



Red Hat Enterprise Linux 5

Visão Geral do Cluster Suite

Red Hat Cluster Suite para o Red Hat Enterprise Linux 5

Edição 3

Red Hat Enterprise Linux 5 Visão Geral do Cluster Suite

Red Hat Cluster Suite para o Red Hat Enterprise Linux 5

Edição 3

Landmann

rlandmann@redhat.com

Nota Legal

Copyright © 2009 Red Hat, Inc.

This document is licensed by Red Hat under the [Creative Commons Attribution-ShareAlike 3.0 Unported License](#). If you distribute this document, or a modified version of it, you must provide attribution to Red Hat, Inc. and provide a link to the original. If the document is modified, all Red Hat trademarks must be removed.

Red Hat, as the licensor of this document, waives the right to enforce, and agrees not to assert, Section 4d of CC-BY-SA to the fullest extent permitted by applicable law.

Red Hat, Red Hat Enterprise Linux, the Shadowman logo, JBoss, OpenShift, Fedora, the Infinity logo, and RHCE are trademarks of Red Hat, Inc., registered in the United States and other countries.

Linux ® is the registered trademark of Linus Torvalds in the United States and other countries.

Java ® is a registered trademark of Oracle and/or its affiliates.

XFS ® is a trademark of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries.

MySQL ® is a registered trademark of MySQL AB in the United States, the European Union and other countries.

Node.js ® is an official trademark of Joyent. Red Hat Software Collections is not formally related to or endorsed by the official Joyent Node.js open source or commercial project.

The OpenStack ® Word Mark and OpenStack logo are either registered trademarks/service marks or trademarks/service marks of the OpenStack Foundation, in the United States and other countries and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation, or the OpenStack community.

All other trademarks are the property of their respective owners.

Resumo

Visão Geral do Red Hat Cluster Suite fornece uma visão geral sobre o Red Hat Cluster Suite para o Red Hat Enterprise Linux 5.

Índice

INTRODUÇÃO	2
1. FEEDBACK	3
CAPÍTULO 1. RED HAT CLUSTER SUITE VISÃO GERAL	4
1.1. CLUSTERS BÁSICOS	4
1.2. RED HAT CLUSTER SUITE INTRODUCTION	5
1.3. CLUSTER INFRASTRUCTURE	6
1.3.1. Gerenciamento de cluster	7
1.3.2. Gerenciamento de bloqueio	8
1.3.3. Fencing	8
1.3.4. Sistema de Configuração de Cluster	12
1.4. GERENCIAMENTO DE SERVIÇO DE ALTA DISPONIBILIDADE	14
1.5. RED HAT GFS	16
1.5.1. Desempenho Superior e Escalabilidade	18
1.5.2. Desempenho, Escalabilidade e Preço Moderado	18
1.5.3. Economia e desempenho	19
1.6. GERENCIADOR DE VOLUME LÓGICO DE CLUSTER	20
1.7. DISPOSITIVO DE BLOQUEIO DA REDE GLOBAL	23
1.8. SERVIDOR VIRTUAL LINUX	24
1.8.1. Two-Tier LVS Topology	26
1.8.2. Three-Tier LVS Topology	28
1.8.3. Métodos Encaminhados	29
1.8.3.1. Roteamento NAT	29
1.8.3.2. Roteamento direto	30
1.8.4. Marcas de Firewall e Persistência	32
1.8.4.1. Persistence	32
1.8.4.2. Marcas Firewall	32
1.9. FERRAMENTAS DE ADMINISTRAÇÃO DE CLUSTER	33
1.9.1. Conga	33
1.9.2. GUI Administração de Cluster	35
1.9.2.1. Cluster Configuration Tool	36
1.9.2.2. Cluster Status Tool	37
1.9.3. Ferramentas de Administração de Linha de Comando	38
1.10. GUI ADMINISTRADOR DE SERVIÇO VIRTUAL LINUX	39
1.10.1. CONTROL/MONITORING	40
1.10.2. GLOBAL SETTINGS	41
1.10.3. REDUNDANCY	43
1.10.4. VIRTUAL SERVERS	44
1.10.4.1. A Subseção do SERVIDOR VIRTUAL	45
1.10.4.2. Subseção do SERVIDOR REAL	47
1.10.4.3. EDIT MONITORING SCRIPTS Subsection	48
CAPÍTULO 2. SUMÁRIO DO COMPONENTE DO RED HAT CLUSTER SUITE	51
2.1. COMPONENTES DE CLUSTER	51
2.2. PÁGINAS MAN	57
2.3. HARDWARE (DISCO RÍGIDO) COMPATÍVEL	60
APÊNDICE A. HISTÓRICO DE REVISÃO	61
ÍNDICE REMISSIVO	62

INTRODUÇÃO

Este documento fornece uma visão geral de alto nível do Red Hat Cluster Suite para a Red Hat Enterprise Linux 5, e é organizado como segue a seguir:

- [Capítulo 1, Red Hat Cluster Suite Visão geral](#)
- [Capítulo 2, Sumário do Componente do Red Hat Cluster Suite](#)

Embora a informação neste documento seja apenas uma visão geral, você deverá ter um conhecimento vantajoso da Red Hat Enterprise Linux e entenderá os conceitos de servidores da computação, para a obtenção de uma boa compreensão sobre esta informação.

Para maiores informações em como se utilizar a Red Hat Enterprise Linux, por favor refira-se aos seguintes recursos:

- *Red Hat Enterprise Linux Guia de instalação*— Fornece informação sobre a instalação da Red Hat Enterprise Linux 5.
- *Red Hat Enterprise Linux Guia de implantação*— Fornece informação sobre a implantação, configuração e administração da Red Hat Enterprise Linux 5.

Para maiores informações sobre o Red Hat Cluster Suite para a Red Hat Enterprise Linux 5, por favor refira-se aos seguintes dados:

- *Configurando e Gerenciando o Red Hat Cluster*— Fornece informação sobre instalação, configuração e gerenciamento dos componentes do Red Hat Cluster.
- *LVM Administrator's Guide: Configuration and Administration*— Provides a description of the Logical Volume Manager (LVM), including information on running LVM in a clustered environment.
- *Sistema de Arquivo Global: Configuração e Administração*— Fornece informação sobre instalação, configuração e manutenção do Red Hat GSF - Red Hat Global File System (Sistema de Arquivo Global Red Hat).
- *Sistema de Arquivo Global 2: Configuração e Administração*— Fornece informação sobre instalação, configuração e manutenção do Red Hat GFS2 (Red Hat Global File System 2 - Sistema de Arquivo Global 2 da Red Hat).
- *Usando o Disparador Mapper de Caminhos Diversos*— Fornece informação sobre o uso do Disparador Mapper de Caminhos Diversos da Red Hat Enterprise Linux 5.
- *Usando GNBD com o Sistema de Arquivo Global*— Fornece uma visão geral de como se utilizar o Global Network Block Device (GNBD, Dispositivo de Bloqueio da Rede Global), com o Red Hat GFS.
- *Administração do Servidor Virtual Linux*— Fornece informação na configuração de sistemas de alta execução e serviços com o Linux Virtual Server (LVS, Servidor Virtual Linux).
- *Notas da Versão do Red Hat Cluster Suite*— Fornece informação sobre a atual versão do Red Hat Cluster Suite.

A documentação do Red Hat Cluster Suite e outros documentos da Red Hat estão disponíveis nas versões HTML, PDF, e RPM na Red Hat Enterprise Linux. A documentação em CD e on-line pode ser encontrada no <http://www.redhat.com/docs/>.

1. FEEDBACK

Caso você localize um erro, ou se você encontrou uma maneira de tornar este documento melhor, nós gostaríamos muito de ouvir a sua opinião. Por favor submeta a informação em Bugzilla (<http://bugzilla.redhat.com/bugzilla/>) contra o componente **Documentation-cluster**.

Be sure to mention the document's identifier:

Cluster_Suite_Overview(EN)-5 (2008-12-11T15:49)

By mentioning this document's identifier, we know exactly which version of the guide you have.

Em caso de sugestões para o aperfeiçoamento deste documento, por favor seja o mais claro possível. E, caso você encontre um erro, por favor insira o número de seção e algumas informações do texto para que possamos localizá-lo facilmente.

CAPÍTULO 1. RED HAT CLUSTER SUITE VISÃO GERAL

Os sistemas de Cluster fornecem confiabilidade, adaptabilidade e disponibilidade para criticar os serviços de produção. Você pode criar um cluster usando o Red Hat Cluster Suite para satisfazer suas necessidades de execução, alta disponibilidade, balanceamento de carga, escalabilidade, arquivo compartilhado e economia. Este capítulo fornece uma visão geral dos componentes e funções do Red Hat Cluster Suite e consiste nas seguintes seções:

- [Seção 1.1, “Clusters Básicos”](#)
- [Seção 1.2, “Red Hat Cluster Suite Introduction”](#)
- [Seção 1.3, “Cluster Infrastructure”](#)
- [Seção 1.4, “Gerenciamento de Serviço de Alta Disponibilidade”](#)
- [Seção 1.5, “Red Hat GFS”](#)
- [Seção 1.6, “Gerenciador de Volume Lógico de Cluster”](#)
- [Seção 1.7, “Dispositivo de Bloqueio da Rede Global”](#)
- [Seção 1.8, “Servidor Virtual Linux”](#)
- [Seção 1.9, “Ferramentas de Administração de Cluster”](#)
- [Seção 1.10, “GUI Administrador de Serviço Virtual Linux”](#)

1.1. CLUSTERS BÁSICOS

Um cluster é formado por dois ou mais computadores (chamados *nós* ou *membros*) dos quais trabalham juntos para executar uma tarefa. Existem quatro tipos principais de clusters:

- Armazenamento
- Alta disponibilidade
- Balanceamento de carga
- Alto desempenho

Os clusters de armazenamento fornecem um sistema de arquivo consistente de imagem, através de servidores de um cluster, permitindo que os servidores leiam e escrevam simultaneamente um único sistema de arquivo compartilhado. Um cluster de armazenamento facilita a administração limitando a instalação e patches de aplicações de um sistema de arquivo. Além disso, em um sistema de arquivo geral de cluster, o cluster de armazenamento elimina a necessidade de cópias redundantes do aplicativo de dados, e facilita o backup e recuperação de desastre. O Red Hat Cluster Suite fornece o cluster de armazenamento através do Red Hat GFS.

Os clusters de alta disponibilidade fornecem disponibilidade contínua de serviços, eliminando pontos isolados de falhas e falhando serviços de um nó de cluster a outro, em caso de um nó se tornar inoperante. Normalmente, os serviços em um cluster de alta disponibilidade lêem e escrevem dados (por leitura e escrita de sistemas de arquivo montados). Portanto, um cluster de alta disponibilidade deve manter a integridade de dados, enquanto que um nó de cluster assume o controle do serviço de outro nó de cluster. As falhas de nós, em um cluster de alta disponibilidade, não são visíveis para

clientes de fora do cluster. (Clusters de alta disponibilidade são às vezes referidos como clusters de falhas.) O Red Hat Cluster Suite fornece cluster de alta disponibilidade através deste componente Gerenciador de Serviço de Alta Disponibilidade.

Os clusters de balanceamento de carga despacham as solicitações de serviço da rede de acordo com os nós de cluster múltiplos, para o balanço da solicitação de carga entre os nós de cluster. O balanceamento de carga fornece escalabilidade de custo efetivo, pois você pode igualar o número de nós de acordo com os requerimentos de carga. Se um nó em um cluster de balanceamento de carga se tornar inoperante, o software de balanceamento de carga detectará a falha e redirecionará as solicitações a outros nós de cluster. As falhas de nós, em um cluster de balanceamento de carga, não são visíveis para clientes de fora do cluster. O Red Hat Cluster Suite fornece o balanceamento de carga através do LVS (Linux Virtual Server).

Os clusters de alto desempenho usam nós de cluster para atuar em avaliações simultâneas. Um cluster de alto desempenho fornece aplicativos para trabalhar em paralelo, portanto eleva a categoria de desempenho dos aplicativos. (Clusters de alto desempenho são também referência para os clusters computacionais ou computação em grade.)



NOTA

Os tipos de cluster resumidos, num texto precedido, refletem em configurações básicas; onde suas necessidades podem requerer uma combinação dos clusters descritos.

1.2. RED HAT CLUSTER SUITE INTRODUCTION

O Red Hat Cluster Suite (RHCS) é um grupo integrado de componentes de software do qual pode ser distribuído em uma variedade de configurações para satisfazer suas necessidades de desempenho, alta disponibilidade, balanceamento de carga, escalabilidade, arquivo compartilhado e economia.

RHCS consists of the following major components (refer to [Figura 1.1, “Red Hat Cluster Suite Introduction”](#)):

- Infraestrutura de cluster – Fornece funções fundamentais para os nós, trabalhando juntamente com um cluster: gerenciamento de configuração de arquivo, gerenciamento de associados e delimitação.
- Gerenciamento de Serviço de Alta disponibilidade – Fornece os serviços de falha de um nó de cluster para outro, caso um nó se torne inoperante.
- Ferramentas de administração de cluster – Ferramentas de configuração e gerenciamento para configurar e gerenciar um cluster Red Hat. As ferramentas são usadas com os componentes de Infraestrutura de Cluster, alta disponibilidade e componentes de Gerenciamento de Serviço, e armazenamento.
- Linux Virtual Server (LVS) – software de rotação que fornece o balanceamento de carga IP. O LVS executa em pares os servidores de redundância e distribui igualmente as solicitações dos clientes para servidores reais, que estão atrás de servidores do LVS.

Você pode suplementar o Red Hat Cluster Suite com os seguintes componentes dos quais fazem parte de um pacote opcional (e não fazem parte do Red Hat Cluster Suite):

- Red Hat GFS (Global File System) Sistema de Arquivo Global – Fornece um sistema de arquivo de cluster para uso com o Red Hat Cluster Suite. O GFS permite que os nós múltiplos compartilhem o armazenamento, num nível de bloqueio, como se o armazenamento fosse conectado localmente com cada nó de cluster.

- Cluster Logical Volume Manager (CLVM) Gerenciador de Volume Lógico de Cluster – Fornece o gerenciamento do volume de armazenamento de cluster.



NOTA

When you create or modify a CLVM volume for a clustered environment, you must ensure that you are running the `clvmd` daemon. For further information, refer to [Seção 1.6, “Gerenciador de Volume Lógico de Cluster”](#).

- Global Network Block Device (GNBD) Dispositivo de Bloqueio da Rede Global – Um componente auxiliar do GFS que exporta o armazenamento de nível de bloqueio para Ethernet. Esta é uma maneira econômica de se fazer disponível o armazenamento de nível de bloqueio no Red Hat GFS.

For a lower level summary of Red Hat Cluster Suite components and optional software, refer to [Capítulo 2, Sumário do Componente do Red Hat Cluster Suite](#)

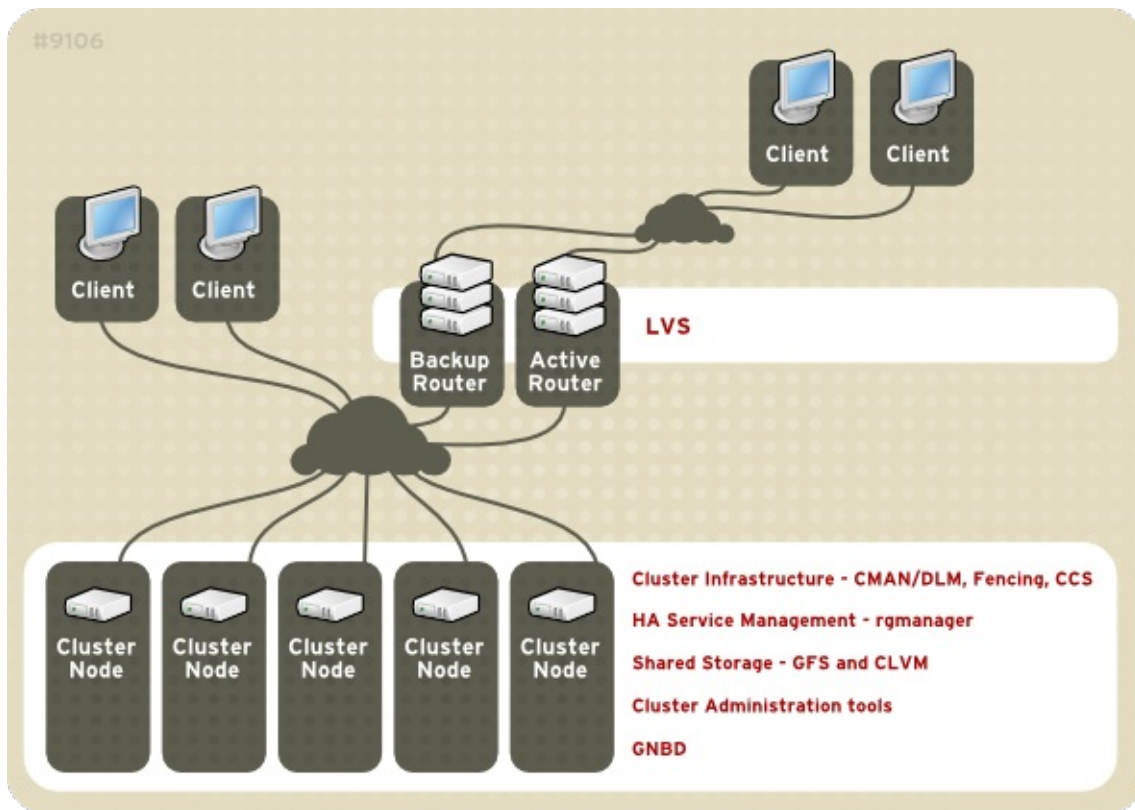


Figura 1.1. Red Hat Cluster Suite Introduction



NOTA

Figura 1.1, “Red Hat Cluster Suite Introduction” includes GFS, CLVM, and GNBD, which are components that are part of an optional package and *not* part of Red Hat Cluster Suite.

1.3. CLUSTER INFRASTRUCTURE

A infraestrutura de cluster do Red Hat Cluster Suite fornece as funções básicas para um grupo de computadores (chamado *nós* ou *membros*), para trabalhar juntamente com um cluster. Uma vez que o cluster é formado utilizando a infraestrutura de cluster, você pode usar outros componentes do Red

Hat Cluster Suite que satisfazem suas necessidades de cluster (por exemplo, configurar um cluster para os arquivos compartilhados num sistema de arquivo do GFS, ou configurar a falha do serviço). A infraestrutura de cluster atua nas seguintes funções:

- Gerenciamento de Cluster
- Gerenciamento de bloqueio
- Fencing
- Gerenciamento de configuração de cluster

1.3.1. Gerenciamento de cluster

Cluster management manages cluster quorum and cluster membership. CMAN (an abbreviation for cluster manager) performs cluster management in Red Hat Cluster Suite for Red Hat Enterprise Linux 5. CMAN is a distributed cluster manager and runs in each cluster node; cluster management is distributed across all nodes in the cluster (refer to [Figura 1.2, “CMAN/DLM Overview”](#)).

CMAN keeps track of cluster quorum by monitoring the count of cluster nodes. If more than half the nodes are active, the cluster has quorum. If half the nodes (or fewer) are active, the cluster does not have quorum, and all cluster activity is stopped. Cluster quorum prevents the occurrence of a "split-brain" condition – a condition where two instances of the same cluster are running. A split-brain condition would allow each cluster instance to access cluster resources without knowledge of the other cluster instance, resulting in corrupted cluster integrity.

O quórum é determinado por comunicação de mensagens entre nós de cluster por meio de Ethernet. Opcionalmente, o quórum pode ser determinado por uma combinação de mensagens de comunicação por meio do Ethernet e através de um disco de quórum. Para obter o quórum via disco de quórum, o quórum consistirá em condições especificadas pelos usuários.



NOTA

Pelo padrão, cada nó tem um voto de quórum. Opcionalmente, você pode configurar cada nó para a obtenção de mais de um voto.

O CMAN controla a associação monitorando as mensagens de um outro cluster de nó. Quando a associação de cluster sofrer alteração, o gerenciador de cluster notificará os outros componentes de infraestrutura, dos quais executarão a ação apropriada. Por exemplo, se um nó A junta-se às montagens de um sistema de arquivo GFS, onde o nó B e C já estiverem montados, um relatório adicional e um gerenciamento de bloqueio serão requeridos pelo nó A, para o uso do sistema de arquivo GFS. Caso um nó de cluster não transmitir a mensagem com uma quantia de tempo prescrita, o gerenciador de cluster removerá o nó do cluster e se comunicará com outro cluster.

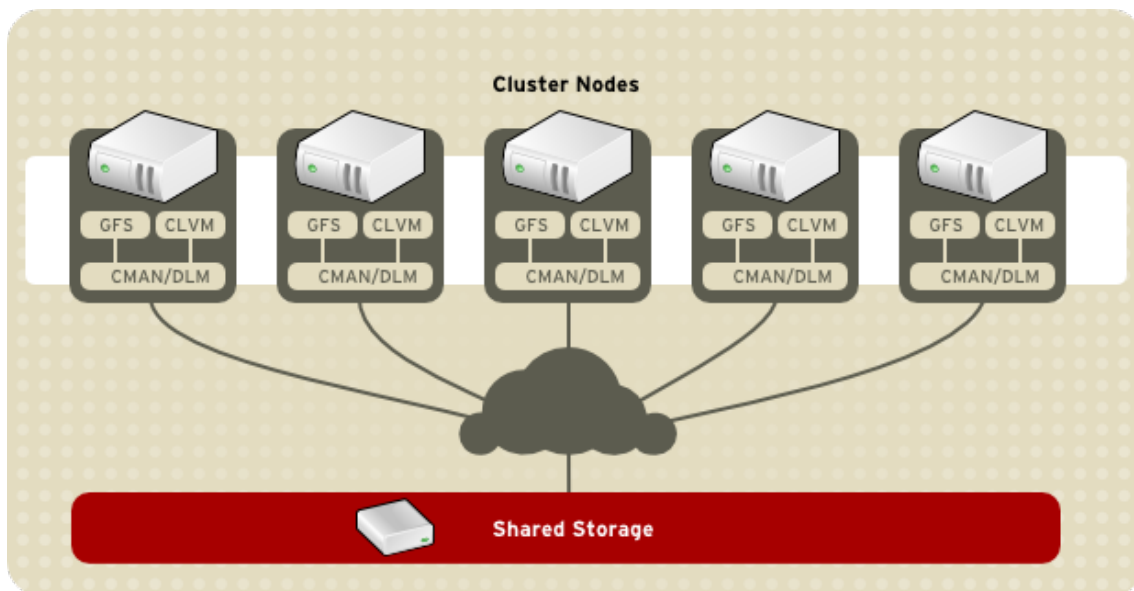


Figura 1.2. CMAN/DLM Overview

1.3.2. Gerenciamento de bloqueio

Lock management is a common cluster-infrastructure service that provides a mechanism for other cluster infrastructure components to synchronize their access to shared resources. In a Red Hat cluster, DLM (Distributed Lock Manager) is the lock manager. As implied in its name, DLM is a distributed lock manager and runs in each cluster node; lock management is distributed across all nodes in the cluster (refer to [Figura 1.2, “CMAN/DLM Overview”](#)). GFS and CLVM use locks from the lock manager. GFS uses locks from the lock manager to synchronize access to file system metadata (on shared storage). CLVM uses locks from the lock manager to synchronize updates to LVM volumes and volume groups (also on shared storage).

1.3.3. Fencing

Fencing is the disconnection of a node from the cluster's shared storage. Fencing cuts off I/O from shared storage, thus ensuring data integrity. The cluster infrastructure performs fencing through the fence daemon, `fenced`.

Quando o CMAN determinar que o nó falhou, ele se comunicará com outros componentes de infraestrutura de cluster para informar que o nó falhou. `Fenced`, quando for notificada a falha, o `fence` atua no nó falhado. Os outros componentes de infraestrutura de cluster determinam quais ações a serem tomadas – eles atuam em qualquer recuperação que necessita ser feita. Por exemplo, quando um nó de falha for notificado pelo DLM e GFS, a atividade será suspensa até eles detectarem que o `fenced` tenha completado o fencing no nó de falha. O DLM e GFS executam a recuperação, sob confirmação de que o nó sofreu fencing. O DLM libera bloqueios de nó de falhas, enquanto que o GFS recupera o relatório de um nó de falha.

O programa fencing determina a partir do arquivo de configuração de cluster qual método fencing a ser usado. Existem dois elementos-chaves num arquivo de configuração de cluster dos quais definem o método fencing: o agente fencing e o dispositivo fencing. O programa fencing efetua uma chamada para um agente fencing especificado num arquivo de configuração de cluster. O agente fencing, por outro lado, limita o nó por meio de um dispositivo fencing. Quando o fencing for completado, o programa fencing notificará o gerenciador de cluster.

O Red Hat Cluster Suite fornece uma variedade de métodos fencing:

- Força fencing – Um método fencing que usa uma força controladora para desligar um nó inoperante.
- Alternador fencing de Canal de Fibra – Um método fencing do qual desativa a porta do Canal de Fibra que conecta o armazenamento a um nó inoperante.
- GNBD fencing – A fencing method that disables an inoperable node's access to a GNBD server.
- Outros fencing – Diversos métodos fencing que desativam o I/O ou força de um nó inoperante, incluindo bladecentres IBM, PAP, DRAC/MC, HP ILO, IPMI, IBM RSA II, e outros.

Figura 1.3, “Power Fencing Example” shows an example of power fencing. In the example, the fencing program in node A causes the power controller to power off node D. Figura 1.4, “Fibre Channel Switch Fencing Example” shows an example of Fibre Channel switch fencing. In the example, the fencing program in node A causes the Fibre Channel switch to disable the port for node D, disconnecting node D from storage.

