

Guia do Administrador do HPCC System

Equipe de documentação de Boca Raton



Guia do Administrador do HPCC Systems®

Equipe de documentação de Boca Raton

Copyright © 2023 HPCC Systems®. All rights reserved

Sua opinião e comentários sobre este documento são muito bem-vindos e podem ser enviados por e-mail para <docfeedback@hpccsystems.com>

Inclua a frase **Feedback sobre documentação** na linha de assunto e indique o nome do documento, o número das páginas e número da versão atual no corpo da mensagem.

LexisNexis e o logotipo Knowledge Burst são marcas comerciais registradas da Reed Elsevier Properties Inc., usadas sob licença.

HPCC Systems® é uma marca registrada da LexisNexis Risk Data Management Inc.

Os demais produtos, logotipos e serviços podem ser marcas comerciais ou registradas de suas respectivas empresas.

Todos os nomes e dados de exemplo usados neste manual são fictícios. Qualquer semelhança com pessoas reais, vivas ou mortas, é mera coincidência.

2023 Version 8.12.8-1

Introdução a Administração do HPCC Systems®	4
Introdução	4
Visão Geral da Arquitetura	5
Requerimento de Hardware e Software	11
Switch de Rede	11
Balanceamento de Carga	12
Dimensionamento do Sistema	13
Software-Nós	14
Hardware e Componentes	16
Hardware Thor	17
Configurações de Hardware do Roxie	18
Configurações de Hardware do Dali e Sasha	19
Outros Componentes HPCC	20
Manutenção da Rotina	21
Manipulação dos Dados	22
Backup dos Dados	22
Arquivo de Log	25
Preflight	28
Preflight do Servidores do Sistema	29
Preflight de Clusters de Destino	34
Preflight Thor	38
Preflight do cluster Roxie	41
Configuração e Gerenciamento do Sistema	44
Utilizando o Gerenciador de Configurações	47
Environment.conf	54
Configurando HPCC para Autenticação	57
Manutenção de Segurança do Usuário	72
Workunits e Active Directory	114
Utilizando o Cassandra para armazenamento de Workunit.	115
Redefinindo nós em um Cluster Thor	123
Melhores práticas:	124
Redundância de Cluster	124
Alto Disponibilidade	126
Considerações sobre Melhores Práticas:	128
Planejamento de Capacidade	132
Dimensionamento da Amostra	134
Recursos do Sistema	136
Recursos do HPCC	136
Recursos Adicionais	137

Introdução a Administração do HPCC Systems®

Introdução

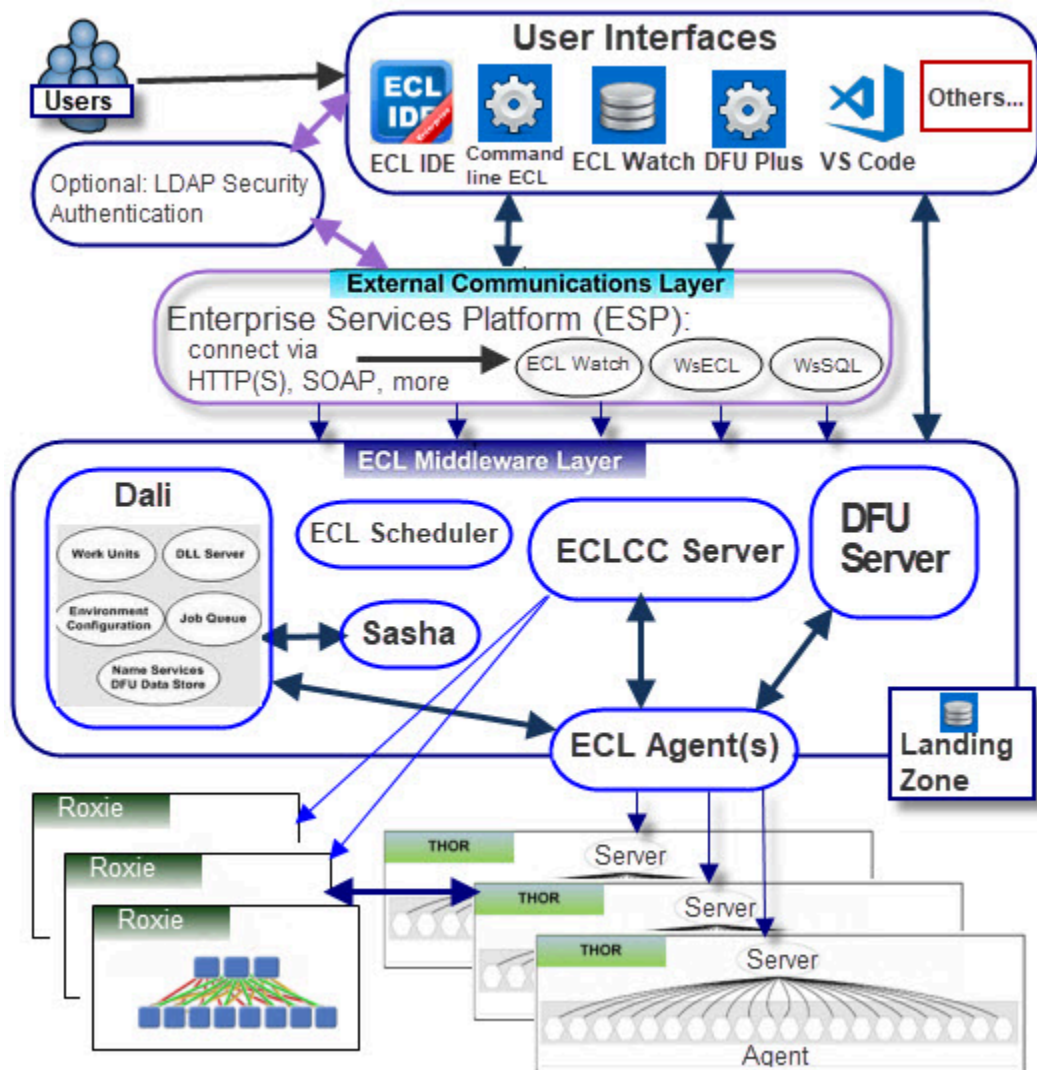
O HPCC (High Performance Computing Cluste) é uma plataforma de computação e de processamento paralelo massivo que soluciona problemas de big data.

O HPCC armazena e processa grandes quantidades de dados, processando bilhões de registros por segundo usando tecnologia de processamento paralelo massivo. Grandes quantidades de dados entre diferentes fontes de informações podem ser acessadas, analisadas e processadas em questão de segundos. O HPCC funciona como um ambiente de processamento e de armazenamento de dados distribuído capaz de analisar terabytes de informações.

Visão Geral da Arquitetura

A plataforma HPCC Systems® é composta pelos seguintes componentes: Thor, Roxie, ESP Server, Dali, Sasha, DFU Server e ECLCC Server. A segurança em LDAP está disponível como opcional.

Figure 1. Diagrama de arquitetura do HPCC



O carregamento de dados é controlado através do servidor de utilitário de arquivos distribuídos (DFU).

Os dados normalmente chegam na zona de entrada de arquivo (por exemplo, por FTP). A transferência de arquivos (entre componentes) é iniciada pelo DFU. Os dados são copiados da zona de entrada de arquivo e distribuídos (repasados aos nós) até a Refinaria de Dados (Thor) pelo código ECL. Os dados podem ser ainda processados mais a fundo por ETL (processo de extração, transformação e de carregamento) na refinaria.

Um único arquivo físico é distribuído em diversos arquivos físicos entre os nós de um cluster. O agregado dos arquivos físicos cria um arquivo lógico que é endereçado pelo código ECL.

Figure 2. Processamento de dados



O processo de recuperação de dados (chamado de despraying, ou consolidação de dados dos nós) retorna o arquivo para a zona de entrada de arquivo.

Clusters

O ambiente do HPCC Systems contém clusters que você define e usa de acordo com as suas necessidades. Os tipos de clusters usados pelo HPCC são:

Thor

Refinaria de Dados (Thor) -- Usado para processar cada um dos bilhões de registros a fim de criar bilhões de registros "aprimorados". O ECL Agent (hThor) também é usado para processar tarefas simples que usariam o cluster Thor de forma ineficiente.

Roxie

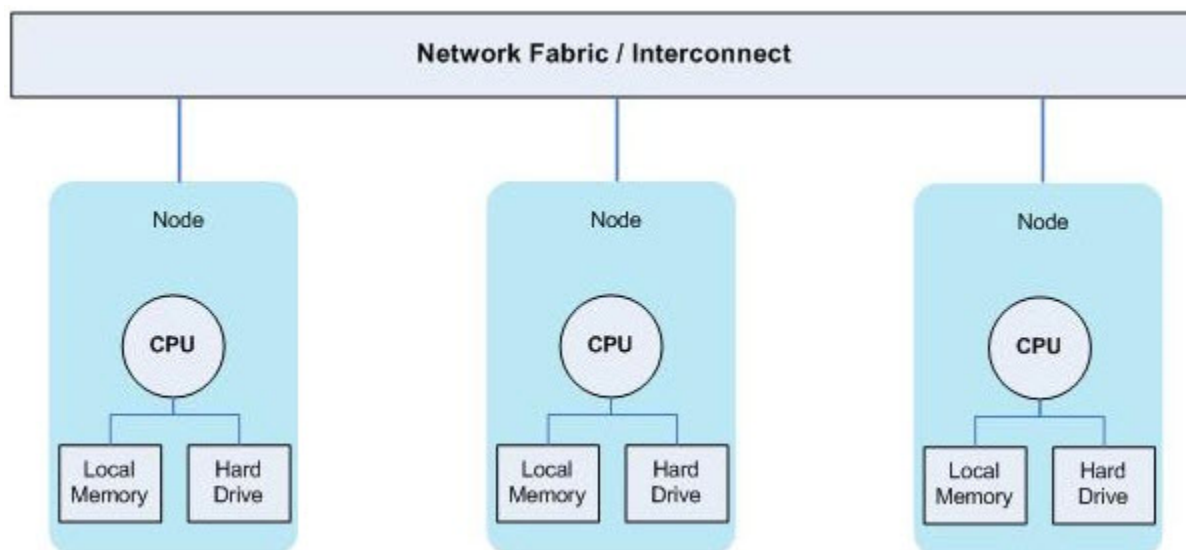
Motor de entrega rápida de dados (Roxie) -- Usado para pesquisas rápidas por um determinado registro ou conjunto de registros.

As consultas são normalmente compiladas e publicadas no ECL Watch. Os dados são transferidos em paralelo dos nós Thor até os nós de recebimento Roxie. A utilização da largura de banda paralela melhora a velocidade com que se coloca novos dados para operar.

ECL Agent

A principal função do ECL Agent é enviar o job para execução no cluster adequado. O ECL Agent pode atuar como um cluster de nó único. Isso é denominado surgimento de um cluster hThor. O hThor é usado para processar tarefas simples que usariam o Thor de forma ineficiente. Para tarefas simples, o ECL Agent determina e realiza a execução sozinho, agindo como um cluster hThor.

Figure 3. Clusters



Servidores do Sistema

Os servidores do sistema são componentes essenciais de middleware de um sistema HPCC. Eles são usados para controlar a comunicação entre componentes e o fluxo de trabalho.

Dali

Dali também é conhecido como o armazenamento de dados do sistema. Ele gerencia registros de tarefas, diretório de arquivo lógico e serviços de objetos compartilhados. Ele mantém as filas de mensagens que cuidam da execução e do agendamento de tarefas.

Dali também realiza a gestão de sessões. Ele rastreia todas as sessões ativas de clientes do Dali registradas no ambiente para que seja possível listar todos os clientes e suas funções. (*consulte `dalidiag -clients`*)

Outra tarefa que o Dali realiza é agir como o gerenciador de bloqueio. O HPCC usa o gerenciador de bloqueio do Dali para controlar bloqueios exclusivos e compartilhados com metadados.

Sasha

O servidor Sasha é um servidor de "organização" complementar ao servidor Dali. O Sasha opera de forma independente, ainda que em conjunto com o Dali. A função principal do Sasha é reduzir a carga no servidor Dali. Sempre que possível, o Sasha reduz a utilização de recursos no Dali. Um aspecto muito importante do Sasha é a coalescência ao salvar o armazenamento em memória para uma nova edição de armazenamento.

O Sasha arquiva workunits tarefas (incluindo DFU workunits) que são então armazenadas em pastas em um disco.

Sasha também realiza organização de rotina, como a remoção de workunits em cache e arquivos de recuperação do DFU.

Sasha também pode executar o XREF para cruzar referências de arquivos físicos com metadados lógicos a fim de determinar se há arquivos perdidos/encontrados/órfãos. Ele então apresenta opções (pelo EclWatch) para sua recuperação ou exclusão.

Sasha é o componente responsável pela remoção de arquivos expirados quando os critérios são atendidos. A opção EXPIRE no OUTPUT do ECL ou PERSIST define essa condição.

Servidor DFU

O servidor DFU controla as operações de distribuição de dados aos nós (spraying) e de consolidação de dados aos nós (despraying) usadas para transferir dados para dentro e fora do Thor.

Os serviços DFU estão disponíveis a partir de:

- Bibliotecas padrão no código ECL.
- Interfaces do Client Eclipse, ECL Playground, ECL IDE, e a interface de linha de comando ECL.
- Interface de linha de comando DFU Plus.

ECLCC Server

ECLCC Server é o compilador que converte o código ECL. Ao enviar o código ECL, o ECLCC Server gera o C++ otimizado, que é então compilado e executado. O ECLCC Server controla todo o processo de compilação.

Ao enviar workunit para execução no Thor, elas são primeiramente convertidas em um código executável pelo ECLCC Server.

Ao enviar uma workunit para o Roxie, o código é compilado e posteriormente publicado no cluster Roxie, onde está disponível para ser executado várias vezes.

O ECLCC Server também é usado quando o ECL IDE solicita uma verificação de sintaxe.

O ECLCC Server também usa uma fila para converter uma workunit de cada vez; no entanto, é possível implementar os servidores ECLCC no sistema para aumentar a produtividade e eles balancearão a carga automaticamente conforme necessário.

ECL Agent

O ECL Agent (hThor) é um processo de nó único para execução de Consultas ECL simples.

O ECL Agent é um mecanismo de execução que processa tarefas ao enviá-las para o cluster adequado. Os processos do ECL Agent são gerados sob demanda quando uma tarefa é enviada.

ESP Server

O ESP (Enterprise Service Platform) Server é o servidor de comunicação de intercomponentes. O ESP Server é um framework que permite que múltiplos serviços sejam "plugados" para oferecer vários tipos de funcionalidades às aplicações dos clientes por meio de diversos protocolos.

Entre os exemplos de serviços que são plugados no ESP, estão:

- **WsECL:** Interface para consultas de ECL publicadas em um cluster Roxie, Thor ou hThor.
- **ECL Watch** Uma interface de gestão de arquivos, monitoramento e execução de consultas baseada na Web. Ela pode ser acessada através do ECL IDE ou de um navegador de Internet. Consulte *Como usar o ECL Watch*.

O servidor ESP suporta formatos XML e JSON.

LDAP

É possível incorporar um servidor de Lightweight Directory Access Protocol (protocolo de acesso aos diretórios leves) LDAP para operar junto ao Dali a fim de reforçar as restrições de segurança para escopo de arquivos, escopos de workunit e acesso a recursos.

Quando o LDAP é configurado, é necessário se autenticar ao acessar o ECL Watch, WsECL, ECL IDE ou quaisquer outras ferramentas de clientes. As credenciais são então usadas para autenticar quaisquer solicitações dessas ferramentas.

Interfaces Client

As seguintes interfaces do client estão disponíveis para interação com a plataforma HPCC.

ECL IDE

ECL IDE é uma GUI completa, com todos os recursos, que oferece acesso ao seu código ECL para desenvolvimento de ECL. O IDE ECL usa diversos serviços de ESP através do SOAP.

O ECL IDE oferece acesso a definições de ECL para desenvolver suas consultas. Essas definições são criadas ao programar uma expressão que define como certos cálculos ou derivações de conjunto de registros devem ser feitos. Depois de definidas, elas podem ser usadas nas próximas definições de ECL.

ECL Watch

ECL Watch é uma interface de gestão de arquivos, monitoramento e execução de consultas baseada na Web. Ela pode ser acessada através do IDE ECL, Eclipse ou de um navegador de Internet. O ECL Watch permite que você veja informações e processe tarefas. Ele também permite que você monitore a atividade de cluster e realize outras tarefas administrativas.

Usando o ECL Watch, é possível:

- Navegar por workunits enviadas anteriormente (WU). Ver uma representação visual (gráficos) completa do fluxo de dados no WU, com estatísticas que são atualizadas a medida em que o trabalho progride.
- Pesquisar por arquivos e ver informações que incluem número de registros e layouts ou registros de amostra.
- Visualizar o status de todos os servidores do sistema.
- Visualizar arquivos de log.
- Adicionar usuários ou grupos e modificar permissões.

Consulte o manual *Utilizando ECL Watch* para obter mais detalhes.

Ferramentas de linha de comando:

Ferramentas de linha de comando: **ECL**, **DFU Plus** e **ECL Plus** oferecem acesso de linha de comando às funcionalidades fornecidas pelas páginas da Web do ECL Watch. Elas funcionam pela comunicação com o serviço ESP correspondente via SOAP.

Consulte o manual *Ferramentas de cliente* para obter mais detalhes.

Requerimento de Hardware e Software

Este capítulo descreve alguns dos requisitos de hardware e software para executar o HPCC System. O HPCC foi projetado para ser executado em hardware padrão, tornando o desenvolvimento e a manutenção de clusters de grande escala (petabytes) economicamente viáveis. Ao planejar o hardware do seu cluster, é necessário colocar na balança diversas considerações específicas às suas necessidades.

Esta seção oferece alguns insights sobre o hardware e a infraestrutura nos quais o HPCC opera de maneira adequada. Este não é um conjunto amplo e completo de instruções, nem uma obrigação sobre qual hardware você precisa usar. Considere este como um guia para usar quando for implementar ou dimensionar seu HPCC System. As sugestões devem ser levadas em conta de acordo com as suas necessidades empresariais específicas.

Switch de Rede

A switch de rede é um componente importante do HPCC System.

Requerimento do Switch

- Número suficiente de portas para permitir que todos os nós sejam conectados diretamente a ele;
- Suporte à IGMP v.2
- Suporte à monitoração IGMP

Seu HPCC System supostamente apresentará um melhor desempenho quando cada nó estiver conectado diretamente a um único switch. Você precisa fornecer uma porta para cada nó em um único switch para otimizar o desempenho do sistema. O tamanho do seu switch deve corresponder ao tamanho do seu sistema. É importante assegurar que a chave utilizada tenha capacidade suficiente para que cada nó seja conectado à sua própria porta.

Recursos adicionais recomendados para o Switch

- Velocidade em Gigabit
- Backplane sem bloqueio e não sobrecarregado
- Baixa latência (menos de 35usec)
- Comutação da camada 3
- Gerenciado e monitorado (SNMP é uma vantagem a mais)
- Suporte de canal de porta (agrupamento de portas)

Normalmente, os switches de maior qualidade e produtividade também oferecerão melhor desempenho. Em sistemas maiores, a melhor opção é usar um switch gerenciado de alta capacidade que possa ser configurada e ajustada com base na eficiência do HPCC Systems.

Balanceamento de Carga

Um balanceador de carga distribui o tráfego da rede entre vários servidores. Cada nó do Roxie é capaz de receber solicitações e de retornar resultados. Consequentemente, um balanceador de carga distribui a carga de maneira eficiente para obter o melhor desempenho e evitar um possível gargalo.

Requerimento do Balanceamento de Carga

Requisitos mínimos de entrada

- Transferência: 1 Gigabit
- Portas Ethernet: 2
- Estratégia de balanceamento: Round Robin

Requisitos padrão

- Transferência: 8Gbps
- Portas Gigabit Ethernet: 4
- Estratégia de balanceamento: Flexível (iRules F5 ou equivalente)

Capacidade Recomendada

- Capacidade de fornecer rotação de carga cíclica (e não o balanceamento de carga)
- Capacidade de encaminhar o tráfego SOAP/HTTP
- Capacidade de fornecer roteamento de triangulação/n-path (tráfego de entrada através do balanceador de carga para o nó, respostas enviadas através do switch).
- Capacidade de tratar um cluster de nós como uma entidade única (para clusters de balanceamento de carga, não os nós)

ou

- Capacidade de empilhar ou estruturar em camadas os balanceadores de carga em vários níveis.

Dimensionamento do Sistema

Esta seção oferece orientações na hora de determinar os requisitos de dimensionamento para uma instalação inicial do HPCC. Abaixo estão alguns guias sugeridos de configuração que podem ser úteis ao planejar o seu sistema.

Mínimo Hardware Sugerido

O HPCC foi projetado para ser executado em hardware comum, podendo funcionar até mesmo em hardware inferior. A lista a seguir contém as especificações de hardware mínimas sugeridas. Como requisito absolutamente mínimo, você deve considerar os seguintes componentes de hardware para seu HPCC System. Essas diretrizes foram estabelecidas com base no uso real de missão crítica (tempo ativo) com alto volume de dados.

Escravo Thor	Processador	4 x 64 bits Processador Intel cada
	RAM	8GB por daemon
	Armazenamento	RAID - Leitura/gravação sequencial por nó de 200MB/seg
	Rede	Largura de banda de 1 Gb/s
Roxie	Processador	4 x 64 bits
	RAM	12 GB por Roxie
	Armazenamento	400 IOPS e 2 Volumes por (RAID opcional)
	Rede	Largura de banda de 1 Gb/s
Dali	Processador	4 x 64 bits cada
	RAM	24GB por Dali
	Armazenamento	RAID 1, 5, 6, 10 volumes de 200GB
Outro	Processador	4 x 64 bits
	RAM	12GB
	Armazenamento	RAID 1, 5, 6, 10 volumes de 200GB
	Rede	Largura de banda de 1 Gb/s

Software-Nós

Todos os nós devem ter sistemas operacionais idênticos. Recomendamos que todos os nós tenham instalados configurações BIOS e pacotes idênticos. Isso diminui significativamente as variáveis em caso de solução de problemas. Embora seja mais fácil gerenciar um sistema onde todos os nós são idênticos, isso não é obrigatório.

Requerimentos do Sistema Operacional

Os pacotes de instalação binária estão disponíveis para diversos sistemas operacionais Linux. Os requisitos da plataforma do HPCC System estão prontamente disponíveis no Portal do HPCC.

<https://hpccsystems.com/training/documentation/system-requirements>

Dependências

Para instalar o HPCC em seu sistema, é preciso ter os pacotes de componentes obrigatórios instalados no sistema. As dependências obrigatórias podem variar de acordo com a sua plataforma. Em alguns casos, as dependências estão incluídas nos pacotes de instalação. Em outras situações, a instalação pode não ser concluída com sucesso, e o utilitário de gerenciamento do pacote solicitará os pacotes obrigatórios. A instalação desses pacotes pode variar de acordo com a sua plataforma. Para detalhes sobre comandos de instalação específicos para obter e instalar esses pacotes, consulte os comandos específicos do seu Sistema Operacional.

Observação: Para instalações CentOS, o repositório Fedora EPEL é obrigatório.

Chaves SSH

Os componentes do HPCC usam chaves ssh para autenticar uns aos outros. Isso é obrigatório para a comunicação entre os nós. Fornecemos um script para geração de chaves. Você precisa executar esse script e distribuir as chaves públicas e privadas a todos os nós após ter instalado os pacotes em todos eles, porém antes de configurar um HPCC de nós múltiplos.

- Como usuário root (ou sudo como mostrado abaixo), gere uma nova chave usando este comando:

```
sudo /opt/HPCCSystems/sbin/keygen.sh
```

- Distribua as chaves para todos os nós. No diretório **/home/hpcc/.ssh**, copie esses três arquivos para o mesmo diretório (**/home/hpcc/.ssh**) em cada nó: Em todos os nós
 - **id_rsa**
 - **id_rsa.pub**
 - **authorized_keys**

Lembre-se de que os arquivos devem reter as permissões ao serem distribuídos. Essas chaves precisam pertencer ao usuário **“hpcc”**.

Hardware e Componentes

Esta seção oferece alguns insights sobre que tipo de hardware e infraestrutura nos quais o HPCC opera da melhor maneira. Este não é um conjunto amplo e completo de instruções, nem uma obrigação sobre qual hardware você precisa usar. Considere este como um guia para usar quando for implementar ou dimensionar seu HPCC System. As sugestões devem ser levadas em conta de acordo com as suas necessidades empresariais específicas.

O HPCC foi projetado para ser executado em hardware padrão, tornando o desenvolvimento e a manutenção de clusters de grande escala (petabytes) economicamente viáveis. Ao planejar o hardware do seu cluster, é necessário colocar na balança diversas considerações, incluindo domínios de fail-over e possíveis problemas de desempenho. O planejamento de hardware deve incluir a distribuição do HPCC entre múltiplos hosts físicos, como um cluster. Geralmente, é uma boa prática executar processos HPCC de um determinado tipo (por exemplo, Thor, Roxie ou Dali) em um host configurado especificamente para aquele tipo de processo.

Hardware Thor

Nós escravos Thor exigem um equilíbrio adequado de CPU, memória RAM, rede e E/S de disco para operar da maneira mais eficiente. Um único nó escravo do Thor funciona de maneira ideal quando alocado em 4 núcleos de CPU, 8GB de memória RAM, I/O de rede de 1Gb/segundo e leitura/gravação de disco sequencial de 200MB/segundo.

A arquitetura de hardware pode oferecer valor superior dentro de um único servidor físico. Em tais casos, é possível usar múltiplos escravos para configurar seus servidores físicos maiores de modo a executar múltiplos nós de escravos Thor por servidor físico.

É importante observar que HPCC, por natureza, é um sistema de processamento paralelo e que todos os nós de escravos Thor serão executados precisamente ao mesmo tempo. Desta forma, ao alocar mais de um escravo HPCC Thor por máquina física, verifique se cada escravo atende aos requisitos recomendados.

Por exemplo, em sua eficiência ideal, 1 servidor físico com 48 núcleos, 96GB de memória RAM, I/O de rede de 10Gb/segundo e sequencial de 2GB/segundo seria capaz de executar dez (10) escravos HPCC Thor. A ordem para otimização do uso de recursos em um nó escravo Thor é I/O de disco de 60%, rede de 30% e CPU de 10%. Qualquer aumento na I/O sequencial terá o maior impacto sobre a velocidade, seguido por melhorias na rede e, depois, por melhorias na CPU.

A arquitetura de rede também é algo importante a ser considerado. Os nós HPCC Thor funcionam idealmente em uma arquitetura de rede dinamizada entre todos os processos escravos Thor.

RAID é recomendado e todos os níveis de RAID adequados para operações de leitura/gravação sequencial e alta disponibilidade são aceitáveis. Por exemplo, RAID1, RAID10, RAID5 (preferido) e RAID6.

Configurações de Hardware do Roxie

Para garantir operações eficientes, os processos Roxie do HPCC exigem um equilíbrio adequado, embora diferente (do Thor), de CPU, memória RAM, rede e de I/O de disco. Um único nó HPCC Roxie funciona de maneira ideal quando alocado a 6 ou mais núcleos de CPU, 24GB de memória RAM, backbone de rede de 1Gb/segundo e IOPS de leitura aleatória 4K de 400/segundo.

Cada nó HPCC Roxie conta com dois discos rígidos, sendo que cada um é capaz de atingir um IOPS de busca aleatória 4K de 200/segundo. As recomendações de disco rígido para a eficiência do Roxie são SAS de 15K ou SSD. Uma boa regra prática é que, quanto maior for o IOPS de leitura aleatória, melhor e mais rápido será o desempenho do seu Roxie.

A execução de múltiplos nós HPCC Roxie em um único servidor físico não é recomendada, exceto em casos de virtualização ou contêineres.

Configure seu sistema para equilibrar o tamanho de seus clusters Thor e Roxie. O número de nós Roxie nunca deve ultrapassar o número de nós Thor. Além disso, o número de nós Thor deve ser uniformemente divisível pelo número de nós Roxie. Isso garante uma distribuição eficiente de partes de arquivo do Thor ao Roxie.

Configurações de Hardware do Dali e Sasha

O Dali do HPCC processa metadados de armazenamento de cluster na memória RAM. Para a melhor eficiência, garanta no mínimo 48GB de memória RAM, 6 ou mais núcleos de CPU, interface de rede de 1Gb/segundo e disco de alta disponibilidade para um único Dali do HPCC. Os processos em Dali do HPCC são um dos poucos componentes nativos ativos/passivos. Recomenda-se usar um clustering de “swinging disk” para uma configuração de alta disponibilidade. Para um único processo em Dali do HPCC, pode ser usado qualquer nível RAID de alta disponibilidade (HA).

Sasha só armazena dados em discos disponíveis localmente, lendo dados do Dali e, depois, processando-os ao arquivar workunits (WUs) em disco. Recomenda-se configurar o Sasha para uma quantidade maior de arquivamento para que o Dali não mantenha um número excessivo de workunit na memória. Isso exige maior espaço em disco.

A alocação de maior espaço em disco para o Sasha é uma prática consagrada, já que configurar o Sasha para mais arquivamento traz melhores benefícios ao Dali. Uma vez que o Sasha auxilia o Dali ao realizar a organização, ele funciona melhor quando está em seu próprio nó. O ideal é evitar colocar o Sasha e o Dali no mesmo nó, uma vez que o nó que executa esses componentes é extremamente importante, especialmente quando se fala em recuperação em caso de perdas. Por isso, ele deve ser o mais sólido possível: unidades RAID, tolerantes a falhas, etc.

Interações Sasha/Dali

Uma função essencial do Sasha é a coalescência. Quando o Dali é desligado, ele salva seu armazenamento em memória para uma nova edição de armazenamento ao criar um novo *dalisdsXXXX.xml*, onde XXXX é incrementado de acordo com a nova edição. A edição atual é gravada pelo armazenamento de nome de arquivo.XXXX

Uma solicitação explícita para salvar usando o *dalidiag*:

```
dalidiag . -save
```

As novas edições, conforme o exemplo acima, são criadas da mesma maneira. Durante uma operação "salvar" explícita, todas as alterações no SDS são bloqueadas. Consequentemente, todos os clientes serão bloqueados se tentarem fazer qualquer alteração até que a operação "salvar" esteja concluída.

Há certas opções (embora não comumente usadas) capazes de configurar o Dali para detectar tempo ocioso/em descanso e forçar uma operação de salvar exatamente da mesma maneira que uma solicitação de salvar explícita opera, o que significa que haverá o bloqueio de quaisquer transações de gravação durante o processo.

Todas as alterações de SDS no Dali são gravadas em um log de transações delta (no formato XML) com uma convenção de nomenclatura de *daliincXXXX.xml*, onde XXXX corresponde a edição de armazenamento atual. Elas também são opcionalmente espelhadas para um local de backup. Esse log de transação cresce indefinidamente até o armazenamento ser salvo.

Na configuração recomendada/normal, o Sasha é principal criador de novas edições de armazenamento de SDS. Ele faz isso de acordo com um cronograma e outras opções de configuração (por exemplo, é possível configurar para um tamanho mínimo de log de transação delta). O Sasha lê o último armazenamento salvo e o log de transação atual, e reproduz o log de transação sobre o último armazenamento salvo para formar uma nova versão em memória e depois o salva. Diferente do processo de salvar do Dali, isso não bloqueia nem interfere no Dali. No caso de um encerramento repentino do processo do Dali (abortado ou em caso de falta de energia), o Dali usa o mesmo log de transações delta na reinicialização para reproduzir o último salvamento e alterações para retornar ao estado operacional mais recente.

Outros Componentes HPCC

ECL Agent, ECLCC Server, DFU Server, o Thor Master e o ECL Watch são processos administrativos usados para auxiliar os componentes dos clusters principais.

Para máxima eficiência, é necessário 24GB de memória RAM, 6 ou mais núcleos de CPU, velocidade de rede de 1Gb/segundo e discos de alta disponibilidade. Esses componentes podem ser altamente disponibilizados no modo ativo/ativo.

Manutenção da Rotina

Para garantir que seu HPCC System continue operando perfeitamente, são necessários alguns cuidados e manutenção. As próximas seções tratam das tarefas de manutenção rotineira para seu HPCC System.

Manipulação dos Dados

Ao começar a trabalhar com o HPCC System, recomenda-se ter alguns dados no sistema para processamento. Os dados são transferidos para o HPCC System através de um processo denominado *spray* (distribuição aos nós). Da mesma forma, os dados são retirados do HPCC System através de um processo denominado *despray* (consolidação aos nós).

Uma vez que o HPCC é um cluster de computador, os dados são implementados sobre os nós que compõem o cluster. Um *spray*, ou importação, é a transferência de um arquivo de dados de um local (como uma zona de entrada de arquivo) para um cluster. O termo *spray* foi adotado devido à natureza da transferência dos arquivos – o arquivo é particionado entre todos os nós em um cluster.

Um *despray*, ou exportação, é a transferência de um arquivo de dados de um cluster de Refinaria de Dados para um único local do computador (como uma zona de entrada de arquivo). O termo *despray* foi adotado em função da natureza da transferência dos arquivos – o arquivo é então remontado a partir de suas peças em todos os nós no cluster e colocado em um único arquivo no destino.

Uma *Zona de entrada de arquivos* (ou zona de chegada) é um local de armazenamento físico definido no ambiente do seu sistema. É possível definir um ou mais desses locais. Um daemon (*dafilesrv*) precisa estar em execução no servidor para possibilitar *sprays* e *desprays* do arquivo. É possível realizar o *spray* ou *despray* de alguns arquivos para sua zona de entrada de arquivo através do ECL Watch. Para enviar arquivos grandes, é necessária uma ferramenta compatível com o protocolo de cópia de segurança – algo como um WinSCP.

Para obter mais informações sobre o processamento de dados do HPCC, consulte os documentos *Manipulação dos Dados do HPCC* e *Tutorial de dados do HPCC*.

Backup dos Dados

O backup de dados essenciais é parte integrante da manutenção de rotina. Elabore uma estratégia de backup para atender às necessidades de sua organização. Esta seção não visa substituir sua estratégia de backup atual – em vez disso, ela a complementa ao destacar as considerações especiais para HPCC Systems.®

Considerações sobre Backup

Você provavelmente já conta com alguma estratégia de backup em vigor; ao adicionar o HPCC Systems® em seu ambiente de operação, há certas considerações adicionais sobre as quais você deve estar ciente. As seções a seguir discutem as considerações de backup para os componentes individuais do HPCC System.

Dali

Dali pode ser configurado para criar seu próprio backup. É altamente recomendado que o backup seja mantido em um servidor ou nó diferente para fins de recuperação de desastres. É possível especificar o local da pasta de backup do Dali usando o Configuration Manager. Recomenda-se manter múltiplas gerações de backups a fim de conseguir realizar a restauração para um determinado período no tempo. Por exemplo, você pode querer salvar instantâneos diariamente ou semanalmente.

Recomenda-se manter cópias de backup no nível de sistema usando métodos tradicionais. Qualquer que seja o método ou esquema, é altamente recomendado manter um backup do Dali.

Deve-se tentar evitar colocar o Dali, Sasha e até mesmo o Thor Master no mesmo nó. O ideal é que cada um desses componentes esteja em nós separados não apenas para reduzir a carga no hardware de sistema (permitindo que ele funcione melhor), como também para que você possa recuperar todo o ambiente, arquivos e tarefas em caso de perda. Isso também influenciaria todos os outros clusters Thor/Roxie no mesmo ambiente se você perder esse nó.

Sasha

Sasha é o componente que cuida da coalescência da SDS. É normalmente o único componente que cria novas edições de armazenamento. Também é o componente que cria os metadados XREF utilizados pelo ECLWatch. Observe que o Sasha pode criar uma grande quantidade de dados de arquivo. Depois de arquivadas, as tarefas não ficam mais disponíveis no armazenamento de dados do Dali. Os arquivos ainda podem ser acessados pelo ECL Watch ao serem restaurados para o Dali.

Se você precisa de alta disponibilidade para tarefas arquivadas, é necessário fazer um backup destas tarefas no nível de sistema usando métodos tradicionais de backup.

Servidor DFU

O DFU Server não possui dados. As tarefas DFU são armazenadas no Dali até serem arquivadas pelo Sasha.

ECLCC Server

O ECLCC Server não armazena dados. As tarefas ECL são armazenadas no Dali e são arquivadas pelo Sasha.

ECL Agent

O ECL Agent não armazena dados.

ECL Scheduler

O ECL Scheduler não armazena dados. As workunits do ECL são armazenadas no Dali.

ESP Server

O ECLCC Server não armazena dados. Se estiver usando certificados SSL, o backup de chaves públicas e privadas deve ser feito por métodos tradicionais.

Thor

Thor, a refinaria de dados, como um dos componentes críticos dos HPCC Systems,® precisa ser salvo em backup. Faça um backup do Thor ao configurar a replicação e definir uma tarefa de backup noturno do cron. Se não houver um RAID configurado, crie um backup do Thor sob demanda antes e/ou depois de qualquer troca de nó ou troca de unidade.

Uma parte muito importante da administração do Thor é verificar os logs para garantir que os backups anteriores foram concluídos com sucesso.

Backupnode

O Backupnode é uma ferramenta que acompanha o HPCC. Ela permite que você crie backups de nós do Thor sob demanda ou por script. Também é possível usar o Backupnode regularmente em um crontab ou adicionar ao seu ambiente um componente do Backupnode com o Gerenciador de Configurações. Sempre é necessário executá-lo no Thor Master do cluster.

O exemplo a seguir é uma maneira sugerida de acionar o Backupnode manualmente.

```
/bin/su - hpcc -c "/opt/HPCCSystems/bin/start_backupnode thor" &
```

O parâmetro de linha de comando deve corresponder ao nome do seu cluster Thor. Em seu ambiente de produção, é provável que você forneça nomes descritivos para seus clusters Thor.

Por exemplo, se o cluster do Thor for denominado thor400_7s, você o chamaria de start_backupnode thor400_7s.

```
/bin/su - hpcc -c "/opt/HPCCSystems/bin/start_backupnode thor400_7s" &
```

O Backupnode é executado regularmente

É possível usar o cron para executar o Backupnode regularmente. Por exemplo, você pode definir uma entrada crontab (para backup do thor400_7s) para ser executada à 1 da manhã diariamente:

```
0 1 * * * /bin/su - hpcc -c "/opt/HPCCSystems/bin/start_backupnode thor400_7s" &
```

O Backupnode grava sua atividade em um arquivo de log. Esse log pode ser encontrado em:

/var/log/HPCCSystems/backupnode/MM_DD_YYYY_HH_MM_SS.log

O (MM) Mês, (DD) Dia, (AAAA) Ano com 4 dígitos, (HH) Hora, (MM) Minutos e (SS) Segundos do backup que inclui o nome do arquivo de log.

O arquivo de log principal existe no nó mestre do Thor. Ele mostra em que nós é executado e se foi concluído. É possível encontrar outros logs do Backupnode em cada um dos nós do Thor mostrando quais arquivos, se aplicável, ele precisou restaurar.

É importante verificar os logs para garantir que os backups anteriores foram concluídos com sucesso. A entrada a seguir é do log do Backupnode mostrando que o backup foi concluído com sucesso:

```
00000028 2014-02-19 12:01:08 26457 26457 "Completed in 0m 0s with 0 errors"
00000029 2014-02-19 12:01:08 26457 26457 "backupnode finished"
```

Roxie

Os dados do Roxie são protegidos por três formas de redundância:

- **Retenção de arquivo de dados de fonte original:** Quando uma consulta é publicada, os dados são normalmente copiados de um local remoto, seja de um Thor ou de um Roxie. Dessa forma, os dados do Thor podem servir como backup, contanto que não sejam removidos nem alterados no Thor. Os dados do Thor normalmente são retidos por um período suficiente para servir como uma cópia de backup.
- **Redundância de nó-par:** Cada nó do agente normalmente tem um ou mais nós pares em seu cluster. Cada par armazena uma cópia dos arquivos de dados que serão lidos.
- **Redundância de cluster irmão:** Embora não seja necessário, as implementações do Roxie podem ser executadas em múltiplos clusters Roxie configurados de modo idêntico. Quando dois clusters são implementados para produção, cada nó possui um gêmeo idêntico em termos de consultas e/ou dados armazenados no nó no outro cluster. Essa configuração oferece múltiplas cópias redundantes de arquivos de dados. Com três clusters Roxie que possuem redundância de nó par, há sempre seis cópias de cada parte do arquivo a qualquer momento. Isso elimina a necessidade de usar procedimentos tradicionais de backup para arquivos de dados Roxie.

Zona de entrada de arquivos

A zona de entrada de arquivos é usada para hospedar arquivos de entrada e saída. Isso deve ser tratado da mesma maneira que em um servidor FTP. Use backups tradicionais no nível de sistema.

Miscelânea

O backup de quaisquer complementos de componentes extras, seus arquivos de ambiente (environment.xml) ou outras configurações personalizadas deve ser feito de acordo com os métodos tradicionais de backup.

Arquivo de Log

É possível conferir as mensagens de sistema e ver quaisquer mensagens de erro no momento em que são reportadas e registradas em arquivos de log. Os arquivos de log podem ajudá-lo a entender o que está ocorrendo no sistema e são úteis na hora de solucionar problemas.

Logs de Componente

Há arquivos de log para cada componente nos diretórios abaixo **/var/log/HPCCSystems** (local padrão). Opcionalmente, é possível configurar o sistema para gravar os logs em um diretório diferente. Você deve saber a localização dos arquivos de log e consultá-los primeiramente ao solucionar quaisquer problemas.

Há arquivos de log que registram a atividade entre vários componentes. É possível encontrar os arquivos de log em subdiretórios nomeados de forma correspondente aos componentes que eles rastreiam. Por exemplo, os logs do Thor são encontrados em um diretório chamado mythor; o log do Sasha deve estar no diretório mysasha, enquanto o log do ESP deve estar no diretório myesp.

