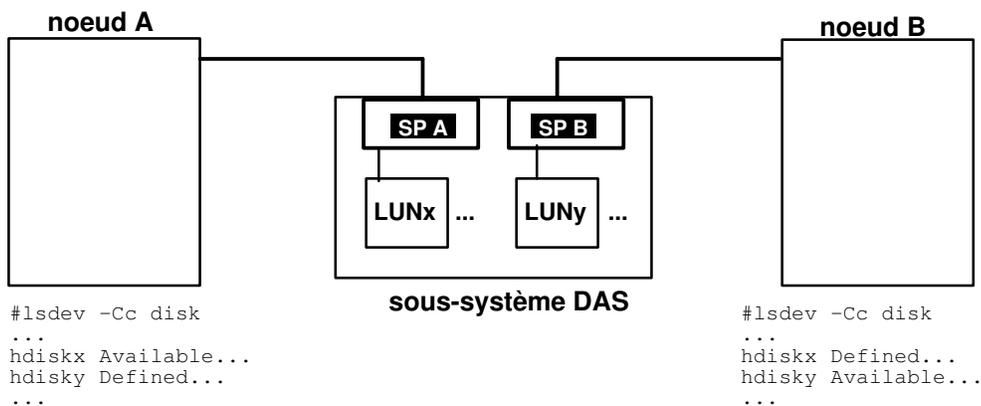
### Prérequis et hypothèses

1. Vérifiez que HACMP est arrêté sur les deux noeuds de la grappe, et fermez toute application en cours sur les disques DAS.
2. Nous supposons que :
  - Vous avez regroupé sur le DAS les modules disque en LUN (avec les niveaux RAID souhaités). Dans l'exemple ci-dessous, nous supposons que deux LUN sont définis, *LUNx* et *LUNy*.
  - Les systèmes sont configurés avec chaque disque doté du même nom sur les deux noeuds. Par exemple, nous supposons que, sur les deux noeuds, *LUNx* est connu sous le nom de *hdiskx*, et *LUNy* sous le nom de *hdisky*.

Notez toutefois qu'en pratique, selon votre configuration, un LUN peut se voir affecter des noms différents par les deux noeuds. Par exemple, un LUN peut être connu sous le nom *hdisk4* par un noeud et sous le nom *hdisk5* sur l'autre.

3. Notez qu'au départ, les disques "vus" comme *available* (disponibles) par un noeud sont "vus" comme seulement *defined* (définis et donc non disponibles) par l'autre. Vous disposez de la commande `lsdev -Cc disk` pour vérifier l'état des **hdisks**.

Dans l'exemple illustré ci-dessous, *hdiskx* est *disponible* du point de vue du noeud A, mais seulement *défini* du point de vue du noeud B (et réciproquement pour *hdisky*).



### Configuration des groupes de volumes et des systèmes de fichiers sur le noeud A

4. Depuis le noeud A, créez un groupe de volume partagé comprenant un disque disponible à partir du noeud A : *hdiskx* dans notre exemple. Utilisez le raccourci **smit mkvg**. N'oubliez pas de donner au champ **Activate volume group AUTOMATICALLY at system restart?** la valeur **No**, et au champ **ACTIVATE volume group after it is created?** la valeur **Yes**.
5. Une fois le groupe de volumes en place, créez les volumes logiques et les systèmes de fichiers souhaités.

6. Démontez le(s) système(s) de fichiers nouvellement créés, mettez hors fonction le groupe de volume nouvellement créé, via la commande `varyoffvg`.

### Configuration des groupes de volumes et des systèmes de fichiers sur le noeud B

7. Depuis le noeud B, créez un groupe de volume partagé comprenant un disque disponible à partir du noeud B : *hdisky* dans notre exemple. Utilisez le raccourci **smit mkvg**. N'oubliez pas de donner au champ **Activate volume group AUTOMATICALLY at system restart?** la valeur **No**, et au champ **ACTIVATE volume group after it is created?** la valeur **Yes**.
8. Une fois le groupe de volumes en place, créez les volumes logiques et les systèmes de fichiers souhaités.
9. Démontez le(s) système(s) de fichiers nouvellement créés, mettez hors fonction le groupe de volume nouvellement créé, via la commande `varyoffvg`.

### Transfert de la propriété du LUN du processeur de stockage (SP) B au SP A

10. Lancez la commande **trespass** pour transférer au *SPA* la propriété du *LUNy* (la commande `trespass` est décrite dans la documentation DAS). Sur le noeud A, entrez :

```
/usr/sbin/cluster/events/utils/trespass -d -c hdisk<n> -l <m>
```

Cette commande indique au SP actuellement propriétaire du `hdisk<n>` de prendre le contrôle du LUN `<m>`. Dans l'exemple, la commande adéquate est :

```
/usr/sbin/cluster/events/utils/trespass -d -c hdiskx -l y
```

(Cette commande indique au SP A de prendre le contrôle du *LUNy*)

Au lieu de la commande **trespass**, vous pouvez utiliser l'option **Change SP Ownership of a Physical Unit** du menu DAS **GridMgr**.

### Mise à disposition du disque sur le noeud A

11. Rendez le disque *hdisky* disponible depuis le noeud A. Sur le noeud A, entrez :

```
mkdev -l hdisky
```

Aucune application ne doit tenter d'accéder au DAS pendant l'exécution de la commande `mkdev`.

Depuis le noeud A, *hdisky* est maintenant vu comme *available* (vous disposez de la commande `lsdev -Cc disk` pour afficher la liste des états `hdisk`).

### Configuration des groupes de volumes pour le disque obtenu

12. Sur le noeud A, lancez la commande `importvg` pour importer le groupe de volumes qui contient le disque obtenu (*hdisky*).
13. Configurez le groupe de volumes importé de sorte qu'il reste inactif à l'initialisation : lancez le raccourci **smit chvg** et donnez au champ **Activate volume group AUTOMATICALLY at system restart?** la valeur **No**.
14. Mettez le groupe de volumes hors fonction via la commande `varyoffvg`.

### Déconfiguration du disque

15. Sur le noeud A, déconfigurez *hdisky* via la commande :

```
rmdev -l hdisky
```

Pour le noeud A, *hdisky* est maintenant vu comme *defined* (et non plus *available*).

## Restitution de la propriété du LUN au SPB

16. Lancez la commande **trespass sur le noeud B** pour transférer au *SP B* la propriété de *LUNy*. Sur le noeud B, entrez :

```
/usr/sbin/cluster/events/utils/trespass -d -c hdisky -l y
```

Avec cette commande, SP B reprend le contrôle de *LUNy*. Du point de vue de la propriété du LUN, la configuration revient à son état initial (voir la figure au début de la procédure).

**Remarque :** Les étapes qui suivent sont identiques aux étapes 10 à 16, à ceci près qu'elles sont exécutées à partir du noeud B (et non du noeud A) et concernent *hdiskx* (au lieu de *hdisky*).

### Etapes à effectuer sur le noeud B

17. Lancez la commande **trespass** pour transférer au *SP B* la propriété du *LUNx* (la commande **trespass** est décrite dans la documentation DAS). Sur le noeud B, entrez :

```
/usr/sbin/cluster/events/utils/trespass -d -c hdisky -l x
```

(Cette commande indique au SP B de prendre le contrôle du *LUNx*).

18. Rendez le disque *hdiskx* disponible depuis le noeud B. Sur le noeud B, entrez :

```
mkdev -l hdiskx
```

Pour le noeud B, *hdiskx* est maintenant vu comme *available* (vous disposez de la commande `lsdev -Cc disk` pour afficher la liste des états *hdisk*).

19. Sur le noeud B, lancez la commande `importvg` pour importer le groupe de volumes qui contient le disque obtenu (*hdiskx*).

20. Configurez le groupe de volumes importé de sorte qu'il reste inactif à l'initialisation : lancez le raccourci **smit chvg** et donnez au champ **Activate volume group AUTOMATICALLY at system restart?** la valeur **No**.

21. Mettez le groupe de volumes hors fonction via la commande `varyoffvg`.

22. Sur le noeud B, déconfigurez *hdiskx* via la commande :

```
rmdev -l hdiskx
```

Pour le noeud B, *hdiskx* est maintenant vu comme *defined* seulement (et non plus *available*).

### Dernière étape sur le noeud A

23. Lancez la commande **trespass sur le noeud A** pour transférer au *SPA* la propriété du *LUNx*. Dans l'exemple, la commande à lancer sur le noeud A est :

```
/usr/sbin/cluster/events/utils/trespass -d -c hdiskx -l x
```

La procédure est terminée. Tous les groupes de volumes sont laissés hors fonction (état "varied off"). Ils seront mis en fonction (état "varied on") par HACMP au démarrage des services de grappe.

---

## Étape suivante

Les tâches décrites dans ce chapitre mettent un terme à la procédure de configuration de la grappe.

- Vous disposez d'une grappe opérationnelle. Si vous souhaitez davantage de précisions sur la configuration de HACMP et le lancement des services de grappe, reportez-vous à la documentation HACMP.
- Vous pouvez consulter le chapitre "Outils et techniques de gestion du système", page 10-1, pour des précisions sur la gestion de grappe.

---

# Chapitre 9. Gestion des ressources de grappe et des modules LVM partagés

Ce chapitre explique comment gérer les ressources de grappe et les modules LVM partagés, par le biais du menu **Cluster Resources Modification** du logiciel BullCluster. Il comporte les sections suivantes :

- Présentation du menu
- Ajout, suppression et affichage d'un groupe de ressources, page 9-4.
- Ajout et suppression d'un groupe de volumes, page 9-8.
- Ajout et suppression d'un volume logique, page 9-11.
- Ajout et suppression d'un système de fichiers, page 9-14.
- Ajout et suppression d'un label IP de service, page 9-17.
- Ajout et suppression d'un serveur d'application, page 9-19.
- Diffusion générale de la définition d'un groupe de volumes, page 9-20.

---

## Présentation du menu

La gestion des ressources de grappe et des modules LVM partagés met en jeu des procédures complexes, qui impliquent l'utilisation de nombreux menus et commandes HACMP et AIX. Ces procédures sont décrites dans *HACMP Administration Guide*.

Pour simplifier ces tâches d'administration, le logiciel BullCluster propose le menu **Cluster Resources Modification**. Ce menu offre des fonctions simplifiées pour gérer ressources de grappe et composants LVM partagés (groupes de volumes, volumes logiques, systèmes de fichiers). Pour chaque tâche, il existe un menu spécifique, unique, qui vous évite de naviguer continuellement entre les différents menus HACMP et AIX.

Utiliser le menu **Cluster Resources Modification** permet non seulement de gagner du temps, mais est également plus sûr, dans la mesure où toutes les opérations de configuration/déconfiguration requises sont effectuées et systématiquement vérifiées.

## Accès au menu Cluster Resources Modification

Pour accéder au menu **Cluster Resources Modification** :

- Sur le menu principal SMIT HACMP (appelé par **smit hacmp**), sélectionnez **Bull Cluster Easy Configuration** puis **Cluster Resources Modification**. Un écran semblable à celui illustré ci-après s'affiche.

## Remarque à l'attention des utilisateurs de PowerConsole

Si votre grappe est équipée d'un PowerConsole, vous pouvez tirer parti du module ClusterAssistant. De fait, le dossier "Configuration" du groupe d'applications ClusterAssistant comprend un dossier "Resources", dont les icônes permettent d'accéder facilement aux différentes fonctions du menu **Cluster Resources Modification**.

# Options

Voici une rapide présentation des options de **Cluster Resources Modification** :

**Add a Resource Group** (voir page 9-4)

Crée un groupe de ressources.

**Add a Volume Group to a Resource Group** (voir page 9-8)

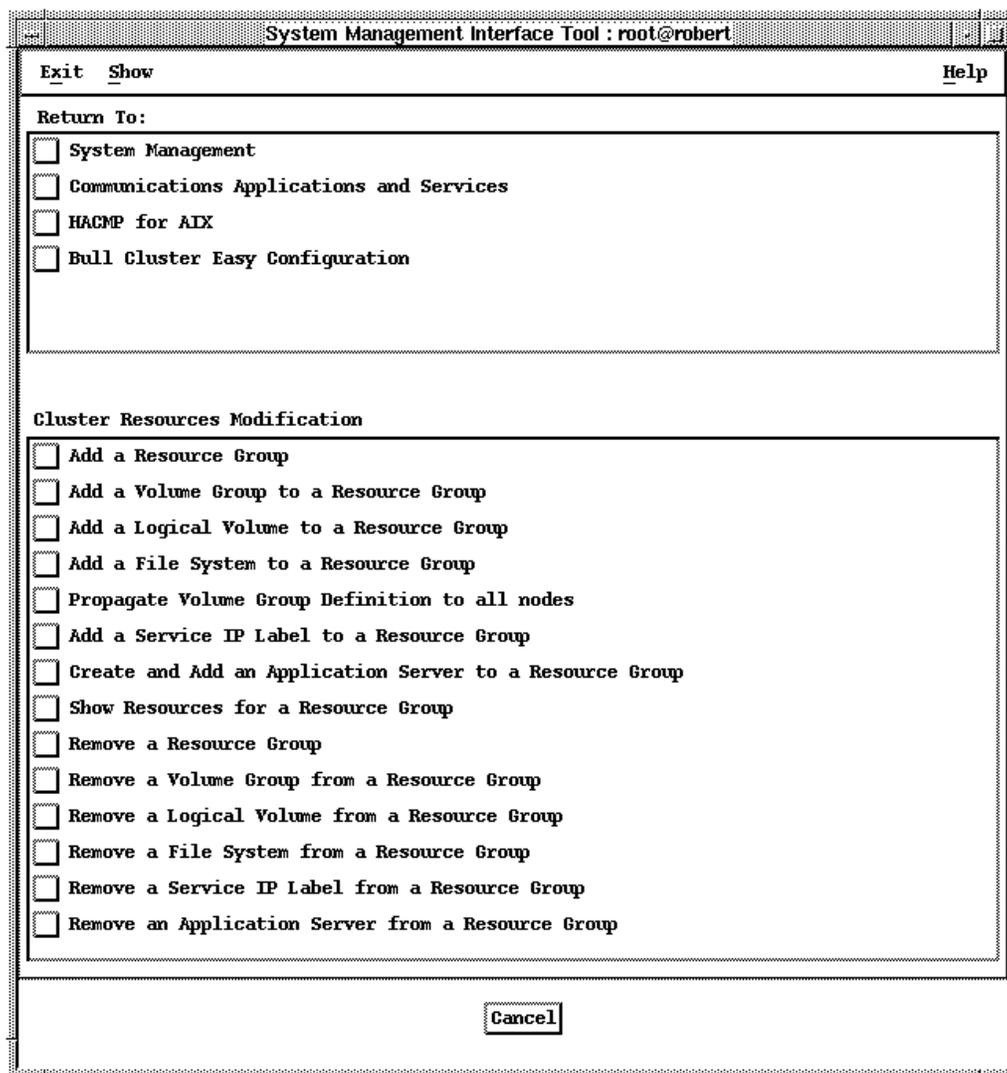
Crée un groupe de volumes partagé (ou utilise un groupe de volumes partagé existant), et le déclare comme ressource d'un groupe de ressources donné. La configuration du groupe (activation/désactivation, export, import) est automatique.

**Add a Logical Volume to a Resource Group** (voir page 9-11)

Crée un volume logique et déclare le groupe de volumes auquel il appartient comme ressource d'un groupe donné. Les paramètres de ce volume logique sont configurables (fonction miroir, taille...). Quant à la configuration du groupe de volumes correspondant (activation/désactivation, export, import), elle est automatique.

**Add a File System to a Resource Group** (voir page 9-14)

Crée un système de fichiers et un volume logique ; les caractéristiques du système de fichiers sont configurables (mise en miroir, dimensions...), et les opérations requises pour configurer le groupe de volume correspondant (vary on/off, export, import, etc.) sont effectuées automatiquement.



**Propagate Volume Group Definition to all nodes** (voir page 9-20)

Diffuse les informations concernant le groupe de volumes sur tous les noeuds de grappe.

**Add a Service IP Label to a Resource Group** (voir page 9-17)

Déclare un label IP de service comme ressource d'un groupe de ressources donné (ce qui active le basculement d'adresse IP).

**Create and Add an Application Server to a Resource Group** (voir page 9-19)

Configure les serveurs d'application et les ressources dédiés à des applications spécifiques.

**Show Resources for a Resource Group** (voir page 9-7)

Affiche les ressources actuellement associées à un groupe de ressources.

**Remove a Resource Group** (voir page 9-6)

Supprime la définition d'un groupe de ressources donné et synchronise les informations sur tous les noeuds de grappe.

**Remove a Volume Group from a Resource Group** (voir page 9-10)

Supprime physiquement un groupe de volumes et met à jour les définitions du groupe de ressources en conséquence.

**Remove a Logical Volume from a Resource Group** (voir page 9-13)

Supprime physiquement un volume logique et met à jour les définitions du groupe de ressources en conséquence.

**Remove a File System from a Resource Group** (voir page 9-16)

Supprime physiquement un système de fichiers et met à jour les définitions du groupe de ressources en conséquence.

**Remove a Service IP Label from a Resource Group** (voir page 9-18)

Supprime un label IP de service de la liste des ressources associées à un groupe de ressources (ce qui désactive le basculement d'adresse IP).

**Remove an Application Server from a Resource Group** (voir page 9-19)

Supprime un serveur d'application de la liste des ressources associées à un groupe de ressources donné.

## Ajout, suppression et affichage d'un groupe de ressources

Cette section traite des trois options suivantes, appelées via le menu **Cluster Resources Modification** :

- **Ajout d'un groupe de ressources**
- **Suppression d'un groupe de ressources**
- **Affichage des ressources d'un groupe de ressources**

**Remarque** : Les fonctions de ces menus sont semblables à celles proposées par le menu HACMP **Define Resource Groups**. En outre, elles exécutent automatiquement la synchronisation des informations sur les noeuds de la grappe.

Pour en savoir plus sur la configuration d'un groupe de ressources, reportez-vous à la section "Configuring Resources" dans le chapitre "Configuring Cluster Resources" du manuel *HACMP Installation Guide*.

### Ajout d'un groupe de ressources

Le menu **Add a Resource Group** permet de :

- créer des groupes de ressources (avec une configuration conforme aux paramètres définis).
- synchroniser les informations sur les noeuds de grappe.

### Procédure

1. Pour utiliser le menu **Add a Resource Group**, les services de grappe doivent être arrêtés. Au besoin, arrêtez-les par le biais du menu HACMP **Stop Cluster Services**.
2. Sur le menu **Bull Cluster Easy Configuration**, sélectionnez **Cluster Resources Modification**, puis **Add a Resource Group**. L'écran suivant s'affiche.

The screenshot shows a dialog box titled "Add a Resource Group". It contains the following fields and controls:

- \* Resource Group Name**: A text input field.
- Node Relationship**: A dropdown menu with "cascading" selected, a "List" button, and up/down arrow icons.
- \* Participating Node Names**: A text input field and a "List" button.
- Inactive Takeover Activated (meaningless for concurrent or rotating)**: A dropdown menu with "false" selected, a "List" button, and up/down arrow icons.

At the bottom of the dialog, there are buttons for "OK", "Command", "Reset", "Cancel", "?", and "Help".

3. Renseignez les champs :

**Resource Group Name**

Entrez le nom souhaité. N'utilisez pas plus de 31 caractères. Les caractères alphanumériques sont autorisés, de même que le trait de soulignement.

**Node Relationship**

Vous avez le choix entre **cascading**, **concurrent** et **rotating**.

**Participating Node Names**

Indiquez le nom des noeuds qui doivent être membres de la chaîne de ressources pour ce groupe de ressources. Entrez ces noms par ordre de priorité décroissante (de gauche à droite). Laissez un espace entre chaque nom.

**Inactive Takeover Activated (meaningless for concurrent or rotating)**

Définissez ce paramètre pour contrôler l'acquisition initiale d'une ressource par un noeud lorsque la relation noeud/ressource est **cascading**. Ce paramètre n'est pas valable pour les groupes de ressources **concurrent** ou **rotating**.

Si la valeur de ce champ est **true**, alors le premier noeud de la chaîne de ressources qui rejoint la grappe acquiert la ressource. Ensuite, la ressource passe, en mode cascade, aux noeuds de la chaîne dont la priorité est la plus élevée au fur et à mesure qu'ils rejoignent la grappe.

Si la valeur de ce champ est **false**, seul le noeud qui a la priorité la plus élevée pour cette ressource acquiert initialement la ressource. Les noeuds à priorité moindre dans la chaîne de ressources n'acquièrent la ressource que lors d'un basculement.

4. Appuyez sur Entrée pour valider et créer le groupe de ressources.
5. Si nécessaire, relancez les services de grappe via le menu HACMP **Start Cluster Services**.

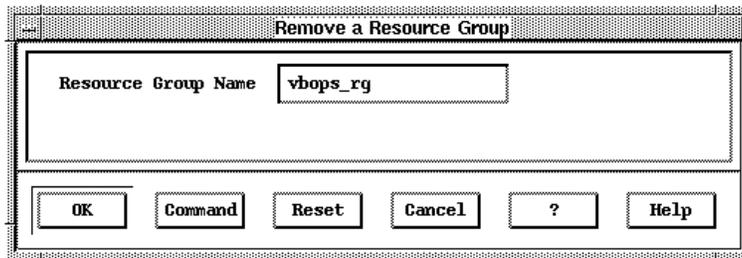
## Suppression d'un groupe de ressources

L'option **Remove a Resource Group** :

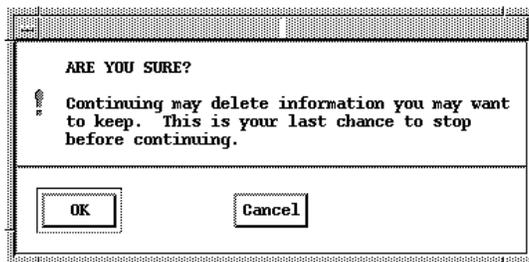
- Supprime un groupe de ressources donné de la configuration de grappe. La définition du groupe de ressources est supprimée de la base de données HACMP ODM, sans impact physique sur les ressources (groupes de volumes, systèmes de fichiers...) associées au groupe de ressources.
- Synchronise les informations sur les noeuds de grappe.

### Procédure

1. Pour utiliser le menu **Remove a Resource Group**, les services de grappe doivent être arrêtés. Au besoin, arrêtez-les par le biais du menu HACMP **Stop Cluster Services**.
2. Sur le menu **Bull Cluster Easy Configuration**, sélectionnez **Cluster Resources Modification**, puis **Remove a Resource Group**.
3. Dans la liste qui s'affiche, sélectionnez le groupe de ressources à supprimer. L'écran **Remove a Resource Group** s'affiche.



4. Le champ **Resource Group Name** n'est pas modifiable. Vérifiez qu'il affiche le groupe de ressource à supprimer.
5. Lorsque vous validez l'écran, un message de confirmation s'affiche, illustré ci-dessous.



6. Appuyez sur Entrée pour supprimer le groupe de ressources.
7. Si nécessaire, relancez les services de grappe via le menu HACMP **Start Cluster Services**.

## Affichage des ressources d'un groupe de ressources

L'option **Show Resources for a Resource Group** répertorie les ressources actuellement associées à un groupe de ressources.

Show Resources for a Resource Group	
Resource Group Name	VB_SalesDB
Node Relationship	cascading
Participating Node Names	java tango
Service IP Label	
Filesystems	
Filesystems to Export	
Filesystems to NFS Mount	
Volume Groups	
Concurrent Volume Groups	
Raw Disk PVIDs	
Application Servers	vbops
Miscellaneous Data	
Inactive Takeover Activated	false
9333 Disk Fencing Activated	false

OK   Command   Reset   Cancel   ?   Help

Les champs de cet écran ne sont pas modifiables. Pour plus de détails sur leur signification, reportez-vous à "Managing ressources" au chapitre "Changing Run-Time Parameters and Ressources Groups" dans *HACMP Administration Guide*.

## Ajout et suppression d'un groupe de volumes

Cette section traite des deux options suivantes, appelées via le menu **Cluster Resources Modification** :

- **Add a Volume Group to a Resource Group**
- **Remove a Volume Group from a Resource Group**

### Ajout d'un groupe de volumes à un groupe de ressources

Le menu **Add a Volume Group to a Resource Group** :

- Crée un groupe de volumes, si celui indiqué n'existe pas. Pour tous les noeuds, le numéro principal du groupe de volumes partagé reste le même. La taille de la partition physique est choisie de façon appropriée.
- Dans tous les cas (que le groupe de volumes partagé soit nouveau ou non) :
  - Le groupe de volumes partagé est désactivé. L'activation de groupes de volumes partagés incombe en effet à HACMP (au moment du lancement des services de grappe).
  - Le groupe de volumes partagé est déclaré à tous les noeuds. Cette fonction est intéressante dans la mesure où la procédure manuelle est assez complexe. (Cette procédure implique les commandes **varyon**, **varyoff**, **exportvg**, **importvg** et **trespass**.)
- Le groupe de volumes est déclaré comme ressource du groupe de ressources spécifié.
- La configuration de grappe est synchronisée sur les noeuds de grappe.

### Procédure

1. Pour utiliser le menu **Add a Volume Group to a Resource Group**, les services de grappe doivent être arrêtés. Au besoin, arrêtez-les par le biais du menu HACMP **Stop Cluster Services**.
2. Sur le menu **Bull Cluster Easy Configuration**, sélectionnez **Cluster Resources Modification**, puis **Add a Volume Group to a Resource Group**.
3. Dans la liste qui s'affiche, sélectionnez le groupe de ressources auquel ajouter un groupe de volumes. L'écran **Add a Volume Group to a Resource Group** s'affiche.

The screenshot shows a dialog box titled "Add a Volume Group to a Resource Group". It contains the following fields and controls:

- Resource Group Name:** VB\_SalesDB
- Node Relationship:** cascading
- Participating Node Names:** java tango
- \* Volume Group Name:** (empty field)
- PHYSICAL VOLUME names:** (empty field) with a **List** button.
- Propagate Volume Group Definition on all nodes:** NO with a **List** button and up/down arrow icons.
- Buttons:** OK, Command, Reset, Cancel, ?, Help.

4. Notez que les champs **Resource Group Name**, **Node Relationship** et **Participating Node Names** ne sont pas modifiables. Ils sont juste indiqués pour rappeler les caractéristiques du groupe de ressources auquel vous ajoutez un groupe de volumes.

5. Renseignez les champs :

**Volume Group Name**

Entrez le nom d'un groupe de volume existant ou du groupe de volume à créer.

**PHYSICAL VOLUME names**

Si vous créez un groupe de volumes, spécifiez le noms des disques partagés dédiés à ce groupe de volume. Vous pouvez en indiquer autant que souhaité. Par exemple, si vous indiquez "hdisk4", le nouveau groupe de volumes ne comportera qu'un seul disque partagé, hdisk4. Si vous spécifiez "hdisk4 hdisk5", il comportera deux disques partagés, hdisk4 et hdisk5.

Ne renseignez pas ce champ si vous définissez un groupe de volumes qui existe déjà.

**Remarques concernant la mise en miroir AIX**

La mise en miroir AIX ne concerne que les disques partagés ordinaires (non RAID). Notez que :

- La mise en miroir n'est possible que si le groupe de volume comporte plusieurs disques. Si vous envisagez la mise en miroir de volumes logiques, vous devez donc spécifier au moins deux disques dans le champ **PHYSICAL VOLUME names**.
- Si vous créez un groupe de volume de deux disques pour mettre en oeuvre la fonction miroir, vous devez dans la mesure du possible, spécifier deux disques installés sur des bus SCSI distincts. Dans ce cas, un disque sur un bus SCSI est doublé par un disque miroir sur l'autre bus SCSI - ce qui assure une plus grande disponibilité des données. Au contraire, lorsque les deux disques se trouvent sur un même bus SCSI, la mise en miroir est déclenchée sur ce bus et, en cas de défaillance, la disponibilité des données n'est plus optimale.
- Si vous créez un groupe de volume de plus de deux disques, la mise en miroir d'un bus SCSI à l'autre n'est pas garantie. Cette restriction vaut même lorsque les disques désignés ne sont pas tous sur le même bus SCSI.
- De même, si la grappe est équipées de sous-systèmes disque SSA avec deux boucles SSA, la mise en miroir doit être implémentée entre les deux boucles (les disques d'une boucle doivent être doublés par des disques de l'autre boucle).
- Pour en savoir plus sur le volumes logiques et le mise en miroir, reportez-vous au chapitre "Volumes logiques" du manuel *AIX System Management Guide*.

**Propagate Volume Group Definition on all nodes**

Entrez YES pour que les modifications de configuration soient immédiatement synchronisées sur tous les noeuds.

Notez toutefois que, dans la mesure où la synchronisation prend du temps, vous pouvez indiquer NO si vous avez plusieurs composants LVM partagés à définir (groupes de volumes, volumes logiques ou systèmes de fichiers). Dans ce cas, une fois que vous avez défini tous les composants LVM, n'oubliez pas de synchroniser la configuration de la grappe via le menu **Propagate Volume Group Definition to all nodes** (voir page 9-20). Dans le doute, entrez YES.

6. Appuyez sur Entrée pour valider et prenez connaissance des messages qui s'affichent.

7. A la fin de la commande, vous pouvez définir d'autres modules LVM partagés par le biais des menus d'ajout de groupe de volumes, de volume logique ou de système de fichiers. Puis, si besoin est, n'oubliez pas de synchroniser la configuration de la grappe via le menu **Propagate Volume Group Definition to all nodes** (voir page 9-20).
8. Si nécessaire, relancez les services de grappe via le menu HACMP **Start Cluster Services**.

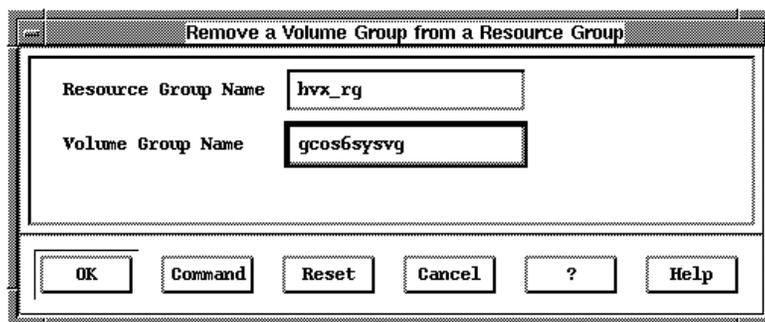
## Suppression d'un groupe de volumes d'un groupe de ressources

Le menu **Remove a Volume Group from a Resource Group** :

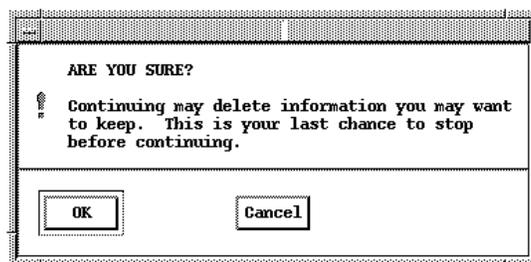
- Détruit physiquement le groupe de volumes partagé spécifié.
- Supprime de chaque noeud de grappe la définition du groupe de volumes détruit.
- Met à jour et synchronise les définitions des ressources de grappe sur tous les noeuds de grappe.

### Procédure

1. Pour utiliser le menu **Remove a Volume Group from a Resource Group**, les services de grappe doivent être arrêtés. Au besoin, arrêtez-les par le biais du menu HACMP **Stop Cluster Services**.
2. Sur le menu **Bull Cluster Easy Configuration**, sélectionnez **Cluster Resources Modification**, puis **Remove a Volume Group from a Resource Group**.
3. Dans la liste qui s'affiche, sélectionnez le groupe de ressources auquel appartient le groupe de volumes à détruire.
4. Si plusieurs groupes de volumes sont déclarés membres de ce groupe de ressources, une liste s'affiche. Sélectionnez-y le groupe de volumes à détruire. L'écran **Remove a Volume Group from a Resource Group** s'affiche.



5. Les champs de cet écran ne sont pas modifiables. Assurez-vous que **Volume Group Name** inclut le groupe de volumes à détruire.
6. Lorsque vous validez l'écran, un message de confirmation s'affiche, illustré ci-dessous.



7. Appuyez sur Entrée sur détruire le groupe de volumes.
8. Si nécessaire, relancez les services de grappe via le menu HACMP **Start Cluster Services**.

---

## Ajout et suppression d'un volume logique

Cette section traite des deux options suivantes, appelées via le menu **Cluster Resources Modification** :

- **Add a Logical Volume to a Resource Group**
- **Remove a Logical Volume from a Resource Group.**

## Ajout d'un groupe logique à un groupe de ressources

### Remarques préliminaires

Le menu **Add a Logical Volume to a Resource Group** est utile si vous implémentez une configuration de grappe à accès concurrent : ce mode utilise des volumes logiques bruts pour stocker les données sur disques partagés. Le système JFS (Journaled File System) d'AIX n'étant pas pris en charge, vous ne pouvez pas utiliser de systèmes de fichiers.

**Remarque** : Pour un complément d'information, reportez-vous à "Maintaining Shared LVM Components in a Concurrent Access Environment", dans le manuel *HACMP Administration Guide*.

Si vous implémentez une configuration non-concurrente et que vous utilisez des systèmes de fichiers (cas habituel), le menu **Add a Logical Volume to a Resource Group**. Faites plutôt appel au menu **Add a File System to a Resource Group** (voir page 9-14).

