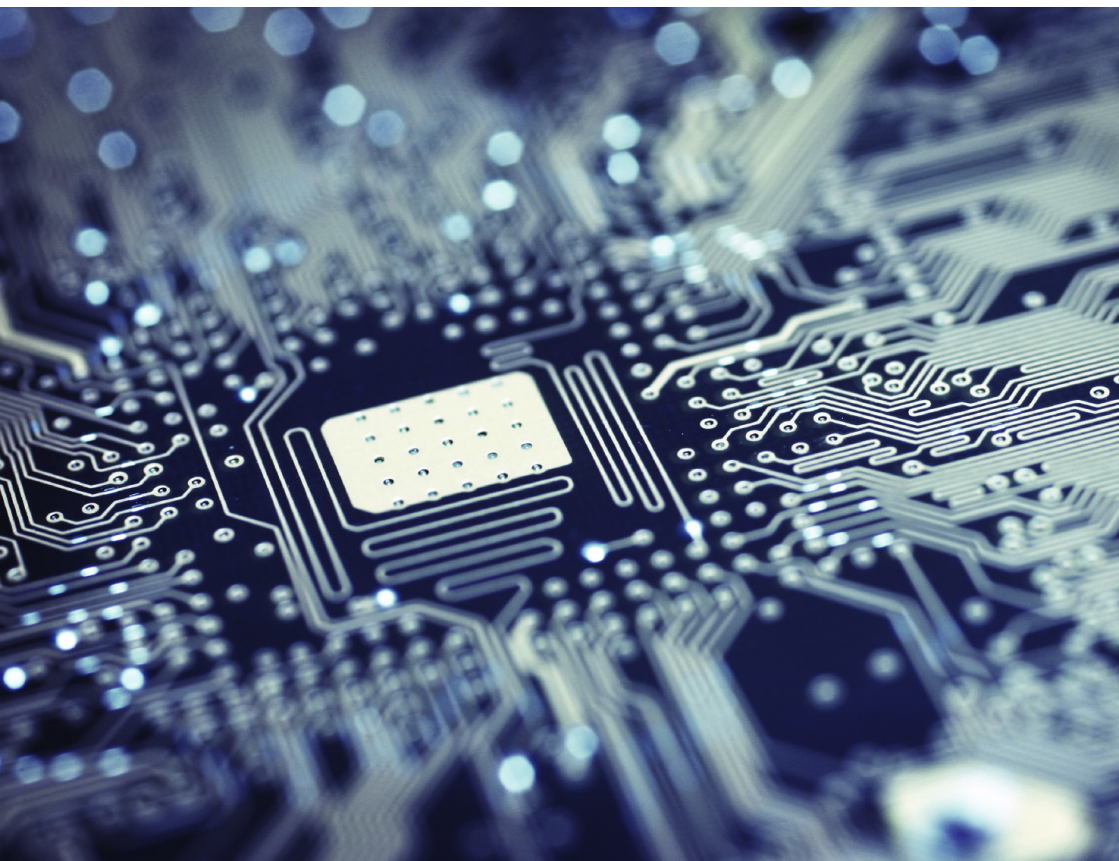


Guia de instalação do Ganglia



*Empresa Brasileira de Pesquisa Agropecuária
Embrapa Informática Agropecuária
Ministério da Agricultura, Pecuária e Abastecimento*

Documentos 125

Guia de instalação do Ganglia

Fábio Danilo Vieira

Embrapa Informática Agropecuária

Av. André Tosello, 209 - Barão Geraldo
Caixa Postal 6041 - 13083-886 - Campinas, SP
Fone: (19) 3211-5700 - Fax: (19) 3211-5754
www.cnptia.embrapa.br
sac@cnptia.embrapa.br

Comitê de Publicações

Presidente: *Silvia Maria Fonseca Silveira Massruhá*

Membros: *Adhemar Zerlotini Neto, Stanley Robson de Medeiros Oliveira, Thiago Teixeira Santos, Maria Goretti Gurgel Praxedes, Adriana Farah Gonzalez, Neide Makiko Furukawa, Carla Cristiane Osawa*