Cada um dos subdiretórios de componente contém vários arquivos de log. A maioria dos arquivos de log usa uma convenção de nomenclatura que inclui o nome do componente, a data e a hora no nome do arquivo de log. Normalmente há também um link para o componente com um nome simples, como esp.log, que é um atalho para o arquivo de log atual mais recente daquele componente.

Entender os arquivos de log, e o que normalmente é reportado neles, ajuda a solucionar problemas no HPCC System.

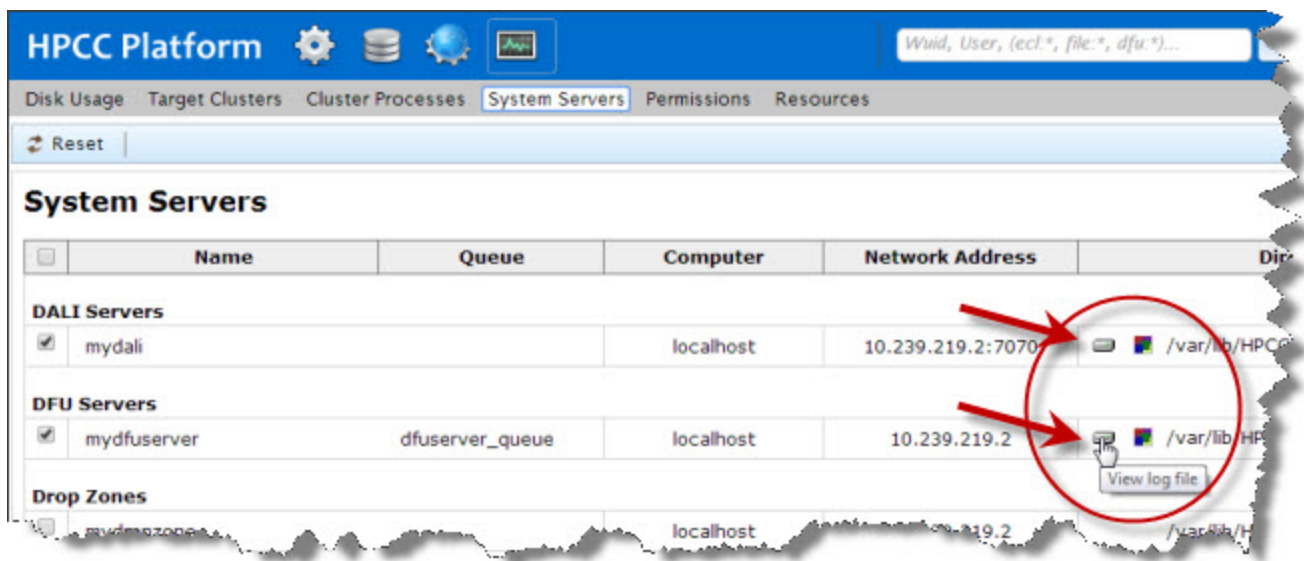
Como parte da manutenção de rotina, recomenda-se fazer um backup, arquivar e remover os arquivos de log mais antigos.

Acessando os Arquivos de Log

É possível acessar e ver arquivos de log diretamente através do diretório de log de componentes de um prompt de comando ou de uma aplicação do terminal. Também é possível ver os arquivos de log de componente através do ECL Watch.

Para ver os logs no ECL Watch, clique no ícone **Operations** e depois no link **System Servers**. Isso abre a página System Servers no ECL Watch. Essa página lista vários componentes do HPCC System. Na coluna **Directory** para cada componente, há um ícone de unidade de computador. Clique no ícone na linha do log de componente que deseja acessar.

Figure 4. Logs no ECL Watch



Também é possível acessar arquivos de log de outros links no ícone Operations no ECL Watch.

1. Clique no link **Target Clusters** para abrir a guia com links para os clusters do seu sistema.
2. Clique no ícone de unidade de computador (circulado em vermelho na figura acima) na linha do cluster e nó do log de componente que você deseja ver.

Para ver os logs de processos de cluster:

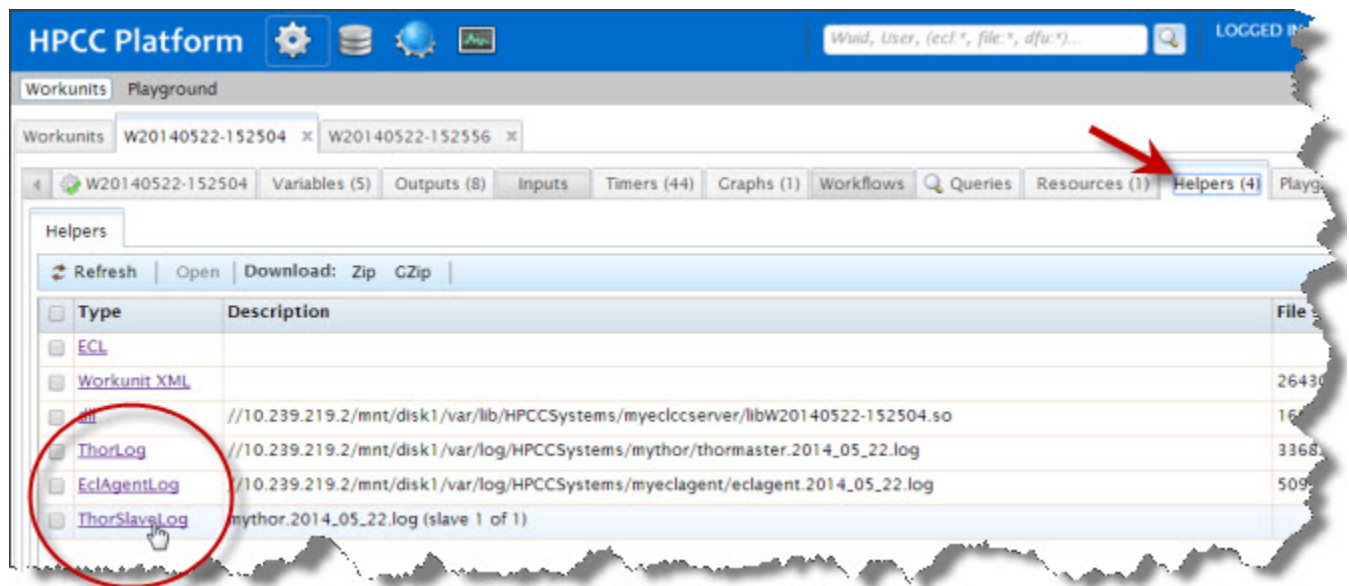
1. Clique no link **Cluster Process** para abrir a guia com links para os processos de clusters do seu sistema.
2. Clique no processo de cluster sobre o qual você deseja ver mais informações.

Por exemplo, clique no link **myroxie** . Você vai ver uma página com todos os nós de componentes. Você vai ver o ícone de unidade de computador na linha de cada nó. Clique nesse ícone para ver os logs do processo de cluster para esse nó.

Arquivos de log em ECL Workunits

Você também pode acessar os arquivos de log do Thor ou do ECL Agent pelas tarefas de ECL. (não disponível para tarefas Roxie) No ECL Watch, ao examinar os detalhes de tarefa, você verá uma aba chamada **Helpers** . Clique na aba Helpers para exibir os arquivos de log relevantes para a determinada workunit.

Figure 5. Logs nas workunits do ECL Watch



Preflight

A primeira etapa da certificação de que a plataforma está instalada e configurada adequadamente é executar uma verificação preflight dos componentes. Isso garante que todas as máquinas e executáveis adequados estão operacionais. Além disso, a verificação confirma que há espaço em disco suficiente, memória disponível e valores percentuais de CPU aceitáveis.

- Abra o ECL Watch em seu navegador usando o seguinte URL:

`http://nnn.nnn.nnn.nnn:pppp` (onde `nnn.nnn.nnn.nnn` é o endereço IP do seu ESP Server e `pppp` é a porta. A porta padrão é 8010)

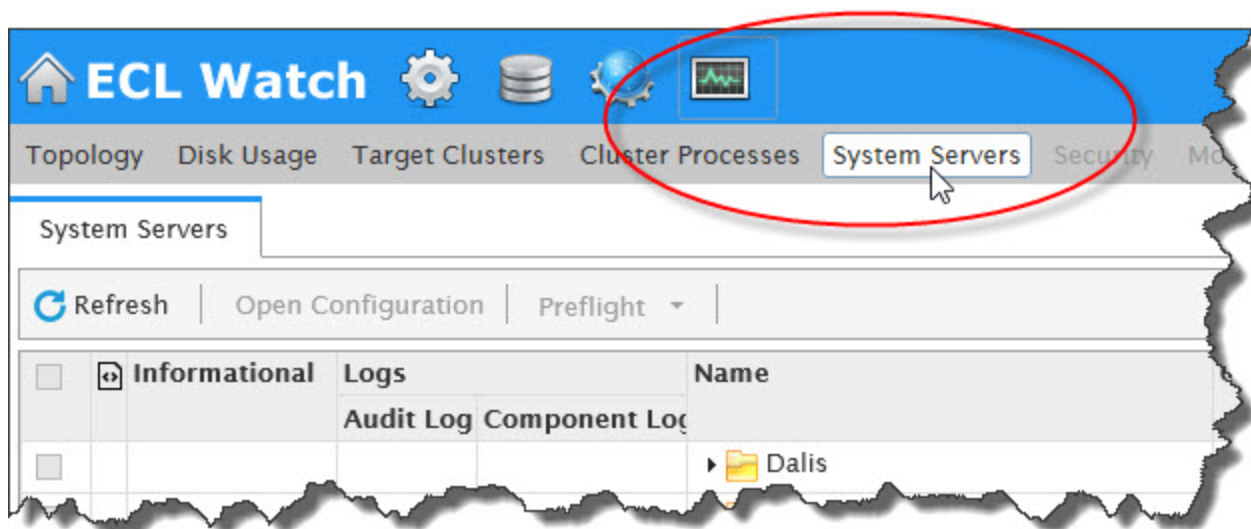


Note: Seu endereço IP poderá ser diferente dos endereços fornecidos nestas imagens. Use o endereço IP fornecido pela sua instalação.

Preflight do Servidores do Sistema

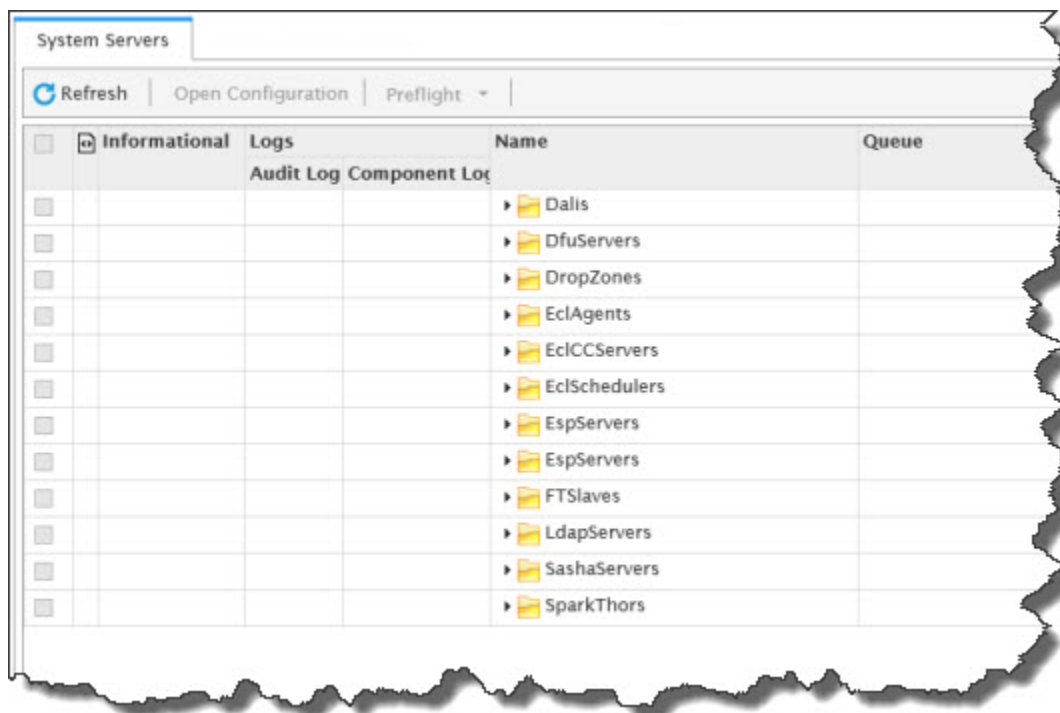
1. Clique no ícone **Operations** e clique no link **System Server**.

Figure 6. Link Servidores do Sistema



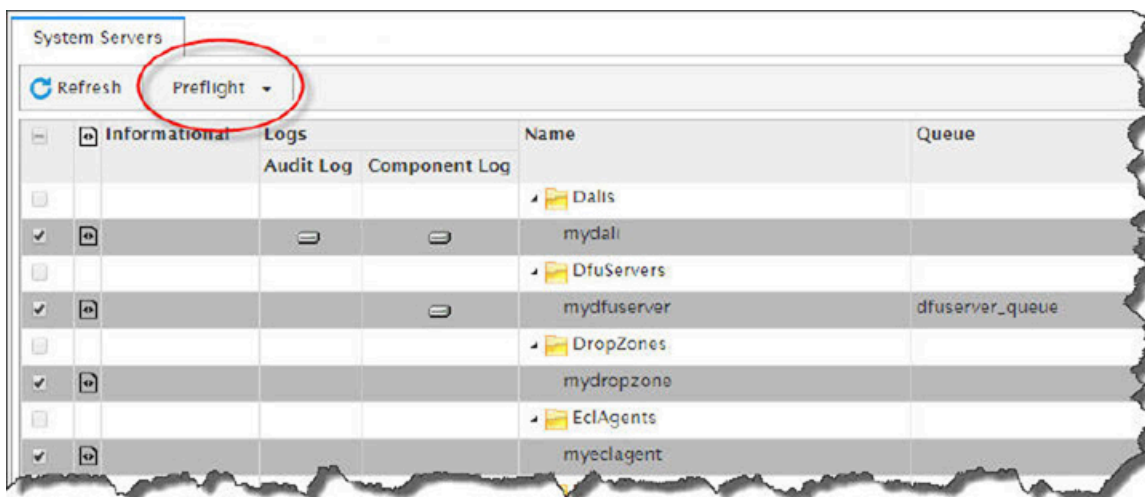
É exibida uma tela semelhante à mostrada abaixo.

Figure 7. Página Servidores do Sistema



2. Expand the folder for the System Server then check the box next to the desired component(s).

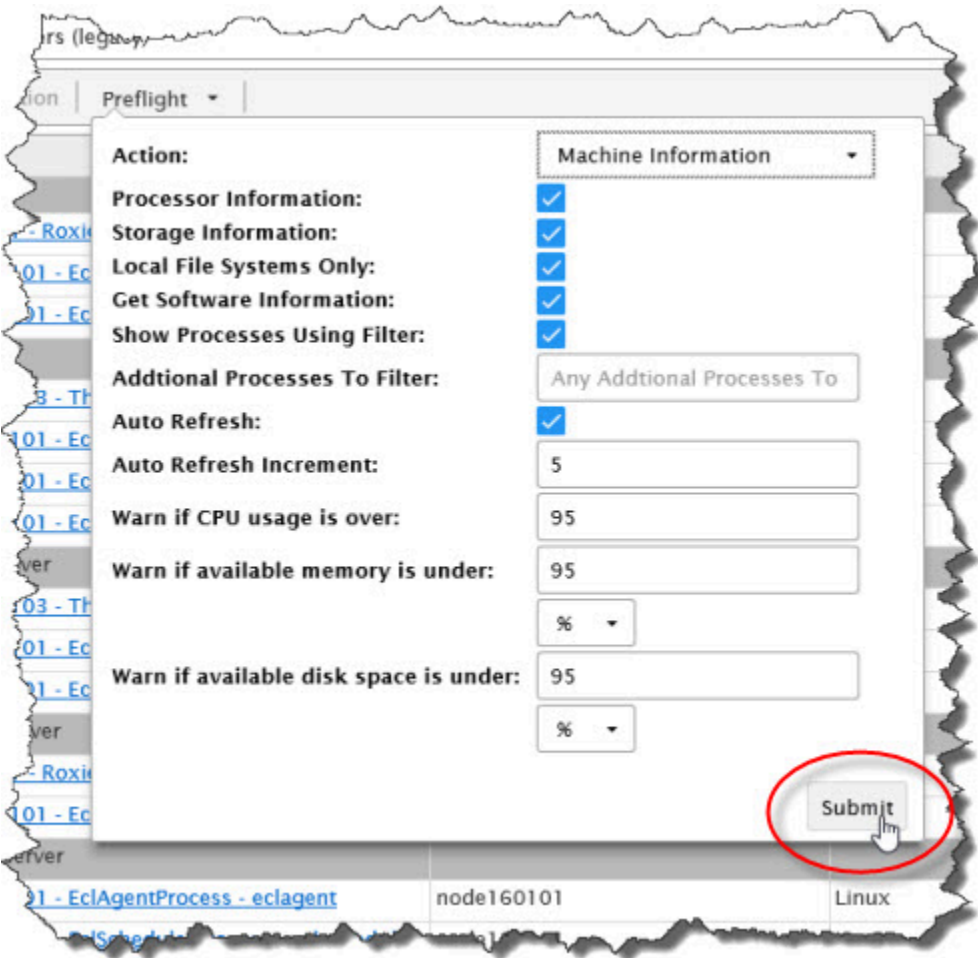
Figure 8. Seleccione os Servidores do Systema



Com os servidores selecionados e o botão de preflight ativado, você pode pressioná-lo para exibir as opções de preflight.

3. Marque ou desmarque qualquer uma das opções, então pressione o botão **Submit** para iniciar o preflight.

Figure 9. Submit



RESULTADOS ESPERADOS:

Depois que o botão Submit é pressionado, será exibida uma tela semelhante à mostrada abaixo.

Figure 10. Informações sobre componentes do sistema

The screenshot shows the HPCC Systems Preflight Results window. The 'Machine Information' tab is selected. The 'Preflight Results' section shows a table of system components and their status.

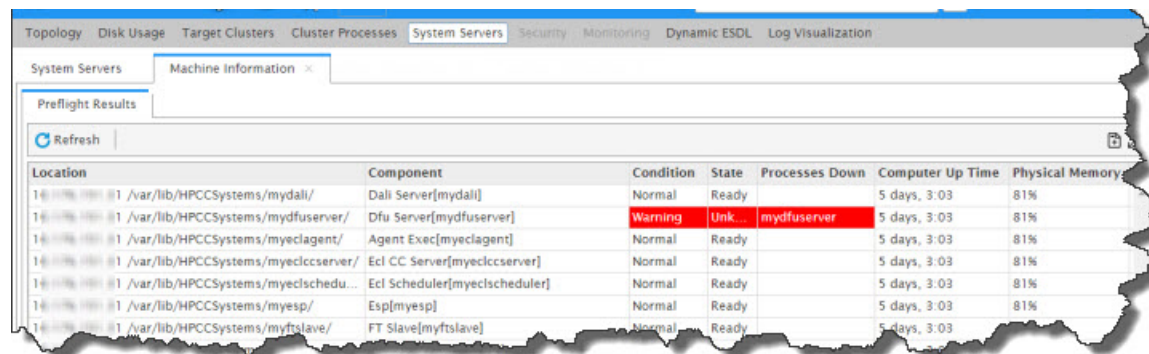
Location	Component	Condition	State	Processes Down	Computer Up Time	Physical M
10.176.151.31 /var/lib/HPCCSystems/mydali/	Dali Server[mydali]	Normal	Ready		4 days, 23:06	76%
10.176.151.31 /var/lib/HPCCSystems/mydfuserver/	Dfu Server[mydfuserver]	Normal	Ready		4 days, 23:06	76%
10.176.151.31 /var/lib/HPCCSystems/myeclagent/	Agent Exec[myeclagent]	Normal	Ready		4 days, 23:06	76%
10.176.151.31 /var/lib/HPCCSystems/myeclagent/	Agent Exec[myeclagent]	Normal	Ready		4 days, 23:06	76%

Essa tela exibe informações sobre vários componentes do sistema. Essas informações indicam se vários componentes estão realmente executando corretamente. A página resultante exibe informações úteis sobre cada componente. O nome, a condição e o estado do componente, há quanto tempo está em execução, uso de disco e memória e outras informações podem ser vistas rapidamente.

Se houver algum alerta, os componentes serão destacados, indicando que requerem mais atenção.

Por exemplo, a imagem a seguir indica que há um incidente com servidor DFU.

Figure 11. Alerta do Servidor do Sistema



The screenshot shows the HPCC Systems Preflight Results window. The 'System Servers' tab is selected, and the 'Machine Information' sub-tab is active. A table lists the components and their status. The 'mydfuserver' component is highlighted in red, indicating a warning.

Location	Component	Condition	State	Processes Down	Computer Up Time	Physical Memory
14.10.10.1 /var/lib/HPCCSystems/mydali/	Dali Server[mydali]	Normal	Ready		5 days, 3:03	81%
14.10.10.1 /var/lib/HPCCSystems/mydfuserver/	Dfu Server[mydfuserver]	Warning	Unk...	mydfuserver	5 days, 3:03	81%
14.10.10.1 /var/lib/HPCCSystems/myeclagent/	Agent Exec[myeclagent]	Normal	Ready		5 days, 3:03	81%
14.10.10.1 /var/lib/HPCCSystems/myeclccserver/	Ecl CC Server[myeclccserver]	Normal	Ready		5 days, 3:03	81%
14.10.10.1 /var/lib/HPCCSystems/myeclscheduler/	Ecl Scheduler[myeclscheduler]	Normal	Ready		5 days, 3:03	81%
14.10.10.1 /var/lib/HPCCSystems/myesp/	Esp[myesp]	Normal	Ready		5 days, 3:03	81%
14.10.10.1 /var/lib/HPCCSystems/myftslave/	FT Slave[myftslave]	Normal	Ready		5 days, 3:03	

Preflight de Clusters de Destino

Use o link do Target Clusters para realizar o preflight de todos os clusters.

1. Clique no ícone **Operations** e clique no link **Target Clusters**.

Figure 12. Link Target Clusters



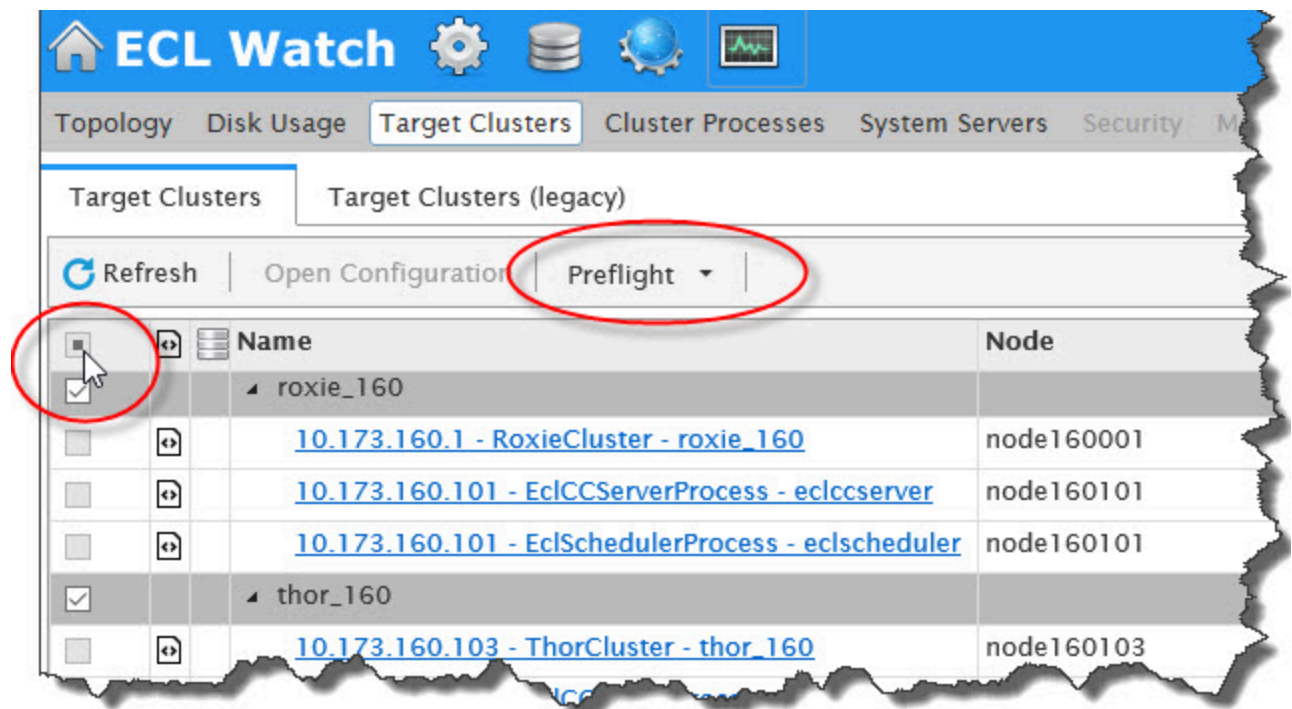
Isto exibe uma lista detalhada de todos os Clusters do sistema.

2. Clique na caixa de seleção selecionar tudo, na linha superior à esquerda, para selecionar todos os clusters de destino.

Opcionalmente, você pode apenas marcar as caixas ao lado do(s) cluster(s) que deseja realizar o preflight. Se você optar por fazer em todos os Clusters de destino, não precisará realizar separadamente no Thor e Roxie, conforme detalhado abaixo.

Com os clusters selecionados e o botão preflight ativo, você pode pressioná-lo para exibir as opções de preflight.

Figure 13. Selecionar Target Clusters



3. Marque ou desmarque qualquer uma das opções desejadas, então pressione o botão **Submit** para iniciar o preflight.

Figure 14. Submit

The screenshot shows the HPCC Systems Preflight configuration window. The 'Preflight' tab is active. The configuration options are as follows:

Option	Value
Action:	Machine Information
Processor Information:	<input checked="" type="checkbox"/>
Storage Information:	<input checked="" type="checkbox"/>
Local File Systems Only:	<input checked="" type="checkbox"/>
Get Software Information:	<input checked="" type="checkbox"/>
Show Processes Using Filter:	<input checked="" type="checkbox"/>
Additional Processes To Filter:	Any Additional Processes To
Auto Refresh:	<input checked="" type="checkbox"/>
Auto Refresh Increment:	5
Warn if CPU usage is over:	95
Warn if available memory is under:	95
	%
Warn if available disk space is under:	95
	%

The 'Submit' button is circled in red at the bottom right of the window. The status bar at the bottom shows 'EclAgentProcess - eclagent', 'node160101', and 'Linux'.

ATENÇÃO: Dependendo do tamanho do seu sistema, poderá haver um pequeno atraso na exibição dos resultados.

RESULTADOS ESPERADOS:

Depois de pressionar **Submit**, será exibida uma tela semelhante à mostrada abaixo.

Figure 15. Informação Target Cluster



Location	Component	Computer Up Time	Physical Mem	Swap	/var/lib/HPCCSys	/run
10.173.160.101 /var/lib/HPCCSyste...	Ecl CC Server[eclccserver]	96 days, 17:03	80%	99%	80%	95%
10.173.160.101 /var/lib/HPCCSyste...	Agent Exec[eclagent]	96 days, 17:03	80%	99%	80%	95%
10.173.160.101 /var/lib/HPCCSyste...	Ecl Scheduler[eclsched...	96 days, 17:03	80%	99%	80%	95%
10.173.160.1 /var/lib/HPCCSystems...	Roxie Server[roxie_160]	132 days, 23:44	20%		31%	91%
10.173.160.2 /var/lib/HPCCSystems...	Roxie Server[roxie_160]	132 days, 23:44	20%		31%	91%
10.173.160.3 /var/lib/HPCCSystems...	Roxie Server[roxie_160]	132 days, 23:44	20%		31%	91%

Essa tela exibe informações sobre os nós dos componentes do sistema. Essas informações podem ajudar a indicar se tudo está operando normalmente e destacar possíveis motivos de preocupação.

Se houver algum alerta pertinente, ele será destacado. Esse alertas exigem uma atenção adicional.

Se houver campos laranja, você deverá examinar os componentes especificados mais detalhadamente. Isso indica algum tipo de problema ou anormalidade.

Preflight Thor

1. Clique no ícone **Operations** e clique no link **Clusters Processes**.

Figure 16. Link Cluster Processes



2. Expanda o cluster Thor clicando na seta ao lado do link **ThorCluster**.

Figure 17. Link Thor Cluster

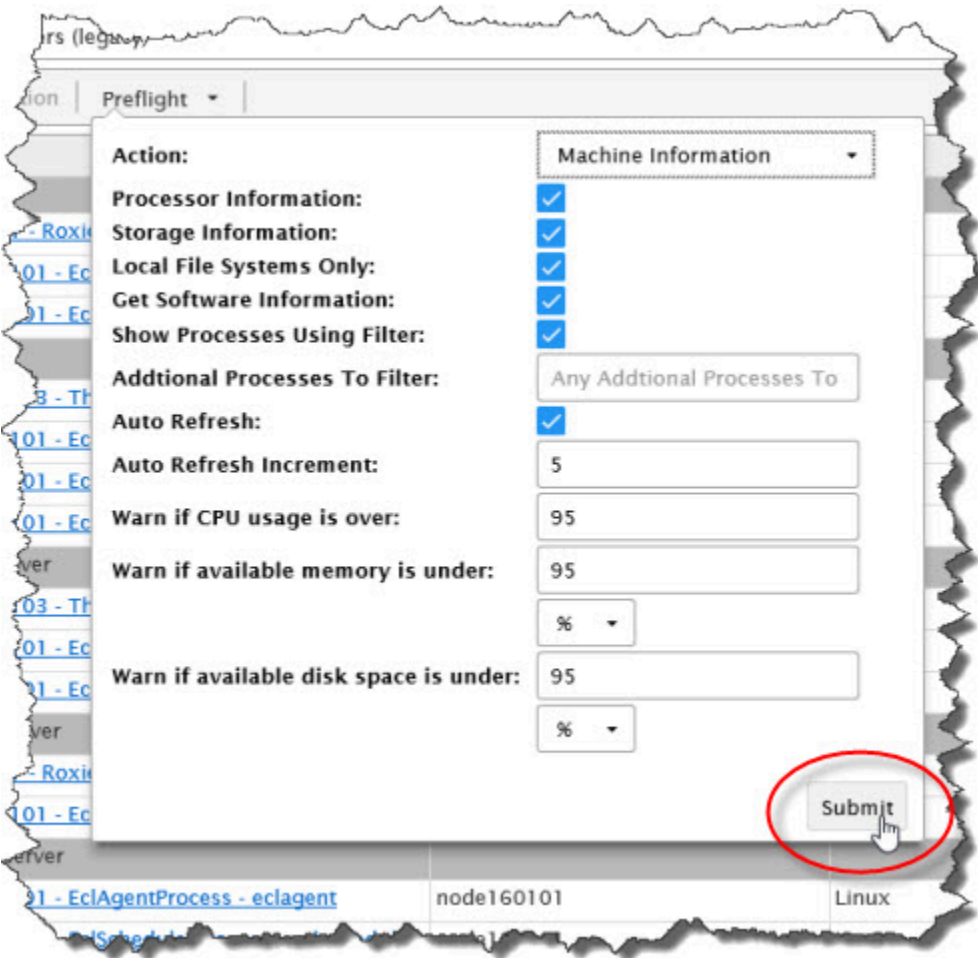


3. Clique na caixa ao lado de cada nós para analisar, ou em **Select All** na primeira linha.
4. Pressione o botão **Submit** para iniciar o preflight

Com sistema selecionado e o botão de preflight ativo, você pode exibir as opções de preflight.

5. Selecione ou desmarque qualquer uma das opções desejadas, então pressione o botão **Submit** para iniciar o preflight.

Figure 18. Submit



RESULTADOS ESPERADOS:

Depois que o botão Submit é pressionado, será exibida uma tela semelhante à mostrada abaixo.

Figure 19. Resultado do Cluster Process

The screenshot shows the HPCC Systems Preflight Results window. The 'Cluster Processes' tab is selected. The table displays the following information:

Location	Component	Condition	State	Processes Down	Computer Up Time	Physical Memory
10.179.140.1 /var/lib/HPCCSystems/thor_160/	Thor Slave[thor_160]	Normal	Ready		91 days, 2:22	23%
10.179.140.2 /var/lib/HPCCSystems/thor_160/	Thor Slave[thor_160]	Normal	Ready		225 days, 3:54	20%
10.179.140.103 /var/lib/HPCCSystems/thor_160/	Thor Master[thor_160]	Normal	Ready		434 days, 12:06	2%

Essa tela exibe informações sobre o cluster selecionado. Essas informações podem ajudar a indicar se tudo está operando normalmente e destacar possíveis motivos de preocupação.

Se houver algum alerta pertinente, ele será destacado. Esses alertas exigem uma atenção adicional.

Preflight do cluster Roxie

1. Clique no ícone **Operations** e clique no link **Clusters Processes**.

Figure 20. Link Cluster Processes



2. Expanda o cluster Roxie clicando na seta ao lado do link **RoxieCluster**.

Figure 21. Link RoxieCluster



3. Marque a caixa ao lado para selecionar os nós individualmente para análise, ou marque **Select All** na primeira linha.
4. Com sistema selecionado e o botão de preflight ativo, você pode exibir as opções de preflight.
5. Selecione ou desmarque qualquer uma das opções desejadas, então pressione o botão **Submit** para iniciar o preflight.

Figure 22. Submit



RESULTADOS ESPERADOS

Depois que o botão Submit é pressionado, será exibida uma tela semelhante à mostrada abaixo.

Figure 23. Informações de sistema do Roxie

ClustersMachine Information

Preflight Results

Refresh

Location	Component	Computer Up Time	Physical
10.173.140.1 /var/lib/HPCCSystems/roxie_160/	Roxie Server[node160001]	4 days, 20:05	
10.173.140.2 /var/lib/HPCCSystems/roxie_160/	Roxie Server[node160002]	138 days, 21:37	31%
10.173.140.3 /var/lib/HPCCSystems/roxie_160/	Roxie Server[node160003]	138 days, 21:37	31%
10.173.140.4 /var/lib/HPCCSystems/roxie_160/	Roxie Server[node160004]	138 days, 21:37	31%

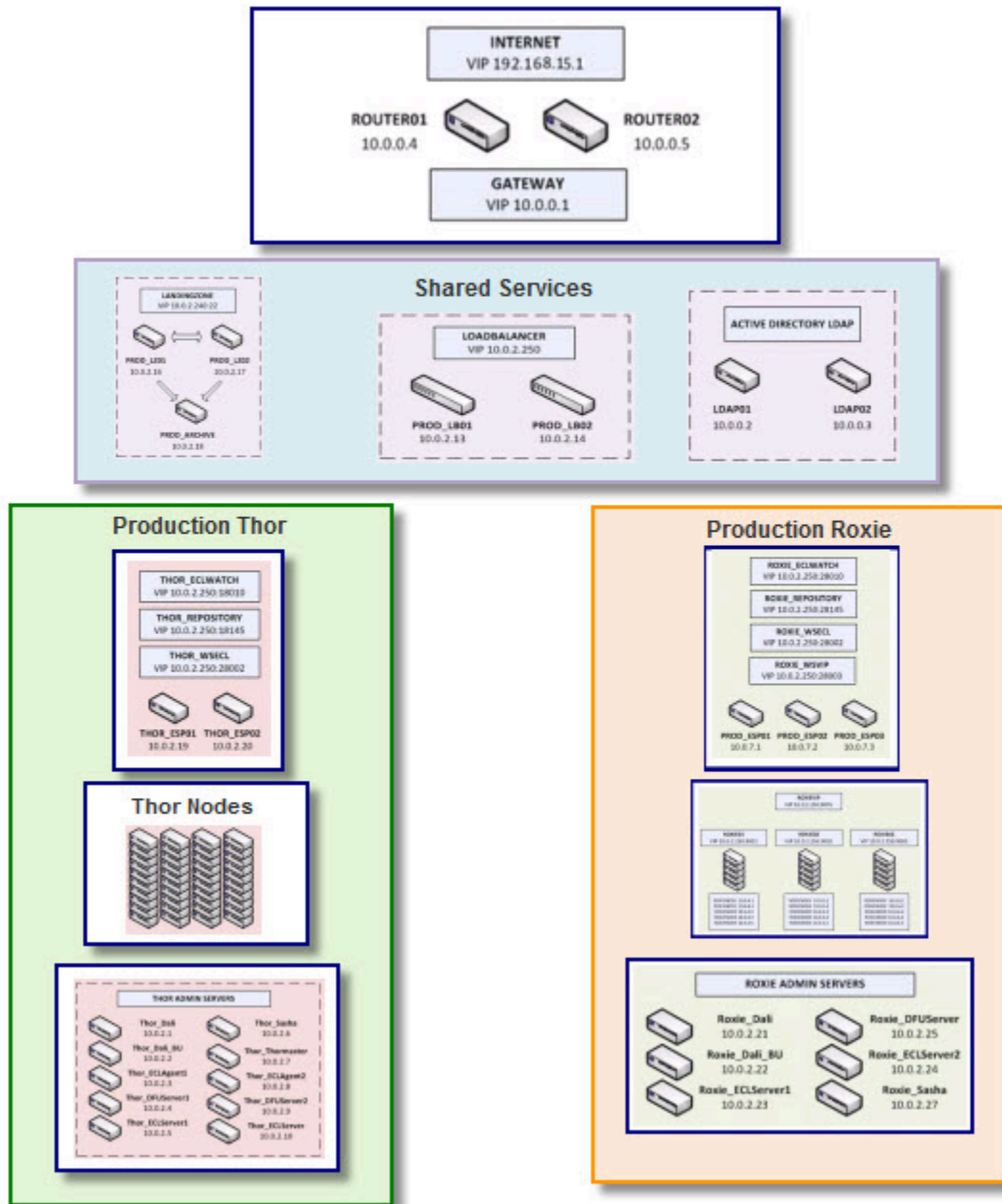
Indicam se os nós Roxie estão em execução e fornecem algumas informações adicionais sobre esses nós.

Se houver algum alerta pertinente, ele será destacado. Esses alertas exigem uma atenção adicional.

Configuração e Gerenciamento do Sistema

O HPCC System requer configuração. A ferramenta Configuration Manager (Gerenciador de Configurações-configmgr) inclusa no software do sistema é um item valioso para configurar seu HPCC System. O Gerenciador de Configurações é uma ferramenta gráfica fornecida para configurar o seu sistema. Ele possui um assistente que pode ser executado para gerar facilmente um arquivo de ambiente que o ajudará a configurar, ajustar e operar o sistema com rapidez. O Gerenciador de Configurações também possui uma opção avançada que permite uma configuração mais específica enquanto ainda usa a interface gráfica. Se quiser, é possível editar os arquivos de ambiente usando qualquer editor de texto ou xml, porém a estrutura do arquivo precisa permanecer válida.

Figure 24. Configuração de amostra de produção



O Configuration Manager é o utilitário no qual configuramos a plataforma HPCC. A configuração da plataforma HPCC é armazenada em um arquivo XML de nome **environment.xml**. Depois de gerar um arquivo de ambiente (xml), ele é salvo em um diretório de origem (o padrão é **/etc/HPCCSystems/source**). Você precisa então parar o sistema para copiá-lo no diretório HPCC ativo e depois distribuí-lo para o local em cada nó e reiniciar o HPCC System. Em nenhum momento durante a configuração você irá trabalhar no arquivo no ambiente em operação.

Ao instalar o pacote do HPCC System, um arquivo **environment.xml** de nó único padrão é gerado. Depois disso, é possível usar o Gerenciador de Configurações para modificá-lo e/ou criar um arquivo de ambiente diferente para configurar componentes ou adicionar nós. Há um assistente de Gerenciador de Configurações para ajudar a criar um arquivo de ambiente. Dê a qualquer arquivo de ambiente criado um nome descritivo que indique para que ele serve na fonte. Por exemplo, você pode criar um ambiente sem o Roxie, chamando o arquivo de **environmentNoRoxie.xml**.

Você então copiaria o novo arquivo de configuração gerado do diretório de origem para o diretório **/etc/HPCCSystems**. Renomeie o arquivo para **environment.xml** e reinicie para reconfigurar o sistema.

O Configuration Manager também oferece um **Advanced View** que possibilita maior granularidade para adicionar instâncias de componentes ou alterar as configurações padrão de componentes para usuários mais avançados. Mesmo se você planejar usar a Advanced View, recomendamos começar com um arquivo de configuração gerado pelo assistente e utilizar a ferramenta para editá-lo.

Mais informações e detalhes específicos para cada componente do Gerenciador de Configurações e atributos desses componentes são detalhadas em *Como usar o Gerenciador de Configurações*.

