

IBM PowerHA SystemMirror for AIX

Standard Edition

Versão 7.2

PowerHA SystemMirror concepts

IBM

IBM PowerHA SystemMirror for AIX

Standard Edition

Versão 7.2

PowerHA SystemMirror concepts

IBM

Nota

Antes de usar estas informações e o produto suportado por elas, leia as informações no “Avisos” na página 107.

Esta edição se aplica ao IBM PowerHA SystemMirror 7.2 Standard Edition for AIX e a todas as liberações e modificações subsequentes até que indicado de outra forma em novas edições.

© Copyright IBM Corporation 2015.

Índice

Sobre este documento	v	Gerenciador do cluster.	47
Destaque	v	Componentes de software do PowerHA SystemMirror.	47
Distinção entre maiúsculas e minúsculas no AIX ..	v	Software de cluster complementar.	54
ISO 9000.	v	Assegurando a disponibilidade do aplicativo	54
Informações relacionadas	v	Visão geral: Disponibilidade de aplicativo	54
		Eliminando pontos únicos de falha em um cluster do PowerHA SystemMirror	55
Conceitos do PowerHA SystemMirror ..	1	Minimizando o tempo de inatividade planejado com o PowerHA SystemMirror	64
O que há de novo nos conceitos do PowerHA		Minimizando o tempo de inatividade não planejado	71
SystemMirror	1	Minimizando o tempo de controle.	72
Recursos novos e descontinuados	1	Maximizando a recuperação de desastre.	72
PowerHA SystemMirror for AIX.	1	Eventos de cluster	73
Clusterização de alta disponibilidade para o AIX	2	Configurações de cluster do PowerHA SystemMirror	76
Uso do Cluster Aware AIX pelo PowerHA		Configurações de espera	76
SystemMirror	4	Configurações de controle	79
Custos e benefícios contínuos da disponibilidade	7	Configurações de cluster com aplicativos com multicamadas	83
Disponibilidade aprimorada com o software AIX	7	Configurações de cluster com dependências de local de grupo de recursos	85
Componentes físicos do cluster do PowerHA		Configurações de espelho do LVM entre sites para recuperação de desastre	86
SystemMirror	9	Configurações de cluster com dynamic LPARs..	87
Objetivo do PowerHA SystemMirror: Eliminar tempo de inatividade planejado	15	Processo e recursos de configuração do PowerHA SystemMirror.	90
Conceitos de nós, redes e pulsação do cluster do		Informações que você fornece ao PowerHA SystemMirror.	90
PowerHA SystemMirror	15	Informações descobertas pelo PowerHA SystemMirror.	91
Nós	16	Opções de configuração de cluster: Padrão e estendido	91
Sites.	16	Segurança do cluster	92
Política de divisão	17	Ferramentas de instalação, configuração e gerenciamento	92
Política de mesclagem	18	Ferramentas de monitoramento.	97
Opção de desempatador para políticas de divisão e de mesclagem	19	Ferramentas de resolução de problemas	102
Soluções de vários sites do PowerHA		Ferramenta de teste do cluster.	104
SystemMirror.	20		
Redes de cluster.	23	Avisos	107
Controle de endereço IP	26	Considerações de política de privacidade	109
Pulsação sobre TCP/IP e redes de área de armazenamento	28	Marcas registradas.	109
Multicasting do PowerHA SystemMirror	29		
Recursos e grupos de recursos do PowerHA		Índice Remissivo	111
SystemMirror.	32		
Identificando e mantendo recursos do cluster disponíveis	32		
Tipos de recursos do cluster	33		
Grupos de recursos de cluster	36		
Políticas e atributos de grupos de recursos	38		
Dependências do grupo de recursos	40		
Sites e grupos de recursos	44		
Hardware suportado pelo PowerHA SystemMirror	45		
Software de cluster PowerHA SystemMirror	45		
Componentes de software de um nó do PowerHA SystemMirror	45		

Sobre este documento

Esse documento apresenta o software PowerHA SystemMirror for AIX. Estas informações também estão disponíveis no CD da documentação enviado com o sistema operacional.

Destaque

As seguintes convenções de destaque são usadas neste documento:

Negrito	Identifica comandos, sub-rotinas, palavras-chave, arquivos, estruturas, diretórios e outros itens cujos nomes são predefinidos pelo sistema. Também identifica objetos gráficos como botões, rótulos e ícones que o usuário seleciona.
<i>Itálico</i>	Identifica parâmetros cujos nomes ou valores reais devem ser fornecidos pelo usuário.
Monospace	Identifica exemplos de valores de dados específicos, de texto semelhante ao que possa estar exibido, de partes de códigos do programa semelhantes às que possam ser gravadas pelo programador, de mensagens do sistema ou informações que devam ser digitadas.

Distinção entre maiúsculas e minúsculas no AIX

Tudo no sistema operacional AIX faz distinção entre maiúsculas e minúsculas, o que significa que ele diferencia letras maiúsculas e minúsculas. Por exemplo, é possível usar o comando **ls** para listar arquivos. Se você digitar **LS**, o sistema responderá que o comando `is not found`. Da mesma forma, **FILEA**, **FiLea** e **filea** são três nomes de arquivos distintos, mesmo que residam no mesmo diretório. Para evitar causar a execução de ações indesejáveis, sempre se assegure de usar maiúsculas e minúsculas corretamente.

ISO 9000

Os sistemas de qualidade registrados ISO 9000 foram utilizados no desenvolvimento e fabricação deste produto.

Informações relacionadas

- Os documentos PDF do PowerHA SystemMirror estão disponíveis no tópico PDFs do PowerHA SystemMirror 7.2.
- As notas sobre a liberação do PowerHA SystemMirror estão disponíveis no tópico Notas sobre a liberação do PowerHA SystemMirror 7.2.

Conceitos do PowerHA SystemMirror

As informações a seguir apresentam conceitos importantes que você deve entender antes de poder usar o software PowerHA SystemMirror para o sistema operacional AIX.

O que há de novo nos conceitos do PowerHA SystemMirror

Leia sobre informações novas ou significativamente alteradas para a coleção de tópicos de conceitos do PowerHA SystemMirror.

Como consultar o que há de novo ou alterado

Neste arquivo PDF, pode haver barras de revisão (|) na margem esquerda que identificam informações novas e alteradas.

Dezembro de 2015

As informações a seguir são um resumo das atualizações que foram feitas nesta coleção de tópico:

- Os tópicos a seguir foram atualizados com informações sobre o número de discos de repositório de backup que podem ser identificados e a função Automatic Repository Disk Replacement (ARR):
 - “Discos de repositório do cluster” na página 12
 - “Falha de disco de repositório” na página 14
 - “Discos de repositório e de caminhos múltiplos” na página 13
 - “Uso do Cluster Aware AIX pelo PowerHA SystemMirror” na página 4
- Os tópicos a seguir foram atualizados com informações sobre a política manual e usando um arquivo NFS para o desempatador para uma política de divisão ou política de mesclagem:
 - “Opção de desempatador para políticas de divisão e de mesclagem” na página 19
 - “Política de divisão” na página 17
 - “Política de mesclagem” na página 18
- Os tópicos a seguir foram atualizados com informações sobre a função Resource Optimized High Availability (ROHA):
 - “Disponibilidade aprimorada com o software AIX” na página 7
 - “Configurações de cluster com dynamic LPARs” na página 87
- Informações incluídas sobre como a reserva no disco desempatador funciona em um cluster com mais de dois nós no tópico “Opção de desempatador para políticas de divisão e de mesclagem” na página 19.

Recursos novos e descontinuados

Para obter uma lista de recursos novos e descontinuados para esta liberação, consulte os materiais de Anúncio de produto IBM®. Informações adicionais também podem ser localizadas no arquivo `/usr/es/sbin/cluster/WhatsNewInThisRelease` que é instalado com o software do produto.

PowerHA SystemMirror for AIX

As informações a seguir discutem os conceitos de alta disponibilidade e de armazenamento em cluster, apresentam o diagrama de cluster do PowerHA SystemMirror e descrevem um cluster do PowerHA SystemMirror a partir de uma perspectiva funcional.

Clusterização de alta disponibilidade para o AIX

O software IBM PowerHA SystemMirror fornece um ambiente de computação comercial de baixo custo que assegura a rápida recuperação de aplicativos essenciais a partir de falhas de hardware e software.

Com o software PowerHA SystemMirror, os recursos críticos permanecem disponíveis. Por exemplo, um cluster do PowerHA SystemMirror pode executar um programa de servidor de banco de dados que atenda os aplicativos clientes. Os clientes enviam consultas para o programa do servidor que responde às solicitações, acessando um banco de dados armazenado em um disco externo compartilhado.

Esse sistema de alta disponibilidade combina o software customizado com o hardware padrão de mercado para minimizar o tempo de inatividade, restaurando serviços rapidamente quando um sistema, um componente ou um aplicativo falha. Embora não seja instantânea, a restauração de serviço é rápida, geralmente dentro de 30 a 300 segundos.

Em um cluster do PowerHA SystemMirror, para assegurar a disponibilidade desses aplicativos, os aplicativos são colocados sob o controle do PowerHA SystemMirror. O PowerHA SystemMirror toma medidas para assegurar que os aplicativos permanecem disponíveis para os processos do cliente, mesmo se um componente em um cluster falhar. Para assegurar a disponibilidade, no caso de falha de componente, o PowerHA SystemMirror move o aplicativo (juntamente com os recursos que asseguram acesso ao aplicativo) para outro nó no cluster.

Alta disponibilidade e disponibilidade de hardware

A *alta disponibilidade* é, às vezes, confundida com a disponibilidade de hardware simples. Sistemas redundantes, tolerantes à falha (como o RAID) e tecnologias de comutador dinâmicas (como o DLPAR) fornecem a recuperação de determinadas falhas de hardware, mas não fornecem o escopo completo de detecção e recuperação necessário para manter um aplicativo complexo altamente disponível.

Um aplicativo moderno e complexo requer acesso a todos estes componentes:

- Nós (CPU, memória)
- Interfaces de rede (incluindo os dispositivos externos na topologia de rede)
- Dispositivos de disco ou de armazenamento.

Pesquisas recentes sobre as causas de tempo de inatividade mostram que as falhas de hardware reais são responsáveis por somente uma pequena porcentagem das indisponibilidades não planejadas. Outros fatores de contribuição incluem:

- Erros do operador
- Problemas ambientais
- Erros do aplicativo e do sistema operacional.

Hardwares confiáveis e recuperáveis simplesmente não podem proteger contra falhas de todos estes aspectos diferentes da configuração. Manter esses elementos variados e, portanto, o aplicativo altamente disponível requer:

- Planejamento total e completo dos procedimentos físicos e lógicos para acesso e operação dos recursos dos quais o aplicativo depende. Esses procedimentos ajudam a evitar falhas, em primeiro lugar.
- Um pacote de monitoramento e recuperação que automatiza a detecção e recuperação de erros.
- Um processo bem controlado para manter os aspectos de hardware e software da configuração de cluster, enquanto mantém o aplicativo disponível.

Alta disponibilidade versus tolerância a falhas

A diferença entre tolerância a falhas e alta disponibilidade, é esta: Um ambiente tolerante a falhas não possui interrupção de serviço, mas possui um custo significativamente mais alto, enquanto um ambiente altamente disponível possui um mínimo de interrupção de serviço.

A tolerância a falhas baseia-se em hardware especializado para detectar uma falha de hardware e alternar instantaneamente para um componente de hardware redundante, seja o componente com falha um processador, uma placa de memória, uma fonte de alimentação, um subsistema de E/S ou um subsistema de armazenamento. Embora essa transição seja aparentemente fácil e ofereça serviço sem interrupção, um valor elevado é gasto para cobrir os custos de hardware e desempenho, porque os componentes redundantes não executam processamento. Mais importante ainda, o modelo tolerante a falhas não abrange falhas de software, que são, de longe, o motivo mais comum para períodos de inatividade.

A alta disponibilidade entende disponibilidade não como uma série de componentes físicos replicados, mas como um conjunto de recursos compartilhados em todo o sistema que cooperam para assegurar os serviços essenciais. A alta disponibilidade combina software com hardware padrão de mercado para minimizar o tempo de inatividade, restaurando rapidamente os serviços essenciais quando um sistema, um componente ou um aplicativo falha. Embora não seja instantâneo, os serviços são restaurados rapidamente, muitas vezes em menos de um minuto.

Muitos sites estão dispostos a absorver uma pequena quantidade de tempo de inatividade com a alta disponibilidade em vez de pagar os custos muito mais altos do fornecimento de tolerância a falhas. Além disso, na maioria das configurações altamente disponíveis, os processadores de backup estão disponíveis para uso durante a operação normal.

Os sistemas de alta disponibilidade são uma excelente solução para aplicativos que devem ser restaurados rapidamente e podem resistir a uma breve interrupção caso ocorra uma falha. Alguns segmentos de mercado possuem aplicativos que dependem tanto do tempo que não podem resistir nem mesmo a alguns segundos de tempo de inatividade. No entanto, muitos outros segmentos de mercado podem resistir a pequenos períodos de tempo em que o banco de dados fica indisponível. Para esses segmentos de mercado, o PowerHA SystemMirror pode proporcionar a continuidade de serviço necessária sem redundância total.

Função do PowerHA SystemMirror

O PowerHA SystemMirror possui muitos benefícios.

O PowerHA SystemMirror ajuda no seguinte:

- O processo de planejamento e a documentação do PowerHA SystemMirror incluem dicas e conselhos sobre as melhores práticas para a instalação e a manutenção de um cluster altamente disponível do PowerHA SystemMirror.
- Quando o cluster está operacional, o PowerHA SystemMirror fornece monitoramento e recuperação automatizados para todos os recursos dos quais o aplicativo depende.
- O PowerHA SystemMirror fornece um conjunto completo de ferramentas para manter o cluster enquanto mantém o aplicativo disponível aos clientes.

O PowerHA SystemMirror permite:

- Configurar um cluster de dois nós básico rápida e facilmente usando o caminho SMIT de configuração de cluster inicial típico ou os assistentes de configuração de aplicativo (Smart Assists).
- Testar a configuração do PowerHA SystemMirror usando a Ferramenta de teste do cluster. É possível avaliar como um cluster se comporta em um conjunto de circunstâncias especificadas, como quando um nó fica inacessível, uma rede fica inacessível e assim por diante.
- Assegurar a alta disponibilidade de aplicativos, eliminando pontos únicos de falha em um ambiente do PowerHA SystemMirror.
- Alavancar os recursos de alta disponibilidade disponíveis no AIX.
- Gerenciar como um cluster manipula falhas de componente.
- Proteger comunicações do cluster.
- Monitorar componentes do PowerHA SystemMirror e diagnosticar os problemas que possam ocorrer.

Informações relacionadas:

Clusters de aplicativo

Um *cluster de aplicativo* é um grupo de máquinas fracamente acopladas juntas em rede, compartilhando recursos de disco.

Em um cluster, várias máquinas servidor cooperam para fornecer um conjunto de serviços ou recursos aos clientes.

A clusterização de dois ou mais servidores para fazer o backup de aplicativos críticos é uma opção de alta disponibilidade com custo reduzido. É possível usar mais da potência de computação do site e ainda assegurar que os aplicativos críticos continuem as operações após uma interrupção mínima causada por uma falha de hardware ou software.

Os clusters de aplicativo também fornecem um caminho de crescimento gradual e escalável. É fácil incluir um processador no cluster para compartilhar a carga de trabalho em crescimento. Também é possível fazer upgrade de um ou mais dos processadores no cluster para obter um modelo mais poderoso. Se você estava usando uma estratégia tolerante a falhas, você deverá incluir *dois* processadores, um como um backup redundante que não executa o processamento durante operações normais.

Uso do Cluster Aware AIX pelo PowerHA SystemMirror

O PowerHA SystemMirror complementa os recursos de armazenamento em cluster principais suportados no sistema operacional AIX. O PowerHA SystemMirror é suportado para todas as edições do AIX que suportam os recursos do Cluster Aware AIX (CAA).

As informações a seguir são sobre os principais componentes do Cluster Aware AIX que são usados como a base para construir uma pilha da solução do PowerHA SystemMirror:

Gerenciamento de pulsação

Por padrão, o PowerHA SystemMirror usa comunicações unicast para pulsação. Como alternativa, as comunicações multicast podem ser configuradas em vez do unicast. Para multicast, é possível selecionar opcionalmente um endereço multicast ou permitir que o Cluster Aware AIX (CAA) designe um automaticamente. É possível especificar um endereço multicast ao configurar o cluster ou permitir que o Cluster Aware AIX (CAA) defina o multicast durante a configuração com base no ambiente de rede. A comunicação do cluster é atingida pela comunicação por meio de vários caminhos de comunicação redundantes. Os seguintes caminhos de comunicação redundantes fornecem uma base de cluster robusta que pode não estar sujeita ao particionamento de cluster:

Redes TCP/IP

O PowerHA SystemMirror e o Cluster Aware AIX usam todas as interfaces de rede que estão disponíveis para comunicação do cluster. Todas essas interfaces são descobertas, por padrão, e usadas para gerenciamento de funcionamento e outras comunicações de cluster. É possível usar as interfaces de gerenciamento do PowerHA SystemMirror para remover as interfaces que você não deseja que sejam usadas para a disponibilidade do aplicativo. Também é possível definir as interfaces que você não deseja que sejam usadas como interfaces privadas com o PowerHA SystemMirror.

Comunicação baseada em SAN

O CAA suporta a comunicação do cluster baseada em fábrica de rede de área de armazenamento (SAN) para um número limitado de adaptadores. Este tipo de pulsação é opcional e pode não funcionar com a maioria dos ambientes, devido a necessidades de zoneamento da rede que permitem que os pacotes viajem de cliente para cliente usando o protocolo Small Computer System Interface (SCSI).

Comunicação baseada em repositório do cluster central

O funcionamento e outras comunicações do cluster também são arquivados no disco do

| armazenador central. PowerHA SystemMirror 7.2.0, ou posterior, fornece uma função
| Automatic Repository Disk Replacement (ARR) que substitui automaticamente um disco
| de repositório com falha por um disco de repositório de backup. A função ARR está
| disponível apenas se configurar e identificar um disco de repositório de backup com o
| PowerHA SystemMirror.

Gerenciamento de evento aprimorado

O CAA gera eventos de armazenamento e de rede de granularidade fina usados pelo PowerHA SystemMirror para fornecer uma melhor capacidade de tomada de decisão para o gerenciamento de alta disponibilidade.

Gerenciamento de armazenamento entre nós

O PowerHA SystemMirror usa os recursos de fence de armazenamento do AIX para obter um gerenciamento de armazenamento melhor entre os nós no cluster. Os recursos de fence são suportados apenas para discos configurados com AIX Multipath I/O (MPIO) nativo. PowerHA SystemMirror gerencia discos compartilhados por meio do método de gerenciamento de volume simultâneo aprimorado.

Nota: O PowerHA SystemMirror tenta usar o recurso de fence de estrutura de armazenamento do CAA para evitar que os discos compartilhados sejam acessados por nós que não possuem acesso ao grupo de volumes compartilhados proprietário. Esse recurso de fence evita a distorção de dados de dados devido a acesso acidental a discos compartilhados a partir de vários nós. Entretanto, o recurso de fence de estrutura de área de armazenamento de CAA é suportado apenas para o AIX MPIO nativo.

Conceitos relacionados:

“Soluções de vários sites do PowerHA SystemMirror” na página 20

O PowerHA SystemMirror suporta diferentes tipos de definições para sites e políticas específicas de site para recuperação de desastre de alta disponibilidade (HADR). É possível definir vários sites no PowerHA SystemMirror Standard Edition for AIX e no PowerHA SystemMirror Enterprise Edition for AIX.

“Cluster vinculado do PowerHA SystemMirror” na página 21

Um cluster vinculado é ideal para situações em que cada site está localizado em uma localização geográfica diferente. Geralmente, os sites estão distantes o suficiente para que não possam compartilhar facilmente uma rede de área de armazenamento (SAN) comum. Cada site deve ter seu próprio disco de repositório ativo e quaisquer repositórios de backup. Os clusters vinculados sempre usam unicast para comunicar-se entre os sites. Os clusters vinculados são uma parte importante da recuperação de desastre de alta disponibilidade (HADR).

“Cluster estendido do PowerHA SystemMirror” na página 22

Um cluster estendido é ideal para situações em que cada site está localizado na mesma localização geográfica. Geralmente, todos os nós estão em uma rede de área de armazenamento (SAN) comum. No mínimo, o repositório ativo e os repositórios de backup são compartilhados entre os sites, assim como os discos em grupos de recursos que podem ser executados em qualquer site. Os clusters estendidos podem suportar o espelhamento do LVM entre sites e o HyperSwap.

“Discos de repositório do cluster” na página 12

O PowerHA SystemMirror usa um disco compartilhado como um armazenador central para gerenciar a configuração do cluster. Esse disco deve ser acessível por todos os nós no cluster usando o método de implementação padrão ou de cluster estendido ou deve ser acessível por todos os nós em um site usando o método de implementação do cluster vinculado.

“Gerenciamento de configuração central” na página 6

Um elemento chave da configuração do cluster é o disco de repositório de cluster. O disco de repositório de cluster é usado como o armazenador central para os dados de configuração do cluster.

Informações relacionadas:

Configurando a comunicação de armazenamento em cluster

Gerenciamento de configuração central

Um elemento chave da configuração do cluster é o disco de repositório de cluster. O disco de repositório de cluster é usado como o armazenador central para os dados de configuração do cluster.

O disco de repositório de cluster deve estar acessível a partir de todos os nós no cluster e deve ter no mínimo 512 MB e menos que 460 GB para uma configuração de cluster do PowerHA SystemMirror. Ele deve ter como backup uma configuração da SAN redundante e altamente disponível. O disco de armazenador central deve ser configurado de forma apropriada para o Redundant Array of Independent Disks (RAID) para evitar que haja um ponto único de falha.

O disco de repositório de cluster é um dispositivo especial para o cluster. Ele possui características especiais e é gerenciado exclusivamente pelo Cluster Aware AIX (CAA) e o PowerHA SystemMirror. Você não deve usar diretamente os comandos e as interfaces de LVM ou de sistema de arquivos porque eles podem causar danos no disco de repositório de cluster. O disco de repositório de cluster é gerenciado pelo CAA de forma privada. O disco de repositório de cluster não deve ser administrado fora do ambiente do CAA ou do ambiente do PowerHA SystemMirror.

O repositório de cluster destina-se aos seguintes propósitos para gerenciamento de cluster:

- A configuração do cluster do CAA é mantida e gerenciada de forma centralizada no disco de repositório.
- Para a comunicação baseada em cluster, uma parte do disco de repositório é reservada para a comunicação de pulsação e mensagem de nó a nó.

Nota: Essa forma de comunicação é usada quando todas as outras formas de comunicação falham.

Conceitos relacionados:

“Uso do Cluster Aware AIX pelo PowerHA SystemMirror” na página 4

O PowerHA SystemMirror complementa os recursos de armazenamento em cluster principais suportados no sistema operacional AIX. O PowerHA SystemMirror é suportado para todas as edições do AIX que suportam os recursos do Cluster Aware AIX (CAA).

Configurando a comunicação da SAN do cluster

O PowerHA SystemMirror e o Cluster Aware AIX suportam um método de comunicação adicional entre os hosts chamado de comunicação da rede de área de armazenamento (SAN).

As comunicações da SAN exploram os links físicos da SAN que existem em um datacenter típico para fornecer comunicações de alta velocidade entre os nós do cluster. Este mecanismo de comunicações fornece uma função de fallback para o caso de ocorrência de falha nos caminhos de comunicações de rede entre os nós do cluster.

A comunicação da SAN funciona para um conjunto de adaptadores de disco. É necessário que haja a configuração de zoneamento da SAN para que os pacotes Small Computer System Interface (SCSI) fluam de host a host por meio de vários comutadores entre cada host.

A configuração da comunicação da SAN envolve a chamada de um conjunto de comandos no adaptador de disco suportado para que ele fique apto para a comunicação da SAN.

No VIOS 2.2.0.11, ou posterior, é possível usar a comunicação da SAN entre as partições lógicas estabelecendo uma rede local virtual por meio de um adaptador Ethernet virtual em cada cliente VIOS. É possível configurar a comunicação da SAN por meio do VIOS para os ambientes NPIV e vSCSI.

Informações relacionadas:

Configurando a comunicação de armazenamento em cluster

Definindo um adaptador Ethernet virtual

Custos e benefícios contínuos da disponibilidade

A disponibilidade do PowerHA SystemMirror tem muitos custos e benefícios.

A figura a seguir mostra os custos e benefícios das tecnologias de disponibilidade.

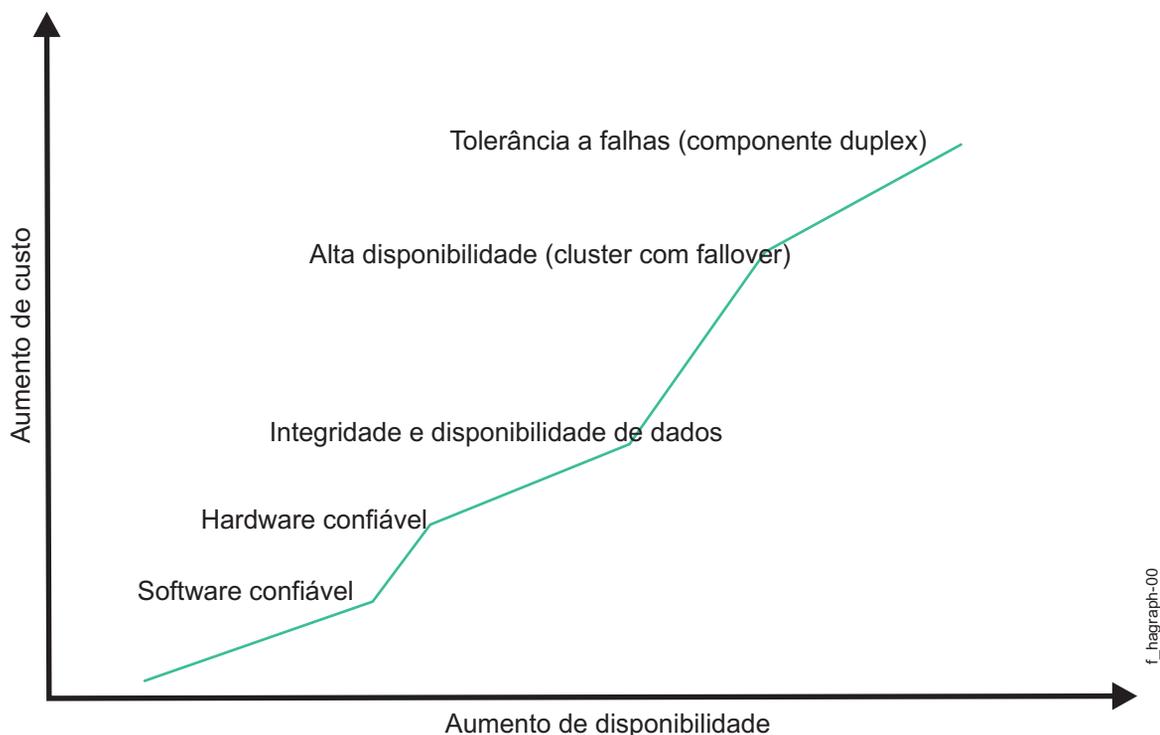


Figura 1. Custos e benefícios das tecnologias disponíveis

Como você pode ver, a disponibilidade não é uma proposição de tudo ou nada. Pense em disponibilidade como um contínuo. Hardwares e softwares confiáveis fornecem o nível básico de disponibilidade. Recursos avançados, como dispositivos RAID, fornecem um nível avançado de disponibilidade. O software de alta disponibilidade fornece acesso quase contínuo a dados e aplicativos. Os sistemas tolerantes a falhas asseguram a disponibilidade constante do sistema inteiro, mas a um custo mais alto.

Disponibilidade aprimorada com o software AIX

O PowerHA SystemMirror usufrui dos recursos no AIX - o sistema operacional UNIX de alto desempenho. O AIX inclui uma nova funcionalidade para aprimorar ainda mais a segurança e a disponibilidade do sistema.

Isso inclui uma melhora na disponibilidade de dados espelhados e aprimoramentos no Workload Manager que ajudam a resolver problemas de cargas de trabalho combinadas, fornecendo dinamicamente a disponibilidade de recursos para aplicativos críticos. O PowerHA SystemMirror fornece escalabilidade horizontal e vertical sem tempo de inatividade para o sistema. O sistema operacional AIX fornece inúmeros recursos designados para aumentar a disponibilidade do sistema, diminuindo o impacto dos tempos de inatividade planejados (backup de dados, administração do sistema) e não planejados (falha de hardware ou software), e a flexibilidade no gerenciamento de recursos de hardware, usando as funções de Capacidade on Demand (CoD).

O sistema operacional AIX fornece os recursos a seguir:

- Sistema de arquivos registrados e sistema de arquivos registrados aprimorado

- Espelhamento de disco
- Controle de processo
- DLPAR e CoD
- Partições de carga de trabalho

Sistema de arquivos registrados e sistema de arquivos registrados aprimorado

O sistema de arquivos nativo do AIX, o sistema de arquivos registrados (JFS), usa técnicas de registro de banco de dados para manter sua integridade estrutural. Falhas do sistema ou de software não deixam o sistema de arquivos em uma condição não gerenciável. Ao reconstruir o sistema de arquivos após uma falha grave, o AIX usa o log do JFS para restaurar o sistema de arquivos para seu último estado consistente. O registro fornece uma recuperação mais rápida do que o utilitário de verificação de consistência de sistema de arquivos (fsck) padrão do UNIX. Além disso, o JFS2 (Sistema de arquivos registrados) está disponível no AIX.

Espelhamento de disco

O software de espelhamento de disco fornece integridade de dados e recursos de backup online. Ele evita a perda de dados devido a falha do disco, mantendo até três cópias dos dados em discos separados para que os dados ainda permaneçam acessíveis após a falha de qualquer um dos discos. O espelhamento de disco é imperceptível para o aplicativo. Nenhuma modificação de aplicativo é necessária porque os discos espelhados e convencionais parecem iguais para o aplicativo.

Controle de processo

O AIX System Resource Controller (SRC) monitora e controla os processos principais. O SRC pode detectar quando um processo é encerrado de forma anormal, registrar o encerramento, transmitir mensagens para um programa de notificação e reiniciar o processo em um processador de backup.

Gerenciamento de dynamic LPAR

| PowerHA SystemMirror pode mover recursos de aplicativos entre LPARs e pode executar os ajustes de recurso dinâmicos necessários por meio da função Resource Optimized High Availability (ROHA). O ROHA usa os recursos que estão disponíveis com os Power Systems para gerenciar dinamicamente os tipos a seguir de recursos de hardware:

- | • As funções de Capacidade on Demand (CoD) (incluindo CoD On/Off e Enterprise Pool CoD) gerenciam recursos de memória e CPU no nível de estrutura (CEC).
- | • As funções da DLPAR gerenciam recursos de memória e CPU no nível de partição lógica.

| Com a função do ROHA, é possível usar o PowerHA SystemMirror para otimizar a quantidade de recursos para cada aplicativo. Por exemplo, durante um controle, os recursos de hardware (CPU e memória) são dinamicamente liberados do nó ativo e são dinamicamente adquiridos e alocados pelo nó de espera.

Partições de carga de trabalho

As partições de carga de trabalho permitem que várias instâncias de um aplicativo sejam executadas no mesmo ambiente de sistema operacional e ainda fornecem isolamento entre esses ambientes, oferecendo assim proteção e isolamento entre as instâncias.

Cluster Aware AIX (CAA)

O PowerHA SystemMirror usa os recursos de cluster do AIX para fornecer uma solução de alta disponibilidade avançada. O CAA fornece o kernel do AIX com gerenciamento de pulsação e de funcionamento. O PowerHA SystemMirror monitora eventos de armazenamento e rede de granularidade fina e manipula as situações críticas no sistema operacional AIX. O PowerHA SystemMirror pode

descobrir componentes de hardware, tornando mais fácil gerenciar e implementar clusters.

Componentes físicos do cluster do PowerHA SystemMirror

O PowerHA SystemMirror fornece um ambiente altamente disponível, identificando um conjunto de recursos essenciais para o processamento sem interrupção. Ele também define um protocolo que os nós usam para colaborar a fim de assegurar que esses recursos estejam disponíveis.

O PowerHA SystemMirror estende o modelo de armazenamento em cluster, definindo relacionamentos entre processadores cooperativos, em que um processador fornece o serviço oferecido por um peer, caso o peer não o possa fornecer. Conforme é mostrado na figura a seguir, um cluster do PowerHA SystemMirror é composto pelos seguintes componentes físicos:

- Nós
- Dispositivos de disco externos compartilhados
- Redes
- Interfaces de rede
- Disco de repositório do cluster
- Clientes.

O software PowerHA SystemMirror permite combinar componentes físicos em uma ampla variedade de configurações de cluster, fornecendo flexibilidade na construção de um cluster que atenda os requisitos de processamento. Esta figura mostra um exemplo de cluster do PowerHA SystemMirror. Outros clusters do PowerHA SystemMirror podem ser muito diferentes, dependendo do número de processadores, da opção de rede e de tecnologia de disco e assim por diante.

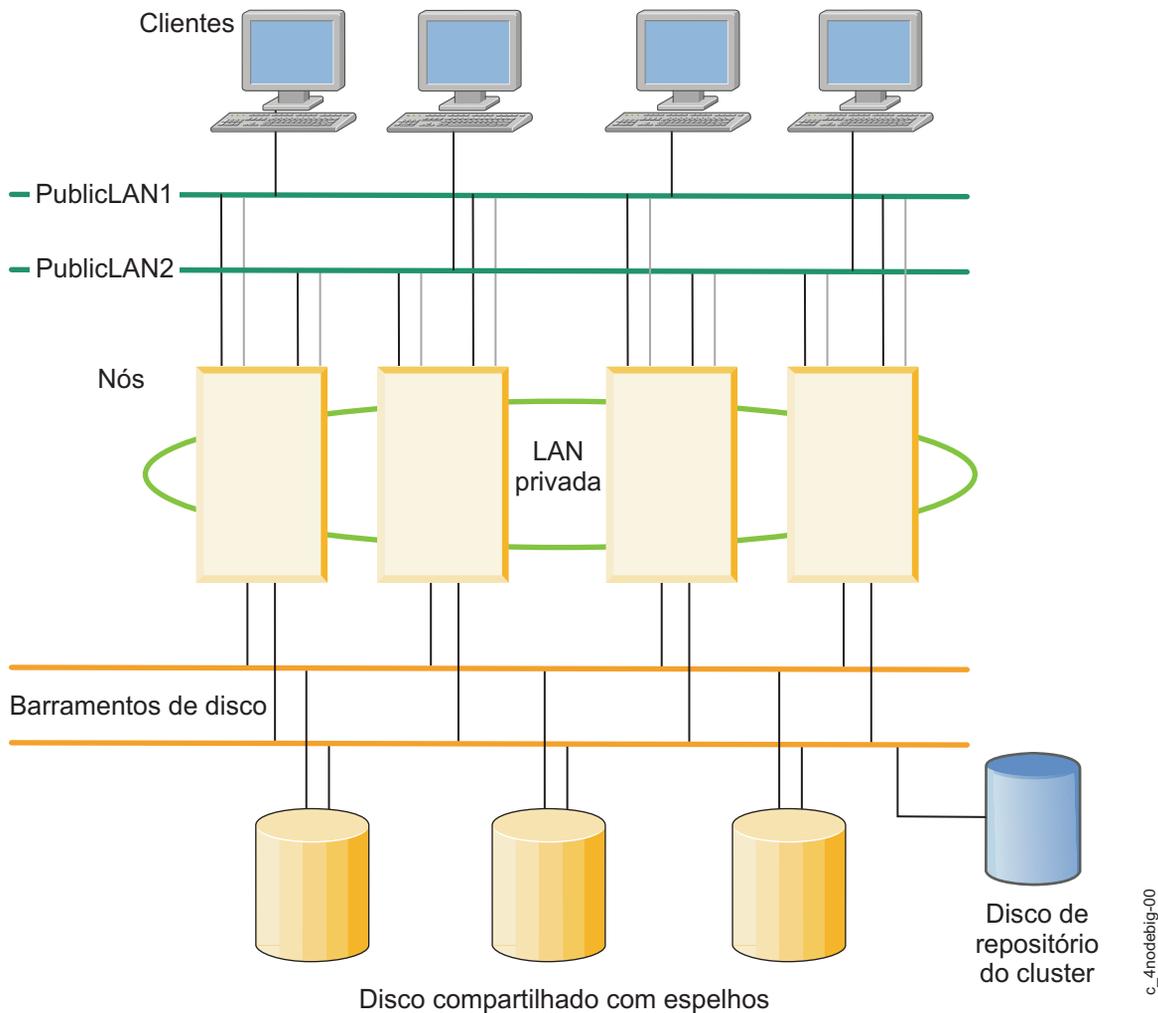


Figura 2. Exemplo de um cluster do PowerHA SystemMirror

As configurações de cluster do PowerHA SystemMirror fornecem exemplos de configurações de cluster suportadas pelo software PowerHA SystemMirror.

Conceitos relacionados:

“Configurações de cluster do PowerHA SystemMirror” na página 76

Este capítulo fornece exemplos dos tipos de configurações de cluster suportadas pelo software PowerHA SystemMirror.

Nós do PowerHA SystemMirror

Os nós formam o núcleo de um cluster do PowerHA SystemMirror. Um nó é um processador que executa o AIX, o software PowerHA SystemMirror e o software de aplicativo.

A clusterização desses servidores para fazer o backup de aplicativos críticos é uma opção de alta disponibilidade com custo reduzido. Uma empresa pode usar mais da sua potência de computação e ainda assegurar que seus aplicativos críticos continuem as operações após uma interrupção breve causada por uma falha de hardware ou software.

Em um cluster do PowerHA SystemMirror, cada nó é identificado por um nome exclusivo. Um nó pode possuir um conjunto de recursos: discos, grupos de volumes, sistemas de arquivos, redes, endereços de rede e aplicativos. Geralmente, um nó executa um servidor ou um aplicativo backend que acessa dados nos discos externos compartilhados.

Conceitos relacionados:

“Assegurando a disponibilidade do aplicativo” na página 54

Esta seção descreve como o software PowerHA SystemMirror assegura a disponibilidade do aplicativo por assegurar a disponibilidade dos componentes do cluster. O PowerHA SystemMirror elimina pontos únicos de falha para todos os principais componentes do sistema e elimina a necessidade de tempo de inatividade planejado para a maioria das tarefas de manutenção de rotina.

“Recursos e grupos de recursos do PowerHA SystemMirror” na página 32

Consulte aqui os conceitos e as definições relacionados a recursos que são usados em toda a documentação e também na interface com o usuário do PowerHA SystemMirror.

Dispositivos de disco externos compartilhados

Cada nó tem acesso a um ou mais dispositivos de disco externos compartilhados. Um *dispositivo de disco externo compartilhado* é um disco conectado fisicamente a vários nós.

O disco compartilhado armazena dados essenciais, geralmente espelhados ou configurados para RAID para redundância de dados. Um nó em um cluster do PowerHA SystemMirror também deve ter discos internos que armazenam o sistema operacional e os binários de aplicativo, mas esses discos não são compartilhados. Dependendo do tipo de disco usado, o software PowerHA SystemMirror suporta os seguintes tipos de acesso a dispositivos de disco externos compartilhados: o acesso não simultâneo e o acesso simultâneo.

- Em *ambientes de acesso não simultâneo*, somente uma conexão está ativa em um determinado momento e o nó com a conexão ativa possui o disco. Quando um nó falha, o nó que atualmente possui o disco é removido do cluster, ocorre o controle de disco e um nó sobrevivente assume a propriedade do disco compartilhado. Esta configuração de cluster típica é usada pela maioria dos aplicativos.
- Em *ambientes de acesso simultâneo*, os discos compartilhados são conectados ativamente a mais de um nó simultaneamente. Portanto, quando um nó falha, o controle de disco não é necessário. Este tipo de acesso é usado somente por aplicativos que podem gerenciar e coordenar o acesso simultâneo a dados compartilhados a partir de vários nós.

Observe que, nesses ambientes, todos os nós definidos no cluster podem fazer parte do acesso simultâneo ou somente um subconjunto de nós do cluster pode ter acesso a discos compartilhados. No segundo caso, você configura os grupos de recursos somente nos nós que possuem acesso ao disco compartilhado. As diferenças entre esses métodos são explicadas mais detalhadamente na seção Acesso a disco externo compartilhado.

Conceitos relacionados:

“Acesso a disco externo compartilhado” na página 51

O software PowerHA SystemMirror suporta dois métodos de acesso a disco externo compartilhado: não simultâneo e simultâneo.

Redes

Como um componente independente, em camadas do Sistema operacional AIX, o software PowerHA SystemMirror é projetado para trabalhar com qualquer rede baseada em TCP/IP.

Os nós em um cluster do PowerHA SystemMirror usam a rede para:

- Permitir que os clientes acessem os nós do cluster
- Permitir que os nós do cluster troquem mensagens de pulsação
- Serializar o acesso a dados (em ambientes de acesso simultâneo)

O software PowerHA SystemMirror foi testado com redes Ethernet e redes de área de armazenamento.

Tipos de redes

O software PowerHA SystemMirror define dois tipos de redes de comunicação, caracterizados pelo uso de interfaces de comunicação com base no subsistema TCP/IP (baseado em TCP/IP) ou nas redes de área de armazenamento.

Rede baseada em TCP/IP

Conecta dois ou mais nós de servidor e, opcionalmente, permite que o cliente acesse esses nós do cluster, usando o protocolo TCP/IP. Por padrão, o PowerHA SystemMirror usa comunicações unicast para pulsação. Opcionalmente, será possível selecionar o uso de comunicações multicast se a rede estiver configurado para suportar multicast.

