

IBM PowerHA SystemMirror for AIX

Standard Edition

Versão 7.2

Planejando o PowerHA SystemMirror

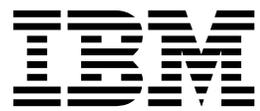
IBM

IBM PowerHA SystemMirror for AIX

Standard Edition

Versão 7.2

Planejando o PowerHA SystemMirror



Nota

Antes de usar estas informações e o produto suportado por elas, leia as informações no “Avisos” na página 111.

Índice

Sobre este documento	v
Destaque	v
Distinção entre maiúsculas e minúsculas no AIX ..	v
ISO 9000.	v
Informações relacionadas	v
Planejando o PowerHA SystemMirror ..	1
O que há de novo no planejamento do PowerHA SystemMirror	1
Visão geral do processo de planejamento	1
Orientações de planejamento	2
Eliminando pontos únicos de falha: configurando componentes redundantes suportados pelo PowerHA SystemMirror	3
Visão geral do processo de planejamento	3
Planejamento de cluster inicial	5
Planejando nós do cluster	5
Planejamento para disco de repositório e endereço IP multicast do cluster	6
Planejamento para fence de disco	8
Planejamento para sites de cluster	9
Planejando a segurança do cluster	10
Planejamento do aplicativo	11
Projetando um diagrama de cluster	16
Requisitos de nome do host	17
Planejando a conectividade de rede do cluster.	17
Considerações gerais da rede para o PowerHA SystemMirror.	18
Monitoramento no PowerHA SystemMirror ..	21
Projetando a topologia de rede	21
Planejamento para controle de endereço IP via aliases de IP	25
Planejamento para outras condições de rede ..	29
Evitando conflitos de rede	33
Incluindo a topologia de rede no diagrama de cluster	33
Planejando dispositivos de disco e de fita compartilhados	33
Visão geral do disco compartilhado e dispositivos de fita	33
Escolhendo uma tecnologia de disco compartilhado	34
Considerações de fonte de alimentação de disco	35
Planejamento para armazenamento em disco não compartilhado	35
Planejando uma instalação de disco compartilhado	36
Incluindo a configuração do disco no diagrama de cluster	37
Planejamento para unidades de fita e recursos de cluster	37
Planejando componentes de LVM compartilhados	39
Planejamento para componentes do LVM ..	40
Planejando espelhamento de LVM.	42
Planejamento para acesso de disco.	46

Usando controle rápido de disco	47
Usando o quorum e a ativação para aumentar a disponibilidade de dados.	49
Usando NFS com o PowerHA SystemMirror ..	52
Planejando grupos de recursos	59
Visão geral para grupos de recursos	59
Regras gerais para recursos e grupos de recursos	60
Tipos de grupos de recursos: simultâneo e não simultâneo	60
Políticas de grupo de recursos para inicialização, fallover e fallback	61
Atributos de grupo de recursos.	62
Movendo grupos de recursos para outro nó ..	70
Planejando redes de cluster e grupos de recursos	71
Planejamento paralelo ou serial da ordem para processamento de grupos de recursos	71
Planejando grupos de recursos nos clusters que possuem sites.	72
Planejamento para recursos replicados	79
Planejamento para o Workload Manager.	80
Planejamento para eventos de cluster.	82
Planejando eventos de site e nó.	83
Planejamento de eventos node_up e node_down	84
Eventos de rede	87
Eventos da interface de rede.	88
Eventos de status em todo o cluster	89
Manipulação e recuperação de evento do grupo de recursos	90
Customizando o processamento de evento de cluster	92
Notificação remota customizada de eventos ..	97
Customizando o tempo de duração do evento até o aviso	98
Eventos definidos pelo usuário	98
Resumos de eventos e prefácio	101
Planejamento para clientes do PowerHA SystemMirror	101
Clientes que executam Clinfo	101
Clientes que não executam Clinfo	102
Componentes de rede	102
Aplicativos e PowerHA SystemMirror	102
Visão geral dos aplicativos e PowerHA SystemMirror	102
Automação de aplicativo: minimizando a intervenção manual	103
Dependências de aplicativo.	106
Interferência de aplicativos	107
Robustez do aplicativo	107
Estratégias de implementação de aplicativo ..	108

Avisos	111
Considerações de política de privacidade	113
Marcas registradas.	113

Índice Remissivo	115
-----------------------------------	------------

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-chaves, 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.
Monoespaçamento	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 se residirem no mesmo diretório. Para evitar que ações indesejadas sejam executadas, verifique sempre se você está usando a diferenciação entre maiúsculas e minúsculas correta.

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.

Planejando o PowerHA SystemMirror

Antes de configurar e instalar o PowerHA SystemMirror, você deve planejar sua implementação para o sistema operacional AIX.

O que há de novo no planejamento do PowerHA SystemMirror

Leia sobre informações novas ou significativamente alteradas para a coleção de tópico Planejamento 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 feitas nesta coleção de tópico:

- Informações incluídas sobre o número de discos de repositório de backup que podem ser identificados no tópico “Planejamento para disco de repositório e endereço IP multicast do cluster” na página 6.
- Informações incluídas sobre a função Resource Optimized High Availability (ROHA) no tópico “Planejamento para capacidade on demand” na página 12.
- Informações incluídas sobre o recurso de fence de disco no tópico “Planejamento para fence de disco” na página 8.

Visão geral do processo de planejamento

É possível planejar a alta disponibilidade em um datacenter com o PowerHA SystemMirror Standard Edition for AIX e para a alta disponibilidade de diversos sites e a recuperação de desastre com o PowerHA SystemMirror Enterprise Edition for AIX.

Seu principal objetivo em todo o processo de planejamento é eliminar pontos únicos de falha. 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 forma de fornecer essa função e o aplicativo ou serviço dependente desse componente torna-se indisponível.

Por exemplo, se todos os dados para um aplicativo crítico residirem em um único disco e esse disco falhar, esse disco será um ponto único de falha para todo o cluster. Os clientes não podem acessar esse aplicativo até que os dados no disco sejam restaurados. Da mesma forma, se os dados do aplicativo dinâmico forem armazenados em discos internos em vez de em discos externos, não será possível recuperar um aplicativo tendo outro nó do cluster controlando os discos. Portanto, a identificação de componentes lógicos necessários exigidos por um aplicativo, como sistemas de arquivos e diretórios (que poderiam conter dados de aplicativo e variáveis de configuração), é um pré-requisito importante para planejar um cluster bem-sucedido.

Perceba que, enquanto seu objetivo é eliminar todos os pontos únicos de falha, é possível ter que assumir alguns compromissos. Geralmente um custo é associado à eliminação de um ponto único de falha. Por exemplo, a compra de um dispositivo de hardware adicional para servir como backup para o dispositivo primário aumenta o custo. O custo da eliminação de um ponto único de falha poderia ser comparado com o custo da perda de serviços se esse componente falhar. Novamente, o propósito do PowerHA SystemMirror é fornecer uma plataforma computacional altamente disponível com custo reduzido que pode atender às demandas de processamento futuras.

Nota: É importante que as falhas dos componentes de cluster sejam solucionadas o mais rápido possível. Dependendo da sua configuração, o PowerHA SystemMirror pode não ser capaz de manipular uma segunda falha, devido a uma falha de recursos.

Orientações de planejamento

Projetar o cluster que fornece a melhor solução para sua organização requer planejamento cuidadoso e pensado. De fato, o planejamento adequado é chave para construir um cluster bem-sucedido do PowerHA SystemMirror. Um cluster bem planejado é mais fácil de instalar, fornece uma disponibilidade de aplicativo mais alta, é executado melhor e requer menos manutenção que um cluster planejado deficientemente.

Pode ser preciso planejar processos adicionais em seu ambiente. Por exemplo, os processos de gerenciamento de correção e de gerenciamento de processo são críticos se desejar que seu ambiente manipule vários tipos de falhas.

Para que um aplicativo crítico esteja altamente disponível, nenhum dos recursos associados deve ser um ponto único de falha. Conforme você projeta um cluster do PowerHA SystemMirror, seu objetivo é identificar e abordar todos os pontos únicos de falha em potencial. As questões a se fazer incluem:

- Que serviços de aplicativo devem estar altamente disponíveis? Qual é a prioridade desses serviços?
- Qual é o custo de uma falha comparado ao hardware necessário para eliminar a possibilidade desta falha?
- Qual é o número máximo de componentes de software e hardware que o PowerHA SystemMirror pode suportar?
- Qual é a disponibilidade necessária desses serviços? Eles precisam estar disponíveis 24 horas por dia, sete dias por semana ou oito horas ao dia, cinco dias por semana é suficiente?
- O que poderia acontecer para interromper a disponibilidade desses serviços?
- Qual é o tempo atribuído para substituir um recurso com falha? Qual é o grau aceitável de degradação de desempenho ao operar após uma falha?
- Quais falhas serão automaticamente detectadas como eventos de cluster? Que falhas precisam ter o código customizado gravado para detectar a falha e acionar um evento de cluster?
- Qual é o nível de qualificação do grupo que está implementando e mantendo o cluster?

Para planejar, implementar e manter um cluster bem-sucedido do PowerHA SystemMirror é necessário continuar a comunicação entre muitos grupos em sua organização. Idealmente, você deve montar os representantes a seguir (conforme aplicável) para auxiliar nas sessões de planejamento do PowerHA SystemMirror:

- Administrador da rede
- Administrador do sistema
- Administrador de base de dados
- Programação de aplicativos
- Equipe de suporte
- Usuários finais

O PowerHA SystemMirror suporta uma variedade de configurações, fornecendo a você uma grande quantidade de flexibilidade. Para obter informações sobre projetar o nível mais alto de disponibilidade para seu cluster, consulte o White Paper da IBM *High Availability Cluster Multiprocessing Best Practices*.

Referências relacionadas:

“Eliminando pontos únicos de falha: configurando componentes redundantes suportados pelo PowerHA SystemMirror” na página 3

O software PowerHA SystemMirror fornece várias opções para evitar pontos únicos de falha.

Informações relacionadas:

➡ Melhores práticas de multiprocessamento de cluster de alta disponibilidade

Eliminando pontos únicos de falha: configurando componentes redundantes suportados pelo PowerHA SystemMirror

O software PowerHA SystemMirror fornece várias opções para evitar pontos únicos de falha.

A tabela a seguir resume os pontos únicos de falha em potencial e descreve como eliminá-los configurando componentes de cluster redundantes de hardware e software.

Componentes de cluster	Para eliminar como um ponto único de falha	Suportes a PowerHA SystemMirror
Nós	Usar diversos nós	Até 16.
Fontes de Power (alimentação)	Usar diversos circuitos ou fontes de alimentação ininterruptas	Quantos forem necessários.
Redes	Usar diversas redes para conectar nós	Até 48.
Interfaces de rede, dispositivos e rótulos	Usar adaptadores de rede redundantes	Até 256.
Subsistemas TCP/IP	Usar redes para conectar unindo nós e clientes	Quantos forem necessários.
Adaptadores de disco	Usar adaptadores de disco redundantes	Quantos forem necessários.
Controladores	Usar controladores de disco redundantes	Quantos forem necessários.
Discos	Usar espelhamento de hardware e disco redundantes, striping ou ambos	Quantos forem necessários.
Aplicações	Designe um nó para o controle do aplicativo, para configurar um monitor de aplicativo e configurar clusters com nós em mais de um site.	Políticas de configuração flexíveis para alta disponibilidade em um site e entre sites.
Sites	Usar mais de um site para recuperação de desastre.	Até dois sites.
Grupos de recursos	Usar grupos de recursos para especificar como um conjunto de entidades devem executar.	Até 64 por cluster.
Recursos de cluster	Usar diversos recursos de cluster.	Até 128 para daemon Clinfo (pode existir mais no cluster).
Servidor de E/S Virtual (VIOS)	Usar VIOS redundante	Quantos forem necessários.
HMC	Usar RMC redundante	2
Sistema gerenciado hospedando um nó do cluster	Usar sistemas gerenciados separados para cada nó do cluster	16 nós
Disco de repositório do cluster	Usar proteção RAID	Um disco de repositório ativo por site que possui a capacidade de substituir o disco após uma falha. Você deve ter um disco sobressalente disponível para substituir o disco de repositório com falha no cluster em tempo real.

Referências relacionadas:

“Orientações de planejamento” na página 2

Projetar o cluster que fornece a melhor solução para sua organização requer planejamento cuidadoso e pensado. De fato, o planejamento adequado é chave para construir um cluster bem-sucedido do PowerHA SystemMirror. Um cluster bem planejado é mais fácil de instalar, fornece uma disponibilidade de aplicativo mais alta, é executado melhor e requer menos manutenção que um cluster planejado deficientemente.

Visão geral do processo de planejamento

Este tópico descreve as etapas para o planejamento de um cluster do PowerHA SystemMirror.

Etapa 1: Planejamento para aplicativos altamente disponíveis

Nessa etapa, você planeja o núcleo do cluster: os aplicativos a se tornarem altamente disponíveis, os tipos de recursos que requerem, o número de nós, endereços IP compartilhados e um modo de compartilhar discos (acesso não simultâneo ou simultâneo). Seu objetivo é desenvolver uma visão de alto nível do

sistema que atende como um ponto de início para o design do cluster. Após tomar essas decisões iniciais, comece a projetar um diagrama do cluster. O planejamento de cluster inicial descreve essa etapa do processo de planejamento.

Etapa 2: Planejamento para topologia de cluster

Nessa etapa, você decide sobre nomes para o cluster e dos nós. Opcionalmente, você também decide sobre os nomes para os sites e decide sobre que nós pertencem a qual site. O planejamento de cluster inicial descreve essa etapa do processo de planejamento.

Etapa 3: Planejamento de sites

Nessa etapa, você decide se o site usará um cluster estendido ou um cluster vinculado. Um cluster estendido contém nós de sites nas mesmas localizações geográficas. Os clusters estendidos devem compartilhar um disco de repositório. Um cluster vinculado contém nós de sites que estão localizados em diferentes localizações geográficas. Os clusters vinculados usam um disco de repositório separado.

Etapa 4: Planejamento de conectividade de rede do cluster

Nessa etapa, você planeja as redes que conectam os nós em seu sistema. Primeiro, você examina problemas em relação a redes TCP/IP e de ponto a ponto em um ambiente do PowerHA SystemMirror. O planejamento da conectividade de rede de cluster descreve esta etapa do processo de planejamento.

Etapa 5: Planejamento de dispositivos de disco compartilhado

Nessa etapa, você planeja os dispositivos de disco compartilhado para o cluster. Você decide quais tecnologias de armazenamento em disco usará em seu cluster e examinará problemas em relação a essas tecnologias no ambiente do PowerHA SystemMirror. O planejamento de disco compartilhado e dispositivos de fita descreve essa etapa do processo de planejamento.

Etapa 6: Planejamento de componentes do LVM compartilhados

Nessa etapa, você planeja os grupos de volumes compartilhados para o cluster. Primeiro, você examina problemas em relação aos componentes do LVM em um ambiente do PowerHA SystemMirror. O planejamento de componentes do LVM compartilhados descreve essa etapa do processo de planejamento.

Etapa 7: Planejamento de grupos de recursos

O planejamento de grupos de recursos incorpora todas as informações que você gerou nas etapas anteriores. Além disso, é necessário decidir sobre o uso de grupos de recursos dependentes ou políticas de tempo de execução específicas para manter certos grupos de recursos relacionados no mesmo nó ou em diferentes nós. O planejamento de grupos de recursos descreve essa etapa do processo de planejamento.

Etapa 8: Planejamento do processamento de evento de cluster

Nessa etapa, você planeja o processamento de eventos para seu cluster. O planejamento para os eventos do cluster descreve esta etapa do processo de planejamento.

Etapa 9: Planejamento de clientes do PowerHA SystemMirror

Nessa etapa, você examina problemas em relação a clientes do PowerHA SystemMirror. O planejamento para clientes do PowerHA SystemMirror descreve esta etapa do processo de planejamento.

Referências relacionadas:

“Planejamento de cluster inicial”

Esta seção descreve as etapas iniciais que você executa para planejar um cluster do PowerHA SystemMirror para tornar os aplicativos altamente disponíveis.

“Planejando a conectividade de rede do cluster” na página 17

Esta seção descreve o planejamento do suporte de rede para um cluster do PowerHA SystemMirror.

“Planejando grupos de recursos” na página 59

Estes tópicos descrevem como planejar grupos de recursos em um cluster do PowerHA SystemMirror.

“Planejando componentes de LVM compartilhados” na página 39

Esta seção descreve o planejamento de grupos de volumes compartilhados para um cluster do PowerHA SystemMirror.

“Planejando dispositivos de disco e de fita compartilhados” na página 33

Esta seção discute informações a considerar antes de configurar discos externos compartilhados em um cluster do PowerHA SystemMirror e fornece informações sobre planejamento e configuração de unidades de fita como recursos de cluster.

“Planejamento para eventos de cluster” na página 82

Estes tópicos descrevem os eventos do cluster do PowerHA SystemMirror.

“Planejamento para clientes do PowerHA SystemMirror” na página 101

Estes tópicos discutem considerações de planejamento para clientes do PowerHA SystemMirror. Esta é a última etapa antes de continuar a instalação do seu software do PowerHA SystemMirror.

“Planejamento para sites de cluster” na página 9

Os clusters do PowerHA SystemMirror podem ser usados em um único site ou diversos sites para recuperação de desastre.

Informações relacionadas:

Comportamento do grupo de recursos durante eventos de cluster

Planejamento de cluster inicial

Esta seção descreve as etapas iniciais que você executa para planejar um cluster do PowerHA SystemMirror para tornar os aplicativos altamente disponíveis.

Antes de iniciar o planejamento do PowerHA SystemMirror, assegure-se de entender os conceitos e terminologia relevante para o PowerHA SystemMirror.

Um cluster do PowerHA SystemMirror fornece um ambiente altamente disponível para aplicativos essenciais. Em muitas organizações, esses aplicativos devem permanecer disponível todas as vezes. Por exemplo, um cluster do PowerHA SystemMirror poderia executar um programa do servidor de banco de dados que atende aos aplicativos do cliente, mantendo-o altamente disponíveis para clientes que enviam consultas ao programa do servidor.

Informações relacionadas:

Conceitos do PowerHA SystemMirror

Planejando nós do cluster

Para cada aplicativo crítico, lembre-se dos recursos necessários pelo aplicativo, incluindo seus requisitos de armazenamento de dados e processamento.

Por exemplo, ao planejar o tamanho de seu cluster, inclua nós suficientes para manipular os requisitos de processamento do seu aplicativo após um nó falhar.

Lembre-se das considerações a seguir ao determinar o número de nós do cluster:

- Um cluster do PowerHA SystemMirror pode ser composto de qualquer combinação de servidores IBM® Power Systems. Assegure-se de que todos os nós do cluster não compartilhem os componentes que poderiam ser um ponto único de falha (por exemplo, fonte de alimentação). Da mesma forma, não coloque nós em um único rack.
- Crie pequenos clusters que consistam em nós que executem funções semelhantes ou compartilhem recursos. Cluster menores e simples são mais fáceis de projetar, implementar e manter.
- Por motivos de desempenho, pode ser desejável usar diversos nós para suportar o mesmo aplicativo. Para fornecer serviços de controle mútuo, o aplicativo deve ser projetado de modo que permita que diversas instâncias do aplicativo sejam executadas no mesmo nó.

Por exemplo, se um aplicativo requerer que os dados dinâmicos residam em um diretório chamado */data*, o aplicativo provavelmente não poderá suportar diversas instâncias no mesmo processador. Para tal aplicativo (em execução em um ambiente não simultâneo), tente particionar os dados para que diversas instâncias do aplicativo possam ser executadas, cada uma acessando um banco de dados exclusivo.

Além disso, se o aplicativo suportar arquivos de configuração que permitam que o administrador especifique que os dados dinâmicos para *instance1* do aplicativo residam no diretório *data1*, *instance2* residirá no diretório *data2*, e assim por diante, então, diversas instâncias do aplicativo serão suportadas provavelmente.

- Certas configurações, incluindo nós adicionais no design do cluster, podem aumentar o nível de disponibilidade fornecido pelo cluster. Certas configurações também lhe dão mais flexibilidade no planejamento de fallover e reintegração de nó.

A configuração de nó do cluster mais confiável é ter pelo menos um nó de espera.

- Escolha nós de cluster que têm slots de E/S suficientes para suportar placas da interface de rede e adaptadores de disco redundantes.

Lembre-se, o cluster composto de diversos nó ainda é mais caro que um único nó, mas sem o planejamento para suportar o hardware redundante, (como slots de E/S suficientes para redes e adaptadores de disco), o cluster não terá melhor disponibilidade.

- Use nós com velocidade de processamento semelhante.
- Use nós com ciclos de CPU suficientes e largura da banda de E/S para permitir que o aplicativo de produção seja executado na carga de pico. Lembre-se, os nós devem ter capacidade suficiente para permitir que o PowerHA SystemMirror opere.

Para planejar isso, faça uma avaliação de desempenho ou modele seu aplicativo de produção e liste os parâmetros das cargas mais pesadas esperadas. Em seguida, escolha nós para um cluster do PowerHA SystemMirror que não excederá 85% ocupado ao executar seu aplicativo de produção.

Ao criar um cluster, designe um nome a ele. O PowerHA SystemMirror associa esse nome ao ID do cluster designado pelo PowerHA SystemMirror.

Planejamento para disco de repositório e endereço IP multicast do cluster

O armazenamento em cluster do PowerHA SystemMirror pode ser feito por meio de rede multicast ou unicast. Para escolher o modo multicast de armazenamento em cluster, é possível planejar e fornecer um endereço IP multicast para comunicação no cluster. Por padrão, se nenhum endereço IP multicast for fornecido ao implementar o cluster, o PowerHA implementa o cluster baseado em unicast (comunicação de soquete TCP/IP normal).

Disco de repositório do cluster

- | Você tem 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.

O PowerHA SystemMirror usa um disco compartilhado para armazenar informações de configuração de cluster do Cluster Aware AIX (CAA). Você deve ter pelo menos 512 MB e não mais que 460 GB de espaço em disco alocado para o disco de repositório de cluster. Essa configuração é mantida altamente disponível automaticamente no disco fornecido. Este recurso requer que um disco compartilhado dedicado esteja disponível para todos os nós que fazem parte do cluster. Esse disco não pode ser usado para armazenamento de aplicativo ou qualquer outro propósito.

Ao planejar os discos que deseja usar como discos de repositório, você deve planejar discos de backup ou de substituição, que poderão ser usados no caso de falha do disco de repositório principal. O disco de backup deve ser do mesmo tamanho e tipo que o disco principal, mas pode estar em um disco de armazenamento físico diferente. Atualize os procedimentos administrativos e a documentação com as informações do disco de backup. Também é possível substituir um disco de repositório ativo por um novo, para aumentar o tamanho, ou alterar para um subsistema de armazenamento diferente. Para substituir um disco de repositório, é possível usar a interface SMIT.

Nota: Se o disco compartilhado que é usado como um disco de repositório for um disco SCSI virtual (vSCSI) mapeado, você deverá mapear o disco como um disco vSCSI para todos os nós no cluster. O mapeamento do disco vSCSI deve ser idêntico em todos os nós no cluster. Por exemplo, não é possível mapear o disco de repositório usando o método vSCSI para um nó no cluster e mapear o mesmo disco usando o método N-Port ID Virtualization (NPIV) para outro nó no cluster.

Endereço IP multicast do cluster

É possível usar um endereço IP multicast para monitoramento e comunicação de cluster. É possível especificar esse endereço ao criar o cluster ou é possível ter um gerado automaticamente ao sincronizar a configuração de cluster inicial.

Nota: O mecanismo padrão usa comunicações unicast e não requer configuração extra. Entretanto, se desejar usar a comunicação multicast, você deve continuar lendo e assegurar que seus dispositivos de rede sejam ativados para comunicações multicast.

Se decidir usar o multicast, o PowerHA SystemMirror usa a comunicação baseada em multicast entre hosts no cluster. A rede do seu ambiente deve permitir que pacotes de IP multicast fluam entre hosts no cluster. Para verificar se os nós em seu ambiente suportam a comunicação baseada em multicast, use o comando **mping**. Execute o comando **mping** antes de iniciar o uso do PowerHA SystemMirror em seu ambiente.

Nota: Alguns dos comutadores de rede permitem que os pacotes de multicast fluam por um tempo antes de pará-los. Portanto, é crítico conduzir o teste **mping** por pelo menos 5 minutos e assegurar que a malha de rede permita que o pacote de multicast flua sem quaisquer problemas. Além disso, quando os comutadores são em cascata, geralmente eles precisam de configuração adicional para rotear pacotes de multicast. Para configurar o fluxo do pacote de multicast, consulte a documentação fornecida pelo fornecedor do comutador para configurar o fluxo de pacote multicast.

Um endereço multicast também é conhecido como um endereço de classe D. Cada datagrama de IP cujo endereço de destino se inicia com 1110 é um datagrama de IP multicast. Os 28 bits restantes identificam o grupo multicast no qual o datagrama é enviado. Você deve configurar seu kernel para receber pacotes enviados para um grupo multicast específico, que faz o host unir o grupo na interface especificada.

Não use os grupos multicast a seguir:

224.0.0.1

Este é o grupo todos os hosts. Se você executar ping desse grupo, todos os hosts compatíveis com multicast na rede devem responder, porque cada host compatível com multicast deve se unir a esse grupo na inicialização em todas as interfaces compatíveis com multicast.

224.0.0.2

Este é o grupo todos os roteadores. Todos os roteadores multicast devem se unir a esse grupo em todas as interfaces compatíveis com multicast.

224.0.0.4

Este é o grupo todos os roteadores DVMRP.

224.0.0.5

Este é o grupo todos os roteadores OSPF.

224.0.0.13

Este é o grupo todos os roteadores PIM.

Nota: O intervalo 224.0.0.0-224.0.0.255 é reservado para fins locais, como tarefas administrativas e de manutenção, e os dados que recebem nunca são encaminhados pelos roteadores multicast. Da mesma forma, o intervalo 239.0.0.0 - 239.255.255.255 é reservado para administrativo administrativo. Esses grupos multicast especiais são regularmente publicados no RFC de números designados.

O PowerHA SystemMirror 7.1.2, ou posterior, suporta IP versão 6 (IPv6), no entanto, não é possível especificar explicitamente o endereço multicast IPv6. O CAA usa um endereço multicast IPv6 que é derivado do endereço multicast IP 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`.

Informações relacionadas:

Falha de disco de repositório

Substituindo um disco de repositório pelo SMIT

Resolução de problemas de multicast

Testando multicast em uma rede

| Planejamento para fence de disco

| O fence de disco é um dos recursos para políticas de quarentena disponível no PowerHA SystemMirror.

| Para usar a opção de fence de disco, você deve identificar o recurso crítico e o subsistema de armazenamento deve suportar a reserva persistente SCSI-3 e ODM `reserve_policy` de `PR_shared`. Essa política é aplicada a todos os discos que são parte do grupo de volumes e do grupo de recursos.

| Tipo de reserva

| O disco deve ter um tipo de reserva de Write Exclusive All Registrant (WEAR). É possível executar os comandos `clmgr` a seguir para verificar se seu disco suporta a capacidade SCSI-3 (WEAR - tipo 7h) para suportar o fence de disco do PowerHA SystemMirror:

```
| clmgr scsibr_capability query physical_volume <disk>  
| clmgr scsibr_capability query volume_group <vg>
```

| em que `disk` é o nome do disco e `vg` é o nome do grupo de volumes.

| `reserve_policy`

| Define se um método de reserva está em execução no disco. A `reserve_policy` de `PR_shared` é a política necessária que se aplica ao método de host compartilhado para o disco. Para visualizar os atributos para o disco, execute o comando `lsattr -RI <diskname> -a reserve_policy`.

| Para usar o recurso de fence de disco, você deve especificar um grupo de recursos crítico. O grupo de recursos crítico que você especificar deve atender aos critérios a seguir:

- | • O grupo de recursos crítico não pode ser incluído como um filho em nenhum relacionamento de dependência `parent_child`, `start_after` ou `stop_after`.
- | • O grupo de recursos crítico deve ter todos os nós no cluster como nós participantes.

- O grupo de recursos crítico não pode usar a política de inicialização **On-line no nó inicial apenas**. O grupo de recursos crítico pode usar qualquer outra política de inicialização.

A reserve_policy para um disco de um grupo de volumes que possui serviços de cluster que estão em execução e o fence de disco ativado deve usar as configurações a seguir:

Tabela 1. Configurações de disco

Opção	Valor
Política de reserva configurada	PR_shared
Política de reserva efetiva	PR_shared
Status de reserva	Reserva SCSI PR (Write Exclusive All Registrant)

A reserva persistente SCSI-3 não é suportada em discos do HyperSwap. Se você estiver usando um disco do HyperSwap, a verificação exibe uma mensagem de erro.

Uma política de mesclagem que usa a opção **Maioria** agora pode funcionar corretamente com a reserva persistente SCSI-3 porque os nós do site principal já podem estar protegidos. Neste caso, uma política de mesclagem que usa a opção **Maioria** não é suportada.

Informações relacionadas:

- Configurando uma política de quarentena
- Resolução de problemas de fence de disco

Planejamento para sites de cluster

Os clusters do PowerHA SystemMirror podem ser usados em um único site ou diversos sites para recuperação de desastre.

Se você tiver diversos sites, é possível usar os recursos do PowerHA SystemMirror Enterprise Edition a seguir para a recuperação de desastre:

- O PowerHA SystemMirror Enterprise Edition for AIX inclui opções para suporte a tecnologias de replicação que estão disponíveis a partir de vários subsistemas de armazenamento.
- O PowerHA SystemMirror Enterprise Edition for AIX para o GLVM fornece a replicação baseada em host sobre redes TCP/IP.

O PowerHA SystemMirror 7.1.2, ou posterior, suporta diferentes tipos de definições para sites e políticas específicas de site para alta disponibilidade e recuperação de desastre (HADR). É possível definir vários sites no PowerHA SystemMirror Standard Edition for AIX e no PowerHA SystemMirror Enterprise Edition for AIX.

O PowerHA SystemMirror usa Cluster Aware AIX (CAA) para comunicação do cluster e gerenciamento de funcionamento de cluster.

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

Clusters estendidos

Contém nós de sites que estão localizados nas mesmas localizações geográficas. Os clusters estendidos devem compartilhar um disco de repositório entre todos os nós no site. Os clusters estendidos não suportam HADR com o Gerenciamento de replicação de armazenamento. Para usar clusters estendidos, seu ambiente de rede deve suportar a comunicação baseada em multicast.

Clusters vinculados

Contém nós de sites que estão localizados em localizações geográficas diferentes. Os clusters vinculados usam um disco de repositório separado. Os clusters vinculados suportam o

espelhamento LVM entre sites e o HyperSwap. Nos clusters vinculados, o CAA usa pacotes unicast para se comunicar e gerenciar os sites entre os clusters CAA independentes.

Planejando recursos e política do site

O PowerHA SystemMirror tenta assegurar que a instância primária de um grupo de recursos seja mantida on-line em um site e a instância secundária seja mantida on-line no outro site. Planeje quais nós configurar em qual site e onde deseja que os aplicativos sejam executados, para que possa planejar as políticas de grupo de recursos de acordo.

Nota: Os sites são suportados apenas no PowerHA SystemMirror 7.1.2 ou posterior, em ambas as Enterprise Edition e Standard Edition. O gerenciamento de replicação é suportado apenas no PowerHA SystemMirror Enterprise Edition.

Todos os recursos definidos para o PowerHA SystemMirror devem ter nomes exclusivos. Os rótulos de IP de serviço, grupos de volumes e nomes de grupos de recursos devem ser exclusivos no cluster e distintos entre si. O nome de um recurso deve estar relacionado ao aplicativo a que atende, bem como qualquer dispositivo correspondente. Por exemplo, um endereço de serviço para um grupo de recursos que executa uma instância do WebSphere poderia ser chamada `websphere_service_address`.

Planejando a segurança do cluster

O PowerHA SystemMirror fornece a segurança do cluster controlando o acesso de usuário ao PowerHA SystemMirror e fornecendo a segurança para comunicações entre nós.

Autenticação de conexão

O PowerHA SystemMirror fornece a autenticação de conexão para proteger comunicações do PowerHA SystemMirror entre nós do cluster. Essa autenticação de conexão também é conhecida como autenticação padrão. A autenticação padrão inclui conexões verificadas por endereço IP e nome do host e limita os comandos que podem ser executados com o privilégio raiz. Este modo usa o princípio de menor privilégio para execução de comando remoto, que assegura que nenhum comando arbitrário possa ser executado em um nó remoto com privilégio de administrador. Um conjunto seletivo de comandos do PowerHA SystemMirror é considerado confiável e permitido para execução como raiz. Todos os outros comandos são executados como usuário `nobody`. A dependência `/.rhosts` para a comunicação entre nós é eliminada.

Também é possível configurar uma rede privada virtual (VPN) para comunicações entre nós. Se você usar uma VPN, use rótulos de IP persistentes para túneis de VPN.

Configuração de segurança

O PowerHA SystemMirror usa a função do Cluster Aware AIX (CAA) para criar o caminho de comunicação segura para pulsações e sincronização entre nós em um cluster.

É possível usar os métodos de CAA a seguir para criar credenciais de segurança de cluster para nós no cluster.

Autoassinado

O PowerHA SystemMirror gera as credenciais de segurança.

Certificado de segurança e pares de chaves privadas

O PowerHA SystemMirror usa o certificado de segurança e pares de chaves privadas existentes que você fornecer.

Secure Shell (SSH)

O PowerHA SystemMirror usa as chaves já configuradas para comunicação SSH em seu ambiente.

Criptografia e autenticação de mensagem

O PowerHA SystemMirror fornece segurança para mensagens do PowerHA SystemMirror enviadas entre os nós de cluster como a seguir:

- A autenticação de mensagem assegura a origem e a integridade de uma mensagem.
- A criptografia de mensagem altera a aparência dos dados enquanto são transmitidos e retorna-os para sua forma original quando recebidos por um nó que autentica a mensagem.
- As mensagens são criptografadas ou com hash executado dependendo de um nível de segurança de Baixo, Médio ou Alto. Um nível de segurança Baixo executa hash apenas de algumas das mensagens em comparação com um nível de segurança Alto em que as mensagens são criptografadas.

O PowerHA SystemMirror suporta os tipos a seguir de chaves de criptografia para autenticação de mensagem e criptografia:

- Message Digest 5 (MD5) com Padrão de Criptografia de Dados (DES)
- MD5 com Triple DES
- MD5 com Padrão de Criptografia Avançado (AES).

Selecione um algoritmo de criptografia que seja compatível com a metodologia de segurança usada por sua organização.

Planejamento do aplicativo

Antes de iniciar o planejamento para um aplicativo, assegure-se de entender os recursos de dados para seu aplicativo e o local desses recursos no cluster a fim de fornecer uma solução que permite que sejam manipulados corretamente se um nó falhar.

Para evitar uma falha, você deve entender completamente como o aplicativo se comporta em um ambiente de único ou vários nós. Não faça suposições sobre o desempenho do aplicativo sob condições adversas.

Use nós com ciclos de CPU suficientes e largura da banda de E/S para permitir que o aplicativo de produção seja executado na carga de pico. Lembre-se, os nós devem ter capacidade suficiente para permitir que o PowerHA SystemMirror opere.

Para planejar isso, faça uma avaliação de desempenho ou modele seu aplicativo de produção e liste os parâmetros das cargas mais pesadas esperadas. Em seguida, escolha nós para um cluster do PowerHA SystemMirror que não excederá 85% ocupado, ao executar seu aplicativo de produção.

É possível configurar diversos monitores de aplicativos para um aplicativo e direcionar o PowerHA SystemMirror a ambos:

- Monitore o término de um processo ou mais problemas sutis que afetam um aplicativo
- Tente automaticamente reiniciar o aplicativo e tome a ação apropriada (notificação ou fallover) se as tentativas de reinicialização falharem.

Esta seção explica como registrar todas as informações chave sobre seu aplicativo e comece a desenhar seu diagrama de cluster.

Lembre-se das orientações a seguir para assegurar que seus aplicativos sejam atendidos corretamente em um ambiente em cluster do PowerHA SystemMirror:

- Projete o aplicativo e seus dados para que apenas os dados residam em discos externos compartilhados. Essa organização não só impede as violações de licença de software, mas também simplifica a recuperação de falha.
- Se estiver planejando incluir aplicativos com multicamadas em grupos de recursos dependentes pai-filho em seu cluster, consulte a seção Considerações de planejamento para aplicativos com

multicamadas. Se estiver planejando usar dependências de local para manter certos aplicativos no mesmo nó ou em nós diferentes, consulte a seção Dependências do grupo de recursos.

- Grave scripts robustos para iniciar e parar o aplicativo nos nós do cluster. O script de inicialização deve especialmente poder recuperar o aplicativo de um término anormal, como uma falha de energia. Assegure-se de que execute adequadamente em um ambiente de único nó antes de incluir o software do PowerHA SystemMirror.
- Confirme os requisitos de licenciamento do aplicativo. Alguns fornecedores requerem uma licença exclusiva para cada processador que executa um aplicativo, o que significa que você deve proteger a licença para o aplicativo incorporando informações específicas do processador no aplicativo quando é instalado. Como resultado, mesmo que o software do PowerHA SystemMirror processe uma falha de nó corretamente, ele pode ser incapaz de reiniciar o aplicativo no nó de failover devido a uma restrição no número de licenças para esse aplicativo disponível no cluster. Para evitar este problema, assegure-se de ter uma licença para cada unidade de sistema no cluster que pode potencialmente executar um aplicativo.
- Assegure-se de que o aplicativo seja executado com sucesso em um ambiente de nó único. Depurar um aplicativo em um cluster é mais difícil que depurá-lo em um único processador.
- Verifique se o aplicativo usa um mecanismo de bloqueio proprietário se precisar de acesso simultâneo.

Referências relacionadas:

“Considerações de planejamento para aplicativos com multicamadas” na página 15

As configurações de negócio que usam aplicativos com multicamadas podem usar grupos de recursos dependentes pais e filhos. Por exemplo, o banco de dados deve estar on-line antes do controlador de aplicativo. Neste 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 teria de ser desativado e ativado em qualquer nó no cluster.

“Dependências do grupo de recursos” na página 65

O PowerHA SystemMirror oferece uma ampla variedade de configurações em que é possível especificar os relacionamentos entre grupos de recursos que deseja manter na inicialização, failover e fallback.

| **Planejamento para capacidade on demand**

| A Capacidade on Demand (CoD) é um dos recursos da função Resource Optimized High Availability (ROHA) que o PowerHA SystemMirror usa para gerenciar dinamicamente os recursos de hardware para aplicativos. Com a CoD, é possível ativar os processadores pré-instalados que estão inativos e não pagos como uma mudança de requisitos de recurso.

| Os recursos CoD são compostos de recursos CoD Ativo/Inativo e recursos Enterprise Pool CoD (EPCoD). Ambos esses recursos podem entregar dinamicamente os recursos complementares para seu ambiente que são usados por meio do gerenciamento DLPAR normal (alocação ou liberação de recursos para um LPAR).

| Os processadores e memória adicionais, embora fisicamente presentes, não são usados até que o PowerHA SystemMirror decida que a capacidade adicional necessária valha a pena. É possível usar a função ROHA para adquirir rápida e facilmente recursos extra para atender a cargas de trabalho de pico ou inesperadas em seu ambiente.

| O PowerHA SystemMirror integra-se às funções DLPAR, On/Off CoD e EPCoD. A coleção dessas funções é chamada de ROHA. O nó ativo é hospedado por um LPAR em um quadro com recursos permanentes suficientes. O nó de espera é hospedado por um LPAR em um quadro com os recursos mínimos permanentes e conta com ROHA para incluir dinamicamente recursos extra.

| Quando é necessário executar o aplicativo no nó de espera, o PowerHA SystemMirror usa a função ROHA. A função ROHA verifica que o nó tem recursos suficientes para executar com sucesso o aplicativo e aloca os recursos necessários. Os recursos podem ser alocados a partir das fontes a seguir:

| **Recursos provisionados On/Off CoD**

| Se o conjunto livre CEC tiver recursos insuficientes que podem ser alocados por meio de DLPAR

para o nó, a função On/Off CoD fornecerá mais recursos ao CEC. Esses recursos adicionais são incluídos no conjunto livre CEC e podem ser usados por meio de uma operação LPAR. Se o aplicativo requerer mais memória ou processador, o PowerHA SystemMirror, por meio da função ROHA, poderá alocar automaticamente esses recursos para o nó de espera.

Recursos provisionados EPCoD

Se o conjunto livre CEC tiver recursos insuficientes que podem ser alocados por meio de DLPAR para o nó, a função EPCoD fornecerá mais recursos ao CEC. Esses recursos adicionais são incluídos no conjunto livre CEC e podem ser usados por meio de uma operação DLPAR. Se o aplicativo requerer mais memória ou processador, o PowerHA SystemMirror, por meio da função ROHA, poderá alocar automaticamente esses recursos para o nó de espera.

O conjunto livre CEC

A função DLPAR fornece os recursos para o nó de espera, alocando os recursos, que estão disponíveis no conjunto livre. Esses recursos são fornecidos a partir do conjunto On/Off CoD ou EPCoD.

Ao configurar o PowerHA SystemMirror para usar recursos por meio da função ROHA, os nós LPAR no cluster não usam mais recursos até que os recursos sejam necessários pelo aplicativo.

O PowerHA SystemMirror não ativa nenhum recurso CoD até que o conjunto livre seja esgotado. Após o conjunto livre estar esgotado, mais recursos de hardware são ativados pelo PowerHA SystemMirror. Os recursos de hardware CoD são ativados e alocados para o LPAR dinamicamente até que os requisitos do aplicativo sejam atendidos. Quando um aplicativo não requer mais os recursos de hardware alocados, eles são liberados para o conjunto livre, ponto no qual o PowerHA SystemMirror desativa os recursos de hardware. Esses recursos de hardware retornam para o conjunto do qual foi originado (conjunto On/Off CoD ou conjunto EPCoD).

Referências relacionadas:

“Monitoramento no PowerHA SystemMirror” na página 21

A tarefa primária do PowerHA SystemMirror é reconhecer e responder a falhas. O PowerHA SystemMirror usa a infraestrutura do Cluster Aware AIX para monitorar a atividade de suas interfaces de rede, dispositivos e rótulos de IP.

Informações relacionadas:

Administrando o PowerHA SystemMirror

Controladores de aplicativos

Para colocar o aplicativo no controle do PowerHA SystemMirror, você cria um recurso de controlador de aplicativo que associa um nome definido pelo usuário aos nomes de scripts gravados especialmente para iniciar e parar o aplicativo.

Ao definir um controlador de aplicativo, o PowerHA SystemMirror pode iniciar outra instância do aplicativo no nó de controle quando uma fallover ocorre. Isso protege seu aplicativo para que não se torne um ponto único de falha.

Após definir o controlador de aplicativo, é possível incluí-lo em um grupo de recursos. Um grupo de recursos é um conjunto de recursos que você define para que o software PowerHA SystemMirror possa tratá-los como uma única unidade.

Referências relacionadas:

“Planejando grupos de recursos” na página 59

Estes tópicos descrevem como planejar grupos de recursos em um cluster do PowerHA SystemMirror.

Aplicativos integrados ao PowerHA SystemMirror

Certos aplicativos, incluindo o Workload Manager, podem ser configurados diretamente como recursos altamente disponíveis, sem controladores de aplicativos ou scripts adicionais. Além disso, a verificação do cluster do PowerHA SystemMirror assegura a correção e consistência de certos aspectos da configuração do seu Workload Manager.