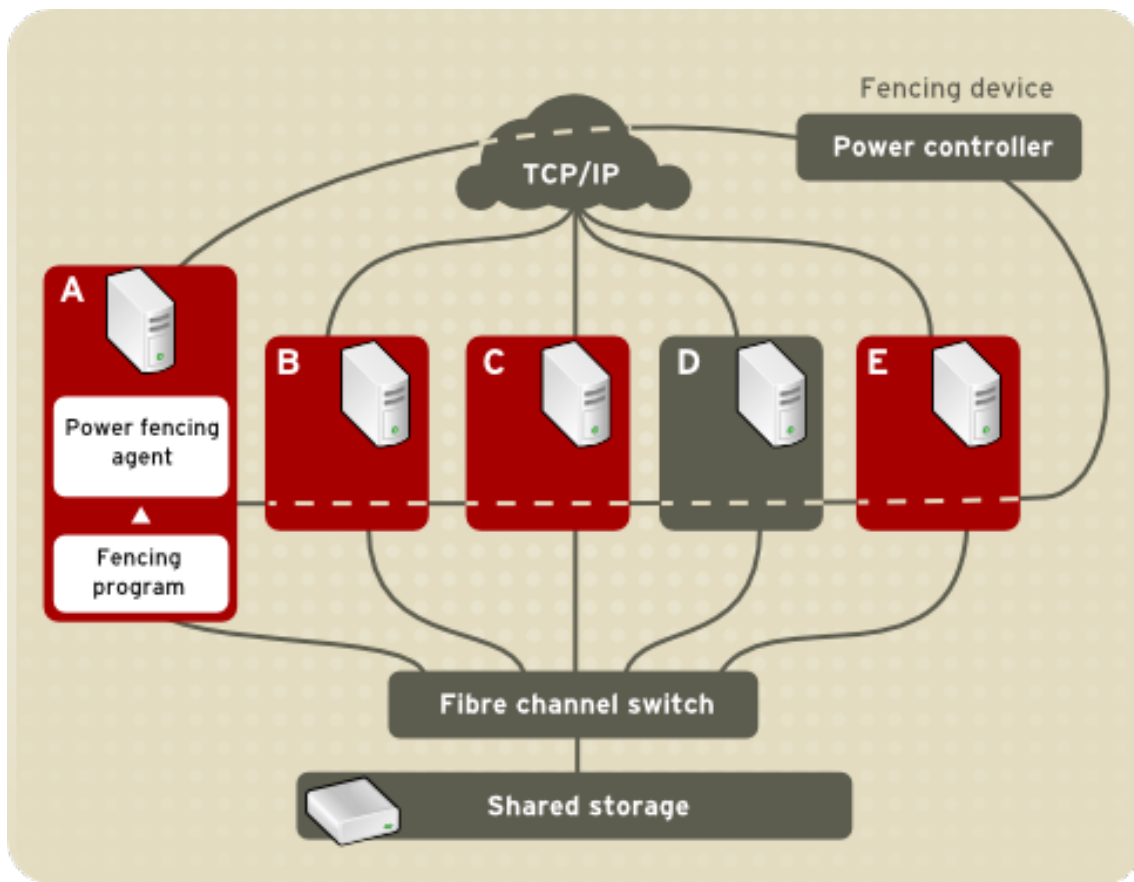


Figura 1.3. Power Fencing Example

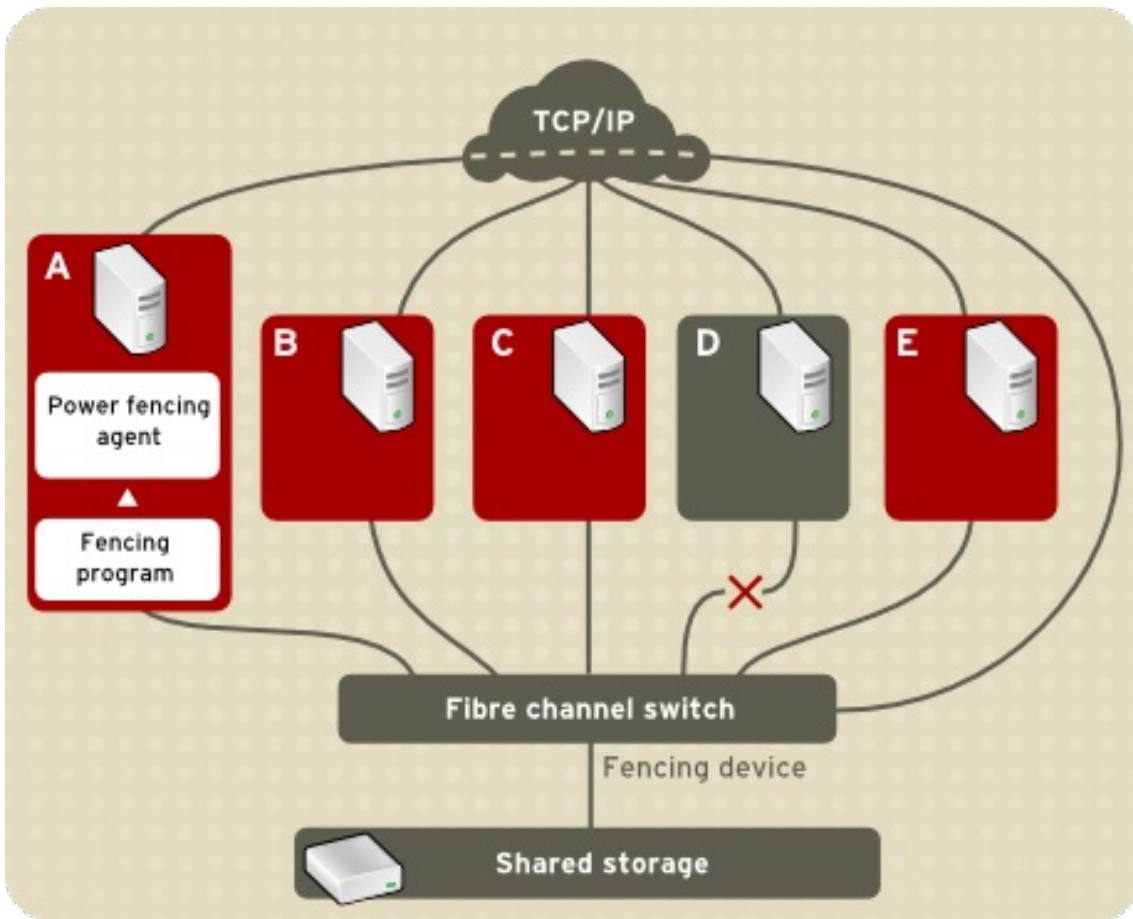


Figura 1.4. Fibre Channel Switch Fencing Example

Especifica que um método fencing consiste na edição de um arquivo de configuração de cluster para determinar o nome do método fencing, o agente fencing, e dispositivo fencing para cada nó num cluster.

The way in which a fencing method is specified depends on if a node has either dual power supplies or multiple paths to storage. If a node has dual power supplies, then the fencing method for the node must specify at least two fencing devices – one fencing device for each power supply (refer to [Figura 1.5, “Fencing a Node with Dual Power Supplies”](#)). Similarly, if a node has multiple paths to Fibre Channel storage, then the fencing method for the node must specify one fencing device for each path to Fibre Channel storage. For example, if a node has two paths to Fibre Channel storage, the fencing method should specify two fencing devices – one for each path to Fibre Channel storage (refer to [Figura 1.6, “Fencing a Node with Dual Fibre Channel Connections”](#)).

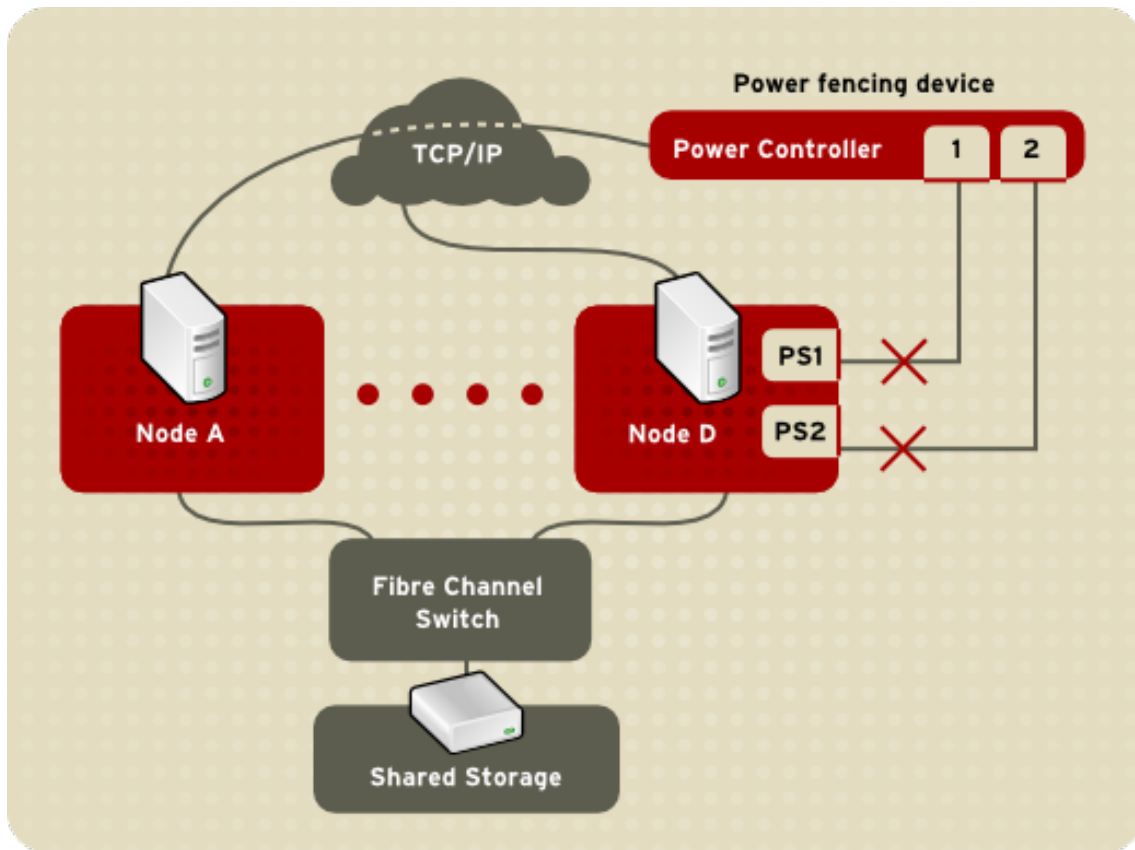


Figura 1.5. Fencing a Node with Dual Power Supplies

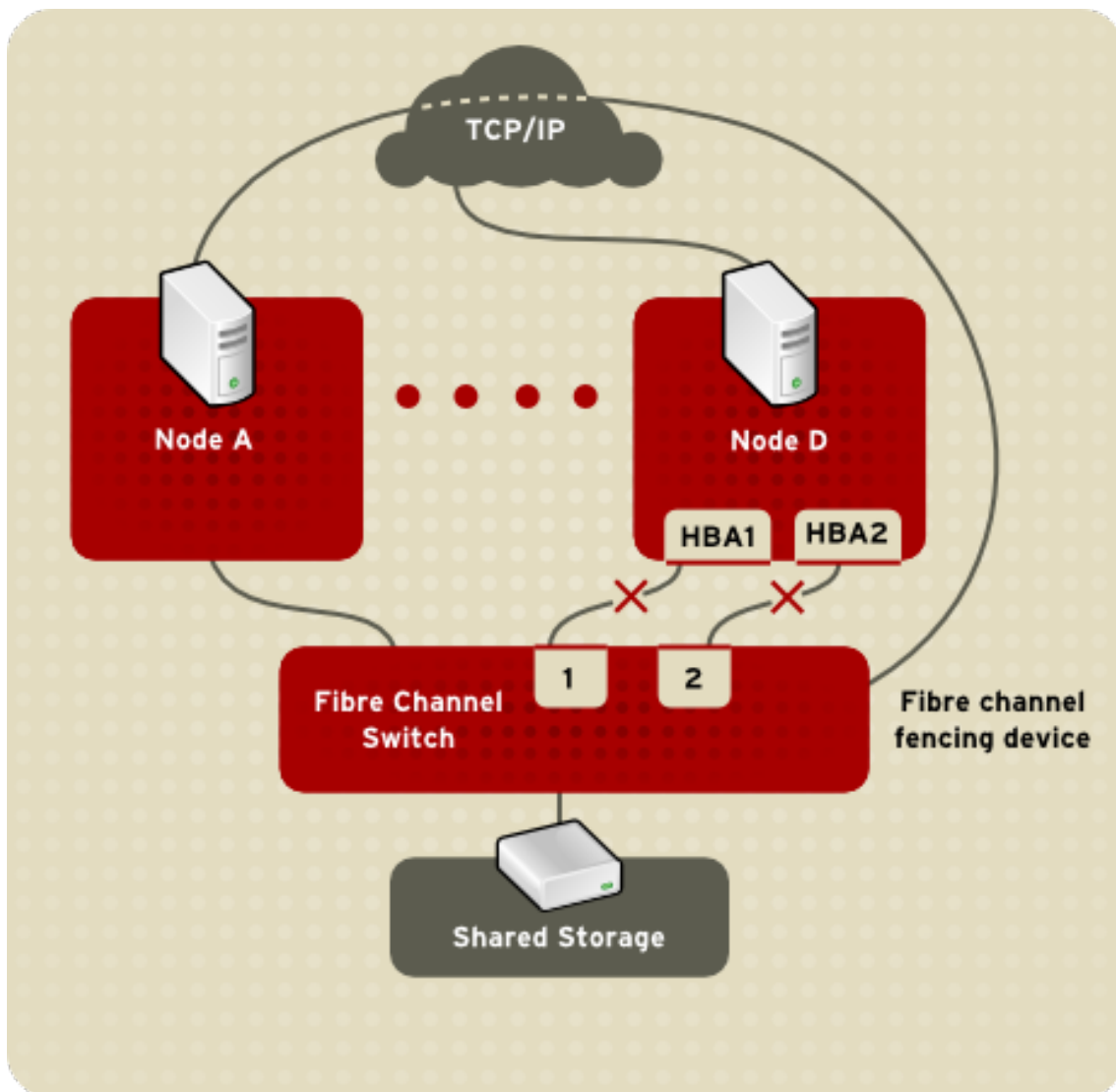


Figura 1.6. Fencing a Node with Dual Fibre Channel Connections

Você pode configurar um nó com um método fencing ou múltiplos métodos fencing. Quando você configurar um nó num método fencing, este será o único método disponível para fencing aquele nó. Quando você configurar um nó para múltiplos métodos fencing, os métodos fencing se apresentam como um efeito *cascata* de um método fencing a outro, de acordo com os métodos fencing especificados num arquivo de configuração de cluster. Se o nó falhar, este será fenced usando-se o primeiro método especificado num arquivo de configuração de cluster para aquele nó. Caso o primeiro método fencing não seja bem sucedido, o próximo método fencing especificado para o nó será usado. Se nenhum dos métodos fencing forem bem sucedidos, então o fencing iniciará novamente o primeiro método fencing especificado, e continuará laçado através de métodos fencing, de maneira especificada no arquivo de configuração de cluster até o nó ser fenced.

1.3.4. Sistema de Configuração de Cluster

The Cluster Configuration System (CCS) manages the cluster configuration and provides configuration information to other cluster components in a Red Hat cluster. CCS runs in each cluster node and makes sure that the cluster configuration file in each cluster node is up to date. For example, if a cluster system administrator updates the configuration file in Node A, CCS propagates the update from Node A to the other nodes in the cluster (refer to [Figura 1.7, “CCS Overview”](#)).

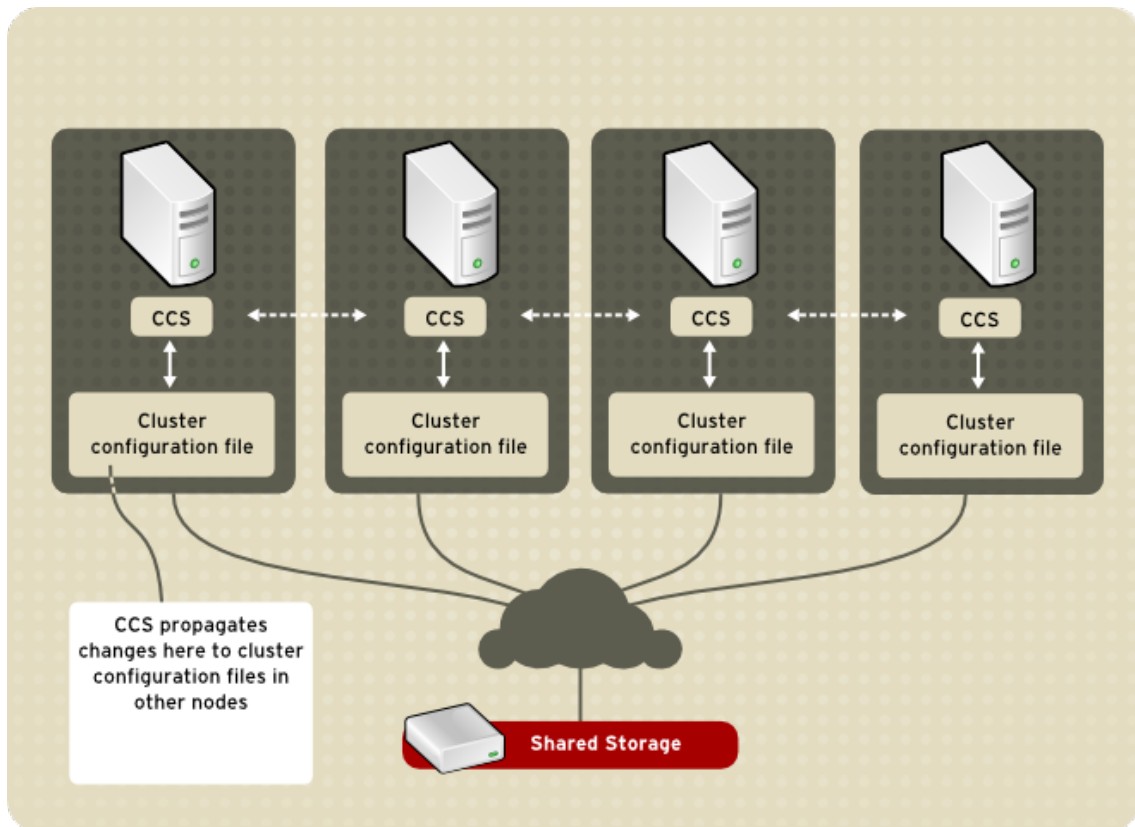


Figura 1.7. CCS Overview

Other cluster components (for example, CMAN) access configuration information from the configuration file through CCS (refer to [Figura 1.7, “CCS Overview”](#)).

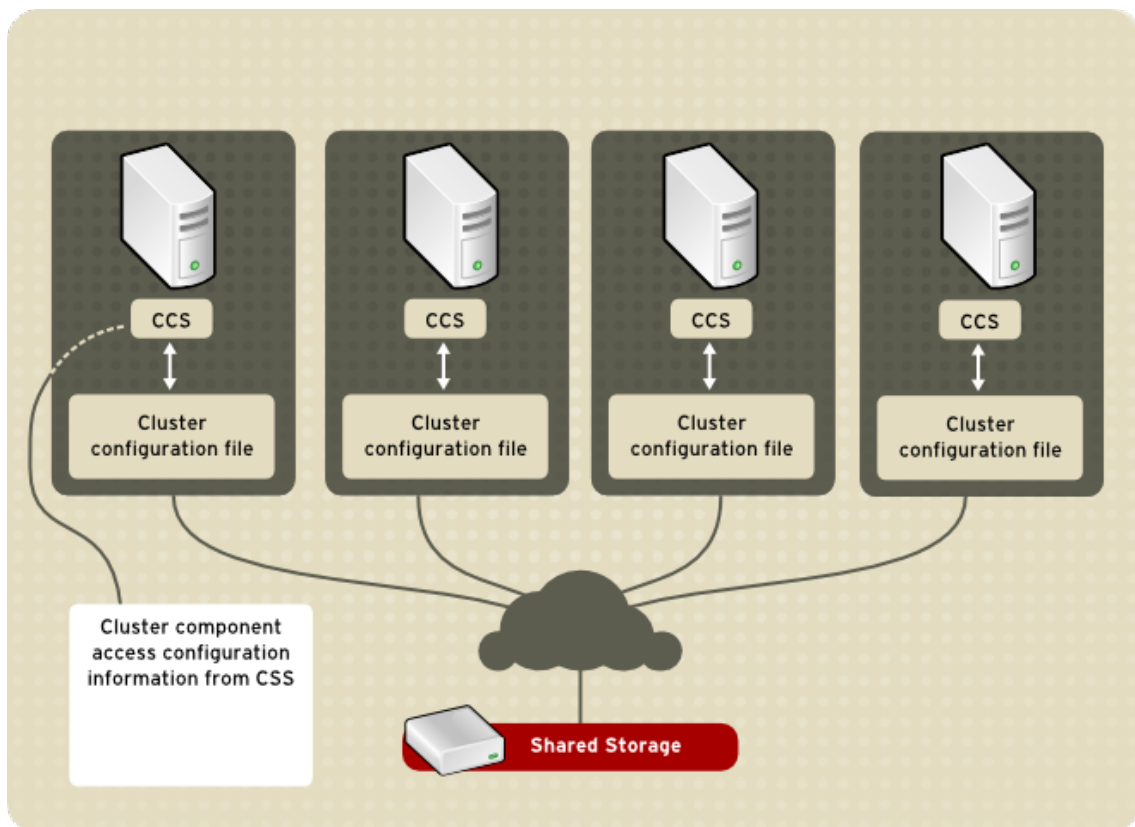


Figura 1.8. Accessing Configuration Information

O arquivo de configuração de cluster (`/etc/cluster/cluster.conf`) é um arquivo XML que descreve as seguintes características de cluster:

- O nome do cluster – Exibe o nome do cluster, o nível de revisão do arquivo de configuração de cluster, e propriedades básicas de tempo de fence usadas quando um nó une-se a um cluster, ou é limitado por um cluster.
- Cluster – Exibe cada nó de um cluster, especificando o nome do nó, a Id do nó, o número de votos do quórum, e o método fencing para aquele nó.
- Dispositivo de fence – Exibe os dispositivos de fence num cluster. Os parâmetros variam de acordo com o tipo de dispositivo de fence. Por exemplo, para um controlador de energia usado como um dispositivo de fence, a configuração de cluster define o nome do controlador de energia, seu endereço IP, logon e senha.
- Recursos Gerenciados – Exibe os recursos requeridos para criação de serviços de cluster. Os recursos gerenciados englobam a definição de domínios de falha, recursos (por exemplo um endereço IP) e serviços. Com isto, os recursos gerenciados definem os serviços de cluster e comportamento de falha dos serviços cluster.

1.4. GERENCIAMENTO DE SERVIÇO DE ALTA DISPONIBILIDADE

O Gerenciamento de Serviço de Alta Disponibilidade fornece a habilidade de criação e gerencia a alta disponibilidade nos *serviços de cluster* em um cluster Red Hat. A chave componente para o gerenciamento de serviço de alta disponibilidade num cluster Red Hat, `rgmanager`, implementa a falha fria para aplicações off-the-shelf (fora da prateleira). Em um cluster Red Hat, a aplicação é configurada com outros recursos para formar um serviço de cluster de alta disponibilidade. Um serviço de cluster de alta disponibilidade pode falhar de um para outro nó de cluster sem nenhuma interrupção aparente para clientes de cluster. A falha do serviço de cluster pode ocorrer caso um nó de cluster falhar, ou se um administrador de sistema de cluster mover o serviço de cluster de um para outro nó de cluster. (por exemplo, para uma interrupção planejada de um nó de cluster).

Para criar um serviço de alta disponibilidade, você deve configurá-lo no arquivo de configuração de cluster. O serviço de cluster compreende os *recursos* de cluster. Os recursos de cluster são obstáculos construídos que você pode criar e gerenciar num arquivo de configuração de cluster – por exemplo, um endereço IP, um script de inicialização de aplicação, ou um Red Hat GFS de partição dividido.

You can associate a cluster service with a *failover domain*. A failover domain is a subset of cluster nodes that are eligible to run a particular cluster service (refer to [Figura 1.9, “Domínios de Falha”](#)).



NOTA

Os domínios de falha *não* são requeridos para a operação.

Um serviço de cluster pode executar apenas um nó de cluster, de cada vez, para manter a integridade de dados. Você pode especificar a prioridade da falha num domínio de falha. A especificação da prioridade da falha consiste em designar o nível de prioridade para cada nó, num domínio de falha. O nível de prioridade determina a ordem de falha – especificando o nó em que o serviço de cluster deve falhar. Caso você não especifique a prioridade de falha, um serviço de cluster poderá falhar em qualquer nó de seu domínio de falha. Além disso, você pode especificar se um serviço de cluster é restrito para atuar nos nós de seu domínio de falha associado. (Quando associado a um domínio de falha sem restrição, o serviço cluster pode iniciar qualquer nó de cluster, num evento onde o membro do domínio de falha não é disponível.)

In [Figura 1.9, “Domínios de Falha”](#), Failover Domain 1 is configured to restrict failover within that

domain; therefore, Cluster Service X can only fail over between Node A and Node B. Failover Domain 2 is also configured to restrict failover with its domain; additionally, it is configured for failover priority. Failover Domain 2 priority is configured with Node C as priority 1, Node B as priority 2, and Node D as priority 3. If Node C fails, Cluster Service Y fails over to Node B next. If it cannot fail over to Node B, it tries failing over to Node D. Failover Domain 3 is configured with no priority and no restrictions. If the node that Cluster Service Z is running on fails, Cluster Service Z tries failing over to one of the nodes in Failover Domain 3. However, if none of those nodes is available, Cluster Service Z can fail over to any node in the cluster.