Utilizando o Gerenciador de Configurações

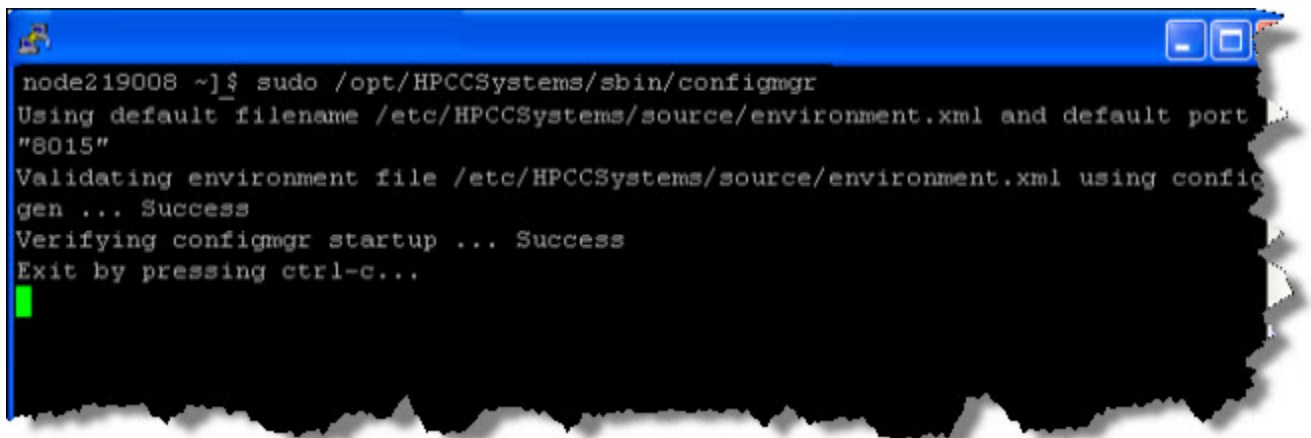
Esta seção irá orientá-lo a configurar um ambiente do HPCC Systems usando o Gerenciador de Configurações.

O pacote HPCC Systems precisa estar instalado em TODOS os nós.

É possível usar qualquer ferramenta ou script de shell desejado.

1. SSH para um nó em seu ambiente e fazer login como usuário com privilégio sudo. Sugerimos que esse seja o primeiro nó, que é um nó de suporte; no entanto, isso fica a seu critério.
2. Inicie o serviço do Gerenciador de Configurações no nó (novamente, sugerimos que seja um nó de suporte e que você use o mesmo nó para iniciar o Gerenciador de Configurações todas as vezes, mas isso fica totalmente a seu critério).

```
sudo /opt/HPCCSystems/sbin/configmgr
```



3. Usando um navegador de Internet, acesse a interface do Gerenciador de Configurações:

```
http://<ip de instalação do sistema>:8015
```

O assistente de inicialização do Gerenciador de Configurações é exibido.

Há diferentes maneiras para configurar o seu HPCC System. É possível usar o **Generate environment wizard** e utilizar esse ambiente ou, para usuários avançados, a **Advanced View** para uma personalização mais específica. Há também a opção de usar o recurso **Create blank environment** para gerar um ambiente vazio que você poderia então acessar e adicionar os componentes desejados.

Environment Wizard

1. Para usar o assistente, selecione o botão **Generate new environment using wizard**.



2. Forneça um nome ao arquivo do ambiente.

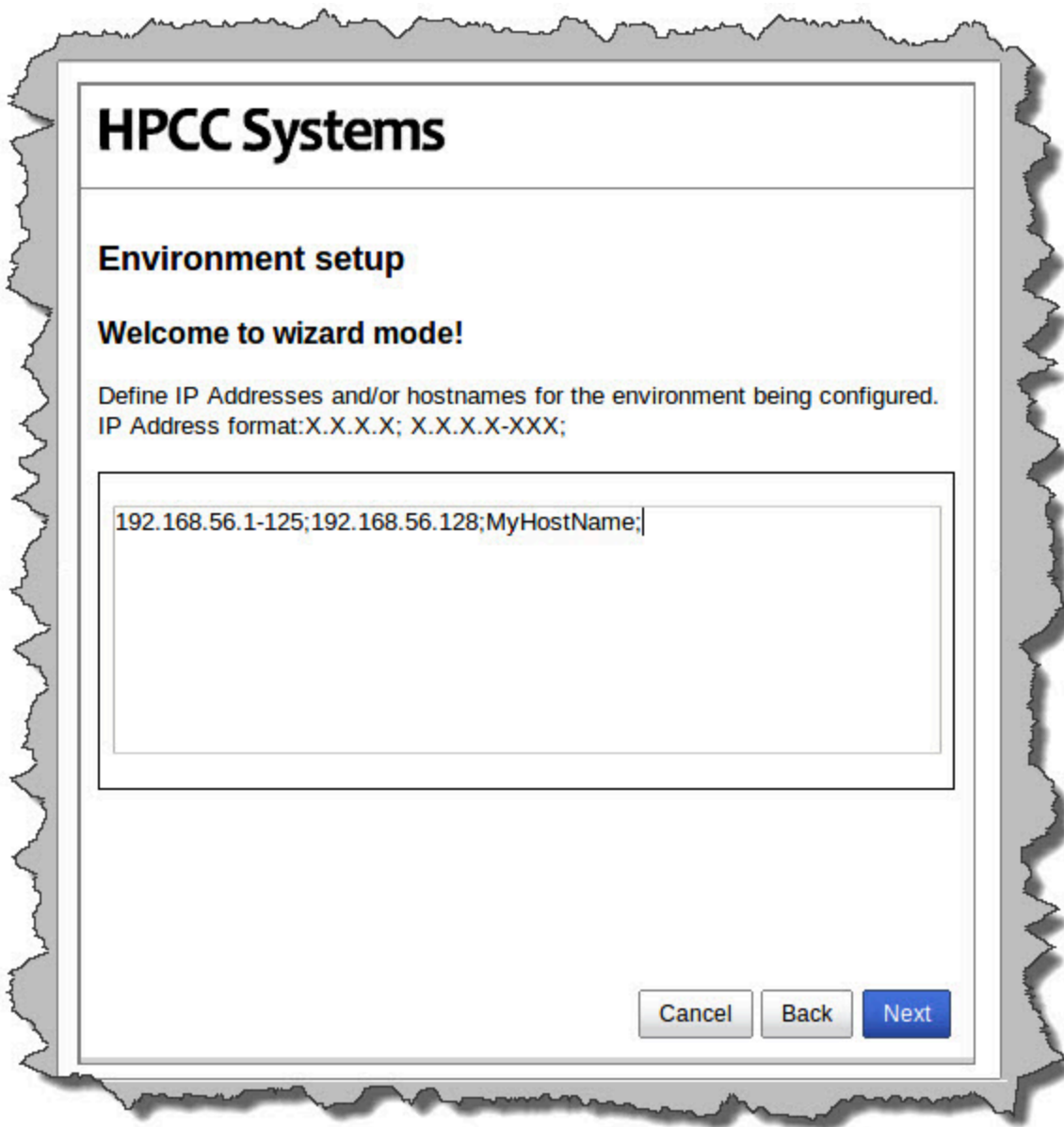
Esse será então o nome do arquivo XML de configuração. Por exemplo, vamos nomear nosso ambiente como *NewEnvironment* e isso criará um arquivo de configuração XML com o nome *NewEnvironment.xml* que iremos usar.

3. Pressione o botão Next.

Em seguida, você precisará definir os endereços IP que seu HPCC System usará.

4. Insira os endereços IP ou os nomes de host.

Os endereços IP podem ser especificados individualmente usando ponto e vírgula como separadores. Também é possível especificar um intervalo de IPs usando um hífen (por exemplo nnn.nnn.nnn.x-y). Na imagem abaixo, especificamos os endereços IP 10.239.219.1 até 10.239.219.100 usando a sintaxe de intervalo, além de um IP único (10.239.219.111).



HPCC Systems

Environment setup

Welcome to wizard mode!

Define IP Addresses and/or hostnames for the environment being configured.
IP Address format: X.X.X.X; X.X.X.X-XXX;

192.168.56.1-125;192.168.56.128;MyHostName;

Cancel Back Next

5. Pressione o botão Next.

Agora você vai definir quantos nós usar para os clusters Roxie e Thor.

6. Insira os valores adequados conforme indicado.

HPCC Systems

Environment setup

Enter number of nodes for Roxie and Thor clusters. No Roxie/Thor cluster will be generated for zero (0) number of nodes.

Number of support nodes	7
Number of nodes for Roxie cluster	20
Number of slave nodes for Thor cluster (A Thor Master will be added to the cluster and assigned to a support node)	100
Number of Thor slaves per node (default 1)	1
Enable Roxie on demand	<input checked="" type="checkbox"/>

Cancel Back Next

Número de nós de suporte.

Especifique o número de nós a serem usados para os componentes de suporte. O padrão é 1.

Número de nós do cluster Roxie:

Especifique o número de nós a serem usados para seu cluster Roxie. Insira zero (0) se você não quiser um cluster Roxie.

Número de nós escravos do cluster Thor

Especifique o número de nós escravos a serem usados para seu cluster Thor. Um nó mestre Thor será adicionado automaticamente. Insira zero (0) se você não quiser nenhum escravo Thor.

Número de escravos Thor por nó (padrão 1)

Especifique o número de processos de escravos Thor para instanciar em cada nó escravo. Insira zero (0) se você não quiser um cluster Thor.

Ativar o Roxie sob demanda

Especifique se você deseja permitir ou não que as consultas sejam executadas imediatamente no Roxie. Isso precisa ser ativado para executar o depurador. (O padrão é true)

7. Pressione o botão **Next** .

O assistente apresenta os parâmetros de configuração.

8. Pressione o botão **Finish** para aceitar esses valores ou o botão **Advanced View** para editar no modo avançado.



Você agora será notificado de que concluiu o assistente.



Neste ponto, você criou um arquivo com o nome NewEnvironment.xml no diretório `/etc/HPCCSystems/source`.



Lembre-se de que as configurações do HPCC podem variar de acordo com as suas necessidades. Por exemplo, você pode não precisar de um Roxie ou pode precisar de vários clusters Roxie menores. Além disso, em um sistema de produção [Thor] seria necessário assegurar que os nós do Thor e Roxie sejam dedicados e que não contenham nenhum outro processo em execução. Este documento visa mostrar como usar as ferramentas de configuração. Planejamento de capacidade e design de sistema estão disponíveis em um módulo de treinamento.

Distribuindo a Configuração

1. Para o HPCC System.

Se estiver em execução, pare o HPCC System (em todos os nós) usando um comando como este:

```
sudo /sbin/service hpcc-init stop
```

Observação: Você pode ter um sistema de múltiplos nós e um script personalizado como aquele ilustrado no Anexo do documento [Instalando e Executando a Plataforma HPCC](#) para iniciar e parar o seu sistema. Se este for o caso, use o comando adequado para parar o sistema em cada nó.



Certifique-se de que o HPCC não esteja em execução antes de tentar copiar o arquivo `environment.xml`.

2. Salve o arquivo `environment.xml` em um backup.

```
# For example  
sudo -u hpcc cp /etc/HPCCSystems/environment.xml /etc/HPCCSystems/source/environment-date.xml
```

Observação: O arquivo `environment.xml` do ambiente em execução está situado em seu **diretório** `/etc/HPCCSystems/`. O Gerenciador de Configurações funciona em arquivos no diretório `etc/HPCCSystems/source`. É necessário copiar deste local para criar um arquivo `environment.xml` ativo.

Também é possível optar por dar um nome mais descritivo ao arquivo do ambiente para ajudar a distinguir quaisquer diferenças.

Contar com arquivos do ambiente no controle de origem é uma boa maneira de arquivar suas configurações de ambiente.

3. Copie o novo arquivo `.xml` do diretório de origem para `/etc/HPCCSystems` e renomeie o arquivo para `environment.xml`.

```
# for example  
sudo -u hpcc cp /etc/HPCCSystems/source/NewEnvironment.xml /etc/HPCCSystems/environment.xml
```

4. Copie o arquivo `/etc/HPCCSystems/environment.xml` para `/etc/HPCCSystems/` em cada nó.

Você pode usar um script para forçar o arquivo XML para todos os nós. Consulte a seção *Scripts de exemplo* no documento anexo [Instalando e Executando a plataforma HPCC](#). É possível usar os scripts como um modelo para criar seu próprio script e copiar o arquivo `environment.xml` para todos os seus nós.

5. Reinicie a plataforma HPCC em todos os nós.

Environment.conf

Outro componente da configuração do HPCC System é o arquivo `environment.conf`. O `environment.conf` contém algumas definições globais que o gerenciador de configuração usa para configurar o HPCC System. Na maioria dos casos, os padrões são suficientes.



ATENÇÃO: Essas configurações são fundamentais para que o sistema possa operar de forma adequada. Apenas os administradores do HPCC com nível de especialista devem tentar alterar qualquer aspecto deste arquivo.

Por padrão, o arquivo `environment.conf` está localizado em:

```
/etc/HPCCSystems
```

O `environment.conf` é obrigatório na inicialização do HPCC. É onde o arquivo de ambiente do HPCC é definido.

```
/opt/HPCCSystems/environment.xml
```

Esse é também onde o diretório de trabalho é definido.

```
path=/opt/HPCCSystems
```

O diretório de trabalho é usado por vários recursos da aplicação e sua modificação pode causar complicações desnecessárias. Por padrão, a aplicação é instalada aqui e define vários recursos para este diretório.

O `environment.conf` padrão:

```
## Default environment configuration file for OpenHPCC

[DEFAULT]
configs=/etc/HPCCSystems
path=/opt/HPCCSystems
classpath=/opt/HPCCSystems/classes
runtime=/var/lib/HPCCSystems
lock=/var/lock/HPCCSystems
# Supported logging fields:
#   AUD,CLS,DET,MID,TIM,DAT,PID,TID,NOD,JOB,USE,SES,COD,MLT,MCT,NNT,COM,QUO,PFX,ALL,STD
logfields=TIM+DAT+MLT+MID+PID+TID+COD+QUO+PFX
pid=/var/run/HPCCSystems
log=/var/log/HPCCSystems
user=hpcc
group=hpcc
#umask=022
#nice=0
home=/home
environment=environment.xml
sourcedir=/etc/HPCCSystems/source
blockname=HPCCSystems
interface=*
# enable epoll method for notification events (true/false)
use_epoll=true
# allow kernel pagecache flushing where enabled (true/false)
allow_pgcache_flush=true
# report UDP network stats
udp_stats=true
mpStart=7101
mpEnd=7500
mpSoMaxConn=128
mpTraceLevel=0
# enable SSL for dafilesrv remote file access
#   (SSLNone/false | SSLOnly/true | SSLFirst | UnsecureFirst)
```

```
# Enabling requires setting the HPCCPassPhrase, HPCCCertFile, and HPCCPrivateKeyFile values
#dfsUseSSL=SSLNone

#Specify location of HPCC PKI public/private key files
# note: if HPCCPassPhrase specified it must be encrypted
#HPCCPassPhrase=
#HPCCCertificateFile=/home/hpcc/.ssh/cert.cer
#HPCCPublicKeyFile=/home/hpcc/.ssh/id_rsa.pub
#HPCCPrivateKeyFile=/home/hpcc/.ssh/id_rsa

jvmoptions=-XX:-UsePerfData
#Options to enable remote debugging of Java service or application
#jvmoptions=-XX:-UsePerfData -agentlib:jdwp=transport=dt_socket,server=y,suspend=n,address=2000
#JNI_PATH=/absolute/path/to/alternative/libjvm.so

# Although HPCC platform includes plugins for both Python2 and Python3,
# only one may be safely enabled at a time
# as the Python libraries export the same symbols for both versions.
# Enabling both may lead to unpredictable results
# including segfaults or undefined symbol errors.
#
# If you would prefer to use python 2 and disable python3, change the line below to read
# additionalPlugins=python2
#
# Multiple paths can be specified (separate with :, or ; on Windows).
# Relative paths are assumed to be relative to /opt/HPCCSystems/versioned
additionalPlugins=python3

# To en-/disable Drop Zone restriction.
# Default is enabled (true).
useDropZoneRestriction=true
# If set, will force matching local file paths to become remote reads, e.g:
#forceRemotePattern=/var/lib/HPCCSystems/hpcc-data/eclagent/*

# Dfilesrv: default client side connection settings (NB: 0 = disable/use system defaults)
#dafsConnectTimeoutSeconds=100
#dafsConnectRetries=2
#dafsMaxReceiveTimeSeconds=0

# Dfilesrv: set to change number of seconds before retrying an unresponsive dfilesrv
# (default 10 seconds)
# NB: for now this only applies to the last cached server
#dafsConnectFailRetrySeconds=10
```

Considerações sobre caminhos

A maioria dos diretórios é definida como caminhos absolutos:

```
configs=/etc/HPCCSystems
path=/opt/HPCCSystems
classpath=/opt/HPCCSystems/classes
runtime=/var/lib/HPCCSystems
lock=/var/lock/HPCCSystems
```

O HPCC não executará corretamente sem os caminhos adequados e, em alguns casos, é necessário ser o caminho absoluto. Se um processo ou componente não conseguir localizar um caminho, será exibida uma mensagem de erro como esta:

```
"There are no components configured to run on the node..."
```

Se o caminho mudar do HPCCSystems, ele NÃO será alterado no arquivo environment.xml. Quaisquer alterações precisarão ser realizadas manualmente no arquivo environment.xml.

O arquivo de log, *hpcc-init.log* é gravado no caminho do HPCCSystems.

Outros itens do Environment.conf

Alguns outros itens usados ou indicados no environment.conf

Use_epoll É um mecanismo de eventos para obter melhor desempenho em aplicações mais exigentes cujo número de descritores de arquivo assistidos é alto.

Logfields Categorias disponíveis para registro. Compostas por hora (TIM), data (DAT), ID de processo (PID), ID de linhas de execução ou "thread" (TID) e afins.

Interface No environment.conf padrão, há um valor para interface. O valor padrão para isso é:

```
interface=*
```

O valor padrão de* atribui à interface um endereço IP aberto em qualquer ordem. A especificação da interface, como Eth0, atribuirá o nó especificado como primário.

Acesso Remoto sobre TLS

A configuração do sistema para acesso remoto de arquivos por Transport Layer Security (TLS) exige a modificação das configurações do **dafilesrv** no arquivo *environment.conf*.

Para fazer isso, retire o comentário (se já estiver inserido) ou adicione as seguintes linhas ao arquivo *environment.conf*. Em seguida, defina os valores adequados para o seu sistema.

```
#enable SSL for dafilesrv remote file access
dfsUseSSL=true
dfsSSLCertFile=/certfilepath/certfile
dfsSSLPrivateKeyFile=/keyfilepath/keyfile
```

Defina o **dfsUseSSL=true** e o valor dos caminhos para indicar os caminhos do arquivo de certificado e do arquivo chave em seu sistema. Em seguida, implemente o arquivo *environment.conf* (e os arquivos de certificado/chave) em todos os nós conforme apropriado.

Quando o dafilesrv for ativado para TLS (porta 7600), ele ainda pode se conectar por uma conexão sem TLS (porta 7100) para permitir que clientes legados funcionem.


Configurando HPCC para Autenticação

Esta seção detalha as etapas para configurar a plataforma HPCC a usar autenticação. Atualmente existem algumas formas de usar a autenticação em seu HPCC Systems: autenticação simples `htpasswd`, LDAP, ou outro método de segurança de plugin.

O método de autenticação `htpasswd` constitui na autenticação simples da senha. Ele concede ou nega acesso a um usuário apenas com base na autenticação de senhas criptografadas por MD5.

Autenticação LDAP oferece mais recursos e opções. LDAP é capaz de autenticar usuários e de adicionar granularidade à autenticação. LDAP permite controlar acessos agrupados a recursos, funções e arquivos.

Você deve levar em conta as necessidades do seu sistema na hora de decidir qual desses métodos é o mais adequado para seu ambiente.

	<p>Ao implementar qualquer forma de autenticação, recomendamos ativar seu servidor ESP a usar HTTPS (SSL) e a configurar TODAS as conexões do serviço a usarem apenas HTTPS. Isso garante que as credenciais sejam transmitidas pela rede usando a criptografia SSL Veja Como configurar o ESP Server para usar HTTPS (SSL) para obter mais informações.</p> <p>Não se deve tentar isso até que o ambiente a ser usado já tenha sido implementado, configurado e certificado.</p>
---	--

Utilizando autenticação `htpasswd`

O modelo `htpasswd` oferece a autenticação simples de senhas para todo o sistema. Esta seção contém informações de instalação e de implementação do modelo de autenticação `htpasswd`.

Conectar ao Configuration Manager

Para alterar a configuração para os componentes do HPCC, conecte-se ao Configuration Manager.

1. Pare todos os componentes do HPCC se estiverem em execução.
2. Verifique se eles não estão mais sendo executados. É possível usar um único comando, como:

```
sudo /opt/HPCCSystems/sbin/hpcc-run.sh status
```

3. Inicie o Gerenciador de Configurações.

```
sudo /opt/HPCCSystems/sbin/configmgr
```

4. Conecte seu navegador de Internet à interface da Web do Gerenciador de Configurações.

(usando o URL `http://<configmgr_IP_Address>:8015`, onde `<configmgr_IP_Address>` é o endereço IP do nó que está executando o Configuration Manager)

5. Selecione o botão de opção **Advanced View**.
6. Use a lista suspensa para selecionar o arquivo de configuração XML adequado.

Observação: O Configuration Manager **nunca** atua no arquivo de configurações ativo. Após terminar a edição, será necessário copiar o arquivo `environment.xml` para o local ativo e forçá-lo a todos os nós.

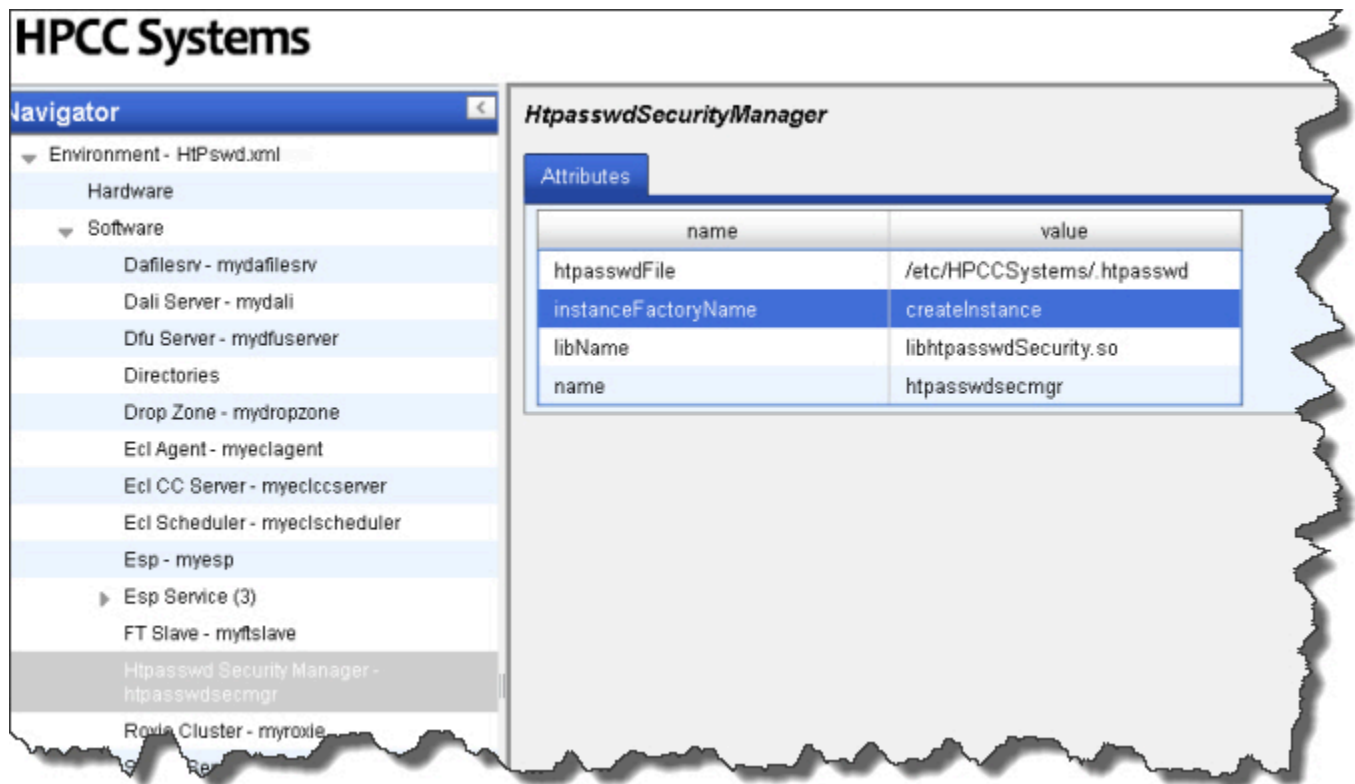
7. Marque a caixa de seleção **Write Access**.

O acesso padrão é somente leitura. Muitas opções estão disponíveis apenas quando o acesso à gravação estiver ativado.

Habilitando a autenticação httpasswd no HPCC

8. Crie uma instância do **Security Manager** Plugin:
 - a. Clique com o botão direito no Pannel de navegação ao lado esquerdo.
 - b. Selecione **New Components**.
 - c. Selecione o componente **httpasswdsecmgr**.
9. Configure o plugin do httpasswd

Figure 25. Página “Security Mgr Configuration” (Configuração do Security Manager)

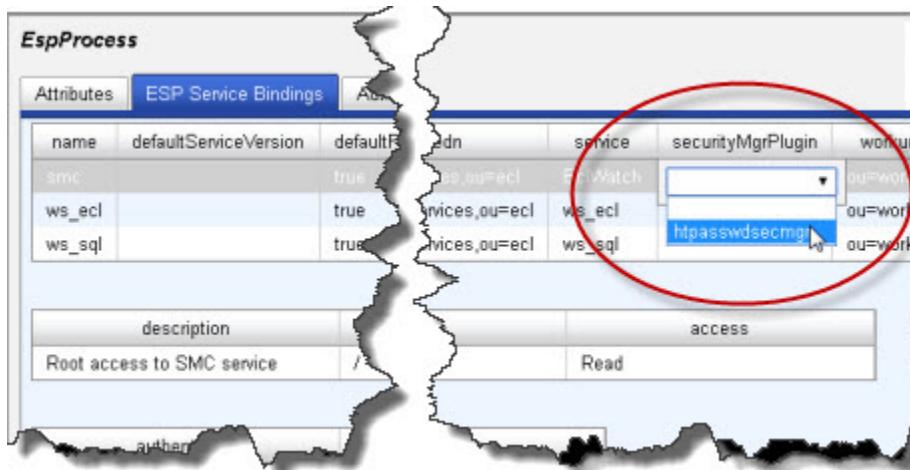


- a. Digite a localização do arquivo Htpasswd que contém o nome do usuário e a senha no sistema de arquivos Linux para **httpasswdFile**
 - b. **InstanceFactoryName** é o nome da função de fábrica do gerenciador de segurança implementado na biblioteca de segurança. O padrão é "createInstance". Use o padrão na implementação do método Htpasswd.
 - c. Forneça um nome da biblioteca para **libName**. Para Htpasswd, use [libhttpasswdSecurity.so](#)
 - d. Forneça um nome da instância para o valor **nome**. Por exemplo, [httpasswdsecmgr](#).
10. Selecione **Esp - myesp** no painel do navegador ao lado esquerdo.

Observação: Se tiver mais de um ESP Server, use apenas um deles para autenticação.

11. Associe o Security Manager Plugin às conexões do ESP.

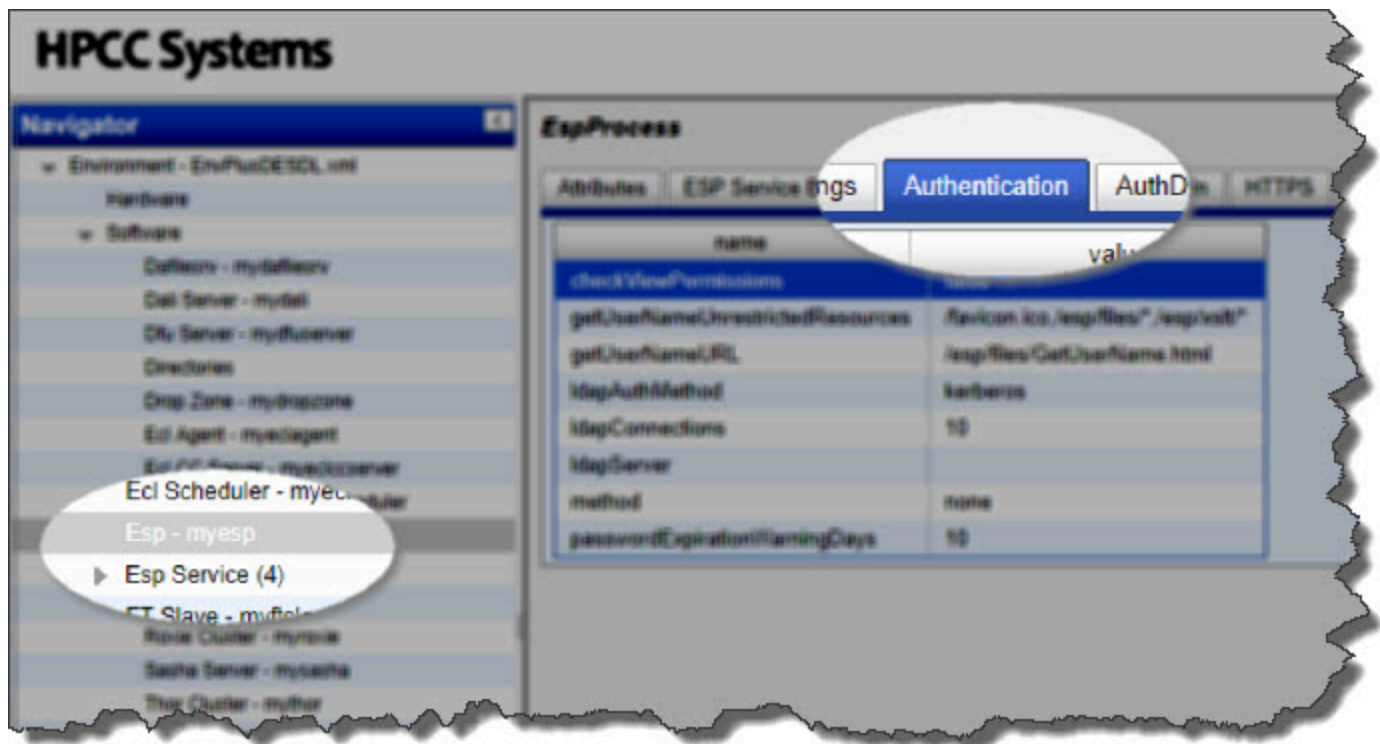
- a. Clique no **Esp** de destino no painel do navegador ao lado esquerdo.
- b. Selecione a **aba de conexões do ESP Service**
- c. Nas ligações de destino, selecione a instância securityMgrPlugin adequada a partir da lista suspensa.



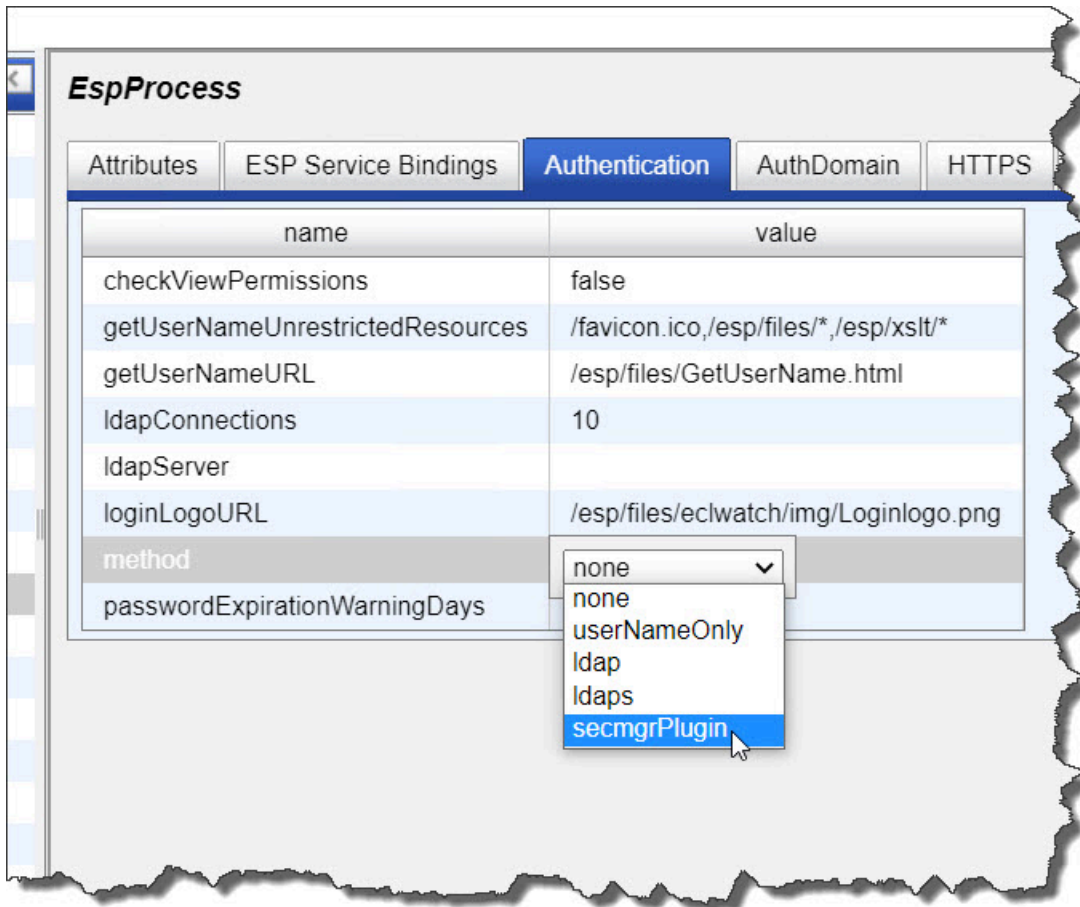
12. Selecione um plugin de segurança para cada serviço que exija um gerenciador de segurança.

Por exemplo, na imagem acima selecione [httpasswdsecmgr](#) para o serviço smc. Em seguida, selecione para ws_ecl e para qualquer outro serviço que deseje usar a segurança httpasswd.

13. Selecione a aba **Authentication**.



14. Clique na lista suspensa da coluna para exibir as opções do **method**.



15. Selecione **secmgrPlugin** na lista suspensa.

16. Clique no ícone de disco para salvar.

Usuário administrador com htpasswd

Usuários e senhas são mantidos no arquivo htpasswd. O arquivo htpasswd deve existir no nó do ESP onde a autenticação está habilitada. HPCC apenas reconhece senhas criptografadas em MD5.

O local padrão é: **/etc/HPCCSystems/.htpasswd** no nó do ESP configurado para autenticação, porém pode também ser configurado no Gerenciador de segurança do Htpasswd como destacado acima (etapa 9).

Você pode usar o utilitário do htpasswd para criar um arquivo de extensão .htpasswd para administrar usuários.

Pode ser que o utilitário do htpasswd já esteja instalado em seu sistema, uma vez que ele faz parte de alguns sistemas Linux. Verifique seu sistema Linux para ver se o utilitário já está instalado. Se não tiver, baixe o utilitário para seu sistema no The Apache Software Foundation.

Para obter mais informações sobre como usar o htpasswd acesse: <http://httpd.apache.org/docs/2.2/programs/htpasswd.html>.

Utilizando Autenticação LDAP

Esta seção contém informações de instalação e de implementação da autenticação baseada em LDAP. A autenticação LDAP oferece o maior número de opções para proteger o seu sistema ou partes de seu sistema. Além dessas definições de configuração, você precisa executar o utilitário **initldap** para criar o usuário padrão Admin do HPCC requerido em seu servidor LDAP.

Se optar por usar a autenticação LDAP, você precisa habilitar o LDAP security em sua configuração do HPCC System. Com a função LDAP security habilitada em seu sistema, você pode optar por ativar a segurança do escopo de arquivos. Há a opção de usar a autenticação LDAP sem habilitar a segurança do escopo de arquivos. As seções a seguir descrevem como habilitar a autenticação LDAP e a segurança do escopo de arquivos em seu HPCC System.

Conectar-se ao Configuration Manager

Para alterar a configuração para os componentes do HPCC, conecte-se ao Configuration Manager.

1. Pare todos os componentes do HPCC se estiverem em execução.
2. Verifique se eles não estão mais sendo executados. É possível usar um comando único, como:

```
sudo /opt/HPCCSystems/sbin/hpcc-run.sh -a hpcc-init status
```

3. Inicie o Gerenciador de Configurações.

```
sudo /opt/HPCCSystems/sbin/configmgr
```

4. Conecte à interface Web do Configuration Manager.

(usando o URL `http://<configmgr_IP_Address>:8015`, where `<configmgr_IP_Address>` é o endereço IP do nó que está executando o Configuration Manager)

5. Selecione o botão de opção **Advanced View**.
6. Use a lista suspensa para selecionar o arquivo de configuração XML adequado.

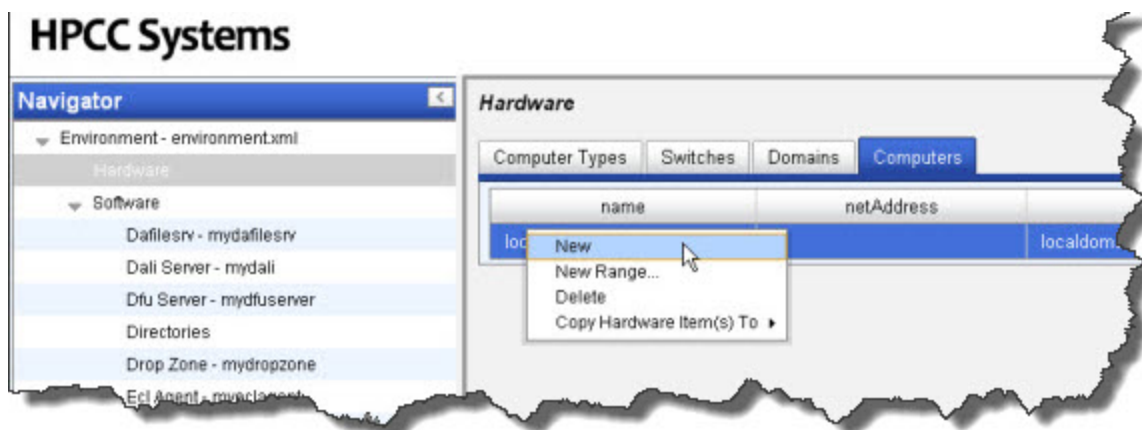
Observação: O Configuration Manager **nunca** atua no arquivo de configurações ativo. Após terminar a edição, será necessário copiar o arquivo `environment.xml` para o local ativo e distribuí-lo a todos os nós.

Modificando a Configuração

Siga as etapas abaixo para modificar sua configuração.

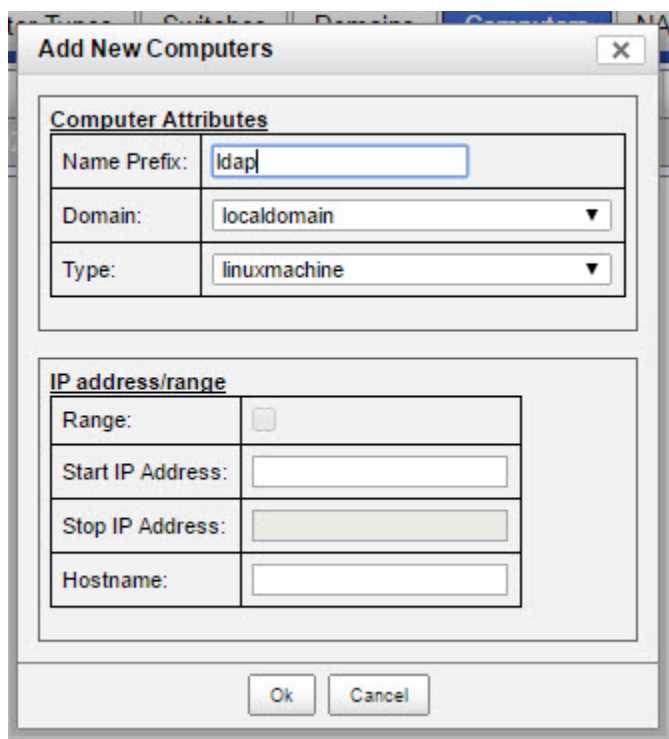
1. Marque a caixa de seleção **Write Access**.
2. No painel do **Navigator**, selecione **Hardware**.
3. Selecione a aba **Computers** no painel à direita.

4. Clique com o botão direito na tabela abaixo de computers e selecione a opção **New** no menu pop-up.



A caixa de diálogo **Add New Computers** será exibida.

5. Preencha a área de **Computer Attributes**



- a. Forneça um **Name Prefix**, como por exemplo: [ldap](#).

Isso ajudará a identificá-lo na lista de computadores.

- b. Preencha as informações de **Domain** e **Type** com o nome do seu domínio e os tipos de máquinas que você está usando.

No exemplo acima, o **Domain** é [localdomain](#), e o **Type** é [linuxmachine](#). Estes devem corresponder ao seu domínio e tipo.

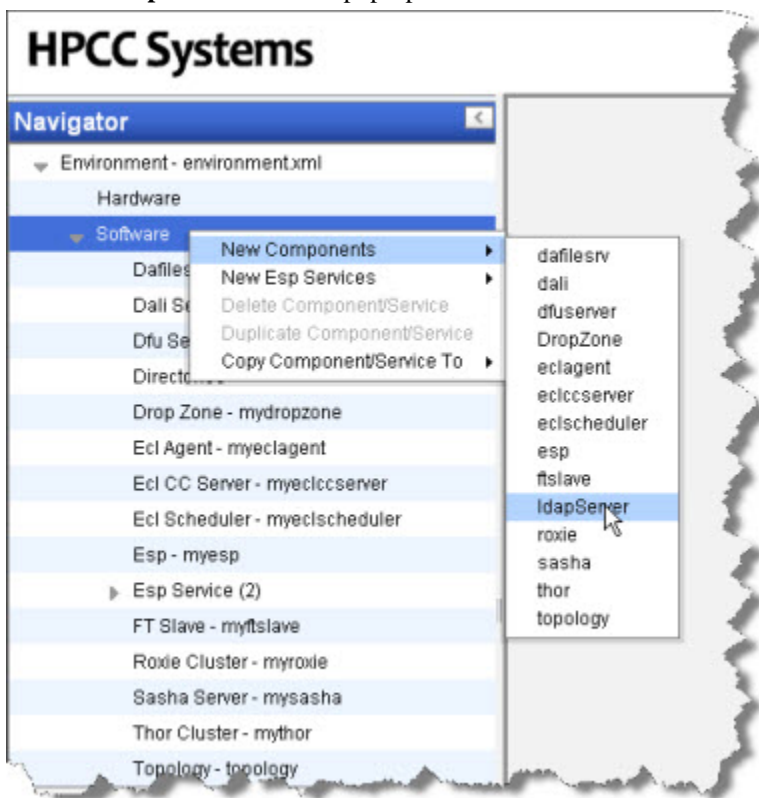
Se for preciso adicionar um novo domínio ou tipo de máquina ao seu sistema para poder definir um servidor LDAP existente, primeiramente é necessário configurar isso nas outras duas abas na seção Hardware.