O PowerHA SystemMirror oferece os aplicativos a seguir do PowerHA SystemMirror Smart Assist para auxiliar a integrar o aplicativo em um cluster do PowerHA SystemMirror:

Smart Assist for WebSphere

Estende uma configuração existente do PowerHA SystemMirror para incluir o monitoramento e o suporte de recuperação para vários componentes do WebSphere.

Smart Assist for DB2

Estende uma configuração existente do PowerHA SystemMirror para incluir o monitoramento e o suporte de recuperação para o DB2 Universal Database (UDB) Enterprise Server Edition.

Smart Assist for Oracle

Fornece assistência aos envolvidos na instalação da solução Oracle Application Server 10g (9.0.4) (AS10g) Cold Failover Cluster (CFC) em um sistema operacional IBM AIX.

Smart Assist for FileNet P8

Oferece a escalabilidade e a flexibilidade de nível corporativo para manipular os desafios de conteúdo de maior demanda, os processos de negócios mais complexos e a integração aos sistemas existentes em seu ambiente.

Smart Assist para SAP MaxDB

Configure instâncias de banco de dados MaxDB e liveCache para alta disponibilidade.

Smart Assist for Lotus Domino Server

Configura automaticamente o PowerHA SystemMirror para o ambiente que já tem o Lotus Domino configurado.

Smart Assist for Tivoli Storage Manager

Usa as três áreas diferentes do Tivoli Storage Manager, servidor, cliente e centro de administração para fornecer uma solução altamente disponível para seu ambiente.

Smart Assist for SAP

Configura o SAP Netweaver 2004s para alta disponibilidade protegendo seu ponto único de falha

Smart Assist for Tivoli Directory Server

Configure automaticamente o PowerHA SystemMirror em que o Tivoli Directory Server já está instalado.

Smart Assist for SAP liveCache

Fornece interfaces de gerenciamento que auxiliam na configuração de uma política do PowerHA SystemMirror e na implementação de métodos de início, métodos de parada e métodos de monitoramento para as pilhas de carga de trabalho em seu ambiente.

Smart Assist for Websphere MQSeries

Permite que os programas se comuniquem entre si em uma rede de componentes que não são semelhantes, como processadores, subsistemas, sistemas operacionais e protocolos de comunicação.

Informações relacionadas:

Aplicativos Smart Assist para o PowerHA SystemMirror

Monitoramento de aplicativo

O PowerHA SystemMirror pode monitorar os aplicativos definidos para os controladores de aplicativos.

O PowerHA SystemMirror monitora aplicativos de uma das duas formas:

- O *Monitoramento de processo* detecta o término de um processo, usando os recursos Reliable Scalable Cluster Technology (RSCT) e Resource Monitoring and Control (RMC).
- O *Monitoramento customizado* monitora o funcionamento de um aplicativo, usando um método de monitor que você definir.

É possível configurar vários monitores de aplicativo e associá-los a um ou mais controladores de aplicativo. É possível designar a cada monitorar um nome exclusivo no SMIT. 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. Caso contrário, é possível configurar um monitor customizado para verificar o funcionamento do banco de dados junto com um processo de parar o monitor para detectar instantaneamente o fim do processo do banco de dados.

É possível usar a ferramenta Análise de disponibilidade do aplicativo para medir a quantia exata de tempo que qualquer um dos aplicativos definidos pelo PowerHA SystemMirror está disponível. O software PowerHA SystemMirror coleta, registra a data e a hora e registra as informações a seguir:

- Um monitor de aplicativo é definido, alterado ou removido
- Um aplicativo é iniciado, interrompido ou falha
- Um nó falha ou é encerrado ou iniciado
- Um grupo de recursos é colocado off-line ou movido
- O monitoramento de aplicativo por meio de diversos monitores é suspenso ou retomado.

Informações relacionadas:

Configurando a topologia de cluster e os recursos (estendidos) do PowerHA SystemMirror

Monitorando um cluster PowerHA SystemMirror

Considerações de planejamento para aplicativos com multicamadas

As configurações de negócio que usam aplicativos com multicamadas podem usar grupos de recursos dependentes pais e filhos. Por exemplo, o banco de dados deve estar on-line antes do controlador de aplicativo. Neste 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 teria de ser desativado e ativado em qualquer nó no cluster.

Os ambientes como pontos de acesso de serviço (SAP) requerem que os aplicativos sejam passem por um novo ciclo (interrompidos e reiniciados) sempre que um banco de dados falhar. Muitos serviços de aplicativos são fornecidos por um ambiente como o SAP e os componentes de aplicativos individuais geralmente precisam ser controlados em uma ordem específica.

Estabelecer interdependências entre grupos de recursos também é útil quando se exige que os serviços do sistema suportem ambientes de aplicativos. Serviços como tarefas **cron** para remoção de arquivos de log ou início de backups precisam mover de um nó para outro junto com um aplicativo, mas geralmente não são iniciados até que o aplicativo seja estabelecido. Esses serviços podem ser construídos em scripts de início e parada do controlador de aplicativo ou podem ser controlados por meio do processamento de pré-evento e pós-evento. Entretanto, grupos de recursos dependentes simplificam a forma de configurar serviços do sistema para que sejam dependentes dos aplicativos a que atendem.

Nota: Para minimizar a chance de perda de dados durante o processo de parada e reinício do aplicativo, customize seus scripts do controlador de aplicativo para assegurar que quaisquer dados não confirmados sejam armazenados em um disco compartilhado temporariamente durante o processo de parada do aplicativo e lidos de volta no aplicativo durante o processo de reinício do aplicativo. É importante usar um disco compartilhado porque o aplicativo pode ser reiniciado em um nó que não aquele no qual ele foi parado.

Também é possível configurar grupos de recursos com dependências de local para que certos grupos de recursos sejam mantidos on-line no mesmo nó ou em diferentes nós na inicialização, fallover e fallback.

Referências relacionadas:

“Planejando grupos de recursos” na página 59

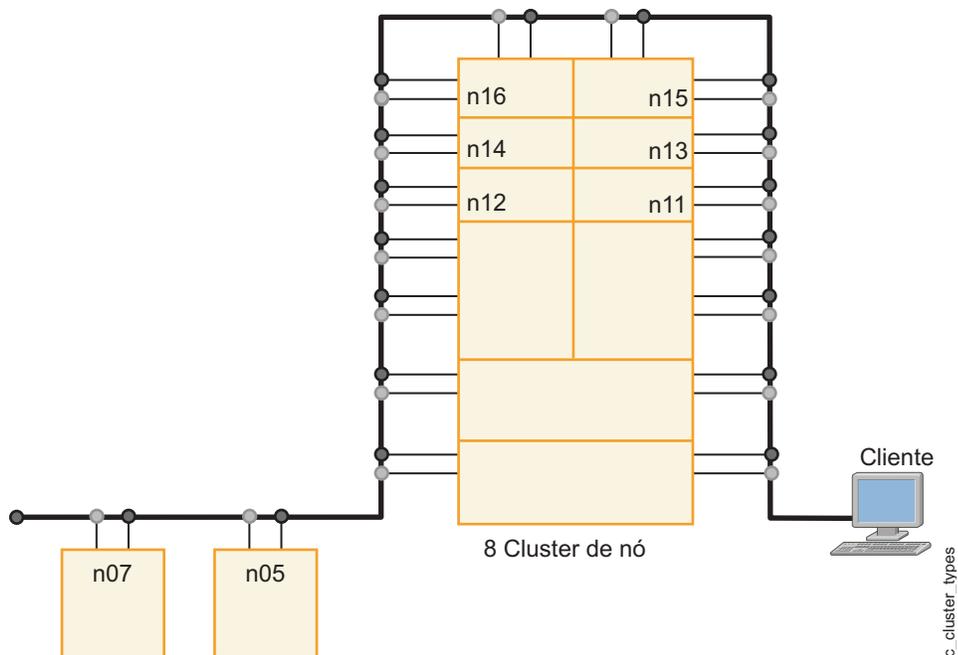
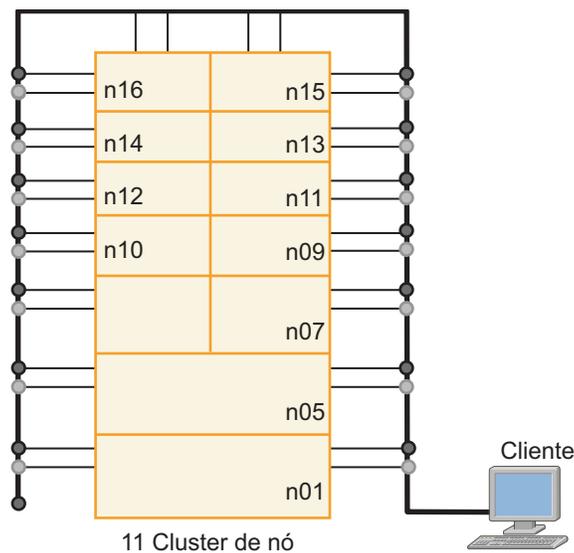
Estes tópicos descrevem como planejar grupos de recursos em um cluster do PowerHA SystemMirror.

Projetando um diagrama de cluster

O diagrama de cluster combina as informações de cada etapa no processo de planejamento em um projeto que mostra a função e a estrutura do cluster.

A ilustração a seguir mostra um cluster combinado que inclui um sistema montado em rack e sistemas independentes. O diagrama usa caixas retangulares para representar os slots suportados pelos nós. Se seu cluster usar nós thin, escureça as bordas dos nós e inclua dois nós em um desenho. Para nós amplos, use o desenho todo. Para nós altos, use o equivalente de dois nós amplos. Lembre-se de que cada nó thin contém uma conexão Ethernet integrada.

Comece projetando este diagrama identificando o nome do cluster e os aplicativos que estão se tornando altamente disponíveis. Em seguida, escureça a descrição dos nós que comporão o cluster. Inclua o nome de cada nó.



Requisitos de nome do host

A partir do PowerHA SystemMirror 7.1.1, há certos requisitos para os quais a interface pode ser o nome do host, devido aos novos requisitos de camada do Cluster Aware AIX (CAA).

Lembre-se dos requisitos a seguir ao selecionar o nome do host:

- O nome do host não pode ser um alias no arquivo `/etc/hosts`.
- A resolução do nome para o nome do host deve funcionar de ambas as formas. Portanto, apenas um conjunto limitado de caracteres pode ser usado.
- O endereço IP que pertence ao nome do host deve ser acessível no servidor, mesmo quando o PowerHA está no estado **DOWN**.
- O nome do host não pode ser um endereço de serviço.
- O nome do host não pode ser um endereço que está localizado em uma rede, que é definida como *privada* no PowerHA.
- O nome do host, o nome do nó CAA e o `COMMUNICATION_PATH` (isto é, o caminho de comunicação para o nó) deve ser o mesmo.
- Por padrão, o nome do nó do PowerHA, o nome do nó CAA e o `COMMUNICATION_PATH` (isto é, o caminho de comunicação para o nó) estão configurado para serem os mesmos.
- O nome do host e o nome do nó do PowerHA podem ser diferentes.
- O nome do host não pode ser alterado após a configuração de cluster ser concluída.

Nota: Esses requisitos deixam os endereços de base e o endereço persistente como candidatos para o nome do host. É possível usar o endereço persistente como o nome do host apenas se você configurar o alias persistente manualmente antes de configurar a topologia de cluster.

Planejando a conectividade de rede do cluster

Esta seção descreve o planejamento do suporte de rede para um cluster do PowerHA SystemMirror.

Pré-requisitos

No tópico Planejamento inicial do cluster, você começou a planejar seu cluster, identificando o número de nós e os principais aplicativos que deseja tornar altamente disponíveis. Você iniciou o desenho do diagrama de cluster. Esse diagrama é o ponto de início para o planejamento que será feito nessa seção.

Além disso, agora você deve ter decidido se usará ou não o controle de endereço IP (IPAT) para manter endereços IP de serviço específicos.

Visão Geral

Seu objetivo principal é usar a redundância para projetar uma topologia de cluster que elimina componentes de rede como pontos únicos de falha em potencial.

A tabela a seguir lista esses componentes de rede com soluções:

Objeto de rede de cluster	Eliminado como um ponto único de falha por...
Rede	Usando diversas redes para conectar nós
Placa da interface de rede (NIC)	Usando NICs redundantes em cada rede

Nesta seção, você conclui as tarefas de planejamento a seguir:

- Projetar a topologia de rede do cluster, isto é, a combinação de redes de IP que conecta seus nós de cluster e o número de conexões que cada nó tem a cada rede.

- Incluir rede em seu diagrama de cluster.

Referências relacionadas:

“Planejamento de cluster inicial” na página 5

Esta seção descreve as etapas iniciais que você executa para planejar um cluster do PowerHA SystemMirror para tornar os aplicativos altamente disponíveis.

“Monitoramento no PowerHA SystemMirror” na página 21

A tarefa primária do PowerHA SystemMirror é reconhecer e responder a falhas. O PowerHA SystemMirror usa a infraestrutura do Cluster Aware AIX para monitorar a atividade de suas interfaces de rede, dispositivos e rótulos de IP.

Considerações gerais da rede para o PowerHA SystemMirror

O PowerHA SystemMirror permite a comunicação entre nós com redes Ethernet.

Aliases de IP

Um alias de IP é um rótulo ou endereço IP que é configurado em um controlador de interface de rede (NIC) além do rótulo ou endereço IP configurado normalmente no NIC. O uso de aliases de IP é uma função do AIX que o PowerHA SystemMirror suporta. O AIX suporta diversos aliases de IP em um NIC. Cada alias de IP em um NIC pode estar em uma sub-rede separada. O AIX também permite que os aliases de IP com diferentes máscaras de sub-rede sejam configurados para uma interface. O PowerHA SystemMirror ainda não suporta esta função.

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

Conexões de rede

O PowerHA SystemMirror requer que cada nó no cluster tenha pelo menos uma conexão direta não roteada com um nó sim, um não. O software usa essas conexões de rede para transmitir mensagens de pulsação entre os nós de cluster para determinar o estado de todos os nós de cluster, redes e interfaces de rede.

O PowerHA SystemMirror requer que todas as interfaces de comunicação para uma determinada rede de cluster sejam definidas na mesma rede física e encaminhe pacotes entre si. Por padrão, o PowerHA SystemMirror usa comunicações unicast para pulsação. Se escolher usar a pulsação multicast no lugar, é necessário assegurar que sua rede suporte multicast. As interfaces de comunicação também devem ser capazes de receber respostas uma da outra sem interferência de nenhum equipamento de rede. O PowerHA SystemMirror também requer que todos os nós em um site tenham pelo menos uma conexão de rede direta com um nó sim, um não no mesmo site.

As comunicações unicast são sempre usadas entre sites em um cluster vinculado. Em um site, é possível selecionar comunicações unicast (padrão) ou multicast.

Entre os nós de cluster, coloque apenas comutadores inteligentes, roteadores ou outro equipamento de rede que passem de forma transparente por meio de multicast e outros pacotes a todos os nós de cluster. Este requisito inclui equipamento que otimiza protocolos.

Se esse equipamento for colocado nos caminhos entre nós de cluster e clientes, é possível que precise configurar uma lista de cliente ping no arquivo **clinfo.rc** para ajudar a informar clientes de movimentações de endereço IP. A topologia de rede específica pode requerer que outras soluções assegurem que os clientes possam continuar a acessar o servidor após o controle de endereço IP.

Pontes, hubs e outros dispositivos passivos que não modificam o fluxo do pacote podem ser colocados com segurança entre os nós do cluster e entres nós e clientes.

Informações relacionadas:

Programando aplicativos cliente para a API Clinfo

Rótulos de IP

Em um ambiente que não está usando o PowerHA SystemMirror, um nome do host geralmente identifica um sistema, com o nome do host também sendo o rótulo de IP de uma das interfaces de rede no sistema. Assim, um sistema pode ser acessado usando seu nome do host como o rótulo de IP para uma conexão.

Resolução de nome do host

No PowerHA SystemMirror, todos os nomes de host devem ser resolvidos com o arquivo `/etc/hosts`. Ao definir nós para o cluster, você deve especificar um endereço IP ou rótulo que resolve localmente para o nome do host e após ter sincronizado a configuração inicial do cluster, o nome do host do nó pode não ser alterado.

Rótulos de IP em redes TCP/IP

Para redes TCP/IP, um rótulo de IP e seu endereço IP associado devem aparecer no arquivo `/etc/hosts`.

O nome do rótulo de IP de serviço ou endereço deve ser exclusivo no cluster e distinto do grupo de volumes e nomes do grupo de recursos. Ele deve estar relacionado ao aplicativo por ele atendido, bem como a qualquer dispositivo correspondente, como `websphere_service_address`.

Ao designar um rótulo de IP de serviço a uma interface, use uma convenção de nomenclatura que ajude a identificar a função da interface no cluster. As entradas associadas no arquivo `/etc/hosts` seriam semelhantes ao seguinte:

```
100.100.50.1 net1_en0
100.100.60.1 net2_en1
```

Você configura o controlador de interface de rede (NIC) seguindo as instruções na documentação relevante do AIX. O AIX designa um nome da interface ao NIC quando é configurado. O nome da interface é composto de 2 ou 3 caracteres que indicam o tipo de NIC, seguido por um número que o AIX designa em sequência para cada adaptador de um certo tipo. Por exemplo, o AIX designa um nome da interface como `en0` para o primeiro NIC Ethernet que configura, `en1` para o segundo e assim por diante.

Informações relacionadas:

Configurando eventos de cluster

Particionamento de cluster

O particionamento, também chamado de isolamento do nó, ocorre quando uma falha de rede ou do controlador de interface de rede (NIC) isola nós do cluster entre si.

Quando um nó do PowerHA SystemMirror para de receber tráfego de rede de outro nó, presume-se que o outro nó tenha falhado. Dependendo da sua configuração do PowerHA SystemMirror, o nó pode começar a adquirir discos do nó com falha e tornar disponíveis aplicativos e rótulos de IP. Se o nó com falha ainda estiver realmente ativo, a distorção de dados poderá ocorrer quando os discos forem retirados dele. Se a rede se tornar disponível novamente, o PowerHA SystemMirror para um dos nós para impedir a contenção de disco adicional e endereços IP duplicados na rede.

O mecanismo de pulsação do PowerHA SystemMirror conta com o subsistema de IP e a infraestrutura de rede. Portanto, se a rede estiver congestionada ou um nó estiver congestionado, o subsistema de IP poderá descartar silenciosamente as pulsações. São feitas tentativas de ajustar características de monitoramento para levar em consideração o congestionamento de rede e evitar o particionamento de cluster.

Referências relacionadas:

“Monitorando clusters” na página 29

A infraestrutura do Cluster Aware AIX monitora todas as interfaces de rede e armazenamento suportadas e disponíveis. Os gerenciadores de cluster nos nós de cluster também enviam mensagens um para o outro por meio de conexões entre essas interfaces.

Exemplo: conexão de rede geral

Um exemplo da rede correta do PowerHA SystemMirror consiste em duas redes Ethernet separadas, cada uma com duas interfaces de rede em cada nó.

Dois roteadores conectam as redes e roteiam pacotes entre o cluster e os clientes, mas não entre as duas redes. Um arquivo `clinfo.rc` é instalado em cada nó no cluster, contendo os endereços IP de vários sistemas do cliente.

Configuração do PowerHA SystemMirror em redes alternadas

Se as redes e os comutadores forem definidos ou configurados incorretamente, os eventos de falha de interface de rede inesperados poderão ocorrer em configurações do PowerHA SystemMirror que usam redes alternadas.

Siga estas orientações ao configurar redes alternadas:

- **Virtual local area network (VLAN).** Se as VLANs forem usadas, todas as interfaces em uma determinada rede deverão ser configuradas na mesma VLAN (uma rede por VLAN). Se estiver usando o multicast para a pulsação, todas as interfaces e dispositivos de rede devem ser capazes de transportar pacotes de multicast entre nós.
- **Configurações de negociação automática.** Algumas placas de interface de rede Ethernet (NIC) são capazes de negociar automaticamente sua velocidade e outras características, como metade duplex ou duplex integral. Configure o NIC não para usar a **autonegociação**, mas para executar na velocidade desejada e valor de duplex. Configure a porta do comutador para a qual o NIC está conectado para a mesma velocidade fixada e valor de duplex.

PowerHA SystemMirror e Ethernet virtual

O PowerHA SystemMirror suporta Ethernet virtual fornecido pelo Virtual I/O Server (VIOS) ou recursos integrados do Ethernet virtual (IVE), com as APARs aplicáveis instaladas. O suporte ao PowerHA SystemMirror é idêntico para VIOS e IVE.

O utilitário PCI Hot Plug no PowerHA SystemMirror não é aplicável a interfaces em uma Ethernet virtual. Esse utilitário apenas manipula placas de interface física. Como o Ethernet virtual usa adaptadores virtuais de E/S, não é possível usar o utilitário.

A lista a seguir contém considerações adicionais para o PowerHA SystemMirror com a Ethernet virtual:

- Se o VIOS tiver diversas interfaces físicas definidas na mesma rede ou se houver dois ou mais nós do PowerHA SystemMirror usando o VIOS no mesmo quadro, o PowerHA SystemMirror não será informado de falhas de interface física única (e portanto, não reagirão). Isso não limita a disponibilidade de todo o cluster, porque o VIOS roteia o tráfego em torno da falha. O suporte ao VIOS é análogo ao EtherChannel em relação a isso. Use os métodos que não são baseados em VIOS para fornecer notificação de falhas de interface física individual.
- Se o VIOS tiver apenas uma interface física única em uma rede, o PowerHA SystemMirror detectará uma falha dessa interface física. Entretanto, a falha isolará o nó da rede.

Nota: No VIOS 2.2.0.11, ou posterior, é possível usar a comunicação de rede de área de armazenamento (SAN) entre 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.

- Em um ambiente VIOS, a falha do adaptador de rede física e dos componentes de rede fora da rede virtualizada podem não ser detectados de forma confiável. Para detectar falhas de rede externa, você deve configurar o arquivo `netmon.cf` com um ou mais endereços fora da rede virtualizada.

Resolução de problemas de conexões de Ethernet virtual

Para solucionar problemas de interfaces de Ethernet virtual definidas para o PowerHA SystemMirror e detectar uma falha de interface, tratar essas interfaces como interfaces definidas em redes de adaptador único.

Nota: Para Ethernet, o PowerHA suporta qualquer combinação de adaptadores virtuais e físicos no mesmo nome da rede.

Em particular, liste as interfaces de rede que pertencem a uma VLAN na variável `etc/cluster/ping_client_list` ou `PING_CLIENT_LIST` no script `/usr/es/sbin/cluster/etc/clinfo.rc` e execute o programa `clinfo`. Desta forma, sempre que um evento de cluster ocorrer, o programa `clinfo` monitorará e detectará uma falha de interfaces de rede listadas. Devido à natureza da VLAN, outros mecanismos para detectar a falha das interfaces de rede não são efetivos.

Conceitos relacionados:

“Redes virtuais no PowerHA SystemMirror” na página 31

No PowerHA SystemMirror Versão 7.1.0, ou posterior, o monitor do adaptador fornecido pelo Cluster Aware AIX (CAA) não pode identificar sempre se um adaptador virtual perdeu seu adaptador físico correspondente.

Monitoramento no PowerHA SystemMirror

A tarefa primária do PowerHA SystemMirror é reconhecer e responder a falhas. O PowerHA SystemMirror usa a infraestrutura do Cluster Aware AIX para monitorar a atividade de suas interfaces de rede, dispositivos e rótulos de IP.

As conexões de monitoramento são necessárias porque permitem que o PowerHA SystemMirror reconheçam a diferença entre uma falha de rede e uma falha de nó. Por exemplo, se a conectividade na rede do PowerHA SystemMirror (os rótulos de IP desta rede são usados em um grupo de recursos) for perdida e você tiver outra rede baseada em TCP/IP, o PowerHA SystemMirror reconhecerá a falha de sua rede de cluster e tomará as ações de recuperação que impedem o cluster de ser particionado.

Para evitar o particionamento de cluster, você deve configurar redes redundantes no cluster do PowerHA SystemMirror.

O PowerHA SystemMirror monitora automaticamente as interfaces nos componentes a seguir:

- Redes TCP/IP
- Redes de área de armazenamento
- Discos de repositório

Projetando a topologia de rede

A combinação de redes IP e não IP (ponto a ponto) que vinculam nós de cluster e clientes é chamada de *topologia de rede* de cluster. O software PowerHA SystemMirror suporta um grande número de dispositivos IP e ponto a ponto em cada nó para fornecer a flexibilidade no projeto de uma configuração de rede.

Ao projetar sua topologia de rede, assegure-se de que os clientes tenham acesso à rede altamente disponível para seus aplicativos. Isso requer que nenhuma das interfaces de rede a seguir sejam um ponto único de falha:

- O subsistema IP
- Uma rede única
- Um NIC único

Eliminando redes como pontos únicos de falha

Em uma configuração de rede única, cada nó no cluster está conectado a apenas uma rede e possui uma interface de serviço única disponível aos clientes. Nessa configuração, a rede é um ponto único de falha para todo o cluster e cada interface de serviço é um ponto único de falha.

O diagrama a seguir mostra uma configuração de rede única.

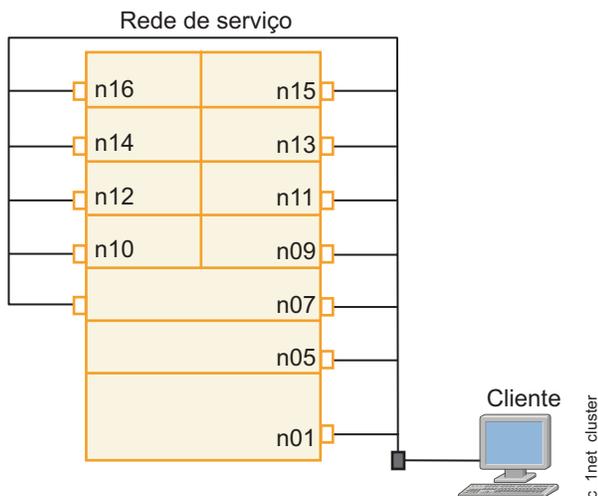


Figura 1. Configuração de rede única e NIC único

Para eliminar a rede como um ponto único de falha, configure diversas redes para que o PowerHA SystemMirror tenha diversos caminhos entre os nós do cluster. Lembre-se de que se um cliente estiver conectado a apenas uma rede, essa rede será um ponto único de falha para o cliente. Em uma configuração de rede múltipla se uma rede falhar, as redes restantes ainda poderão funcionar para conectar os nós e fornecer acesso para os clientes.

Quanto mais redes puder configurar para transportar pulsações e outras informações entre os nós do cluster, maior o grau de disponibilidade do sistema.

O diagrama a seguir ilustra uma configuração de rede dupla com mais de um caminho para cada nó do cluster.

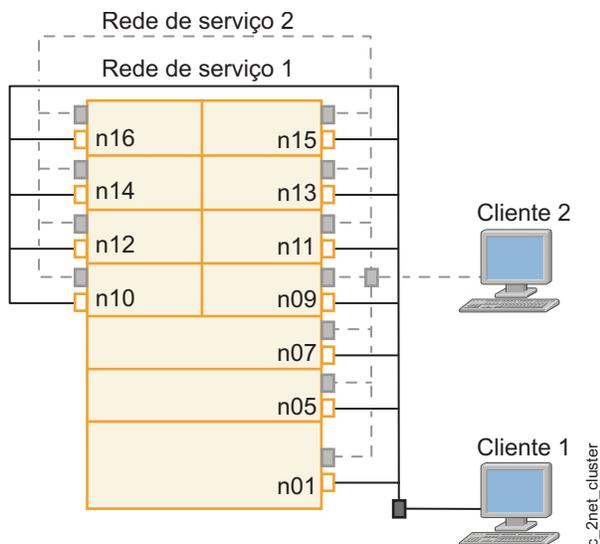


Figura 2. Configuração de rede dupla

Nota: A substituição automática do adaptador Ethernet de porta dupla usada para configurar duas interfaces para uma rede IP do PowerHA SystemMirror não é suportada atualmente.

Eliminando placas da interface de rede como um ponto único de falha

Uma placa da interface de rede (NIC) conecta fisicamente um nó a uma rede.

Quando configurado com uma única NIC por rede, a NIC se torna o ponto único de falha em potencial. Para corrigir este problema, configure um nó com pelo menos duas NICs para cada rede aos quais ele se conecta. Na figura a seguir, cada nó do cluster possui duas conexões com cada rede.

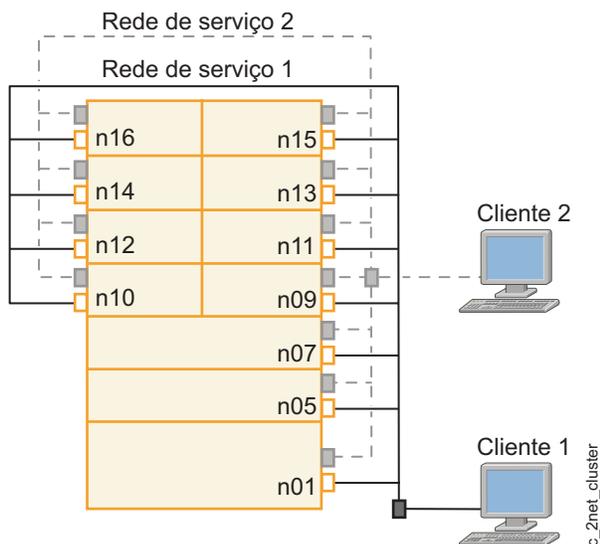


Figura 3. Configuração de rede dupla e NIC dupla

Nota: A substituição automática do adaptador Ethernet de porta dupla usada para configurar duas interfaces para uma rede IP do PowerHA SystemMirror não é suportada atualmente.

Informações relacionadas:

Comportamento do grupo de recursos durante eventos de cluster

Funções de interface de rede:

Quando um nó é configurado com diversas conexões com uma única rede, as interfaces de rede atendem a diferentes funções no PowerHA SystemMirror.

Interface de serviço

Uma interface de serviço é uma interface de rede configurada com um rótulo de IP de serviço do PowerHA SystemMirror. O rótulo de IP de serviço é usado por clientes para acessar os programas de aplicativos. O IP de serviço só está disponível quando o grupo de recursos correspondente está on-line.

Rótulo de IP de nó persistente

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. Um rótulo de IP de nó persistente sempre fica no mesmo nó (limite de nó) e coexiste em um NIC que já possui um serviço ou rótulo de IP de inicialização definido. Um rótulo de IP de nó persistente não requer a instalação de um NIC físico adicional nesse nó e não é parte de nenhum grupo de recurso.

Designar um rótulo de IP de nó persistente fornece um endereço de limite de nó que pode ser usado para propósitos administrativos, porque uma conexão com o rótulo de IP do nó persistente sempre vai para um nó específico no cluster. Você pode ter um rótulo de IP de nó persistente por rede por nó.

Como uma das melhores práticas para o PowerHA SystemMirror, você deve configurar um rótulo de IP persistente para cada nó do cluster. Isso é útil, por exemplo, se você deve acessar um nó específico em um cluster do PowerHA SystemMirror para fins de execução de relatórios ou diagnósticos. Ter um rótulo de IP persistente configurado tem a vantagem de que o PowerHA SystemMirror pode acessar o rótulo de IP persistente no nó apesar das falhas de NIC individuais, presumindo que haja NICs sobressalentes na rede.

Após um rótulo de IP de nó persistente ser configurado em um nó de rede especificado, ele se torna disponível no tempo de inicialização e permanece configurado mesmo se o PowerHA SystemMirror for encerrado nesse nó.

É possível criar rótulos de IP de nó persistente nas redes Ethernet.

A lista a seguir descreve como o PowerHA SystemMirror responde a uma falha quando um rótulo de nó persistente é configurado:

- Se um NIC que possui um rótulo de IP de serviço configurado falhar e também houver um rótulo persistente nesse NIC, o rótulo persistente cairá para mesma interface de inicialização à qual o rótulo de IP de serviço cai.
- Se todos os NICs na rede do cluster em um nó especificado falharem, o rótulo de IP de nó persistente se tornará indisponível. Um rótulo de IP de nó persistente sempre permanece na mesma rede e no mesmo nó. Ele não se move entre os nós no cluster.

Informações relacionadas:

Configurando a topologia de cluster e os recursos (estendidos) do PowerHA SystemMirror

Controle de endereço IP por meio de aliases de IP:

O PowerHA SystemMirror usa o IPAT via aliases de IP para manter os endereços IP de serviço altamente disponíveis.

Quando o PowerHA SystemMirror é iniciado no nó, o rótulo de IP do serviço tem o alias executado em uma das interfaces de inicialização que está definida para o PowerHA SystemMirror. Se essa interface

falhar, o rótulo de IP de serviço que tem o alias executado em outra interface se uma estiver disponível na mesma rede. Para usar o IPAT vi aliases de IP, a rede deve suportar ARP gratuito.

Considerações de sub-rede para rótulos de IP de nó persistente

Ao configurar um rótulo de IP de nó persistente em uma rede de cluster, o endereço IP associado ao rótulo de IP persistente deve estar em uma sub-rede diferente de qualquer endereço de serviço com o qual pode compartilhar uma interface. As redes com uma interface por nó não requerem sub-redes separadas.

Em algumas situações, é possível que precise configurar um rótulo de IP persistente na mesma sub-rede que o rótulo de IP do serviço. Neste caso, para evitar problemas com pacotes de rede enviados de um dos endereços, considere configurar a preferência de distribuição para aliases de IP de serviço. Esta preferência permite configurar o tipo de preferência de distribuição adequado para os requisitos de conectividade externa do firewall da VPN.

Nota: As considerações de sub-rede são diferentes se você estiver planejando configurar um Network File System (NFS).

Referências relacionadas:

“Montagem cruzada do NFS e rótulos IP” na página 55

Para permitir a montagem cruzada do NFS, cada nó do cluster pode atuar como um cliente NFS. Cada um desses nós deve ter uma rota válida para o rótulo de IP de serviço do nó do servidor NFS. Isto é, para permitir a montagem cruzada do NFS, um rótulo IP deve existir nos nós cliente e esse rótulo de IP deve ser configurado na mesma sub-rede que o rótulo de IP do serviço do nó do servidor NFS.

“Tipos de distribuição para aliases de rótulo de IP de serviço” na página 27

É possível especificar no SMIT diferentes preferências de distribuição para a disposição de aliases de rótulo de IP de serviço.

“Planejamento para controle de endereço IP via aliases de IP”

A designação de aliases de IP a NICs permite criar mais de um rótulo de IP na mesma interface de rede.

Planejamento para controle de endereço IP via aliases de IP

A designação de aliases de IP a NICs permite criar mais de um rótulo de IP na mesma interface de rede.

Durante o controle de endereço IP via aliases de IP, quando um rótulo de IP se move de um NIC para outro, o NIC de destino recebe o novo rótulo de IP como um alias de IP e mantém o rótulo de IP original e o endereço de hardware.

Configurar redes para IPAT por aliases de IP simplifica a configuração de rede no PowerHA SystemMirror. Você configura um endereço de serviço e um ou mais endereços de inicialização para NICs.

Designando rótulos de IP para IPAT via aliases de IP

O PowerHA SystemMirror usa uma tecnologia mencionada como controle de endereço IP via aliases de IP para manter os endereços IP altamente disponíveis.

Revise as informações a seguir ao planejar o controle do endereço IP via aliases de IP:

- Cada interface de rede deve ter um rótulo de IP de inicialização definido para o PowerHA SystemMirror. As interfaces que são definidas para o PowerHA SystemMirror são usadas para manter os endereços IP do serviço altamente disponíveis.
- O Hardware Address Takeover (HWAT) não pode ser configurado para redes que estão usando o controle de endereço IP via aliasing de IP.
- Os requisitos de sub-rede a seguir se aplicam quando há diversas interfaces em um nó anexas à mesma rede:
 - Todos os endereços de inicialização devem ser definidos em diferentes sub-redes.

- Os endereços de serviço devem estar em uma sub-rede diferente de todos os endereços de inicialização e endereços persistentes.

Nota: Esses requisitos de sub-rede evitam a função de striping da rota de IP do sistema operacional AIX, o que permite que diversas rotas para a mesma sub-rede e possa fazer com que o tráfego do aplicativo seja enviado a uma interface com falha. Esses requisitos não se aplicam a diversas interfaces que são combinadas em uma interface lógica única usando a Agregação de Ethernet ou EtherChannel.

- Os rótulos de endereço de serviço configurados para controle de endereço IP via aliases de IP podem ser incluídos em todos os grupos de recursos não simultâneos.
- Diversos rótulos de serviço podem coexistir como aliases em uma determinada interface.
- A máscara de rede para todos os rótulos de IP em uma rede do PowerHA SystemMirror deve ser a mesma.
- Se houver diversos serviços e rótulos persistentes, o PowerHA SystemMirror tentará distribuí-los igualmente entre todas as interfaces de rede disponíveis. É possível especificar uma preferência de local, de modo que os aliases persistentes e de serviço sejam sempre mapeados para a mesma interface. Para obter mais informações sobre os rótulos de serviço, consulte Preferência de distribuição para aliases de rótulo de IP de serviço.

Os endereços de inicialização em um nó (o endereço de base designado pelo sistema operacional AIX após uma reinicialização do sistema e antes do software PowerHA SystemMirror ser iniciado) são definidos para o PowerHA SystemMirror como endereços de inicialização. O PowerHA SystemMirror descobre e configura automaticamente endereços de inicialização quando você configura pela primeira vez seu cluster. Quando o software PowerHA SystemMirror é iniciado em um nó, o rótulo do IP do serviço do nó é incluído como um alias em um dos NICs que está definido como um endereço de inicialização. Se o NIC que hospeda o IP do serviço falhar, o PowerHA SystemMirror tentará movê-lo para outro NIC de inicialização ativo no mesmo nó.

Durante um evento de fallover de nó, o rótulo de IP do serviço que é movido é posicionado como um alias no NIC do nó de destino além de quaisquer outros rótulos de serviço que possam já estar configurados nesse NIC.

Por exemplo, se o Nó A falhar, o PowerHA SystemMirror tentará mover todos os grupos de recursos para o Nó B. Se os grupos de recursos contiverem endereços IP do serviço, o PowerHA SystemMirror posicionará o IP do serviço como um alias no NIC apropriado no Nó B e quaisquer outros rótulos existentes permanecerão intactos no NIC do Nó B. Logo, um NIC no Nó B agora pode receber o tráfego do cliente que antigamente era direcionado ao endereço de serviço que estava no Nó A. Posteriormente, quando o Nó A é reiniciado, ele inicia em seus endereços de inicialização e está disponível para hospedar o IP do serviço se o Nó B falhar. Quando o Nó B liberar o rótulo do IP do serviço solicitado, o alias para os rótulos de IP do serviço será excluído no Nó B. O Nó A novamente coloca o rótulo de IP do serviço como um alias em uma das interfaces do endereço de inicialização na rede apropriada.

Ao usar o IPAT por meio de aliases de IP, os rótulos de IP do serviço serão adquiridos usando todas as interfaces disponíveis e apropriadas. Se diversas interfaces estiverem disponíveis para hospedar o rótulo de IP do serviço, a interface será escolhida de acordo com a política de distribuição configurada. Por padrão, uma política anticolocação é usada e uma tentativa é feita para distribuir igualmente os rótulos de IP do serviço entre todas as interfaces disponíveis.

No PowerHA SystemMirror, não há impacto em nenhum rótulo de IP do serviço com alias em uma interface se você remover o endereço de tempo de inicialização para a interface com o comando **ifconfig**. Entretanto, se você usar o comando **chdev** (ou o atalho **chinet**) para substituir um endereço de inicialização em uma interface, esse comando exclui quaisquer aliases de IP do serviço e não há indicação que isso aconteceu no PowerHA SystemMirror. Se você precisar alterar ou cancelar a designação temporariamente o endereço de inicialização de uma interface que esteja hospedando ativamente os endereços de serviço, é melhor mover os endereços de serviço para outra interface primeiro. Se nenhuma

outra interface estiver disponível, use o alias do comando **ifconfig** para substituir um endereço de inicialização em uma interface sem interromper o acesso ao aplicativo.

Referências relacionadas:

“Montagem cruzada de NFS no PowerHA SystemMirror” na página 54

Uma montagem cruzada de NFS é uma configuração de NFS do PowerHA SystemMirror específica em que cada nó no cluster pode atuar como o servidor NFS e o cliente NFS. Enquanto um sistema de arquivos está sendo exportado de um nó, o sistema de arquivos é montado com o NFS em todos os nós para o grupo de recursos, incluindo um que o esteja exportando. Outro sistema de arquivos também pode ser exportado de outro nó e pode ser montado com o NFS em todos os nós.

Planejamento para a disposição de alias de rótulo de IP de serviço

Se você usar o IPAT via aliases de IP, é possível configurar uma preferência de distribuição para a disposição de rótulos de IP que são configurados no PowerHA SystemMirror.

O PowerHA SystemMirror permite especificar a preferência de distribuição para os aliases de rótulo de IP de serviço. Esses são os rótulos de IP de serviço que são parte dos grupos de recursos do PowerHA SystemMirror e que pertencem ao IPAT via redes de aliases de IP.

Uma preferência de distribuição para aliases de rótulo de IP de serviço é um atributo de toda a rede usado para controlar a disposição de aliases de rótulo de IP de serviço nas placas de interface de rede física nos nós no cluster.

Use a preferência de distribuição para aliases de IP para abordar os requisitos de cluster a seguir:

- Considerações de firewall
- As configurações de cluster que usam VLANs (quando os aplicativos estão esperando receber pacotes de uma interface de rede específica)
- Requisitos específicos para a disposição de rótulos de IP no cluster

Preferência de distribuição para aliases de rótulo de IP de serviço

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, levando em conta 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 adequada para os requisitos de conectividade externa de firewall de 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, fallback ou 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 é exercida no cluster contanto que as interfaces de rede aceitáveis estejam disponíveis. O PowerHA SystemMirror sempre mantém os rótulos de IP de serviço ativos, mesmo que a preferência não possa ser satisfeita.

Tipos de distribuição para aliases de rótulo de IP de serviço

É possível especificar no SMIT diferentes preferências de distribuição para a disposição de aliases de rótulo de IP de serviço.

Os tipos de preferências de distribuição estão a seguir.

Tipo de preferência de distribuição	Descrição
Antidisposição	Este é o padrão. O PowerHA SystemMirror distribui todos os aliases de rótulo de IP de serviço entre todos os rótulos de IP de inicialização usando um processo de seleção menos carregado.
Disposição	O PowerHA SystemMirror aloca todos os aliases de rótulo de IP de serviço na mesma placa da interface de rede (NIC).
Antidisposição com rótulos persistentes	O PowerHA SystemMirror distribui todos os aliases de rótulo de IP de serviço entre todas as interfaces físicas ativas que não estão hospedando o rótulo de IP persistentes. O PowerHA SystemMirror coloca o alias de rótulo de IP de serviço na interface que está hospedando o rótulo persistente apenas se nenhuma outra interface de rede estiver disponível. Nota: Se você não configurou rótulos de IP persistentes, o PowerHA SystemMirror permite selecionar a antidisposição com preferência de distribuição persistente, mas emite um aviso e usa a preferência antidisposição regular por padrão.
Disposição com rótulos persistentes	Todos os aliases de rótulo de IP de serviço são alocados na mesma NIC que está hospedando o rótulo de IP persistente. Esta opção pode ser útil em configurações de firewall da VPN em que apenas uma interface recebe a conectividade externa e todos os rótulos de IP (persistente e serviço) devem ser alocados na mesma placa da interface. Nota: Se você não configurou rótulos de IP persistentes, é possível usar o PowerHA SystemMirror para selecionar a disposição com preferência de distribuição persistente, mas emite um aviso e usa a preferência de disposição regular por padrão.
Antidisposição com origem	Os rótulos de serviço são mapeados usando a preferência de antidisposição. Se não houver adaptadores suficientes, mais de um rótulo de serviço pode ser colocado em um adaptador. Com esta opção, um rótulo é escolhido como o endereço de origem para a comunicação de saída. O rótulo de interface escolhido no campo Rótulo de IP de origem para pacotes de saída é o endereço de origem.
Disposição com origem	Os rótulos de serviço são mapeados usando a preferência Disposição. Com esta opção, um rótulo de serviço é escolhido como o endereço de origem para a comunicação de saída. O rótulo de interface escolhido no campo Rótulo de IP de origem para pacotes de saída é o endereço de origem.
Antidisposição com rótulo persistente e origem	Os rótulos de serviço são mapeados usando a antidisposição com preferência persistente. Um endereço de serviço pode ser escolhido como um endereço de origem para o caso em que há mais endereços de serviço que adaptadores de inicialização.