### Actions

Le menu **Add a Logical Volume to a Resource Group** :

- Crée un volume logique (dans un groupe de volumes partagé existant) avec les caractéristiques indiquées dans le menu.
- Déclare le volume logique à tous les noeuds. Cette fonction est intéressante dans la mesure où la procédure manuelle est assez complexe. (Cette procédure implique les commandes **varyon**, **varyoff**, **exportvg**, **importvg** et **trespass**.)
- Déclare le groupe de volumes dans lequel est créé le nouveau volume logique comme ressource du groupe de ressources spécifié.
- Met à jour et synchronise les définitions de ressource de grappe sur tous les noeuds de grappe.

### Procédure

1. Pour utiliser le menu **Add a Logical Volume to a Resource Group**, les services de grappe doivent être arrêtés. Au besoin, arrêtez-les par le biais du menu **HACMP Stop Cluster Services**.
2. Sur le menu **Bull Cluster Easy Configuration**, sélectionnez **Cluster Resources Modification**, puis **Add a Logical Volume to a Resource Group**.
3. Dans la liste qui s'affiche, sélectionnez le groupe de ressources auquel appartient le groupe de volumes dans lequel créer un volume logique. L'écran **Add a Logical Volume to a Resource Group** s'affiche.
4. Notez que les champs **Resource Group Name**, **Node Relationship**, **Participating Node Names**, **Volume Group Name** et **Volume Group Disk(s)** ne sont pas modifiables. Ils sont juste indiqués pour rappeler les caractéristiques du groupe de ressources et du groupe de volumes impliqués.

5. Renseignez les champs :

#### **Logical Volume Name**

Entrez un nom unique (15 caractères au maximum) pour le volume logique à créer. Ne renseignez pas ce champ si vous souhaitez conserver le nom généré par défaut à la création du volume logique. (Notez que, dans une grappe, le nom d'un volume logique partagé doit être unique. De plus, il doit être identique sur tous les noeuds le partageant. Des vérifications sont automatiquement effectuées lors de l'application de ce menu.)

#### **Logical Volume Disks**

Lorsqu'un volume logique est créé, la règle d'affectation par défaut est d'utiliser le moins de disques possibles par copie de volume logique et de placer les partitions physiques d'une même copie de façon aussi contiguë que possible.

- Pour accepter cette règle d'affectation par défaut, laissez ce champ vide.

- Si vous préférez créer un volume logique dont les partitions physiques sont réparties sur plusieurs disques, spécifiez les disques souhaités.

Laissez un espace entre les noms des disques  
(exemple : "hdisk5 hdisk6").

#### **AIX Mirroring**

Indiquez `yes` ou `no` selon que vous souhaitez ou non implémenter la mise en miroir.

- Si les disques partagés sont ordinaires (non RAID) : choisissez `yes` pour activer la mise en miroir et accroître ainsi la disponibilité des données. Opter pour cette valeur signifie que vous souhaitez maintenir deux exemplaires du journal JFS partagé et des volumes logiques.

- Si les disques partagés sont des disques RAID : spécifiez `no`. En effet, les unités de disque RAID intègrent leur propre système de redondance de données : la mise en miroir est inutile.

Sur la mise en miroir, qui ne concerne que les disques partagés ordinaires (non RAID), reportez-vous également à Remarques concernant la mise en miroir AIX, page 9-9

#### **Logical Volume Size (in PPs)**

Indiquez le nombre de partitions physiques à affecter au volume logique.

#### **Propagate Volume Group Definition on all nodes**

Entrez `YES` pour que les modifications de configuration soient immédiatement synchronisées sur tous les noeuds.

Notez toutefois que, dans la mesure où la synchronisation prend du temps, vous pouvez indiquer `NO` si vous avez plusieurs composants LVM partagés à définir (groupes de volumes, volumes logiques ou systèmes de fichiers).

Dans ce cas, une fois que vous avez défini tous les composants LVM, n'oubliez pas de synchroniser la configuration de la grappe via le menu **Propagate Volume Group Definition to all nodes** (voir page 9-20). Dans le doute, entrez `YES`.

6. Appuyez sur Entrée pour valider et prenez connaissance des messages qui s'affichent.
7. A la fin de la commande, vous pouvez définir d'autres modules LVM partagés par le biais des menus d'ajout de groupe de volumes, de volume logique ou de système de fichiers. Puis, si besoin est, n'oubliez pas de synchroniser la configuration de la grappe via le menu **Propagate Volume Group Definition to all nodes** (voir page 9-20).
8. Si nécessaire, relancez les services de grappe via le menu HACMP **Start Cluster Services**.

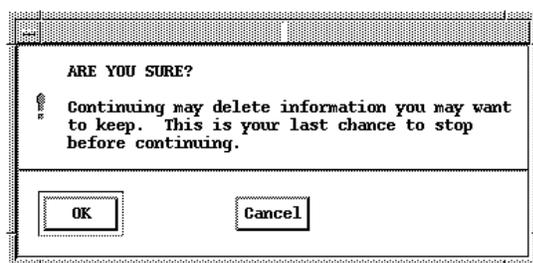
## Suppression d'un groupe logique d'un groupe de ressources

Le menu **Remove a Logical Volume from a Resource Group** :

- Détruit physiquement le volume logique partagé spécifié.
- Supprime de chaque noeud de grappe la définition du volume logique détruit.
- Met à jour et synchronise les définitions des ressources de grappe sur tous les noeuds de grappe.

### Procédure

1. Pour utiliser le menu **Remove a Volume Group from a Resource Group**, les services de grappe doivent être arrêtés. Au besoin, arrêtez-les par le biais du menu HACMP **Stop Cluster Services**.
2. Sur le menu **Bull Cluster Easy Configuration**, sélectionnez **Cluster Resources Modification**, puis **Remove a Logical Volume from a Resource Group**.
3. Dans la liste qui s'affiche, sélectionnez le groupe de ressources auquel appartient le groupe de volumes qui inclut le volume logique à détruire.
4. La liste des volumes logiques s'affiche. Sélectionnez le volume à détruire. L'écran **Remove a Logical Volume Group from a Resource Group** s'affiche.
5. Les champs de cet écran ne sont pas modifiables. Assurez-vous que le champ **Logical Volume Name** inclut le volume logique à détruire.
6. Lorsque vous validez l'écran, un message de confirmation s'affiche, illustré ci-dessous.



7. Appuyez sur Entrée pour détruire le volume logique.
8. Si nécessaire, relancez les services de grappe via le menu HACMP **Start Cluster Services**.

---

## Ajout et suppression d'un système de fichiers

Cette section traite des deux options suivantes, appelées via le menu **Cluster Resources Modification** :

- **Add a File System to a Resource Group**
- **Remove a File System from a Resource Group**

### Ajout d'un système de fichiers à un groupe de ressources

Le menu **Add a File System to a Resource Group** :

- Crée un système de fichiers et un volume logique (dans un groupe de volumes partagé existant), avec les caractéristiques indiquées dans le menu.
- Déclare le système de fichiers à tous les noeuds. Cette fonction est intéressante dans la mesure où la procédure manuelle est assez complexe. (Cette procédure implique les commandes **varyon**, **varyoff**, **exportvg**, **importvg** et **trespass**, reportez-vous à "Groupes de volumes partagés sur sous-systèmes DAS", page 8-5).
- Déclare le système de fichiers partagé comme ressource du groupe de ressources spécifié.
- Met à jour et synchronise les définitions de ressources de grappe sur les noeuds de grappe.

### Procédure

1. Pour utiliser le menu **Add a File System to a Resource Group**, les services de grappe doivent être arrêtés. Au besoin, arrêtez-les par le biais du menu HACMP **Stop Cluster Services**.
2. Sur le menu **Bull Cluster Easy Configuration**, sélectionnez **Cluster Resources Modification**, puis **Add a File System to a Resource Group**.
3. Dans la liste qui s'affiche, sélectionnez le groupe de ressources auquel appartient le groupe de volumes dans lequel vous souhaitez créer un nouveau système de fichiers.
4. Un nouvel écran s'affiche. Sélectionnez le groupe de volumes dans lequel créer le système de fichiers. L'écran **Add a File System to a Resource Group** s'affiche.
5. Notez que les champs **Resource Group Name**, **Node Relationship**, **Participating Node Names**, **Volume Group Name** et **Volume Group Disk(s)** ne sont pas modifiables. Ils sont juste indiqués pour rappeler les caractéristiques du groupe de ressources et du groupe de volumes impliqués.
6. Renseignez les champs :

#### **Logical Volume Name**

Entrez un nom unique (15 caractères au maximum) pour le volume logique à créer pour le nouveau système de fichiers. Ne renseignez pas ce champ si vous souhaitez conserver le nom généré par défaut à la création du volume logique.

### Logical Volume Disks

Lorsqu'un volume logique est créé, la règle d'affectation par défaut est d'utiliser le moins de disques possibles par copie de volume logique et de placer les partitions physiques d'une même copie de façon aussi contiguë que possible.

– Pour accepter cette règle d'affectation par défaut, laissez ce champ vide.

– Si vous préférez créer un volume logique dont les partitions physiques sont réparties sur plusieurs disques, spécifiez les disques souhaités.

Laissez un espace entre les noms des disques (exemple :

"hdisk5 hdisk6").

### AIX Mirroring

Indiquez `yes` ou `no` selon que vous souhaitez ou non implémenter la mise en miroir.

– Si les disques partagés sont ordinaires (non RAID) : choisissez `yes` pour activer la mise en miroir et accroître ainsi la disponibilité des données. Opter pour cette valeur signifie que vous souhaitez maintenir deux exemplaires du journal JFS partagé et des volumes logiques.

– Si les disques partagés sont des disques RAID : spécifiez `no`. En effet, les unités de disque RAID intègrent leur propre système de redondance de données : la mise en miroir est inutile.

Sur la mise en miroir, qui ne concerne que les disques partagés ordinaires (non RAID), reportez-vous également à "Remarques concernant la mise en miroir AIX", page 9-9

### Log Size (in PPs)

Indiquez la taille du journal JFS, en nombre de partitions physiques (PP).

### Logical Volume Size (in PPs)

Indiquez le nombre de partitions physiques à affecter au volume logique.

### File System Mount Point

Indiquez le point de montage souhaité (c'est-à-dire le répertoire dans lequel le système de fichiers est disponible).

### Propagate Volume Group Definition on all nodes

Entrez `YES` pour que les modifications de configuration soient immédiatement synchronisées sur tous les noeuds.

Notez toutefois que, dans la mesure où la synchronisation prend du temps, vous pouvez indiquer `NO` si vous avez plusieurs composants LVM partagés à définir (groupes de volumes, volumes logiques ou systèmes de fichiers).

Dans ce cas, une fois que vous avez défini tous les composants LVM, n'oubliez pas de synchroniser la configuration de la grappe via le menu

**Propagate Volume Group Definition to all nodes** (voir page 9-20). Dans le doute, entrez `YES`.

7. Appuyez sur Entrée pour valider et prenez connaissance des messages qui s'affichent.
8. A la fin de la commande, vous pouvez définir d'autres modules LVM partagés par le biais des menus d'ajout de groupe de volumes, de volume logique ou de système de fichiers. Puis, si besoin est, n'oubliez pas de synchroniser la configuration de la grappe via le menu **Propagate Volume Group Definition to all nodes** (voir page 9-20).
9. Si nécessaire, relancez les services de grappe via le menu **HACMP Start Cluster Services**.

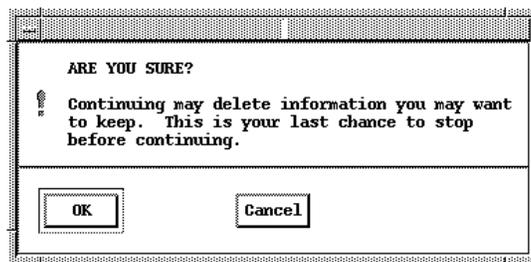
## Suppression d'un système de fichiers d'un groupe de ressources

Le menu **Remove a File System from a Resource Group** :

- Détruit physiquement le système de fichiers spécifié.
- Supprime de chaque noeud de grappe la définition du système de fichiers détruit.
- Met à jour et synchronise les définitions des ressources de grappe sur tous les noeuds de grappe.

### Procédure

1. Pour utiliser le menu **Remove a File System from a Resource Group**, les services de grappe doivent être arrêtés. Au besoin, arrêtez-les par le biais du menu HACMP **Stop Cluster Services**.
2. Sur le menu **Bull Cluster Easy Configuration**, sélectionnez **Cluster Resources Modification**, puis **Remove a File System from a Resource Group**.
3. Dans la liste qui s'affiche, sélectionnez un groupe de ressources pour lequel le système de fichiers à détruire est déclaré comme ressource.
4. La liste des systèmes de fichiers s'affiche. Sélectionnez le système de fichiers à détruire. L'écran **Remove a File System from a Resource Group** s'affiche.
5. Les champs de cet écran ne sont pas modifiables. Vérifiez que le champ **File System Name** comprend le système de fichiers à détruire.
6. Lorsque vous validez l'écran, un message de confirmation s'affiche, illustré ci-dessous.



7. Appuyez sur Entrée pour détruire le système de fichiers.
8. Si nécessaire, relancez les services de grappe via le menu HACMP **Start Cluster Services**.

## Ajout et suppression d'un label de service IP

Cette section traite des deux options suivantes, appelées via le menu **Cluster Resources Modification** :

- **Add a Service IP Label to a Resource Group**
- **Remove a Service IP Label from a Resource Group**

**Remarque** : Pour en savoir plus sur les concepts liés aux labels IP et au basculement d'adresse IP, reportez-vous à la documentation HACMP.

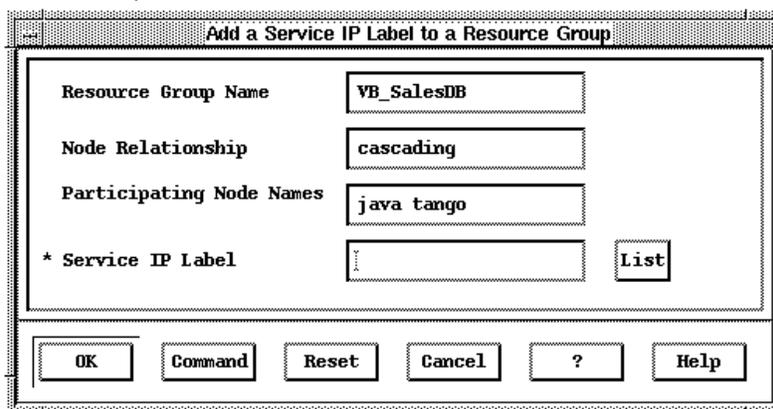
### Ajout d'un label de service IP à un groupe de ressources

Le menu **Add a Service IP Label to a Resource Group** permet de déclarer un label IP de service comme ressource d'un groupe de ressources existant. Autrement dit, il permet d'implémenter le basculement d'adresse IP. Le menu **Add a Service IP Label to a Resource Group** :

- Déclare le label IP de service spécifié comme ressource du groupe de ressources sélectionné.
- Met à jour et synchronise les définitions de ressources de grappe sur les noeuds de grappe.

### Procédure

1. Pour utiliser le menu **Add a Service IP Label to a Resource Group**, les services de grappe doivent être arrêtés. Au besoin, arrêtez-les par le biais du menu HACMP **Stop Cluster Services**.
2. Sur le menu **Bull Cluster Easy Configuration**, sélectionnez **Cluster Resources Modification**, puis **Add a Service IP Label to a Resource Group**.
3. Dans la liste qui s'affiche, sélectionnez le groupe de ressources auquel ajouter un label IP de service. L'écran **Add a Service IP Label to a Resource Group** s'affiche, comme illustré ci-après :



4. Notez que les champs **Resource Group Name**, **Node Relationship** et **Participating Node Names** ne sont pas modifiables. Ils sont juste indiqués pour rappeler les caractéristiques du groupe de ressources sélectionné.
5. Indiquez le **Service IP Label** à basculer lorsque le groupe de ressources bascule. Appuyez sur F4 pour visualiser la liste des labels IP valides.
6. Appuyez sur Entrée pour ajouter le label IP de service sélectionné au groupe de ressources.
7. Si nécessaire, relancez les services de grappe via le menu HACMP **Start Cluster Services**.

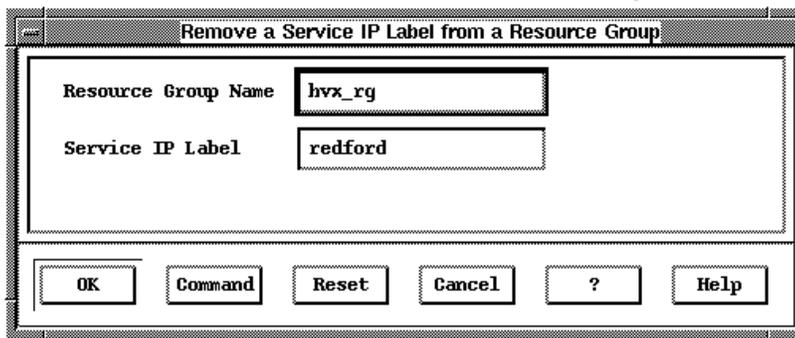
## Suppression d'un label IP de service d'un groupe de ressources

Le menu **Remove a Service IP Label to a Resource Group** :

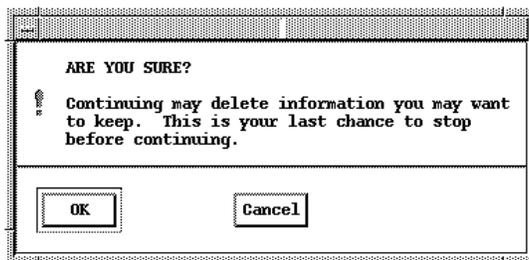
- Supprime un label IP de service de la liste des ressources associées au groupe de ressources spécifié (autrement dit, il désactive le basculement d'adresse IP).
- Met à jour et synchronise les définitions de ressources de grappe sur les noeuds de grappe.

### Procédure

1. Pour utiliser le menu **Remove a Service IP Label from a Resource Group**, les services de grappe doivent être arrêtés. Au besoin, arrêtez-les par le biais du menu **HACMP Stop Cluster Services**.
2. Sur le menu **Bull Cluster Easy Configuration**, sélectionnez **Cluster Resources Modification**, puis **Remove a Service IP Label from a Resource Group**.
3. Sur les deux listes qui s'affichent successivement, choisissez le groupe de ressources duquel supprimer le label IP de service associé, puis le label concerné. Le menu **Remove a Service IP Label from a Resource Group** s'affiche :



4. Les champs de cet écran ne sont pas modifiables. Assurez-vous que le champ **Service IP Label** comprend le label IP de service à supprimer du groupe de ressources ad hoc.
5. Lorsque vous validez l'écran, un message de confirmation s'affiche, illustré ci-dessous.



6. Appuyez sur Entrée pour supprimer le label IP de service du groupe de ressources.
7. Si nécessaire, relancez les services de grappe via le menu **HACMP Start Cluster Services**.

---

## Ajout et suppression d'un serveur d'application

Cette section traite des deux options suivantes, appelées via le menu **Cluster Resources Modification** :

- **Create and Add an Application Server to a Resource Group**
- **Remove an Application server from a Resource Group**

### Création et ajout d'un serveur d'application à un groupe de ressources

Le logiciel BullCluster propose des exemples de scripts de début et de fin pour vous aider à implémenter des applications spécifiques dans la grappe. Le menu **Create and Add an Application Server to a Resource Group** permet de configurer les serveurs d'application dédiés à ces applications. Le menu crée un serveur d'application pour HACMP, et le configure avec les scripts de début et de fin correspondant à l'application choisie.

Les exemples de scripts de début et de fin comme le menu **Create and Add an Application Server to a Resource Group** sont principalement destinés au personnel technique Bull.

- Pour une présentation générale des exemples de scripts, reportez-vous à "Planification d'applications spécifiques", page 3-1.
- Pour des informations plus détaillées, reportez-vous au manuel *Powercluster & solutions HA - Utilisation des scripts d'exemple HA*.

### Suppression d'un serveur d'application d'un groupe de ressources

Pour supprimer un système de fichiers créé via **Add an Application Server to a Resource Group**, passez par **Remove an Application Server from a Resource Group**.

Cette procédure garantit que toutes les opérations de déconfiguration requises sont effectuées correctement.

**Remarque** : Si vous avez créé des serveurs d'application X25 ou TUXEDO via le menu **Create and Add an Application Server to a Resource Group**, il est impératif d'utiliser l'option **Remove an Application Server from a Resource Group** pour garantir que les objets spécifiques définis dans la base de données ODM sont déconfigurés.

## Diffusion générale de la définition d'un groupe de volumes

Le menu **Cluster Resources Modification** comprend plusieurs options vous permettant de créer des modules LVM partagés :

- **Add a Volume Group to a groupe de ressources**
- **Add a Logical Volume to a Resource Group**
- **Add a File System to a Resource Group.**

Exécuter l'option **Propagate Volume Group Definition on all nodes** du menu **Cluster Resources Modification** n'est nécessaire que si :

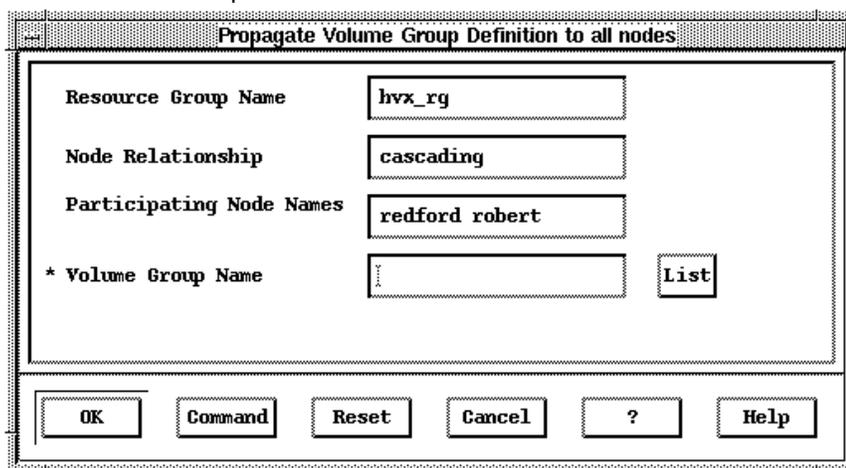
vous avez utilisé l'un des trois menus précités

ET

vous avez, dans ces menus, spécifié NO dans le champ **Propagate Volume Group Definition to all nodes**.

Dans ce cas, vous devez déclarer le(s) groupe(s) de volumes aux différents noeuds de grappe. Pour ce faire :

1. Sur le menu **Bull Cluster Easy Configuration**, sélectionnez **Cluster Resources Modification**, puis **Propagate Volume Group Definition on all nodes**.
2. La liste des groupes de ressources s'affiche. Sélectionnez un groupe de ressources auquel vous avez ajouté un module LVM par le biais de l'un des trois menus mentionnés plus haut. Le menu **Propagate Volume Group Definition on all nodes** s'affiche, comme illustré ci-après.



3. Notez que les champs **Resource Group Name**, **Node Relationship** et **Participating Node Names** ne sont pas modifiables. Ils sont juste indiqués pour rappeler les caractéristiques du groupe de ressources sélectionné.
4. Dans le champ **Volume Group Name**, indiquez le groupe de volumes à déclarer à tous les noeuds de grappe.
5. Appuyez sur Entrée pour lancer l'opération.

---

# Chapitre 10. Outils et techniques de gestion du système

Les deux premières sections traitent de deux utilitaires BullCluster liés à la gestion de grappe :

- Outil de diagnostic BullCluster
- Utilitaire clusterview, page 10-3.

Les deux sections suivantes traitent de deux utilitaires BullCluster liés à la configuration de grappe :

- Implémentation d'alias d'adresses IP, page 10-5.
- Implémentation du transfert d'applications (Rolling Applications), page 10-7.
- Implémentation de HA native, page 10-17.
- Implémentation du stockage partagé HA étendue, page 10-24.
- Personnalisation du traitement des événements de grappe, page 10-27.

**Remarque :** BullCluster intègre d'autres utilitaires précieux pour la configuration et la gestion d'une grappe.

Les autres sections (à partir de la page 10-27) donnent des conseils relatifs à des outils utiles pour la gestion d'une grappe, non intégrés au logiciel BullCluster :

- Synchronisation des horloges avec timed, page 10-27.
- Clés de licence iFOR/LS Nodelocked, page 10-28.
- Conseils relatifs à DSMIT, page 10-28.
- Gestion des sous-systèmes DAS avec ArrayGUIDe, page 10-29.
- Gestion des sous-systèmes DAS avec Navisphere, page 10-29.

---

## Outil de diagnostic BullCluster (bclerrdiag)

Le logiciel BullCluster comporte un outil de diagnostic qui permet de générer des informations de diagnostic. Il est conseillé à l'administrateur de grappe d'utiliser régulièrement cet outil (une fois par jour, par exemple) pour vérifier la "santé" de la grappe.

L'outil de diagnostic BullCluster est fourni sous forme de commande UNIX, **bclerrdiag**. Voici la page man concernant cette commande.

### Remarque à l'attention des utilisateurs de PowerConsole

Si votre grappe est équipée d'un PowerConsole, vous pouvez tirer parti du module ClusterAssistant, qui intègre un utilitaire basé sur les fenêtres (icône "Cluster Diagnostics") qui simplifie l'utilisation de l'outil de diagnostic BullCluster.

## Objet

L'outil de diagnostic **bclerrdiag** BullCluster rend compte des opérations de grappe.

## Description

**bclerrdiag** analyse l'état et la cohérence des ressources de grappe. En outre (et, le cas échéant, conformément aux indicateurs spécifiés), il effectue des diagnostics spécifiques du sous-système en balayant les fichiers journaux sur les noeuds de grappe concernés.

### Analyse de la cohérence des ressources de grappe

**bclerrdiag** analyse l'état et la cohérence des ressources de. Il vérifie la présence des définitions des groupes de ressources HACMP ainsi que l'état des ressources impliquées dans ces groupes. Après analyse, il signale tout éventuel problème ou incohérence. Par exemple, si aucun des noeuds de grappe ne prend en charge une ressource qui est supposée hautement disponible (par exemple, un groupe de volume ou une adresse IP), cette anomalie est signalée.

### Diagnostics spécifiques du sous-système

Selon les indicateurs spécifiés, l'outil **bclerrdiag** peut générer des diagnostics relatifs aux trois sous-système : AIX, HACMP et DAS (Disk Array Subsystems) :

- **Sous-système AIX :**  
L'outil recherche des erreurs matérielles consignées dans le journal d'erreur AIX (**errlog**) de chaque noeud concerné. Toute erreur matérielle consignée est signalée.
- **Sous-système HACMP :**  
L'outil explore les messages d'erreur consignés par HACMP dans le fichier **/var/adm/cluster.log** et signale les messages qui contiennent une des chaînes `error`, `fail` ou `network_down`.
- **Sous-système DAS :**  
L'outil interroge le journal d'événement (**DASlog**) du DAS (Disk Array Subsystems) et signale les messages significatifs, c'est-à-dire les messages identifiés par les codes série `0x900` et `0xa00`.

### Consolidation des informations de diagnostic

**bclerrdiag** collecte les informations extraites des fichiers journaux concernés des différents noeuds. Il consolide ensuite les données collectées, en tenant compte des horodates et en récapitulant les informations dans le cas d'erreurs répétitives.

## Syntaxe

```
/usr/sbin/bullcluster/monitoring/diagtool/bclerrdiag  [ -b sous-systèmes ]  
                                                    [ -n noeuds ]  
                                                    [ -p période ]
```

La commande **bclerrdiag** peut être exécutée sur n'importe quel noeud, tant que ce noeud appartient à la grappe dont le fonctionnement doit être vérifié. **bclerrdiag** rapporte les informations de diagnostic, vers sa sortie standard et vers des fichiers journaux spécifiques (voir "Fichiers", ci-dessous).

## Indicateurs

### [ -b *sous-systèmes* ]

Liste les sous-systèmes auxquels appliquer les diagnostics. Le paramètre *sous-système* est une liste de noms de sous-systèmes, séparés par des virgules. Les noms admis sont `AIX`, `HACMP` et `DAS`. Par défaut (indicateur **-b** non spécifié), les diagnostics sont exécutés sur les trois sous-systèmes. Notez en outre que, indicateur **-b** spécifié ou non, **bclerrdiag** analyse toujours l'état et la cohérence des ressources de la grappe.

- [ -n *noeuds* ] Liste les noeuds à prendre en compte pour exécuter les diagnostics. Le paramètre *noeuds* est une liste de noms de noeuds, séparés par des virgules. Par défaut (indicateur **-n** non spécifié), **bclerrdiag** prend en compte tous les noeuds de la grappe.

[ **-p période** ] **bclerrdiag** examine les journaux du sous-système à la recherche des messages générés dans la période spécifiée par le paramètre *période*. La période par défaut est de 1 jour (les messages d'erreur vieux de plus d'un jour sont ignorés). La *période* doit être spécifiée comme un nombre entier d'heures.

## Fichiers

`/usr/sbin/bullcluster/monitoring/diagtool/bclerrdiag` est le fichier exécutable.

Le répertoire `/var/bullcluster/logs` reçoit les divers fichiers journaux générés par **bclerrdiag** (ces journaux sont écrasés à chaque nouvelle exécution de **bclerrdiag**). Le fichier journal **cluster\_global\_report** contient les informations qui étaient affichés à la dernière exécution de la commande. **cluster\_detailed\_report** contient des informations plus détaillées, intéressantes notamment pour les dépannages. Les deux fichiers sont stockés sur le noeud sur lequel a été exécutée la commande **bclerrdiag**. Notez que **bclerrdiag** stocke les autres fichiers journaux intermédiaires sur les différents noeuds concernés.

---

## Utilitaire clusterview

Le logiciel BullCluster intègre **clusterview**, un utilitaire qui rend compte de l'état des grappes. Cet utilitaire est fourni sous forme de commande UNIX. Il est comparable au programme **clstat** livré avec HACMP.

Voici la page man relative à **clusterview**.

## Objet

L'utilitaire clusterview a pour fonction de reporter l'état des grappes HACMP.

## Syntaxe

`/usr/sbin/cluster/clusterview`

## Description

L'utilitaire **clusterview** est un programme client Clinfo qui utilise l'API Clinfo et les informations du fichier `/usr/sbin/cluster/etc/clhosts` pour afficher des informations sur les noeuds et les interfaces de toutes les grappes détectées en mode d'affichage ASCII. Le programme **clinfo** doit être actif pour que la commande clusterview puisse être exécutée.

**clusterview** indique si la grappe est active, inactive ou instable. Il précise également si un noeud est actif, inactif, partant ou arrivant, ainsi que le nombre de noeuds de la grappe.

Pour chaque noeud, **clusterview** affiche le label et l'adresse IP de toutes les interfaces réseau qui lui sont associées, ainsi que leur état (actives ou inactives).

Pour exécuter l'utilitaire, entrez `/usr/sbin/cluster/clusterview`. Vous pouvez également appeler **clusterview** via SMIT, par le raccourci **smit clusterview**, ou en sélectionnant l'option **Show Cluster State** du menu **Bull Cluster Easy Configuration**.

Un exemple de sortie est donné ci-après.

## Exemple de sortie

**clusterview** affiche des informations semblables à :

```
clinfo daemon is running.
Number of clusters active: 2

----- CLUSTER blues -----

ID          NAME          STATE          SUBSTATE          PRIMARY NODE
4           blues         UP             STABLE            otis

Cluster blues (4) has 2 nodes:

NODE          STATE          CLUSTER ID     NB INTERFACES
otis         UP             4              2
  INTERFACE   STATE          ADDRESS
otis_srv     UP             130.183.1.95
otis_tty1    UP             0.0.0.0

NODE          STATE          CLUSTER ID     NB INTERFACES
redding      UP             4              2
  INTERFACE   STATE          ADDRESS
redding_srv  UP             130.183.1.85
redding_tty1 UP             0.0.0.0

----- CLUSTER cluster_tests -----

ID          NAME          STATE          SUBSTATE          PRIMARY NODE
27          cluster_tests UP             STABLE            mars_srv

Cluster cluster_tests (27) has 2 nodes:

NODE          STATE          CLUSTER ID     NB INTERFACES
mars_srv     UP             27             2
  INTERFACE   STATE          ADDRESS
mars_srv     UP             130.183.5.2
mars_tty0    UP             0.0.0.0

NODE          STATE          CLUSTER ID     NB INTERFACES
mercure_srv  UP             27             2
  INTERFACE   STATE          ADDRESS
mercure_srv  UP             130.183.5.1
mercure_tty1 UP             0.0.0.0
```

Les informations affichées indiquent le nom, l'ID, l'état et le sous-état de la grappe. Dans cet exemple, les 2 grappes sont actives et équipées de 2 noeuds, lesquels sont également actifs. En outre, chaque noeud est équipé de 2 cartes réseau.

## Fichiers

**/usr/sbin/cluster/clusterview** est le fichier exécutable.

**/usr/sbin/cluster/etc/clhosts** est le fichier contenant la liste des noms hôte disponibles.

## Voir aussi

La commande **clinfo** dans la documentation HACMP.

## Implémentation d'alias d'adresses IP

Le logiciel BullCluster comporte une option de menu qui active la fonction d'alias d'adresse IP.

### Objet

La fonction d'alias d'adresse IP est une fonction facultative, qui ne présente d'intérêt que si le basculement d'adresse IP (IPAT) est mis en oeuvre sur la grappe.

Si IPAT est actif, une carte de service utilise deux adresses IP différentes, selon l'état en cours :

- à l'amorçage et jusqu'au démarrage des services de grappe, l'adresse IP d'amorçage est affectée à la carte de service;
- quand les services de grappe sont démarrés sur le noeud, celui-ci prend l'adresse IP de service.