- c. Adicione o endereço IP como apropriado para o servidor LDAP.
- d. Pressione o botão **Ok**.
- e. Clique no ícone de disco para salvar.

Adicionando o componente IdapServer

Após o nó do LDAP Server ter sido adicionado às configurações de Hardware, configure a definição de Software do servidor LDAP.

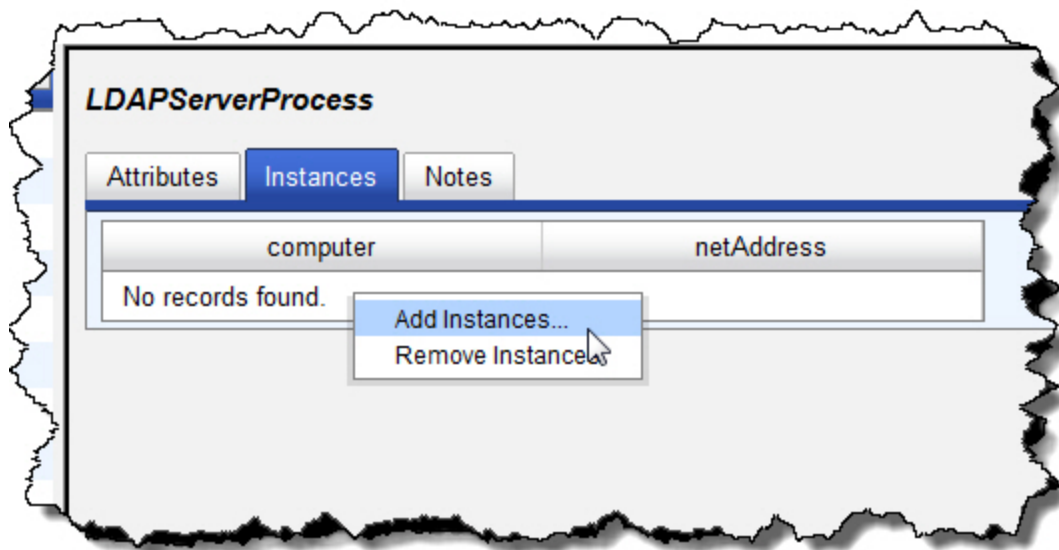
1. Clique com o botão direito no painel **Navigator** e selecione **New Components** no menu pop-up; em seguida, selecione **IdapServer** no menu pop-up.



Observação: O componente IdapServer é meramente uma definição que especifica um servidor LDAP existente. Ele não instala um servidor.

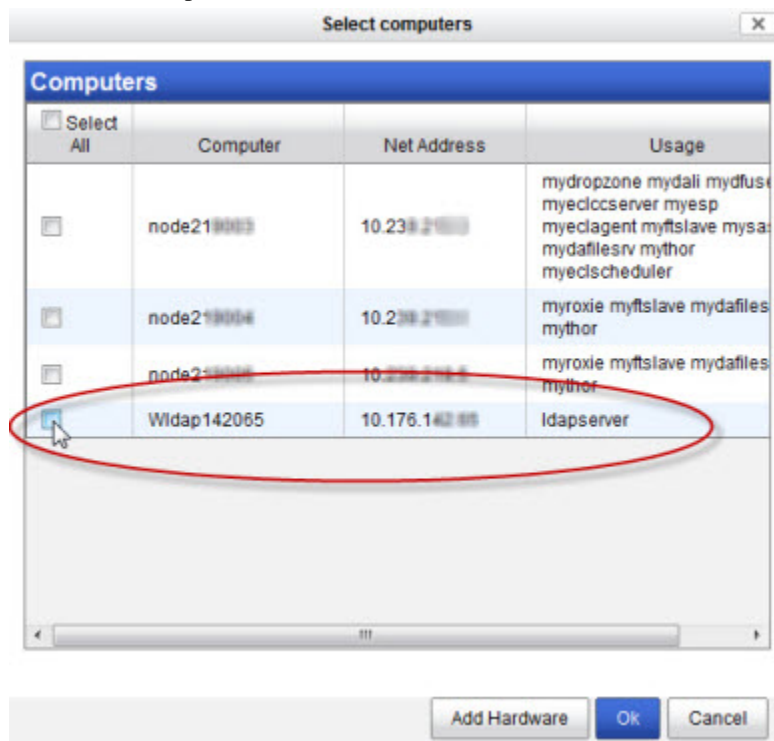
2. Preencha as **propriedades do LDAP Server Process**:

- a. Na aba **Instances** , clique com o botão direito na tabela à direita e selecione **Add Instances...**



A caixa de diálogo **Select Computers** aparecerá.

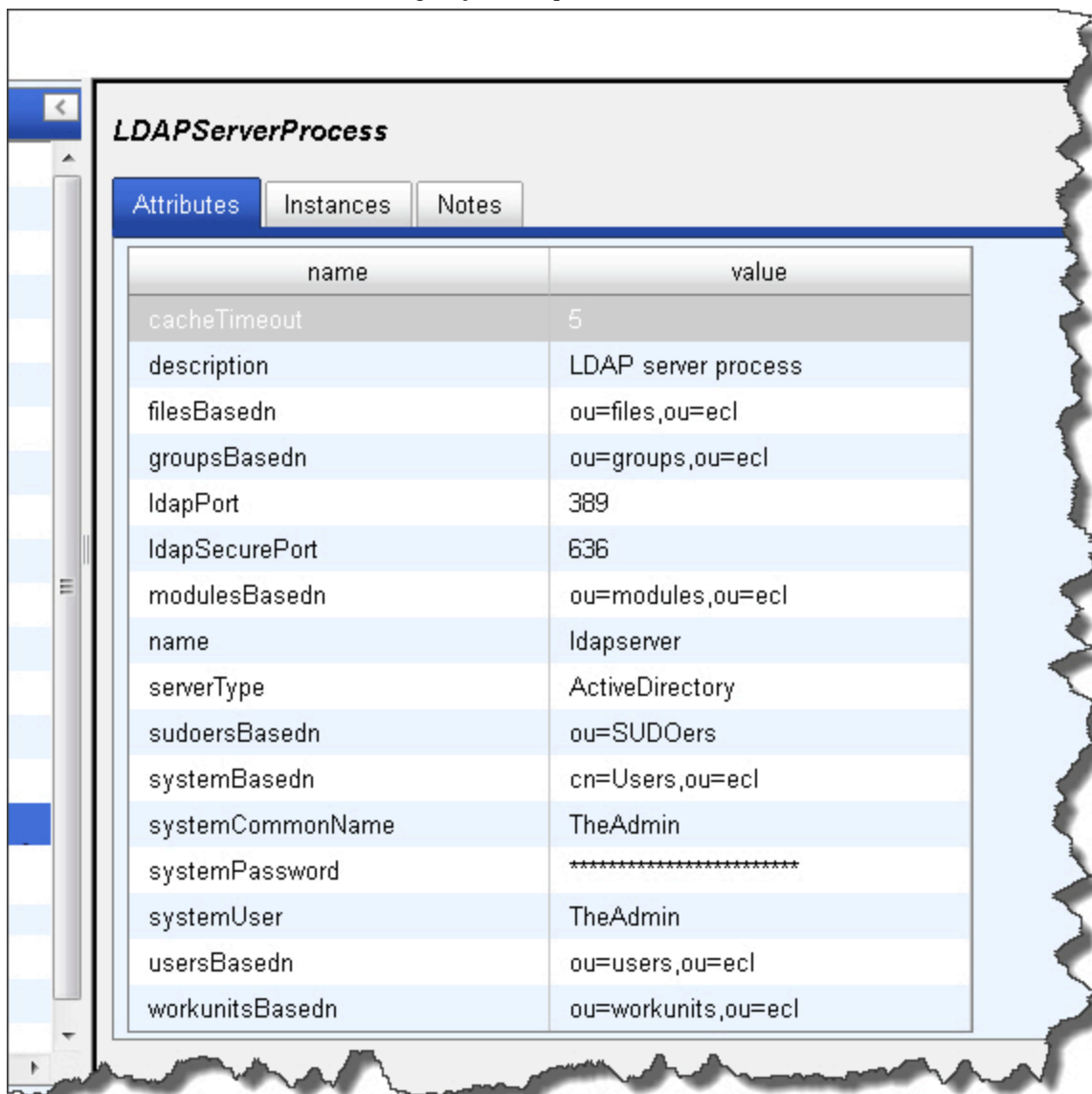
- b. Selecione o computador a ser usado clicando na caixa ao lado dele.



Este é o computador que foi adicionado anteriormente na parte **Hardware / Add New Computer** .

- c. Pressione o botão **OK** .

- d. Preencha a aba **Attributes** com as configurações adequadas de seu LDAP Server existente.



The screenshot shows a configuration window titled "LDAPServerProcess". It has three tabs: "Attributes" (selected), "Instances", and "Notes". Below the tabs is a table with two columns: "name" and "value". The table contains the following configuration data:

name	value
cacheTimeout	5
description	LDAP server process
filesBasedn	ou=files,ou=ecl
groupsBasedn	ou=groups,ou=ecl
ldapPort	389
ldapSecurePort	636
modulesBasedn	ou=modules,ou=ecl
name	ldapserver
serverType	ActiveDirectory
sudoersBasedn	ou=SUDOers
systemBasedn	cn=Users,ou=ecl
systemCommonName	TheAdmin
systemPassword	*****
systemUser	TheAdmin
usersBasedn	ou=users,ou=ecl
workunitsBasedn	ou=workunits,ou=ecl

- e. Selecione tipo de servidor LDAP no atributo serverType da caixa suspensa.



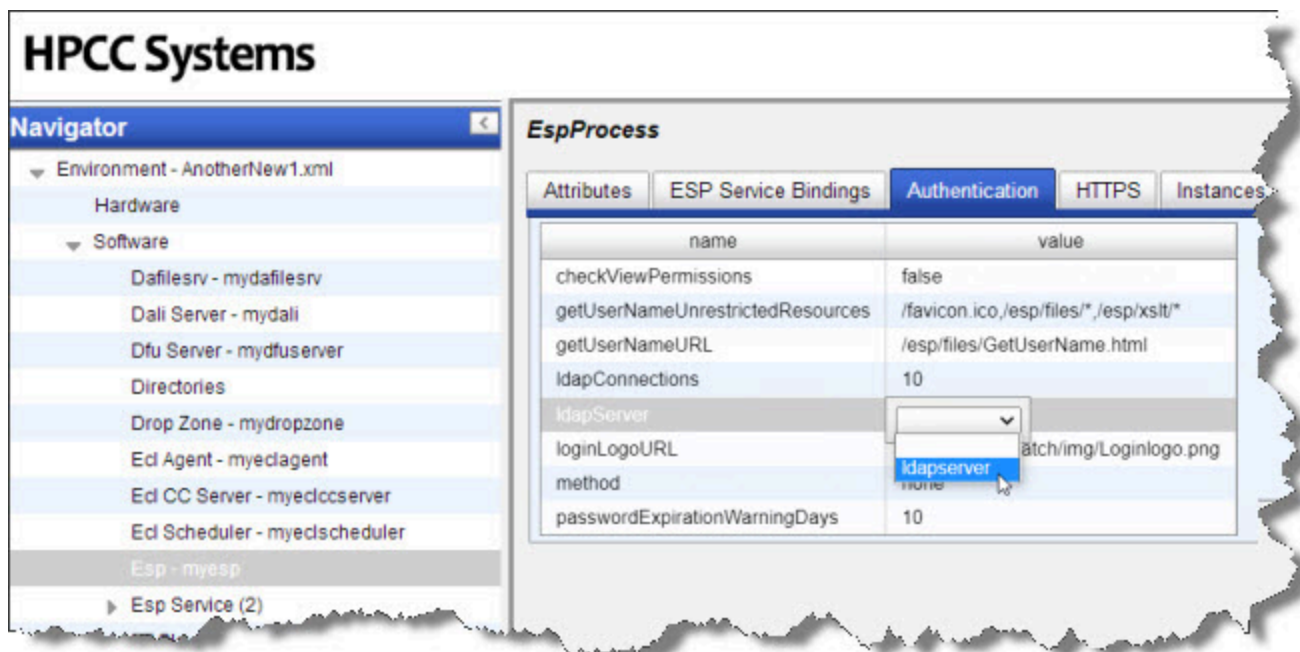
OBSERVAÇÃO: O suporte para OpenLDAP foi descontinuado. Esta opção foi incluída apenas para necessidades legadas.

- f. Clique no ícone de disco para salvar.

Observação: O valor do **cacheTimeout** corresponde ao número de minutos em que as permissões estão em cache no ESP. Ao alterar qualquer permissão no LDAP, as novas configurações não estarão em vigor até que o ESP e o Dali atualizem. Isso pode demorar a mesma quantidade de tempo do cacheTimeout. A definição disso para 0 significa sem cache, porém sobrecarrega o desempenho, assim não deve ser usado em produção.

3. No painel do navegador, clique em **ESP -- myesp**

4. Na página **EspProcess** ao lado direito, selecione a aba **Authentication** .

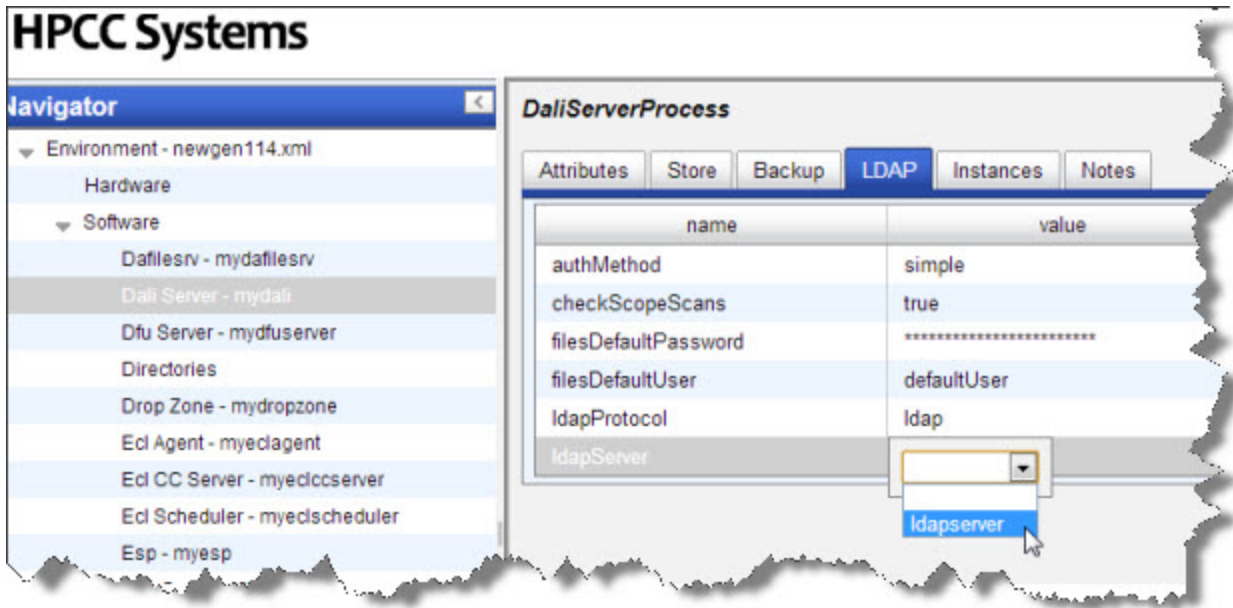


Preencha as informações adequadas:

- Altere o **ldapConnections** para o número adequado ao seu sistema (o número 10 é apenas um exemplo).
- Selecione o componente **ldapServer** adicionado anteriormente da lista suspensa, por exemplo: [ldapserver](#).
- Altere a informação do **método** para [ldap](#).
- Selecione a aba **ESP Service Bindings**. Certifique-se de que as configurações do LDAP apareçam em **resources-Basedn** e **workunitsBasedn**
- Clique no ícone de disco para salvar.

5. Para habilitar as permissões do escopo de arquivos, realize a configuração no servidor Dali.

No painel do navegador, clique em **Dali Server -- mydali**



Preencha com as informações apropriadas:

- Selecione a aba **LDAP**.
- Altere o **authMethod** para **simple** (simples).
- Defina o **checkScopeScans** para **true** (verdadeiro).

Defina esse campo para “true” apenas quando quiser habilitar a segurança do escopo de arquivos. As configurações de segurança podem ter três estados.

- Nenhum, sem autenticação e sem segurança do escopo de arquivos.
 - LDAP segurança apenas para autenticação, sem habilitar a segurança do escopo de arquivos.
 - LDAP autenticação e segurança do escopo de arquivos habilitados.
- Altere as informações do LDAP como apropriado para que correspondam às configurações do componente de seu servidor LDAP no Configuration Manager.

Exemplo: altere o **ldapServer** para o mesmo valor do seu LDAP Server. Ness caso, o valor é: *ldapservice*.

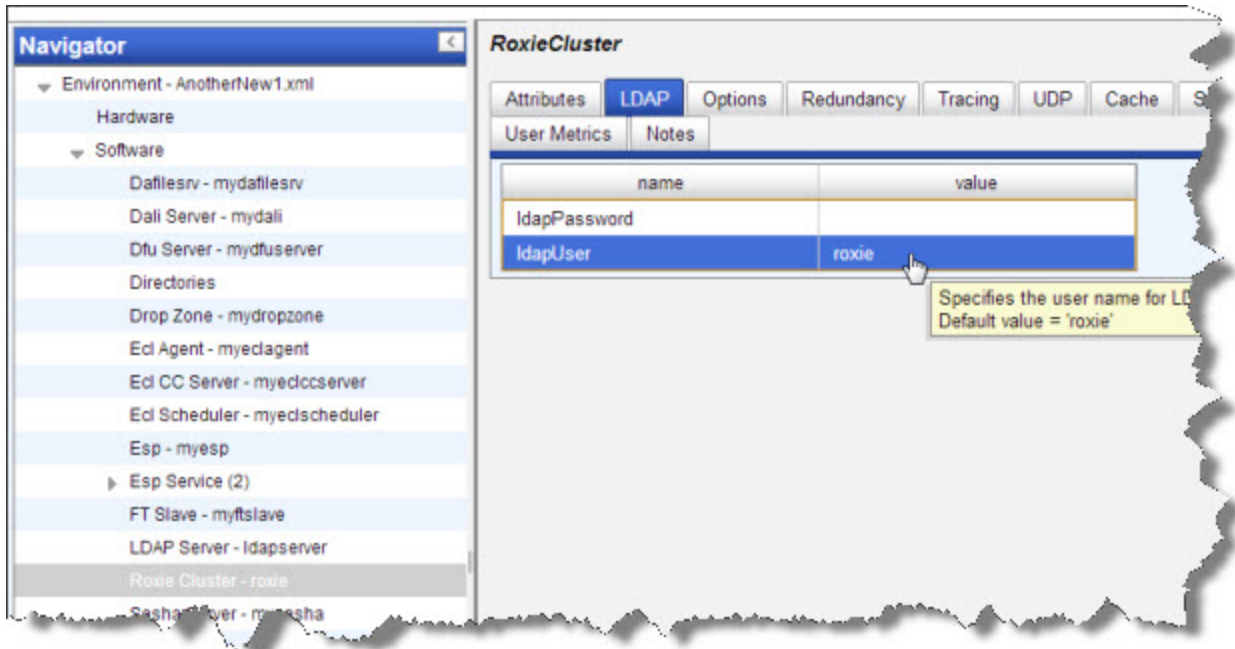
Confirme a alteração quando solicitado.

O **filesDefaultUser** é uma conta do LDAP usada para acessar arquivos quando nenhuma não há nenhuma credencial. É semelhante a conta “convidado”, por isso deve ter acesso **bastante** limitado em seu uso. Deixe o **filesDefaultUser** em branco para desabilitar esse tipo de acesso.

O **filesDefaultPassword** corresponde a senha dessa conta.

- Clique no ícone de disco para salvar.

6. No painel do navegador, clique em **Roxie Cluster -- myroxie**



- Na página **RoxieCluster** ao lado direito, selecione a aba **LDAP**.
- Localize o campo **ldapUser** e verifique se há um usuário válido do HPCC que seja membro do grupo de Usuários autenticados em seu servidor LDAP. Por exemplo, o usuário "roxie" assume que usuário "roxie" é um usuário autenticado válido do HPCC.
- Adicione a segurança de senha para Roxie, adicionando-a ao campo **ldapPassword** na mesma guia.

Para executar consultas no Roxie através da segurança do escopo de arquivos, verifique se um usuário do Roxie foi criado na lista de usuários autenticados.
Na seção seguinte, *Adicionar e editar usuários*, adicione o usuário *roxie* e verifique se a senha é a mesma que foi inserida no Configuration Manager.

Instalando o usuário de Admin padrão

Após habilitar suas configurações do LDAP Security, é preciso copiar seu arquivo de ambiente para o diretório /etc/HPCCSystems. Ver a seção *Como configurar um sistema de múltiplos nós* para obter mais informações sobre como configurar seu sistema. Com o arquivo environment.xml correto em vigor, é preciso executar o utilitário **initldap** para inicializar os componentes de segurança e os usuários padrão.

O utilitário initldap

O utilitário **initldap** cria a conta de usuário de Administrador do HPCC e as OUs do HPCC para um servidor LDAP recém-definido. O utilitário **initldap** extrai essas configurações dos componentes do LDAP Server no environment.xml ligado aos ESPs configurados.

Você pode executar o utilitário **initldap** após ter concluído a configuração com componentes do LDAP ativados e depois de ter distribuído o arquivo environment.xml para todos os nós.

```
sudo /opt/HPCCSystems/bin/initldap
```

O utilitário **initldap** solicitará as credenciais de administrador do LDAP. Insira os valores apropriados quando solicitado.

Segue abaixo um exemplo de initldap na implementação do 389DirectoryServer.

```
Enter the '389DirectoryServer' LDAP Admin User name on '10.123.456.78'...Directory Manager
Enter the LDAP Admin user 'Directory Manager' password...*****

Ready to initialize HPCC LDAP Environment, using the following settings
  LDAP Server      : 10.123.456.78
  LDAP Type        : 389DirectoryServer
  HPCC Admin User  : HPCCAdmin389
Proceed? y/n
```

Utilizando a ferramenta addScopes

Quando uma nova conta de usuário do ESP é criada, um escopo de arquivo privado "hpccinternal::<user>" também é criado concedendo aos novos usuários o acesso total àquele escopo e acesso restrito aos outros usuários. Este escopo de arquivo é usado para armazenar temporariamente arquivos do HPCC como os arquivos de despejo e temporário.

Se você estiver habilitando a segurança do escopo de arquivos do LDAP e já tiver contas de usuários, execute o programa de utilitário addScopes para criar um escopo hpccinternal::<user> para esses usuários existentes.

Usuários que já pertençam a esse escopo são ignorados o que permite o uso seguro dessa solução tanto em contas de usuários ESP novas como pré-existentes.

A ferramenta está localizada na pasta **/opt/HPCCSystems/bin/** e, para executá-la, é preciso especificar a localização do **daliconf.xml**, por exemplo:

```
/opt/HPCCSystems/bin/addScopes /var/lib/HPCCSystems/mydali/daliconf.xml
```

Execute o addScopes no nó do Dali.

Manutenção de Segurança do Usuário

Configurar o HPCC System para usar o Active Directory ou segurança baseada em LDAP permite definir permissões para o controle de acesso aos Recursos, Escopo de arquivos e Escopos da Workunit.

Introdução

HPCC systems® preserva sua segurança de diversas formas. HPCC Systems® pode ser configurado para gerenciar os direitos de segurança dos usuários direcionando para o Active Directory da Microsoft no sistema Windows ou para o 389Directory Server no sistema Linux.

Ao usar a interface Permissões no ECL Watch, os administradores podem controlar o acesso aos recursos no ECL IDE, ECL Watch, ECL Plus, DFU Plus, e nos módulos ECL no atributo Repositório. Você também pode optar por implementar o controle de acesso a arquivos e workunit habilitando essa configuração no servidor Dali.

Estabeleça as permissões por grupo ou por usuário e determine-as por associação com um recurso específico do HPCC System. As permissões podem ser determinadas para cada combinação única de um grupo e de um recurso. As permissões são divididas entre as seguintes categorias:

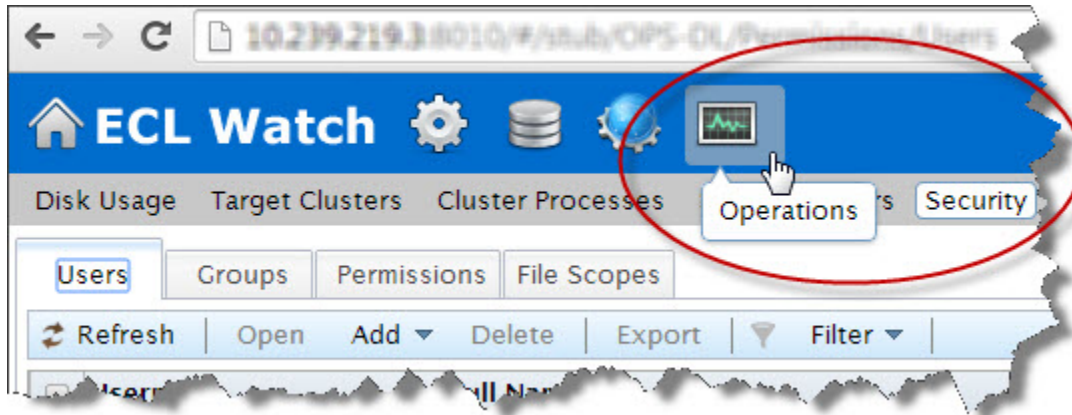
Esp Features for SMC	Controla o acesso a recursos no ECLWatch e a recursos similares acessados a partir do ECL IDE.
Esp Features for WsEclAccess	Controla o acesso ao serviço Web WS-ECL
Esp Features for EclDirectAccess	Controla o acesso ao serviço Web do ECLDirect
File Scopes	Controla o acesso aos arquivos de dados aplicando permissões aos escopo de arquivos
Workunit Scopes	Controla o acesso às workunits aplicando permissões aos Escopos de tarefa
Repository Modules	Controla o acesso ao atributo Repositório e aos módulos no repositório (antigo)

Administração de Segurança utilizando o ECL Watch

É preciso ter direitos de administrador para administrar as permissões. Após obter direitos de administrador, abra o ECL Watch em seu navegador usando o seguinte URL:

- **<http://nnn.nnn.nnn.nnn:pppp> (onde nnn.nnn.nnn.nnn é o endereço IP do seu ESP Server e pppp é a porta. A porta padrão é 8010).**

A administração da segurança é controlada através da área **Security** do ECL Watch. Para acessar a área de Security, clique no ícone **Operations**, e em seguida clique no link **Security** a partir do submenu de navegação.



As três áreas nas quais as permissões devem ser definidas são:

- **Users.** Mostra a configuração atual de todos os usuários. Use esta área para adicionar ou remover um usuário, editar as informações do usuário, definir/redefinir a senha do usuário e visualizar as permissões que estão atualmente atribuídas para o usuário.
- **Groups.** Mostra a configuração atual de todos os grupos. Use esta área para adicionar ou remover um grupo, visualizar e editar os membros do grupo, visualizar e editar as permissões que foram determinadas para o grupo.
- **Permissions.** Mostra os recursos do HPCC System onde as permissões devem ser determinadas. Use esta área para visualizar as permissões atualmente determinadas para qualquer área do HPCC System, para adicionar grupos e usuários e para definir ou modificar permissões em relação a um recurso específico.



OBSERVAÇÃO: É preciso ter cautela ao determinar qualquer configuração de permissão para **negar um direito**. A permissão mais restritiva sempre se aplica.

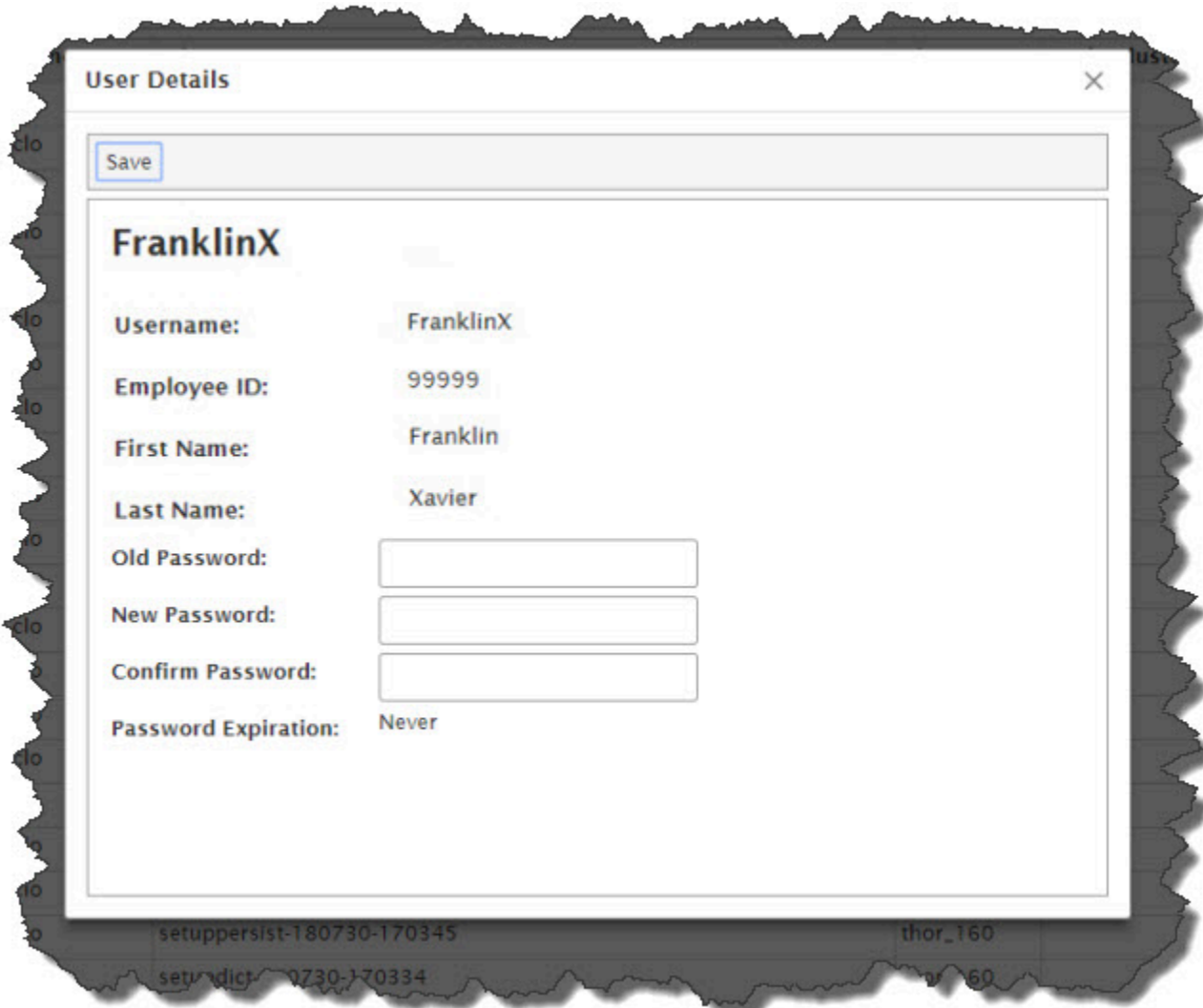
Informações sobre sua conta

Para obter mais informações sobre sua conta no ECL Watch, clique no link **LOGGED IN AS:** localizado no topo da página do ECL Watch .



1. Clique no link **LOGGED IN AS:**

A aba User Details será exibida com as informações de sua conta.



2. Confirme o User Name que você usou para entrar no sistema.

Observe que são necessários direitos de administrador para gerenciar usuários e permissões.

Verifique se você está usando uma conta com direitos de administrador se precisar gerenciar usuários ou permissões.

3. Verifique a data de validade da senha ou se a senha está prestes a expirar.

Se desejar, você também pode mudar sua senha aqui.

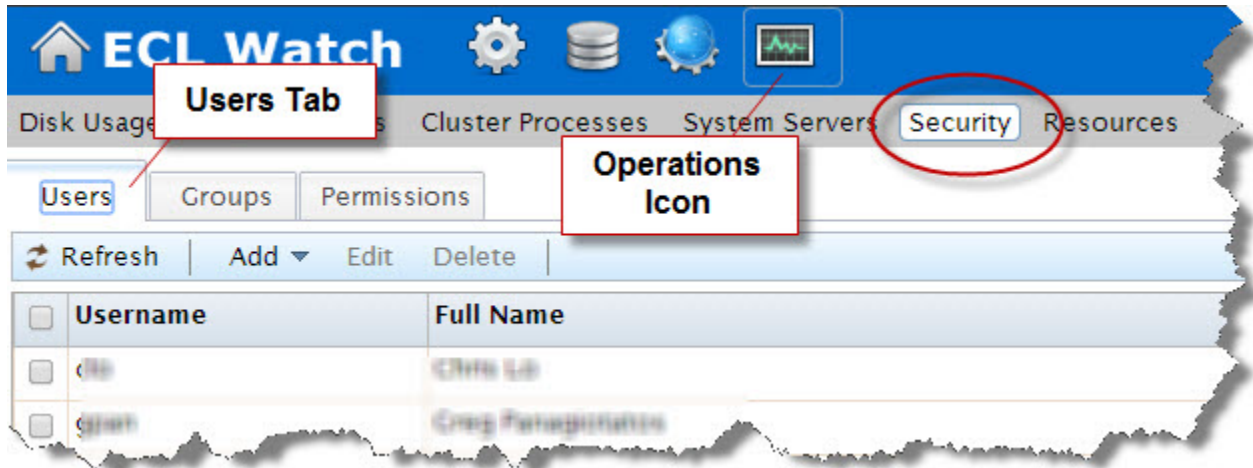
Configurando e modificando permissões de usuários

Em um ambiente habilitado para segurança, o acesso ao ECL Watch e seus recursos é controlado com o uso de um login e senha. A área **Users** permite controlar quem acessa o ECL Watch e os recursos do seu HPCC System para os quais esses usuários têm acesso. As permissões dos usuários podem ser definidas com base nas necessidades individuais de cada usuário, e também é possível adicionar usuários aos grupos que já tenham sido configurados. Use o item **Users** do menu para:

- Adicionar um novo usuário (**observação:** o Username não pode ser alterado)
- Remover um usuário
- Adicionar o usuário a um grupo
- Alterar a senha do usuário
- Modificar os detalhes ou as permissões de um usuário

Adicionando e editando usuários

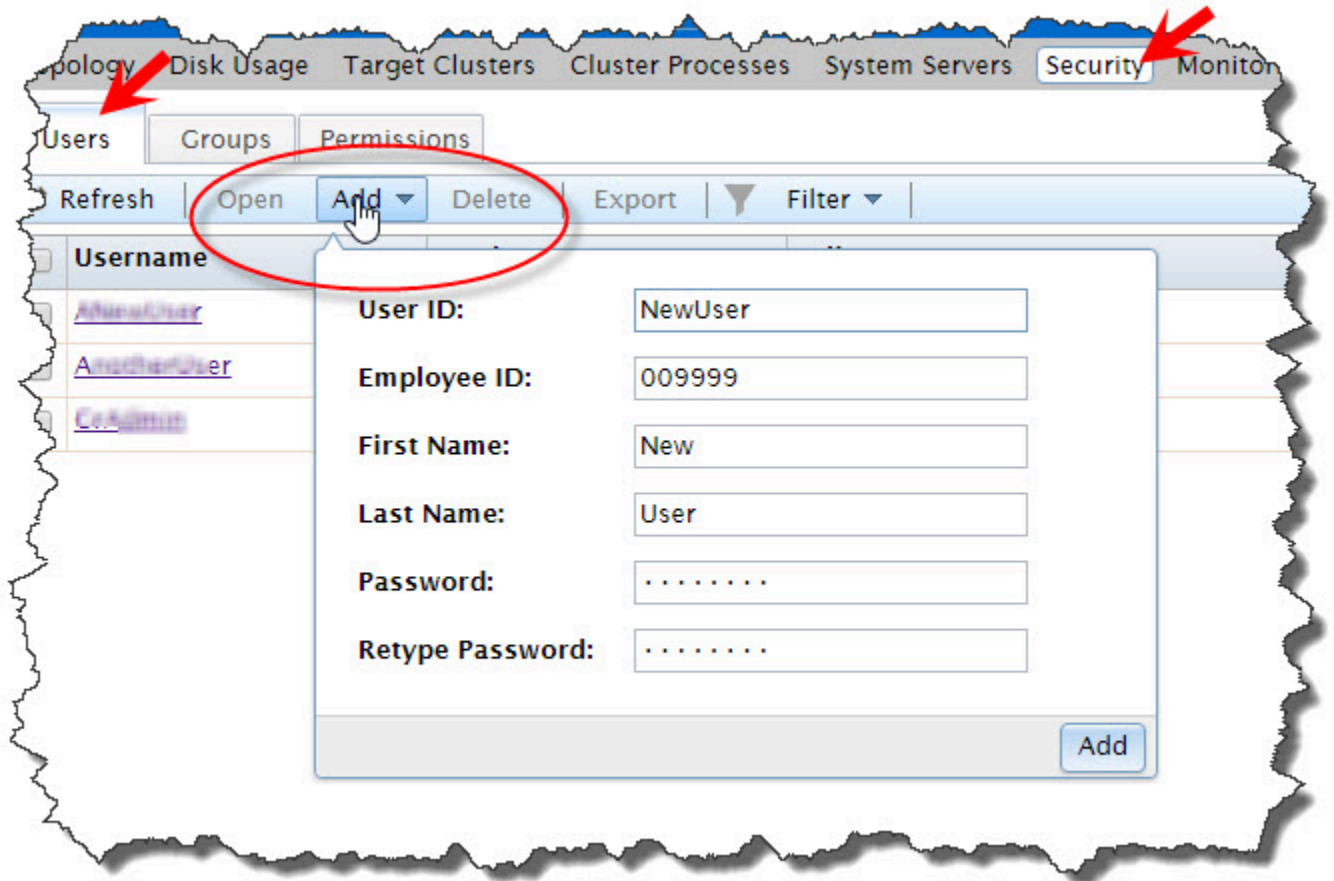
Para acessar as seções de administração do usuário, clique no ícone **Operations** e em seguida clique no link **Security** no submenu de navegação. Clique na guia **Users** para adicionar ou editar usuários.



Todos os usuários atuais são identificados na lista pelo seu Username e Full Name.

Para adicionar um novo usuário a lista de usuários autenticados:

Você precisa ter privilégio de administrador para adicionar um novo usuário.



1. Pressione o botão **Add** .

A caixa de diálogo Adicionar usuário será exibida.

2. Insira um **User ID**.

Este é o login que será usado no ECL Watch, ECL IDE, WsECL, etc.

3. Insira o **First Name** e o **Last Name** do usuário.

Estas informações ajudam a identificar o usuário e são exibidas no campo **Full Name** na janela principal **User** .

4. Insira uma **senha** para o usuário e confirme-a no campo **Retype Password** .

OBSERVAÇÃO: A senha deve estar em conformidade com a política do servidor do gerenciador de segurança.

5. Pressione o botão **Add** .

Após ter adicionado essas informações com sucesso, uma nova guia será aberta para que você possa verificar as informações do novo usuário.

6. Pressione o botão **Save** .

Após ter sido adicionado, o novo usuário será exibido na lista e você poderá modificar os detalhes e determinar as permissões conforme exigido.

Modificar detalhes do usuário:

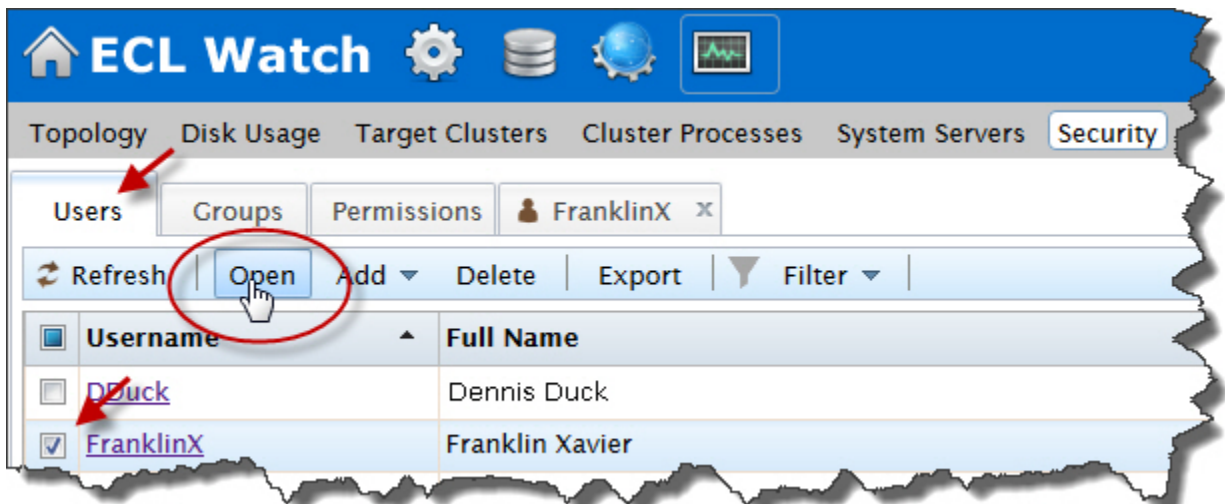
No submenu de navegação, clique no ícone **Operations** e em seguida no link **Security** .