As regras a seguir se aplicam à preferência de distribuição:

- Se houver interfaces insuficientes disponíveis para atender à preferência, o PowerHA SystemMirror aloca aliases de rótulo de IP de serviço e rótulos de IP persistentes para uma placa da interface de rede ativo existente.
- Se você não configurou rótulos persistentes, o PowerHA SystemMirror permite selecionar a disposição com persistente e antidisposição com preferência de distribuição persistentes, mas emite um aviso e usa a disposição regular ou preferências antidisposição por padrão.
- É possível alterar a preferência de distribuição de rótulos de IP dinamicamente: a nova seleção torna-se ativa durante os eventos do cluster subsequentes. O PowerHA SystemMirror não interrompe o processamento realocando os rótulos de IP do serviço ativo atualmente no momento que a preferência é alterada.

Quando um rótulo de IP de serviço falha e outro está disponível no mesmo nó, o PowerHA SystemMirror recupera os aliases de rótulo de IP de serviço movendo-os para outra NIC no mesmo nó. Durante este evento, a preferência de distribuição especificada permanece em vigor.

Informações relacionadas:

Administrando o PowerHA SystemMirror

Planejamento para rótulos de IP de serviço específicos ao site

É possível ter um rótulo de IP de serviço configurável em diversos nós e estar associado a um grupo de recursos que possa mover-se entre nós ou sites.

Nota: Os sites são suportados apenas no PowerHA SystemMirror 7.1.2 ou posterior, em ambas as Enterprise Edition e Standard Edition. O gerenciamento de replicação é suportado apenas no PowerHA SystemMirror Enterprise Edition.

Um endereço IP que é válido em um site pode não ser válido no outro site devido a problemas de sub-rede, é possível associar o rótulo de IP de serviço configurável em diversos nós com um site específico. Os rótulos de IP de serviço específicos ao site são configurados no PowerHA SystemMirror e podem ser usados com ou sem redes do PowerHA SystemMirror Enterprise Edition for AIX. Esse rótulo está associado a um grupo de recursos e é ativo apenas quando o grupo de recursos está em um estado Primário on-line no site associado.

Planejamento para outras condições de rede

A consideração de planejamento mais típica envolve a entrega de nomes em um ambiente do PowerHA SystemMirror e a configuração do monitoramento do cluster e a detecção de falha.

Usando o PowerHA SystemMirror com NIS e DNS

Alguns dos comandos usados para solucionar problemas de rede e interface após uma falha requerem a consulta de IP para determinar o endereço IP associado a um rótulo de IP especificado.

Se o network information service (NIS) ou servidor de nomes de domínio (DNS) estiver em operação, a consulta de IP será padronizada para um sistema de servidor de nomes para resolução de nome e endereço. Entretanto, se o servidor de nomes foi acessado por meio de uma interface que falhou, a solicitação não será concluída e, finalmente, atingirá o tempo limite. Esse tempo limite pode reduzir significativamente o processamento de eventos do PowerHA SystemMirror.

Para assegurar que o evento de cluster seja concluído com sucesso e rapidamente, o PowerHA SystemMirror desativa a resolução de nome do host de NIS ou DNS configurando a variável de ambiente AIX a seguir durante a troca de rótulo de IP de serviço:

```
NSORDER = local
```

Como resultado, o arquivo `/etc/hosts` de cada nó do cluster deve conter todos os rótulos de IP definidos do PowerHA SystemMirror para todos os nós de cluster.

Solicitações de DNS enviadas de processos que não o PowerHA SystemMirror

A desativação da resolução de nome do host de NIS ou DNS é específica para o ambiente de script do evento do PowerHA SystemMirror. O PowerHA SystemMirror configura a variável NSORDER para `local` quando anexa um rótulo de IP de serviço e quando troca rótulos de IP em uma interface.

Outros processos continuam usando as configurações de resolução de nome do sistema padrão (por exemplo, aplicativos fora do PowerHA SystemMirror que requerem resolução de endereço IP de DNS). Se esses processos solicitarem a consulta de IP, durante os eventos de reconfiguração de interface de rede no PowerHA SystemMirror, os processos ainda não podem entrar em contato com um servidor de nome externo. A solicitação para o DNS terá êxito após o PowerHA SystemMirror concluir o evento de reconfiguração de interface de rede.

Monitorando clusters

A infraestrutura do Cluster Aware AIX monitora todas as interfaces de rede e armazenamento suportadas e disponíveis. Os gerenciadores de cluster nos nós de cluster também enviam mensagens um para o outro por meio de conexões entre essas interfaces.

Todas as interfaces de rede e armazenamento disponíveis e suportadas no cluster serão usadas para monitorar interfaces, assegurar a conectividade aos peers de cluster e relatar quando uma conexão falhar.

Entre em contato com seu representante IBM para obter a lista atual de adaptadores, discos e drivers de caminhos múltiplos suportados. O PowerHA SystemMirror suporta o monitoramento e a comunicação dos tipos de interfaces a seguir:

- Ethernet
- Adaptadores Fibre Channel Emulex de 4 GB e 8 GB
- O disco de repositório (SAN e discos SAN são suportados)
- FCoE (Fibre Channel over Ethernet)

Informações relacionadas:

 Matriz de suporte de hardware PowerHA

Planejamento para configurações de rede de firewall de VPN no PowerHA SystemMirror

O PowerHA SystemMirror permite especificar a preferência de distribuição para os aliases de rótulo de IP de serviço. Esses são os rótulos de IP de serviço que são parte dos grupos de recursos do PowerHA SystemMirror e que pertencem ao IPAT via redes de aliasing de IP.

Certas configurações de firewall de VPN permitem a conectividade externa com apenas uma NIC por vez. Se seu firewall estiver configurado desta forma, aloque todos os rótulos de IP persistente e de serviço do PowerHA SystemMirror na mesma interface.

Para que o PowerHA SystemMirror gerencie os rótulos de IP para atender aos requisitos de um firewall de VPN:

- Especifique o rótulo de IP persistente para cada nó no cluster. O rótulo de IP persistente é mapeado para uma interface disponível na rede selecionada.
- Especifique a disposição com a preferência de distribuição persistente para a rede que contém os rótulos de IP de serviço. Isso assegura que todos os aliases de rótulo de IP de serviço sejam alocados na mesma interface física que está hospedando o rótulo de IP persistente.

Referências relacionadas:

“Tipos de distribuição para aliases de rótulo de IP de serviço” na página 27

É possível especificar no SMIT diferentes preferências de distribuição para a disposição de aliases de rótulo de IP de serviço.

Informações relacionadas:

Administrando o PowerHA SystemMirror

Planejamento para endereço IP versão 6 com PowerHA SystemMirror

O Protocolo da Internet versão 6 (IPv6) é suportado no PowerHA SystemMirror 7.1.2 ou posterior.

Antes de implementar IPv6 em seu ambiente, você deve considerar as áreas a seguir do PowerHA SystemMirror:

- O Cluster Aware AIX (CAA) usa automaticamente pulsações para qualquer endereço IP configurado em um nó. Para impedir o CAA de pulsar qualquer endereço IPv6 em uma determinada interface, sua rede que usa o PowerHA SystemMirror deve ser identificada como uma rede privada.
- O IPv6 usa a configuração dinâmica dos endereços de adaptador e outros atributos de rede. Por padrão, os endereços IPv6 não persistem entre uma operação de reinicialização do sistema, no entanto, é possível configurar seu sistema para executar o comando **autoconf6** durante a inicialização. É necessário planejar como e onde o comando **autoconf6** é executado em seu ambiente do sistema.
- O PowerHA SystemMirror usa endereços locais de link como endereços de inicialização. É possível configurar um segundo endereço de alias para ser usado como o endereço de inicialização do PowerHA SystemMirror. Os endereços locais de link são convenientes para usar para endereços de inicialização, porque sempre estão configurados para qualquer interface que usa IPv6. Você deve planejar os endereços de inicialização IPv6 que podem ser usados em seu ambiente.

Informações relacionadas:

Pulsção sobre redes de Internet Protocol
Persistindo endereços IPv6 entre a reinicialização do sistema

Planejando redes para comunicação entre nós com a Oracle

A Oracle usa a configuração de atributo de rede **privada** para selecionar redes para comunicações entre nós da Oracle. Esse atributo não é usado pelo PowerHA SystemMirror e não afetará o PowerHA SystemMirror de forma alguma. O atributo padrão é **público**.

Alterar o atributo de rede para **privada** torna a rede Oracle compatível, alterando todas as interfaces para o serviço.

Após criar suas redes de cluster (seja manualmente ou usando descoberta), é possível alterar o atributo de rede seguindo este caminho do SMIT:

Nós de cluster e redes > Gerenciar redes e interfaces de rede > Redes > Alterar/mostrar uma rede

Selecione a rede a ser alterada e altere a configuração Atributo de rede para **privada**. Sincronize o cluster após fazer essa mudança.

Regras para configurar redes privadas

Siga estas etapas para configurar redes privadas para uso pela Oracle:

1. Configure a rede e todas as interfaces. Não é possível alterar o atributo se a rede não tiver interfaces.
2. Altere o atributo de rede para **privado**.
3. Verifique se as redes privadas têm todas as interfaces de inicialização e serviço. Se a rede tiver todas as interfaces de inicialização (o padrão ao usar a descoberta), o PowerHA SystemMirror converterá essas interfaces para o serviço (a Oracle apenas observa interfaces de serviço).
4. Sincronize o cluster após alterar o atributo.

Nota: Após definir o atributo de rede como **privado**, não será possível alterá-lo novamente para **público**. Você tem que excluir a rede e redefini-la para PowerHA SystemMirror (padrão **público**).

Informações relacionadas:

Verificando e sincronizando um cluster do PowerHA SystemMirror

Redes virtuais no PowerHA SystemMirror

No PowerHA SystemMirror Versão 7.1.0, ou posterior, o monitor do adaptador fornecido pelo Cluster Aware AIX (CAA) não pode identificar sempre se um adaptador virtual perdeu seu adaptador físico correspondente.

Por exemplo, se um cabo de rede estiver desconectado de um Virtual I/O Server (VIOS), ele não poderá se comunicar com a rede externa. Assim, as partições do VIOS podem relatar suas interfaces virtuais individuais conforme disponível, quando não puderem atingir nenhuma LAN externa além da rede virtual. É possível corrigir este problema com a APAR IV14422. Este problema é semelhante a APAR IZ01331, que é para HACMP 6.1. Se estiver migrando do HACMP 6.1, ou anterior, e aplicou o APAR IZ01331, não será necessário alterar seu arquivo netmon.cf existente. Entretanto, você deve aplicar a APAR IV14422 após a migração.

Se você estiver configurando uma rede virtual ou uma rede Integrated Virtual Ethernet (IVE) no PowerHA SystemMirror pela primeira vez, você deve criar um arquivo netmon.cf no diretório /usr/es/sbin/cluster. O conteúdo dentro do arquivo netmon.cf deve seguir o formato usado no HACMP 6.1 e descrito na APAR IZ01331. No arquivo netmon.cf, você deve ter pelo menos uma linha para cada interface virtual usando o formato a seguir:

```
!REQD owner target
```

A lista a seguir descreve as variáveis usadas no arquivo `netmon.cf`.

!REQD

Uma sequência explícita que deve estar no início da linha sem espaço extra.

proprietário

A interface cujo status on-line ou off-line é determinada se pode realizar ping de qualquer um dos destinos especificados. O proprietário pode ser especificado como um nome do host, endereço IP ou nome da interface. Se você usar um nome do host, ele deve resolver para um endereço IP ou a linha é ignorada. É possível especificar a sequência `!ALL` para indicar que todos os adaptadores usam o destino especificado.

target O endereço IP ou nome do host que deseja que o proprietário tente realizar ping. Para usar o nome do host, o destino deve poder ser resolvido para um endereço IP.

Quando você está criando ou alterando o arquivo `netmon.cf`, considere as informações a seguir:

- Você deve selecionar os destinos que estão fora do ambiente de rede virtual.
- Os destinos que você identifica devem ser mantidos por meio das mudanças em seu ambiente de rede.
- É possível fornecer apenas um destino por linha, mas até 32 linhas diferentes no arquivo podem especificar a mesma entrada do proprietário. Para um adaptador possuído listado em mais de uma linha, o adaptador é considerado disponível se puder realizar ping de quaisquer destinos fornecidos.
- Não use destinos que estão todos no mesmo sistema físico. Além disso, não faça que todos os seus destinos sejam adaptadores do mesmo cluster do PowerHA SystemMirror. Caso contrário, qualquer nó nesse cluster não pode manter seus adaptadores disponíveis quando é o único nó on-line.
- Cada adaptador virtual deve ter pelo menos uma linha dentro do arquivo `netmon.cf` que especifica um destino que pode ter o ping executado a partir do endereço IP de inicialização nessa interface ou um alias de IP persistente se estiver configurado.
- O hardware de rede que pode ter o ping executado, como gateways e roteadores, é útil como endereços de destino porque os nós do PowerHA SystemMirror já os utilizam.

Se alguns adaptadores na mesma rede forem virtuais e outros não, é perfeitamente aceitável usar o formato **!REQD**. Para o adaptador virtual `en0` no arquivo `netmon.cf`, use o formato **!REQD**. Para o adaptador físico `en1`, é opcional incluir **!REQD** no arquivo `netmon.cf` e também é opcional usar o formato **!REQD**.

No PowerHA 7.1 (CAA), apenas as entradas **!REQD** são usadas. Se houver quaisquer outras entradas no arquivo `netmon.cf`, elas serão inúteis e ignoradas. Mas a forma como o valor **!REQD** é usado ainda é igual. Você deve poder executar ping de pelo menos um destino (se houver mais de uma linha para o mesmo adaptador).

Nota: Esse formato também se aplica a redes IVE. Entretanto, não é possível usar um destino que seja um membro da rede IVE no mesmo sistema físico.

Exemplos

Os exemplos a seguir explicam o conteúdo no arquivo `netmon.cf`:

1. Neste exemplo, o adaptador que possui `host1.ibm` só fica disponível se puder executar ping de `100.12.7.9` ou para o que o `host4.ibm` for resolvido. O adaptador que possui `100.12.7.20` só fica disponível se puder executar ping de `100.12.7.10` ou para o que o `host5.ibm` for resolvido. Se `100.12.7.20` for o endereço IP para o que o `host1.ibm` é resolvido, todos os quatro destinos pertencem a esse mesmo adaptador.

```
!REQD host1.ibm 100.12.7.9
!REQD host1.ibm host4.ibm
!REQD 100.12.7.20 100.12.7.10
!REQD 100.12.7.20 host5.ibm
```

2. Neste exemplo, todos os adaptadores ficam disponíveis apenas se puderem executar ping de endereços IP 100.12.7.9, 110.12.7.9 ou 111.100.1.10. A entrada do proprietário en1 possui um destino adicional de 9.12.11.10.

```
!REQD !ALL 100.12.7.9
!REQD !ALL 110.12.7.9
!REQD !ALL 111.100.1.10
!REQD en1 9.12.11.10
```

Referências relacionadas:

“PowerHA SystemMirror e Ethernet virtual” na página 20

O PowerHA SystemMirror suporta Ethernet virtual fornecido pelo Virtual I/O Server (VIOS) ou recursos integrados do Ethernet virtual (IVE), com as APARs aplicáveis instaladas. O suporte ao PowerHA SystemMirror é idêntico para VIOS e IVE.

Informações relacionadas:

➡ APAR IV14422: o pull de cabo VIOS não leva a eventos executados pelo PowerHA

➡ APAR IZ01331: nova funcionalidade netmon para suportar o PowerHA SystemMirror para AIX no VIO

Evitando conflitos de rede

É possível evitar conflitos de rede em endereços IP. Uma verificação notificará você de endereços IP duplicados. Corrija o endereço duplicado e resincronize o cluster.

Incluindo a topologia de rede no diagrama de cluster

Rascunhe as redes para incluir todas as redes TCP/IP. Identifique cada rede por nome e atributo. Nas caixas em cada nó que representa slots, registre o rótulo da interface.

É possível agora incluir a rede no diagrama de cluster de amostra iniciado na Visão geral do processo de planejamento.

Referências relacionadas:

“Visão geral do processo de planejamento” na página 3

Este tópico descreve as etapas para o planejamento de um cluster do PowerHA SystemMirror.

Planejando dispositivos de disco e de fita compartilhados

Esta seção discute informações a considerar antes de configurar discos externos compartilhados em um cluster do PowerHA SystemMirror e fornece informações sobre planejamento e configuração de unidades de fita como recursos de cluster.

Pré-requisitos

Você concluiu as etapas de planejamento nas seções de planejamento de conectividade de rede de cluster e de planejamento de aplicativos e controladores de aplicativos.

Consulte a documentação do AIX para a configuração geral de hardware e software para seu disco e dispositivos de fita.

Visão geral do disco compartilhado e dispositivos de fita

Em um cluster do PowerHA SystemMirror, os discos compartilhados são discos externos conectados a mais de um nó do cluster usado para armazenamento compartilhado do aplicativo.

Em uma configuração não simultânea, apenas um nó por vez possui os discos. Se o nó do proprietário falhar, o nó do cluster com a próxima prioridade mais alta na lista de nós do grupo de recursos adquirirá

a propriedade de discos compartilhados e reiniciará os aplicativos para restaurar serviços críticos aos clientes. Isso assegura que os dados armazenados nos discos permaneçam acessíveis aos aplicativos cliente.

Geralmente, o controle ocorre dentro de 30 a 300 segundos. Este intervalo depende do número e dos tipos de discos usados, do número de grupos de volumes, dos sistemas de arquivos (seja compartilhado ou montado cruzado do Network File System (NFS)) e do número de aplicativos críticos na configuração do cluster.

Ao planejar o disco externo compartilhado para seu cluster, o objetivo é eliminar pontos únicos de falha no subsistema de armazenamento em disco. A tabela a seguir lista os componentes de subsistema de armazenamento em disco, com formas sugeridas de eliminá-las como pontos únicos de falha.

Objeto do cluster	Eliminado como o ponto único de falha por...
Adaptador de disco	Usando adaptadores de disco redundantes
Controlador	Usando controladores de disco redundantes
Disk	Usando hardware redundante e espelhamento de disco do LVM ou espelhamento do RAID

Nesta seção, você executa as tarefas de planejamento a seguir:

- Escolher uma tecnologia de disco compartilhado.
- Planejar a instalação do armazenamento de disco compartilhado. Isso inclui:
 - Determinar o número de discos necessários para manipular a capacidade de armazenamento projetada. São necessários diversos discos físicos nos quais colocar os volumes lógicos espelhados. Colocar cópias de um volume lógico espelhado no mesmo dispositivo físico anula o propósito de fazer cópias. Para obter mais informações sobre criar volumes lógicos espelhados, consulte *Planejando componentes compartilhados do LVM*.
 - Determinar o número de adaptadores de disco que cada nó conterà para conectar aos discos ou subsistema de disco.

Os discos físicos que contêm cópias de volume lógico devem estar em adaptadores separados. SE todas as cópias de volume lógico forem conectadas a um único adaptador, o adaptador será potencialmente um ponto único de falha. Se o único adaptador falhar, o PowerHA SystemMirror moverá o grupo de volumes para um nó alternativo. Os adaptadores separados evitam a necessidade desta movimentação.
 - Entender os requisitos de cabeamento para cada tipo de tecnologia de disco.
- Incluir a configuração de disco selecionado no diagrama de cluster.
- Planejar a configuração de uma conexão de unidade de fita de Fibre Channel direta como um recurso de cluster.

Referências relacionadas:

“Planejando componentes de LVM compartilhados” na página 39

Esta seção descreve o planejamento de grupos de volumes compartilhados para um cluster do PowerHA SystemMirror.

Escolhendo uma tecnologia de disco compartilhado

O software PowerHA SystemMirror suporta tecnologias de disco como discos externos compartilhados do aplicativo em um cluster altamente disponível.

Para obter informações específicas sobre quais tecnologias de disco são suportadas na versão específica do PowerHA SystemMirror e do sistema operacional AIX, consulte *Matriz de suporte de hardware PowerHA*.

Informações relacionadas:

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

Considerações de fonte de alimentação de disco

As fontes de alimentação confiáveis são críticas para um cluster altamente disponível. Cada cadeia de disco espelhada no cluster deve ter uma fonte de alimentação separada. Conforme planejar o cluster, assegure-se de que a falha de nenhuma fonte de alimentação (PDU, fonte de alimentação ou circuito do prédio) não desative mais de um nó ou cadeia espelhada.

O IBM DS4000 series é menos propenso a problemas de fonte de alimentação porque têm fontes de alimentação redundantes.

Planejamento para armazenamento em disco não compartilhado

Mantenha algumas considerações em mente em relação ao armazenamento em disco não compartilhado.

Essas considerações incluem:

- Discos internos. Os discos internos em cada nó em um cluster devem fornecer espaço suficiente para:
 - Software AIX (aproximadamente 500 MB)
 - Software PowerHA SystemMirror (aproximadamente 50 MB para um nó de servidor)
 - Módulos executáveis de aplicativos altamente disponíveis
- Grupo de volume raiz. O grupo de volume raiz para cada nó não deve residir no barramento SCSI compartilhado.
- Recurso de notificação de erro do AIX. Use o Recurso de notificação de erro do AIX para monitorar os discos e adaptadores em cada nó. É possível ativar a notificação automática de erro no PowerHA SystemMirror para monitorar ambos os discos compartilhados e não compartilhados inserindo `smi t sysmirror` da linha de comandos e selecionando **Ferramentas de determinação de problemas > PowerHA SystemMirror Error Notification > Configurar notificação automática de erro > Incluir métodos de notificação de erro para recursos do cluster**.
- Uso de adaptador de disco. Como os discos compartilhados requerem seus próprios adaptadores, não é possível usar o mesmo adaptador para ambos os discos compartilhados e não compartilhados. Os discos internos em cada nó requerem um adaptador SCSI separado de quaisquer outros adaptadores no cluster.
- Uso de grupo de volumes. Os discos internos devem estar em um grupo de volumes diferente dos discos compartilhados externos.

Os módulos executáveis dos aplicativos altamente disponíveis devem estar nos discos internos e não nos discos externos compartilhados, pelos motivos a seguir:

- Licensing
- Inicialização do aplicativo

Informações relacionadas:

Configurando AIX para PowerHA SystemMirror

Licensing

Os fornecedores que requerem a compra de uma cópia separada de cada aplicativo para cada processador ou multiprocessador que possa executá-la e a proteção do aplicativo incorporando informações específicas de processador no aplicativo quando for instalado.

Assim, se você estiver executando seu executável de aplicativo a partir de um disco compartilhado, é possível que após um failover, o PowerHA SystemMirror não possa reiniciar o aplicativo em outro nó, porque, por exemplo, o ID do processador no novo nó não corresponde ao ID do nó no qual o aplicativo foi instalado.

O aplicativo também pode requerer que compre o que é chamado de licença ligada ao nó, ou seja, um arquivo de licença em cada nó que contém informações específicas ao nó.

Também pode haver uma restrição no número de licenças flutuantes (disponível a qualquer nó do cluster) disponível no cluster para esse aplicativo. Para evitar este problema, assegure-se de que haja licenças suficientes para todos os processadores no cluster que possam executar em potencial um aplicativo ao mesmo tempo.

Iniciando aplicativos

Os aplicativos podem conter arquivos de configuração que você pode customizar durante a instalação e armazenar os arquivos de aplicativo. Esses arquivos de configuração geralmente armazenam informações, como nomes de caminho e arquivos de log, que são usados quando o aplicativo é iniciado.

Pode ser preciso customizar seus arquivos de configuração se sua configuração requerer ambos:

- Você planeja armazenar esses arquivos de configuração em um sistema de arquivo compartilhado.
- O aplicativo não puder usar a mesma configuração em cada nó de fallover.

Por exemplo, em uma configuração de controle mútuo de dois nós, ambos os nós podem executar diferentes instâncias do mesmo aplicativo e apoiar um ao outro. Cada nó deve estar ciente do local dos arquivos de configuração para ambas as instâncias do aplicativo e devem poder acessá-los após um fallover. Caso contrário, o fallover falhará, deixando aplicativos críticos indisponíveis aos clientes.

Para diminuir quanto precisará customizar seus arquivos de configuração, coloque arquivos de inicialização levemente diferentes para aplicativos críticos em sistemas de arquivos locais em um dos nós. Isso permite que os parâmetros de aplicativo iniciais permaneçam estáticos. O aplicativo não precisará recalcular os parâmetros toda vez que for chamado.

Planejando uma instalação de disco compartilhado

Esta seção resume os componentes de hardware básicos necessários para configurar um cluster PowerHA SystemMirror.

Seus requisitos de cluster dependem da configuração especificada. Para assegurar que você considere todos os componentes necessários, conclua um diagrama para seu sistema. Além disso, consulte as informações de hardware para obter informações detalhadas sobre cabeamento e conexão para determinados dispositivos que estiver configurando.

PowerHA SystemMirror e SCSI virtual

O PowerHA SystemMirror suporta o SCSI virtual (VSCSI) com as APARs aplicáveis instaladas.

As restrições a seguir se aplicam ao uso de VSCSI em uma configuração de cluster:

- Se os sistemas de arquivos forem usados nos nós de espera, eles não serão montados até o ponto de fallover para que o uso acidental dos dados nos nós de espera seja impossível.
- Se os volumes compartilhados forem acessados diretamente (sem sistemas de arquivos) em modo simultâneo aprimorado, esses volumes serão acessíveis a partir de diversos nós, portanto, o acesso deve ser controlado em uma camada superior, como bancos de dados.
- A partir do ponto de visão do Virtual I/O Server (VIOS), os discos físicos são compartilhados, não volumes lógicos ou grupos de volumes.
- Toda a construção e manutenção do grupo de volumes nesses discos compartilhados são feitas a partir dos nós do PowerHA SystemMirror, não do VIOS.

Adaptadores de disco

Remova quaisquer terminadores SAS no adaptador. Use terminadores externos em um cluster do PowerHA SystemMirror. Se finalizar o barramento SAS compartilhado no adaptador, você perderá o término quando o nó do cluster que contém o adaptador falhar.

Cabos

Os cabos necessários para conectar nós em seu cluster dependem do tipo de barramento SCSI que está configurando. Selecione os cabos que são compatíveis com seus adaptadores de disco e controladores. Para obter informações sobre o tipo e o comprimento do cabo SCSI necessário, consulte a documentação de hardware que acompanha cada dispositivo que deseja incluir no barramento SCSI.

Exemplo: configuração do DS4000 Storage Server

Este exemplo mostra uma configuração para a alta disponibilidade ao usar o IBM DS4000 Storage Server em um ambiente PowerHA SystemMirror.

Consulte a figura a seguir.

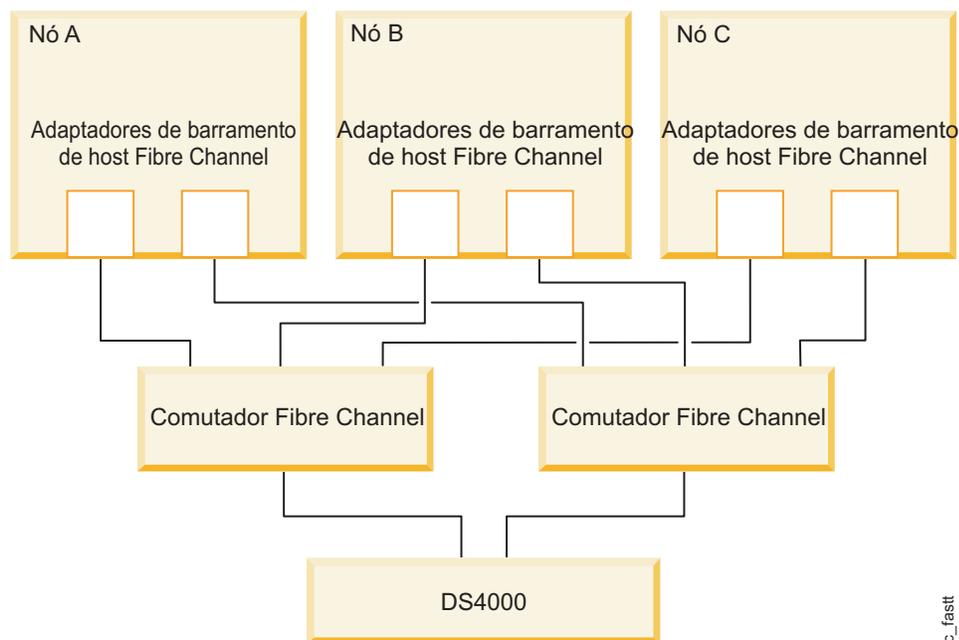


Figura 4. Ambiente do DS4000 Storage Server

Incluindo a configuração do disco no diagrama de cluster

Após ter escolhido uma tecnologia de disco, inclua sua configuração de disco no diagrama de cluster que iniciou no Planejamento inicial do cluster.

Para o diagrama de cluster, desenhe uma caixa representando cada disco compartilhado. Em seguida, cada caixa com um nome de disco compartilhado.

Referências relacionadas:

“Planejamento de cluster inicial” na página 5

Esta seção descreve as etapas iniciais que você executa para planejar um cluster do PowerHA SystemMirror para tornar os aplicativos altamente disponíveis.

Planejamento para unidades de fita e recursos de cluster

É possível configurar uma unidade de fita como um recurso de cluster, tornando-o altamente disponível a diversos nós em um cluster.

Os anexos de unidade de fita de Fibre Channel direto são suportados. O gerenciamento de unidades de fita compartilhadas é simplificado seguindo as funções do PowerHA SystemMirror:

- Configuração de unidades de fita usando o 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 a reconfiguração dinâmica

Limitações

Ao planejar incluir as unidades de fita como recursos de cluster, lembre-se do seguinte:

- Um carregador ou empilhador de fita é tratado como uma unidade de fita simples pelo PowerHA SystemMirror.
- Não mais de dois nós do cluster podem compartilhar o recurso de fita.
- Os recursos de fita não podem ser parte de grupos de recursos simultâneos.
- A unidade de fita deve ter o mesmo nome (por exemplo, `/dev/rmt0`) em ambos os nós que compartilham o dispositivo de fita.
- Quando um arquivo especial de fita é encerrado, a ação padrão é liberar a unidade de fita. O PowerHA SystemMirror não é responsável pelo estado da unidade de fita após um aplicativo ter aberto a fita.
- Nenhum meio de sincronizar operações de fita e controladores de aplicativos é fornecido. Se você decidir que uma operação de reserva e liberação de fita deve ser feita de forma assíncrona, forneça uma forma de notificar o controlador de aplicativos para aguardar até que a operação de reserva e liberação seja concluída.

Reservando e liberando unidades de fita compartilhadas

Quando um grupo de recursos com recursos de fita é ativado, a unidade de fita é reservada para permitir seu uso exclusivo.

Essa reserva é mantida até que um aplicativo libere-a, ou o nó seja removido do cluster:

- Quando o arquivo especial para a fita é encerrado, a ação padrão é liberar a unidade de fita. Um aplicativo pode abrir uma unidade de fita com uma sinalização para não liberar no encerramento. O PowerHA SystemMirror não será responsável por manter a reserva após um aplicativo ser iniciado.
- Mediante a parada dos serviços de cluster em um nó e colocar os grupos de recursos off-line, a unidade de fita é liberada, permitindo o acesso dos outros nós.
- Mediante a falha de nó inesperada, uma liberação forçada é feita no nó de controle. A unidade de fita é reservada como parte da ativação do grupo de recursos.

Configurando unidades de fita para operar síncrona ou assincronamente

Se uma operação de fita estiver em andamento quando uma reserva ou liberação de fita for iniciada, pode levar alguns minutos antes da operação de reserva ou liberação ser concluída. O PowerHA SystemMirror permite as operações de reserva e liberação síncrona ou assíncrona. A operação assíncrona e síncrona é especificada separadamente para reserva e liberação.

Operação síncrona

Com a operação síncrona, (o valor padrão), o PowerHA SystemMirror aguarda a operação de reserva ou liberação, incluindo a execução de um procedimento de recuperação definido pelo usuário, para que conclua antes de continuar.

Operação assíncrona

Com operação assíncrona, o PowerHA SystemMirror cria um processo-filho para executar a operação de reserva ou liberação, incluindo a execução de um procedimento de recuperação definido pelo usuário e continua imediatamente.

Procedimentos de recuperação

Os procedimentos de recuperação são altamente dependentes do aplicativo que acessa a unidade de fita.

Em vez de tentar prever cenários prováveis e desenvolver procedimentos de recuperação, o PowerHA SystemMirror fornece a execução dos scripts de recuperação definidos para as operações a seguir:

- Início de fita
- Parada de fita

Scripts de início e scripts de parada de fita

As operações de início e parada de fita ocorrem durante o início e a parada do nó, fallover e reintegração de nó e reconfiguração dinâmica. Esses scripts são chamados quando um grupo de recursos é ativado (início de fita) ou quando um grupo de recursos é desativado (parada de fita). Os scripts de início e parada de amostra podem ser localizados no diretório `/usr/es/sbin/cluster/samples/tape:`

`tape_resource_stop_example`

- Durante o início de fita, o PowerHA SystemMirror reserva a unidade de fita, forçando uma liberação se necessário e, em seguida, chama o script de início de fita fornecida pelo usuário.
- Durante a parada de fita, o PowerHA SystemMirror chama o script de parada de fita fornecido pelo usuário e libera a unidade de fita.

Nota: Você é responsável por posicionar corretamente a fita, finalizando processos ou aplicativos, gravando na unidade de fita e gravando marcas de fim de fita nesses scripts.

Outros procedimentos específicos de aplicativos devem ser incluídos como parte dos scripts de servidor de início e de servidor de parada.

Fallover e recuperação do adaptador

As unidades de fita com mais de uma interface SCSI não são suportadas. Portanto, apenas uma conexão existe entre um nó e uma unidade de fita. A noção usual de fallover de adaptador não se aplica.

Fallover e recuperação de nó

Se um nó que tiver recursos de fita que são parte de um grupo de recursos do PowerHA SystemMirror falhar, o nó de controle reserva a unidade de fita, forçando uma liberação se necessário e, em seguida, chama o script de início de fita fornecido pelo usuário.

Na reintegração de um nó, o nó de controle executa o script de parada de fita e libera a unidade de fita. O nó que está sendo reintegrado reserva a unidade de fita e chama o script de início de fita fornecido pelo usuário.

Fallover e recuperação de rede

O PowerHA SystemMirror não fornece procedimentos de fallover e recuperação para a falha de rede.

Planejando componentes de LVM compartilhados

Esta seção descreve o planejamento de grupos de volumes compartilhados para um cluster do PowerHA SystemMirror.

Pré-requisitos

Você também deve estar familiarizado com o uso do Gerenciador de Volume Lógico (LVM).

Visão Geral

O planejamento de componentes do gerenciador de volume lógico (LVM) compartilhados para um cluster do PowerHA SystemMirror depende do tipo de dispositivo de disco compartilhado e do método de acesso do disco compartilhado.

Para evitar um ponto único de falha para o armazenamento de dados, use a redundância de dados conforme suportado pelo LVM ou seu sistema de armazenamento.

Informações relacionadas:

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

Gerenciamento do Sistema Operacional e de Dispositivos

Planejamento para componentes do LVM

O gerenciador de volume lógico (LVM) controla os recursos de disco mapeando dados entre armazenamento físico e lógico.

O *armazenamento físico* refere-se ao local real dos dados em um disco. O *armazenamento lógico* controla quantos dados são disponibilizados ao usuário. O armazenamento lógico pode ser distante, expandido e replicado e pode abranger diversos discos físicos. Esses recursos fornecem a disponibilidade melhorada dos dados.

Volumes físicos

Um volume físico é um único disco físico ou unidade lógica apresentado por uma matriz de armazenamento.

O volume físico é particionado para fornecer o sistema operacional AIX com uma forma de gerenciar como os dados são mapeados para o volume. A figura a seguir mostra um uso convencional de partições físicas em um volume físico.

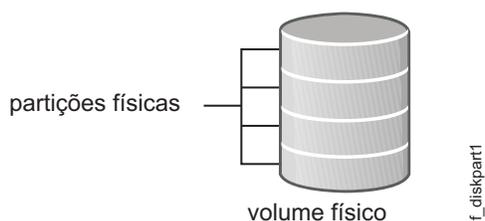


Figura 5. Partições físicas em um volume físico

Ao planejar volumes físicos compartilhados, assegure-se de que:

- A lista de PVIDs para um grupo de volumes é idêntica em todos os nós do cluster que têm acesso ao volume físico compartilhado
- A configuração para o atributo simultâneo do grupo de volumes é consistente entre todos os nós do cluster relacionados.

Grupos de volume

Um grupo de volumes é um conjunto de volumes físicos que o sistema operacional AIX trata como uma região de disco contígua e endereçável. É possível colocar diversos volumes físicos no mesmo grupo de volumes. O número real depende de como o grupo de volumes é criado.

A figura a seguir mostra um grupo de volumes de três volumes físicos:

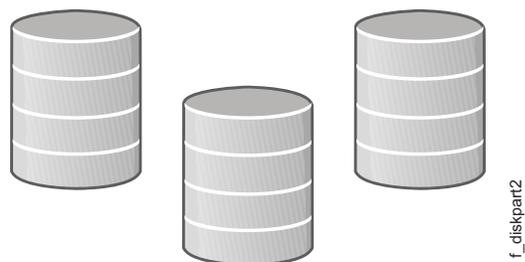


Figura 6. Grupo de volumes de três volumes físicos

No ambiente do PowerHA SystemMirror, um *grupo de volumes compartilhados* é um grupo de volumes que reside completamente em discos externos que são compartilhados pelos nós do cluster. Um grupo de volumes compartilhados não simultâneo pode ser variado apenas por um nó por vez.

Ao trabalhar com um grupo de volumes compartilhados:

- Não inclua um disco interno em um grupo de volumes compartilhados, porque não pode ser acessado por outros nós. Se você incluir um disco interno em um grupo de volumes compartilhados, o comando **varyonvg** falhará.
- Não ative os grupos de volumes compartilhados em um cluster do PowerHA SystemMirror manualmente na inicialização do sistema. Use scripts de evento de cluster para fazer isso.
- Assegure-se de que o atributo de ativação automática no AIX ODM esteja configurado como **No** para os grupos de volumes compartilhados listados em um grupo de recursos. O utilitário de verificação de cluster do PowerHA SystemMirror corrige automaticamente esse atributo para você mediante a verificação dos recursos de cluster e configura o atributo de variação automática como **No**.
- Se você definir um grupo de volumes para PowerHA SystemMirror, não gerencie-o manualmente em nenhum nó fora do PowerHA SystemMirror enquanto o PowerHA SystemMirror está em execução em outros nós. Isso pode levar a resultados imprevisíveis. Se desejar executar ações em um grupo de volumes independente do PowerHA SystemMirror, pare os serviços de cluster, execute uma tarefa de gerenciamento de grupo de volumes manual, deixe o grupo de volumes desativado e reinicie o PowerHA SystemMirror. Para facilitar o planejamento do uso de volumes físicos do PowerHA SystemMirror, o utilitário de verificação verifica:
 - Consistência de grupo de volumes
 - Disponibilidade de disco

Informações relacionadas:

Comando **mkvg**

Comando **varyon**

Volumes Lógicos

Um *volume lógico* é um conjunto de partições lógicas que o sistema operacional AIX torna disponível como uma única unidade de armazenamento, ou seja, a visualização lógica de um disco.

Uma partição lógica é a visualização lógica de uma partição física. 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 ou um dispositivo bruto.

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.

No sistema PowerHA SystemMirror, um sistema de arquivos compartilhados é um sistema de arquivos registrados que reside inteiramente em um volume lógico compartilhado.

Você deseja planejar sistemas se arquivo compartilhado a serem colocados em discos externos que são compartilhados pelos nós do cluster. Os dados residem em sistemas de arquivos nesses discos compartilhados externos para que fiquem altamente disponíveis.

A ordem na qual os sistemas de arquivos são montados geralmente não é importante. Entretanto, se isso for importante para o seu cluster, será necessário planejar algumas coisas:

- Os sistemas de arquivos que existem em um único grupo de recursos são montados na ordem alfanumérica quando o grupo de recursos fica on-line. Eles também são desmontados na ordem alfanumérica reversa quando o grupo de recursos é colocado off-line.
- Se você tiver sistemas de arquivos compartilhados e aninhados, será necessário um cuidado adicional. Se tiver sistemas de arquivos compartilhados e aninhados em um único grupo de recursos, você deve configurar o método de recuperação de sistemas de arquivos para que o grupo de recursos como sequencial para garantir a ordem de montagem correta.
- Se você tiver sistemas de arquivos aninhados que residam em diferentes grupos de recursos, você deve planejar adicionalmente um relacionamento-pai/filho para que esses grupos de recursos garantam a ordem de montagem correta.

Planejando espelhamento de LVM

O espelhamento do gerenciador de volume lógico (LVM) fornece a capacidade de alocar mais de uma cópia de uma partição física para aumentar a disponibilidade dos dados. Quando um disco falha e suas partições físicas se tornam indisponíveis, você ainda tem acesso aos dados espelhados em um disco disponível. O LVM executa o espelhamento no volume lógico.

Em um cluster do PowerHA SystemMirror, é possível espelhar o seguinte:

- Dados de volume lógico em um grupo de volumes compartilhados
- O volume lógico do log para cada grupo de volumes compartilhados com sistemas de arquivos

Espelhamento de partições físicas

Para melhorar a disponibilidade do volume lógico, é possível alocar uma, duas ou três cópias de uma partição física para espelhar dados contidos na partição.

Se uma cópia for perdida devido a um erro, as outras cópias não danificadas serão acessadas e o sistema operacional AIX continuará processando com uma cópia precisa. Após o acesso ser restaurado para a partição física com falha, o AIX ressincronizará o conteúdo (dados) da partição física com o conteúdo (dados) de uma cópia espelhada consistente.

A figura a seguir mostra um volume lógico composto de duas partições lógicas com três cópias espelhadas. No diagrama, cada partição lógica mapeia para três partições físicas. Cada partição física deve ser designada para residir em um volume físico separado em um grupo de volumes único. Esta configuração fornece o número máximo de caminhos alternativos para as cópias espelhadas e, portanto, a maior disponibilidade.