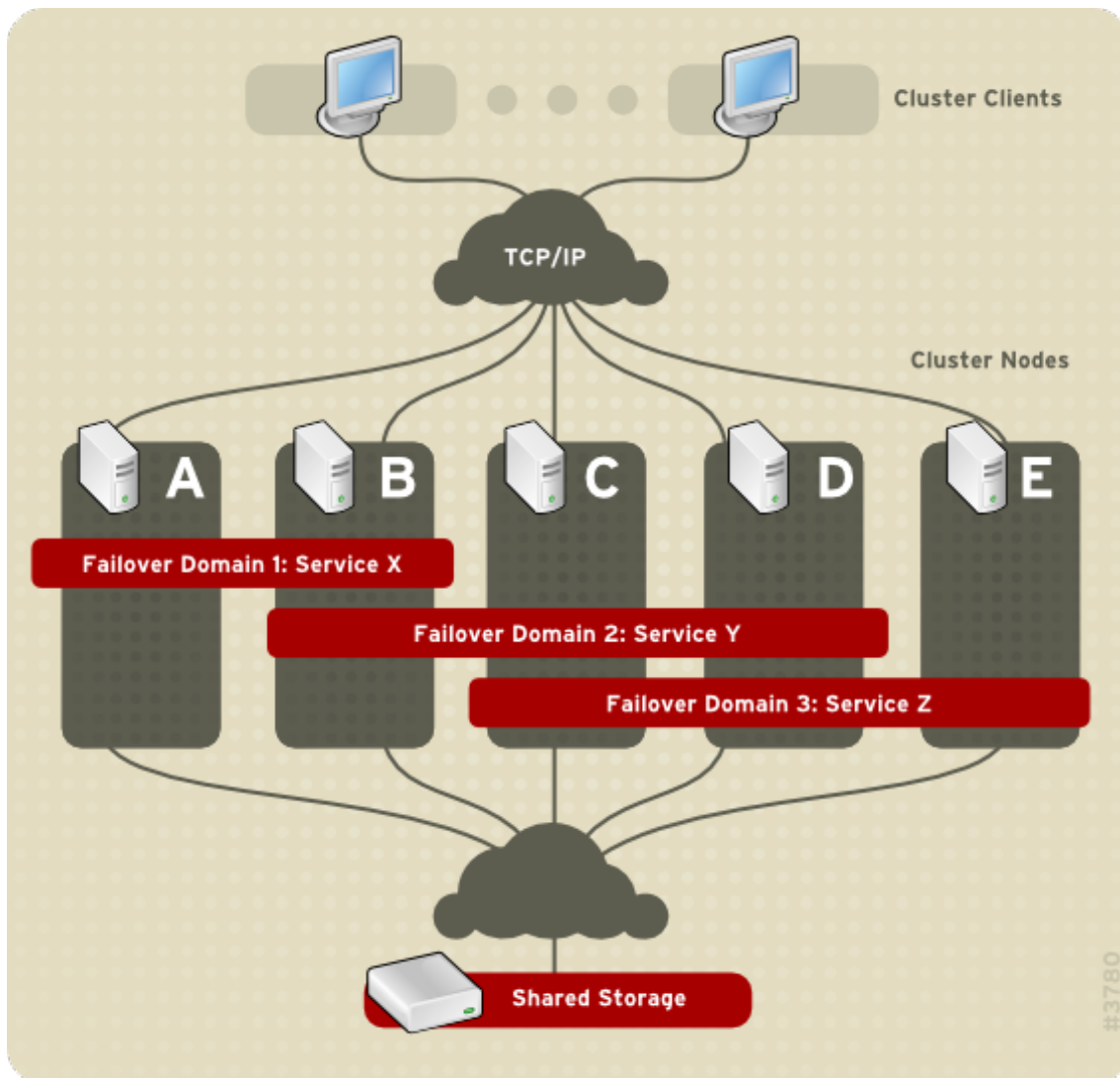


Figura 1.9. Domínios de Falha

Figura 1.10, “Web Server Cluster Service Example” shows an example of a high-availability cluster service that is a web server named "content-webserver". It is running in cluster node B and is in a failover domain that consists of nodes A, B, and D. In addition, the failover domain is configured with a failover priority to fail over to node D before node A and to restrict failover to nodes only in that failover domain. The cluster service comprises these cluster resources:

- Recurso do endereço IP – endereço IP 10.10.10.201.
- An application resource named "httpd-content" – a web server application init script `/etc/init.d/httpd` (specifying httpd).
- A file system resource – Red Hat GFS named "gfs-content-webserver".

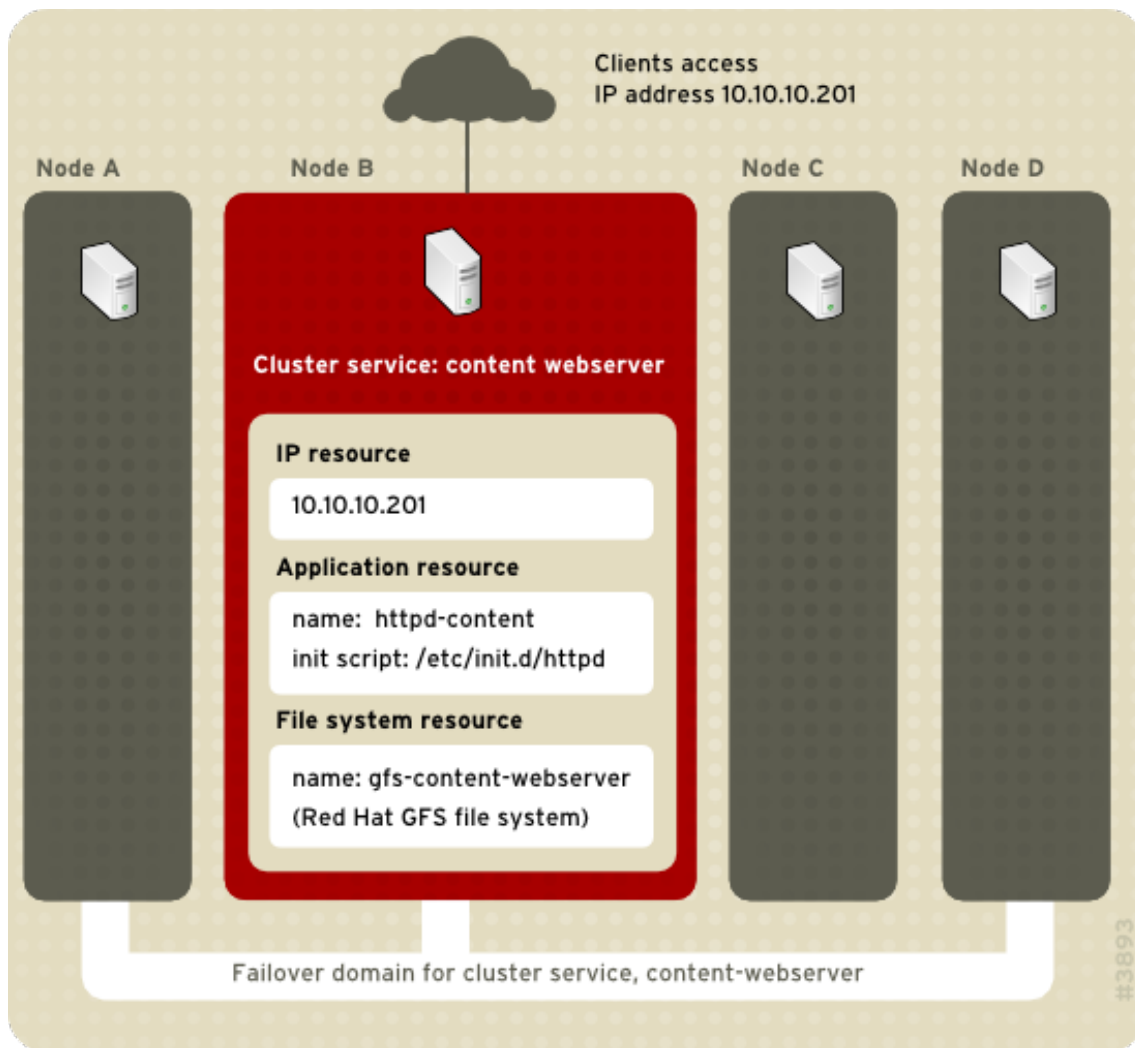


Figura 1.10. Web Server Cluster Service Example

Os clientes podem acessar o serviço de cluster através do endereço IP 10.10.10.201, capacitando a interação com o aplicativo do servidor da web, httpd-content. O aplicativo httpd-content usa o sistema de arquivo gfs-content-webserver. Caso o nó B falhasse, o serviço de cluster content-webserver falharia o nó D. Se o nó D não estivesse disponível e também falhasse, o serviço falharia o nó A. A falha poderia aparecer sem nenhuma interrupção aparente para os clientes de cluster. O cliente de cluster poderia ser acessível de um outro nó de cluster por meio do mesmo endereço, pois este foi falhado anteriormente.

1.5. RED HAT GFS

Red Hat GFS é um sistema de arquivo de cluster que permite um cluster de nós acessar simultaneamente um disparador de obstáculo do qual é dividido entre os nós. O GFS é um sistema de arquivo nativo que se conecta diretamente por meio da interface com a camada VFS da interface de sistema de arquivo Kernel Linux. O GFS emprega os metadados distribuídos e diários múltiplos para uma operação mais eficiente em um cluster. Para manter a integridade do sistema de arquivo, o GFS usa o gerenciador de bloqueio para coordenar o I/O. Quando um nó altera os dados em um sistema de arquivo GFS, a alteração é imediatamente visível para outros nós de cluster usando esse sistema de arquivo.

Utilizando o Red Hat GFS, você pode atingir o tempo de funcionamento do aplicativo máximo através dos seguintes benefícios:

- Simplificar a sua infraestrutura de dados

- Instalar e patch as aplicações de uma única vez para todo o cluster.
- Eliminar a necessidade de cópias redundantes dos dados de aplicação (duplicação).
- Permitir o acesso ler/escrever coexistente para dados de diversos clientes.
- Simplificar o backup e recuperação de desastre (apenas um sistema de arquivo para backup ou recuperação).
- Aumentar o uso de recursos de armazenamento; diminuindo os custos de administração do armazenamento.
 - Gerenciar o armazenamento como um todo, ao contrário de apenas por partição.
 - Diminuir, de uma maneira geral, as necessidades do armazenamento eliminando as replicações de dados.
- Escalar o cluster sem interrupção adicionando servidores ou armazenamento no fly.
 - A partição de armazenamento através de técnicas complicadas não é existente.
 - Adicionar servidores para o cluster num fly, montando-os ao sistema de arquivo comum.

Os nós que executam o Red Hat GFS são configurados e gerenciados com o Red Hat Cluster Suite de configuração e gerenciamento de ferramentas. O gerenciamento de volume é administrado através do CLVM (Cluster Logical Volume Manager - Gerenciador de Volume Lógico de Cluster). O Red Hat GFS fornece os dados divididos no meio de nós em um cluster Red Hat. O GFS fornece uma única visão consistente de espaço do nome do sistema de arquivo, através dos nós GFS num Red Hat cluster. O GFS permite os aplicativos instalarem e executarem, sem possuir muito conhecimento da infraestrutura de armazenamento subjacente. Além disso, o GFS fornece características que são tipicamente requeridas em ambientes empresariais, como cotas, jornais múltiplos e suporte de vários caminhos.

O GFS fornece um método versátil de armazenamento de rede de acordo com o desempenho, a escalabilidade, e necessidade econômica de seu ambiente de armazenamento. Este capítulo fornece algumas informações básicas abreviadas de fundo, para melhor entendimento do GFS.

You can deploy GFS in a variety of configurations to suit your needs for performance, scalability, and economy. For superior performance and scalability, you can deploy GFS in a cluster that is connected directly to a SAN. For more economical needs, you can deploy GFS in a cluster that is connected to a LAN with servers that use *GNBD* (Global Network Block Device) or to *iSCSI* (Internet Small Computer System Interface) devices. (For more information about GNBD, refer to [Seção 1.7, “Dispositivo de Bloqueio da Rede Global”](#).)

As seguintes seções fornecem exemplos de como o GFS pode ser desdobrado para satisfazer sua necessidade de desempenho, escalabilidade e economia:

- [Seção 1.5.1, “Desempenho Superior e Escalabilidade”](#)
- [Seção 1.5.2, “Desempenho, Escalabilidade e Preço Moderado”](#)
- [Seção 1.5.3, “Economia e desempenho”](#)



NOTA

Os exemplos de desenvolvimento do GFS refletem em configurações básicas; suas necessidades talvez solicitem uma combinação de configurações apresentadas nos exemplos.

1.5.1. Desempenho Superior e Escalabilidade

You can obtain the highest shared-file performance when applications access storage directly. The GFS SAN configuration in [Figura 1.11, “GFS with a SAN”](#) provides superior file performance for shared files and file systems. Linux applications run directly on cluster nodes using GFS. Without file protocols or storage servers to slow data access, performance is similar to individual Linux servers with directly connected storage; yet, each GFS application node has equal access to all data files. GFS supports over 300 GFS nodes.

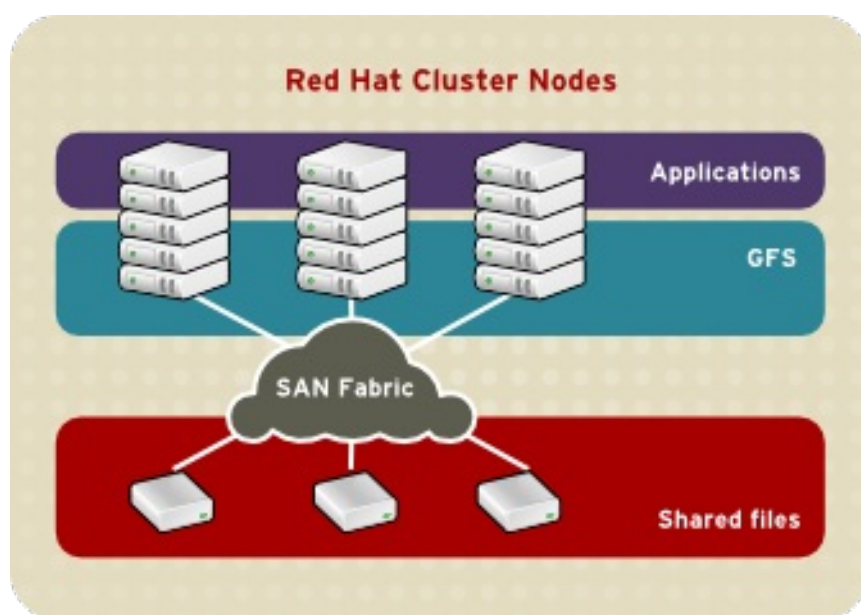


Figura 1.11. GFS with a SAN

1.5.2. Desempenho, Escalabilidade e Preço Moderado

Multiple Linux client applications on a LAN can share the same SAN-based data as shown in [Figura 1.12, “GFS and GNBD with a SAN”](#). SAN block storage is presented to network clients as block storage devices by GNBD servers. From the perspective of a client application, storage is accessed as if it were directly attached to the server in which the application is running. Stored data is actually on the SAN. Storage devices and data can be equally shared by network client applications. File locking and sharing functions are handled by GFS for each network client.

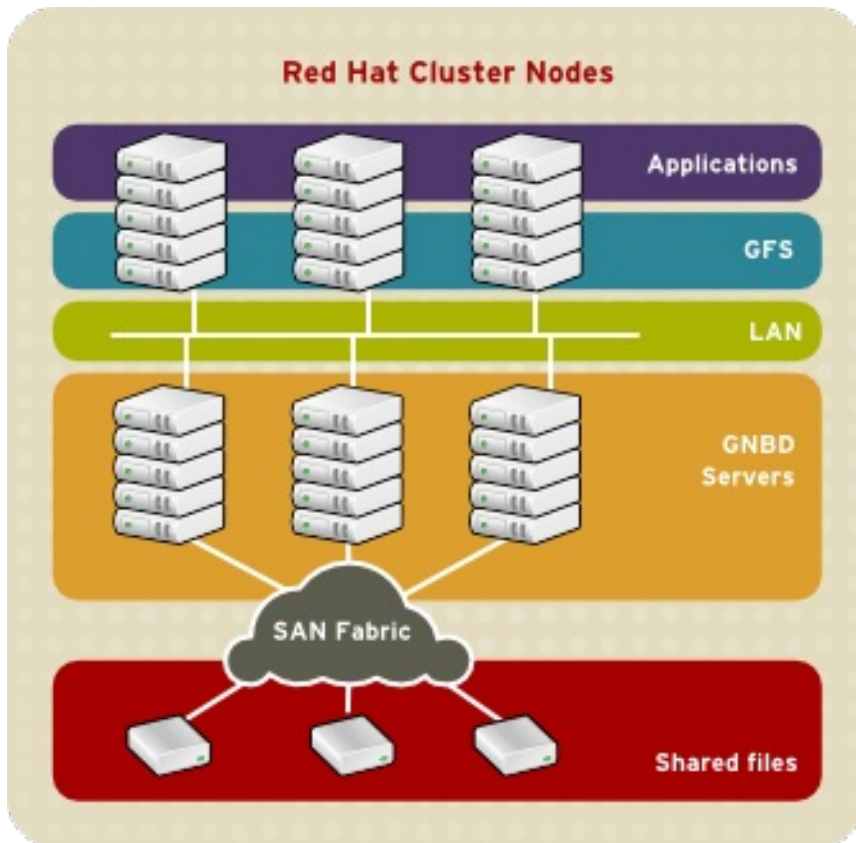


Figura 1.12. GFS and GNBD with a SAN

1.5.3. Economia e desempenho

Figura 1.13, “GFS e GNBD com Armazenamento Conectado Diretamente” shows how Linux client applications can take advantage of an existing Ethernet topology to gain shared access to all block storage devices. Client data files and file systems can be shared with GFS on each client. Application failover can be fully automated with Red Hat Cluster Suite.

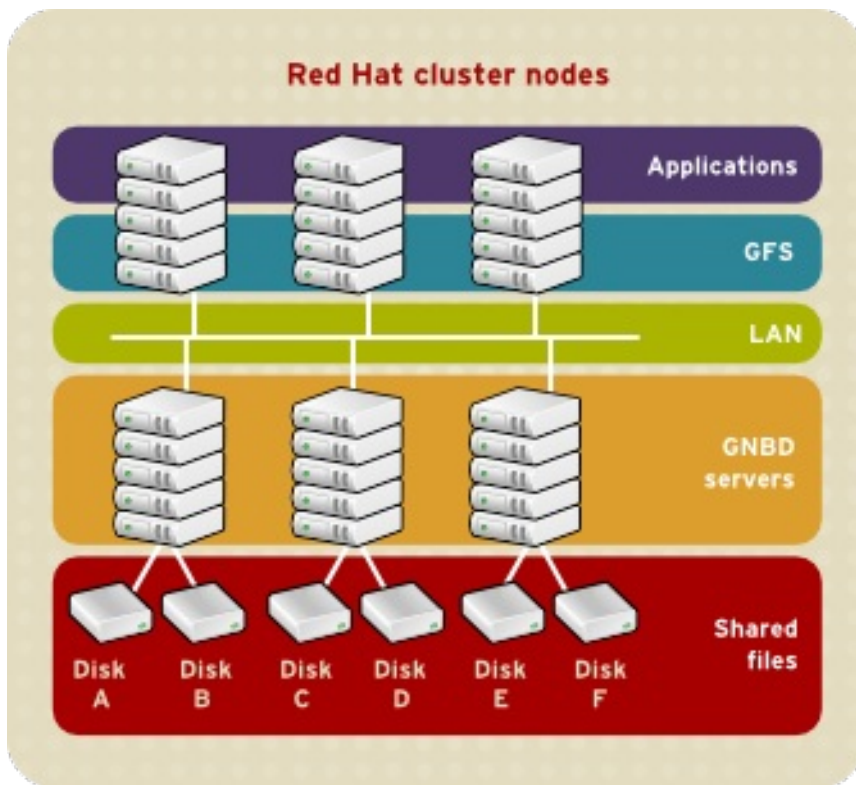


Figura 1.13. GFS e GNBD com Armazenamento Conectado Diretamente

1.6. GERENCIADOR DE VOLUME LÓGICO DE CLUSTER

O Cluster Logical Volume Manager (CLVM), (Gerenciador de Volume Lógico de Cluster) fornece uma versão ampla de cluster de LVM2. O CLVM fornece a mesma capacidade de LVM2 num único nó, mas faz a disponibilidade dos volumes de todos os nós num cluster Red Hat. Os volumes lógicos criados com o CLVM, tornam disponíveis os volumes lógicos para todos os nós em um cluster.

The key component in CLVM is `clvmd`. `clvmd` is a daemon that provides clustering extensions to the standard LVM2 tool set and allows LVM2 commands to manage shared storage. `clvmd` runs in each cluster node and distributes LVM metadata updates in a cluster, thereby presenting each cluster node with the same view of the logical volumes (refer to [Figura 1.14, “CLVM Overview”](#)). Logical volumes created with CLVM on shared storage are visible to all nodes that have access to the shared storage. CLVM allows a user to configure logical volumes on shared storage by locking access to physical storage while a logical volume is being configured. CLVM uses the lock-management service provided by the cluster infrastructure (refer to [Seção 1.3, “Cluster Infrastructure”](#)).



NOTA

Armazenamento compartilhado para uso no Red Hat Cluster Suite, requer que você esteja executando o daemon do gerenciador do volume lógico do cluster (`clvmd`) ou os agentes do Gerenciamento de Volume Lógico de Alta Disponibilidade (HA-LVM). Se você não conseguir usar o daemon do `clvmd` ou o HA-LVM por motivos operacionais ou porque você não possui os direitos corretos, não use o single-instance LVM no disco compartilhado, pois isto pode danificar os dados. Caso tenha qualquer dúvida, entre em contato com o representante de serviços da Red Hat.



NOTA

O uso do CLVM requer pequenas mudanças de acordo com o `/etc/lvm/lvm.conf` para bloqueamento de um cluster amplo.

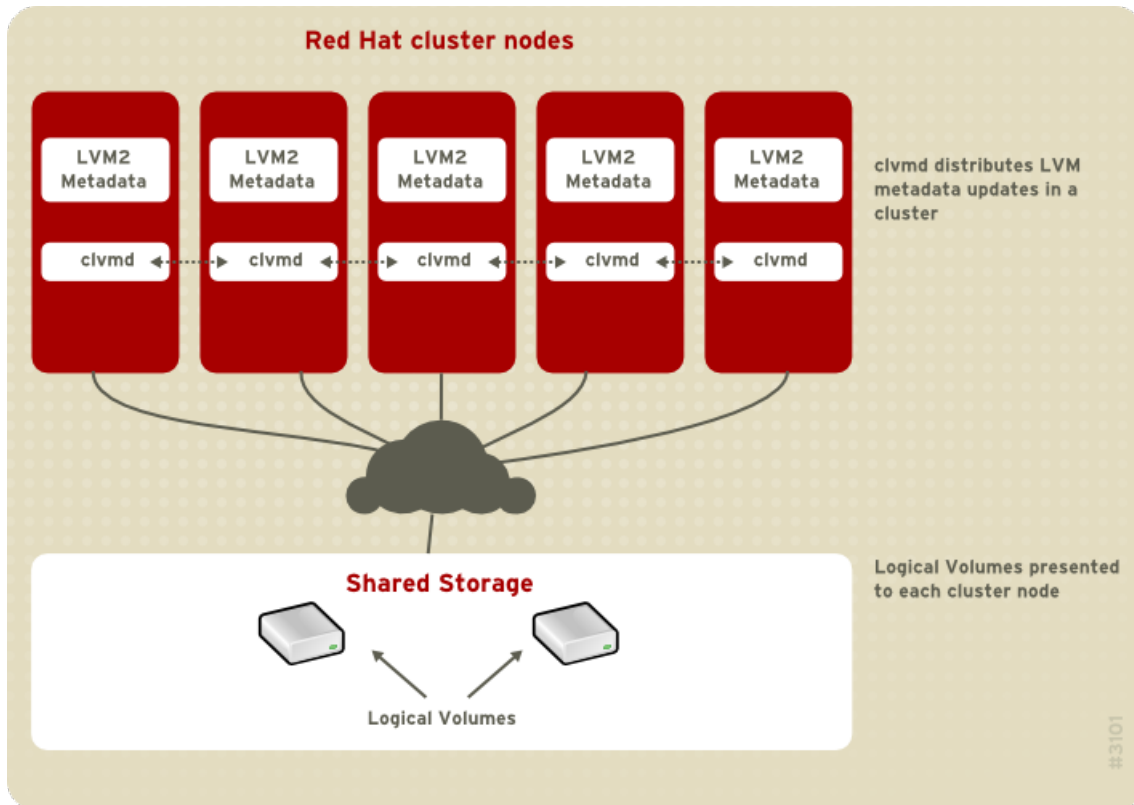


Figura 1.14. CLVM Overview

You can configure CLVM using the same commands as LVM2, using the LVM graphical user interface (refer to [Figura 1.15, “LVM Graphical User Interface”](#)), or using the storage configuration function of the **Conga** cluster configuration graphical user interface (refer to [Figura 1.16, “Conga LVM Graphical User Interface”](#)). [Figura 1.17, “Creating Logical Volumes”](#) shows the basic concept of creating logical volumes from Linux partitions and shows the commands used to create logical volumes.