Rede de área de armazenamento (SAN)

Fornece comunicação entre nós do cluster do PowerHA SystemMirror para controle de mensagem e de tráfego de pulsação. Este é um caminho de comunicações alternativo para a rede TCP/IP.

| Discos de repositório do cluster

| O PowerHA SystemMirror usa um disco compartilhado como um armazenador central para gerenciar a configuração do cluster. Esse disco deve ser acessível por todos os nós no cluster usando o método de implementação padrão ou de cluster estendido ou deve ser acessível por todos os nós em um site usando o método de implementação do cluster vinculado.

| Use a interface do System Management Interface Tool (SMIT) para gerenciar e configurar o disco do repositório do cluster e os discos de repositório de backup.

| Deve-se ter pelo menos um disco de repositório ativo por cluster para clusters padrão e clusters estendidos. É possível identificar até seis discos de repositório de backup por cluster para clusters padrão e clusters estendidos. Deve-se ter um disco de repositório ativo por site para clusters vinculados. É possível identificar até seis discos de repositório de backup por site para clusters vinculados.

| Com o Cluster Aware AIX, o disco do repositório do cluster é usado para os propósitos a seguir:

- | • Gerenciamento de configuração em todo o cluster
- | • Sistema de mensagens e pulsação do cluster. O disco de repositório é usado como outro caminho redundante da comunicação entre os nós.

| Deve-se ter pelo menos 512 MB e não mais de 460 GB de espaço em disco alocado para o disco do repositório do cluster.

| Verifique se o disco que selecionou como o disco de repositório não tem uma política de reserva. Para verificar a política de reserva de discos, execute o comando **lsattr -El hdisk -a reserve_policy**. Para alterar a política de reserva do disco para no_reserve, execute o comando **chdev -a reserve_policy=no_reserve -l hdisk**.

| Após designar um disco como o disco de repositório, o disco não pode ser usado para nenhum outro propósito. Verifique se o disco que usou como o disco de repositório não contém dados do usuário. Quando o disco é identificado como um disco de repositório pelo PowerHA SystemMirror, todas as informações no disco são apagadas.

| Ao planejar usar os discos como discos de repositório, deve-se planejar discos de backup ou substituição, que podem ser usados no caso de o disco de repositório primário falhar. O disco de backup deve ter o mesmo tamanho e tipo que o disco primário, mas pode estar em um disco de armazenamento físico diferente. Atualize os procedimentos administrativos e a documentação com as informações do disco de backup. O cluster tolera a perda ou falha do disco de repositório. As operações críticas continuam sem um disco de repositório. Entretanto, deve-se corrigir rapidamente o problema com o disco de repositório. Se você estiver usando o PowerHA SystemMirror 7.2.0, ou anterior, deve-se substituir manualmente o disco de repositório com falha por um disco de repositório da lista de backup. Também é possível

| substituir um disco de repositório em funcionamento por um novo disco para aumentar o tamanho ou
| alterar para um subsistema de armazenamento diferente.

| É possível configurar o PowerHA SystemMirror 7.2.0, ou posterior, para usar a função Automatic
| Repository Disk Replacement (ARR) no Cluster Aware AIX (CAA) durante a falha de um disco de
| repositório ativo. A função ARR substitui automaticamente o disco de repositório com falha por um disco
| a partir dos discos de repositório de backup. O primeiro disco de repositório de backup na lista substitui
| o disco de repositório com falha.

| Deve-se configurar todos os discos de repositório com AIX Multiple Path I/O (MPIO) nativo. O AIX
| MPIO ativa o PowerHA SystemMirror para monitorar e responder melhor a falhas de disco usando os
| discos de repositório.

| **Conjuntos de reflexos e discos de repositório**

| Se você usar o PowerHA SystemMirror and Logical Volume Manager (LVM) para gerenciar conjuntos de
| reflexos entre diferentes locais e o disco de repositório falhar, o novo disco de repositório não deverá ser
| do mesmo local onde a falha ocorreu.

| **Conceitos relacionados:**

| “Cluster vinculado do PowerHA SystemMirror” na página 21

| Um cluster vinculado é ideal para situações em que cada site está localizado em uma localização
| geográfica diferente. Geralmente, os sites estão distantes o suficiente para que não possam compartilhar
| facilmente uma rede de área de armazenamento (SAN) comum. Cada site deve ter seu próprio disco de
| repositório ativo e quaisquer repositórios de backup. Os clusters vinculados sempre usam unicast para
| comunicar-se entre os sites. Os clusters vinculados são uma parte importante da recuperação de desastre
| de alta disponibilidade (HADR).

| “Cluster estendido do PowerHA SystemMirror” na página 22

| Um cluster estendido é ideal para situações em que cada site está localizado na mesma localização
| geográfica. Geralmente, todos os nós estão em uma rede de área de armazenamento (SAN) comum. No
| mínimo, o repositório ativo e os repositórios de backup são compartilhados entre os sites, assim como os
| discos em grupos de recursos que podem ser executados em qualquer site. Os clusters estendidos podem
| suportar o espelhamento do LVM entre sites e o HyperSwap.

| “Uso do Cluster Aware AIX pelo PowerHA SystemMirror” na página 4

| O PowerHA SystemMirror complementa os recursos de armazenamento em cluster principais suportados
| no sistema operacional AIX. O PowerHA SystemMirror é suportado para todas as edições do AIX que
| suportam os recursos do Cluster Aware AIX (CAA).

| **Informações relacionadas:**

| Planejando o disco de repositório

| Falha de disco de repositório

| Substituindo um disco de repositório pelo SMIT

| **Discos de repositório e de caminhos múltiplos:**

| O Sistema operacional AIX gerencia discos de repositório de uma maneira especial durante a
| configuração de um cluster.

| Os serviços de cluster do AIX reconhecem o disco de repositório durante o processo de configuração do
| dispositivo do AIX e podem executar ações especiais no disco. Por exemplo, o disco do repositório é
| usado pelos serviços de cluster para ativar a comunicação baseada em disco entre hosts. A comunicação
| de disco envolve um host que grava no disco e o outro host ou outros hosts que recuperam a mensagem
| por meio de um mecanismo de pesquisa.

| Os serviços de cluster também geram eventos detalhadas em relação ao funcionamento do disco de
| repositório. Se um disco de repositório falhar, o CAA usará uma função Automatic Repository

| Replacement (ARR) que substitui automaticamente um disco de repositório com falha por um repositório de backup. A função ARR está disponível apenas se identificar um disco de repositório de backup com o PowerHA SystemMirror.

| É possível configurar os tipos de discos a seguir como discos de repositório:

| **Discos de caminhos múltiplos do AIX (AIX local MPIO)**

| Esses discos podem ser criados automaticamente e usados como discos de repositório. Este é o tipo recomendado de disco para usar como um disco de repositório.

| **Disco SCSI virtual (vSCSI)**

| Esses discos podem ser mapeados por meio do Virtual I/O Server (VIOS) como disco vSCSI para a partição lógica cliente (LPAR). Quando o disco vSCSI é usado como um disco de repositório, ele deve ser mapeado para um nó diferente do PowerHA SystemMirror ou um nó diferente do Cluster Aware AIX (CAA) usando os mesmos métodos vSCSI.

| **Discos de caminhos múltiplos de terceiros**

| Esses discos seguem as recomendações dos conceitos de caminhos múltiplos do AIX, mas fornecem seus próprios drivers e softwares de dispositivo de caminhos múltiplos. Esses discos podem ser configurados como discos de repositório quando as informações relevantes do Gerenciador de Dados do Objeto (ODM) estão disponíveis. Os discos gerenciados pelo EMC PowerPath e o Hitachi Dynamic Link Manager (HDLM) pode ser configurados como discos de repositório usando esse método.

| Para atualizar o software MPIO ou software de caminhos múltiplos de terceiro, deve-se parar os serviços de cluster CAA inserindo o comando **clmgr offline cluster STOP_CAA=yes**.

| **Falha de disco de repositório:**

| Deve-se planejar corretamente a falha de disco de repositório e o que é necessário para corrigir problemas relacionados à falha de disco de repositório.

| A falha de disco de repositório é tolerada por um cluster do PowerHA SystemMirror. Se qualquer nó no cluster encontrar erros com o disco de repositório ou ao acessar o disco de repositório, o cluster entra em um modo limitado ou restrito. Nesse modo de operação, não é possível usar a maioria das operações relacionadas a topologia. Por exemplo, um nó não pode ser incluído ou adicionado no cluster. Entretanto, as funções críticas do cluster podem ser executadas. Por exemplo, é possível mover um grupo de recursos de um nó ativo para um nó de espera.

| Quando o disco de repositório falha, o administrador é notificado sobre a falha de disco. O PowerHA SystemMirror continua a notificar o administrador sobre a falha do disco de repositório até que seja resolvida.

| PowerHA SystemMirror e Cluster Aware AIX (CAA) suportam a substituição de disco de repositório em tempo real, que você pode usar para substituir um disco de repositório com falha ou em funcionamento. O CAA preenche novamente o novo disco com informações do cluster e inicia o uso do disco como disco de repositório.

| O PowerHA SystemMirror 7.2.0, ou posterior, suporta a função Automatic Repository Disk Replacement (ARR). ARR usa CAA e substitui automaticamente um disco de repositório com falha por um disco de repositório de backup. A função ARR está disponível apenas se configurar um disco de repositório de backup com PowerHA SystemMirror.

| Para usar a função ARR, seu ambiente deve atender aos requisitos a seguir:

- | • Um cluster ou site tem um disco de repositório de backup identificado.
- | • PowerHA SystemMirror Versão 7.2.0, ou posterior, está instalado.
- | • Uma das versões a seguir do sistema operacional do AIX está instalada:

- | - AIX Versão 7.1.4, ou posterior
- | - AIX Versão 7.2.0, ou posterior

| O CAA monitora a falha do disco de repositório verificando erros de E/S e verificando se o disco está no estado ativo. Essas verificações ocorrem periodicamente e não são executadas toda vez que o disco de repositório é lido ou gravado. Não grave diretamente no disco de repositório, mesmo que para propósitos de teste. A gravação direta no disco de repositório de forma assíncrona pode fazer com que o sistema operacional e as operações do CAA sejam interrompidas de forma brusca, gerando resultados imprevisíveis.

| **Informações relacionadas:**

- | Comando dd
- | Planejando o disco de repositório
- | Falha de disco de repositório
- | Substituindo um disco de repositório pelo SMIT

Cientes

Um cliente é um processador que pode acessar os nós em um cluster sobre uma rede local.

Cada cliente executa um aplicativo "front-end" ou cliente que consulta o aplicativo do servidor que está em execução no nó do cluster. O software PowerHA SystemMirror fornece um ambiente altamente disponível para dados e aplicativos críticos em nós do cluster. O software PowerHA SystemMirror não torna os próprios clientes altamente disponíveis. Os clientes do AIX podem usar os serviços de Informação do cluster (Clinfo) para receber aviso de eventos de cluster. O Clinfo fornece uma API que exibe informações de status do cluster. O PowerHA SystemMirror fornece um utilitário de status do cluster, o `/usr/es/sbin/cluster/clstat`. Ele é baseado no Clinfo e relata o status dos principais componentes do cluster: o próprio cluster, os nós no cluster, as interfaces de rede conectadas aos nós e os grupos de recursos em cada nó. A interface de status do cluster do utilitário `clstat` inclui as versões baseadas na web, baseadas em Motif e baseadas em ASCII.

Conceitos relacionados:

"Cluster information program" na página 50

O Cluster information program (Clinfo), o daemon `clinfo`, é um monitor baseado em SNMP.

Objetivo do PowerHA SystemMirror: Eliminar tempo de inatividade planejado

O objetivo principal do software de armazenamento em cluster de alta disponibilidade é minimizar ou, de modo ideal, eliminar a necessidade de colocar os recursos fora de serviço durante as atividades de manutenção e reconfiguração.

O software PowerHA SystemMirror otimiza a disponibilidade, permitindo a reconfiguração dinâmica de clusters em execução. A maioria das tarefas de manutenção de cluster de rotina, como incluir ou remover um nó ou alterar a prioridade de nós que participam de um grupo de recursos, pode ser aplicada a um cluster ativo sem parar e reiniciar os serviços de cluster. Além disso, é possível manter um cluster do PowerHA SystemMirror online ao fazer mudanças na configuração usando o recurso de Ponto único de controle do cluster (C-SPOC). O C-SPOC facilita o gerenciamento do cluster porque torna possível alterar grupos de volumes compartilhados, usuários e grupos em todo o cluster a partir de um único nó. As mudanças são propagadas de forma transparente para outros nós do cluster.

Conceitos de nós, redes e pulsação do cluster do PowerHA SystemMirror

Esta seção apresenta os principais conceitos relacionados à topologia de cluster e as definições usadas em toda a documentação e na interface com o usuário do PowerHA SystemMirror.

Nós

Um nó é um processador que executa o AIX e o software PowerHA SystemMirror.

Os nós podem compartilhar um conjunto de recursos como discos, grupos de volumes, sistemas de arquivos, redes, endereços IP de rede e aplicativos. O software PowerHA SystemMirror suporta de dois a trinta e dois nós em um cluster. Em um cluster do PowerHA SystemMirror, cada nó é identificado por um nome exclusivo. No PowerHA SystemMirror, um nome do nó e um nome do host, geralmente, podem ser iguais. Os nós funcionam como os componentes físicos principais de um cluster do PowerHA SystemMirror. Para obter mais informações sobre os nós e hardware, consulte a seção Nós. Dois tipos de nós são definidos:

- Os nós do servidor constituem o núcleo de um cluster do PowerHA SystemMirror. Os nós do servidor executam serviços ou aplicativos de backend que acessam dados nos discos externos compartilhados.
- Os nós clientes executam aplicativos de front-end que recuperam dados dos serviços fornecidos pelos nós do servidor. Os nós clientes podem executar o software PowerHA SystemMirror para monitorar o funcionamento dos nós e para reagir a falhas.

Nós do servidor

Um nó de servidor de cluster geralmente executa um aplicativo que acessa dados nos discos externos compartilhados. Os nós do Servidor executam daemons do PowerHA SystemMirror e mantêm recursos altamente disponíveis. Geralmente, os aplicativos são executados, o armazenamento é compartilhado entre esses nós e os clientes conectam-se aos nós do servidor por meio de um endereço IP de serviço.

Nós cliente

Uma solução de alta disponibilidade geralmente inclui a máquina cliente que usa os serviços fornecidos pelos servidores. Os nós clientes podem ser divididos em duas categorias: simples e inteligente.

- Um cliente simples visualiza o cluster como uma entidade única. Se um servidor falhar, o cliente deverá ser reiniciado ou pelo menos deverá reconectar-se ao servidor.
- Um cliente inteligente reconhece o cluster. Um cliente que reconhece o cluster reage corretamente diante de uma falha do servidor, conectando-se a um servidor alternativo, talvez mascarando a falha do usuário. Um cliente inteligente como esse deve ter conhecimento do estado do cluster.

O PowerHA SystemMirror amplia o paradigma de cluster aos clientes, fornecendo relatório de configuração do cluster dinâmico e notificação de mudanças de estado do cluster, como mudanças em subsistemas ou falha do nó.

Sites

É possível definir um grupo de um ou mais nós do servidor como pertencentes a um site.

O site torna-se um componente, como um nó ou uma rede, conhecido pelo software PowerHA SystemMirror. O PowerHA SystemMirror suporta clusters divididos em dois sites.

É possível configurar o espelhamento do Gerenciador de Volume Lógico (LVM) de divisão de site, associando conjuntos de espelhos específicos do LVM a um site físico. Se, em seguida, você especificar para o PowerHA SystemMirror, quais volumes físicos estão localizadas em cada site, o C-SPOC exibirá as informações do site ao selecionar volumes nos sites para os conjuntos de espelhos do LVM. Durante a verificação do cluster, o PowerHA SystemMirror executa verificações adicionais para assegurar-se de que as definições de espelho sejam consistentes com as definições do site.

Além disso, a função Distância estendida do PowerHA SystemMirror Enterprise Edition fornece duas soluções de software distintas para recuperação de desastre. Essas soluções permitem que um cluster do PowerHA SystemMirror opere em distâncias estendidas em dois sites.

O PowerHA SystemMirror Enterprise Edition para Metro Mirror aumenta a disponibilidade de dados para volumes DS8000, volumes DS6000 e volumes IBM TotalStorage Enterprise Storage Server (ESS) que usam Peer-to-Peer Remote Copy (PPRC) para copiar dados para um site remoto para propósitos de recuperação de desastre. O PowerHA SystemMirror Enterprise Edition para Metro Mirror usufrui das funções de fallover e fallback de PPRC e do gerenciamento de cluster do PowerHA SystemMirror para reduzir o tempo de inatividade e o tempo de recuperação durante a recuperação de desastre.

Quando o PPRC é usado para espelhamento de dados entre sites, a distância física entre os sites fica limitada aos recursos do hardware ESS.

O PowerHA SystemMirror Enterprise Edition para Gerenciador de Volume Lógico (GLVM) aumenta a disponibilidade de dados para os volumes IBM que usam GLVM para copiar dados para um site remoto para propósitos de recuperação de desastre. O PowerHA SystemMirror Enterprise Edition para GLVM usufrui dos seguintes componentes para reduzir o tempo de inatividade e de recuperação durante a recuperação de desastre:

Espelhamento e sincronização de dados do AIX e do PowerHA SystemMirror Enterprise Edition para GLVM. Ambos os grupos de volumes simultâneos padrão e aprimorado podem ser espelhados geograficamente com os utilitários do GLVM.

A rede de distância ilimitada baseada em TCP/IP suporta até quatro redes de espelhamento de dados XD_data que podem ser configuradas.

Se um componente falhar, o PowerHA SystemMirror assegurará que uma cópia espelhada dos dados esteja acessível em um site local ou remoto. Ambos os grupos de recursos simultâneos e não simultâneos podem ser configurados em um cluster do PowerHA SystemMirror com o GLVM. No entanto, a política entre sites não pode ser simultânea.

Informações relacionadas:

Guia de planejamento

Metro Mirror: Planejamento e administração

Planejamento e administração de LVM geográfico

Política de divisão

Um evento de divisão do cluster pode ocorrer entre sites quando um grupo de nós não pode se comunicar com os nós restantes em um cluster. Por exemplo, em um cluster vinculado, ocorrerá uma divisão se todos os links de comunicação entre os dois sites falharem. Um evento de divisão de cluster divide o cluster em duas ou mais partições.

É possível usar o PowerHA SystemMirror para configurar uma política de divisão que especifica a resposta para um evento de divisão de cluster.

As opções a seguir estão disponíveis para configurar uma política de divisão:

Nenhum

Uma opção de **Nenhum** indica que nenhuma ação ocorre quando um evento de divisão de cluster é detectado. Cada partição criada pelo evento de divisão do cluster se tornará um cluster independente. Cada partição poderá iniciar uma carga de trabalho independente da outra partição. Se os grupos de volumes compartilhados estiverem em uso, possivelmente isso levará à distorção de dados. Essa opção é a configuração padrão, pois é necessária configuração manual para estabelecer uma política alternativa. Não use essa opção se o ambiente estiver configurado para usar o HyperSwap for PowerHA SystemMirror.

Desempatador

Uma opção de **Desempatador** indica que um disco ou NFS é usado para determinar qual site particionado tem permissão para continuar a operar quando um evento de divisão de cluster

ocorre. Cada partição tenta adquirir o desempatador posicionando um bloqueio no disco desempatador ou arquivo NFS. O desempatador é um disco SCSI ou arquivo NFS que é acessível a todos os nós no cluster. A partição que não pode bloquear o disco ou arquivo NFS é reinicializada. Se você usar essa opção, a configuração da política de mesclagem também deverá usar a opção do desempatador.

Manual

Uma opção de **Manual** indica que você deseja corrigir manualmente o problema quando uma divisão de cluster ocorre. Cada nó na partição apresenta uma mensagem para escolher continuar executando serviços de cluster ou recuperar serviços de cluster (o que reinicializa o nó).

Esta política está disponível apenas para clusters vinculados. Deve-se especificar informações para ambos os sites vencedor ou perdedor quando a divisão de cluster ocorre. É possível especificar qual site é o site vencedor e perdedor usando a interface SMIT. Com esta política, é possível especificar o número de tentativas e a frequência de tentativas que requerem sua entrada. Também é possível especificar uma ação padrão para que ocorra após o número de tentativas que requerem sua entrada ser atingida e você não ter fornecido nenhuma entrada.

Conceitos relacionados:

“Política de mesclagem”

Dependendo da política de divisão do cluster, o cluster poderá ter duas partições que são executadas independentemente uma da outra. É possível usar o PowerHA SystemMirror para configurar uma política de mesclagem que permite que as partições operem juntas novamente após as comunicações serem restauradas entre as partições.

“Opção de desempatador para políticas de divisão e de mesclagem” na página 19

É possível usar a opção do desempatador para especificar um disco SCSI ou um arquivo Network File System (NFS) que é usado pelas políticas de divisão e mesclagem.

“Cluster vinculado do PowerHA SystemMirror” na página 21

Um cluster vinculado é ideal para situações em que cada site está localizado em uma localização geográfica diferente. Geralmente, os sites estão distantes o suficiente para que não possam compartilhar facilmente uma rede de área de armazenamento (SAN) comum. Cada site deve ter seu próprio disco de repositório ativo e quaisquer repositórios de backup. Os clusters vinculados sempre usam unicast para comunicar-se entre os sites. Os clusters vinculados são uma parte importante da recuperação de desastre de alta disponibilidade (HADR).

Informações relacionadas:

Configurando uma política de divisão e mesclagem

Política de mesclagem

Dependendo da política de divisão do cluster, o cluster poderá ter duas partições que são executadas independentemente uma da outra. É possível usar o PowerHA SystemMirror para configurar uma política de mesclagem que permite que as partições operem juntas novamente após as comunicações serem restauradas entre as partições.

As opções a seguir estão disponíveis para configurar uma política de mesclagem:

Maioria

A partição com o número mais alto de nós permanece online. Se cada partição tiver o mesmo número de nós, a partição que tiver o menor ID do nó será escolhida. A partição que não permanece online é reinicializada ou os serviços de cluster são reiniciados, de acordo com a especificação do plano de ação escolhido.

Desempatador

É possível usar um disco ou um arquivo NFS para um desempatador. Se você usar um disco para o desempatador, cada partição tentará adquirir o desempatador posicionando um bloqueio no disco desempatador. O desempatador é um disco SCSI acessível a todos os nós no cluster. A partição que não pode bloquear o disco é reinicializada ou os serviços de cluster são reiniciados, conforme especificado pelo plano de ação escolhido.

Se você usar um arquivo NFS para o desempatador, a montagem de NFS deverá existir em cada um dos nós no cluster a partir do servidor NFS selecionado. A partição que primeiro reserva o arquivo NFS continua a funcionar. A partição que não pode bloquear o arquivo NFS é reinicializada ou os serviços de cluster são reiniciados, conforme especificado pelo plano de ação escolhido. Se você usar essa opção, a configuração da política de divisão também deverá usar a opção do desempatador.

Manual

Esta opção de política requer que selecione o site vencedor durante um evento de mesclagem.

Esta política está disponível apenas para clusters vinculados. Deve-se especificar informações para ambos os sites vencedor ou perdedor quando a divisão de cluster ocorre. É possível especificar qual site é o site vencedor e perdedor usando a interface SMIT. Com esta política, é possível especificar o número de tentativas e a frequência de tentativas que requerem sua entrada. Também é possível especificar uma ação padrão para que ocorra após o número de tentativas que requerem sua entrada ser atingida e você não ter fornecido nenhuma entrada.

Prioridade

Esta política indica que o site de prioridade mais alta continua a operar quando um evento de mesclagem de cluster ocorre. É designada uma prioridade aos sites com base na ordem em que são listados na lista do site. O primeiro site na lista de sites é o site de prioridade mais alta. Essa política é disponível apenas para clusters vinculados.

Conceitos relacionados:

“Política de divisão” na página 17

Um evento de divisão do cluster pode ocorrer entre sites quando um grupo de nós não pode se comunicar com os nós restantes em um cluster. Por exemplo, em um cluster vinculado, ocorrerá uma divisão se todos os links de comunicação entre os dois sites falharem. Um evento de divisão de cluster divide o cluster em duas ou mais partições.

“Opção de desempatador para políticas de divisão e de mesclagem”

É possível usar a opção do desempatador para especificar um disco SCSI ou um arquivo Network File System (NFS) que é usado pelas políticas de divisão e mesclagem.

“Cluster vinculado do PowerHA SystemMirror” na página 21

Um cluster vinculado é ideal para situações em que cada site está localizado em uma localização geográfica diferente. Geralmente, os sites estão distantes o suficiente para que não possam compartilhar facilmente uma rede de área de armazenamento (SAN) comum. Cada site deve ter seu próprio disco de repositório ativo e quaisquer repositórios de backup. Os clusters vinculados sempre usam unicast para comunicar-se entre os sites. Os clusters vinculados são uma parte importante da recuperação de desastre de alta disponibilidade (HADR).

Informações relacionadas:

Configurando uma política de divisão e mesclagem

Opção de desempatador para políticas de divisão e de mesclagem

É possível usar a opção do desempatador para especificar um disco SCSI ou um arquivo Network File System (NFS) que é usado pelas políticas de divisão e mesclagem.

Um disco desempatador ou um arquivo NFS é usado quando os sites no cluster não podem mais se comunicar entre si. Essa falha de comunicação faz com que o cluster divida os sites em duas partições independentes. Se a falha ocorrer porque os links da comunicação do cluster não estão respondendo, ambas as partições tentarão bloquear o disco desempatador ou o arquivo NFS. A partição que adquire o disco desempatador continua a funcionar, enquanto a outra partição é reinicializada ou tem os serviços de cluster reiniciados, dependendo do plano de ação selecionado.

O disco ou arquivo montado por NFS que é identificado como o desempatador deve ser acessível a todos os nós no cluster.

Quando as partições que eram parte do cluster são colocadas on-line novamente após a falha na comunicação, elas devem estar aptas a se comunicar com a partição que possui o disco desempatador ou arquivo NFS. Se uma partição que é colocada on-line novamente não puder se comunicar com o disco desempatador ou o arquivo NFS, ela não será incluída no cluster. O disco desempatador ou arquivo NFS é liberado quando todos os nós na configuração se unem novamente no cluster.

| **Limitação na opção do desempatador com mais de dois nós em um cluster**

| Em um cluster com mais de dois nós, após uma divisão de cluster ocorrer, o disco desempatador é reservado pela partição vencedora. Na partição vencedora, um único nó faz uma reserva no disco desempatador. Se esse nó falhar sem liberar a reserva no disco desempatador, os nós restantes não poderão obter a reserva no disco desempatador e perderão. Portanto, todos os nós são reinicializados.

| Após a divisão de cluster ocorrer, deve-se resolver rapidamente o problema com o nó com falha para que a reserva do disco desempatador seja liberada.

Conceitos relacionados:

“Política de divisão” na página 17

Um evento de divisão do cluster pode ocorrer entre sites quando um grupo de nós não pode se comunicar com os nós restantes em um cluster. Por exemplo, em um cluster vinculado, ocorrerá uma divisão se todos os links de comunicação entre os dois sites falharem. Um evento de divisão de cluster divide o cluster em duas ou mais partições.

“Política de mesclagem” na página 18

Dependendo da política de divisão do cluster, o cluster poderá ter duas partições que são executadas independentemente uma da outra. É possível usar o PowerHA SystemMirror para configurar uma política de mesclagem que permite que as partições operem juntas novamente após as comunicações serem restauradas entre as partições.

Soluções de vários sites do PowerHA SystemMirror

O PowerHA SystemMirror suporta diferentes tipos de definições para sites e políticas específicas de site para recuperação de desastre de alta disponibilidade (HADR). É possível definir vários sites no PowerHA SystemMirror Standard Edition for AIX e no PowerHA SystemMirror Enterprise Edition for AIX.

É possível usar as interfaces de gerenciamento do PowerHA SystemMirror para criar as seguintes soluções de vários sites:

Cluster estendido

Contém nós de sites que estão localizados nas mesmas localizações geográficas. Os clusters estendidos não suportam HADR com gerenciamento de replicação de armazenamento.

Cluster vinculado

Contém nós de sites que estão localizados em localizações geográficas diferentes. Os clusters vinculados suportam o espelhamento LVM entre sites e o HyperSwap.

A tabela a seguir exhibe a diferença entre clusters estendidos e clusters vinculados.

Tabela 1. Diferenças entre os clusters estendidos e os clusters vinculados

Função	Clusters estendidos	Clusters vinculados
Comunicação de site	Multicast	Unicast
Disco de repositório	Compartilhado	Separado
Comunicação do cluster	<ul style="list-style-type: none"> • Rede • Rede de área de armazenamento (SAN) • Disco 	Rede
Espelhamento de Gerenciador de Volume Lógico entre sites	Disponível	Disponível com SAN

Tabela 1. Diferenças entre os clusters estendidos e os clusters vinculados (continuação)

Função	Clusters estendidos	Clusters vinculados
HyperSwap	Disponível	Disponível com SAN
Grupo de recursos simultâneos com oHyperSwap	Disponível	Disponível com SAN

Conceitos relacionados:

“Uso do Cluster Aware AIX pelo PowerHA SystemMirror” na página 4

O PowerHA SystemMirror complementa os recursos de armazenamento em cluster principais suportados no sistema operacional AIX. O PowerHA SystemMirror é suportado para todas as edições do AIX que suportam os recursos do Cluster Aware AIX (CAA).

Cluster vinculado do PowerHA SystemMirror

Um cluster vinculado é ideal para situações em que cada site está localizado em uma localização geográfica diferente. Geralmente, os sites estão distantes o suficiente para que não possam compartilhar facilmente uma rede de área de armazenamento (SAN) comum. Cada site deve ter seu próprio disco de repositório ativo e quaisquer repositórios de backup. Os clusters vinculados sempre usam unicast para comunicar-se entre os sites. Os clusters vinculados são uma parte importante da recuperação de desastre de alta disponibilidade (HADR).

O Cluster Aware AIX (CAA) pode usar clusters vinculados para suportar vários sites geograficamente distantes (em cidades diferentes). Os clusters vinculados vinculam os clusters individuais do CAA aos sites que estão localizados em diferentes locais. Os links entre os sites são usados para pulsação e comunicação do cluster.

Os clusters vinculados podem ser usados para comunicação em distâncias geográficas curtas, como entre prédios diferentes na mesma cidade.

A figura a seguir exibe um cluster vinculado, em que os sites estão localizados em cidades diferentes e cada site possui seu próprio disco de repositório.

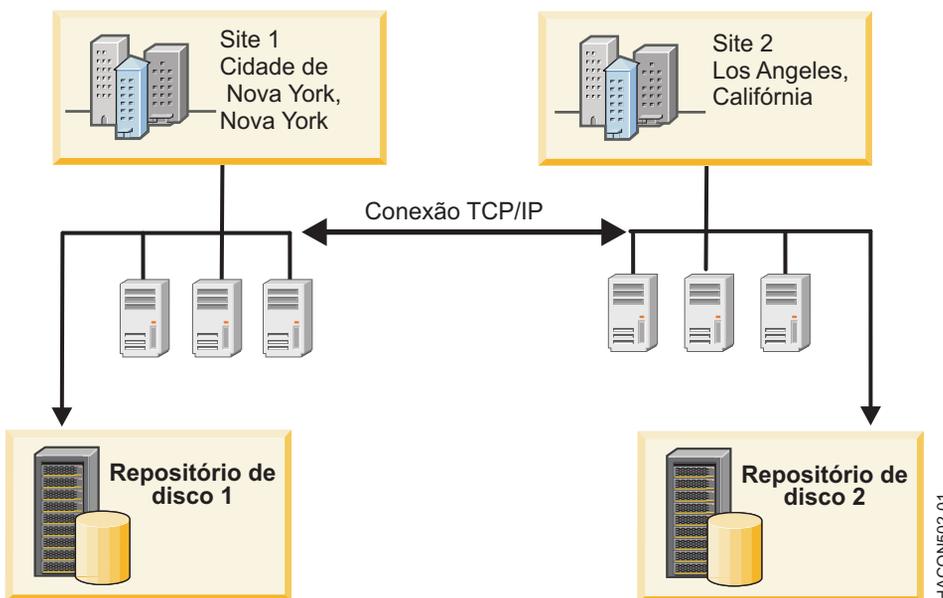


Figura 3. Configuração típica de cluster vinculado

Conceitos relacionados:

“Uso do Cluster Aware AIX pelo PowerHA SystemMirror” na página 4

O PowerHA SystemMirror complementa os recursos de armazenamento em cluster principais suportados no sistema operacional AIX. O PowerHA SystemMirror é suportado para todas as edições do AIX que suportam os recursos do Cluster Aware AIX (CAA).

“Cluster estendido do PowerHA SystemMirror”

Um cluster estendido é ideal para situações em que cada site está localizado na mesma localização geográfica. Geralmente, todos os nós estão em uma rede de área de armazenamento (SAN) comum. No mínimo, o repositório ativo e os repositórios de backup são compartilhados entre os sites, assim como os discos em grupos de recursos que podem ser executados em qualquer site. Os clusters estendidos podem suportar o espelhamento do LVM entre sites e o HyperSwap.

“Discos de repositório do cluster” na página 12

O PowerHA SystemMirror usa um disco compartilhado como um armazenador central para gerenciar a configuração do cluster. Esse disco deve ser acessível por todos os nós no cluster usando o método de implementação padrão ou de cluster estendido ou deve ser acessível por todos os nós em um site usando o método de implementação do cluster vinculado.

Cluster estendido do PowerHA SystemMirror

Um cluster estendido é ideal para situações em que cada site está localizado na mesma localização geográfica. Geralmente, todos os nós estão em uma rede de área de armazenamento (SAN) comum. No mínimo, o repositório ativo e os repositórios de backup são compartilhados entre os sites, assim como os discos em grupos de recursos que podem ser executados em qualquer site. Os clusters estendidos podem suportar o espelhamento do LVM entre sites e o HyperSwap.

Ao usar um cluster estendido no ambiente, um único cluster do Cluster Aware AIX (CAA) é implementado nos sites no cluster.

Se a comunicação de IP falhar e o ambiente estiver usando clusters estendidos, todos os nós no cluster poderão usar a função de pulsação de discos para comunicar-se e manter o ambiente operacional.

Os cluster estendidos devem atender aos seguintes requisitos:

- A comunicação de site pode usar endereços IP unicast ou multicast.
- Compartilhar pelo menos um disco de repositório entre os sites.

A figura a seguir é um exemplo de um cluster estendido que possui dois sites que estão em diferentes prédios, geograficamente próximos, na mesma cidade e que compartilham um disco de repositório.

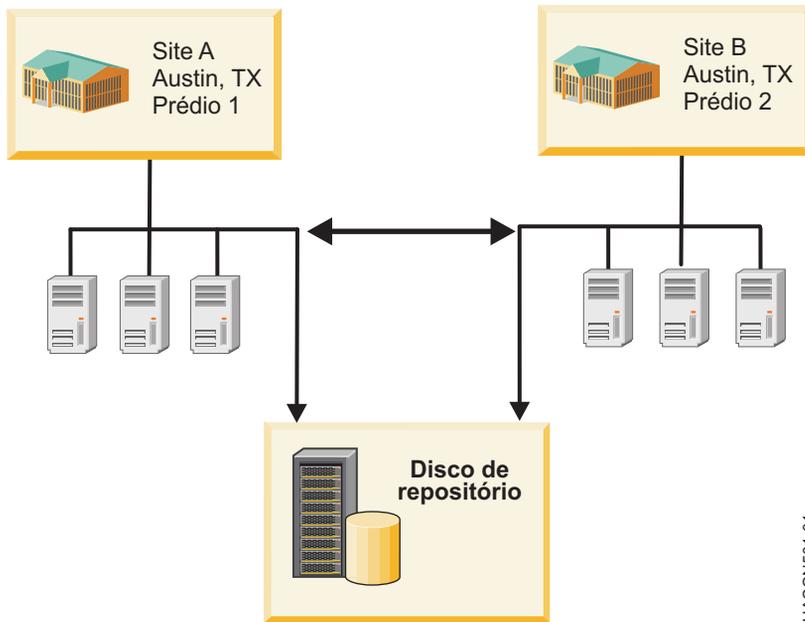


Figura 4. Configuração típica de cluster estendido

Conceitos relacionados:

“Cluster vinculado do PowerHA SystemMirror” na página 21

Um cluster vinculado é ideal para situações em que cada site está localizado em uma localização geográfica diferente. Geralmente, os sites estão distantes o suficiente para que não possam compartilhar facilmente uma rede de área de armazenamento (SAN) comum. Cada site deve ter seu próprio disco de repositório ativo e quaisquer repositórios de backup. Os clusters vinculados sempre usam unicast para comunicar-se entre os sites. Os clusters vinculados são uma parte importante da recuperação de desastre de alta disponibilidade (HADR).

“Uso do Cluster Aware AIX pelo PowerHA SystemMirror” na página 4

O PowerHA SystemMirror complementa os recursos de armazenamento em cluster principais suportados no sistema operacional AIX. O PowerHA SystemMirror é suportado para todas as edições do AIX que suportam os recursos do Cluster Aware AIX (CAA).

“Discos de repositório do cluster” na página 12

O PowerHA SystemMirror usa um disco compartilhado como um armazenador central para gerenciar a configuração do cluster. Esse disco deve ser acessível por todos os nós no cluster usando o método de implementação padrão ou de cluster estendido ou deve ser acessível por todos os nós em um site usando o método de implementação do cluster vinculado.

Redes de cluster

Os nós do cluster comunicam-se entre si sobre redes de comunicação.

Se uma das placas da interface de rede física em um nó em uma rede falhar, o PowerHA SystemMirror preservará a comunicação com o nó, transferindo o tráfego para outra placa da interface de rede física no mesmo nó. Se uma "conexão" com o nó falhar, o PowerHA SystemMirror transferirá os recursos para outro nó ao qual possui acesso.

Além disso, o RSCT envia pulsações entre os nós sobre as redes do cluster para verificar periodicamente o funcionamento dos próprios nós do cluster. Se o PowerHA SystemMirror não detectar nenhuma pulsação a partir de um nó, esse nó será considerado como com falha e os recursos serão transferidos automaticamente para outro nó.

É altamente recomendável configurar vários caminhos de comunicação entre os nós no cluster. A presença de várias redes de comunicação impede o particionamento do cluster, no qual os nós dentro de cada partição formam a sua própria entidade. Em um cluster particionado, é possível que os nós em cada partição permitam o acesso simultâneo, não sincronizado aos mesmos dados. Isso pode, potencialmente, levar a diferentes visualizações de dados de diferentes nós.

Conceitos relacionados:

“Pulsção sobre TCP/IP e redes de área de armazenamento” na página 28

Uma *pulsção* é um tipo de pacote de comunicação enviado entre nós. As pulsções são usadas para monitorar o funcionamento dos nós, das redes e interfaces de rede e para evitar o particionamento do cluster.

Redes físicas e lógicas

Uma rede física conecta duas ou mais interfaces de rede física.

Conforme mencionado na seção anterior, a configuração de várias redes baseadas em TCP/IP ajuda a evitar particionamento do cluster. O PowerHA SystemMirror comunica-se pela rede de armazenamento, quando necessário. Esse caminho de comunicação adicional ajuda a evitar que clusters particionados forneçam caminhos de comunicação adicionais nos casos em que as conexões de rede baseadas em TCP/IP se tornem congestionadas ou danificadas entre os nós de cluster.

Nota: Se você estiver considerando um cluster no qual as redes físicas usam dispositivos de rede externos para rotear pacotes de uma rede para outra, considere o seguinte: Ao configurar um cluster do PowerHA SystemMirror, o PowerHA SystemMirror verifica a conectividade e o acesso a todas as interfaces definidas em uma rede física específica. No entanto, o PowerHA SystemMirror não pode determinar a presença de dispositivos de rede externos, como pontes e roteadores, em um caminho de rede entre nós do cluster. Se as redes tiverem dispositivos de rede externos, assegure-se de estar usando dispositivos que sejam altamente disponíveis e redundantes para que eles não criem um ponto único de falha no cluster do PowerHA SystemMirror.

Uma *rede lógica* é uma parte de uma rede física que conecta duas ou mais interfaces ou dispositivos de rede lógicos. Uma interface ou dispositivo de rede lógico é a entidade de software conhecida para um sistema operacional. Há um mapeamento um-para-um entre uma interface/dispositivo de rede física e uma interface/dispositivo de rede lógica. Cada interface de rede lógica pode trocar pacotes com cada interface de rede lógica na mesma rede lógica.

Se em um subconjunto de interfaces de rede lógica na rede lógica, as interfaces precisarem se comunicar entre si (mas com nenhuma outra) enquanto compartilham a mesma rede física, serão usadas sub-redes. Uma máscara de sub-rede define a parte do endereço IP que determina se uma interface de rede lógica pode enviar pacotes a outra interface de rede lógica na mesma rede lógica.

Redes lógicas no PowerHA SystemMirror:

O PowerHA SystemMirror possui seu próprio conceito semelhante de uma rede lógica.

Todas as interfaces de rede lógicas em uma rede do PowerHA SystemMirror podem comunicar pacotes do PowerHA SystemMirror entre si, diretamente. Cada rede lógica é identificada por um nome exclusivo. Se você usar uma função de descoberta automática para a configuração de cluster do PowerHA SystemMirror, o PowerHA SystemMirror designará um nome a cada rede lógica do PowerHA SystemMirror que descobrir, como *net_ether_01*. Uma rede lógica do PowerHA SystemMirror pode conter uma ou mais sub-redes. O RSCT cuida do roteamento de pacotes entre sub-redes lógicas.

Conceitos relacionados:

“Hardware suportado pelo PowerHA SystemMirror” na página 45

É possível usar diferentes tipos de hardware IBM com o PowerHA SystemMirror para implementar o nível base de um ambiente altamente disponível.

Falhas de rede local e global:

Quando ocorre uma falha em uma rede de clusters, o PowerHA SystemMirror usa os *eventos de falha de rede* para gerenciar esses casos. O PowerHA SystemMirror verifica e distingue entre dois tipos de eventos de falha de rede: eventos de falha de rede local e de falha de rede global.

Falha de rede local

Uma *falha de rede local* é um evento do PowerHA SystemMirror no qual os pacotes não podem ser enviados ou recebidos por um nó por meio de uma rede lógica do PowerHA SystemMirror. Isso pode ocorrer, por exemplo, se todas as placas da interface de rede do nó que participam da rede lógica particular do PowerHA SystemMirror falharem. No caso de uma falha de rede local, a rede ainda estará em uso por outros nós. Para manipular falhas da rede local, o PowerHA SystemMirror move seletivamente os recursos (nessa rede) de um nó para outro. Esta operação é mencionada como failover seletivo.

Falha de rede global

Uma *falha de rede global* é um evento do PowerHA SystemMirror no qual os pacotes não podem ser enviados ou recebidos por nenhum nó por meio de uma rede lógica do PowerHA SystemMirror. Isso pode ocorrer, por exemplo, se a rede física é danificada. É importante distinguir entre uma rede global e um evento de falha de rede global. Um *rede global* é uma combinação de redes do PowerHA SystemMirror. Um *evento de falha de rede global* refere-se a uma falha que afeta todos os nós conectados a qualquer rede lógica do PowerHA SystemMirror, não necessariamente uma rede global.

Interfaces de comunicação do PowerHA SystemMirror

Uma interface de comunicação do PowerHA SystemMirror é um agrupamento de uma interface de rede lógica, um endereço IP do serviço e um rótulo de IP do serviço definidos no PowerHA SystemMirror.

As interfaces de comunicação do PowerHA SystemMirror combinam-se para criar redes baseadas em IP. Uma interface de comunicação do PowerHA SystemMirror é uma combinação de:

- *Interface de rede lógica* é o nome para o qual o AIX resolve uma porta (por exemplo, en0) de uma placa de interface de rede física.
- *Endereço IP de serviço* é um endereço IP (por exemplo, 129.9.201.1) sobre o qual serviços, como um aplicativo, são fornecidos e sobre o qual os nós de cliente se comunicam.
- *Rótulo de IP de serviço* é um rótulo (por exemplo, um nome do host no arquivo `/etc/hosts` ou um equivalente lógico de um endereço IP de serviço, como `node_A_en_service`) que é mapeado para o endereço IP de serviço.

As interfaces de comunicação no PowerHA SystemMirror são usadas das seguintes maneiras:

- Interface de comunicação refere-se às redes baseadas em IP e à placa da interface de rede (NIC). As NICs conectadas a uma rede física comum são combinadas em redes lógicas usadas pelo PowerHA SystemMirror.
- Cada NIC é capaz de hospedar vários endereços TCP/IP. Ao configurar um cluster, você deve definir os endereços IP que o PowerHA SystemMirror monitora (endereços IP base ou de inicialização), e os endereços IP que o PowerHA SystemMirror mantém altamente disponíveis (os endereços IP de serviço).
- A pulsação no PowerHA SystemMirror ocorre sobre as interfaces de comunicação. O PowerHA SystemMirror usa o recurso de pulsação do Cluster Aware AIX (CAA) para monitorar suas interfaces de rede e seus endereços IP. O CAA fornece a topologia de rede criada para o RSCT, enquanto o RSCT fornece notificações de falha ao PowerHA SystemMirror.