1. Clique na **aba Users**.

Os usuários serão exibidos em uma lista.

2. Selecione o(s) usuário(s) que deseja modificar. Clique no link **Username** para abrir a aba de detalhes do usuário.

Para selecionar vários usuários, marque a caixa de seleção ao lado do User Name. Isso ativará os botões de ação para Users. Pressione o botão de ação **Open** .



Uma aba será aberta para cada usuário selecionado. Cada aba de usuário contém várias sub-abas.

Os detalhes do usuário estão localizados na aba **Summary**.

3. Modifique os detalhes do usuário como solicitado (caso tenha selecionado mais de um usuário, repita a operação para cada um deles).

Observação: O **Username** não pode ser alterado.

4. Pressione o botão **Save**.

Uma mensagem de confirmação será exibida.

Para adicionar um usuário para um grupo:

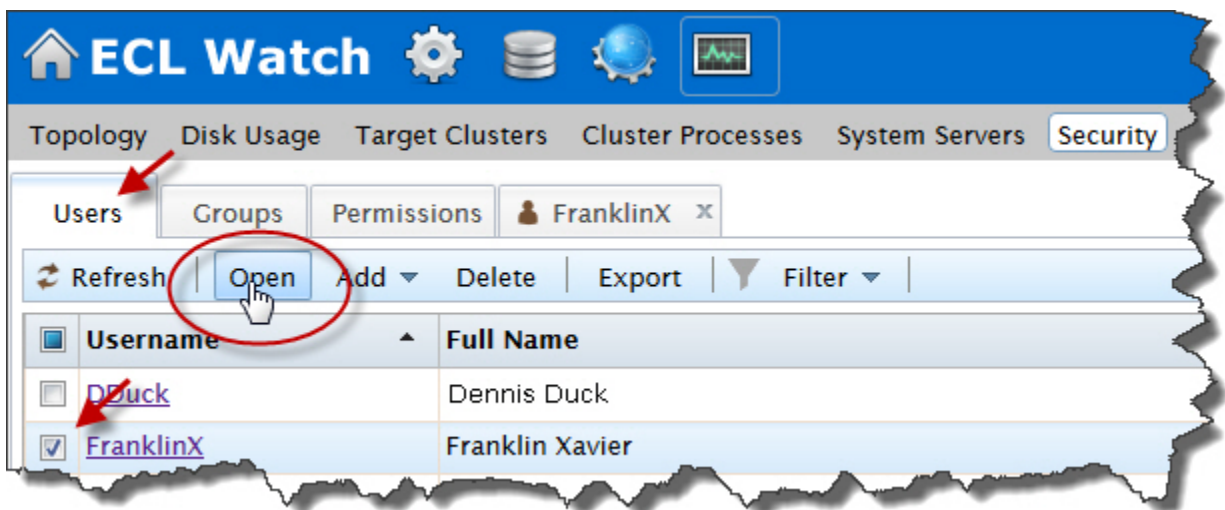
No submenu de navegação, clique no ícone **Operations** e em seguida no link **Security**.

1. Clique na aba **Users**.

Os usuários serão exibidos em uma lista.

2. Selecione o(s) usuário(s) que deseja modificar. Clique no link **User Name** para abrir a aba de detalhes do usuário.>

Para selecionar vários usuários, marque a caixa de seleção ao lado do User Name. Isso ativará os botões de ação para Users. Pressione o botão de ação **Open**.

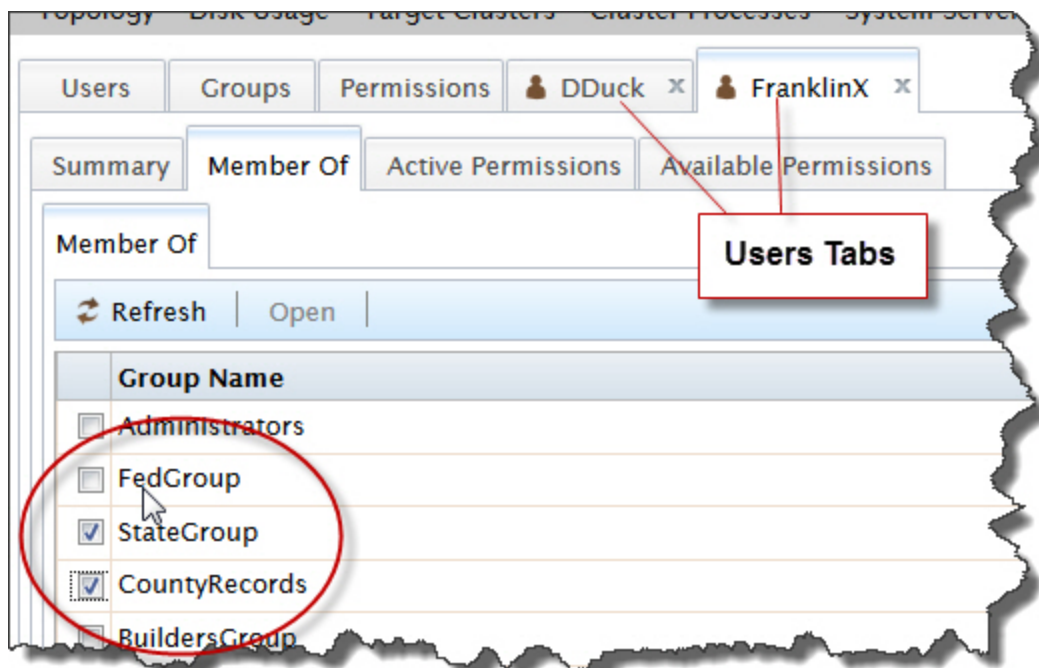


Uma aba será aberta para cada usuário selecionado. Cada aba de usuário contém várias sub-abas.

Os detalhes do usuário estão localizados na aba **Summary**.

3. Clique na aba do usuário para fazer a modificação desejada (caso tenha selecionado mais de um usuário, repita a operação para cada um deles).

A aba do usuário contém várias sub-abas.



Clique na subaba **Member of** para modificar os grupos do usuário.

4. Uma lista dos grupos disponíveis será exibida na aba **Member of** desse usuário.

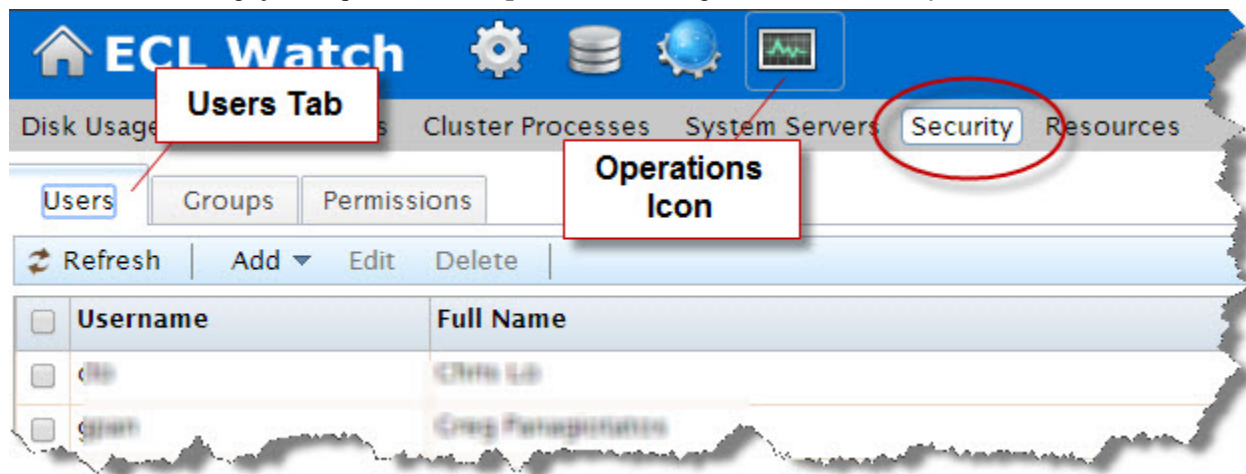
Para adicionar o usuário ao grupo, marque a caixa de seleção ao lado do grupo desejado.

5. As alterações serão salvas automaticamente. Feche a aba.

Promover um usuário para Administrador

Para modificar as credenciais de usuário você precisa ter acesso de administrador. Você pode designar a conta do Administrador do HPCC para permissões limitadas apenas relacionadas aos elementos do HPCC e não a direitos de administrador no LDAP.. Para promover um usuário a um administrador do HPCC, adicione o usuário ao grupo **Administrators** .

No submenu de navegação, clique no ícone **Operations** e em seguida no link **Security** .

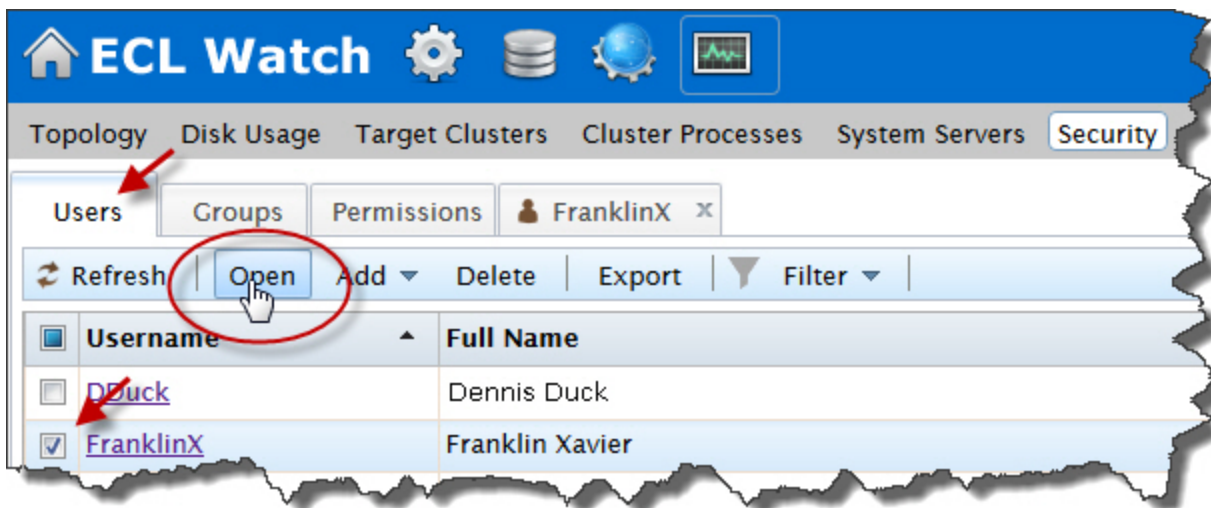


1. Clique na aba **Users**.

Os usuários serão exibidos em uma lista.

2. Selecione o(s) usuário(s) que deseja mudar de função. Clique no link **Username** para abrir a aba de detalhes do usuário.

Para selecionar vários usuários, marque a caixa de seleção ao lado do User Name. Isso ativará os botões de ação para Users. Pressione o botão de ação **Open**.



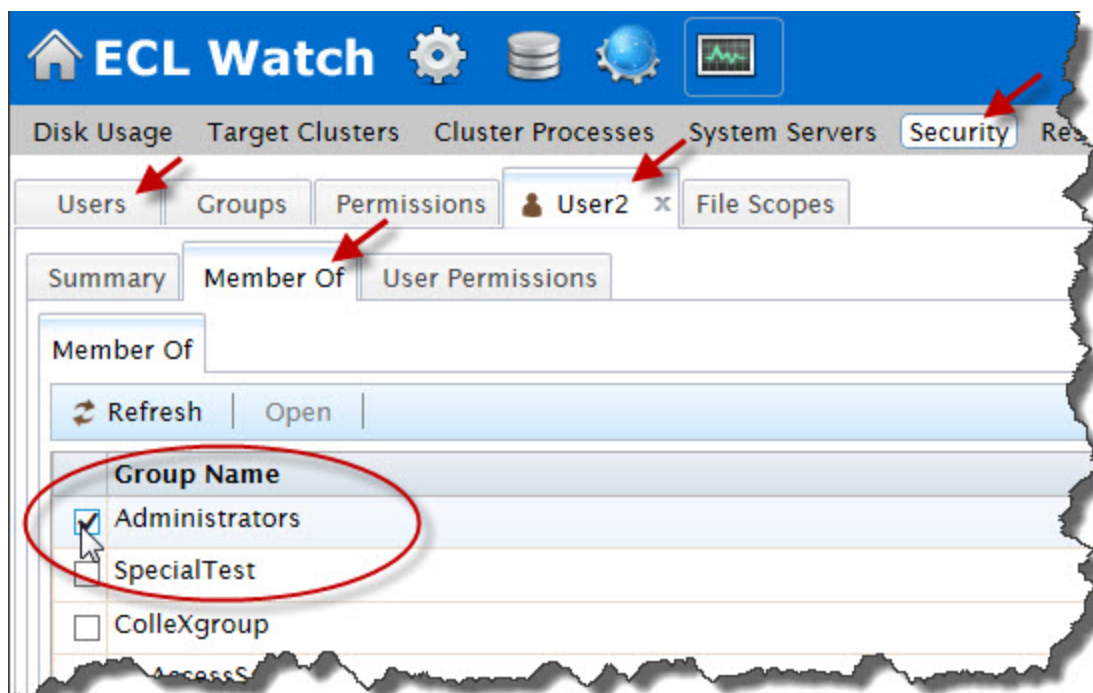
Uma aba será aberta para cada usuário selecionado. Cada aba de usuário contém várias sub-abas.

Os detalhes do usuário estão localizados na aba **Summary**.

3. Clique na aba do usuário para fazer a modificação desejada (caso tenha selecionado mais de um usuário, repita a operação para cada um deles).

A aba do usuário contém várias sub-abas.

Clique na subaba **Member of**.



4. Selecione **Administrators** marcando a caixa de seleção.

OBSERVAÇÃO: O nome do grupo padrão Administrador pode variar. É um valor configurável definido em **adminGroupName**. Por exemplo, se você configurar no ambiente o **adminGroupName** para "HPCCAdministrators", então a opção HPCCAdministrators será exibida na lista

5. As alterações serão salvas automaticamente. Feche a(s) aba(s).

Excluir um usuário de um grupo:

Você precisa ter acesso em nível de administrador para remover o usuário de um grupo.

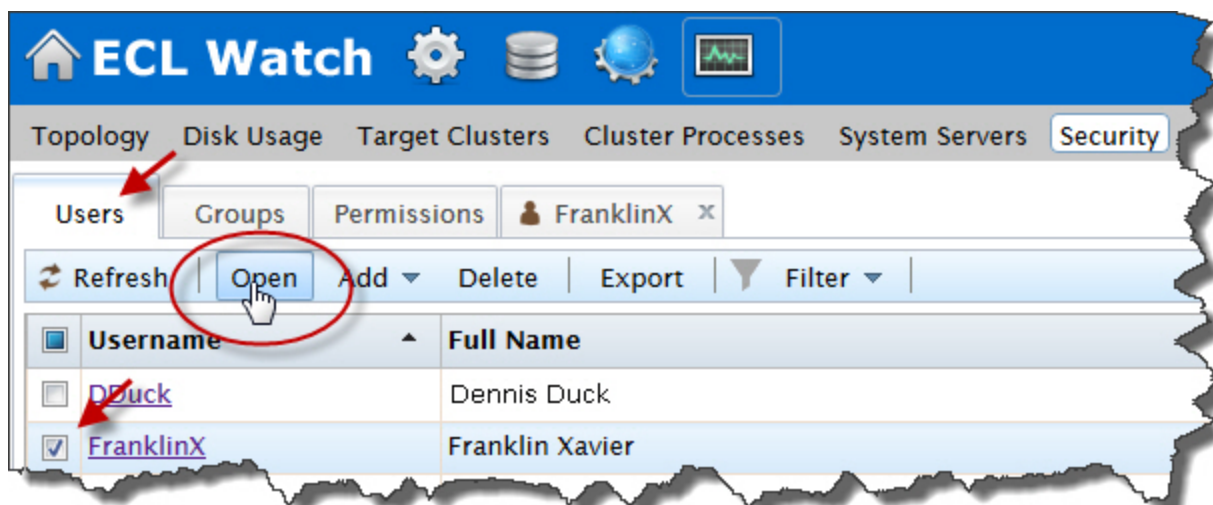
No submenu de navegação, clique no ícone **Operations** e em seguida no link **Security**.

1. Clique no **hiperlink Users**.

Os usuários serão exibidos em uma lista.

2. Selecione o(s) usuário(s) que deseja remover. Clique no link **Username** para abrir a aba de detalhes do usuário.

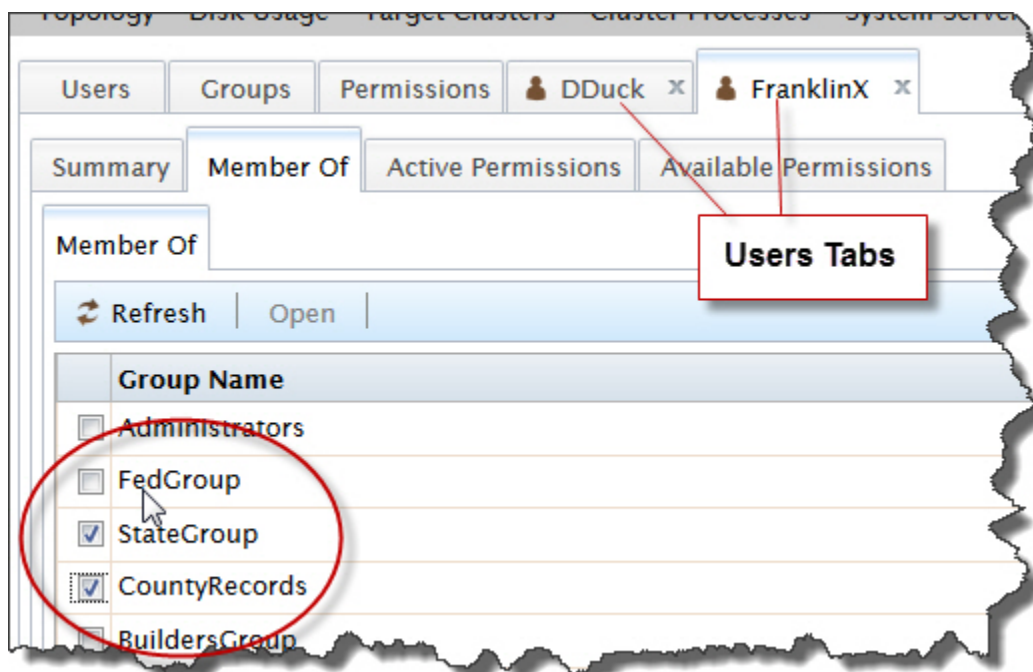
Para selecionar vários usuários, marque a caixa de seleção ao lado do User Name. Isso ativará os botões de ação para Users. Pressione o botão de ação **Open**.



Uma aba será aberta para cada usuário selecionado. Cada aba de usuário contém várias sub-abas.

3. Clique na guia do usuário que deseja modificar (caso tenha selecionado múltiplos usuários, repita a operação para cada um deles).

A aba do usuário contém várias sub-abas.



Clique na subaba **Member of** para modificar os grupos do usuário.

4. Há uma lista dos grupos disponíveis na aba **Member of** desse usuário.

Há uma caixa de seleção marcada ao lado de cada grupo ao qual o usuário pertence.

Para remover o usuário de um grupo, desmarque a caixa de seleção ao lado do grupo desejado.

5. As alterações serão salvas automaticamente. Feche a aba.

Alterar a senha do usuário:

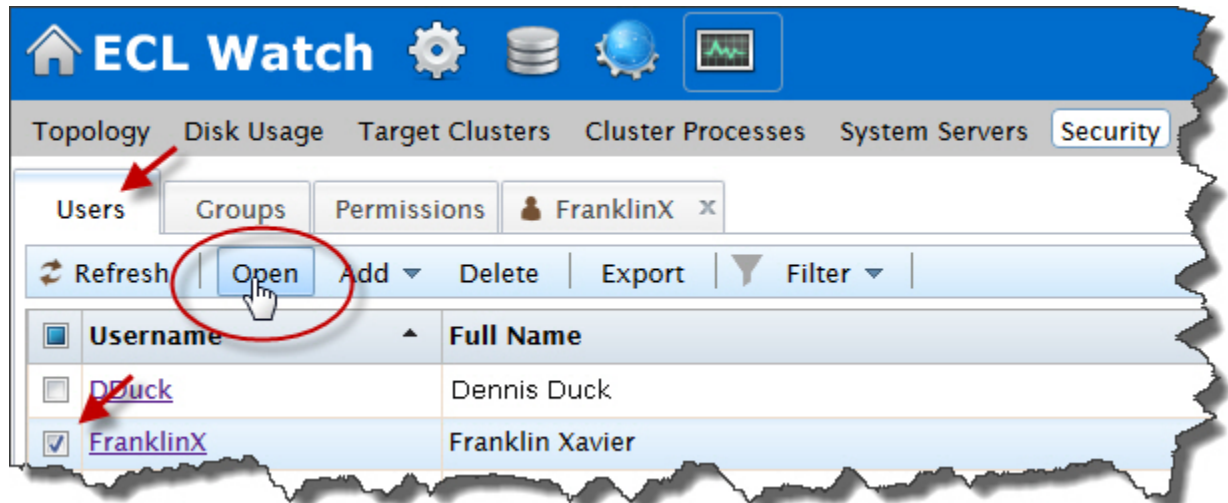
No submenu de navegação, clique no ícone **Operations** e em seguida no link **Security**.

1. Clique na aba **Users**.

Os usuários serão exibidos em uma lista.

2. Selecione o(s) usuário(s) que deseja modificar. Clique no link **Username** para abrir a aba de detalhes do usuário.

Para selecionar vários usuários, marque a caixa de seleção ao lado do User Name. Isso ativará os botões de ação para Users. Pressione o botão de ação **Open**.



Uma aba será aberta para cada usuário selecionado. Essa aba contém várias sub-abas.

Os detalhes do usuário estão localizados na aba **Summary**.

3. Selecione a aba **Summary**.
4. Altere a senha nos campos **Password** e **Retype new Password** na aba de detalhes do usuário conforme solicitado (caso tenha selecionado mais de um usuário, repita o procedimento para cada um dos demais).

Observação: O **Username** não pode ser alterado.

5. Pressione o botão **Save**.

Uma mensagem de confirmação será exibida.

Excluir um usuário da lista de usuários autenticados:

No submenu de navegação, clique no ícone **Operations** e em seguida no link **Security**.

1. Clique na aba **Users**.

Os usuários serão exibidos em uma lista.

2. Marque a caixa à esquerda do nome do(s) usuário(s) que deseja remover.

Observação: Estes usuários não terão mais acesso ao ECL Watch.

3. Pressione o botão de ação **Delete**.

A confirmação será exibida.

Configurar permissões para um usuário individual

Haverá casos em que você precisará modificar as permissões para usuários individuais. Por exemplo, os usuários podem ter necessidades individuais de segurança que não sejam totalmente cobertas em nenhum grupo; ou poderá haver situações em que um usuário solicitará acesso temporário a um recurso do HPCC Systems. As permissões configuradas nesta área do ECL Watch afetam apenas o usuário selecionado. A maioria das permissões individuais configuradas aqui substitui as que foram configuradas em qualquer grupo ao qual o usuário pertença, exceto em casos de negação explícita.

Configurando permissões para um usuário individual:

No submenu de navegação, clique no ícone **Operations** e em seguida no link **Security**.

1. Clique na aba **Users**.

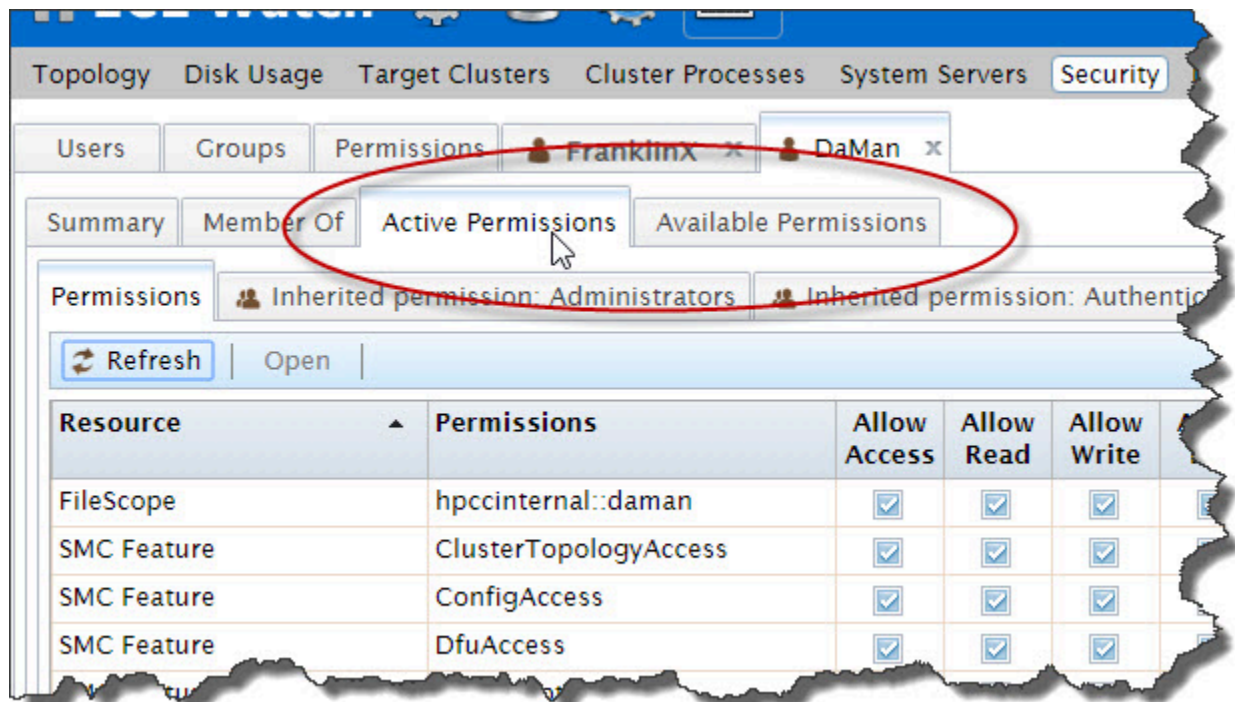
Os usuários serão exibidos em uma lista.

2. Selecione o(s) usuário(s) que deseja modificar. Clique no link **Username** para abrir a guia de detalhes do usuário.

Para selecionar vários usuários, marque a caixa de seleção ao lado do Username. Isso ativará os botões de ação para Users. Pressione o botão de ação **Open**.

3. Clique na aba do nome do usuário para modificar (caso tenha selecionado múltiplos usuários, repita a operação para cada um deles).

A aba do usuário contém várias sub-abas.

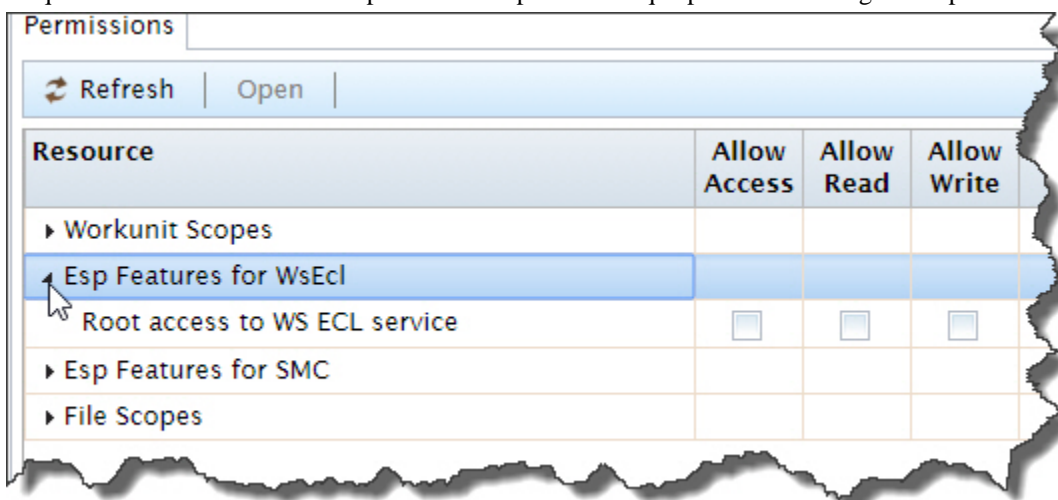


Clique na subaba **Active Permissions** para visualizar as permissões atuais do usuário.

4. Clique na aba **Available Permissions** para ver todos os conjuntos de permissões disponíveis para esse usuário.

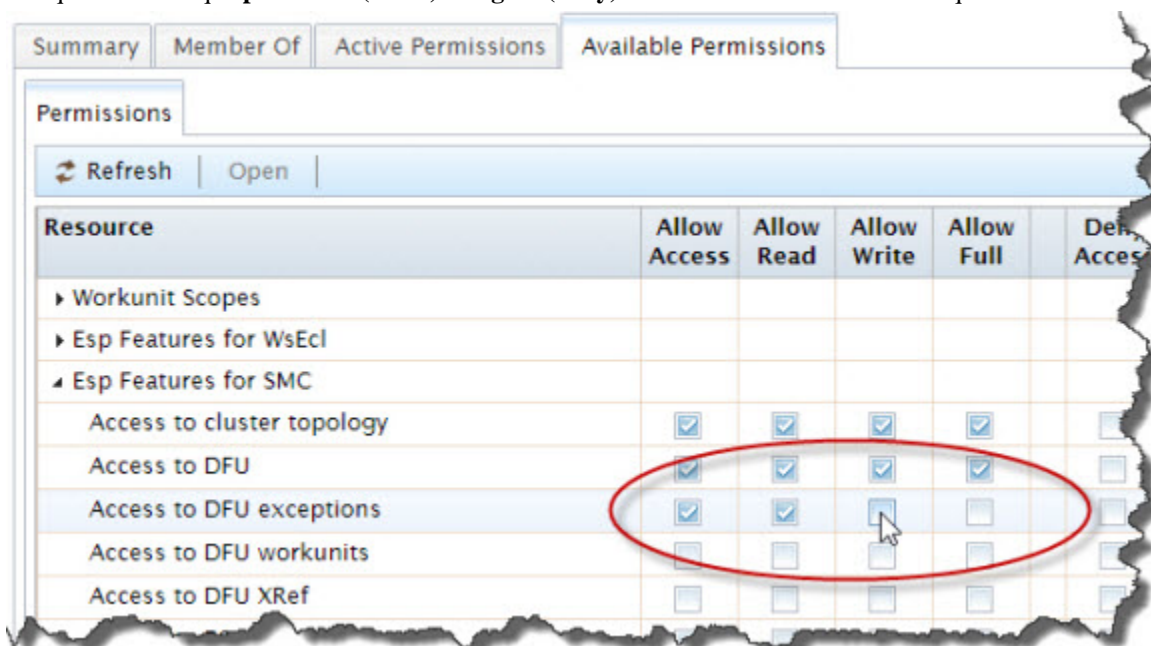
Ao selecionar as permissões na aba Available Permissions disponíveis, elas serão exibidas e podem ser configuradas na aba Active Permissions.

5. Clique na seta ao lado do recurso para exibir as permissões que podem ser configuradas para esse recurso.



A lista dos grupos de permissão atualmente configurados para este usuário e dos grupos que foram herdados pelo usuário também está listada. Clique na seta para permitir a definição das configurações individuais do recurso.

6. Pode haver mais de uma configuração de recurso disponível em cada grupo. Por isso, não se esqueça de definir as permissões para cada configuração conforme requerido.
7. Marque as caixas que **permitem (allow)** e **negam (deny)** acesso ao usuário conforme requerido.



OBSERVAÇÃO: É preciso ter cautela ao determinar qualquer configuração de permissão para **negar** um direito. A permissão mais restritiva sempre se aplica.

8. As alterações serão salvas automaticamente. Feche a aba.

Configurando e modificando grupos de permissões

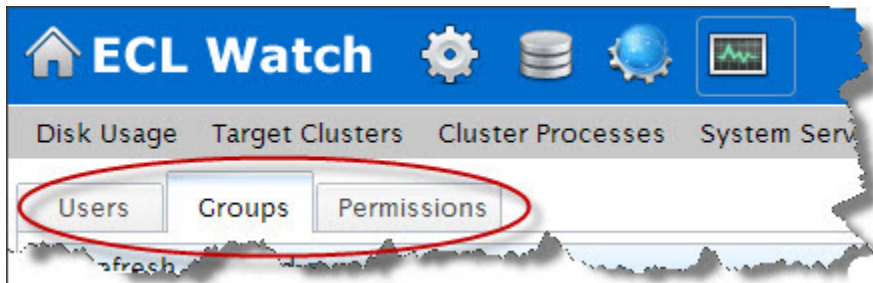
A organização dos grupos garante que todos os usuários com as mesmas necessidades de permissão tenham as mesmas configurações de permissão. Você pode fornecer aos usuários o acesso necessário às áreas de recursos do HPCC. Não há limite quanto ao número de grupos que podem ser criados. Você pode criar quantos grupos forem precisos para controlar o acesso de todos os seus usuários, independentemente das workunit desempenhadas por eles.

Use o item **Groups** do menu para:

- Adicionar um novo grupo.
- Remover um grupo.
- Adicionar membros a um grupo
- Modificar as permissões de um grupo

Adicionando e editando grupos

Ao adicionar ou alterar as permissões de um grupo, todos os membros desse grupo receberão essas configurações de permissão. Por isso, é importante ter certeza de estar concedendo ou negando acesso aos recursos apropriados para os membros desse grupo. Se precisar fazer alterações para um único usuário (ou para um pequeno número de usuários), será melhor fazer tais alterações para cada usuário individual como ilustrado nas seções anteriores.



Para modificar grupos, clique no ícone **Operations**, e em seguida no link **Security** do submenu de navegação. Clique na aba **Groups**.

Adicionando um novo grupo:

No submenu de navegação, clique no ícone **Operations** e em seguida no link **Security**

1. Clique na aba **Groups**.
2. Pressione o botão de ação **Add**.



Isso abrirá a caixa de diálogo onde é possível inserir um nome para o grupo.

3. Insira o **Group Name**.
4. Insira um nome completamente distinto para o dono do grupo no campo **Manager by**.
5. Insira uma descrição para o grupo. (opcional)
6. Pressione o botão **Add**.

Isso abrirá uma nova aba e várias subabas para o grupo

A subaba **Summary** exibe o nome do grupo.

A aba **Members** exibe a lista dos usuários; marque a caixa de seleção ao lado de cada usuário para adicioná-lo ao grupo.

A aba **Active Group Permissions** exibe as permissões aplicadas ao grupo.

A aba **Available Groups Permissions** exibe todas as permissões disponíveis; a seleção a partir da aba Permissions disponíveis aplica as permissões à aba Permissão de grupo ativo.

Você pode definir as permissões e adicionar membros a esse grupo nas respectivas sub-abas do grupo.

Excluir um grupo:

No submenu de navegação, clique no ícone **Operations** e em seguida no link **Security**.

1. Clique na aba **Groups**.
2. Localize o grupo na lista e marque a caixa de seleção ao lado dele.

3. Pressione o botão de ação **Delete** .

4. Pressione o botão de confirmação **OK**

O grupo não será mais exibido na lista.

Adicionar novos membros para um grupo:

No submenu de navegação, clique no ícone **Operations** e em seguida no link **Security** .

1. Clique na aba **Groups** .

2. Localize o grupo na lista e marque a caixa de seleção ao lado dele.

3. Pressione o botão de ação **Open** .

Isso abrirá uma nova aba para o grupo.

As sub-abas exibem: Summary , Members, **Active Group Permissions**, e **Available Group Permission**.

4. Selecione a aba **Members**

A aba Members exibirá uma lista de todos os usuários no sistema. Aqueles que pertencem ao grupo selecionado terão a caixa de seleção marcada ao lado.

5. Marque a(s) caixa(s) à esquerda do nome dos usuários que deseja adicionar ao grupo.

6. As alterações serão salvas automaticamente. Feche a aba.

Excluir membros de um grupo:

No submenu de navegação, clique no ícone **Operations** e em seguida no link **Security** .

1. Clique na aba **Groups** .

2. Localize o grupo na lista e marque a caixa de seleção ao lado dele.

3. Pressione o botão de ação **Open** .

Isso abrirá uma nova aba para o grupo.

A aba Grupos possui diversas sub-abas: **Summary**, **Members**, **Active Group Permissions**, e **Available Group Permissions**.

4. Selecione a aba **Members** .

A aba Members exibirá uma lista de todos os usuários no sistema. Aqueles que pertencem ao grupo selecionado terão a caixa de seleção marcada ao lado.

5. Desmarque a(s) caixa(s) à esquerda para todos os usuários que deseja remover do grupo.

6. As alterações serão salvas automaticamente. Feche a aba.

Configurar Permissões para Grupo

Por padrão, todos os usuários são membros do grupo **Authenticated Users** . O grupo **Authenticated Users** possui direitos de acesso a quase todos os recursos. Para definir controles mais restritos, é preciso criar grupos específicos com permissões mais limitadas.

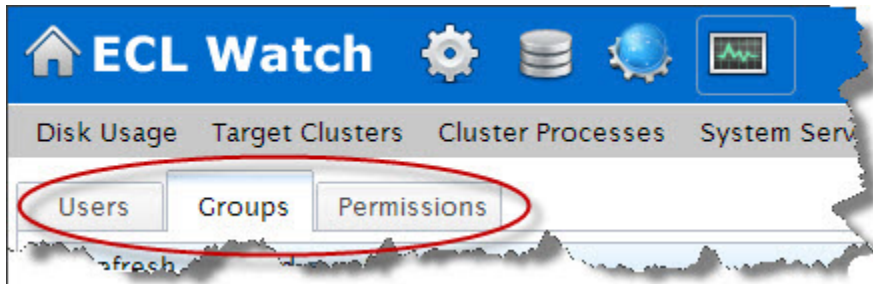
Você pode criar grupos apenas com os direitos de acesso que deseja conceder. Essa abordagem oferece maior flexibilidade, já que um único ID de usuário pode estar associada a vários grupos.

Como prática recomendada, use **Allow** em vez de **Deny** para controlar o acesso. Quando possível, use a função Deny “negar” apenas como exceção. Caso queira negar o acesso de um usuário a algum controle específico, recomenda-se criar um grupo para isso e adicionar o(s) usuário(s) neste grupo para, então, negar o acesso somente para esse grupo.

Lembre-se de que o controle mais restritivo tem precedência. Por exemplo, se um usuário faz parte de um grupo que não dá permissão de acesso a um determinado arquivo, porém este mesmo usuário também faz parte de outro grupo cujo acesso a tal arquivo é permitido, o usuário continuará sem permissão para acessar o arquivo.

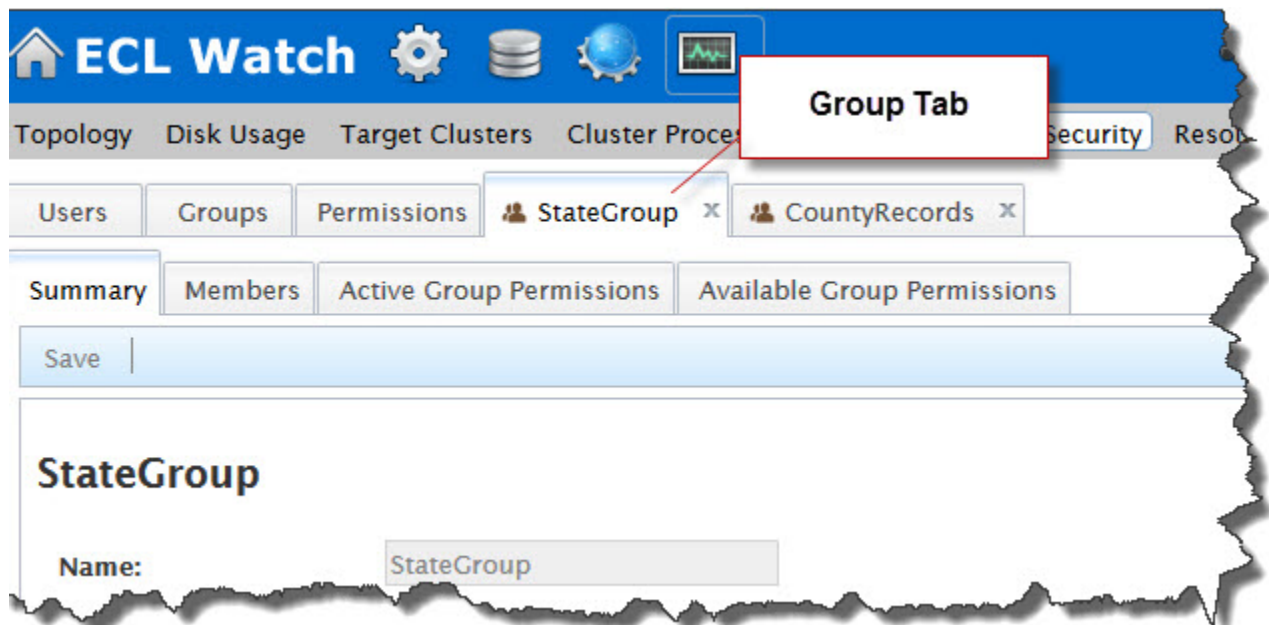
Configurando Permissões para Grupo:

No submenu de navegação, clique no ícone **Operations** e em seguida no link **Security**.