Membros suplentes: *Felipe Rodrigues da Silva, José Ruy Porto de Carvalho, Eduardo Delgado Assad, Fábio César da Silva*

Supervisor editorial: *Stanley Robson de Medeiros Oliveira, Neide Makiko Furukawa*

Revisor de texto: *Adriana Farah Gonzalez*

Normalização bibliográfica: *Maria Goretti Gurgel Praxedes*

Editoração eletrônica/Capa: *Diana Breitenbach*

Imagem da capa: <http://www.hcbe.net/tsc/technology-services-center.aspx>

Secretária: *Carla Cristiane Osawa*

1ª edição on-line 2012

Todos os direitos reservados.

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei no 9.610).

Dados Internacionais de Catalogação na Publicação (CIP) Embrapa Informática Agropecuária

Vieira, Fábio Danilo.

Guia de instalação do Ganglia / Fábio Danilo Vieira. - Campinas :

Embrapa Informática Agropecuária, 2012.

21 p. : il. ; (Documentos / Embrapa Informática Agropecuária,

ISSN 1677-9274; 125)

1. *Cluster*. 2. Computação em nuvem. 3. Sistemas distribuídos.I.
Embrapa Informática agropecuária. II. Título. III. Série.

CDD 005.15 (21.ed.)

Autor

Fábio Danilo Vieira

Bacharel em Tecnologia da Informação
Analista da Embrapa Informática Agropecuária
Av. André Tosello, 209, Barão Geraldo
Caixa Postal 6041 - 13083-886 - Campinas, SP
Fone: (19) 3211-5798
e-mail: fabio.vieira@embrapa.br

Apresentação

Nos últimos anos, as redes (sejam locais ou a própria internet) deixaram de ser apenas uma forma de se conectar vários computadores e fornecer transporte de dados entre eles. Atualmente, diversos serviços são oferecidos por meio de redes de computadores, como, por exemplo, computação em nuvem, processamentos de grandes volumes de dados em *clusters* e *grids* etc.

Com o oferecimento desses diversos serviços, cresceu também a quantidade de dados que trafega pela rede, o número de processos executados nas Central Processing Units (CPUs) e o volume de memória ocupado nos computadores (principalmente servidores). Como consequência, surgiu a necessidade do desenvolvimento de ferramentas que auxiliassem administradores de redes no gerenciamento dos recursos utilizados por tais serviços.

Para auxiliar no gerenciamento efetivo de todos estes recursos computacionais, este guia mostrará como instalar o software Ganglia, que faz o monitoramento de sistemas distribuídos, tais como clusters e grids. Sua principal característica é permitir a visualização remota de dados estatísticos do conjunto de máquinas monitoradas.

Espera-se, assim, que administradores de redes que ofereçam serviços que demandam alto processamento de dados possam se beneficiar dos muitos recursos que o Ganglia oferece.

Kleber Xavier Sampaio de Souza

Chefe-Geral

Embrapa Informática Agropecuária

Sumário

Introdução	9
Ganglia Monitoring Daemon (gmond)	10
Ganglia Meta Daemon (gmetad)	10
Ganglia PHP Web Front-end	11
Instalação	12
Regras de Firewall	12
Configuração de um cluster	13
Multicast	13
Unicast	13
Configuração de múltiplos clusters	15
Conclusão	19
Referências	19
Literatura recomendada	20

Guia de instalação do Ganglia

Fábio Danilo Vieira

Introdução

Ganglia é uma ferramenta para monitoramento de sistemas distribuídos escaláveis, tais como *clusters* e *grids* de computadores. Sua principal característica é permitir a visualização remota de dados estatísticos (*online* ou armazenados num banco de dados próprio) do conjunto de máquinas monitoradas. Esses dados estatísticos incluem utilização de rede, memória e da utilização da (s) CPU (s).