Dispositivos de comunicação não IP do PowerHA SystemMirror

O PowerHA SystemMirror também monitora dispositivos de rede que não podem efetuar comunicações IP. Esses dispositivos incluem a rede de área de armazenamento (SAN) e discos compartilhados.

Requisitos de roteamento de sub-rede no PowerHA SystemMirror

Uma rota sub-rede define um caminho, definido por uma sub-rede, para enviar pacotes por meio da rede lógica para um endereço em outra rede lógica.

O AIX permite incluir várias rotas para o mesmo destino na tabela de roteamento do kernel. Se várias rotas correspondentes tiverem critérios iguais, o roteamento poderá ser executado de forma alternativa usando uma das várias rotas de sub-rede. É importante considerar o roteamento de sub-rede no PowerHA SystemMirror por causa das seguintes considerações:

- O PowerHA SystemMirror não faz distinção entre as interfaces de rede lógica que compartilham a mesma rota de sub-rede. Se uma interface de rede lógica compartilhar uma rota com outra interface, o PowerHA SystemMirror não terá meios de determinar seu funcionamento. Para obter mais informações sobre rotas de rede, consulte a man page do AIX sobre o comando **route**.
- Várias restrições são impostas frequentemente às redes baseadas em IP por um administrador de rede ou pelos requisitos de TCP/IP. As sub-redes e rotas também são restrições dentro das quais o PowerHA SystemMirror deve ser configurado para operação.

Nota: Cada interface de comunicação deve estar em um nó que pertença a uma sub-rede exclusiva, para que o PowerHA SystemMirror possa monitorar cada interface. Esse não é um requisito obrigatório para todos os casos e depende de vários fatores. Por exemplo, quando há somente uma interface de rede por nó, como em um ambiente de rede virtual, os endereços IP de serviço e de inicialização podem estar na mesma sub-rede, porque há uma única rota de sub-rede. Caso sejam necessárias sub-redes exclusivas, a verificação de cluster do PowerHA SystemMirror sinalizará isso como um erro.

Rótulo e endereço IP de serviço

Um *rótulo de IP de serviço* é um rótulo que é mapeado para o endereço IP de serviço e é usado para estabelecer a comunicação entre os nós clientes e o nó de servidor.

Os serviços, como um aplicativo de banco de dados, são fornecidos usando a conexão sobre o rótulo de IP de serviço. Um rótulo de IP de serviço pode ser colocado em um grupo de recursos como um recurso, o que permite que o PowerHA SystemMirror monitore seu funcionamento e mantenha-o altamente disponível, em um nó ou, se o controle de endereço IP estiver configurado, entre os nós do cluster, transferindo-o para outro nó no caso de falha.

Alias de IP

Um *alias de IP* é um rótulo de IP ou endereço IP que é configurado em uma placa da interface de rede além do rótulo de IP ou endereço IP configurado originalmente na placa da interface de rede (NIC).

Os aliases de IP são uma função do AIX suportada pelo PowerHA SystemMirror. O Sistema operacional AIX suporta vários aliases de IP em uma NIC. Cada alias de IP em uma NIC pode ser configurado em uma sub-rede separada.

Os aliases de IP são usados no PowerHA SystemMirror como endereços de serviço para controle de endereço IP.

Os tópicos a seguir contêm informações sobre como o PowerHA SystemMirror liga um rótulo de IP de serviço a uma interface de comunicação, dependendo de qual mecanismo é usado para recuperar um rótulo de IP de serviço.

Controle de endereço IP

O *controle de endereço IP* é um mecanismo para recuperar um rótulo de IP de serviço movendo-o para outra placa da interface de rede (NIC) em outro nó, quando a NIC inicial falha.

O IPAT ocorrerá se a placa da interface de rede física em um nó falhar e se não houver nenhuma outra placa da interface de rede física acessível na mesma rede, no mesmo nó. Portanto, a troca de rótulos de IP dessas NICs no mesmo nó não poderá ser executada e o PowerHA SystemMirror usará o IPAT para

recuperar o endereço IP de serviço usando uma NIC em um nó de backup. O controle de endereço IP mantém o endereço IP altamente disponível por recuperar o endereço IP após falhas. O PowerHA SystemMirror usa um método chamado IPAT por meio de aliases de IP.

IPAT e rótulos de IP de serviço

O IPAT manipula o rótulo de IP de serviço.

Quando o IPAT por meio de aliases de IP é usado, o rótulo de IP ou o endereço IP de serviço recebe um alias ou é incluído como um endereço adicional na mesma interface de rede. Ou seja, vários endereços IP ou rótulos de IP são configurados na mesma interface de rede ao mesmo tempo. Nessa configuração, todos os endereços IP ou rótulos de IP definidos devem ser configurados em sub-redes diferentes, a menos que haja uma única interface de rede por nó. Este método pode economizar hardware, mas requer sub-redes adicionais.

Controle de endereço IP por meio de aliases de IP

É possível configurar o controle de endereço IP em determinados tipos de redes usando os recursos de rede de aliasing de IP do sistema operacional AIX.

A definição de aliases de IP para as interfaces de rede permite a criação de mais de um rótulo e endereço IP na mesma interface de rede. O IPAT por meio de aliases de IP usa os recursos do Protocolo de Resolução de Endereço (ARP) disponíveis em muitos tipos de redes.

Quando um grupo de recursos que contém o rótulo de IP de serviço efetua fallover a partir do nó primário para o nó de destino, os rótulos de IP de serviço são incluídos (e removidos) como endereços de alias sobre os endereços IP base em uma NIC disponível. Isso permite que uma única NIC suporte mais de um rótulo de IP de serviço colocado como um alias. Portanto, o mesmo nó pode hospedar mais de um grupo de recursos ao mesmo tempo.

Quando há várias interfaces no mesmo nó conectadas à mesma rede e essas interfaces não estão combinadas em uma Agregação de Ethernet, todos os endereços de inicialização devem estar em sub-redes diferentes. Além disso, todos os endereços persistentes ou endereços de serviço devem estar em sub-redes diferentes do que os endereços de inicialização.

Como o aliasing de IP cria a coexistência de vários rótulos de serviço na mesma interface de rede, é possível usar menos placas da interface de rede físicas no cluster. Em um fallover, o PowerHA SystemMirror distribui igualmente os aliases entre as placas da interface de rede disponíveis.

Preferência de distribuição para aliases de rótulo de IP de serviço

O PowerHA SystemMirror usa o IP address takeover (IPAT) por meio do método de aliases de IP para manter os rótulos de IP de serviço nos grupos de recursos altamente disponíveis.

Na inicialização do cluster, o PowerHA SystemMirror (por padrão) distribui todos os aliases de rótulo de IP de serviço entre todas as interfaces de inicialização disponíveis em uma rede usando o princípio da menos carregada. O PowerHA SystemMirror designa um novo endereço de serviço à interface que possui o menor número de aliases ou rótulos de IP persistentes já designados.

No entanto, em alguns casos, pode ser desejável especificar outros tipos de alocação ou assegurar que os rótulos continuem a ser alocados de uma maneira específica, não somente durante a inicialização, mas também durante os eventos de cluster subsequentes.

Por exemplo, é possível alocar todos os aliases de rótulo de IP de serviço para a mesma interface de inicialização que hospeda atualmente o rótulo de IP persistente para esse nó. Essa opção pode ser útil em configurações de firewall da VPN em que somente uma interface recebe conectividade externa e todos os rótulos de IP (aliases de rótulo de IP persistentes e de serviço) devem ser colocados na mesma interface para permitir a conectividade.

É possível configurar uma preferência de distribuição para os aliases dos rótulos de IP de serviço que são colocados sob o controle do PowerHA SystemMirror.

Uma preferência de distribuição de aliases de rótulo de IP de serviço é um atributo de toda a rede usado para controlar o posicionamento dos aliases de rótulo de IP de serviço nas placas da interface de rede física, nos nós no cluster. A configuração de uma preferência de distribuição de aliases de rótulo de IP de serviço executa o seguinte:

- Permite customizar o balanceamento de carga para rótulos de IP de serviço no cluster, considerando os rótulos de IP persistentes designados anteriormente nos nós
- Permite que o PowerHA SystemMirror redistribua os rótulos de IP de serviço de acordo com a preferência especificada.
- Permite configurar o tipo de preferência de distribuição adequado aos requisitos de conectividade externa do firewall da VPN.
- Embora os rótulos de IP de serviço possam ser movidos para outra interface de rede, o PowerHA SystemMirror assegura que os rótulos continuem a ser alocados de acordo com a preferência de distribuição especificada. Ou seja, a preferência de distribuição é mantida durante a inicialização e os eventos de cluster subsequentes, como um failover, um fallback ou uma mudança da interface no mesmo nó. Por exemplo, se você tiver especificado os rótulos para serem mapeados à mesma interface, os rótulos permanecerão mapeados na mesma interface, embora o rótulo de IP de serviço configurado inicialmente seja movido para outro nó.
- A preferência de distribuição será executada desde que as interfaces de rede aceitáveis estejam disponíveis no cluster. O PowerHA SystemMirror sempre mantém os rótulos de IP de serviço ativos, mesmo que a preferência não possa ser satisfeita.

Conceitos relacionados:

“Rótulos de IP de nó persistente” na página 99

Um *rótulo de IP de nó persistente* é uma ferramenta administrativa útil que permite entrar em contato com um nó mesmo se os serviços de cluster do PowerHA SystemMirror estiverem inativos nesse nó.

Informações relacionadas:

Administrando o PowerHA SystemMirror

Planejando o PowerHA SystemMirror

Pulsção sobre TCP/IP e redes de área de armazenamento

Uma *pulsção* é um tipo de pacote de comunicação enviado entre nós. As pulsções são usadas para monitorar o funcionamento dos nós, das redes e interfaces de rede e para evitar o particionamento do cluster.

Para que o cluster do PowerHA SystemMirror reconheça e responda a falhas, ele deverá verificar o funcionamento do cluster continuamente. Algumas dessas verificações são fornecidas pela função de pulsção.

Cada nó do cluster envia mensagens de pulsção em intervalos específicos para outros nós do cluster e espera receber mensagens de pulsção dos nós em intervalos específicos. Se as mensagens pararem de ser recebidas, o PowerHA SystemMirror reconhecerá que ocorreu uma falha. As pulsções podem ser enviadas sobre:

- Redes TCP/IP
- Rede de área de armazenamento
- Disco de repositório do cluster

O Cluster Aware AIX (CAA) usa a comunicação de pulsção em todas as redes TCP/IP e redes de área de armazenamento (SAN) disponíveis. Se as redes TCP/IP e as redes SAN falharem, o CAA tentará usar o disco de repositório como um mecanismo de pulsção alternativo. O caminho de pulsção para o disco

de repositório de backup é exibido como uma interface dpcom na saída do comando **lscluster**. Se as redes TCP/IP e as redes SAN estiverem funcionando, o comando **lscluster -i** exibirá a interface dpcom como restricted.

A função de pulsação é configurada para usar caminhos entre nós. Isso permite que as pulsações monitorem o funcionamento de todas as redes e interfaces de rede do PowerHA SystemMirror, além dos próprios nós do cluster.

Os caminhos de pulsação são configurados automaticamente pelo CAA. Existe a opção de configurar caminhos ponto a ponto e caminhos de disco como parte da configuração do PowerHA SystemMirror.

Conceitos relacionados:

“Redes de cluster” na página 23

Os nós do cluster comunicam-se entre si sobre redes de comunicação.

Pulsação sobre redes de Internet Protocol

O PowerHA SystemMirror baseia-se no Cluster Aware AIX (CAA) para fornecer pulsação entre os nós do cluster sobre redes de Internet Protocol. Por padrão, o CAA usa comunicações unicast para pulsação. Opcionalmente, será possível selecionar o uso de comunicações multicast se a rede estiver configurado para suportar multicast.

O PowerHA SystemMirror 7.1.2, ou posterior, suporta o Protocolo da Internet versão 6 (IPv6). No entanto, não é possível especificar explicitamente o endereço multicast IPv6. O CAA usa um endereço multicast IPv6 derivado do endereço multicast do Protocolo da Internet versão 4 (IPv4). Para determinar o endereço multicast IPv6, o prefixo padrão **0xFF05** é combinado usando o operador lógico OR com o equivalente hexadecimal do endereço IPv4. Por exemplo, o endereço multicast IPv4 é **228.8.16.129** ou **0xE4081081**. A transformação pela operação lógica OR com o prefixo padrão é **0xFF05:: | 0xE4081081**. Portanto, o endereço multicast IPv6 resultante é **0xFF05::E408:1081**.

Por padrão, o CAA fornece pulsação sobre todas as interfaces do AIX configuradas (e não somente sobre as interfaces configuradas na topologia do PowerHA SystemMirror).

Nota: Para parar a pulsação do CAA sobre uma rede específica, é possível defini-la como **privada**, em vez de **pública**. Esta nota aplica-se a ambas as interfaces, IPv4 e IPv6.

Conceitos relacionados:

“Requisitos de roteamento de sub-rede no PowerHA SystemMirror” na página 26

Uma rota sub-rede define um caminho, definido por uma sub-rede, para enviar pacotes por meio da rede lógica para um endereço em outra rede lógica.

Informações relacionadas:

Etapas para alterar uma rede baseada em IP

Pulsação sobre a rede de área de armazenamento

O Cluster Aware AIX (CAA) troca pulsações automaticamente sobre a rede de área de armazenamento (SAN), abrangendo os nós do cluster. Essas conexões podem fornecer um caminho de pulsação alternativo para um cluster que use uma única rede baseada em TCP/IP.

Multicasting do PowerHA SystemMirror

O PowerHA SystemMirror- usa uma camada de gerenciamento de funcionamento do cluster integrada como parte do sistema operacional AIX chamada Cluster Aware AIX (CAA). O CAA usa o código no nível do kernel do AIX para trocar pulsações sobre uma rede, usando adaptadores Fibre Channel e usando o sistema de mensagens baseado em disco por meio do armazenador central.

Por padrão, o CAA usa comunicações unicast para pulsação. Opcionalmente, será possível selecionar o uso de comunicações multicast se a rede estiver configurado para suportar multicast. Para a comunicação do cluster, é possível configurar manualmente um endereço multicast ou permitir que o CAA selecione automaticamente o endereço multicast.

A comunicação multicast não é usada quando um cluster é criado inicialmente, mas é necessária para as operações normais do cluster. Você deve testar e verificar se a comunicação multicast pode percorrer a rede entre todos os nós em um cluster.

Informações relacionadas:

- ☞ IP multicast com escopo definido de forma administrativa
- ☞ Intervalo e designações de endereço multicast
- ☞ Informações de roteador multicast

Comunicação de pacote de multicast

Multicasting é uma forma de endereçamento, na qual os nós formam um grupo e trocam mensagens.

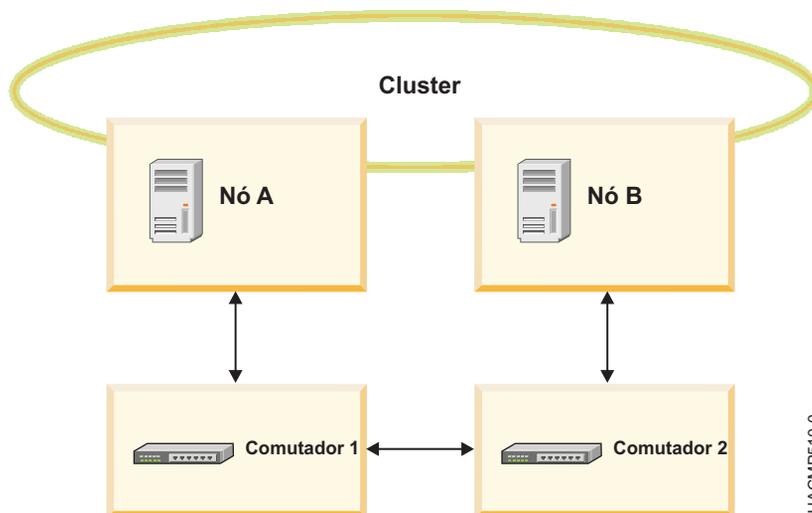
Uma mensagem multicast enviada por um nó no grupo é recebida por todos os outros nós no grupo. Essa função permite a comunicação eficiente do cluster. Por exemplo, um nó em um cluster pode enviar um único pacote multicast que pode notificar os outros nós sobre um evento crítico.

Comutações de rede multicast

Um comutador conecta diferentes nós e segmentos de rede e envia dados da rede para o nó correto. Um *comutador* é uma ponte de rede de várias portas que processa e roteia os dados na camada de link de dados (camada 2) do modelo Interconexão de sistemas abertos (OSI). Alguns comutadores também podem processar dados na camada de rede (camada 3).

Geralmente, em um ambiente de datacenter, vários nós são interconectados por comutadores. Quando os comutadores estão em cascata, um pacote de multicast é enviado a partir de um nó por meio de um comutador e, em seguida, a outro comutador até atingir o nó de destino final no cluster. Os comutadores manipulam pacotes de multicast de forma diferente do que a comunicação de rede regular. Portanto, a comunicação de comutador a comutador poderá não ocorrer para pacotes de multicast, caso a rede não esteja configurada corretamente.

A figura a seguir exibe uma configuração na qual o Nó A faz parte de um cluster que se conecta ao Nó B por meio do Comutador 1 e do Comutador 2. Para que os pacotes de multicast fluam entre o Nó A e o Nó B, a conexão de rede entre o Comutador 1 e Comutador 2 deverá ser ativada para comunicação de multicast.



Internet Group Management Protocol

O Internet Group Management Protocol (IGMP) é um protocolo de comunicações que permite que um nó (destinatário) informe um roteador multicast (consultor IGMP) sobre a intenção do nó de receber um determinado tráfego multicast.

O IGMP é executado entre um roteador e um nó que permite as seguintes ações:

- Os roteadores perguntam aos nós se eles precisam de um fluxo de multicast específico (consulta IGMP).
- Os nós respondem ao roteador se eles estão buscando um fluxo de multicast específico (relatórios IGMP).

O protocolo de comunicação IGMP é usado pelos nós e os roteadores adjacentes em redes IP para interagir e estabelecer regras de aterramento para comunicação multicast e estabelecer associação de grupo de multicast.

Snooping IGMP

O snooping IGMP é uma atividade executada por comutadores para controlar as comunicações IGMP relacionadas a trocas de pacote e adaptar à filtragem dos pacotes de multicast. Os comutadores que apresentam snooping IGMP derivam informações úteis, observando as transações IGMP entre os nós e roteadores. Essa função permite que os comutadores encaminhem corretamente os pacotes de multicast, quando necessário, ao próximo comutador no caminho de rede.

Os comutadores monitoram o tráfego IGMP e somente enviam pacotes de multicast quando necessário. Um comutador normalmente cria uma tabela de snooping IGMP que possui uma lista de todas as portas que solicitaram um grupo de multicast específico. A tabela de snooping IGMP é usada para permitir que os pacotes de multicast percorram a rede ou para desaprovar que eles percorram a rede. É possível configurar o comutador para evitar o snooping IGMP.

Roteamento multicast

O *roteamento multicast* é o processo pelo qual os dispositivos de rede decidem como os pacotes multicast são entregues a todos os destinatários-alvo do tráfego multicast.

Leia a documentação do roteador para determinar se ele implementa o roteamento multicast.

As entidades de rede que encaminham pacotes multicast usando algoritmos de roteamento especiais são chamadas *mrollers*. Os nós e outros elementos de rede implementam *mrollers* para permitir que o tráfego multicast percorra a rede.

Nota: Alguns roteadores também suportam roteadores de pacote de multicast.

Quando os comutadores estiverem em cascata, poderá ser necessário configurar comutações de rede para encaminhar os pacotes, conforme o necessário, para implementar *mrollers*. Se estiver ocorrendo problemas com as comunicações multicast no cluster, verifique se os dispositivos de rede possuem as configurações corretas para IGMP e snooping. Para obter mais informações sobre a configuração da rede para o tráfego multicast, consulte a documentação dos dispositivos de rede.

Recursos e grupos de recursos do PowerHA SystemMirror

Consulte aqui os conceitos e as definições relacionados a recursos que são usados em toda a documentação e também na interface com o usuário do PowerHA SystemMirror.

Identificando e mantendo recursos do cluster disponíveis

O software PowerHA SystemMirror fornece um ambiente altamente disponível.

O software PowerHA SystemMirror faz isso da seguinte forma:

- Identificando o conjunto de recursos do cluster que são essenciais para a operação de um aplicativo e combinando esses recursos em um grupo de recursos.
- Definindo as políticas e os atributos do grupo de recursos que ditam como PowerHA SystemMirror gerencia os recursos para mantê-los altamente disponíveis em diferentes estágios de operação do cluster (inicialização, failover e fallback).

Identificando recursos e definindo políticas de grupo de recursos, o software PowerHA SystemMirror torna possível inúmeras configurações de cluster, oferecendo enorme flexibilidade na definição de um ambiente em cluster customizado para requisitos individuais.

Identificando os recursos do cluster

Os seguintes recursos de cluster podem incluir os recursos de hardware e os recursos de software:

- Discos
- Grupos de volume
- Volumes lógicos
- Sistemas de arquivos
- Rótulos ou endereços IP de serviço
- Aplicativos
- Recursos de fita

Um processador que executa o PowerHA SystemMirror possui um conjunto de recursos definido pelo usuário: discos, grupos de volume, sistemas de arquivos, endereços IP e aplicativos. Para o propósito de manter os recursos altamente disponíveis, conjuntos de recursos interdependentes podem ser configurados nos grupos de recursos. Os *grupos de recursos* permitem combinar recursos relacionados em uma única entidade lógica para facilitar a configuração e o gerenciamento. O software PowerHA SystemMirror manipula o grupo de recursos como uma unidade, mantendo os recursos interdependentes juntos em um nó e mantendo-os altamente disponíveis.

Conceitos relacionados:

“Grupos de volumes” na página 33

Um grupo de volumes é um conjunto de volumes físicos que o AIX trata como uma região de disco contígua e endereçável.

“Volumes lógicos”

Um *volume lógico* é um conjunto de partições lógicas que o AIX disponibiliza como uma única unidade de armazenamento, ou seja, o volume lógico é a visualização lógica de um disco físico. As partições lógicas podem ser mapeadas para uma, duas ou três partições físicas para implementar o espelhamento.

“Sistemas de arquivos”

Um sistema de arquivos é gravado em um único volume lógico. Normalmente, você organiza um conjunto de arquivos como um sistema de arquivos para praticidade e velocidade no gerenciamento de dados.

“Aplicativos” na página 34

A finalidade de um sistema altamente disponível é assegurar que os serviços críticos estejam acessíveis aos usuários. Os aplicativos geralmente não precisam de nenhuma modificação para serem executados no ambiente do PowerHA SystemMirror. Qualquer aplicativo que possa ser reiniciado com sucesso após um encerramento inesperado é um candidato para o PowerHA SystemMirror.

“Rótulos de IP e endereços IP de serviço” na página 35

Um rótulo de IP de serviço é usado para estabelecer a comunicação entre os nós clientes e o nó de servidor. Os serviços, como um aplicativo de banco de dados, são fornecidos usando a conexão sobre o rótulo de IP de serviço.

“Recursos de fita” na página 35

É possível configurar uma unidade de fita SCSI ou Fibre Channel como um recurso de cluster em um grupo de recursos não simultâneos, tornando-a altamente disponível a dois nós em um cluster.

Tipos de recursos do cluster

Esta seção fornece uma visão geral breve dos recursos que podem ser configurados no PowerHA SystemMirror e incluídos em grupos de recursos para permitir que o PowerHA SystemMirror os mantenha altamente disponíveis.

Grupos de volumes

Um grupo de volumes é um conjunto de volumes físicos que o AIX trata como uma região de disco contígua e endereçável.

Os grupos de volumes são configurados para o sistema operacional AIX e podem ser incluídos em grupos de recursos no PowerHA SystemMirror. No ambiente do PowerHA SystemMirror, um grupo de volumes compartilhados é um grupo de volumes que reside inteiramente nos discos externos compartilhados pelos nós do cluster. Os discos compartilhados são conectados fisicamente aos nós do cluster e configurados logicamente em todos os nós do cluster.

Volumes lógicos

Um *volume lógico* é um conjunto de partições lógicas que o AIX disponibiliza como uma única unidade de armazenamento, ou seja, o volume lógico é a visualização lógica de um disco físico. As partições lógicas podem ser mapeadas para uma, duas ou três partições físicas para implementar o espelhamento.

No ambiente do PowerHA SystemMirror, os volumes lógicos podem ser usados para suportar um sistema de arquivos registrados (acesso não simultâneo) ou um dispositivo bruto (acesso simultâneo). O acesso simultâneo não suporta sistemas de arquivos. Os bancos de dados e aplicativos em ambientes de acesso simultâneo devem acessar os volumes lógicos brutos (por exemplo, `/dev/rsharedlv`).

Um volume lógico compartilhado deve ter um nome exclusivo em um cluster do PowerHA SystemMirror.

Nota: Um grupo de volumes compartilhados não pode conter um espaço de paginação ativo.

Sistemas de arquivos

Um sistema de arquivos é gravado em um único volume lógico. Normalmente, você organiza um conjunto de arquivos como um sistema de arquivos para praticidade e velocidade no gerenciamento de dados.

Sistemas de arquivos compartilhados

No sistema PowerHA SystemMirror, um sistema de arquivos compartilhados é um sistema de arquivos registrados que reside inteiramente em um volume lógico compartilhado. Para acesso não simultâneo, é possível planejar sistemas de arquivos compartilhados para que eles sejam colocados em discos externos compartilhados por nós do cluster. Os dados residem em sistemas de arquivos nesses discos compartilhados externos para que fiquem altamente disponíveis. Para acesso simultâneo, não é possível usar sistemas de arquivos registrados. Use volumes lógicos brutos.

Sistema de arquivos registrados e sistema de arquivos registrados aprimorado

Um sistema de arquivos registrados aprimorado (JFS2) fornece a capacidade de armazenar arquivos muito maiores do que o sistema de arquivos registrados (JFS). O JFS2 é o sistema de arquivos padrão do kernel de 64 bits. É possível escolher implementar o JFS, que é o sistema de arquivos recomendado para ambientes de 32 bits, ou o JFS2, que oferece a funcionalidade de 64 bits. O JFS2 é mais flexível do que o JFS porque permite aumentar e diminuir dinamicamente o número de arquivos que pode haver em um sistema de arquivos. O JFS2 também permite incluir o log de sistema de arquivos no mesmo volume lógico que os dados, em vez de alocar um volume lógico separado para os logs de todos os sistemas de arquivos no grupo de volumes.

Aplicativos

A finalidade de um sistema altamente disponível é assegurar que os serviços críticos estejam acessíveis aos usuários. Os aplicativos geralmente não precisam de nenhuma modificação para serem executados no ambiente do PowerHA SystemMirror. Qualquer aplicativo que possa ser reiniciado com sucesso após um encerramento inesperado é um candidato para o PowerHA SystemMirror.

Por exemplo, todos os produtos comerciais DBMS fornecem um ponto de verificação quanto ao estado do disco em uma espécie de diário da transação. No caso de uma falha do servidor, o servidor de failover reinicia o DBMS, que restabelece a consistência do banco de dados e, em seguida, continua o processamento.

É possível usar o AIX Fast Connect para compartilhar recursos com sistemas que não estejam executando o AIX como o sistema operacional. Se você configurar o Fast Connect como um recurso SystemMirror, o PowerHA SystemMirror o manterá altamente disponível e o recuperará de falhas da interface de nó e de rede. O PowerHA SystemMirror também verifica a configuração do Fast Connect durante a verificação do cluster. Os aplicativos são gerenciados ao definir o aplicativo para o PowerHA SystemMirror como um recurso do controlador de aplicativo. O controlador de aplicativo inclui scripts de inicialização e de interrupção de aplicativos. O PowerHA SystemMirror usa esses scripts quando o aplicativo precisa ser colocado online ou offline em um nó específico, para manter o aplicativo altamente disponível.

Nota: Os scripts de inicialização e de interrupção são os pontos principais de controle do PowerHA SystemMirror sobre um aplicativo. É muito importante que os scripts que você especificar operem corretamente para iniciar e parar todos os aspectos do aplicativo. Se os scripts falharem ao controlar o aplicativo corretamente, outras partes da recuperação do aplicativo poderão ser afetadas. Por exemplo, se o script usado para parar falhar ao parar o aplicativo completamente e um processo continuar a acessar um disco, o PowerHA SystemMirror não poderá colocar o grupo de volume offline no nó que falhou ou recuperá-lo no nó de backup.

Somente inclua o controlador de aplicativo em um grupo de recursos do PowerHA SystemMirror após testar completamente os scripts de inicialização e de interrupção de aplicativos.

O grupo de recursos que contém o controlador de aplicativo também deve conter todos os recursos dos quais o aplicativo depende, incluindo endereços IP de serviço, grupos de volumes e sistemas de arquivos. Quando esse grupo de recursos for criado, o PowerHA SystemMirror gerenciará o grupo de recursos inteiro e, portanto, todos os recursos interdependentes dentro dele como uma única entidade. O

PowerHA SystemMirror coordena a recuperação de aplicativo e gerencia os recursos na ordem que assegura a ativação de todos os recursos interdependentes antes de outros recursos.

Além disso, o PowerHA SystemMirror inclui o recurso de monitoramento de aplicativo, no qual é possível definir um monitor para detectar o encerramento inesperado de um processo ou para pesquisar periodicamente o encerramento de um aplicativo e tomar uma ação automática após a detecção de um problema.

É possível configurar vários monitores de aplicativo e associá-los a um ou mais controladores de aplicativo. Por suportar vários monitores por aplicativo, o PowerHA SystemMirror pode suportar configurações mais complexas. Por exemplo, é possível configurar um monitor para cada instância de um servidor Oracle paralelo em uso. Ou, é possível configurar um monitor customizado para verificar o funcionamento de um banco de dados e um monitor de encerramento de processo para detectar instantaneamente o encerramento do processo do monitor.

Também é possível especificar um modo para um monitor de aplicativo. Ele pode controlar como o aplicativo está sendo executado (modo de execução) ou se o aplicativo foi iniciado com sucesso (modo de inicialização de aplicativo). O uso de um monitor para assistir a inicialização do aplicativo é útil principalmente para configurações de cluster complexas.

Rótulos de IP e endereços IP de serviço

Um rótulo de IP de serviço é usado para estabelecer a comunicação entre os nós clientes e o nó de servidor. Os serviços, como um aplicativo de banco de dados, são fornecidos usando a conexão sobre o rótulo de IP de serviço.

Um rótulo de IP de serviço pode ser colocado em um grupo de recursos como um recurso, o que permite que o PowerHA SystemMirror monitore seu funcionamento e mantenha-o altamente disponível, em um nó ou, se o controle de endereço IP estiver configurado, entre os nós do cluster, transferindo-o para outro nó no caso de falha.

Nota: Determinados requisitos de sub-rede aplicam-se à configuração de rótulos de IP de serviço como recursos, em diferentes tipos de grupos de recursos.

Conceitos relacionados:

“Rótulo e endereço IP de serviço” na página 26

Um *rótulo de IP de serviço* é um rótulo que é mapeado para o endereço IP de serviço e é usado para estabelecer a comunicação entre os nós clientes e o nó de servidor.

Informações relacionadas:

Planejando o PowerHA SystemMirror

Recursos de fita

É possível configurar uma unidade de fita SCSI ou Fibre Channel como um recurso de cluster em um grupo de recursos não simultâneos, tornando-a altamente disponível a dois nós em um cluster.

O gerenciamento de unidades de fita compartilhadas é simplificado pela seguinte funcionalidade do PowerHA SystemMirror:

- Configuração de unidades de fita usando a ferramenta de configuração do SMIT
- Verificação de configuração adequada de unidades de fita
- Gerenciamento automático de unidades de fita durante as operações de inicialização e encerramento do grupo de recursos
- Realocação de unidade de fita em falha de nó e recuperação de nó
- Realocação controlada de unidades de fita em encerramento de cluster
- Realocação controlada de unidades de fita durante uma reconfiguração dinâmica de recursos de cluster.

Grupos de recursos de cluster

Para se tornar altamente disponível pelo software PowerHA SystemMirror, cada recurso deve ser incluído em um grupo de recursos. Os grupos de recursos permitem combinar recursos relacionados em uma única entidade lógica para facilitar o gerenciamento.

Lista de nós participantes

Uma lista de nós participantes define uma lista de nós que podem hospedar um grupo de recursos específico.

Você define uma lista de nós ao configurar um grupo de recursos. A lista de nós participantes pode conter alguns ou todos os nós no cluster.

Geralmente, essa lista contém todos os nós que compartilham os mesmos dados e discos.

Prioridade do nó padrão

A prioridade do nó padrão é identificada pela posição de um nó na lista de nós para um determinado grupo de recursos.

O primeiro nó na lista de nós tem a prioridade de nó mais alta. Esse nó também é chamado de nó inicial para um grupo de recurso. O nó que está listado antes de outro nó possui uma prioridade de nó superior com relação ao nó atual.

Dependendo da política de fallback para um grupo de recursos, quando um nó com uma prioridade superior de um grupo de recursos (que esteja atualmente sendo controlado por um nó de prioridade inferior) é inserido ou reintegrado no cluster, ele assume o controle do grupo de recursos. Ou seja, o grupo de recursos é movido dos nós com prioridades mais baixas para o nó com prioridade mais alta.

Em um determinado momento, o grupo de recursos pode ter uma prioridade do nó padrão especificada pela lista de nós participantes. No entanto, várias políticas do grupo de recursos selecionadas podem substituir a prioridade do nó padrão e "criar" a prioridade do nó real, de acordo com a qual um determinado grupo de recursos seria movido no cluster.

Prioridade dinâmica do nó

Ao configurar uma política dinâmica de prioridade de nó, você usa uma variável de recurso RSCT, como a menor carga de CPU, para selecionar o nó de controle para um grupo de recursos não simultâneos. Com a política de prioridade dinâmica ativada, a ordem da lista de nó de controle é determinada pelo estado do cluster no momento do evento, medido pela variável de recurso RSCT selecionada. É possível configurar diferentes políticas para diferentes grupos ou a mesma política para vários grupos.

Nó inicial

O nó inicial (o nó com prioridade mais alta para este grupo de recursos) é o primeiro nó listado na lista de nós participantes de um grupo de recursos não simultâneos.

O nó inicial geralmente é aquele que possui o grupo de recursos. Um grupo de recursos não simultâneos pode ter um nó inicial ou não. Se um grupo de recursos não simultâneos possui um nó inicial depende dos comportamentos de inicialização, failover e fallback do grupo de recursos.

Devido às diferentes mudanças no cluster, o grupo nem sempre pode ser iniciado no nó inicial. É importante diferenciar o nó inicial de um grupo de recursos do nó que atualmente o possui.

O termo nó inicial não é usado para grupos de recursos simultâneos porque eles pertencem a vários nós.

Inicialização, failover e fallback

O PowerHA SystemMirror assegura a disponibilidade de recursos do cluster, movendo grupos de recursos de um nó para outro quando as condições no cluster são alteradas.

O PowerHA SystemMirror gerencia grupos de recursos, ativando-os em um determinado nó ou em vários nós na inicialização do cluster ou movendo-os para outro nó, quando as condições no cluster são alteradas. Estes são os estágios em um ciclo de vida do cluster que afetam a forma como o PowerHA SystemMirror gerencia um grupo de recursos específico:

Inicialização do cluster

Os nós estão ativos e os grupos de recursos são distribuídos entre eles de acordo com a política de inicialização do grupo de recursos selecionada.

Falha do nó

Os grupos de recursos que estão ativos neste nó efetuam failover para outro nó.

Recuperação de nó

Um nó é incluído novamente no cluster e os grupos de recursos podem ser readquiridos, dependendo das políticas de grupo de recursos selecionadas.

Falha e recuperação de recurso

Um grupo de recursos pode efetuar failover para outro nó e ser readquirido quando o recurso ficar disponível.

Encerramento do cluster

Há diferentes maneiras de encerrar um cluster e uma delas assegura que os grupos de recursos sejam movidos para outro nó.

Durante cada um desses estágios do cluster, o comportamento dos grupos de recursos no PowerHA SystemMirror é definido pelo seguinte:

- Qual nó, ou nós, ativa o grupo de recursos na inicialização do cluster
- Quantos grupos de recursos podem ser adquiridos em um nó durante a inicialização do cluster
- Qual nó controlará o grupo de recursos quando o nó que possuía o grupo de recursos falhar e o PowerHA SystemMirror precisar mover o grupo de recursos para outro nó
- Se um grupo de recursos efetuará fallback para um nó que acabou de ser incluído no cluster ou continuará no nó que o possui atualmente.

As políticas de grupo de recursos selecionadas determinam qual nó do cluster controla originalmente um grupo de recursos e quais nós do cluster controlarão o grupo de recursos quando o nó original abandonar o controle.

Cada combinação dessas políticas permite especificar graus variados de controle sobre qual nó, ou nós, controla um grupo de recursos.

Para resumir, o foco do PowerHA SystemMirror na propriedade de grupo de recursos possibilita inúmeras configurações de cluster e fornece enorme flexibilidade na definição do ambiente em cluster que atenda às necessidades específicas do aplicativo. A combinação de políticas de inicialização, failover e fallback resume todas as políticas de gerenciamento disponíveis nas liberações anteriores sem o requisito de especificar o conjunto de opções que modificavam o comportamento dos tipos de grupo "predefinidos".

Ao definir comportamentos de grupo de recursos, tenha em mente que um grupo de recursos pode ser controlado por um ou mais nós no cluster.

Inicialização, failover e fallback são comportamentos específicos que descrevem como os grupos de recursos se comportam em diferentes estágios do cluster. É importante ter em mente as diferenças entre failover e fallback. Esses termos aparecem frequentemente em discussões sobre as variadas políticas de grupo de recursos.

Inicialização

Inicialização é a ativação de um grupo de recursos em um nó (ou em vários nós) no qual ele reside atualmente ou no nó inicial desse grupo de recursos. A inicialização do grupo de recursos ocorre durante a inicialização do cluster ou a aquisição inicial do grupo em um nó.

Fallover

Fallover é o movimento de um grupo de recursos a partir do nó que atualmente o possui para outro nó ativo, quando o nó atual sofre uma falha. O novo proprietário não é um nó de junção ou adicionado novamente.

O fallover somente é válido para grupos de recursos não simultâneos.

Fallback

Fallback é o movimento de um grupo de recursos a partir do nó no qual reside atualmente (que não é um nó inicial desse grupo de recursos) para um nó que está sendo inserido ou adicionado novamente no cluster.

Por exemplo, quando um nó com uma prioridade mais alta para esse grupo de recursos é inserido ou adicionado novamente no cluster, ele assume o controle do grupo de recursos. Ou seja, o grupo de recursos efetua fallback a partir dos nós com prioridade mais baixa para o nó com prioridade mais alta.

A definição de um comportamento de fallback somente é válida para grupos de recursos não simultâneos.

Políticas e atributos de grupos de recursos

É possível configurar os grupos de recursos para usar políticas específicas de inicialização, fallover e fallback. Esta seção descreve os atributos e cenários de grupo de recursos e ajuda a decidir quais grupos de recursos se adequam aos seus requisitos de cluster.

As políticas de grupo de recursos podem ser usadas das seguintes maneiras:

- Configure grupos de recursos para assegurar que eles sejam colocados novamente online ao serem novamente incluídos nos nós durante as horas de menor atividade.
- Especifique que um grupo de recursos que contenha um determinado aplicativo seja o único a receber preferência e a ser adquirido durante a inicialização em um nó específico. Isso é feito especificando a política de distribuição de nó. Isso será relevante se vários grupos de recursos não simultâneos puderem, potencialmente, ser adquiridos em um nó, mas um grupo de recursos específico tiver um aplicativo mais importante a ser mantido disponível.
- Especifique que grupos de recursos específicos devem ser mantidos juntos online no mesmo nó, ou mantidos online separados em nós diferentes, na inicialização, no fallover e no fallback.

Inicialização, fallover e fallback de grupo de recursos

Existem várias políticas para grupos de recursos individuais.

Essas políticas são:

Tabela 2. Políticas de grupo de recursos

Grupo de recursos	Política
Inicialização	<ul style="list-style-type: none"> • Somente online no nó inicial. O grupo de recursos fica online somente em seu nó inicial (prioridade mais alta) durante a inicialização do grupo de recursos. Isso requer que a prioridade mais alta do nó esteja disponível (primeiro nó na lista de nós do grupo de recursos). • Online no primeiro nó disponível. O grupo de recursos fica online no primeiro nó participante que fica disponível. • Online em todos os nós disponíveis. O grupo de recursos fica online em todos os nós. • Online usando a política de distribuição. Somente um grupo de recursos fica online em cada nó.
Fallover	<ul style="list-style-type: none"> • Fallover para o nó com a próxima prioridade na lista. O grupo de recursos segue a ordem de prioridade de nó padrão especificada na lista de nós do grupo de recursos. • Fallover usando a prioridade de nó dinâmica. Antes de selecionar essa opção, selecione uma das três políticas de prioridade de nó dinâmica predefinidas. Elas são baseadas em variáveis RSCT, como o nó com mais memória disponível. O fallover para essa opção não é suportado quando os sites são configurados. • Ficar offline (somente em erro do nó). Selecione esta opção para colocar um grupo de recursos offline em um nó durante uma condição de erro.
Fallback	<ul style="list-style-type: none"> • Fallback para o nó de prioridade mais alta na lista. Um grupo de recursos efetua fallback quando um nó de prioridade mais alta é incluído no cluster. Se você selecionar esta opção, o cronômetro de fallback atrasado poderá ser usado. Se você não configurar uma política de fallback atrasado, o grupo de recursos efetuará fallback imediatamente quando um nó de prioridade mais alta for incluído no cluster. • Nunca efetuar fallback. Um grupo de recursos não efetuará fallback para um nó de prioridade mais alta quando ele for incluído no cluster.

Informações relacionadas:

Administrando o PowerHA SystemMirror

Planejando o PowerHA SystemMirror

Tempo de acomodação, prioridade dinâmica do nó e cronômetro de fallback

É possível configurar alguns parâmetros adicionais para os grupos de recursos que indicam como o grupo de recursos se comporta na inicialização, no fallover e no fallback.

A seguir, os parâmetros adicionais para grupos de recursos que você pode configurar:

Tempo de acomodação

É possível configurar um comportamento de inicialização de um grupo de recursos, especificando o tempo de acomodação para um grupo de recursos que está offline atualmente. Quando o tempo de acomodação não está configurado, o grupo de recursos é iniciado no primeiro nó de alta prioridade disponível que é incluído no cluster. Quando o tempo de acomodação está configurado, o PowerHA SystemMirror espera durante o período de tempo de acomodação para que um nó com prioridade mais alta seja incluído no cluster, antes de ativar um grupo de recursos. A especificação do tempo de acomodação permite que um grupo de recursos seja adquirido em um nó que tenha uma prioridade mais alta, quando vários nós forem incluídos no cluster simultaneamente. O tempo de acomodação é um atributo que abrange todo o cluster e que, se estiver configurado, afetará o comportamento de inicialização de todos os grupos de recursos no cluster para os quais o comportamento de inicialização Online no primeiro nó disponível estiver selecionado.

Política de distribuição

É possível configurar o comportamento de inicialização de um grupo de recursos para usar a política de distribuição baseada em nó. Essa política assegura que durante a inicialização, um nó adquira somente um grupo de recursos. Consulte a seção a seguir para obter informações adicionais.