Supposons, par exemple, que le noeud `foo` est doté d'une carte de service associée à l'adresse de service IP `foo` et à l'adresse d'amorçage IP `foo_boot`. Selon l'état de la grappe, les applications (ou les utilisateurs) souhaitant se connecter à cette carte de service doivent parfois utiliser l'adresse `foo` et d'autres fois l'adresse `foo_boot`. De plus, une application basée sur l'adresse de service `foo` peut en réalité être connectée à un autre noeud de grappe, si cette adresse a été basculée sur ce noeud.

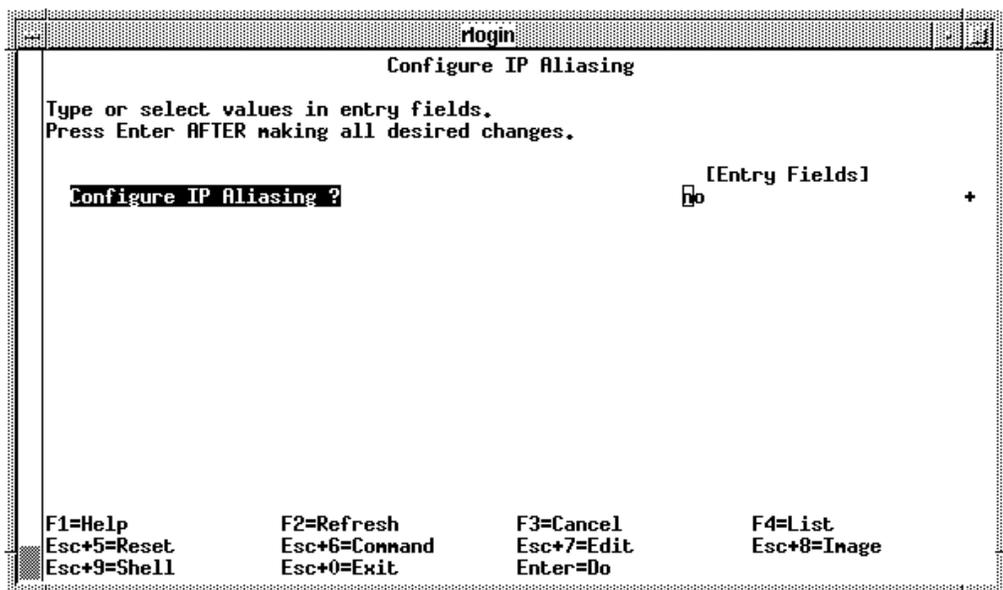
Ces changements d'adresse sont peu pratiques du point de vue de l'utilisateur ou de l'application. Par exemple, pour se connecter à distance au noeud `foo` via sa carte de service, vous devez tantôt indiquer `rlogin foo`, tantôt `rlogin foo_boot`, sans connaître au départ l'adresse correcte.

L'alias d'adresses IP permet de contourner ce problème, par le biais d'alias des adresses de service et d'amorçage.

### Procédure

Pour implémenter la fonction d'alias IP (facultative) :

1. **Arrêtez, le cas échéant, les services de grappe sur tous les noeuds.**
2. Sur le menu principal SMIT HACMP, sélectionnez successivement **Bull Cluster Easy Configuration** et **IP Aliasing Configuration**. L'écran suivant s'affiche :



3. Le champ **Configure IP Aliasing?** affiche la valeur définie en cours ("yes" ou "no" selon que l'alias IP est configuré ou non). Basculez sur la valeur voulue puis validez en appuyant sur Entrée.
4. Dès l'écran a été validé, la fonction d'alias IP est activée (ou désactivée) sur le noeud courant, et sur les autres noeuds de la grappe (noeuds dont la carte de service est associée à la fois à une adresse IP de service et à une adresse IP d'amorçage).

**Remarques :**

Sachez que la fonction d'alias IP décrite ci-dessus est basée sur la commande AIX `ifconfig`, qui permet de déclarer des alias pour une adresse IP.

En outre, notez que lorsque cette fonction est activée :

- Dans les fichiers journaux HACMP, les messages relatifs à l'événement `swap_adapter` apparaissent deux fois.
- Un post-événement est personnalisé pour `acquire_service_addr` (sur la base de la gestion de la personnalisation des événements BullCluster).
- En cas de basculement d'adresse IP, l'adresse IP de service du noeud défaillant est transférée sur la carte de secours du noeud de secours et la carte de secours du noeud défaillant reprend son adresse d'amorçage.

# Implémentation du transfert d'applications (Rolling Applications)

## Objet

Lorsque plusieurs applications sont exécutées sur un même noeud afin d'effectuer des opérations d'optimisation des performances ou de maintenance, vous avez besoin de transférer des applications d'un noeud vers un autre sans procéder à un basculement.

Le transfert d'applications permet d'effectuer cette opération pour une application (ou un groupe d'applications).

Pour utiliser cette fonction, les ressources doivent être configurées de manière spécifique à l'aide de HACMP :

- Un groupe de ressources est créé avec un label IP de service configuré comme ressource, sans que le serveur d'application, le système de fichiers ou le groupe de volume ne soient configurés.

Voici un exemple :

Configure Resources for a Resource Group : root@toi

Resource Group Name	foo_rg	
Node Relationship	cascading	
Participating Node Names	foo bar	
Service IP label	foo	List
Filesystems		List
Filesystems Consistency Check	fsck	List ▲ ▼
Filesystems Recovery Method	sequential	List ▲ ▼
Filesystems to Export		List
Filesystems to NFS mount		List
Volume Groups		List
Concurrent Volume groups		List
Raw Disk PVIDs		List
AIX Connections Services		List
AIX Fast Connect Services		List
Application Servers		List
Highly Available Communication Links		List
Miscellaneous Data		
Inactive Takeover Activated	false	List ▲ ▼
9333 Disk Fencing Activated	false	List ▲ ▼
SSA Disk Fencing Activated	false	List ▲ ▼
Filesystems mounted before IP configured	false	List ▲ ▼

OK Command Reset Cancel ?

- Pour chaque application (ou groupe d'applications) qui peut faire l'objet d'une migration, un groupe de ressources est créé, contenant un (ou des) serveur(s) qui permettent de démarrer le(s) application(s) et les ressources lvm (groupes de volume, systèmes de fichiers). Toutefois, les labels IP de service ne sont pas configurés.

Voici un exemple :

Configure Resources for a Resource Group : root@toi

Resource Group Name	foo_rg	
Node Relationship	cascading	
Participating Node Names	foo bar	
Service IP label		List
Filesystems	foofs	List
Filesystems Consistency Check	fscck	List ▲ ▼
Filesystems Recovery Method	sequential	List ▲ ▼
Filesystems to Export		List
Filesystems to NFS mount		List
Volume Groups	foovg	List
Concurrent Volume groups		List
Raw Disk PVIDs		List
AIX Connections Services		List
AIX Fast Connect Services		List
Application Servers	fooapps	List
Highly Available Communication Links		List
Miscellaneous Data		
Inactive Takeover Activated	false	List ▲ ▼
9333 Disk Fencing Activated	false	List ▲ ▼
SSA Disk Fencing Activated	false	List ▲ ▼
Filesystems mounted before IP configured	false	List ▲ ▼

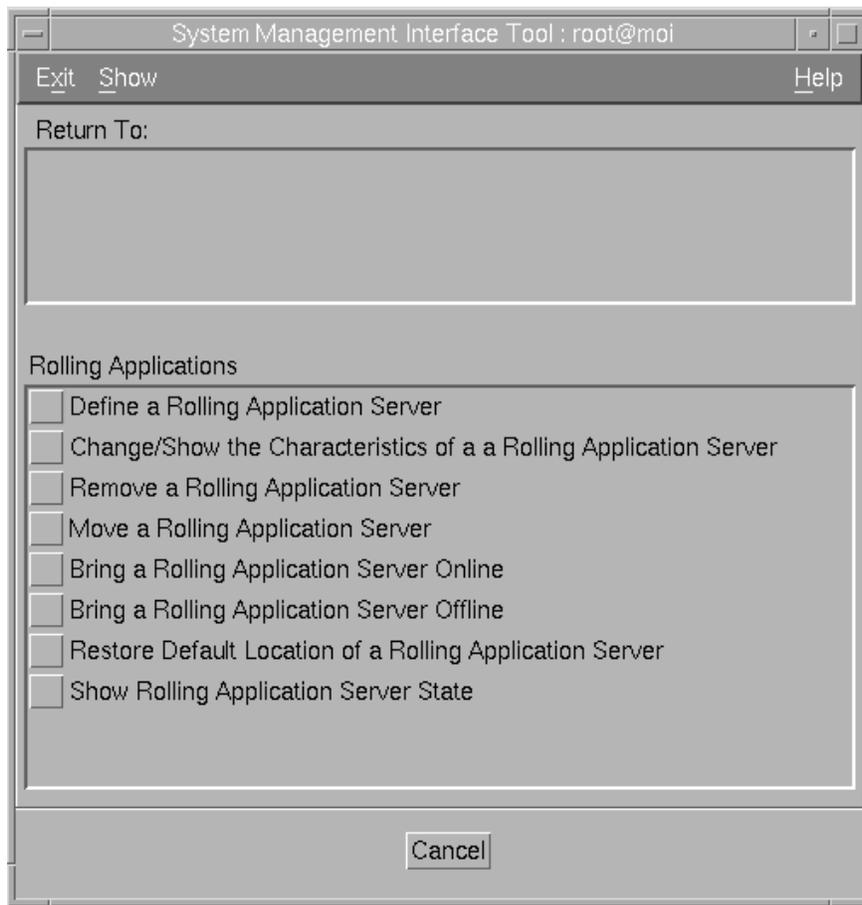
OK Command Reset Cancel ?

L'objet de cette fonction étant de faire migrer une application (ou un groupe d'applications) sans arrêter HACMP, le mécanisme de configuration dynamique doit être utilisé (commande *clRGdare*).

## Procédure

Pour implémenter les fonctions de transfert d'applications (facultatif) :

- A partir du menu SMIT HACMP, sélectionnez successivement les options **Bull Cluster Easy Configuration** et **Rolling Applications**. L'écran suivant s'affiche :



## Présentation du menu Rolling Applications

**Le menu Rolling Applications permet d'effectuer les opérations suivantes :**

- Configuration d'un serveur de transfert d'applications
  - Définition d'un serveur de transfert d'applications
  - Modification/Affichage des caractéristiques d'un serveur de transfert d'applications
  - Retrait d'un serveur de transfert d'applications
- Gestion d'un serveur de transfert d'applications
  - Mise en ligne d'un serveur de transfert d'application
  - Mise hors ligne d'un serveur de transfert d'applications
  - Déplacement d'un serveur de transfert d'applications
  - Restauration de l'emplacement par défaut d'un serveur de transfert d'applications

### **ATTENTION :**

**Avant d'effectuer une opération de gestion sur un serveur de transfert d'applications, assurez-vous que celui-ci est configuré dans un groupe de ressources HACMP.**

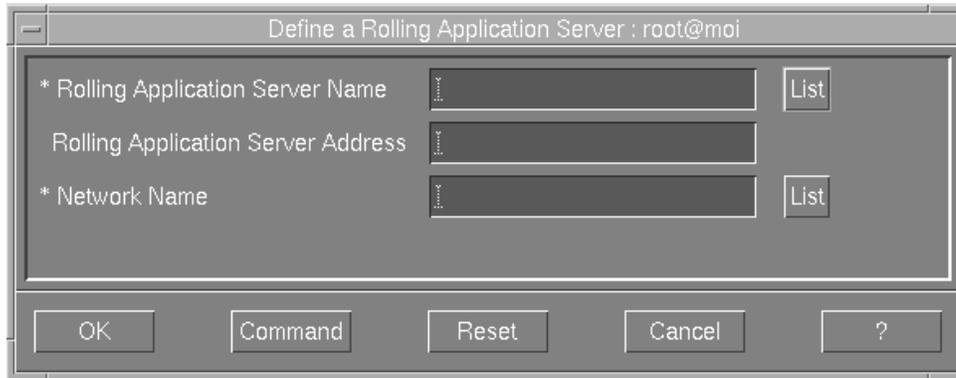
- Affichage de l'état du serveur de transfert d'applications
  - Visualisation de l'état du serveur de transfert d'applications

## Définition d'un serveur de transfert d'applications

Cette action permet d'ajouter un serveur de transfert d'applications dans une configuration existante.

A partir d'un noeud quelconque, activez le menu **Bullcluster Rolling Applications** et sélectionnez **Define a Rolling Application Server**.

La fenêtre **Define a Rolling Application Server** s'affiche.



1. Sélectionnez le nom du serveur de transfert d'applications dans la liste appropriée, ou saisissez-le. Le bouton List permet d'afficher la liste des serveurs de transfert d'applications configurés dans un groupe de ressources HACMP.
2. Entrez l'adresse IP du serveur de transfert d'applications sélectionné.
3. Sélectionnez le nom du réseau dans la liste appropriée, ou saisissez-le.
4. Validez l'écran pour appliquer vos paramètres.

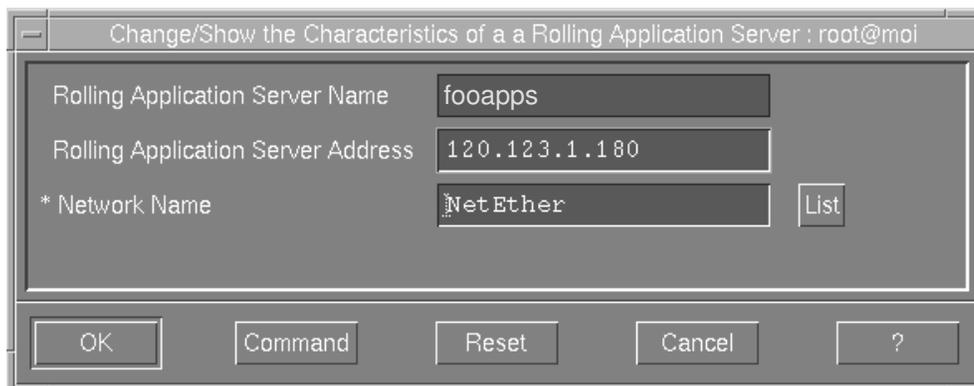
**Remarque :** Des données relatives à la configuration du transfert d'applications sont automatiquement transmises à tous les noeuds de la grappe.

## Modification/Affichage des caractéristiques d'un serveur de transfert d'applications

Cette action permet de modifier ou d'afficher les caractéristiques d'un serveur de transfert d'applications.

A partir d'un noeud quelconque, activez le menu **Bullcluster Rolling Applications**, puis sélectionnez **Change/Show the Characteristics of a Rolling Application Server**. Une liste des serveurs de transfert d'applications s'affiche.

1. Sélectionnez le serveur de transfert d'applications à afficher ou à modifier. La fenêtre **Change/Show the Characteristics of a Rolling Application Server** s'affiche.



2. Dans ce menu, le champ Rolling Application Server Name n'apparaît qu'à titre indicatif. Vous ne pouvez pas modifier les données qui y sont affichées.

Pour modifier l'adresse du serveur, saisissez-la dans le champ Rolling Application Server Address.

Pour modifier le nom du réseau, sélectionnez-le dans la liste appropriée ou saisissez-le dans le champ Network Name.

3. Une fois que vous avez modifié les champs nécessaires, validez (en cliquant sur **OK**) pour appliquer les nouveaux paramètres.

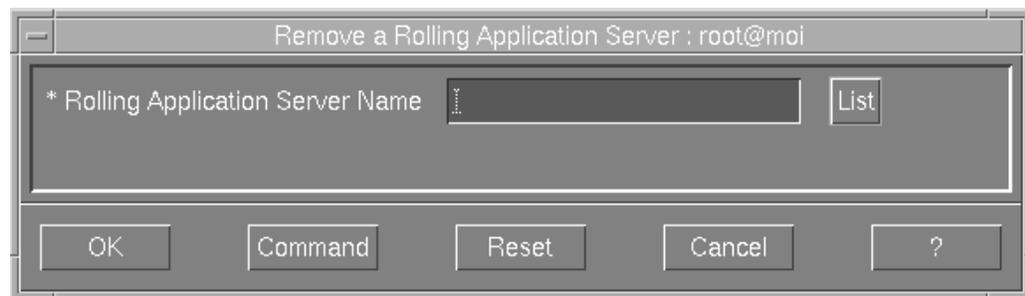
**Remarque :** Les paramètres sont automatiquement transmis à tous les noeuds de la grappe.

## Retrait d'un serveur de transfert d'applications

Cette action permet de retirer un serveur de transfert d'applications d'une configuration existante.

A partir d'un noeud quelconque, activez le menu **Bullcluster Rolling Applications** et sélectionnez **Remove a Rolling Application Server**.

La fenêtre **Remove a Rolling Application Server** s'affiche.



1. Sélectionnez dans la liste le nom du serveur à retirer. Le bouton List permet d'afficher la liste des serveurs de transfert d'applications.
2. Validez la fenêtre en cliquant sur **OK** pour appliquer les nouveaux paramètres.

**Remarque :** Les paramètres sont automatiquement transmis à tous les noeuds de la grappe.

## Mise en ligne d'un serveur de transfert d'applications

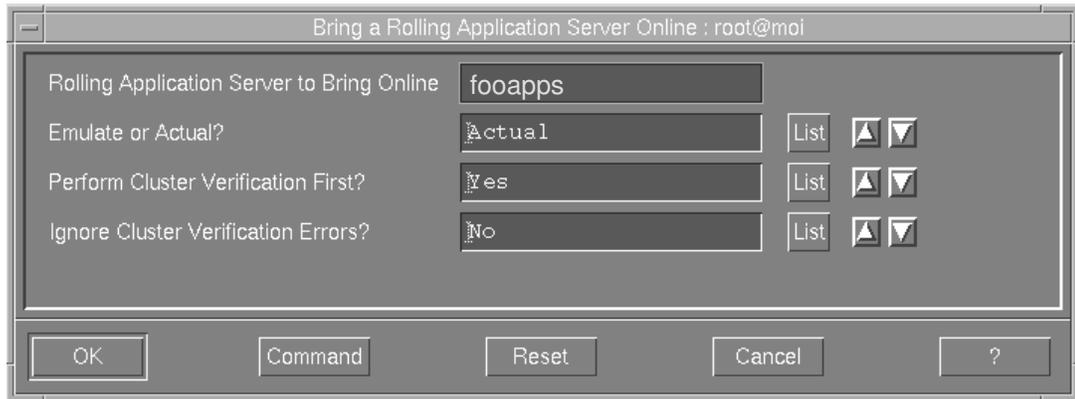
Cette action permet de lancer le groupe de ressources associé au serveur de transfert d'applications défini sur le noeud comme prioritaire.

A partir d'un noeud quelconque, activez le menu **Bullcluster Rolling Applications**, puis sélectionnez **Bring a Rolling Application Server Online**.

Une liste de serveurs de transfert d'applications s'affiche.

1. Dans cette liste, sélectionnez le serveur à mettre en ligne.

La fenêtre **Bring a Rolling Application Server Online** s'affiche.



2. Sélectionnez l'option **Actual** (par défaut) pour mettre le serveur en ligne.

L'option **Emulate** permet d'exécuter la commande dans un mode permettant de ne pas modifier la grappe, mais qui affiche les résultats/messages comme si l'action avait eu lieu.

3. Indiquez si vous souhaitez ou non exécuter l'utilitaire **clverify** pour vérifier la configuration avant de procéder à la migration.

Par défaut, l'exécution de l'utilitaire est activée (**Yes**). Il est recommandé de procéder à la vérification de toute modification dans la configuration avant de la mettre en oeuvre.

Toutefois, si vous êtes certain que la nouvelle configuration conviendra, vous pouvez passer l'étape de vérification pour accélérer le processus.

4. Pendant le processus de vérification, vous pouvez prolonger l'action DARE même si des erreurs de configuration sont détectées.

Par défaut cette option est désactivée (**No**). Il est recommandé de procéder à la vérification de toute modification dans la configuration avant de la mettre en oeuvre.

Toutefois, dans certains cas, il est préférable de poursuivre la procédure même si la vérification a échoué. Sélectionnez alors l'option **Yes**.

5. Validez la fenêtre pour appliquer les nouveaux paramètres.

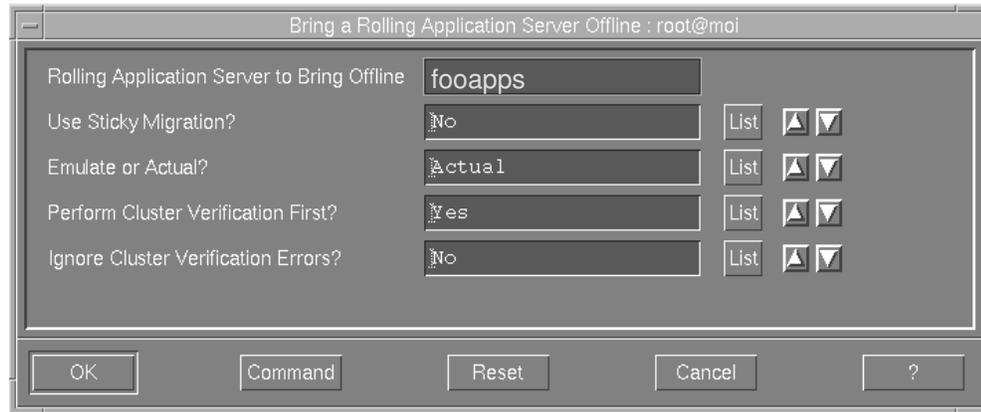
## Mise hors ligne d'un serveur de transfert d'applications

Cette action permet de désactiver ou d'arrêter le groupe de ressources associé au serveur de transfert d'applications spécifié.

A partir d'un noeud quelconque, activez le menu **Bullcluster Rolling Applications**, puis sélectionnez **Bring a Rolling Application Server Offline**. Une liste de serveurs de transfert d'applications s'affiche.

1. Dans cette liste, sélectionnez le serveur à mettre hors ligne.

La fenêtre **Bring a Rolling Application Server Offline** s'affiche.



2. Indiquez si vous souhaitez effectuer une migration sticky ou normale.

Par défaut, la migration sticky est désactivée (**No**). Le groupe de ressources associé au serveur de transfert d'applications est arrêté, mais son emplacement prioritaire demeure inchangé. Enfin, le groupe de ressources associé peut être relancé lorsque l'événement suivant se produit sur la grappe.

Si vous sélectionnez **Yes**, le groupe de ressources associé au serveur de transfert d'applications s'arrête et tente de rester hors ligne pendant toute la durée du processus de relève après panne ou de réintégration de la grappe.

3. Sélectionnez **Actual** (option par défaut) pour mettre le serveur hors ligne.

L'option **Emulate** permet d'exécuter la commande dans un mode permettant de ne pas modifier la grappe, mais qui affiche les résultats/messages comme si l'action avait eu lieu.

4. Indiquez si vous souhaitez ou non exécuter l'utilitaire **clverify** pour vérifier la configuration avant de procéder à la migration.

Par défaut, l'exécution de l'utilitaire est activée (**Yes**). Il est recommandé de procéder à la vérification de toute modification dans la configuration avant de la mettre en oeuvre.

Toutefois, si vous êtes certain que la nouvelle configuration conviendra, vous pouvez passer l'étape de vérification pour accélérer le processus.

5. Pendant le processus de vérification, vous pouvez prolonger l'action DARE même si des erreurs de configuration sont détectées.

Par défaut cette option est désactivée (**No**). Il est recommandé de procéder à la vérification de toute modification dans la configuration avant de la mettre en oeuvre.

Toutefois, dans certains cas, il est préférable de poursuivre la procédure même si la vérification a échoué. Sélectionnez alors l'option **Yes**.

6. Validez la fenêtre pour appliquer les nouveaux paramètres.

## Déplacement d'un serveur de transfert d'applications

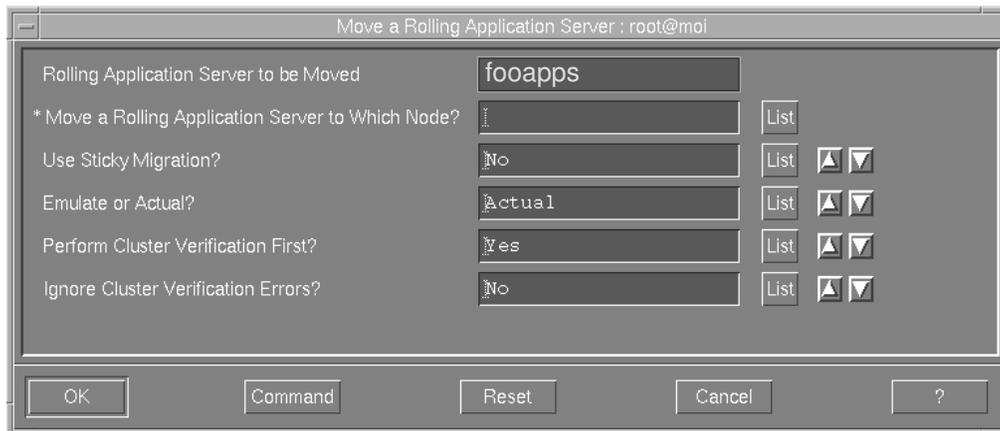
Cette action permet de déplacer le groupe de ressources associé au serveur de transfert d'applications spécifié.

A partir d'un noeud quelconque, activez le menu **Bullcluster Rolling Applications**, puis sélectionnez **Move a Rolling Application Server**.

Une liste de serveurs de transfert d'applications s'affiche.

1. Dans cette liste, sélectionnez le serveur à déplacer.

La fenêtre **Move a Rolling Application Server** s'affiche.



2. Dans la liste appropriée, sélectionnez le noeud vers lequel vous souhaitez déplacer le serveur, ou saisissez son nom.

3. Indiquez si vous souhaitez effectuer une migration sticky ou normale.

Par défaut, la migration sticky est désactivée (**No**). Le groupe de ressources associé au serveur de transfert d'applications est arrêté, mais son emplacement prioritaire demeure inchangé. Enfin, le groupe de ressources associé peut être relancé lorsque l'événement suivant se produit sur la grappe.

Si vous sélectionnez **Yes**, le groupe de ressources associé au serveur de transfert d'applications s'arrête et tente de rester hors ligne pendant toute la durée du processus de relève après panne ou de réintégration de la grappe.

4. Sélectionnez **Actual** (option par défaut) pour mettre le serveur hors ligne.

L'option **Emulate** permet d'exécuter la commande dans un mode permettant de ne pas modifier la grappe, mais qui affiche les résultats/messages comme si l'action avait eu lieu.

5. Indiquez si vous souhaitez ou non exécuter l'utilitaire **clverify** pour vérifier la configuration avant de procéder à la migration.

Par défaut, l'exécution de l'utilitaire est activée (**Yes**). Il est recommandé de procéder à la vérification de toute modification dans la configuration avant de la mettre en oeuvre.

Toutefois, si vous êtes certain que la nouvelle configuration conviendra, vous pouvez passer l'étape de vérification pour accélérer le processus.

6. Pendant le processus de vérification, vous pouvez prolonger l'action DARE même si des erreurs de configuration sont détectées.

Par défaut cette option est désactivée (**No**). Il est recommandé de procéder à la vérification de toute modification dans la configuration avant de la mettre en oeuvre.

Toutefois, dans certains cas, il est préférable de poursuivre la procédure même si la vérification a échoué. Sélectionnez alors l'option **Yes**.

7. Validez la fenêtre pour appliquer les nouveaux paramètres.

## Restauration de l'emplacement par défaut d'un serveur de transfert d'applications

Cette action permet de replacer le groupe de ressources associé au serveur spécifié sur le noeud initial ; elle supprime toute fonction sticky définie.

A partir d'un noeud quelconque, activez le menu **Bullcluster Rolling Applications**, puis sélectionnez **Restore Default Location of a Rolling Application Server**. Une liste de serveurs de transfert d'applications s'affiche.

1. Dans cette liste, sélectionnez le serveur à restaurer.

La fenêtre **Restore Default Location of a Rolling Application Server** s'affiche.



2. Sélectionnez l'option **Actual** (par défaut) pour mettre le serveur en ligne.

L'option **Emulate** permet d'exécuter la commande dans un mode permettant de ne pas modifier la grappe, mais qui affiche les résultats/messages comme si l'action avait eu lieu.

3. Indiquez si vous souhaitez ou non exécuter l'utilitaire **clverify** pour vérifier la configuration avant de procéder à la migration.

Par défaut, l'exécution de l'utilitaire est activée (**Yes**). Il est recommandé de procéder à la vérification de toute modification dans la configuration avant de la mettre en oeuvre.

Toutefois, si vous êtes certain que la nouvelle configuration conviendra, vous pouvez passer l'étape de vérification pour accélérer le processus.

4. Pendant le processus de vérification, vous pouvez prolonger l'action DARE même si des erreurs de configuration sont détectées.

Par défaut cette option est désactivée (**No**). Il est recommandé de procéder à la vérification de toute modification dans la configuration avant de la mettre en oeuvre.

Toutefois, dans certains cas, il est préférable de poursuivre la procédure même si la vérification a échoué. Sélectionnez alors l'option **Yes**.

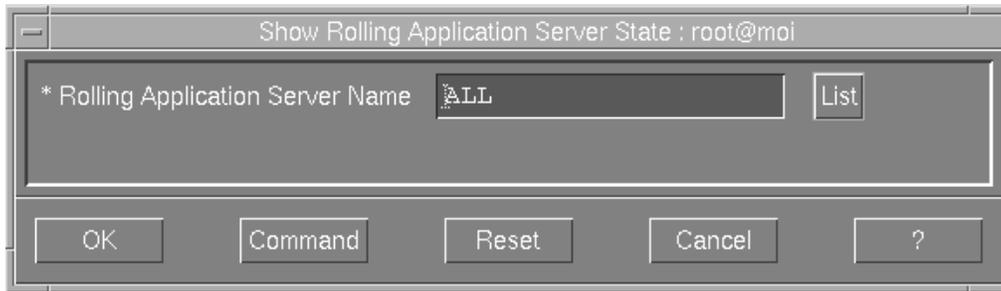
5. Validez la fenêtre pour appliquer les nouveaux paramètres.

## Afficher l'état du serveur de transfert d'applications

Cette action permet d'afficher l'état d'un ou de plusieurs serveurs de transfert d'applications.

A partir d'un noeud quelconque, activez le menu **Bullcluster Rolling Applications**, puis sélectionnez **Show Rolling Application Server State**.

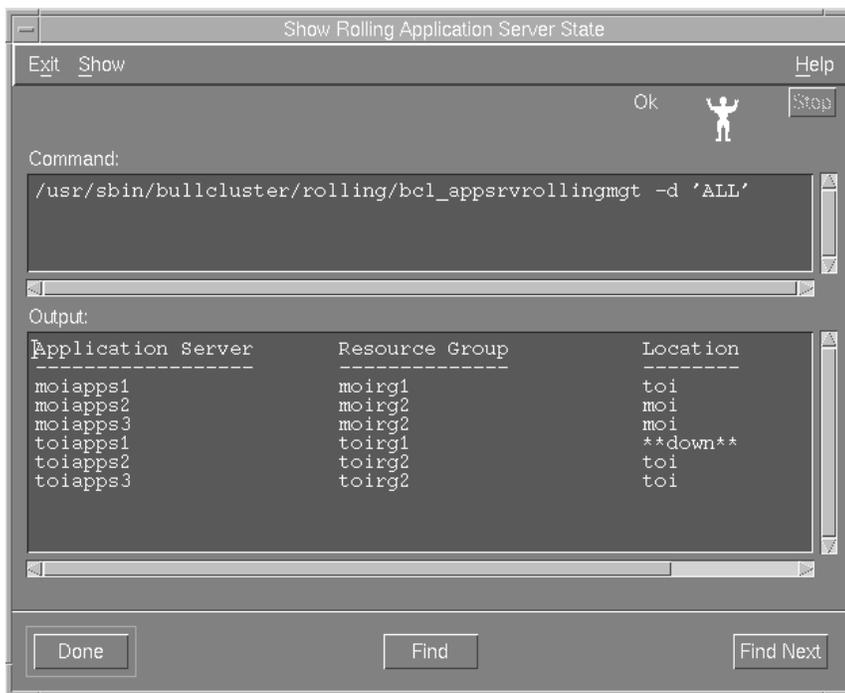
La fenêtre **Show Rolling Application Server State** s'affiche.



1. Dans la liste appropriée, sélectionnez le ou les serveur(s) dont vous souhaitez afficher l'état. L'option **ALL** permet de sélectionner la totalité des serveurs de la liste.
2. Validez la fenêtre pour afficher l'état des serveurs sélectionnés.

Chaque serveur apparaît avec le groupe de ressources associé et son emplacement (ou l'état : **\*\*down\*\***, si l'emplacement n'est pas indiqué).

Voici un exemple de résultat que vous pouvez obtenir en utilisant ce menu :



---

# Implémentation de HA native (avec ou sans armoire d'extension disque)

## Objet

HA native est une fonction en option qui doit être activée si votre configuration est de type stockage partagé HA native. Ce type de configuration est décrit dans **Planification du stockage partagé "HA native" pour les solutions Escala HA**, page 3-6.

Si une défaillance se produit, aucun basculement automatique n'est possible dans l'environnement de stockage partagé HACMP standard, notamment en cas de défaillance d'un noeud.