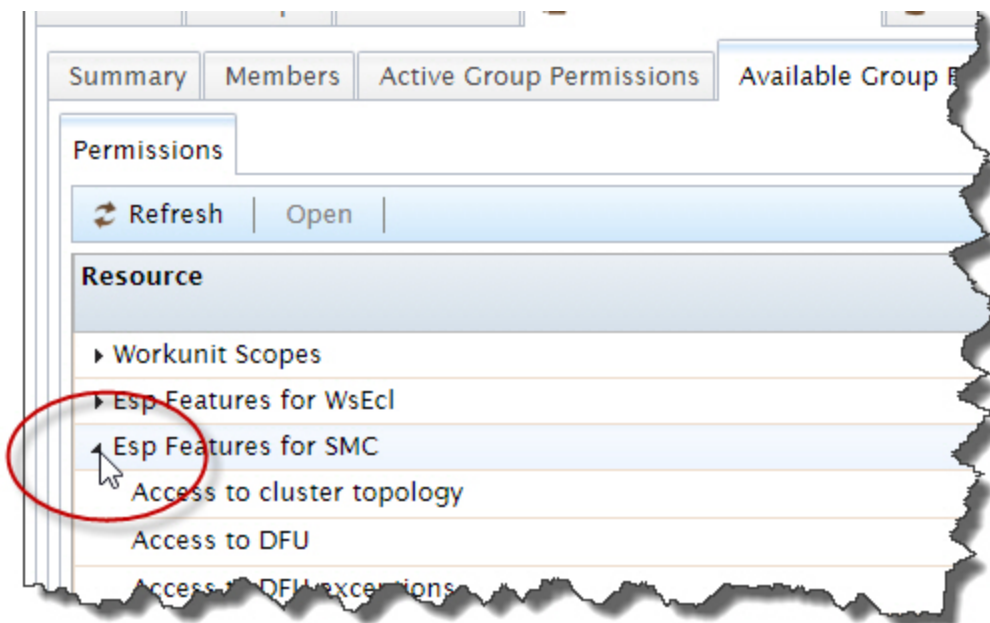
1. Clique na aba **Groups**.
2. Localize o grupo na lista e marque a caixa de seleção ao lado dele.
3. Pressione o botão de ação **Open**.

Isso abrirá uma nova aba para o grupo.



A guia do grupo exibirá as sub-abas: **Summary**, **Members**, **Active Group Permissions**, e **Available Group Permissions**.

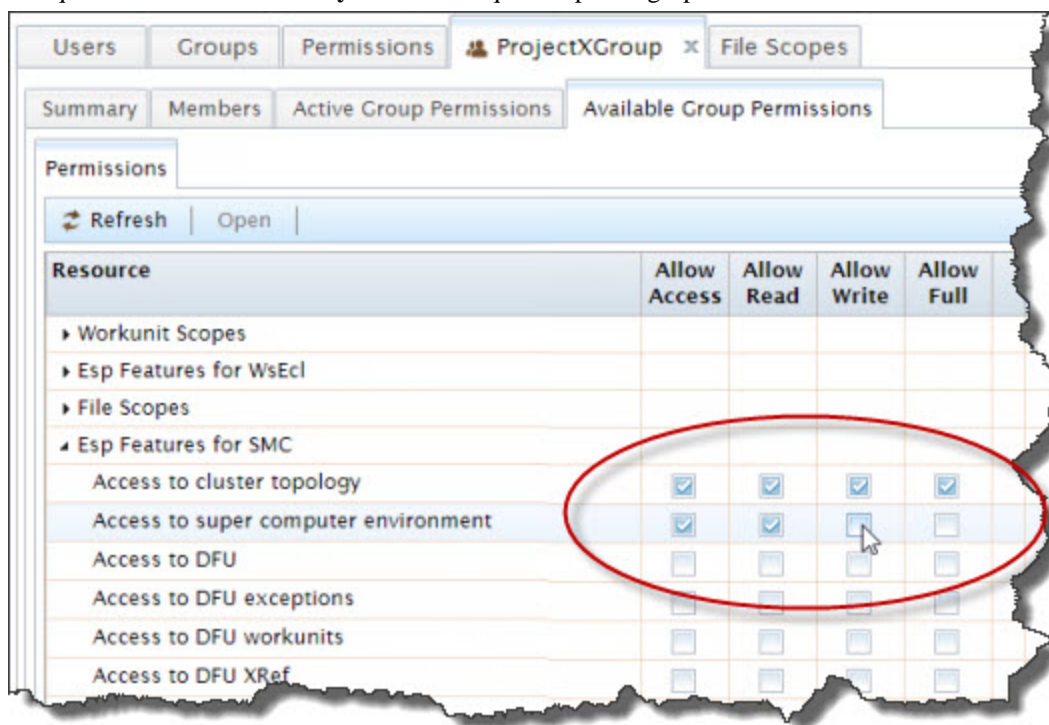
4. Selecione a sub-aba **Available Group Permissions** . Isso exibirá todos os recursos de permissão disponíveis.
5. Clique na seta à esquerda de **Resource** para expandir e mostrar as configurações de permissão para os recursos.



Os recursos de permissão dos grupos serão exibidos.

6. Pode haver mais de uma configuração de recurso disponível em cada grupo. Por isso, não se esqueça de definir as permissões para cada configuração conforme requerido.

7. Marque as caixas **Allow** e **Deny** conforme requerido para o grupo.



OBSERVAÇÃO: É preciso ter cautela ao determinar qualquer configuração de permissão para **negar** um acesso. A permissão mais restritiva sempre se aplica.

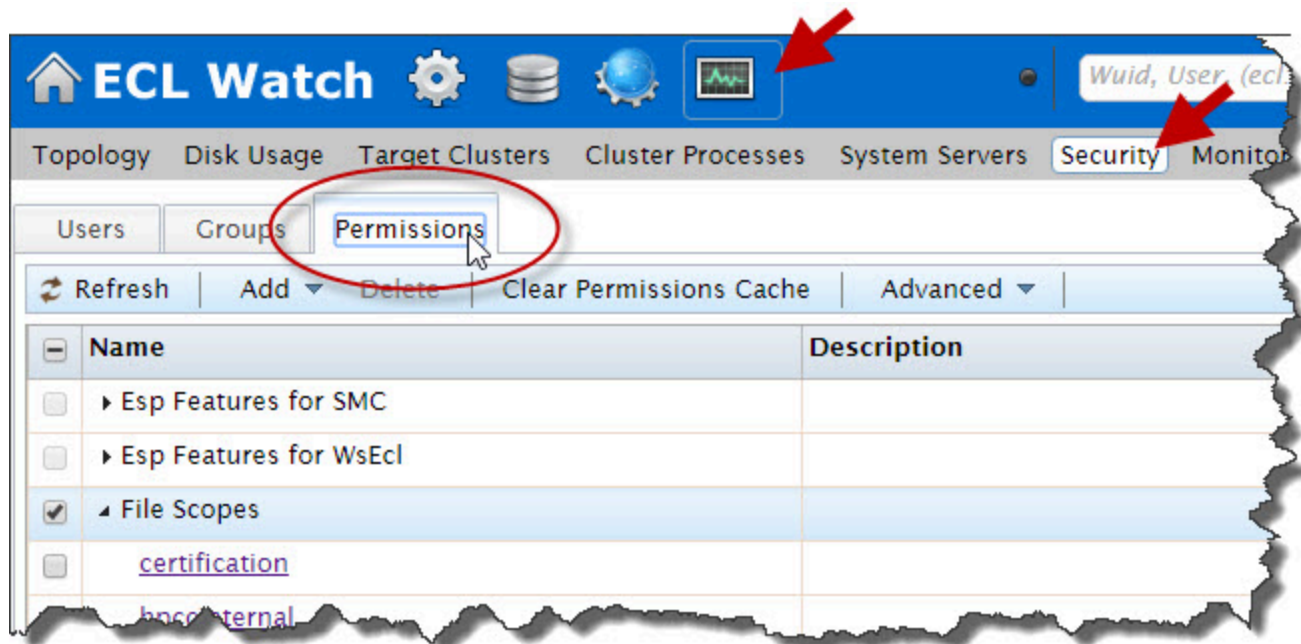
8. Pode haver mais de uma configuração de recurso disponível, selecione o(s) recurso(s) necessário(s) na lista suspensa.

Repita o procedimento para cada recurso aplicável.

9. As alterações serão salvas automaticamente. Feche a aba.

Controle de acesso em nível de recurso

O acesso às permissões específicas está disponível através do ECL Watch. Para modificar as permissões específicas, é preciso ter acesso em nível de administrador. Para acessar as permissões específicas, clique no ícone **Operations**, e em seguida clique no link **Security** a partir do submenu de navegação.

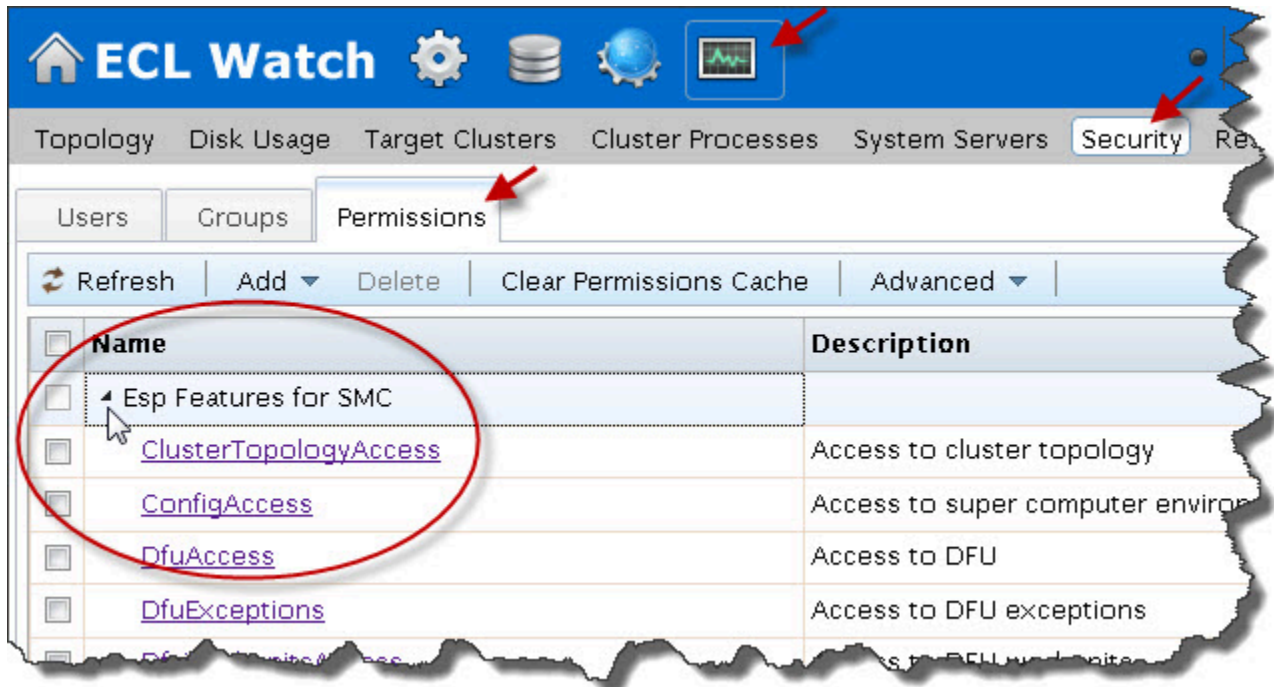


Para usar os controles de nível específico, aplique o recurso específico da aba **Available Permissions** para a aba **Active Permissions** para usuários e grupos. O uso de controles de nível específico permitirá que você:

- Veja as funcionalidades e as permissões para qualquer recurso
- Edite as permissões para qualquer recurso
- Atualize as permissões de usuários e grupos para um recurso específico

Recursos

Há muitos outros recursos para os quais você pode configurar o controle de acesso no HPCC. O acesso aos recursos do HPCC system é controlado através dos Recursos do **ESP Features for SMC**.



Os recursos disponíveis estão listados abaixo da aba **Permissions**. Você pode visualizar e obter acesso aos controles específicos aqui. No entanto, os controles específicos devem ser aplicados aos usuários e não aos grupos. Ao clicar no link do nome específico, será aberta a guia que mostra os usuários e grupos onde essas permissões específicas são aplicadas.

As configurações de permissão dos recursos do ECL Watch que não estão listadas são irrelevantes e não devem ser usadas.

Aplicar permissões para um recurso:

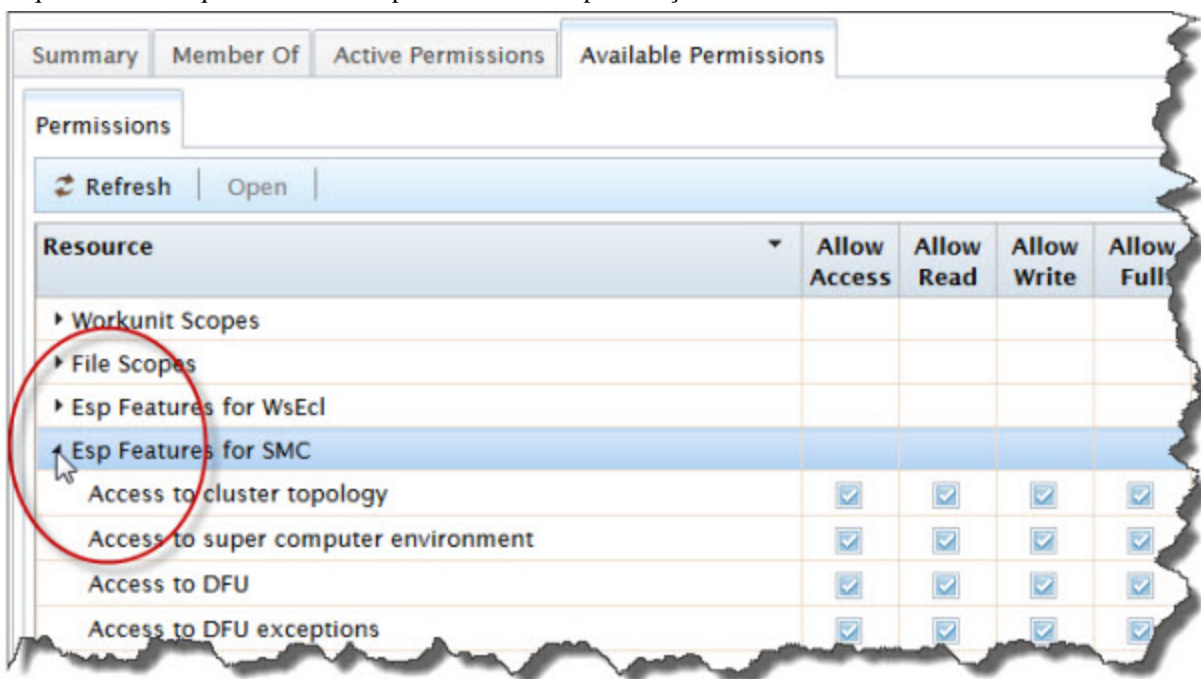
Para usar permissões específicas, é preciso aplicá-las a um usuário ou grupo(s). Para acessar as permissões específicas, clique no ícone **Operations**, e em seguida clique no link **Security** a partir do submenu de navegação.

1. Identifique o(s) usuário(s) ou grupo(s) que deseja modificar as permissões específicas.

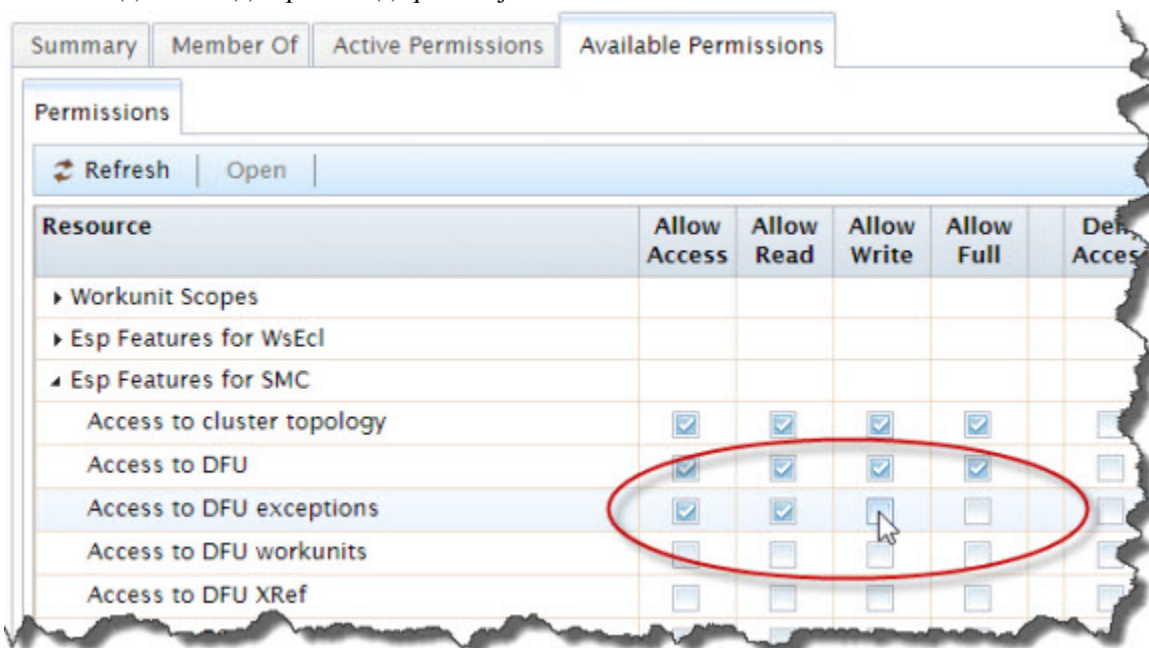
Selecione a aba apropriada. (Users or Groups)

2. Marque a(s) caixa(s) de seleção ao lado do(s) usuário(s) ou grupo(s) que deseja modificar.
3. Pressione o botão de ação **Open**. Uma aba será aberta para cada usuário ou grupo.
4. Clique na sub-aba **Available Permissions**.

5. Clique na seta à esquerda do recurso para mostrar as especificações desse recurso.



6. Localize o(s) recurso(s) específico(s) que deseja atualizar.



7. Clique nas colunas “allow” e “deny” na(s) caixa(s) de seleção.
8. As alterações serão salvas automaticamente. Feche a(s) aba(s).

Observação: Esse processo deve ser realizado individualmente para cada usuário ou grupo.

Recursos de Permissões SMC

A tabela a seguir descreve o nível de acesso exigido para que esses recursos do ECL Watch do HPCC possam ser usados.

Nome	Descrição	Acesso
ClusterTopologyAccess	Acesso à Topologia do cluster	Leitura
	Ver arquivos de log.	Completo
DfuAccess	Acesso aos arquivos lógicos DFU	Leitura
	Remover arquivos, Adicionar aos superarquivos e remover dos superarquivos	Escrita
	Apagar metadados do histórico do arquivo	Completo
DfuExceptions	Acesso à leitura de Exceções DFU	Leitura
DfuWorkunitsAccess	Acesso à leitura de Workunit DFU	Leitura
	Acesso para Criar, Excluir, Atualizar, Enviar e Abortar DFU Workunits	Gravação
DfuXrefAccess	Acesso à leitura de DFU XREF	Leitura
	Limpar diretório	Gravação
	Fazer alterações e gerar relatórios XREF	Completo
EclDirectAccess	Acesso ao serviço ECL Direct.	Completo
ESDLConfigAccess	ESDL Acesso à configuração	Leitura
	Publicar definição e conexão ESDL, configurar método de conexão ESDL.	Gravação
	Apagar definições ESDL, apagar conexões ESDL	Completo
FileDesprayAccess	Permite que o usuário faça o despray (consolidar dados dos nós) dos arquivos lógicos.	Gravação
FileIOAccess	Acesso à leitura de arquivos na Zona de entrada de arquivos	Leitura
	Acesso à gravação de arquivos na Zona de entrada de arquivos	Gravação
PackageMapAccess	Acesso à(ao) ListPackage, ListPackages, GetPackage, GetPackageMapById, ValidatePackage, GetQueryFileMapping, GetPackageMapSelectOptions, GetPartFromPackageMap	Leitura
	Access a(ao) AddPackage, CopyPackageMap, ActivatePackage, DeActivatePackage, AddPartToPackageMap, RemovePartFromPackageMap	Gravação
	Apagar Pacote	Completo
FileScopeAccess	Permite acesso à consulta, configuração e remoção de permissões do escopo de arquivos	Completo
FileDesprayAccess	Acesso ao spraying (processo de distribuição dos dados aos nós) e cópia	Leitura
	Renomear, spray, copiar e replicar arquivos	Gravação
	Fazer o download Apagar da Zona de entrada de arquivos	Completo
MachineInfoAccess	Acesso às informações da máquina/preflight	Leitura


Nome	Descrição	Acesso
MetricsAccess	Acesso às informações sobre métricas SNMP (Métricas Roxie)	Leitura
OthersWorkunitsAccess	Acesso à visualização de workunit de outro usuário	Leitura
	Acesso à Modificar ou reenviar workunit do usuário	Gravação
	Acesso à Remover workunit de outros usuários	Completo
OwnWorkunitsAccess	Acesso à visualização da própria workunit	Leitura
	Acesso à Criar ou modificar a própria workunit	Gravação
	Acesso a remoção da própria workunit	Completo
RoxieControlAccess	Acesso aos comandos de controle do Roxie	Leitura
SmcAccess	Acesso ao ECL Watch (Serviço SMC)	Leitura
ThorQueueAccess	Acesso ao controle da fila de workunit do Thor	Completo
WsEclAccess	Acesso ao serviço WS ECL	Completo
WsLogAccess	Habilita a função de leitura de logs dos componentes	Leitura

Algumas Notas de Permissões de Recursos

- Para o SMCAccess é obrigatório ter feito o login no ECL Watch.
- ThorQueueAccess permite manipular a fila promovendo ou rebaixando as workunit de acordo com a prioridade.
- ThorQueueAccess também permite pausar ou limpar a fila do Thor. Você também pode ver as estatísticas de uso do Thor.
- Dependendo do nível de acesso do usuário, é possível visualizar, modificar e remover suas próprias workunit ou as workunit de outros usuários. Trata-se do OwnWorkunitsAccess e OthersWorkunitsAccess, respectivamente.
- As permissões do DfuWorkunitsAccess permitem que os usuários visualizem ou manipulem as workunit DFU .
- Os usuários precisam ter autorização para ver os arquivos na zona de entrada de arquivos, assim como também para inserir arquivos lá. Também é preciso obter permissões adicionais para fazer o spray (distribuir aos nós) e copiar arquivos da zona de entrada de arquivos para o cluster, assim como para fazer o despray (consolidar dados dos nós) dos arquivos do cluster para a zona de entrada de arquivos.

DFU Xref

XREF é usado para monitorar os arquivos nos clusters. Os relatórios gerados mostram onde a organização é necessária nos clusters, e os usuários precisam obter permissão adicional para usar este recurso.

	Em um sistema maior, sugerimos limitar o número de usuários que têm permissão para gerar relatórios XREF configurando o acesso ao DfuXrefAccess para FULL (Completo) apenas para esses usuários.
---	--

Usuários/Permissões

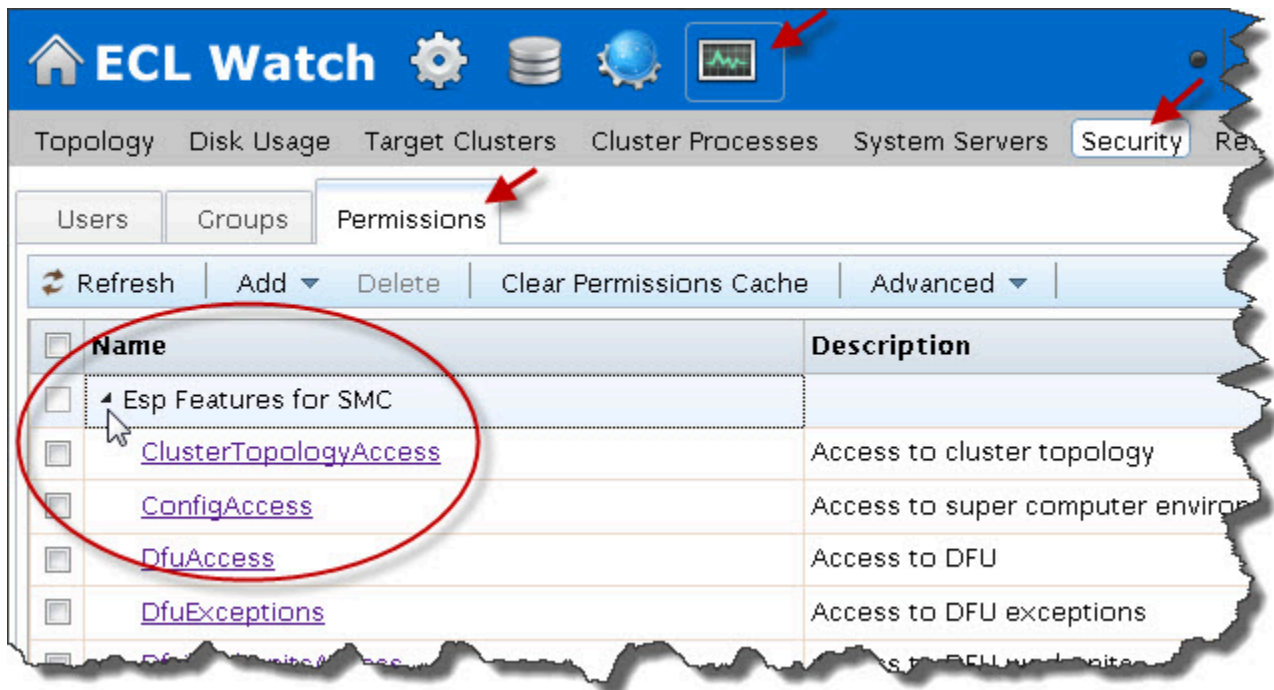
Para poder visualizar a área **Users/Permissions** no ECL Watch, o usuário precisa ser um membro do grupo Administradores (ou de nome similar) com permissões adequadas no servidor LDAP ou Active Directory.

Controle de Acesso a Arquivos

A tecnologia do **servidor LDAP Dali** do HPCC permite configurar permissões de acesso seguro às pastas de arquivo de dados (ou escopos de arquivo). Isso é controlado pelo uso dos recursos de escopo de arquivos.

Uma OU denominada **Files** é criado automaticamente quando o servidor Dali é inicializado. Para proteger as pastas de dados, crie escopos de arquivo para essa pasta e aplique os direitos para cada escopo.

Figure 26. Permissões de escopo de arquivos



Por exemplo, abaixo de **Files** há uma unidade (OU) representando o cluster, como **thor** (ou o nome configurado para seu cluster). Além disso, logo abaixo poderia ter uma unidade denominada **collectionx** que contém duas unidades: **publicdata** e **securedata**. A pasta **publicdata** possui direitos concedidos a um grande grupo de usuários; já para a pasta **securedata** foi concedido o acesso limitado. Isso permite impedir que usuários não autorizados acessem os arquivos da pasta **securedata** folder.

A estrutura descrita acima corresponde a essa estrutura lógica:

collectionx::securedata

A qual corresponde a essa estrutura física:

/var/lib/HPCCSystems/hpcc-data/thor/collectionx/securedata

Todos os componentes e ferramentas HPCC respeitam a segurança de acesso a arquivos definidos no LDAP. As seguintes exceções são consideradas a nível de sistema ou para usuários administradores:

- Acesso aos arquivos de rede usando UNCs, Serviços de Terminal, ou SSH.
- Utilitários administrativos

A tentativa de acesso a um arquivo em uma pasta sem que a permissão tenha sido concedida, resultará em um dos seguintes erros:

```
DFS Exception: 4 Create access denied for scope <filepath>
```

ou

DFS Exception: 3 Lookup access denied for scope <filepath>

(onde <filepath> corresponde ao caminho completo do escopo de arquivo lógico)

Criando um escopo de arquivo

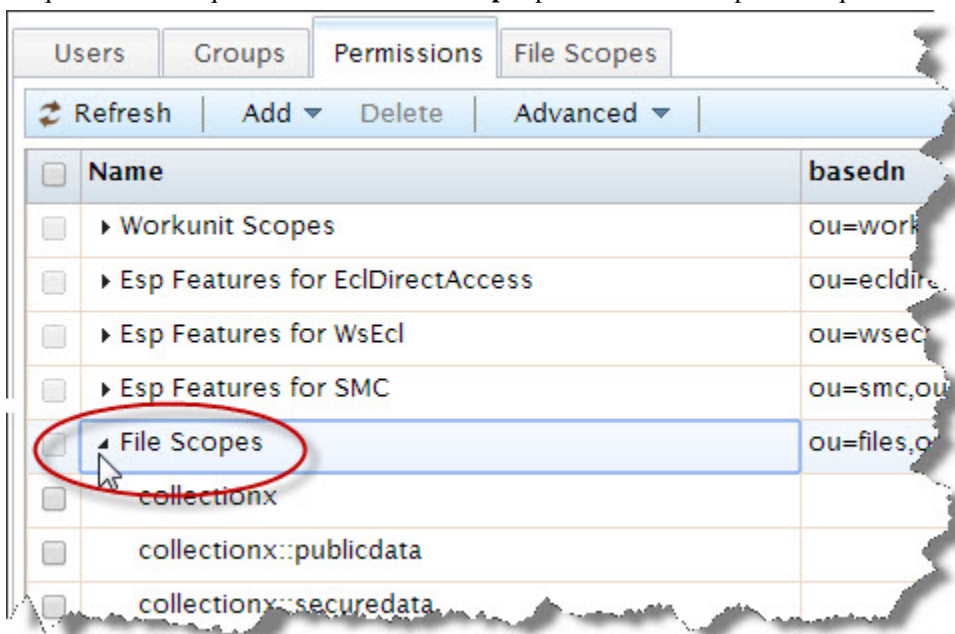
Para aplicar as permissões a um escopo de arquivos, primeiramente é preciso criar o(s) escopo(s) de arquivos.

Para criar o(s) escopo(s) de arquivos, clique no ícone **Operations** , e em seguida no link **Security** localizados no submenu de navegação.

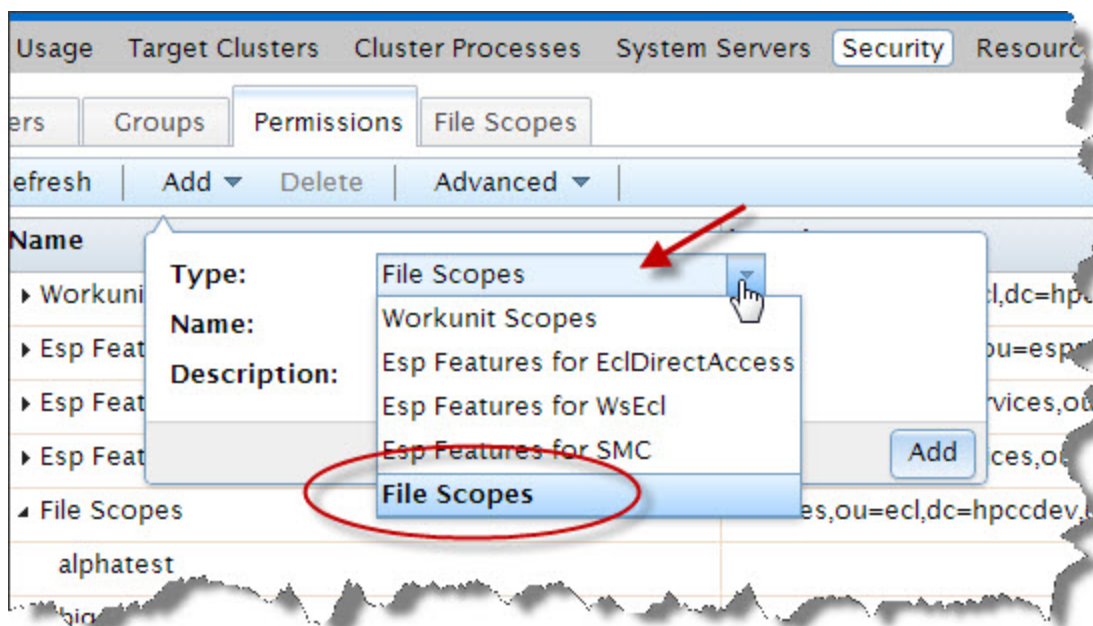
1. Clique na aba **Permissions** .

Os recursos específicos serão exibidos.

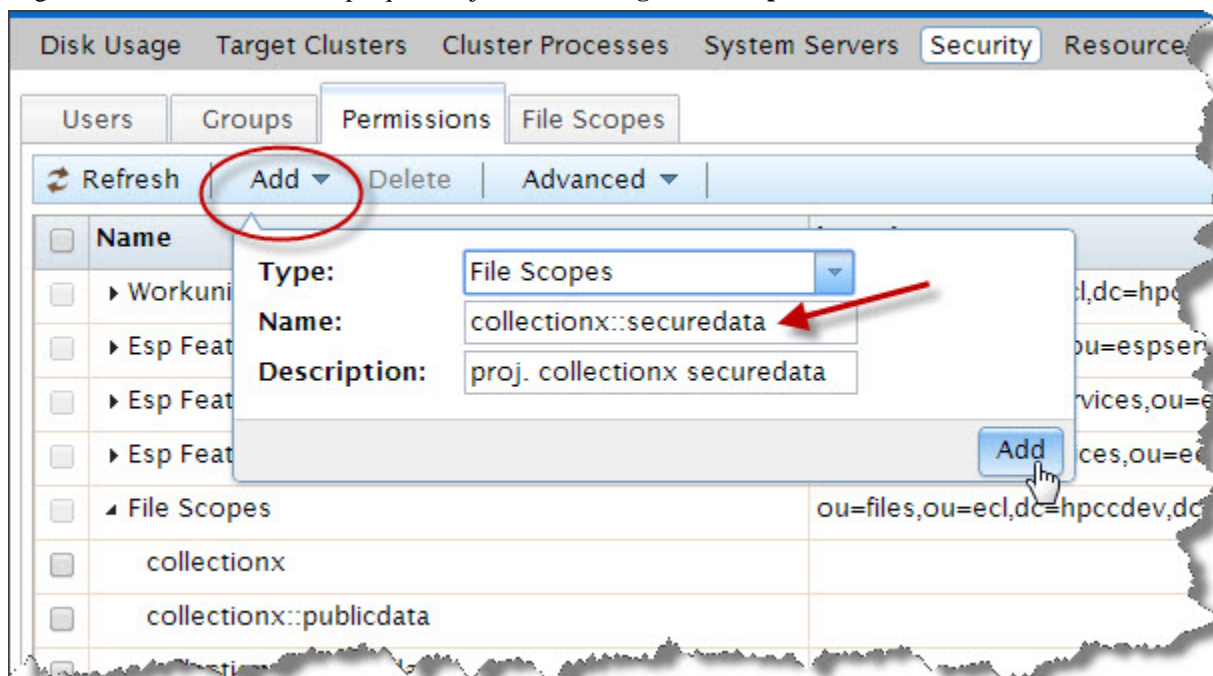
2. Clique na seta à esquerda do recurso **File Scopes** para exibir os escopos do arquivo.



3. Pressione o botão **Add** .
4. Selecione **File Scopes** na lista suspensa.



5. Digite no o nome exato do escopo que deseja adicionar. **Digite no campo** Name



o nome exato do escopo que deseja adicionar. **Digite uma breve descrição no campo** Description.

6. Pressione o botão **Add**.

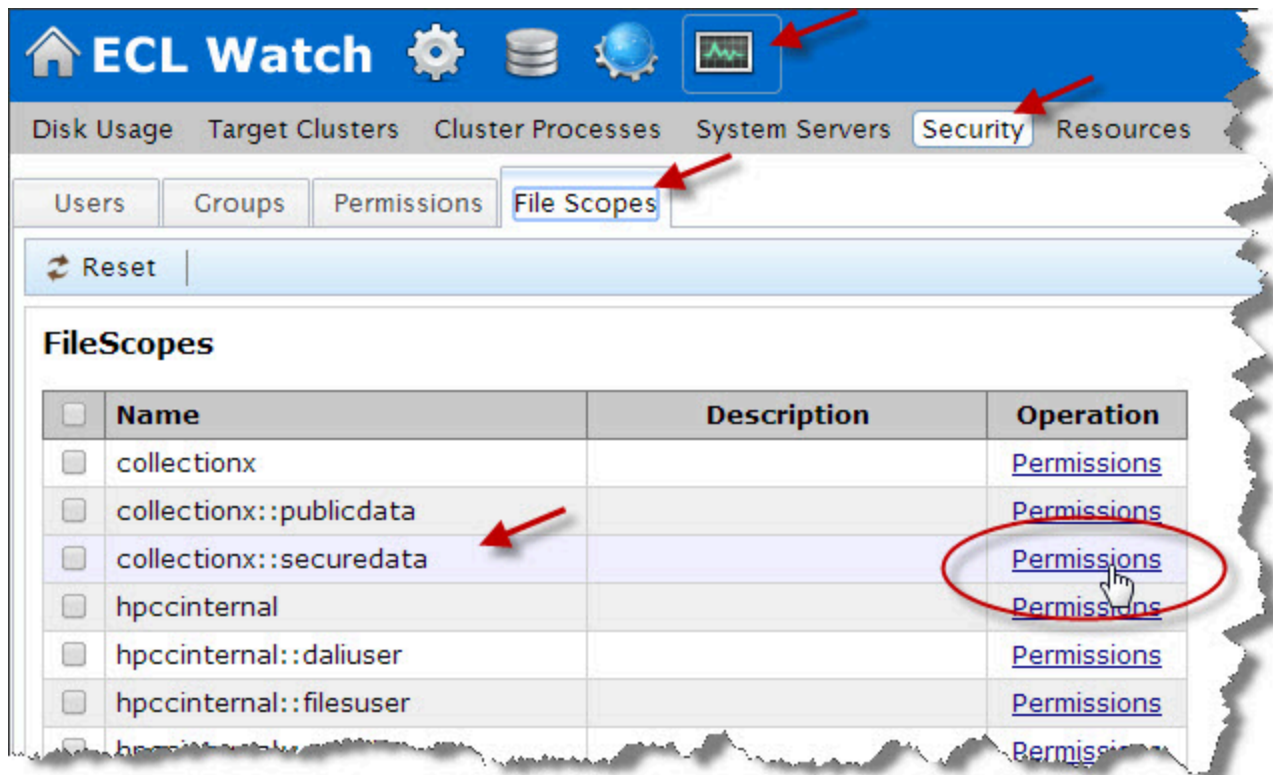
O novo escopo será exibido na lista.

Configurando permissões para escopos de arquivos

Você precisa estabelecer permissões para os escopos de arquivo dos usuários ou grupos. Se desejar aplicar o escopo em um novo grupo, crie o(s) grupo(s) como requerido.

Para configurar as permissões de escopo de arquivos, clique no ícone **Operations** e no link **Security** a partir do submenu de navegação.

1. Selecione a aba **File Scopes**.
2. Selecione o escopo a ser modificado. Clique no link **Permissions** desse escopo.



3. As permissões definidas para os usuários e grupos desse escopo serão exibidas.

Disk Usage
Target Clusters
Cluster Processes
System Servers
Security
Resources

Users
Groups
Permissions
File Scopes

Reset

Permissions of collectionx::securedata

Account	allow				deny				Operation
	access	read	write	full	access	read	write	full	
Administrators	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	delete update
Authenticated Users	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	delete update
EmilyKate	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	delete update
Jimmy	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	delete update

Add

- Marque (ou desmarque) as caixas de seleção nas colunas **allow** e **deny** dos usuários ou grupos exibidos.
- Para adicionar usuários ou grupos ao escopo, pressione o botão **Add**.
- A caixa de diálogo Adicionar permissão será exibida.
- Selecione o usuário ou o grupo que deseja adicionar a permissão a partir da lista suspensa.

Disk Usage Target Clusters Cluster Processes System Servers **Security**

Users Groups Permissions **File Scopes**

Reset

Add Permission for collectionx::securedata

Select user: none

Or group: none

allow:

access	read	write	full
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

deny:

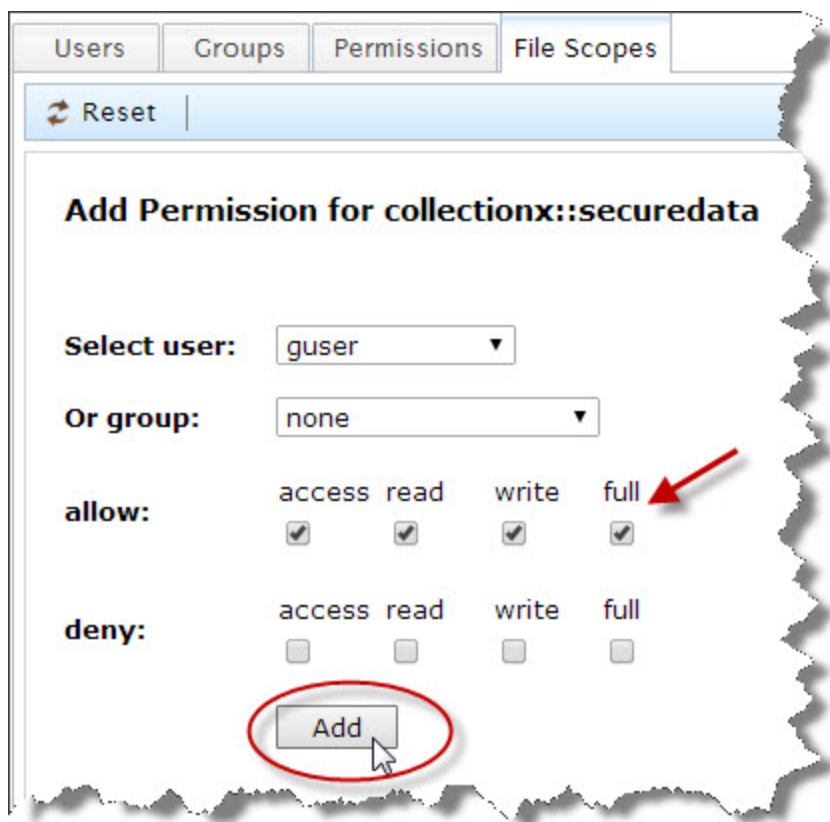
access	read	write	full
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Add

Add user or group permission drop list

Após ter selecionado o usuário ou grupo, o botão Adicionar e as caixas de seleção “permitir” e “negar” serão ativadas.