O ganglia é baseado em um *design* hierárquico dirigido a federações de *clusters*. A ferramenta também se aproveita de tecnologias mais utilizadas atualmente, tais como XML para representação de dados, *eXtendedDataRepresentation* (XDR) para o transporte de dados e *Round Robin Database Tool* (RRDTool) para o armazenamento e visualização de dados.

O ganglia foi codificado por meio de algoritmos e estruturas de dados muito bem projetados para não sobrecarregar os nós onde está instalado. A implementação é robusta e portátil, podendo ser instalado num amplo conjunto de sistemas operacionais e em diferentes arquiteturas de processador. Ele tem sido utilizado, por exemplo, para conectar *clusters* de várias universidades, lidando com grupos de até 2000 nós.

O software ganglia é composto por três partes principais: *Ganglia Monitoring Daemon* (gmond), *Ganglia Meta Daemon* (gmetad) e uma interface web codificada em *Personal Home Pag* (PHP), chamada de ganglia-web. A seguir, uma explicação mais detalhada desses componentes.

Ganglia Monitoring Daemon (gmond)

O gmond é um daemon que é executado em cada nó do *cluster* que se deseja monitorar do *cluster*. Sua instalação não necessita que se tenha um sistema de arquivos *Network File System* (NFS) em comum ou mesmo um banco de dados relacional instalado. Suas principais funções são: monitorar as alterações de estado de uma máquina, avisar sobre mudanças relevantes, “ouvir” os estados de outros *daemons* do ganglia instalados em outros nós e responder a pedidos da descrição XML do estado do *cluster*. Em suma:

- Se for um remetente de dados (mute=no), irá coletar alguns indicadores básicos do *host*, tais como carga do sistema *load_one* e utilização de CPU, e enviá-los a outros computadores. Além desses dados, também pode ser configurado pelo usuário para enviar outras informações específicas por meio da adição de módulos C e Python.
- Se for um ouvinte (deaf=no), agregará todos os indicadores enviados de outros computadores e irá mantê-los em memória.

Cada processo gmond pode transmitir as informações de 2 formas distintas:

- Enviando-as no formato XDR usando *User Datagram Protocol* (UDP).
- Enviando-as no formato XML por meio de uma conexão *Transmission Control Protocol* (TCP).

Ganglia Meta Daemon (gmetad)

A hierarquia no Ganglia é obtida por meio de uma árvore de conexões ponto-a-ponto representando os nós do (s) *cluster* (s) e seus estados. O

daemon gmetad instalado em um nó da árvore faz o controle periódico das fontes de dados de seus “filhos”, analisando os XMLs coletados, armazenando os indicadores obtidos numa base de dados *round-robin* e exportando o XML agregado por meio de conexões TCP para os clientes. As fontes de dados podem ser tanto *daemons* gmond, representando computadores específicos, ou outros *daemons* gmetad, representando outros *clusters*. Observe a figura a seguir ilustrando o funcionamento do ganglia e seus *daemons*.