HA native pallie cette difficulté en :

- modifiant l'activation des groupes de volumes partagés,
- modifiant la désactivation des groupes de volumes partagés,
- ajoutant un pré-événement pour la synchronisation miroir (sur la base de la gestion de la personnalisation des événements BullCluster).

## Configuration

Pour configurer HA native, servez-vous de la liste suivante :

- Sur les deux noeuds, tous les disques internes ou externes partagés doivent être à l'état AVAILABLE (disponible). (Les disques externes sont ceux de l'armoire d'extension).

Ces disques doivent être **tous** visibles par chaque noeud. Une erreur dans le total des disques à l'état disponible indique un problème de câblage. Dans ce cas, vérifiez votre installation à l'aide du schéma de câblage. Pour en savoir plus, reportez-vous à **Câblage des cartes SCSI et des bus**, page 3-8.

A partir d'un noeud, créez le(s) groupe(s) de volumes (VG) partagé(s), en utilisant les disques partagés par les deux systèmes (internes et externes) pour la fonction HA native. Suivez les instructions du manuel *AIX - Guide d'administration : système d'exploitation et unités* :

- Chaque groupe de volumes doit comporter au moins deux disques (un sur chaque noeud).
- Vérifiez les paramètres QUORUM et auto-varyon de chaque : **TOUS DEUX** doivent avoir la valeur **NO**.
- Chaque volume logique des groupes de volumes partagés doit être intégralement créé sur un seul disque et dupliqué (2 copies) sur un autre disque de l'autre noeud ou de l'armoire d'extension sur un bus SCSI distinct.
- A la création d'un volume logique, donnez au paramètre SCHEDULING POLICY for writing logical partitions copies la valeur **SEQUENTIAL** (dans SMIT, la valeur par défaut est PARALLEL).
- Mettez en miroir le volume logique **jfslog** (valeur par défaut) de votre groupe de volumes.

Importez les informations des groupes de volumes depuis l'autre noeud :

- Soyez prudent au niveau des noms des unités de disques, qui peuvent être différents sur ce noeud.

Tenez compte de l'identificateur de volume logique (PVID) pour importer le bon groupe de volumes.

**Remarque :** Pour déterminer le disque à importer, entrez :

```
lspv
```

- Vérifiez à nouveau les paramètres **QUORUM** et **auto-varyon** sur ce noeud. Après exécution de la commande **importvg**, redonnez à QUORUM et à auto-varyon leur valeur par défaut, **YES**.

```
# chvg -a n -Qn <votre VG>
```

**Il est important de noter sur la fiche de PLANIFICATION de HACMP** le nom des disques et leur PVID (ID de volume physique) sur chaque noeud. Notez également en regard de chaque disque leur code d'emplacement sur chaque noeud. Pour chaque disque, ajoutez une colonne pour inscrire le nom du système où réside réellement le disque. Indiquez en outre les informations suivantes sur l'utilisation des deux cartes SCSI-SE par la machine :

- emplacement PCI de la carte donnant accès aux disques de l'autre noeud,
- emplacement PCI de la carte donnant accès aux disques internes de ce noeud.

Ces informations aident à identifier les problèmes matériels : voir les sections REPRISE, ci-après.

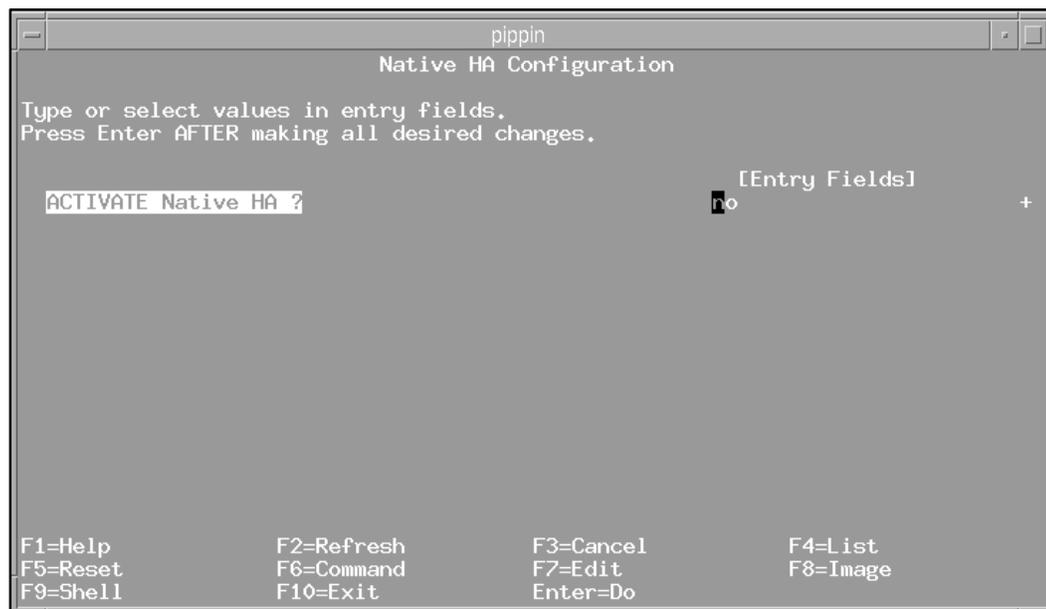
Pour la configuration de grappe HACMP, procédez comme suit :

- configurez AIX : adresses IP, ttys, etc.
- créez la topologie : noeuds, cartes.
- créez les groupes de ressources et leurs ressources : groupes de volumes, systèmes de fichiers, label IP, serveurs d'application, etc.

## Procédure

Pour implémenter la fonction HA native (facultatif) :

1. **Le cas échéant, arrêtez les services de grappes sur tous les noeuds.**
2. Sur le menu principal SMIT HACMP, sélectionnez successivement **Bull Cluster Easy Configuration, Native HA Configuration**. L'écran suivant s'affiche :



3. Le champ **ACTIVATE Native HA** affiche le paramétrage actuel ("no" si HA native n'est pas activé, "yes" s'il l'est). Au besoin, modifiez la valeur et validez avec Entrée.
4. Dès validation, HA native est activé (ou désactivé) sur le noeud courant et sur les autres noeuds de grappe.

# Reprise sur incident matériel

## Unité d'extension disque - installation matérielle

Si vous utilisez une unité d'extension disque avec haute disponibilité, toutes les unités externes partagées dans une configuration HA et connectées au tiroir CPU **NE DOIVENT PAS** être mise hors tension pendant les procédures d'installation matérielles.

## Remplacement du disque partagé d'une configuration HA native

La procédure indique les étapes requises pour remplacer un disque interne ou externe partagé dans une configuration HA native est configuré et HACMP installé. Une défaillance de disque est totalement transparente pour les applications, dans la mesure où les données restent disponibles sur le miroir. Cette section traite de la procédure à suivre pour réintroduire le nouveau disque dans le plan miroir.

Un groupe de volume VG1 dispose de 2 disques mis en miroir (chacun sur un noeud distinct) dont l'un, hdisk4, par exemple, doit être remplacé.

**Remarque :** Dans cet exemple, les disques ont le même nom sur les deux systèmes. Ce n'est pas toujours le cas : le même disque physique peut être nommé différemment sur 2 systèmes (hdisk4 sur le noeud A et hdisk5 sur le noeud B, par exemple). Vérifiez donc les disques partagés sur les différents noeuds, via les ID des PV (volumes physiques).

### Contrôle initial

1. Lancez toutes les commandes à partir du système où le groupe de volume VG1 est actif. Aucune commande administrative ne doit être exécutée sur les autres systèmes jusqu'à la fin de la procédure (sauf indication contraire explicite).
2. Identifiez le disque à remplacer et le groupe de volumes auquel il appartient :

```
# errpt | grep hdisk or errpt -a | pg
```

3. Vérifiez le câblage avant de poursuivre. Si un câble en est débranché, passez à la section **Reprise : connexion de câble défectueuse**, sinon, continuez.
4. S'il n'y a pas eu de basculement, vérifiez les cartes utilisées sur les bus SCSI sur les deux noeuds. Si un enregistrement dans le journal des erreurs apparaît concernant l'une des cartes sur l'un des noeuds au moment de l'enregistrement du disque, passez directement à la section **Reprise : défaillance d'une carte**.
5. Localisez le groupe de volumes et vérifiez les volumes logiques concernés sur le disque défaillant :

```
# lspv
hdisk4          000002634f03cc16    vg1
hdisk5          000002634e995408    vg1
...

# lsvg -M VG1 (or lsvg -p VG1)
hdisk2:1        sharedlv:1:1
hdisk2:2        sharedlv:2:1
hdisk2:3-50
hdisk4:1        sharedlv:1:2    stale
hdisk4:2        sharedlv:1:2    stale
hdisk4:3-50
```

La sortie "stale" indique un problème sur hdisk4.

Notez le nom de l'unité de disque :

```
Disk drive:      hdisk4
```

## 6. Trouvez le code d'emplacement de l'unité :

```
# lsdev -Cc disk
hdisk4 Available 04-07-00-8,0 Bull 2.1 GB 16 Bit SCSI Disk Drive.
hdisk5 Available 04-08-00-8,0 Bull 2.1 GB 16 Bit SCSI Disk Drive.
...
```

Dans cet exemple, le code d'emplacement est 04-07-00-8,0 ce qui signifie que l'emplacement de carte PCI 7 avec l'ID SCSI 8 comme disque se trouve dans la zone de données supérieure (les disques se trouvant dans le boîtier de disque partagé supérieur sont identifiés par les ID SCSI : 8,9,a - de haut en bas, et dans le boîtier inférieur par : 0,1,2)

### Procédure

#### 1. Identifiez tous les volumes logiques de hdisk4 :

```
# lspv -l hdisk4
List of logical volumes:
```

#### 2. Réduisez le nombre de copies de tous les volumes logiques mis en miroir sur ce disque hdisk4, en spécifiant que la copie à supprimer se trouve sur le disque hdisk4 défaillant :

```
# rmlvcopy sharedlv 1 hdisk4
etc... (pour tous les volumes logiques du groupe de volumes)
```

N'oubliez pas de supprimer également la copie dans le volume logique JFSLOG, s'il en existe un sur ce disque.

#### 3. Retirez le disque du groupe de volumes :

```
# reducevg VG1 hdisk4
```

Vous pouvez effectuer cette opération même si vos systèmes de fichiers sont montés et vos applications en cours.

#### 4. SUR CHAQUE NOEUD, retirez le disque défaillant de la base de données de configuration :

```
# rmdev -l hdisk4 -d
```

**Remarque :** N'oubliez pas que le nom du disque peut ne pas être identique sur tous les noeuds. Lancez la commande **lspv** pour déterminer l'ID PV du disque sur chaque noeud.

```
# lspv
```

#### 5. MAINTENANT, retirez le disque du boîtier, sur le système identifié précédemment, et insérez le nouveau disque.

**Remarque :** Sur certains serveurs Escala, vous pouvez ôter les disques "à chaud".

#### 6. Sur chaque noeud, configurez le nouveau disque :

```
# cfgmgr
```

#### 7. Déterminez le nom du nouveau disque :

```
# lsdev -Cc disk
hdisk4      Available      Location      Comments
# lspv
hdisk4      PVID
```

Mettez à jour en conséquence les informations de la fiche de PLANIFICATION (groupes de volumes partagés et disques).

Dans les étapes suivantes, nous avons supposé que le même nom logique (hdisk4) a été attribué au nouveau disque.

Si le disque n'a pas de PVID,

sur un noeud, entrez :

```
chdev -l hdisk4 -a pv=yes
```

et sur l'autre :

```
rmdev -l hdisk4  
mkdev -l hdisk4
```

8. Ajoutez le disque au groupe de volumes :

```
# extendvg VG1 hdisk4
```

Mettez à jour le PVID du nouveau disque dans la fiche de PLANIFICATION (groupe de volumes partagés et disques, si différents).

9. Ajoutez (et synchronisez) une copie miroir à tous les volumes logiques sur le disque, sans oublier la copie miroir dans JFSLOG :

```
# mklvcopy -k sharedlv 2 hdisk4  
etc...
```

Sur tous les autres noeuds, le nouveau disque sera pris en compte à la défaillance suivante.

## Reprise sur connexion de câble défectueuse

Vous êtes arrivé à ce point après détection d'un enregistrement d'erreur non suivi d'un basculement. Un des câbles en Y est mal branché ou débranché.

### Contrôle initial

Vérifiez le branchement du câble ainsi que des câbles plats SCSI internes et des fiches de bouclage.

### Procédure

1. Une fois les connexions des câbles SCSI vérifiées, tentez d'accéder à votre disque en lecture, sur chaque système (avec droits de l'utilisateur racine).

```
# dd if=/dev/rhdisk4 of=/dev/null bs=128k count=1
```

S'il se produit une erreur d'E/S avec nouvel enregistrement dans errlog, cela signifie que soit votre carte soit votre disque est défaillant et doit être remplacé. Quittez cette procédure et suivez la procédure usuelle pour remplacer la carte sur le noeud où est exécutée la commande. S'il s'agit d'un disque interne, passez à la section traitant de la reprise HA NATIVE après défaillance de carte.

2. Sur le noeud où VG1 est actif et où la commande dd a abouti, relancez la commande **varyonvg** sans arrêter les applications :

```
# varyonvg VG1
```

Cette commande synchronise les copies périmées.

## Reprise sur défaillance de noeud

Le noeud où VG1 était actif est en panne (888) ou hors tension, et VG1 est désormais actif sur l'autre noeud (après basculement HACMP).

### Contrôle initial

Sur le noeud restant, le journal errlog contient des entrées concernant la défaillance du bus SCSI et les disques internes de l'autre système, devenus inaccessibles.

Dans `/tmp/hacmp.out` ou `/var/adm/cluster.log`, il apparaît un événement `node_down_remote` car l'autre noeud a échoué.

### Procédure

Les applications sont relancées automatiquement sur le noeud restant, bien que les disques du système défaillant soient désormais absents du groupe de volumes VG1.

Le groupe de volumes de ces disques inaccessibles peut indiquer NONE lorsque VG1 est pris en charge par le noeud restant.

```
# lspv VG1
hdisk4          PVID          NONE
hdisk5          PVID          VG1
```

Ici, `hdisk4` est un disque interne de la machine en panne.

Cette situation, normale, n'empêche pas l'exploitation de l'application, dans la mesure où VG1 reste accessible via `hdisk5`.

### Procédure post-réparation

Vérifiez le fonctionnement de la machine et la présence de tous les disques.

Lancez le gestionnaire de grappe sur la machine réparée :

```
# smit clstart
```

Cette procédure, automatique, ne demande aucune intervention.

**Remarque** : Une fois HACMP lancé sur le noeud défaillant, l'état NONE pour `hdisk4` sur le noeud restant n'est plus visible. En fait, le nouveau script `cl_deactivate_vgs` (appelé pour libérer VG1) livré avec cette version, désactive, exporte et ré-importe les informations de VG1 avant de le libérer. Ceci est possible dans la mesure où les applications sont déjà arrêtées et les systèmes de fichiers partagés de VG1 déjà démontés. Sur ce noeud restant, tous les disques accessibles dans VG1 n'affichent plus l'état NONE.

Sur le noeud restant, si un autre groupe de volumes, VG2, par exemple, est actif avant la défaillance de l'autre noeud, il devient inutile de synchroniser manuellement le miroir périmé. Le script pré-événement personnalisé ajouté à l'événement `node_up_remote_complete` se charge de synchroniser les éventuelles copies périmées, une fois tous les disques redevenus opérationnels.

Dès que la machine réparée a effectivement rejoint la grappe, vérifiez que tous les miroirs de chaque groupe de volumes actif sur chaque noeud sont synchronisés et qu'aucun disque partagé n'affiche l'état NONE :

```
# lsvg -l VG1 (VG2)
# lspv
```

## Reprise sur défaillance d'une carte

Si le noeud où VG1 est actif comporte une carte défaillante qu'il convient de remplacer, procédez comme indiqué à CAS 1.

Si aucun groupe de volumes n'est actif sur le noeud comportant une carte défaillante, procédez comme indiqué à CAS 2.

### CAS 1 : procédure de reprise

Il convient d'arrêter les applications sur ce noeud et de les relancer sur l'autre, car remplacer une carte suppose de mettre le système hors tension.

1. Débranchez le câble de la chaîne SCSI au niveau de sa terminaison.

**Remarque** : veillez à ce que le câble débranché reste équipé de sa terminaison.

2. Arrêtez HACMP en mode basculement.
3. Patientez jusqu'au redémarrage de l'application et à la synchronisation des volumes logiques sur le noeud de secours.
4. Mettez la machine hors tension.
5. Remplacez la carte défaillante.

### Procédure post-réparation

Procédez comme indiqué au paragraphe **Reprise sur défaillance d'un noeud**.

### CAS 2 : procédure de reprise

Aucune application n'est exécutée sur ce noeud (PAS de groupe de ressources actif) :

1. Débranchez le câble de la chaîne SCSI au niveau de sa terminaison.

**Remarque** : veillez à ce que le câble débranché reste équipé de sa terminaison.

2. Mettez la machine hors tension.
3. Changez la carte.
4. Remettez la machine sous tension.

### Procédure post-réparation

Sur ce système, lancez **hacmp** :

```
# smit clstart
```

Sur l'autre noeud, le script de pré-événement personnalisé ajouté à l'événement *node\_up\_remote\_complete* synchronise les copies périmées sur la machine restante.

---

# Implémentation du stockage partagé HA étendue

## Précautions

Chaque noeud d'une grappe définie sur deux sites doit appartenir au même sous-réseau pour respecter les spécifications HA régissant les connexions de réseau public (possibilité de basculement IP).

Les réseaux série utilisés pour HACMP heartbeat entre des noeuds de grappe doivent être étendus à l'aide de micro-modems RS232.

Les caractéristiques de configuration du réseau doivent être vérifiées sur le site du client par un technicien de maintenance Bull avant toute installation.

## Configuration

Pour mettre en oeuvre la configuration HA étendue, vérifiez les points suivants :

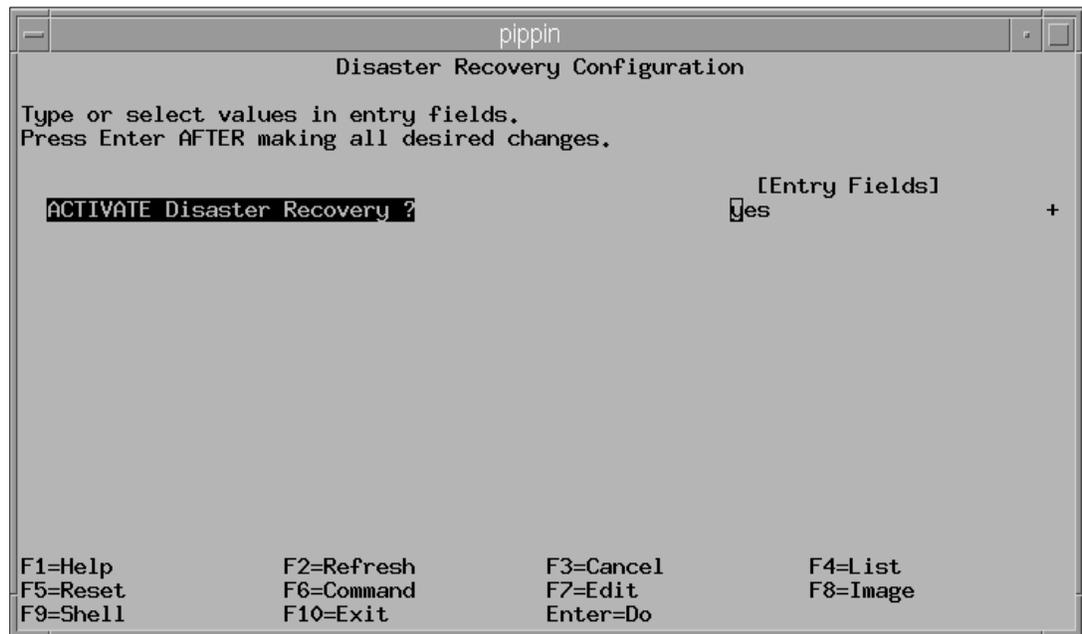
- Les objets array/atf et al\_slot doivent être disponibles sur les deux noeuds (voir la documentation relative à la carte ou au pilote AIX Emulex).
- Tous les disques doivent être disponibles sur les deux noeuds.
- Créez les groupes de volumes partagés selon la procédure habituelle, en utilisant les lun du premier DAS. Pour chaque volume logique dans un groupe de volumes, créez une copie sur un disque (lun) du second système DAS.
- Choisissez les deux lun (au minimum) de même type RAID (RAID1, RAID3 ou RAID5 pour une haute disponibilité) et de même capacité disque.
- Choisissez les copies sur différentes boucles FC-AL. Par exemple, si vous prenez disk1 (lun1) sur DAS1 lié à SPA, prenez disk2 (lun2) sur DAS2 lié à SPB.

Suivez les instructions fournies dans la configuration HA native.

## Procédure

Pour mettre en oeuvre l'environnement haute disponibilité (HA) étendue :

1. **Arrêtez, le cas échéant, les services de grappe sur tous les noeuds.**
2. Sur le menu principal HACMP SMIT, sélectionnez successivement **Bull Cluster Easy Configuration** et **Disaster Recovery Configuration**.



3. Le champ **ACTIVATE Disaster Recovery** affiche la valeur définie en cours ("yes" ou "no" selon que la reprise après sinistre est active ou non). Basculez sur la valeur voulue puis validez en appuyant sur Entrée.
4. Dès que l'écran a été validé, la fonction de reprise après sinistre (HA étendue) est activée (ou désactivée) sur le noeud courant et sur l'autre noeud de la grappe.

**Remarque :** Vous pouvez accéder à ce dernier menu en ajoutant smit dans le chemin d'accès rapide après incident.

## Reprise sur défaillance matérielle

Lorsqu'ATF (Application Transparent Failover) et/ou le gestionnaire de grappes HACMP détecte une défaillance matérielle (noeud, carte, SP, alimentation générale, hub, système DAS) sur un site, un dispositif de relève après panne s'enclenche sur l'élément redondant (carte, hub, noeud, disques miroirs AIX) pour assurer la continuité du service (applications). C'est ce qui doit se produire dans un environnement HACMP.

Appliquez, selon la nature de la défaillance, l'une des procédures, générale ou spécifique, de dépannage décrites ci-après afin de revenir à l'état précédant l'incident.

## Reprise sur noeud défaillant

HACMP sur le noeud restant reprend toutes les ressources du noeud défaillant.

Une fois le noeud défaillant dépanné, mettez-le sous tension et lancez HACMP. Les ressources seront automatiquement réactivées sur ce noeud.

## Reprise sur carte défectueuse

Lorsque la carte défectueuse est détectée sur un noeud, ATF (Application Transparence Failure) transfère le contrôle de tous les lun, liés au SP raccordé à cette carte via un hub, sollicités par les applications.

### Procédure de dépannage

Les mesures à prendre sont les suivantes :

- Arrêtez HACMP avec basculement sur ce noeud, mettez le noeud hors tension et changez la carte.
- Sur l'autre noeud, restaurez ATF (atf0 ou atf1) si des intrusions se sont produites. Sélectionnez le menu smit atf, puis **restore LUNs to originated SPs**.
- Mettez sous tension le premier noeud, vérifiez que le voyant vert est allumé sur les deux cartes FC Emulex et que le voyant jaune clignote à l'issue du processus d'amorçage.
- Redémarrez HACMP sur ce noeud.

## Reprise sur SP défaillant

ATF transfère à l'autre SP le contrôle des lun sollicités, sans interrompre les applications sur les deux noeuds.

### Procédure de dépannage

Les mesures à prendre sont les suivantes :

- Retirez le SP défaillant et remplacez-le sans arrêter les applications.
- Une fois le SP initialisé, vérifiez son état : le LED du port HUB auquel il est connecté doit être éteint.
- Restaurez ATF sur chaque noeud (atf0 ou atf1) si atfx (x : 0 ou 1) s'est introduit illicitement dans les lun de ce SP.

## Reprise sur panne générale d'alimentation électrique sur site

Le site restant (site 1) reprend toutes les ressources activées sur le site victime de la panne d'alimentation (site 2).

Une fois l'alimentation générale rétablie, exécutez la procédure de dépannage décrite ci-après.

Il est impératif de **préparer** cette procédure, car dans certains cas il peut être nécessaire d'arrêter toutes les applications dans la grappe.

### Procédure de dépannage

Les mesures à prendre sont les suivantes :

1. Mettez sous tension les hubs sur ce site (site 2).
2. Mettez sous tension le système DAS sur ce site, attendez que le voyant de service s'allume et s'éteigne.

Si SP présente un état hors ligne, éteignez puis rallumez son hub pour réinitialiser l'insertion du SP dans la boucle FC-AL.

3. Mettez sous tension le noeud sur le site (site 2).
4. Vérifiez l'état d'ATF sur les deux sites au cas où ATF se serait introduit illicitement.

En cas d'intrusion par atfx, restaurez les lun concernés à leur SP par défaut.

5. Démarrez HACMP sur le noeud du site 2. Toutes les ressources seront alors récupérées sur ce site.

Il est possible qu'Atfx se soit introduit illicitement sur les deux sites lors du démarrage du gestionnaire HACMP sur le site dépanné, dans ce cas :

- attendez la fin de l'événement `node_up_complete` sur le site (site 2) ;
- restaurez les lun de chaque atfx sur ce site (site 2) ;
- sur l'autre site (site 1), restaurez de nouveau les lun à leur SP.

## Reprise sur défaillance de disques DAS

Contactez votre technicien de maintenance Bull pour connaître la procédure d'échange de matériel ainsi que les procédures AIX à mettre en oeuvre pour rétablir la disponibilité de tous les VG avec tous les (nouveaux) disques miroirs.

---

## Personnalisation du traitement des événements de grappe

### Objet

Dans HACMP, tout événement de grappe peut être associé à des scripts de traitement exécutés lorsque l'événement en question a lieu (pour des informations de fond sur les concepts et caractéristiques HACMP relatifs aux événements de grappe, reportez-vous à la documentation HACMP).

---

## Synchronisation des horloges avec `timed`

Le temps est une donnée essentielle pour nombre d'applications, qui prennent des "décisions" en fonction de l'heure actuelle ou d'horodateurs enregistrés. Les logiciels de bases de données comme les outils d'analyse, de comptabilité et d'audit sont des applications de ce type. Une désynchronisation des horloges des noeuds conduit à des résultats peu fiables.

C'est pourquoi il est vivement conseillé d'installer le démon **timed** dans la grappe de sorte que les horloges soient synchronisées en permanence sur tous les noeuds.

### timed en bref

**timed** est un serveur de temps standard fourni avec AIX. Il synchronise l'horloge d'une machine avec celles des autres machines du réseau sur lesquelles **timed** est également actif. En ralentissant certaines horloges et en accélérant d'autres, **timed** maintient une heure uniforme sur le réseau.

**timed** est un démon exécutable en mode maître ou en mode esclave. Via ce mécanisme, **timed** assure la haute disponibilité du service horloge.

- Lorsque le démon **timed** est lancé avec l'indicateur `-M` (mode maître) :
  - Si un **timed** maître est déjà exécuté sur le réseau, le **timed** que vous venez de lancer devient un *sous-maître*. Les serveurs de sous-maître d'horloge ont pour seule fonction de prendre le relais du **timed** maître, si celui-ci vient à faillir.
  - Si aucun **timed** maître n'est déjà actif sur le réseau, le **timed** que vous venez de lancer devient le serveur d'horloge maître pour ce réseau. Le **timed** maître est responsable des transactions avec les autres serveurs **timed** pour synchroniser les horloges. Si le **timed** maître meurt (suite, par exemple, à une défaillance de l'hôte sur lequel il est actif), un nouveau serveur maître est "élu" parmi les serveurs sous-maîtres (c'est-à-dire les autres serveurs **timed** activés avec l'indicateur `-M`).
- Lorsque le démon **timed** est lancé sans l'indicateur `-M`, il est exécuté comme serveur esclave (jamais candidat au rôle de serveur de temps maître).

## Conseils pour implémenter timed dans votre grappe

Pour implémenter le service d'horloge dans votre grappe :

- Activez **timed** sur chaque noeud. Considérez qu'au moins deux serveurs **timed** (chacun sur un noeud) doivent être activés en mode maître (indicateur  $-M$ ).
- Le plus simple consiste à activer **timed** comme serveur maître (indicateur  $-M$ ) sur tous les noeuds de grappe. Vous vous assurez ainsi de la continuité du service d'horloge même s'il ne reste qu'un seul noeud actif.

Pour en savoir plus sur **timed** et sur son activation automatique au démarrage du système, reportez-vous à la documentation AIX.

### Remarque à l'attention des utilisateurs de PowerConsole

Si vous configurez un Powercluster équipé d'un PowerConsole, nous vous conseillons d'utiliser l'utilitaire de configuration intégré au module ClusterAssistant. Notez que :

- Pour accéder à cet utilitaire, activez l'icône **AIX Tools** (dans le dossier "Configuration" du groupe d'applications ClusterAssistant), puis lancez la fonction "timed Configuration".
- Avec cette fonction, **timed** est automatiquement configuré :
  - comme maître synchronisé sur l'horloge du PowerConsole,
  - comme esclave synchronisé sur les différents noeuds de la grappe.

Notez que cette configuration, bien que non conforme aux conseils donnés plus haut, est recommandée lorsque la grappe installée est équipée d'un PowerConsole.

---

## Clés de licence iFOR/LS Nodelocked

iFOR/LS, via l'utilisation de clés cryptées, contrôle les licences des logiciels utilisés par des machines autonomes ou reliées en réseau.

Si vous éprouvez des difficultés concernant les clés de licences de vos produits, reportez-vous à la documentation iFOR/LS. Notez qu'une annexe du manuel *AIX iFOR/LS Tips and Techniques* traite en particulier de l'exploitation de iFOR/LS dans un environnement HACMP.

---

## Conseils relatifs à DSMIT

DSMIT (Distributed System Management Interface Tool) est un module en option de la famille AIX. DSMIT est semblable à l'outil SMIT standard, à ceci près qu'il peut exécuter des commandes de gestion sur plusieurs systèmes en réseau simultanément.

Notez ceci :

- **N'utilisez pas DSMIT pour activer les différents menus SMIT HACMP.** DSMIT sert à démarrer et arrêter les services de grappe HACMP. Presque toutes les autres opérations SMIT HACMP entraînent des résultats imprévisibles si vous les lancez via DSMIT.
- Mais vous pouvez tirer parti de DSMIT pour d'autres tâches de gestion du système qui ne font pas appel aux menus HACMP (pour gérer des comptes utilisateur ou installer des logiciels sur plusieurs noeuds, par exemple).

Voici quelques conseils complémentaires :

- Les noeuds d'une même grappe doivent être rassemblés dans un "collectif de travail" ou un "domaine".
- L'implémentation DSMIT la plus simple consiste à configurer un seul noeud de grappe comme serveur DSMIT. Mais vous pouvez rendre DSMIT hautement disponible, en configurant plusieurs noeuds de grappe comme serveurs DSMIT. Vous n'aurez plus alors à vous soucier de noeud à utiliser pour exécuter les commandes DSMIT.

Pour en savoir plus, reportez-vous à la documentation DSMIT.

---

## Gestion des sous-systèmes DAS avec ArrayGUIde

**Remarque** : Sous AIX 4.3.3, tous les sous-systèmes DAS doivent être administrés à partir de Navisphere.

Les sous-systèmes disque DAS (2900, 3x00 et 3500) sont livrés avec les utilitaires *ArrayGUIde* et **dassmgr**. Les sous-systèmes disque DAS sont fournis avec les utilitaires *ArrayGUIde* et **dassmgr**.

- *ArrayGUIde* est une interface utilisateur graphique qui permet de gérer, de configurer et de contrôler des sous-systèmes DAS multiples, connectés à plusieurs serveurs du réseau. Vous pouvez exécuter *ArrayGUIde* à partir de n'importe quel système du réseau, et pas nécessairement à partir d'un noeud auquel est connecté le DAS.
- **dassmgr** communique avec le système DAS via une connexion série reliant un processeur de stockage à un port RS-232 sur un noeud.

Vous trouverez peut-être *ArrayGUIde* plus convivial que **dassmgr** pour exécuter les tâches de contrôle quotidiennes. Pour en savoir plus, reportez-vous aux manuels *Configuring and Managing a DAS and Using ATF* et *Using the ArrayGUIde Utility*.

### **Remarque à l'attention des utilisateurs de PowerConsole**

Si votre grappe est équipée d'un PowerConsole, vous pouvez tirer parti du module ClusterAssistant, qui intègre des utilitaires permettant configurer et d'exécuter facilement *ArrayGUIde*.

---

## Gestion des sous-systèmes DAS avec Navisphere

**Remarque** : Sous AIX 4.3.3, tous les sous-systèmes DAS doivent être gérés à partir de Navisphere.

La gestion des sous-systèmes de stockage DAS s'effectue via Navisphere Manager, interface graphique interactive de Powerconsole.

Navisphere Manager fournit des informations classées dans une hiérarchie relatives à l'état des disques physiques qui constituent le système de stockage.

Cet outil offre également des fonctions de configuration et de gestion permettant de définir la mémoire de stockage, de localiser des modules de disque dans des unités de disque logiques et de modifier les paramètres du système de stockage.



---

# Chapitre 11. Télécontrôle de grappe avec ISM

Ce chapitre explique comment télécontrôler une grappe HACMP 4.3 via ISM 4.54. Il comporte les sections suivantes :

- Concepts.
- Configuration de l'environnement SNMP sur les noeuds de grappe, page 11-3.
- Installation des fichiers requis sur la station ISM Manager, page 11-5.
- Tracé d'une image ISM de base pour représenter une grappe, page 11-6.
- Interprétation des informations de base, page 11-9.
- Examen de la base MIB HACMP et ajout d'objets à l'image, page 11-12.
- Remarques sur les objets MIB HACMP, page 11-14.
- Remarques sur les interruptions, page 11-17.