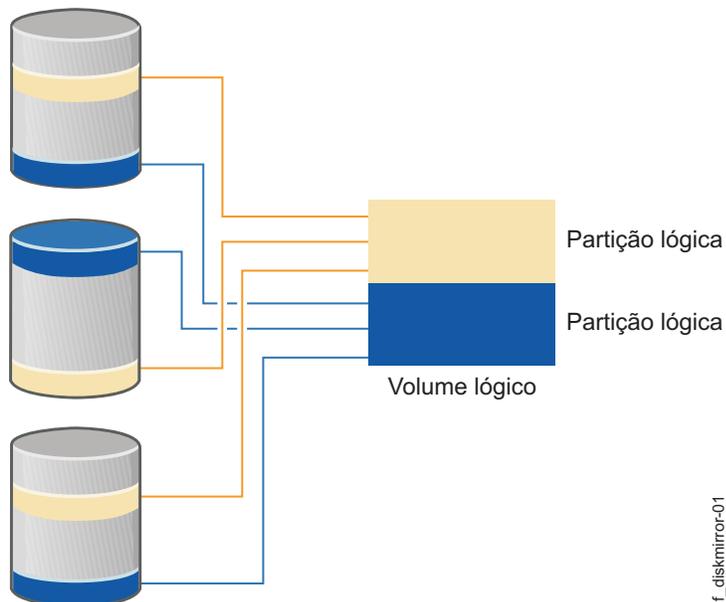


Figura 7. Volume lógico de duas partições lógicas com três cópias espelhadas

As cópias espelhadas são transparentes, o que significa que não é possível isolar uma dessas cópias. Por exemplo, se você excluir um arquivo de um volume lógico com diversas cópias, o arquivo excluído será removido de todas as cópias do volume lógico.

As configurações a seguir aumentam a disponibilidade de dados:

- Alocar três cópias de uma partição lógica em vez de alocar uma ou duas cópias
- Alocar as cópias de uma partição lógica em volumes físicos diferentes em vez de alocar as cópias no mesmo volume físico
- Alocar as cópias de uma partição lógica entre diferentes gabinetes do disco físico em vez do mesmo gabinete, se possível
- Alocar as cópias de uma partição lógica entre diferentes adaptadores de disco em vez de usar um adaptador de disco único

Embora o uso de cópias espelhadas que abrangem diversos discos (nas fontes de alimentação separadas) junto com diversos adaptadores de disco assegura que nenhum disco seja um ponto único de falha para seu cluster, essas configurações podem aumentar o tempo para operações de gravação.

Especifique a política de alocação de disco **superstrict** para os volumes lógicos nos grupos de volumes para os quais a ativação forçada é especificada. Esta configuração:

- Garante que as cópias de um volume lógico sempre residam em discos separados
- Aumenta as chances de que a ativação forçada seja bem-sucedida após a falha de um ou mais discos.

Se você planeja usar a ativação forçada para o volume lógico, aplique a política de alocação de disco **superstrict** para os gabinetes de disco no cluster.

Para obter mais informações sobre a ativação forçada, consulte a seção Usando quorum e ativação para aumentar a disponibilidade de dados.

Referências relacionadas:

“Usando o quorum e a ativação para aumentar a disponibilidade de dados” na página 49

Como você configura o quorum e a ativação para grupos de volumes pode aumentar a disponibilidade dos dados de espelhamento.

Logs de diário de espelhamento

As configurações de acesso não simultâneo suportam sistemas de arquivos registrados e sistemas de arquivos registrados aprimorados.

O sistema operacional AIX usa o registro no diário para seus sistemas de arquivos. Em geral, isso significa que o estado interno de um sistema de arquivos na inicialização (em termos de lista de bloqueios e lista livre) é o mesmo estado que no encerramento. Em termos práticos, isso significa que quando o AIX é inicializado, a extensão de qualquer distorção de arquivo não pode ser pior que no encerramento.

Cada grupo de volumes contém um log **jfslog** ou **jfs2log**, que em si é um volume lógico. Esse log geralmente reside em um disco físico diferente no grupo de volumes do que o sistema de arquivos registrados. Entretanto, se o acesso a esse disco for perdido, mudanças nos sistemas de arquivos após esse ponto estarão em risco.

Para evitar a possibilidade de esse disco físico ser um ponto único de falha, é possível especificar cópias espelhadas de cada log **jfslog** ou **jfs2log**. Coloque essas cópias em volumes físicos separados.

Planejamento para espelhamento de site de divisão do LVM

É possível configurar discos localizados em dois ou três locais diferentes para o espelhamento do Logical Volume Manager (LVM) remoto usando uma rede de área de armazenamento (SAN). Por exemplo, o espelhamento de site de divisão usa o LVM para replicar dados entre o subsistema de disco em cada local diferente para a recuperação de desastre.

Nota: O PowerHA SystemMirror apenas suporta duas configurações de site.

Uma SAN é uma rede de alta velocidade que permite que seu ambiente estabeleça conexões diretas entre os dispositivos de armazenamento e sistemas (nós). Assim, dois ou mais sistemas localizados em diferentes locais podem acessar os mesmos discos físicos usando uma conexão de rede SAN. É possível combinar discos remotos em um grupo de volumes usando o LVM. É possível importar esse grupo de volumes para os nós localizados em locais diferentes.

Os volumes lógicos no grupo de volumes que contêm os discos remotos podem ter até três espelhos remotos. É possível configurar pelo menos um espelho remoto em cada local. Os dados armazenados no volume lógico são altamente disponíveis. Portanto, em caso de falha em seu ambiente, por exemplo, todos os nós não estão disponíveis em um local, o espelho remoto em outro local possui os dados mais recentes.

O PowerHA SystemMirror sincroniza automaticamente todos os espelhos remotos após uma falha de disco ou nó ocorrer e os nós serem colocados novamente on-line. A sincronização automática acontece até mesmo se um dos discos estiver no estado PVREMOVED ou PVMISSING. A sincronização automática não está disponível para todos os casos de espelhamento de site de divisão do LVM. Se não estiver disponível, é possível usar C-SPOC para sincronizar os dados.

Ao planejar uma configuração de espelhamento de site de divisão do LVM, você também deve planejar o disco de repositório usado no cluster. Você deve verificar se há um segundo disco pronto para ser usado como um disco de repositório quando o disco primário falhar.

Nota: O Cluster Aware AIX suporta a substituição de repositório em tempo real sem impactar funções de cluster críticas.

Exemplo:

A figura a seguir é um exemplo de configuração de espelhamento de site de divisão do LVM usando uma SAN.

É possível espelhar os discos conectados a pelo menos um nó em cada um dos dois locais. Neste

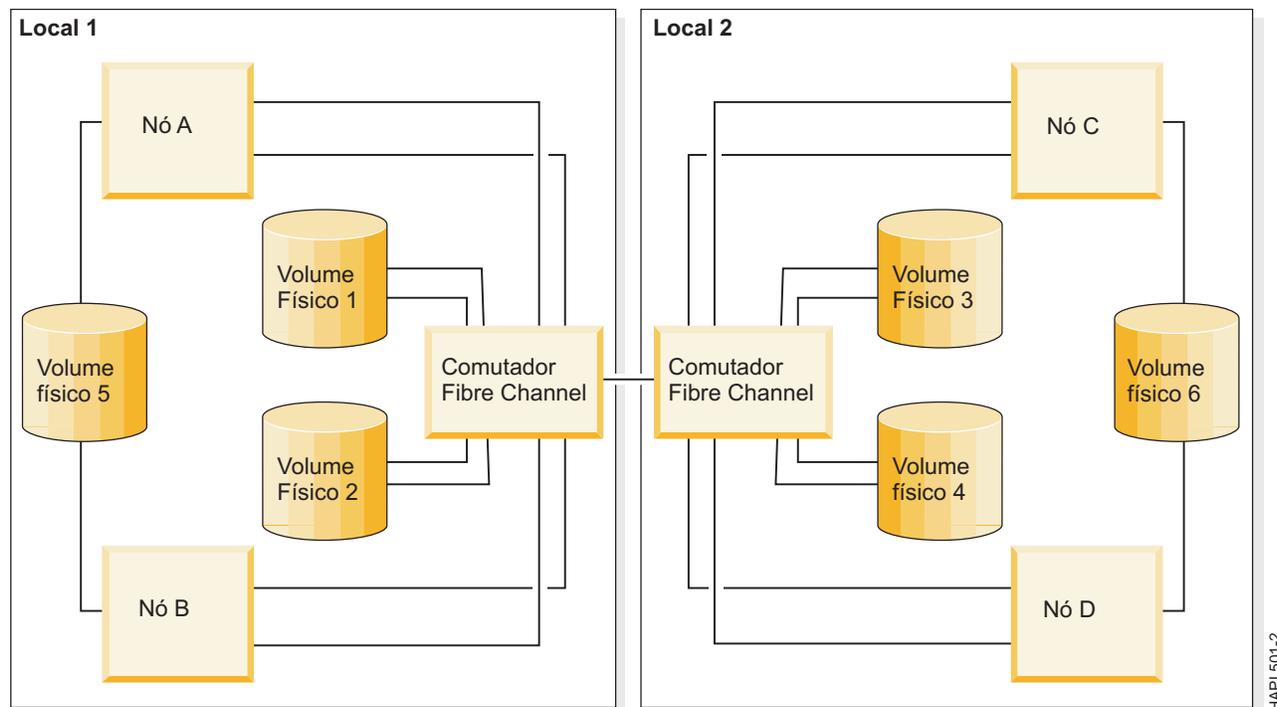


Figura 8. Configuração de espelhamento de site de divisão do LVM usando uma SAN

exemplo, o PV4 está disponível para o Nó A e Nó B no Local 1 e Nó C no Local 2 usando a conexão Fibre Channel Switch 1 e Fibre Channel Switch 2. É possível ter um espelho de PV4 no Local 1. Os discos conectados aos nós em apenas um local (PV5 e PV6) não podem ser espelhados entre locais.

É possível usar a função de conjuntos de reflexos espelhados do LVM do AIX para assegurar que os dados sejam corretos e completamente espelhados entre os dois locais. Se PV1 e PV2 estiverem em um conjunto espelhado e PV3 e PV4 estiverem em um conjunto de reflexos separado, o LVM permitirá que uma cópia completa dos dados esteja presente em cada local. Você deve usar conjuntos espelhados superestrictos para garantir que uma cópia completa dos dados esteja em cada local.

Neste exemplo, é possível usar conjuntos espelhados para ajudar a manter o espelhamento até mesmo conforme os discos forem incluídos, removidos ou substituídos em um grupo de volumes. Também é possível usar a função C-SPOC para definir conjuntos espelhados e associá-los aos discos em cada local.

Informações relacionadas:

Gerenciando o espelhamento de site de divisão do LVM

Espelhamento entre sites

É possível configurar discos localizados em dois sites diferentes para usar o espelhamento do Gerenciador de Volume Lógico (LVM), usando uma rede de área de armazenamento (SAN). Os dados são replicados entre o subsistema de disco em cada site para a recuperação de desastre.

Uma SAN é uma rede de alta velocidade que permite o estabelecimento de conexões diretas entre os dispositivos de armazenamento e processadores. Assim, dois ou mais nós localizados em diferentes sites podem acessar os mesmos discos físicos, que podem ser separados por alguma distância, por meio da SAN comum. Esses discos remotos podem ser combinados em um grupo de volumes usando o LVM e esse grupo de volumes pode ser importado para os nós localizados em diferentes sites. 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 nesse volume lógico são mantidas

altamente disponíveis e no caso de certas falhas, o espelho remoto em outro site ainda terá as informações mais recentes para que as operações possam ser continuadas no 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 para todos os casos, mas é possível usar o C-SPOC para sincronizar os dados a partir dos espelhos sobreviventes para os espelhos estagnados após uma falha de disco ou site e reintegração subsequente.

Nota: No PowerHA SystemMirror Enterprise Edition, também é possível usar o espelhamento em um cluster que abrange dois sites usando a função de espelhamento do Gerenciador de Volume Lógico (GLVM).

Planejamento para acesso de disco

É possível configurar discos para ter acesso simultâneo aprimorado ou acesso não simultâneo.

- **Acesso simultâneo aprimorado.** Os dados nos discos estão disponíveis a todos os nós conectados simultaneamente e todos os nós têm acesso aos metadados nos discos. Este modo de acesso permite um controle rápido de disco, porque o grupo de volumes pode ser colocado on-line antes de os metadados serem lidos.

Todos os grupos de volumes compartilhados devem ser configurados como grupos de volumes de modo simultâneo aprimorado, se serão acessados simultaneamente ou não. É possível migrar grupos de volumes existentes para o modo simultâneo aprimorado.

O uso de sistema de arquivos registrados (JFS) e do sistema de arquivos registrados aprimorado (JFS2) não é suportado se você configurar o grupo de volumes aprimorado correspondente (na política Grupo de recursos) para ser acessado simultaneamente a partir de dois ou mais nós.

As configurações de acesso simultâneo que usam o IBM TotalStorage DS Series ou IBM 2105 Enterprise Storage Servers não usam o espelhamento de gerenciador de volume lógico (LVM). Em vez disso, esses sistemas fornecem sua própria redundância de dados.

- **Acesso não simultâneo.** Apenas um nó por vez pode acessar as informações sobre os discos.

Se o grupo de recursos que contém esses discos se mover para outro nó, o novo nó poderá acessar os discos, ler os metadados (informações sobre o estado atual dos grupos de volumes e outros componentes), execute o comando **varyon** nos grupos de volumes e monte quaisquer sistemas de arquivos associados.

As configurações de acesso não simultâneo geralmente usam sistemas de arquivos registrados. Em alguns casos, um aplicativo de banco de dados que está em execução em um ambiente não simultâneo pode efetuar bypass do sistema de arquivos registrados e acessar o volume lógico bruto diretamente.

Referências relacionadas:

“Acesso simultâneo aprimorado”

Qualquer disco suportado pelo PowerHA SystemMirror para conexão com diversos nós deve ser colocado em um grupo de volumes de modo simultâneo aprimorado e pode ser usado em ambientes simultâneos e não simultâneos (conforme especificado pelo tipo de grupo de recursos)

Acesso simultâneo aprimorado

Qualquer disco suportado pelo PowerHA SystemMirror para conexão com diversos nós deve ser colocado em um grupo de volumes de modo simultâneo aprimorado e pode ser usado em ambientes simultâneos e não simultâneos (conforme especificado pelo tipo de grupo de recursos)

- Simultâneo. Um aplicativo é executado em todos os nós de cluster ativos ao mesmo tempo.

Para permitir que tais aplicativos acessem seus dados, os grupos de volumes simultâneos são variados para todos os nós de cluster ativos. O aplicativo tem a responsabilidade de assegurar o acesso a dados consistente.

- Não simultâneo. Um aplicativo é executado em um nó por vez.

Os grupos de volumes não são acessados simultaneamente. Eles ainda são acessados por apenas um nó a qualquer horário dado.

Ao ativar o grupo de volumes no modo simultâneo aprimorado em todos os nós que possuem o grupo de recursos em um cluster, o LVM permite o acesso ao grupo de volumes em todos os nós. Entretanto, isso restringe as conexões de nível mais alto, como montagens NFS e montagens JFS, em todos os nós e permite-os apenas no nó que possui atualmente o grupo de volumes no PowerHA SystemMirror.

É possível usar a função AIX MPIO para acessar subsistemas de disco por meio de diversos caminhos. Diversos caminhos fornecem mais rendimento e maior disponibilidade que o uso de um único caminho. Em particular, quando diversos caminhos são usados, a falha de um único caminho devido a um adaptador ou uma falha de cabo ou comutador não fará com que os aplicativos percam o acesso aos dados. Embora o PowerHA SystemMirror tentará recuperar-se da perda completa do acesso a um grupo de volumes, essa perda em si será temporariamente disruptiva. A função do AIX MPIO pode impedir que uma falha de componente único cause uma indisponibilidade de aplicativo.

Quando o controle rápido de disco é usado, a função de reserva de disco não é usada. Se o cluster se tornar particionado, os nós em cada partição poderiam ativar acidentalmente o grupo de volumes no estado ativo. Como a ativação do estado ativo do grupo de volumes permite a montagem de sistemas de arquivos e a mudança dos volumes físicos, esta situação pode resultar em cópias diferentes do mesmo grupo de volumes. Para obter mais informações sobre o controle rápido de disco e usar diversas redes, consulte a seção Usando o controle rápido de disco.

Requisitos de acesso simultâneo para discos MPIO acessados

O modo simultâneo aprimorado é a única opção para criar grupos de volumes simultâneos. No PowerHA SystemMirror, os grupos de volumes de modo simultâneo aprimorado não usam reservas de disco. O acesso simultâneo que é necessário para discos MPIO acessados é automaticamente fornecido no PowerHA SystemMirror.

Sobre o modo simultâneo aprimorado

Todos os grupos de volumes simultâneos são criados como grupos de volumes de modo simultâneo aprimorado por padrão. Para grupos de volumes simultâneos aprimorados, o Concurrent Logical Volume Manager (CLVM) coordena as mudanças entre os nós por meio do componente de Serviços de grupo da função Reliable Scalable Cluster Technology (RSCT) no sistema operacional AIX. Os protocolos dos Serviços de Grupo fluem pelos links de comunicação entre os nós do cluster.

Tarefas relacionadas:

Convertendo grupos de volumes para o modo simultâneo aprimorado

Referências relacionadas:

“Usando controle rápido de disco”

O PowerHA SystemMirror automaticamente detecta grupos de volumes com falha e inicia um controle rápido de disco para grupos de volumes de modo simultâneo aprimorado que são incluídos como recursos em grupos de recursos não simultâneos.

Usando controle rápido de disco

O PowerHA SystemMirror automaticamente detecta grupos de volumes com falha e inicia um controle rápido de disco para grupos de volumes de modo simultâneo aprimorado que são incluídos como recursos em grupos de recursos não simultâneos.

O controle rápido de disco é especialmente útil para o failover de grupos de volumes de modo simultâneo aprimorado compostos de um grande número de discos. Este mecanismo de controle de disco é mais rápido que o controle de disco usado para grupos de volumes padrão incluídos em grupos de recursos não simultâneos. Durante o controle rápido de disco, o PowerHA SystemMirror ignora o

processamento extra necessário para quebrar as reservas de disco ou atualizar e sincronizar as informações do gerenciador de volume lógico (LVM) executando a atualização tardia.

Foi observado que o controle rápido de disco não leve mais de 10 segundos para um grupo de volumes com dois discos. Espera-se que esse tempo aumente muito lentamente para números maiores de discos e grupos de volumes. O tempo real observado em qualquer configuração depende dos fatores fora do controle do PowerHA SystemMirror, como a energia de processamento dos nós e a quantidade de atividade não relacionada no momento no failover. O tempo real observado para a conclusão do processamento de failover depende dos fatores adicionais, quanto a uma verificação de sistema de arquivos ser necessária ou não, e a quantidade de tempo necessária para reiniciar o aplicativo.

Nota: Os grupos de volumes de modo simultâneo aprimorado não são acessados simultaneamente. Eles só são acessados por um nó em um determinado horário. O mecanismo de controle rápido de disco funciona no nível do grupo de volumes e é, portanto, independente do número de discos usados.

Controle rápido de disco e ativação ativa e passiva

Um grupo de volumes simultâneo aprimorado pode ser ativado em um nó ou variado como ativo ou passivo.

Para permitir o controle rápido de disco, o PowerHA SystemMirror ativa os grupos de volumes simultâneos aprimorados nos estados ativos e passivos.

Ativação ativa

A ativação ativa comporta-se da mesma forma que a ativação comum e torna os volumes lógicos disponíveis. Quando um grupo de volumes simultâneo aprimorado é variado no estado ativo em um nó:

- Operações em sistemas de arquivos, como montagens do sistema de arquivos
- Operações em aplicativos
- Operações em volumes lógicos, como criar volumes lógicos
- Sincronizar grupos de volumes

Ativação passiva

Quando um grupo de volumes simultâneo aprimorado é variado em estado passivo, o LVM fornece o equivalente do fencing de disco para o grupo de volumes no nível do LVM.

A ativação do estado passivo só permite um número limitado de operações somente leitura no grupo de volumes:

- Acesso somente leitura do LVM ao arquivo especial do grupo de volumes
- Acesso somente leitura do LVM aos primeiros 4 kb de todos os volumes lógicos possuídos pelo grupo de volumes.

As operações a seguir não são permitidas quando um grupo de volumes é variado em estado passivo:

- Operações nos sistemas de arquivos, como montagem de sistemas de arquivos
- Quaisquer operações em volumes lógicos, como ter volumes lógicos abertos
- Sincronizar grupos de volumes

PowerHA SystemMirror e ativação ativa e passiva

PowerHA SystemMirror varia corretamente no grupo de volumes em estado ativo no nó que possui o grupo de recursos e altera os estados ativo e passivo apropriadamente como o estado e o local das mudanças de grupo de recursos.

- Mediante inicialização do cluster:

- No nó que possui o grupo de recursos, o PowerHA SystemMirror ativa o grupo de volumes no estado ativo. O PowerHA SystemMirror ativa um grupo de volumes no estado ativo apenas em um nó por vez.
- O PowerHA SystemMirror ativa o grupo de volumes em estado passivo em todos os outros nós no cluster.
- Mediante fallover:
 - Se um nó liberar um grupo de recursos ou se o grupo de recursos estiver sendo movido para outro nó por qualquer outro motivo, o PowerHA SystemMirror alterna o estado de ativação para o grupo de volume de ativo para passivo no nó que libera o grupo de recursos. O PowerHA SystemMirror ativa o grupo de volumes em estado ativo no nó que adquire o grupo de recursos.
 - O grupo de volumes permanece em estado passivo em todos os outros nós no cluster.
- O PowerHA SystemMirror faz os processos a seguir quando a reintegração de nó ocorre:
 - Altera o estado de ativação do grupo de volumes de ativo para passivo no nó que libera o grupo de recursos.
 - Varia o grupo de volumes em estado ativo no nó de junção.
 - Ativa esse grupo de volumes em estado passivo em todos os outros nós no cluster.

Nota: O comutador entre estados ativos e passivos é necessário para impedir a montagem dos sistemas de arquivo em mais de um nó por vez.

Usando o quorum e a ativação para aumentar a disponibilidade de dados

Como você configura o quorum e a ativação para grupos de volumes pode aumentar a disponibilidade dos dados de espelhamento.

Usando o quorum

O quorum assegura que mais da metade dos discos físicos em um grupo de volumes estão disponíveis.

O quorum não mantém o controle dos espelhos do volume lógico e, portanto, não é uma forma útil de assegurar a disponibilidade de dados. É possível perder o quorum quando ainda se tem todos os seus dados. Por outro lado, é possível perder acesso a alguns dos seus dados e não perder o quorum.

O quorum é benéfico para os grupos de volumes em matrizes RAID, como ESS e IBM TotalStorage DS Series. O dispositivo RAID fornece a disponibilidade de dados e a recuperação da perda de um único disco. O espelhamento geralmente não é usado para grupos de volumes contidos completamente em um único dispositivo RAID. Se um grupo de volumes for espelhado entre dispositivos RAID, a ativação forçada pode colocar um grupo de volumes on-line apesar da perda de um dos dispositivos RAID.

Decida se deve ativar ou desativar o quorum para cada grupo de volumes. A tabela a seguir mostra como o quorum é afetado quando os grupos de volumes são ativados ou não:

	Condição para ativação de grupo de volumes	Condição para desativação de grupo de volumes
Quorum ativado	Mais de 50% dos discos no grupo de volumes estão disponíveis	O acesso é perdido para 50% ou mais dos discos
Quorum desativado	Todos os discos no grupo de volumes estão disponíveis	O acesso é perdido a todos os discos

A verificação de quorum é ativada por padrão. É possível desativar o quorum usando o comando **chvg -Qn vgname** ou usando o atalho **smit chvg**.

Informações relacionadas:

Comando **chvg**

Quorum em configurações de acesso simultâneo:

O quorum deve ser ativado para uma configuração de acesso simultâneo do PowerHA SystemMirror. Desativar o quorum poderia resultar na distorção de dados. Qualquer configuração de acesso simultâneo em que diversas falhas poderiam resultar em nenhum disco compartilhado comum entre os nós de cluster possui o potencial para a distorção de dados ou inconsistência.

A figura a seguir mostra um cluster com dois conjuntos de subsistemas de disco da IBM configurados para nenhum ponto único de falha. Os volumes lógicos são espelhados entre subsistemas e cada subsistema de disco é conectado a cada nó com NICs separadas.

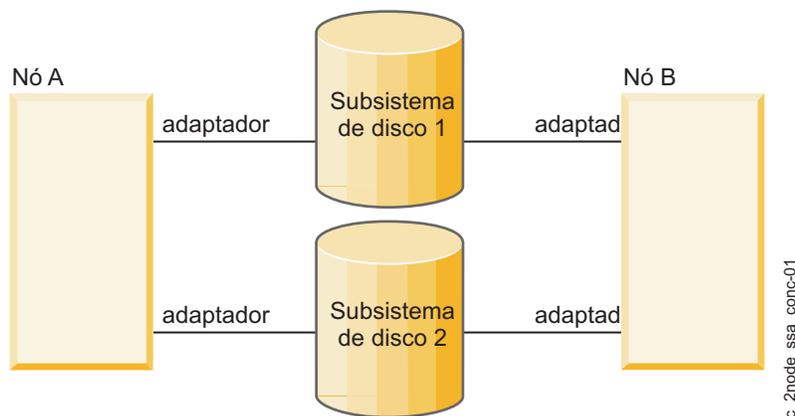


Figura 9. Configuração de acesso simultâneo do subsistema de disco da IBM

Se diversas falhas resultarem em uma perda de comunicações entre cada nó e um conjunto de discos de tal forma que o nó A possa acessar o subsistema 1, mas não o subsistema 2, e o nó B pode acessar o subsistema 2, mas não o subsistema 1. Ambos os nós continuam a operarem na mesma linha de base dos dados da cópia espelhada que possam acessar. Entretanto, cada nó não vê modificações feitas pelo outro nó nos dados no disco. Como resultado, os dados se tornam inconsistentes entre os nós.

Com a proteção de quorum ativada, a falha na comunicação resulta em um ou ambos os nós que variam do grupo de volumes. Embora um aplicativo não tenha acesso aos dados no grupo de volumes que são variados, a consistência de dados é preservada.

Fallover seletivo acionado pela perda do quorum:

O PowerHA SystemMirror fornece seletivamente a recuperação para grupos de recursos não simultâneos (com a política de inicialização não On-line em todos os nós disponíveis) que são afetados por falhas de recursos específicos. O PowerHA SystemMirror reage automaticamente a um erro de perda de quorum LVM_SA_QUORCLOSE associado a um grupo de volumes que vai ficar off-line em um nó do cluster. Na resposta a esse erro, um grupo de recursos não simultâneos fica off-line no nó em que o erro ocorreu.

Se o Gerenciador de Volume Lógico do AIX colocar um grupo de volumes no grupo de recursos off-line devido a uma perda de quorum para o grupo de volumes no nó, o PowerHA SystemMirror moverá seletivamente o grupo de recursos para outro nó. É possível alterar este comportamento padrão customizando a recuperação de recursos para usar um método de notificação em vez do fallover.

O PowerHA SystemMirror ativa o failover seletivo e recupera os grupos de recursos afetados em resposta a um erro LVM_SA_QUORCLOSE. Esse erro é gerado pelo AIX LVM para condições de erro específicas, mesmo se o grupo de volumes não estiver definido como quorum ativado. O AIX LVM também pode gerar outros tipos de notificações de erro; no entanto, o PowerHA SystemMirror não reage a eles por

padrão. Nestes casos, ainda é necessário configurar métodos de notificação de erro customizados ou usar os métodos de notificação automática de erro do AIX para reagir a falhas do grupo de volumes.

É possível usar o evento do sistema rootvg para monitorar a perda de acesso ao rootvg. Se o sistema perder o acesso, o PowerHA SystemMirror registrará um evento no log de erro de sistema e reinicializará o sistema por padrão. É possível alterar esta configuração usando o SMIT para registrar um evento, mas não reinicializar o sistema.

Os eventos de rootvg podem ser monitorados apenas quando o disco rootvg está usando a unidade MPIO (Multipath I/O) nativa do AIX e o disco rootvg não é um disco SCSI paralelo interno. Para verificar se o disco rootvg está usando a unidade MPIO, na linha de comandos, digite `lspath -l hdiskname`, em que *hdiskname* é o nome do disco rootvg. Se o disco rootvg não estiver usando a unidade MPIO, a seguinte mensagem de erro será exibida:

```
lspath: 0514-538 Cannot perform the requested function because the
        specified device does not support multiple paths.
```

Informações relacionadas:

Método de notificação de erro usado para perda do grupo de volumes

Eventos de sistema de monitoramento do PowerHA SystemMirror

Fallover seletivo para manipulação de grupos de recursos

Usando ativação forçada

O PowerHA SystemMirror fornece uma função de ativação forçada para usar em conjunto com os métodos de notificação automática de erro do AIX. A função de ativação forçada permite ter a disponibilidade de dados mais alta possível.

Forçar uma ativação de um grupo de volumes permite manter um grupo de volumes on-line contanto que haja uma cópia válida dos dados disponível. Use uma ativação forçada apenas para grupos de volumes que possuem volumes lógicos espelhados.

Nota: Tome cuidado ao usar esta função para evitar a criação de um cluster particionado.

É possível usar o SMIT para forçar uma ativação de um grupo de volumes em um nó se o comando **varyon** normal falhar nesse grupo de volumes devido à falta de quorum, mas com uma cópia válida dos dados disponível. Usar o SMIT para forçar uma ativação é muito útil para a recuperação de desastre local, quando os dados são espelhados entre dois gabinetes de disco e um dos gabinetes de disco se torna indisponível.

Nota: É possível especificar um atributo de ativação forçada para grupos de volumes em discos SCSI que usam o espelhamento de gerenciador de volume lógico (LVM) e para grupos de volumes que são espelhados entre dispositivos RAID ou ESS separados.

Se desejar forçar o grupo de volumes para ativação quando os discos estiverem indisponíveis, use **varyonvg -f**, que força o grupo de volumes para ativação, se há cópias ou não de seus dados. É possível especificar a ativação forçada no SMIT para grupos de volumes em um grupo de recursos.

Ativação forçada e particionamento de cluster

Se estiver usando a ativação forçada, é importante que diversas conexões de rede entre os nós que compartilha a área de armazenamento existam. Diversas conexões de rede ajudam a assegurar que cada nó sempre tenha um caminho de comunicação para outro nó, mesmo de uma rede falhar. Ter diversas conexões de rede impede seu cluster de ser particionado. Caso contrário, uma falha de rede pode fazer com que os nós tentem controlar grupos de recursos que ainda estão ativos em outros nós. Nesta situação, se tiver configurado uma configuração de ativação forçada, é possível passar pela perda de dados ou divergência.

Usando NFS com o PowerHA SystemMirror

O software do PowerHA SystemMirror fornece aprimoramentos de disponibilidade para a manipulação do Network File System (NFS).

Esses aprimoramentos incluem:

- Um recurso confiável do servidor NFS que permite que um processador de backup recupere a atividade do atual do NFS caso o servidor NFS primário falhe, preservando os bloqueios nos sistemas de arquivos do NFS e o cache de solicitação duplicado. Esta funcionalidade está restrita a Grupos de recursos de dois nós se contiver exportações do NFS versão 2 ou 3. Os grupos de recursos com apenas o NFS versão 4, ou posterior, podem suportar configurações de até 16 nós.
- Auxiliar de configuração do NFS para facilitar a instalação e a configuração.
- Controlador de aplicativos e monitor de aplicativo pré-configurados (clam_nfsv4) para monitorar exportações do NFS versão 4 e os daemons do NFS.
- Capacidade de especificar uma rede para montagem do NFS.
- Capacidade de definir exportações do NFS e montagens no nível do diretório.
- Capacidade de especificar opções de exportação para diretórios exportados do NFS e sistemas de arquivos.

Para que o NFS funcione conforme esperado em um cluster do PowerHA SystemMirror, há requisitos de configuração específicos. Portanto, você deve planejar as tarefas a seguir:

- Criar grupos de volumes compartilhados
- Exportar sistemas de arquivos do NFS
- Montagem e failover do NFS

Os scripts do PowerHA SystemMirror manipulam o comportamento do NFS padrão. Pode ser necessário modificar os scripts para manipular sua configuração específica.

É possível configurar o NFS em todos os grupos de recursos que se comportam como não simultâneos; isto é, eles não têm uma política de inicialização On-line em todos os nós disponíveis.

Renunciar o controle sobre sistemas de arquivos NFS em um cluster do PowerHA SystemMirror

Após configurar grupos de recursos que contenham sistemas de arquivos NFS, você renuncia o controle sobre sistemas de arquivos NFS para o PowerHA SystemMirror.

Após os sistemas de arquivos NFS fazerem parte dos grupos de recursos que pertencem a um cluster ativo do PowerHA SystemMirror, o PowerHA SystemMirror cuida da montagem cruzada e desmontagem dos sistemas de arquivos durante os eventos de cluster (como failover de um grupo de recursos que contém o sistema de arquivos para outro nó no cluster).

Se, por algum motivo, você parar os serviços de cluster e tiver que gerenciar os sistemas de arquivos NFS manualmente, os sistemas de arquivos deverão ser desmontados antes de reiniciar os serviços do cluster. Isso permite o gerenciamento dos sistemas de arquivos NFS pelo PowerHA SystemMirror após os nós se associarem ao cluster.

Recurso de servidor NFS confiável

Um cluster do PowerHA SystemMirror pode aproveitar as extensões do AIX para as funções NFS padrão que permitem manipular solicitações duplicadas corretamente e restaurar o estado de bloqueio durante o failover e a reintegração do servidor NFS.

Quando os clientes NFS usam o bloqueio de NFS para arbitrar o acesso ao sistema de arquivos NFS compartilhado, existe um limite de dois nós por grupo de recursos. Cada grupo de recursos que usa o NFS confiável contém um par de nós do PowerHA SystemMirror.

Os pares independentes de nós no cluster podem fornecer serviços NFS confiáveis. Por exemplo, em um cluster de quatro nós, é possível configurar dois pares de cliente e servidor NFS (por exemplo, Nó A e Nó B fornecem um conjunto de serviços NFS confiáveis e o Nó C e o Nó D podem fornecer outro conjunto de serviços NFS confiáveis). O Par 1 pode fornecer serviços NFS confiáveis para um conjunto de sistemas de arquivos NFS e o par 2 pode fornecer serviços NFS confiáveis para outro conjunto de sistemas de arquivos NFS. Isso é verdadeiro se a montagem cruzada do NFS estiver ou não configurada. O PowerHA SystemMirror não impõe um limite para o número de grupos de recursos ou sistemas de arquivos NFS contanto que os nós que participam dos grupos de recursos sigam as restrições descritas neste exemplo.

Especificando um endereço IP para o NFS

O nome do host também deve poder ser resolvido para um endereço IP que esteja sempre presente no nó e sempre ativo em uma interface. O nome do host não pode ser um endereço IP do serviço que pode se mover para outro nó.

Para assegurar que o endereço IP que será usado pelo NFS sempre resida no nó, é possível:

- Usar um endereço IP associado a um rótulo persistente
- Para um IPAT por meio da configuração de aliases, usar o endereço IP usado no tempo de inicialização
- Usar um endereço IP que resida em uma interface que não é controlada pelo PowerHA SystemMirror

Grupos de volumes compartilhados

Ao criar grupos de volumes compartilhados, geralmente é possível deixar o campo **Número principal** em branco e deixar que o sistema forneça um padrão. Entretanto, o NFS usa os números principais do grupo de volumes para ajudar a identificar exclusivamente sistemas de arquivos exportados. Portanto, todos os nós a serem incluídos em um grupo de recursos que contém um sistema de arquivos exportado pelo NFS devem ter o mesmo número principal para o grupo de volumes no qual o sistema de arquivos reside.

No caso de uma falha do nó, os clientes do NFS conectados a um cluster do PowerHA SystemMirror operam da mesma forma que quando um servidor NFS padrão falha e reinicializa. Os acessos aos sistemas de arquivos são interrompidos e recuperados quando os sistemas de arquivos se tornam disponíveis novamente. Entretanto, se os números principais não forem os mesmos, quando outro nó do cluster controlar o sistema de arquivos e exportar novamente o sistema de arquivos, o aplicativo cliente não será recuperado. O aplicativo cliente não se recuperará porque o sistema de arquivos exportado pelo nó parecerá diferente do exportado pelo nó com falha.

Exportando sistemas de arquivos NFS e diretórios

O processo de exportar sistemas de arquivos NFS e diretórios no PowerHA SystemMirror é diferente daquele no sistema operacional AIX.

Lembre-se dos pontos a seguir ao planejar a exportação de sistemas de arquivos NFS e diretórios do PowerHA SystemMirror:

- Sistemas de arquivos NFS e diretórios a exportar:
No AIX, você especifica os sistemas de arquivos NFS e diretórios a exportar usando o comando **smit mknfsexp** (que cria o arquivo **/etc/exports**). No PowerHA SystemMirror, você especifica sistemas de arquivos NFS e diretórios a exportar incluindo-os em um grupo de recursos no PowerHA SystemMirror.
- Opções de exportação para sistemas de arquivos exportados pelo NFS e diretórios:
Se desejar especificar opções especiais para exportar o NFS no PowerHA SystemMirror, é possível criar um arquivo **/usr/es/sbin/cluster/etc/exports**. Esse arquivo possui o mesmo formato que o arquivo regular do AIX **/etc/exports**.

Nota: O uso desse arquivo de exportações alternativo é opcional. O PowerHA SystemMirror verifica o arquivo **/usr/es/sbin/cluster/etc/exports** ao exportar um sistema de arquivos NFS ou diretório. Se houver uma entrada para o sistema de arquivos ou diretório nesse arquivo, o PowerHA SystemMirror utilizará as opções listadas. Se o sistema de arquivos NFS ou diretório para exportação não estiver

listado no arquivo ou se o arquivo alternativo não existir, o sistema de arquivos ou diretório será exportado do NFS com a opção padrão de acesso raiz para todos os nós do cluster.

- Um grupo de recursos que especifica sistemas de arquivos para exportação:

No SMIT, configure o campo **Sistemas de arquivos montados antes do IP configurado** para o grupo de recursos como **true**. Configurar o campo **Sistemas de arquivos montados antes do IP configurado** como **true** assegura que o controle de endereço IP seja executado após exportar os sistemas de arquivos. Se os endereços IP foram gerenciados primeiro, o servidor NFS rejeita as solicitações do cliente até que os sistemas de arquivos tenham sido exportados.

- Armazenamento estável para exportações de NFS versão 4:

Use o armazenamento estável das exportações do NFS versão 4 e tenha-o acessível a todos os nós participantes do grupo de recursos. O NFS versão 4 usa esse espaço do sistema de arquivos para armazenar as informações de estado relacionadas às transações do cliente NFS. As informações de estado no armazenamento estável são cruciais para as opções de failover, fallback e movimentação adequadas do grupo de recursos de um nó para outro, ao manter o estado do cliente do NFS versão 4 não afetado.

Enquanto o grupo de recursos está on-line, o local do armazenamento estável não pode ser alterado.

Referências relacionadas:

“Montagem cruzada de NFS no PowerHA SystemMirror”

Uma montagem cruzada de NFS é uma configuração de NFS do PowerHA SystemMirror específica em que cada nó no cluster pode atuar como o servidor NFS e o cliente NFS. Enquanto um sistema de arquivos está sendo exportado de um nó, o sistema de arquivos é montado com o NFS em todos os nós para o grupo de recursos, incluindo um que o esteja exportando. Outro sistema de arquivos também pode ser exportado de outro nó e pode ser montado com o NFS em todos os nós.

NFS e failover

Para que o PowerHA SystemMirror e o NFS funcionem juntos, o endereço IP para o servidor NFS devem ser configurados como um recurso em um grupo de recursos para alta disponibilidade.

Para assegurar o melhor desempenho do NFS, os sistemas de arquivos do NFS usados pelo PowerHA SystemMirror devem incluir a entrada `vers = <version number>` no campo **opções** no arquivo `/etc/filesystems`.

Referências relacionadas:

“Montagem cruzada do NFS e rótulos IP” na página 55

Para permitir a montagem cruzada do NFS, cada nó do cluster pode atuar como um cliente NFS. Cada um desses nós deve ter uma rota válida para o rótulo de IP de serviço do nó do servidor NFS. Isto é, para permitir a montagem cruzada do NFS, um rótulo IP deve existir nos nós cliente e esse rótulo de IP deve ser configurado na mesma sub-rede que o rótulo de IP do serviço do nó do servidor NFS.

Montagem cruzada de NFS no PowerHA SystemMirror

Uma montagem cruzada de NFS é uma configuração de NFS do PowerHA SystemMirror específica em que cada nó no cluster pode atuar como o servidor NFS e o cliente NFS. Enquanto um sistema de arquivos está sendo exportado de um nó, o sistema de arquivos é montado com o NFS em todos os nós para o grupo de recursos, incluindo um que o esteja exportando. Outro sistema de arquivos também pode ser exportado de outro nó e pode ser montado com o NFS em todos os nós.

Quando seu ambiente usa o NFS versão 2 e versão 3 para exportar em grupos de recursos, o recurso de montagem cruzada é restrito a apenas dois grupos de recursos do nó. Se o grupo de recursos contiver apenas exportações do NFS versão 4 (ou posterior), o recurso de montagem cruzada é estendido a até qualquer número de nós que estejam suportando um grupo de recursos.

Cada nó no grupo de recursos é parte de uma configuração de cluster de controle mútuo (ou ativo-ativo), fornecendo e montando um sistema de arquivos do NFS.

Por padrão, os grupos de recursos que contêm os sistemas de arquivos do NFS exportados automaticamente faz a montagem cruzada desses sistemas de arquivos (se a exportação e a importação estiverem configuradas) como a seguir:

- No nó que está hospedando atualmente o grupo de recursos, todos os sistemas de arquivos do NFS no grupo são exportados pelo NFS.
- Cada nó que pode hospedar esse NFS do grupo de recursos monta todos os sistemas de arquivos do NFS no sistema de arquivos.

Os aplicativos acessam os sistemas de arquivos do NFS em qualquer nó que seja parte do grupo de recursos.

Quando um failover ocorre para um grupo de recursos que seja configurado com o controle do endereço IP, o sistema de arquivos do NFS é montado localmente pelo nó de controle e exportado novamente. Todos os outros nós no grupo de recursos mantêm sua montagem de sistema de arquivos do NFS.

Em uma configuração de montagem cruzada do NFS, qualquer montagem do NFS que seja definida em um grupo de recursos deve ter uma exportação de NFS correspondente em um grupo de recursos. Se suas montagens cruzadas do NFS não seguirem esta configuração, a mensagem a seguir será exibida:

```
claddress: WARNING: NFS mounts were specified for the resource group
'<RG name>'; however no NFS exports have been specified.
```

Se um campo de montagem cruzada do NFS contiver um valor, o arquivo exportado pelo NFS correspondente também deverá conter um valor.

Montagem cruzada do NFS e rótulos IP:

Para permitir a montagem cruzada do NFS, cada nó do cluster pode atuar como um cliente NFS. Cada um desses nós deve ter uma rota válida para o rótulo de IP de serviço do nó do servidor NFS. Isto é, para permitir a montagem cruzada do NFS, um rótulo IP deve existir nos nós cliente e esse rótulo de IP deve ser configurado na mesma sub-rede que o rótulo de IP do serviço do nó do servidor NFS.

Se os nós do cliente NFS tiverem rótulos de IP do serviço na mesma rede, isso não será um problema. Entretanto, em certas configurações de cluster, é necessário criar uma rota válida.

Formas de criar uma rota para o servidor NFS:

A forma mais fácil de assegurar o acesso ao servidor NFS é ter um rótulo de IP no nó cliente que está na mesma sub-rede que o rótulo de IP do serviço do nó do servidor NFS.