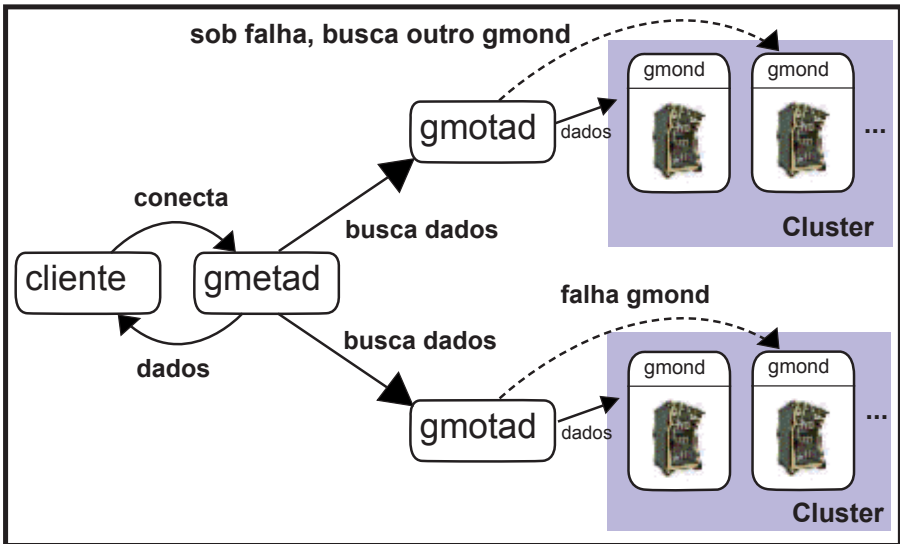


Figura 1. Funcionamento ganglia e seus *daemons*.

Fonte: Adaptado de Massie et al. (2004).

Ganglia PHP Web Front-end

O sistema web (*front-end*) do Ganglia fornece informações obtidas do *daemon* gmetad em tempo real por meio de páginas dinâmicas. O sistema exibe os dados de uma forma eficaz para administradores de sistema e usuários da infraestrutura monitorada, para os quais disponibiliza diversos serviços, como, por exemplo, a possibilidade de se ver a utilização da CPU

da última hora, de um dia específico e até mesmo por intervalos de data. Também gera gráficos de uso de memória, uso de disco, estatísticas de rede, número de processos em execução, e todos os outros indicadores do Ganglia.

Como dito, o sistema web depende da existência do *daemon* gmetad que lhe fornece dados a partir de fontes diversas do Ganglia. Por padrão, o sistema irá abrir a porta local 8651 para receber a árvore XML do Ganglia. Além disso, esse *front-end* é muito dinâmico, ou seja, qualquer alteração de estado numa das máquinas é imediatamente vista no site. Esse comportamento torna o sistema muito eficiente.

O sistema foi escrito na linguagem de script PHP, e usa gráficos gerados pelo *daemon* gmetad para exibir informações de histórico. Ele foi testado em muitas versões do Linux com o servidor web Apache, e demonstrou que funciona muito bem.

Instalação

O melhor caminho para a instalação é usar os pacotes binários. No Ubuntu ou Debian pode se usar o comando `apt-get`, como segue:

```
apt-get install ganglia-monitor gmetad
```

Já no Red-Hat ou Fedora, pode ser usado o comando `yum`, como segue:

```
yum install ganglia-gmond ganglia-gmetad
```

Regras de Firewall

O *daemon* gmond, por padrão, se comunica na porta 8649 por intermédio do protocolo UDP e o *daemon* gmetad faz o *download* dos dados por meio da porta TCP 8649 (como especificado no parâmetro `tcp_accept_channel`). Se existir quaisquer regras que bloqueiam o tráfego nessas portas, os dados obtidos pelo Ganglia não serão mostrados no sistema web.

Configuração de um *cluster*

Para apenas um pequeno conjunto de máquinas, é recomendado a utilização de um único *cluster*, que é algo mais simples de instalar e configurar. A única decisão que se precisa tomar é se será usado transporte *multicast* ou *unicast*. O modo *multicast* é a configuração padrão quando se instala o *daemon gmond*, assim é o mais simples de se configurar. Já os ambientes onde pode haver "instabilidade" é recomendável a criação do Ganglia no modo *unicast*, o que reduz significativamente o tráfego, apesar de ser um pouco mais complexo de ser configurado.