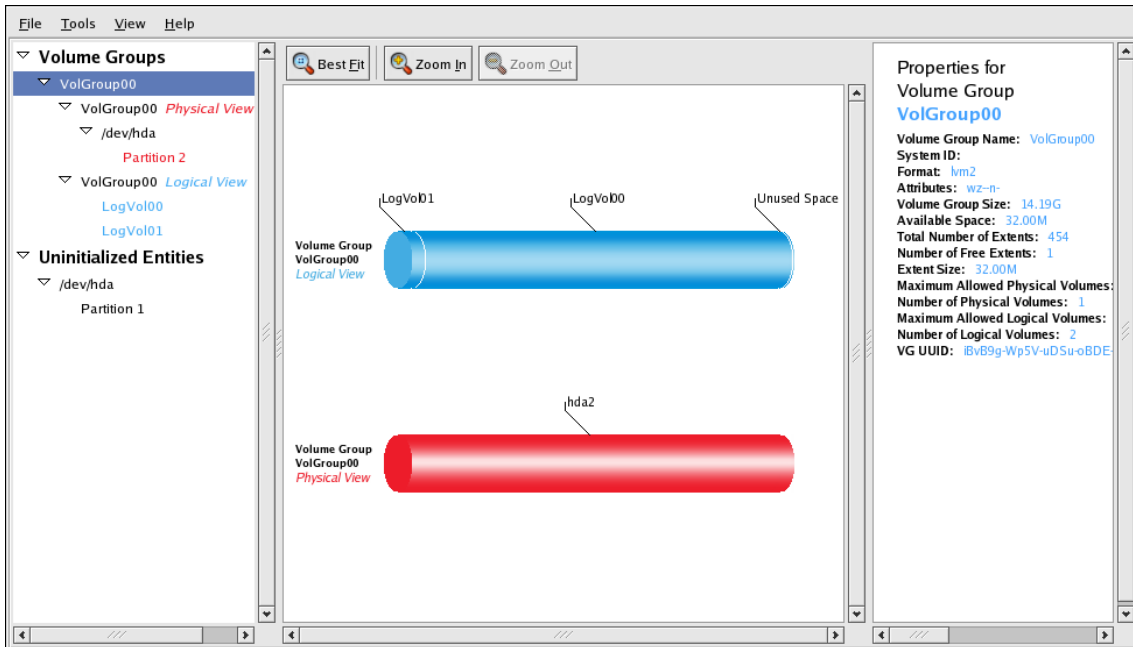


Figura 1.15. LVM Graphical User Interface

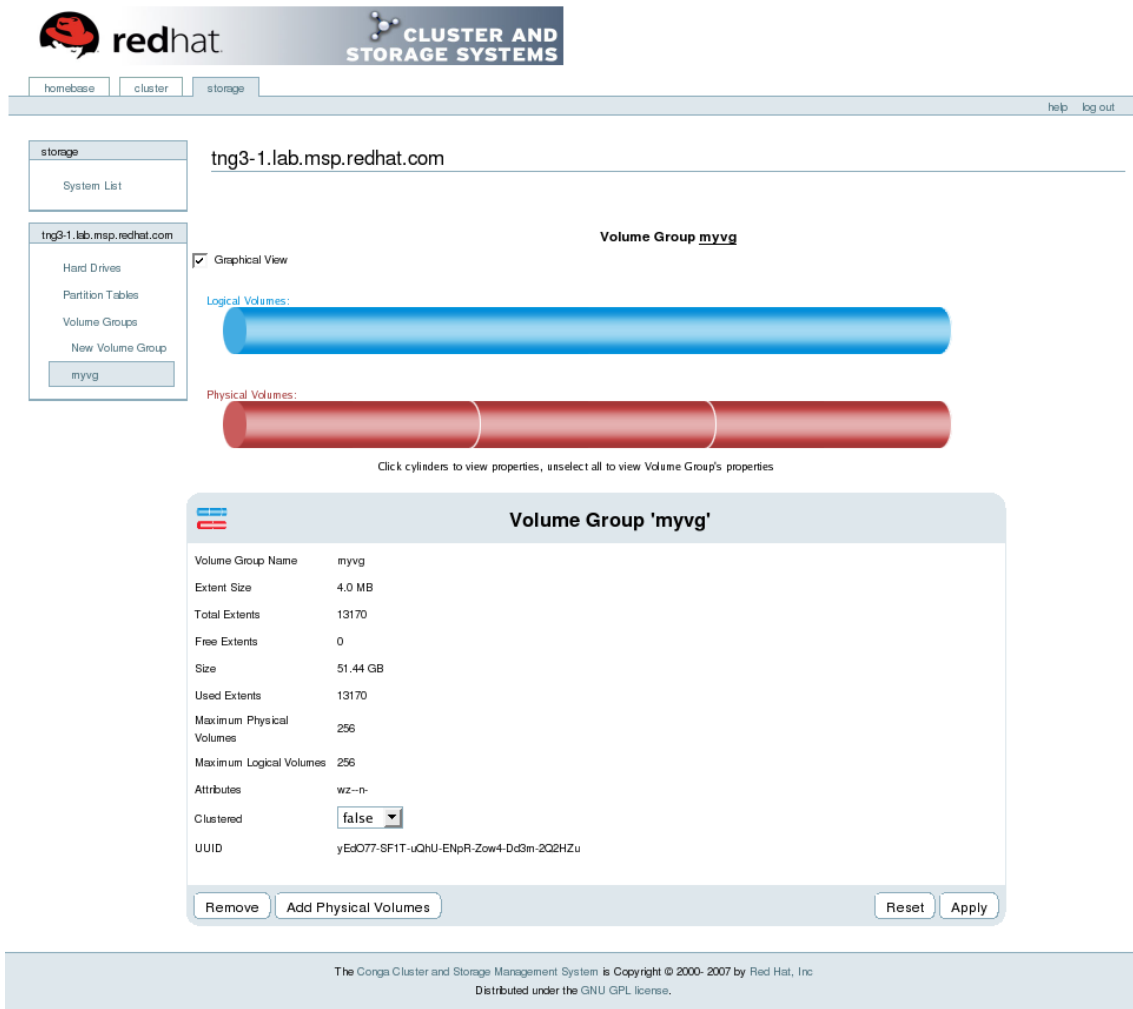


Figura 1.16. Conga LVM Graphical User Interface

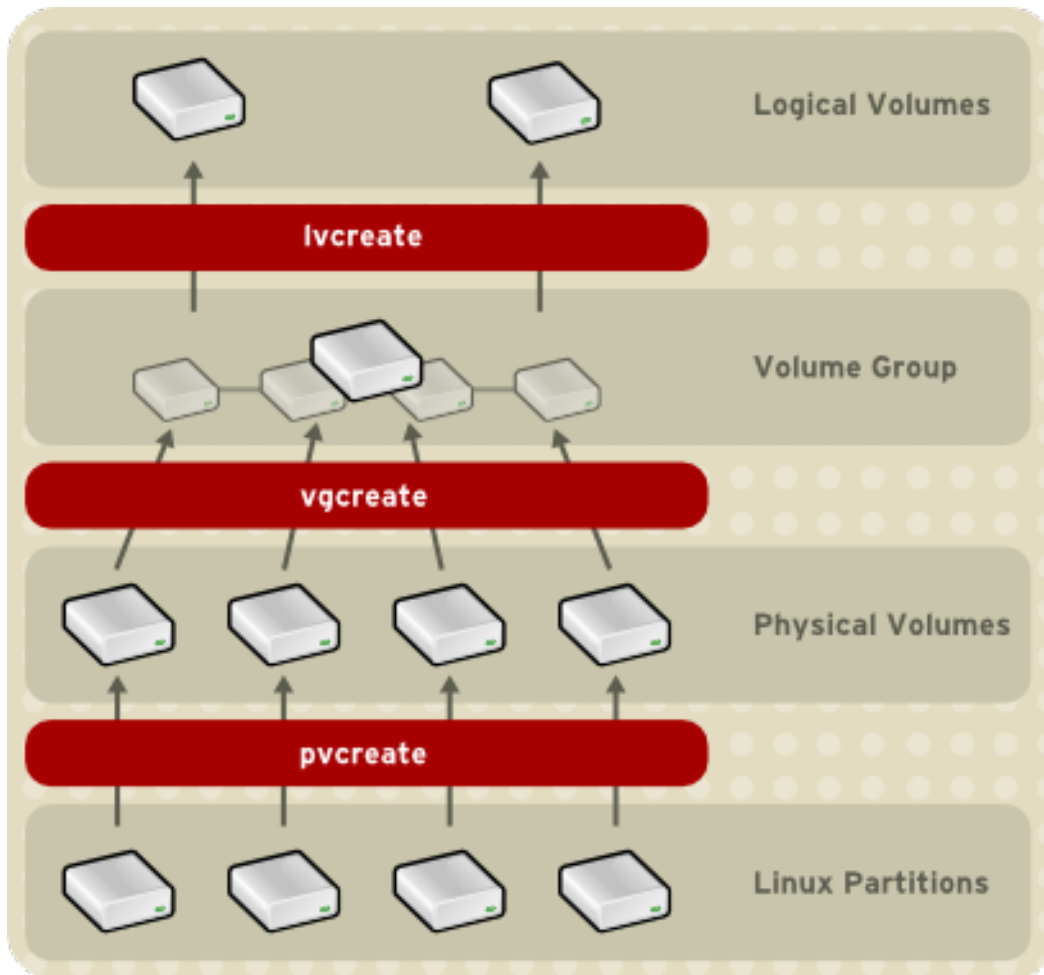


Figura 1.17. Creating Logical Volumes

1.7. DISPOSITIVO DE BLOQUEIO DA REDE GLOBAL

O Dispositivo de Bloqueio da Rede Global (GNBD - Global Global Network Block Device) fornece o acesso ao dispositivo de bloqueio para o Red Hat GFS sobre o TCP/IP. O GNBD possui um conceito similar ao NBD. No entanto, o GNBD é um GFS-específico e somente direcionado para uso com o GFS. O GNBSD é útil quando a necessidade de mais tecnologias robustas – Canal de Fibra ou iniciador único SCSI – não são necessárias ou são de custo proibidos.

GNBD consists of two major components: a GNBD client and a GNBD server. A GNBD client runs in a node with GFS and imports a block device exported by a GNBD server. A GNBD server runs in another node and exports block-level storage from its local storage (either directly attached storage or SAN storage). Refer to [Figura 1.18, “Visualização GNBD”](#). Multiple GNBD clients can access a device exported by a GNBD server, thus making a GNBD suitable for use by a group of nodes running GFS.

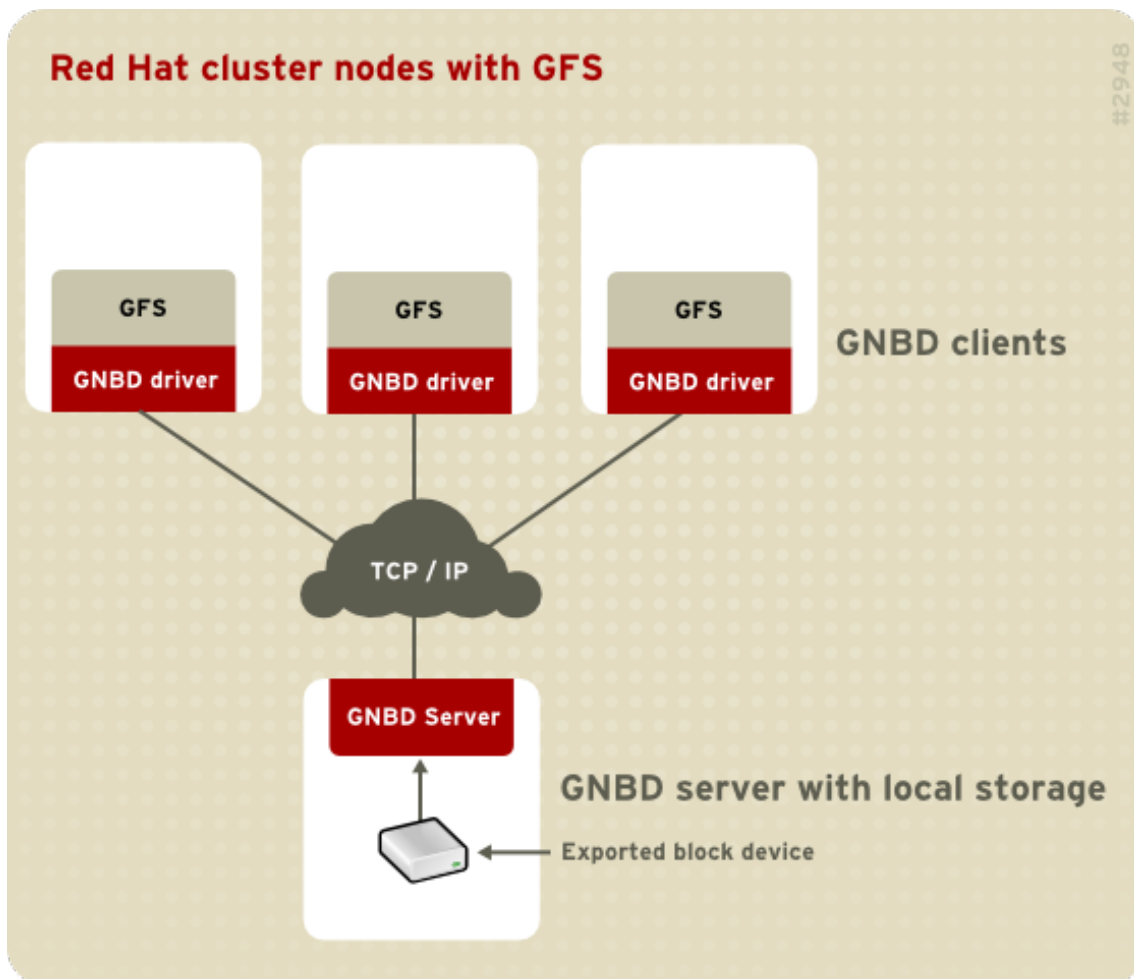


Figura 1.18. Visualização GNBD

1.8. SERVIDOR VIRTUAL LINUX

O Servidor Virtual Linux (LVS - Linux Virtual Server) é um conjunto de componentes de software integrados para balanceamento da carga IP através de um conjunto de servidores reais. O LVS atua num par de igualdade de computadores configurados: o roteador LVS ativo e o roteador LVS de backup. O roteador LVS ativo apresenta duas funções:

- Balancear a carga através de servidores reais.
- Checar a integridade de serviços em cada servidor real.

O roteador LVS de backup monitora o roteador LVS ativo e se responsabiliza por isto, em caso do roteador LVS ativo falhar.

Figura 1.19, “Components of a Running LVS Cluster” provides an overview of the LVS components and their interrelationship.

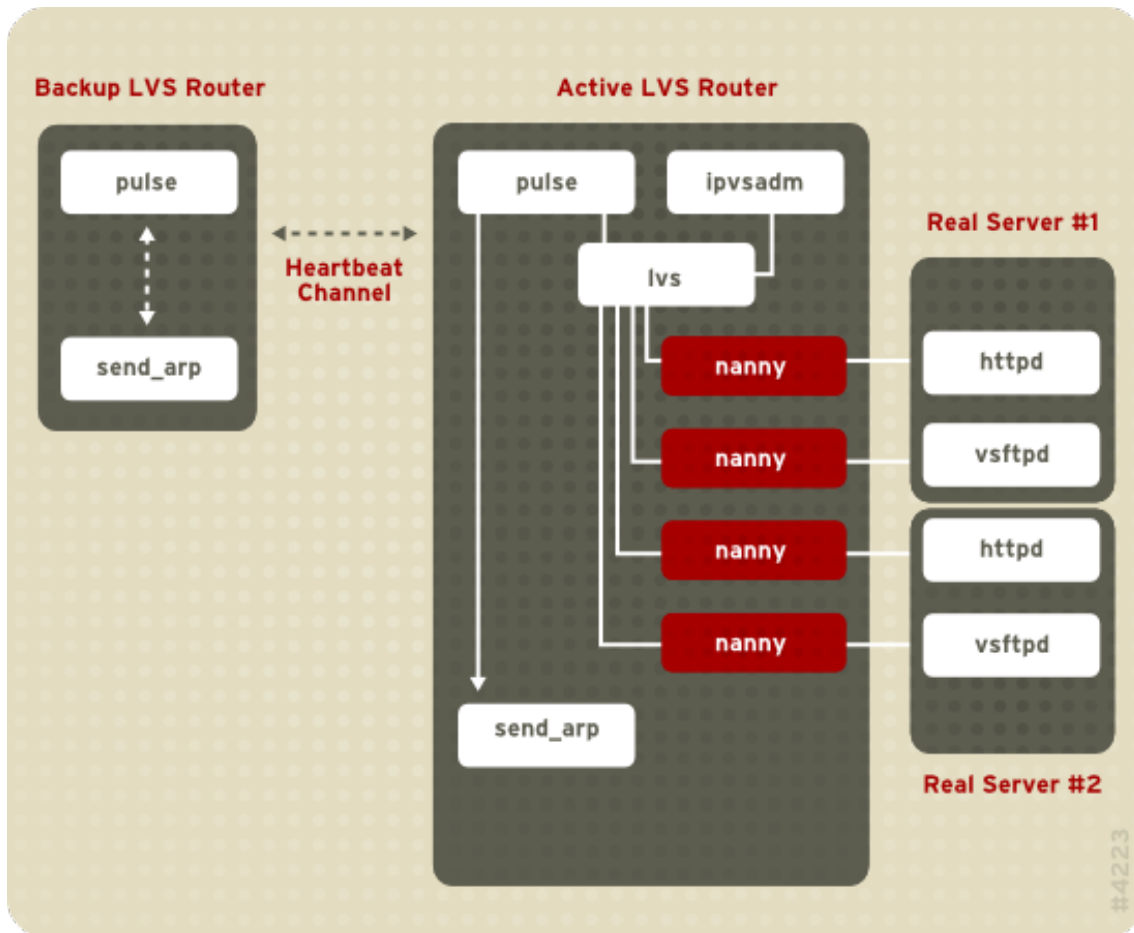


Figura 1.19. Components of a Running LVS Cluster

O `pulse` daemon atua nos roteadores LVS ativos e passivos. Em um roteador LVS de backup, o `pulse` envia um *heartbeat* para a interface pública de um roteador ativo para garantir que o roteador LVS esteja funcionando apropriadamente. Num roteador LVS ativo, o `pulse` inicia o `lvs` daemon e responde ao *heartbeat* as perguntas do roteador LVS de backup.

Um vez iniciado, o `lvs` daemon chama a utilidade `ipvsadm` para configurar e manter o IPVS (Servidor Virtual IP) de tabela ampla no kernel, e inicia um processo `nanny` para cada servidor virtual configurado em cada servidor real. Cada processo `nanny` verifica o estado de um serviço configurado em um servidor real, e informa o daemon `lvs`, caso o serviço daquele servidor real possuir mal funcionamento. Se o mal funcionamento for detectado, o daemon `lvs` induz o `ipvsadm` a remover aquele servidor real de uma tabela ampla IPVS.

Caso um roteador LVS de backup não receber uma resposta de um roteador LVS, ele iniciará a falha chamando o `send_arp` para reinstalar todos os endereços IP virtuais nos endereços de hardware NIC (endereços MAC) de roteador LVS de backup. Isto enviará um comando para o roteador LVS ativo por meio das interfaces da rede pública e privada, desativando o daemon `lvs` num roteador LVS ativo e iniciando o daemon `lvs`, no roteador LVS de backup, para aceitação das solicitações de servidores virtuais configurados.

No caso de um usuário de fora estar acessando um serviço anfitrião (como um website ou aplicação de banco de dados), o LVS se apresenta como um servidor. No entanto, o usuário estará acessando os servidores reais por trás dos roteadores LVS.

Devido ao LVS não possuir o componente construído para compartilhamento dos dados entre os servidores reais, você terá duas opções básicas:

- Sincronizar os dados através de servidores reais.

- Adicionar a terceira camada da topologia para acesso de dados compartilhados.

A primeira opção é preferida pelos servidores que não permitem um número grande de usuários para carregar ou mudar os dados em servidores reais. Se um servidor real permitir um número grande de usuários modificar os dados, como um website de e-comércio, é preferível a adição da terceira camada.

Existem diversas maneiras de sincronizar dados de servidores reais. Por exemplo, você pode utilizar os scripts shell para enviar simultaneamente as páginas atualizadas da web de usuários reais. Além disso, você pode usar programas como o `rsync` para replicar a mudança de dados através de todos os nós em um intervalo definido. No entanto, a sincronização de dados não funciona de maneira otimista para os ambientes onde os usuários freqüentemente fazem o carregamento do arquivo ou têm problemas em transações de datadados, usando o script ou o comando `rsync`. Portanto, para os servidores reais com um alto número de carregamento, transações de datadados, ou tráfego similar, a *topologia de três camadas* é a mais apropriada para a sincronização de dados.

1.8.1. Two-Tier LVS Topology

[Figura 1.20, “Two-Tier LVS Topology”](#) shows a simple LVS configuration consisting of two tiers: LVS routers and real servers. The LVS-router tier consists of one active LVS router and one backup LVS router. The real-server tier consists of real servers connected to the private network. Each LVS router has two network interfaces: one connected to a public network (Internet) and one connected to a private network. A network interface connected to each network allows the LVS routers to regulate traffic between clients on the public network and the real servers on the private network. In [Figura 1.20, “Two-Tier LVS Topology”](#), the active LVS router uses *Network Address Translation (NAT)* to direct traffic from the public network to real servers on the private network, which in turn provide services as requested. The real servers pass all public traffic through the active LVS router. From the perspective of clients on the public network, the LVS router appears as one entity.

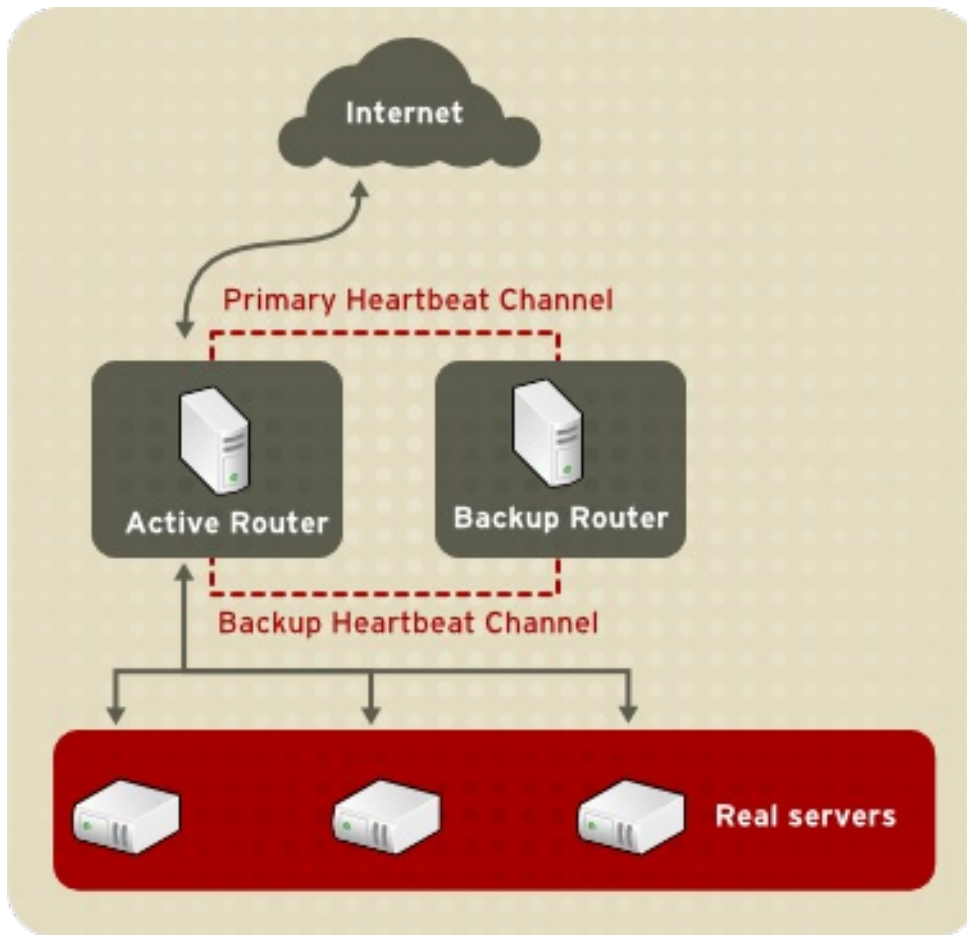


Figura 1.20. Two-Tier LVS Topology

O serviço solicita que a chegada no roteador LVS seja endereçada a um endereço *IP virtual* ou VIP. Este é um endereço roteado publicamente do administrador do site de associados, com o nome dominante inteiramente qualificado. Por exemplo: `www.example.com`, que é determinado para um ou mais *servidores virtuais*^[1]. Perceba que um endereço VIP migra de um para outro roteador LVS durante a falha, deste modo mantendo a presença no endereço IP, também conhecido como *endereço IP flutuante*.

Os endereços VIP podem ser alias no mesmo dispositivo que conecta o roteador LVS à rede pública. Por exemplo, se o `eth0` é conectado à Internet, então diversos servidores virtuais podem ser alias ao `eth0:1`. Alternativamente, cada servidor virtual pode ser associado com o dispositivo separado por serviço. Por exemplo, o tráfego HTTP pode ser tratado em `eth0:1`, e o tráfego FTP pode ser tratado em `eth0:2`.

Apenas o roteador LVS está ativo desta vez. A função do roteador LVS ativo é de redirecionar os serviços solicitados de um endereço virtual IP aos servidores reais. A redireção é baseada em um destes oito algoritmo de carga balanceada:

- Round-Robin Scheduling – Distribui cada solicitação seqüencial em volta de um pool de servidores reais. Usando este algoritmo, todos os servidores reais são tratados como equivalentes sem considerar a capacidade ou carga.
- Weighted Round-Robin Scheduling – Distribui cada solicitação seqüencial em volta do pool de servidores reais, mas gera mais trabalho para servidores de grande capacidade. A capacidade é indicada por um usuário determinado pelo fator peso, que então é ajustado acima e abaixo por informação de carga dinâmica. Esta é a escolha preferida se houver diferenças significantes entre a capacidade de servidores reais em um pool de servidor. No entanto, se a solicitação de carga variar dramaticamente, um servidor de carga mais pesada, poderá responder além de sua própria repartição de solicitações.

- **Least-Connection** – Distribui mais solicitações de servidores reais com poucas conexões ativas. Essa conexão é um algoritmo de agendamento dinâmico, executando uma melhor escolha em caso de existir um alto grau de variação nesta carga solicitada. É, também, a conexão mais adequada para um pool de servidor real, onde cada nó de servidor possui aproximadamente a mesma capacidade. Em caso de servidores reais possuírem capacidades variáveis, o agendamento de conexão mínima carregada será a melhor escolha.
- **Weighted Least-Connections (padrão)** – Distribui mais solicitações para servidores com poucas conexões ativas relativas às suas capacidades. A capacidade é indicada pela carga determinada pelo usuário, o que é então ajustada para acima ou para baixo por uma informação de carga dinâmica. A adição de carregamento faz com que este algoritmo seja perfeito, quando o pool de servidor real conter o hardware de capacidade variável.
- **Locality-Based Least-Connection Scheduling** – Distribui mais solicitações aos servidores com poucas conexões ativas relativas as suas destinações IPs. Esse algoritmo é para uso de um cluster de servidor proxy-cache. Ele roteia os pacotes dos endereços IP para o servidor daquele endereço, a não ser que este servidor esteja acima da capacidade e possua um servidor de baixa carga. Neste caso, isto determinará o endereço IP do servidor real de carregamento mínimo.
- **Locality-Based Least-Connection Scheduling with Replication Scheduling** – Distribui mais solicitações aos servidores com poucas conexões ativas relativas as suas destinações IPs. Esse algoritmo é também para o uso de um cluster de servidor proxy-cache. Isto diferencia um Locality-Based Least-Connection Scheduling pelo mapeamento do endereço alvo IP ao subset do servidor real.
- **Source Hash Scheduling** – Distribui solicitações ao pool de servidores reais, checando o recurso IP numa mesa hash estática. Este algoritmo é para os roteadores LVS com múltiplos firewalls.

Além disso, o roteador LVS ativo monitora dinamicamente o health, em geral, de serviços específicos nos servidores reais através de um simples *script enviado/esperado*. Para assistir na detecção do health dos servidores que requerem os dados dinâmicos, como os HTTPS ou SSL, você pode também chamar por execuções externas. Caso um serviço em um servidor real não funcione corretamente, o roteador LVS ativo interrompe o envio de trabalhos para o servidor até que isto retorne à operação normal.

O roteador LVS de backup executa a função de um sistema standby (estado de prontidão). Periodicamente, os roteadores LVS mudam as mensagens hearbeat através da interface pública externa primária. No caso do roteador LVS de backup falhar em receber a mensagem heartbeat no intervalo esperado, a falha será iniciada e assumirá a função do roteador LVS ativo. Durante a falha, o roteador LVS de backup é responsável pelos serviços de endereços VIP por um roteador falhado usando uma técnica conhecida como *ARP spoofing* – onde o roteador LVS de backup anuncia-se como destino dos pacotes endereçados para os nós falhados. Quando os nós falhados retornarem ao serviço ativo, o roteador LVS de backup assumirá sua função de backup novamente.