Para criar uma rota válida entre o nó cliente NFS e o nó que está exportando o sistema de arquivos, é possível configurar seu ambiente das formas a seguir:

- Uma NIC separada com um rótulo de IP configurado na rede de IP do serviço e sub-rede
- Um rótulo de IP de nó persistente na rede e sub-rede do IP do serviço.

Esteja ciente de que essas soluções não forneçam permissões-raiz automáticas aos sistemas de arquivos porque as opções de exportação para os sistemas de arquivos NFS que estão configuradas no PowerHA SystemMirror por padrão.

Para permitir o acesso do nível raiz aos sistemas de arquivos montados pelo NFS no nó cliente, inclua todos os rótulos e endereços IP do nó na opção `root = no` no arquivo de exportações de cluster: `/usr/es/sbin/cluster/etc/exports`. É possível fazer isso em um nó, porque sincronizar os recursos de cluster propaga essas informações para os outros nós do cluster.

Referências relacionadas:

“Exportando sistemas de arquivos NFS e diretórios” na página 53

O processo de exportar sistemas de arquivos NFS e diretórios no PowerHA SystemMirror é diferente

daquele no sistema operacional AIX.

Criando e configurando o armazenamento estável para exportações do NFS versão 4:

O armazenamento estável é um espaço do sistema de arquivos que é usado para salvar as informações de estado pelo servidor NFS versão 4. Isso é muito crucial para manter as informações de estado do cliente do NFS versão 4 para facilitar um failover/fallback/movimentação tranquila e transparente do Grupo de recursos de um nó para outro.

Requisitos do armazenamento estável:

- O tamanho sugerido é 512 MB.
- Sugere-se que tenha um sistema de arquivos dedicado para o armazenamento estável, mas um subdiretório de um sistema de arquivos existente é aceitável.
- O sistema de arquivos e o grupo de volumes em que o Armazenamento estável reside deve ser parte do grupo de recursos.
- O PowerHA SystemMirror faz uma tentativa de melhor esforço para verificar se o caminho especificado para o armazenamento estável pertence a um sistema de arquivos no grupo de recursos; no entanto, essas verificações não podem levar links simbólicos em consideração uma vez que os sistemas de arquivos não podem ser montados localmente quando a verificação ocorre. Evitar usar links simbólicos que podem interferir na verificação no PowerHA SystemMirror.
- Assegure-se de que o diretório de armazenamento estável esteja vazio (livre de quaisquer dados anteriores) antes de incluí-lo no grupo de recursos.
- Enquanto o armazenamento estável deve ser armazenado nos sistemas de arquivos gerenciados pelo grupo de recursos, ele não deve ser armazenado em um diretório do NFS que é exportado pelo grupo de recursos.
- O assistente de configuração oferece uma opção AUTO_SELECT. Se você selecionar esta opção, o PowerHA SystemMirror usará um grupo de volumes na lista de grupos de volumes que são parte do grupo de recursos fornecido. O PowerHA SystemMirror cria um volume lógico e um sistema de arquivos para usar como local de armazenamento estável.

Exemplo: configuração de montagem cruzada do NFS de dois nós:

Neste exemplo, o Nó A hospeda atualmente um grupo de recursos não simultâneos, RG1, que inclui /fs1 como um sistema de arquivos NFS e service1 como um rótulo de IP de serviço.

Neste exemplo, o Nó B hospeda atualmente um grupo de recursos não simultâneos, RG2, que inclui /fs2 conforme o sistema de arquivos NFS exportado e service2 como um rótulo de IP de serviço. Na reintegração, /fs1 é transmitido de volta para o Nó A, montado localmente e exportado. O Nó B monta-o sobre o NFS novamente.

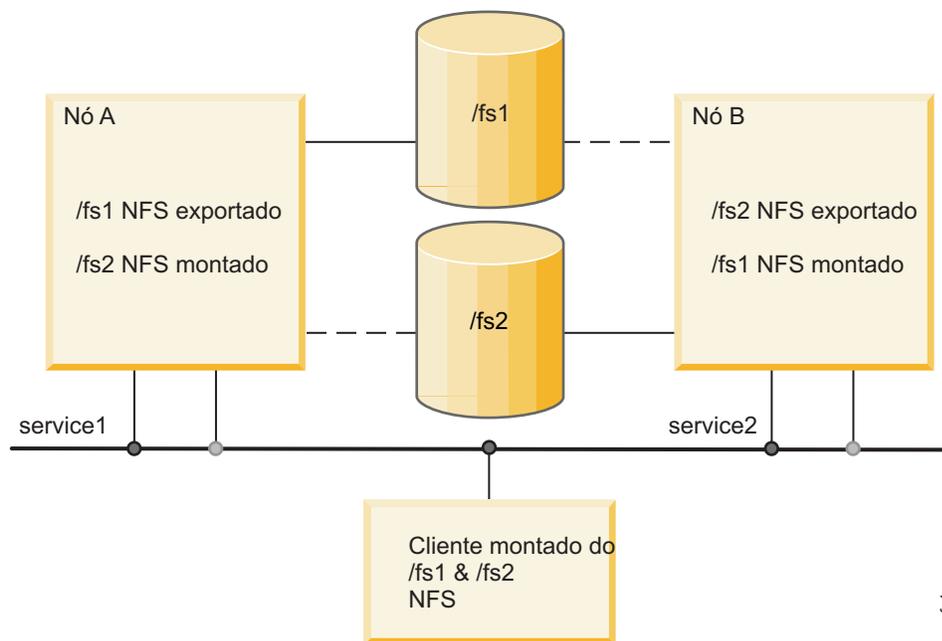


Figura 10. Montagem cruzada do NFS de dois nós

Os dois grupos de recursos seriam definidos no SMIT como a seguir.

Grupo de recursos	RG1	RG2
Nomes de nós participantes	Nó A Nó B	Nó B Nó A
Sistemas de arquivos Os sistemas de arquivos a serem montados localmente pelo nó que possui atualmente o grupo de recursos.	/fs1	/fs2
Sistemas de arquivos a exportar O sistema de arquivos para exportação do NFS pelo nó que atualmente possui o grupo de recursos. O sistema de arquivos é um subconjunto do sistema de arquivos listado acima.	/fs1	/fs2
Sistemas de arquivos para montar NFS Os sistemas de arquivos e diretórios a serem montados pelo NFS por todos os nós no grupo de recursos. O primeiro valor é o ponto de montagem de NFS. O segundo valor é o ponto de montagem local.	/mnt1;/fs1	/mnt2;/fs2
Sistemas de arquivos montados antes do IP configurado	true	true

Neste cenário:

- O Nó A monta e exporta /fs1 localmente, então, monta no /mnt1.
- O Nó B monta o NFS /fs1, em /mnt1 a partir do Nó A.

A configuração de um grupo de recursos como esse assegura o comportamento padrão de NFS de nó para nó.

Quando o Nó A falha, o Nó B fecha quaisquer arquivos abertos no Nó A: /fs1, desmonta-o, monta-o localmente e o exporta novamente para clientes que estão esperando.

Após o controle, o Nó B tem:

- /fs2 montado localmente
- /fs2 exportado por NFS
- /fs1 montado localmente
- /fs1 exportado por NFS
- service1:/fs1 montado por NFS sobre /mnt1
- service2:/fs2 montado por NFS sobre /mnt2b

Ambos os grupos de recursos contêm ambos os nós como possíveis proprietários dos grupos de recursos.

Controle de grupo de recursos com sistemas de arquivos do NFS de montagem cruzada

Esta seção descreve como configurar grupos de recursos não simultâneos com sistemas de arquivos do NFS de montagem cruzada para que os sistemas de arquivos do NFS sejam manipulados corretamente durante o controle e a reintegração. Além disso, os grupos de recursos não simultâneos suportam a montagem automática do NFS entre diversos servidores durante o failover.

Configurando o ponto de montagem do NFS a partir do ponto de montagem local:

O PowerHA SystemMirror manipula a montagem do NFS em grupos de recursos não simultâneos de várias formas.

Estão inclusos:

- O nó que possui atualmente o grupo de recursos monta o sistema de arquivos sobre o ponto de montagem local do sistema de arquivos e esse nó exporta o sistema de arquivos do NFS.
- Todos os nós no grupo de recursos (incluindo o proprietário atual do grupo) montam o sistema de arquivos do NFS sobre um ponto de montagem diferente.

Portanto, o proprietário do grupo possui o sistema de arquivos montado duas vezes. Um sistema de arquivos é montado como uma montagem local e o outro é uma montagem do NFS.

Nota: O ponto de montagem NFS deve estar fora da árvore de diretórios do ponto de montagem local.

Como o IPAT é usado em grupos de recursos que possuem sistemas de arquivos montados por NFS, os nós não desmontarão e remontarão os sistemas de arquivos do NFS durante um failover. Quando o grupo de recursos cai sobre um novo nó, o nó de aquisição monta localmente o sistema de arquivos e o NFS o exporta. (O sistema de arquivos montado pelo NFS está temporariamente indisponível para os nós de cluster durante o failover.) Assim que o novo nó adquire o rótulo IPAT, o acesso ao sistema de arquivos do NFS é restaurado.

Todos os aplicativos devem se referir ao sistema de arquivos por meio do sistema de arquivos montado por NFS. Se os aplicativos usados devem sempre referenciar o sistema de arquivos pelo mesmo nome do ponto de montagem, é possível alterar o ponto de montagem para a montagem do sistema de arquivos local (por exemplo, altere-o para a montagem de point_local e use o ponto de montagem local anterior como o novo ponto de montagem do NFS).

Opções de montagem de NFS padrão para o PowerHA SystemMirror:

As opções padrão usadas pelo PowerHA SystemMirror ao executar montagens de NFS são hard, intr.

Para configurar montagens soft ou quaisquer outras opções nas montagens de NFS:

1. Digite `smit mknfsmnt`.
2. No campo **MONTAR agora, incluir entrada em /etc/filesystems ou ambos?**, selecione a opção **sistemas de arquivos**.

3. No campo **A entrada /etc/filesystems montará o diretório na REINICIALIZAÇÃO do sistema**, aceite o valor padrão **no**.
Este procedimento inclui as opções escolhidas na entrada **/etc/filesystems** criada. Os scripts do PowerHA SystemMirror leem essa entrada para usar quaisquer opções que possa ter selecionado.

Criando e configurando pontos de montagem NFS em clientes:

Um ponto de montagem de NFS é necessário para montar um sistema de arquivos usando o NFS. Em um grupo de recursos não simultâneos, todos os nós no grupo de recursos montam o sistema de arquivos no NFS. Você cria um ponto de montagem do NFS em cada nó no grupo de recursos. O ponto de montagem NFS deve estar fora da árvore de diretórios do ponto de montagem local.

Após o ponto de montagem do NFS ser criado em todos os nós no grupo de recursos, configure o atributo **Sistema de arquivos do NFS para montagem de NFS** para o grupo de recursos.

Para criar pontos de montagem NFS e para configurar o grupo de recursos para a montagem NFS:

1. Em cada nó no grupo de recursos, crie um ponto de montagem NFS, executando o seguinte comando:

```
mkdir /mount point
```

em que *mount point* é o nome do ponto de montagem de NFS local sobre o qual o sistema de arquivos remoto é montado.
2. No painel do SMIT **Alterar/mostrar recursos e atributos para um grupo de recursos**, o campo **Sistema de arquivos para montagem de NFS** deve especificar ambos os pontos de montagem. Especifique o ponto de montagem nfs e o ponto de montagem local, separando os dois com um ponto-e-vírgula. Por exemplo:

```
/nfspoint;/localpoint
```

Se houver mais entradas, separe-as com um espaço:

```
/nfspoint1;/local1 /nfspoint2;/local2
```
3. Opcional: Se houver pontos de montagem aninhados, aninhe os pontos de montagem de NFS da mesma forma que os pontos de montagem local para que correspondam corretamente.
4. Opcional: Quando houver montagem cruzada de sistemas de arquivos do NFS, configure o campo **Sistemas de arquivos montados antes do IP configurado** no SMIT para o grupo de recursos como **true**.

Planejando grupos de recursos

Estes tópicos descrevem como planejar grupos de recursos em um cluster do PowerHA SystemMirror.

Visão geral para grupos de recursos

O PowerHA SystemMirror organiza recursos em grupos de recursos. Cada grupo de recursos é manipulado como uma unidade que contém recursos compartilhados, como rótulos de IP, aplicativos, sistemas de arquivos e grupos de volume. Você define as políticas para cada grupo de recursos que define quando e como ele será adquirido ou liberado.

No Planejamento de cluster inicial, você fez opções preliminares sobre as políticas de grupo de recursos e a prioridade de controle para cada nó nas listas de nó do grupo de recursos. Nesta seção, você faz o seguinte:

- Identifique os recursos individuais que constituem cada grupo de recursos.
- Para cada grupo de recursos, identifique qual tipo de grupo é: simultâneo ou não simultâneo.
- Defina a lista de nós participantes para os grupos de recursos. A lista de nós consiste nos nós designados a participar do controle de um determinado grupo de recursos.
- Identifique a política de inicialização, failover e fallback do grupo de recursos.

- Identifique aplicativos e seus grupos de recursos para os quais deseja configurar dependências de local, dependências pai-filho ou ambas.
- Identifique as políticas de gerenciamento entre sites dos grupos de recursos. Há recursos replicados a serem considerados?
- Identifique outros atributos e políticas de tempo de execução para refinar o comportamento do grupo de recursos.

Os termos a seguir são usados nesta seção:

- *Lista de nós participantes*. Uma lista de nós que podem hospedar um determinado grupo de recursos, conforme definido nos nomes de nó participantes para um grupo de recursos no SMIT.
Esteja ciente de que a combinação das diferentes políticas de grupo de recursos e as condições atuais de cluster também afetam a colocação do grupo de recursos nos nós no cluster.
- *Nó inicial (ou o nó de maior prioridade para este grupo de recursos)*. O primeiro nó que está listado na lista de nós participantes para quaisquer grupos de recursos não simultâneos.

Os grupos de recursos do PowerHA SystemMirror suportam sistemas de arquivos de NFS.

Referências relacionadas:

“Planejamento de cluster inicial” na página 5

Esta seção descreve as etapas iniciais que você executa para planejar um cluster do PowerHA SystemMirror para tornar os aplicativos altamente disponíveis.

“Montagem cruzada do NFS e rótulos IP” na página 55

Para permitir a montagem cruzada do NFS, cada nó do cluster pode atuar como um cliente NFS. Cada um desses nós deve ter uma rota válida para o rótulo de IP de serviço do nó do servidor NFS. Isto é, para permitir a montagem cruzada do NFS, um rótulo IP deve existir nos nós cliente e esse rótulo de IP deve ser configurado na mesma sub-rede que o rótulo de IP do serviço do nó do servidor NFS.

Regras gerais para recursos e grupos de recursos

Há algumas regras gerais e restrições para recursos e grupos de recursos.

As regras e restrições a seguir se aplicam aos recursos e grupos de recursos:

- Para que o PowerHA SystemMirror mantenha um recurso de cluster altamente disponível, ele deve ser parte de um grupo de recursos. Se desejar que um recurso seja mantido separado, defina um grupo para esse recurso sozinho. Um grupo de recursos pode ter um ou mais recursos definidos.
- Um recurso não pode ser incluso em mais de um grupo de recursos.
- Os componentes de um grupo de recursos devem ser exclusivos. Coloque o aplicativo junto com os recursos que ele requer no mesmo grupo de recursos.
- Os rótulos de IP de serviço, grupos de volumes e nomes de grupos de recursos devem ser exclusivos no cluster e distintos entre si. O nome de um recurso deve estar relacionado ao aplicativo a que atende, bem como qualquer dispositivo correspondente, como `websphere_service_address`.
- Se você incluir o mesmo nó nas listas de nós participantes para mais de um grupo de recursos, assegure-se de que o nó tenha a memória e as interfaces de rede, necessárias para gerenciar todos os grupos de recursos simultaneamente.

Tipos de grupos de recursos: simultâneo e não simultâneo

Para categorizar e descrever o comportamento do grupo de recursos, você deve primeiro dividir os grupos de recursos em dois tipos: simultâneo e não simultâneo.

Grupos de recursos simultâneos

Um grupo de recurso simultâneo pode ficar on-line em diversos nós. Todos os nós na lista de nós do grupo de recursos adquirem esse grupo de recursos quando unem-se ao cluster. Não há prioridades entre

os nós. Os grupos de recursos simultâneos podem ser configurados para executar em todos os nó no cluster.

Os únicos recursos incluídos em um grupo de recurso simultâneo são grupos de volumes com volumes lógicos brutos, discos brutos e controladores de aplicativos que usam os discos. O dispositivo no qual essas entidades de armazenamento lógico estão definidas deve suportar o acesso simultâneo.

Os grupos de recursos simultâneos têm a política de inicialização On-line em todos os nós disponíveis e não executam fallover ou fallback de um nó para outro.

Grupos de recursos não simultâneos

Os grupos de recursos não simultâneos podem ficar on-line em diversos nós. É possível definir uma variedade de políticas de inicialização, fallover e fallback para esses grupos de recursos.

É possível ajustar o comportamento do grupo de recursos não simultâneos para preferências de nós durante uma inicialização de nó, fallover de grupo de recursos para outro nó no caso de uma falha de nó ou quando o grupo de recursos realiza fallback para o nó de reintegração.

Referências relacionadas:

“Atributos de grupo de recursos para inicialização, fallover e fallback” na página 62

Cada atributo afeta a inicialização do grupo de recursos, o fallover do grupo de recursos para outro nó no caso de uma falha de nó ou fallback do grupo de recursos para o nó de reintegração.

Políticas de grupo de recursos para inicialização, fallover e fallback

Os comportamentos do grupos de recursos são separados em três tipos de políticas de nós.

Essas políticas são:

- A política de *inicialização* define em qual nó o grupo de recursos será ativado quando um nó se unir ao cluster e o grupo de recursos não estiver ativo em nenhum nó.
- A política de *fallover* define para qual nó o grupo de recursos realizará fallover quando o grupo de recursos tiver que deixar o nó onde está atualmente on-line devido a uma condição de falha (ou se você parar os serviços de cluster em um nó usando a opção de fallover).
- A política de *fallback* define para qual nó o grupo de recursos realizará fallback quando um nó se unir e o grupo de recursos já estiver ativo em outro nó.

O PowerHA SystemMirror permite configurar apenas combinações válidas de comportamentos de inicialização, fallover e fallback para grupos de recursos. A tabela a seguir resume os comportamentos básicos de inicialização, fallover e fallback que pode configurar para grupos de recursos no PowerHA SystemMirror.

Comportamento de inicialização	Comportamento de fallover	Comportamento de fallback
On-line apenas no nó inicial (primeiro nó na lista de nós)	Qualquer um destes: <ul style="list-style-type: none">• Fallover para o próximo nó de prioridade na lista• Fallover usando a prioridade do nó dinâmico	Qualquer um destes: <ul style="list-style-type: none">• Nunca realizar fallback• Realizar fallback para um nó de maior prioridade na lista
On-line usando a política de distribuição de nós	Qualquer um destes: <ul style="list-style-type: none">• Fallover para o próximo nó de prioridade na lista• Fallover usando a prioridade do nó dinâmico	Nunca realizar fallback

Comportamento de inicialização	Comportamento de fallover	Comportamento de fallback
On-line no primeiro nó disponível	Qualquer um destes: <ul style="list-style-type: none"> Fallover para o próximo nó de prioridade na lista Fallover usando a prioridade do nó dinâmico 	Qualquer um destes: <ul style="list-style-type: none"> Nunca realizar fallback Realizar fallback para um nó de maior prioridade na lista
On-line em todos os nós disponíveis	Colocar off-line (no nó de erro apenas)	Nunca realizar fallback

Além das políticas de nó descritas na tabela anterior, outros problemas podem determinar os grupos de recursos que um nó adquire.

Referências relacionadas:

“Planejamento para eventos de cluster” na página 82

Estes tópicos descrevem os eventos do cluster do PowerHA SystemMirror.

Atributos de grupo de recursos

Esta seção fornece uma visão geral dos atributos de grupo de recursos que podem ser usados para ajustar as políticas de inicialização, fallover e fallback dos grupos de recursos.

Atributos de grupo de recursos para inicialização, fallover e fallback

Cada atributo afeta a inicialização do grupo de recursos, o fallover do grupo de recursos para outro nó no caso de uma falha de nó ou fallback do grupo de recursos para o nó de reintegração.

A tabela a seguir resume quais políticas de inicialização, fallover ou fallback do grupo de recursos serão afetadas por um determinado atributo ou política de tempo de execução. Nem todos os grupos de recursos disponíveis estão listados abaixo.

Atributo	Política de inicialização	Política de Fallover	Política de fallback
Tempo de acomodação	X		
Política de distribuição de nó	X		
Prioridade dinâmica do nó		X	
Cronômetro de fallback atrasado			X
Dependência pai e filho de grupos de recursos	X	X	X
Dependências de local de grupos de recursos	X	X	X

Referências relacionadas:

“Grupos de recursos dependentes pais e filhos” na página 66

Os aplicativos relacionados em diferentes grupos de recursos são configurados para serem processados na ordem lógica.

“Dependências de local do grupo de recursos” na página 67

Certos aplicativos em diferentes grupos de recursos ficam on-line juntos em um nó ou em diferentes nós.

Informações relacionadas:

Recursos e grupos de recursos do PowerHA SystemMirror

Tempo de acomodação para inicialização

É possível modificar o comportamento de inicialização de um grupo de recursos especificando um tempo de acomodação para um grupo de recursos que está atualmente off-line e possui uma política de inicialização de On-line no Primeiro nó disponível.

Com o tempo de acomodação especificado, é possível evitar ter um grupo de recursos ativado no primeiro nó disponível quando diversos nós estiverem iniciando serviços de cluster ao mesmo tempo. Além disso, um nó de prioridade mais alta para o grupo de recursos pode se associar ao cluster durante este período.

O tempo de acomodação permite que o gerenciador do cluster aguarde uma quantidade especificada de tempo antes de ativar um grupo de recursos. Use este atributo para assegurar que um grupo de recursos não seja devolvido entre os nós, enquanto os nós com prioridade aumentando para o grupo de recursos sejam colocados on-line.

Se o nó que está iniciando for o primeiro nó na lista de nós para este grupo de recursos, o período de tempo de acomodação será ignorado e o PowerHA SystemMirror tentará imediatamente adquirir o grupo de recursos nesse nó.

O tempo de acomodação possui as características a seguir:

- Afeta somente esses grupos de recursos que estão atualmente off-line e para os quais você especificou a política de inicialização para ficar On-line no Primeiro nó disponível. Você configura um tempo de acomodação para todos esses grupos de recursos.
- Ativa quando o primeiro nó que pode adquirir o grupo de recursos se associa ao cluster, a menos que esse seja o primeiro nó na lista de nós (o tempo de acomodação é ignorado e o grupo é adquirido).
Se o primeiro nó que se associa ao cluster e puder adquirir em potencial o grupo de recursos falhar, esta condição poderá cancelar o tempo de acomodação ou reconfigurá-lo.
- Atraza a ativação do grupo durante eventos **node_up**, em caso de nós de maior prioridade se associarem ao cluster.

Nota: Se um período de tempo de acomodação for especificado para um grupo de recursos e um grupo de recursos estiver atualmente no estado ERRO, o gerenciador do cluster aguardará o período de tempo de acomodação antes de tentar colocar o grupo de recursos on-line durante um evento **node_up**.

Reintegração de nó com tempo de acomodação configurado

Em geral, quando um nó se associa ao cluster, ele pode adquirir grupos de recursos. A lista a seguir descreve a função do tempo de acomodação neste processo:

- Se o nó for o nó de maior prioridade para um grupo de recursos específico, o nó adquirirá imediatamente esse grupo de recursos e o tempo de acomodação será ignorado. Esta é a única circunstância sob a qual o PowerHA SystemMirror ignora a configuração.
- Se o nó puder adquirir alguns grupos de recursos, mas não for o nó de maior prioridade para esses grupos, os grupos de recursos não serão adquiridos nesse nó. Em vez disso, eles aguardam durante o intervalo de tempo de acomodação para ver se um nó de maior prioridade se associa ao cluster.

Quando o intervalo de tempo de acomodação termina, o PowerHA SystemMirror move o grupo de recursos para o nó de maior prioridade que está atualmente disponível e pode usar o grupo de recursos. Se o PowerHA SystemMirror não localizar nós apropriados, o grupo de recursos permanecerá off-line.

Política de distribuição de nó

É possível configurar um comportamento de inicialização de um grupo de recursos para usar a política de distribuição de nó durante a inicialização. Essa política assegura que apenas um grupo de recursos que possua essa política ativada seja adquirido em um nó durante a inicialização.

É possível usar a *política de distribuição de nó* para inicialização de cluster a fim de assegurar que o PowerHA SystemMirror ative apenas um grupo de recursos que possua essa política ativada em cada nó. Essa política ajuda a distribuir seus aplicativos de uso intenso de CPU em diferentes nós.

Os fatos sobre a política de distribuição de nó são:

- Se planejar usar uma única rede de adaptador que será configurada com o IPAT por meio da substituição, a política de inicialização para seu grupo de recursos deverá ser configurada para On-line usando a Política de distribuição.
- Se dois grupos de recursos com essa política ativada estiverem off-line no momento em que um determinado nó se associar, apenas um dos grupos de recursos será adquirido em um nó. O PowerHA SystemMirror dá preferência ao grupo de recursos que tiver menos nós na lista de nós e classificará a lista de grupos de recursos de forma independente.
- Se um dos grupos de recursos for um grupo de recursos pai (tiver um grupo de recursos filho), o PowerHA SystemMirror dará preferência ao grupo de recursos pai e será ativado em um nó.
- Para assegurar que seus grupos de recursos sejam distribuídos não só na inicialização, mas também para eventos de recuperação (failover e fallback), use as dependências de local.

Referências relacionadas:

“Dependências do grupo de recursos” na página 65

O PowerHA SystemMirror oferece uma ampla variedade de configurações em que é possível especificar os relacionamentos entre grupos de recursos que deseja manter na inicialização, failover e fallback.

Política dinâmica de prioridade de nó

É possível configurar o comportamento de failover de um grupo de recursos para usar a prioridade de nó dinâmico. Permite usar uma variável de recurso predefinida Reliable Scalable Cluster Technology (RSCT) como **menor carga de CPU**, para selecionar o nó de controle ou uma variável de prioridade de nó dinâmico definida pelo usuário, como `cl_highest_udscript_rc` para selecionar o nó de controle.

Configurar uma política dinâmica de prioridade de nó permite usar uma variável de recurso predefinida Resource Monitoring and Control (RMC), como **menor carga de CPU**, para selecionar o nó de controle. Com uma política de prioridade dinâmica ativada, a ordem da lista de nós de controle é determinada pelo estado do cluster no horário do evento, conforme medido pela variável de recurso do RMC selecionada. É possível configurar diferentes políticas para diferentes grupos ou a mesma política para vários grupos.

Outra opção é usar uma variável de prioridade de nó dinâmico definida pelo usuário, com `cl_highest_udscript_rc`. Se usar essa opção, você deve fornecer um script e um valor de tempo limite de execução, que é chamado em todos os nós de failover candidato no horário do evento. Os valores de retorno são coletados de cada um e o nó de controle é selecionado com base nos valores de retorno do script e a variável de prioridade de nó dinâmico selecionada.

Se decidir definir políticas dinâmicas de prioridade de nó usando variáveis de recurso do RMC para determinar o nó de failover para um grupo de recursos, considere os pontos a seguir:

- A política dinâmica de prioridade de nó é mais útil em um cluster em que todos os nós têm energia de processamento e memória iguais.
- A política dinâmica de prioridade de nó é irrelevante para clusters de menos de três nós.
- A política dinâmica de prioridade de nó é irrelevante para grupos de recursos simultâneos.
- A política dinâmica de prioridade de nó não é suportada quando os sites são configurados.

Lembre-se de que selecionar um nó de controle também depende dessas condições como a disponibilidade de uma interface de rede nesse nó.

Cronômetro de fallback atrasado

É possível configurar o comportamento de fallback do grupo de recursos para que ocorra em uma das vezes recorrentes predefinidas (diária, semanal, mensal e anualmente ou em uma data e hora específica) especificando e designando um cronômetro de fallback atrasado.

É possível usar um *cronômetro de fallback atrasado* para configurar o tempo para que um grupo de recursos realize fallback para um nó de maior prioridade. É possível configurar o comportamento de fallback para que um grupo de recursos ocorra em um horário recorrente predefinido (diária, semanal, mensalmente ou em uma data específica).

O cronômetro de fallback atrasado possui as características a seguir:

- Especifica o tempo quando um grupo de recursos que está on-line e reside em um nó de prioridade não inicial ou baixa realiza fallback para seu nó inicial ou um nó de maior prioridade.
- Afeta o movimento do grupo de recursos para outro nó. Por exemplo, se você mover o grupo de recursos não simultâneos (que possui um atributo de cronômetro de fallback) para outro nó usando o utilitário de gerenciamento de grupo de recursos (RGM) (**clRGMmove**), o grupo permanece no nó de destino (a menos que reinicialize o cluster, o que raramente é feito). Se o nó de destino ficar inativo e for reintegrado, o grupo de recursos também realiza fallback para este nó no horário especificado.

Reintegração de nó com um cronômetro de fallback atrasado definido

O grupo de recursos não realiza fallback imediatamente para seu nó de maior prioridade sob a condição a seguir.

- Você configurou um cronômetro de fallback atrasado para um grupo de recursos.
- Um nó de maior prioridade se associa ao cluster.

No momento especificado no atributo do Cronômetro de fallback atrasado, um dos dois cenários ocorre:

- *Um nó de maior prioridade é localizado.* Se um nó de maior prioridade estiver disponível para o grupo de recursos, o PowerHA SystemMirror tentará mover o grupo de recursos para este nó quando o cronômetro de fallback expirar. Se a aquisição for bem-sucedida, o grupo de recursos será adquirido nesse nó.

Entretanto, se a aquisição do grupo de recursos no nó falhar, o PowerHA SystemMirror tentará mover o grupo de recursos para o próximo nó de maior prioridade na lista de nós do grupo e assim por diante. Se a aquisição do grupo de recursos no último nó que estiver disponível falhar, o grupo de recursos entrará em um estado de erro. Você deve tomar uma ação para corrigir e colocar um grupo de recursos novamente on-line.

- *Um nó de maior prioridade não foi localizado.* Se não houver nós de maior prioridade disponíveis para um grupo de recursos, o grupo de recursos permanecerá on-line no mesmo nó até que o cronômetro de fallback expire novamente. Por exemplo, se um cronômetro de fallback diário expirar às 23h e não houver nós de maior prioridade disponíveis para que o grupo de recursos tenha o fallback executado, o cronômetro de fallback ocorrerá novamente no próximo horário às 23h.

Um cronômetro de fallback que é configurado para uma data específica não ocorre novamente.

Dependências do grupo de recursos

O PowerHA SystemMirror oferece uma ampla variedade de configurações em que é possível especificar os relacionamentos entre grupos de recursos que deseja manter na inicialização, fallover e fallback.

É possível configurar as dependências dos grupos de recursos a seguir:

- Grupos de recursos dependentes pais e filhos. Os aplicativos relacionados e outros recursos em diferentes grupos de recursos são configurados para serem processados na ordem correta.
- Inicie após as dependências. Especifique se um grupo de recursos pode ser iniciado apenas após outro grupo de recursos ou grupos estarem ativos no cluster.
- Pare após as dependências. Especifique se um grupo de recursos pode ser interrompido apenas após outro grupo de recursos ou grupos serem colocados off-line.
- Dependências de local do grupo de recursos. Certos aplicativos em diferentes grupos de recursos ficam on-line juntos em um nó ou em diferentes nós.

Mantenha os pontos a seguir em mente ao planejar como configurar essas dependências:

- Embora, por padrão, todos os grupos de recursos sejam processados em paralelo, o PowerHA SystemMirror processa os grupos de recursos dependentes de acordo com a ordem ditada pela dependência e não necessariamente em paralelo. As dependências de grupo de recursos são honradas em todo cluster e substituem qualquer customização para ordem serial de processamento de quaisquer grupos de recursos inclusos na dependência.
- As dependências entre os grupos de recursos oferecem uma forma previsível e confiável de construir clusters com aplicativo com multicamadas.

As limitações a seguir se aplicam a configurações que combinam dependências. A verificação falhará se você não tiver apenas um grupo de recursos pertencente a uma dependência On-line no mesmo nó definida e uma dependência On-line em diferentes nós definida ao mesmo tempo.

Grupos de recursos dependentes pais e filhos:

Os aplicativos relacionados em diferentes grupos de recursos são configurados para serem processados na ordem lógica.

A configuração de uma dependência de grupo de recursos permite um melhor controle para clusters com aplicativos com multicamadas em que um aplicativo depende da inicialização bem-sucedida de outro aplicativo e ambos os aplicativos são adquiridos para serem mantidos altamente disponíveis com o PowerHA SystemMirror.

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

- Se o grupo de recursos A depender do grupo de recursos B, o grupo de recursos B deverá ser colocado on-line antes do grupo de recursos A ser adquirido em qualquer nó no cluster. Observe que o grupo de recursos A está definido como um *grupo de recursos filho* e o grupo de recursos B como um *grupo de recursos pai*.
- Se o grupo de recursos filho A depender do grupo de recursos pai B, durante uma inicialização de ó ou reintegração de nó, o grupo de recursos filho A não poderá ficar on-line antes do grupo de recursos pai B ficar on-line. Se o grupo de recursos pai B for colocado off-line, o grupo de recursos filho A será colocado off-line primeiro, uma vez que depende do grupo de recursos B.

As configurações de negócios que usam aplicativos com multicamadas podem usar grupos de recursos dependentes pai-filho. Por exemplo, o banco de dados deve estar on-line antes do controlador de aplicativo. Neste caso, o banco de dados é movido para um nó diferente, o grupo de recursos que contém o controlador de aplicativo teria de ser desativado e ativado novamente em qualquer nó no cluster.

Se um grupo de recursos filho contiver um aplicativo que depende de recursos no grupo de recursos pai e o grupo de recursos pai cai para outro nó, o grupo de recursos filho é temporariamente interrompido e automaticamente reiniciado. Da mesma forma, se o grupo de recursos filho for simultâneo, o PowerHA SystemMirror o colocará off-line temporariamente em todos os nós e novamente on-line em todos os nós disponíveis. Se o failover do grupo de recursos pai não for bem-sucedido, os grupos de recursos pai e filho entrarão em estado ERROR.

Considere o seguinte ao planejar grupos de recursos dependentes pai-filho:

- Planeje os aplicativos que precisa manter altamente disponíveis e considere se seu ambiente de negócios requer que um aplicativo esteja em execução antes de outro aplicativo poder ser iniciado.
- Assegure-se de que esses aplicativos que requerem o sequenciamento sejam incluídos em diferentes grupos de recursos. Desta forma, é possível estabelecer dependências entre esses grupos de recursos.
- Planeje monitores de aplicativos para cada aplicativo que está planejando incluir em um grupo de recursos pai ou filho. Para um aplicativo em um grupo de recursos pai, configure um monitor no modo **inicialização de monitoramento**.

Para minimizar a chance de perda de dados durante o processo de parada e reinício do aplicativo, customize seus scripts do controlador de aplicativo para assegurar que quaisquer dados não confirmados sejam armazenados em um disco compartilhado temporariamente durante o processo de

parada do aplicativo e lidos de volta no aplicativo durante o processo de reinício do aplicativo. É importante usar um disco compartilhado, pois o aplicativo pode ser reiniciado em um nó que não aquele no qual foi interrompido.

Referências relacionadas:

“Considerações de planejamento para aplicativos com multicamadas” na página 15

As configurações de negócio que usam aplicativos com multicamadas podem usar grupos de recursos dependentes pais e filhos. Por exemplo, o banco de dados deve estar on-line antes do controlador de aplicativo. Neste 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 teria de ser desativado e ativado em qualquer nó no cluster.

Dependências de local do grupo de recursos:

Certos aplicativos em diferentes grupos de recursos ficam on-line juntos em um nó ou em diferentes nós.

Se ocorrerem falhas durante esse tempo, o PowerHA SystemMirror distribuirá grupos de recursos para que permaneçam disponíveis, mas não necessariamente nos nós originalmente especificados, a menos que tenham o mesmo nó inicial e o mesmo failover e as políticas de fallback.

A dependência de local do grupo de recursos oferece uma forma explícita de especificar que certos grupos de recursos sempre estarão on-line no mesmo nó ou que certos grupos de recursos sempre estarão on-line em diferentes nós. É possível combinar essas políticas de localização com dependências pai-filho ou iniciar após dependências e parar após dependências para que todos os grupos de recursos filho ou de origem on-line no mesmo nó enquanto o pai ou destino está on-line em um nó diferente. Também é possível ter todos os grupos de recursos filho ou de origem on-line em diferentes nós para melhor desempenho.

Nota: Os sites são suportados apenas no PowerHA SystemMirror 7.1.2 ou posterior, em ambas as Enterprise Edition e Standard Edition. O gerenciamento de replicação é suportado apenas no PowerHA SystemMirror Enterprise Edition.

Se você tiver recursos replicados, é possível combinar grupos de recursos em uma dependência de site para mantê-los on-line no mesmo site.

O PowerHA SystemMirror suporta os tipos a seguir de dependências de local do grupo de recursos entre grupos de recursos:

- **On-line no mesmo nó**

As regras e restrições a seguir se aplicam a On-line no mesmo conjunto de dependência de nó de grupos de recursos. A verificação falhará se você não seguir estas recomendações:

- Todos os grupos de recursos configurados como parte de um determinado conjunto de dependência de nó deve ter a mesma lista de nó (os mesmos nós na mesma ordem).
- Todos os grupos de recursos não simultâneos no mesmo conjunto de dependência de nó devem ter as mesmas políticas de inicialização, failover e fallback.
- A Política On-line usando a distribuição de nó não é permitida para inicialização.
- Se uma Política dinâmica de prioridade de nó é configurada como uma Política de failover, todos os grupos de recursos no conjunto devem ter a mesma política.
- Se um grupo de recursos tiver o cronômetro de fallback configurado, ela se aplicará ao conjunto de grupos de recursos. Todos os grupos de recursos no conjunto devem ter a mesma configuração de tempo de fallback.
- Ambos os grupos de recursos simultâneos e não simultâneos podem ser incluídos.
- É possível ter mais de um mesmo conjunto de dependência no cluster.

- O PowerHA SystemMirror reforça a condição de que todos os grupos de recursos no mesmo conjunto de dependência de nó que estão ativos (ONLINE) devem estar no mesmo nó. Alguns grupos de recursos no conjunto podem estar off-line ou no estado de erro.
- Se um ou mais grupos de recursos no mesmo conjunto de dependência de nó falharem, o PowerHA SystemMirror tentará colocar todos os grupos de recursos no conjunto no nó que pode hospedar todos os grupos de recursos que estejam atualmente ONLINE (aqueles que ainda estão ativos) mais um ou mais grupos de recursos com falha.

- **On-line no mesmo site**

As regras e restrições a seguir são aplicáveis a On-line no mesmo conjunto de dependência de site do grupo de recursos. A verificação falhará se você não seguir estas recomendações:

- Todos os grupos de recursos em um mesmo conjunto de dependência de site deve ter a mesma política de Gerenciamento entre sites, mas pode ter diferentes políticas de inicialização, políticas de fallover e políticas de fallback. Se os cronômetros de fallback forem usados, deverão ser idênticos para todos os grupos de recursos no conjunto.
- O cronômetro de fallback não se aplica à movimentação de um grupo de recursos entre os limites de site.
- Todos os grupos de recursos no mesmo conjunto de dependência de site devem ser configurados, para que os nós que possam possuir os próprios grupos de recursos designados aos mesmos sites primário e secundário.
- Ambos os grupos de recursos simultâneos e não simultâneos podem ser incluídos.
- É possível ter mais de um conjunto de dependência de site no cluster.
- Todos os grupos de recursos no mesmo conjunto de dependência que estão ativos (ONLINE) precisam estar ONLINE no mesmo site, mesmo que alguns grupos de recursos no conjunto possam estar no estado OFFLINE ou ERROR.
- Se você incluir um grupo de recursos que esteja incluído em um mesmo conjunto de dependência de nó em um mesmo conjunto de dependência de site, todos os outros grupos de recursos no mesmo conjunto de dependência de nó deverão ser incluídos no mesmo conjunto de dependência de site.

- **On-line em nós diferentes**

As regras e restrições a seguir se aplicam a on-line em diferentes conjuntos de dependência de nós dos grupos de recursos. A verificação falhará se você não seguir estas recomendações:

- Apenas um on-line em diferentes conjuntos de dependência de nós é permitido por cluster.
- Planeje políticas de inicialização para que cada grupo de recursos no conjunto seja inicializado em um nó diferente.
- Se uma dependência pai ou uma dependência filha for especificada, o grupo de recursos filho não poderá ter uma maior prioridade que seu grupo de recursos pai.
Após o cluster estar em execução com esses grupos configurados, esteja ciente que:
- Se um grupo de recursos com alta prioridade estiver on-line em um nó, nenhum outro grupo de recursos de prioridade inferior em diferentes conjuntos de dependência de nós poderá ficar on-line nesse nó.
- Se um grupo de recursos com uma prioridade mais alta executar fallover ou fallback para um determinado nó, o grupo de recursos com a prioridade mais alta ficará on-line e o gerenciador do cluster colocará o grupo de recursos de prioridade inferior off-line e o moverá para outro nó se for possível.
- Os grupos de recursos com a mesma prioridade não podem ficar on-line (inicialização) no mesmo nó. A prioridade de um grupo de recursos para um nó no mesmo nível de prioridade é determinada pela ordem alfabética do grupo no conjunto.
- Os grupos de recursos com a mesma prioridade não fazem com que outro seja movido do nó após um fallover ou fallback.

Referências relacionadas:

“Considerações especiais para o uso de sites” na página 74

As informações a seguir são sobre considerações especiais se um grupo de recursos tiver dependências especificadas.

Iniciar após dependências:

Ao iniciar após dependências, o grupo de recurso de destino deve estar on-line em qualquer nó no cluster antes de um grupo de recursos de origem (dependente) poder ser ativado em um nó. Não há dependência ao liberar grupos de recursos e os grupos liberados em paralelo.

A seguir, orientações e limitações para iniciar após dependências.

- Um grupo de recursos pode servir como um grupo de recursos de destino e de origem, dependendo de qual extremidade de um link de dependência for colocada.
- É possível especificar três níveis de dependências para grupos de recursos.
- Não é possível especificar dependências circulares entre grupos de recursos.
- Esta dependência se aplica apenas ao horário da aquisição do grupo de recursos. Não existe dependência entre esses grupos de recursos durante a liberação do grupo de recursos.
- Um grupo de recursos de origem não pode ser adquirido em um nó até que seu grupo de recursos de destino esteja completamente funcional. Se o grupo de recursos de destino não se tornar completamente funcional, o grupo de recursos de origem entrará em estado off-line devido a um estado off-line de destino. Se você observar que um grupo de recursos está nesse estado, pode ser necessário solucionar problemas dos recursos que podem precisar estar no estado on-line manualmente para resolver a dependência do grupo de recursos.
- Quando um grupo de recursos em uma função de destino cair de um nó para outro, não haverá efeito nos grupos de recursos que dependem dele.
- Após o grupo de recursos de origem ficar on-line, nenhuma operação (colocar off-line, mover grupo de recursos) no grupo de recursos de destino entrará em vigor do grupo de recursos de origem.
- Se o grupo de recursos de destino estiver off-line, não será possível mover manualmente um grupo de recursos ou colocar um grupo de recursos on-line no grupo de recursos de origem.