Multicast

Se o Ganglia for utilizar transporte *multicast*, então será preciso pouca customização, pois esta é a configuração padrão nos pacotes de instalação. A única coisa que se precisará fazer é apontar o *daemon gmetad* para uma ou algumas das máquinas que estejam executando *gmond*. Vale ressaltar que não há necessidade de se listar os computadores que estejam em modo "ouvinte" (*deaf=no*). Abaixo, a linha que deve ser alterada no arquivo *gmetad.conf*:

```
# /etc/gmetad.conf on host1
data_source "Cluster_Exemplo" host1 host2 host3
```

Unicast

O modo *unicast* deve ser configurado fazendo-se a designação de uma ou mais máquinas como receptores (ouvintes), na (s) qual (is) deve (m) ter o *daemon gmetad*. Por exemplo, para se designar o *host servidor1* como ouvinte, o arquivo ***gmond.conf*** dos remetentes (menos o da própria máquina ***servidor1***) deve ser configurado assim:

```
globals {
  daemonize = yes
  setuid = yes
  user = nobody
  debug_level = 0
  max_udp_msg_len = 1472
  mute = no
  deaf = yes
```

```

allow_extra_data = yes
host_dmax = 86400
cleanup_threshold = 300 /*secs */
gexec = no
send_metadata_interval = 30 /*secs */
}

cluster {
  name = "LMB"
  owner = "LBA"
  latlong = "unspecified"
  url = "unspecified"
}

host {
  location = "unspecified"
}

udp_send_channel {
  host = servidor1
  port = 8649
  ttl = 1
}

```

Já na própria máquina **servidor1**, o arquivo **gmond.conf** sofre pequenas alterações, e fica assim:

```

globals {
  daemonize = yes
  setuid = yes
  user = nobody
  debug_level = 0
  max_udp_msg_len = 1472
  mute = no
  deaf = no
  allow_extra_data = yes
  host_dmax = 86400
  cleanup_threshold = 300 /*secs */
  gexec = no
  send_metadata_interval = 30 /*secs */
}

cluster {
  name = "LMB"
  owner = "LBA"
  latlong = "unspecified"
  url = "unspecified"
}

host {
  location = "unspecified"
}

udp_send_channel {
  host = servidor1
  port = 8649
  ttl = 1
}

udp_rcv_channel {
  port = 8649
}

tcp_accept_channel {
  port = 8649
}

```

Os indicadores (tráfego de rede, uso de memória, uso de CPU, etc) no Ganglia são enviados separadamente dos metadados. Os metadados contém informações como tipo de indicador, grupo de indicadores, etc., e são transportados em pacotes de dados diferentes dos pacotes contendo os dados sobre os indicadores. Caso se reinicie o *daemon* gmond, o recebimento dos metadados será perdido, podendo resultar em gráficos em branco.

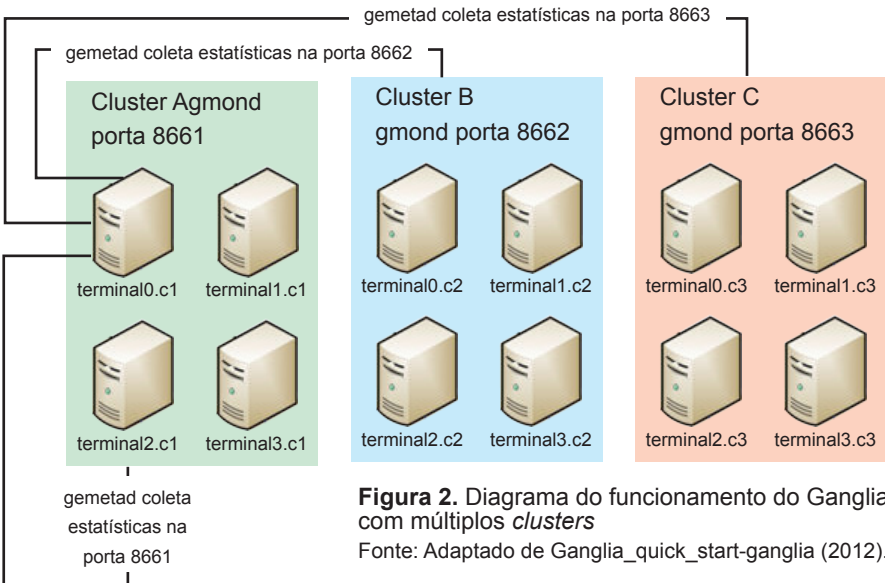
No arquivo gmetad.conf se altera a linha abaixo:

```
# /etc/gmetad.conf on servidor1
data_source "LMB" servidor1
```