Prioridade de nó dinâmico

É possível configurar um comportamento de fallover de um grupo de recursos para usar uma entre três políticas de prioridade de nó dinâmica. Essas políticas são baseadas em variáveis RSCT,

como a maior memória livre ou menor uso de CPU. Para recuperar o grupo de recursos, o PowerHA SystemMirror seleciona o nó que melhor se adequa à política no momento do failover.

Nota: O failover que usa a opção Prioridade de nó dinâmico não é suportado quando os sites são configurados.

Cronômetro de fallback atrasado

É possível configurar que um comportamento de fallback de um grupo de recursos ocorra em uma das frequências recorrentes predefinidas: diariamente, semanalmente, mensalmente e anualmente ou em uma data e hora específicas, especificando e designando um cronômetro de fallback atrasado. Isso é útil, por exemplo, para planejar que os fallbacks de grupo de recursos ocorram durante as horas de negócios de menor atividade.

Política de distribuição

Na inicialização do cluster, é possível usar uma política de distribuição baseada em nó para grupos de recursos.

Se você selecionar essa política para vários grupos de recursos, o PowerHA SystemMirror tentará fazer com que cada nó adquira somente um desses grupos de recursos durante a inicialização. Isso permite distribuir os aplicativos com uso intenso de CPU em nós diferentes.

Informações relacionadas:

Administrando o PowerHA SystemMirror

Planejando o PowerHA SystemMirror

Redes e grupos de recursos

Todos os grupos de recursos suportam rótulos de IP de serviço configurados em redes com alias.

Um rótulo de IP de serviço pode ser incluído em qualquer grupo de recursos não simultâneos. Esse grupo de recursos pode ter qualquer uma das políticas de inicialização permitidas, exceto Online em todos os nós disponíveis.

Dependências do grupo de recursos

O PowerHA SystemMirror suporta a ordenação de grupo de recursos e o processamento serial customizado de recursos para acomodar as configurações do cluster nas quais existe uma dependência entre os aplicativos que residem em grupos de recursos diferentes.

Com o processamento serial customizado, é possível especificar que um determinado grupo de recursos seja processado antes de outro grupo de recursos. O PowerHA SystemMirror oferece uma maneira fácil de configurar dependências de pai e filho entre grupos de recursos (e os aplicativos que pertencem a eles) para assegurar o processamento adequado durante eventos de cluster.

As políticas de dependência de local estão disponíveis para configurar grupos de recursos para que eles sejam distribuídos da maneira esperada não somente ao iniciar o cluster, mas também durante o failover e o fallback. É possível configurar dependências para que os grupos especificados fiquem online em nós diferentes ou nos mesmos nós. O PowerHA SystemMirror processa os grupos de recursos dependentes na ordem adequada usando o processamento paralelo sempre que possível e o serial quando necessário. Não é necessário customizar o processamento. É possível configurar diferentes tipos de dependências entre os grupos de recursos:

- Dependências de pai e filho
- Dependências de local.

As dependências entre os grupos de recursos que podem ser configuradas são:

- Especificada explicitamente usando a interface SMIT
- Estabelecida em todo o cluster, não somente no nó local

- Garantida para ser respeitada no cluster.

Dependências de grupos de recursos filhos e pais

A configuração de uma dependência de grupo de recursos pai e filho permite facilitar a configuração e o controle de clusters com aplicativos com multicamadas, nos quais um aplicativo depende do sucesso da inicialização de outro aplicativo e ambos os aplicativos devem ser mantidos altamente disponíveis com o PowerHA SystemMirror.

O exemplo a seguir ilustra o comportamento da dependência de pai e filho:

- Se o grupo de recursos A depender do grupo de recursos B, na inicialização do nó, o grupo de recursos B deverá ser colocado online antes que o grupo de recursos A seja adquirido em qualquer nó no cluster. No failover, a ordem é inversa: O grupo de recursos A deve ser colocado offline antes que o grupo de recursos B seja colocado offline.
- Além disso, se o grupo de recursos A depender do grupo de recursos B, durante uma inicialização do nó ou uma reintegração do nó, o grupo de recursos A não poderá ser colocado online antes que o grupo de recursos B seja colocado online. Se o grupo de recursos B for colocado offline, o grupo de recursos A também será colocado offline, pois ele depende do grupo de recursos B.

As dependências entre os grupos de recursos oferecem uma maneira previsível e confiável de construir clusters com aplicativos com multicamadas. Para obter mais informações sobre ambientes em cluster típicos que podem usar grupos de recursos dependentes, consulte Configurações de cluster com aplicativos com multicamadas. Estes termos descrevem as dependências de pai e filho entre grupos de recursos:

- Um *grupo de recursos pai* precisa estar no estado online antes que o grupo de recursos que depende dele (filho) possa ser iniciado.
- Um *grupo de recursos filho* depende de um grupo de recursos pai. Ele somente será ativado em qualquer nó no cluster depois que o grupo de recursos pai for ativado. Geralmente, o grupo de recursos filho depende de alguns serviços de aplicativo que o grupo de recursos pai fornece.

Após a liberação do grupo de recursos (durante o failover ou a interrupção de serviços de cluster, por exemplo), o PowerHA SystemMirror coloca o grupo de recursos filho offline antes que o grupo de recursos pai seja colocado offline. O gráfico a seguir ilustra o relacionamento de dependência de pai e filho entre grupos de recursos.

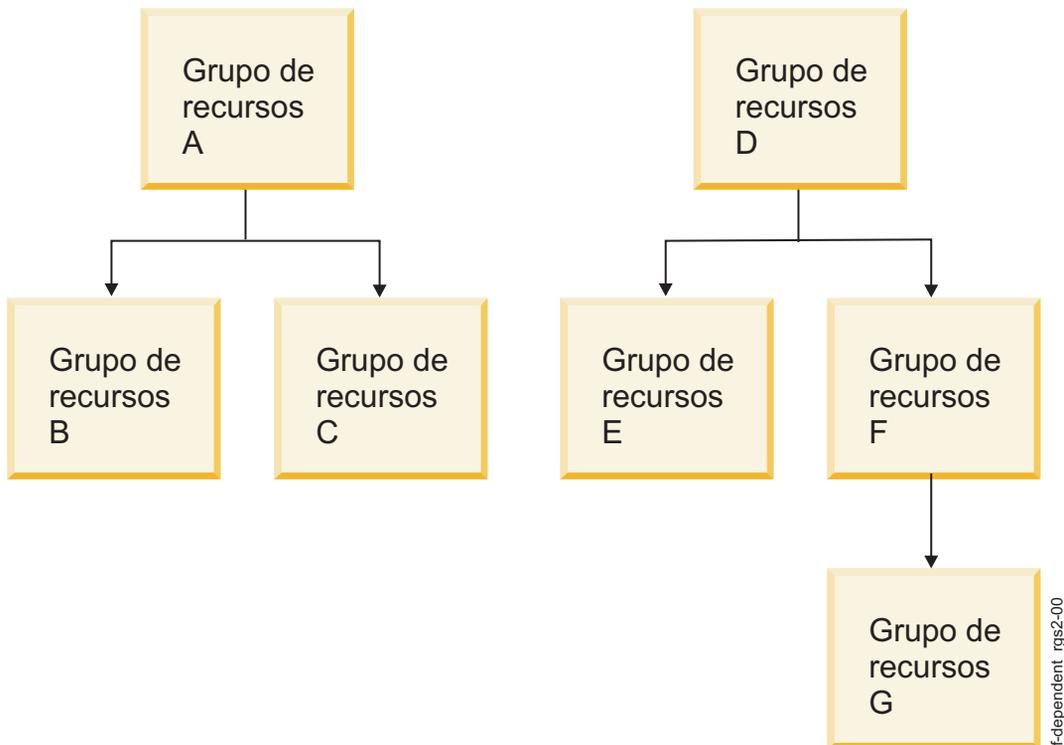


Figura 5. Exemplo de dois e três níveis de dependências entre os grupos de recursos

O exemplo mostra os relacionamentos que foram estruturados nestas recomendações e limitações:

- É possível configurar um tipo de dependência em que um grupo de recursos pai deve estar online em qualquer nó no cluster antes que um grupo de recursos filho (dependente) possa ser ativado em um nó.
- Um grupo de recursos pode servir como um grupo de recursos pai ou filho, dependendo de qual extremidade de um determinado link de dependência seja colocada.
- É possível especificar três níveis de dependências para grupos de recursos.
- Não é possível especificar dependências circulares entre grupos de recursos.

Essas recomendações e limitações também se aplicam a dependências de pai e filho entre grupos de recursos:

- É possível incluir, alterar ou excluir uma dependência entre grupos de recursos, enquanto os serviços de cluster estão em execução.
- Quando uma dependência entre dois grupos de recursos é excluída, somente o link entre esses grupos de recursos é removido do Banco de dados de configuração do PowerHA SystemMirror. Os grupos de recursos não são excluídos.
- Durante um failover de um grupo de recursos pai, um grupo de recursos filho contendo o aplicativo fica offline temporariamente e, em seguida, online em cada nó disponível. O aplicativo que pertença ao grupo de recursos filho também é parado e reiniciado.

Conceitos relacionados:

“Configurações de cluster com aplicativos com multicamadas” na página 83

Uma configuração de cluster típica que pode usar grupos de recursos com dependência de pai e filho é o ambiente no qual um aplicativo como o WebSphere depende de outro aplicativo como o DB2.

Dependências de local do grupo de recursos

Além de várias políticas para grupos de recursos individuais e dependências de pai e filho, o PowerHA SystemMirror oferece políticas para manipular as interdependências gerais dos grupos de recursos. O PowerHA SystemMirror reconhece esses relacionamentos e processa os grupos de recursos na ordem apropriada.

É possível configurar grupos de recursos para que:

- Dois ou mais grupos de recursos especificados estejam sempre online no mesmo nó. Sua inicialização, fallover e fallback serão para o mesmo nó.
- Dois ou mais grupos de recursos especificados sempre estejam online em nós diferentes. Sua inicialização, fallover e fallback serão para nós diferentes. É possível designar prioridades aos grupos de recursos para que os mais críticos sejam manipulados primeiro no caso de fallover e fallback.
- Dois ou mais grupos de recursos especificados (com recursos replicados) sempre estejam online no mesmo site.

Após a configuração de grupos de recursos individuais com uma dependência de local determinada, eles formarão um conjunto que será tratado como uma unidade pelo Gerenciador do cluster. As regras a seguir se aplicam quando um grupo de recursos é movido explicitamente com o comando **cIRGmove**:

- Se um grupo de recursos participar de um conjunto de *Dependência online no mesmo nó*, ele somente poderá ser colocado online no nó no qual todos os outros grupos de recursos do mesmo conjunto de nós estejam online no momento. Essa é a mesma regra para o Gerenciador do cluster.
- Se um grupo de recursos participar de um conjunto de *Dependência online no mesmo site*, ele somente poderá ser colocado online no site no qual os outros grupos de recursos do mesmo conjunto de site estejam online no momento. Essa é a mesma regra para o Gerenciador do cluster.
- Se um grupo de recursos participar de um conjunto de *Dependência online em nós diferentes*, ele somente poderá ser colocado online em um nó que não hospede nenhum outro grupo de recursos no conjunto de dependência de nó diferente. (Essa é a mesma regra para o Gerenciador de cluster). No entanto, quando um grupo de recursos que pertence a este conjunto é movido, as prioridades são tratadas como de igual valor, ao passo que quando o PowerHA SystemMirror coloca esses grupos online, ele considera as prioridades.

Modelo de dependência de local de amostra

A empresa fictícia XYZ Publishing segue um modelo de continuidade de negócios que envolve a priorização de diferentes plataformas usadas para desenvolver o conteúdo da web. Políticas de local são usadas para manter alguns grupos de recursos estritamente em nós separados e outros juntos no mesmo nó.

A figura a seguir exibe um exemplo de modelo de dependência de local que possui três nós que são usados para três aplicativos diferentes e seus bancos de dados associados. O banco de dados de cada aplicativo sempre deve ser executado no mesmo nó que o aplicativo. Por exemplo, o banco de dados de produção deve conter o aplicativo de produção.

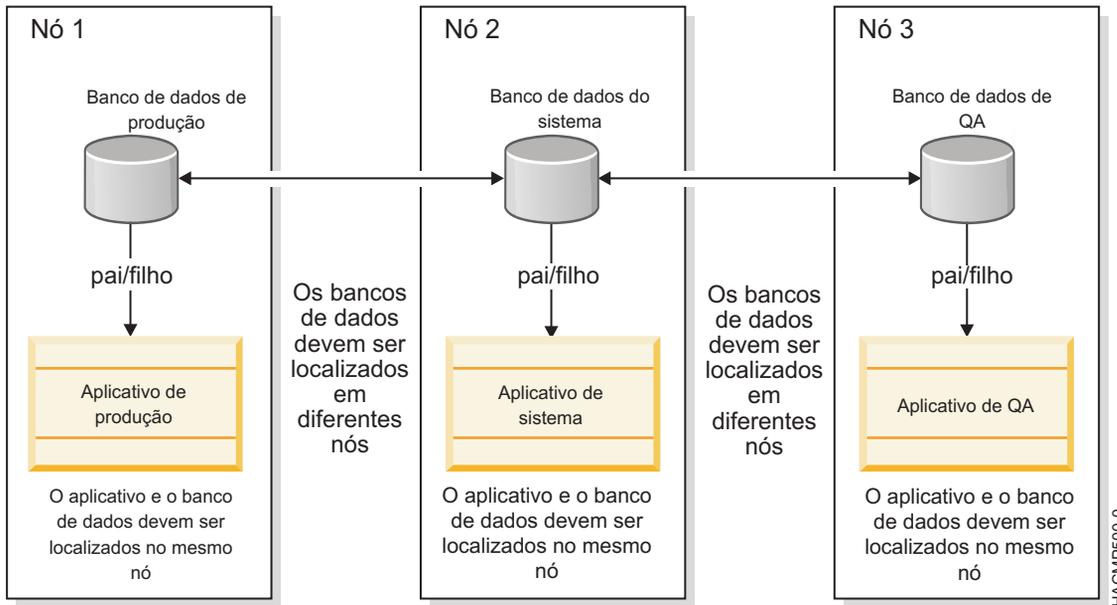


Figura 6. Exemplo de nós 1, 2 e 3 sendo usados para bancos de dados separados

Informações relacionadas:

Administrando o PowerHA SystemMirror

Planejando o PowerHA SystemMirror

Sites e grupos de recursos

A maioria das configurações do PowerHA SystemMirror não incluem sites e usam a política de gerenciamento IGNORE, que é o padrão. Se você tiver instalado um componente do PowerHA SystemMirror Enterprise Edition para recuperação de desastre, distribua os nós do cluster entre os sites separados geograficamente e selecione uma das políticas de gerenciamento entre sites.

Inclua os recursos que deseja replicar nos grupos de recursos. Defina as políticas de inicialização, failover e fallback para a instância primária de um grupo de recursos replicado. A instância primária é onde o grupo de recursos está online. O nó com a instância primária do grupo de recursos ativa todos os recursos do grupo. A instância secundária (a replicação) é ativada em um nó que está localizado no outro site como um backup. A política de gerenciamento entre sites combinada com as políticas de inicialização, failover e fallback do grupo de recursos determina o local onde a instância primária é localizada primeiro e como o failover e o fallback entre sites são manipulados.

As seguintes opções existem para configurar as política de gerenciamento intersites do grupo de recursos:

- Preferir site primário
- Online em qualquer site
- Online em ambos os sites

Se você definir sites para o cluster, ao definir as políticas de inicialização, failover e fallback para cada grupo de recursos que deseja replicar, será possível designar o grupo de recursos a um nó no site primário e a um nó no outro site (secundário). A instância primária do grupo de recursos é executada no site primário e a instância secundária no site secundário.

Se você tiver um grupo de recursos simultâneos, defina-o para ser executado em todos os nós. Neste caso, é possível selecionar a opção Online em ambos os sites para a política de gerenciamento entre sites. Em seguida, as instâncias em ambos os sites ficam ativas (não há instâncias secundárias). Também é

possível selecionar as outras políticas de gerenciamento entre sites para que um grupo de recursos simultâneos fique online em todos os nós em um site e possua instâncias de backup no outro site.

Nota: Quando houver sites definidos e um grupo de recursos não simultâneos, a interrupção dos serviços de cluster com a opção não gerenciada em um nó do site colocará o grupo de recursos no estado UNMANAGED em todos os nós desse site.

Também é possível mover a instância primária de um grupo de recursos para além dos limites do site com o utilitário **clRGmove**. Em seguida, o PowerHA SystemMirror redistribuirá as instâncias secundárias peers, conforme o necessário (ou fornecerá um aviso se a operação de movimentação for desaprovada devido a um requisito de configuração).

Ao mover a instância primária de um grupo de recursos enquanto sua instância secundária estiver offline, a instância secundária permanecerá offline. O PowerHA SystemMirror não coloca online a instância secundária do grupo de recursos automaticamente.

Informações relacionadas:

Guia de Administração

Guia de planejamento

Hardware suportado pelo PowerHA SystemMirror

É possível usar diferentes tipos de hardware IBM com o PowerHA SystemMirror para implementar o nível base de um ambiente altamente disponível.

Para obter uma sumarização dos hardwares suportados para o PowerHA SystemMirror, consulte IBM Techdocs: PowerHA hardware support matrix (<http://www.ibm.com/support/techdocs/atmastr.nsf/WebIndex/TD105638>).

Informações relacionadas:

Instalando o PowerHA SystemMirror

Acomodação de disco, grupo de volumes e sistemas de arquivos OEM

Software de cluster PowerHA SystemMirror

Esta seção descreve o software PowerHA SystemMirror que implementa um ambiente altamente disponível.

Componentes de software de um nó do PowerHA SystemMirror

Os componentes de software de um nó do PowerHA SystemMirror possuem muitas camadas.

A figura a seguir mostra as camadas de software em um nó, em um cluster do PowerHA SystemMirror:



Figura 7. Modelo de um nó do cluster do PowerHA SystemMirror

A lista a seguir descreve cada camada:

Camada de aplicativo

Todos os aplicativos tornados altamente disponíveis por meio de serviços fornecidos pelo PowerHA SystemMirror for AIX.

Camada do PowerHA SystemMirror for AIX

Software que reconhece as mudanças dentro de um cluster e coordena o uso dos recursos do AIX para criar um ambiente altamente disponível para dados e aplicativos críticos. O PowerHA SystemMirror complementa as camadas inferiores fornecendo mais serviços para permitir que aplicativos com suas comunicações associadas e serviços de armazenamento de dados fiquem altamente disponíveis.

Camada do RSCT

Os serviços de IBM Reliable Scalable Cluster Technology (RSCT) são empacotados com o AIX. A camada do RSCT fornece recursos para sincronização e coordenação de atividade de cluster usando o sistema de mensagens confiável, em ambientes distribuídos ou de cluster. O RSCT inclui os componentes Resource Monitoring and Control (RMC), Gerenciador de recursos de configuração e Serviços de grupo.

Camada do Cluster Aware AIX

Essa extensão do Sistema operacional AIX fornece recursos para monitorar a associação de nó, o funcionamento da interface de rede e a notificação de evento.

Camada do AIX

Software que fornece o suporte subjacente para um cluster do PowerHA SystemMirror.

Camada do subsistema Gerenciador de Volume Lógico (LVM)

Gerencia o armazenamento de dados no nível lógico.

Camada do subsistema TCP/IP

Fornecer suporte de comunicações para um cluster do PowerHA SystemMirror.

Conceitos relacionados:

“Serviços de disponibilidade do IBM Reliable Scalable Cluster Technology” na página 48

Os serviços de alta disponibilidade do IBM Reliable Scalable Cluster Technology (RSCT) fornecem melhor escalabilidade, notificam os sistemas distribuídos sobre falhas de software e coordenam a recuperação e a sincronização entre todos os subsistemas na pilha de software. O RSCT relata falhas de nó e de rede ao PowerHA Cluster Manager.

Gerenciador do cluster

As mudanças no estado do cluster são mencionadas como eventos de cluster. Em cada nó, o Gerenciador do cluster monitora os subsistemas de hardware e software locais para esses eventos, como um evento de falha de aplicativo.

Em resposta a esses eventos, o Gerenciador de cluster executa um ou mais scripts de evento, como um script de reinicialização de aplicativo. Os Gerenciadores de cluster em todos os nós trocam mensagens para coordenar as ações necessárias em resposta a um evento.

O Gerenciador do cluster é um daemon que é executado em cada nó do cluster. A principal tarefa do Gerenciador do cluster é responder a eventos não planejados, como a recuperação de falhas de software e hardware ou eventos iniciados pelo usuário, como um evento de inclusão de nó. O subsistema RSCT informa o Gerenciador do cluster sobre eventos do nó e relacionados à rede.

Conexão do gerenciador de cluster para outros daemons do PowerHA SystemMirror

O Gerenciador do cluster reúne informações relacionadas a mudanças de estado de nós e interfaces. O Cluster Information Program (Cinfo) obtém essas informações a partir do Gerenciador do cluster e permite que os clientes se comuniquem com o Cinfo para ficarem cientes das mudanças de estado do cluster. Essas informações de estado do cluster são armazenadas no PowerHA SystemMirror Management Information Base (MIB).

Conceitos relacionados:

“Processo e recursos de configuração do PowerHA SystemMirror” na página 90

Estes tópicos fornecem uma visão geral do processo de configuração de cluster do PowerHA SystemMirror e as ferramentas administrativas fornecidas com o software PowerHA SystemMirror.

“Serviços de disponibilidade do IBM Reliable Scalable Cluster Technology” na página 48

Os serviços de alta disponibilidade do IBM Reliable Scalable Cluster Technology (RSCT) fornecem melhor escalabilidade, notificam os sistemas distribuídos sobre falhas de software e coordenam a recuperação e a sincronização entre todos os subsistemas na pilha de software. O RSCT relata falhas de nó e de rede ao PowerHA Cluster Manager.

Componentes de software do PowerHA SystemMirror

O software PowerHA SystemMirror possui muitos componentes.

Subsistemas de comunicação segura do cluster

O PowerHA SystemMirror possui uma infraestrutura de comunicação comum que aumenta a segurança de comunicações intersistêmicas. Os utilitários do cluster usam o daemon Comunicações do cluster que é executado em cada nó para comunicação entre os nós.

Como há somente um caminho de comunicação comum, todas as comunicações são protegidas com segurança. Embora a maioria dos componentes se comuniquem por meio do daemon Comunicações do cluster, os componentes a seguir usam outro mecanismo para comunicações entre nós:

Componente	Método de comunicação
Gerenciador do cluster	RSCT
Sistema de mensagens de pulsação	Cluster Aware AIX
Cluster Information Program (Clinfo)	SNMP

Para usuários que requerem segurança adicional, o PowerHA SystemMirror fornece autenticação de mensagem e criptografia para mensagens enviadas entre os nós do cluster.

Autenticação de conexão

O modo de segurança padrão verifica o endereço IP de origem com relação a uma lista de acesso, verifica se o valor da porta de origem está entre 571 e 1023 e usa o princípio de privilégio mínimo para a execução de comando remoto. A segurança padrão é o modo de segurança padrão. Para obter segurança adicional, é possível configurar um VPN para conexões entre nós para comunicações entre nós do PowerHA SystemMirror.

Criptografia e autenticação de mensagem

A autenticação de mensagem e a criptografia de mensagem fornecem segurança adicional para mensagens do PowerHA SystemMirror enviadas entre os nós do cluster. A autenticação de mensagem assegura a origem e a integridade de uma mensagem. A criptografia de mensagens altera a aparência dos dados enquanto eles são transmitidos e converte-o para seu formato original quando eles são recebidos por um nó que autentica a mensagem. É possível configurar as opções de segurança e opções para distribuir chaves de criptografia usando a interface SMIT.

Serviços de disponibilidade do IBM Reliable Scalable Cluster Technology

Os serviços de alta disponibilidade do IBM Reliable Scalable Cluster Technology (RSCT) fornecem melhor escalabilidade, notificam os sistemas distribuídos sobre falhas de software e coordenam a recuperação e a sincronização entre todos os subsistemas na pilha de software. O RSCT relata falhas de nó e de rede ao PowerHA Cluster Manager.

A pilha de software do PowerHA SystemMirror e do RSCT é executada em cada nó do cluster. O PowerHA SystemMirror Cluster Manager obtém indicações de possíveis falhas a partir de várias origens:

- O RSCT monitora o estado dos dispositivos de rede
- O AIX LVM monitora o estado dos grupos de volumes e discos
- Os monitores de aplicativo monitoram o estado dos aplicativos.

O PowerHA SystemMirror Cluster Manager gerencia as ações de recuperação do cluster no caso de falha de algum componente. O RSCT em execução em cada nó troca uma pulsação com seus peers para que possa monitorar a disponibilidade dos outros nós no cluster. Se a pulsação parar, os sistemas peers gerenciaram o processo de recuperação. Os peers executam as ações necessárias para colocar os aplicativos críticos em execução e para assegurar que os dados não tenham sido corrompidos ou perdidos. O serviços do RSCT incluem os seguintes componentes:

- Resource Monitoring and Control (as versões anteriores do PowerHA SystemMirror usam o subsistema Gerenciamento de evento). Um subsistema distribuído que fornece um conjunto de serviços de alta

disponibilidade. Ele cria eventos pela correspondência de informações sobre o estado dos recursos do sistema com informações sobre as condições dos recursos de interesse para os programas do cliente. Os programas clientes, por sua vez, pode usar notificações de evento para acionar a recuperação de falhas do sistema.

- Serviços de grupo. Um recurso altamente disponível em todo o sistema para coordenar e monitorar as mudanças no estado de um aplicativo em execução em um conjunto de nós. Os Serviços de grupos ajudam no design e na implementação de aplicativos altamente disponíveis e na recuperação consistente de vários aplicativos. Ele executa estas duas tarefas distintas em uma estrutura integrada.

A figura a seguir mostra os componentes principais que compõe a arquitetura do PowerHA SystemMirror:

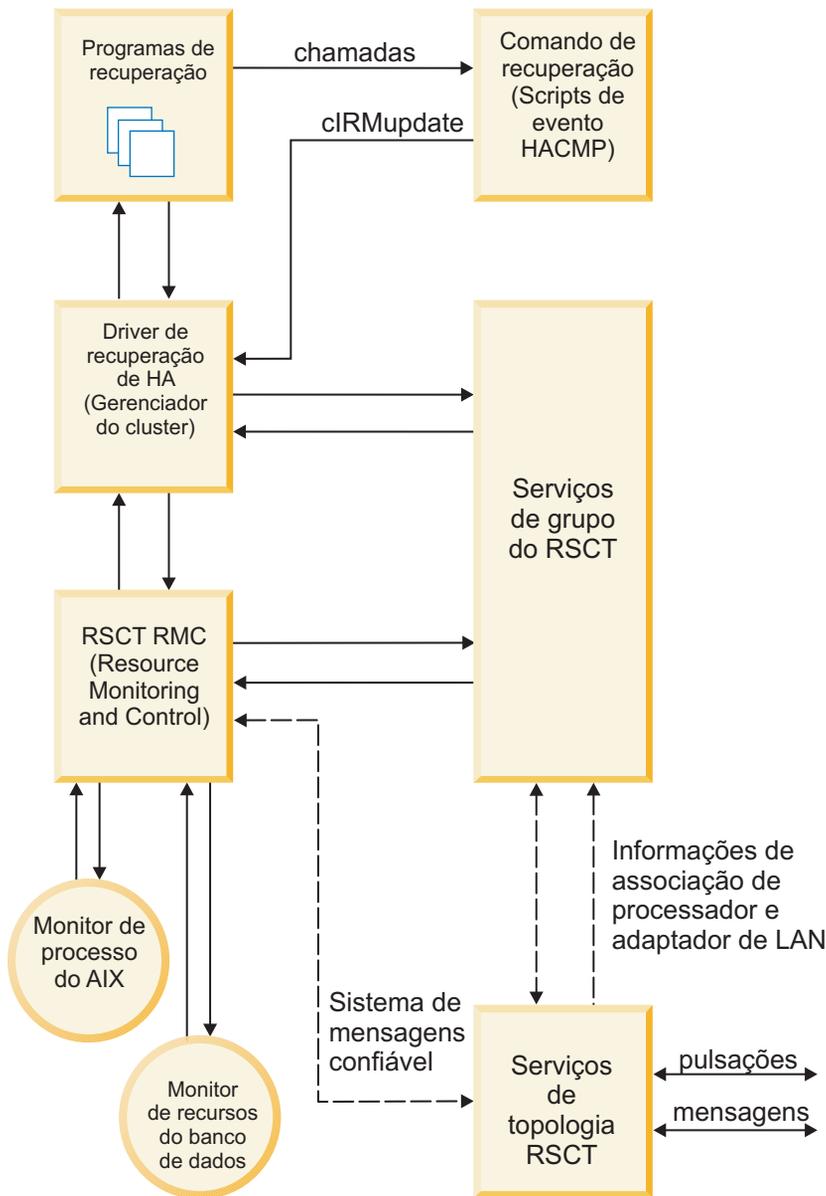


Figura 8. Componentes da arquitetura do PowerHA SystemMirror

Gerenciador de cluster e programas de monitoramento SNMP

Um cluster do PowerHA SystemMirror é dinâmico e pode passar por várias transições de estado ao longo do tempo.

Por exemplo, um nó pode ser incluído ou removido do cluster, ou outro rótulo de IP pode substituir um rótulo de IP de serviço em uma placa da interface de rede física. Cada uma dessas mudanças afeta a composição do cluster, principalmente quando clientes e aplicativos altamente disponíveis precisam usar os serviços fornecidos pelo cluster.

Suporte de SNMP

O Gerenciador de cluster fornece o suporte de Protocolo Simples de Gerenciamento de Rede (SNMP) para aplicativos clientes. O SNMP é uma especificação padrão de mercado para monitoramento e gerenciamento de redes baseadas em TCP/IP. O SNMP inclui um protocolo, uma especificação de banco de dados e um conjunto de objetos de dados. Esse conjunto de objetos de dados forma uma Management Information Base (MIB). O SNMP fornece uma MIB padrão que inclui informações como endereços IP e o número de conexões TCP ativas. O agente SNMP padrão é o daemon **snmpd**.

É possível estender o SNMP para incluir MIBs específicas da empresa contendo informações relacionadas a um ambiente ou aplicativo discreto. No PowerHA SystemMirror, o Gerenciador do cluster mantém informações sobre os objetos definidos na MIB e transmite essas informações para uma estação de monitoramento de rede ou de gerenciamento de rede especializada.

MIB do PowerHA SystemMirror

O Gerenciador de cluster mantém as informações de status do cluster em uma MIB especial do PowerHA SystemMirror. Quando o Gerenciador do cluster é iniciado em um nó do cluster, ele é registrado com o daemon **snmpd** do SNMP e, em seguida, passa a reunir as informações do cluster de forma contínua. O Gerenciador do cluster mantém uma topologia atualizada do cluster na MIB do PowerHA SystemMirror, conforme rastreia os eventos e os estados resultantes do cluster.

Informações relacionadas:

Programando aplicativos clientes

Cluster information program

O Cluster information program (Cinfo), o daemon **clinfo**, é um monitor baseado em SNMP.

O Cinfo, em execução em uma máquina cliente ou em um nó do cluster, consulta o MIB para obter informações do cluster atualizadas. Por meio do Cinfo, as informações sobre o estado de um cluster do PowerHA SystemMirror, dos nós e das redes podem ser disponibilizadas aos clientes e aplicativos. Os clientes podem ser divididos em duas categorias: simples e inteligente.

- Um cliente *simples* visualiza o complexo do cluster como uma entidade única. Se um nó do cluster falhar, o cliente deverá ser reiniciado (ou pelo menos deverá reconectar-se ao nó), caso o controle de endereço IP (IPAT) não esteja ativado.
- Um cliente *inteligente*, por outro lado, reconhece o cluster e reage da maneira apropriada à falha do nó, conectando-se a um nó alternativo e talvez, mascarando a falha para o usuário. Um cliente inteligente como esse deve ter conhecimento do estado do cluster.

O software PowerHA SystemMirror estende os benefícios de servidores, dados e aplicativos altamente disponíveis aos clientes, fornecendo notificação de mudanças de estado do cluster aos clientes por meio do Gerenciador de cluster e das funções da API do Cinfo.

Respondendo a mudanças do cluster

O Cinfo chama o script `/usr/es/sbin/cluster/etc/clinfo.rc` sempre que ocorre um evento de cluster, rede ou nó. Por padrão, o script **clinfo.rc** limpa o cache do ARP do sistema para refletir as mudanças aos endereços IP da rede e não atualiza o cache até que outro endereço responda a uma solicitação *ping*. Geralmente, a limpeza do cache ARP não será necessária se o recurso **troca de endereço de hardware** do PowerHA SystemMirror estiver ativado, porque a troca de endereço de hardware mantém o relacionamento entre um endereço IP e um endereço de hardware.

Em uma rede Ethernet comutada, talvez seja necessário limpar o ARP do cache para assegurar-se de que o novo endereço MAC seja comunicado ao comutador.

Será possível incluir uma lógica no script **clinfo.rc** caso alguma ação adicional seja desejada.

APIs do Clinfo

As APIs do Clinfo fornecem aos desenvolvedores de aplicativos uma interface de linguagem C e C++ para acessar informações de status do cluster. O software PowerHA SystemMirror inclui duas versões de APIs do Clinfo: uma para aplicativo de encadeamento único e uma para aplicativos multiencadeados. O Clinfo e suas APIs associadas permitem que os desenvolvedores gravem aplicativos que reconhecem e respondem a mudanças em um cluster.

Conceitos relacionados:

“Conceitos de nós, redes e pulsação do cluster do PowerHA SystemMirror” na página 15

Esta seção apresenta os principais conceitos relacionados à topologia de cluster e as definições usadas em toda a documentação e na interface com o usuário do PowerHA SystemMirror.

“Assegurando a disponibilidade do aplicativo” na página 54

Esta seção descreve como o software PowerHA SystemMirror assegura a disponibilidade do aplicativo por assegurar a disponibilidade dos componentes do cluster. O PowerHA SystemMirror elimina pontos únicos de falha para todos os principais componentes do sistema e elimina a necessidade de tempo de inatividade planejado para a maioria das tarefas de manutenção de rotina.

Informações relacionadas:

Programando aplicativos clientes

Servidor NFS altamente disponível

A funcionalidade do servidor NFS altamente disponível é incluída no subsistema do produto PowerHA SystemMirror.

Um servidor NFS altamente disponível permite que um processador de backup recupere a atividade atual do NFS caso o servidor NFS primário falhe. A funcionalidade especial do servidor NFS inclui modificações e bloqueios altamente disponíveis nos Network File Systems (NFS). É possível fazer o seguinte:

- Usar o recurso do servidor NFS confiável que preserva bloqueios e dupcache (clusters de 2 nós somente se estiver usando o NFS versão 2 e versão 3)
- Especificar uma rede para montagem cruzada do NFS
- Definir as exportações e montagens cruzadas do NFS no nível do diretório
- Especificar opções de exportação para os diretórios e sistemas de arquivos exportados do NFS
- Configurar dois nós para usar o NFS.

O clusters do PowerHA SystemMirror podem conter até 16 nós. Os clusters que usam o NFS versão 2 e versão 3 pode ter no máximo dois nós e os clusters que usam o NFS versão 4 podem ter no máximo 16 nós.

Acesso a disco externo compartilhado

O software PowerHA SystemMirror suporta dois métodos de acesso a disco externo compartilhado: não simultâneo e simultâneo.

Acesso a disco externo compartilhado não simultâneo:

Em um ambiente não simultâneo, somente um nó possui acesso a um disco externo compartilhado em um determinado momento.

Se esse nó falhar, um dos nós peer adquirirá o disco, montará os sistemas de arquivos definidos como recursos e reiniciará os aplicativos para restaurar os serviços críticos. Geralmente, isso leva de 30 a 300 segundos, dependendo do número e do tamanho dos sistemas de arquivos.

Uma configuração não simultânea pode usar:

- Discos SCSI
- Matrizes de disco SCSI
- Discos seriais
- Discos de conexão direta Fibre Channel
- Discos de conexão SAN Fibre Channel

Para evitar que um disco com falha se torne um ponto único de falha, cada volume lógico em um grupo de volumes compartilhados deve ser espelhado usando o recurso LVM do AIX. Se estiver usando um IBM Enterprise Storage System ou outra matriz RAID suportada, não use o espelhamento do LVM. Os dispositivos RAID fornecem seus próprios dados de redundância.

A maioria dos softwares que podem ser executados no modo de máquina única podem ser gerenciados pelo software PowerHA SystemMirror, sem modificação.

Geralmente, o acesso não simultâneo não requer nenhuma mudança de código para programas de servidor (um sistema de gerenciamento de banco de dados, por exemplo) ou para aplicativos, para fornecer uma solução altamente disponível. Para os usuários finais, a falha do nó parece com uma reinicialização da máquina muito rápida. Um dos nós sobreviventes assume a propriedade dos grupos de recursos do nó com falha e reinicia os aplicativos altamente disponíveis. O Sistema de arquivos registrados, o sistema de arquivos nativo do AIX, garante a integridade do sistema de arquivos. O programa do servidor garante a integridade dos dados da transação.

Os usuários simplesmente efetuam logon em um dos nós sobreviventes e reiniciam o aplicativo. Os procedimentos de logon e reinicialização de aplicativo podem ser orientados pelo software PowerHA SystemMirror. Em algumas configurações do PowerHA SystemMirror, os usuários podem continuar sem precisar executar nenhuma ação, eles somente percebem um atraso durante fallover.

Acesso a disco externo compartilhado simultâneo:

O recurso de acesso simultâneo melhora os benefícios fornecidos por um cluster do PowerHA SystemMirror. O acesso simultâneo permite a simultaneidade de acesso a um grupo de volumes em um subsistema de disco conectado a vários nós (até 16).

Usando o acesso simultâneo, um cluster pode oferecer uma disponibilidade praticamente contínua de dados que se assemelha à tolerância a falhas, mas a um custo bem mais baixo. Além disso, o acesso simultâneo fornece maior desempenho, facilita o desenvolvimento de aplicativos e permite crescimento horizontal. Como o acesso simultâneo fornece a simultaneidade de acesso a dados de vários nós, ferramentas adicionais podem ser necessárias para evitar que vários nós modifiquem o mesmo bloco de dados de maneira conflituosa. O software PowerHA SystemMirror fornece o programa Clinfo que prepara um aplicativo para ser executado em um ambiente de acesso simultâneo. A API do Clinfo fornece uma API por meio da qual os aplicativos podem se tornar "cientes do cluster". A ferramenta Clinfo está descrita anteriormente neste capítulo.

A lista a seguir inclui os benefícios do acesso a disco externo compartilhado simultâneo:

- *A recuperação transparente aumenta a disponibilidade.* O acesso simultâneo reduz significativamente o tempo para um fallover - à vezes somente alguns segundos - porque os sistemas peer já possuem acesso físico ao disco compartilhado e executam suas próprias instâncias do aplicativo.

Em um ambiente de acesso simultâneo, o fallover envolve basicamente a restauração de transações em andamento a partir do processador com falha. O software do servidor em execução nos nós

sobreviventes é responsável por recuperar todas as transações parciais causadas pela paralisação. Como todos os nós possuem acesso simultâneo aos dados, um aplicativo cliente/do servidor pode tentar novamente, imediatamente, uma solicitação que falhou nos nós sobreviventes, que continuam a processar as transações recebidas.

- *A proteção de vários processadores aumenta o rendimento.* Os aplicativos não são mais limitados ao rendimento de um único processador. Agora, múltiplas instâncias de um aplicativo podem ser executadas simultaneamente em vários processadores. À medida que mais energia de processamento for necessária, mais sistemas poderão ser incluídos no cluster para aumentar o rendimento.
- *A imagem única do banco de dados facilita o desenvolvimento e a manutenção de aplicativo.* Em um ambiente não simultâneo, o único meio para melhorar o desempenho é particionar um aplicativo e seus dados. Dividir o código e os dados em partes torna o desenvolvimento e a manutenção de aplicativo mais complexos.

A divisão de um banco de dados requer um alto grau de conhecimento para assegurar que os dados e carga de trabalho sejam distribuídos igualmente entre os processadores. O particionamento de código e de dados não é necessário em um ambiente de acesso simultâneo. Para aumentar o rendimento, várias instâncias do mesmo aplicativo em execução em processadores diferentes podem acessar simultaneamente um banco de dados em um disco externo compartilhado. Uma configuração simultânea pode usar:

- Discos SCSI
- Matrizes de disco SCSI
- Discos seriais
- Discos de conexão direta Fibre Channel
- Discos de conexão SAN Fibre Channel

Ao criar volumes lógicos de acesso simultâneo, use o espelhamento do LVM para evitar que o disco seja um ponto único de falha, exceto para os subsistemas de disco do RAID que fornecem seus próprios espelhamentos. O acesso simultâneo não suporta o uso do Sistema de arquivos registrados. Portanto, o gerenciador de banco de dados deve gravar diretamente para os volumes lógicos brutos ou hdisks no grupo de volumes compartilhados. Um aplicativo deve usar o mesmo método para arbitrar todas as solicitações de dados compartilhados. A maioria dos bancos de dados UNIX comerciais fornecem um modelo de bloqueio que os torna compatíveis com o software PowerHA SystemMirror. Verifique com o fornecedor de banco de dados para determinar se um aplicativo específico suporta o processamento de acesso simultâneo.

Conceitos relacionados:

“Modo simultâneo aprimorado” na página 54

O AIX fornece uma nova forma de modo simultâneo: o *modo simultâneo aprimorado*. Em um modo simultâneo aprimorado, as instâncias do Concurrent Logical Volume Manager (CLVM) coordenam mudanças entre nós por meio do componente Serviços de grupo do recurso Reliable Scalable Cluster Technology (RSCT) no AIX. O PowerHA SystemMirror cria automaticamente grupos de volumes compartilhados como o modo simultâneo aprimorado e converte os grupos de volumes compartilhados existentes para o modo simultâneo aprimorado

Gerenciador de recurso simultâneo

O Gerenciador de recurso simultâneo do PowerHA SystemMirror fornece acesso simultâneo a discos compartilhados em um cluster altamente disponível, permitindo que ações customizadas sejam executadas durante o controle para a adequação às necessidades de negócios.

O Gerenciador de recurso simultâneo inclui o suporte simultâneo aprimorado para grupos de volumes compartilhados em todos os tipos de discos e o gerenciamento de acesso compartilhado simultâneo para os subsistemas de disco do RAID suportados.

Modo simultâneo aprimorado:

O AIX fornece uma nova forma de modo simultâneo: o *modo simultâneo aprimorado*. Em um modo simultâneo aprimorado, as instâncias do Concurrent Logical Volume Manager (CLVM) coordenam mudanças entre nós por meio do componente Serviços de grupo do recurso Reliable Scalable Cluster Technology (RSCT) no AIX. O PowerHA SystemMirror cria automaticamente grupos de volumes compartilhados como o modo simultâneo aprimorado e converte os grupos de volumes compartilhados existentes para o modo simultâneo aprimorado.

Os protocolos dos Serviços de grupo fluem pelos links de comunicações entre os nós do cluster. Qualquer disco suportado pelo PowerHA SystemMirror para conexão com vários nós pode ser incluído em grupo de volumes de modo simultâneo aprimorado.

Software de cluster complementar

Uma grande variedade de ferramentas adicionais auxilia na construção, no gerenciamento e na expansão eficazes de clusters de alta disponibilidade em ambientes AIX.