The simple, two-tier configuration in [Figura 1.20, “Two-Tier LVS Topology”](#) is suited best for clusters serving data that does not change very frequently – such as static web pages – because the individual real servers do not automatically synchronize data among themselves.

1.8.2. Three-Tier LVS Topology

[Figura 1.21, “Three-Tier LVS Topology”](#) shows a typical three-tier LVS configuration. In the example, the active LVS router routes the requests from the public network (Internet) to the second tier – real servers. Each real server then accesses a shared data source of a Red Hat cluster in the third tier over the private network.

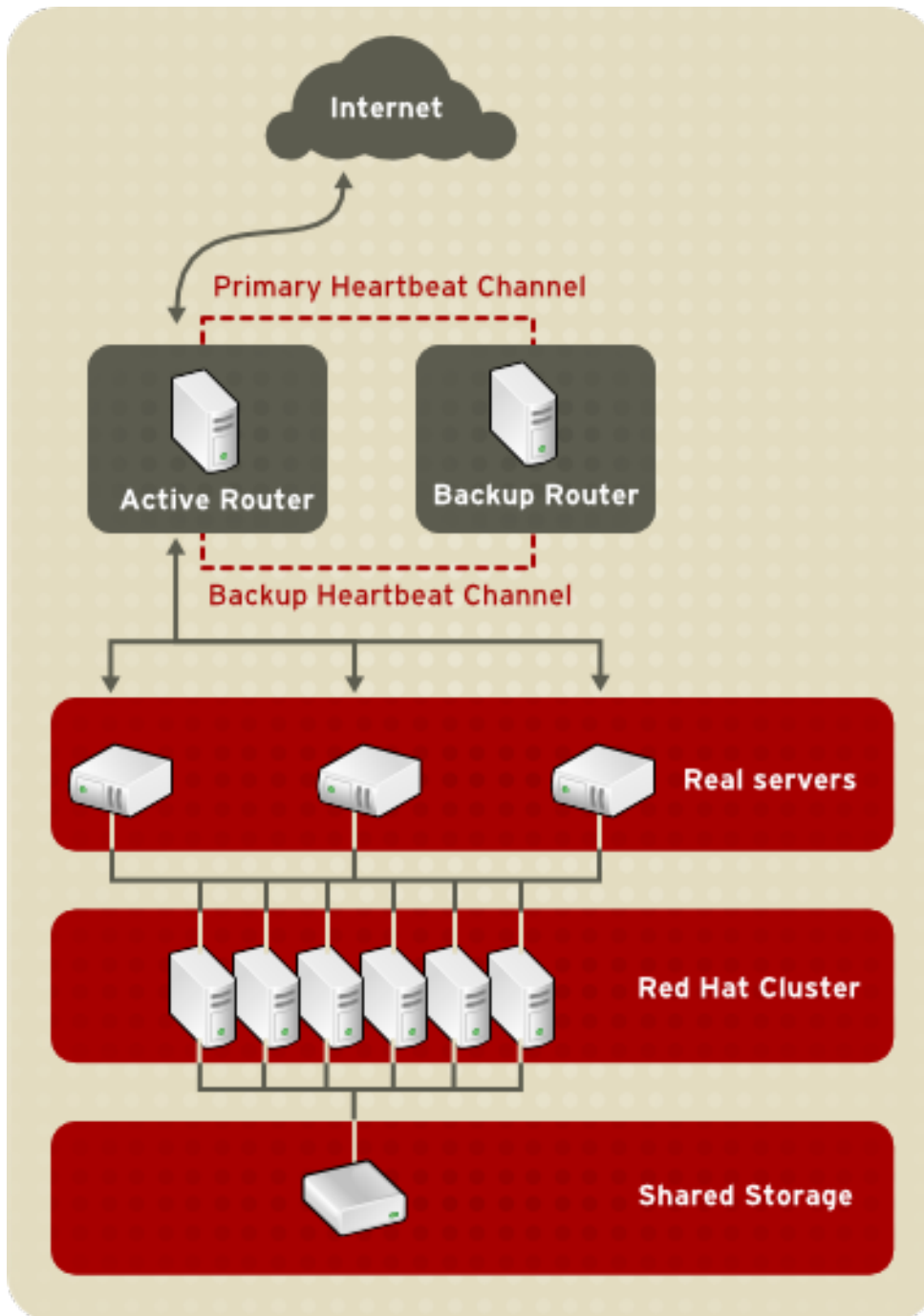


Figura 1.21. Three-Tier LVS Topology

Armazenada num servidor central, altamente disponível e acessado por cada servidor real por meio de um diretório NFS exportado ou Samba compartilhado. Essa topologia é também compartilhada por websites que acessam um banco de dados central e de alta disponibilidade para transações. Adicionalmente, usando uma configuração ativa-ativa com um cluster Red Hat, você pode configurar um cluster de alta disponibilidade para satisfazer ambas funções simultaneamente.

1.8.3. Métodos Encaminhados

Você pode usar a Network Address Translation (NAT, Tradução de Endereço da Rede) roteando ou roteando diretamente o LVS. As seguintes seções descreverão rapidamente o roteamento ou roteamento direto do LVS.

1.8.3.1. Roteamento NAT

Figura 1.22, “LVS Implemented with NAT Routing”, illustrates LVS using NAT routing to move requests between the Internet and a private network.

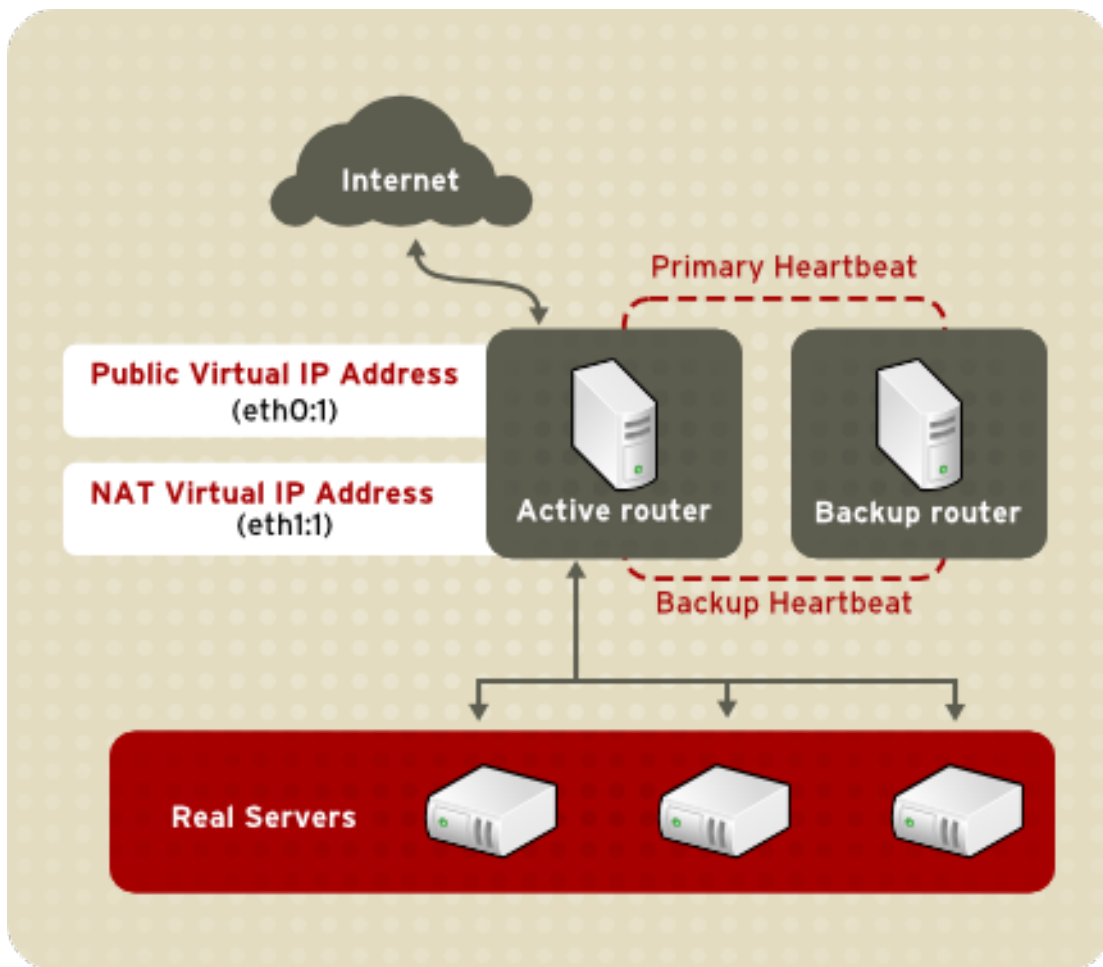


Figura 1.22. LVS Implemented with NAT Routing

Neste exemplo, existem dois NICs no roteador LVS ativo. O NIC para a Internet possui um *endereço IP real* no eth0, e também possui um alias de endereço IP flutuante para o eth0:1. O NIC para a interface de rede privada possui um endereço IP real no eth1, e possui um alias de endereço IP flutuante para o eth1:1. No caso de uma falha, a interface virtual de cobertura da Internet e a interface virtual de cobertura privada são comandadas simultaneamente pelo roteador LVS de backup.

In the example, the LVS router's public LVS floating IP address and private NAT floating IP address are aliased to two physical NICs. While it is possible to associate each floating IP address to its physical device on the LVS router nodes, having more than two NICs is not a requirement.

Usando esta topologia, o roteador LVS ativo recebe a solicitação e encaminha isto ao servidor apropriado. O servidor real então processa a solicitação e retorna aos pacotes do roteador LVS. O roteador LVS usa a tradução do endereço da rede para substituir o endereço do servidor real, nos pacotes de roteadores LVS públicos de endereço VIP. Este processo é chamado *IP masquerading* devido aos endereços IP atuais, de servidores reais, estarem ocultos para os clientes solicitantes.

Usando o roteamento NAT, os servidores reais podem ser qualquer tipo de computadores atuando numa variedade de sistemas operantes. A maior desvantagem do roteamento NAT é que o roteador LVS pode iniciar um funilamento de grande extensão, pois ele deve processar as solicitações de entrada e saída.

1.8.3.2. Roteamento direto

O roteamento direto fornece um benefício no aumento do desempenho comparado ao roteamento NAT. O roteamento direto permite os servidores reais processarem e encaminharem os pacotes diretamente a um usuário requerido, em vez de passar os pacotes de saída através do roteador LVS. O roteamento direto reduz a possibilidade dos problemas de desempenho da rede, apenas transmitindo o trabalho do roteador LVS para o processamento dos pacotes de início.

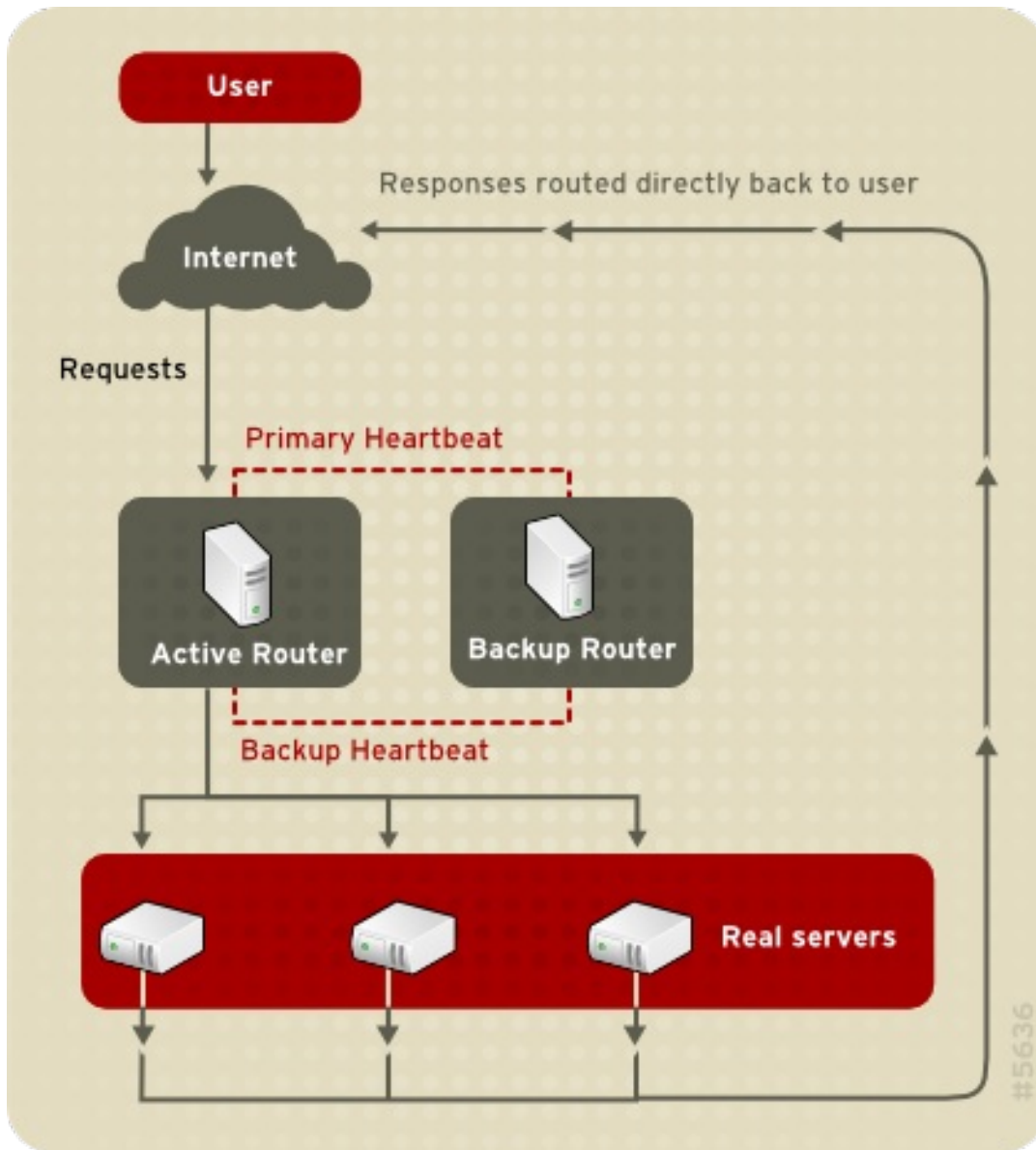


Figura 1.23. LVS Implemented with Direct Routing

Numa configuração típica de roteamento direto LVS, o roteador LVS recebe solicitações de servidores de entrada através de um IP virtual (VIP), e utiliza um algoritmo agendado para encaminhar a solicitação aos servidores reais. Cada servidor processa as solicitações e envia respostas diretamente aos clientes, contornando os roteadores LVS. O roteamento direto permite escalabilidade nos servidores reais que podem ser aderidos sem o volume adicionado, no roteador LVS para pacotes de saída de um servidor real para cliente. Este pode iniciar um funilamento sob carga pesada da rede.

Enquanto existem muitas vantagens na utilização do roteamento direto em LVS, há também limitações. O problema mais comum no roteamento direto e LVS é o *Protocolo de Resolução de Endereço* (ARP).

In typical situations, a client on the Internet sends a request to an IP address. Network routers typically send requests to their destination by relating IP addresses to a machine's MAC address with ARP. ARP requests are broadcast to all connected machines on a network, and the machine with the correct

IP/MAC address combination receives the packet. The IP/MAC associations are stored in an ARP cache, which is cleared periodically (usually every 15 minutes) and refilled with IP/MAC associations.

O problema com as solicitações ARP em uma configuração de roteamento direto LVS, é devido a solicitação de um cliente a um endereço IP ser associada ao endereço MAC para a solicitação ser executada. O endereço IP virtual de roteador LVS deve ser associado ao MAC. No entanto, devido ao roteador LVS e os mesmos servidores possuírem o mesmo VIP, a solicitação ARP é transmitida a todos os nós associados com VIP. Isto pode causar diversos problemas, como o VIP ser associado diretamente a um dos servidores reais e as solicitações serem processadas diretamente, contornando o roteador LVS completamente e derrotando o propósito da configuração LVS. A utilização de um roteador LVS com o poderoso CPU, do qual responde rapidamente às solicitações de clientes, não é a solução para este problema. Se o roteador LVS estiver sob alto carregamento, isto poderá responder a solicitação IP mais vagarosamente do que um servidor real subutilizado. Este responderá mais rapidamente e determinará o VIP num cache ARP do cliente solicitante.

Para resolver este problema, as solicitações seguintes devem *apenas* associar o VIP ao roteador LVS, do qual irá processar propriamente as solicitações e enviá-las para o pool de servidor real. Isto pode ser feito usando a ferramenta de filtração do pacote *arptables*.

1.8.4. Marcas de Firewall e Persistência

Em algumas situações, talvez seja desejável um cliente reconectar repetidamente o mesmo servidor real, em vez de ter um algoritmo de balanceamento de carga LVS enviando uma solicitação ao melhor servidor disponível. Alguns exemplos disto incluem os formulários da web de tela múltipla, cookies, SSL e conexões FTP. Nestes casos, o cliente talvez não trabalhe propriamente, a não ser que as transações sejam executadas pelo mesmo servidor para reter o contexto. O LVS fornece duas características diferentes para cuidar disto: *persistência* e *marcas firewall*.

1.8.4.1. Persistence

Uma vez ativada, a persistência atua como um marcador de tempo. Quando um cliente se conectar a um serviço, o LVS lembra da última conexão por um período específico. Caso o mesmo cliente de endereço IP conectar novamente no mesmo período, isto será enviado ao mesmo servidor que foi conectado anteriormente – contornando os mecanismos de balanceamento de carga. Quando a conexão ocorrer fora da janela de tempo, isto será tratada de acordo com as regras de agendamento do local.

A persistência também permite que você especifique um subnet mask (*marcas subnet*), para aplicação ao teste de endereço IP do cliente, como uma ferramenta de controle dos endereços que possuem um alto nível de persistência, agrupando as conexões ao subnet.

As conexões agrupadas destinam-se a portos diferentes, podendo ser importantes para protocolos que usam mais de um porto para comunicação, como o FTP. No entanto, a persistência não é a maneira mais eficiente para cuidar do problema de agrupamento juntamente com conexões destinadas para portos diferentes. No caso destas situações, é aconselhável utilizar as *marcas firewall*.

1.8.4.2. Marcas Firewall

As marcas Firewall são uma maneira fácil e eficiente no grupo de portos utilizados para um protocolo ou grupo de protocolos relatados. Por exemplo, se o LVS for implantado para executar um site de comércio eletrônico, as marcas firewall podem ser usadas para juntar as conexões HTTP no porto 80 e segurar as conexões HTTPS no porto 443. Determinando a mesma marca firewall ao servidor virtual de cada protocolo a informação de estado para cada transação pode ser preservada devido ao roteador LVS seguir em frente com todas as solicitações do mesmo servidor real após a conexão ser aberta.

Devido a sua eficiência e uso fácil, os administradores do LVS devem usar marcas firewall ao invés de persistência, assim que possível nas conexões de agrupamento. No entanto, você deve continuar adicionando persistência aos servidores virtuais em conjunção com as marcas firewall, para garantir que os clientes sejam reconectados ao mesmo servidor por um período adequado de tempo.

1.9. FERRAMENTAS DE ADMINISTRAÇÃO DE CLUSTER

O Red Hat Cluster Suite fornece uma variedade de ferramentas para configurar e gerenciar o seu Cluster Red Hat. Esta seção fornece uma visão geral das ferramentas de administração disponíveis com o Red Hat Cluster Suite;

- [Seção 1.9.1, “Conga”](#)
- [Seção 1.9.2, “GUI Administração de Cluster”](#)
- [Seção 1.9.3, “Ferramentas de Administração de Linha de Comando”](#)

1.9.1. Conga

Conga é um grupo integrado de componentes de software que fornece a configuração e o gerenciamento de cluster, e armazenamento do Red Hat. O **Conga** fornece as principais características:

- Uma interface da Web para gerenciamento e armazenamento de cluster
- Implantação Automática de dados de Cluster e Pacotes de Apoio
- Fácil Integração com Clusters Existentes
- Não Necessita da Re-Authenticação
- Integração de Status de Clusters e Logs
- Controle Fino-Granulado sobre Permissão do Usuário

Os componentes primários do **Conga** são **luci** e **ricci**, dos quais são instalados separadamente. O **luci** é um servidor que atua em um computador e se comunica com múltiplos clusters e computadores por meio do **ricci**. O **ricci** é um agente que atua em cada computador (tanto um membro de cluster ou um computador autônomo) gerenciado pelo **Conga**.

O **luci** é acessado através do navegador da Web e fornece três funções principais das quais são acessadas através das seguintes tabelas:

- **homebase** – fornece ferramentas para a adição e deleção de computadores, adição e deleção de usuários, e configuração de usuários privilegiados. Apenas um sistema administrador é permitido a acessar esta tab.
- **cluster** – Fornece ferramentas para a criação e configuração de clusters. Cada instância das listas de clusters **luci** das quais foram montadas com o **luci**. Um administrador de sistemas pode administrar todos os clusters listados nesta tab. Outros usuários podem administrar apenas os clusters em que o usuário possui permissão para gerenciar. (concedido por um administrador).
- **storage** – Fornece ferramentas para uma administração remota de armazenamento. Com as ferramentas desta tab, você pode gerenciar o armazenamento em computadores, sendo estes pertencentes a um cluster ou não.

Para administrar um cluster ou armazenamento, um administrador adiciona (ou *registradores*) o cluster ou um computador a um servidor **luci**. Quando um cluster ou um computador é registrado com o **luci**, o nome anfitrião FQDN ou endereço IP de cada computador é armazenado num banco de dados **luci**.

Você pode propagar um banco de dados de uma instância **luci** à outra instância **luci**. Esta capacidade fornece um significado de replicação de uma instância de servidor e fornece um melhoramento eficiente e avaliação do caminho. Quando você instalar uma instância do **luci**, este banco de dados estará vazio. No entanto, você pode importar parte ou todo banco de dados **luci** de um servidor existente **luci**, quando implantando um novo servidor **luci**.

Cada instância **luci** possui um usuário na instalação inicial – **admin**. Apenas o usuário **adm** poderia adicionar sistemas ao servidor **luci**. Além disso, o usuário **admin** pode criar contas de usuários adicionais e determinar quais são os usuários que possuem acesso permitido aos clusters e computadores registrados, no banco de dados **luci**. É possível importar usuários como uma operação em lote de um novo servidor **luci**, assim como é possível importar clusters e computadores.

Quando um computador é adicionado a um servidor **luci** a ser administrado, a autenticação é feita apenas uma vez. Então, a autenticação não será mais necessária (a não ser que o certificado usado seja anulado por um CA). Após isto, você pode remotamente configurar e gerenciar os clusters e os armazenamento através da interface do usuário **luci**.

As figuras seguintes apresentam um exemplo exibido das três principais tabs **luci**: **homebase**, **cluster**, e **storage**.

Para maiores informações sobre o **Conga**, refira-se a *Configurando e Gerenciando um Cluster Red Hat* à ajuda disponível online com o servidor **luci**.