**Remarque :** ISM doit être installé et exécuté sur une station de travail AIX dédiée du réseau, et non sur un noeud de grappe. Une station de travail dédiée à l'exécution d'ISM est appelée "station de gestion ISM".

---

## Concepts

### Qu'est-ce que ISM ?

ISM (Integrated System Management) est un produit Bull conçu pour télécontrôler et gérer globalement les systèmes distribués et les réseaux.

ISM offre une interface graphique conviviale configurable. Il utilise en outre une architecture sophistiquée, dans laquelle toutes les ressources réseau sont définies comme des objets, structure qui facilite l'intégration des différents éléments dans une vue unique et cohérente.

La figure, page suivante, illustre un écran ISM.

### ISM et le contrôle de grappe

Une grappe est appelée à évoluer : un noeud peut, par exemple, en être exclu ou y être inclus, une carte de secours remplacer une carte de service, etc. Ces changements génèrent des événements qui, rapportés à ISM, vous permettent de les contrôler.

Avec ISM, vous pouvez créer une image, c'est-à-dire une représentation graphique d'une grappe avec ses noeuds et leurs interfaces réseau. Après configuration, vous pouvez afficher l'image pour contrôler visuellement la grappe.

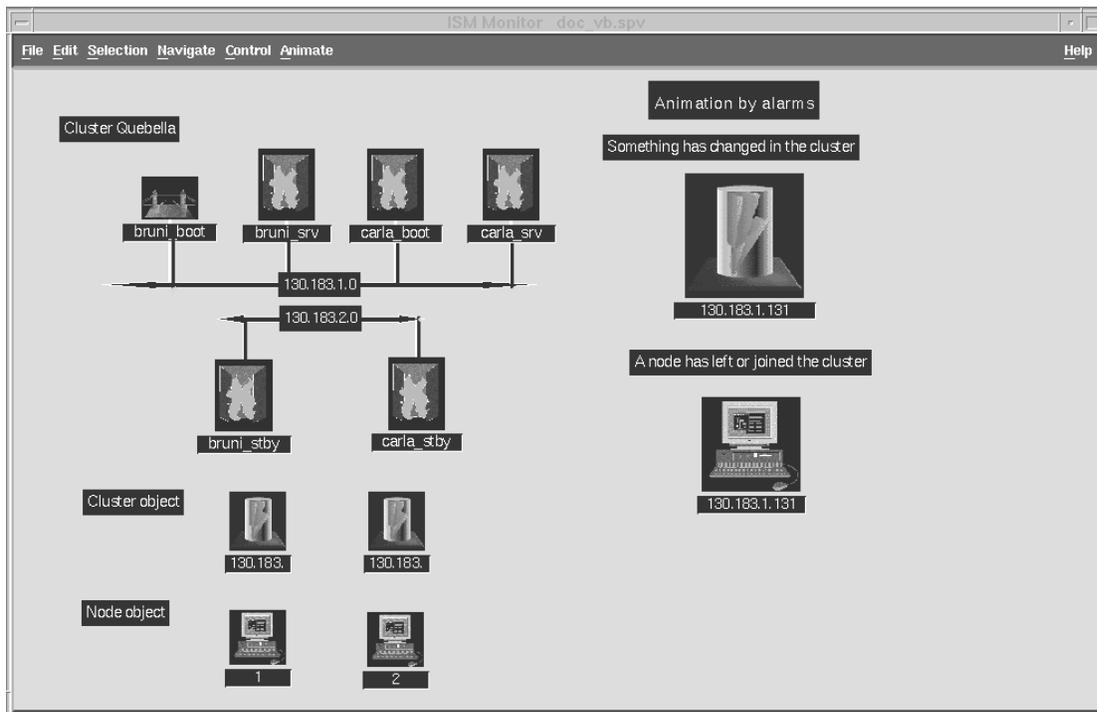
Les événements qui se produisent dans la grappe sont répercutés sur l'image par ISM. Ainsi, certains labels verts deviennent rouges pour signaler des événements. Vous pouvez ensuite consulter la base MIB HACMP (*Management Information Base*) ou afficher les interruptions reçues pour obtenir plus d'informations sur l'état de la grappe.

HACMP pour AIX contient une base MIB propre à l'entreprise, dédiée aux environnements de grappe. HACMP signale les changements d'état de la grappe via le démon pair de grappe SMUX, **clsmuxpd**, qui doit être actif sur chaque noeud de grappe. Pour contrôler une grappe, ISM se base sur la MIB HACMP.

#### Voir aussi

- La suite de ce chapitre est destinée aux administrateurs responsables de la configuration d'ISM, qui connaissent bien ISM. Les indications fournies sont d'ordre général. Pour en savoir plus sur la configuration, la personnalisation et l'exploitation d'ISM, reportez-vous à la documentation ISM.

- Pour en savoir plus sur les concepts et les utilitaires relatifs au contrôle de grappe, reportez-vous à la documentation HACMP. Vous pouvez également consulter les entrées "SNMP", "clsmuxpd" et "clinfo" du *HACMP Master Index and Glossary* (**clinfo**, fourni avec le logiciel HACMP pour AIX, est un moniteur de grappe basé sur SNMP).



Exemple d'image ISM

## Applications ISM requises

Trois modules ISM servent à configurer l'environnement de grappe : ISM Monitor, ISM IP Discovery et ISM Alarm.

### ISM Monitor

Il s'agit du principal utilitaire de contrôle des objets gérés. Il permet de :

- Créer et afficher des images - représentations graphiques des objets gérés.
- Animer les images pour refléter les changements d'état des objets gérés. Ces changements peuvent être matérialisés par des changements de couleur des icônes animées.
- Parcourir les objets et les attributs stockés dans les bases MIB concernées.

### ISM IP Discovery

Cet utilitaire permet d'exploiter les réseaux à la recherche des réseaux et unités compatibles avec le protocole IP. Vous pouvez intégrer les objets trouvés à une image, via ISM Monitor.

### ISM Alarm

Cet utilitaire permet de contrôler les alarmes, c'est-à-dire les interruptions émises par les démons SNMP (**snmpd** et **clsmuxpd**). Vous pouvez, au sein d'une image ISM, animer les icônes en fonction des interruptions. Lorsque les alarmes (c'est-à-dire les interruptions) se produisent, vous pouvez alors examiner leur contenu.

## Récapitulatif

Pour contrôler une grappe via ISM, vous devez configurer l'environnement. La procédure est la suivante :

- Configurer l'environnement SNMP sur les noeuds de grappe.
- Installer les utilitaires requis sur la station ISM Manager.
- Créer les images ISM.

Ces tâches sont décrites ci-après.

---

## Configuration de l'environnement SNMP sur les noeuds de grappe

Cette tâche doit être effectuée sur chacun des noeuds de la grappe. Elle consiste essentiellement à effectuer des vérifications.

### Définition des fichiers de configuration snmpd

Pendant l'installation de HACMP, deux fichiers de configuration, `/etc/snmpd.conf` et `/etc/snmpd.peers`, sont mis à jour, avec chaînes d'identification et mots de passe ad hoc. Pour contrôler et mettre à jour ces fichiers, procédez comme suit sur chaque noeud de grappe.

1. Editez le fichier `/etc/snmpd.conf`.

Vérifiez que la ligne suivante a été ajoutée à la fin du fichier (ajoutez-la sinon):

```
smux 1.3.6.1.4.1.2.3.1.2.1.5 clsmuxpd_password # HACMP
```

Ajoutez également, si besoin est, les lignes spécifiant les machines devant recevoir les interruptions générées par **clsmuxpd**. Ajoutez une ligne pour chaque gestionnaire ISM à affecter au contrôle de grappe. Dans l'exemple ci-dessous, deux lignes correspondant à deux stations ISM, **ism\_foo** et **ism\_bar**, ont été ajoutées - dont les adresses IP sont 130.183.1.1 et 130.183.1.51. Les deux stations **ism\_foo** et **ism\_bar** recevront les interruptions.

```
trap public 130.183.1.1 #ism_foo
trap public 130.183.1.51 #ism_bar
```

2. Editez le fichier `/etc/snmpd.peers`, et vérifiez que la ligne suivante a été ajoutée à la fin du fichier (ajoutez-la sinon) :

```
clsmuxpd 1.3.6.1.4.1.2.3.1.2.1.5 "clsmuxpd_password" # HACMP
```

3. Arrêtez et relancez les processus **snmpd** et **clsmuxpd** sur les noeuds modifiés :

```
stopsrc -s clsmuxpd
stopsrc -s snmpd
startsrc -s snmpd
startsrc -s clsmuxpd
```

## Mise à jour des fichiers `/etc/hosts` et `/etc/inittab`

### Fichier `/etc/hosts`

Pour que le démon `clsmuxpd` soit correctement enregistré avec `snmpd`, vous devez inclure `loopback` comme alias de l'entrée `127.0.0.1` dans `/etc/hosts`.

Sur chaque noeud de grappe, éditez le fichier `/etc/hosts` pour y ajouter la ligne :

```
127.0.0.1      loopback
```

**Remarque :** Le fichier `/etc/hosts` contient généralement une ligne `"127.0.0.1 localhost"`. Le cas échéant, ne la supprimez pas. (Ajoutez de toutes façons la ligne `"127.0.0.1 loopback"` comme expliqué plus haut.)

### Fichier `/etc/inittab`

Mettez à jour `/etc/inittab` (sur tous les noeuds de grappe) de sorte que les démons `clsmuxpd` et `snmpd` soient automatiquement activés au lancement du noeud.

## Vérification de l'installation des fichiers de définition MIB

La MIB HACMP est définie dans le fichier source `/usr/sbin/cluster/hacmp.my`. Sa version compilée est `/usr/sbin/cluster/hacmp.defs`.

Sur chaque noeud de la grappe, vérifiez que ces deux fichiers sont effectivement installés dans le répertoire `/usr/sbin/cluster/hacmp.defs`.

## Vérification des opérations SNMP

Pour vérifier les opérations SNMP, demandez, en local et à partir d'une station ISM, le démon `clsmuxpd` via la commande `snmpinfo`. Par exemple :

1. A partir d'un noeud de grappe, demandez en local un cliché du groupe `clsmuxpd` de l'arborescence MIB HACMP :

```
snmpinfo -o /usr/sbin/cluster/hacmp.defs -m dump -v clsmuxpd
```

La commande ci-dessus doit afficher le cliché du groupe `clsmuxpd` de l'arborescence MIB HACMP. Il est semblable à :

```
clsmuxpdGets.0 = 0
clsmuxpdGetNexts.0 = 2
clsmuxpdSets.0 = 0
nodeId.1 = 1
clsmuxpdTraps.0 = 15
...
```

2. Sur la station ISM, entrez la commande :

```
snmpinfo -h <NomNoeud> -o /usr/sbin/cluster/hacmp.defs
          -m dump -v clsmuxpd
```

Cette commande doit également afficher le cliché du groupe `clsmuxpd`.

### Remarques :

- Si rien ne s'affiche, vérifiez votre installation comme indiqué plus loin dans ce chapitre.
- Répétez cette étape de vérification sur chaque noeud de grappe.

---

# Installation des fichiers requis sur la station ISM Manager

## Installation de l'ensemble de fichiers dédiés à ISM

Le logiciel BullCluster intègre des ensembles de fichiers spécifiques dédiés à ISM. Pour les installer, reportez-vous au *Bulletin de logiciel* (SRB) fourni avec le logiciel. Ce bulletin contient des instructions d'installation et indique les éventuels prérequis.

Une fois l'ensemble de fichiers installé, installez la base MIB HACMP dans ISM comme expliqué ci-après.

### Répertoires d'installation

Sous ISM 4.54, les fichiers sont installés dans le répertoire **\$ISMROOT/var/mibspack/snmp/enterprise/ibm/hacmp**.

### Fichiers fournis

L'ensemble de fichiers dédiés à ISM fournit notamment :

- Le fichier **hacmp.mib**, version compilée de MIB HACMP, prêt à l'emploi sous ISM.
- Les fichiers **hacmp.ext/default.mnc**, qui définissent les icônes associées à chaque objet, ainsi que leur mode d'animation.
- Le répertoire **hacmp.icons**, qui contient des fichiers icône.

Pour une liste exhaustive, affichez le contenu des répertoires d'installation.

## Installation de la base MIB HACMP au sein d'ISM

**Avertissement** : Les fichiers de personnalisation de MIB HACMP se trouvent dans le répertoire **hacmp.ext/default.mnc**. Lors de l'installation de HACMP MIB, la personnalisation HACMP est automatiquement ajoutée au fichiers de configuration par défaut de l'application ISM\_Monitor, située dans le répertoire **\$ISMROOT/var/config/ISM\_Monitor/default.mnc**.

Il est recommandé d'enregistrer la configuration par défaut de ISM\_Monitor avant d'installer MIB HACMP, à l'aide de l'application ISM\_Configurator/ISM\_Monitor.

Nous supposons ici que vous avez lancé ISM à partir de la station ISM Manager. Pour installer la base MIB HACMP dans ISM (répétez la procédure sur chaque station ISM Manager prévue pour contrôler la grappe) :

1. Sur la fenêtre ISM Application Board, exécutez ISM Configurator.
2. Sur le configurateur affiché, ouvrez le menu **MIB Schema** et sélectionnez **LAUNCH**.
3. Sur la fenêtre MIB qui s'affiche, cliquez sur **Install/Remove private MIB** puis indiquez le nom de la base MIB HACMP :

hacmp

Suivez les instructions affichées (reportez-vous au besoin à la documentation ISM).

4. Une fois la base MIB installée, arrêtez ISM : sélectionnez **Exit** sur le menu **File** de Application Board, puis arrêtez le Framework par la commande **fmkstop**.
5. Relancez ISM et le Framework.

Ces étapes préliminaires accomplies, vous pouvez tracer des images ISM pour représenter et contrôler la grappe.

---

## Tracé d'une image ISM de base pour représenter une grappe

Cette section vous donne quelques conseils pour créer une image ISM permettant de télécontrôler une grappe. ISM est personnalisable à volonté selon les besoins spécifiques de chacun.

La création d'une image ISM comporte les étapes suivantes :

- Recherche des réseaux et des unités associés à la grappe
- Tracé d'une image comportant les unités découvertes
- Animation des icônes

Cette image de base constituée, vous pouvez y ajouter des détails et des critères d'animation spécifiques. Ces tâches sont décrites à partir de la section "Examen de la base MIB HACMP et ajout d'objets à l'image", page 11-12.

**Remarque :** Vous pouvez effectuer ces étapes pour vous familiariser avec les particularités de l'intégration d'une grappe dans ISM. Il vous appartiendra ensuite de déterminer vos besoins réels et d'adapter les images en conséquence.

### Recherche des réseaux et des unités associés à la grappe

Il s'agit de déclarer à ISM les noeuds de grappe, les cartes et les réseaux, via ISM IP Discovery.

#### Remarques préliminaires

##### Fonctionnement d'ISM IP Discovery

ISM IP Discovery identifie les systèmes par leur adresse IP. A chaque adresse IP correspond un objet icône. Chaque fois qu'il repère un réseau ou un système, ISM IP Discovery crée un objet (réseau ou unité IP) dans la base ISM CMIS.

##### Spécificités des grappes

Les noeuds de grappe sont dotées au minimum de deux carte réseau : une carte de service et une carte de secours, dont les adresses renvoient à deux sous-réseaux distincts. En outre, si la grappe fait appel à la fonction de basculement d'adresse IP (cas le plus fréquent), la carte de service est associée à deux adresses IP : l'adresse d'amorçage de la carte (utilisée lors de l'amorçage et jusqu'au lancement des services de grappe) et l'adresse de la carte de service (utilisée une fois les services de grappe lancés sur le noeud).

##### Adresses IP principale et secondaire

ISM IP Discovery associe une icône à chaque adresse IP repérée et lui attribue pour titre l'adresse IP en question. Aussi une carte de service se voit-elle associer deux icônes, l'une correspondant à son adresse de service et l'autre à son adresse d'amorçage.

Lorsque le processus de recherche IP est exécuté sur le réseau auquel sont connectées les cartes de service, la carte de service (adresse de service ou d'amorçage) est représentée comme **Adresse IP principale**, et la carte de secours comme **Adresse IP secondaire**.

## Procédure

1. Sur chaque noeud de grappe que vous souhaitez "découvrir", vérifiez que le démon **snmpd** est actif. Au besoin, lancez-le via la commande `startsrc -s snmpd`.
2. A partir de la station ISM, lancez l'application ISM IP Discovery. Sélectionnez l'option **Network**. IP Discovery affiche les réseaux découverts.

### Découverte des labels de carte d'amorçage des noeuds de grappe

3. Sur la liste affichée par ISM IP Discovery, sélectionnez le réseau utilisé par les cartes de service/d'amorçage des noeuds. Lancez la recherche en sélectionnant **On Selected Networks** sur le menu **Devices**.

Lorsque les services de grappe sont actifs sur un noeud, celui-ci se sert de son adresse de service (et non de son adresse d'amorçage). IP Discovery découvrira donc les adresses de service des noeuds.

La recherche terminée (ce qui peut prendre plusieurs minutes), les unités IP découvertes sont affichées : repérez celles qui correspondent aux labels de la carte de service des noeuds de grappe.

### Découverte des labels de carte d'amorçage des noeuds de grappe

4. Si vous pouvez arrêter HACMP sur les noeuds, faites-le. Ces derniers utilisent alors leur adresse d'amorçage qui peut être découverte par IP Discovery. Cette étape est facultative car les objets HACMP du démon **clsmuxpd** sont accessibles grâce aux adresses IP découvertes au cours de l'étape précédente. Elle permet toutefois d'obtenir un tracé des cartes d'amorçage sur un plan de contrôle.

### Découverte des cartes de secours des noeuds de grappe

5. Remarque : cette étape peut être effectuée, que les services de grappe soient ou non actifs sur les noeuds de grappe. A partir de ISM IP Discovery, sélectionnez le sous-réseau utilisé par les cartes de secours des noeuds (il s'agit d'un sous-réseau distinct du précédent, les cartes de service/amorçage et les cartes de secours n'utilisant pas le même sous-réseau). Lancez la recherche en sélectionnant **On Selected Networks** sur le menu **Devices**.

La recherche terminée (ce qui peut prendre plusieurs minutes), les unités IP découvertes sont affichées : repérez celles qui correspondent aux labels de la carte de secours des noeuds de la grappe.

Ces étapes exécutées, vous pouvez commencer à tracer une image représentant les unités découvertes, comme expliqué ci-après.

## Tracé d'une image comportant les unités découvertes

Il s'agit de tracer des images contenant les réseaux et les unités de grappe découverts, via ISM Monitor. Ces images serviront à télécontrôler graphiquement la grappe.

Vous allez créer une image représentant les différentes adresses des noeuds de grappe (voir exemple page 11-8). Les instructions ci-après supposent que vous traitez une grappe simple à deux noeuds. Adaptez-les à vos besoins.

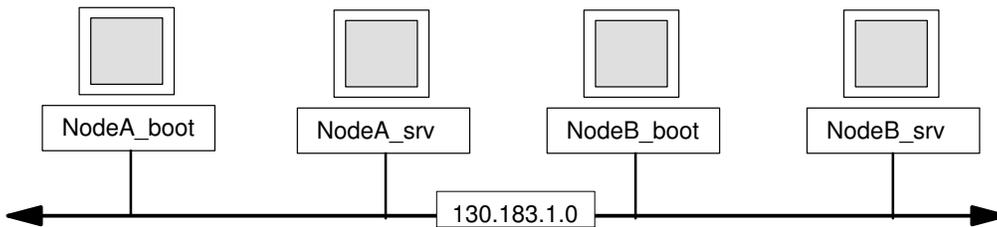
## Tracé des labels des cartes de service et d'amorçage

1. Sur la fenêtre IP Devices d'ISM IP Discovery, sélectionnez les labels de carte d'amorçage et de service. Par exemple, pour une grappe à 2 noeuds utilisant 2 labels de service, vous devez sélectionner 4 lignes : 2 labels d'amorçage et 2 labels de service. Appelez le menu **Control** et sélectionnez **Draw Selection**. Une nouvelle fenêtre s'affiche.
2. Sélectionnez **New**, de façon que la sélection soit représentée par une image nouvelle, et l'algorithme **Bus** (vous pouvez choisir d'autres options, pour ajouter les labels de carte à une image existante et/ou utiliser d'autres algorithmes de dessin).
3. Cliquez sur **Do** pour dessiner l'image. ISM Monitor crée une nouvelle image, contenant les unités IP sélectionnées. Pour l'instant, l'image ne représente que les labels de carte d'amorçage et de service.

### Remarque :

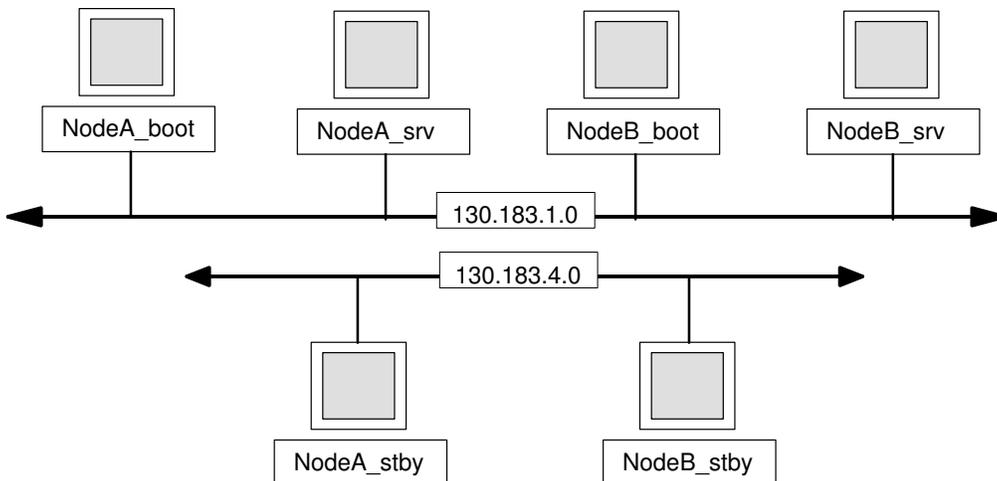
Le tracé des cartes de service ne doit être effectué que si les cartes d'amorçage n'ont pas été découvertes. Il est possible que les labels de carte soient représentés par des icônes de type "pare-feu" ou passerelle. Les noeuds de grappe ayant plusieurs cartes sur des réseaux distincts, ISM peut considérer les noeuds comme des machines "pare-feu" ou passerelle. Vous pouvez modifier ces icônes à votre guise s.

Ces tâches effectuées, votre image ressemble à :



## Tracé des labels de la carte de secours

Vous voulez sans doute ajouter les labels de carte de secours, pour que l'image devienne :



Procédez comme suit :

1. Sur la fenêtre IP Devices d'ISM IP Discovery, sélectionnez les labels de la carte de secours (c'est-à-dire, pour une grappe de 2 noeuds, vous devez sélectionner 2 lignes). Appelez le menu **Control** et sélectionnez **Draw Selection**. Une nouvelle fenêtre s'affiche.
2. ISM ne vous autorise pas à dessiner les labels de carte de secours dans l'image courante, car les mêmes hôtes sont déjà représentés par d'autres labels de carte (nous aurons à utiliser le presse-papiers, comme expliqué plus loin). Sélectionnez donc **New** pour les dessiner dans une nouvelle image. Sélectionnez également l'algorithme de dessin **Bus**.
3. Cliquez sur **Do** pour dessiner l'image. ISM Monitor affiche une nouvelle image, contenant les labels de carte de secours et de réseau sélectionnés. Vous devez à présent les copier dans l'image précédente, pour obtenir une image contenant tous les labels de carte.
4. Sélectionnez les cartes de secours et le réseau correspondant, via l'option **Select All** à partir du menu **Selection** d'ISM Monitor. Sur le menu Edit, sélectionnez **Copy** : les objets sélectionnés sont copiés dans le presse-papiers.
5. Ouvrez le presse-papiers, via l'option **Clipboard** du menu **Edit**. Chargez ensuite la première image (qui, pour le moment, ne contient que les icônes des labels des cartes d'amorçage et de service).
6. Dans le presse-papiers, sélectionnez le réseau que vous venez de copier, puis l'option **Paste** du menu **Edit**. Cliquez sur l'emplacement souhaité pour coller le réseau.
7. Répétez l'étape précédente pour coller les différents objets : labels de carte de secours et liens réseau. Vous ne pouvez pas sélectionner plusieurs objets dans le presse-papiers, d'où la nécessité de répéter l'opération. Pour les lignes réseau, notez que vous devez les coller deux fois, sur les deux objets à lier : icône de réseau et icône de label de carte.

## Animation des icônes

Cette étape consiste à animer l'icône de sorte que les changements d'état des objets représentés soient signalés par un changement de couleur de leur icône.

A partir d'une fenêtre ISM Monitor, sélectionnez toutes les icônes via l'option **Select All** du menu **Selection**. Sur le menu **Animate**, sélectionnez ensuite l'option **Live**.

## Étape suivante

Vous pouvez sauvegarder l'image. (Vous pouvez supprimer celle provisoirement créée pour les labels de carte de secours).

Nous vous engageons ensuite à lire les conseils ci-après avant de passer à "Examen de la base MIB HACMP et ajout d'objets à l'image", page 11-12.

---

## Interprétation des informations de base

### Couleur des icônes

Chaque icône de label de carte est animée en fonction de l'état des adresses IP principale et secondaire représentées. (Concernant ces adresses, lisez les remarques page 11-6.)

- Pour une icône représentant une adresse de service, l'adresse principale est l'adresse de service.
- Pour une icône représentant une adresse d'amorçage, l'adresse principale est l'adresse d'amorçage.

- Dans tous les cas, l'adresse secondaire est l'adresse de secours.

Si l'icône est :

- **verte** : les adresses principale et secondaire sont activées.
- **rouge** : les adresses principale et secondaire sont désactivées.

**Remarque** : Une icône rouge peut également signifier que le démon **snmpd** n'est pas actif que le système concerné.

## Recherche du noeud contenant une adresse de service donnée

Lorsque le basculement d'une adresse IP se produit dans une grappe, un noeud prend l'adresse d'un noeud sortant.

Lorsque vous contrôlez une grappe, vous pouvez déterminer le noeud qui porte une adresse de service donnée, en balayant la base MIB comme suit :

1. Sur l'image affichée, sélectionnez l'icône correspondant à l'adresse de service concernée.
2. Appelez le menu **Control** et sélectionnez **Detail**. Cliquez ensuite sur le bouton MIB-2 et sur le bouton système.
3. Consultez le champ **sysName** : il indique l'hôte sur lequel se trouve effectivement l'adresse de service que vous avez sélectionnée.

Dans l'exemple ci-après, le champ **sysName** indique l'hôte `liz`.

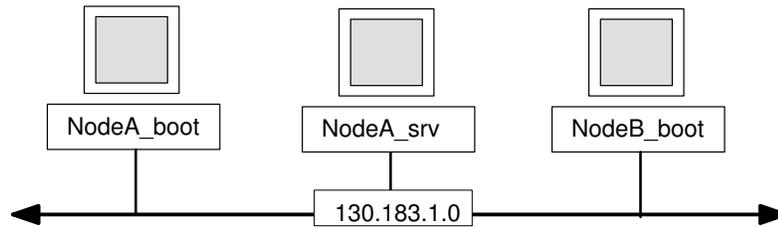
The screenshot shows a window titled 'system:systemId=taylor'. At the top, there are buttons for 'Exit' and 'Control'. Below is a table of system parameters:

<b>sysDescr</b>	Architecture RISC System/6000 Type de machine : 0x0800 ID prc Base Operating System Runtime TCP/IP Client Support version :
<b>sysObjectID</b>	1.3.6.1.4.1.2.3.1.2.1.1.3
<b>sysUpTime</b>	21:55:53.20
<b>sysContact</b>	
<b>sysName</b>	liz
<b>sysLocation</b>	
<b>sysServices</b>	7 4
<b>systemId</b>	taylor
<b>objectClass</b>	mib-2.system
<b>nameBinding</b>	mib-2.systemId
<b>packages</b>	CMIS NULL
<b>allomorphs</b>	CMIS NULL

## Exemple de scénario

Considérons un scénario impliquant une grappe 2 noeuds. Supposons qu'un groupe de ressources est défini, associé au noeud A et configuré comme groupe de secours en mode rotatif.

Soit une image contenant les objets ci-dessous :



Supposons réunies les conditions initiales suivantes :

- Le noeud A contient l'adresse de service :
  - Dans ce cas, l'icône `nodeA_srv` est verte, tandis que `nodeA_boot` est rouge.
  - Si vous affichez le détail de `nodeA_srv`, le champ **sysName** affiche le nom hôte `nodeA_boot`.
- Pour l'instant, le noeud B utilise son adresse d'amorçage. Il est prêt à prendre la relève en cas de défaillance du noeud A. L'icône `NodeB_boot` est verte.

En cas de défaillance du noeud A, un basculement d'adresse IP se produit (sous réserve que le basculement soit implémenté sur la grappe). En résultat :

- Le noeud B contient l'adresse de service :
  - L'icône `NodeA_srv` reste verte, tandis que `NodeB_boot` devient rouge.
  - Si vous affichez le détail de `NodeA_srv`, le champ **sysName** affiche le nom hôte `NodeB_boot` (et non plus `NodeA_boot`).
- L'icône `nodeA_boot` devient verte.

Le tableau ci-dessous récapitule ces changements :

	<b>NodeA_boot</b>	<b>NodeA_srv</b>	<b>NodeB_boot</b>
Avant le basculement d'adresse IP	rouge	vert (le noeud A contient l'adresse de service)	vert
Après le basculement d'adresse IP	vert	vert (le noeud B contient l'adresse de service)	rouge

---

## Examen de la base MIB HACMP et ajout d'objets à l'image

Une fois que vous avez dessiné une image de base avec les labels de carte, vous pouvez parcourir la base MIB HACMP et ajouter d'autres objets, pour finaliser votre image.

### Prérequis : vérification des opérations `clsmuxpd`

ISM obtient des informations de la grappe par l'intermédiaire du démon pair Cluster SMUX, `clsmuxpd`. Vérifiez que ce démon est actif sur chaque noeud de grappe.

#### Remarques :

- Normalement, le démon `clsmuxpd` est automatiquement activé lorsque vous lancez les services de grappe sur un noeud. Au besoin, faites appel au menu HACMP SMIT **Start Cluster Services** (raccourci : `smit clstart`).

- Vous pouvez également lancer `clsmuxpd` via la commande :

```
startsrc -s clsmuxpd
```

- Pour en savoir plus, reportez-vous à "Configuration de l'environnement SNMP sur les noeuds de grappe", page 11-3.

### Examen de la base MIB HACMP

Pour parcourir la base MIB HACMP :

1. L'image étant affichée dans une fenêtre ISM Monitor, sélectionnez une icône représentant une de carte de service.
2. Appelez le menu **Control** et sélectionnez **Detail**. Une nouvelle fenêtre s'affiche, indiquant les valeurs courantes de la base MIB SNMP, avec des boutons d'accès à un complément d'informations.
3. Cliquez sur le bouton privé puis sur `risc6000clsmuxpd` pour accéder à la base MIB HACMP. Une nouvelle fenêtre apparaît, avec un bouton pour chaque objet HACMP géré (address, application, clinfo...).
4. Cliquez sur l'un des boutons pour afficher les attributs de l'objet correspondant.

### Tracé des objets de la base MIB HACMP

Pour dessiner des objets MIB HACMP :

1. L'image étant affichée dans une fenêtre ISM Monitor, parcourez la base MIB comme expliqué ci-dessus. Notez que, **lorsque vous parcourez la base MIB dans le but de dessiner des objets, vous devez commencer par un label de carte de service** (et non d'amorçage ou de secours). Ceci parce que l'adresse de service est normalement toujours valide lorsque HACMP est actif sur la grappe.

Vous devez donc, avant de parcourir la base, sélectionner une icône représentant une carte de service.

2. Continuez à parcourir la base MIB jusqu'à atteindre l'objet que vous voulez dessiner et dont vous voulez afficher les attributs. Par exemple, pour atteindre un objet "grappe", cliquez sur le bouton **cluster**.
3. Appelez le menu **Control** et sélectionnez **Draw**, puis cliquez sur l'emplacement souhaité pour l'icône (une grappe si vous avez sélectionné un objet "grappe", par exemple).

## Objets MIB HACMP intéressants

Les labels de carte de secours, d'amorçage et de service précédemment dessinés sont des objets `snmpSystem`.

Sont également intéressants au niveau du contrôle de grappe:

- l'objet `cluster`
- l'objet `node`.

**Remarque :** Pour afficher le type d'objet associé à une icône, sélectionnez l'icône, exécutez la commande **Detail** et consultez la valeur de **objectClass**.

Pour en savoir plus, reportez-vous à "Remarques sur les objets MIB HACMP", page 11-14.

## Choix des attributs pour l'animation d'icône

Un objet étant dessiné, vous pouvez sélectionner l'option ISM **Live** pour l'animer - selon le mode d'animation par défaut (ainsi, un objet `cluster` est animé selon son attribut `clusterState`). Notez que, par défaut, un objet `cluster` est animé en fonction de son attribut `clusterState`.