Após as configurações, basta apenas se reiniciar os *daemons* gmond e gmetad (quando for o caso):

```
sudo /etc/init.d/ganglia-monitor restart
sudo /etc/init.d/gmetad restart
```

Configuração de múltiplos clusters



Considerando o diagrama, suponha-se que tenha três *clusters* na mesma rede, mas ao invés de ser três *front-ends* web do Ganglia e três *daemons* coletores *gmetad* separados, o computador **terminal0.c1** será configurado como responsável central por recolher as estatísticas de três canais *multicast* (neste caso).

Assim, os componentes do Ganglia necessários em cada computador desta rede serão:

- o *daemon* *gmond* deverá ser instalado em cada nó da rede;
- o *daemon* *gmetad* e o *front-end* web serão necessários apenas no computador **terminal0.c1** (pois será o servidor para coletar as estatísticas dos *clusters*).

Abaixo, são exibidos os trechos mais importantes dos arquivos de configuração. O primeiro deles é o arquivo */etc/gmond.conf* do **ClusterA**, configurado igualmente nos computadores do *cluster* (*terminal0.c1*, *terminal1.c1*, *terminal2.c1*, *terminal3.c1*):

```
# /etc/gmond.conf - ClusterA
cluster {
    name = "ClusterA"
    owner = "LBA"
    latlong = "unspecified"
    url = "unspecified"
}

udp_send_channel {
    mcast_join = 192.168.0.10
    port = 8661
    ttl = 1
}

udp_recv_channel {
    mcast_join = 192.168.0.10
    port = 8661
    bind = 192.168.0.10
}

tcp_accept_channel {
    port = 8661
}
```

O segundo arquivo é o */etc/gmond.conf* do **ClusterB**, também configurado igualmente nos computadores do *cluster* (*terminal0.c2*, *terminal1.c2*,

terminal2.c2, terminal3.c2):

```
# /etc/gmond.conf - ClusterB
cluster {
  name = "ClusterB"
  owner = "LBA"
  latlong = "unspecified"
  url = "unspecified"
}
udp_send_channel {
  mcast_join = 192.168.0.10
  port = 8662
  ttl = 1
}
udp_rcv_channel {
  mcast_join = 192.168.0.10
  port = 8662
  bind = 192.168.0.10
}
tcp_accept_channel {
  port = 8662
}
```

O terceiro arquivo é o **/etc/gmond.conf** do **ClusterC**, configurado igualmente nos computadores do *cluster* (terminal0.c3, terminal1.c3, terminal2.c3, terminal3.c3):

```
# /etc/gmond.conf - ClusterC
cluster {
  name = "ClusterC"
  owner = "LBA"
  latlong = "unspecified"
  url = "unspecified"
}
udp_send_channel {
  mcast_join = 192.168.0.10
  port = 8663
  ttl = 1
}
udp_rcv_channel {
  mcast_join = 192.168.0.10
  port = 8663
  bind = 192.168.0.10
}
tcp_accept_channel {
  port = 8663
}
```

Já o arquivo `/etc/gmetad.conf`, presente apenas no computador `terminal0.c1`, ficará assim (apenas o trecho relevante do arquivo):

```
# /etc/gmetad.conf - terminal0.c1
data_source "ClusterA" terminal0.c1:8661 terminal1.c1:8661
data_source "ClusterB" terminal0.c2:8662 terminal1.c2:8662
data_source "ClusterC" terminal3.c2:8663 terminal1.c3:8663
```

Observe que não foram incluídos todos os computadores como fontes de dados para cada *cluster*, até porque isso seria inviável no caso de existirem milhares de máquinas por *cluster*, por exemplo. Isso ocorre porque os *daemons* `gmond` dos computadores da rede trocam estatísticas entre eles, assim, se um deles “morre”, outro será capaz de fornecer dados para o *daemon* `gmetad`.