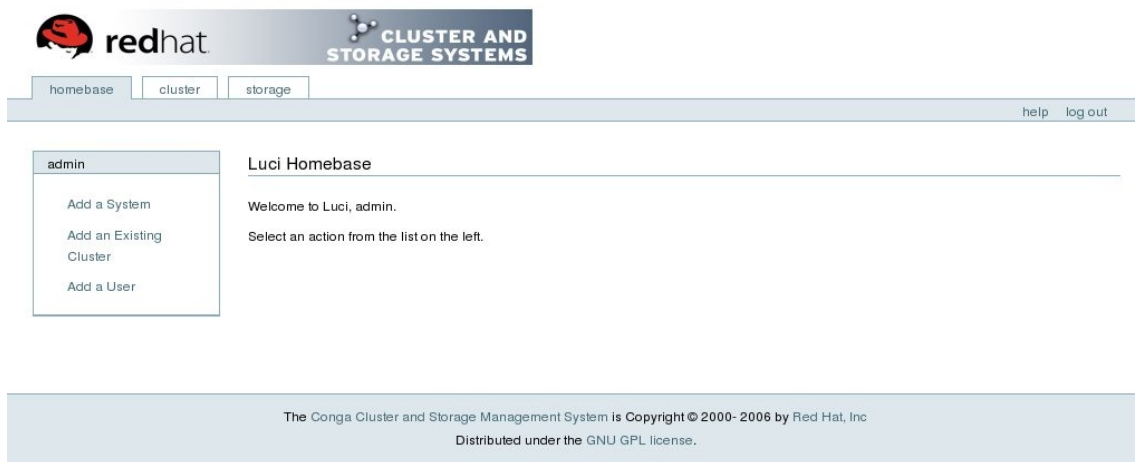


Figura 1.24. luci homebase Tab

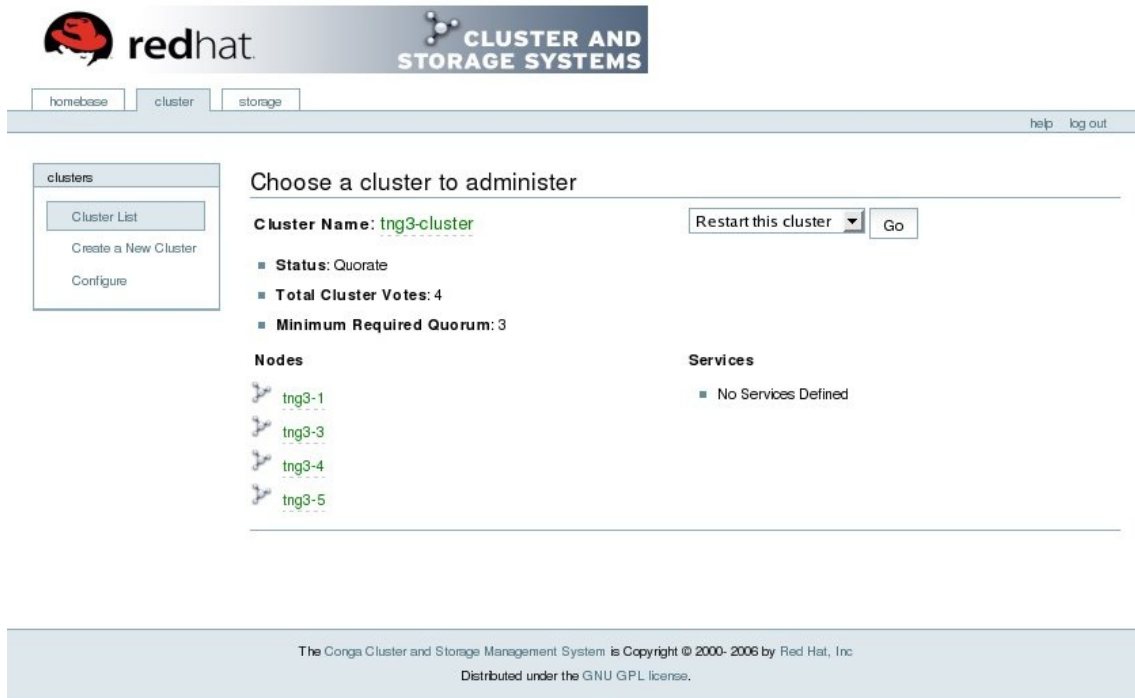


Figura 1.25. luci cluster Tab

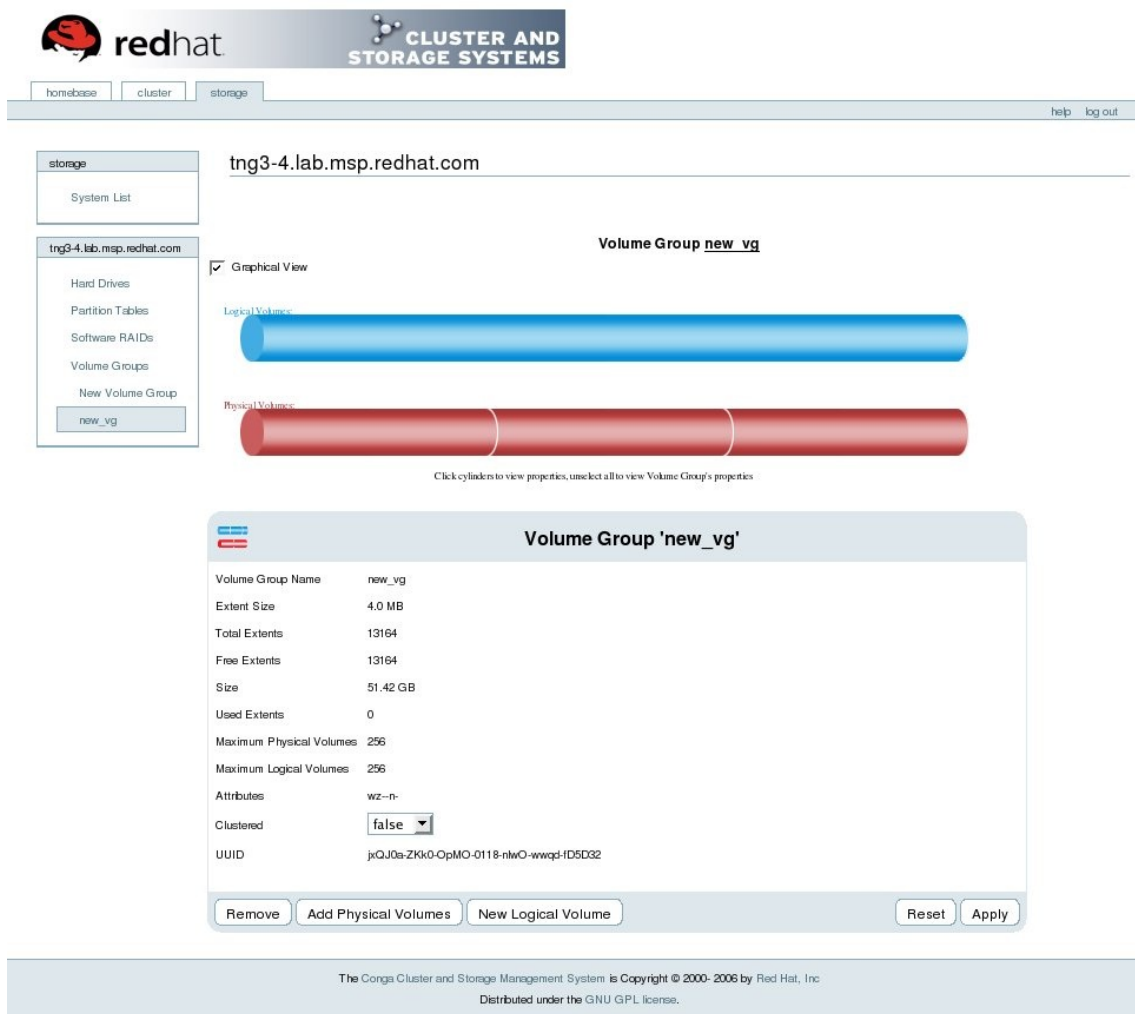


Figura 1.26. luci storage Tab

1.9.2. GUI Administração de Cluster

This section provides an overview of the `system-config-cluster` cluster administration graphical user interface (GUI) available with Red Hat Cluster Suite. The GUI is for use with the cluster infrastructure and the high-availability service management components (refer to [Seção 1.3, “Cluster Infrastructure”](#) and [Seção 1.4, “Gerenciamento de Serviço de Alta Disponibilidade”](#)). The GUI consists of two major functions: the **Cluster Configuration Tool** and the **Cluster Status Tool**. The **Cluster Configuration Tool** provides the capability to create, edit, and propagate the cluster configuration file (`/etc/cluster/cluster.conf`). The **Cluster Status Tool** provides the capability to manage high-availability services. The following sections summarize those functions.

- [Seção 1.9.2.1, “Cluster Configuration Tool”](#)
- [Seção 1.9.2.2, “Cluster Status Tool”](#)

1.9.2.1. Cluster Configuration Tool

You can access the **Cluster Configuration Tool** ([Figura 1.27, “Cluster Configuration Tool”](#)) through the **Cluster Configuration** tab in the Cluster Administration GUI.

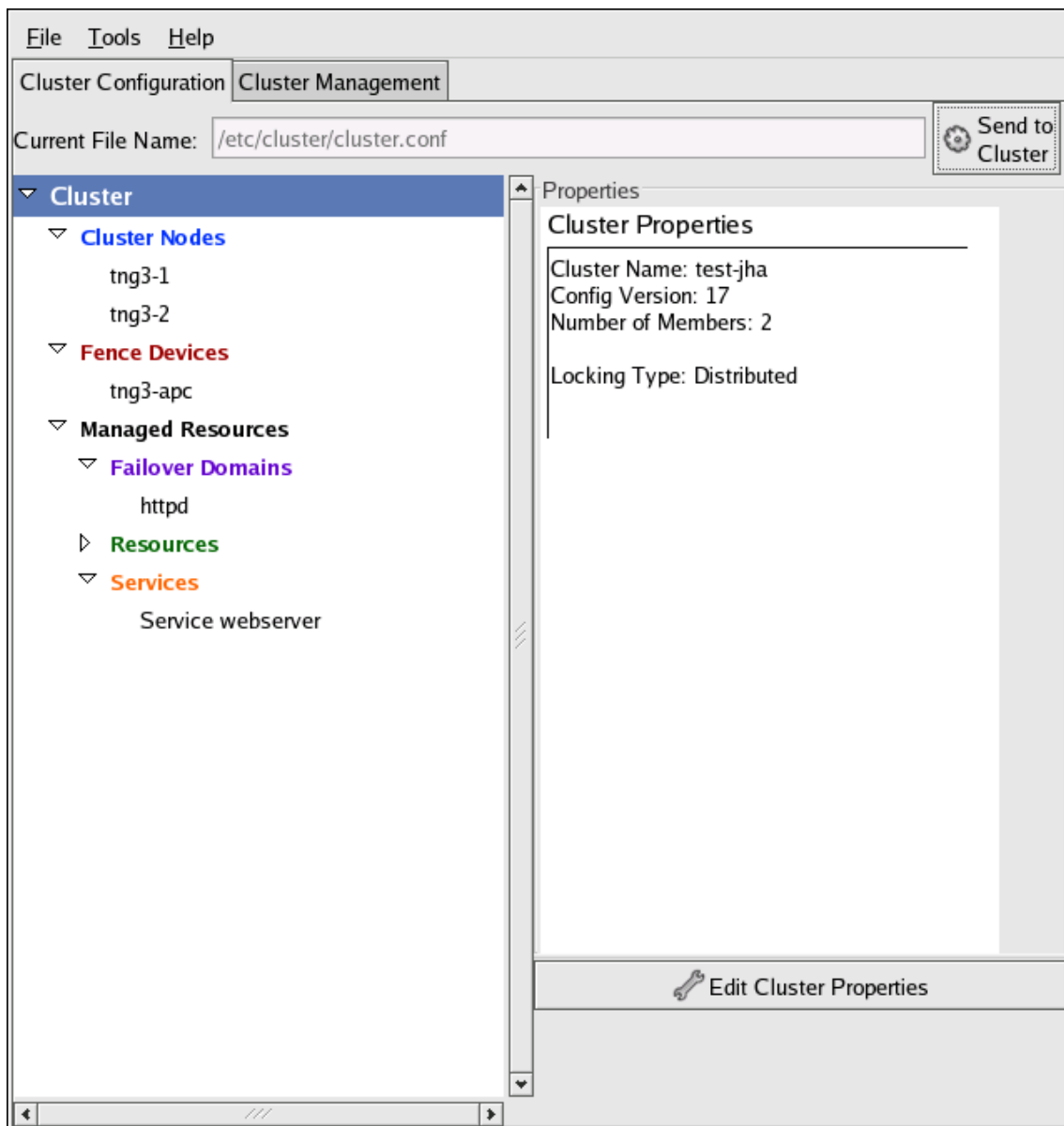


Figura 1.27. Cluster Configuration Tool

A **Cluster Configuration Tool** representa os componentes de configuração de cluster no arquivo de

configuração (`/etc/cluster/cluster.conf`), com um gráfico hierárquico exibido à esquerda do painel. Um ícone no formato de um triângulo à esquerda do nome de um componente, indica que o componente possui um ou mais componentes subordinados designados para isto. Clicando no ícone de triângulo, a porção de árvore abaixo de um componente se expande e colapsa. Os componentes exibidos no GUI são resumidos como a seguir:

- **Nós de Cluster** – Exibe os nós de cluster. Os nós são representados pelo nome assim como os elementos subordinados sob os **Nós de Cluster**. Usando os botões de configuração no botão direito do quadro (abaixo das **Propriedades**), você pode adicionar nós, deletar nós, editar propriedades de nós e configurar os métodos de fencing para cada nó.
- **Dispositivos de Fence** – Exibe dispositivos de fence. Os dispositivos de fence são representados como elementos subordinados sob os **Dispositivos de Fence**. Usando os botões de configuração no canto direito abaixo do quadro (abaixo das **Propriedades**), você pode adicionar os dispositivos de fence, deletar os dispositivos de fence e editar as propriedades de dispositivos de fence. Os dispositivos de fence devem ser definidos antes de você configurar o fencing (com o botão **Manage Fencing For This Node**) para cada nó.
- **Recursos Gerenciados** – exibe as falhas dominantes, os recursos e os serviços.
 - **Domínios de Falha** – Para configurar um ou mais subconjuntos de nós de cluster usados para executar um serviço de alta disponibilidade, num evento de uma falha de nó. Os domínios de falhas são representados como elementos subordinados sob os **Domínios de Falhas**. Usando botões de configuração no botão ao lado direito do quadro (abaixo de **Propriedades**), você pode criar as falhas dominantes (quando forem selecionados os **Domínios de Falha**) ou editar propriedades de domínio de falha (quando um domínio de falha for selecionado).
 - **Recursos** – Para configurar os recursos compartilhados a serem usados por serviços de alta disponibilidade. Os recursos compartilhados consistem nos sistemas de arquivos, endereços IP, montagem e exportação NFS e scripts de criação do usuário dos quais estão disponíveis a um serviço de alta disponibilidade num cluster. Os recursos são representados como elementos subordinados sob os **Recursos**. Usando os botões de configuração no canto direito do quadro (abaixo das **Propriedades**), você pode criar recursos (quando forem selecionados os **Recursos**) ou editar propriedades de recursos (quando um recurso é selecionado).



NOTA

A **Cluster Configuration Tool** também fornece a capacidade de configurar os recursos privados. O recurso privado é um recurso do qual é configurado para o uso de apenas um serviço. Você pode configurar um recurso privado com um componente **Service** em GUI.

- **Serviços** – Para criar e configurar serviços de alta disponibilidade. Um serviço é configurado apenas determinando recursos (compartilhados e privados), determinando o domínio de falha, e definindo uma política de recuperação para o serviço. Os serviços são representados como elementos subordinados sob os **Serviços**. Usando botões de configuração do lado direito do quadro (abaixo de **Propriedades**), você pode criar serviços (quando forem selecionados os **Serviços**) ou editar propriedades de serviço (quando um serviço for selecionado)

1.9.2.2. Cluster Status Tool

You can access the **Cluster Status Tool** (Figura 1.28, “Cluster Status Tool”) through the **Cluster Management** tab in Cluster Administration GUI.

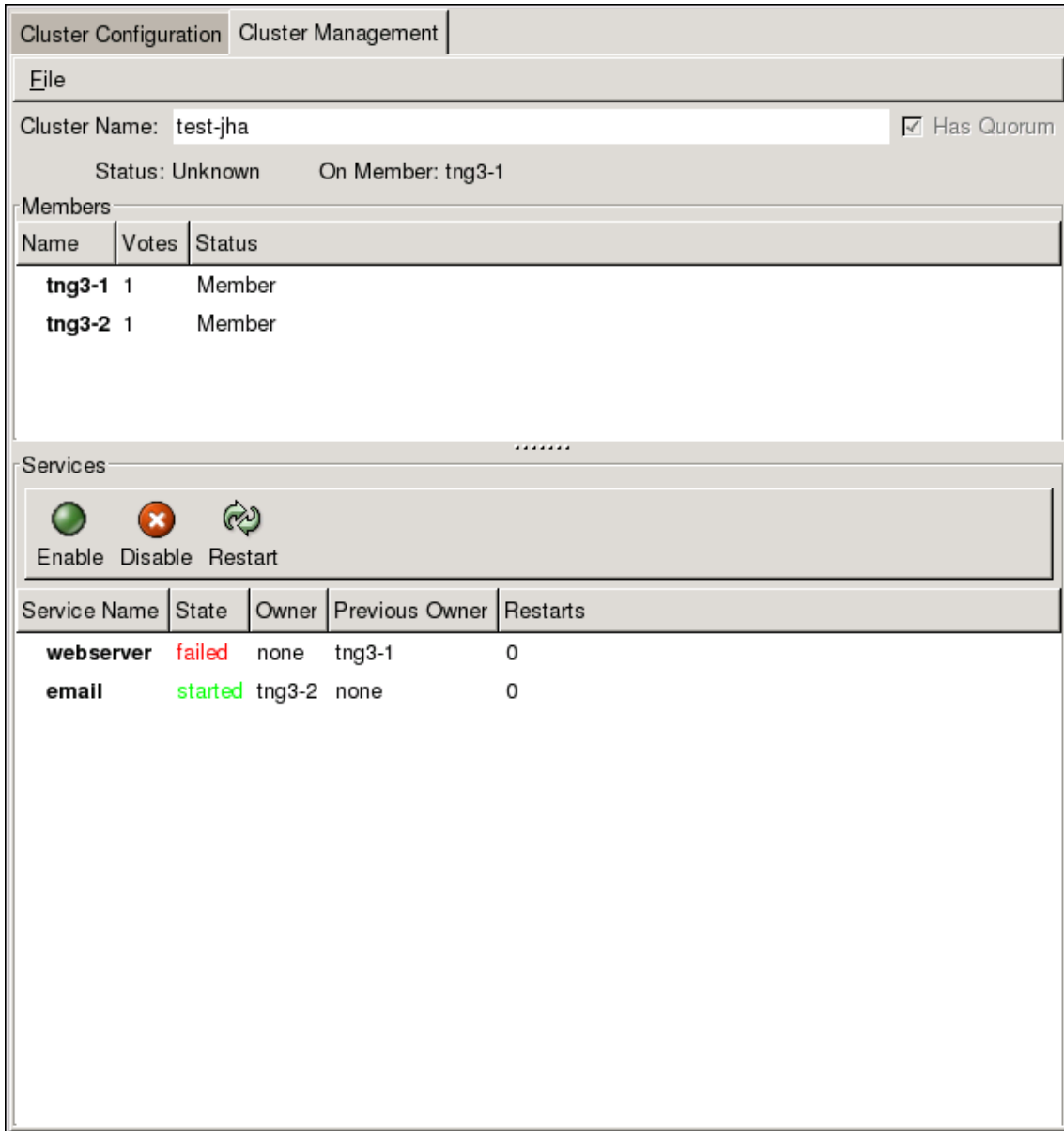


Figura 1.28. Cluster Status Tool

Os nós e serviços exibidos na **Cluster Status Tool** são determinados pelo arquivo de configuração de cluster (/etc/cluster/cluster.conf). Você pode usar a **Cluster Status Tool** para ativar, desativar, reiniciar ou redirecionar o serviço de alta disponibilidade.

1.9.3. Ferramentas de Administração de Linha de Comando

In addition to **Conga** and the **system-config-cluster** Cluster Administration GUI, command line tools are available for administering the cluster infrastructure and the high-availability service management components. The command line tools are used by the Cluster Administration GUI and init scripts supplied by Red Hat. [Tabela 1.1, “Ferramentas de Linha de Comando”](#) summarizes the command line tools.

Tabela 1.1. Ferramentas de Linha de Comando

Ferramenta de Linha de Comando	Usado com	Propósito
ccs_tool – Ferramenta de Sistema de Configuração de Cluster	Cluster Infrastructure	ccs_tool é um programa para realizar atualizações on-line do arquivo de configuração de cluster. Isto fornece a capacidade de criar e modificar os componentes de infraestrutura de cluster (por exemplo, criação do cluster, adição e remoção do nó). Para maiores informações sobre esta ferramenta, refira-se à página <code>man ccs_tool(8)</code> .
cman_tool – Ferramenta de Gerenciamento de Cluster	Cluster Infrastructure	cman_tool é o programa que administra o gerenciador de cluster CMAN. Isto fornece a capacidade de unir-se ao cluster, abandonar o cluster, cancelar o nó, ou mudar os votos de quórum esperados de um nó num cluster. Para maiores informações sobre esta ferramenta, refira-se à pagina <code>man cman_tool(8)</code> .
fence_tool(fenced) – Ferramenta de Fence	Cluster Infrastructure	fence_tool é um programa usado para unir ou abandonar o domínio de fence padrão. Especificamente, isto inicia o fence daemon (fenced) para juntar-se ao domínio e eliminar o fenced para abandonar o domínio. Para maiores informações sobre esta ferramenta, refira-se à página <code>man fence_tool(8)</code> .
clustat – Utilidade de Status de Cluster	Componentes de Gerenciamento de Serviço de Alta Disponibilidade	O comando clustat exibe o status do cluster. Isto apresenta a informação de associação, verificação do quórum e estado de todos os serviços de usuários configurados. Para maiores informações sobre esta ferramenta, refira-se a página <code>man clustat(8)</code> .
clusvcadm – Utilidade de Administração de Serviço do Usuário de Cluster	Componentes de Gerenciamento de Serviço de Alta Disponibilidade	O comando clusvcadm permite que você ative, desative, redirecione e reinicie os serviços de alta disponibilidade num cluster. Para maiores informações sobre esta ferramenta, refira-se à página <code>man clusvcadm(8)</code> .

1.10. GUI ADMINISTRADOR DE SERVIÇO VIRTUAL LINUX

Esta seção fornece uma visão geral da ferramenta de configuração LVS disponível com o Red Hat Cluster Suite – a **Piranha Configuration Tool**. A **Piranha Configuration Tool** é uma interface gráfica de usuário (GUI) da qual fornece uma estrutura de acesso à criação do arquivo de configuração em LVS – `/etc/sysconfig/ha/lvs.cf`.

Para acessar a **Piranha Configuration Tool** você precisa que o serviço **piranha-gui** seja executado no roteador LVS ativo. Você pode acessar a **Piranha Configuration Tool** localmente ou remotamente com o navegador da Web. É possível acessar isto localmente com o URL: `http://localhost:3636`. Como também, é possível acessar isto remotamente, tanto com o nome do anfitrião como o endereço IP real seguido pelo `:3636`. Caso você esteja acessando a **Piranha Configuration Tool** remotamente, você precisará de uma conexão ssh para ativar o roteador LVS ativo como o usuário de root.

Starting the **Piranha Configuration Tool** causes the **Piranha Configuration Tool** welcome page to be displayed (refer to [Figura 1.29, “The Welcome Panel”](#)). Logging in to the welcome page provides access

to the four main screens or *panels*: **CONTROL/MONITORING**, **GLOBAL SETTINGS**, **REDUNDANCY**, and **VIRTUAL SERVERS**. In addition, the **VIRTUAL SERVERS** panel contains four *subsections*. The **CONTROL/MONITORING** panel is the first panel displayed after you log in at the welcome screen.

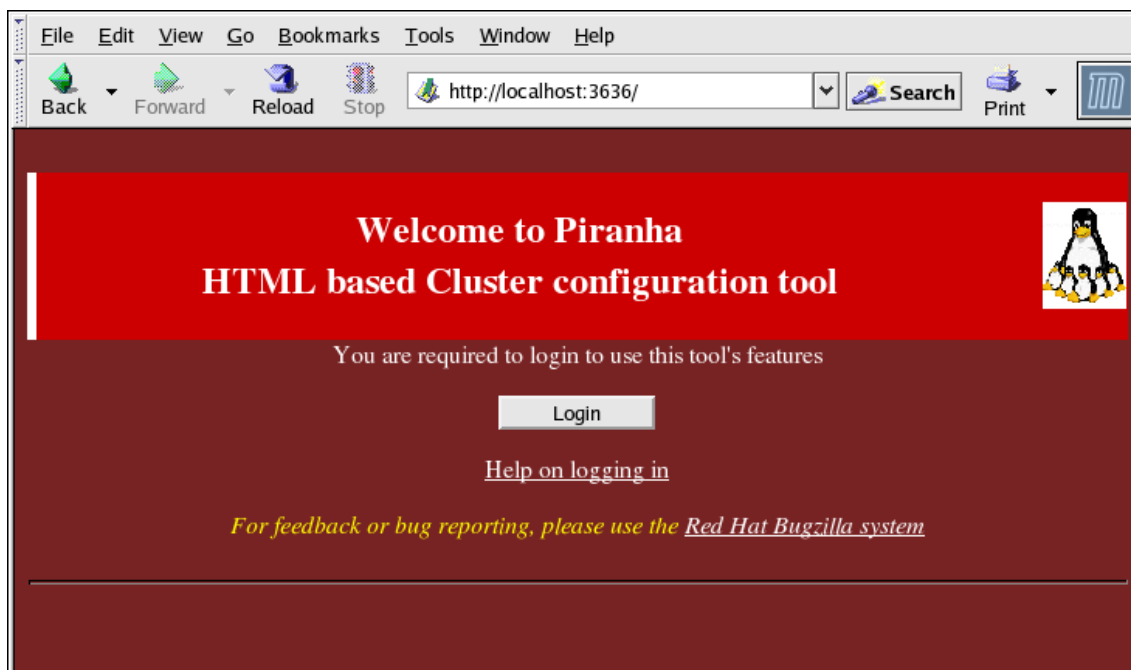


Figura 1.29. The Welcome Panel

As seções seguintes fornecem uma descrição breve das páginas de configuração da Piranha Configuration Tool.

1.10.1. CONTROL/MONITORING

O painel de **CONTROLE/MONITORAMENTO** exibe o status de tempo de execução. Isto exibe o status do daemon **puLse**, a tabela de roteamento LVS e os processos **nanny LVS** gerados.

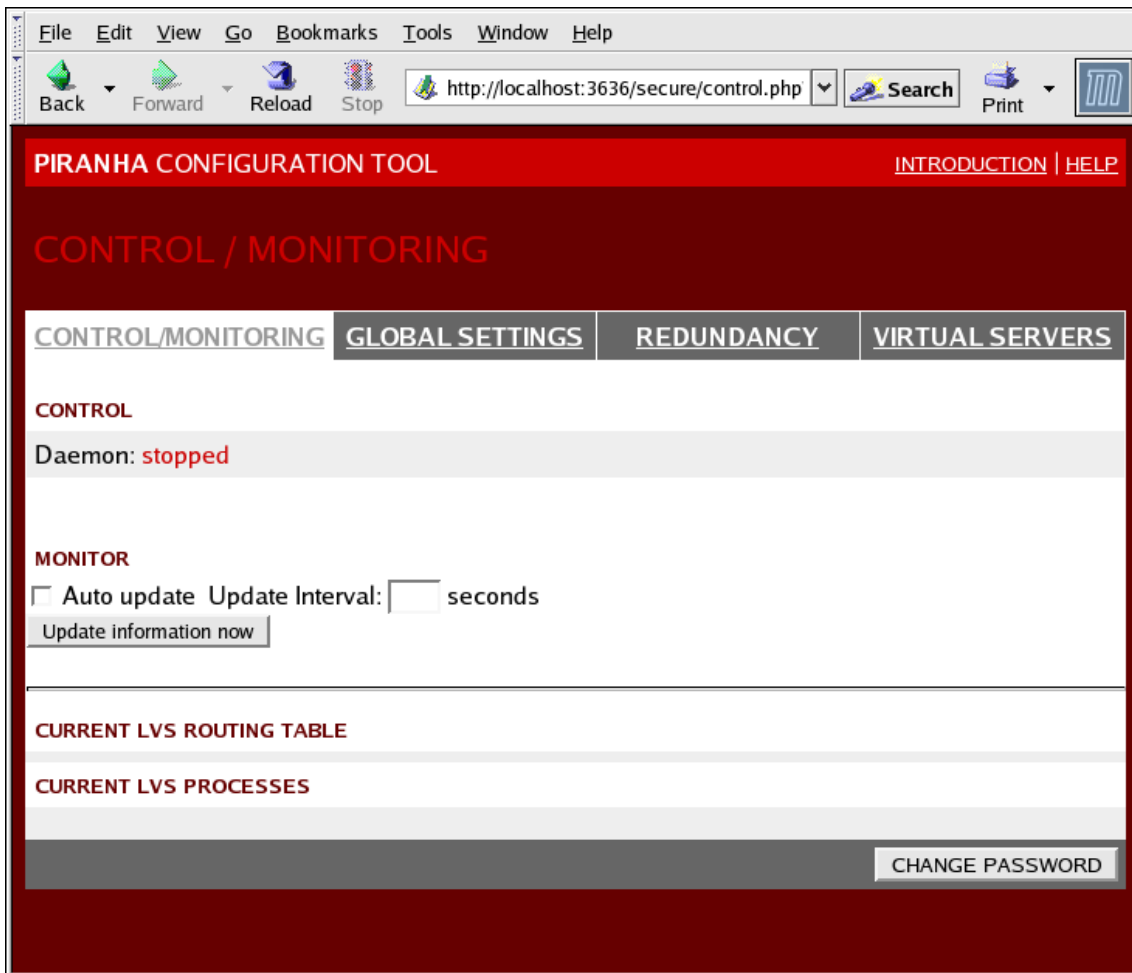


Figura 1.30. The CONTROL/MONITORING Panel

Auto update

Ativa o status exibido para ser automaticamente atualizado no intervalo configurável do usuário definido na caixa de texto de **frequência de atualização em segundos** (o valor do padrão é de 10 segundos).