Vous pouvez changer l'attribut d'animation d'un objet. Vous pouvez, par exemple, dessiner un objet `cluster` et l'animer selon l'attribut `clusterSubState`. Cet attribut se révèle très utile pour le contrôle de grappe (reportez-vous à "Remarques sur les objets MIB HACMP", page 11-14).

### Procédure

Pour créer un objet `cluster` animé selon son attribut `clusterSubState` :

1. L'image étant affichée dans une fenêtre ISM Monitor, sélectionnez une icône représentant un label de carte de service, parcourez la base MIB HACMP jusqu'à l'objet `cluster` et dessinez-le sur l'image.
2. Sélectionnez l'icône dessinée. Appelez le menu **Animate**, et sélectionnez **Animate by Attribute** pour changer de critère d'animation : animez l'objet `cluster` en fonction de son attribut `clusterSubState` (au lieu de l'attribut `clusterState` par défaut).

STATUS	VALUES
Bad	down
Warning	unknown
Good	up

STATUS	VALUES
Bad	unknown error reconfig
Warning	unstable
Good	stable

**Remarque :** Vous pouvez également animer les icônes en fonction des interruptions SNMP reçues. Voir plus loin dans ce chapitre.

## Remarques sur les objets MIB HACMP

Nous avons déjà souligné l'intérêt des objets MIB HACMP `cluster` et `node`. Cette section récapitule les informations concernant ces deux types d'objets. Les autres objets sont décrits page 11-16.

### Remarque sur le répertoire `hacmp.ext`

Lorsque vous dessinez un objet MIB HACMP, l'icône et son mode d'animation sont, par défaut, définis en fonction des paramètres de `hacmp.ext`.

Une fois familiarisé avec le contrôle de grappe via ISM, vous pourrez définir, par exemple, d'autres icônes par défaut ou d'autres critères d'animation. Utilisez l'application de configuration ISM (menus Applications/Monitor).

### Objet `cluster`

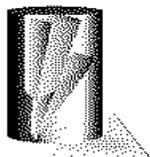
Une grappe HACMP est un groupe de processeurs coopérant pour fournir un environnement hautement disponible.

Pour l'objet `cluster`, les attributs les plus intéressants sont `clusterState` et `clusterSubState`.

#### Objets `cluster` dans une image ISM

Les deux attributs étant significatifs pour le contrôle de grappe, vous pouvez inclure dans votre image ISM deux icônes de grappe : la première animée selon `clusterState`, l'autre selon `clusterSubState`.

Pour accéder à un objet `cluster`, parcourez la base MIB HACMP et cliquez sur le bouton **cluster** (vous pouvez alors le dessiner). L'icône par défaut d'un objet `cluster`, telle que définie dans le répertoire `hacmp.ext`, est la suivante :



#### Attribut `clusterState`

Le tableau ci-dessous indique les valeurs possibles de l'attribut `clusterState` et leur signification. Il indique également les couleurs utilisées pour afficher le label d'icône, selon les définitions par défaut du répertoire `hacmp.ext`.

<code>clusterState</code>	Signification	Couleur
up	Il existe au moins un noeud actif dans la grappe et un noeud principal est défini.	vert
down	Il existe au moins un noeud actif dans la grappe, mais pas encore de noeud principal défini.	rouge
unknown	<code>clsmuxpd</code> ne peut communiquer ou ne communique pas encore avec un Cluster Manager (gestionnaire de grappe) actif.	orange

### Attribut clusterSubState

Le tableau ci-dessous indique les valeurs possibles de l'attribut `clusterSubState`.

clusterSubState	Signification	Couleur
stable	La grappe est stable (pas de reconfiguration en cours).	vert
unstable	La grappe est instable (une modification de topologie est en cours).	orange
error	Un script a échoué, la grappe est restée trop longtemps en mode configuration (instable).	rouge
unknown	<b>clsmuxpd</b> ne peut communiquer avec un Cluster Manager (gestionnaire de grappe).	rouge
reconfig	La grappe est en cours de reconfiguration.	rouge

## Objet Node

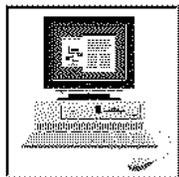
Un noeud (node) est l'un des serveurs qui constituent la grappe. Chaque noeud exécute les gestionnaires de grappe et les programmes des démons pairs SMUX et SNMP.

`nodeState` est l'attribut qui importe pour l'objet `node`.

### Objets node dans une image ISM

Vous pouvez insérer dans votre image ISM les icônes correspondant aux différents noeuds de la grappe HACMP.

Pour accéder aux objets `node`, parcourez la base MIB HACMP et cliquez sur le bouton **node**, puis "détaillez" l'instance du tableau et cliquez sur le bouton **node Entry**. Un tableau des noeuds de la grappe que vous pourrez tracer dans l'image s'affiche. L'icône par défaut d'un objet `node` est la suivante :



### Attribut nodeState

Le tableau ci-dessous indique les valeurs possibles de l'attribut `nodeState` et leur signification. Il indique également les couleurs utilisées pour afficher le label d'icône.

nodeState	Signification	Couleur
up	Le noeud est actif.	vert
down	Le noeud est inactif.	rouge
joining	Le noeud est en passe de rejoindre la grappe.	orange
leaving	Le noeud est en passe de quitter la grappe.	orange

## Autres objets MIB HACMP

La base MIB HACMP maintient des objets autres que `cluster` et `node` (application, clinfo, cllockd...).

Ces objets n'intervenant quasiment pas dans le contrôle de grappe, nous n'entrerons pas dans le détail. Il vous appartient de déterminer, en fonction de vos besoins spécifiques, s'ils présentent un intérêt pour vous.

Pour en savoir plus sur ces objets, vous pouvez consulter le fichier **hacmp.mib** (qui est commenté), ou parcourir la base MIB HACMP à partir d'une image ISM. Les différents objets maintenus dans la base sont les suivants :

address	Attributs des adresses réseau des noeuds (indique, par exemple, le rôle de l'adresse : secours, amorçage ou adresse).
application	Informations sur les applications enregistrées avec la grappe par SMUX HACMP.
clinfo	Informations sur le démon <b>clinfo</b> .
cllockd	Informations sur le démon <b>cllockd</b> (cluster lock manager).
clsmuxpd	Informations statistiques sur le démon <b>clsmuxpd</b> (cluster SMUX Peer).
clstrmgr	Informations sur le démon <b>clstrmgr</b> .
cluster	Attributs d'une grappe.
event	Informations sur les 1000 derniers événements de grappe.
network	Attributs des réseaux prenant en charge la grappe.
node	Attributs des noeuds d'une grappe.

---

# Traitement des interruptions

## Prérequis

### clsmuxpd

Le démon pair SMUX de grappe, **clsmuxpd**, doit être actif sur chaque noeud de grappe. En cas de doute, vérifiez-le comme indiqué page 11-12.

Lorsqu'un changement intervient au sein d'une grappe, le gestionnaire de grappe en avertit le démon **clsmuxpd**. Celui-ci émet alors des interruptions pour diffuser l'information. Pour que le gestionnaire de station ISM puisse recevoir ces interruptions, le fichier **/etc/snmpd.conf** doit être défini en conséquence sur chaque noeud de grappe. Au besoin, reportez-vous à "Définition des fichiers de configuration snmpd", page 11-3.

### clinfo

Si vous prévoyez d'activer l'utilitaire HACMP clinfo sur la station ISM Manager, sachez que vous ne pouvez pas l'affecter à la réception des interruptions. Autrement dit, n'activez pas **clinfo** avec l'option **-a**, faute de quoi vous risquez de provoquer des conflits entre ISM et **clinfo** sur la station ISM.

Ceci est dû au fait qu'un seul gestionnaire SNMP à la fois peut être actif sur une station réseau donnée (un seul programme TCP/IP peut "écouter" un port donné).

## Remarques sur les interruptions HACMP

ISM peut traiter les interruptions émises par les démons **clsmuxpd**. Vous pouvez décider d'animer les icônes uniquement en fonction des interruptions HACMP reçues par ISM. L'application "ISM Alarm" permet d'afficher ces interruptions.

Utilisez ISM Configurator (menu de personnalisation des interruptions) pour afficher les définitions d'interruption HACMP.

## Alarmes : remarque de terminologie

Dans la terminologie ISM, les interruptions sont appelées "alarmes".

## Les interruptions HACMP

Le démon **clsmuxpd** peut générer les interruptions suivantes (l'objet associé à l'interruption apparaît entre parenthèses) :

TrapClusterState	(clsmuxpd)
TrapClusterSubState	(clsmuxpd)
TrapNodeState	(node)
TrapNetworkState	(network)
TrapAddressState	(address)
TrapNewPrimary	(clsmuxpd)
TrapAppState	(application)
TrapAdapterSwap	(address)
TrapAddressTakeover	(address)

Toutes ces interruptions ont un niveau de gravité "important".

Ces interruptions signalent un changement d'état des principaux objets maintenus dans la base MIB HACMP. Les interruptions les plus significatives sont :

- TrapAddressTakeover, qui indique qu'un basculement IP s'est produit.
- TrapAdapterSwap, qui indique qu'une commutation de carte a eu lieu.

Il est recommandé d'affecter à ces deux interruptions le niveau de gravité le plus élevé.

Il existe 27 autres interruptions HACMP dont le niveau de gravité est indéterminé et qui sont toutes associées à l'objet `node`. La liste ci-dessous en répertorie quelques-unes liées à un échec de la grappe HACMP, dont le niveau de gravité est "critique".

```
trapFailNetwork
trapFailNode
trapClusterConfigToolong
trapClusterUnstableToolong
trapEventError
```

Toutes les interruptions HACMP ont été associées à un incident spécifique. Il s'agit d'un OID portant l'alias du texte supplémentaire HACMP. L'incident permet de filtrer chaque alarme. Par exemple, pour animer une icône qui deviendrait rouge dès que le basculement de carte se produit.

Pour afficher les paramètres personnalisés des interruptions HACMP (niveau de gravité des alarmes et incidents spécifiques), utiliser les outils ISM Configurator Trap Custom et Alias Server.

## Affichage des interruptions reçues

### Affichage de la liste des alarmes

Pour afficher les interruptions (alarmes) reçues par ISM :

1. Dans la fenêtre ISM Application Board, sélectionnez l'application ISM Alarm.
2. Appelez le menu **Execute** et sélectionnez l'option **from file....** Une liste de fichiers de configuration s'affiche, parmi lesquels le fichier **default.ALcfg**.
3. Sélectionnez le fichier **default.ALcfg** et cliquez sur **OK** pour le charger. Une fenêtre Alarm affiche le tableau des alarmes reçues par ISM Manager. Chaque ligne correspond à un ensemble d'occurrences de la même alarme émise par le même système.

### Modification de la présentation par défaut de la liste des alarmes

Les colonnes **EventType** et **ProblemType** apparaissent en premier (à gauche du tableau). Mais toutes les interruptions HACMP sont de type (**EventType**) `processing` et `sfwrEnvironmental` (**ProblemType**) - ce qui est défini dans le fichier `hacmp.trapdb`. En conséquence, les informations de ces deux colonnes présentent peu d'intérêt et vous voudrez sans doute changer de présentation par défaut.

Il est recommandé de modifier la présentation en plaçant la colonne **ProblemText** à gauche du tableau. L'affichage de **ProblemText** dans le(s) fichier(s) de configuration permet de distinguer facilement les interruptions HACMP de celles qui sont émises par d'autres agents SNMP, `processing` / `sfwrEnvironmental` par exemple.

Pour réorganiser les colonnes, sélectionnez l'option **Preferences...** du menu **Control** de la fenêtre ISM Alarm.

### Génération d'interruptions HACMP pour test

Pour tester les caractéristiques décrites ci-dessus, il suffit de provoquer l'émission d'interruptions HACMP.

Pour ce faire, faites une modification au niveau de l'état de la grappe : déconnectez une carte de service ou arrêtez les services de grappe sur un noeud, par exemple. Au bout de quelques secondes, les interruptions HACMP s'affichent dans la fenêtre ISM Alarm.

## Animation des images ISM en fonction des interruptions reçues

Vous pouvez finaliser votre image en y ajoutant des objets animés en fonction des interruptions HACMP reçues.

Voici une procédure type.

1. Sur l'image ISM, dessinez une nouvelle icône, un objet de type `cluster`, par exemple. Sélectionnez-le.
2. Appelez le menu **Animate** d'ISM Monitor, puis le sous-menu **Animate By**. Sur le menu qui s'affiche, activez la case **Alarms** et désactivez la case **Attribute**.
3. Rappelez le menu **Animate**. Sélectionnez successivement **Animate By**, **Configure** et **Alarms**. Un panneau de configuration s'affiche.
4. Sur ce panneau, sélectionnez les alarmes qui doivent animer l'icône sélectionnée (un objet `cluster`, dans notre exemple). Vous pouvez spécifier différents critères de filtrage.

Par exemple, pour ne prendre en compte que les interruptions HACMP issues d'une grappe donnée, appelez l'éditeur de liste Source Object et ajoutez des lignes de la forme :

```
sysNameId=_node_nom/risc6000clsmuxpdId=node_nom/*
```

Ainsi, si la grappe concernée contient les deux noeuds **node\_foo** et **node\_bar**, ajoutez les lignes :

```
snmpSystemId=node_foo/privateId=node_foo/risc6000clsmuxpdId=node_name/*
snmpSystemId=node_bar/privateId=node_bar/risc6000clsmuxpdId=node_name/sysNameId=node_foo/risc6000clsmuxpdId=node_foo/*
sysNameId=node_bar/risc6000clsmuxpdId=node_bar/*
```

**Remarque :** Un moyen simple de déterminer le *nom\_noeud* consiste à localiser les alarmes HACMP dans une fenêtre ISM Alarm (voir l'aide en ligne ISM Alarm).

Pour animer l'icône associée à l'occurrence d'une interruption HACMP ou un sous-ensemble d'interruptions HACMP, utilisez la fonction de filtre Specific Problem filtering.

5. Animez l'icône : appelez le menu **Animate** et sélectionnez **Live**. Le label de l'icône passe au rouge ou au vert.

### Couleur du label d'icône

- Si le label passe au vert, aucune interruption HACMP n'a été reçue jusqu'ici.  
Si vous générez des interruptions HACMP (pour le test) en modifiant la grappe, le label passe au rouge dans les secondes qui suivent.

- Si le label passe au rouge, des interruptions HACMP ont été reçues.

Pour afficher les interruptions, sélectionnez l'icône `cluster` concernée, appelez le menu associé et sélectionnez **ISM Alarm**. Une fenêtre ISM Alarm s'ouvre, affichant toutes les interruptions HACMP émises par les démons **clsmuxpd** actifs sur les différents noeuds de grappe.

## Contrôle des erreurs liées au disque

Pour contrôler les erreurs liées au disque, utilisez l'agent **alixd** d'ISM UNIX. Lorsqu'une erreur se produit, ce dernier génère automatiquement des interruptions dans le journal des erreurs AIX.



---

# Chapitre 12. Intégration d'un tiroir PCI

Ce chapitre indique comment intégrer un tiroir PCI dans un rack EPC Escala au sein d'une configuration classique.

---

## Intégration d'un tiroir PCI

### Exemple de configuration

Soit un noeud de grappe avec trois interfaces Ethernet (maintenance, secours, administration) et deux bus SCSI-DE.

La configuration de base se présente comme suit :

Ethernet en0	Interface de maintenance
Ethernet en1	Interface de secours
Ethernet en2 (native)	Interface d'administration
scsi0	
scsi1	

Les adaptateurs **Ethernet de secours** et **scsi1** doivent être déplacés et installés dans le tiroir PCI.

**Remarque :** Pour éviter que les définitions d'interface ne "polluent" la base de données ODM et rendre possible la création d'une interface supplémentaire (ent3..., scsi2), il faut retirer toutes les interfaces Ethernet et disque avant de lancer la procédure d'intégration.

### Procédure d'intégration

Effectuez les étapes suivantes :

1. Arrêtez HACMP sur le noeud qui doit recevoir le tiroir PCI, avec "basculement" sur l'autre noeud, afin que le service ne soit pas interrompu. Invalidez le redémarrage automatique d'HACMP au réamorçage.
2. Entrez la commande suivante :

```
lsattr -E -l scsi1
```

puis notez le numéro ID scsi de l'interface.
3. Arrêtez les interfaces Ethernet et disque et déconfigurez-les comme suit :
  - a. `smit chinet`  
sélectionnez en <x> pour l'interface en0, en1, en2  
Current STATE [detach]
  - b. `rmdev -dl en <x>` pour l'interface en0, en1, en2
  - c. `rmdev -dl et <x>` pour l'interface et0, et1, et2
  - d. `rmdev -dl ent <x>` pour les interfaces ent0, ent1, ent2
  - e. `rmdev -dl hdisk <x>` pour hdisk x, y ...
  - f. `rmdev -dl atf <x>` pour atf x, y ...
  - g. `rmdev -dl sp <x>` pour sp x, y ...
  - h. `rmdev -dl scsi 1` (pour l'adaptateur SCSI à enlever).

Ne déconfigurez pas l'adaptateur scsi0, car il doit rester en place. Sinon, n'oubliez pas de mettre à jour son numéro ID scsi après avoir isolé le bus correspondant, comme décrit ci-dessus.

4. Arrêtez le noeud : `shutdown -F` et mettez-le hors tension.

5. Déconnectez le bus SCSI de l'interface à retirer (scsi 1), puis débranchez les câbles Ethernet des adaptateurs.
6. Enlevez les adaptateurs appropriés et installez l'adaptateur de l'interface dans le tiroir PCI.
7. Installez le tiroir PCI dans le rack.
8. Branchez tous les câbles Ethernet, puis reliez un adaptateur de terminaison au connecteur arrière de l'adaptateur SCSI installé dans le tiroir PCI.
9. Déclarez le tiroir PCI (à l'aide du menu 9 du menu "sbb"). La réponse attendue doit ressembler à la suivante.

```
CONFIGURATION in progress
WAIT #
UNIT FOUND : 1
```

10. Mettez sous tension le noeud et son tiroir PCI, puis réamorcer AIX.
11. Reconfigurez l'interface Ethernet à l'aide de **smit chinet**

sélectionnez en <x>

x = 0 pour l'interface d'amorçage

x = 1 pour l'interface de secours

x = 2 pour l'interface d'administration

```
INTERNET ADDRESS      [@IP]
Network MASK          [défini par l'utilisateur]
Current STATE         [up]
```

12. Mettez à jour la table de routage, si nécessaire, à l'aide de **smit mkroot** :

```
DESTINATION Address   [default]
GATEWAY Address       [@IP_gateway]
METRIC                 [défini par l'utilisateur]
Network MASK          [défini par l'utilisateur]
```

13. Pour l'interface SCSI, remplacez l'ID scsi par la valeur fournie à l'étape 2 à l'aide de **smit chgscsi** :

sélectionnez scsi1

```
Adapter card SCSI ID  [x]
```

14. Déconfigurez SCSI, atf and sp, comme suit :

a. `rmdev -dl atf<x>`

b. `rmdev -dl sp<x>`

15. Arrêtez la machine : `shutdown -F` et mettez-la hors tension.

16. Connectez le bus SCSI commun à la place de l'adaptateur de terminaison.

17. Mettez sous tension le noeud et réamorcer AIX.

18. Vérifiez que les éléments suivants sont présents :

```
disques
pilotes de disque array
groupe de volumes
pv -id correspondants
réseau
```

à l'aide des commandes suivantes :

```
lsdev -Cc disk      pour vérifier "hdisk x, y, ..."
lsdev -Cc array     pour vérifier "sp x, y, ..."
lsdev -Cc driver    pour vérifier "atf x, y, ..."
lspv                pour vérifier pvid et leurs groupes de volumes associés
netstat -r          pour vérifier l'interface et la passerelle Ethernet
```

19. Lorsque vous aurez constaté que tous ces éléments fonctionnent correctement, lancez HACMP sur ce noeud.

---

# Annexe A. Mise en oeuvre des solutions console

Cette annexe traite des différentes solutions qui peuvent être mises en oeuvre pour gérer les consoles des noeuds de grappe. Il comporte les sections suivantes :

- Présentation des solutions console.
- Solutions console pour solutions HA, page A-2.
- Solutions console pour Escala Powerclusters (y compris le PowerConsole), page A-5.
- Configuration et utilisation du concentrateur de console, page A-13.
- Configuration et exploitation de l'utilitaire pwcons, page A-19.

---

## Présentation des solutions console

Un noeud de grappe, comme tout serveur UNIX, requiert une console système. Si votre grappe comporte 4 noeuds, 4 consoles sont utilisées dans la grappe.

Pour gérer les consoles des noeuds d'une grappe, vous avez le choix entre plusieurs solutions, basées sur les différents composants matériels.

Le plus simple consiste à raccorder un terminal ASCII à chaque noeud. Toutefois, si la grappe possède plusieurs noeuds, vous pouvez juger que tous ces terminaux encombrant l'espace de votre salle informatique. En outre, utiliser plusieurs terminaux pour contrôler une grappe (qui constitue, du point de vue administratif, une entité unique) peut se révéler peu pratique. De plus, les terminaux ASCII sont peu attrayants et nombre d'administrateurs système leur préfèrent des écrans graphiques.

C'est pourquoi d'autres solutions console sont proposées.

### Solutions console pour solutions HA

Pour les solutions HA, les solutions proposées pour la gestion des consoles sont :

- Solution console traditionnelle, reportez-vous page A-2.
- Solution basée sur une carte multiport série, reportez-vous page A-3.

### Solutions console pour Escala Powerclusters

Pour gérer les consoles des noeuds d'un Powercluster, vous avez le choix entre plusieurs solutions, basées sur les différents composants matériels. La solution la plus puissante, appelée PowerConsole, est sur une station de travail AIX et un concentrateur de console. La station de travail AIX peut être une plate-forme Escala série S100. Le PowerConsole et son logiciel graphique ClusterAssistant associé offrent à l'administrateur de grappe un point de contrôle unique très intéressant. Il existe encore d'autres solutions possibles.

- Pour en savoir plus, reportez-vous à "Solutions console pour Escala Powerclusters", page A-5.

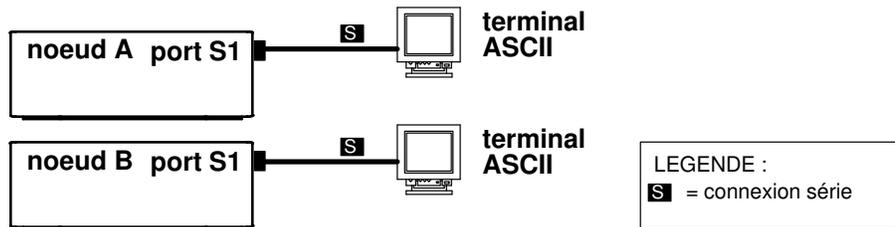
---

# Solutions console pour solutions HA

## Solution console traditionnelle

Avec cette solution, chaque noeud est directement connecté à sa propre console. Chaque console est généralement un terminal ASCII relié au port console S1 du noeud à contrôler. La mise en oeuvre de cette solution console est identique à l'installation de plusieurs serveurs indépendants.

La figure ci-dessous illustre un exemple de solution console traditionnelle :



## Variante de la solution console traditionnelle

### Consoles graphiques

Certaines solutions HA sont basées sur des serveurs équipés de consoles graphiques et non de terminaux ASCII. Les serveurs traitent alors les entrées et sorties via la console graphique et le port console S1 est inutilisé.

### Ordinateurs PC

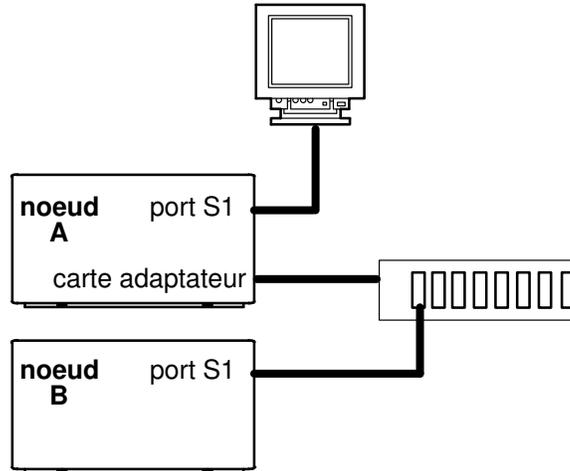
A la place de terminaux ordinaires (ASCII ou consoles graphiques), vous pouvez opter, comme consoles système, pour des ordinateurs PC connectés aux ports série S1 des noeuds de grappe. Ensuite, sur l'écran du PC, vous pouvez afficher plusieurs fenêtres d'émulation de terminal. Chaque fenêtre fait office de console système d'un noeud. Vous disposez des programmes d'émulation de terminal livrés avec MS-Windows 3.1 et avec MS-Windows 95, ou d'autres programmes de fournisseurs tiers.

Cette solution console basée sur PC peut être mise en oeuvre via le module "PC Operator's Console Kit", qui comporte les câbles et la documentation adéquats. Pour en savoir plus, reportez-vous au manuel *PC Operator's Console Facility User's Guide* (86 A2 94AT).

## Solutions basées sur une carte multiport série

Si vous souhaitez limiter le nombre de terminaux console de votre configuration, vous pouvez définir une configuration dans laquelle un ou deux terminaux gèrent plusieurs consoles. Dans une configuration de ce type, un des noeuds est équipé d'une carte série 8 ports, laquelle gère les ports console S1 des autres noeuds. Les deux figures ci-dessous illustrent des exemples de ce type de configuration.

**Exemple 1 :** Exemple de configuration semblable à la première, mais prévue pour une grappe à 2 noeuds.

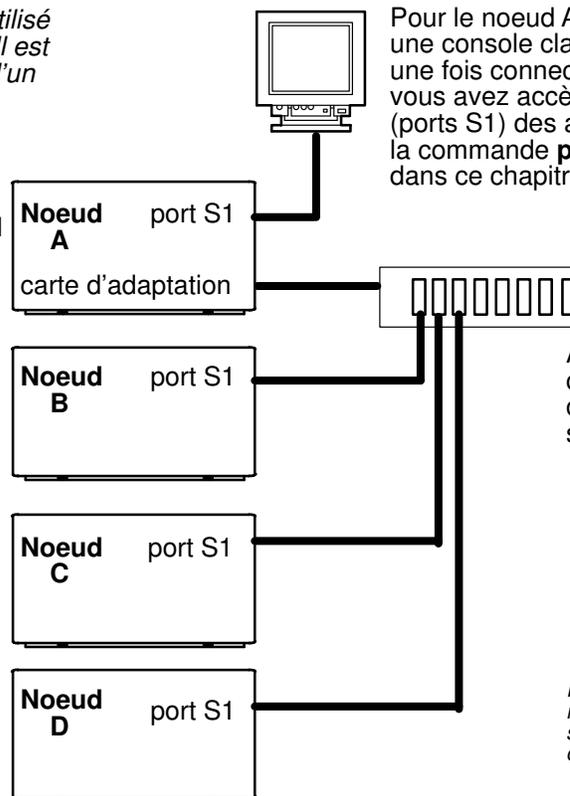


**Exemple 2 :** la figure suivante représente une configuration à 4 noeuds et un seul terminal.

*Un seul terminal est utilisé pour toute la grappe. Il est connecté au port S1 d'un noeud.*

*Un seul noeud bénéficie d'un terminal associé. Ce noeud est équipé d'un adaptateur 8 ports.*

*Les autres noeuds ne sont pas associés à un terminal. Leurs ports S1 sont connectés à l'adaptateur 8 ports du noeud auquel est associé un terminal.*



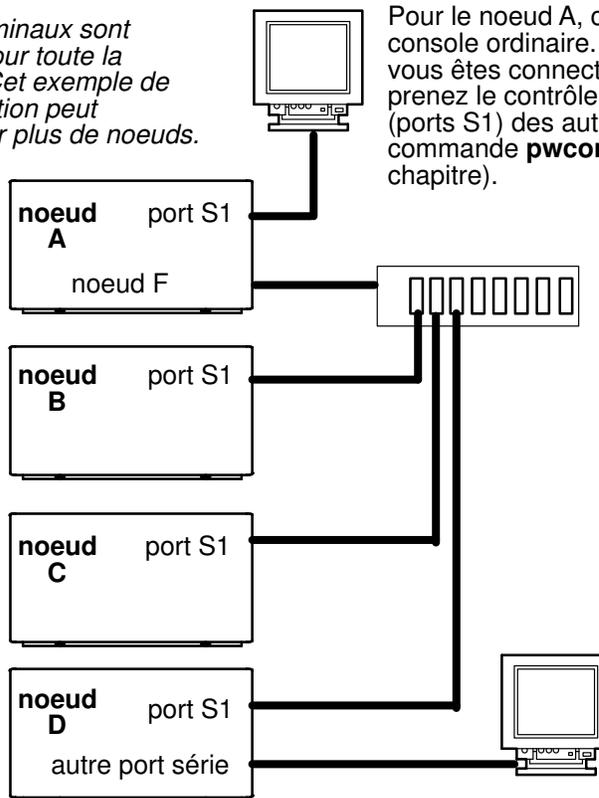
Pour le noeud A, ce terminal est une console classique. En outre, une fois connecté au noeud A, vous avez accès aux consoles (ports S1) des autres noeuds (voir la commande **pwcons**, plus loin dans ce chapitre).

Adaptateur 8 ports connecté à sa carte d'adaptation sur le noeud A.

*Remarque : Sur tous les modèles Escala, le port série S1 est dédié aux connexions console.*

**Exemple 3** : la figure suivante représente la même configuration que ci-dessus exceptée l'utilisation de deux terminaux physiques.

*Deux terminaux sont utilisés pour toute la grappe. Cet exemple de configuration peut comporter plus de noeuds.*



Pour le noeud A, ce terminal est une console ordinaire. En outre, dès que vous êtes connecté au noeud A, vous prenez le contrôle des consoles (ports S1) des autres noeuds (voir la commande **pwcons**, plus loin dans ce chapitre).

Un noeud (appelé "noeud de référence"), ici le noeud A, est équipé d'un terminal et d'une carte multiport qui répartit les liaisons vers les ports S1 des autres noeuds.

Le second terminal est raccordé à un autre noeud quelconque.

Le second terminal est utilisé comme console ordinaire par le noeud D. Il ne peut servir à accéder aux ports S1 des autres noeuds.

## Avantages et inconvénients des solutions basées sur une carte série multiport

Les configurations où un seul terminal gère plusieurs consoles (comme dans les trois exemples ci-dessus) permettent de gagner de l'espace, et d'économiser quelques mètres de câbles. Le choix d'une configuration de ce type semble donc judicieux.

Toutefois, sur des configurations où chaque noeud n'est pas associé à un terminal physique, ce choix présente quelques inconvénients. En particulier, si le noeud auquel est reliée la carte multiport vient à faillir, vous ne pouvez plus l'exploiter pour accéder aux ports S1 (ports console) des autres noeuds : vous devez, le cas échéant, connecter physiquement un terminal au port S1 concerné. Ce qui n'est pas très simple. Par ailleurs, il est rarement nécessaire en pratique d'exploiter un système à partir de sa console.

## Utilitaire pwcons

Lorsque les consoles sont gérées par le biais d'une carte multiport, **pwcons** (un utilitaire livré avec le module BullCluster), peut être utilisé à partir du terminal unique pour contrôler les consoles (ports S1) des différents noeuds. Pour en savoir plus, reportez-vous à "Configuration et exploitation de l'utilitaire **pwcons**", page A-19.

---

# Solutions console pour Escala Powerclusters

## Composants des solutions console pour Escala Powerclusters

Pour gérer les consoles des noeuds d'une grappe, vous avez le choix entre plusieurs solutions, basées sur les différents composants matériels, rappelés ci-après.

- Tout Powercluster est livré avec un terminal ASCII conventionnel (modèle BQ306), appelé **console système**, voir page A-5.
- Le terminal ASCII peut être remplacé par un **écran graphique**, voir page
- Outre ce terminal ASCII, suivant les options que vous avez commandées, les configurations Powercluster peuvent comporter :
  - soit un terminal X, appelé **console de grappe**, voir page **A-6**,
  - soit une station de travail Escala série S100 appelée **PowerConsole**, voir page A-6.
- Un **concentrateur de console** (serveur de communication 3Com CS/2600) peut également être utilisé :
  - il est fourni en standard avec les configurations de plus de deux noeuds, et avec les configurations 2 noeuds équipées d'une station de travail PowerConsole, voir page A-7.
- Un **concentrateur (hub) Ethernet** (Hub 10 3Com SuperStack II TP 12 ports) peut être fourni avec toute configuration Powercluster, pour implémenter un réseau d'administration dédié, voir page A-7.

## Rôle des composants des solutions console

Nombre de configurations peuvent être réalisées via les composants des solutions console. Voici quelques indications sur leur utilisation possible. Il vous appartient de déterminer la configuration correspondant à vos besoins. A titre d'exemple, reportez-vous aux configurations illustrées ci-après.

### Console système (terminal ASCII)

La **console système** (terminal ASCII) est disponible dans les configurations en grappe suivantes :

- Si la configuration intègre un concentrateur de console, le terminal ASCII peut servir de console de secours. Dans ce cas, maintenez le terminal connecté au concentrateur. Si le concentrateur vient ensuite à faillir (ce qui entraîne la perte de l'accès à tous les ports de console S1), ou qu'une panne se produit au niveau de la console graphique, vous pouvez connecter provisoirement le terminal de secours au port console S1 d'un noeud, et l'utiliser comme console jusqu'à ce que l'unité défailante soit réparée.
- Si la configuration est dépourvue de concentrateur de console (voir ci-dessous), le terminal ASCII peut être utilisé comme console système ordinaire de l'un des noeuds de grappe. Dans ce cas, connectez-le port console S1 du noeud concerné.
- L'exemple 1, page A-3, est une solution de base applicable à un powercluster à 2 noeuds sans console de grappe ni PowerConsole.