Nota: Você deve configurar vários monitores de aplicativos, especialmente um monitor que verifica a inicialização do aplicativo para o aplicativo que está incluído nos grupos de recursos de destino. Este processo verifica que o aplicativo no grupo de recursos de destino é iniciado com sucesso.

Parar após dependências:

Ao parar após as dependências, o grupo de recursos de destino deve estar off-line em qualquer nó no cluster antes de um grupo de recursos de origem (dependente) poder ser colocado off-line em um nó. Não há dependência ao adquirir grupos de recursos e os grupos serem adquiridos em paralelo.

A seguir, estão as limitações e orientações para parar após as dependências.

- Um grupo de recursos pode servir como um grupo de recursos de destino e de origem, dependendo de qual extremidade de um link de dependência for colocada.
- É possível especificar três níveis de dependências para grupos de recursos.
- Não é possível especificar dependências circulares entre grupos de recursos.
- Esta dependência se aplica apenas no momento em que um grupo de recursos é liberado. Não há dependência entre esses grupos de recursos durante a aquisição do grupo de recursos.
- Um grupo de recursos de origem não pode ser liberado em um nó até que seu grupo de recursos de destino esteja off-line.
- Quando um grupo de recursos em uma função de origem cai de um nó para outro, primeiro o grupo de recursos de destino é liberado e depois o grupo de recursos de origem é liberado. Após isso, ambos os grupos de recursos são adquiridos em paralelo, presumindo que não haja início após ou dependência pai-filho entre esses grupos de recursos.

- Se o grupo de recursos de destino estiver off-line, não será possível mover manualmente um grupo de recursos ou colocar um grupo de recursos off-line no grupo de recursos de origem.

Movendo grupos de recursos para outro nó

Quando o PowerHA SystemMirror move o grupo de recursos para outro nó, você tem algumas opções.

Essas opções incluem:

- Os grupos de recursos permanecem nos nós para os quais são movidos.
É possível mover grupos de recursos que têm a política Nunca realizar fallback para outro nó. Quando o fizer, é possível informar o PowerHA SystemMirror para deixar o grupo de recursos no nó de destino até que decida mover o grupo novamente.
- Ao mover um grupo de recursos com o comando **RG_move**, ele permanece no nó ao qual foi movido indefinidamente (até informar que o PowerHA SystemMirror mova-o para outro nó) ou até que reinicialize o cluster.

Se você parar os serviços de cluster, que raramente tem de ser feito e desejar alterar permanentemente a lista de nós do grupo de recursos e o nó de maior prioridade, altere os atributos do grupo de recursos e reinicie o cluster.

- Se você colocar o grupo de recursos on-line ou off-line em qualquer um dos nós, ele permanecerá on-line ou off-line até a próxima reinicialização do cluster ou até que coloque manualmente o grupo on-line em outro lugar no cluster.
- Se seu grupo de recursos tiver o fallback para a política do nó de prioridade mais alta, o grupo realizará fallback para seu nó de destino, após movê-lo.

Por exemplo, se o grupo tiver o nó A configurado como seu nó de prioridade mais alta e movê-lo para o nó B, esse grupo permanecerá no nó B e tratará esse nó agora como seu nó de prioridade mais alta. Sempre é possível escolher mover o grupo novamente para o nó A. Ao usar o SMIT para isso, o PowerHA SystemMirror o informa se o nó de prioridade mais alta original (nó A) agora está disponível para hospedar o grupo.

É possível manter controle de todos os grupos de recursos que foram movidos manualmente usando o comando **clRGinfo -p**.

Movendo grupos de recursos usando o comando **clRGmove**

É possível usar o comando **clRGmove** para mover um grupo de recursos para outro nó ou colocar um grupo de recursos on-line ou off-line. É possível executar o comando **clRGmove** usando o SMIT ou a partir da linha de comandos.

Se você usar **clRGmove** com grupos de recursos que têm a política de fallback Fallback, o grupo de recursos permanecerá nesse nó até que mova-o para outro lugar.

Os parágrafos a seguir descrevem as regras que se aplicam ao usar o **clRGmove** para gerenciar grupos de recursos com diferentes políticas.

Movendo grupos de recursos dependentes pai-filho

- Se os grupos de recursos pai estiverem off-line devido a sua solicitação feita por meio do comando **clRGmove**, o PowerHA SystemMirror rejeitará tentativas manuais de colocar os grupos de recursos filho que dependem desses grupos de recursos on-line. A mensagem de erro lista os grupos de recursos pai que devem ser colocados on-line primeiro.
- Se você tiver um grupo de recursos pai e um grupo de recursos filho e desejar mover o grupo de recursos pai para outro nó ou colocá-lo off-line, o PowerHA SystemMirror o impedirá de fazê-lo antes de um grupo de recursos filho ser colocado off-line.

Movendo início após grupos de recursos dependentes

Neste tipo de dependência, o recurso de destino deve ficar on-line em qualquer nó no cluster antes de um grupo de recursos de origem (dependente) poder ser ativado em um nó. As regras a seguir se aplicam a grupos de recursos com um início após a dependência:

- Se os grupos de recursos de destino estiverem off-line devido a sua solicitação feita por meio do comando **cIRGmove**, o PowerHA SystemMirror rejeitará tentativas manuais de colocar os grupos de recursos de origem que dependem desses grupos de recursos on-line. A mensagem de erro lista os grupos de recurso de destino que devem ser colocados on-line primeiro.

Movendo parada após grupos de recursos dependentes

Neste tipo de dependência, o grupo de recursos de destino deve ficar off-line em qualquer nó no cluster antes de um grupo de recursos de origem (dependente) poder ser colocado off-line em um nó. As regras a seguir se aplicam aos grupos de recursos com uma parada após a dependência:

- Se você tiver um grupo de recursos de destino e um de origem on-line e desejar mover o grupo de recursos de origem para outro nó ou colocá-lo off-line, o PowerHA SystemMirror o impedirá de fazê-lo antes de um grupo de recursos de destino ser colocado off-line.

Movendo grupos de recursos dependentes de local

- Se você mover um grupo de recursos dependente de mesmo site para o outro site, todo o conjunto de grupos de recursos na dependência será movido para o outro site.
- Se você mover um grupo de recursos dependente de mesmo nó para outro nó, todo o conjunto de grupos de recursos na dependência será movido.
- Não é possível mover um grupo de recursos para nenhum nó que hospede um grupo de recursos on-line que seja parte de uma dependência de nó diferente. Você deve primeiro deixar o grupo de recursos off-line que está incluído na dependência de nó diferente no nó selecionado.

Planejando redes de cluster e grupos de recursos

A maioria dos grupos de recursos não simultâneos usam rótulos de IP de serviço para fornecer acesso ao aplicativo para clientes pela rede. O PowerHA SystemMirror usa o IP address takeover (IPAT) via aliasing de IP para manter esses endereços de serviços disponíveis no cluster.

O IPAT não se aplica a grupos de recursos simultâneos ou on-line em todos os grupos de recursos de nós disponíveis.

É importante considerar como os clientes acessam endereços de aplicativo pela rede do cluster, quando há firewalls ou VPNs configurados no ambiente.

Referências relacionadas:

“Planejamento para controle de endereço IP via aliases de IP” na página 25

A designação de aliases de IP a NICs permite criar mais de um rótulo de IP na mesma interface de rede.

Planejamento paralelo ou serial da ordem para processamento de grupos de recursos

Por padrão, o PowerHA SystemMirror adquire e libera todos os recursos individuais configurados em seu cluster em paralelo. Entretanto, é possível especificar uma ordem serial específica com a qual alguns ou todos os grupos de recursos individuais são adquiridos ou liberados.

Os processos a seguir são concluídos durante a aquisição:

1. O PowerHA SystemMirror adquire os grupos de recursos serialmente na ordem especificada na lista.
2. O PowerHA SystemMirror adquire os grupos de recursos restantes em paralelo.

Durante a liberação de grupos de recursos, o processo é reservado:

1. O PowerHA SystemMirror libera os grupos de recursos para os quais não definiu uma ordem serial específica em paralelo.
2. Os grupos de recursos restantes no cluster serão processados na ordem especificada para esses grupos de recursos na lista.
3. Se você atualizou seu cluster a partir de uma versão anterior do PowerHA SystemMirror, para obter mais informações sobre qual ordem de processamento é usada neste caso, consulte Fazendo upgrade de um cluster do PowerHA SystemMirror.

Nota: Mesmo se você especificar a ordem de processamento do grupo de recursos em um único nó, o failover real dos grupos de recursos poderá ser acionado por diferentes políticas. Portanto, não é garantido que os grupos de recursos sejam processados em todo o cluster na ordem especificada porque a ordem de processamento customizada serial dos grupos de recursos se aplica ao processamento em um determinado nó apenas.

4. Quando os grupos de recursos são processados em paralelo, menos eventos de cluster ocorrem no cluster. Em particular, eventos como `node_up_local` ou `get_disk_vg_fs` não ocorrem se os grupos de recursos forem processados em paralelo.
5. Como resultado, o uso do processamento paralelo reduz o número de eventos de cluster específicos para os quais é possível criar scripts de pré e pós-evento customizados. Se você iniciar o uso do processamento paralelo para alguns dos grupos de recursos em sua configuração, esteja ciente de que seus scripts de pré e pós-evento existentes podem não funcionar para esses grupos de recursos.
6. O processamento paralelo e serial dos grupos de recursos é refletido nos resumos de eventos no arquivo `hacmp.out`.

Para obter mais informações sobre como configurar a aquisição serial customizada e a ordem de liberação dos grupos de recursos, consulte Configurando a ordem de processamento para os grupos de recursos.

Grupos de recursos dependentes e ordem paralela ou serial

Embora, por padrão, o PowerHA SystemMirror processe grupos de recursos em paralelo, se você estabelecer dependências entre alguns dos grupos de recursos no cluster, o processamento poderá demorar mais que para os clusters sem grupos de recursos dependentes, pois há mais processamento a ser feito para manipular um ou mais eventos `rg_move`

Mediante a aquisição, primeiro os grupos de recursos pai ou de maior prioridade são adquiridos, em seguida, os grupos de recursos filho são adquiridos. Mediante a liberação, a ordem é revertida. Os grupos de recursos restantes no cluster (aqueles que não têm dependências) são processados em paralelo.

Além disso, se você especificar a ordem serial ou o processamento e tiver grupos de recursos dependentes configurados, assegure-se de que a ordem serial não contradiga a dependência especificada. A dependência dos grupos de recursos substitui qualquer ordem serial no cluster.

Referências relacionadas:

“Planejamento para eventos de cluster” na página 82

Estes tópicos descrevem os eventos do cluster do PowerHA SystemMirror.

Informações relacionadas:

Configurando a ordem de processamento para grupos de recursos

Fazendo upgrade de um cluster do PowerHA SystemMirror

Planejando grupos de recursos nos clusters que possuem sites

A combinação de política de gerenciamento entre sites e a inicialização do nó, políticas de failover e políticas de fallback que você selecionar determina o comportamento de inicialização, failover e fallback do grupo de recursos.

Nota: Os sites são suportados apenas no PowerHA SystemMirror 7.1.2 ou posterior, em ambas as Enterprise Edition e Standard Edition. O gerenciamento de replicação é suportado apenas no PowerHA SystemMirror Enterprise Edition.

O suporte do site no PowerHA SystemMirror permite uma variedade de configurações do grupo de recursos.

Grupos de recursos simultâneos e sites

É possível usar as políticas a seguir para grupos de recursos simultâneos:

Tipo de política	Descrição
Política de gerenciamento entre sites	On-line em ambos os sites Online em qualquer site Preferir site primário Ignorar
Política de inicialização	On-line em todos os nós disponíveis
Política de fallover	Colocar off-line (no nó Erro apenas)
Política de fallback	Nunca fazer fallback

Grupos de recursos não simultâneos e sites

Para grupos de recursos não simultâneos, é possível usar as políticas a seguir:

Tipo de política	Descrição
Política de gerenciamento entre sites	Online em qualquer site Preferir site primário Ignorar
Política de inicialização	On-line no nó inicial On-line no primeiro nó disponível On-line usando a política de distribuição de nós
Política de fallover	Fallover para o próximo nó de prioridade (na lista de nós)
Política de fallback	Fallback para o nó de maior prioridade (na lista de nós) Nunca fazer fallback

Comportamento de grupo de recurso geral em clusters com sites

Os grupos de recursos não simultâneos definidos com uma política de gerenciamento entre sites do site primário preferencial ou online em um dos sites tem duas instâncias quando o cluster está em execução.

Nota: Os sites são suportados apenas no PowerHA SystemMirror 7.1.2 ou posterior, em ambas as Enterprise Edition e Standard Edition. O gerenciamento de replicação é suportado apenas no PowerHA SystemMirror Enterprise Edition.

A instância a seguir está em execução em grupos de recursos não simultâneos:

- Uma instância primária em um nó no site primário
- Uma instância secundária em um nó no site secundário

O comando **clRGinfo** mostra essas instâncias como:

- ONLINE
- ONLINE SECONDARY

Os grupos de recursos simultâneos (on-line em todos os nós) com uma política de gerenciamento entre sites de on-line em ambos os sites têm diversas instâncias ONLINE e nenhuma instância ONLINE SECONDARY quando o cluster está em execução em ambos os locais.

Os grupos de recursos simultâneos com uma política de gerenciamento entre sites do site primário preferencial ou on-line em um dos sites têm instâncias primárias em cada nó no site primário e instâncias secundárias em nós no site secundário.

Os grupos de recursos com recursos replicados são processados em paralelo em um evento **node_up**, de acordo com seu gerenciamento de site e políticas de inicialização de nó, levando em conta quaisquer dependências. Quando os grupos de recursos caírem entre sites, as instâncias secundárias serão adquiridas e colocadas em estado ONLINE SECONDARY no nó de prioridade mais alta que estiver disponível no novo site de backup. Mais de uma instância secundária pode estar no mesmo nó. A instância primária do grupo de recursos é adquirida e colocada ONLINE no nó de prioridade mais alta que estiver disponível para hospedar esse grupo de recursos no novo site ativo. Esta ordem de processamento assegura que o site de backup esteja pronto para receber os dados de backup quando a instância primária for iniciada.

Se a instância secundária não puder ser colocada no estado ONLINE SECONDARY, a instância primária ainda será colocada ONLINE, se possível.

Considerações especiais para o uso de sites

As informações a seguir são sobre considerações especiais se um grupo de recursos tiver dependências especificadas.

Os grupos de recursos e sites dependentes

Nota: Os sites são suportados apenas no PowerHA SystemMirror 7.1.2 ou posterior, em ambas as Enterprise Edition e Standard Edition. O gerenciamento de replicação é suportado apenas no PowerHA SystemMirror Enterprise Edition.

É possível especificar uma dependência entre dois ou mais grupos de recursos que residem em nós, que estão localizados em sites diferentes. Neste caso, se um pai ou filho se mover para o outro site, o grupo dependente também será movido. Se o grupo-pai não puder ser ativado no site de fallover, o grupo de recursos filho também permanecerá inativo.

A dependência se aplica apenas ao estado da instância primária do grupo de recursos. Se a instância primária do grupo-pai estiver OFFLINE e a instância secundária for ONLINE SECONDARY em um nó, a instância primária do grupo filho estará OFFLINE.

Durante a recuperação do grupo de recursos, os grupos de recursos podem realizar fallover para os nós em um dos sites. A sequência para adquirir grupos de recursos dependentes é a mesma que dos clusters sem sites, na qual o grupo de recursos pai é adquirido primeiro e o grupo de recursos filho é adquirido. A lógica de liberação é revertida quando o grupo de recursos filho é liberado antes de um grupo de recursos pai ser liberado.

Se você tiver sites definidos em clusters sem sites, você deve configurar o monitoramento de aplicativo para aplicativos inclusos em grupos de recursos com dependências.

Referências relacionadas:

“Dependências de local do grupo de recursos” na página 67

Certos aplicativos em diferentes grupos de recursos ficam on-line juntos em um nó ou em diferentes nós.

Exemplos: comportamento do grupo de recursos nos clusters que possuem sites

A política de fallback se aplica a instâncias ONLINE e ONLINE SECONDARY do grupo de recursos.

Nota: Os sites são suportados apenas no PowerHA SystemMirror 7.1.2 ou posterior, em ambas as Enterprise Edition e Standard Edition. O gerenciamento de replicação é suportado apenas no PowerHA SystemMirror Enterprise Edition.

A política de gerenciamento entre sites para um grupo de recursos determina o comportamento de fallback das instâncias ONLINE dos grupos de recursos entre sites, o que controla o local da instância secundária.

A instância ONLINE SECONDARY está localizada no site que não tem a instância ONLINE. A tabela a seguir mostra o comportamento esperado dos grupos de recursos durante eventos de site de acordo com as políticas de inicialização e gerenciamento entre sites.

Nota: A maioria das restrições na tabela a seguir se aplicam apenas quando diversas interfaces por nó estão na mesma rede. Quando apenas uma interface por nó está na rede, o que pode ser comum ao usar o Fibre Channel ou Ethernet virtual, a maioria dessas restrições não se aplica.

Política de inicialização de nó (aplica-se no site)	Política de gerenciamento entre sites	Comportamento de inicialização, fallover ou fallback
On-line no nó inicial apenas	Preferir site primário	<p>Inicialização do Cluster</p> <p>Site primário O nó inicial adquire o grupo de recursos no estado ONLINE. Nós não iniciais deixam o grupo de recursos.</p> <p>Site secundário O primeiro nó que se une a este site adquire o grupo de recursos no estado ONLINE SECONDARY.</p> <p>Fallover entre sites A instância ONLINE fica entre sites quando nenhum nó no site local pode adquirir o grupo de recursos. A instância secundária se move para outro site e é levada para o estado ONLINE SECONDARY no nó com mais alta prioridade que está disponível, se possível.</p> <p>Fallback entre sites A instância ONLINE realiza fallback para o site primário quando um nó a partir do site primário se une. A instância secundária se move para outro site e é levada para o estado ONLINE SECONDARY no nó com mais alta prioridade que está disponível, se possível.</p>

Política de inicialização de nó (aplica-se no site)	Política de gerenciamento entre sites	Comportamento de inicialização, fallover ou fallback
<p>On-line no primeiro nó disponível</p> <p>ou</p> <p>On-line usando a política de distribuição de nós</p>	<p>Preferir site primário</p>	<p>Inicialização do Cluster</p> <p>Site primário O nó que se une primeiro a partir do site primário, atende aos critérios e adquire o grupo de recursos no estado ONLINE. O grupo de recursos é OFFLINE em todos os outros nós no site primário. A política de distribuição de nó se aplica apenas à instância primária do grupo de recursos.</p> <p>Site secundário O primeiro nó a se unir ao cluster neste site adquire todas as instâncias secundárias dos grupos de recursos com esta política de inicialização no estado ONLINE_SECONDARY (sem distribuição).</p> <p>Fallover entre sites A instância ONLINE fica entre sites quando nenhum nó no site local pode adquirir o grupo de recursos. A instância secundária se move para outro site e é levada para o estado ONLINE SECONDARY no nó com mais alta prioridade que está disponível, se possível.</p> <p>Fallback entre sites A instância ONLINE realiza fallback para o site primário quando um nó a partir do site primário se une. A instância secundária se move para outro site e é levada para o estado ONLINE SECONDARY no nó com mais alta prioridade que está disponível, se possível.</p>
<p>On-line em todos os nós disponíveis</p>	<p>Preferir site primário</p>	<p>Inicialização do Cluster</p> <p>Site primário Todos os nós adquirem o grupo de recursos no estado ONLINE.</p> <p>Site secundário Todos os nós adquirem o grupo de recursos no estado ONLINE_SECONDARY.</p> <p>Fallover entre sites As instâncias ONLINE caem entre sites quando todos os nós no site local ficam OFFLINE ou falham ao iniciar o grupo de recursos. As instâncias secundárias se movem para o outro site e são colocadas no estado ONLINE SECONDARY quando possível.</p> <p>Fallback entre sites As instâncias ONLINE realizam fallback para o site primário quando um nó no site primário se une novamente. Os nós no site secundário adquirem o grupo de recursos no estado ONLINE_SECONDARY.</p>

Política de inicialização de nó (aplica-se no site)	Política de gerenciamento entre sites	Comportamento de inicialização, fallover ou fallback
On-line no nó inicial apenas	Online em qualquer site	<p>Inicialização do Cluster</p> <p>Site primário O nó inicial que se une ao cluster (de qualquer site) adquire o grupo de recursos no estado ONLINE. Os nós não iniciais deixam o grupo de recursos OFFLINE.</p> <p>Site secundário O primeiro nó a se unir a partir do outro site adquire o grupo de recursos no estado ONLINE_SECONDARY.</p> <p>Fallover entre sites A instância ONLINE fica entre sites quando nenhum nó no site local pode adquirir o grupo de recursos. A instância secundária se move para o outro site e é colocada no estado ONLINE_SECONDARY no nó de maior prioridade disponível, se possível.</p> <p>Fallback entre sites A instância ONLINE não realiza fallback para o site primário quando um nó no site primário se une novamente. A prioridade mais alta que se une novamente ao nó adquire o grupo de recursos no estado ONLINE_SECONDARY.</p>
On-line no primeiro nó disponível ou On-line usando a política de distribuição de nós	Online em qualquer site	<p>Inicialização do Cluster</p> <p>Site primário O nó que se une primeiro de qualquer site, que atenda aos critérios de distribuição, adquire o grupo de recursos no estado ONLINE.</p> <p>Site secundário Após o grupo de recursos ficar ONLINE, o primeiro nó que se une ao nó a partir de outro site adquire o grupo de recursos no estado ONLINE_SECONDARY.</p> <p>Fallover entre sites A instância ONLINE fica entre sites quando nenhum nó no site local pode adquirir o grupo de recursos.</p> <p>Fallback entre sites A instância ONLINE não realiza fallback para o site primário quando o site primário se une. Um nó que se une novamente adquire o grupo de recursos no estado ONLINE_SECONDARY.</p>
On-line em todos os nós disponíveis	Online em qualquer site	<p>Inicialização do Cluster</p> <p>Site primário O nó que se une primeiro de qualquer site adquire o grupo de recursos no estado ONLINE. Após uma instância do grupo estar ativa, os nós restantes no mesmo site também ativam o grupo no estado ONLINE.</p> <p>Site secundário Todos os nós adquirem o grupo de recursos no estado ONLINE_SECONDARY.</p> <p>Fallover entre sites A instância ONLINE cai entre sites quando todos os nó no site local entram OFFLINE ou falham ao iniciar o grupo de recursos.</p> <p>Fallback entre sites A instância ONLINE não realiza fallback para o site primário quando o site primário se une. Os nós que se unem novamente adquirem o grupo de recursos no estado ONLINE_SECONDARY.</p>

Política de inicialização de nó (aplica-se no site)	Política de gerenciamento entre sites	Comportamento de inicialização, fallover ou fallback
On-line em todos os nós disponíveis	On-line em ambos os sites	<p>Inicialização do Cluster Todos os nós em ambos os sites ativam o grupo de recursos no estado ONLINE.</p> <p>Fallover entre sites Não ocorre fallover. O grupo de recursos está no estado OFFLINE ou ERROR.</p> <p>Fallback entre sites Não ocorre fallback.</p>

Customizando a recuperação do grupo de recursos entre sites

A recuperação do grupo de recursos entre sites pode ser configurada para a notificação automática do processamento ou falha apenas.

Uma determinada instância de um grupo de recursos pode cair em um site, mas não pode se mover entre sites. Se nenhum nó estiver disponível no site em que a instância afetada reside, essa instância entrará no estado ERROR ou ERROR_SECONDARY. Ele não fica no nó em que falhou. Este comportamento se aplica tanto às instâncias primárias quanto secundárias.

O Gerenciador do cluster move o grupo de recursos se um evento **node_down** ou **node_up** ocorrer, mesmo se o fallover entre os sites estiver desativado. Também é possível mover manualmente um grupo de recursos entre sites.

Ativando ou desativando o fallover entre sites

Se você migrou de uma liberação anterior do PowerHA SystemMirror, poderá alterar a política de recuperação de grupo de recursos para permitir que o Gerenciador do cluster mova um grupo de recursos para outro site para evitar que o grupo de recursos entre no estado ERROR.

Recuperação de instâncias primárias de grupos de recursos replicados entre sites

Quando o fallover entre os sites é ativado, o PowerHA SystemMirror tenta recuperar a instância primária de um grupo de recursos em situações em que uma interface que está conectada a uma rede entre sites falha ou se torna disponível.

Recuperação de instâncias secundárias de grupos de recursos replicados entre sites

Quando o fallover entre sites é ativado, o PowerHA SystemMirror tenta recuperar a instância secundária e a instância primária de um grupo de recursos nestas situações:

- Se uma falha de aquisição ocorrer enquanto a instância secundária de um grupo de recursos estiver sendo adquirida, o Gerenciador do cluster tentará recuperar a instância secundária do grupo de recursos, como o faz para a instância primária. Se nenhum nó estiver disponível para a aquisição, a instância secundária do grupo de recursos entrará no estado global ERROR_SECONDARY.
- Se a perda de quorum for acionada e o grupo de recursos tiver sua instância secundária no estado ONLINE no nó afetado, o PowerHA SystemMirror tentará recuperar a instância secundária em outro nó disponível.
- Se todas as redes XD_data falharem, o PowerHA SystemMirror moverá todos os grupos de recursos ONLINE que tiverem recursos do GLVM para outro nó disponível nesse site. Esta função da instância primária é espelhada para a instância secundária para que as instâncias secundárias possam ser recuperadas por meio do fallover seletivo.

Usando o SMIT para ativar ou desativar a recuperação do grupo de recursos entre sites

Para ativar ou desativar a recuperação do grupo de recursos entre sites, conclua as etapas a seguir:

1. Na linha de comandos, insira `smit sysmirror`.
2. Na interface do SMIT, selecione **Configurações de cluster customizadas > Recursos > PowerHA SystemMirrorConfiguração de recursos estendidos > Customizar recuperação do grupo de recursos entre sites** e pressione Enter.

Planejamento para recursos replicados

O PowerHA SystemMirror oferece o suporte integrado para recursos replicados.

Com recursos replicados do PowerHA SystemMirror, é possível usar as funções a seguir:

- Permite reconfigurar dinamicamente os grupos de recursos que contêm recursos replicados com as configurações de site do PowerHA SystemMirror Enterprise Edition for AIX e PowerHA SystemMirror.
- Consolida a verificação do PowerHA SystemMirror Enterprise Edition for AIX em uma verificação de cluster padrão detectando automaticamente e chamando os utilitários de verificação do produto PowerHA SystemMirror Enterprise Edition for AIX instalado.

Nota: Os sites são suportados apenas no PowerHA SystemMirror 7.1.2 ou posterior, em ambas as Enterprise Edition e Standard Edition. O gerenciamento de replicação é suportado apenas no PowerHA SystemMirror Enterprise Edition.

Configurando recursos replicados

Se você instalou um produto do PowerHA SystemMirror Enterprise Edition for AIX, poderá usar o suporte integrado para configurar as diferentes tecnologias de replicação.

Os recursos replicados a seguir são configurações do PowerHA SystemMirror Enterprise Edition for AIX suportadas:

- Os grupos de recursos com a política de nó simultânea pode ter a política de gerenciamento do site não simultânea.
- A recuperação entre sites dos grupos de recursos que contêm os recursos replicados do PowerHA SystemMirror Enterprise Edition for AIX é permitido, por padrão, para novas instalações do PowerHA SystemMirror. As configurações que são atualizadas e migradas de liberações anteriores mantêm o comportamento pré-existente. É possível configurar esse comportamento para que seja uma opção **fallover** ou **notificar** no movimento de grupo de recursos iniciado pelo cluster. Se você selecionar a opção **notificar**, será necessário configurar um script pré-evento ou pós-evento ou um método de notificação remoto.
- Configurações de dependência pai, filho e local para grupos de recursos replicados.
- Política de inicialização de distribuição de grupo de recursos baseada em nó com os sites do PowerHA SystemMirror.

Nota: Os sites são suportados apenas no PowerHA SystemMirror 7.1.2 ou posterior, em ambas as Enterprise Edition e Standard Edition. O gerenciamento de replicação é suportado apenas no PowerHA SystemMirror Enterprise Edition.

Não é possível configurar um grupo de recursos para usar uma política de nó não simultânea e uma política de gerenciamento entre sites simultânea.

Processando recursos replicados

O processamento de eventos do PowerHA SystemMirror automaticamente suporta os recursos replicados configurados.

Os recursos replicados incluem as funções a seguir:

- Sempre que possível, o PowerHA SystemMirror processa eventos em paralelo. Os eventos são dinamicamente divididos em fases para que o PowerHA SystemMirror processe as instâncias primárias e secundárias de um grupo de recursos na ordem adequada (release_primary, release_secondary, acquire_secondary, acquire_primary).
- O PowerHA SystemMirror recupera as instâncias secundárias, bem como as instâncias primárias dos grupos de recursos replicados durante as perdas do grupo de volumes, falhas de aquisição e eventos **local_network_down** se outro nó ou rede estiver disponível.
- O PowerHA SystemMirror tem mais probabilidade de adquirir a instância secundária de um grupo de recursos mediante o failover do site que qualquer outra instância. PowerHA SystemMirror pode considerar todos os nós no site secundário como destinos, não apenas o nó que hospedou anteriormente a instância primária, que era o caso na versão anterior do PowerHA SystemMirror.

Nota: Os sites são suportados apenas no PowerHA SystemMirror 7.1.2 ou posterior, em ambas as Enterprise Edition e Standard Edition. O gerenciamento de replicação é suportado apenas no PowerHA SystemMirror Enterprise Edition.

Movendo grupos de recursos que possuem recursos replicados

É possível mover a instância primária de um grupo de recursos que possui recursos replicados para outro site e o PowerHA SystemMirror automaticamente processará as instâncias secundárias do grupo na mesma operação.

Nota: Os sites são suportados apenas no PowerHA SystemMirror 7.1.2 ou posterior, em ambas as Enterprise Edition e Standard Edition. O gerenciamento de replicação é suportado apenas no PowerHA SystemMirror Enterprise Edition.

O Gerenciador de Cluster usa a fase de evento dinâmico que move para a instância secundária desse site para outro se um nó estiver disponível para hospedá-lo. Cada tentativa é feita para manter a instância secundária no estado SECONDARY_ONLINE. Mesmo se um nó em um determinado estado estiver configurado para que não possa hospedar mais de uma instância primária, poderá hospedar mais de uma instância secundária para manter todos no estado SECONDARY_ONLINE.

Planejamento para o Workload Manager

A IBM oferece o AIX Workload Manager (WLM) como um recurso de administração do sistema incluído com o sistema operacional AIX.

O WLM permite configurar destinos e limites na CPU, uso de memória física e largura da banda de E/S do disco para diferentes processos e aplicativos. Isso fornece um melhor controle sobre o uso de recursos críticos do sistema em cargas de pico. O PowerHA SystemMirror permite configurar classes do WLM para grupos de recursos do PowerHA SystemMirror para que o início e a parada do WLM e a configuração do WLM ativa possam ser sob o controle de cluster.

O PowerHA SystemMirror não verifica cada aspecto da configuração do seu WLM; portanto, é sua responsabilidade assegurar a integridade dos arquivos de configuração do WLM. Após incluir as classes do WLM em um grupo de recursos do PowerHA SystemMirror, o utilitário de verificação verifica apenas se as classes do WLM necessárias existem. Portanto, você deve entender como o WLM funciona e configurá-lo cuidadosamente. Os parâmetros de configuração incorretos, porém aceitáveis, podem impedir a produtividade e a disponibilidade de seu sistema.

Para obter informações completas sobre como configurar e usar o Workload Manager, consulte a publicação IBM *AIX Workload Manager (WLM) Redbooks*.

O Workload Manager distribui recursos do sistema entre os processos que solicitam-nos de acordo com a classe em que estão. Os processos são designados a classes específicas de acordo com as regras de designação de classe. O planejamento para a integração do WLM com o PowerHA SystemMirror inclui duas etapas básicas:

1. O uso de painéis SMIT do AIX para definir as classes do WLM e regras de designação de classe relacionadas a aplicativos altamente disponíveis.
2. O uso de painéis SMIT do PowerHA SystemMirror para estabelecer a associação entre a configuração do WLM e os grupos de recursos do PowerHA SystemMirror.

Informações relacionadas:

 [Redbooks Workload Manager \(WLM\) do AIX](#)

Classes do Workload Manager

O Workload Manager distribui recursos de sistema entre os processos que os solicita de acordo com a classe à qual os processos são designados.

As propriedades de uma classe incluem:

- Nome da classe. Uma sequência alfanumérica exclusiva de não mais que 16 caracteres de comprimento.
- Camada da classe. Um número de 0 a 9. Esse número determina a importância relativa de uma classe do mais importante (camada 0) para o menos importante (camada 9).
- Número da CPU e compartilhamentos de memória física. A quantia real de recursos atribuídos a cada classe depende do número total de compartilhamentos em todas as classes (logo, se duas classes forem definidas em um sistema, uma com dois compartilhamentos de uso de CPU de destino e a outra com três compartilhamentos, a primeira classe receberá 2/5 e a segunda classe receberá 3/5 do tempo de CPU).
- Limites de configuração. Limites de porcentagem máxima e mínima de tempo de CPU, memória física e largura da banda de E/S de disco acessível ao processo.

Você configura as regras de designação de classe que informam ao WLM como classificar todos os novos processos (bem como aqueles já em execução no momento da inicialização do WLM) de acordo com seu ID de grupo (GID), ID do usuário (UDI) e o nome do caminho integral.

Reconfiguração, inicialização e encerramento do Workload Manager

Esta seção descreve a forma que o WLM é reconfigurado, iniciado ou interrompido após colocar o WLM sob o controle do PowerHA SystemMirror.

Reconfiguração do Workload Manager

Após as classes do WLM serem incluídas em um grupo de recursos do PowerHA SystemMirror, no momento da sincronização do cluster no nó, o PowerHA SystemMirror reconfigurará o WLM para usar as regras necessárias pelas classes associadas ao nó. No caso da reconfiguração de recurso dinâmica no nó, o WLM é reconfigurado de acordo com quaisquer mudanças feitas nas classes do WLM associadas a um grupo de recursos.

Inicialização do Workload Manager

A inicialização do WLM ocorre quando o nó se associa ao cluster ou quando uma reconfiguração dinâmica da configuração do WLM ocorre.

A configuração é específica ao nó e depende dos grupos de recursos dos quais o nó participa. se o nó não puder adquirir nenhum grupo de recursos associado às classes do WLM, o WLM não será iniciado.

Para todos os grupos de recursos não simultâneos que não têm a política de inicialização On-line usando distribuição de nó, o script de inicialização determina se o grupo de recursos está em execução em um nó primário ou secundário e inclui as regras de designação de classe do WLM correspondentes na configuração do WLM. Para todos os outros grupos de recursos não simultâneos e para os grupos de recursos de acesso simultâneo que o nó adquire, a classe do WLM primária associada a cada grupo de recursos é colocada na configuração do WLM. As regras correspondentes são incluídas na tabelas de regras.

Por fim, se o WLM estiver atualmente em execução e não tiver sido iniciado pelo PowerHA SystemMirror, o script de inicialização reiniciará o WLM da configuração especificada pelo usuário, salvando a configuração anterior. Quando o PowerHA SystemMirror é interrompido, ele retorna o WLM de volta para sua configuração anterior.

A falha ao inicializar o WLM gera uma mensagem de erro registrada no arquivo de log **hacmp.out**, mas a inicialização do nó e a reconfiguração do recurso continua.

Encerramento do Workload Manager

O encerramento do WLM ocorre quando o nó deixa o cluster ou na reconfiguração do cluster dinâmico. Se o WLM estiver atualmente em execução, o script de encerramento determina se o WLM estava em execução antes de ser iniciado pelo PowerHA SystemMirror e que configuração estava usando. Ele não faz nada (se o WLM não estiver em execução atualmente) ou interrompe o WLM (se não estava em execução antes da inicialização do PowerHA SystemMirror) ou o interrompe e o reinicializa na configuração anterior (se o WLM estava em execução anteriormente).

Limitações e considerações

Lembre-se de algumas limitações e considerações ao planejar sua configuração do Workload Manager

Essas limitações e considerações incluem:

- Algumas configurações do WLM que impedem o desempenho do PowerHA SystemMirror. Seja cuidadoso ao projetar suas classes e regras e esteja alerta quanto ao possível impacto no PowerHA SystemMirror.
- É possível ter não mais que 27 classes do WLM não padrão entre o cluster, porque uma configuração é compartilhada entre os nós do cluster.
- Uma configuração do PowerHA SystemMirror Workload Manager não suporta subclasses, mesmo que o WLM permita. Se configurar subclasses para quaisquer classes do WLM que estiverem colocadas em um grupo de recursos, um aviso será emitido mediante a verificação do cluster e as subclasses não serão propagadas para outros nós durante a sincronização.
- Em um determinado nó, apenas as regras para classes associadas aos grupos de recursos que podem ser adquiridas por um nó estão ativas nesse nó.

Planejamento para eventos de cluster

Estes tópicos descrevem os eventos do cluster do PowerHA SystemMirror.

Visão Geral

O PowerHA SystemMirror fornece as formas a seguir de gerenciar o processamento de eventos:

- Customizar eventos predefinidos
- Definir novos eventos

No PowerHA SystemMirror, os grupos de recursos são processados em paralelo por padrão, se possível, a menos que especifique uma ordem de processamento serial customizada para todos ou alguns dos grupos de recursos no cluster.

A lógica e a sequência de eventos conforme descrito em exemplos pode não listar todos os eventos.

Consulte Iniciando e parando serviços de cluster para obter informações sobre:

- Etapas a executar para iniciar e parar serviços do cluster
- Interação com o comando do AIX **shutdown** e interação de serviços de cluster do PowerHA SystemMirror com Reliable Scalable Cluster Technology (RSCT)

Referências relacionadas:

“Planejamento paralelo ou serial da ordem para processamento de grupos de recursos” na página 71
Por padrão, o PowerHA SystemMirror adquire e libera todos os recursos individuais configurados em seu cluster em paralelo. Entretanto, é possível especificar uma ordem serial específica com a qual alguns ou todos os grupos de recursos individuais são adquiridos ou liberados.

Planejando eventos de site e nó

É possível definir sites no PowerHA SystemMirror Standard Edition for AIX ou PowerHA SystemMirror Enterprise Edition. Você deve definir sites para permitir o suporte de replicação de armazenamento do PowerHA SystemMirror Enterprise Edition, incluindo o Gerenciador de Volume Lógico (GLVM) e Metro Mirror. Ao associar nós e dispositivos de armazenamento a um site, é possível usar o PowerHA SystemMirror para implementar uma configuração de espelhamento do Gerenciador de Volume Lógico (LVM) de site de divisão. O PowerHA SystemMirror identifica seleções apropriadas que são baseadas em informações do site e verifica a consistência da configuração de espelhamento em um nível do site.

Nota: Os sites são suportados apenas no PowerHA SystemMirror 7.1.2 ou posterior, em ambas as Enterprise Edition e Standard Edition. O gerenciamento de replicação é suportado apenas no PowerHA SystemMirror Enterprise Edition.

Os scripts de evento de site são incluídos no software PowerHA SystemMirror. Se os sites não forem definidos, nenhum evento de site será gerado. Os scripts **site_event** do PowerHA SystemMirror são executados como a seguir se os sites forem definidos:

- O primeiro nó em um site executa o evento **site_up** antes de concluir o processamento de eventos **node_up**. O evento **site_up_complete** é executado após o evento **node_up_complete**.
- Quando o último nó em um site ficar inativo, o evento **site_down** será executado antes do evento **node_down** e o evento **site_down_complete** é executado após o evento **node_down_complete**.

Sem instalar o PowerHA SystemMirror Enterprise Edition, é possível definir pré-eventos e pós-eventos para execução quando um site altera o estado. Neste caso, é possível definir todos os processos relacionados ao site.

Os eventos de site (incluindo o evento **check_for_site_up** e o evento **check_for_site_down**) são registrados no arquivo de log **hacmp.out**.

Se os sites forem definidos, o evento **site_up** será executado quando o primeiro nó no site ficar ativo e o evento **site_down** será executado quando o último nó no site ficar inativo. A sequência de script do evento para manipulação de grupos de recursos em geral é:

site_up

site_up_remote

node_up

Eventos rg_move para processar ações de grupo de recursos

node_up_complete

site_up_complete

site_up_remote_complete

site_down

site_down_remote

node_down

Eventos rg_move para processar ações de grupo de recursos

node_down_complete

site_down

site_down_remote_complete

Planejamento de eventos `node_up` e `node_down`

Um evento `node_up` é iniciado por um nó que se une ao cluster na inicialização do cluster ou se unindo novamente ao cluster posteriormente.

Estabelecendo associação de cluster inicial

Este tópico descreve as etapas realizadas pelo gerenciador do cluster em cada nó quando o cluster inicia e a associação inicial do cluster é estabelecida. Ele mostra como os gerenciadores de cluster estabelecem a comunicação entre os nós do membro e como os recursos do cluster são distribuídos conforme a associação de cluster aumenta.

Primeiro nó se une ao cluster

1. Os serviços de cluster do PowerHA SystemMirror são iniciados no nó A. O subsistema Reliable Scalable Cluster Technology (RSCT) examina o estado das interfaces de rede e começa a se comunicar com o subsistema RSCT nos outros nós do cluster. O gerenciador do cluster no nó A acumula as informações de estado inicial e transmite uma mensagem indicando que está pronto para se unir ao cluster em todas as redes configuradas às quais está conectado (`node_up`).
2. O Nó A interpreta a falta de resposta para significar que é o primeiro nó no cluster.
3. O Nó A inicia um script `process_resources`, que processa as informações de configuração do recurso do nó.
4. Quando o processamento de eventos está concluído, o nó A se torna um membro do cluster. PowerHA SystemMirror executa `node_up_complete`.

Todos os grupos de recursos definidos para o nó A estão disponíveis para os clientes neste momento. Se a política de inicialização **On-line no primeiro nó disponível** for especificada como um comportamento de inicialização para grupos de recursos, o nó A assumirá o controle de todos esses grupos de recursos.

Se o nó A estiver definido como parte de um grupo de recursos não simultâneos que possui a política de inicialização **On-line usando distribuição de nó**, esse nó assumirá o controle do primeiro grupo de recursos listado no ambiente do nó.

Se o nó A estiver definido como parte de uma configuração de recurso de acesso simultâneo, ele tornará esses recursos simultâneos disponíveis.

Para os grupos de recursos que possuem a política de inicialização **On-line no primeiro nó disponível** e o tempo de acomodação configurados, o nó A aguarda o intervalo de tempo de acomodação antes de adquirir tais grupos de recursos. O tempo de acomodação facilita a espera por um nó de maior prioridade para se unir ao cluster.

O segundo nó se une ao cluster

5. Os Serviços de cluster do PowerHA SystemMirror são iniciados no nó B. O Nó B transmite uma mensagem indicando que está pronto para se unir ao cluster em todas as redes configuradas às quais está conectado (`node_up`).
6. O Nó A recebe a mensagem e envia uma confirmação.
7. O Nó A inclui o Nó B em uma lista de nós ativos e inicia comunicações keep-alive com o nó B.
8. O Nó B recebe a confirmação do nó A. A mensagem inclui informações que identificam o nó A como o único outro membro do cluster. (Se havia outros membros, o nó B receberia a lista de membros.)
9. O Nó B processa o script `process_resources` e envia uma mensagem para deixar que outros nós saibam quando está concluído.