Estão inclusos:

- O General Parallel File System (GPFS) for AIX, um sistema de arquivos que permeia todo o cluster e permite o acesso compartilhado de usuários a arquivos que abrangem várias unidades de disco e vários nós.
- O Workload Manager for AIX fornece o equilíbrio de recursos entre aplicativos.
- O software Smart Assist para configuração do Oracle, DB2, WebSphere e de outros aplicativos nos clusters do PowerHA SystemMirror.

Informações relacionadas:

Desenvolvendo aplicativos Smart Assist para o PowerHA SystemMirror
Smart Assist for PowerHA SystemMirror

Assegurando a disponibilidade do aplicativo

Esta seção descreve como o software PowerHA SystemMirror assegura a disponibilidade do aplicativo por assegurar a disponibilidade dos componentes do cluster. O PowerHA SystemMirror elimina pontos únicos de falha para todos os principais componentes do sistema e elimina a necessidade de tempo de inatividade planejado para a maioria das tarefas de manutenção de rotina.

Visão geral: Disponibilidade de aplicativo

O principal aspecto de um cluster altamente disponível é sua capacidade de detectar, e responder a, mudanças que poderiam interromper os serviços fundamentais que ele fornece. O software PowerHA SystemMirror permite que um cluster continue a fornecer serviços de aplicativos críticos para uma instalação, mesmo que um componente principal do sistema, por exemplo, a placa de interface de rede não esteja mais disponível.

Quando um componente fica indisponível, o software PowerHA SystemMirror é capaz de detectar e perda e deslocar a carga de trabalho desse componente para outro componente no cluster. Ao planejar um cluster altamente disponível, tente assegurar-se de que os componentes principais não se tornem pontos únicos de falha.

É possível usar o Smart Assists for PowerHA SystemMirror para gerenciar aplicativos altamente disponíveis como DB2, SAP, Oracle e WebSphere. É possível visualizar esses aplicativos como uma única entidade do PowerHA SystemMirror representada como um nome de aplicativo no software PowerHA SystemMirror.

O software PowerHA SystemMirror permite que um cluster continue fornecendo serviços de aplicativos enquanto tarefas de manutenção de rotina são executadas, usando um processo chamado de reconfiguração dinâmica. Na reconfiguração dinâmica, é possível alterar os componentes em um cluster em execução, como incluir ou remover uma interface de nó ou de rede, sem precisar parar e reiniciar os serviços de cluster. A configuração alterada se torna a configuração ativa dinamicamente. Também é possível substituir dinamicamente um disco com falha. As seções a seguir descrevem conceitualmente como usar o software PowerHA SystemMirror para:

- Eliminar pontos únicos de falha em um cluster.
- Minimizar o tempo de inatividade planejado em um cluster do PowerHA SystemMirror com a reconfiguração dinâmica, o gerenciamento de grupo de recursos e os utilitários de gerenciamento de cluster (C-SPOC).
- Minimizar o tempo de inatividade não planejado com o recurso de recuperação rápida e especificando uma política de cronômetro de fallback atrasado para grupos de recursos.
- Minimizar o tempo gasto para executar o controle de disco.
- Interpretar e emular eventos de cluster.

Nota: Você pode precisar monitorar a atividade do cluster enquanto um componente principal falhar e o cluster continuar a fornecer a disponibilidade de um aplicativo.

Conceitos relacionados:

“Processo e recursos de configuração do PowerHA SystemMirror” na página 90

Estes tópicos fornecem uma visão geral do processo de configuração de cluster do PowerHA SystemMirror e as ferramentas administrativas fornecidas com o software PowerHA SystemMirror.

Informações relacionadas:

Smart Assits for PowerHA SystemMirror

Eliminando pontos únicos de falha em um cluster do PowerHA SystemMirror

O software PowerHA SystemMirror permite construir clusters que são ambos altamente disponível e escalável, eliminando pontos únicos de falhas (SPOF). Um ponto único de falha existe quando uma função crítica do cluster é fornecida por um único componente.

Se esse componente falhar, o cluster não terá outra maneira de fornecer essa função e os serviços essenciais ficarão indisponíveis. Por exemplo, se todos os dados de um aplicativo crítico residirem em um único disco que não seja espelhado e esse disco falhar, o disco será um ponto único de falha para o sistema inteiro. Os nós clientes não poderá acessar esse aplicativo até que os dados no disco sejam restaurados.

Pontos únicos de falha potenciais em um cluster do PowerHA SystemMirror

Para ser altamente disponível, um cluster não deve ter nenhum ponto único de falha.

O PowerHA SystemMirror fornece opções de recuperação para os seguintes componentes do cluster:

- Nós
- Aplicativos
- Redes e interfaces de rede
- Discos e adaptadores de disco.

Como o objetivo é eliminar todos os pontos único de falha, compromissos devem ser feitos. Geralmente, há um custo associado à eliminação de um ponto único de falha. Por exemplo, um hardware redundante aumenta o custo. O custo da eliminação de um ponto único de falha deve ser comparado ao custo da perda de serviços caso um componente falhe. O propósito do software PowerHA SystemMirror é fornecer um ambiente de computação altamente disponível, com custo reduzido que possa crescer para atender a demandas de processamento futuras.

Eliminando nós como um ponto único de falha

Os nós são removidos do cluster por meio de uma transição planejada (um encerramento de nó ou uma interrupção de serviços de cluster em um nó) ou por causa de uma falha.

A falha do nó começa quando um nó que monitora um nó vizinho deixa de receber o tráfego de pulsação durante um período de tempo definido. Se os outros nós do cluster concordarem que a falha é uma falha de nó, o nó com falha será removido do cluster e seus recursos serão controlados pelos nós configurados para fazer isso. Um nó ativo pode, por exemplo, assumir o controle dos discos compartilhados configurados no nó com falha. Ou, um nó ativo pode ser mascarado como o nó com falha (adquirindo seu endereço IP de serviço) e executar os processos do nó com falha enquanto ainda mantém seus próprios processos. Assim, os aplicativos clientes podem alternar para um nó sobrevivente para obter serviços de disco e processador compartilhados. O software PowerHA SystemMirror fornece os seguintes recursos para o processamento de falha de nó:

- Controle de disco
- Controle de endereço IP por meio de aliases de IP
- Controle de endereço IP por meio de substituição de IP (com ou sem controle endereço de hardware)

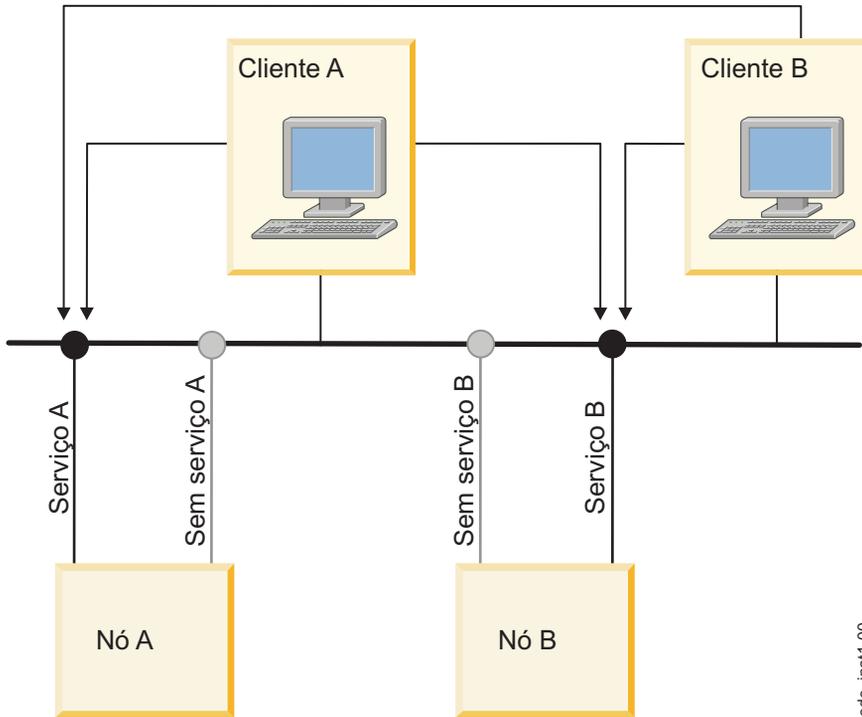
Controle de disco:

Em um ambiente do PowerHA SystemMirror, os discos compartilhados são conectados fisicamente a vários nós.

Controle de disco em ambientes simultâneos

Em configurações de acesso simultâneo, os discos compartilhados são conectados ativamente a vários nós ao mesmo tempo. Portanto, o controle de disco não é necessário quando um nó é removido do cluster.

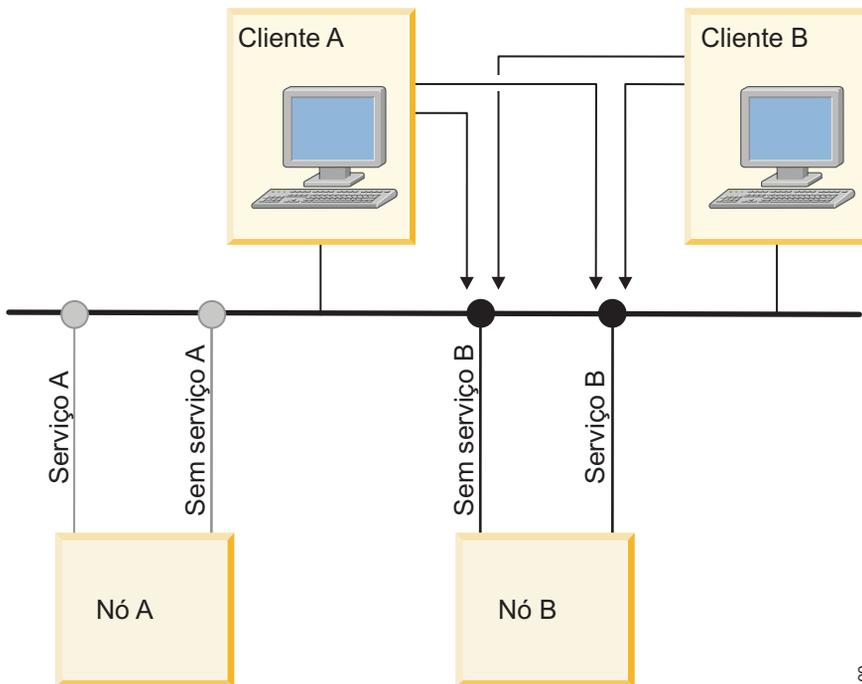
As figuras a seguir ilustram o controle de disco em ambientes simultâneos.



Cada nó fornece um serviço de rede separado

c_2node_ipat1-00

Figura 9. Configuração de acesso simultâneo antes do controle de disco



O Nó B assume o rótulo de IP do nó A em sua interface sem serviço e fornece os serviços de aplicativo do nó A aos clientes nesta interface.

c_2node_ipat2-00

Figura 10. Configuração de acesso simultâneo após o controle de disco

Controle de endereço IP:

O controle de endereço IP (IPAT) é um recurso de rede que permite que um nó adquira o endereço de rede de um nó que foi removido do cluster.

O controle de endereço IP é necessário em um cluster do PowerHA SystemMirror quando um serviço que está sendo fornecido aos clientes está ligado a um endereço IP específico, ou seja, quando um rótulo de IP de serviço, por meio do qual os serviços são fornecidos aos clientes, é incluído como um recurso em um grupo de recursos do cluster. Se, em vez de executar um IPAT, um nó sobrevivente simplesmente executou um simples controle de disco e aplicativo, os clientes não seriam capazes de continuar usando o aplicativo ao endereço IP do servidor especificado. O PowerHA SystemMirror usa o IPAT por meio de Aliases.

As figuras a seguir ilustra o controle de endereço IP.

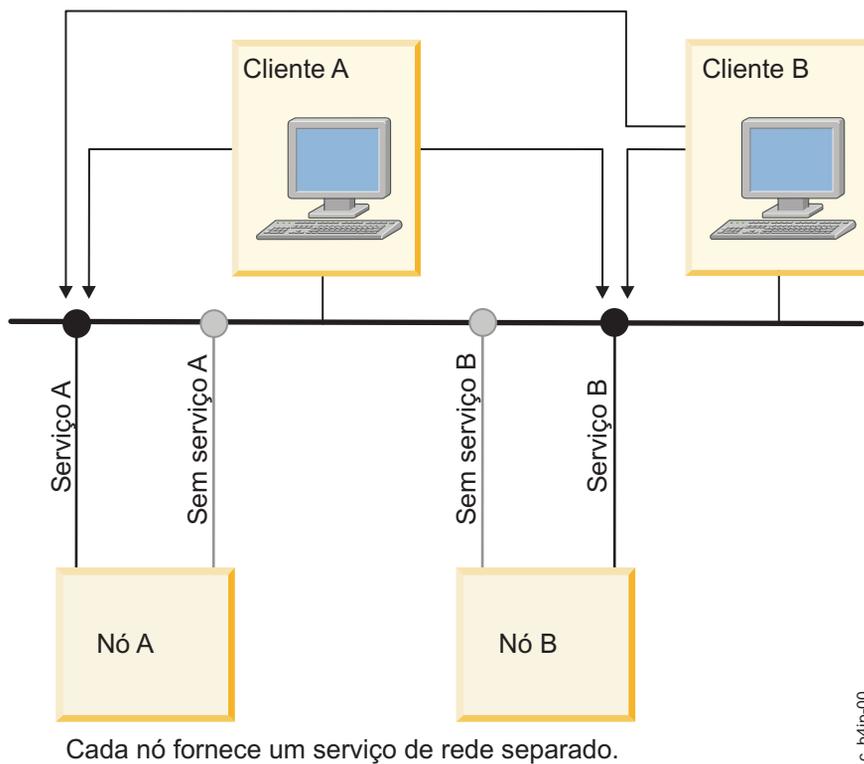
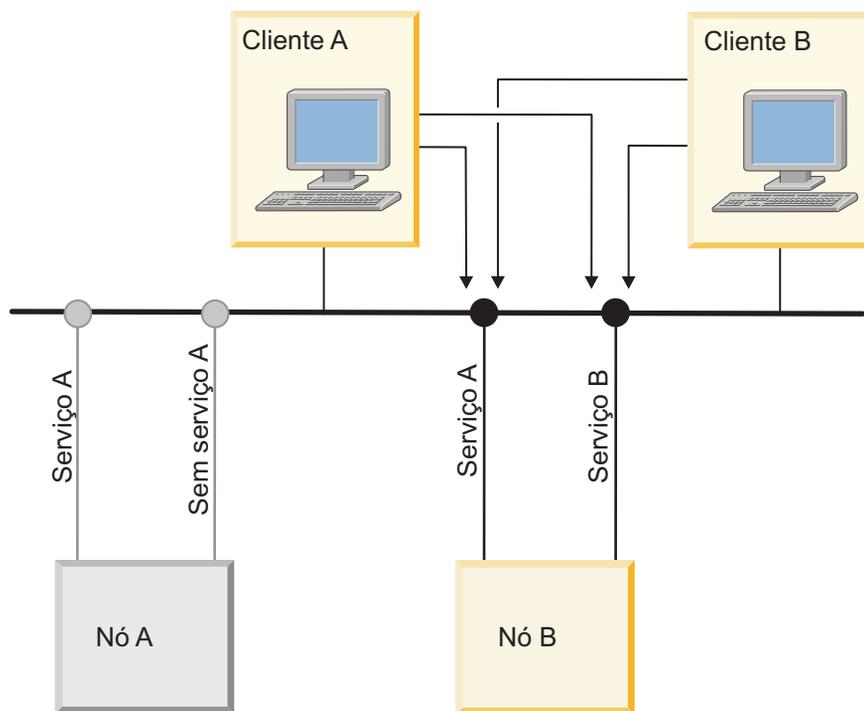


Figura 11. Configuração antes do controle de endereço IP por meio de substituição de IP



O Nó B assume o endereço do nó A em sua interface sem serviço e fornece o serviço de rede de A aos clientes.

c_affip-00

Figura 12. Configuração após controle de endereço IP por meio de substituição de IP

Conceitos relacionados:

“Controle de endereço IP por meio de aliases de IP” na página 27

É possível configurar o controle de endereço IP em determinados tipos de redes usando os recursos de rede de aliasing de IP do sistema operacional AIX.

Eliminando aplicativos como um ponto único de falha

O motivo principal para criar clusters do PowerHA SystemMirror é fornecer um ambiente altamente disponível para aplicativos essenciais.

Por exemplo, um cluster do PowerHA SystemMirror pode executar um programa de servidor de banco de dados que atenda os aplicativos clientes. Os clientes enviam consultas para o programa do servidor que responde às solicitações, acessando um banco de dados armazenado em um disco externo compartilhado.

Em um cluster do PowerHA SystemMirror, esses aplicativos críticos podem ser um ponto único de falha. Para assegurar a disponibilidade desses aplicativos, o nó configurado para controlar os recursos do nó que está sendo removido do cluster também deve reiniciar estes aplicativos para que eles permaneçam disponíveis para os processos do cliente. É possível tornar um aplicativo altamente disponível usando:

- Controlador de aplicativo
- Controle de cluster
- Monitores de aplicativo
- Ferramenta de análise de disponibilidade de aplicativo

Para colocar o aplicativo sob o controle do PowerHA SystemMirror, crie um recurso de cluster de controlador de aplicativo que associe um nome definido pelo usuário do servidor aos nomes de scripts

gravados para iniciar e parar o aplicativo fornecidos pelo usuário. Ao definir um controlador de aplicativo, o PowerHA SystemMirror pode iniciar outra instância do aplicativo no nó de controle quando uma fallover ocorre.

Determinados aplicativos podem se tornar altamente disponíveis sem controladores de aplicativos. É possível colocar esses aplicativos sob o controle de cluster, configurando um aspecto do aplicativo como parte de um grupo de recursos. Por exemplo, os serviços Fast Connect podem ser todos incluídos como recursos em um grupo de recursos do cluster, o que os torna altamente disponíveis no caso de falha do nó ou da interface de rede.

Nota: O controle de aplicativo é geralmente associado ao controle de endereço IP. Se o nó que reiniciar o aplicativo também adquirir o endereço IP do serviço no nó com falha, os clientes somente precisarão ser reconectados ao mesmo endereço IP de serviço. Se o endereço IP não tiver sido controlado, o cliente precisará conectar-se ao novo servidor para continuar acessando o aplicativo.

Além disso, é possível usar o AIX System Resource Controller (SRC) para monitorar a presença ou a ausência de um daemon do aplicativo e para responder de forma adequada.

Monitores de aplicativo

Também é possível configurar um monitor de aplicativos para verificar a falha do processo ou outras falhas de aplicativo e executar ações automaticamente para reiniciar o aplicativo.

É possível configurar vários monitores de aplicativo e associá-los a um ou mais controladores de aplicativo. Por suportar vários monitores por aplicativo, o PowerHA SystemMirror pode suportar configurações mais complexas. Por exemplo, é possível configurar um monitor para cada instância de um servidor Oracle paralelo em uso. Ou, é possível configurar um monitor customizado para verificar o funcionamento de um banco de dados e um monitor de encerramento de processo para detectar instantaneamente o encerramento do processo do monitor.

Ferramenta Análise de disponibilidade de aplicativo

A ferramenta Análise de disponibilidade de aplicativo mede a quantia exata de tempo que cada um dos aplicativos ficou disponível. O software PowerHA SystemMirror coleta, faz o registro de data e hora e registra informações abrangentes sobre os aplicativos escolhidos para serem monitorados com esta ferramenta. Usando o SMIT, é possível selecionar um período de tempo e a ferramenta exibe estatísticas de tempo de atividade e de tempo de inatividade para um aplicativo específico durante esse período.

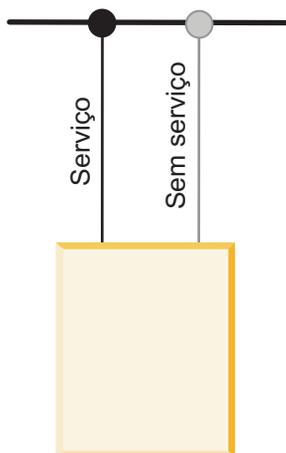
Eliminando interfaces de comunicação como um ponto único de falha

O software PowerHA SystemMirror manipula as falhas de interfaces de rede nas quais um rótulo de IP de serviço está configurado.

Dois tipos dessas falhas são:

- Entre duas interfaces de rede configuradas em um nó, a interface de rede com um rótulo de IP de serviço falha, mas uma placa da interface de rede de backup adicional permanece disponível no mesmo nó. Nesse caso, o Gerenciador do cluster troca as funções desses duas placas da interface nesse nó. Esse tipo de falha de interface de rede é imperceptível, com exceção de um pequeno atraso enquanto o sistema reconfigura a interface de rede em um nó.
- Entre duas interfaces de rede configuradas em um nó, uma interface de rede de backup adicional falha, mas a interface de rede com um rótulo de IP de serviço configurado permanece disponível. Nesse caso, o Gerenciador do cluster detecta uma falha da interface de rede de backup, registra o evento e envia uma mensagem ao console do sistema. Se você desejar algum processamento adicional, será possível customizar o processamento para esse evento.

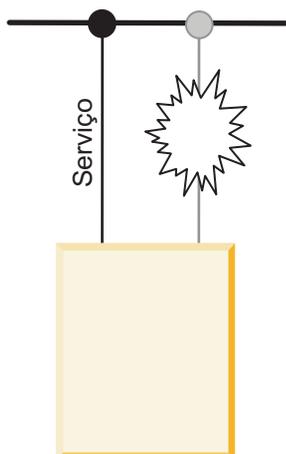
As figuras a seguir ilustram a troca de interface de rede que ocorre no mesmo nó:



Aqui, a interface de serviço fornece a conexão com a rede. A interface sem serviço deve ser ocultada dos aplicativos e somente ser conhecida para o Gerenciador do cluster.

c_adapswap1-00

Figura 13. Configuração antes da troca do adaptador de rede



Aqui, a interface de serviço falhou e o Gerenciador do cluster designa a interface sem serviço antiga como a nova interface de serviço.

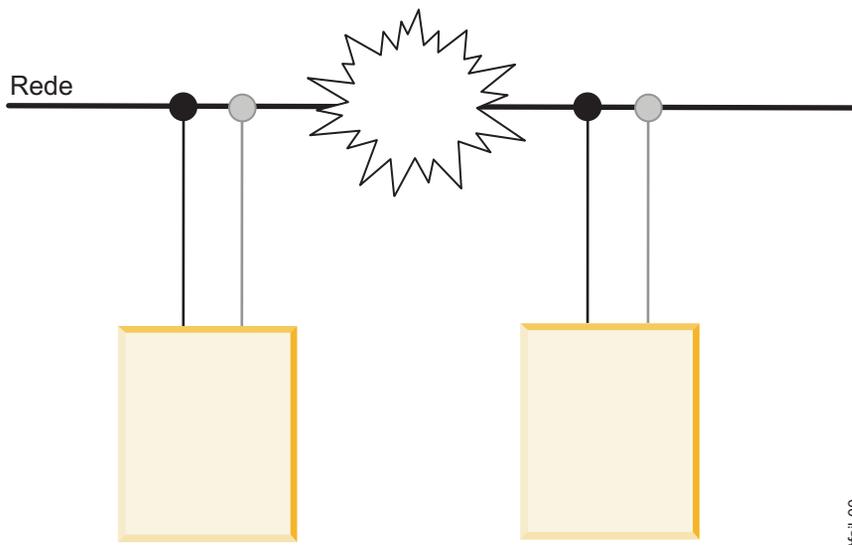
c_adapswap2-00

Figura 14. Configuração após a troca do adaptador de rede

Eliminando redes como um ponto único de falha

A falha de rede ocorre quando uma rede do PowerHA SystemMirror falha para todos os nós em um cluster. Este tipo de falha ocorre quando os nós do cluster não podem se acessar entre si usando as placas da interface de rede configuradas para uma determinada rede do PowerHA SystemMirror.

A figura a seguir ilustra uma falha de rede:



Aqui, a rede que conecta os nós falhou. Os nós não podem mais se comunicar por esta rede.

c_2node_netfail-00

Figura 15. Falha de rede

A primeira linha de defesa do software PowerHA SystemMirror contra uma falha de rede é possuir os nós no cluster conectados por várias redes. Se uma rede falhar, o software PowerHA SystemMirror usará uma rede que ainda esteja disponível para o tráfego de cluster e para monitorar o status dos nós.

É possível especificar ações adicionais para processar uma falha de rede - por exemplo, rotear novamente por meio de uma rede alternativa. A presença de pelo menos duas redes para proteção contra falha de rede é altamente recomendada.

Quando ocorre um evento de falha de rede local, o Gerenciador do cluster toma ações de recuperação seletivas para os grupos de recursos que contêm um rótulo de IP de serviço conectado a essa rede. O Gerenciador do cluster tenta mover somente os grupos de recursos afetados pelo evento de falha de rede local, em vez de todos os grupos de recursos em um determinado nó.

Isolamento do nó e clusters particionados:

O isolamento do nó é quando todas as redes que conectam duas ou mais partes do cluster falham. Cada grupo de nós (um ou mais) fica completamente isolado dos outros grupos. Um cluster particionado é aquele no qual determinados grupos de nós ficam impossibilitados de se comunicarem com outros grupos de nós.

Na ilustração a seguir de um cluster particionado, o Nó A e o Nó C estão em um lado da partição e o Nó B e o Nó D estão no outro lado da partição.

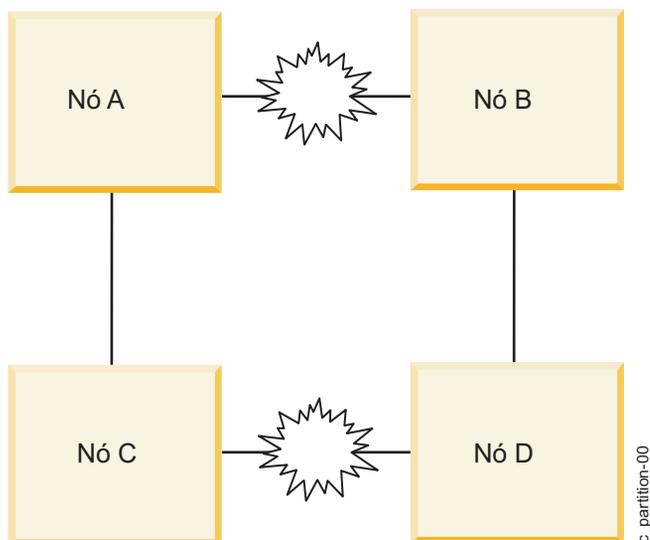


Figura 16. Cluster particionado

O problema com um cluster particionado é que os nós em um lado da partição interpretam a ausência de pulsações dos nós no outro lado da partição como uma falha desses nós, em seguida, geram eventos de falha do nó para eles. Quando isso ocorre, os nós em cada lado do cluster (se estiver configurado assim) tentam controlar os recursos de um nó que ainda está ativo e, portanto, ainda possui esses recursos legitimamente. Essas tentativas de controle podem causar resultados imprevisíveis no cluster, por exemplo, a distorção de dados devido a um disco que está sendo reconfigurado.

Usando redes de área de armazenamento para evitar o particionamento:

Para evitar que a perda de uma rede TCP/IP cause um particionamento do cluster, o software Cluster Aware AIX pulsará sobre a rede de área de armazenamento que abrange os nós do cluster.

Informações relacionadas:

Planejando o PowerHA SystemMirror

Eliminando discos e adaptadores de disco como um ponto único de falha

O software PowerHA SystemMirror não manipula sozinho diretamente as falhas de disco e de adaptador de disco. Em vez disso, o AIX manipula essas falhas pelo espelhamento do LVM nos discos e pela redundância de dados internos nos discos que possuem essa capacidade.

Por exemplo, ao configurar o sistema com vários adaptadores de armazenamento, pulsação por meio da rede de área de armazenamento e espelhamento do LVM, qualquer componente no subsistema de disco (adaptador, cabeamento, discos) poderá falhar sem causar indisponibilidade de dados no disco.

Se você estiver usando o IBM 2105 ESS ou outros subsistemas de armazenamento de RAID, a própria matriz de disco será responsável por fornecer a redundância de dados.

Recurso de notificação de erro do AIX

Com o recurso Notificação de erro do AIX é possível detectar um evento que não seja monitorado especificamente pelo software PowerHA SystemMirror. Por exemplo, um adaptador de disco falha ao programar uma resposta ao evento.

Erros de hardware permanentes em unidades de disco, controladores ou adaptadores podem afetar a resiliência de dados com falha. Ao monitorar esses erros por meio de métodos de notificação de erro, é possível avaliar o impacto de uma falha na capacidade do cluster de fornecer alta disponibilidade. Uma

implementação simples de notificação de erro seria enviar uma mensagem de correio ao administrador do sistema para que ele investigue melhor o problema. Uma implementação mais complexa poderia incluir um lógica ao analisar a falha e decidir continuar o processamento, parar o processamento ou escalar a falha para uma falha de nó e permitir que o nó de controle disponibilize os recursos do grupo de volumes aos clientes.

É altamente recomendado implementar um método de notificação de erro para todos os erros que afetam o subsistema de disco. Isso garante que a resiliência comprometida pela falha não permaneça sem ser detectada.

Os métodos de notificação de erro do AIX são usados automaticamente no PowerHA SystemMirror para monitorar determinados erros recuperáveis do LVM, como erros de perda do grupo de volumes.

Notificação automática de erro

É possível configurar a notificação de erro automaticamente para determinados recursos do cluster, usando uma opção específica no SMIT. Se você selecionar essa opção, a notificação de erro será ativada automaticamente em todos os nós no cluster para determinados dispositivos.

Determinados tipos de erros não recuperáveis, como erros de adaptador de disco, são suportados pela notificação automática de erro. Esse recurso não suporta erros de mídia, erros recuperados ou erros temporários. Um dos dois métodos de notificação de erro é designado a todos os tipos de erros suportados pela notificação automática de erro.

Além disso, se você incluir um grupo de volumes em um grupo de recursos, o PowerHA SystemMirror criará um método de Notificação de erro do AIX para ele. Quando um grupo de volumes perde o quorum, o PowerHA SystemMirror usa esse método para mover de forma seletiva o grupo de recursos afetado para outro nó. Não edite ou altere os métodos de notificação de erro geradas pelo PowerHA SystemMirror.

O recurso de notificação de erro do PowerHA SystemMirror não suporta discos MPIO.

Minimizando o tempo de inatividade planejado com o PowerHA SystemMirror

O software PowerHA SystemMirror permite executar a maioria das tarefas de manutenção de rotina em um cluster ativo dinamicamente, sem precisar parar e, em seguida, reiniciar os serviços de cluster para fazer com que a configuração alterada seja a configuração ativa.

Iniciando serviços de cluster sem parar aplicativos

É possível iniciar os serviços de cluster do PowerHA SystemMirror no(s) nó(s) sem parar os aplicativos.

Informações relacionadas:

Iniciando e parando serviços do cluster

Reconfiguração automática e dinâmica

A reconfiguração automática e dinâmica (DARE) é acionada quando você sincroniza a configuração de cluster após ter feito mudanças em um cluster ativo. A aplicação de uma captura instantânea de cluster usando o SMIT também aciona um evento de reconfiguração dinâmica.

Por exemplo, para incluir um nó em um cluster em execução, simplesmente conecte o nó no cluster, inclua o nó na topologia de cluster em qualquer um dos nós de cluster existentes e sincronize o cluster. O novo nó é incluído na definição de topologia de cluster em todos os nós do cluster e a configuração alterada se torna a configuração ativa atualmente. Após a conclusão do evento de reconfiguração dinâmica, é possível iniciar os serviços de cluster no novo nó.

O PowerHA SystemMirror verifica a configuração modificada antes de torná-la a configuração atualmente ativa, para assegurar que as mudanças resultem em uma configuração válida.

Como a reconfiguração dinâmica funciona:

Quando a reconfiguração automática e dinâmica (DARE) é usada em um cluster em execução e o PowerHA SystemMirror é iniciado, o DARE cria uma cópia privada das classes de objeto específicas do PowerHA SystemMirror armazenadas no Object Data Model (ODM) padrão do sistema.

A partir de agora, o ODM é mencionado como o Banco de dados de configuração do PowerHA SystemMirror. Dois diretórios armazenam dados do banco de dados de configuração:

- O Active Configuration Directory (ACD), um diretório privado, armazena os dados do Banco de dados de configuração do PowerHA SystemMirror a serem referenciados por todos os daemons, scripts e utilitários do PowerHA SystemMirror em um nó em execução.
- O Default Configuration Directory (DCD), o diretório padrão do sistema, armazena os dados e o banco de dados de configuração do PowerHA SystemMirror.

Nota: A operação do DARE é descrita aqui para complementação. Não é necessária intervenção manual para assegurar que o PowerHA SystemMirror realize essas operações. O PowerHA SystemMirror gerencia corretamente todas as operações de reconfiguração dinâmica no cluster.

O DCD é o diretório chamado `/etc/objrepos`. Esse diretório contém as classes de objeto do sistema padrão, como o banco de dados de dispositivo customizado (CuDv) e o banco de dados de dispositivo predefinido (PdDv), além das classes de objeto específicas do PowerHA SystemMirror. O ACD é `/usr/es/sbin/cluster/etc/objrepos/active`.

Nota: Ao configurar um cluster, modifique os dados do banco de dados de configuração do PowerHA SystemMirror armazenados no DCD, não os dados no ACD. O SMIT e todos os outros utilitários de configuração do PowerHA SystemMirror modificam os dados do banco de dados de configuração do PowerHA SystemMirror no DCD. Além disso, todos os comandos de usuário que exibem os dados do banco de dados de configuração do PowerHA SystemMirror, como o comando `cllsif`, leem os dados a partir do DCD.

A figura a seguir ilustra como todos os daemons, scripts e utilitários do PowerHA SystemMirror fazem referência ao ACD ao acessar as informações de configuração.

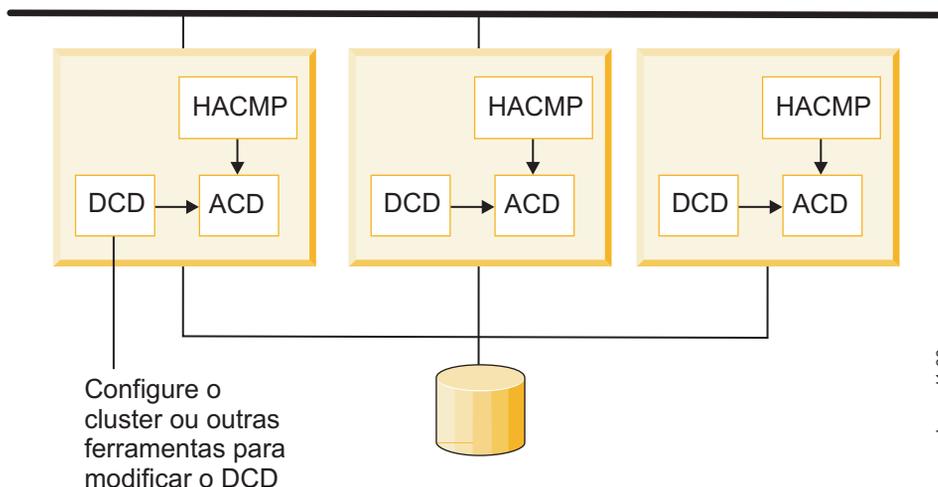


Figura 17. Relacionamento do PowerHA SystemMirror com o ACD quando um cluster é iniciado.

Reconfigurando um cluster dinamicamente:

O software PowerHA SystemMirror depende do local de determinados repositórios do banco de dados de configuração do PowerHA SystemMirror para armazenar dados de configuração.

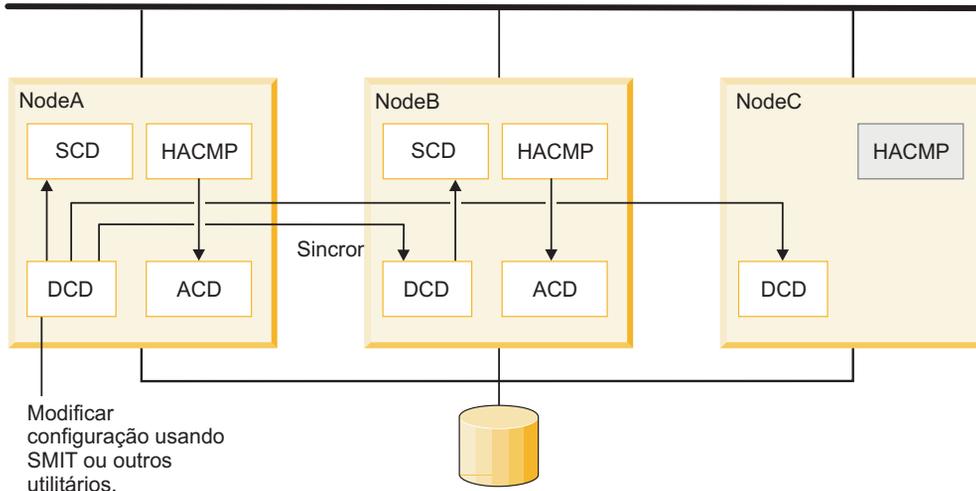
A presença ou a ausência desses repositórios, às vezes, é usada para determinar as etapas executadas durante a configuração e a operação do cluster. A variável de ambiente ODMPATH permite que os comandos e as sub-rotinas do banco de dados de configuração do PowerHA SystemMirror consultem locais que não sejam o local padrão (mantido na variável de ambiente ODMDIR) se o objeto consultado não existir no local padrão. É possível configurar esta variável, mas ela não deve ser configurada para incluir o diretório `/etc/objrepos` ou a integridade das informações de configuração do PowerHA SystemMirror será perdida.

Para alterar a configuração de um cluster ativo, modifique a definição de cluster armazenada nas classes específicas do banco de dados de configuração do PowerHA SystemMirror armazenadas no default configuration directory (DCD) usando o SMIT. Ao alterar a configuração de cluster em um cluster ativo, são usados os mesmos caminhos do SMIT para fazer as mudanças, mas elas não entram em vigor imediatamente. Portanto, é possível fazer várias mudanças em uma operação. Ao sincronizar a configuração entre todos os nós do cluster, ocorre um evento de reconfiguração dinâmica em todo o cluster. Quando o PowerHA SystemMirror processa um evento de reconfiguração dinâmica, ele atualiza as classes de objeto do banco de dados de configuração do PowerHA SystemMirror armazenadas no DCD em cada cluster e substitui os dados do banco de dados de configuração do PowerHA SystemMirror armazenados no ACD pelos novos dados do banco de dados de configuração do PowerHA SystemMirror no DCD, em uma transição coordenada em todo o cluster. Ele também atualiza os daemons do cluster para que eles façam referência aos novos dados de configuração.

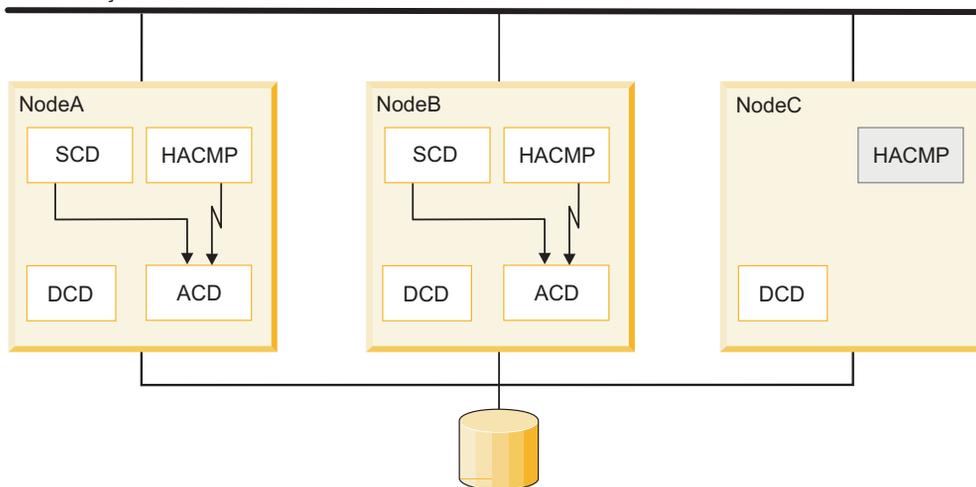
Após esse processamento, a pulsação do cluster será brevemente suspensa e o cluster ficará em um estado instável. A configuração alterada se tornará a configuração ativa. Quando os serviços de cluster forem iniciados no nó recém-incluído, ele será integrado automaticamente no cluster.

A figura a seguir ilustra o processamento envolvido com a inclusão de um nó em um cluster ativo usando a reconfiguração dinâmica.

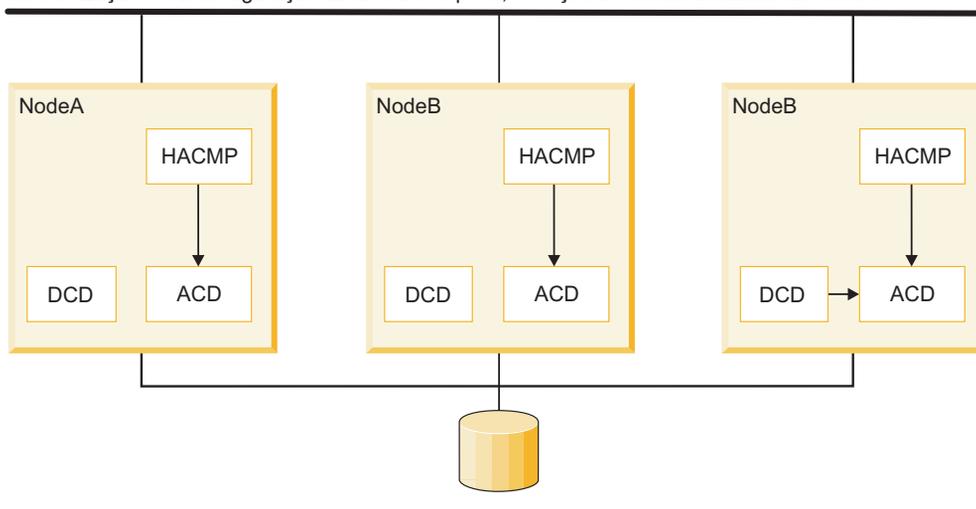
Sincronização 1: Os DCDs são sincronizados



Sincronização 2: Os daemons são atualizados



Sincronização 3: Reconfiguração dinâmica completa; serviços de cluster iniciados no NodeC



c_odm_dynreconfig-00

Figura 18. Processamento de reconfiguração dinâmica

O nó a ser incluído é conectado a um cluster em execução, mas os serviços de cluster estão inativos nesse nó. A configuração é redefinida no NodeA. Quando as mudanças na configuração são sincronizadas, os dados do banco de dados de configuração do PowerHA SystemMirror armazenados no DCD no NodeA

são copiados para os DCDs nos outros nós do cluster e um evento de reconfiguração dinâmica é acionado. O PowerHA SystemMirror copia os novos dados do banco de dados de configuração do PowerHA SystemMirror no DCD para um local temporário em cada nó, chamado de Staging Configuration Directory (SCD). O local do SCD é `/usr/es/sbin/cluster/etc/objrepos/stage`. Usando esse local temporário, o PowerHA SystemMirror permite que você comece a fazer mudanças adicionais na configuração enquanto uma reconfiguração dinâmica está em andamento. Antes de copiar os novos dados do banco de dados de configuração do PowerHA SystemMirror no SCD sobre os dados atuais do banco de dados de configuração do PowerHA SystemMirror no ACD, o PowerHA SystemMirror verifica a nova configuração.

Nota: É possível iniciar uma segunda reconfiguração enquanto uma reconfiguração dinâmica está em andamento, mas não é possível sincronizá-la. A presença de um SCD em qualquer nó do cluster age como um bloqueio, impedindo o início de uma nova reconfiguração dinâmica.