## Ecran graphique

Un **écran graphique** peut être utilisé comme une alternative à la console système dans les configurations en grappe suivantes :

- Dans une Powercluster à un seul noeud, il peut directement remplacer une console système.
- Dans une Powercluster à deux noeuds, une console système est connectée à un noeud, et un écran graphique au second. C'est la solution mise en oeuvre dans l'architecture avec **reprise après incident**.

**Remarque** : Aucun écran graphique n'est disponible pour un EPC800.

## Console de grappe (terminal X)

La **console de grappe** (terminal X) est équipée d'un port série qui peut être exploité à partir d'une fenêtre d'émulation de terminal, donnant ainsi accès aux consoles système. Le terminal X peut être amorcé à partir de sa propre mémoire résidente (cette capacité accroît sa disponibilité, car le terminal X n'a pas à télécharger son logiciel via le réseau et il peut être lancé même si les ressources réseau ne sont pas disponibles). Le terminal X peut être utilisé comme suit :

- Il peut être directement raccordé via son port série au port console S1 d'un noeud.
- Dans une mise en oeuvre à deux noeuds, un **hub d'administration** est exigé.
- Si la configuration intègre un **concentrateur de console** (voir ci-dessous), le terminal X peut être connecté via son port série au concentrateur, lequel à son tour assure la gestion des ports console S1 des différents noeuds.

Dans les deux cas, il faut également connecter le terminal X au réseau, pour bénéficier des fonctions de multifenêtrage et de calcul inhérentes au réseau. Ce peut être soit *réseau d'administration* dédié, installé via le concentrateur fourni, soit le réseau d'entreprise.

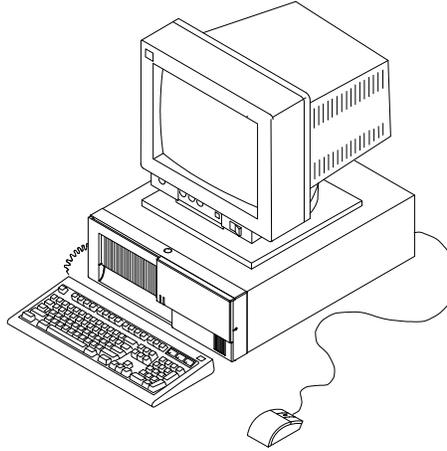
En l'absence d'un hub d'administration (pas de réseau d'administration dédié), le concentrateur de console et la console de grappe doivent être connectés au réseau d'entreprise.

## PowerConsole

Le **PowerConsole**, disponible en option avec Powerclusters, est une station de travail AIX qui fournit un point de contrôle unique pour la gestion des grappes. Il est normalement utilisé en conjonction avec une console.

La station de travail sert à la fois de point d'accès (via le concentrateur) aux différentes consoles système, et de puissante station de travail capable d'exécuter n'importe quel utilitaire d'administration. Le PowerConsole est livré avec ClusterAssistant, un ensemble d'applications graphiques exécutées sur PowerConsole pour configurer, gérer et contrôler les grappes. ClusterAssistant est intégré à CDE (Common Desktop Environment), qui propose un environnement basé sur la trilogie fenêtres /icônes /souris. Les opérations en deviennent plus simples et plus "intuitives".

## Station de travail Escala série S



### Concentrateur de console

Le **concentrateur de console** (serveur de communication 3Com CS/2600) permet de regrouper les ports de console S1 de tous les noeuds, de sorte qu'ils sont ainsi accessibles à partir d'un seul point d'accès. Ce qui évite de devoir disposer d'un terminal ASCII par noeud de grappe - impératif lourd pour les grappes comportant de plus de deux noeuds. A défaut, le PowerConsole est connecté au réseau public du client, pour accéder au noeud pour la gestion.

### Concentrateur (hub)

Le **concentrateur (Hub 10 3Com SuperStack II TP 12 ports)** offre douze ports 10Base-T (paire torsadée). Il est principalement conçu pour implémenter un *réseau d'administration* dédié, qui interconnecte les unités concernant la console (concentrateur de console, console de grappe, PowerConsole) et les noeuds de grappe. Disposer d'un réseau d'administration indépendant est une garantie de sécurité.

## Exemples de configuration console pour PowerClusters

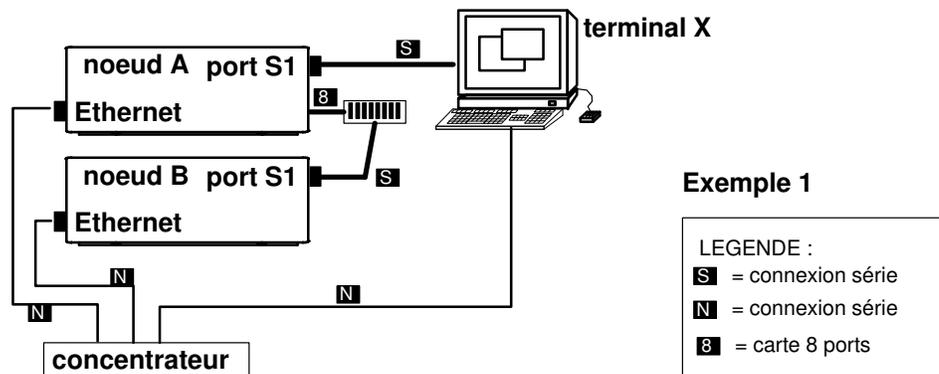
Cette section présente quelques configurations classiques. Notez qu'il s'agit de simples exemples, et qu'il existe bien d'autres configurations possibles.

Pour connaître des configurations de console réelles, reportez-vous au manuel *EPC Guide de câblage*.

### Exemple 1 (grappe 2 noeuds, console de grappe)

La figure ci-dessous décrit une grappe à deux noeuds (A et B), dont les consoles sont gérées via une console de grappe (terminal X). Dans cette configuration :

- Seul le noeud A dispose d'un terminal directement connecté. Le noeud A est équipé d'une carte série 8 ports, et le port S1 du noeud B est connecté à cette carte 8 ports. Remarque : les configurations basées sur une carte 8 ports sont également décrites à la section "Solutions console basées sur une carte multiport", page A-3).
- Le contrôle de la console du noeud A est effectué via la ligne série et la fenêtre d'émulation de terminal du terminal X.
- Une fois connecté au noeud A (via la ligne série), vous pouvez prendre le contrôle de la console du noeud B, via la commande `pwcons`. Cette commande est décrite à la section "Configuration et exploitation de l'utilitaire pwcons", page A-19.
- Pour une sécurité maximale, nous avons choisi d'utiliser le concentrateur pour implémenter un *réseau d'administration* indépendant, qui interconnecte les noeuds et le terminal X. A partir de celui-ci, vous avez accès aux deux noeuds via le réseau, et vous pouvez à volonté exécuter les utilitaires d'administration, en profitant de la fonction de multifenêtrage du terminal X.

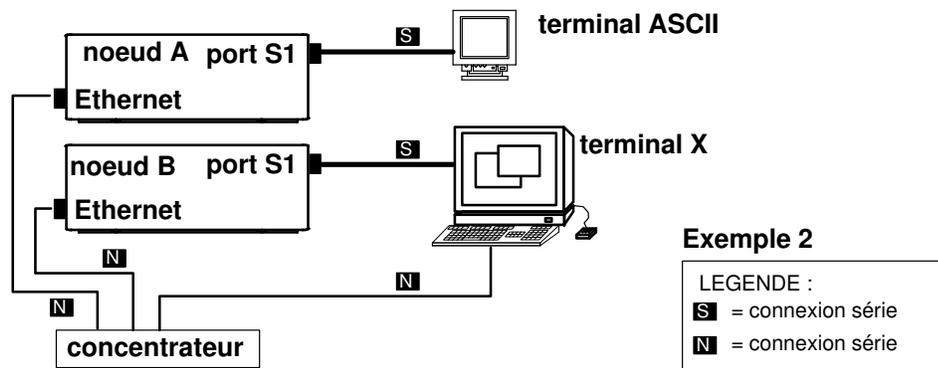


### Exemple 2 (variante de l'exemple 1)

La figure ci-dessous illustre une variante de l'exemple 1, où le terminal ASCII est utilisé. Il illustre une grappe à deux noeuds (A et B), dont les consoles sont gérées par une console système (terminal ASCII) et une console de grappe (terminal X). Dans cette configuration :

- Le terminal ASCII permet de contrôler la console (port S1) du noeud A.
- Le terminal X, via une ligne série et une fenêtre d'émulation de terminal, permet de contrôler la console du noeud B.

- Comme dans l'exemple 1, le concentrateur est utilisé pour implémenter un *réseau d'administration* indépendant, qui interconnecte les noeuds et le terminal X.



### Exemple 3 - PowerConsole (grappe de N noeuds, station de travail AIX et concentrateur de console)

La figure ci-dessous décrit une grappe à trois noeuds (A, B et C), dont les consoles sont gérées via un concentrateur de console CS/2500 et une station de travail AIX (PowerConsole). Notez que, bien que cet exemple concerne une grappe de 3 noeuds, des configurations similaires peuvent être réalisées sur des grappes à deux noeuds ou à plus deux noeuds.

Dans cette configuration :

- La station de travail est connectée à un port série (J4, par exemple) du concentrateur. Dans ce cas, la station doit être à proximité du PowerCluster.
- Un terminal ASCII est connecté au port série J0 du concentrateur. Sur un concentrateur CS/2600, ce port J0 est spécial en ce sens qu'il sert à accéder aux menus et aux commandes qui permettent de gérer le concentrateur.
- Les ports console S1 des noeuds sont connectés aux ports série J1, J2 et J3 du concentrateur. Ainsi, tous ces ports console sont accessibles via le concentrateur.
- A partir de la station de travail, vous avez à tout moment accès, via les fonctions du concentrateur, à n'importe lequel des ports du concentrateur. Vous pouvez ainsi contrôler la console (S1) de n'importe quel noeud. Pour en savoir plus sur la configuration et l'utilisation du concentrateur, reportez-vous à "Configuration et utilisation du concentrateur de console", page A-13.

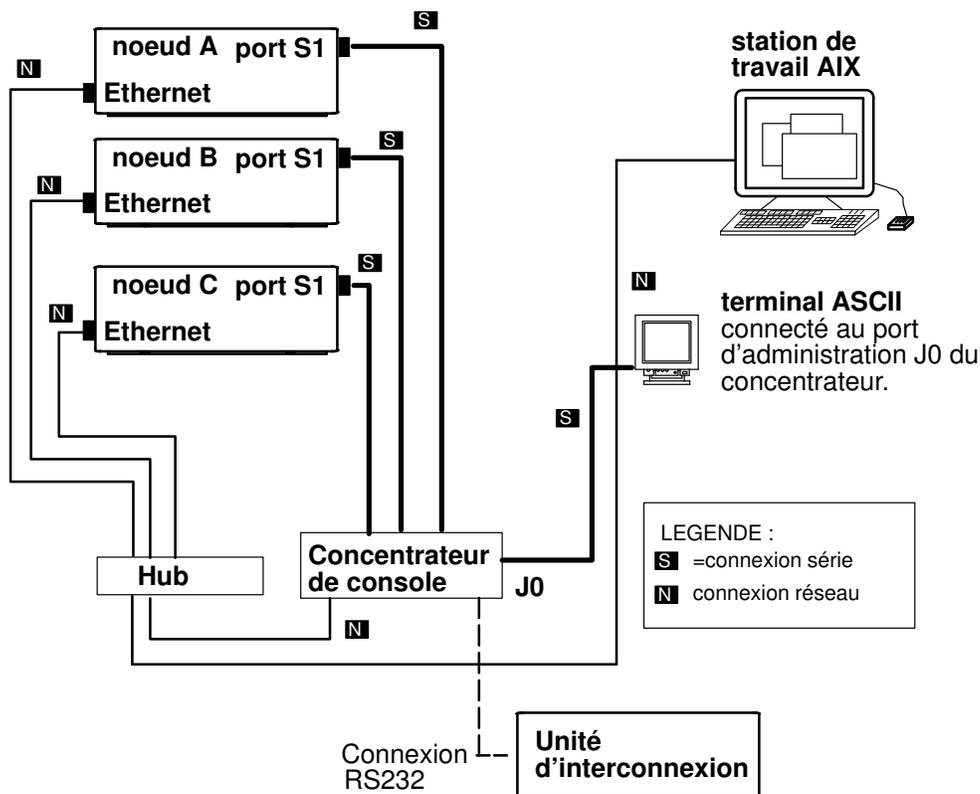
Notez en outre ce qui suit :

- Si le concentrateur vient ensuite à faillir (ce qui entraîne la perte de l'accès à tous les ports de console S1), vous pouvez connecter provisoirement le terminal ASCII au port console S1 d'un noeud, et l'utiliser comme console jusqu'à ce que l'unité défaillante soit réparée.
- Pour une sécurité maximale, nous avons choisi d'utiliser le concentrateur pour implémenter un *réseau d'administration* indépendant, qui interconnecte les noeuds, le concentrateur et la station de travail. A partir de celle-ci, vous avez accès aux noeuds via le réseau, et vous pouvez à volonté exécuter les utilitaires d'administration, en profitant des fonctions informatiques de la station de travail. En l'absence de concentrateur, le PowerConsole et le concentrateur de console sont connectés au réseau Ethernet de l'entreprise.

- Si vous implémentez RSF (Remote Services Facilities, un logiciel dédié au contrôle des erreurs), connectez le modem RSF à la ligne série S2 de la station de travail, configurez RSF sur la station de travail et sur les noeuds.

**Remarque :** La station de travail Escala série S100 est dotée d'un modem interne.

- Pour en savoir plus sur la configuration de RSF, reportez-vous à la documentation RSF.



#### Exemple 4 (grappe de N noeuds, terminal X et concentrateur de console)

Vous pouvez implémenter une variante de l'exemple 3 ci-dessus, consistant à utiliser un terminal X (console de grappe) au lieu d'une station de travail (PowerConsole). Le schéma logique est identique. Dans cette configuration :

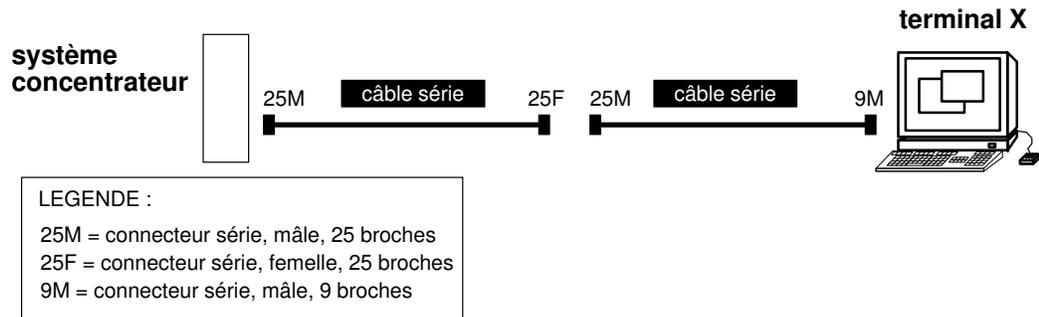
- Le terminal X peut être connecté, via son port série, à un port série (J4, par exemple) du concentrateur.
- Un terminal ASCII est connecté au port série J0 du concentrateur. Sur un concentrateur CS/2600, ce port J0 est spécial en ce sens qu'il sert à accéder aux menus et aux commandes qui permettent de gérer le concentrateur.
- Les ports console S1 des noeuds sont connectés aux ports série J1, J2 et J3 du concentrateur. Ainsi, tous ces ports console sont accessibles via le concentrateur.
- A partir du terminal X, vous avez à tout moment accès, via les fonctions du concentrateur, à n'importe lequel des ports du concentrateur. Vous pouvez ainsi contrôler la console (S1) de n'importe quel noeud. Pour en savoir plus sur la configuration et l'utilisation du concentrateur, reportez-vous à "Configuration et utilisation du concentrateur de console", page A-13.

## Guides pour le câblage du concentrateur

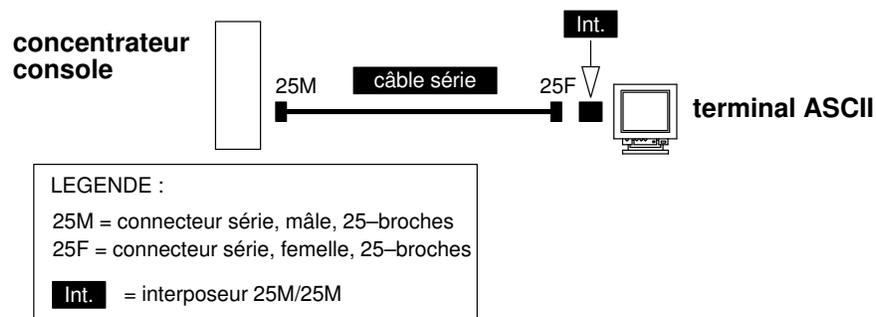
**Remarque :** Cette section ne traite que de la connexion des ports série du concentrateur aux différentes consoles possibles, et explique comment implémenter une connexion Ethernet entre le concentrateur et le hub. Pour des informations sur le câblage des ports console S1 des serveurs (noeuds de grappe), reportez-vous à la documentation matérielle du serveur et au manuel *EPC – Guide de câblage*.

**Remarque :** Un câble série est un câble droit (non torsadé).

La figure ci-dessous illustre les pièces requises pour implémenter une connexion série entre le port série d'une console de grappe (terminal X) et un port série du concentrateur.

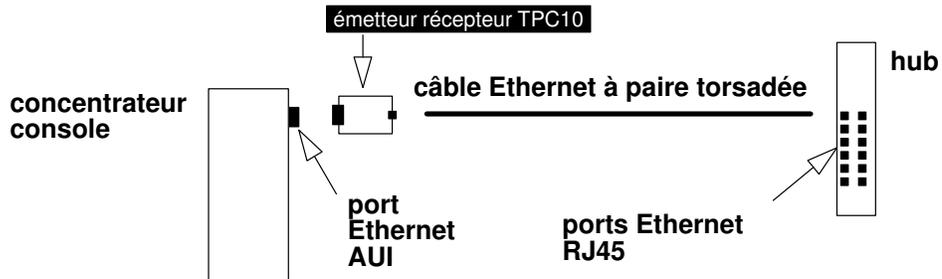


La figure ci-dessous illustre les pièces requises pour implémenter une connexion série entre un terminal ASCII BQ306 et un port série (J0, généralement) du concentrateur.



## Hub (concentrateur)

Le concentrateur de console est équipé d'un port AUI (Ethernet épais). Si vous prévoyez de connecter le concentrateur au hub fourni, lequel est équipé de ports RJ45 (Ethernet à paire torsadée), vous devez utiliser un émetteur-récepteur Ethernet TPC10, comme décrit ci-après.



## Remarques sur le montage dans une armoire

Il existe un kit en option permettant de monter le concentrateur de console dans l'armoire.

### ATTENTION :

Lorsque vous installez le concentrateur de console dans l'armoire, la température extérieure ne doit pas dépasser 35 °C.

---

# Configuration et utilisation du concentrateur de console

## Remarques préliminaires

La procédure ci-dessous ne présente que les grandes lignes des opérations. Pour en savoir plus sur la configuration, l'utilisation et la gestion du concentrateur de console, reportez-vous aux manuels du *serveur de communication multiprotocole CS/2600*.

Dans la procédure exemple ci-dessous, nous supposons que votre configuration comporte un terminal ASCII (BQ306) connecté au port d'administration J0 du concentrateur et dédié à la gestion du concentrateur. Le débit de transmission des données doit être de 9600 bauds.

**Remarque :** D'autres configurations non standard sont possibles (par exemple, le port administratif J0 peut être connecté à la station de travail AIX).

Munissez-vous d'une disquette (3½", double face/densité) pour copier la disquette 3Com d'origine.

Avant de poursuivre, lisez le *Bulletin de logiciel (SRB)* livré avec le logiciel PowerConsole. Il indique les contraintes et les restrictions relatives à l'environnement, ainsi que les informations de dernière minute, le cas échéant.

## Connexions réseau du PowerConsole

### Réseau d'administration dédié

Lorsqu'elle se trouve sur un réseau d'administration indépendant (dédié), la configuration HACMP des grappes (gérées par le PowerConsole ou la console de grappe) doit être modifiée pour y inclure ce réseau. Ce réseau doit être déclaré "public" dans HACMP.

**Remarque :** Il est important que l'administration HACMP "voit" l'attribut du réseau comme "public".

Sur chaque noeud, l'interface du réseau Ethernet (avec le réseau d'administration) doit être configurée comme suit :

- Configuration AIX

Lancez la commande **smit chinnet** sur chaque noeud pour configurer les interfaces Ethernet du réseau d'administration dédié. Sur chaque noeud, mettez à jour le fichier */etc/host* (noms et adresses des interfaces).

- Configuration HACMP
- Connectez-vous à l'un des noeuds
- Modifiez la configuration HACMP pour y inclure la carte utilisée par le réseau d'administration sur chaque noeud :
  - **smit hacmp**
  - **Cluster configuration**
  - **Cluster Topology**
  - **Configure Adapters**
  - **Add an Adapter**

- La fenêtre suivante s'affiche :

Add an Adapter			
Type or select values in entry fields. Press Enter AFTER making all desired changes.			
[Entry Fields]			
* Adapter IP Label	[ ]		
* Network Type	[ether]		+
* Network Name	[<name>]		+
* Network Attribute	public		+
* Adapter function	service		+
Adapter Identifier	[ ]		
Adapter Hardware Ad	[ ]		
Node Name	[ ]		+
F1=Help	F2=Refresh	F3=Cancel	F6=Command
F8=Image	F9=Shell	F10=Exit	/=Find
n=Find Next			

- Effectuez les sélections et appuyez sur Entrée.
- Répétez les trois dernières étapes pour inclure la carte sur chaque nœud.
- Synchronisez la topologie de la grappe sur tous les nœuds.

### Réseau public de grappe

Les adresses IP du PowerConsole et des ports CS2600 doivent appartenir au même sous-réseau que les adresses d'amorçage et de service des nœuds.

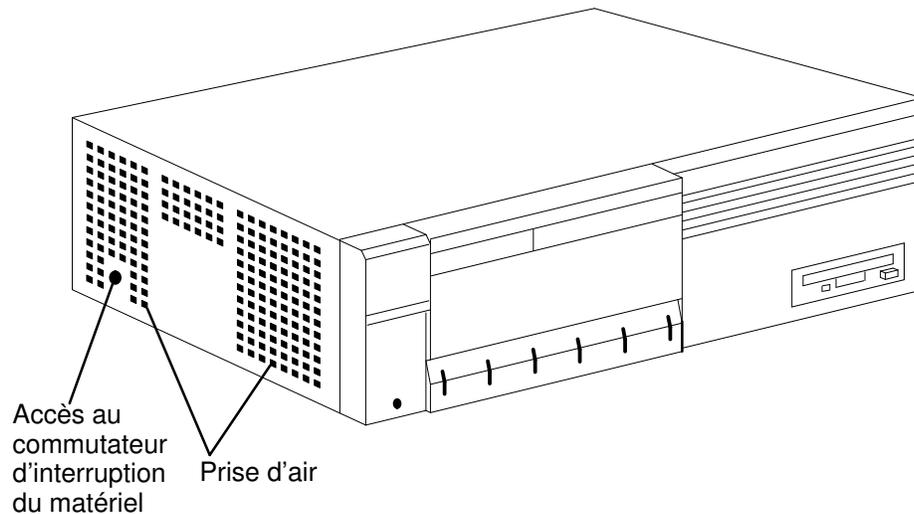
### Amorçage du concentrateur de console

- Le terminal ASCII (connecté au port d'administration J0 du concentrateur) doit être correctement configuré, et présenter les caractéristiques suivantes : débit de 9600 bauds, 8 bits par caractère, 1 bit d'arrêt, pas de parité, DTR ignoré, duplex intégral et echo désactivé.
- Insérez la disquette d'amorçage dans l'unité de disquette du concentrateur (la disquette est libellée "83-0377-004 CS/2600 SW/2500-TO-3270-LOCAL").

Mettez ensuite le concentrateur sous tension, et attendez la fin des autotests et de l'initialisation : voyant "autotest" jaune et voyant disquette éteints (au moins 3 minutes).

## Configuration des paramètres de base du concentrateur

- Lancez un réamorçage matériel du concentrateur, comme expliqué dans le manuel *CS/2500 Series Communication Server Installation Guide*. L'accès au commutateur d'interruption du matériel est représenté sur la figure.



Utilisez uniquement un objet non conducteur, comme une pointe en plastique, pour appuyer sur le commutateur d'interruption du matériel. N'utilisez pas la pointe d'un crayon. Les particules en graphite sont susceptibles de produire des décharges électriques et d'endommager des composants sur les cartes de circuit du serveur.

- Patientez un instant et appuyez deux ou trois fois sur Entrée (à intervalle de 1 seconde), jusqu'à afficher "3Com Corporation CS/2600 Series Monitor", suivi de l'invite de commande > sur la console ASCII.

- Sur la nouvelle disquette 3½", effectuez une copie de la disquette d'origine. A l'invite >, entrez :

```
co <Enter>
```

Pour afficher la liste des commandes disponibles, à l'invite >, entrez :

```
?
```

Reportez-vous au besoin aux instructions du manuel *CS/2600 Installation Guide*.

Placez la disquette d'origine en lieu sûr (utilisez désormais la copie).

- A l'invite >, entrez ;

```
fc <Enter>
```

Le menu "Firmware Configuration Utility Commands" s'affiche.

```
C - Change parameters
D - Display parameters
Esc - Exit to monitor.
```

Sélectionnez Change parameters, à l'invite ?, entrez :

```
C
```

Le menu "Change parameters" s'affiche. Modifiez les paramètres comme suit :

```
5. Monitor           Disabled
6. Initial boot source Local Floppy
A. Boot protocol     TFTP Boot
B. Change IP/TFTP parameters
```

Sélectionnez le menu **B** qui apparaît lorsque vous sélectionnez le protocole d'amorçage TFTP. Modifiez les paramètres comme suit :

Addr discover protocol	<b>Local Information</b>
Client Ip address	CS2600 Ethernet address for example 1.0.0.10
Gateway address	for example 1.0.0.10 (if necessary, public network administrator)
Subnet mask value	subnet mask of administration network for example: 255.0.0.0

"Addr discover protocol" est conforme au protocole utilisé sur votre site.

"Client IP address" spécifie l'adresse IP choisie pour le concentrateur.

"Subnet mask value" spécifie le masque de sous-réseau du réseau LAN d'administration (dédié ou public).

- Appuyez plusieurs fois sur Echap pour quitter les menus et revenir à l'invite > du moniteur.
- Effectuez une réinitialisation logicielle du concentrateur en appuyant sur le bouton de réinitialisation situé sur le panneau avant. Attendez la fin des autotests (voyants "autotest" et "état d'amorçage" éteints – au moins 1 minute).

Patiencez encore deux minutes, que le voyant disquette s'éteigne.

Appuyez plusieurs fois sur Retour, jusqu'à afficher l'invite :

```
Welcome to the 3Com Communication  
[1] CS>
```

- Pour effectuer des tâches de configuration plus poussées, vous devez passer au niveau de privilège "Network Manager". Entrez la commande :

```
[1] CS> set pri = nm <Enter>
```

Le concentrateur demande un mot de passe. Appuyez sur Entré (au départ, aucun mot de passe n'est défini).

Une nouvelle invite apparaît, indiquant que vous vous trouvez au niveau de privilège Network Manager (administrateur de réseau) :

```
[2] cs#
```

- A partir de là, vous pouvez configurer les paramètres de base du concentrateur (date, heure, nom système, mot de passe...), comme décrit dans la documentation du *serveur de communication multiprotocole CS/2600*. Une fois les paramètres définis, mettez à jour la liste des services autorisés, via la commande :

```
cs# set cs = all <Enter>
```

- Une nouvelle invite s'affiche :

```
[3] cs#
```

- Déclarez à nouveau l'adresse IP client affectée (plus haut) au concentrateur, par exemple :

```
cs# setd -ip net = 1.0.0.10 <Enter> (CS/2600 Ethernet address)
```

## Configuration des ports du concentrateur de console

Vous devez à présent configurer les ports série du concentrateur (J1, J2, J3...) connectés aux ports console S1 ou COM1 des noeuds de grappe.

**Remarque :** Lorsque vous entrez une commande à l'invite `cs#`, pour désigner un port du concentrateur, vous devez spécifier une expression de la forme `!n` où `n` est le numéro du port. Par exemple, l'expression `!2` désigne le port série J2 du concentrateur.

- Le port J0 du concentrateur doit être configuré comme un port de **terminal**. Par ailleurs, les ports série (J1, J2, J3...), connectés aux ports S1 ou COM1 des noeuds de grappe ou des unités interconnectées, doivent être définis comme ports **hôte**.

Par défaut, tous les ports sont configurés comme **ports de terminal**. Vous n'avez qu'à modifier les ports connectés aux noeuds de grappe ou aux unités interconnectées.

Par défaut, le paramètre débit est défini à **autobaud**. Les ports hôte doivent être configurés à 9600 bauds.

A titre d'exemple, voici les deux commandes à entrer pour configurer le port J1 comme port hôte, avec le débit en bauds approprié :

```
[3] cs# setd !1 -term dv = host <Enter>
[4] cs# setd !1 -term baud = 9600 <Enter>
[5] cs# setd !1 -term autd = <x> <Enter>
```

où `<x>` = 1 à 16 000 minutes. Cette option définit le délai d'inactivité du port au bout duquel il est "auto-déconnecté".

Répétez ces commandes pour les autres ports (!2, !3...) à configurer et achever la configuration.

- Affectez une adresse IP à chaque port série connecté à un noeud de grappe ou "switch". Par exemple, pour affecter à l'adresse IP du port J1 la valeur 1.0.0.11, entrez :

```
[7] cs# setd !1 -tcpappl porm = <@IP1> <Enter>
```

où `<@IP1>` est l'adresse IP du port J1 connecté au port S1 ou COM1 du noeud 1, par exemple : 1.0.0.11.

Répétez cette étape pour chaque port connecté au port console S1 ou COM1 d'un noeud ou d'unités interconnectées.

## Vérification des paramètres du concentrateur

- Pour vérifier les adresses IP affectées aux ports du concentrateur, entrez :

```
[8] cs# sh -tcpappl porm <Enter>
```

- Pour vérifier les paramètres du port hôte J1, entrez :

```
[9] cs# sh !1 dp <Enter>
```

Lancez des commandes similaires pour les autres ports (!2, !3...) dont les paramètres doivent être vérifiés.

## Vérification de la connexion réseau du concentrateur

Pour vérifier que le concentrateur peut communiquer via le réseau, faites appel à la commande `ping`.

Vous pouvez, par exemple, utiliser l'adresse IP de la station PowerConsole comme argument de la commande `ping`.

```
[10] cs# ping @IP_<nom> <Enter> (par exemple, ping 1.0.0.20)
```

Si la connexion au réseau est OK, la commande ping renvoie le message :

```
pinging... 140.184.33.122 is alive
```

## Vérification de l'état LISTEN sur les ports du concentrateur de console

Pour vérifier que les ports du concentrateur, connectés à S1 sur les noeuds, sont à l'état LISTEN, entrez :

```
[6] cs# sh -term all <Enter>
```

Le port 0 est à l'état "command". Les autres ports, connectés à S1, doivent être à l'état "LISTEN". Si ce n'est pas le cas, déconnectez les ports via la commande **logout** :

```
[10] cs# logout !1 <Enter>
```

pour le port J1, par exemple.

## Mise à jour du fichier /etc/hosts sur le PowerConsole

Sur une station de travail PowerConsole, vous devez mettre à jour le fichier `/etc/hosts` en y intégrant les adresses IP définies pour le concentrateur. Ces adresses comprennent l'adresse IP du concentrateur lui-même, ainsi que les différentes adresses IP affectées aux différents ports hôte (J1, J2...) du concentrateur, connectés aux noeuds.

Par exemple, si vous configurez une grappe HACMP de 3 noeuds (appelés par exemple `jazz`, `java` et `bebop`), vous avez attribué des adresses IP aux ports J1, J2 et J3 du concentrateur, auquel cas vous devez ajouter les entrées appropriées ci-après (en adaptant les noms et les adresses à votre configuration) dans le fichier `/etc/hosts` :

```
120.154.33.10 concentrator
120.154.33.54 EstrellaLink
120.154.33.51 jazz_console
120.154.33.52 java_console
120.154.33.53 bebop_console
```

**Remarque** : Suivant votre configuration, vous pouvez être amené à mettre à jour `/etc/hosts` sur d'autres machines également ou, le cas échéant sur le serveur de noms utilisé sur votre site.

## Accès aux ports console des noeuds via le concentrateur

Une fois les éléments du PowerConsole configurés comme indiqué dans ce chapitre, le PowerConsole peut se connecter aux noeuds via le port console série (au moyen du concentrateur de console). Vous pouvez ainsi accéder aux consoles des noeuds pour effectuer d'autres tâches de configuration.

Pour accéder au port console d'un noeud, appelez la commande `telnet` avec l'adresse IP souhaitée comme argument.