7. Marque as caixas “permitir” e “negar” para configurar as permissões para este escopo.



8. Pressione o botão **Adicionar** .
9. As alterações serão salvas automaticamente. Feche a(s) aba(s).

Permissões de escopo de arquivos

Abaixo da lista de escopo de arquivos há botões que permitem:

- Redefinir o(s) arquivo(s) selecionado(s) para **Default Permissions** .

Isso permite remover rapidamente quaisquer configurações de permissão adicionadas a um arquivo e redefinir para o acesso padrão.

- Permitir ou negar acesso aos arquivos físicos na zona de entrada de arquivos

Isso oferece uma maneira de permitir ou negar acesso ao escopo de arquivo principal.

Por padrão, apenas os administradores têm acesso a esse escopo.

- Verificar permissões do arquivo para um usuário ou grupo

Isso oferece uma maneira de verificar o acesso de um usuário ou grupo a um arquivo lógico.



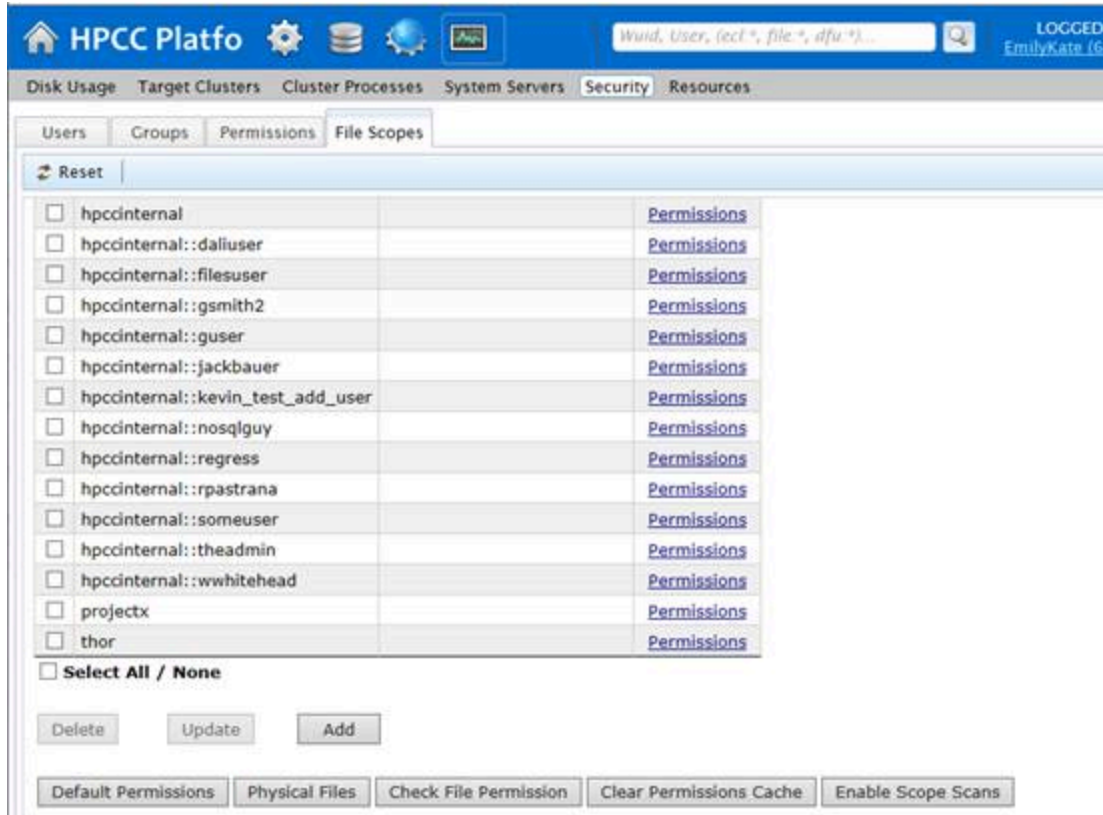
- Limpar cache de permissões

Isso limpa o cache de permissões e permite que quaisquer outras configurações de permissão passem a vigorar de imediato.

- Habilitar/Desabilitar busca de escopo

Isso fornece meios de habilitar ou desabilitar busca de escopo. Habilite a busca de escopos para verificar permissões de acesso aos escopos para os usuários. Este procedimento afetará o desempenho. A função Disable busca de escopo ignora quaisquer permissões de escopo e remove todos os controles de acesso, mas melhora o desempenho. Desabilitar o controle de acesso não é recomendado.

Mudar essa configuração através do ECL Watch, como descrito aqui, é apenas uma substituição temporária. Quando o Dali é reinicializado, essa configuração será revertida para seu estado definido na configuração environment.xml.



Controle de Acesso a Workunit

Existem 2 aspectos sobre a segurança da tarefa (WU) security:

- A Autenticação de recursos para workunit permite configurar permissões para controlar se os usuários podem visualizar suas próprias WUs e/ou as WUs de outros usuários.
- A segurança do escopo do workunit permite configurar permissões para escopos individuais da WU. Todas as workunit possuem um valor de escopo. Todas as tarefas possuem um valor de escopo.

Ambos os métodos podem ser usados (separadamente ou em conjunto), e a restrição mais estrita sempre se aplica.

Em outras palavras, se alguém tiver permissão para ver as WUs no escopo *john DOE*, mas não tem permissão para ver as WUs de outros usuários nas permissões de Autenticação de recursos, esse usuário terá o acesso à visualização de WUs no escopo *john DOE* negado.

Por outro lado, se o usuário tiver permissão de acesso à visualização de WUs de outros usuários, mas não tiver permissão de acesso ao escopo *john DOE* da WU, esse usuário poderá ver outras WUs nesse escopo.

Observação: Caso não tenha acesso à WU, você nunca poderá visualizá-la ou sequer saber da sua existência.

Por padrão, uma WU enviada possui o escopo da ID do usuário. Por exemplo, a WU que JohnDoe enviou possui o valor *scope=johndoe* na WU. Este valor em uma WU permite que ESP e seus serviços LDAP verifiquem as permissões e as coloquem em vigor.

Você pode substituir o escopo padrão usando o código ECL:

```
#workunit('scope', 'MyScopeValue');
```

Protegendo os escopos dos workunits

ESP (na inicialização) cria automaticamente uma OU no LDAP denominada **Workunits** (a menos que ele já exista). Se uma OU de escopo específico não existir no LDAP (p.ex., o escopo johndoe usado no exemplo anterior), as permissões da OU primária serão usadas. Todos os escopos da WU estão localizados abaixo da OU das *workunit*, tanto de forma implícita quanto explícita.

Se uma OU de escopo específico não existir no LDAP (p.ex., o escopo johndoe usado no exemplo anterior), as permissões da OU primária serão usadas. Em outras palavras, o escopo de *johndoe* está implicitamente abaixo da OU *workunit* mesmo não estando listado de maneira explícita na estrutura do LDAP; consequentemente, ele usaria as permissões concedidas às *workunit primárias*.

Permissões de Recursos de Workunits

Ao usar o recurso de **escopos da Workunit** na área **Permissions** do ECL Watch, as permissões de qualquer escopo podem ser redefinidas para as configurações padrão de permissão para seu sistema. As configurações de permissão para os Escopos de tarefa podem ser definidas da seguinte forma:

Descrição	Acesso
View WUs in that scope (Visualização de WUs nesse escopo)	Leitura
Create/modify a WU in that scope (Criar ou modificar uma WU nesse escopo)	Gravação
Delete a WU in that scope (Remover uma WU nesse escopo)	Completo

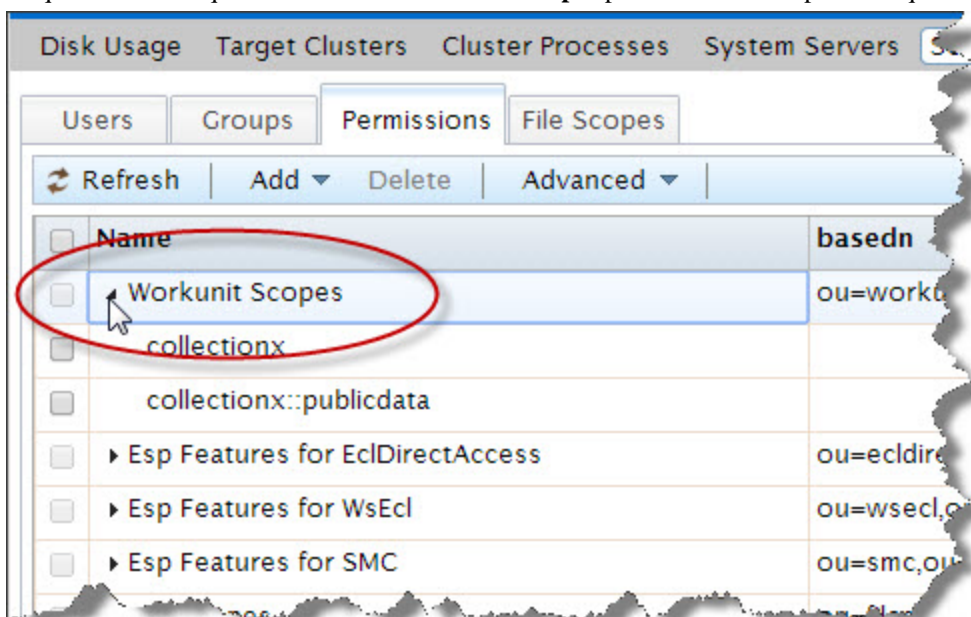
Permissões de Recursos das Workunits

Para adicionar permissões ao escopo de tarefa, clique no ícone **Operations** e no link **Security** a partir do submenu de navegação.

1. Clique na **aba** Permissions.

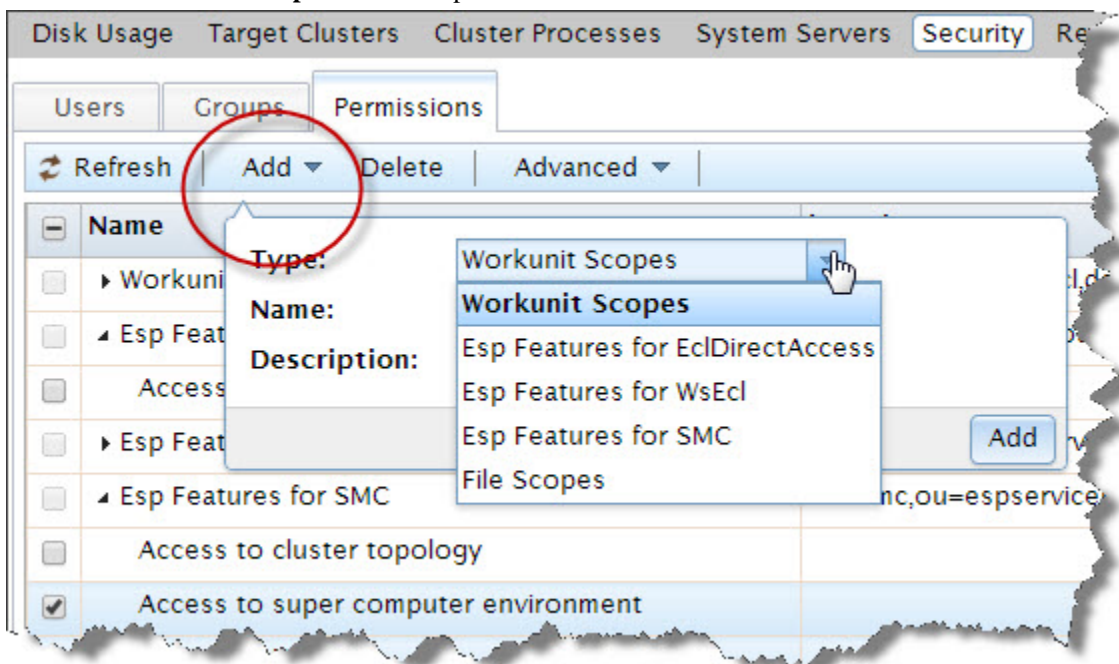
Os recursos específicos serão exibidos.

2. Clique na seta à esquerda do recurso **Workunit Scopes** para exibir os escopos do arquivo.

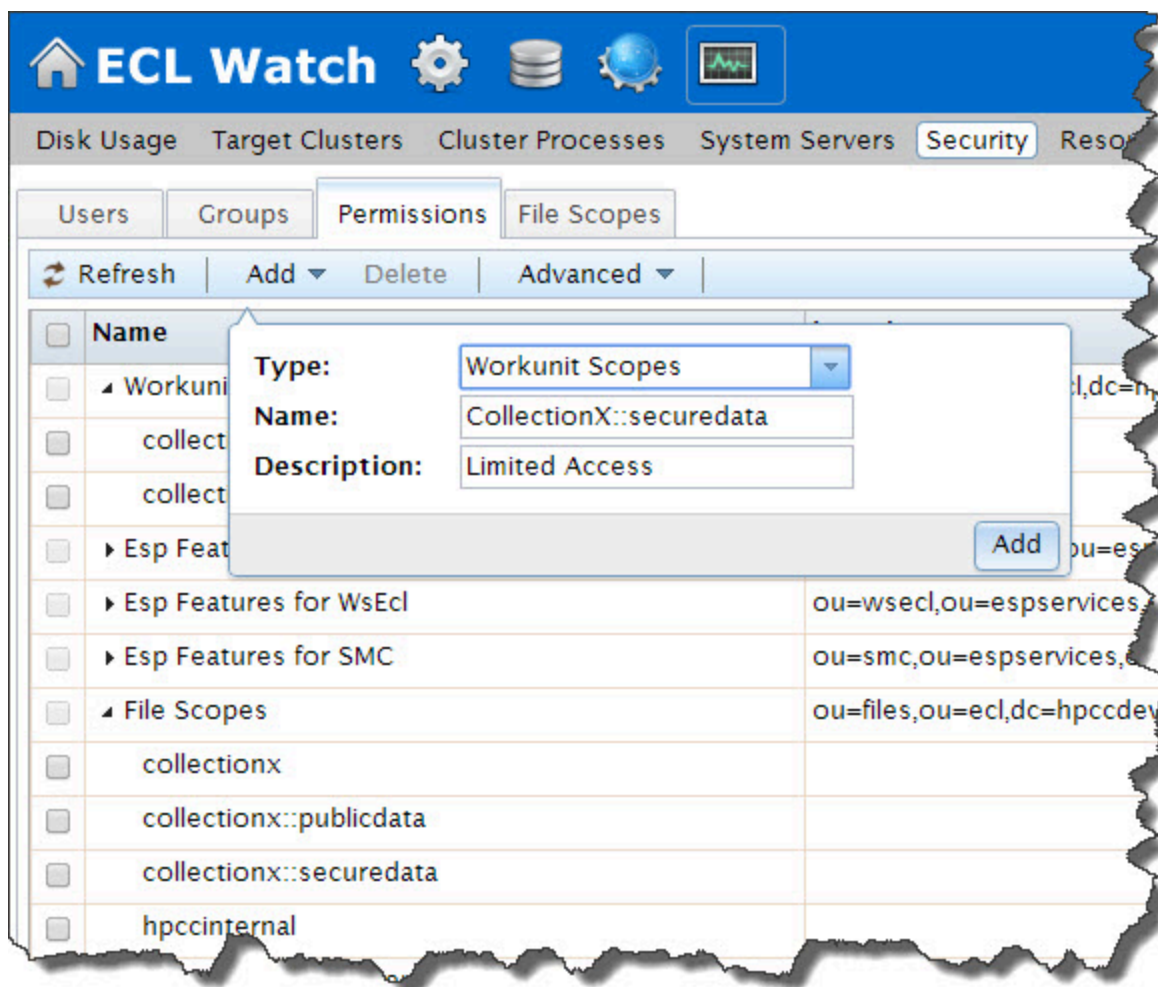


3. Pressione o botão **Add**.

4. Selecione **Workunits Scopes** na lista suspensa.



5. Digite no o nome exato do escopo que deseja adicionar. **Digite no campo Name** o nome exato do escopo que deseja adicionar.



Digite uma breve descrição no campo **Description** .

6. Pressione o botão **Add** .

O novo escopo será exibido na lista.

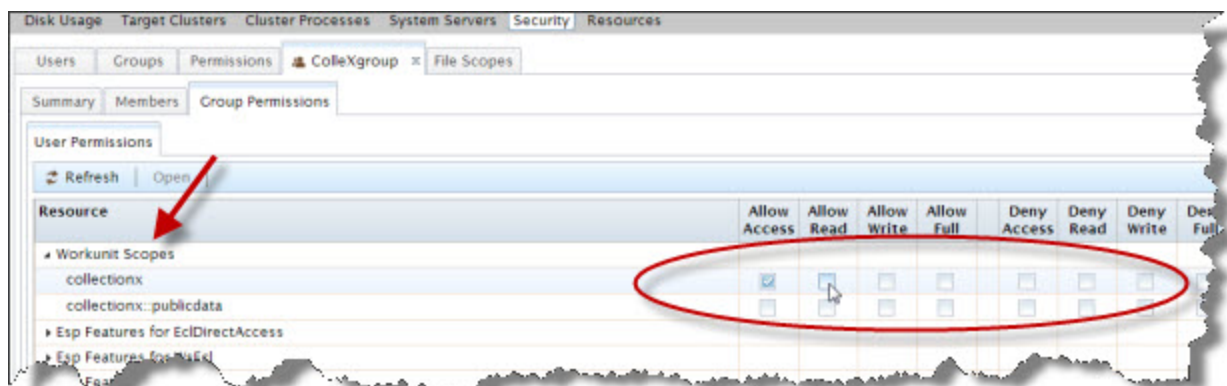
Ajustar permissões do Escopo.

Aplique os escopos de workunit a um grupo. Se desejar aplicar o escopo em um novo grupo, crie o(s) grupo(s) como requerido.

1. Vá para a aba **Groups** .
2. Selecione um grupo no qual deseja aplicar o escopo marcando a caixa ao lado do nome do grupo.

Pressione o botão de ação **Open** . Você pode selecionar múltiplos grupos; uma guia será aberta para cada um deles.

3. Selecione a aba **Group Permissions** para esse grupo. (caso tenha selecionado mais de um grupo, realize esse mesmo procedimento para cada grupo)
4. clique na seta à esquerda de Workunit Scopes para exibir os escopos disponíveis.



Os escopos de tarefa serão exibidos. Marque as caixas de forma adequada para configurar as permissões para este escopo.

5. Para configurar as permissões neste escopo para outro grupo, abra e acesse a guia do grupo desejado.
6. Para configurar permissões neste escopo para um usuário, selecione a guia.
7. Selecione o usuário e pressione o botão de ação Edit.

Uma nova guia será aberta para esse usuário.

8. Nessa aba, clique na sub-aba **User Permissions**.
9. Localize o novo escopo listado no recurso apropriado.

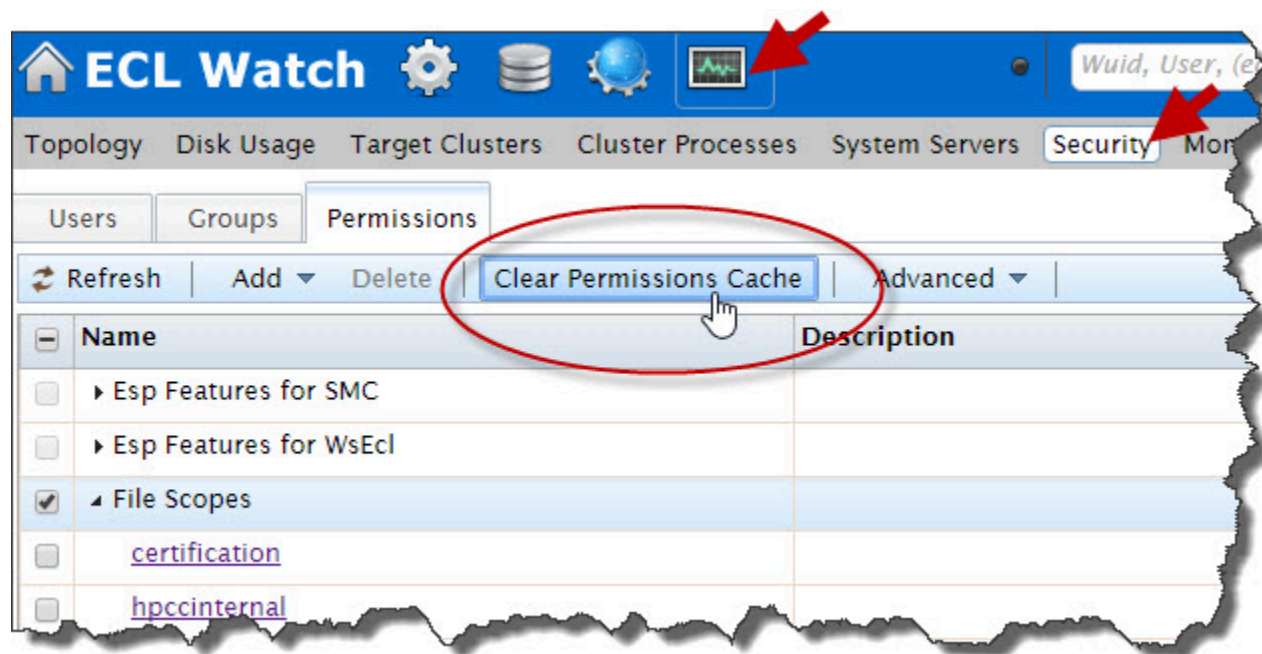
Configure as permissões de acesso da forma adequada para esse usuário.

10. As alterações serão salvas automaticamente. Feche a(s) aba(s).

Cache de Permissão

O botão *Clear Permissions Cache* é um recurso bastante útil e pode ser encontrado na aba Permissions. O botão *Clear Permissions Cache* apaga as permissões armazenadas em cache do Dali e da ESP.

Ao alterar uma permissão no ECL Watch, as configurações são armazenadas em cache no servidor da ESP e armazenadas no servidor Dali. As informações armazenadas em cache são atualizadas em um intervalo configurável. Esse intervalo pode ser definido no Configuration Manager, na aba *LDAP Server settings Attributes*. O tempo limite padrão do cache é 5 minutos.

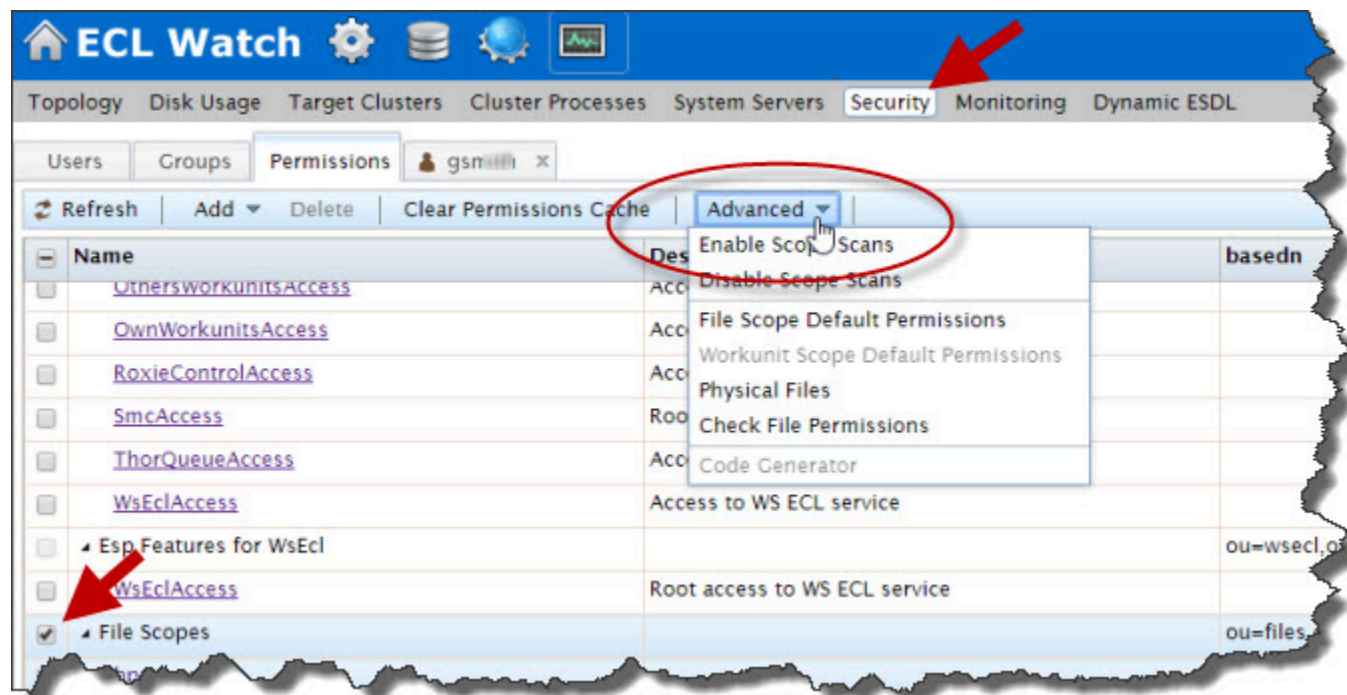


O Cache de permissões pode ser apagado de qualquer lugar na aba ECL Watch Permissions.

Se quiser que uma alteração de permissão comece a vigorar de imediato, apague o cache e force o Dali a atualizar as configurações de permissão pressionando o botão **Clear Permission Cache**. Esta ação transfere as configurações ao pressionar o botão. Use esse recurso criteriosamente, pois o desempenho geral do sistema será temporariamente afetado enquanto as configurações do LDAP são preenchidas novamente na Armazenagem de dados do Dali System.

Permissões Avançadas

Trata-se do botão **Advanced** (Permissions) localizado na aba Permissions. O botão/menu Advanced oferece acesso ao gerenciamento da segurança do escopo de arquivos e workunit. O botão Advanced é habilitado apenas quando você selecionar Files Scope ou Workunit Scopes na aba Permissions.

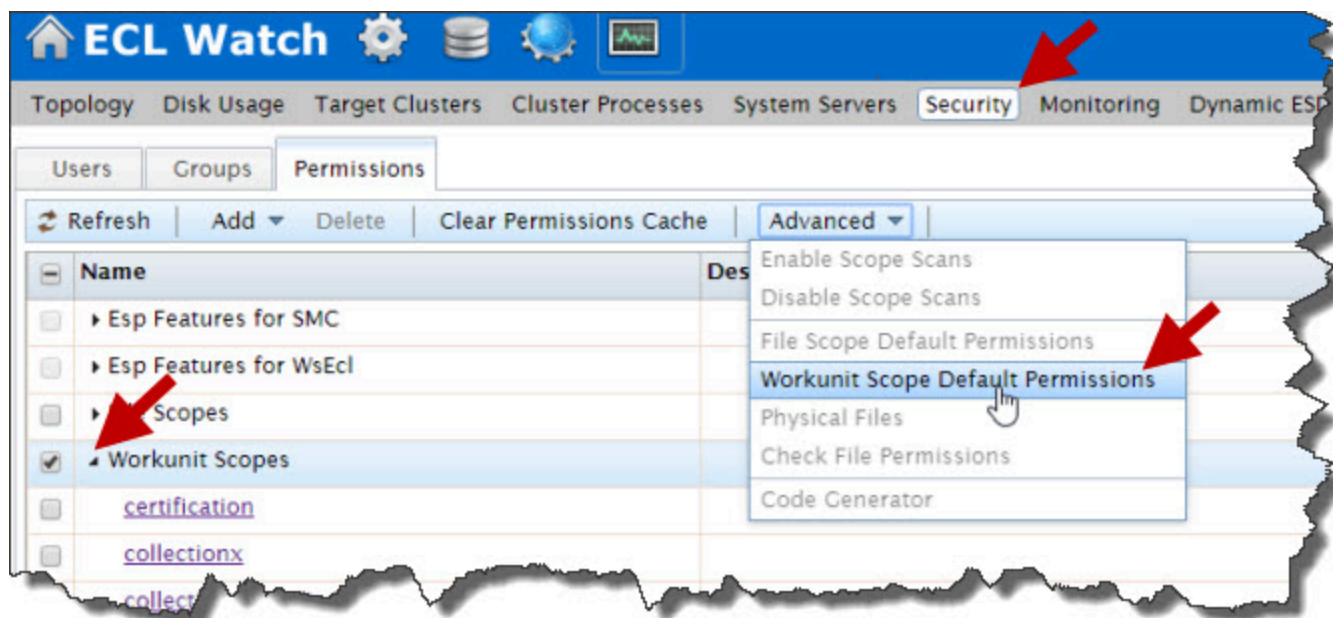


Pressione o botão Advanced para exibir o menu de permissões avançadas. O menu Advanced é sensível ao contexto, portanto, se você selecionar Escopo de arquivos na aba Permissions, poderá optar apenas por aplicar as permissões dos escopos de arquivo que são relevantes; isso também se aplica ao selecionar escopos de tarefa.

Verificações de arquivos

Usando o menu Avançado com a opção Escopo de arquivos selecionada:

- Habilite ou desabilite a segurança do escopo de arquivos
- Acesse a página de permissão padrão do escopo de arquivos
- Acesse a página de permissão para Arquivos físicos.
- Verificar as permissões do arquivo – Essa opção abre uma caixa de diálogo onde é possível inserir um nome de arquivo e selecionar usuários ou grupos para o escopo de segurança do arquivo.



Verificação de tarefa

O uso do menu Avançado com a opção Escopos de tarefa selecionada abre apenas a guia Permissões padrão das Permissões dos Escopos de tarefa (Padrão).

OBSERVAÇÃO: A segurança dos escopos de tarefa ou de arquivo precisa estar habilitada nas configurações do seu sistema para que você possa usar o recurso de segurança dos escopos de tarefa ou de arquivo no sistema.

Workunits e Active Directory

O desempenho do seu sistema pode variar dependendo da forma de interação de alguns componentes. Uma área que poderia influenciar o desempenho é a relação com usuários, grupos e o Active Directory. Se possível, pode ser uma boa política contar com um Active Directory individual e específico para o HPCC. Pode haver alguns casos onde apenas um Active Directory que atende aos vários aplicativos diversos não ofereça um desempenho ideal.

O HPCC torna a configuração da OU do Active Directory relativamente fácil. Quando iniciado, o ESP cria todas as UOs para você com base nas configurações definidas no Configuration Manager. É possível iniciar o Dali/ESP e usar o ECLWatch para adicionar ou modificar usuários ou grupos.

É possível atribuir permissões para cada usuário individualmente, mas é mais gerenciável atribuir essas permissões a grupos e depois adicionar usuários a esses grupos, conforme apropriado. Crie um grupo para desenvolvedores e usuários avançados (pessoas com acesso completo de leitura/gravação/exclusão), outro grupo para usuários que terão apenas acesso de leitura, e talvez outro para aqueles com acesso de leitura e gravação. Adicione quaisquer outros grupos em seu ambiente conforme apropriado. Agora é possível atribuir usuários aos seus grupos adequados.

Active Directory e semelhança com LDAP

Há componentes que são comuns para o Active Directory e para o LDAP. Há certos termos relevantes que podem ainda precisar de uma explicação mais detalhada.

filesBasedn	Trata da restrição de acesso aos arquivos. Também é citado como "escopo de arquivo".
groupsBasedn	Controla os grupos associados com o ambiente. Por exemplo: administradores, desenvolvedores, apenas ws_ecl e afins.
modulesBasedn	Específico para sistemas que usam um repositório central de legado e controla o acesso a módulos específicos. Qualquer módulo criado na aplicação criará uma entrada em Eclwatch>>User/Permissions>>Repository Modules
sudoersBasedn	Obsoleto.
workunitsBasedn	Controla o acesso às workunits.

Utilizando o Cassandra para armazenamento de Workunit.

A partir da versão 6.0.0, é possível configurar seu sistema para usar o banco de dados Cassandra™ da Apache para armazenamento de workunits. O Cassandra NoSQL data store da Apache é um framework de software com licença gratuita que suporta aplicações de distribuição intensa de dados.

A configuração padrão do HPCC continua idêntica – as Workunits são armazenadas no data store do Dali e o Sasha é normalmente configurado para arquivar Workunits em disco a fim de economizar o uso de memória.

Alternar para o armazenamento de Workunits no Cassandra elimina a necessidade de arquivá-las e permite que você gerencie seus dados de forma independente.

Esta seção detalha as etapas para configurar a plataforma HPCC a usar Cassandra como um servidor de Workunits.

Cassandra como um servidor Workunit

Prós e contras

Dali	Cassandra
Armazenamento no Dali não requer trabalho ou instalação adicional de software	Exige a instalação e a configuração de um servidor Cassandra
Sasha fornece serviço de limpeza simples	Exige sua própria organização e manutenção
Limitado pela memória do Dali	Ilimitado (escalabilidade elástica)
Workunits arquivadas são difíceis de acessar, devem ser restauradas antes que o resultado seja visto	Não é necessário arquivar, de forma que todas as workunit permanecer "em tempo real"
Localização única	Distribuído, escalonável e altamente disponível.
Backup único de dados suportados. Fail over manual	Tolerância a falhas inerente com replicação de dados, projetada para alta disponibilidade.

Instalação e Visão Geral da Configuração

1. Instale o Cassandra e configure-o como desejar. Os únicos requisitos para trabalhar com uma plataforma HPCC são:
 - a. Crie um keyspace para usar para o armazenamento das workunits no Cassandra. O padrão é hpcc.
 - b. Se quiser segurança, é necessário configurar a autenticação do Cassandra e criar um superusuário para o HPCC usar. Essas credenciais precisam então ser inseridas na configuração do DaliPlugin no Gerenciador de Configurações.
 - c. Verifique se é possível acessar o servidor Cassandra pelos nós do HPCC .
 - d. Conheça as configurações do Cassandra (IP, porta, keyspace, credenciais de usuário etc.)
2. Use o Gerenciador de Configurações para definir a plataforma HPCC para usar o Cassandra para armazenamento das Workunits.
3. Certifique-se de que o Cassandra esteja sendo executado.

Depois de ter configurado o HPCC para usar o Cassandra para armazenamento das workunits, é necessário garantir que o Cassandra esteja em execução antes de iniciar o sistema.
4. Force o arquivo environment.xml e reinicie a plataforma HPCC.
5. Se possuir Workunits existentes, exporte-as do Dali e importe-as no Cassandra usando a wutool.
6. Se possuir Workunits arquivadas e quiser ativá-las, importe-as para o Cassandra.
 - a. Depois que as Workunits forem armazenadas no armazém de Workunits do Cassandra, o Sasha não vai mais arquivar Workunits (independentemente das configurações que possa ter)
 - b. Quaisquer Workunits previamente arquivadas permanecerão armazenadas no servidor Sasha e continuarão disponíveis da mesma forma que antes.

Configure sua plataforma para utilizar o Cassandra para o armazenamento de Workunit

Essa opção é configurada ao adicionar um componente DaliPlugin ao seu sistema.

Conectar-se ao Gerenciador de Configurações (Configuration Manager)

Para alterar a configuração para os componentes do HPCC , conecte-se ao Configuration Manager.

1. Pare todos os componentes do HPCC se estiverem em execução.
2. Verifique se eles não estão mais sendo executados. É possível usar um único comando, como:

```
sudo /opt/HPCCSystems/sbin/hpcc-run.sh -a hpcc-init status
```

3. Inicie o Gerenciador de Configurações.

```
sudo /opt/HPCCSystems/sbin/configmgr
```

4. Conecte seu navegador de Internet à interface da Web do Gerenciador de Configurações.

(usando o URL `http://<configmgr_IP_Address>:8015`, onde `<configmgr_IP_Address>` , é o endereço IP do nó que está executando o Gerenciador de Configuração)

5. Selecione o **botão Advanced View** .
6. Use a lista suspensa para selecionar o arquivo de configuração XML adequado.

Observação: Gerenciador de Configurações. **nunca** funciona no arquivo de configurações ativo. Após terminar a edição, será necessário copiar o arquivo `environment.xml` para o local ativo e forçá-lo a todos os nós.

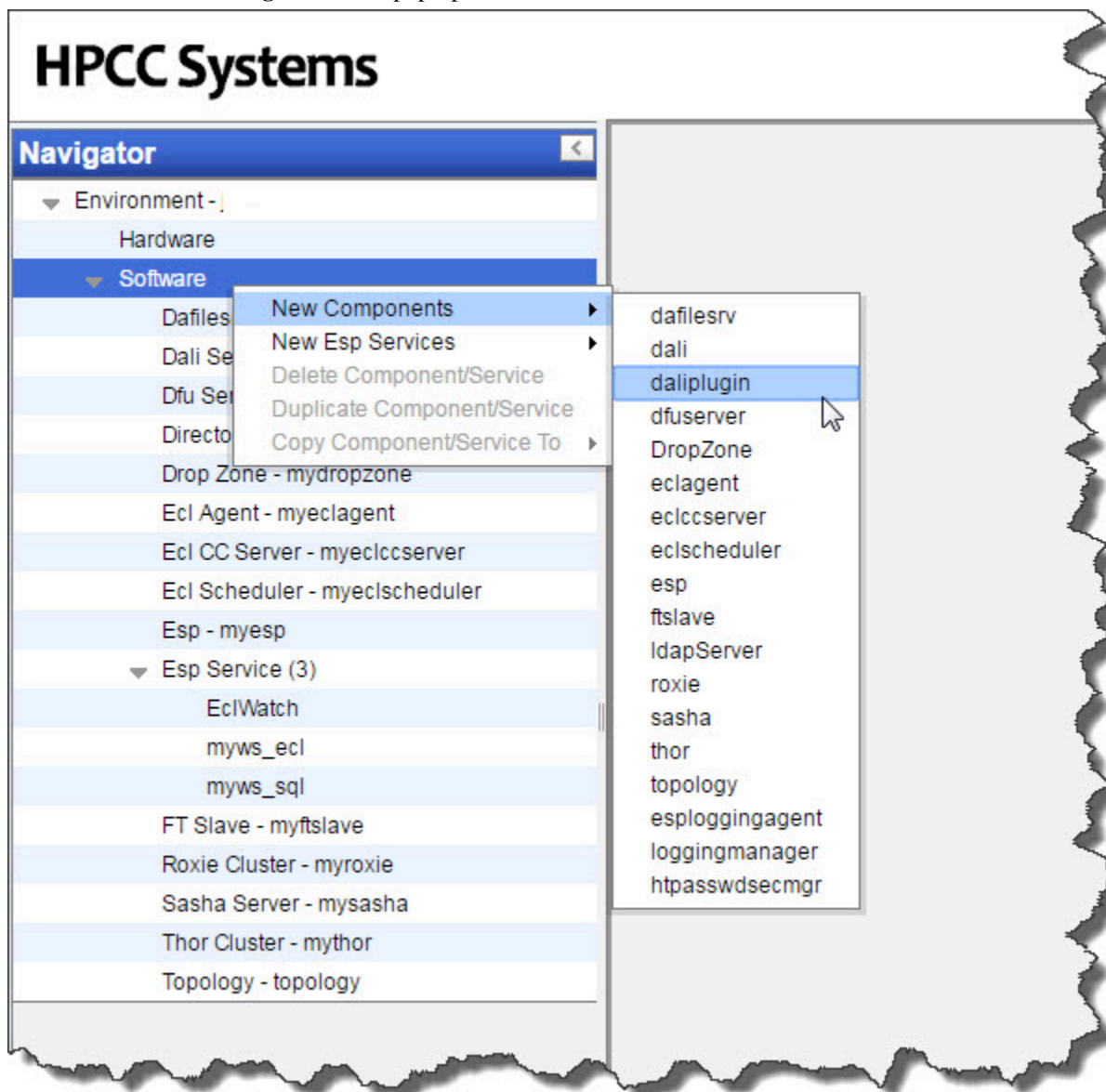
7. Marque a caixa de seleção **Write Access** .

O acesso padrão é somente leitura. Muitas opções estão disponíveis apenas quando o acesso à gravação estiver ativado.

Adicionando o componente DaliServerPlugin

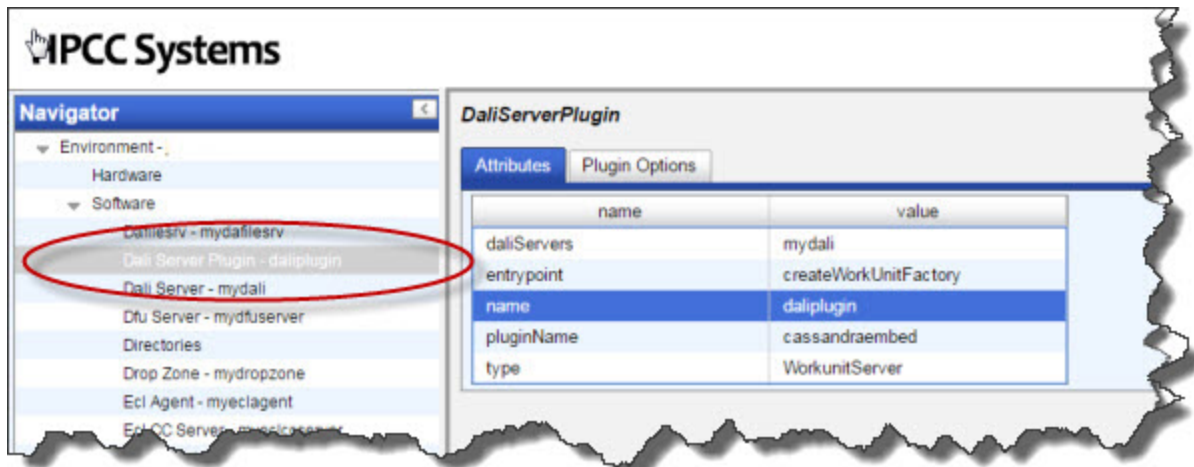
O DaliServerPlugin controla suas opções de armazenamento de workunits no Cassandra.