Gerenciamento de grupo de recursos

É possível usar o utilitário de gerenciamento de grupo de recursos (`clRGmove`) para mover grupos de recursos para outros nós do cluster e sites ou colocá-los online ou offline sem parar os serviços de cluster.

Fornece flexibilidade para gerenciar os grupos de recursos e seus aplicativos. Também é possível usar esse utilitário para liberar o nó dos grupos de recursos para executar manutenção do sistema em um nó do cluster específico. O utilitário de gerenciamento de grupo de recursos (`clRGmove`) do PowerHA SystemMirror foi melhorado significativamente, o que facilita a movimentação dos grupos de recursos para o gerenciamento do cluster. Além disso, é possível entender com mais facilidade as consequências da movimentação manual dos grupos de recursos. Por exemplo, é possível prever se os grupos permanecerão nos nós para os quais forem movidos. O PowerHA SystemMirror segue este princípio: Em todos os casos, quando os grupos de recursos são movidos, eles permanecem nos nós até que sejam movidos novamente. O PowerHA SystemMirror move-os quando precisa recuperá-los.

Se você deseja verificar se o grupo de recursos está hospedado atualmente no nó com prioridade mais alta que está disponível no momento, o PowerHA SystemMirror apresentará opções de listas de seleção inteligentes para nós e sites. Por exemplo, se o grupo estiver hospedado atualmente em um nó e o PowerHA SystemMirror localizar outro nó com uma prioridade mais alta, a lista de seleção do SMIT com nós de destino indicará qual nó possui a prioridade mais alta. Dessa forma, sempre é possível escolher mover esse grupo para esse nó. Quando os grupos são movidos para outros nós, estas regras se aplicam:

- Para grupos de recursos com a política de fallback Nunca efetuar fallback, a movimentação de um grupo não terá efeito sobre o comportamento desse grupo durante os eventos de cluster futuros. O mesmo é verdadeiro para os grupos de recursos com a política de fallback de site Online em qualquer site.
- Para os grupos de recursos com uma política de fallback que não seja Nunca efetuar fallback ou Preferir o site primário, a movimentação de um grupo fará com que um nó de destino se torne um nó com uma prioridade mais alta agindo até que seja movido novamente. Depois que o nó for movido, ele se tornará um nó de prioridade mais alta agindo.

Uma consideração importante sobre esse comportamento tange os grupos de recursos que possuem a política Efetuar fallback para o nó de prioridade mais alta ou a política Preferir o site primário. Quando um grupo de recursos desse tipo é movido para um nó que não seja seu nó de prioridade mais alta ou para um site primário, o nó para o qual ele é movido se torna seu nó preferencial temporariamente, embora não seja seu nó de prioridade mais alta (conforme está configurado). Esses grupos permanecem nos nós para os quais foram movidos até que sejam movidos novamente. Os grupos também efetuem fallback para esses nós ou sites.

Conceitos relacionados:

“Processo e recursos de configuração do PowerHA SystemMirror” na página 90
Estes tópicos fornecem uma visão geral do processo de configuração de cluster do PowerHA SystemMirror e as ferramentas administrativas fornecidas com o software PowerHA SystemMirror.

Informações relacionadas:

Gerenciando grupos de recursos em um cluster
Planejando grupos de recursos

Gerenciamento de grupo de recursos solicitado pelo usuário versus gerenciamento de grupo de recursos automático:

Em geral, para manter os aplicativos altamente disponíveis, o PowerHA SystemMirror gerencia automaticamente (e, às vezes, move) grupos de recursos e os aplicativos incluídos.

Por exemplo, quando é necessário recuperar um grupo de recursos, o PowerHA SystemMirror pode tentar recuperá-lo automaticamente em outro nó durante as operações de failover ou fallback. Ao mover um grupo, o PowerHA SystemMirror segue as políticas de grupo de recursos especificadas e outras configurações (por exemplo, em vez de recuperar automaticamente um grupo de recursos com falha em outro nó, é possível solicitar que o PowerHA SystemMirror somente notifique sobre a falha do grupo).

Quando você solicitar que o PowerHA SystemMirror execute o gerenciamento de grupo de recursos, ele usará o utilitário *clRGmove*, que move os grupos de recursos, chamando um evento *rg_move*.

Nota: Ao solucionar problemas de arquivos de log, é importante distinguir entre um evento *rg_move* que, em alguns casos, é acionado automaticamente pelo PowerHA SystemMirror e um evento *rg_move* que ocorre quando você solicita que o PowerHA SystemMirror gerencie os grupos de recursos. Para identificar as causas das operações executadas nos grupos de recursos no cluster, procure a saída de comando no SMIT e informações no arquivo *hacmp.out*.

Operações de gerenciamento de grupo de recursos:

É possível usar o gerenciamento de grupo de recursos para executar várias tarefas.

Use o gerenciamento de grupo de recursos para:

- Mover um grupo de recursos do nó em um site para o nó em outro site.
- Mover um grupo de recursos de um nó para outro.

Em um cluster em funcionamento, mova temporariamente um grupo de recursos não simultâneos a partir do nó que ele reside atualmente para qualquer nó de destino. Os grupos de recursos movidos continuam a comportar-se de forma consistente com a maneira em que foram configurados, ou seja, seguem as políticas de inicialização, failover e fallback especificadas para eles. A interface com o usuário SMIT permitirá especificar e prever claramente o comportamento do grupo de recursos, se você decidir movê-lo para outro nó.

Se você usar o SMIT para mover um grupo de recursos para outro nó, ele permanecerá em seu novo nó de destino até que seja movido manualmente novamente. O PowerHA SystemMirror poderá precisar movê-lo durante um failover.

- Mover o grupo de recursos de volta para o nó que tinha originalmente sua prioridade mais alta.

O grupo de recursos pode ter uma política de fallback ou não. Se um grupo de recursos tiver a política de fallback Efetuar fallback para o nó de prioridade mais alta, depois que for movido, o grupo assumirá que agora o "novo" nó é o seu local temporário preferencial e efetuará fallback para esse nó. Para alterar este comportamento, sempre será possível mover o grupo de volta para o nó que tinha originalmente a prioridade mais alta de nó.

Da mesma forma, se houver um grupo de recursos que possua uma política de fallback Nunca efetuar fallback, quando este grupo de recursos for movido, ele não será movido de volta para o nó a partir do qual foi movido, mas permanecerá em seu novo nó de destino, até que seja movido novamente para outro nó. Dessa forma, é possível assegurar-se de que o grupo sempre siga a política Nunca efetuar fallback especificada para ele.

- Colocar um grupo de recursos online ou offline em um nó ou em todos os nós no cluster.

Informações relacionadas:

Ponto único de controle do cluster

Com o utilitário Ponto único de controle do cluster (C-SPOC), é possível fazer mudanças em todo o cluster a partir de um único nó do cluster.

Em vez de executar tarefas administrativas em cada nó do cluster, é possível usar a interface SMIT para emitir um comando C-SPOC uma única vez, em um único nó e a mudança será propagada para todos os nós do cluster.

Conceitos relacionados:

“Gerenciamento de sistemas do PowerHA SystemMirror com C-SPOC” na página 94

Para facilitar o gerenciamento de um cluster, o PowerHA SystemMirror fornece maneiras de executar comandos a partir de um nó e, em seguida, verificar e sincronizar as mudanças para todos os outros nós. É possível usar a ferramenta de Gerenciamento de sistema do PowerHA SystemMirror, o ponto único de controle do cluster (C-SPOC) para incluir usuários, arquivos e hardware automaticamente sem parar as tarefas essenciais.

Troca de adaptador dinâmico

A funcionalidade de troca dinâmica de adaptador permite trocar o endereço IP de uma placa da interface de rede (NIC) ativa pelo endereço IP de uma placa da interface de rede "de backup" disponível e ativa, especificada pelo usuário, no mesmo nó e na mesma rede.

Os serviços de cluster não precisam ser interrompidos para a troca.

Esse recurso pode ser usado para remover um endereço IP de uma placa da interface de rede que está se comportando de forma errada para outra NIC, sem encerrar o nó. Ele também pode ser usado se uma NIC conectável a quente estiver sendo substituída no nó. As NICs conectáveis a quente podem ser removidas e substituídas fisicamente sem desligar o nó. Quando a NIC (conectável a quente) que você deseja substituir for removida do nó, o PowerHA SystemMirror deixará a NIC indisponível como um backup.

É possível configurar a troca do adaptador usando o SMIT. O endereço IP do serviço é movido da sua NIC atual para uma NIC especificada pelo usuário. Em seguida, o endereço IP do serviço torna-se um endereço "de backup" disponível. Quando a nova placa é colocada no nó, a NIC é incorporada no cluster como um "backup" disponível novamente. Em seguida, é possível trocar o endereço IP da NIC de backup para a NIC original.

Nota: Esse tipo de troca dinâmica de adaptador somente pode ser executado em um único nó. Não é possível trocar o endereço IP pelo endereço IP em um nó diferente com essa funcionalidade. Para mover um endereço IP de serviço para outro nó, mova seu grupo de recursos usando o utilitário Gerenciamento de grupo de recursos.

Verificação e sincronização automáticas

A verificação e sincronização automáticas minimizam o tempo de inatividade quando um nó é incluído ao cluster.

Este processo é executado antes de iniciar os serviços de cluster e verifica e assegura que os nós a serem incluídos em um cluster sejam sincronizados corretamente. Este processo verifica os nós inserindo configurações ativas ou inativas. Alguns exemplos das inconsistências de configuração típicas corrigidas pela verificação e sincronização automáticas são:

- Os endereços IP estão configurados nas interfaces de rede que o RSCT espera
- Os grupos de volumes compartilhados não são configurados para serem ativados automaticamente
- Os Sistemas de arquivos não são configurados para serem montados automaticamente.

Se algum erro de configuração adicional for localizado, os serviços de cluster não serão iniciados no nó e mensagens de erro detalhadas permitirão resolver as inconsistências.

Informações relacionadas:

Verificando e sincronizando um cluster do PowerHA SystemMirror

Minimizando o tempo de inatividade não planejado

Outro objetivo importante do PowerHA SystemMirror é minimizar o tempo de inatividade não planejado em resposta a falhas de componente de cluster não planejadas.

O software PowerHA SystemMirror fornece os seguintes recursos para minimizar o tempo de inatividade não planejado:

- Recuperação rápida para acelerar o fallover em clusters grandes
- Um cronômetro de fallback atrasado para permitir que um grupo de recursos customizado efetue fallback em um horário especificado
- IPAT por meio de aliases de IP para acelerar o processamento durante a recuperação de rótulos de IP de serviço
- Recuperação automática de grupos de recursos que estão no estado ERROR, sempre que um nó do cluster fica ativo. Para obter mais informações, consulte a seção a seguir.

Recuperando grupos de recursos na inicialização do nó

O Gerenciador do cluster tenta colocar os grupos de recursos que estão atualmente no estado ERROR no estado online (ativo) no nó que está sendo incluído. Isto aumenta mais as chances de colocar os aplicativos novamente online. Quando um nó for inicializado, se um grupo de recursos estiver no estado ERROR em qualquer nó no cluster, esse nó tentará adquirir o grupo de recursos.

Observe que o nó deve ser incluído na lista de nós para o grupo de recursos.

A recuperação do grupo de recurso na inicialização do nó é diferente para grupos de recursos não simultâneos e simultâneos:

- Se o nó inicial falhar ao ativar um grupo de recursos não simultâneos que esteja no estado ERROR, o grupo de recursos continuará a efetuar fallover para outro nó na lista de nós, caso haja um nó disponível. A ação de fallover continuará até que todos os nós disponíveis na lista de nós tenham sido tentados.
- Se o nó inicial falhar ao ativar um grupo de recursos simultâneos que esteja no estado ERROR no nó, o grupo de recursos simultâneos será deixado no estado ERROR nesse nó. Observe que o grupo de recursos ainda pode permanecer online em outros nós.

Recuperação rápida

O recurso de recuperação rápida do PowerHA SystemMirror acelera o fallover em clusters grandes.

A recuperação rápida permite selecionar uma verificação de consistência de sistemas de arquivos e um método de recuperação de sistemas de arquivos:

- Se você configurar um sistema de arquivos para usar uma verificação de consistência e um método de recuperação, ele economizará tempo, executando logredo em vez de fsck em cada sistema de arquivos. Se a montagem subsequente falhar, ele executará um fsck completo.

Se um sistema de arquivos sofrer danos em uma falha, mas ainda puder ser montado, o logredo poderá não ter sucesso ao corrigir o dano, produzindo um erro durante o acesso a dados.

- Além disso, ele economizará tempo adquirindo, liberando e efetuando fallover de todos os grupos de recursos e sistemas de arquivos em paralelo, em vez de em série.

Não configure o sistema para executar esses comandos em paralelo se você tiver sistemas de arquivos compartilhados, aninhados. Eles devem ser recuperados de forma sequencial. (Observe que o utilitário de

verificação de cluster não relata inconsistências do sistema de arquivos e de recuperação rápida.) Os comandos `varyonvg` e `varyoffvg` sempre são executados em grupos de volumes em paralelo, independentemente da configuração do método de recuperação.

Cronômetro de fallback atrasado para grupos de recursos

O cronômetro de fallback atrasado permite que um grupo de recursos efetue fallback para o nó de prioridade mais alta em um momento especificado.

O grupo de recursos que possui um cronômetro de fallback atrasado configurado e que reside atualmente em um nó não inicial efetua fallback para o nó de prioridade mais alta no horário recorrente (diariamente, semanalmente, mensalmente ou anualmente) ou em uma data especificada.

Informações relacionadas:

Planejando o PowerHA SystemMirror

Minimizando o tempo de controle

No caso de uma falha do cluster, grupos de volumes simultâneos aprimorados são controlados com mais rapidez do que nas liberações anteriores do PowerHA SystemMirror devido ao mecanismo de controle de disco aprimorado.

No caso de uma falha do cluster, grupos de volumes simultâneos aprimorados são controlados com mais rapidez do que nas liberações anteriores do PowerHA SystemMirror devido ao mecanismo de controle de disco aprimorado. O PowerHA SystemMirror detecta automaticamente os grupos de volumes simultâneos e assegura que a opção mais rápida para o controle do grupo de volume seja ativada no caso de uma falha do nó, desde que os grupos de volumes de modo simultâneo aprimorados tenham sido incluídos (ou que os grupos de volumes existentes tenham sido convertidos) nos grupos de recursos não simultâneos.

Esta funcionalidade é especialmente útil para fallover de grupos de volumes compostos de um grande número de discos.

Durante o controle rápido de disco, o PowerHA SystemMirror ignora o processamento extra necessário para quebrar as reservas de disco ou para atualizar e sincronizar as informações do LVM executando uma atualização lenta. Como resultado, o mecanismo de controle de disco do PowerHA SystemMirror usado para grupos de volumes simultâneos aprimorados é mais rápido que o controle de disco usado para grupos de volumes padrão incluídos em grupos de recursos não simultâneos.

Maximizando a recuperação de desastre

O PowerHA SystemMirror pode ser uma parte integrante de um plano de recuperação de desastres abrangente da empresa.

A seguir há várias maneiras possíveis de distribuir cópias de backup de dados em diferentes sites, para possíveis operações de recuperação de desastre:

- PowerHA SystemMirror Enterprise Edition para Gerenciador de Volume Lógico (GLVM)
- PowerHA SystemMirror Enterprise Edition para Metro Mirror (PPRC síncrono com os sistemas ESS e DS)
- Espelhamento do LVM entre sites

Espelhamento do LVM entre sites

É possível configurar discos que estejam localizados em dois sites diferentes para o espelhamento do LVM remoto, usando uma rede de área de armazenamento (SAN), por exemplo. O espelhamento do LVM entre sites replica dados entre o subsistema de disco em cada site para recuperação de desastre.

Uma SAN é uma rede de alta velocidade que permite o estabelecimento de conexões diretas entre dispositivos de armazenamento e processadores (servidores) dentro da distância suportada pelo Fibre Channel. Assim, dois ou mais servidores (nós) localizados em sites diferentes podem acessar os mesmos discos físicos, que podem estar separados por alguma distância, por meio da SAN comum. Os discos podem ser combinados em um grupo de volumes usando o Gerenciador de Volume Lógico do AIX e esse grupo de volumes pode ser importado para os nós que estão localizados em sites diferentes. Os volumes lógicos contidos neste grupo de volumes podem ter até três espelhos. Assim, é possível configurar pelo menos um espelho em cada site. As informações armazenadas neste volume lógico são mantidas altamente disponíveis e, em caso de certas falhas, o espelho remoto em outro site possuirá as informações mais recentes, portanto, as operações poderão ser continuadas em outro site. O PowerHA SystemMirror sincroniza os espelhos automaticamente após uma falha do disco ou de nó e a reintegração subsequente.

O PowerHA SystemMirror manipula a sincronização de espelho automática mesmo se um dos discos estiver no estado PVREMOVED ou PVMISSING. A sincronização automática não é possível em todos os casos, mas é possível usar o C-SPOC para sincronizar os dados manualmente a partir dos espelhos sobreviventes para os espelhos estagnados após uma falha do disco ou site e a reintegração subsequente.

Eventos de cluster

Esta seção descreve como o software PowerHA SystemMirror responde a mudanças em um cluster para manter a alta disponibilidade.

O software de cluster PowerHA SystemMirror monitora todos os componentes que compõem o aplicativo altamente disponível, incluindo discos, interfaces de rede, nós e os próprios aplicativos. O Gerenciador de cluster usa métodos diferentes para monitorar recursos diferentes:

- O subsistema RSCT é responsável por monitorar redes e nós.
- O subsistema AIX LVM produz notificações de erro para perda de quorum do grupo de volume.
- O próprio Gerenciador do cluster despacha os monitores de aplicativo.

O ambiente em cluster do PowerHA SystemMirror é acionado por evento. Um evento é uma mudança de status dentro de um cluster que o Gerenciador do cluster reconhece e processa. Um evento de cluster pode ser acionado por uma mudança que afete uma placa da interface de rede, uma rede ou um nó, ou se o processo de reconfiguração de cluster exceder seu tempo limite. Quando o Gerenciador do cluster detecta uma mudança no status do cluster, ele executa um script designado para manipular o evento e seus subeventos.

Nota: A lógica de eventos de cluster é descrita aqui para complementação. Nenhuma intervenção manual é necessária para assegurar que o PowerHA SystemMirror transporte os eventos de cluster corretamente.

Os exemplos a seguir mostram alguns eventos que o Gerenciador de cluster reconhece:

- Eventos node_up e node_up_complete (um nó sendo inserido no cluster)
- Eventos node_down e node_down_complete (um nó saindo do cluster)
- Evento network_down local ou global (uma rede falhou)
- Evento network_up (uma rede foi conectada)
- Evento swap_adapter (um adaptador de rede falhou e um novo foi colocado em seu lugar)
- Eventos de reconfiguração dinâmica
- site_up e site_up_complete (um site sendo incluído no cluster)
- site_down e site_down_complete (um site sendo removido do cluster)

Quando ocorre um evento de cluster, o Gerenciador do cluster executa o script de evento correspondente a esse evento. Conforme o script do evento é processado, uma série de scripts de subevento pode ser executada. O software PowerHA SystemMirror fornece um script para cada evento e subevento. Os scripts padrão são localizados no diretório /usr/es/sbin/cluster/events.

Por padrão, o Gerenciador de cluster chama o script de evento correspondente fornecido com o software PowerHA SystemMirror para um evento específico.

É possível especificar informações de processamento adicionais para customizar a manipulação de eventos para o site, se necessário.

Processando eventos de cluster

Os dois eventos de cluster principal que o software PowerHA SystemMirror manipula são o fallover e a reintegração.

- Fallover refere-se às ações executadas pelo software PowerHA SystemMirror quando um componente do cluster falha ou um nó é removido do cluster.
- Reintegração refere-se às ações que ocorrem dentro do cluster quando um componente que foi removido anteriormente do cluster retorna ao cluster.

Scripts de evento controlam ambos os tipos de ações. Durante o processamento do script do evento, programas de aplicativo que reconhecem o cluster entendem o estado do cluster como instável.

Fallover

Um fallover ocorre quando um grupo de recursos é movido de seu nó inicial para outro nó porque seu nó inicial foi removido do cluster.

Os nós são removidos do cluster por uma transição planejada (um encerramento do nó ou uma interrupção dos serviços de cluster em um nó) ou por falha. No caso anterior, o Gerenciador do cluster controla a liberação de recursos mantidos pelo nó existente e a aquisição desses recursos por nós ainda ativos no cluster. Quando necessário, é possível ignorar a liberação e a aquisição de recursos (por exemplo, para executar a manutenção do sistema). Também é possível adiar a aquisição dos recursos, integrando os nós (configurando o cronômetro de fallback atrasado para grupos de recursos customizados).

A falha do nó começa quando um nó que monitora um nó vizinho deixa de receber o tráfego keep-alive durante um período de tempo definido. Se os outros nós do cluster concordarem que a falha é uma falha de nó, o nó com falha será removido do cluster e seus recursos serão controlados pelos nós ativos configurados para fazer isso.

Se outros componentes, como uma placa da interface de rede falhar, o Gerenciador do cluster executará um script de evento para alternar o tráfego de rede para uma placa da interface de rede de backup (se estiver presente).

Reintegração

A reintegração ou um fallback ocorre quando um grupo de recursos é movido para um nó que acabou de ser incluído no cluster. Quando um nó é incluído em um cluster em execução, o cluster torna-se temporariamente instável. Os nós membros coordenam o início do processo de inclusão e, em seguida, executam scripts de evento para liberar todos os recursos que o nó incluído está configurado para assumir. Em seguida, o nó incluído executa um script de evento para controlar esses recursos. Finalmente, o nó torna-se um membro do cluster. Neste ponto, o cluster fica estável novamente.

Customizando o processamento de eventos

O software PowerHA SystemMirror possui um recurso de customização de evento que pode ser usado para customizar o processamento de eventos.

A capacidade do Gerenciador do cluster de reconhecer uma série de eventos e subeventos específicos permite um esquema de customização muito flexível. A customização do processamento de eventos permite fornecer o caminho mais eficiente para os recursos críticos no caso de ocorrência de falha.

É possível definir vários pré-eventos e pós-eventos para compor uma lista de eventos que apareça na lista de seleção no painel do SMIT **Alterar/Mostrar eventos predefinidos do PowerHA SystemMirror** ou por meio do atalho do SMIT `clcsclv.select`. A customização de um evento poderia incluir uma notificação ao administrador do sistema antes e após o processamento do evento, além de comandos ou scripts definidos pelo usuário antes e após o processamento do evento, conforme é mostrado na lista:

- Notificação ao administrador do sistema sobre o evento a ser processado
- Script ou comando de pré-evento
- Script de evento do PowerHA SystemMirror for AIX
- Script ou comando pós-evento
- Notificação ao administrador do sistema de que o processamento do evento foi concluído.

Use este recurso para os seguintes tipos de customização:

- Processamento pré-evento e pós-evento
- Notificação de evento
- Recuperação e nova tentativa de evento.

Nota: No PowerHA SystemMirror, as informações de customização de evento armazenadas no banco de dados de configuração do PowerHA SystemMirror são sincronizadas em todos os nós do cluster quando os recursos do cluster são sincronizados. Assim, a pré-notificação e a pós-notificação e os nomes de script de evento de recuperação devem ser iguais em todos os nós, embora o processamento real executado por esses scripts possa ser diferente.

A verificação de cluster inclui uma função para monitorar a configuração de cluster automaticamente por meio de um novo evento chamado **cluster_notify**. Esse evento pode ser usado para configurar um método de notificação remoto do PowerHA SystemMirror (página numérica ou alfanumérica ou mensagem de texto) para enviar uma mensagem caso sejam localizados erros na configuração do cluster. A saída desse evento também é registrada no **hacmp.out** em cada nó do cluster que esteja executando serviços de cluster.

Também é possível enviar uma notificação por email para telefones celulares por meio de scripts de notificação de evento, no entanto, o uso da notificação remota possui vantagens. Se você é a pessoa responsável por responder às mudanças de notificações de eventos, você deve alterar manualmente o endereço em cada script de notificação de eventos. Defina para cada pessoa, métodos de notificação remotos que contenham todos os eventos e nós, para que seja possível alternar os métodos de notificação como uma unidade quando os respondentes forem alterados.

Definindo novos eventos

No PowerHA SystemMirror, é possível definir novos eventos além de customizar os que já existem.

Processamento pré-evento e pós-evento

Para customizar o processamento de eventos para o ambiente, especifique comandos ou scripts definidos pelo usuário a serem executados antes e depois que um evento específico seja gerado pelo Gerenciador de cluster. Para o pré-processamento, por exemplo, é possível enviar uma mensagem a usuários específicos, informando-os para permanecer em espera enquanto um determinado evento ocorre. Para o pós-processamento, é possível desativar o login para um grupo de usuários específicos, caso uma determinada rede falhe.

Notificação de evento

É possível especificar um comando ou um script definido pelo usuário que forneça uma notificação (por exemplo, um email) de que um evento está prestes a ocorrer e de que um evento acabou de ocorrer, juntamente com a informação de sucesso ou falha do evento. Também é possível definir um método de

notificação por meio da interface SMIT para emitir um método de notificação remoto customizado em resposta a um evento de cluster.

Recuperação e nova tentativa de evento

É possível especificar um comando que tente recuperar a partir de uma falha de comando de evento. Se a contagem de novas tentativas for maior que zero e o comando de recuperação for bem-sucedido, o comando de script de evento será executado novamente. Também é possível especificar o número de vezes para tentar executar o comando de recuperação.

Customizando a duração do evento

O software PowerHA SystemMirror emite um aviso do sistema sempre que um evento de cluster demora mais para ser concluído do que um período de tempo limite especificado.

Usando a interface SMIT, é possível customizar o período de tempo permitido para que um evento de cluster seja concluído antes que o PowerHA SystemMirror emita um aviso do sistema.

Configurações de cluster do PowerHA SystemMirror

Este capítulo fornece exemplos dos tipos de configurações de cluster suportadas pelo software PowerHA SystemMirror.

Esta lista não é, de maneira alguma, um catálogo completo de configurações possíveis que podem ser definidas usando o software PowerHA SystemMirror. Ela deve ser usada como um ponto de início para pensar sobre a configuração de cluster que melhor se adequará ao ambiente.

Configurações de espera

Configurações de espera são as configurações de hardware redundantes tradicionais nas quais um ou mais nós de espera ficam inativos, esperando que um nó de servidor seja removido do cluster.

Os grupos de recursos simultâneos requerem que todos os nós tenham acesso simultâneo ao grupo de recursos e não podem ser usados em uma configuração de espera.

Exemplo: Configuração de espera com política de inicialização somente online no nó inicial

Este exemplo mostra os grupos de recursos com a política de inicialização somente online no nó inicial, fallover para o nó com a próxima prioridade na lista e fallback para o nó de prioridade mais alta na política de fallback na lista.

Na configuração de espera a seguir, os grupos de recursos possuem estas políticas:

- Política de inicialização: Somente online no nó inicial
- Política de fallover: Fallover para o nó com a próxima prioridade na lista
- Política de fallback: Fallback para o nó com prioridade mais alta na lista.

Na figura, um número mais baixo indica uma prioridade mais alta:

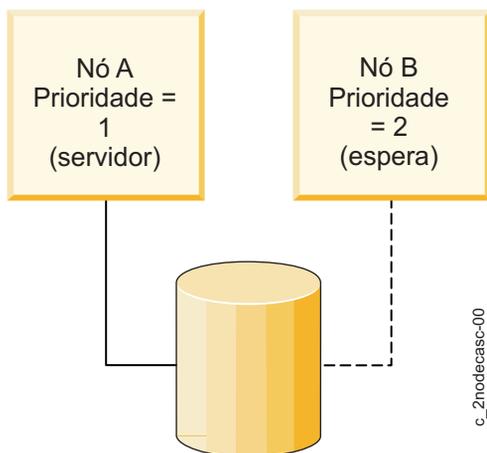


Figura 19. Configuração de espera de um para um, em que o rótulo de IP retorna para o nó inicial

Nesta configuração, os recursos de cluster são definidos como parte de um único grupo de recursos. Em seguida, uma lista de nós é definida consistindo em dois nós. O primeiro nó, o Nó A, recebe a prioridade de controle (propriedade) 1. O segundo nó, o Nó B, recebe a prioridade de controle 2.

Na inicialização do cluster, o Nó A (que possui a prioridade 1) assume a propriedade do grupo de recursos. O Nó A é o nó de "servidor". O Nó B (que possui a prioridade 2) permanece inativo, pronto para quando o Nó A falhar ou for removido do cluster. O Nó B é, na verdade, o "de espera".

Se o nó de servidor for removido do cluster, o nó de espera assumirá o controle do grupo de recursos pertencente ao servidor, iniciará os aplicativos altamente disponíveis e os clientes de serviço. O nó de espera permanecerá ativo até que o nó com a prioridade de controle mais alta seja incluído novamente no cluster. Nesse ponto, o nó de espera liberará os grupos de recursos que assumiu e o nó de servidor os solicitará novamente. Em seguida, o nó de espera retornará para o estado inativo.

Ampliando configurações de espera

A configuração de espera do exemplo descrito anteriormente pode ser facilmente ampliada para clusters maiores. A vantagem desta configuração é que ela faz melhor uso do hardware. A desvantagem é que o cluster pode sofrer degradação grave de desempenho se mais de um nó de servidor for removido do cluster.

A figura a seguir ilustra uma configuração de espera de três nós usando os grupos de recursos com estas políticas:

- Política de inicialização: Somente online no nó inicial
- Política de failover: Failover para o nó com a próxima prioridade na lista
- Política de fallback: Fallback para o nó com prioridade mais alta na lista

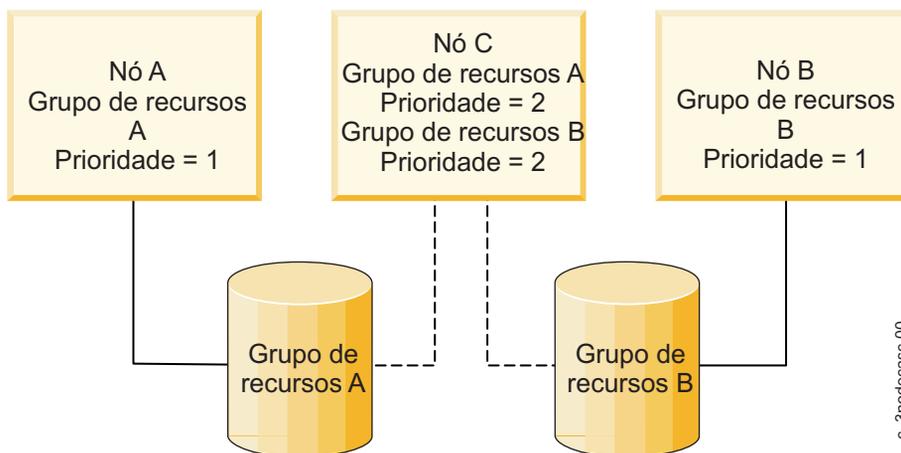


Figura 20. Configuração de espera de um para dois com três grupos de recursos

Nesta configuração, existem dois grupos de recursos separados (A e B) e uma lista de nós separada para cada grupo de recursos. A lista de nós para o Grupo de recursos A consiste de Nó A e Nó C. O Nó A tem uma prioridade de controle de 1, enquanto o Nó C tem a prioridade de controle de 2. A lista de nós para o Grupo de recursos B consiste de Nó B e Nó C. O Nó B tem uma prioridade de controle de 1; o Nó C novamente tem uma prioridade de controle de 2. (Um grupo de recursos pode ser possuído apenas por um único nó em uma configuração não concorrente.)

Como cada grupo de recursos possui um nó diferente em primeiro lugar em sua lista de nós, a carga de trabalho do cluster é dividida, ou particionada, entre esses dois grupos de recursos. Ambos os grupos de recursos, no entanto, possuem o mesmo nó como o de espera em suas listas de espera. Se um nó de servidor for removido do cluster, o nó de espera assumirá o controle desse grupo de recursos do nó de servidor e funcionará como o nó removido.

Neste exemplo, o nó de espera possui três interfaces de rede (não mostradas) e conexões físicas separadas para cada disco externo de nó de servidor. Portanto, o nó de espera pode, se necessário, controlar ambos os nós de servidor simultaneamente. No entanto, o desempenho do cluster provavelmente será comprometido enquanto o nó de espera estiver funcionando como ambos os nós de servidor.

Exemplo: Configurações de espera com inicialização online usando política de distribuição

Este exemplo explica as políticas de grupo de recursos que você deve usar com configurações de espera.

Na configuração de espera a seguir, os grupos de recursos possuem estas políticas:

- Política de inicialização: Online usando política de distribuição (baseada em rede ou baseada em nó)
- Política de failover: Nó com a próxima prioridade na lista
- Política de fallback: Nunca efetuar fallback.

Esta configuração difere de uma configuração de espera na qual a prioridade de propriedade de grupos de recursos não é fixa. Nesse caso, o grupo de recursos é associado a um endereço IP que pode girar entre os nós. Isso torna as funções de servidor e espera fluidas, alterando ao longo do tempo.

A figura a seguir ilustra a configuração de espera de um para um, descrita nesta seção:

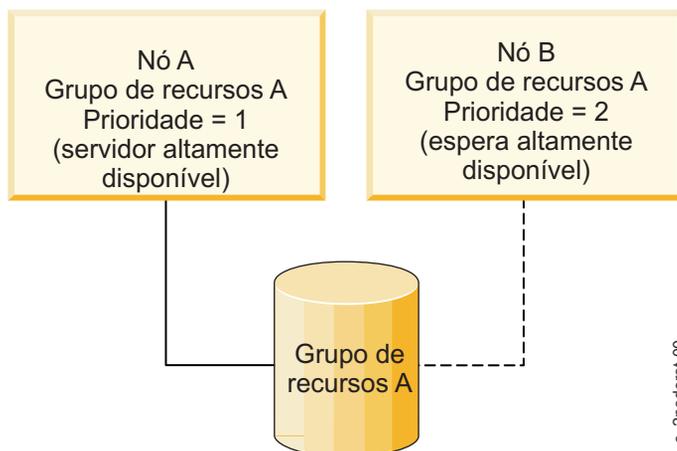


Figura 21. Configuração de espera de um para um com grupos de recursos em que o rótulo de IP gira

Na inicialização do sistema, o grupo de recurso conecta-se ao nó que solicita o endereço IP compartilhado. Esse nó possuirá o grupo de recursos durante todo o período em que ele permanecer no cluster. Se esse nó for removido do cluster, o nó peer assumirá o endereço IP compartilhado e solicitará a propriedade desse grupo de recursos. Agora, o nó peer "possuirá" o grupo de recursos durante todo o período em que ele permanecer no cluster.

Quando o nó que inicialmente solicitou o grupo de recursos for incluído no cluster novamente, ele não tomará o grupo de recursos de volta. Ele permanecerá inativo enquanto o nó ligado atualmente ao endereço IP compartilhado estiver ativo no cluster. Somente se o nó peer for removido do cluster, o nó que inicialmente "possuía" o grupo de recursos o solicitará novamente. Assim, a propriedade de recursos gira entre os nós.

Ampliando configurações de espera

Como com o primeiro exemplo de configuração de espera, as configurações deste exemplo podem ser facilmente ampliadas para clusters maiores. Por exemplo, nesta configuração de espera de um para dois do exemplo, o cluster poderia ter dois grupos de recursos separados, cada uma incluindo um endereço IP compartilhado diferente.

Na inicialização do cluster, os primeiros dois nós solicitam um endereço IP compartilhado e assumem a propriedade do grupo de recursos associado a esse endereço IP compartilhado. O terceiro nó permanece inativo. Se um nó ativo é removido do cluster, o nó inativo solicita esse endereço IP compartilhado e assume o controle desse grupo de recursos.

Configurações de controle

Nas configurações de controle, todos os nós do cluster executam algum trabalho útil, processando parte da carga de trabalho do cluster. Não há nós de espera. As configurações de controle usam recursos de hardware com mais eficiência do que as configurações de espera, pois não há nenhum processador inativo. No entanto, o desempenho pode ser comprometido após a desconexão do nó, pois o carregamento nos nós restantes aumenta.

Controle unilateral

Essa configuração possui dois nós processando trabalho ativamente, mas somente um nó fornecendo serviços altamente disponíveis aos clientes do cluster. Ou seja, embora haja dois conjuntos de recursos no cluster (por exemplo, dois aplicativos do servidor que manipulam solicitações do cliente), somente um conjunto de recursos precisa ser altamente disponível.

A figura a seguir ilustra uma configuração de controle unilateral de dois nós. Na figura, um número mais baixo indica uma prioridade mais alta.

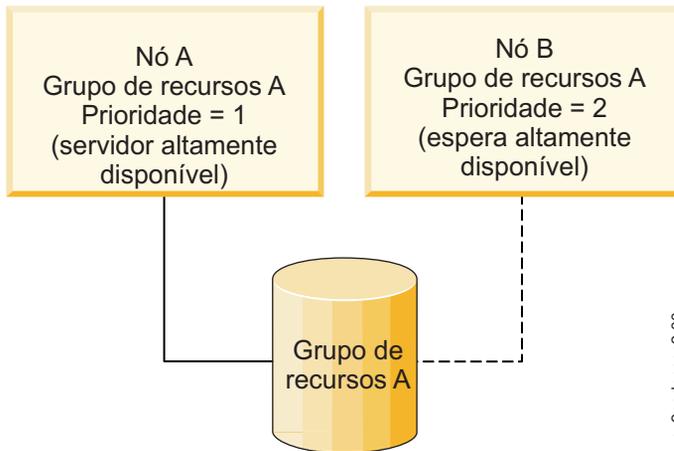


Figura 22. Configuração de controle unilateral com grupos de recursos nos quais o rótulo de IP retorna ao nó inicial

Este conjunto de recursos é definido como um grupo de recursos do PowerHA SystemMirror e possui uma lista de nós que inclui ambos os nós. O segundo conjunto de recursos não é definido como um grupo de recursos e, portanto, *não* é altamente disponível.

Na inicialização do cluster, o Nó A (que possui a prioridade 1) assume a propriedade do Grupo de recursos A. O Nó A, em vigor "possui" o Grupo de recursos A. O Nó B (que possui a prioridade 2 para o Grupo de recursos A) processa sua própria carga de trabalho, independentemente desse grupo de recursos.

Se o Nó A sair do cluster, o Nó B assumirá o controle dos recursos compartilhados. Quando o Nó A retorna ao cluster, o Nó B libera os recursos compartilhados.

No entanto, se o Nó B sair do cluster, o Nó A não assumirá nenhum de seus recursos, pois os recursos do Nó B não são definidos como parte de um grupo de recursos altamente disponíveis de cuja cadeia esse nó faz parte.

Essa configuração é apropriada quando um único nó pode executar todos os aplicativos críticos que precisam ser altamente disponíveis aos clientes do cluster.

Controle mútuo

O controle mútuo para configuração de acesso não simultâneo possui vários nós e cada um fornece serviços distintos altamente disponíveis para clientes do cluster. Por exemplo, cada nó pode executar sua própria instância de um banco de dados e acessar seu próprio disco.

Além disso, cada nó possui capacidade de controle. Se um nó for removido do cluster, um nó sobrevivente controlará os grupos de recursos que pertenciam ao nó removido.

O controle mútuo para configuração de acesso não simultâneo é apropriado quando cada nó no cluster está executando aplicativos críticos que precisam estar altamente disponíveis e quando cada processador é capaz de manipular o carregamento de mais de um nó.

A figura a seguir ilustra uma configuração de controle mútuo de dois nós para acesso não simultâneo. Na figura, um número mais baixo indica uma prioridade mais alta.

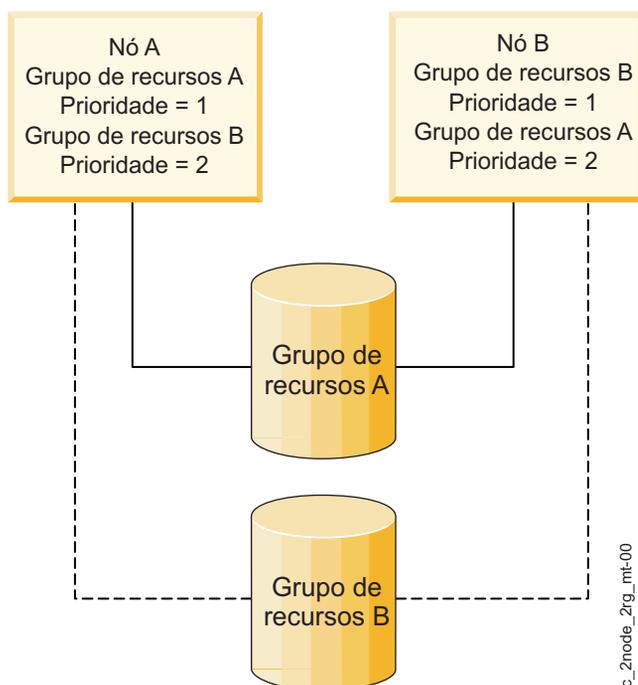


Figura 23. Configuração de controle mútuo para acesso não simultâneo

O recurso-chave dessa configuração é que a carga de trabalho do cluster é dividida, ou particionada, entre os nós. Existem dois grupos de recursos, além de uma cadeia de recursos separada para cada grupo de recursos. Os nós que participam das cadeias de recursos são os mesmos. É a diferença de prioridades dentro das cadeias que designa essa configuração como controle mútuo.

As cadeias para ambos os grupos de recursos consistem de Nó A e Nó B. Para o Grupo de recursos A, o Nó A tem uma prioridade de controle 1 e Nó B tem uma prioridade de controle de 2. Para o Grupo de recursos B, as prioridades de controle são invertidas. Aqui, o Nó B tem a prioridade de controle 1 e o Nó A tem a prioridade de controle 2.

Na inicialização do cluster, o Nó A assume propriedade do Grupo de recursos A, enquanto o Nó B assume propriedade do Grupo de recursos B.

Se algum dos nós for removido do cluster, seu nó peer assumirá o controle do grupo de recursos do nó removido. Quando o nó "proprietário" desse grupo de recursos for incluído no cluster novamente, o nó de controle liberará os recursos associados e eles serão readquiridos pelo nó de alta prioridade que foi incluído novamente.

Configuração de controle mútuo de dois nós

Nesta configuração, ambos os nós possuem acesso simultâneo aos discos compartilhados e possuem os mesmos recursos de disco. Não haverá nenhum controle de discos compartilhados se um nó sair do cluster, porque o nó peer já possui o grupo de volumes compartilhados ativado.

A figura a seguir ilustra uma configuração de controle mútuo de dois nós para acesso simultâneo:

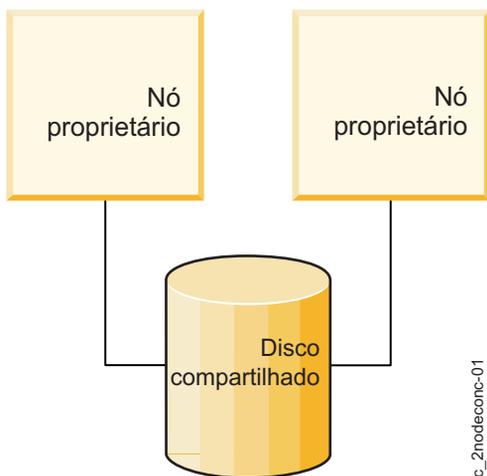


Figura 24. Configuração de controle mútuo de dois nós para acesso simultâneo

Neste exemplo, ambos os nós estão executando uma instância de um aplicativo do servidor que acessa o banco de dados no disco compartilhado. O modelo de bloqueio do proprietário do aplicativo é usado para arbitrar solicitações de aplicativo para recursos de disco.