## Utilisation de Telnet

- Par exemple, pour accéder à la console du noeud connecté au port J1 du concentrateur, en supposant que vous avez affecté au port J1 l'adresse IP 120.154.33.51, entrez la commande :

```
telnet 120.154.33.51 <Enter>
```

Une fois cette commande lancée, vous accédez à la console système du noeud, comme vous l'auriez fait à partir d'un terminal ASCII conventionnel directement raccordé au port console S1 du noeud.

- Pour fermer une session telnet, appuyez simultanément sur **Ctrl** et sur **J**. Une fois déconnecté, vous revenez à l'invite `telnet>`. A ce stade, entrez :

```
q <Enter>
```

ou

```
quit <Enter>
```

---

# Configuration et exploitation de l'utilitaire pwcons

## Présentation de pwcons

L'utilitaire **pwcons** est livré avec le module BullCluster. Cet utilitaire est particulièrement utile pour les configurations équipées d'un seul terminal où les consoles sont gérées par le biais d'une carte multiport (voir exemples de configurations à la section "Solutions console basées sur une carte multiport", page A-3, et à la section "Exemple 1", page A-8).

**pwcons** est un script shell, résidant dans le répertoire **/usr/sbin**, qui repose sur des fichiers de configuration que vous devez définir, comme indiqué à la section "Définition des fichiers de configuration pwcons", ci-après.

### Scénario et commentaires

Supposons que le noeud A soit relié à un terminal, et que les ports S1 des autres noeuds de la grappe (noeud B, noeud C, etc.) soient connectés à une carte multiport sur le noeud A. Pour accéder aux consoles des autres noeuds :

1. A partir de la console du noeud A, connectez-vous au noeud A.
2. La connexion établie, entrez une commande du type :

```
pwcons -c ttyx
```

où *ttyx* est le nom de l'unité correspondant au port série (sur la carte multiport le plus souvent) connecté au port S1 du noeud à atteindre.

*Par exemple, la commande `pwcons -c tty4` indique que vous souhaitez utiliser l'unité `tty4`. Si `tty4` correspond au port (sur la carte multiport) auquel est connecté le port S1 du noeud C, cette commande vous connecte au noeud C.*

Vous pouvez, au lieu du nom d'une unité *ttyx*, spécifier un nom symbolique (chaîne mnémorique arbitraire, nom de noeud le plus souvent). Ceci suppose, toutefois, de définir le fichier **/etc/pwcons.conf** pour déclarer la correspondance entre noms d'unité et noms symboliques. Voir plus bas.

3. Supposons que vous avez entré `pwcons -c tty4`, et que vous êtes donc connecté au port S1 d'un noeud donné, disons le noeud C. A partir de maintenant, tout se passe comme si vous étiez directement connecté à la console système du noeud C :
  - Vous recevez les messages destinés à la console du noeud C (mais vous n'avez pas accès aux messages reçus avant que vous soyez connecté à la console).
  - Si le noeud C est actif, sa bannière de connexion est affichée, vous invitant à entrer votre nom et votre mot de passe. Vous pouvez ainsi vous connecter au noeud C si vous le souhaitez.
4. Une fois terminé votre travail sur le noeud C, déconnectez-vous de sa console :
  - Si vous étiez connecté, déconnectez-vous.
  - Entrez la séquence `~.` (tilde et point), puis Retour, pour mettre fin à la connexion.
5. Une fois la connexion fermée, vous revenez à la console d'origine (noeud A). Si vous le souhaitez, vous pouvez alors relancer la commande **pwcons** pour vous connecter à un autre noeud.

Notez que vous ne pouvez utiliser **pwcons** qu'à partir du noeud auquel est associée la console (noeud C dans l'exemple). Vous ne pouvez, par exemple, prendre le contrôle de la console du noeud C, puis lancer la commande **pwcons** pour accéder à la console d'un troisième noeud. Vous devez revenir au noeud A avant d'établir une connexion avec une troisième console.

Pour en savoir plus sur la configuration et l'utilisation de **pwcons**, voir ci-dessous.

## Définition des fichiers de configuration pwcons

Commencez par configurer TTY avec désactivation de la fonction de connexion.

Entrez :

```
smit maktty
```

L'ensemble des paramètres tty s'affiche.

1. Spécifiez **disable** dans le champ **Enable LOGIN**.
2. Tous les autres champs conservent leur paramètres par défaut.
3. Appuyez sur Entrée pour appliquer les paramètres.

## Configuration de /etc/uucp/Devices

- Si vous avez commandé un Powercluster équipé d'une carte multiport, les ports série (TTY) de cette carte ont été configurés en usine. Le fichier **/etc/uucp/Devices** a été également défini en usine. Vous pouvez néanmoins lire les paragraphes qui suivent pour vérifier la configuration du fichier **/etc/uucp/Devices**, ou pour comprendre son fonctionnement.
- Pour les autres configurations, consultez les paragraphes qui suivent pour configurer correctement le fichier **/etc/uucp/Devices**.

**/etc/uucp/Devices** (qui est décrit en détail dans la documentation AIX standard) contient des informations sur les unités du système local susceptibles d'établir une connexion à un système distant via des utilitaires tels que **cu**. Le script shell **pwcons** est basé sur **cu**, et donc sur ce fichier de configuration.

Le fichier **/etc/uucp/Devices** doit contenir une description de chaque unité TTY utilisée pour la connexion aux ports S1 des autres noeuds. Aussi, vérifiez que ce fichier est configuré comme souhaité (créez-le s'il n'existe pas, modifiez-le si nécessaire).

Les descriptions TTY doivent être spécifiées par des lignes au format :

```
Direct ttyx - 9600 direct
```

où *ttyx* est un nom d'unité correspondant à un port série (sur la carte multiport généralement) connecté au port S1 d'un noeud que vous souhaitez atteindre.

Notez que le fichier peut comporter des commentaires (lignes introduites par un #) et des lignes vides (elles sont ignorées).

### Exemple

Supposons que vous implémentiez une grappe 4 noeuds, où noeud A est équipé d'une carte multiport pour la connexion aux ports S1 des trois autres noeuds (comme illustré dans l'exemple 2, page A-3). Supposons également que la carte multiport est configurée de sorte que ses ports connectés aux autres noeuds correspondent aux noms d'unité *tty4*, *tty5* et *tty6*.

Le fichier **/etc/uucp/Devices** doit également comporter les trois lignes suivantes :

```
Direct tty4 - 9600 direct
Direct tty5 - 9600 direct
Direct tty6 - 9600 direct
```

## Configuration de `/etc/pwcons.conf`

Cette tâche est facultative. Le fichier `/etc/pwcons.conf`, simple outil de confort, permet de remplacer les noms d'unité (tels que `tty4`) par des noms symboliques, lors de la connexion à un noeud de console via **pwcons**.

Le fichier `/etc/pwcons.conf` décrit les correspondances entre noms symboliques et nom d'unité TTY, pour les unités connectées aux ports S1 des autres noeuds. Il doit comporter une ligne par correspondance à définir (créez le fichier s'il n'existe pas, modifiez-le si nécessaire).

Chaque ligne est au format :

```
mnémonique ttyx
```

où *mnémonique* est un nom quelconque que la commande **pwcons** doit interpréter comme synonyme du nom de l'unité *ttyx*. Le plus souvent, vous spécifierez un *mnémonique* correspondant au nom du noeud accessible via *ttyx*.

Notez que le fichier peut comporter des commentaires (introduits par un #).

### Exemple

Supposons que vous implémentiez une grappe 4 noeuds, où *alpha* est un noeud équipé d'une carte multiport utilisée pour la connexion aux ports S1 des trois autres noeuds. Supposons que la carte multiport est configurée de sorte que ses `tty4`, `tty5` et `tty6` soient connectés aux ports S1 des noeuds *beta*, *gamma* et *delta*.

Si `/etc/pwcons.conf` n'existe pas ou n'est pas configuré, vous pouvez vous connecter à la console de *beta* via la commande **pwcons -c tty4**. Ce qui suppose de retenir les numéros des tty.

Par contre, si vous configurez `/etc/pwcons.conf` comme suit :

```
beta tty4
gamma tty5
delta tty6
```

Vous pouvez alors vous connecter à la console de *beta* avec la commande **pwcons -c beta**. Ce qui est plus simple, les noms de noeud étant plus aisés à mémoriser que les numéros de tty.

## Syntaxe de la commande **pwcons**

L'objet de **pwcons**, expliqué ci-dessus, est supposé connu.

### Syntaxe

```
/usr/sbin/pwcons [ -c ttyx | mnémonique ] | [ -s ttyx | mnémonique ] | [ -S ] |  
[ -d ttyx | mnémonique ] | [ -D ]
```

### Indicateurs

[ -c *ttyx* | *mnémonique* ]

Etablit la connexion avec un tty spécifique, identifié soit par son nom d'unité *ttyx*, soit par le *mnémonique* correspondant, tel que défini dans le fichier `/etc/pwcons.conf` (décrit plus haut).

A l'ouverture d'une connexion, un fichier de verrouillage est créé, qui déclare le tty verrouillé. Le mécanisme de verrouillage de fichier est identique au fichier de verrouillage de **cu** (**cu** est un programme de la famille uucp - pour en savoir plus, reportez-vous à la documentation AIX).

Une fois connecté, vous pouvez mettre fin à la connexion à tout moment, via la séquence de touches `~.` (tilde et point), puis Retour. (Si vous étiez connecté, déconnectez-vous avant de clore la connexion.)

[ **-s ttyx** | *mnémonique* ]

Affiche l'état d'un tty spécifique. Il est soit ouvert, soit fermé.

Un tty fermé n'est pas verrouillé : vous pouvez vous y connecter.

Un tty ouvert est verrouillé : vous ne pouvez pas vous connecter à un tty verrouillé sauf à forcer une déconnexion à le déverrouiller (voir les indicateurs `-d` et `-D`). Si un tty est ouvert, l'indicateur `-s` affiche les numéros de processus des commandes qui l'ont ouvert.

[ **-S** ]

Affiche l'état de tous les tty de la liste `/etc/pwcons.conf`.

[ **-d ttyx** | *mnémonique* ]

Force la déconnexion d'un tty spécifique, de sorte qu'il devient fermé et déverrouillé. Vous pouvez utiliser cette fonction si la connexion à un tty n'aboutit pas parce qu'il est verrouillé. Normalement, un tty n'est verrouillé que lorsque quelqu'un y est connecté et l'utilise. Toutefois, il peut également être verrouillé parce qu'une session de communication a été interrompue anormalement, sans fermeture correcte de la connexion.

[ **-D** ]

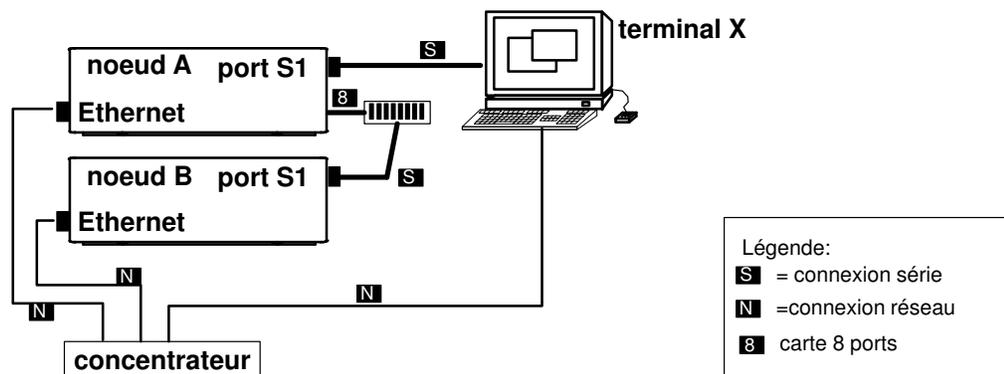
Force la déconnexion de tous les tty de la liste `/etc/pwcons.conf`.

# Annexe B. Implémentation de la solution console de grappe

Cette annexe traite de l'installation du terminal X.

## Procédure d'installation du terminal X

Le terminal X est connecté comme illustré sur la figure.



### ATTENTION :

**Ne branchez pas les cordons d'alimentation au boîtier du terminal X et à l'avant du moniteur avant d'y être invité.**

1. Installez l'extension mémoire et la carte PCMCIA, préalablement configurée pour l'écriture.

Reportez-vous au chapitre 5 du manuel "*Installing Your Explora Family System*".

2. Connectez le câble vidéo entre le boîtier du terminal X et le moniteur.
3. Connectez le terminal X au concentrateur Superstack II Hub 10 via un câble RJ45/RJ45 (CBF5410).
4. Connectez le boîtier du terminal X:  
au concentrateur de console via des câbles RS232 (CBL 1912, CBLG106–2000),  
sinon, à la fiche S1 d'un nœud Powercluster via des câbles RS232 (CBL 1912, CBLG105–1800).
5. Connectez le clavier et la souris au boîtier du terminal X.
6. Branchez le câble d'alimentation au boîtier du terminal X et au moniteur.
7. Mettez sous tension le boîtier du terminal X.
8. Mettez sous tension le moniteur (voyant vert allumé). Pour en savoir plus sur les voyants et commutateurs, reportez-vous au manuel "*17 Professional Color Monitor – User's Guide*".
9. Vous pouvez interrompre le démarrage automatique du terminal X en tapant ESC une fois les tests de mise sous tension terminés.
10. A l'invite >, tapez la commande `se` et appuyez sur ENTREE pour appeler le menu principal.
11. Sélectionnez « keyboard » pour définir le type de clavier (IBM PS/2 ou N101).

12. Sélectionnez « Monitor » pour définir la fréquence de résolution du moniteur.

**ATTENTION :**

**Sélectionner une résolution inappropriée peut endommager le moniteur. Vérifiez que la résolution adoptée est prise en charge par le moniteur. Reportez-vous au manuel "17 Professional Color Monitor – User's Guide".**

13. Sélectionnez « Network », puis définissez les paramètres suivants :

```
Get IP Address from:          NVRAM
Terminal IP Address:        @IP of XT
Subnet Mask:                Subnet Mask
```

14. Sélectionnez « Boot », puis définissez les paramètres suivants :

```
TFTP Boot Directory:        /local/
TFTP Order:                 Disabled
NFS Order:                  Disabled
MOP Order                   Disabled
Local Order                 1
```

15. Sélectionnez « Done », puis « Reboot ».

16. A l'invite >, si l'amorçage n'est pas automatique, tapez :

>BL et appuyez sur ENTREE.

17. Une fois l'amorçage terminé, deux ou trois fenêtres s'affichent :

une fenêtre Setup & Configuration (en haut à gauche)

une fenêtre telnet

une fenêtre console système correspondant à la ligne série RS232 (prise S1 d'un noeud Powercluster), si le terminal X est directement connecté à la prise S1 d'un noeud.

18. Dans la fenêtre Setup & Configuration, sélectionnez « Window Manager » et lancez « NCD Window Manager » en cliquant sur l'icône qui lui est associée.

19. Dans la fenêtre telnet, ou après ouverture de "New Telnet" dans "TERMINAL" de la fenêtre de définition :

Tapez, dans le champ Service, l'adresse IP d'un autre noeud et cliquez sur OK pour établir la session telnet.

Ouvrez autant de fenêtres New Telnet qu'il y a de noeuds restants : passez à la fenêtre Setup & Configuration, sélectionnez Terminal, puis New Telnet et tapez l'adresse IP du noeud dans le champ Service.

20. Pour que l'amorçage à la mise sous tension soit entièrement automatique, sélectionnez le menu SETUP dans la fenêtre Setup and Configuration, puis sélectionnez :

```
Change Setup Parameters
Booting
```

Validez « Boot automatically at power-up ».

Sauvegardez la configuration en sélectionnant « Apply ».

21. Pour que l'établissement d'autres sessions telnet sur les autres noeuds soit automatique, sélectionnez, dans la fenêtre Setup :

SETUP

Change Setup Parameter

Commands and Startup

Puis, pour chaque noeud, indiquez l'adresse IP dans la ligne de commande :

```
term -ctype telnet <@IP> -geometry ...-title <node name>...-n<node abbr>...
```

Sauvegardez la configuration en cliquant sur « Apply ».

22. A la mise sous tension suivante, désactivez la configuration "écriture admise" de la carte PCMCIA.



---

# Index

## Symboles

./rhosts, 7-4  
/etc/hosts, 4-6, 7-4, 11-3, 11-4, 11-17  
    mise à jour du fichier, A-18  
/etc/pwcons.conf, A-21  
/etc/snmpd.conf, 11-3  
/etc/uucp/Devices, A-20  
/usr/sbin/bullcluster/install/samples, 3-2  
/usr/sbin/cluster/clusterview, 10-3  
/usr/sbin/cluster/events/utills/mailltoroot, 8-4

## A

Add a File System to a Resource Group, 9-14  
Add a Logical Volume to a Resource Group, 9-11  
Add a Resource Group, 9-4  
Add a Service IP Label to a Resource Group, 9-17  
Add a Volume Group to a Resource Group, 9-8  
Add an Adapter (menu), A-13  
administration. *Voir* outil de gestion  
adresse d'amorçage, 7-12  
adresse d'amorçage IP, 4-5  
adresse de service IP, 4-5  
adresse IP, 4-4  
AIX, configuration, A-13  
alias d'adresse IP, 10-5  
Alternate Hardware Address, 7-10  
Apply Configuration Definitions, 7-11  
Architecture de stockage en série (SSA). *Voir* SSA  
    (carte)  
armoire d'extension disque, 3-7, 3-8  
armoire d'extension partagée, 3-9  
ArrayGUIDe, 10-29  
ATF (Application Transparent Failover), 3-12  
ATM, 3-3

## B

basculement d'adresse IP, 4-4, 4-7  
    ajout d'un label de service IP, 9-17  
    alias, 10-5  
bibliothèques de bandes, 1-11  
Boot Address, 7-10  
BQ306, A-13  
Bull Cluster Easy Configuration, 7-5  
BullCluster  
    outil de gestion, 9-1  
    présentation du module, 1-4  
    SRB, 5-1

## C

câblage  
    guide, A-11  
    montage dans une armoire, A-12  
    spécifique du hub (concentrateur), A-12  
câble, 3-13  
câble en Y, 3-13  
Cable Type, 7-9  
Carte. *Voir* SCSI  
carte  
    ID SCSI, 4-7  
    PCI Fibre Channel, 6-4  
    SSA, 3-22  
carte de secours, 4-5, 7-12  
carte de service, 7-12  
carte SCSI, 6-2  
Cartes  
    PCI Fibre Channel, 6-4, 6-5  
    SCSI, terminaison simple, 6-3  
Cartes pour bus MCA, 6-2  
Chapter, title, more information, X-1  
clé de licence, 10-28  
clsmuxpd, 11-3  
cluster configuration (menu), A-13  
Cluster ID, 7-9  
Cluster Name, 7-9  
Cluster Resources Modification, 9-1  
cluster topology (menu), A-13  
cluster.clvm, 5-1  
ClusterWatch, 1-5  
commandes  
    maktty, A-20  
    pwcons, A-21  
composant, grappe, 1-4  
concentrateur, PCI Fibre Channel, 6-5  
concentrateur de console, A-6  
    configuration et utilisation, A-13  
    état LISTEN des ports, A-18  
concepts, reprise après sinistre, 1-1  
concurrent (accès), 1-2  
Concurrent Resource Manager (logiciel), 1-4  
configuration  
    concentrateur de console, A-13  
    paramètres du concentrateur de console, A-15  
    utilitaire pwcons, A-19  
configuration AIX, A-13  
configuration de la topologie. *Voir* configuration  
    initiale rapide  
configuration de reprise après incident, 10-24  
configuration HA native simple, partagée, 3-8  
configuration HACMP, A-13

- configuration initiale rapide, 7-1
  - actions exécutées, 7-12
  - annulation, 7-14
  - composants LVM (groupes de volumes...), 8-2
  - configurations acceptées, 7-2
  - finalisation, 8-1
  - menu Quick Initial Configuration, 7-5
  - tâches de préconfiguration (choix de la méthode), 7-1
- configuration jumelée (FDDI), 3-5
- Configure Adapters (menu), A-13
- Configure IP Aliasing? (menu), 10-6
- connexion, réseau PowerConsole, A-13
- console, grappe, A-6
- console ASCII, A-15
- console de grappe, A-6
- console système, A-5
- consoles
  - configurations classiques, A-8
  - écran graphique, A-6
  - PowerConsole, A-6
- Create and add an Application Server to a Resource Group, 9-19
- CRM. *Voir* Concurrent Resource Manager (logiciel)

## D

- Default GATEWAY Address, 7-9
- démon timed, 10-27
- dépannage, outil de diagnostic, 10-1
- diagnostic (outil), 10-1
- disque
  - Voir aussi* sous-système DAS
  - configuration, 6-6
  - modèles acceptés, 1-7
- Disques de sous-systèmes EMC Symmetrix partagés, Configuration de base avec HACMP, 3-26
- DSMIT, 10-28

## E

- écran graphique, A-6
  - restriction relative au EPC800, A-6
  - utilisé avec la fonction de reprise après incident, A-6
- Emplacements de disques en grappe, 3-19
- Escala Powercluster, consoles, A-5
- événement, notification par messagerie électronique, 8-3
- événement de grappe, traitement, 10-27

## F

- FDDI, 3-3, 3-4
- Fibre Channel
  - carte PCI, 6-4
  - cartes PCI, 6-4, 6-5
  - concentrateur, 6-5
  - configurations, 3-17
- fichier de configuration, 7-8
  - vérification, 7-15
- fichier de configuration snmpd, 11-3
- fichiers de configuration, pwcons, A-20

- fonction HA native, 3-6

## G

- Gestion des sous-systèmes DAS, 10-29
- grappe, 1-1
  - Voir aussi* outil de gestion
  - noeuds, 1-4
  - personnalisation du traitement, 10-27
- groupe de ressources, ajout, 9-4
- groupe de volumes
  - Voir aussi* système de fichiers; volume logique
  - brut
  - ajout, 9-8
  - configuration, 8-2
  - diffusion générale des définitions, 9-20
  - sous-système DAS, 8-5

## H

- HA Native, 10-18
- HA native, 10-17
  - implémentation, 10-17
- HACMP
  - accès au menu principal, 7-5
  - configuration, A-13
  - configuration des serveurs d'application et des ressources, 8-1
  - généralités, 1-4
  - interruptions, 11-17
  - MIB, 11-5, 11-14
  - SRB, 5-1
- HACMP heartbeat, 10-24
- haute disponibilité
  - concept, 1-1
  - duplication du matériel, 1-4
- hôte, 4-7, 7-12
- hub d'administration, A-6

## I

- iFOR/LS, 10-28
- Inactive Takeover Activated, 9-5
- Input Configuration File, 7-8
- installation, logiciel, 5-1
- installation logicielle
  - finalisation, 8-1
  - finalisation de l'installation matérielle, 6-1
  - généralités, 2-1
  - matériel de base, 4-1
  - méthode de configuration, 7-1
- Integrated System Management (ISM), 11-1
- intégration
  - procédure, 12-1
  - tiroir PCI, 12-1
- interruption, 11-17
- IP (adresse), 4-4
- IP Address, 7-10
- IP Aliasing Configuration (menu), 10-5
- IPAT. *Voir* basculement d'adresse IP
- ISM
  - concepts, 11-1
  - télécontrôle de grappe, 11-1

## K

kit console, 4-2

## L

libellé WSA1, 6-2  
Logical Volume Disks, 9-12  
logiciel ATF, désactivation d'auto-assign, 6-8  
LUN, 3-13, 6-7  
LVM (composants). *Voir* groupe de volumes

## M

Mailing List, 8-4  
mailtoroot, 8-4  
matériel, installation logicielle. *Voir* installation logicielle  
MIB, 11-4  
mise en miroir, 9-9, 9-15

## N

Navisphere, 10-29  
Network Interface, 7-10  
Network MASK, 7-9  
Network Type, 7-9  
noeud  
    composants d'une grappe, 1-3  
    nom de noeud, 4-6  
nom de noeuds, 7-6

## O

Outil d'instantané BullCluster, utilisation, 7-15  
outil de gestion  
    conseils, 10-1  
    ressources de grappe, 9-1  
    télécontrôle de grappe, 11-1  
Output Configuration File, 7-8, 7-11

## P

PCI drawer, Configuration, Example, 12-1  
permutation d'adresse, 7-13  
planification, 3-1  
    ports série, 3-2  
    réseaux, 3-3  
    stockage partagé HA étendue, 3-20  
    stockage partagé HA native, 3-6  
Planning, Shared EMC Symmetrix Subsystem Disks, 3-25  
PORT Number, 7-10  
port série, 3-2  
ports console des noeuds, accès, A-18  
Powercluster, composants, 1-3  
PowerConsole, A-6  
    connexion réseau, A-13  
privé, réseau, A-13  
procédure d'installation, terminal X, B-1  
processeur de stockage, 1-9  
Propagate Volume Group Definition on all nodes, 9-9, 9-20

pwcons, A-19  
    syntaxe de la commande, A-21

## R

RAID, 3-13, 6-7  
    *Voir aussi* sous-système DAS; systèmes de stockage Symmetrix  
Remove a File System from a Resource Group, 9-16  
Remove a Logical Volume from a Resource Group, 9-13  
Remove a Resource Group, 9-6  
Remove a Service IP Label from a Resource Group, 9-18  
Remove a Volume Group from a Resource Group, 9-10  
Remove an Application Server from a Resource Group, 9-19  
Remove Cluster Configuration, 7-14  
Remove hacmp only or all definitions?, 7-14  
reprise après incident, A-6  
réseau, 3-10, 4-4, 7-13  
    *Voir aussi* FDDI; réseau connexion, 6-1  
    installation logicielle, 4-2  
    public, 3-3  
réseau d'administration, dédié, A-13  
réseau de grappe, public, A-14  
réseau privé, A-13  
réseau public, A-14  
Rolling Applications, 10-7  
RS-232, 7-13

## S

SCSI  
    *Voir aussi* câble  
    adresse (ID), 4-7  
    cartes à terminaison simple, 6-3  
    libellé 4-6, 6-2  
    terminaison interne, 6-3  
    terminaisons SCSI, 6-2  
SCSI (carte), 6-2  
Send Mail for Cluster Events Notification, 8-3  
Serial Adapter, 7-10  
Serial Network Type, 7-9  
série, 3-10  
serveur, 1-4  
serveur d'application, ajout, 9-19  
Service Address, 7-10  
Service Interface, 7-9  
Show Cluster State, 10-3  
Show Resources for a Resource Group, 9-7  
smit hacmp, 7-5  
smit hacmp (menu), A-13  
smit quickconf, 7-5  
snmpd, 11-3  
snmpd.conf, 11-17  
snmpinfo -o /usr/sbin/cluster/hacmp.defs -m dump -v clsmuxpd, 11-4  
solution console, traditionnelle, A-2

- solutions console
  - basées sur une carte multiport, A-3
  - pour Escala Powerclusters, A-5
- solutions console pour solutions HA, A-2
- solutions HA
  - composants, 1-3
  - console, A-2
- sous-système DAS, caractéristiques, 1-8
- sous-système DAS
  - ArrayGUIDe, 10-29
  - auto-assign et ATF, 6-8
  - câble, 3-13
  - configuration, 6-7
  - groupes de volumes partagés, 8-5
  - logiciel ATF, 3-12
  - planification, 3-10
- SPOF, 1-4
- SRB, 5-1
- SSA (carte)
  - "Serial Storage Architecture", 1-10
  - caractéristiques, 1-10
  - topologie en boucle, 3-22
- Standby Address, 7-10
- Standby Interface, 7-10
- stockage EMC, configuration complémentaire, 6-8
- stockage partagé HA étendue, 3-20, 10-24
- Symmetrix (EMC), 1-11
- Symmetrix CDA (systèmes de stockage), 1-11
- synchronisation, 10-27
- système, console, A-5
- système de fichiers, ajout, 9-14

## T

- Telnet, A-18
- template.cfg, 7-6, 7-8
- terminal X, A-6
  - procédure d'installation, B-1
- tiroir PCI
  - intégration, 12-1
  - procédure d'intégration, 12-1
- trespass, 8-7
- TUXEDO, suppression de serveurs d'application, 9-19

## U

- unité d'extension disque, HA, 10-19
- UPS, 3-2
- Utilisation, outil d'instantané BullCluster, 7-15
- utilisation
  - concentrateur de console, A-13
  - pwcons, A-4
  - Telnet, A-18
  - utilitaire pwcons, A-19
- utilitaire, 10-3
- utilitaire, exemple, 3-2

## V

- Verify Configuration Definitions, 7-15
- volume logique brut, ajout, 9-11

## X

- X25, suppression de serveurs d'application, 9-19

## Vos remarques sur ce document / Technical publication remark form

**Titre / Title :** Bull Escala série EPC EPC et solutions HA - Guide d'installation

**N° Référence / Reference N° :** 86 F2 79HX 05

**Daté / Dated :** Septembre 1999

### ERREURS DETECTEES / ERRORS IN PUBLICATION

### AMELIORATIONS SUGGEREES / SUGGESTIONS FOR IMPROVEMENT TO PUBLICATION

Vos remarques et suggestions seront examinées attentivement.

Si vous désirez une réponse écrite, veuillez indiquer ci-après votre adresse postale complète.

Your comments will be promptly investigated by qualified technical personnel and action will be taken as required.

If you require a written reply, please furnish your complete mailing address below.

NOM / NAME : \_\_\_\_\_ Date : \_\_\_\_\_

SOCIETE / COMPANY : \_\_\_\_\_

ADRESSE / ADDRESS : \_\_\_\_\_

Remettez cet imprimé à un responsable BULL ou envoyez-le directement à :

Please give this technical publication remark form to your BULL representative or mail to:

**BULL ELECTRONICS ANGERS  
CEDOC  
34 rue du Nid de Pie – BP 428  
49004 ANGERS CEDEX 01  
FRANCE**

# Technical Publications Ordering Form

## Bon de Commande de Documents Techniques

To order additional publications, please fill up a copy of this form and send it via mail to:

Pour commander des documents techniques, remplissez une copie de ce formulaire et envoyez-la à :

**BULL ELECTRONICS ANGERS**  
**CEDOC**  
**ATTN / MME DUMOULIN**  
**34 rue du Nid de Pie – BP 428**  
**49004 ANGERS CEDEX 01**  
**FRANCE**

**Managers / Gestionnaires :**  
**Mrs. / Mme :** C. DUMOULIN +33 (0) 2 41 73 76 65  
**Mr. / M :** L. CHERUBIN +33 (0) 2 41 73 63 96  
**FAX :** +33 (0) 2 41 73 60 19  
**E-Mail / Courrier Electronique :** srv.Cedoc@franp.bull.fr

Or visit our web site at: / Ou visitez notre site web à:

<http://www-frec.bull.com> (PUBLICATIONS, Technical Literature, Ordering Form)

CEDOC Reference # N° Référence CEDOC	Qty Qté	CEDOC Reference # N° Référence CEDOC	Qty Qté	CEDOC Reference # N° Référence CEDOC	Qty Qté
__ __ - - - - - [__]		__ __ - - - - - [__]		__ - - - - - [__]	
__ __ - - - - - [__]		__ - - - - - [__]		__ - - - - - [__]	
__ __ - - - - - [__]		__ - - - - - [__]		__ - - - - - [__]	
__ __ - - - - - [__]		__ - - - - - [__]		__ - - - - - [__]	
__ __ - - - - - [__]		__ - - - - - [__]		__ - - - - - [__]	
__ __ - - - - - [__]		__ - - - - - [__]		__ - - - - - [__]	
__ __ - - - - - [__]		__ - - - - - [__]		__ - - - - - [__]	

[\_\_] : **no revision number means latest revision** / pas de numéro de révision signifie révision la plus récente

NOM / NAME : \_\_\_\_\_ Date : \_\_\_\_\_

SOCIETE / COMPANY : \_\_\_\_\_

ADRESSE / ADDRESS : \_\_\_\_\_

PHONE / TELEPHONE : \_\_\_\_\_ FAX : \_\_\_\_\_

E-MAIL : \_\_\_\_\_

**For Bull Subsidiaries / Pour les Filiales Bull :**

Identification: \_\_\_\_\_

**For Bull Affiliated Customers / Pour les Clients Affiliés Bull :**

**Customer Code / Code Client :** \_\_\_\_\_

**For Bull Internal Customers / Pour les Clients Internes Bull :**

**Budgetary Section / Section Budgétaire :** \_\_\_\_\_

**For Others / Pour les Autres :**

**Please ask your Bull representative. / Merci de demander à votre contact Bull.**



**BULL ELECTRONICS ANGERS  
CEDOC  
34 rue du Nid de Pie – BP 428  
49004 ANGERS CEDEX 01  
FRANCE**

**86 F2 79HX 05**

PLACE BAR CODE IN LOWER  
LEFT CORNER



Utiliser les marques de découpe pour obtenir les étiquettes.  
Use the cut marks to get the labels.



**Escala  
série EPC**

AIX  
EPC et solutions HA  
Guide d'installation

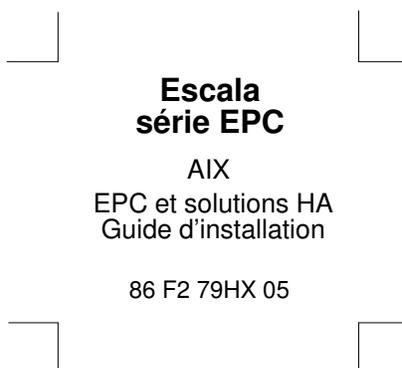
86 F2 79HX 05



**Escala  
série EPC**

AIX  
EPC et solutions HA  
Guide d'installation

86 F2 79HX 05



**Escala  
série EPC**

AIX  
EPC et solutions HA  
Guide d'installation

86 F2 79HX 05