O processamento do script `process_resources` pode incluir os recursos de liberação do nó A que mantém atualmente. Se ambos os nós A e B estiverem na lista de nós do grupo de recursos para um ou mais grupos de recursos e o nó B tiver uma prioridade mais alta para um ou mais desses recursos. Isso é verdadeiro apenas para grupos de recursos que suportam fallback.

Observe que se o cronômetro de fallback atrasado estiver configurado, qualquer grupo de recursos que estiver on-line no nó A para o qual o nó B é um nó de prioridade mais alta cairá novamente no nó B no momento especificado pelo cronômetro de fallback atrasado.

10. Enquanto isso, o nó B esteve monitorando e enviando keep-alive e aguardando receber mensagens sobre mudanças na associação de cluster. Quando o nó B conclui seu próprio script **process_resources**, ele notifica o nó A.

Durante seu processamento **node_up**, o nó B solicita todos os grupos de recursos configurados para ele (consulte a etapa 3). Observe que se o cronômetro de fallback atrasado estiver configurado, o grupo de recursos cairá novamente em um nó de maior prioridade no momento especificado pelo cronômetro.

11. Ambos os nós processam um evento **node_up_complete** simultaneamente.
Neste ponto, o nó B inclui o nó A na sua lista de nós do membro e sua lista de keepalive.
12. O Nó B envia uma mensagem de "novo membro" a todos os possíveis nós no cluster.
13. Quando o nó A obtém a mensagem, ele move o nó B de sua lista de nós ativos para sua lista de nós de membro.

Neste momento, todos os grupos de recurso reconfigurados para o nó A e o nó B ficam disponíveis para os clientes de cluster.

Os nós restantes se unem ao cluster

14. Conforme os serviços de cluster do PowerHA SystemMirror iniciam em cada nó do cluster restante, as etapas de 4 a 9 se repetem, com cada nó de membro enviando e recebendo mensagens de controle e processando eventos na ordem descrita. Observe especialmente que todos os nós devem confirmar o evento **node_up_complete** antes de concluir o processamento do evento e mover o novo nó para a lista de membro de cluster.

Conforme os novos nós se unem, o subsistema RSCT em cada nó estabelece comunicações e começa enviando pulsações. Os nós e os adaptadores unem sinais de heartbeat de RSCT que são formados de acordo com a definição na configuração do PowerHA SystemMirror. Quando o status de uma placa da interface de rede (NIC) ou nó é alterado, o gerenciador do cluster recebe a mudança de estado e gera o evento apropriado.

Unindo novamente o cluster

Quando um nó une novamente o cluster, os gerenciadores de cluster que estão em execução nos nós existentes iniciam um evento **node_up** para reconhecer que o nó retornado está ativo. Quando esses nós tiverem concluído o processamento de seu script **process_resources**, o novo nó processa um evento **node_up** para que possa retomar o fornecimento de serviços de cluster.

Este processamento é necessário para assegurar o equilíbrio adequado de recursos de cluster. Contudo que os gerenciadores de cluster existentes primeiro reconheçam um nó que está unindo novamente o cluster, eles podem liberar quaisquer grupos de recursos pertencentes a esse nó se necessário. Se os grupos de recursos forem realmente liberados ou não nesta situação depende de como os grupos de recursos são configurados para o controle (ou dependências). O novo nó pode iniciar suas operações.

Sequência de eventos node_up

A lista a seguir descreve a sequência de eventos **node_up**:

node_up

Este evento ocorre quando um nó une ou une novamente o cluster.

process_resources

Este script chama os subeventos necessários para que o nó adquira o endereço de serviço (ou endereço compartilhado), obtém todos os seus recursos próprios (ou compartilhados) e controla os recursos. Isso inclui tornar os discos disponíveis, variando em grupos de volume, sistemas de arquivo de montagem, sistemas de arquivo de exportação, sistemas de arquivos NFS de montagem e variando em grupos de volume de acesso simultâneo.

process_resources_complete

Cada nó executa este script quando os recursos foram processados.

node_up_complete

Este evento ocorre após os recursos serem processados e após o evento **node_up** ser concluído com sucesso. Dependendo se o nó é local ou remoto, esse evento chama o script **start_server** para iniciar controladores de aplicativos no nó local ou permite que o nó local monte um sistema de arquivos NFS apenas após o nó remoto estar completamente ativo.

Eventos node_up com grupos de recursos dependentes

Se as dependências entre quaisquer grupos de recursos estiverem configurados no cluster, o PowerHA SystemMirror processará todos os eventos relacionados aos grupos de recursos no cluster com o uso de eventos **rg_move** que são ativados para todos os grupos de recursos quando os eventos **node_up** ocorrerem.

O gerenciador do cluster leva em consideração todas as políticas de nó, especialmente a configuração de dependências para grupos de recursos e a distribuição atual e o estado dos grupos de recursos em todos os nós a fim de manipular adequadamente grupos de recursos ao serem adquiridos, liberados, colocados on-line ou off-line antes que um evento **node_up_complete** possa ser executado.

As dependências de local ou pai e filho entre os grupos de recursos oferecem uma forma previsível e confiável de construir clusters com aplicativos com multicamadas. Entretanto, **node_up** o processamento em cluster com dependências poderia levar mais tempo que o processamento paralelo em clusters sem dependências dos grupos de recursos. É possível que seja preciso ajustar o cronômetro de aviso **config_too_long** para eventos **node_up**.

Eventos node_down

Quando todas as interfaces de rede estiverem inativas ou um nó não responder a pulsações, o gerenciador do cluster executará um evento **node_down**. Dependendo da configuração de cluster, os nós de peer tomam as ações necessárias para que os aplicativos críticos funcionem e para assegurar que os dados permaneçam disponíveis.

Um evento **node_down** pode ser iniciado por um nó:

- Parando os serviços de cluster e colocando os grupos de recursos off-line
- Parando os serviços de cluster e movendo os grupos de recursos para outro nó
- Parando os serviços de cluster e colocando os grupos de recursos em um estado não gerenciado.
- Falhando.

Parando os serviços de cluster e colocando os grupos de recursos off-line

Ao parar os serviços de cluster e colocar os grupos de recursos off-line, o PowerHA SystemMirror para no nó local após o evento **node_down_complete** liberar os recursos do nó interrompido. Os outros nós executam o evento **node_down_complete** e não controlam os recursos do nó interrompido.

Parando os serviços de cluster e movendo os grupos de recursos

Ao parar os serviços de cluster e mover os grupos de recursos para outros nó, o PowerHA SystemMirror para após o evento **node_down_complete** no nó local liberar seus grupos de recursos. Os nós sobreviventes na lista de nós do grupo de recursos controlam esses grupos de recursos.

Parando os serviços de cluster e colocando os grupos de recursos em um estado não gerenciado

Ao parar os serviços de cluster e colocar os grupos de recursos em um estado não gerenciado, o software PowerHA SystemMirror para imediatamente no nó local. O evento **node_down_complete** é executado no nó interrompido. Os gerenciadores de cluster nos nós remotos processam eventos **node_down**, mas não controlam nenhum grupo de recursos. O nó interrompido não libera seus grupos de recursos.

Falha do nó

Quando um nó falha, o gerenciador do cluster nesse nó não tem tempo de gerar um evento **node_down**. Neste caso, os gerenciadores de cluster nos nós sobreviventes reconhecem que um evento **node_down** ocorreu (quando percebem que o nó com falha não está mais se comunicando) e eles acionam eventos **node_down**.

Isso inicia uma série de subeventos que reconfiguram o cluster para lidar com esse nó com falha. Com base na configuração de cluster, os nós sobreviventes na lista de nós de grupo de recursos controlam os grupos de recursos.

Sequência de eventos **node_down**

A lista a seguir descreve a sequência paralela padrão de eventos **node_down**:

1. **node_down**
2. Este evento ocorre quando um nó deixa intencionalmente o cluster ou falha.
3. Em alguns casos, o evento **node_down** recebe o parâmetro **forced**.
4. Todos os nós executam o evento **node_down**.
5. Todos os nós executam o evento **node_down**.
6. Todos os nós executam o script **process_resources**. Após o gerenciador do cluster avaliar o status dos grupos de recursos afetados e a configuração, ele inicia uma série de subeventos para redistribuir recursos conforme configurado para failover ou fallback.
7. Todos os nós executam o script **process_resources_complete**.
8. **node_down_complete**

Eventos de rede

O PowerHA SystemMirror distingue entre dois tipos de falha de rede, *local* e *global* e usa eventos de falha de rede diferentes para cada tipo de falha. O script do evento de falha de rede geralmente é customizado para enviar mensagens.

Sequência de eventos de rede

A tabela a seguir mostra os eventos de rede.

Tabela 2. Sequência de eventos de rede

Nome do evento	Descrição
network_down (local)	<p>Este evento ocorre quando apenas um determinado nó perdeu contato com uma rede. O evento possui o formato a seguir:</p> <pre>network_down node_name network_name</pre> <p>O gerenciador do cluster toma uma ação de recuperação seletiva para mover os grupos de recursos mais afetados para outros nós se o IP de serviço for configurado como parte do grupo de recursos. Os resultados das ações de recuperação são registrados para hacmp.out.</p>
network_down (global)	<p>Este evento ocorre quando todos os nós conectados a uma rede perderam contato com uma rede. Presume-se que, neste caso, essa falha relacionada à rede ocorreu em vez de uma falha relacionada ao nó. Esse evento possui o formato a seguir:</p> <pre>network_down -1 network_name</pre> <p>Nota: O argumento -1 é <i>-one</i>. Esse argumento indica que o evento network_down é global.</p> <p>O evento de falha de rede global envia uma notificação para o administrador do sistema, mas não toma ação posterior uma vez que ações apropriadas dependem da configuração de rede local.</p>

Tabela 2. Sequência de eventos de rede (continuação)

Nome do evento	Descrição
network_down_complete (local)	<p>Este evento ocorre após um evento de falha de rede local ser concluído. Possui o formato a seguir:</p> <pre>network_down_complete node_name network_name</pre> <p>Quando um evento de falha de rede local ocorrer, o gerenciador do cluster tomará ações de recuperação seletivas para grupos de recursos que contêm uma placa da interface de rede (NIC) de serviço conectada a essa rede.</p>
network_down_complete (global)	<p>Este evento ocorre após um evento de falha de rede global ter sido concluído. Possui o formato a seguir:</p> <pre>network_down_complete -1 network_name</pre> <p>O processamento padrão para esse evento não toma ações porque as ações apropriadas dependem da configuração de rede.</p>
network_up	<p>Este evento ocorre quando o gerenciador do cluster determina que uma rede se tornou disponível para uso. Sempre que uma rede se tornar disponível novamente, o PowerHA SystemMirror tentará colocar os grupos de recursos que contêm os rótulos de IP de serviço nessa rede novamente on-line.</p>
network_up_complete	<p>Este evento ocorre apenas após um evento network_up ser concluído com sucesso. Esse evento é geralmente customizado para notificar o administrador do sistema que um evento demanda atenção manual. Sempre que uma rede se tornar disponível novamente, o PowerHA SystemMirror tentará colocar os grupos de recursos que contêm os rótulos de IP de serviço nessa rede novamente on-line.</p>

Eventos da interface de rede

O gerenciador do cluster reage à falha, indisponibilidade ou junção de interfaces de rede iniciando um evento.

A tabela a seguir mostra os eventos de interface de rede.

Tabela 3. Eventos da interface de rede

Eventos da interface de rede	Descrição do evento
swap_adapter	<p>Este evento ocorre quando a interface que hospeda um rótulo de IP de serviço em um nó falha. O evento swap_adapter move o rótulo de IP de serviço para uma interface de inicialização na mesma rede do PowerHA SystemMirror e reconstrói a tabela de roteamento. Se o rótulo de IP de serviço for um alias de IP, ele será colocado na interface de inicialização como um rótulo de IP adicional. Caso contrário, o rótulo de IP de inicialização é removido da interface e colocado na interface com falha. Se a interface que agora está mantendo o rótulo de IP de serviço posteriormente falhar, swap_adapter poderá alternar para outra interface de inicialização se existir uma. Se um rótulo de IP de nó persistente foi designado à interface com falha, ele se moverá com o rótulo de serviço para a interface de inicialização.</p> <p>Nota: O PowerHA SystemMirror remove aliases de IP das interfaces no encerramento. Ele cria os aliases novamente quando a rede se torna operacional. O arquivo hacmp.out registra essas mudanças.</p>
swap_adapter_complete	<p>Este evento ocorre apenas após um evento swap_adapter ter sido concluído com sucesso. O evento swap_adapter_complete assegura que o cache de protocolo de resolução de endereço local (ARP) é atualizado excluindo entradas e realizando ping de endereços IP de cluster.</p>
fail_standby	<p>Este evento ocorre se uma interface de inicialização falhar ou se tornar indisponível como o resultado de um controle de endereço IP. O evento fail_standby exibe uma mensagem do console, que indica que uma interface de inicialização falhou ou não está mais disponível.</p>
join_standby	<p>Este evento ocorre se uma interface de inicialização se tornar disponível. O evento join_standby exibe uma mensagem do console, que indica que uma interface de inicialização se tornou disponível. No PowerHA SystemMirror, sempre que uma interface de rede se torna disponível, o PowerHA SystemMirror tenta colocar os grupos de recursos novamente on-line.</p>

Tabela 3. Eventos da interface de rede (continuação)

Eventos da interface de rede	Descrição do evento
fail_interface	Este evento ocorre se uma interface falhar e não há interface de inicialização disponível para recuperar o endereço de serviço. Os endereços de serviço de controle são monitorados. É possível ter uma falha de interface e nenhuma interface disponível para recuperação e outra interface na mesma rede. Esse evento se aplica a todas as redes, incluindo aquelas que usam aliasing de IP para recuperação. Quando um NIC de inicialização falha em uma rede que está configurada para IPAT via aliases de IP, o evento fail_interface é executado. Se a interface que falhou era um rótulo de serviço, um evento rg_move é acionado.
join_interface	Este evento ocorre se uma interface de inicialização se tornar disponível ou recuperar. Esse evento se aplica a todas as redes, incluindo aquelas que usam IPAT via aliases de IP para recuperação. As redes que usam aliases de IP por definição não têm interfaces de inicialização definidas, então o evento join_interface que é executado neste caso apenas indica que uma interface de inicialização se une ao cluster.

A falha de uma única interface de rede não gera eventos

Se apenas uma interface de rede estiver ativa em uma rede, o gerenciador do cluster não poderá gerar um evento de falha para essa interface de rede, porque não possui peers com os quais se comunicar determinando o funcionamento da interface. Situações que possuem uma única interface de rede são:

- Clusters de um nó
- Clusters de diversos nós com apenas um nó ativo
- Clusters de diversos nós com interfaces de Ethernet virtual
- Falha de todas menos uma interface em uma rede, uma por vez

Por exemplo, e iniciar um cluster com todas as interfaces de serviço ou inicialização produz os resultados a seguir:

- Primeiro nó ativo: nenhum evento de falha é gerado.
- Segundo nó ativo: um evento de falha é gerado.
- Terceiro nó ativo: um evento de falha é gerado.

Eventos de status em todo o cluster

Por padrão, o gerenciador do cluster reconhece um limite de tempo para reconfigurar um cluster e processar mudanças de topologia. Se o limite de tempo for atingido, o gerenciador do cluster iniciará um evento **config_too_long**.

Os eventos de status do cluster total incluem:

Tabela 4. Eventos de status em todo o cluster

Nomes de eventos de status de todo o cluster	Descrição do evento
config_too_long	Este aviso de sistema ocorre toda vez que um evento de cluster leva mais tempo para concluir que um período de tempo limite especificado. Esta mensagem é registrada no arquivo hacmp.out . O período de tempo limite para todos os eventos é configurado como 360 segundos por padrão. É possível usar o SMIT para customizar o período permitido para a conclusão de um evento de cluster antes de o PowerHA SystemMirror emitir um config_too_long para ele.
reconfig_topology_start	Este evento marca o início de uma reconfiguração dinâmica da topologia de cluster.
reconfig_topology_complete	Este evento indica que uma reconfiguração dinâmica de topologia de cluster foi concluída.
reconfig_resource_acquire	Este evento indica que os recursos de cluster afetados pela reconfiguração dinâmica estão sendo adquiridos pelos nós apropriados.
reconfig_resource_release	Este evento indica que os recursos de cluster afetados pela reconfiguração dinâmica estão sendo liberados pelos nós apropriados.

Tabela 4. Eventos de status em todo o cluster (continuação)

Nomes de eventos de status de todo o cluster	Descrição do evento
reconfig_resource_complete	Este evento indica que uma reconfiguração dinâmica de recurso de cluster foi concluída com sucesso.
cluster_notify	Este evento é acionado pela verificação quando o monitoramento de configuração de cluster automático detecta erros na configuração do cluster. A saída desse evento é registrada no arquivo de log hacmp.out em todo o cluster em cada nó que está executando serviços de cluster.
event_error	Todos os nós de cluster executam o evento event_error se qualquer nó tiver um erro fatal. Todos os nós registram o erro e chamam o nome do nó com falha no arquivo de log hacmp.out .

Manipulação e recuperação de evento do grupo de recursos

O gerenciador do cluster mantém controle das políticas de prioridade de nó do grupo de recursos e quaisquer dependências configuradas, bem como as informações de topologia necessárias e o status do grupo de recursos para que possa tomar uma variedade maior de ações de recuperação, geralmente evitando a necessidade de intervenção do usuário. A criação de log de eventos inclui um resumo detalhado para cada evento de alto nível a fim de ajudá-lo a entender exatamente que ações foram realizadas para cada grupo de recursos durante a manipulação de falhas.

Para obter mais informações sobre como os grupos de recursos são manipulados no PowerHA SystemMirror, consulte Comportamento do grupo de recursos durante eventos de cluster. Este tópico contém informações sobre as funções do PowerHA SystemMirror a seguir:

- Fallover seletivo para manipulação de grupos de recursos
- Manipulação de falhas de aquisição do grupo de recursos
- Manipulação de grupos de recursos configurados com recursos de IP de serviço
- Manipulação de grupos de recursos do PowerHA SystemMirror Enterprise Edition

Informações relacionadas:

Comportamento do grupo de recursos durante eventos de cluster

Eventos do grupo de recursos

O gerenciador do cluster pode mover os grupos de recursos como um resultado das ações de recuperação tomadas durante o processamento de eventos como nó inativo.

Notas:

- Se as dependências entre os grupos de recursos ou sites forem especificadas, o PowerHA SystemMirror processará eventos em uma sequência diferente do comum.
- As listas na tabela a seguir não incluem todos os estados do grupo de recursos possíveis. Além disso, as instâncias do grupo de recursos poderiam estar no processo de aquisição ou liberação. Os estados do grupo de recursos correspondente não são listados aqui, mas há nomes descritivos que explicam quais ações ocorrem.

Tabela 5. Eventos do grupo de recursos

Nome do evento do grupo de recursos	Descrição do evento
rg_move	Este evento move um grupo de recursos especificado de um nó para o outro.
rg_move_complete	Esta ação indica que o evento rg_move foi concluído com sucesso.
resource_state_change	Este evento acionador é usado para a recuperação do grupo de recursos se as dependências do grupo de recursos forem configuradas no cluster. Esta ação indica que o gerenciador do cluster precisa alterar o estado de um ou mais grupos de recursos ou existe uma mudança no estado de um recurso gerenciado pelo gerenciador do cluster. Esse evento é executado em todos os nós se uma das situações a seguir ocorrer: <ul style="list-style-type: none"> • Falha de monitoramento de aplicativo • Fallover seletivo para perda de grupo de volumes • Rede local inativa • Falha de WAN • Falha de aquisição do grupo de recursos • Recuperação do grupo de recursos na disponibilidade da interface IP • Expiração do cronômetro de acomodação para um grupo de recursos • Expiração do cronômetro de fallback para um grupo de recursos.
resource_state_change_complete	Este evento é executado quando o evento resource_state_change é concluído com sucesso. É possível incluir pré e pós-eventos aqui se necessário. Você pode querer ser notificado sobre as mudanças do estado de recurso, por exemplo.
external_resource_state_change	Este evento é executado quando você move um grupo de recursos e o PowerHA SystemMirror usa o caminho do processamento usar para manipular a solicitação porque as dependências do grupo de recursos são configuradas no cluster.
external_resource_state_change_complete	Este evento é executado quando o evento external_resource_state_change é concluído com sucesso.

Subeventos do grupo de recursos

A manipulação dos recursos individuais durante o processamento de um evento pode incluir as ações a seguir. Por exemplo, quando um sistema de arquivos está no processo de ser desmontado e montado, ele é colocado off-line e liberado por um nó. Em seguida, é adquirido por outro nó e colocado on-line.

A tabela a seguir inclui algum, mas não todos os estados possíveis de grupos de recursos:

Tabela 6. Subeventos do grupo de recursos

Subeventos do grupo de recursos	Descrição do evento
releasing	Esta ação indica que um grupo de recursos está sendo liberado para ser colocado off-line ou adquirido em outro nó.
acquiring	Esta ação é usada quando um grupo de recursos está sendo adquirido em um nó.
rg_up	Esta ação indica que o grupo de recursos está on-line.
rg_down	Esta ação indica que o grupo de recursos está off-line.
rg_error	Esta ação indica que o grupo de recursos está em estado de erro.
rg_acquiring_secondary	Esta ação indica que o grupo de recursos está entrando on-line no site de destino (apenas os recursos replicados estão on-line).
rg_up_secondary	Esta ação indica que o grupo de recursos está on-line na função secundário no site de destino (apenas recursos replicados estão on-line).
rg_error_secondary	Esta ação indica que o grupo de recursos no site que está recebendo os dados de espelhamento está em estado de erro.
rg_temp_error_state	Esta ação indica que o grupo de recursos está em um estado de erro temporário. Por exemplo, ocorre devido a uma falha de rede local ou aplicativo. Esse estado informa o gerenciador do cluster para iniciar um evento rg_move para este grupo de recursos. Os grupos de recursos não devem estar nesse estado quando o cluster estiver estável.

Após a conclusão de um evento, o gerenciador do cluster tem o estado dos recursos e grupos de recursos envolvidos no evento. O gerenciador do cluster analisa as informações do grupo de recursos que mantém internamente e determina se os eventos de recuperação precisam ser enfileirados para qualquer um dos grupos de recursos. O gerenciador do cluster também usa o status dos recursos individuais nos grupos de recursos para imprimir um resumo de evento abrangente para o arquivo de log **hacmp.out**.

Para cada grupo de recursos, o gerenciador do cluster mantém controle dos nós nos quais o grupo de recursos tentou ficar on-line e falhou. Essas informações são atualizadas quando os eventos de recuperação são processados. O gerenciador do cluster reconfigura a lista de nós para um grupo de recursos assim que o grupo de recursos mover para estados on-line ou de erro.

No PowerHA SystemMirror, os estados ERROR do grupo de recursos são exibidos com detalhes:

Tabela 7. Grupos de recursos em um estado ERROR

Causas para estados ERROR do grupo de recursos	O PowerHA SystemMirror exibe esta mensagem
O grupo-pai fica NOT ONLINE; como resultado, o grupo de recursos filho, indisponível	OFFLINE devido ao pai off-line
O grupo de dependência de nó diferente de maior prioridade fica ONLINE	OFFLINE devido à falta de nó disponível
Outro grupo distribuído foi adquirido	OFFLINE
O grupo está caindo e no estado OFFLINE temporariamente	OFFLINE

A intervenção manual só é necessária quando um grupo de recursos permanece no estado ERROR após o processamento de eventos ser concluído.

Quando um grupo de recursos está no processo de ser movido, o monitoramento de aplicativo é suspenso e retomado apropriadamente. O monitor de aplicativo reconhece que o aplicativo está no estado de recuperação enquanto o evento está sendo processado.

Tabela 8. Eventos de monitoramento de aplicativos

Eventos de monitoramento de aplicativos	Descrição do evento
resume_appmon	Esta ação é usada pelo monitor de aplicativo para retomar o monitoramento de um aplicativo.
suspend_appmon	Esta ação é usada pelo monitor de aplicativo para suspender o monitoramento de um aplicativo.

Informações relacionadas:

Comportamento do grupo de recursos durante eventos de cluster

Customizando o processamento de evento de cluster

A capacidade do gerenciador do cluster de reconhecer uma série específica de eventos e subeventos permite um esquema de customização flexível. Com o recurso de customização de evento do PowerHA SystemMirror, é possível customizar o processamento de evento de cluster para seu site. Com o processamento de eventos customizado, é possível fornecer um caminho altamente eficiente para os recursos mais críticos no evento de uma falha. Entretanto, esta eficiência depende da sua configuração.

Como parte do processo de planejamento, é necessário decidir se customizará o processamento de eventos. Se as ações tomadas pelos scripts padrão forem suficientes para seus fins, não é necessário fazer nada para configurar eventos durante o processo de configuração.

Se você decidir customizar o processamento de eventos para o seu ambiente, use o recurso de customização de evento do PowerHA SystemMirror descrito nesta seção. Se você customizar o processamento de eventos, registre scripts definidos pelo usuário com o PowerHA SystemMirror durante o processo de configuração.

O recurso de customização de evento inclui as funções a seguir:

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

A customização completa de um evento inclui uma notificação ao administrador do sistema (antes e após o processamento de eventos) e os comandos ou scripts definidos pelo usuário que são executados antes e após o processamento de eventos, conforme mostrado no exemplo a seguir:

```
Notify sysadmin of event to be processed
Script ou comando de pré-evento
PowerHA SystemMirror event script
Script ou comando pós-evento
Notify sysadmin that event processing is complete
```

Notificação de evento

É possível especificar um comando **notify** que envie mensagens para indicar que um evento está prestes a acontecer (ou acabou de ocorrer) e que um script do evento foi bem-sucedido ou falhou.

Você configura métodos de notificação para eventos de cluster no SMIT no menu **Configuração de cluster customizada > Eventos > Eventos de cluster > Alterar/mostrar eventos pré-definidos**. Por exemplo, um cluster pode querer usar um evento de notificação de falha de rede para informar aos administradores do sistema que o tráfego pode ter de ser roteado novamente. Depois disso, é possível usar um evento de notificação **network_up** para informar os administradores de sistema que o tráfego pode ser atendido novamente por meio da rede restaurada.

A notificação de evento em um cluster do PowerHA SystemMirror também pode ser feita usando scripts pré-evento e pós-evento.

Também é possível configurar uma notificação remota customizada em resposta aos eventos.

Referências relacionadas:

“Notificação remota customizada de eventos” na página 97

É possível definir um método de notificação por meio da interface SMIT para emitir uma página customizada em resposta a um evento de cluster. É possível enviar a notificação de mensagem de texto para qualquer número que inclui um telefone celular ou é possível enviar a notificação para um endereço de email.

Scripts pré-evento e pós-evento

É possível especificar comandos ou diversos scripts definidos pelo usuário que executam antes e após o gerenciador do cluster chamar um script do evento.

Por exemplo, é possível especificar um ou mais scripts pré-evento que são executados antes do script do evento **node_down** ser processado. Quando o gerenciador do cluster reconhece que um nó remoto está inativo, primeiro ele processa esses scripts definidos pelo usuário. Esse script pode designar que uma mensagem seja enviada a todos os usuários para indicar que esse desempenho pode ser afetado (quando os adaptadores são trocados e quando os controladores de aplicativos são interrompidos e reiniciados). Depois do script do evento **node_down**, um script de evento de pós-processamento para notificação **network_up** pode ser incluído para transmitir uma mensagem a todos os usuários de que um certo sistema agora está disponível em outro endereço de rede.

Os cenários a seguir são outros exemplos de onde o processamento pré-evento e pós-evento é útil:

- Se um evento **node_down** ocorrer, esse script poderia notificar os usuários no servidor que está prestes a controlar para o controlador de aplicativo inativo que o desempenho pode variar ou que devem buscar sistemas alternativos para certos aplicativos.
- Devido a uma rede que está ficando inativa, uma instalação customizada pode ser capaz de rotear novamente o tráfego por meio de outras máquinas criando novas rotas de IP. Os scripts de evento **network_up** e **network_up_complete** poderiam reverter o procedimento, assegurando que as rotas corretas existam após todas as redes estarem funcionando.
- É possível parar os serviços de cluster e mover grupos de recursos para outro nó como um script pós-evento se uma rede falhou no nó local (mas, por outro lado, a rede está funcionando).

Observe que ao gravar seus scripts pré-evento e pós-evento do PowerHA SystemMirror, nenhuma das variáveis de ambiente de shell definidas em **/etc/environment** estão disponíveis ao seu programa. Se precisar usar quaisquer uma dessas variáveis, acesse-as explicitamente incluindo esta linha em seu script:

```
". /etc/environment"
```

Se você planejar criar scripts pré-evento ou pós-evento para seu cluster, esteja ciente de que seus scripts serão transmitidos para os mesmos parâmetros usados pelo script de evento do PowerHA SystemMirror que especificar. Para scripts pré-evento e pós-evento, os argumentos transmitidos ao comando de evento são o nome do evento, status de saída do evento e os argumentos à direita transmitidos ao comando de evento.

Todos os scripts de evento do PowerHA SystemMirror são mantidos no diretório **/usr/es/sbin/cluster/events**. Os parâmetros transmitidos ao seu script são listados nos cabeçalhos de script do evento.

CUIDADO:

Tenha cuidado para não eliminar quaisquer processos do PowerHA SystemMirror como parte do seu script. Se estiver usando a saída do comando **ps** e usando **grep** para procurar um certo padrão, assegure-se de que o padrão não corresponda a nenhum dos processos do PowerHA SystemMirror, Cluster Aware AIX (CAA) ou Reliable Scalable Cluster Technology (RSCT).

Os scripts pré-evento e pós-evento podem não ser necessários

Se você migrou de uma versão anterior do PowerHA SystemMirror, alguns de seus scripts pré-evento e pós-evento existentes podem não ser mais necessários. O PowerHA SystemMirror em si manipula mais situações.

Usando o atributo de ativação forçada em vez de scripts pré-evento ou pós-evento

Se o atributo ativação forçada for especificado para um grupo de volumes, scripts especiais para forçar uma operação de ativação não são mais necessários.

O evento `event_error` indica falha em um nó remoto

Historicamente, falhas de script do evento não recuperáveis resultam no evento **event_error** sendo executado no nó do cluster em que a falha ocorreu. Os nós do cluster restantes não indicaram a falha. Com o PowerHA SystemMirror, todos os nós do cluster executam o evento **event_error** se qualquer nó tiver um erro irrecuperável. Todos os nós registram o erro e o nome do nó com falha no arquivo de log **hacmp.out**.

Se você tiver incluído pré-evento ou pós-evento para o evento **event_error**, esteja ciente de que esses métodos de eventos sejam chamados em cada nó, não só o nó com falha.

Uma variável de ambiente de shell Korn indica o nó em que o script do evento falhou: **EVENT_FAILED_NODE** está configurado para o nome do nó em que o evento falhou. Use esta variável em seu script pré-evento ou pós-evento para determinar onde a falha ocorreu.

A variável LOCALNODENAME identifica o nó local. Se LOCALNODENAME não for o mesmo que EVENT_FAILED_NODE, a falha ocorreu em um nó remoto.

Grupos de recursos processados em paralelo e usando scripts de pré-evento e pós-evento

Os grupos de recursos são processados em paralelo por padrão no PowerHA SystemMirror a menos que especifique uma ordem de processamento serial customizada para todos ou alguns dos grupos de recursos no cluster.

Quando os grupos de recursos são processados em paralelo, menos eventos de cluster ocorrem no cluster e aparecem nos resumos de evento.

O uso de processamento paralelo reduz o número de eventos de cluster específicos para os quais é possível criar scripts pré-evento ou pós-evento customizados. Se você começar a usar o processamento paralelo para uma lista de grupos de recursos em sua configuração, esteja ciente de que alguns dos scripts pré-evento e pós-evento existentes podem não funcionar nesses grupos de recursos.

Em particular, apenas os eventos a seguir ocorrem durante o processamento paralelo dos grupos de recursos:

- acquire_svc_addr**
- acquire_takeover_addr**
- node_down**
- node_up**
- release_svc_addr**
- release_takeover_addr**
- start_server**
- stop_server**

Nota: No processamento paralelo, esses eventos se aplicam a toda uma lista de grupos de recursos que estão sendo processados em paralelo e não a um único grupo de recursos, como no processamento serial. Se você tiver scripts pré-evento e pós-evento configurados para esses eventos, então após a migração, esses scripts de eventos são ativados não para um único grupo de recursos, mas para uma lista de grupos de recursos e podem não funcionar conforme o esperado.

Os eventos a seguir não ocorrem no processamento paralelo dos grupos de recursos:

- get_disk_vg_fs**
- node_down_local**
- node_down_remote**
- node_down_local_complete**
- node_down_remote_complete**
- node_up_local**
- node_up_remote**
- node_up_local_complete**
- node_up_remote_complete**
- release_vg_fs**

Considere esses eventos que não ocorrem no processamento paralelo se você tiver scripts pré-evento e pós-evento e planeje fazer upgrade para a versão atual.

Se desejar continuar usando scripts de pré-evento e pós-evento, você poderia ter um dos casos a seguir.

Cenário	O que você deve fazer
<p>Você gostaria de usar os scripts pré-evento e pós-evento para grupos de recursos recém-incluídos.</p>	<p>Todos os grupos de recursos recém-incluídos são processados em paralelo, o que resulta em menos eventos de cluster. Portanto, existe uma opção limitada de eventos para os quais é possível criar scripts de pré-evento e pós-evento.</p> <p>Neste caso, se você tiver recursos nos grupos de recursos que requerem a manipulação de scripts de pré-evento e pós-evento gravados para eventos de cluster específicos, inclua esses grupos de recursos nas listas de processamento serial no SMIT para assegurar que os scripts de pré-evento e pós-evento específicos possam ser usados para esses recursos.</p> <p>Para obter informações sobre especificar o processamento serial ou paralelo de grupos de recursos, consulte a seção Configurando ordem de processamento para grupos de recursos.</p>
<p>Você faz upgrade para o PowerHA SystemMirror 4.5 ou posterior e escolhe o processamento paralelo para alguns grupos de recursos pré-existent em sua configuração.</p>	<p>Se, antes da migração você configurou scripts de pré-evento ou pós-evento customizados em seu cluster, agora que esses grupos de recursos são processados em paralelo após a migração, os scripts de evento para um número de eventos não podem ser usados para esses grupos de recursos, uma vez que esses eventos não ocorrem no processamento paralelo.</p> <p>Se desejar que os scripts de evento existente continuem funcionando para os grupos de recursos, inclua esses grupos de recursos nas listas de ordens seriais no SMIT, para assegurar que os scripts pré-evento e pós-evento possam ser usados para esses recursos.</p> <p>Para obter informações sobre especificar o processamento serial ou paralelo dos grupos de recursos, consulte a seção Configurando a ordem de processamento para grupos de recursos.</p>

Referências relacionadas:

“Usando ativação forçada” na página 51

O PowerHA SystemMirror fornece uma função de ativação forçada para usar em conjunto com os métodos de notificação automática de erro do AIX. A função de ativação forçada permite ter a disponibilidade de dados mais alta possível.

Grupos de recursos dependentes e scripts pré-evento e pós-evento

Historicamente, para atingir o grupo de recursos e o sequenciamento de aplicativo, os administradores de sistema tinham que construir a lógica de recuperação do aplicativo em seus scripts de processamento pré e pós-evento. Cada cluster seria configurado com um script pré-evento para todos os eventos do cluster e um script de pós-evento para todos os eventos de cluster. As liberações mais recentes do PowerHA SystemMirror incluem muito mais opções de configuração que permitem que os clientes usem políticas integradas para atingir o mesmo sequenciamento e disposição de grupos de recursos. Se estiver usando atualmente eventos pré e pós-eventos para atingir o sequenciamento de grupo de recursos, poderá querer revisar sua implementação e transição para os mecanismos integrados.

Tais scripts poderiam todos incluir instruções de caso. Por exemplo, se desejar tomar uma ação para um evento específico em um nó específico, é necessário editar esse caso individual, incluir o código necessário para scripts pré-evento e pós-evento e também assegurar que os scripts são os mesmos entre todos os nós.

Para resumir, mesmo que a lógica de tais scripts capture o comportamento desejado do cluster, eles podem ser difíceis de customizar e ainda mais difíceis de manter no futuro, quando a configuração de cluster for alterada.

Se estiver usando scripts pré-evento e pós-evento ou outros métodos, como ordem de processamento do grupo de recursos para estabelecer dependências entre aplicativos suportados por seu cluster, esses métodos podem não ser mais necessários ou podem ser simplificados significativamente. Em vez disso, é possível especificar dependências entre grupos de recursos em um cluster. Para obter mais informações sobre o planejamento de grupos de recursos dependentes, consulte Dependências do grupo de recursos.

Se tiver aplicativos incluídos nos grupos de recursos dependentes e ainda planejar usar scripts pré-evento e pós-evento além das dependências, a customização adicional de scripts pré-evento e pós-evento pode ser necessária. Para minimizar a chance de perda de dados durante o processo de parada e reinício do aplicativo, customize seus scripts do controlador de aplicativo para assegurar que quaisquer dados não confirmados sejam armazenados em um disco compartilhado temporariamente durante o processo de parada do aplicativo e lidos de volta no aplicativo durante o processo de reinício do aplicativo. É importante usar um disco compartilhado porque o aplicativo pode ser reiniciado em um nó que não aquele no qual ele foi parado.

Referências relacionadas:

“Dependências do grupo de recursos” na página 65

O PowerHA SystemMirror oferece uma ampla variedade de configurações em que é possível especificar os relacionamentos entre grupos de recursos que deseja manter na inicialização, failover e fallback.

Recuperação e nova tentativa de evento

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

Por exemplo, um comando de recuperação poderia incluir a nova tentativa de desmontar um sistema de arquivos após desconectar um usuário e assegurar que nenhum estava acessando atualmente o sistema de arquivos.

Se uma condição que afeta o processamento de um determinado evento em um cluster for identificada, como um problema de sincronização, é possível inserir um comando de recuperação com uma contagem de nova tentativa alta o bastante para assegurar a recuperação do problema.

Notificação remota customizada de eventos

É possível definir um método de notificação por meio da interface SMIT para emitir uma página customizada em resposta a um evento de cluster. É possível enviar a notificação de mensagem de texto para qualquer número que inclui um telefone celular ou é possível enviar a notificação para um endereço de email.

É possível usar o evento de monitoramento automático de verificação **cluster_notify** para configurar um método de notificação remoto do PowerHA SystemMirror para enviar uma mensagem em caso de erros detectados na configuração do cluster. A saída desse evento é registrada no arquivo **hacmp.out** em todo o cluster em cada nó que esteja executando serviços de cluster.

É possível configurar qualquer número de métodos de notificação, para diferentes eventos e com mensagens de texto ou numéricas diferentes e números de telefone para discagem. 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.

Após configurar o método de notificação, é possível enviar uma mensagem de teste para assegurar que tudo esteja configurado corretamente e que a mensagem esperada será enviada para um determinado evento.

Planejamento para notificação remota customizada

A notificação remota requer as condições a seguir:

- Qualquer porta especificada deve ser definida para o sistema operacional AIX e deve estar disponível.
- Cada nó que enviar uma página ou mensagens de texto deve ter um modem apropriado instalado e ativado.

Nota: O PowerHA SystemMirror verifica a disponibilidade da porta quando o método de notificação é configurado e antes de uma página ser emitida. O status do modem não é verificado.

- Cada nó que enviar mensagens de email do painel do SMIT usando o correio do AIX deve ter uma conexão TCP/IP com a Internet.
- Cada nó que pode enviar mensagens de texto a um celular deve ter um modem discador compatível com Hayes apropriado instalado e ativado.

Customizando o tempo de duração do evento até o aviso

Dependendo da configuração de cluster, a velocidade dos nós do cluster e o número de tipos de recursos que precisam ser movidos durante eventos de cluster, certos eventos podem usar diferentes intervalos de tempo para concluir. Para tais eventos, você pode querer customizar o período que o PowerHA SystemMirror aguarda para que um evento seja concluído antes de emitir a mensagem de aviso **config_too_long**.

Os eventos de cluster que incluem adquirir e liberar grupos de recursos levam um tempo maior para conclusão. Os eventos de cluster a seguir são considerados eventos *lentos*:

- **node_up**
- **node_down**
- **reconfig_resource**
- **rg_move**

Customize o tempo de duração de evento para os eventos de cluster *lentos* para evitar obter avisos de sistema desnecessários durante a operação normal do cluster.

Todos os outros eventos de cluster são considerados eventos *rápidos*. Esses eventos geralmente levam um tempo menor para concluir e não envolvem adquirir ou liberar recursos. Exemplos de eventos rápidos incluem:

- **swap_adapter**
- Eventos que não manipulam grupos de recursos

É possível customizar o tempo de duração do evento antes de receber um aviso para eventos rápidos para tomar a ação corretiva mais rapidamente.

Considere customizar **Tempo de duração de evento até o aviso** se, no caso de eventos de cluster lentos, o PowerHA SystemMirror emitirá mensagens de aviso muito frequentemente. No caso de eventos rápidos, você deseja acelerar a detecção de um possível evento do problema.

Nota: As dependências entre os grupos de recursos oferecem uma forma previsível e confiável de construir clusters com aplicativos com multicamadas. Entretanto, o processamento de alguns eventos de cluster (como **node_up**) em clusters com dependências poderia levar mais tempo que o processamento desses eventos em que todos os grupos de recursos são processados em paralelo. Sempre que uma dependência de grupo de recursos permite, o PowerHA SystemMirror processa diversos grupos de recursos não simultâneos em paralelo e processa diversos grupos de recursos simultâneos em todos os nós de uma vez. Entretanto, um grupo de recursos que é dependente em outros grupos de recursos não pode ser iniciado até que os outros tenham iniciado primeiro. O cronômetro de aviso **config_too_long** para eventos **node_up** deve ser configurado grande o bastante para permitir isso.

Eventos definidos pelo usuário

É possível definir seus próprios eventos para os quais o PowerHA SystemMirror pode executar programas de recuperação especificados. Este processo inclui uma nova dimensão no recurso de customização de script pré e pós-evento predefinido do PowerHA SystemMirror.

Você especifica o mapeamento entre eventos que você define e os programas de recuperação que definem as ações de recuperação de evento por meio da interface SMIT. Com esse mapeamento, você controla o escopo de cada ação de recuperação e o número de etapas do evento sincronizado entre todos os nós.

Um *recurso* do RMC refere-se a uma instância de uma entidade física ou lógica que fornece serviços para algum outro componente do sistema. O recurso de termo é usado muito amplamente para se referir a entidades de software e hardware. Por exemplo, um recurso poderia ser um sistema de arquivos específico ou uma máquina host específica. Uma *classe de recurso* refere-se a todos os recursos do mesmo tipo, como processadores ou máquinas host.

Um *gerenciador de recursos* (daemon) mapeia entidades reais para abstrações do RMC. Cada gerenciador de recursos representa um conjunto específico de tarefas administrativas ou funções do sistema. O gerenciador de recursos identifica os tipos de entidade física ou lógica chave relacionados a esse conjunto de tarefas administrativas ou funções do sistema e define classes de recursos para representar esses tipos de entidade.