A execução de várias instâncias do mesmo aplicativo do servidor permite que o cluster distribua a carga de processamento. Conforme a carga aumenta, nós adicionais podem ser incluídos para distribuir mais a carga.

Configuração de controle mútuo de oito nós

Nesta configuração, todos os nós possuem acesso simultâneo aos discos compartilhados e possuem os mesmos recursos de disco.

A figura a seguir ilustra uma configuração de controle mútuo de oito nós para acesso simultâneo:

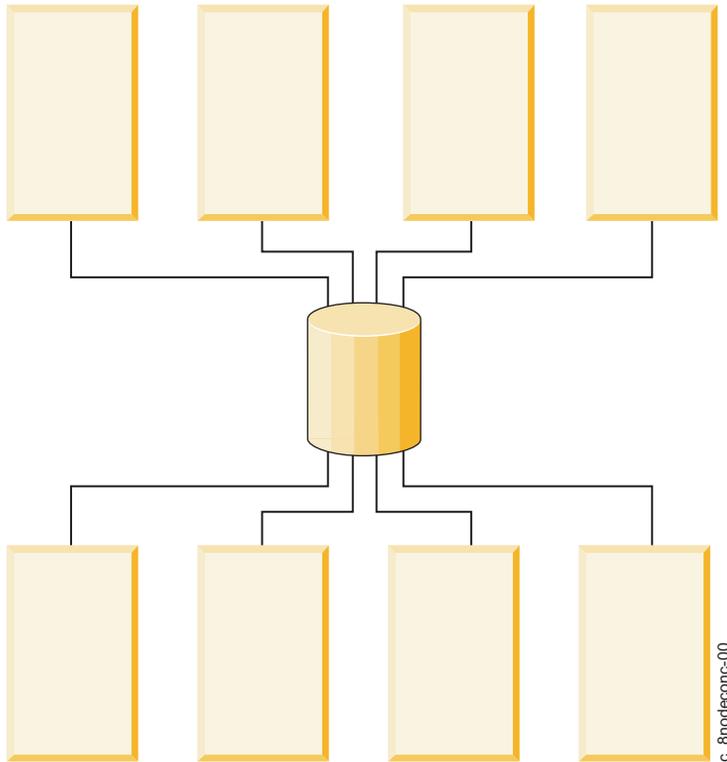


Figura 25. Controle mútuo de oito nós

Cada nó está executando um aplicativo do servidor diferente. Os clientes consultam um aplicativo específico em um endereço IP específico. Portanto, cada controlador de aplicativo e seu endereço IP associado deve ser definido como parte de um grupo de recursos não simultâneo e todos os nós que sejam proprietários potenciais desse grupo de recursos devem ser incluídos em uma lista de nós correspondentes.

Os grupos de recursos de acesso simultâneo são suportados em clusters com até 16 nós no PowerHA SystemMirror.

Configurações de cluster com aplicativos com multicamadas

Uma configuração de cluster típica que pode usar grupos de recursos com dependência de pai e filho é o ambiente no qual um aplicativo como o WebSphere depende de outro aplicativo como o DB2.

Para satisfazer as necessidades de negócios, uma dependência de pai e filho em todo o cluster deve ser definida entre dois ou mais grupos de recursos.

A figura a seguir ilustra o cenário de negócios que usa dependências entre aplicativos:

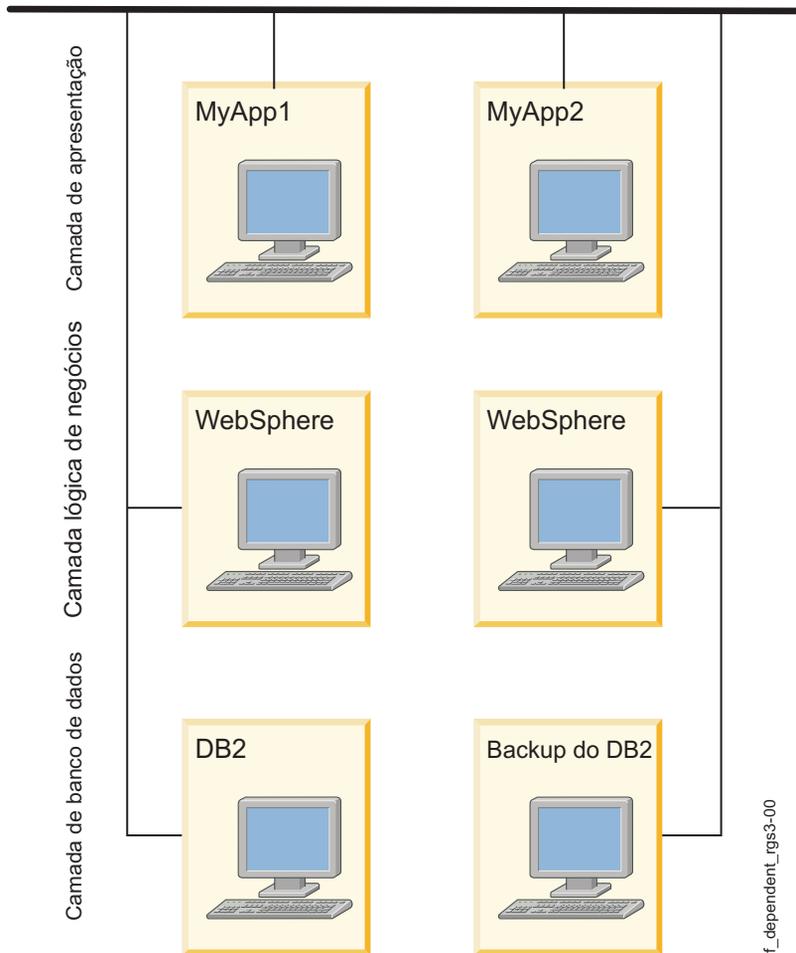


Figura 26. Ambiente em cluster com multicamadas típico com dependências entre aplicativos

Aplicativos com multicamadas

As configurações de negócios que usam aplicativos em camadas ou com multicamadas também podem usar grupos de recursos dependentes. Por exemplo, o banco de dados de backend deve ficar online antes do controlador de aplicativo. Nesse caso, se o banco de dados ficar inativo e for movido para um nó diferente, o grupo de recursos que contém o controlador de aplicativo deverá ser desativado e ativado novamente em qualquer nó no cluster.

Ambientes como o SAP requerem que os aplicativos sejam reciclados (parados e reiniciados) sempre que um banco de dados falha. Um ambiente como o SAP fornece muitos serviços de aplicativos e os componentes de aplicativo individuais geralmente precisam ser controlados em uma ordem específica.

Outra área em que é útil estabelecer interdependências entre os grupos de recursos é quando os serviços do sistema são necessários para suportar ambientes de aplicativos. Serviços como as tarefas **cron** para remover arquivos de log ou iniciar backups precisam ser movidos de nó a nó juntamente com os aplicativos, mas, geralmente, eles não são iniciados até que o aplicativo seja estabelecido. Esses serviços podem ser construídos nos scripts de inicialização e de interrupção do controlador de aplicativo. Quando uma maior granularidade for necessária, eles poderão ser controlados por meio de pré e pós-processamento de eventos. Os grupos de recursos com dependência de pai/filho possibilitam uma maneira mais fácil de configurar os serviços do sistema como dependentes do aplicativo que eles atendem.

Conceitos relacionados:

“Dependências do grupo de recursos” na página 40

O PowerHA SystemMirror suporta a ordenação de grupo de recursos e o processamento serial customizado de recursos para acomodar as configurações do cluster nas quais existe uma dependência entre os aplicativos que residem em grupos de recursos diferentes.

Configurações de cluster com dependências de local de grupo de recursos

É possível configurar o cluster para que determinados aplicativos fiquem no mesmo nó ou em nós diferentes, não somente na inicialização, mas durante eventos de fallover e fallback. Para fazer isso, configure os grupos de recursos selecionados como parte de um conjunto de dependência de local.

Modelo de publicação com dependências do mesmo nó e de nós diferentes

A empresa fictícia XYZ Publishing segue um modelo de continuidade de negócios que envolve a separação de diferentes plataformas usadas para desenvolver o conteúdo da web. A XYZ Publishing usa políticas de dependência de local para manter alguns grupos de recursos estritamente em nós separados e outros juntos no mesmo nó.

O Banco de dados de produção (PDB) e o Aplicativo de produção (PApp) são hospedados no mesmo nó para facilitar a manutenção (e talvez o nó de prioridade mais alta desses grupos de recursos tenha o processador com mais memória ou mais rápido). Também faz sentido criar uma relação pai/filho entre eles, pois o aplicativo depende do banco de dados. O banco de dados deve estar online para que o aplicativo funcione. As mesmas condições são verdadeiras para o Banco de dados de sistema (SDB) e o Aplicativo de sistema (Sapp) e para o Banco de dados de QA (QADB) e o Aplicativo de QA (QAapp).

Como manter o banco de dados de produção e o aplicativo e em execução é a prioridade mais alta, faz sentido configurar o cluster para que os três grupos de recursos do banco de dados fiquem em nós diferentes (torná-los conjuntos de dependência Online em nós diferentes) e designar o grupo de recursos do PDB com a prioridade mais alta. O SDB é a prioridade intermediária e o QADB é a prioridade baixa.

Os bancos de dados e seus aplicativos relacionados são configurados para pertencer a um conjunto de dependência Online no mesmo nó.

O PowerHA SystemMirror manipula esses grupos de forma um pouco diferente, dependendo de como as políticas de inicialização, fallover e fallback estão configuradas. Faz sentido diferenciar as listas de nós participantes para cada conjunto de banco de dados e aplicativo para facilitar a manutenção desses grupos de recursos nos nós preferenciais.

A figura abaixo mostra a configuração básica dos três nós e seis grupos de recursos.

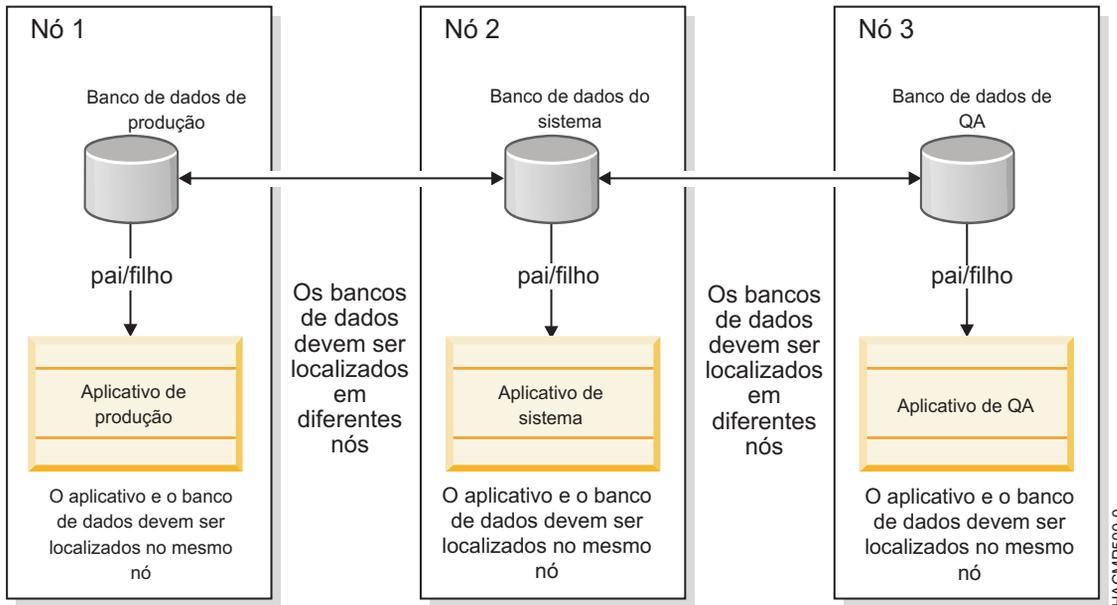


Figura 27. Modelo de publicação com dependência de pai ou dependências de filho e dependências de local

Políticas de grupo de recursos

Para propósito da ilustração deste caso, todos os seis grupos de recursos podem ter as seguintes políticas de comportamento:

- Política de inicialização: Online no primeiro nó disponível
- Política de failover: Failover para o nó com a próxima prioridade
- Política de fallback: Nunca efetuar fallback

Nós participantes	Dependência de local	Dependência de pai/filho
<ul style="list-style-type: none"> • PApp: 1, 2, 3 • PDB: 1, 2, 3 • SApp: 2, 3 • SDB: 2, 3 • QAApp: 3 • QADB: 3 	<p>Grupos de dependência Online no mesmo nó:</p> <ul style="list-style-type: none"> • PApp com PDB • SApp com SDB • QAApp com QADB <p>Conjunto de dependência Online em nós diferentes: [PDB SDB QADB] Prioridade: PDB > SDB > QADB</p>	<ul style="list-style-type: none"> • PApp (filho) depende do PDB (pai) • SApp (filho) depende do SDB (pai) • QAApp (filho) depende do QADB (pai)

Informações relacionadas:

Comportamento do grupo de recursos durante eventos de cluster

Configurações de espelho do LVM entre sites para recuperação de desastre

É possível configurar discos localizados em dois sites diferentes para o espelhamento do LVM remoto, usando uma rede de área de armazenamento (SAN).

Uma SAN é uma rede de alta velocidade que permite o estabelecimento de conexões diretas entre dispositivos de armazenamento e processadores (servidores) dentro da distancia suportada pelo Fibre Channel. Assim, dois ou mais servidores (nós) separados por uma distancia significativa, que estejam localizados em sites diferentes podem acessar os mesmos discos físicos, que podem estar separados por uma distancia significativa usando a SAN comum. Esses discos remotos podem ser combinados em grupos de volumes, usando o C-SPOC.

Os volumes lógicos em um grupo de volumes pode ter até três espelhos ou cópias, por exemplo, um espelho em cada site. Assim, as informações armazenadas nesse volume lógico podem ser mantidas altamente disponíveis no caso de falha. Por exemplo, se todos os nós em um site (incluindo o subsistema de disco nesse site) tiverem alguma falha, o espelho remoto em outro site ainda terá as informações mais recentes e as operações poderão continuar nesse site.

A principal intenção desse recurso é suportar clusters de dois sites, nos quais o espelhamento do LVM por meio de uma SAN replica os dados entre o subsistema de disco em cada site para recuperação de desastre.

Outra vantagem do espelhamento do LVM entre sites é que após uma falha do site/disco e a reintegração do site subsequente, o PowerHA SystemMirror tentará sincronizar automaticamente os dados dos discos sobreviventes com os discos sendo inseridos. A sincronização ocorre no plano de fundo e não causa impacto significativo no tempo de reintegração.

A figura a seguir ilustra uma configuração de espelhamento do LVM entre sites com uma SAN:

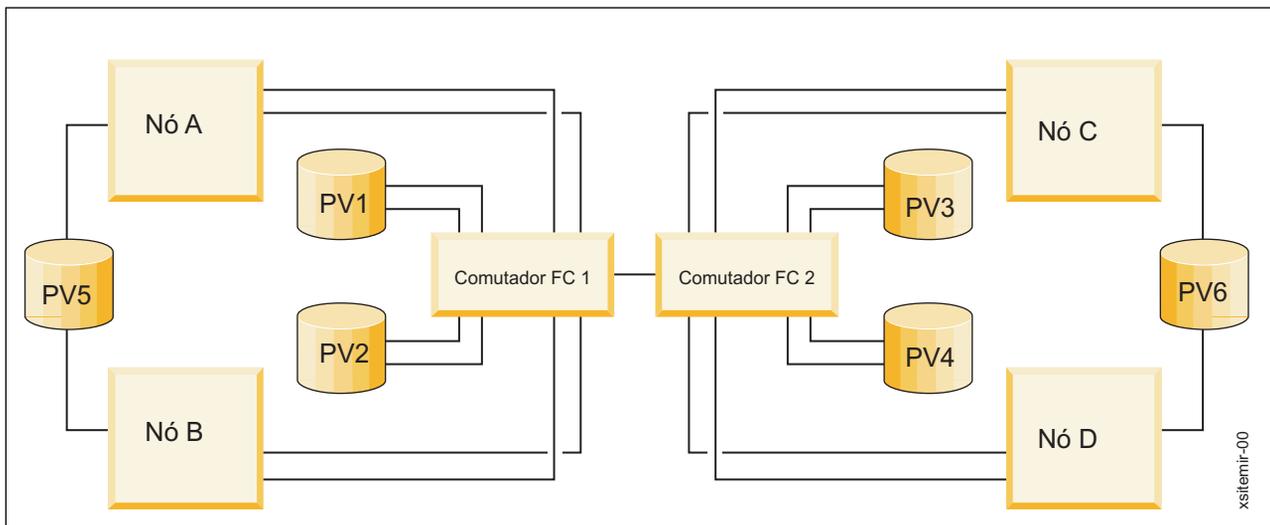


Figura 28. Configuração de espelhamento do LVM entre sites com uma SAN

Os discos que estão conectados com pelo menos um nó em cada um dos dois sites podem ser espelhados. Neste exemplo, PV4 é visto pelo Nós A e B no Site 1 usando o Comutador de Fibre Channel 1 que o Comutador Fibre Channel 2 e também é visto no Nó C usando o Comutador de Fibre Channel 2. Poderia haver um espelhamento do PV4 no Site 1. O discos que estão conectados aos nós somente em um site (PV5 e PV6) não podem ser espelhados entre sites.

As informações do disco são replicadas a partir de um site local para um site remoto. A velocidade dessa transferência de dados depende das características físicas do canal, da distância e do desempenho do espelhamento do LVM.

Configurações de cluster com dynamic LPARs

Os recursos de particionamento avançados do AIX fornecem a capacidade de alocar dinamicamente recursos de CPU, memória e slot de E/S do sistema (*dynamic LPAR*).

O uso do PowerHA SystemMirror combinado com LPARs permite:

`<Nome do adaptador físico>_Large_Send`

Sinalização para indicar a opção de grande envio para este adaptador

<Nome de adaptador físico>_Large_receive

Sinalização para indicar a opção de grande recebimento para este adaptador. Observação: quando configurada e se o adaptador real suportar, os pacotes recebidos pelo adaptador real são agregados antes de serem transmitidos para a camada superior, resultando em melhor desempenho.

<Nome do adaptador físico>_DMA_Errors

O número de pacotes recebidos eliminados pelo hardware devido ao erro Sem recursos.

Observação: esse erro geralmente ocorre porque os buffers de recebimento no adaptador foram esgotados. Alguns adaptadores podem ter o tamanho dos buffers de recebimento como um parâmetro configurável.

<Nome de adaptador físico>_Speed

Indica o atributo de velocidade de mídia deste adaptador

- Executar upgrades de rotina do sistema por meio da alocação dinâmica de recursos do sistema. Quando usado com dynamic LPARs, o PowerHA SystemMirror pode reduzir a quantidade de tempo de inatividade para upgrades bem planejados do sistema, automatizando a transição da carga de trabalho do aplicativo de uma partição lógica para outra, para que a primeira partição lógica possa ser atualizada sem risco para o aplicativo.
- Redistribuir com eficiência recursos de CPU e memória para gerenciar a carga de trabalho. A combinação do PowerHA SystemMirror com uma dynamic LPAR permite usar scripts de inicialização e interrupção de aplicativo customizados para redistribuir dinamicamente os recursos de CPU e memória para partições lógicas que estejam executando atualmente a carga de trabalho do aplicativo, para suportar melhor a transição do aplicativo dentro de um único quadro. Dessa forma, são mantidas a potência de processamento e os recursos necessários para suportar os aplicativos, enquanto recursos mínimos são dedicados ao upgrade, uma tarefa que precisa de menos recursos.

Nota: Não configure todos os nós do cluster como LPARs dentro do mesmo servidor físico. Essa configuração poderia ser, potencialmente, um ponto único de falha significativo. O exemplo a seguir ilustra uma configuração de cluster que usa três LPARs:

- A LPAR 1 está executando um banco de dados de backend (DB2 UDB)
- A LPAR 2 está executando um WebSphere Application Server (WAS)
- A LPAR 3 está sendo executada como um backup (espera) para ambas as LPARs, do DB2 e do WAS. Esta LPAR somente contém recursos mínimos de CPU e memória.

Quando chegar a hora de mover o aplicativo DB2 ou WAS para a terceira LPAR (devido a um upgrade planejado ou a uma falha de recurso nessas LPARs, por exemplo), será possível usar os scripts de inicialização e de interrupção de aplicativo customizados no PowerHA SystemMirror para automatizar a realocação dinâmica de CPU e memória da LPAR primária para a LPAR de espera. Essa operação permite que a LPAR de terceiro adquira os recursos de CPU e memória necessários para atender às necessidades de desempenho dos negócios. Quando o PowerHA SystemMirror move o grupo de recursos que contém o aplicativo de volta para sua LPAR inicial, os recursos de CPU e memória são movidos automaticamente com ele. Observe: Em geral, as dynamic LPARs permitem alocação dinâmica de CPU, memória e recursos de slot de E/S. O PowerHA SystemMirror e os recursos de slot de E/S de dynamic LPAR não são compatíveis (embora seja possível alocar dinamicamente os recursos do slot de E/S para fora do cluster do PowerHA SystemMirror).

A figura a seguir ilustra esse ambiente em cluster:

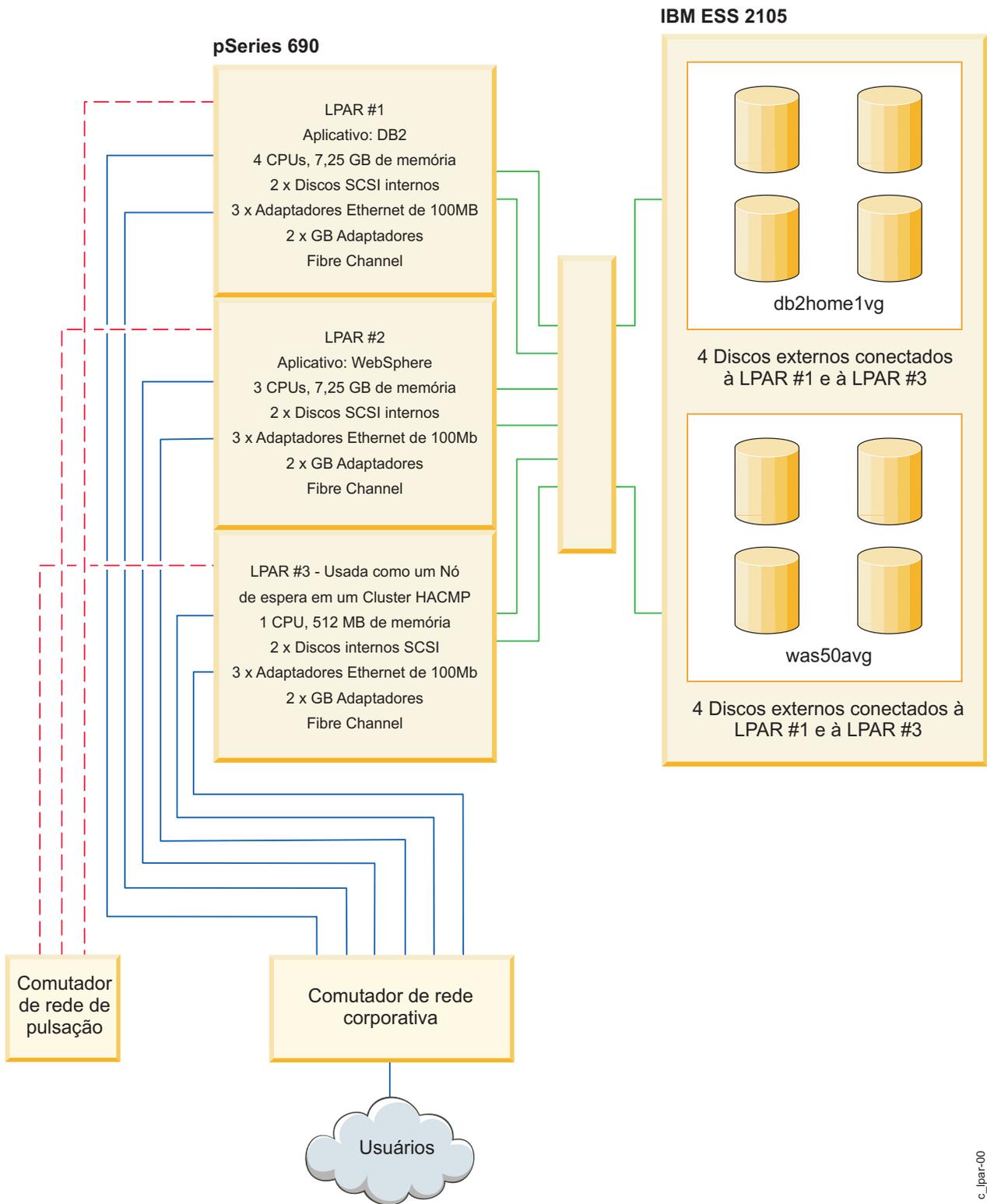


Figura 29. Cluster com três LPARs

c_lpar-00

| **DLPARs e Capacity on Demand**

| PowerHA SystemMirror pode mover recursos de aplicativos entre LPARs e pode executar os ajustes de recurso dinâmico necessários por meio da função Resource Optimized High Availability (ROHA). O ROHA usa os recursos que estão disponíveis com os Power Systems para gerenciar dinamicamente os tipos a seguir de recursos de hardware:

- | • As funções de Capacidade on Demand (CoD) (incluindo CoD On/Off e Enterprise Pool CoD) gerenciam recursos de memória e CPU no nível de estrutura (CEC).
- | • As funções da DLPAR gerenciam recursos de memória e CPU no nível de partição lógica.

| A função de CoD pode ativar recursos pré-instalados (CPU e memória) que estão inativos. O ROHA usa a função do CoD para ativar esses recursos quando os requisitos de recurso para seu ambiente mudam. Por exemplo, durante um processo de controle, seu ambiente requer recursos extras na estrutura de espera. Neste exemplo, os recursos adicionais são dinamicamente fornecidos e não há necessidade de incluir recursos de hardware adicionais na estrutura de espera.

| O nó ativo é hospedado por uma LPAR em uma estrutura com recursos permanentes suficientes. O nó de espera é hospedado por uma LPAR em uma estrutura com recursos mínimos permanentes e depende do ROHA para incluir dinamicamente os recursos extras. É possível usar a função do ROHA para adquirir de forma rápida e fácil recursos extras para atender ao peek ou cargas de trabalho inesperadas em seu ambiente.

Processo e recursos de configuração do PowerHA SystemMirror

Estes tópicos fornecem uma visão geral do processo de configuração de cluster do PowerHA SystemMirror e as ferramentas administrativas fornecidas com o software PowerHA SystemMirror.

Informações que você fornece ao PowerHA SystemMirror

Antes de configurar um cluster, assegure-se de que os blocos de construção estejam planejados e configurados e que o caminho de comunicação inicial exista para o PowerHA SystemMirror para atingir cada nó. Esta seção cobre as tarefas básicas que devem ser executadas para configurar um cluster.

Informações sobre a configuração física de um cluster

A configuração física de um cluster consiste nas várias tarefas de planejamento e configuração.

Essas tarefas incluem:

- Assegurar o suporte de rede TCP/IP para o cluster.
- Configurar os dispositivos de disco compartilhados para o cluster.
- Configurar os grupos de volumes compartilhados para o cluster.
- Considerar os aplicativos essenciais para os quais o PowerHA SystemMirror está sendo usado. Além disso, considere o controlador de aplicativo e qual tipo de gerenciamento de grupo de recursos é melhor para cada aplicativo.
- Examinar os problemas relativos aos clientes do PowerHA SystemMirror.
- Assegurar a redundância física usando vários circuitos ou fontes de alimentação que não possam ser interrompidos, placas da interface de rede físicas redundantes, várias redes para conectar nós e espelhamento de disco.

Informações relacionadas:

Planejando o PowerHA SystemMirror

Informações de configuração do AIX

Os componentes do cluster devem ser configurados corretamente no nível do AIX.

Para essa tarefa, assegure-se de que:

- A comunicação básica com os nós do cluster exista.
- Os grupos de volume, os volumes lógicos, o espelhamento e os sistemas de arquivo estejam configurados e definidos. Para assegurar a redundância lógica, considere diferentes os tipos de grupos de recursos e planeje como os recursos serão agrupados nos grupos de recursos.

Informações relacionadas:

Instalando o PowerHA SystemMirror

Informações descobertas pelo PowerHA SystemMirror

É possível definir os componentes básicos do cluster em apenas algumas etapas. Para ajudar na configuração do cluster, o PowerHA SystemMirror pode recuperar automaticamente as informações necessárias para a configuração a partir de cada nó.

Nota: Para realizar uma configuração do cluster mais fácil e rápida, também é possível usar um assistente de configuração de cluster. Para obter mais informações, consulte o Assistente de configuração do cluster de dois nós.

Para que o processo de descoberta automática funcione, as condições a seguir devem ser atendidas no PowerHA SystemMirror:

- Você configurou anteriormente os componentes físicos e realizou todas as configurações necessárias do AIX.
- Existem caminhos de comunicação ativos para cada nó. Essas informações serão usadas para configurar automaticamente a topologia de TCP/IP do cluster quando o caminho de configuração padrão for usado.

Depois que essas tarefas forem executadas, o PowerHA SystemMirror descobrirá automaticamente os componentes físicos predefinidos dentro do cluster e selecionará os comportamentos padrão. Além disso, o PowerHA SystemMirror executará a descoberta de informações do cluster caso haja mudanças durante o processo de configuração.

A execução da descoberta recupera as informações de configuração atuais do AIX a partir de todos os nós do cluster. Essas informações aparecem nas listas de seleção para ajudar a fazer seleções precisas de componentes existentes.

O processo de descoberta automático do PowerHA SystemMirror é fácil, rápido e não coloca uma carga "de espera" sobre o administrador do cluster.

Opções de configuração de cluster: Padrão e estendido

Nesta seção, o processo de configuração é significativamente simplificado. Enquanto os detalhes do processo de configuração são abordados no *Guia de administração*, esta seção fornece uma breve visão geral de duas formas de configurar um cluster do PowerHA SystemMirror.

Informações relacionadas:

Administrando o PowerHA SystemMirror

Configurando um cluster do PowerHA SystemMirror usando o caminho de configuração padrão

É possível incluir os componentes básicos de um cluster no banco de dados de configuração do PowerHA SystemMirror em algumas etapas. O caminho de configuração de cluster padrão simplifica e acelera o processo de configuração, pois o PowerHA SystemMirror ativa automaticamente a descoberta para coletar as informações e para selecionar comportamentos padrão.

Se você usar este caminho:

- A descoberta automática de informações de cluster será executada por padrão. Antes de iniciar o processo de configuração do PowerHA SystemMirror, será necessário configurar as interfaces ou os

dispositivos de rede no AIX. No PowerHA SystemMirror, estabeleça caminhos de comunicação iniciais para outros nós. Assim que isso for feito, o PowerHA SystemMirror coletará essas informações e configurará automaticamente os nós do cluster e as redes com base na conectividade física. Todas as redes descobertas serão incluídas na configuração do cluster.

- O aliasing de IP é usado como o mecanismo padrão para ligar os rótulos ou endereços IP com as interfaces de rede.
- É possível configurar os tipos mais comuns de recursos. No entanto, a customização do comportamento de fallback e failover de grupo de recursos é limitada.

Configurando um cluster do PowerHA SystemMirror usando o caminho de configuração estendido

Para configurar os elementos de cluster menos comuns, ou se a conectividade com cada um dos nós do cluster não estiver estabelecida, será possível inserir manualmente as informações de forma semelhante às liberações anteriores do software PowerHA SystemMirror.

Ao usar os caminhos do SMIT de configuração estendida do PowerHA SystemMirror, se houver componentes em nós remotos, você deverá iniciar manualmente a descoberta de informações do cluster. Ou seja, a descoberta é opcional (em vez de automática, como é ao usar o caminho do SMIT de configuração padrão do PowerHA SystemMirror).

Usando as opções no menu de configuração estendido, é possível incluir os componentes básicos de um cluster no banco de dados de configuração do PowerHA SystemMirror, além de vários tipos de recursos adicionais. Use o caminho de configuração estendido para customizar o cluster para todos os componentes, políticas e opções que não estão incluídos nos menus de configuração padrão.

Segurança do cluster

Toda a comunicação entre os nós é enviada por meio do daemon de Comunicações do cluster, o **clcomd**, que é executado em cada nó.

O daemon **clcomd** gerencia a autenticação de conexão entre os nós e toda a autenticação ou criptografia de mensagem configurada. O daemon do PowerHA SystemMirror Cluster Communications usa o arquivo **/etc/cluster/rhosts** confiável e remove a reliance em um arquivo **/.rhosts**. O daemon fornece suporte para autenticação e criptografia de mensagens.

Ferramentas de instalação, configuração e gerenciamento

O PowerHA SystemMirror inclui as ferramentas descritas nestas seções para instalar, configurar e gerenciar clusters.

Smart Assists para integrar aplicativos específicos com o PowerHA SystemMirror

O Smart Assist para um determinado aplicativo examina a configuração no sistema para determinar os recursos que o PowerHA SystemMirror deve monitorar (Rótulo de IP de serviço, grupos de volumes). Em seguida, o Smart Assist configura um ou mais grupos de recursos para tornar os aplicativos e seus recursos altamente disponíveis.

O Smart Assist executa as seguintes ações:

- Descobre a instalação do aplicativo e, se necessário, os recursos atualmente configurados, como endereço IP do serviço, sistemas de arquivos e grupos de volumes
- Fornece uma interface SMIT para obter ou alterar as informações de configuração do usuário, incluindo um novo endereço IP de serviço
- Define o aplicativo para o PowerHA SystemMirror e fornece scripts de inicialização e de interrupção customizados para ele
- Fornece um monitor de aplicativos para o aplicativo
- Configura um grupo de recursos a ser contido.

- Configura dependências de local e de tempo do grupo de recursos, caso a solução do aplicativo requeira
- Especifica os arquivos que precisam ser sincronizados usando o recurso Coleção de arquivo do PowerHA SystemMirror
- Modifica os aplicativos configurados anteriormente, conforme necessário
- Verifica a configuração
- Testa a configuração de cluster do aplicativo.

Aplicativos suportados

O PowerHA SystemMirror fornece o Smart Assist para os seguintes aplicativos e modelos de configuração:

- DB2
- DB2 - Espera a quente
- DB2 - Controle mútuo
- WebSphere 6.0
- WebSphere Application Server 6.0
- Recuperação de log de transação de cluster do WebSphere
- Deployment Manager
- Tivoli Directory Server
- IBM HTTP Server
- Oracle 10G

Smart Assist de aplicativo geral

O Smart Assist de aplicativo geral ajuda os usuários a configurarem aplicativos instalados que não possuem seu próprio Smart Assist.

O usuário fornece algumas informações básicas como:

- Nó primário - por padrão, o nó local
- Nós de controle - por padrão, todos os nós configurados com exceção do nó local
- Nome do aplicativo
- Script de início do aplicativo
- Script de parada do aplicativo
- Rótulo de IP de serviço

Em seguida, o Smart Assist geral conclui a configuração de cluster quase que da mesma maneira que o Assistente de configuração de cluster de dois nós (mas a configuração pode ter mais de dois nós). O usuário pode modificar, testar ou remover o aplicativo ao usar o Smart Assist de aplicativo geral.

API do Smart Assist

O PowerHA SystemMirror inclui um *guia de desenvolvedores do Smart Assist* para que os OEMs possam desenvolver Smart Assists para integrar seus próprios aplicativos com o PowerHA SystemMirror.

Informações relacionadas:

Desenvolvendo aplicativos Smart Assist para o PowerHA SystemMirror

Iniciando, parando e reiniciando serviços de cluster

Após instalar o PowerHA SystemMirror e configurar o cluster, será possível iniciar os serviços de cluster. No PowerHA SystemMirror, as opções para iniciar, parar e reiniciar os serviços de cluster foram simplificadas e melhoradas. O PowerHA SystemMirror manipula as solicitações para iniciar e parar os serviços de cluster sem interromper os aplicativos, permitindo o controle total.

No PowerHA SystemMirror, é possível:

- *Iniciar e reiniciar os serviços de cluster.* Ao iniciar os serviços de cluster ou reiniciá-los após um encerramento, o PowerHA SystemMirror, por padrão, ativa automaticamente os recursos de acordo com a maneira em que eles foram definidos, considerando as dependências do aplicativo, os scripts de inicialização e de interrupção de aplicativo, os atributos dinâmicos e outros parâmetros. Ou seja, o PowerHA SystemMirror gerencia automaticamente (e ativa, se necessário) grupos de recursos e os aplicativos dentro deles.

Também é possível iniciar os serviços de cluster do PowerHA SystemMirror e informá-lo para não iniciar automaticamente nenhum grupo de recursos (e aplicativos). Se um aplicativo já estiver em execução, não será mais necessário pará-lo antes de iniciar os serviços de cluster.

Nota: O PowerHA SystemMirror baseia-se nos scripts de monitor de aplicativo e de inicialização de aplicativo para verificar se ele precisa iniciar o aplicativo ou se o aplicativo já está em execução (o PowerHA SystemMirror tenta não iniciar uma segunda instância do aplicativo). O PowerHA SystemMirror baseia-se nos monitores de aplicativo configurados para detectar falhas do aplicativo. Os monitores de aplicativo devem ser configurados para que o PowerHA SystemMirror detecte um cluster em execução, durante a inicialização, para que ele não inicie instâncias duplicadas do aplicativo. A abordagem alternativa é executar scripts que assegurem que instâncias duplicadas do controlador de aplicativo não sejam iniciadas.

- *Encerrar os serviços de cluster.* Durante um encerramento do PowerHA SystemMirror, é possível selecionar uma das três ações a seguir para os grupos de recursos:
 - Colocar offline
 - Mover para outros nós
 - Colocar os grupos de recursos no estado UNMANAGED

O Gerenciador do cluster lembra o estado de todos os nós e responde da maneira adequada quando os usuários tentam reiniciar os nós.

Informações relacionadas:

Administrando o PowerHA SystemMirror

Interface SMIT

É possível usar os painéis do SMIT fornecidos com o software PowerHA SystemMirror para executar várias tarefas.

Essas tarefas incluem:

- Configurar clusters, nós, redes, recursos e eventos.
- Capturar e restaurar capturas instantâneas de configurações de cluster.
- Ler arquivos de log.
- Diagnosticar problemas do cluster.
- Gerenciar um cluster usando o utilitário C-SPOC.
- Executar as tarefas de gerenciamento de grupo de recursos.
- Configurar a notificação automática de erro.
- Execute o adaptador de troca dinâmica.
- Configurar o ajuste de desempenho do cluster.
- Configurar métodos de disco customizados.

Gerenciamento de sistemas do PowerHA SystemMirror com C-SPOC

Para facilitar o gerenciamento de um cluster, o PowerHA SystemMirror fornece maneiras de executar comandos a partir de um nó e, em seguida, verificar e sincronizar as mudanças para todos os outros nós. É possível usar a ferramenta de Gerenciamento de sistema do PowerHA SystemMirror, o ponto único de controle do cluster (C-SPOC) para incluir usuários, arquivos e hardware automaticamente sem parar as tarefas essenciais.

É possível executar as seguintes tarefas usando o C-SPOC:

- Iniciar ou parar serviços do PowerHA SystemMirror
- Gerenciamento da interface de comunicação do PowerHA SystemMirror
- Gerenciamento de grupos de recursos e de aplicativo do PowerHA SystemMirror
- Gerenciamento de coleção de arquivos do PowerHA SystemMirror
- Visualização e gerenciamento de log do PowerHA SystemMirror
- Segurança e gerenciamento de usuários do PowerHA SystemMirror
- Gerenciamento de volume lógico do PowerHA SystemMirror
- Gerenciamento de volume lógico simultâneo do PowerHA SystemMirror
- Gerenciamento de volume físico do PowerHA SystemMirror
- Abrir uma sessão do SMIT em um Nó.

O utilitário C-SPOC simplifica a manutenção dos componentes do LVM compartilhados em clusters de até 16 nós. Os comandos do C-SPOC fornecem funções em um ambiente em cluster comparáveis aos comandos padrão do AIX que funcionam em um único nó. Ao automatizar tarefas repetitivas, o C-SPOC elimina uma origem de erros potencial e acelera o processo.

Sem a funcionalidade do C-SPOC, o administrador do sistema deve executar as tarefas administrativas individualmente em cada nó do cluster. Por exemplo, a tarefa para incluir um usuário geralmente deve ser executada em cada nó do cluster. Usando o utilitário C-SPOC, um comando que é executado em um nó, também é executado nos outros nós do cluster. Assim, o C-SPOC minimiza a sobrecarga administrativa e reduz a possibilidade de estados de nó inconsistentes. Usando o C-SPOC, você emite um comando do C-SPOC uma vez em um único nó e o usuário é incluído em todos os nós do cluster especificados.

O C-SPOC também torna mais eficientes o gerenciamento de componentes do volume lógico e o controle de serviços do cluster. É possível usar o utilitário C-SPOC para iniciar ou parar os serviços do cluster nos nós a partir de um único nó.

A figura a seguir ilustra uma configuração de dois nós e a interação de comandos, scripts e nós quando os serviços de cluster são iniciados a partir de um único nó do cluster. Observe que o prefixo *cl_* iniciará todos os comandos do C-SPOC.

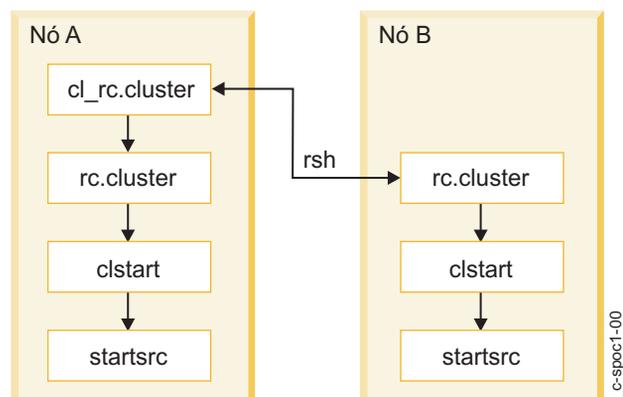


Figura 30. Fluxo de comandos usados na inicialização do cluster pelo utilitário C-SPOC

O C-SPOC fornece essa funcionalidade por meio de seu próprio conjunto de comandos de administração de cluster, acessíveis por meio de menus e painéis do SMIT. Para usar o C-SPOC, selecione a opção *Cluster System Management* no menu SMIT PowerHA SystemMirror.

Utilitário de captura instantânea do cluster

O utilitário de captura instantânea do cluster permite salvar as configurações de cluster que você gostaria de restaurar posteriormente.

Também é possível salvar informações adicionais do sistema e do cluster que podem ser úteis para diagnosticar problemas de configuração do sistema ou do cluster. É possível criar seus próprios métodos de captura instantânea customizados para armazenar informações adicionais sobre o cluster.

Uma captura instantânea de cluster permite ignorar o salvamento dos arquivos de log na captura instantânea. As capturas instantâneas do cluster são usadas para registrar as informações de configuração do cluster, enquanto que os logs do cluster somente registram a operação do cluster e não as informações de configuração. Por padrão, o PowerHA SystemMirror não coleta mais os arquivos de log do cluster quando a captura instantânea do cluster é criada, embora ainda seja possível especificar a coleta de logs no SMIT. Ignorar a coleção de logs acelera o tempo de execução do utilitário de captura instantânea e reduz o tamanho da captura instantânea.