Agora tudo que se tem a fazer é iniciar o *daemon* `gmetad` no `terminal0.c1` e os *daemons* `gmond` em todos os *clusters*.

A seguir, uma figura ilustrando, como exemplo, o *front-end* web do Ganglia (*single cluster*) instalado no Laboratório Multusuário de Bioinformática (LMB):

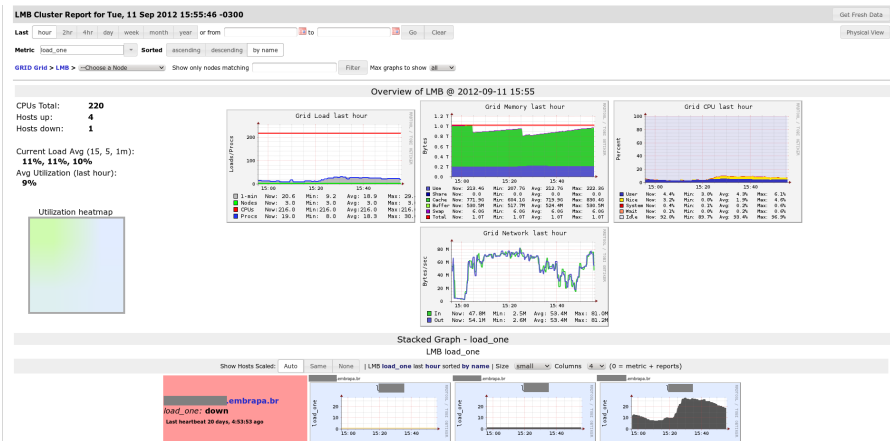


Figura 3. Interface web Ganglia (*single cluster*) do LMB.

Conclusão

Como visto, o ganglia é uma ferramenta que foi desenvolvida para auxiliar administradores de redes e sistema no monitoramento de sistemas distribuídos escaláveis, tais como *clusters* e *grids* de computadores. Sua principal característica é permitir a visualização remota de dados estatísticos (*online* ou armazenados num banco de dados próprio) do conjunto de máquinas monitoradas. Estes dados estatísticos incluem utilização de rede, memória e da utilização da (s) CPU (s). O ganglia está sendo implementado em diversas universidades para conectar os *clusters* nelas instalados.

O software ganglia é composto o por três partes principais: gmond, gmetad e uma interface web codificada em PHP, chamada de ganglia-web. Sua instalação pode ser feita para um ou múltiplos *clusters*. A única decisão que se precisa realizar é em relação ao transporte de dados, se será *multicast* ou *unicast*. O modo *multicast* é a configuração padrão quando se instala o *daemon* gmond, assim é o mais simples de se configurar. Já os ambientes onde pode haver "instabilidade" é recomendável a criação do Ganglia no modo *unicast*, o que reduz significativamente o tráfego.

Diante dos diversos recursos oferecidos pelo ganglia, acredita-se que os administradores de redes que gerenciam serviços que demandam alto processamento de dados possam se beneficiar de forma significativa com os resultados e informações fornecidos por esta ferramenta.

Referências

GANGLIA MONITORING SYSTEM. Disponível em <<http://ganglia.info/>>. Acesso em: 14 de ago. 2012.

GANGLIA_QUICK_START. 2012. Disponível em: <http://sourceforge.net/apps/trac/ganglia/wiki/ganglia_quick_start>. Acesso em: 14 de ago. 2012.

MASSIE, M. L.; CHUN, B. N.; CULLER, D. E. The Ganglia distributed monitoring system: design, Implementation, and experience. **Parallel Computing**, v. 30, n. 7, p. 817-840, 2004.

Literatura recomendada

WIKIPEDIA. 2012. Disponível em: <http://en.wikipedia.org/wiki/Ganglia_%28software%29>. Acesso em: 15 de ago. 2012.



Informática Agropecuária

Ministério da
**Agricultura, Pecuária
e Abastecimento**



CGPE 10232