Por exemplo, o gerenciador de recurso de host contém um conjunto de classes de recursos para representar aspectos de uma máquina host individual. Ele define classes de recursos para representar:

- Máquinas individuais (IBM.Host)
- Dispositivos de paginação (IBM.PagingDevice)
- Volumes físicos (IBM.PhysicalVolume)
- Processadores (IBM.Processor)
- Um token de identificador de host (IBM.HostPublic)
- Programas em execução no host (IBM.Program)
- Cada tipo de adaptador Ethernet suportado pelo host

O *monitor de recurso* AIX gera eventos para condições de recurso relacionadas ao S.O. como a porcentagem de CPU que está inativa (IBM.Host.PctTotalTimeIdle) ou a porcentagem de espaço de disco em uso (IBM.PhysicalVolume.PctBusy). O *monitor de recurso de programa* gera eventos para ocorrências relacionadas a processo, como o fim inesperado de um processo. O monitor de recurso de programa usa o atributo de recurso IBM.Program.ProgramName.

Gravando programas de recuperação

Um programa de recuperação tem a sequência de especificações de comando de recuperação, possivelmente intercalada com comandos **barrier**.

O formato dessas especificações segue:

```
:node_set recovery_command expected_status NULL
```

em que:

- *node_set* é um conjunto de nós nos quais o programa de recuperação deve ser executado
- *recovery_command* é uma sequência delimitada por aspas que especifica um caminho completo para o programa executável. O comando não pode incluir argumentos. Qualquer programa executável que requer argumentos deve ser um script separado. O programa de recuperação deve estar nesse caminho em todos os nós no cluster. O programa deve especificar um status de saída.
- *expected_status* é um status de número inteiro a ser retornado quando o comando de recuperação é concluído com sucesso. O gerenciador do cluster compara o status real retornado para o status esperado. Uma incompatibilidade indica uma recuperação malsucedida. Se você especificar o caractere X no campo de status esperado, o gerenciador do cluster omite a comparação.
- *NULL* não é usado atualmente.

Você especifica conjuntos de nós por relacionamentos dinâmicos. O PowerHA SystemMirror suporta os relacionamentos dinâmicos a seguir:

Todos O comando de recuperação é executado em todos os nós na associação atual.

Evento

O nó em que o evento ocorreu.

Outro Todos os nós exceto aquele no qual o evento ocorreu.

O relacionamento dinâmico especificado gera um conjunto de comandos de recuperação idêntico ao original, exceto que um ID do nó substitui *node_set* em cada conjunto de comandos.

A sequência de caracteres de comando para comandos de eventos definidos pelo usuário devem iniciar com uma barra (/). O comando **clcallev** executa comandos que não iniciam com uma barra.

Úteis e referências úteis para informações de RMC

Para listar todas as definições de atributo persistente para o recurso do RMC IBM.Host (campo *sequência de seleção*):

```
lsrsrcdef -e -A p IBM.Host
```

Para listar todas as definições de atributo dinâmico para o recurso do RMC IBM.Host (campo *Expressão*):

```
lsrsrcdef -e -A d IBM.Host
```

Exemplo: programa de recuperação

Um programa de amostra envia uma mensagem para **/tmp/r1.out** que o espaço de paginação está baixo no nó em que o evento ocorreu. Para o programa de recuperação **r1.rp**, os campos do SMIT seriam preenchidos como a seguir.

Tabela 9. Exemplo: campos de programa de recuperação

Campo	Valor
Nome do evento	E_page_space (nome definido pelo usuário)
Caminho do programa de recuperação	/r1.rp
Nome do recurso	IBM.Host (nó do cluster)
Sequência de seleção	Name = ?" (nome do nó)
Expressão	TotalPgSpFree < 256000 (VMM está em 200 MB do nível de aviso de espaço de paginação).
Expressão Rearm	O atributo de recurso mais a condição que deseja sinalizar. TotalPgSpFree >256000 O atributo de recurso mais a condição ajustada.

Em que o programa de recuperação **r1.rp** é como a seguir:

```
#format:  
#relationship >command to run >expected status NULL  
#  
event "/tmp/checkpagingspace" 0 NULL
```

O programa de recuperação não executa um comando com argumentos sozinho. Em vez disso, aponta para um shell script, **/tmp/checkpagingspace**, que contém:

```
#!/bin/ksh  
/usr/bin/echo "Paging Space LOW!" > /tmp/r1.out  
exit 0
```

Programa de recuperação para o exemplo de evento node_up

O exemplo a seguir é um programa de recuperação para o evento **node_up**:

```
#format:
#relationshipcommand to run expected status NULL
#
other "node_up" 0 NULL
#
barrier
#
event "node_up" 0 NULL
#
barrier
#
all "node_up_complete" X NULL
```

Comandos de barreira

É possível colocar qualquer número de comandos de barreira no programa de recuperação. Todos os comandos de recuperação antes de uma barreira iniciar em paralelo. Após um nó encontrar um comando de barreira, todos os nós devem atingi-lo antes de o programa de recuperação continuar.

A sintaxe do comando de barreira é **barrier**.

Acumulação de evento

Se diversos eventos estiverem pendentes simultaneamente, você só verá o evento de prioridade mais alta. Os eventos de nó são de prioridade mais alta que os eventos de rede. Mas os eventos definidos pelo usuário, da prioridade mais baixa, não acumulam de forma alguma, então você vê todos.

Resumos de eventos e prefácio

Quando os eventos são registrados para um arquivo de log **hacmp.out** do nó, a saída detalhada contém várias linhas de detalhes de eventos seguidos por um resumo conciso do evento. Os resumos de evento facilitam varrer o log para importantes eventos de cluster.

É possível visualizar uma compilação de apenas partes de resumo do evento dos últimos sete dias dos arquivos de log **hacmp.out** usando a opção **Visualizar resumos de eventos** no painel **Ferramentas de determinação de problemas** do SMIT. Os resumos de eventos podem ser compilados mesmo se você tiver redirecionado o arquivo **hacmp.out** para um local não padrão. O relatório **Exibir resumos de eventos** também inclui informações do grupo de recursos geradas pelo comando **clRGinfo**. Também é possível salvar os resumos de eventos em um arquivo especificado em vez de visualizá-los por meio do SMIT.

Quando os eventos manipularem grupos de recursos com dependências, um prefácio será gravado no arquivo de log **hacmp.out** que lista o plano dos subeventos para manipulação dos grupos de recursos.

Planejamento para clientes do PowerHA SystemMirror

Estes tópicos discutem considerações de planejamento para clientes do PowerHA SystemMirror. Esta é a última etapa antes de continuar a instalação do seu software do PowerHA SystemMirror.

Os clientes do PowerHA SystemMirror são dispositivos de usuário final que podem acessar os nós em um cluster do PowerHA SystemMirror. Para fins de planejamento, é importante avaliar o cluster a partir do ponto de vista dos clientes.

Clientes que executam Clinfo

O programa Clinfo chama o script `/usr/es/sbin/cluster/etc/clinfo.rc` sempre que um evento de rede ou nó ocorre. Por padrão, esta ação atualiza o cache do protocolo de resolução de endereço do sistema (ARP) para refletir as mudanças nos endereços de rede. É possível customizar esse script se for necessária uma ação posterior.

Reconectando-se ao cluster

Os clientes que estejam executando o daemon Clinfo podem se reconectar ao cluster rapidamente após um evento de cluster. Se tiver hardware diferente do IBM System p entre o cluster e os clientes, assegure-se de que é possível atualizar o cache ARP desses componentes de rede após um evento de cluster ocorrer.

Se você configurar o cluster para trocar endereços de hardware, bem como endereços IP, não é necessário preocupar-se sobre a atualização do cache ARP. Entretanto, esteja ciente de que esta opção causa um atraso mais longo.

Se estiver usando IPAT via aliases de IP, assegure-se de que todos os clientes suportem ARP gratuito TCP/IP.

Customizando o script clinfo.rc

Para clientes que estejam executando o daemon Clinfo, decida se customizará o script `/usr/es/sbin/cluster/etc/clinfo.rc` para fazer mais da atualização do cache ARP quando um evento do cluster ocorrer.

Clientes que não executam Clinfo

Em clientes que não estejam executando o daemon Clinfo, é possível ter que atualizar o cache do protocolo de resolução de endereço local (ARP) indiretamente realizando ping do cliente a partir do nó de cluster.

Nos nós do cluster, inclua o nome ou endereço de um host do cliente que deseja notificar para a variável `PING_CLIENT_LIST` no script `clinfo.rc`. Quando um evento de cluster ocorre, o script `clinfo.rc` executa o comando a seguir para cada host que é especificado na variável `PING_CLIENT_LIST`:

```
ping -c1 $host
```

Isso presume que o cliente esteja conectado diretamente a uma das redes do cluster.

Componentes de rede

Se você configurou a rede para que os clientes conectem às redes no outro lado de um roteador, ponte ou gateway em vez de às redes locais do cluster, assegure-se de que seja possível atualizar o cache ARP desses componentes de rede após um evento de cluster ocorrer.

Aplicativos e PowerHA SystemMirror

Este tópico aborda alguns dos principais problemas a se considerar ao tornar seus aplicativos altamente disponíveis em PowerHA SystemMirror.

Com o PowerHA SystemMirror, é possível configurar os clusters com aplicativos com multicamadas estabelecendo dependências entre grupos de recursos que contêm aplicativos diferentes. Esses tópicos descrevem dependências de grupo de recursos e como podem ajudar a manter os aplicativos dependentes altamente disponíveis.

Referências relacionadas:

“Planejamento de cluster inicial” na página 5

Esta seção descreve as etapas iniciais que você executa para planejar um cluster do PowerHA SystemMirror para tornar os aplicativos altamente disponíveis.

Visão geral dos aplicativos e PowerHA SystemMirror

Além de entender o hardware e o software, é necessário tornar um cluster altamente disponível, você deve estar ciente das considerações de disponibilidade do aplicativo ao planejar seu ambiente do

PowerHA SystemMirror. O objetivo do armazenamento em cluster é manter seus aplicativos importantes disponíveis apesar de qualquer ponto único de falha. Para atingir esse objetivo, considere os aspectos de um aplicativo que o tornam recuperável em PowerHA SystemMirror.

Há alguns requisitos que um aplicativo deve atender para recuperar-se bem em PowerHA SystemMirror. Algumas características necessárias, bem como um número de sugestões, são discutidas aqui. Elas são agrupadas de acordo com os principais pontos que se aplicam a todos os ambientes do PowerHA SystemMirror. Este tópico abrange as considerações de aplicativo a seguir:

- *Automação*. Assegure-se de que os aplicativos iniciem e parem sem intervenção do usuário
- *Dependências*. Saiba quais fatores fora do PowerHA SystemMirror afetam os aplicativos
- *Interferência*. Saber que os próprios aplicativos podem obstruir o funcionamento do PowerHA SystemMirror
- *Robustez*. Escolha aplicativos fortes e estáveis
- *Implementação*. Use scripts apropriados, locais de arquivos e planejamentos de cron.

Você deve incluir um monitor de aplicativo para detectar um problema com a inicialização do aplicativo. Um monitor de aplicativo no modo de monitoramento de inicialização verifica uma inicialização bem-sucedida do controlador de aplicativo no intervalo de estabilização especificado e sai após o período de estabilização terminar.

É possível iniciar os serviços de cluster PowerHA SystemMirror do nos nós sem parar seus aplicativos, selecionando uma opção de um painel do SMIT Serviços **PowerHA SystemMirror do > Iniciar serviços de cluster**. Ao iniciar, o PowerHA SystemMirror conta com os scripts de inicialização de aplicativo e monitores de aplicativos configurados para assegurar que o PowerHA SystemMirror esteja ciente do aplicativo em execução e não inicie uma segunda instância do aplicativo.

Da mesma forma, é possível parar os serviços de cluster do PowerHA SystemMirror e deixe os aplicativos em execução nos nós. Quando nó que tiver sido interrompido e colocado em um estado não gerenciado se une novamente ao cluster, presume-se que o estado dos recursos seja o mesmo a menos que um usuário inicie um comando do grupo de recursos do PowerHA SystemMirror para colocar o grupo de recursos em outro estado (por exemplo, on-line em um nó ativo).

Automação de aplicativo: minimizando a intervenção manual

Um requisito principal para que um aplicativo funcione com sucesso sob o PowerHA SystemMirror é que o aplicativo possa iniciar e parar sem qualquer intervenção manual.

Scripts de início de aplicativo

Crie um script de início que inicie o aplicativo. O script de início deve executar qualquer limpeza ou preparação necessária para assegurar a inicialização adequada do aplicativo e também para gerenciar adequadamente o número de instâncias do aplicativo que precisam ser iniciadas. Quando o controlador de aplicativo for incluído em um grupo de recursos. O PowerHA SystemMirror chama esse script para colocar o aplicativo on-line como parte do processamento do grupo de recursos. Como os daemons do cluster chamam o script de início, não há opção para interação. Além disso, mediante um failover do PowerHA SystemMirror, o processo de recuperação chama esse script para colocar o aplicativo on-line em um nó de espera. Isso permite uma recuperação totalmente automatizada e é por isso que qualquer limpeza ou preparação necessária deve ser incluída nesse script.

O PowerHA SystemMirror chama o script de início como o usuário raiz. Pode ser necessário alterar para um usuário diferente a fim de iniciar o aplicativo. O comando **su** pode realizar isso. Além disso, pode ser necessário executar o comando **nohup** em comandos iniciados no segundo plano e ter o potencial de serem encerrados mediante a saída do shell.

Por exemplo, um nó do cluster do PowerHA SystemMirror pode ser um cliente em um ambiente de Serviço de Informações de Rede (NIS). Se este for o caso e você precisar usar o comando **su** para alterar o ID do usuário, deverá haver uma rota para o servidor NIS todas as vezes. No caso de uma rota não existir e haver uma tentativa do comando **su**, o script do aplicativo será interrompido. É possível evitar esta situação ativando o nó do cluster do PowerHA SystemMirror para que seja um cliente NIS. Dessa forma, um nó do cluster tem a capacidade de acessar seus próprios arquivos de mapa do NIS para validar um ID do usuário.

O script de início também deve verificar a presença dos recursos ou processos necessários. Isso assegurará que um aplicativo possa ser iniciado com sucesso. Se os recursos necessários não estiverem disponíveis, uma mensagem poderá ser enviada à equipe de administração para corrigir isso e reiniciar o aplicativo.

Os scripts de início devem ser gravados de modo que determinem se uma instância do aplicativo já está em execução e não iniciar outra instância a menos que diversas instâncias sejam desejadas. Lembre-se de que o script de início pode ser executado após um nó primário ter falhado. Pode haver ações de recuperação necessárias no nó de backup a fim de reiniciar um aplicativo. Isso é comum em aplicativos de banco de dados. Novamente, a recuperação deve poder ser executada sem qualquer interação dos administradores.

Scripts de parada do aplicativo

O aspecto mais importante de um script de parada do aplicativo é que ele para completamente um aplicativo. Não fazê-lo pode impedir o PowerHA SystemMirror de concluir com sucesso um controle de recursos pelos nós de backup. Na parada, o script pode precisar abordar alguns dos interesses que o script de início aborda, como NIS e o comando **su**.

O script de parada do aplicativo deve usar uma abordagem em fases. A primeira fase deve ser uma tentativa de parar os serviços de cluster e colocar os grupos de recursos off-line. Se os processos recusarem o término, a segunda fase deverá ser usada para assegurar forçosamente que todo o processamento seja interrompido. Por fim, uma terceira fase pode usar um loop para repetir quaisquer etapas necessárias para assegurar que o aplicativo tenha sido encerrado completamente.

Assegure-se de que seu script de parada do aplicativo saia com o valor 0 quando o aplicativo tiver sido interrompido com sucesso. Especialmente, examine o que acontece se executar seu script de parada quando o aplicativo já estiver interrompido. Seu script deve sair com 0 neste caso também. Se seu script de parada sair com um valor diferente, isso informa ao PowerHA SystemMirror que o aplicativo ainda está em execução, embora possivelmente em um estado danificado. O evento **event_error** será executado e o cluster entrará em estado de erro. Esta verificação alerta os administradores que o cluster não está funcionando adequadamente.

Lembre-se de que o PowerHA SystemMirror permite 360 segundos por padrão para que os eventos concluam o processamento. É exibida uma mensagem que indica que o cluster esteve na reconfiguração por muito tempo até que o cluster concluísse sua reconfiguração e retornasse para um estado estável. Este aviso pode ser uma indicação que um script está interrompido e requer a intervenção manual. Se esta for uma possibilidade, você pode querer considerar parar um aplicativo manualmente antes de parar o PowerHA SystemMirror.

É possível alterar o período antes de o evento **config_too_long** ser chamado.

Scripts de início e parada de aplicativo e grupos de recursos dependentes

No PowerHA SystemMirror, o suporte para grupos de recursos dependentes permite configurar as opções a seguir:

- Três níveis de dependências entre os grupos de recursos, por exemplo, uma configuração na qual o nó A depende do nó B e o nó B depende do nó C. O PowerHA SystemMirror o impede de configurar dependências circulares.
- Um tipo de dependência na qual um grupo de recursos pai deve estar on-line em qualquer nó no cluster antes de um grupo de recursos filho (dependente) poder ser ativado em um nó.

Se dois aplicativos tiverem que ser executados no mesmo nó, ambos os aplicativos deverão residir no mesmo grupo de recursos.

Se um grupo de recursos filho contiver um aplicativo que dependa dos recursos no grupo de recursos pai e, mediante condições de failover, e se o grupo de recursos pai realizar failover para outro nó, o grupo de recursos filho será temporariamente interrompido e automaticamente reiniciado. Da mesma forma, se o grupo de recursos filho for simultâneo, o PowerHA SystemMirror o colocará off-line temporariamente em todos os nós e novamente on-line em todos os nós disponíveis. Se o failover do grupo de recursos pai não for bem-sucedido, os grupos de recursos pai e filho entrarão em estado ERROR.

Observe que quando o grupo de recursos filho é temporariamente interrompido e reiniciado, o aplicativo que pertence a ele também é interrompido e reiniciado. Portanto, para minimizar a chance de perda de dados durante o processo de parada e reinicialização do aplicativo, customize seus scripts de controlador de aplicativo para assegurar que quaisquer dados não confirmados sejam armazenados em um disco compartilhado temporariamente durante o processo de parada do aplicativo e lidos novamente no aplicativo durante o processo de reinicialização do aplicativo. É importante usar um disco compartilhado porque o aplicativo pode ser reiniciado em um nó que não aquele no qual ele foi parado.

Problemas de camada do aplicativo

Geralmente, os aplicativos têm uma arquitetura com multicamadas (por exemplo, uma camada de banco de dados, uma camada do aplicativo e uma camada de cliente). Considere todas as camadas de uma arquitetura se uma ou mais estiver altamente disponível pelo uso do PowerHA SystemMirror.

Por exemplo, se o banco de dados estiver altamente disponível e ocorrer um failover, considere se as ações devem ser realizadas em camadas mais altas a fim de retornar automaticamente o aplicativo para o serviço. Em caso positivo, pode ser necessário parar e reiniciar as camadas do aplicativo ou cliente. Isso pode ser facilitado de uma das duas formas. Uma forma é executar o comando `cli_on_node` nas camadas e outra é usar um comando de execução remota, como `rsh`, `rexec` ou `ssh`.

Nota: Certos métodos, como o uso de arquivos `/.rhosts`, apresentam um risco de segurança.

Usando grupos de recursos dependentes

Para configurar clusters complexos com aplicativos com multicamadas, é possível usar grupos de recursos dependentes pai-filho. Também é possível que queira considerar o uso de dependências de local.

Usando a API Clinfo

A API Clinfo é o daemon de informações do cluster. É possível gravar um programa usando a API Clinfo para executar em quaisquer camadas que parariam e reiniciariam um aplicativo após um failover ter sido concluído com sucesso. Neste sentido, a camada ou o aplicativo, torna-se ciente quanto ao cluster, respondendo a eventos que ocorrem no cluster.

Usando scripts pré-evento e pós-evento

Outra forma de abordar o problema das arquiteturas com multicamadas é usar script pré-evento e pós-evento em um evento de cluster. Esses scripts chamariam um comando de execução remota, como `rsh`, `rexec` ou `ssh`, para parar e reiniciar o aplicativo.

Conceitos relacionados:

“Aplicativos e PowerHA SystemMirror” na página 102

Este tópico aborda alguns dos principais problemas a se considerar ao tornar seus aplicativos altamente disponíveis em PowerHA SystemMirror.

Referências relacionadas:

“Gravando scripts efetivos” na página 108

Gravar scripts de início de aplicativo inteligente também pode ajudar a reduzir a probabilidade dos problemas ao colocar os aplicativos on-line.

“Planejando grupos de recursos” na página 59

Estes tópicos descrevem como planejar grupos de recursos em um cluster do PowerHA SystemMirror.

“Dependências de aplicativo”

Historicamente, para atingir o grupo de recursos e o sequenciamento de aplicativo, os administradores de sistema tinham que construir a lógica de recuperação do aplicativo em seus scripts de processamento pré e pós-evento. Cada cluster seria configurado com um script de pré-evento para todos os eventos do cluster e com um script de pós evento para todos os eventos do cluster.

Dependências de aplicativo

Historicamente, para atingir o grupo de recursos e o sequenciamento de aplicativo, os administradores de sistema tinham que construir a lógica de recuperação do aplicativo em seus scripts de processamento pré e pós-evento. Cada cluster seria configurado com um script de pré-evento para todos os eventos do cluster e com um script de pós evento para todos os eventos do cluster.

Tais scripts poderiam todos incluir instruções de caso. Por exemplo, se desejar tomar uma ação para um evento específico em um nó específico, é necessário editar esse caso individual, incluir o código para scripts de pré e pós-eventos e também assegurar que os scripts sejam os mesmos entre todos os nós.

Para resumir, mesmo que a lógica de tais scripts capture o comportamento desejado do cluster, os scripts poderão ser difíceis de customizar e até mais difíceis de manter posteriormente quando a configuração de cluster for alterada.

Se estiver usando scripts de pré e pós-evento ou outros métodos, como a ordenação do processamento do grupo de recursos para estabelecer as dependências entre os aplicativos suportados por seu cluster, esses métodos não serão mais necessários ou poderão ser simplificados de forma significativa. Em vez disso, é possível especificar dependências entre grupos de recursos em um cluster.

Nota: Em muitos casos, os aplicativos dependem mais dos dados e de um endereço IP. Para o sucesso de qualquer aplicativo sob o PowerHA SystemMirror, é importante saber que o aplicativo não deve depender para funcionar adequadamente. Este tópico descreve muitos dos problemas de dependência principais. Lembre-se de que essas dependências podem vir de fora do PowerHA SystemMirror e do ambiente de aplicativos. Elas podem ser produtos incompatíveis ou conflitos de recurso externo. Olhe além do aplicativo em si para ver problemas em potencial na empresa.

Dispositivos conectados localmente

Os dispositivos conectados localmente podem apresentar um problema claro de dependência. No caso de um failover, se esses dispositivos não estiverem conectados e acessíveis para o nó de espera, um aplicativo poderá falhar ao executar adequadamente. Podem incluir um dispositivo de CD-ROM, um dispositivo de fita ou uma juke box ótica. Considere se seu aplicativo depende de algum desses e se podem ser compartilhados entre os nós do cluster.

Código permanente

O código permanente de um aplicativo para um dispositivo específico em um local específico cria um problema de dependência em potencial. Por exemplo, o console é geralmente designado como `/dev/tty0`.

Embora este nome designado seja comum, não é garantido de nenhuma forma. Se seu aplicativo assumir o nome `/dev/tty0`, assegure-se de que todos os nós de espera possíveis tenham a mesma configuração.

Dependências de nome do host

Alguns aplicativos são gravados para serem dependentes do nome do host AIX. Eles emitem um comando para validar licenças ou sistemas de arquivos de nomes. O nome do host não é um rótulo de endereço IP. O nome do host é específico a um nó e não passou por failover pelo PowerHA SystemMirror. O nome do host não pode ser alterado. Qualquer aplicativo que alterar o nome do host não é suportado pelo PowerHA SystemMirror.

Licenciamento de software

Outro problema possível é o licenciamento de software. O software pode ser licenciado para um ID de CPU específico. Se este for o caso com seu aplicativo, um failover do software não será reinicializado com sucesso. É possível que seja capaz de evitar esse problema tendo uma cópia do software em todos os nós do cluster. Saiba se seu aplicativo usa o software que é licenciado a um ID de CPU específico.

Referências relacionadas:

“Considerações de planejamento para aplicativos com multicamadas” na página 15

As configurações de negócio que usam aplicativos com multicamadas podem usar grupos de recursos dependentes pais e filhos. Por exemplo, o banco de dados deve estar on-line antes do controlador de aplicativo. Neste 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 teria de ser desativado e ativado em qualquer nó no cluster.

Interferência de aplicativos

Às vezes, um aplicativo ou um ambiente de aplicativos pode interferir com o funcionamento adequado do PowerHA SystemMirror. Um aplicativo pode ser executado adequadamente em ambos os nós primários e de espera. Entretanto, quando o PowerHA SystemMirror é iniciado, um conflito com o aplicativo ou ambiente poderia surgir impedindo o PowerHA SystemMirror de funcionar com sucesso.

Produtos que manipulam rotas de rede

Além disso, produtos que manipulam rotas de rede podem impedir o PowerHA SystemMirror de funcionarem como foi projetado. Esses produtos podem localizar um caminho secundário por meio de uma rede que teve uma falha inicial. Esse roteamento pode impedir o PowerHA SystemMirror de diagnosticar adequadamente uma falha e tomar ações de recuperação apropriadas.

Referências relacionadas:

“Planejamento de cluster inicial” na página 5

Esta seção descreve as etapas iniciais que você executa para planejar um cluster do PowerHA SystemMirror para tornar os aplicativos altamente disponíveis.

Robustez do aplicativo

De importância primordial ao sucesso de qualquer aplicativo é o funcionamento, ou robustez, do aplicativo. Se o aplicativo estiver instável ou travando intermitentemente, resolva esses problemas antes de colocá-lo em um ambiente de alta disponibilidade.

Além da estabilidade básica, um aplicativo no PowerHA SystemMirror deve atender a outras características de robustez.

Início bem-sucedido após a falha no hardware

Um bom candidato a aplicativo para o PowerHA SystemMirror deve poder ser reiniciado com sucesso após uma falha no hardware. Execute um teste em um aplicativo antes de gerenciá-lo com o PowerHA

SystemMirror. Execute o aplicativo sob uma carga pesada e faça o nó falhar. O que é preciso para recuperar-se após o nó ficar novamente on-line? Essa recuperação pode ser completamente automatizada? Se não, o aplicativo poderá não ser um bom candidato para alta disponibilidade.

Sobrevivência de perda de memória real

Os aplicativos devem salvar regularmente no disco quaisquer informações necessárias para a reinicialização. Se ocorrer uma falha, o aplicativo poderá escolher de onde estava antes, em vez de iniciar do zero completamente.

Estratégias de implementação de aplicativo

Há vários aspectos de um aplicativo a se considerar conforme você planeja sua implementação em PowerHA SystemMirror.

Considere as características, como tempo para iniciar, tempo para reinicialização após falha e tempo para parar. Suas decisões em várias áreas, como composição de script, armazenamento de arquivo, arquivo `/etc/inittab` e problemas de planejamento de cron, podem melhorar a probabilidade de implementação de aplicativo bem-sucedida.

Gravando scripts efetivos

Gravar scripts de início de aplicativo inteligente também pode ajudar a reduzir a probabilidade dos problemas ao colocar os aplicativos on-line.

Uma boa prática para scripts de início é verificar condições obrigatórias antes de iniciar um aplicativo. Ela pode incluir o acesso a um sistema de arquivos, espaço de paginação adequado e espaço de sistema de arquivos livre. O script de início deve sair e executar um comando para notificar os administradores de sistema se os requisitos não forem atendidos.

Ao iniciar um banco de dados, é importante considerar se há diversas instâncias no mesmo cluster. Se esse for o caso, inicie apenas as instâncias aplicáveis para cada nó. Certos comandos de inicialização do banco de dados leem um arquivo de configuração e iniciam todos os bancos de dados conhecidos ao mesmo tempo. Este comportamento pode não ser uma configuração desejada para todos os ambientes.

Tenha cuidado para não eliminar quaisquer processos do PowerHA SystemMirror como parte do seu script. Se estiver usando a saída do comando `ps` e usando um `grep` para procurar por um certo padrão, assegure-se de que o padrão não atenda a nenhum dos processos do PowerHA SystemMirror ou Reliable Scalable Cluster Technology (RSCT).

Considerando locais de armazenamento de arquivo

Pense sobre onde os arquivos de configuração residem. Eles poderiam estar em um disco compartilhado e, portanto, possivelmente acessados por qualquer nó que tenha o grupo de volumes variado ou nos discos internos de cada nó. Isso é verdadeiro para todos os aspectos de um aplicativo. Certos arquivos devem estar em unidades compartilhadas. Esses arquivos incluem dados, logs e qualquer coisa que poderia ser atualizada pela execução do aplicativo. Arquivos como arquivos de configuração ou binários de aplicativos poderiam residir em qualquer local.

Há vantagens e desvantagens para armazenar arquivos opcionais em qualquer local. Ter arquivos armazenados nos discos internos de cada nó implica que você tem diversas cópias, e possivelmente várias licenças, do aplicativo. Isso poderia requerer o custo adicional bem como a manutenção em manter desses arquivos sincronizados. Entretanto, no caso de um aplicativo precisar ser atualizado, todo o cluster não precisará ser retirado de produção. Um nó poderia ser atualizado enquanto o outro permanece na produção. A melhor solução é a que funciona melhor para um determinado ambiente.

Considerando problemas de `/etc/inittab` e tabela cron

Pense também sobre os aplicativos ou recursos necessários por um aplicativo que começa do arquivo `/etc/inittab` ou da tabela cron.

O arquivo `inittab` inicia aplicativos ao iniciar o sistema. Se os recursos de cluster forem necessários para que um aplicativo funcione, eles não se tornarão disponíveis até após o PowerHA SystemMirror ser iniciado. É melhor usar o recurso do controlador de aplicativo do PowerHA SystemMirror que permite que o aplicativo seja um recurso iniciado apenas após todos os recursos dependentes ficarem online.

Nota: É importante que as configurações a seguir estejam corretas no arquivo `/etc/inittab`:

```
hacmp:2:once:/usr/es/sbin/cluster/etc/rc.init
```

- As entradas `clinit` e `pst_clinit` devem ser as últimas entradas do nível de execução "2".
- A entrada `clinit` deve continuar a entrada `pst_clinit`.

Uma entrada incorreta para essas impede que o PowerHA SystemMirror inicie.

Na tabela `cron`, as tarefas são iniciadas de acordo com o planejamento definido na tabela e a configuração da data em um nó. Essas informações são mantidas em discos internos e, portanto, não podem ser compartilhados por um nó de espera. Sincronize essas tabelas `cron` para que um nó de espera possa executar a ação necessária no momento apropriado. Além disso, assegure-se que a mesma data esteja definida no nó primário e em qualquer um de seus nós de espera.

Exemplos: banco de dados Oracle e SAP R/3

Aqui estão dois exemplos que ilustram problemas a se considerar a fim de fazer com que os aplicativos Oracle Database e SAP R/3 funcionem bem no PowerHA SystemMirror.

Exemplo 1: banco de dados Oracle

O banco de dados Oracle, como muitos bancos de dados, funciona bem no PowerHA SystemMirror. É um aplicativo robusto que manipula falhas bem. Ele pode recuperar transações não confirmadas após um failover e retornar para o serviço de forma oportuna. Entretanto, há algumas coisas para se manter em mente ao usar o banco de dados Oracle no PowerHA SystemMirror.

Iniciando o Oracle

O Oracle deve ser iniciado pelo ID do usuário Oracle. Assim, o script de início deve conter o código a seguir `su - oracleuser`. O traço (-) é importante desde que o comando `su` precisa assumir todas as características do usuário Oracle e residir no diretório inicial dos usuários Oracle. O comando pareceria com algo assim:

```
su - oracleuser -c /apps/oracle/startup/dbstart
```

O comando `dbstart` e comando `dbshut` leem o arquivo `/etc/oratabs` para instruções sob as quais as instâncias de banco de dados são conhecidas e devem ser iniciadas. Em certos casos, é inapropriado iniciar todas as instâncias, porque elas podem ser possuídas por outro nó. Isso seria o caso no controle mútuo de duas instâncias Oracle. O arquivo `oratabs` geralmente reside no disco interno e não pode ser compartilhado. Se apropriado, considere outras formas de iniciar diferentes instâncias do Oracle.

Parando o Oracle

A parada do Oracle é um processo de interesse especial. Há várias formas diferentes de assegurar que o Oracle tenha sido completamente interrompido. A sequência sugerida é esta: primeiro, implemente um encerramento normal; segundo, chame um encerramento imediato, que é mais um método forçado; finalmente, crie um loop para verificar a tabela de processo para assegurar que todos os processos Oracle tenham saída.

Armazenamento de arquivo Oracle

O banco de dados de produto Oracle contém vários arquivos, bem como dados. É necessário que os dados e os logs de recuperação sejam armazenados no disco compartilhado para que ambos os nós

possam ter acesso às informações. Entretanto, os binários e os arquivos de configuração Oracle poderiam residir nos discos interno ou compartilhado. Considere qual solução é melhor para o seu ambiente.

Exemplo 2: SAP R/3, um aplicativo com multicamadas

O SAP R/3 é um exemplo de um aplicativo de terceira camada. Ele possui uma camada de banco de dados, uma camada do aplicativo e uma camada de cliente. Mais frequentemente, é a camada de banco de dados que é feita altamente disponível. Neste caso, quando um fallover ocorre e o banco de dados é reiniciado, é necessário parar e reiniciar a camada do aplicativo SAP. É possível fazer isso de uma das duas formas:

- Usar um comando de execução remota, como **rsh**, **rexec** ou **ssh**

Nota: Certos métodos, como o uso de arquivos `~/.rhosts`, colocam um risco à segurança.

- Tornando o cluster de nós da camada do aplicativo cliente.

Usando um comando de execução remota

A primeira forma de parar e iniciar a camada do aplicativo SAP é criar um script que desempenha a execução de comando remoto nos nós de aplicativos. A camada do aplicativo do SAP é interrompida e depois reiniciada. Isso é feito para cada nó na camada do aplicativo. Usar um comando de execução remota requer um método de permissão do acesso do nó de banco de dados ao nó do aplicativo.

Nota: Certos métodos, como o uso de arquivos `~/.rhosts`, colocam um risco à segurança.

Tornando o cluster dos nós de camada do aplicativo cliente

Um segundo método para parar e iniciar a camada do aplicativo é tornar o cluster de nós de camada do aplicativo cliente. Isso significa que os nós de camada do aplicativo estão clientes do banco de dados do cluster e sabem quando um fallover ocorre. É possível implementar isso tornando os nós da camada do aplicativo servidores ou clientes do PowerHA SystemMirror. Se o nó do aplicativo for um servidor, ele executará os mesmos eventos de cluster que os nós de banco de dados para indicar uma falha. Scripts de pré e pós-evento poderiam ser gravados para parar e reiniciar a camada do aplicativo SAP. Se o nó do aplicativo for um cliente PowerHA SystemMirror, ele será notificado do fallover do banco de dados usando o SNMP por meio do daemon de informações do cluster (Clnfo). Um programa poderia ser gravado usando a API Clnfo para parar e reiniciar a camada do aplicativo SAP.

Informações relacionadas:

Programando aplicativos cliente para a API Clnfo

Avisos

Estas informações foram desenvolvidas para produtos e serviços oferecidos nos Estados Unidos.

É possível que a IBM não ofereça os produtos, serviços ou recursos discutidos nesta publicação 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 poderá ser utilizado em substituição a este produto, programa ou serviço. 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 descritos neste documento. O fornecimento deste documento não garante ao Cliente nenhum direito 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
CEP 22290-240
Rio de Janeiro, RJ
Brasil

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
2-31 Roppongi 3-chome
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 MERCADO OU ADEQUAÇÃO A UM DETERMINADO PROPÓSITO. Alguns países não permitem a exclusão de garantias expressas ou implícitas em certas transações; portanto, esta instrução pode não se aplicar ao Cliente.

Estas informações podem conter imprecisões técnicas ou erros tipográficos. São feitas alterações periódicas nas informações aqui contidas; tais alterações serão incorporadas em futuras edições desta publicação. A IBM pode, a qualquer momento, aperfeiçoar e/ou alterar os produtos e/ou programas descritos nesta publicação, sem aviso prévio.

As referências nestas informações a websites que não sejam da IBM são fornecidas apenas por conveniência e não representam de forma alguma um endosso a esses Web sites. Os materiais contidos nesses websites não fazem parte dos materiais para este produto IBM e o uso desses websites é de total 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:

Gerência de Relações Comerciais e Industriais da IBM Brasil
Av. Pasteur, 138-146
Botafogo
Rio de Janeiro, RJ
CEP 22290-240

Tais informações podem estar disponíveis, sujeitas a termos e condições apropriadas, incluindo em alguns casos o pagamento de uma taxa.

O programa licenciado descrito nesta publicação e todo o material licenciado disponível são fornecidos pela IBM sob os termos do Contrato com o Cliente IBM, do Contrato Internacional de Licença do Programa IBM ou de qualquer outro contrato equivalente.

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 em nível de desenvolvimento e não há garantia de que estas medidas serão iguais em sistemas geralmente disponíveis. Além disso, algumas medidas podem ter sido estimadas por extrapolação. Os resultados reais podem 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 em relação à direção ou intenção futuras da IBM estão sujeitas a mudanças ou retiradas, sem aviso prévio, e representam apenas metas e objetivos.

Todos os preços da IBM mostrados são preços de varejo da IBM sugeridos, são atuais e estão sujeitos a mudanças sem aviso prévio. Os preços dos revendedores podem variar.

Estas informações foram projetadas apenas para o propósito de planejamento. As informações aqui contidas estão sujeitas a mudança antes da disponibilização dos produtos.

Essas informações contêm exemplos de dados e relatórios usados 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 estes nomes são fictícios e qualquer semelhança com nomes e endereços usados por uma empresa real é mera coincidência.

LICENÇA DE COPYRIGHT:

Estas informações contêm programas de aplicativos de exemplo no idioma de origem, ilustrando as técnicas de programação em diversas plataformas operacionais. O Cliente pode copiar, modificar e distribuir esses programas de amostra de qualquer maneira, sem pagamento à IBM, para o propósito de desenvolvimento, utilização, marketing ou distribuição dos programas aplicativos conforme a interface de programação de aplicativos da plataforma operacional para a qual os programas de amostra são gravados. Esses exemplos não foram testados completamente em todas as condições. Portanto, a IBM não pode garantir ou subentender confiabilidade, capacidade de manutenção ou funcionamento desses programas. Os programas de amostra são fornecidos "NO ESTADO EM QUE SE ENCONTRAM", sem garantia de qualquer tipo. A IBM não poderá ser responsabilizada por nenhum dano decorrente do uso dos programas de amostra.

Todas as cópias ou qualquer parte desses programas de amostra, ou qualquer trabalho derivado, devem incluir um aviso de copyright como este:

Partes deste código são derivadas dos Programas de Amostras da 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.

Índice Remissivo

A

- acesso a disco 46
 - simultâneo aprimorado 46
- adaptadores virtuais 31
- AIX Workload Manager
 - visão geral 80
- Aliases de IP 18
- amostra
 - banco de dados Oracle e SAP R/3 109
 - IBM DS4000 Storage Server 37
- aplicativo 11, 102
 - com multicamadas 15
 - dependências 106
 - gravando scripts 108
 - interferência 107
 - visão geral 103
- ativação 49
 - forçada 51

C

- cliente 101
 - clinfo 102
 - componentes de rede 102
 - não executando clinfo 102
- clinfo 102
 - não executando 102
- clRGmove
 - grupos de recursos 70
- cluster
 - diagrama 16
 - particionamento 19
- CoD 12
- componente LVM 40
- componente LVM compartilhado 40
- concurrent
 - grupo de recursos 60
- configuração de disco
 - incluindo 37
- controlador de aplicativo 13
- Controle de endereço IP
 - Aliases de IP 25
- controle rápido de disco 47

D

- disco
 - adaptadores 37
 - armazenamento em disco não compartilhado 35
 - cabos 37
 - considerações de fonte de alimentação 35
 - IBM DS4000 Storage Server
 - amostra 37
 - instalação de disco compartilhado 36
 - SCSI virtual 36
 - tecnologia de disco compartilhado 34
 - visão geral 33
- disco compartilhado 33
- disco SCSI compartilhado
 - instalação 36

- dispositivo de fita 33
- DNS 29

E

- espelhamento
 - log de diário 44
 - partição física 42
 - site 45
- espelhamento de LVM 42
- espelhamento de site de divisão
 - planejamento 44
- espelhamento de site de divisão do LVM
 - planejamento 44
- estado
 - event 89
- Ethernet virtual 20
- event
 - definidos pelo usuário 99
 - grupo de recursos 90
 - interface de rede 88
 - nó 83
 - notificação 93
 - rede 87
 - resumo 101
 - scripts pré-evento e pós-evento 93
 - site 83
 - status em todo o cluster 89
 - visão geral 82
- evento de cluster
 - Veja* event
- exemplo
 - conexão de rede 20

F

- firewall de VPN 30

G

- grupo de recursos 59
 - attributes 62
 - event 90
 - movendo 70
 - movendo usando clRGmove 70
 - networks 71
 - ordem de processamento 71
 - políticas 61
 - recursos replicados 79
 - sites 73
 - tipos 60
 - visão geral 59
- grupo de volume 41

H

- hacmp.out 101

I

- IBM DS4000 Storage Server
 - amostra 37
- incluindo
 - configuração de disco 37
 - topologia de rede 33
- interface de rede
 - event 88
- IPv6
 - Planejando 30
- isolamento do nó 19

L

- log de diário
 - espelhamento 44

M

- monitoramento
 - clusters 29
- monitoramento de aplicativo 14
- movendo
 - grupos de recursos 70

N

- não simultâneo
 - grupo de recursos 60
- netmon.cf 31
- NFS 52
- NIS 29
- nó 5
 - eventos 83
 - Eventos node_down 84
 - eventos node_up 84

O

- Oracle
 - planejando redes 31
- ordem de processamento
 - grupos de recursos 71
- orientações 2

P

- partição física
 - espelhamento 42
- planejamento
 - espelhamento de site de divisão do LVM 44
 - planejamento de cluster inicial 5
- Planejando
 - IPv6 30
- processo de planejamento
 - visão geral 3
- pulsção 21

Q

- quorum 49

R

- recurso replicado 79
- rede
 - Aliases de IP 18
 - cliente 102
 - conectividade 17
 - conexões 18
 - Controle de endereço IP por meio de aliases de IP 25
 - DNS 29
 - Ethernet virtual 20
 - event 87
 - evitando conflitos 33
 - exemplo 20
 - firewall de VPN 30
 - grupos de recursos 71
 - incluindo topologia 33
 - monitorando clusters 29
 - NIS 29
 - Oracle 31
 - particionamento de cluster 19
 - pulsção 21
 - redes alternadas 20
 - rótulos de IP 19
 - topologia 21
- redes virtuais 31
- Rótulo de IP 19

S

- script
 - gravando 108
 - pré-evento e pós-evento 93
- SCSI virtual 36
- security 10
- sistema de arquivos 42
- site 9
 - espelhamento 45
 - eventos 83
 - grupos de recursos 73

T

- topologia
 - rede 21

U

- unidade de fita 37
- upgrade de capacidade sob demanda
 - Veja CoD

V

- visão geral
 - AIX Workload Manager 80
 - aplicativo 103
 - disco 33
 - eventos de cluster 82
 - grupos de recursos 59
 - processo de planejamento 3
- volume físico 40
- volume lógico 41

W

WLM

Veja AIX Workload Manager



Impresso no Brasil