Não é recomendável que você defina a atualização automática a um intervalo menor que 10 segundos. Do contrário, será difícil reconfigurar o intervalo **Auto atualização**, pois a página irá atualizar-se com uma frequência muito grande. Caso você se depare com este problema, apenas clique num outro painel e então retorne ao **CONTROLE/MONITORAMENTO**.

Update information now

Fornece a atualização do manual de status de informação

CHANGE PASSWORD

Clicando neste botão você encontrará uma tela de ajuda com a informação de como mudar a senha administrativa para a **Piranha Configuration Tool**.

1.10.2. GLOBAL SETTINGS

The **GLOBAL SETTINGS** panel is where the LVS administrator defines the networking details for the primary LVS router's public and private network interfaces.

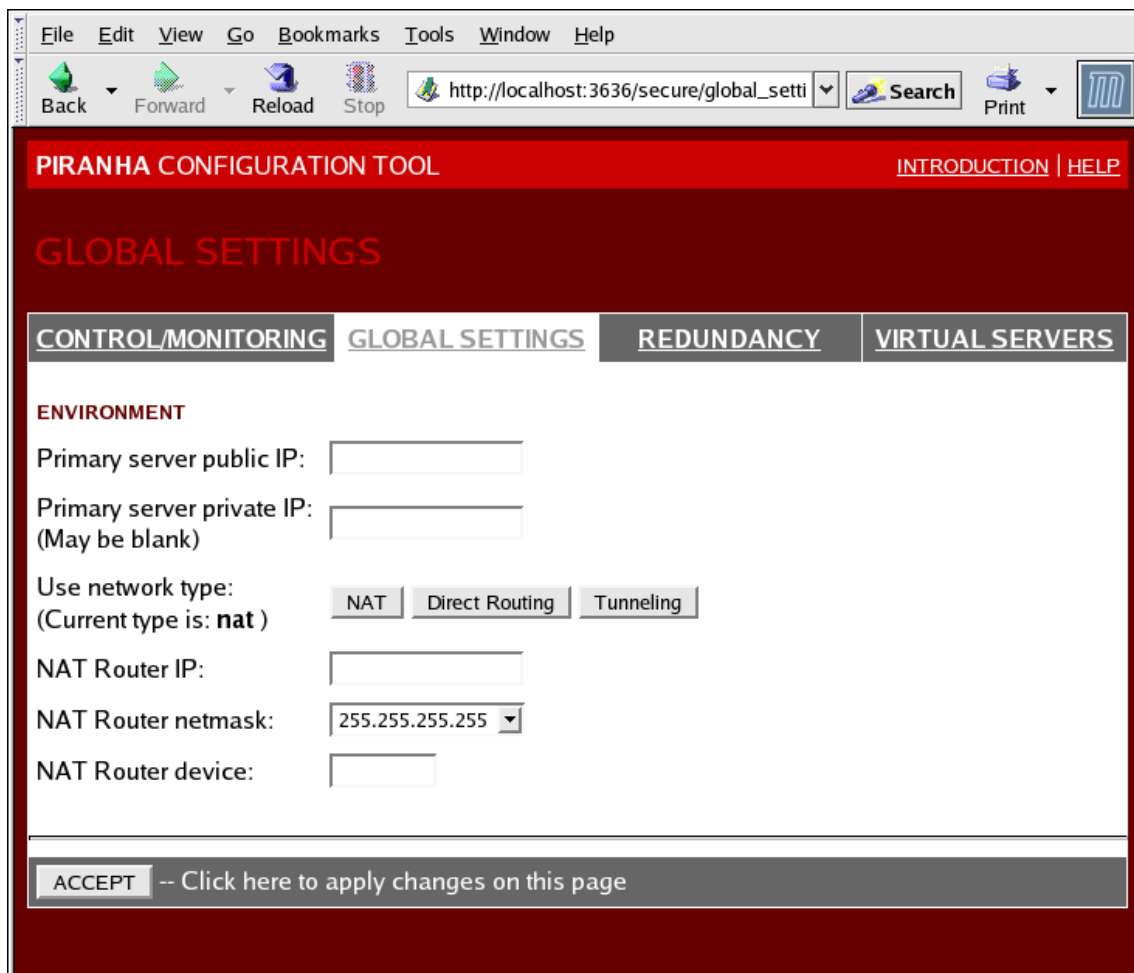


Figura 1.31. The GLOBAL SETTINGS Panel

The top half of this panel sets up the primary LVS router's public and private network interfaces.

Primary server public IP

O endereço IP real de roteamento em comum para o nó LVS primário

Primary server private IP

O endereço IP real para uma interface da rede alternativa no nó LVS primário. Este endereço é somente usado como uma alternativa do canal de heartbeat para o roteador de backup.

Use network type

Seleciona o roteamento NAT de seleção.

The next three fields are specifically for the NAT router's virtual network interface connected the private network with the real servers.

NAT Router IP

O IP flutuante privado no campo do texto. O IP flutuante deve ser usado como porta de comunicação para os servidores reais.

NAT Router netmask

If the NAT router's floating IP needs a particular netmask, select it from drop-down list.

NAT Router device

Fornece o nome do dispositivo da interface da rede para o endereço IP flutuante, como `eth1:1`.

1.10.3. REDUNDANCY

O painel de **REDUNDÂNCIA** permite que você configure o nó de roteador LVS de backup e prepara várias opções de monitoramento de heartbeat.

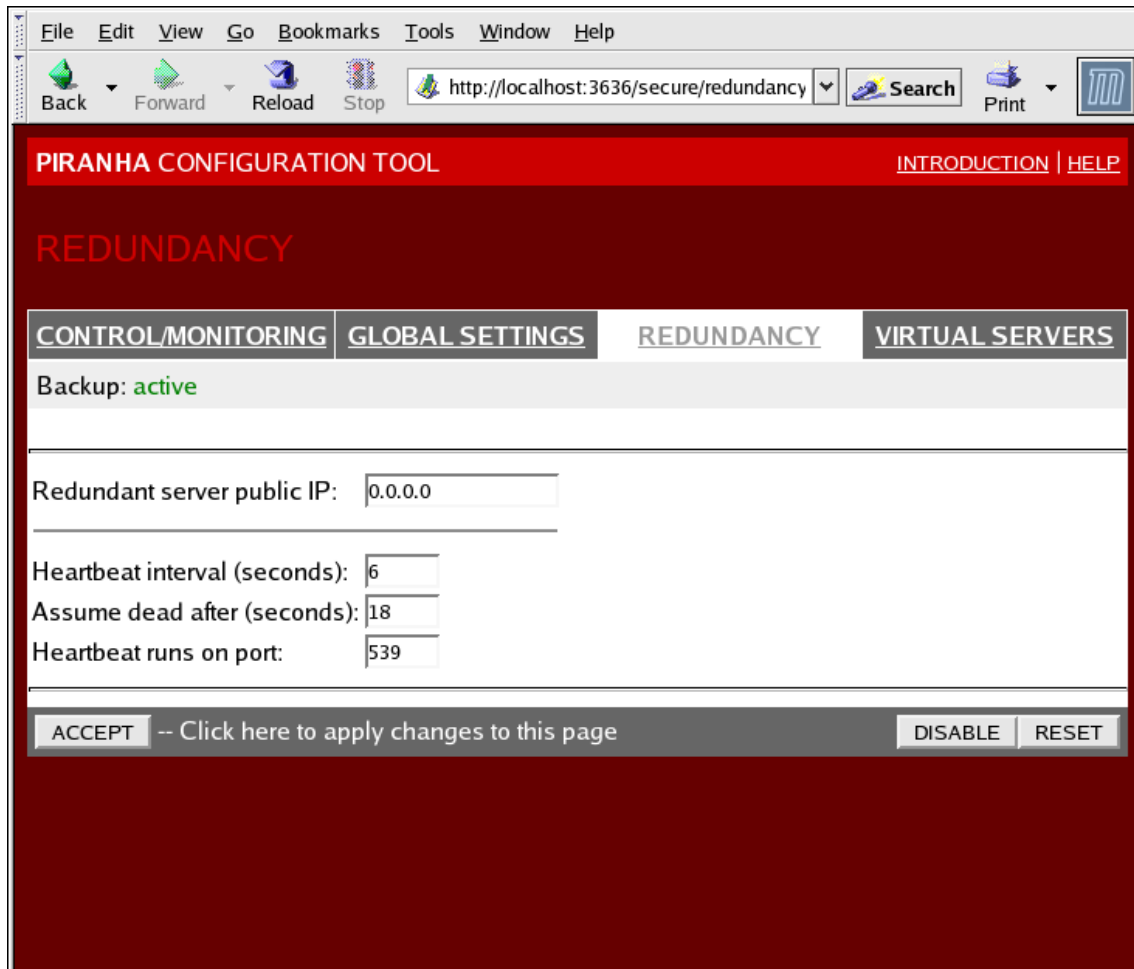


Figura 1.32. The REDUNDANCY Panel

Redundant server public IP

O endereço IP real público para o roteador de backup.

Redundant server private IP

The backup router's private real IP address.

O resto do painel serve para configuração do canal de heartbeat, que é usado pelo nó de backup para a monitoração do nó primário de falha.

Heartbeat Interval (seconds)

Marca o número de segundos entre heartbeats – o intervalo em que o nó de backup irá checar o status funcional do nó primário LVS.

Assume dead after (seconds)

Caso o nó LVS primário não responda após este número de segundos, então o nó de roteador LVS de backup iniciará a falha.

Heartbeat runs on port

Prepara o portal do qual o heartbeat se comunica com o nó LVS primário. O padrão é programado para 539 se o campo estiver em branco.

1.10.4. VIRTUAL SERVERS

O painel de **SERVIDORES VIRTUAIS** exibe informação para cada servidor virtual definido atualmente. Cada tabela de entrada apresenta o status do servidor virtual, o nome do servidor, o determinado IP virtual ao servidor, o netmask do IP virtual, o número do portal do qual o serviço comunica-se, o protocolo usado e a interface do dispositivo virtual.

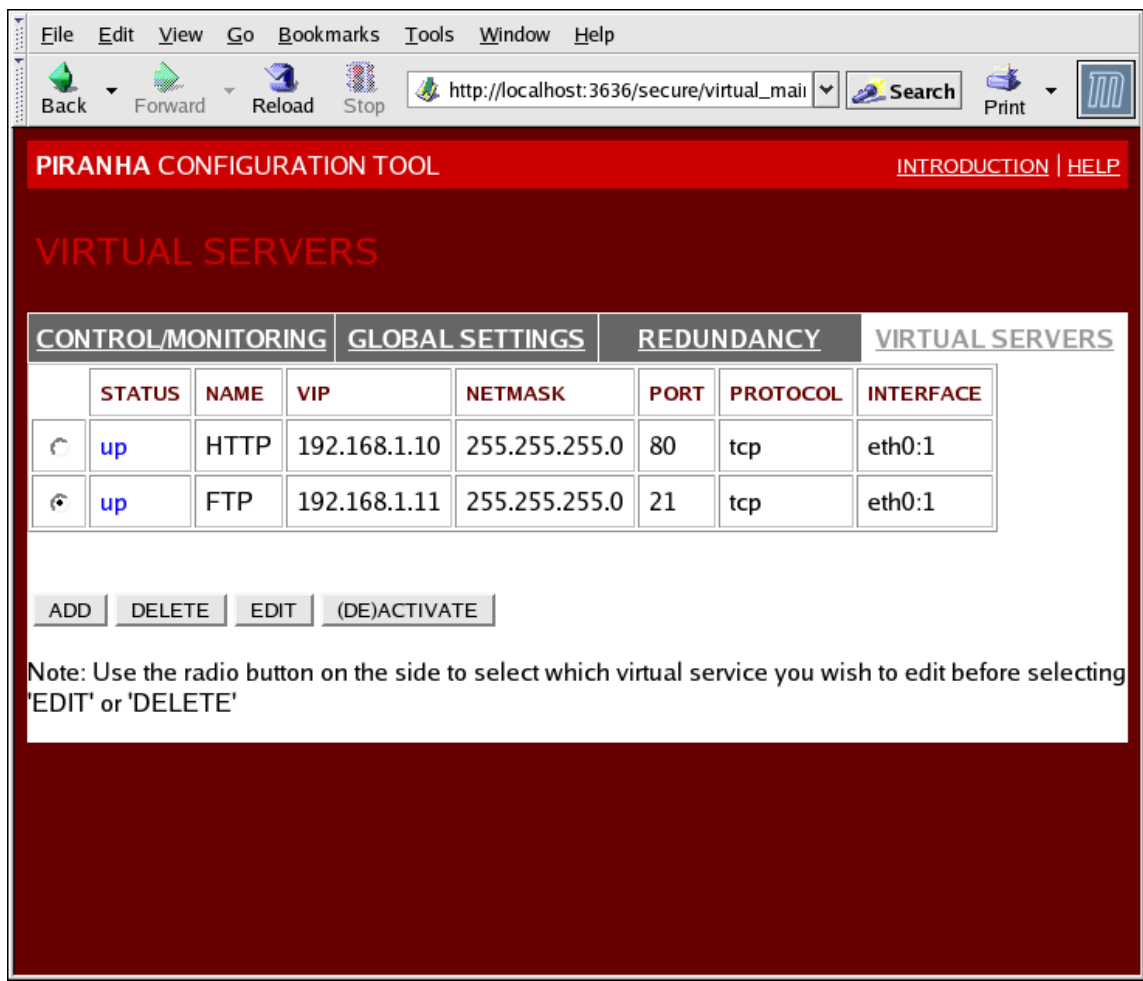


Figura 1.33. The VIRTUAL SERVERS Panel

Cada servidor exibido no painel **SERVIDORES VIRTUAIS** pode ser configurado nas telas subseqüentes ou *subseqções*.

Para adicionar um serviço, clique no botão **ADICIONAR**. Para remover um serviço, faça a seleção clicando no botão rádio, perto do servidor virtual, e clique no botão **DELETAR**.

Para ativar ou desativar um servidor virtual na tabela, clique no botão rádio e clique no botão **(DES)ATIVAR**.

Após adicionar um servidor virtual, você pode configurar isto clicando no botão rádio para o lado esquerdo e clicando no botão **EDITAR** para exibir a subseção **SERVIDOR VIRTUAL**.

1.10.4.1. A Subseção do SERVIDOR VIRTUAL

The **VIRTUAL SERVER** subsection panel shown in [Figura 1.34, “The VIRTUAL SERVERS Subsection”](#) allows you to configure an individual virtual server. Links to subsections related specifically to this virtual server are located along the top of the page. But before configuring any of the subsections related to this virtual server, complete this page and click on the **ACCEPT** button.

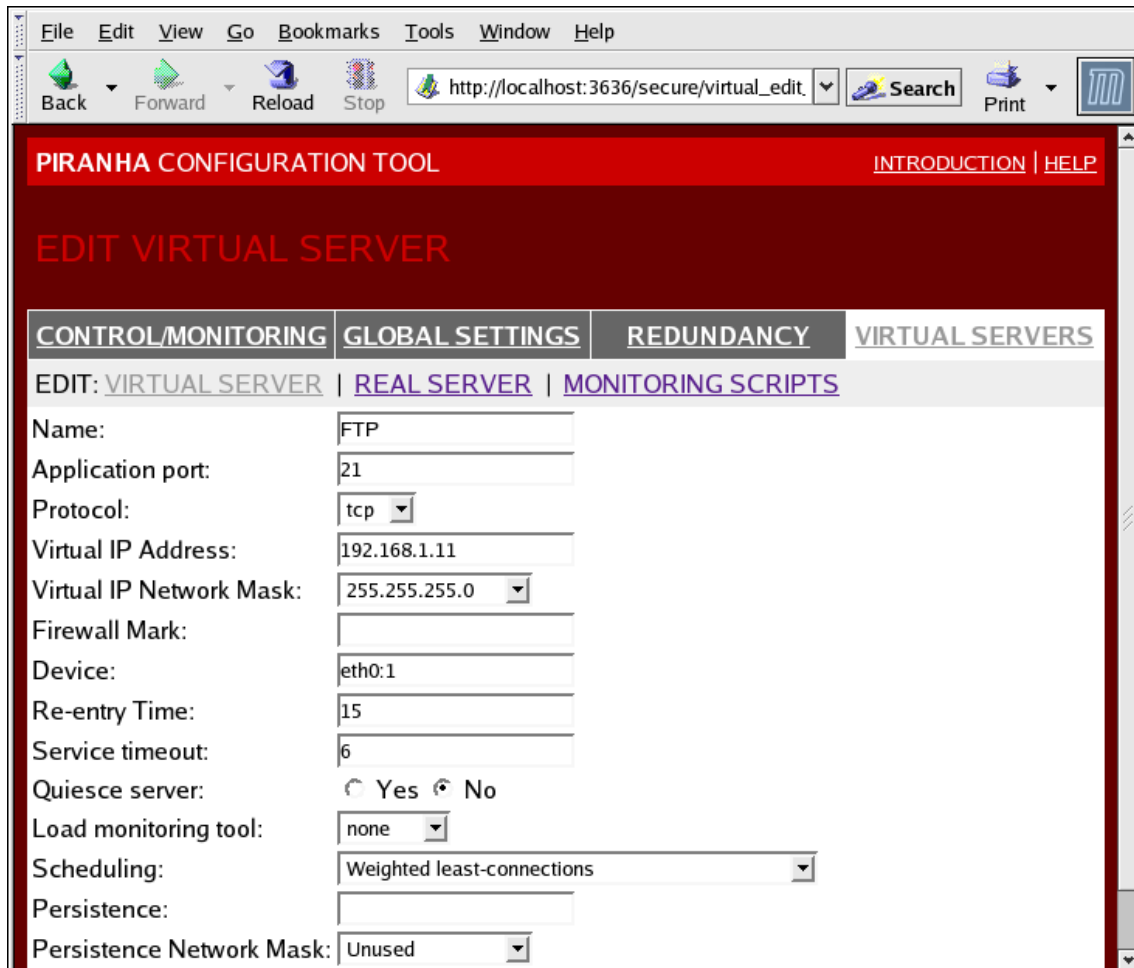


Figura 1.34. The VIRTUAL SERVERS Subsection

Name

O nome descritivo para identificar o servidor virtual. Este nome *não* é o nome do anfitrião para a máquina, sendo então descritivo e facilmente identificável. Você pode também referenciar o protocolo usado pelo servidor virtual, como por exemplo o HTTP.

Application port

O número do portal através do qual o aplicativo do serviço irá escutá-lo.

Protocol

Fornece uma escolha de UDP ou TCP, num menu suspenso.

Virtual IP Address

The virtual server's floating IP address.

Virtual IP Network Mask

A netmask para este servidor virtual, num menu suspenso.

Firewall Mark

A entrada do valor integral da marca do firewall, quando construindo os protocolos de portais múltiplos ou criando os servidores virtuais de portal múltiplo de separação, no entanto os protocolos relatados.

Device

O nome do dispositivo de rede, do qual você queira o endereço IP flutuante definido no campo do **Endereço IP Virtual**, para a vinculação.

Você pode alias o endereço IP flutuante a uma interface Ethernet conectada à rede pública.

Re-entry Time

Um valor integral que define o número de segundos antes do roteador LVS ativo tentar usar o servidor real, após a falha do servidor real.

Service Timeout

Um valor integral que define o número de segundos antes de um servidor real ser considerado inativo e não disponível.

Quiesce server

Quando o botão de rádio do **servidor Quiesce** for selecionado, um novo nó do servidor real aparecerá on-line a qualquer instante. A tabela mínima de conexão é regulada para zero, assim o roteador LVS ativo encaminha as solicitações como se todos os servidores reais fossem adicionados recentemente ao cluster. Esta opção evita um novo servidor começar a atolar-se com o alto número de conexões pela entrada do cluster.

Load monitoring tool

O roteador LVS pode monitorar a carga em vários servidores reais usando tanto o **rup** ou **ruptime**. Se você selecionar o **rup** do menu suspenso, cada servidor real deverá executar o serviço **rstatd**. Se você selecionar o **ruptime**, cada servidor real deverá executar o serviço **rwhod**.

Scheduling

O algoritmo de agendamento preferido do menu suspenso. O padrão é de **Weighted least-connection**.

Persistência

Utilizado caso você precise de conexões persistentes para o servidor virtual durante as transações dos clientes. Especifica o número de segundos da inatividade permitida para o lapso, antes do tempo limite neste campo do texto.

Persistence Network Mask

Para limitar a persistência de um subnet particular, selecione a máscara da rede apropriada para o menu suspenso.

1.10.4.2. Subseção do SERVIDOR REAL

Clicando no link de subseção no topo do painel, exibirá a subseção **EDITAR O SERVIDOR REAL**. Isto exibe o status dos anfitriões de servidor físico para um serviço virtual particular.

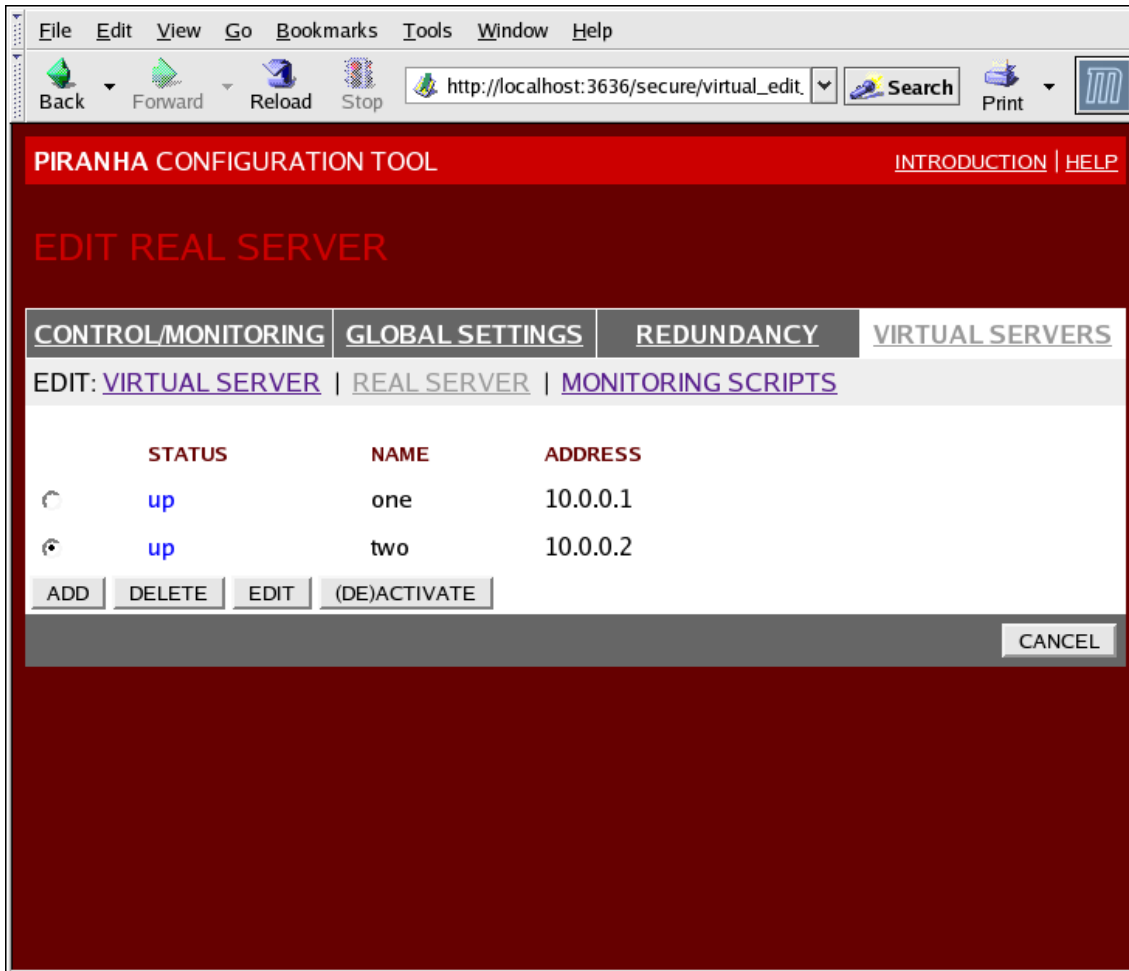


Figura 1.35. The REAL SERVER Subsection

Click the **ADD** button to add a new server. To delete an existing server, select the radio button beside it and click the **DELETE** button. Click the **EDIT** button to load the **EDIT REAL SERVER** panel, as seen in [Figura 1.36, "The REAL SERVER Configuration Panel"](#).