1. Clique com o botão direito no Pannel de **Navigator** e selecione **New Components** no menu pop-up; em seguida, selecione **DaliServerPlugin** no menu pop-up.

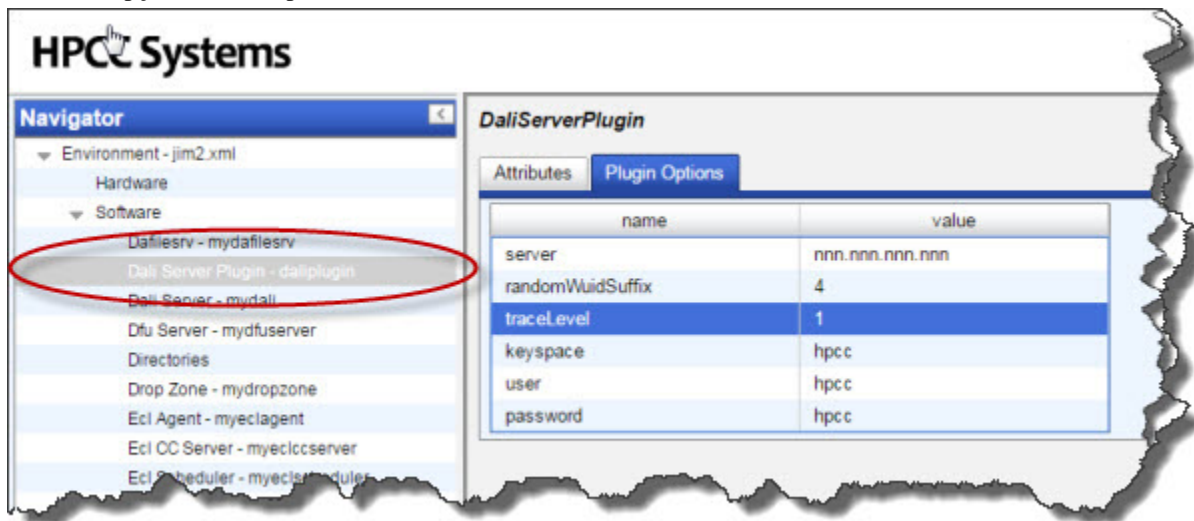


2. Insira as propriedades do **DaliServerPlugin** :

- a. Preencha a aba **Attributes** com as configurações adequadas de seu Cassandra Server existente.



- b. Adicione opções na aba Options.



Observação: O nome de usuário e senha são necessários apenas caso seu servidor Cassandra esteja configurado para exigir credenciais.

- c. Clique no ícone de disco para salvar.

3. Copie o novo arquivo xml do diretório de origem para /etc/HPCCSystems e renomeie o arquivo para environment.xml

```
# for example sudo cp /etc/HPCCSystems/source/NewEnvironment.xml  
/etc/HPCCSystems/environment.xml
```



Certifique-se de que você possua direitos suficientes para gravar arquivos no diretório de destino antes de copiar. Se solicitado a substituir o arquivo de destino, responda **yes**.

4. Copie o arquivo /etc/HPCCSystems/environment.xml para /etc/HPCCSystems/ em cada nó.

É possível criar um script para distribuir o arquivo XML para todos os nós. Um script de amostra é fornecido com o HPCC. Os comandos a seguir copiam os arquivos XML para todos os nós como exigido:

```
sudo /opt/HPCCSystems/sbin/hpcc-push.sh <sourcefile> <destinationfile>
```

5. Reinicie o HPCC System em **todos** os nós.

Utilizando o wutool

wutool *action* [WUID=nnn] [DALISERVER=ip | CASSANDRASERVER=ip] [option=value]

O wutool é um utilitário de linha de comando usado para manter seu armazenamento das workunits. Ele pode ser encontrada em /opt/HPCCSystems/bin/ em qualquer servidor onde a plataforma tenha sido instalada. É possível usar esse utilitário para exportar suas Workunits de um servidor Dali e depois importar para um data storage do Cassandra. Também é possível usá-lo para importar workunits arquivadas em um servidor Sasha.

Ações	
list <workunits>	Listar Workunits
dump <workunits>	Descarregar o xml para Workunits especificadas.
delete <workunits>	Remover Workunits.
results <workunits>	Descarregar resultados para as workunits especificadas.
archive <workunits>	Arquivar Workunits especificadas em arquivos xml. Há suporte para as seguintes opções: [TO=<directory>] [DEL=1] [DELETERESULTS=1] [INCLUDEFILES=1]
restore <filenames>	Restaurar a partir de arquivos xml. [INCLUDEFILES=1]
orphans	Remover informações órfãs do armazenamento
cleanup [days=NN]	Remover Workunits com mais de NN dias
validate	Verificar se há erros no conteúdo do repositório de Workunits. [FIX=1] tentará reparar quaisquer problemas encontrados.
clear	Remover todo o repositório da workunit (exige entire=1 repository=1)
initialize	Inicializar o novo repositório de workunit

<workunits> pode ser especificado na linha de comando ou usando um filter owner=XXXX. Se omitido, todas as workunits são selecionadas.

Opções	
(Essas opções são válidas se o CASSANDRASERVER estiver especificado)	
CASSANDRA_KEYSPACE	Keyspace do Cassandra, o padrão é HPCC.
CASSANDRA_USER	Nome de usuário do Cassandra a ser usado (se necessário).
CASSANDRA_PASSWORD	Senha do Cassandra a ser usada (se necessário).
TRACELEVEL	Trace level (1-10)

Exemplos:

```
/opt/HPCCSystems/bin/wutool archive DALISERVER=. del=1  
  
/opt/HPCCSystems/bin/wutool restore *.xml  
CASSANDRASERVER=192.168.56.120 CASSANDRA_KEYSPACE=hpcc2
```

Redefinindo nós em um Cluster Thor

Para configurar um cluster Thor onde você substitui os nós atuais (com novos IPs) ou adiciona ou remove nós, é necessário adotar uma etapa adicional para reestruturar o grupo. O Dali não reestruturará automaticamente um grupo existente.

Isso porque os arquivos publicados atuais referem-se ao estado de grupo de cluster anterior pelo nome e, por isso, alterar sua estrutura invalidaria esses arquivos e tornaria os arquivos físicos inacessíveis.

Há algumas situações nas quais você optaria por redefinir seu cluster Thor.

Substituindo nó(s) defeituoso(s)

Se os arquivos de dados forem replicados, pode ser útil substituir um nó e forçar o uso do novo grupo pelos arquivos existentes. Nesta situação, a leitura de um arquivo existente provocará um failover para encontrar uma parte no nó replicado ao tentar localizar um arquivo físico no novo nó de reposição.

Para forçar o uso do novo grupo, utilize o seguinte comando:

```
updt dalienv <environment_file> -f
```

Em casos onde não há replicação, a perda de dados pode ser inevitável e forçar o novo grupo ainda pode ser a melhor opção.

Redimensionando o cluster

No caso da adição ou remoção de nós de cluster Thor, mas *todos os nós originais continuam no ambiente e acessíveis*, é necessário **renomear** o grupo associado ao cluster Thor (ou nome do cluster se não houver nome do grupo).

Isso garantirá que todos os arquivos anteriores continuem a usar a estrutura de grupo antiga, enquanto novos arquivos usam a nova estrutura de grupo.

Em resumo, se o cluster Thor mudar, ele precisa ser atualizado no Dali.

Melhores práticas:

Este capítulo descreve as diversas formas das boas práticas traçadas por usuários e administradores de longa data do HPCC que executam o HPCC em um ambiente de produção exigente e de alta disponibilidade. No entanto, não é necessário executar seu ambiente dessa forma, uma vez que seus requisitos específicos podem variar. Esta seção oferece algumas recomendações de boas práticas traçadas após vários anos de execução do HPCC em um ambiente de produção intenso e exigente.

Redundância de Cluster

Há vários aspectos de redundância de cluster que devem ser considerados ao configurar seu HPCC System.



Lembre-se de alocar recursos amplos para seus principais componentes. Dali exige bastante memória RAM. ECL Agent e ECL Server dependem do processador. Thor deve ter pelo menos 4GB de memória RAM por nó.

Dali

Dali deve ser executado em uma configuração ativa/passiva. Os termos ativo/passivo significam que há dois Dalis em execução: um primário, ou ativo, e outro passivo. Nesta situação, todas as ações são executadas no Dali ativo, mas duplicadas no passivo. Se o Dali ativo falhar, é possível realizar o fail over para o passivo.

Outra boa prática sugerida é usar o cluster padrão com um quórum e um VIP de tomada de controle (um tipo de balanceador de carga). Se o Dali primário falhar, transfere-se o VIP e o diretório de dados para o nó passivo e reinicia-se o serviço Dali.

Servidor DFU

É possível executar múltiplas instâncias do DFU Server. É possível executar todas as instâncias como ativas, ao contrário de uma configuração ativa/passiva. Não há necessidade de um balanceador de cargas ou VIP. Cada instância consulta rotineiramente o Dali por workunits. Caso uma falhe, as outras continuarão extraindo novas workunits.

ECLCC Server

É possível executar múltiplas instâncias ativas do ECLCC Server para redundância. Também não há necessidade de um balanceador de cargas ou VIP para isso. Cada instância verificará tarefas rotineiramente. Caso uma falhe, as outras continuarão realizando a compilação.

ESP/ECL Watch/WsECL

Para garantir a redundância, coloque os ESP Servers em um VIP. Para um design ativo/ativo, é necessário usar um balanceador de cargas. Para o ativo/passivo, é possível usar um precursor/contador (pacemaker/heartbeat). Se você optar pela configuração ativa/ativa, é necessário manter uma única conexão de cliente com apenas um servidor durante a sessão do ECL Watch (porta 8010). Outros serviços, como o WsECL (porta 8002) não exigem uma conexão persistente com um único servidor.

ECL Agent

É possível executar múltiplas instâncias ativas do ECL Agent. Não é necessário um balanceador de cargas ou VIP. Cada instância consulta tarefas rotineiramente. Caso uma falhe, as outras continuarão extraindo novas workunits.

Sasha

Sasha deve ser executado em uma configuração ativa/passiva. Os termos ativo/passivo significam que há dois Sashes configurados, um primário (ativo) e outro em espera.

ECL Scheduler

Não é necessário um balanceador de carga na execução do design ativo/ativo. Cada instância consulta workunits rotineiramente. Caso uma falhe, as outras continuarão agendando as workunits.

Thormaster

Defina o Thor em uma configuração ativa/passiva. Os termos ativo/passivo significam que há duas instâncias em execução, uma primária (ativa) e outra passiva. Nenhum balanceador de cargas é necessário. Se a instância ativa falhar, é possível realizar o fail over para a passiva. O fail over então usa o VIP (um tipo de balanceador de cargas) para distribuir quaisquer solicitações recebidas.

Dropzone

Esse é apenas um servidor de arquivos que executa o processo dafilesrv. Configure da mesma maneira como qualquer servidor de arquivos ativo/passivo. Um primário, ou ativo, e outro passivo. Nenhum balanceador de cargas é necessário. Se a instância ativa falhar, é possível realizar o fail over para a passiva.

Alto Disponibilidade

Se precisar de alta disponibilidade para seu HPCC System, há algumas considerações adicionais que devem ser analisadas.

Esta não é uma lista completa e não visa fornecer instruções passo a passo para configuração de recuperação de desastres. Em vez disso, a seção apenas oferece algumas informações adicionais que devem ser levadas em conta ao incorporar o HPCC em seu plano de recuperação de desastres.

Thor

Ao projetar um cluster Thor para alta disponibilidade, considere como ele de fato funciona – um cluster Thor aceita tarefas a partir de uma fila de tarefas. Se houver dois clusters Thor atendendo a fila de tarefas, um deles continuará aceitando tarefas se o outro falhar.

Com a replicação ativada, o Thor que ainda funciona poderá ler os dados do local de backup do Thor que falhou. Outros componentes (como ECL Server ou ESP) também podem ter múltiplas instâncias. Os componentes restantes, como Dali ou DFU Server, funcionam em um modelo tradicional de fail over de alta disponibilidade de armazenamento compartilhado.

Outra consideração importante é manter o ESP e o Dali em nós separados do Thor Master. Desta forma, se o mestre Thor falhar, você pode substituí-lo, mantendo a substituição com o mesmo IP (endereço). Uma vez que o Thor não armazena dados de workunit, o DALI e o ESP podem providenciar os metadados de arquivo para recuperar suas tarefas.

A Desvantagem

Os custos iniciais são até duas vezes maiores, uma vez que será necessário adquirir basicamente tudo em dobro.

A Vantagem

É possível utilizar a capacidade de processamento adicional em quase 100% do tempo. É possível executar mais jobs, ter mais espaço e afins.

Observações sobre Disaster Recovery (Recuperação de Desastres)

O fator importante a ser considerado para a recuperação de desastres (DR) é a largura de banda necessária para replicar os dados. Seu administrador de rede deve avaliar esse aspecto com atenção.

Se possui dezenas de gigabytes de deltas todos os dias, então uma replicação do tipo rsync ou algum tipo de modelo híbrido deve ser suficiente. Se você possui centenas de gigabytes ou petabytes de deltas, o limite real é o seu orçamento.

Recomenda-se encontrar onde os dados são menores (na ingestão, após a normalização, no Roxie), replicá-los deste ponto e reexecutar o processamento em ambos os locais.

O segredo para realizar a recuperação de desastres corretamente é conhecer o seu fluxo de dados. Por exemplo, se estiver ingerindo 20TB de dados brutos diariamente e realizar sua implementação, classificação e indexação e afins. Seria melhor replicar um conjunto de dados intermediário (que chamamos de arquivos de base) em vez de replicar a ingestão elevada. Se a situação for contrária (pequena ingestão diária e depois o aumento do tamanho dos dados), seria melhor ingerir na entrada e depois reexecutá-los.

Thor tem a capacidade de realizar uma "cópia Thor", que copia dados de um cluster para outro. Também é possível fazer isso através de um código ECL. Além disso, pode-se optar por não ter, ou precisar de ter um DR Thor "moderno". Neste caso, as falhas leves mais comuns causariam apenas um desastre relativamente pequeno e inferior a 1 dia. Uma

vez que o Thor é responsável por criar atualizações de dados, ele pode levar um ou alguns dias para recuperar. Os dados apenas não são tão novos, mas continuarão fluindo contanto que os Roxies ainda sejam replicados. No caso de um desastre grave, como um grande terremoto, uma onda gigantesca, perda de energia total prolongada, múltiplos cortes de fibra ótica, onde os sistemas ficarão fora do ar por um dia ou mais. A probabilidade disso ocorrer não justifica os custos da prevenção.

Conclusão

A recuperação de desastres é uma equação. O custo de falha multiplicado pela probabilidade de o evento acontecer anualmente ser menor ou maior do que o custo para evitá-lo. Levar tudo isso em consideração pode ajudar a colocar um plano de DR razoável em ação.

Roxie

No caso do Roxie, recomenda-se contar com múltiplos clusters Roxie e usar um proxy para manter o equilíbrio. Para manter os dados sincronizados, a abordagem de extração é a mais recomendada. O Roxie extrai automaticamente os dados de que precisa da "fonte" listada no arquivo do pacote. Os dados também podem ser extraídos de outro Roxie ou de um Thor. Na maioria dos casos, seu DR Roxie seria extraído do Roxie primário do balanceador de carga, mas ele também pode ser extraído de um Thor no local primário.

Middleware

A replicação de alguns componentes (ECL Agent, ESP/Eclwatch, DFU Server, etc.) é bastante simples, uma vez esses componentes não possuem nada para replicar. O Dali é o principal elemento quando se fala em replicação. No caso do Dali, tem-se o Sasha como o backup local. Os arquivos Dali podem ser replicados usando o rsync. Uma melhor abordagem seria usar um dispositivo de sincronização (sincronização de WAN de cluster, replicação de bloco de SAN, etc.) e simplesmente colocar os armazenamentos Dali nele e permitir que a replicação seja feita conforme projetado.

Não há uma abordagem universal. Para garantir uma estratégia efetiva de DR que não realize uma "sincronização excessiva" entre conexões WAN lentas, mas que ainda lhe ofereça um nível aceitável de redundância para suas necessidades de negócios, é necessário ter cuidado, design e planejamento especiais.

Considerações sobre Melhores Práticas:

Há vários outros aspectos para as considerações de boas práticas, e esses aspectos mudarão de acordo com os requisitos do seu sistema. As seções a seguir englobam algumas considerações de boas práticas para determinados aspectos do HPCC System. Tenha em mente que as boas práticas sugeridas são apenas recomendações e podem não ser adequadas às suas necessidades. Uma análise minuciosa das considerações destacadas nesta seção pode ser útil se as suas necessidades se alinharem às considerações informadas.

Múltiplos Thors

É possível executar múltiplos Thors em um mesmo hardware físico. Os múltiplos Thors em um mesmo hardware são independentes e não sabem da existência do outro. Os Thors executam jobs à medida em que os recebem, independentemente do que os outros estejam fazendo. A velocidade de uma única tarefa nunca será superior com múltiplos Thors, porém a produtividade poderá ser maior. É possível executar duas tarefas de coleta de Thors a partir de duas filas diferentes ou em uma mesma fila.

A desvantagem de executar múltiplos Thors no mesmo hardware é que a memória física nos nós precisa ser compartilhada entre cada um dos Thors. Isso precisa ser configurado segundo a definição de cluster Thor.

Não é possível colocar múltiplos Thors em um hardware que não possua núcleos de CPU suficientes para suportá-los. Você não deve ter uma quantidade maior de Thors em relação ao número de núcleos. Uma boa regra é usar uma fórmula na qual o número de núcleos dividido por dois equivale ao número máximo de clusters Thor a ser usado.

Múltiplos nós

Se possível, tente manter os recursos em execução nos seus próprios nós para um ou múltiplos clusters Thor. Se estiver executando algum tipo de alta disponibilidade ativa/passiva, não mantenha seu Master ativo e passivo no mesmo nó. Tente manter o Dali e o ESP em nós separados. Mesmo que você não possa se dar ao luxo de ter diversos nós, recomenda-se que o Thor Master e o Dali (no mínimo) estejam em nós separados. A boa prática é manter o máximo de componentes possível em seus próprios nós.

Outra consideração para um sistema de múltiplos nós é evitar colocar qualquer um dos componentes em nós com escravos. Essa não é uma boa prática por acarretar em um cluster desequilibrado, resultando em escravos com menos memória/CPU que são mais lentos do que os demais e, consequentemente, reduzindo o desempenho de todo o cluster.

Tempo-limite do Thor

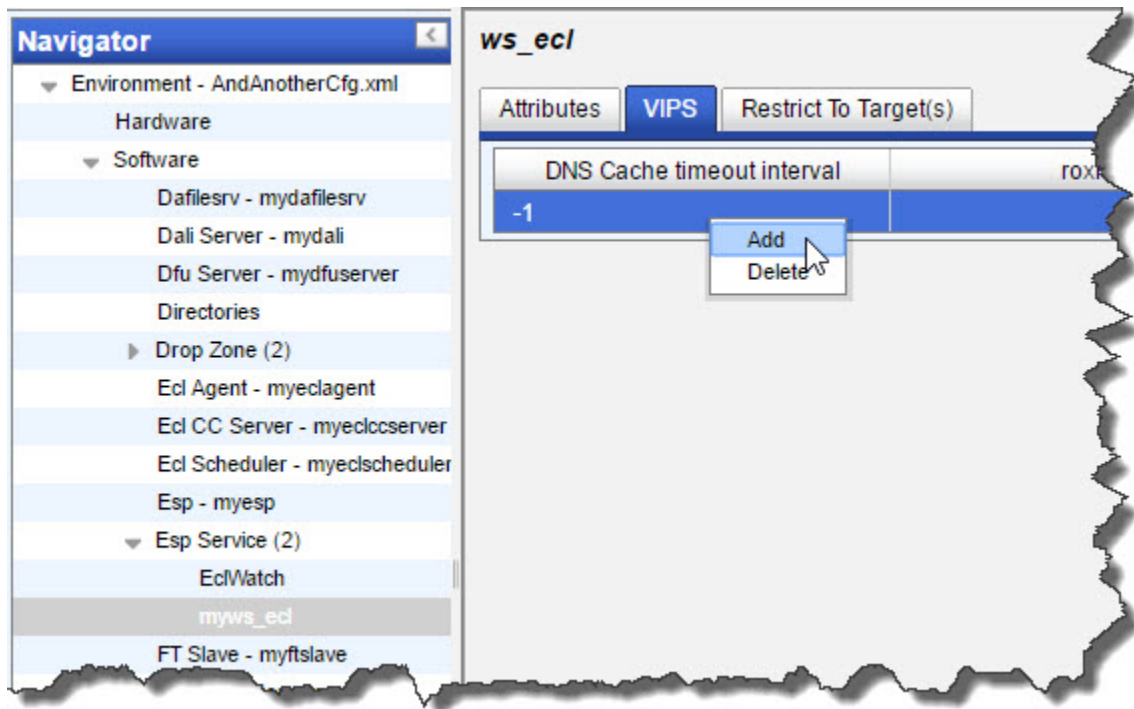
Há um caso no qual uma política ou prática de sistema poderia causar um problema com os nós Thor. Se um cluster Thor travar na inicialização e o tempo-limite acabar expirando. Depois, se o seu log Thormaster mostrar que o Master não apresenta problemas, mas indica que está aguardando se conectar aos escravos. Você pode então ter problemas com a configuração do daemon SSH.

Há um recurso de segurança chamado *"AllowUsers"* que cria uma lista de permissões no sshd (o processo de servidor OpenSSH) que vai bloquear conexões de qualquer um que não estiver presente na lista. Esse não é um padrão para sshd, mas uma opção que precisa ser ativada. Se essa opção for ativada, isso pode fazer com que os nós Thor travem na maneira descrita. Se essa opção for ativada, é necessário desativá-la ou adicionar o usuário hpcc à lista de AllowUsers.

Múltiplos Clusters Roxie

É possível configurar múltiplos clusters Roxie. Quando se tem múltiplos clusters Roxie, é melhor usar um balanceador de cargas com eles. Para configurar múltiplos clusters Roxie, comece adicionando seu Roxie na guia VIPS no Configuration Manager.

Figure 27. Configurar o VIP



Abra o Configuration Manager do HPCC e prossiga para a opção Advanced View. Para obter mais informações sobre o uso do ConfigMgr, consulte "Como usar o Configuration Manager".

1. Selecione seu ESP Service (o padrão é o **myws_ecl**) no painel Navegador ao lado esquerdo da tela.
2. Selecione a guia **VIPS**.
3. Clique com o botão direito sobre a tabela e selecione *Add*. (veja a imagem acima)
4. Defina o valor **Send Target To** para *False*.

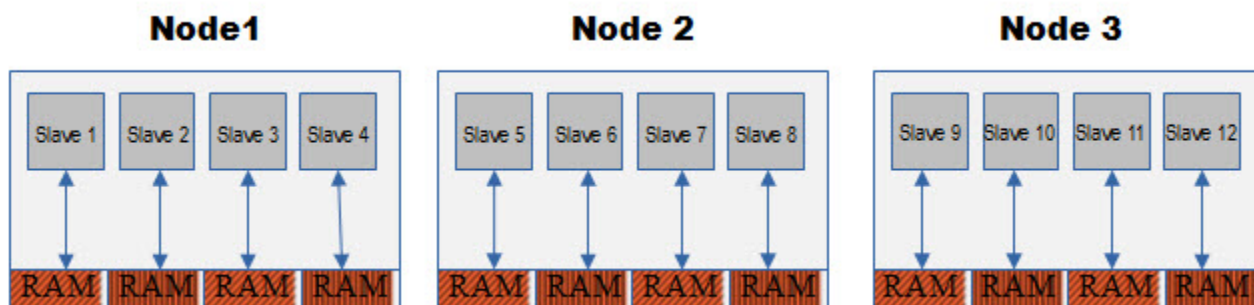
Essa configuração (a opção *includeTargetInURL* >) precisa ser falsa se você executar múltiplos clusters Roxie.

Virtual Thor escravos

A partir da versão 6.0.0, os clusters Thor podem ser configurados de forma a obter todos os benefícios oferecidos pelos recursos disponíveis por nó ao usar os escravos Thor virtuais.

Nas versões do HPCC anteriores à 6.0.0, as configurações de cluster eram normalmente definidas para um número N de **slavesPerNode**, onde N é igual ou próximo ao número de núcleos por máquina.

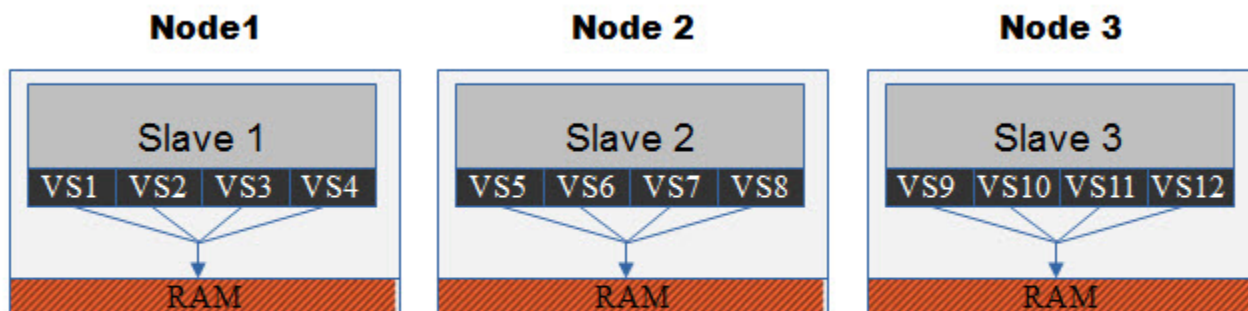
Isso resultou em N processos de escravo independentes por nó, como visto abaixo:



A prática apresentou diversas desvantagens significativas:

- Cada processo de escravo nessa configuração tem uma mesma divisão fixa da memória física disponível ao nó.
- Os escravos não compartilham memória RAM ou outros recursos.
- Os escravos transmitem mensagens através da interface de loopback para se comunicarem.

Atualmente emprega-se uma nova abordagem, na qual os escravos virtuais são criados com um único processo de escravo, como mostrado abaixo:



* VS = Virtual Slave (also known as a 'channel')

- Nesta configuração, cada nó físico possui um único processo de escravo Thor.
- Cada processo de escravo possui N escravos virtuais. Isso é configurado através da opção de configuração do Thor denominada **channelsPerSlave**. Nessa arquitetura, escravos no mesmo processo podem se comunicar diretamente entre si e compartilhar recursos.

Observação: A configuração **slavesPerNode** ainda existe e ambas podem ser usadas de forma combinada, se necessário.

Principais vantagens:

- Cada escravo virtual compartilha recursos em cache, como chaves de páginas de indexação, etc.
- Os escravos podem solicitar e compartilhar toda a memória RAM disponível.
- A inicialização e o gerenciamento do cluster são mais rápidos e simples.

- Permitem futuras melhorias para melhor coordenação/gerenciamento de núcleos de CPU.

O fato de se ter acesso a toda a memória disponível é de suma importância para determinadas atividades. O exemplo mais claro é um SMART ou LOOKUP JOIN.

Exemplo de SMART/LOOKUP JOIN

Um LOOKUP JOIN funciona aproximadamente da seguinte maneira:

- Transmite um dataset de RHS escravo local para todos os outros escravos.
- Todos os escravos reúnem o RHS global em uma única tabela.
- Uma tabela hash baseada nos campos de correspondência de chave física é criada.
- Depois que todos os escravos estão prontos, o LHS é transmitido e avaliado em relação à tabela hash para produzir os resultados unidos.

Observação: A tabela de RHS completa e a tabela hash precisam caber na memória; caso contrário, o JOIN falhará através de um erro de falta de memória.

SMART JOIN é uma evolução do LOOKUP JOIN. Se ele não puder encaixar o RHS global na memória, ele fará o HASH PARTITION, do RHS e o HASH DISTRIBUTE, do LHS e executará um LOCAL LOOKUP JOIN.

Se não for possível alocar o conjunto de RHS local na memória em um determinado nó, então irá reunir e classificar ambos os datasets locais e realizar uma JOIN padrão.

A principal vantagem da LOOKUP JOIN é a velocidade. Se o RHS couber na memória, ele pode realizar com rapidez uma operação de coleta e JOIN transmitida de um grande conjunto LHS sem a necessidade de coletar e classificar nada.

As vantagens de uma configuração do Thor com escravo virtual para o código ECL usando o LOOKUP/SMART JOIN são que, na prática, ele terá N vezes mais memória antes de falhar ou aplicar o fail over no caso do SMART JOIN.

Ele também é mais rápido: em vez de transmitir o RHS local para o processo de N escravos por nó, ele precisa realizar esta comunicação para apenas um. Esse único escravo pode compartilhar diretamente a mesma tabela e o mesmo HT com os outros escravos virtuais.

As principais vantagens de um LOOKUP/SMART JOIN em uma configuração do Thor com escravo virtual:

- N vezes mais memória disponível para o RHS. Em outras palavras, o RHS pode ser N vezes maior antes de falhar ou acionar o fail over no caso do SMART JOIN. (No exemplo ilustrado acima, o JOIN teria 4 vezes mais memória disponível)
- Comunicação significativamente menor de dados de linha – o que resulta em um processamento mais rápido para conjuntos de RHS maiores.

Huge Pages

O Linux usa páginas como suas unidades básicas de memória. Seu sistema pode operar de forma mais rápida e se beneficiar do suporte para huge pages. Huge pages de tipo e tamanho adequados precisam ser alocadas a partir do sistema operacional. Quase todos os sistemas Linux são configurados com Transparent Huge Pages (THP) disponíveis por padrão.

Thor, Roxie e clusters ECL Agent possuem opções na configuração para possibilitar suporte para huge pages. As huge pages transparentes são habilitadas para clusters Thor, Roxie e ECL Agent no ambiente padrão do HPCC. Os clusters Thor podem se beneficiar mais das huge pages do que o Roxie.

Consulte o arquivo `/sys/kernel/mm/transparent_hugepage/enabled` para verificar qual é a sua configuração de OS (SO). Com THP, não é necessário definir um tamanho de forma explícita. Se seu sistema não estiver configurado para usar o THP, você pode implementar huge pages.

Configurando Huge Pages

Para configurar o suporte a huge pages, consulte a documentação do seu OS (SO) e determine como ativar o suporte. Por exemplo, o administrador pode alocar huge pages persistentes (para o OS (SO) adequado) na linha de comando de inicialização de kernel especificando o parâmetro `"hugepages=N"` na inicialização. Com huge pages, também é necessário alocar o tamanho de forma explícita.

No Configuration Manager do HPCC, há três locais de definição dos atributos para uso de huge pages.

Há atributos em cada componente, nos atributos do ECL Agent, nos atributos do Roxie e nos atributos do Thor. Há dois valores em cada componente:

```
heapUseHugePages heapUseTransparentHugePages
```

Ative Huge Pages em seu sistema operacional e depois configure o HPCC para os componentes desejados.

Planejamento de Capacidade

Os clusters Roxie são clusters de computação de alto desempenho baseado em disco (HPCC) que normalmente usam arquivos indexados. Um cluster é capaz de armazenar e manipular a mesma quantidade de dados contida em seu espaço em disco rígido combinado, porém, isso não produz um desempenho ideal.

Para maximizar o desempenho, é preciso configurar seu cluster de modo que os nós dos agentes desempenhem a maioria das jobs na memória.

Por exemplo, se uma consulta utiliza três arquivos de dados com um arquivo combinado de tamanho igual a 60 GB, um cluster de 40 canais está de bom tamanho, embora um de 60 canais provavelmente seja melhor.

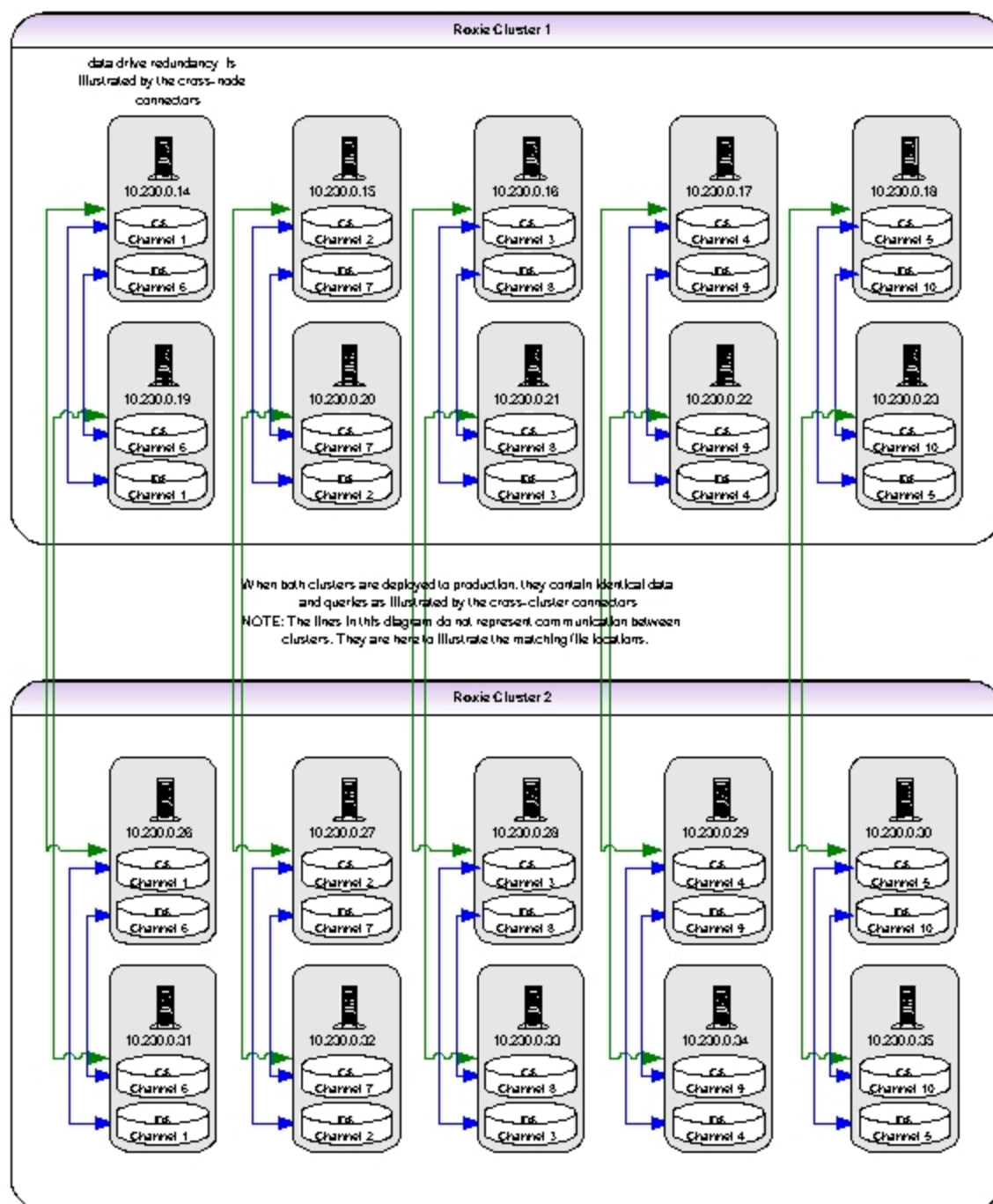
Outro ponto a considerar é o tamanho do cluster Thor que cria os arquivos de dados e que faz a indexação dos dados a serem carregados. Seu cluster Roxie de destino deve ter o mesmo tamanho do Thor, no qual os arquivos de dados e índices são criados ou um número que seja uniformemente divisível pelo tamanho do seu cluster Roxie. Por exemplo, um Thor de 100 nós para um Roxie de 20 nós seria considerado aceitável.

Isso se deve a forma na qual os dados são carregados e processados pelos agentes Roxie. Se os dados forem copiados para os nós dos agentes, as partes do arquivo são copiadas diretamente do local fonte para os locais de destino. Elas não são divididas ou redimensionadas para que caibam em um cluster com tamanho diferente. Consequentemente, ao carregar 50 partes de arquivo em um cluster de 40 canais, a parte um irá para o canal um, a parte dois para o canal dois e assim por diante. As partes 41 a 50 irão novamente para o topo, de modo que a parte 41 vá para o canal 1, a parte 42 para o canal 2, etc. O resultado será uma carga de trabalho desigualmente distribuída e que resultaria em menor desempenho. A velocidade de desempenho de um cluster é igual a de seu nó mais lento.

A consideração final é o número de processos do servidor em um cluster. Cada agente também deve ser um servidor, mas você pode dedicar nós adicionais para que sejam apenas processos do servidor. Isso é útil para consultas que exigem o processamento no servidor após os agentes terem retornado os resultados. Essas consultas intensivas de servidor poderiam ser enviadas apenas para endereços IP do servidor dedicado para que a carga seja removida dos nós atuando como servidor e agente.

Configurando os Canais

Na ilustração abaixo, os nós são configurados usando um esquema N+5 para compartilhar canais. Os canais podem ser configurados de várias formas, este é apenas um exemplo.



Nesta representação, cada chassi contém cinco blades de agentes Roxie (uma fila de servidores na imagem). Usaremos este exemplo para o restante deste manual.

Dimensionamento da Amostra

Esta seção ilustra exemplos dos dimensionamentos de sistema para vários ambientes de trabalho. Diferentemente dos requisitos de sistema, os exemplos a seguir são sugestões para configurar seu sistema nas diversas condições de operação.

Tamanho de amostra para alto volume de dados (típico)

A situação mais comum para o HPCC é utilizá-lo com um alto volume de dados. Esse exemplo de dimensionamento sugerido seria adequado para um local com grandes volumes de dados. Uma boa política é definir o tamanho do Thor para 4 vezes o dos dados de origem em seu HPCC. Normalmente, o Roxie tem ¼ do tamanho do Thor. Isso porque os dados são compactados e o sistema não retém dados transitórios no Roxie. Lembre-se de que não é recomendado que o número de nós do Roxie ultrapasse o número de nós do Thor.

Considerações sobre Dimensionamento de dados Thor

Cada nó do Thor pode conter cerca de 2,5 TB de dados (MAX - no máximo), por isso, planeje o número de nós do Thor adequadamente para os seus dados.

Se possível, use unidades SAS para ambos Thor e Roxie uma vez que elas são quase iguais às unidades SATA atualmente. Caso não use para ambas, tenha unidades SAS pelo menos para seu cluster Roxie.

O Thor replica os dados e é normalmente configurado para duas cópias.

Consideração sobre Dimensionamento de Dados Roxie

O Roxie mantém a maior parte de seus dados na memória, por isso é preciso alocar bastante memória para ele. Calcule o tamanho aproximado de seus dados e faça a alocação adequadamente. Você deve aumentar o número de nós ou a quantidade de memória.

Recomendamos alocar um Dali para cada cluster Roxie.

O Roxie deve ter um espelho. Isso é útil no momento em que você precisar atualizar os dados. O espelho é então atualizado, transformando-se posteriormente em primário no momento em que o outro é desativado. Essa é uma prática recomendada, mas não uma obrigação, exceto no caso de alta disponibilidade.

Dimensionamento de Amostra para Alto Processamento com Baixo Volume de Dados

A seção a seguir fornece exemplos de dimensionamento para processamento pesado, com a quantidade de dados aproximada indicada.

750 GB de Dados Brutos

Thor = 3 (escravos) + 2 (gerenciamento) = 5 nós

Roxie = 3 (agentes) + 1 (Dali) = 4 nós (isso significará que o ambiente ficará fora do ar durante a implementação de consultas)

Reservas = 2

Total = 13 nós

1250 GB de Dados Brutos

Thor = 6 (escravos) + 2 (gerenciamento) = 8 nós

Roxie = 4 (agentes) + 1 (Dali) = 5 nós (isso significará que o ambiente ficará fora do ar durante a implementação de consultas)

Reservas = 2

Total = 17 nós

2000 GB de Dados Brutos

Thor = 8 (escravos) + 3 (gerenciamento) = 11 nós

Roxie = 4 (agentes) + 1 (Dali) = 5 nós (isso significará que o ambiente ficará fora do ar durante a implementação de consultas)

Reservas = 2

Total = 20 nós

3500 GB de Dados Brutos

Thor = 12 (escravos) + 5 (gerenciamento) = 17 nós

Roxie = 6 (agentes) + 1 (Dali) = 7 nós (isso significará que o ambiente ficará fora do ar durante a implementação de consultas)

Reservas = 2

Total = 28 nós

Recursos do Sistema

Há recursos adicionais disponíveis para o HPCC System.

Recursos do HPCC

O link de recursos está disponível abaixo do link do ícone Operations. O link de recursos no ECL Watch oferece um link para o portal da Web do HPCC Systems®. Visite o portal da Web do HPCC Systems® em <http://hpccsystems.com/> para acessar atualizações de software, plugins, suporte, documentação e muito mais. Lá você encontrará recursos úteis para execução e manutenção do HPCC no portal da Web.

O ECL Watch oferece um link para a página de download do portal do HPCC: <http://hpccsystems.com/download>. Essa é uma página onde é possível baixar pacotes de instalação, imagens virtuais, código fonte, documentação e tutoriais.

Recursos Adicionais

Outras opções de ajuda com o HPCC e de aprendizado para o ECL também estão disponíveis. Há cursos on-line disponíveis. Acesse:

<https://learn.lexisnexus.com/hpcc>

Pode ser necessário se registrar no site. Há vários vídeos de treinamento e outras informações bastante úteis.