Processamento de evento customizado

É possível definir vários pré-eventos e pós-eventos para customizar o processamento de eventos para necessidades exclusivas do site.

Informações relacionadas:

Administrando o PowerHA SystemMirror

Utilitário de gerenciamento de grupo de recursos

O utilitário de gerenciamento de grupo de recursos, `clRGmove`, fornece um meio para gerenciar grupos de recursos no cluster e aprimora os recursos de recuperação de falha do PowerHA SystemMirror.

Ele permite colocar qualquer tipo de grupo de recursos (juntamente com seus recursos – endereços IP, aplicativos e discos) online, offline ou em outro nó, sem parar os serviços de cluster. O gerenciamento de grupo de recursos ajuda gerenciar o cluster com mais eficiência, possibilitando um melhor aproveitamento dos recursos de hardware do cluster.

O gerenciamento de grupo de recurso também permite executar manutenção seletiva sem reinicializar o cluster ou perturbar os nós operacionais. Por exemplo, esse utilitário pode ser usado para liberar o nó dos grupos de recursos para executar manutenção do sistema em um nó do cluster específico.

O uso do utilitário de gerenciamento de grupo de recursos não afeta os outros grupos de recursos que pertencem atualmente a um nó. O nó atual libera-o e o nó de destino adquire-o, da mesma forma que durante um failover de nó. (Se houver dependências de local configuradas entre grupos de recursos, o PowerHA SystemMirror verificará e assegurará que elas sejam respeitadas).

Use o gerenciamento de grupo de recursos para concluir as seguintes tarefas:

- Mover temporariamente um grupo de recursos não simultâneos de um nó para outro (e de um site para outro) em um cluster em funcionamento.
- Colocar um grupo de recursos online ou offline em um nó ou em todos os nós no cluster.

Quando um grupo é movido, ele permanece no nó para o qual foi movido, até que seja movido novamente. Se você mover um grupo que possua a política de fallback Efetuar fallback para o nó com prioridade mais alta, o grupo efetuará fallback ou retornará para seu "novo" nó de prioridade mais alta temporário (nos casos em que o PowerHA SystemMirror precisa recuperá-lo em outros nós durante eventos de cluster subsequentes).

Se você desejar mover o grupo novamente, o PowerHA SystemMirror informará inteligentemente (nas listas de seleção com os nós de destino) se ele perceber que existe um nó com uma prioridade mais alta que possa hospedar um grupo. Será haver a opção de mover o grupo para esse nó.

Gerenciamento de coleção de arquivos do PowerHA SystemMirror

Como os grupos de volumes, determinados arquivos localizados em cada nó do cluster precisam ser mantidos em sincronização para que o PowerHA SystemMirror (e outros aplicativos) se comportem corretamente. Esses arquivos incluem scripts de evento, scripts de aplicativo e alguns arquivos de configuração do AIX e do PowerHA SystemMirror. O gerenciamento de Coleção de arquivo do PowerHA SystemMirror fornece uma maneira fácil de solicitar que uma lista de arquivos seja mantida em sincronização em todo o cluster.

Com o uso da coleção de arquivos do PowerHA SystemMirror, não é necessário copiar manualmente um arquivo atualizado para cada nó do cluster, verificar se o arquivo foi copiado corretamente e confirmar se cada nó tem a mesma versão do arquivo.

Além disso, se um ou mais desses arquivos for excluído ou danificado acidentalmente em um ou mais nós do cluster, poderá ser necessário tempo e esforço para determinar o problema. Com o uso da coleção de arquivos do PowerHA SystemMirror, esse cenário é minimizado. O PowerHA SystemMirror detecta quando um arquivo em uma coleção de arquivos é excluído ou se o tamanho do arquivo é alterado para zero e registra uma mensagem para informar o administrador. Duas coleções de arquivos do PowerHA SystemMirror são instaladas, por padrão:

- **Configuration_Files.** Um contêiner para arquivos essenciais do sistema, como `/etc/hosts` e `/etc/services`.
- **PowerHA SystemMirror_Files.** Um contêiner para todos os arquivos configuráveis pelo usuário na configuração do PowerHA SystemMirror. Esta é uma coleção de arquivos especial que o utilitário de propagação de coleção de arquivos subjacente usa para fazer referência a todos os arquivos configuráveis pelo usuário nas classes de configuração de banco de dados (ODM) do PowerHA SystemMirror.

Informações relacionadas:

Administrando o PowerHA SystemMirror

Instalando o PowerHA SystemMirror

Ferramentas de monitoramento

O PowerHA SystemMirror fornece várias ferramentas diferentes para monitoramento.

Muitos dos utilitários descritos aqui usam o arquivo `clhosts` para ativar a comunicação entre os nós do cluster do PowerHA SystemMirror.

Conceitos relacionados:

“Entendendo o arquivo `clhosts`” na página 101

Muitos dos monitores descritos neste tópico, incluindo os utilitários `Cinfo`, `HAView` e `clstat` contam com o uso de um arquivo `clhosts`. O arquivo `clhosts` contém informações de endereço IP que ajudam a ativar a comunicação entre os nós do cluster do PowerHA SystemMirror. O arquivo `clhosts` reside em todos os servidores e clientes do cluster do PowerHA SystemMirror.

Informações relacionadas:

Administrando o PowerHA SystemMirror

Gerenciador do cluster

O Gerenciador do cluster fornece informações e traps SNMP para clientes SNMP.

Ele reúne informações do cluster relativas a mudanças de estado do cluster de nós e interfaces. As informações do cluster podem ser recuperadas usando comandos SNMP ou programas clientes baseados em SNMP.

Conceitos relacionados:

“Gerenciador de cluster e programas de monitoramento SNMP” na página 49

Um cluster do PowerHA SystemMirror é dinâmico e pode passar por várias transições de estado ao longo do tempo.

Cluster information program

O Cluster Information Program (Cinfo) reúne informações do cluster a partir do SNMP e permite que os clientes se comuniquem com esse programa para ficarem cientes das mudanças de estado do cluster.

Conceitos relacionados:

“Cluster information program” na página 50

O Cluster information program (Cinfo), o daemon **cinfo**, é um monitor baseado em SNMP.

Monitoramento de aplicativo

O monitoramento de aplicativo permite configurar vários monitores para que um controlador de aplicativo monitore aplicativos e processos específicos e definir a ação a ser executada após a detecção de um encerramento inesperado de processo ou de outras falhas de aplicativo.

Conceitos relacionados:

“Eliminando aplicativos como um ponto único de falha” na página 59

O motivo principal para criar clusters do PowerHA SystemMirror é fornecer um ambiente altamente disponível para aplicativos essenciais.

Monitor de aplicativo clam_nfsv4

O monitor de aplicativo clam_nfsv4 é incluído automaticamente pelo PowerHA SystemMirror Versão 7.1 quando os recursos do Network File System versão 4 (NFSV4) são especificados em um grupo de recursos.

O script do monitor de aplicativo clam_nfsv4 é executado uma vez por minuto, por padrão. A contagem de reinicialização dos monitores é zero. O monitor de aplicativo clam_nfsv4 possui uma ação de falha de failover. Quando ocorre um failover, qualquer erro registrado pelo monitor faz com que o servidor de aplicativos NFSV4 fique no estado offline e os grupos de recursos no servidor de aplicativos sejam movidos para o próximo nó com prioridade mais alta.

Não é possível remover o monitor de aplicativo clam_nfsv4, mas é possível desativá-lo alterando a ação de falha para notificar.

O monitor de aplicativo clam_nfsv4 verifica se os nós com configurações de montagem cruzada estão montados corretamente. O monitor também verifica se o servidor NFS pode responder à solicitação do cliente feita por meio de configurações de montagem cruzada.

O monitor de aplicativo clam_nfsv4 verifica se as seguintes funções do subsistema estão em execução:

- nfsrgyd
- nfsd
- rpc.mountd
- rpc.statd
- rpc.lockd

Se alguma das funções do subsistema não estiver sendo executada, o monitor de aplicativo clam_nfsv4 identificará o servidor de aplicativos NFSV4 como offline.

Você deve registrar os nós do NFS que contêm grupos de recursos. Você também deve configurar o NFS para exportar os grupos de recursos corretamente.

Informações relacionadas:

Usando o NFS com o PowerHA SystemMirror

Montagem cruzada do NFS no PowerHA SystemMirror

Resolução de problemas do monitor de aplicativo clam_nfsv4

Opção do SMIT Mostrar aplicativos de cluster

A opção do SMIT Mostrar aplicativos de cluster fornece uma visualização centrada em aplicativo da configuração de cluster.

Este utilitário exibe as interfaces e informações existentes em um tipo de visualização "aplicativo inativo".

Utilitário Status do cluster (clstat)

O utilitário Status do cluster, `/usr/es/sbin/cluster/clstat`, monitora o status do cluster. O utilitário relata o status dos principais componentes: o próprio cluster, os nós no cluster, as interfaces de rede conectadas aos nós e os grupos de recursos em cada nó.

Ele relata se o cluster está ativo, inativo ou instável. Ele também relata se um nó está ativo, inativo, sendo inserido ou removido do cluster, ou reconfigurando e o número de nós no cluster. O utilitário **clstat** fornece as interfaces ASCII, Motif, X Window System e HTML. É possível executar o **clstat** a partir do ASCII SMIT.

Para o cluster como um todo, o **clstat** indica o estado do cluster e o número de nós do cluster. Para cada nó, o **clstat** exibe o rótulo e o endereço IP de cada interface de rede de serviço conectada ao nó e se essa interface está ativa ou inativa. O **clstat** também exibe o estado do grupo de recursos.

É possível visualizar informações de status no modo de exibição ASCII ou X Window System ou por meio de um navegador da web.

Nota: O utilitário **clstat** usa a API Clinfo para recuperar informações sobre o cluster. Portanto, assegure-se de que o Clinfo esteja em execução no sistema cliente para visualizar a exibição do **clstat**.

Ferramenta Análise de disponibilidade de aplicativo

A ferramenta Análise de disponibilidade de aplicativo mede estatísticas de tempo de atividade para aplicativos com controladores de aplicativos definidos para o PowerHA SystemMirror.

O software PowerHA SystemMirror coleta, faz o registro de data e hora e registra informações abrangentes sobre os aplicativos escolhidos para serem monitorados com esta ferramenta. Usando SMIT, é possível selecionar um período de tempo e a ferramenta exibe as estatísticas de tempo de atividade e de tempo de inatividade para um determinado aplicativo durante esse período.

Rótulos de IP de nó persistente

Um *rótulo de IP de nó persistente* é uma ferramenta administrativa útil que permite entrar em contato com um nó mesmo se os serviços de cluster do PowerHA SystemMirror estiverem inativos nesse nó.

Quando são definidos rótulos de IP de nó persistente, o PowerHA SystemMirror tenta colocar um endereço IP no nó. A designação de um rótulo de IP de nó persistente a uma rede em um nó permite que haja um endereço IP de limite de nó em uma rede de cluster que pode ser usado para propósitos administrativos para acessar um nó específico no cluster. Um rótulo de IP de nó persistente é um alias de IP que pode ser designado a um nó específico em uma rede de cluster e que:

- Sempre permanece no mesmo nó (*é limite de nó*)
- Coexiste em uma placa da interface de rede que já possui um rótulo de IP de serviço definido
- Não requer a instalação de uma placa da interface de rede física adicional nesse nó
- Não faz parte de nenhum *grupo de recursos*.

Pode haver um rótulo de IP de rede persistente por rede, por nó.

Verificação e sincronização do PowerHA SystemMirror

O processo de verificação e sincronização do PowerHA SystemMirror verifica se modificações específicas do PowerHA SystemMirror para os arquivos do sistema AIX estão corretas, se o cluster e seus recursos

estão configurados corretamente, se a segurança (caso esteja configurada) está configurada corretamente, se todos os nós concordam quanto à topologia do cluster, à configuração de rede e à propriedade e ao controle de recursos do PowerHA SystemMirror.

A verificação também indica se existem métodos de captura instantânea de cluster customizados e se eles são executáveis em cada nó do cluster.

Sempre que configurar, reconfigurar ou atualizar um cluster, em seguida, você deverá executar o procedimento de verificação do cluster. Se a verificação for bem-sucedida, a configuração será automaticamente sincronizada. A sincronização entra em vigor imediatamente em um cluster ativo.

O utilitário de verificação mantém um registro detalhado das informações no banco de dados de configuração do PowerHA SystemMirror em cada um dos nós após sua execução. Os subdiretórios para cada nó contêm informações sobre a última verificação bem-sucedida (aprovada), a penúltima verificação bem-sucedida (`pass.prev`) e a última verificação malsucedida (com falha).

Mensagens emitidas pelo utilitário indicam onde o erro ocorreu (por exemplo, no nó, no dispositivo, no comando e assim por diante).

Verificação com o monitoramento automático de configuração de cluster:

O PowerHA SystemMirror fornece o monitoramento automático de configuração de cluster. Por padrão, o PowerHA SystemMirror executa automaticamente a *verificação* no nó que aparece primeiro na ordem alfabética uma vez a cada 24 horas à meia-noite.

O administrador do cluster é notificado se a configuração de cluster se torna inválida. Quando a verificação do cluster é concluída no nó do cluster selecionado, esse nó notifica os outros nós do cluster. Cada nó armazena as informações sobre a data e a hora em que o nó executou a verificação e os resultados da verificação no arquivo `/var/hacmp/log/clutils.log`. Se o nó selecionado se tornar indisponível ou não puder concluir a verificação do cluster, isso poderá ser detectado pela ausência de um relatório no arquivo `// var hacmp/log/clutils.log`. Se a verificação do cluster for concluída e detectar alguns erros de configuração, você será notificado sobre os problemas potenciais:

- O status de saída da verificação será publicado no cluster juntamente com as informações sobre a conclusão do processo de verificação do cluster.
- Mensagens transmitidas serão enviadas no cluster e exibidas no `stdout`. Essas mensagens informarão sobre os erros de configuração detectados.

Verificação com ações corretivas:

A verificação do cluster consiste em uma série de verificações executadas com relação a vários componentes do servidor PowerHA SystemMirror configurados pelo usuário. Cada verificação tenta detectar um problema de consistência de cluster ou um erro de configuração.

Algumas condições de erro resultam quando informações importantes para a operação do PowerHA SystemMirror, mas não parte do próprio software PowerHA SystemMirror, não são propagadas corretamente a todos os nós do cluster.

Por padrão, a verificação é executada com o modo de ações corretivas automáticas ativado para as configurações Padrão e Estendida. Esse é o modo recomendado para executar a verificação. Se necessário, o modo de ações corretivas automáticas pode ser desativado para a configuração Estendida. No entanto, observe que a execução da verificação no modo de ação corretiva automática permite automatizar muitas tarefas de configuração, como a criação de um arquivo `clhosts` baseado em cliente, que é usado por muitos dos monitores

Quando a verificação detecta alguma das condições a seguir, é possível autorizar uma ação corretiva antes que a verificação de erro continue:

- Os registros de data e hora do grupo de volumes compartilhados do PowerHA SystemMirror não correspondem em todos os nós.
- O arquivo `/etc/hosts` em um nó não contém todos os rótulos de IP e endereços IP gerenciados pelo PowerHA SystemMirror.
- Um sistema de arquivos não é criado em um nó que faz parte do grupo de recursos, embora haja discos disponíveis.
- Há discos disponíveis, mas um grupo de volumes não foi importado para um nó.
- As entradas `/etc/services` necessárias estão ausentes em um nó.
- As entradas `snmpd` do PowerHA SystemMirror necessárias estão ausentes em um nó.

Se um erro encontrado durante a verificação acionar quaisquer ações corretivas, o utilitário executará todas as verificações novamente após concluir a primeira vez. Se a mesma verificação falhar novamente e o problema original for um erro, o erro será registrado e a verificação falhará. Se a condição original for um aviso, a verificação será bem-sucedida.

Conceitos relacionados:

“Entendendo o arquivo `clhosts`”

Muitos dos monitores descritos neste tópico, incluindo os utilitários `Clinfo`, `HAView` e `clstat` contam com o uso de um arquivo `clhosts`. O arquivo `clhosts` contém informações de endereço IP que ajudam a ativar a comunicação entre os nós do cluster do PowerHA SystemMirror. O arquivo `clhosts` reside em todos os servidores e clientes do cluster do PowerHA SystemMirror.

Métodos de verificação customizados:

Por meio do SMIT, também é possível incluir, alterar ou remover métodos de verificação definidos customizados que executam verificações específicas na configuração de cluster. É possível executar a verificação a partir da linha de comandos ou por meio da interface SMIT para emitir um método de notificação remoto em resposta a um evento de cluster.

Entendendo o arquivo `clhosts`:

Muitos dos monitores descritos neste tópico, incluindo os utilitários `Clinfo`, `HAView` e `clstat` contam com o uso de um arquivo `clhosts`. O arquivo `clhosts` contém informações de endereço IP que ajudam a ativar a comunicação entre os nós do cluster do PowerHA SystemMirror. O arquivo `clhosts` reside em todos os servidores e clientes do cluster do PowerHA SystemMirror.

Há diferenças para o arquivo `clhosts`, dependendo de onde o arquivo reside, conforme é resumido na tabela a seguir:

Tabela 3. Arquivo `clhosts`

arquivo <code>clhosts</code>	Descrição
arquivo baseado em servidor	Esse arquivo fica no diretório <code>/usr/es/sbin/cluster/etc/</code> em todos os nós de servidor do PowerHA SystemMirror. Durante a instalação do PowerHA SystemMirror, o endereço de loopback Protocolo da Internet versão 4 (IPv4) padrão 127.0.0.1 é incluído automaticamente no arquivo. Se houver algum endereço Protocolo da Internet versão 6 (IPv6) configurado no momento da instalação, o endereço de loopback IPv6 <code>::1</code> também será incluído. O loopback de nome e o host local de alias que esse endereço IP geralmente define não são necessários.
arquivo baseado em cliente	Esse arquivo fica no diretório <code>/usr/es/sbin/cluster/etc/</code> em todos os nós clientes do PowerHA SystemMirror. Ao executar a verificação com a função de correções automáticas, este arquivo é gerado automaticamente no servidor e é denominado <code>clhosts.client</code> . Você deve copiar esse arquivo para os nós clientes manualmente e renomeá-lo como <code>clhosts</code> . Este arquivo contém todos os endereços IP conhecidos e nunca deve conter os endereços 127.0.0.1, <code>::1</code> , de loopback ou de host local.

Quando um daemon do monitor é inicializado, ele lê no arquivo local `/usr/es/sbin/cluster/etc/clhosts` para determinar quais nós estão disponíveis para comunicação, da seguinte forma:

- Para os daemons em execução em um *nó de servidor* do PowerHA SystemMirror, o arquivo `clhosts` baseado no servidor local somente requer o endereço de loopback (`127.0.0.1` and `::1`), que é incluído automaticamente no arquivo `clhosts` baseado no servidor, quando a parte do servidor do PowerHA SystemMirror é instalada.
- Para os daemons em execução em um *nó cliente* do PowerHA SystemMirror, o arquivo `clhosts` baseado no cliente local deve conter uma lista de endereços IP dos nós de servidor do PowerHA SystemMirror. Dessa forma, se um determinado nó de servidor do PowerHA SystemMirror estiver indisponível (por exemplo, desligado), o daemon no nó cliente ainda poderá comunicar-se com outros nós de servidor do PowerHA SystemMirror.

O utilitário de verificação do PowerHA SystemMirror ajuda a preencher o arquivo `clhosts` baseado no cliente da seguinte maneira, localizando todos os nós de servidor do PowerHA SystemMirror disponíveis, criando um arquivo `/usr/es/sbin/cluster/etc/clhosts.client` nos nós de servidor e preenchendo o arquivo com os endereços IP desses nós de servidor do PowerHA SystemMirror.

Após concluir a verificação e a sincronização do PowerHA SystemMirror no cluster, você deverá copiar manualmente este arquivo **clhosts.client** para cada nó cliente, como `/usr/es/sbin/cluster/etc/clhosts` (renomeie-o removendo a extensão `.client`).

Conceitos relacionados:

“Verificação e sincronização do PowerHA SystemMirror” na página 99

O processo de verificação e sincronização do PowerHA SystemMirror verifica se modificações específicas do PowerHA SystemMirror para os arquivos do sistema AIX estão corretas, se o cluster e seus recursos estão configurados corretamente, se a segurança (caso esteja configurada) está configurada corretamente, se todos os nós concordam quanto à topologia do cluster, à configuração de rede e à propriedade e ao controle de recursos do PowerHA SystemMirror.

Ferramentas de resolução de problemas

Geralmente, um cluster do PowerHA SystemMirror em funcionamento requer o mínimo de intervenção. No entanto, se ocorrer um problema serão essenciais as qualificação de diagnóstico e recuperação. Portanto, a resolução de problemas requer a identificação rápida do problema e a aplicação do entendimento sobre o software PowerHA SystemMirror para restaurar o cluster para operação completa.

Conceitos relacionados:

“Ferramenta de teste do cluster” na página 104

A Ferramenta de teste do cluster é um utilitário que permite testar uma configuração de cluster do PowerHA SystemMirror para avaliar como um cluster se comporta sob um conjunto de circunstâncias especificadas, como quando um nó fica inacessível, uma rede fica inacessível, um grupo de recursos é movido de um nó para outro e assim por diante.

Arquivos de log

O software PowerHA SystemMirror grava as mensagens que gera para o console do sistema e para vários arquivos de log. Como cada arquivo de log contém um nível diferente de detalhes, os administradores do sistema podem focar em aspectos diferentes do processamento do PowerHA SystemMirror, visualizando diferentes arquivos de log.

Os arquivos de log principais incluem:

- O arquivo `/var/hacmp/adm/cluster.log` controla eventos de cluster.
- O arquivo `/var/hacmp/log/hacmp.out` registra a saída gerada por scripts de configuração à medida que eles são executados. As sumarizações de evento aparecem após a saída detalhada para eventos iniciados pelo Gerenciador de cluster, facilitando a varredura do arquivo **hacmp.out** para obter informações importantes. Além disso, as sumarizações do evento fornecem links HTML para os eventos correspondentes no arquivo **hacmp.out**.
- O arquivo de log `/var/hacmp/adm/history/cluster.mmdyyyy` registra o histórico diário do cluster.

- O arquivo `// var hacmp/clverify/clverify.log` contém as mensagens detalhadas emitidas durante a verificação. A verificação de cluster consiste em uma série de verificações executadas com relação a várias configurações do PowerHA SystemMirror. Cada verificação tenta detectar algum problema ou erro de consistência do cluster. As mensagens emitidas pelo utilitário de verificação indicam onde ocorreu o erro (por exemplo, no nó, no dispositivo, no comando e assim por diante).

O PowerHA SystemMirror permite visualizar, redirecionar, salvar e alterar os parâmetros dos arquivos de log, para que seja possível customizá-los para suas necessidades específicas.

Informações relacionadas:

Administrando o PowerHA SystemMirror

Reconfigurando os valores ajustáveis do PowerHA SystemMirror

Ao configurar e testar um cluster, é possível alterar qualquer valor entre os valores ajustáveis do PowerHA SystemMirror que afetam o desempenho do cluster. Ou, é possível reconfigurar os valores ajustáveis para as configurações padrão sem alterar nenhum outro aspecto da configuração.

Um administrador ou consultor de cluster de terceiros pode ser solicitado para controlar a administração de um cluster que ele não configurou e pode ser necessário reconfigurar os valores ajustáveis para seus padrões. É possível reconfigurar os valores ajustáveis do cluster usando a interface SMIT. O PowerHA SystemMirror obtém uma captura instantânea do cluster, antes de reinicializar. Depois que os valores forem reconfigurados para os padrões, se você desejar retornar às configurações de cluster customizadas, será possível aplicar a captura instantânea do cluster. A reconfiguração dos valores ajustáveis do cluster reconfigura as informações no banco de dados de configuração do cluster. As informações que são reconfiguradas ou removidas abrangem as seguintes categorias:

- Informações fornecidas pelos usuários (por exemplo, scripts pré e pós-evento e parâmetros da rede, como máscaras de rede). Observe que a reconfiguração dos valores ajustáveis do cluster não remove os scripts pré e pós-evento que já estavam configurados. No entanto, se você reconfigurar os valores ajustáveis, o conhecimento dos scripts pré e pós-eventos do PowerHA SystemMirror será removido da configuração e esses scripts não serão mais usados pelo PowerHA SystemMirror para gerenciar recursos no cluster. Será possível reconfigurar o PowerHA SystemMirror para usar esses scripts novamente, se necessário.
- Informações geradas automaticamente pelo PowerHA SystemMirror durante a configuração e sincronização. Elas incluem os IDs de nó e rede e as informações descobertas a partir do sistema operacional, como máscaras de rede. Geralmente, os usuários não podem ver as informações geradas.

Informações relacionadas:

Administrando o PowerHA SystemMirror

Instalando o PowerHA SystemMirror

Arquivo de informações de status do cluster

Ao usar o utilitário Captura instantânea do PowerHA SystemMirror para salvar um registro de uma configuração de cluster (conforme visto a partir de cada nó do cluster), opcionalmente, você faz com que o utilitário execute vários comandos do AIX e comandos do PowerHA SystemMirror padrão para obter informações de status sobre o cluster. Essas informações são armazenadas em um arquivo, identificado pela extensão `.info`, no diretório de capturas instantâneas.

O diretório de capturas instantâneas é definido pelo valor da variável de ambiente `SNAPSHOTPATH`. Por padrão, o utilitário de captura instantânea de cluster inclui a saída dos comandos, como `clssif`, `cllsnw`, `df`, `ls` e `netstat`. É possível criar métodos de captura instantânea customizados para especificar informações adicionais que você deseja armazenar no arquivo `.info`.

Uma captura instantânea de cluster permite ignorar o salvamento dos arquivos de log na captura instantânea. As capturas instantâneas de cluster são usadas para registrar informações de configuração do cluster, enquanto os logs de cluster somente registram a operação do cluster e *não* as informações de configuração. Por padrão, o PowerHA SystemMirror não coleta mais os arquivos de log do cluster

quando a captura instantânea do cluster é criada, embora ainda seja possível especificar a coleta de logs no SMIT. Ignorar a coleção de logs reduz o tamanho da captura instantânea e acelera a execução do utilitário de captura instantânea. O tamanho da captura instantânea de cluster depende da configuração. Por exemplo, uma configuração básica de dois nós requer aproximadamente 40KB.

Notificação automática de erro

O recurso Notificação de erro do AIX pode ser usado para detectar eventos que não são monitorados especificamente pelo software PowerHA SystemMirror. Por exemplo, é possível especificar que ocorra uma resposta para eventos como uma falha do adaptador de disco.

Normalmente, os métodos de notificação de erro são definidos manualmente, um por um. O PowerHA SystemMirror fornece um conjunto de métodos de notificação pré-especificados para erros importantes que podem ser "ativados" automaticamente, em uma única etapa, por meio da interface SMIT, economizando tempo e esforço consideráveis, por não ser necessário definir cada método de notificação manualmente.

Notificação remota customizada

É possível definir um método de notificação por meio da interface SMIT para emitir um método de notificação customizado em resposta a um evento de cluster. Também é possível enviar uma notificação de mensagem de texto para qualquer endereço, incluindo um telefone celular, ou um email para um endereço de email.

Após a configuração de um método de notificação remoto, é possível enviar uma mensagem de teste para confirmar se a configuração está correta.

É possível configurar qualquer número de métodos de notificação, para eventos diferentes e com diferentes mensagens de texto e números de telefone para discar. O mesmo método de notificação pode ser usado para vários eventos diferentes, desde que a mensagem de texto associada transmita informações suficientes para responder a todos os possíveis eventos que acionarem a notificação.

Prefácios e sumarizações de evento

Os detalhes de eventos de cluster são registrados no arquivo **hacmp.out**. A saída detalhada deste arquivo contém muitas linhas de informações de evento. Aparece uma sumarização concisa no final de cada detalhe de evento. Para obter uma verificação rápida e eficiente sobre o que aconteceu no cluster recentemente, é possível visualizar uma compilação somente de partes sumarizadas do evento dos arquivos **hacmp.out** atuais e anteriores, usando o painel **Exibir sumarizações do evento** no SMIT.

Também é possível selecionar para salvar as sumarizações de evento compiladas em um arquivo de sua escolha. Opcionalmente, os resumos de eventos fornecem links HTML para os eventos correspondentes no arquivo **hacmp.out**. O Gerenciador de cluster também imprime uma prefácio que informa quais grupos de recursos estão enfileirados para processamento para cada evento. É possível visualizar a ordem de processamento que será seguida.

Informações relacionadas:

Resolução de problemas do PowerHA SystemMirror

Ferramenta de teste do cluster

A Ferramenta de teste do cluster é um utilitário que permite testar uma configuração de cluster do PowerHA SystemMirror para avaliar como um cluster se comporta sob um conjunto de circunstâncias especificadas, como quando um nó fica inacessível, uma rede fica inacessível, um grupo de recursos é movido de um nó para outro e assim por diante.

É possível iniciar o teste, permitir que ele seja executado de maneira não assistida e retornar posteriormente para avaliar os resultados do teste.

Se desejar executar um conjunto de testes do cluster básicos e automatizados para o gerenciamento de topologia e de grupo de recursos, será possível executar o conjunto de teste automatizado a partir do SMIT. Se você for um administrador do PowerHA SystemMirror experiente e desejar customizar o teste de cluster para o ambiente, também será possível criar testes customizados que possam ser executados a partir do SMIT.

É recomendado executar a ferramenta após configurar o PowerHA SystemMirror inicialmente e antes de colocar o cluster em um ambiente de produção, após fazer mudanças na configuração de cluster enquanto o cluster estiver fora de serviço ou em intervalos regulares, embora o cluster pareça estar funcionando bem.

Avisos

Estas informações foram desenvolvidas para produtos e serviços que são oferecidos nos Estados Unidos.

É possível que a IBM não ofereça os produtos, serviços ou recursos discutidos neste documento em outros países. Consulte um representante IBM local para obter informações sobre produtos e serviços disponíveis atualmente em sua área. Qualquer referência a produtos, programas ou serviços IBM não significa que apenas produtos, programas ou serviços IBM possam ser utilizados. Qualquer produto, programa ou serviço funcionalmente equivalente, que não infrinja nenhum direito de propriedade intelectual da IBM pode ser utilizado em substituição. Entretanto, a avaliação e verificação da operação de qualquer produto, programa ou serviço não-IBM são de responsabilidade do Cliente.

A IBM pode ter patentes ou solicitações de patentes pendentes relativas a assuntos tratados nesta publicação. O fornecimento desta publicação não lhe garante direito algum sobre tais patentes. Pedidos de licença devem ser enviados, por escrito, para:

Gerência de Relações Comerciais e Industriais da IBM Brasil
Av. Pasteur, 138-146
Botafogo
Rio de Janeiro, RJ
CEP 22290-240

Para pedidos de licença relacionados a informações de DBCS (Conjunto de Caracteres de Byte Duplo), entre em contato com o Departamento de Propriedade Intelectual da IBM em seu país ou envie pedidos de licença, por escrito, para:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

O parágrafo a seguir não se aplica a nenhum país em que tais disposições não estejam de acordo com a legislação local: A INTERNATIONAL BUSINESS MACHINES CORPORATION FORNECE ESTA PUBLICAÇÃO "NO ESTADO EM QUE SE ENCONTRA", SEM GARANTIA DE NENHUM TIPO, SEJA EXPRESSA OU IMPLÍCITA, INCLUINDO, MAS A ELAS NÃO SE LIMITANDO, AS GARANTIAS IMPLÍCITAS DE NÃO INFRAÇÃO, COMERCIALIZAÇÃO OU ADEQUAÇÃO A UM DETERMINADO PROPÓSITO. Alguns países não permitem a exclusão de garantias expressas ou implícitas em certas transações; portanto, esta disposição pode não se aplicar ao Cliente.

Estas informações podem conter imprecisões técnicas ou erros tipográficos. Periodicamente, são feitas alterações nas informações aqui contidas; tais alterações serão incorporadas em futuras edições desta publicação. A IBM pode, a qualquer momento, aprimorar e/ou alterar os produtos e/ou programas descritos nesta publicação, sem aviso prévio.

Quaisquer referências nestas informações a websites não IBM são fornecidas para conveniência apenas e não representam de forma alguma um endosso a estes websites. Os materiais nesses websites não fazem parte dos materiais desse produto IBM e a utilização desses websites é de inteira responsabilidade do Cliente.

A IBM pode utilizar ou distribuir as informações fornecidas da forma que julgar apropriada, sem incorrer em qualquer obrigação para com o Cliente.

Licenciados deste programa que desejam obter informações sobre este assunto com objetivo de permitir: (i) a troca de informações entre programas criados independentemente e outros programas (incluindo este) e (ii) a utilização mútua das informações trocadas, devem entrar em contato com:

*IBM Corporation
Dept. LRAS/Bldg. 903
11501 Burnet Road
Austin, TX 78758-3400
USA*

Tais informações podem estar disponíveis, sujeitas a termos e condições apropriadas, incluindo em alguns casos o pagamento de uma taxa.

O programa licenciado descrito neste documento e todo o material licenciado disponível são fornecidos pela IBM sob os termos do Contrato com o Cliente IBM, do Contrato de Licença do Programa Internacional IBM ou de qualquer outro contrato equivalente entre as partes.

Todos os dados de desempenho aqui contidos foram determinados em um ambiente controlado. Portanto, os resultados obtidos em outros ambientes operacionais podem variar significativamente. Algumas medidas podem ter sido tomadas em sistemas de nível de desenvolvimento e não há garantia de que essas medidas serão iguais em sistemas geralmente disponíveis. Além disso, algumas medidas podem ter sido estimadas por extrapolação. Os resultados reais poderão variar. Os usuários deste documento devem verificar os dados aplicáveis para seu ambiente específico.

As informações relativas a produtos não-IBM foram obtidas junto aos fornecedores dos respectivos produtos, de seus anúncios publicados ou de outras fontes disponíveis publicamente. A IBM não testou estes produtos e não pode confirmar a precisão de seu desempenho, compatibilidade nem qualquer outra reivindicação relacionada a produtos não-IBM. Dúvidas sobre os recursos de produtos não-IBM devem ser encaminhadas diretamente a seus fornecedores.

Todas as declarações relacionadas aos objetivos e intenções futuras da IBM estão sujeitas a alterações ou cancelamento sem aviso prévio e representam apenas metas e objetivos.

Todos os preços IBM mostrados são preços de varejo da IBM sugeridos, são atuais e estão sujeitos a mudanças sem aviso prévio. Os preços do revendedor podem variar.

Estas informações são para fins de planejamento apenas. As informações aqui contidas estão sujeitas a alterações antes que os produtos descritos estejam disponíveis.

Estas informações contêm exemplos de dados e relatórios utilizados nas operações diárias de negócios. Para ilustrá-los da forma mais completa possível, os exemplos podem incluir nomes de indivíduos, empresas, marcas e produtos. Todos esses nomes são fictícios e qualquer semelhança com nomes e endereços utilizados por uma empresa real é mera coincidência.

LICENÇA DE COPYRIGHT:

Estas informações contêm programas de aplicativos de amostra no idioma de origem, ilustrando as técnicas de programação em diversas plataformas operacionais. O Cliente pode copiar, modificar e distribuir esses exemplos de programas de qualquer forma, sem pagamento à IBM, com o objetivo de desenvolver, utilizar, vender ou distribuir programas aplicativos de acordo com a interface de programação de aplicativo da plataforma operacional para a qual os exemplos de programas são escritos. Esses exemplos não foram completamente testados sob todas as condições. A IBM não pode, portanto, garantir ou inferir confiabilidade, possibilidade de assistência ou função desses programas. Os programas de amostra são fornecidos "NO ESTADO EM QUE SE ENCONTREM", sem qualquer garantia de nenhum tipo. A IBM não será responsável por quaisquer danos decorrentes do uso pelo Cliente dos programas de amostra.

Cada cópia ou parte desses programas de amostra ou qualquer trabalho derivado deve incluir um aviso de copyright com os dizeres:

Partes deste código são derivadas dos Programas de Amostra do IBM Corp.

© Copyright IBM Corp. _digite o ano ou anos_. Todos os direitos reservados.

Considerações de política de privacidade

Os Produtos de software IBM, incluindo soluções de software como serviço, (“Ofertas de Software”) podem usar cookies ou outras tecnologias para coletar informações de uso do produto, para ajudar a melhorar a experiência do usuário final, para customizar as interações com o usuário final ou para outros fins. Em muitos casos, nenhuma informação pessoalmente identificável é coletada pelas Ofertas de Software. Algumas de nossas Ofertas de Software podem ajudar a coletar informações pessoalmente identificáveis. Se esta Oferta de Software usar cookies para coletar informações pessoalmente identificáveis, informações específicas sobre o uso de cookies desta oferta serão definidas abaixo.

Esta Oferta de Software não usa cookies ou outras tecnologias para coletar informações pessoalmente identificáveis.

Se as configurações implementadas para esta Oferta de Software fornecerem a você como cliente a capacidade de coletar informações pessoalmente identificáveis de usuários finais via cookies e outras tecnologias, você deve buscar seu próprio aconselhamento jurídico sobre quaisquer leis aplicáveis a tal coleta de dados, incluindo requisitos para aviso e consento.

Para obter mais informações sobre o uso de várias tecnologias, incluindo cookies, para estes fins, consulte a Política de Privacidade da IBM em <http://www.ibm.com/privacy> e Declaração de Privacidade Online da IBM na <http://www.ibm.com/privacy/details> seção intitulada “Cookies, Web Beacons and Other Technologies” e “IBM Software Products and Software-as-a-Service Privacy Statement” em <http://www.ibm.com/software/info/product-privacy>.

Marcas registradas

IBM, o logotipo IBM e [ibm.com](http://www.ibm.com) são marcas comerciais ou marcas registradas da International Business Machines Corp., registradas em vários países no mundo todo. Outros nomes de produtos e serviços podem ser marcas comerciais da IBM ou de outras empresas. Uma lista atual de marcas comerciais da IBM está disponível na web em Copyright and trademark information em www.ibm.com/legal/copytrade.shtml.

UNIX é uma marca registrada do The Open Group nos Estados Unidos e em outros países.

Índice Remissivo

Caracteres Especiais

/.rhosts 92

A

acesso a disco
 não simultâneo 51
 simultâneo 51
adaptador de disco
 eliminando como falha 63
Alias de IP 26
aplicativo 34
 eliminando como falha 59
 monitoramento 98
arquivo de informações de status do cluster 103

C

C-SPOC 70, 95
captura instantânea 96
clam_nfsv4 98
clhosts 101
cliente 15
clinfo 50
Cinfo 98
clRGmove 68, 96
clstat 99
cluster
 C-SPOC 70
 captura instantânea 96
 cliente 15
 componentes físicos 9
 Controle de endereço IP 26
 dispositivos de disco 11
 estendido 22
 eventos 73
 ferramenta de teste 104
 grupos de recursos 36
 hardware 45
 nó 10, 16
 rede 11, 23
 segurança 92
 site 16
 software 45
 soluções de vários sites 20
 vinculado 21
cluster information program 50
coleção de arquivos 97
comutações de rede
 multicasting 30
configuração
 aplicativos com multicamadas 83
 controle 79
 mútua 80
 mútuo de dois nós 81
 mútuo de oito nós 82
 unilateral 80
 dependência de local de grupo de recursos 85
 dynamic LPAR 87
 espelho do LVM entre sites 86

configuração (*continuação*)
 espera 76
 exemplo 76, 78
 opções 91
 smart assist 92
Configurações de divisão 17
configurando
 caminho de configuração estendido 92
 caminho de configuração padrão 91
Controle de endereço IP 26
controle rápido de disco 72

D

DARE
 Veja reconfiguração automática e dinâmica
disco
 eliminando como falha 63
disco de repositório
 caminhos múltiplos 13
 falha 14
disponibilidade de aplicativo
 visão geral 54
dispositivo de comunicação 26
dispositivo de disco 11

E

eliminando
 pontos únicos de falha 55
 adaptador de disco 63
 aplicativo 59
 disco 63
 interface de comunicação 60
 nó 56
 rede 62
endereço IP de serviço 26, 35
espelhamento do LVM entre sites 72
exemplo
 configuração de espera 76, 78

F

ferramenta de análise de disponibilidade de aplicativo 59, 99
fita 35

G

gerenciador de recurso simultâneo 53
gerenciador do cluster 47
 SNMP 50
grupo de recursos 36
 dependências de local 43
 fallback 37, 38
 fallover 37, 38
 gerenciamento 68
 inicialização 37, 38
 política 38
 prioridade dinâmica do nó 39

- grupo de recursos (*continuação*)
 - site 44
 - tempo de acomodação 39
 - timerfallback de fallback 39
 - utilitário de gerenciamento 96
- grupo de volumes 33

H

- hacmp.out 104
- hardware 45

I

- interface de comunicação 25
 - eliminando como falha 60

L

- log 102

M

- maximizando
 - recuperação de desastre 72
- minimizando
 - tempo de controle 72
 - tempo de inatividade não planejado 71
 - tempo de inatividade planejado 64
- monitor de aplicativos 59
 - clam_nfsv4 98
- monitoramento
 - aplicativo 98
 - arquivo clhosts 101
 - cluster information program 98
 - ferramenta de análise de disponibilidade de aplicativo 99
 - gerenciador do cluster 97
 - rótulos de IP de nó persistente 99
 - utilitário de status do cluster 99
- multicasting 30
 - comunicação de pacote 30
 - comutações de rede 30
 - Internet Group Management Protocol (IGMP) 31
 - roteamento 31

N

- NFS 51
- nó 10, 16
 - eliminando como falha 56
- notificação automática de erro 104
- notificação remota customizada 104

P

- Política de mesclagem 18
- PowerHA SystemMirror
 - multicasting 30
- pulsação 28
 - rede ponto a ponto 29
 - rede TCP/IP 29

R

- reconfiguração automática e dinâmica 64
- reconfigurando
 - valores ajustáveis 103
- recuperação de desastre 72
- recurso
 - aplicativos 34
 - endereço IP de serviço 35
 - fita 35
 - grupos de volumes 33
 - rótulo de IP de serviço 35
 - sistemas de arquivos 34
 - volumes lógicos 33
- rede 11, 23
 - Alias de IP 26
 - dispositivos de comunicação 26
 - eliminando como falha 62
 - endereço IP de serviço 26
 - física 24
 - interfaces de comunicação 25
 - lógica 24
 - pulsação 28
 - rótulo de IP de serviço 26
 - sub-rede 26
- rede física 24
- rede lógica 24
- resolução de problemas
 - arquivo de informações de status do cluster 103
 - arquivos de log 102
 - ferramenta de teste do cluster 104
 - notificação automática de erro 104
 - notificação remota customizada 104
 - prefácios do evento 104
 - reconfigurando valores ajustáveis 103
 - sumarizações de evento 104
- rótulo de IP de nó persistente 99
- rótulo de IP de serviço 26, 35

S

- segurança 92
- serviços de cluster
 - iniciando 94
 - parando 94
 - reiniciando 94
- sistema de arquivos 34
- site 16, 44
 - soluções 20
- smart assist 92
- SMIT 94
- SNMP 50, 97
- software 45
 - complementar 54
 - componentes 48
- sub-rede 26

V

- valores ajustáveis
 - reconfigurando 103
- verificando
 - com ações corretivas 100
 - com monitoramento automático de configuração de cluster 100
 - métodos customizados 101

visão geral
 disponibilidade de aplicativo 54
volume lógico 33



Impresso no Brasil