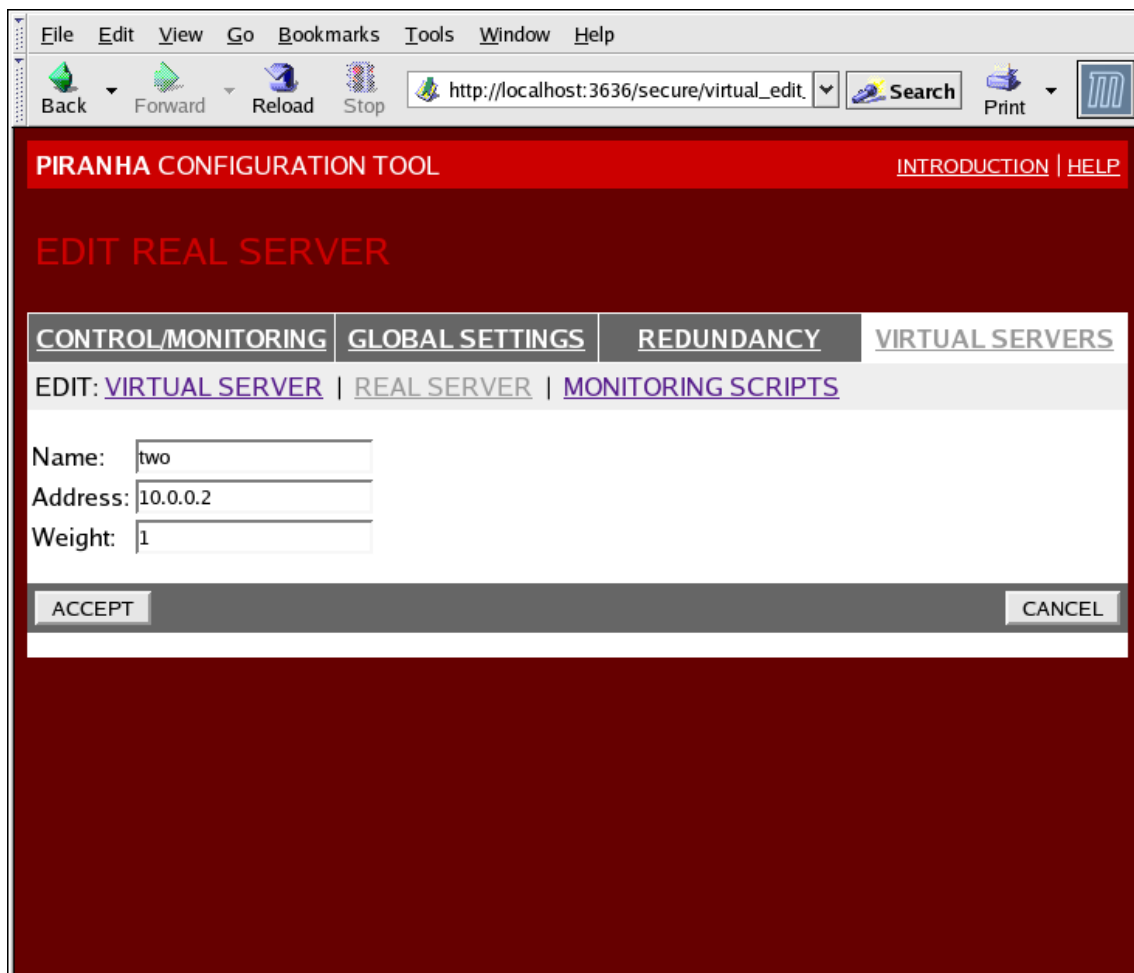
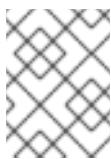


Figura 1.36. The REAL SERVER Configuration Panel

Este painel consiste em três campos de entrada:

Name

O nome descritivo para o servidor real.



NOTA

Este nome *não* é o nome anfitrião para a máquina, então faça com que ele seja descritivo e facilmente identificado.

Address

The real server's IP address. Since the listening port is already specified for the associated virtual server, do not add a port number.

Weight

An integer value indicating this host's capacity relative to that of other hosts in the pool. The value can be arbitrary, but treat it as a ratio in relation to other real servers.

1.10.4.3. EDIT MONITORING SCRIPTS Subsection

Clique no link **SCRIPTS DE MONITORAMENTO** no topo da página. A subseção **EDITAR OS SCRIPTS DE MONITORAMENTO** permite o administrador especificar a seqüência em série enviada/esperada,

para verificação de que o serviço para o servidor virtual é funcional em cada servidor real. Este é também o lugar onde o administrador pode especificar os scripts personalizados para checar os serviços requeridos dinamicamente dos dados de mudança.

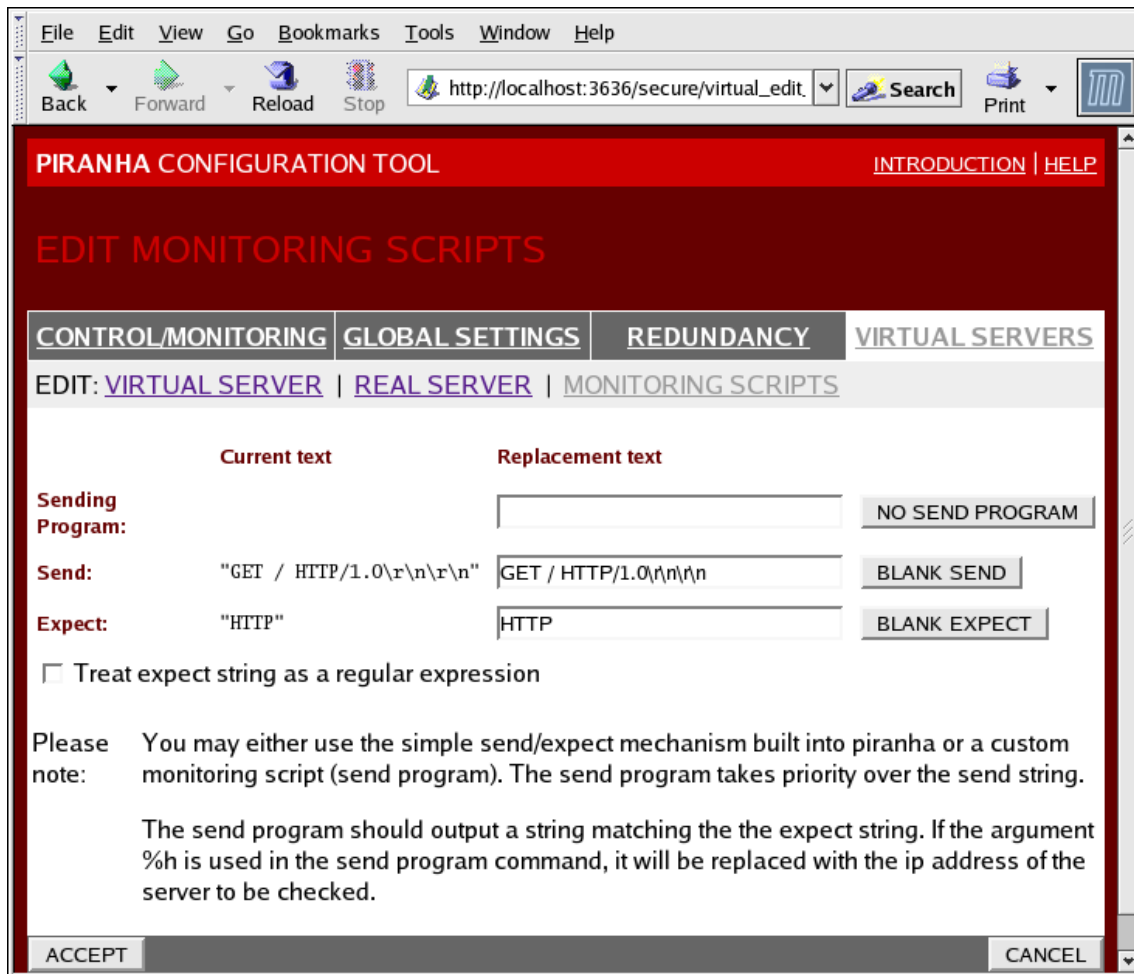


Figura 1.37. The EDIT MONITORING SCRIPTS Subsection

Sending Program

Para uma verificação de serviço avançado, você pode usar este campo para especificar o caminho ao script de serviço de checagem. Esta função é de grande auxílio aos serviços que requerem dinamicamente os dados de mudança, como por exemplo os HTTPS or SSL.

Para uso desta função, você deve escrever um script que retorna a uma resposta de texto, preparar isto para ser executado e digitar o caminho dele no campo **Programa de Envio**.



NOTA

Caso um programa externo seja inserido no campo **Programa de Envio**, o campo **Enviar** será ignorado.

Send

Uma seqüência do daemon **nanny** a ser enviada a cada servidor real deste campo. Por padrão, o campo enviado é completado pelo HTTP. Você pode alterar este valor dependendo de suas necessidades. Caso, você deixe este campo em branco, o daemon **nanny** tentará abrir o portal e assumirá o serviço executado, caso isto se suceda.

Apenas uma seqüência enviada é permitida neste campo, e pode apenas conter imprimíveis caracteres ASCII como também os seguintes caracteres de fuga:

- \n para uma nova linha
- \r para o retorno do carro.
- \t para tab.
- \ para escapar para o próximo capítulo.

Expect

A resposta de texto ao servidor deve retornar, se isto estiver funcionando propriamente. Caso você escreva seu próprio programa de envio, entre a resposta que você elaborou para envio no caso desta ser bem sucedida.

[1] Um servidor virtual é um serviço configurado para escutar um específico IP virtual

CAPÍTULO 2. SUMÁRIO DO COMPONENTE DO RED HAT CLUSTER SUITE

Este capítulo fornece um resumo dos componentes do Red Hat Cluster Suite e consistência das sessões seguintes:

- [Seção 2.1, “Componentes de Cluster”](#)
- [Seção 2.2, “Páginas Man ”](#)
- [Seção 2.3, “Hardware \(disco rígido \) compatível ”](#)

2.1. COMPONENTES DE CLUSTER

Tabela 2.1, “Red Hat Cluster Suite Software Subsystem Components (Componentes de Subistema de Software)” summarizes Red Hat Cluster Suite components.

Tabela 2.1. Red Hat Cluster Suite Software Subsystem Components (Componentes de Subistema de Software)

Função	Componentes	Descrição
Conga	luci	Sistema de Gerenciamento Remoto - Estação de Gerenciamento
	ricci	Sistema de Gerenciamento Remoto - Estação de Gerenciamento
Cluster Configuration Tool	system-config-cluster	Comando usado para gerenciar uma configuração de cluster numa configuração gráfica.
Cluster Logical Volume Manager (CLVM, Gerenciador de Volume Lógico de Cluster)	clvmd	O daemon distribui as atualizações do LVM de metadados em volta de um cluster. Isto deve ser executado em todos os nós de um cluster e pode gerar erro, caso um nó de um cluster não possua a execução daemon.
	lvm	Ferramentas LVM2. Fornece as ferramentas de linha de comando do LVM2.
	system-config-lvm	Fornece a interface gráfica do usuário para o LVM2.
	lvm.conf	O arquivo da configuração do LVM. O caminho direto é o seguinte: /etc/lvm/lvm.conf .

Função	Componentes	Descrição
Cluster Configuration System (CCS, Sistema de Configuração de Cluster).	ccs_tool	ccs_tool faz parte do Cluster Configuration System (CCS). Isto é utilizado para fazer atualizações on-line de arquivos de configuração CCS. Além disso, isto pode ser usado para atualizar os arquivos de configuração de cluster de arquivos CCS criados com GFS 6.0 (e anteriormente) à configuração de formato XML, formato usado com esta liberação de Red Hat Cluster Suite.
	ccs_test	Diagnóstico e comando testados dos quais são utilizados para resgatar informação dos arquivos de configuração através do ccsd .
	ccsd	O CCS daemon que atua em todos os nós de cluster e fornece dados do arquivo de configuração a um software de cluster.
	cluster.conf	Este é um arquivo de configuração de cluster. O caminho direto é /etc/cluster/cluster.conf .
Cluster Manager (CMAN) - Gerenciador de Cluster (CMAN)	cman.ko	O módulo Kernel para o CMAN.
	cman_tool	Este é o fim da frente administrativa para o CMAN. Inicia e interrompe o CMAN, podendo mudar alguns parâmetros internos como os votos.
	dlm_controld	Daemon iniciado pelo cman script de iniciação para gerenciar o dlm em kernel, não é utilizado pelo usuário.
	gfs_controld	Daemon iniciado pelo cman script de iniciação para gerenciar o gfs em kernel, não é usado pelo usuário.
	group_tool	Usado para gerar uma lista de grupos relacionada ao fercing, DLM, GFS, e adquirindo informação do depurador. Inclui os cman_tool services fornecido em RHEL 4.

Função	Componentes	Descrição
	groupd	Daemon iniciado por cman script de iniciação para a interface entre openais/cman e dlm_controlld/gfs_controlld/fenced ; não é usado pelo usuário.
	libcman.so.<version number>	Biblioteca para programas que precisam interagir com o cman.ko .
Resource Group Manager (rgmanager) - Gerenciador de Grupo de Recurso - (rgmanager)	clusvcadm	Comando usado para manualmente capacitar, desativar, relocar e restaurar os serviços do usuário num cluster.
	clustat	Comando usado para exibir o status de um cluster, incluindo o nó associado e serviços de execução.
	clurgmgrd	Daemon usado para manusear as solicitações de serviço do usuário incluindo o serviço iniciado, serviço desativado, serviço realocado e serviço restaurado.
	clurmtabd	Daemon usado para manusear as tabelas de elevação NFS Clustered.
Fence	fence_apc	Agente fence para a tomada elétrica APC.
	fence_bladecenter	Agente fence para o IBM Bladecenters com interface Telnet.
	fence_bullpap	Agente fence para a Interface Bull NovaScale Plataforma Administration Processor (PAP).
	fence_drac	Agente fencing para o Cartão de Acesso Remoto Dell.
	fence_ipmilan	Agente fence para máquinas controladas pela IPMI - Intelligent Platform Management Interface (Interface de Gerenciamento da Plataforma Inteligente) sobre LAN.
	fence_wti	Agente fence para a tomada eletrônica WTI.

Função	Componentes	Descrição
	fence_brocade	Agente fence para a tomada de Canal de Fibra Brocada.
	fence_mcddata	Agente fence para a tomada de Canal de Fibra McData.
	fence_vixel	Agente fence para a tomada de Canal de Fibra Vixel.
	fence_sanbox2	Agente fence para a tomada de Canal de Fibra SANBox2.
	fence_ilo	Agente fence para as interfaces HP ILO (anteriormente fence_rib).
	fence_rsa	Agente fence I/O para o IBM RSA II.
	fence_gnbd	Agente fence usado com o armazenamento GNBD.
	fence_scsi	Agente fencing I/O para as reservas de persistência SCSI.
	fence_egera	Agente fence usado para o sistema Egenera BladeFrame.
	fence_manual	Agente fence de interação manual. <i>NOTA</i> Este componente não é suportado para ambientes de produção.
	fence_ack_manual	Interface do usuário para o agente fence_manual .
	fence_node	Um programa que apresenta o fence I/O num único nó.
	fence_xvm	Agente fence I/O para as máquinas virtuais Xen
	fence_xvmd	Agente fence I/O hospedeiro para as máquinas virtuais Xen.
	fence_tool	Um programa para unir-se e sair do domínio fence.
	fenced	O fence I/O daemon.

Função	Componentes	Descrição
DLM	libdlm.so.<version number>	Biblioteca para o suporte de Distributed Lock Manager (DLM, Gerenciador de Bloqueio Distribuído).
GFS	gfs.ko	Módulo Kernel que implementa o sistema de arquivo GFS, e é carregado nos nós de cluster GFS.
	gfs_fsck	Comando que repara um sistema de arquivo GFS desmontado
	gfs_grow	Comanda o que cresce no sistema de arquivo GFS montado.
	gfs_jadd	Comando que adiciona diários a um sistema de arquivo GFS montado.
	gfs_mkfs	Comando que cria um sistema de arquivo GFS num dispositivo armazenado.
	gfs_quota	Comando que gerencia as cotas num sistema de arquivo GFS montado.
	gfs_tool	Comando que configura ou forma um sistema de arquivo GFS. Este comando pode também concentrar uma variedade de informações sobre o sistema de arquivo.
	mount.gfs	Auxílio de montagem chamado por mount (8) ; não é usado pelo usuário.
GNBD	gnbd.ko	Módulo Kernel que implementa o driver de dispositivo GNBD nos clientes.
	gnbd_export	Comando para criação, exportação e gerenciamento GNBDs num servidor GNBD.
	gnbd_import	Comando para importar e gerenciar os GNBDs num cliente GNBD.
	gnbd_serv	Um servidor daemon que permite um nó exportar o armazenamento local na rede.

Função	Componentes	Descrição
LVS	pulse	<p>This is the controlling process which starts all other daemons related to LVS routers. At boot time, the daemon is started by the <code>/etc/rc.d/init.d/pulse</code> script. It then reads the configuration file <code>/etc/sysconfig/ha/lvs.cf</code>. On the active LVS router, pulse starts the LVS daemon. On the backup router, pulse determines the health of the active router by executing a simple heartbeat at a user-configurable interval. If the active LVS router fails to respond after a user-configurable interval, it initiates failover. During failover, pulse on the backup LVS router instructs the pulse daemon on the active LVS router to shut down all LVS services, starts the <code>send_arp</code> program to reassign the floating IP addresses to the backup LVS router's MAC address, and starts the <code>lvs</code> daemon.</p>
	lvsd	<p>O daemon <code>lvs</code> executa o roteador LVS ativo, uma vez chamado pelo pulse. Ele lê o arquivo de configuração <code>/etc/sysconfig/ha/lvs.cf</code>, chamando a utilidade <code>ipvsadm</code> para construir e manter a tabela de roteamento IPVS, e atribuir um processo <code>nanny</code> para cada serviço LVS configurado. Se o <code>nanny</code> reportar que o servidor real está fora de operação, o <code>lvs</code> instrui a utilidade <code>ipvsadm</code> para remover o servidor real de uma tabela de roteamento IPVS.</p>
	ipvsadm	<p>Este serviço atualiza a tabela de roteamento IPVS no kernel. O daemon <code>lvs</code> configura e administra o LVS pela chamada <code>ipvsadm</code> para adicionar, mudar, ou deletar as entradas numa tabela de roteamento IPVS.</p>
	nanny	<p>O monitoramento daemon <code>nanny</code> executa o roteador LVS ativo. Através deste daemon, o roteador LVS ativo determina o health de cada servidor e, opcionalmente, monitora esta carga de trabalho. Um processo separado atua em cada servidor definido para cada servidor real.</p>

Função	Componentes	Descrição
	lvs.cf	Este é o arquivo de configuração LVS. O caminho completo para este arquivo é o seguinte: /etc/sysconfig/ha/lvs.cf . Todos os daemons adquirem a informação de configuração por este arquivo diretamente ou indiretamente.
	Piranha Configuration Tool	Esta é a ferramenta baseada na web para monitoramento, configuração e administração do LVS. Esta é a ferramenta padrão para manter o arquivo /etc/sysconfig/ha/lvs.cf de configuração LVS.
	send_arp	Este programa envia difusões seletivas quando o endereço IP flutuante muda de um nó para outro durante a falha.
Disco Quórum	qdisk	Um disco baseado em quórum daemon para o CMAN / Linux-Cluster.
	mkqdisk	Utilidade de Disco Quórum de Cluster
	qdiskd	Disco Quórum de Cluster Daemon

2.2. PÁGINAS MAN

Esta seção lista as páginas man (homem) que são relevantes ao Red Hat Cluster Suite, como um recurso adicional.

- Infraestrutura de Cluster
 - **ccs_tool (8)** - A ferramenta usada para realizar atualizações on-line de arquivos de configuração CCS;
 - **ccs_test (8)** - A ferramenta de diagnóstico para a execução de Sistema de Configuração de Cluster;
 - **ccsd (8)** - O daemon usado para acessar os arquivos de configuração de cluster;
 - **ccs (7)** - Sistema de Configuração Cluster;
 - **cman_tool (8)** - Ferramenta de Gerenciamento de Cluster;
 - **cluster.conf [cluster] (5)** - O arquivo de configuração para produtos de cluster;
 - **qdisk (5)** - um daemon baseado no disco quórum para o CMAN / Linux-Cluster;
 - **mkqdisk (8)** - Utilidade de Disco Quórum de Cluster;

- `qdiskd (8)` - Disco Quórum de Cluster Daemon;
 - `fence_ack_manual (8)` - programa executado por um operador como parte do manual de fence I/O;
 - `fence_apc (8)` - Agente fence para o APC MasterSwitch;
 - `fence_bladecenter (8)` - Agente fence I/O para o IBM Bladecenter;
 - `fence_brocade (8)` - Agente fence I/O para as tomadas Brocade FC;
 - `fence_bullpap (8)` - Agente fence I/O para a arquitetura controlado por um console gerenciado PAP;
 - `fence_drac (8)` - Agente fence para o Cartão de Acesso Remoto Dell;
 - `fence_egenera (8)` - Agente fence I/O para a Egenera BladeFrame;
 - `fence_gnbd (8)` - Agente fence I/O para o GNBD baseado em clusters GFS;
 - `fence_ilo (8)` - Agente fence I/O para o HP Integrated Lights Out card;
 - `fence_ipmilan (8)` - Agente fence I/O para as máquinas controladas por IPMI sobre LAN;
 - `fence_manual (8)` - programa executado pelo fence como parte do manual de fence I/O;
 - `fence_mcddata (8)` - Agente fence I/O para as tomadas McData FC;
 - `fence_node (8)` - Um programa que atua no fence I/O para um único nó;
 - `fence_rib (8)` - Agente fence I/O para o Compaq Remote Insight Lights Out card;
 - `fence_rsa (8)` - Agente fence I/O para o IBM RSA II;
 - `fence_sanbox2 (8)` - Agente fence I/O para as tomadas QLogic SANBox2 FC;
 - `fence_scsi (8)` - Agente fence I/O para as reservas de persistência SCSI;
 - `fence_tool (8)` - Um programa para unir e sair do domínio fence;
 - `fence_vixel (8)` - Agente fence I/O para as tomadas Vixel FC;
 - `fence_wti (8)` - Agente fence I/O para o WTI Network Power Switch;
 - `fence_xvm (8)` - Agente fence I/O para as máquinas virtuais Xen;
 - `fence_xvmd (8)` - Agente fence I/O anfitrião para as máquinas virtuais Xen;
 - `fenced (8)` - o fence I/O daemon.
- Gerenciador de Serviço de Alta Disponibilidade
 - `clusvcadm (8)` - Utilidade de Administração do Serviço de Usuário de Cluster;
 - `clustat (8)` - Utilidade de Status de Cluster;
 - `Clurgmgrd [clurgmgrd] (8)` - Grupo de Recurso Gerenciador (Serviço de Cluster) Daemon;

- clurmtabd (8) - Tabela de Montagem Remota NFS de Cluster Daemon.
- GFS
 - gfs_fsck (8) - Sistema de checagem do arquivo GFS off-line;
 - gfs_grow (8) - Expande o sistema de arquivo GFS;
 - gfs_jadd (8) - Adiciona relatórios ao sistema de arquivo GFS;
 - gfs_mount (8) - opções de montagem GFS;
 - gfs_quota (8) - Manipula as cotas de disco GFS;
 - gfs_tool (8) - interface para as chamadas gfs ioctl;
- Gerenciador de Volume Lógico de Cluster
 - clvmd (8) - o LVM de cluster daemon;
 - lvm (8) - ferramentas LVM2;
 - lvm.conf [lvm] (5) - Arquivo de configuração para o LVM2;
 - lvmchange (8) - muda as características do gerenciador de volume lógico;
 - pvcreate (8) - inicia um disco ou partição para ser usado pelo LVM;
 - lvs (8) - relata informação sobre os volumes lógicos.
- Dispositivo de Bloqueio da Rede Global
 - gnbd_export (8) - a interface para exportar os GNBDs;
 - gnbd_import (8) - manipula os dispositivos de bloqueio de um cliente;
 - gnbd_serv (8) - servidor gnbd daemon.
- LVS
 - pulse (8) - heartbeat daemon para monitoração do health de nós de cluster;
 - lvs.cf [lvs] (5) - arquivo de configuração para o lvs;
 - lvscan (8) - rastreia (todos os discos) de volumes lógicos;
 - lvsd (8) - daemon para o controle dos serviços de clustering Red Hat;
 - ipvsadm (8) - administração do Servidor Virtual Linux;
 - ipvsadm-restore (8) - restaura a tabela IPVS do stdin;
 - ipvsadm-save (8) - salva a tabela IPVS para stdout;
 - nanny (8) - ferramenta para monitorar o status do serviço num cluster;
 - send_arp (8) - ferramenta para notificar a rede sobre um novo endereço IP / endereço de mapeamento MAC;

2.3. HARDWARE (DISCO RÍGIDO) COMPATÍVEL

Para maiores informações sobre o hardware que é compatível com os componentes do Red Hat Cluster Suite (por exemplo, disparadores de fence de suporte, disparadores de armazenamento e Canal de Fibra de troca), refira-se ao guia de configuração de hardware em http://www.redhat.com/cluster_suite/hardware/.

APÊNDICE A. HISTÓRICO DE REVISÃO

Revisão 3-7.400 Rebuild with publican 4.0.0	2013-10-31	Rüdiger Landmann
Revisão 3-7 Rebuild for Publican 3.0	2012-07-18	Anthony Towns
Revisão 1.0-0 Consolidação dos lançamentos dos pontos	Tue Jan 20 2008	Paul Kennedy

ÍNDICE REMISSIVO

C

cluster

displaying status, [Cluster Status Tool](#)

cluster administration

displaying cluster and service status, [Cluster Status Tool](#)

cluster component compatible hardware, [Hardware \(disco rígido \) compatível](#)

cluster component man pages, [Páginas Man](#)

cluster components table, [Componentes de Cluster](#)

Cluster Configuration Tool

accessing, [Cluster Configuration Tool](#)

cluster service

displaying status, [Cluster Status Tool](#)

command line tools table, [Ferramentas de Administração de Linha de Comando](#)

compatible hardware

cluster components, [Hardware \(disco rígido \) compatível](#)

Conga

overview, [Conga](#)

Conga overview, [Conga](#)

F

feedback, [Feedback](#)

I

introduction, [Introdução](#)

other Red Hat Enterprise Linux documents, [Introdução](#)

L

LVS

direct routing

requirements, hardware, [Roteamento direto](#)

requirements, network, [Roteamento direto](#)

requirements, software, [Roteamento direto](#)

routing methods

NAT, [Métodos Encaminhados](#)

three tiered

high-availability cluster, [Three-Tier LVS Topology](#)

M

man pages

cluster components, [Páginas Man](#)

N

NAT

routing methods, LVS, [Métodos Encaminhados](#)

network address translation (ver NAT)

O

overview

economy, [Red Hat GFS](#)

performance, [Red Hat GFS](#)

scalability, [Red Hat GFS](#)

P

Piranha Configuration Tool

CONTROL/MONITORING, [CONTROL/MONITORING](#)

EDIT MONITORING SCRIPTS Subsection, [EDIT MONITORING SCRIPTS Subsection](#)

GLOBAL SETTINGS, [GLOBAL SETTINGS](#)

login panel, [GUI Administrador de Serviço Virtual Linux](#)

necessary software, [GUI Administrador de Serviço Virtual Linux](#)

REAL SERVER subsection, [Subseção do SERVIDOR REAL](#)

REDUNDANCY, [REDUNDANCY](#)

VIRTUAL SERVER subsection, [VIRTUAL SERVERS](#)

Firewall Mark , [A Subseção do SERVIDOR VIRTUAL](#)

Persistence , [A Subseção do SERVIDOR VIRTUAL](#)

Scheduling , [A Subseção do SERVIDOR VIRTUAL](#)

Virtual IP Address , [A Subseção do SERVIDOR VIRTUAL](#)

VIRTUAL SERVERS, [VIRTUAL SERVERS](#)

R

Red Hat Cluster Suite

components, [Componentes de Cluster](#)

T

table

cluster components, [Componentes de Cluster](#)

command line tools, [Ferramentas de Administração de Linha de Comando](#)