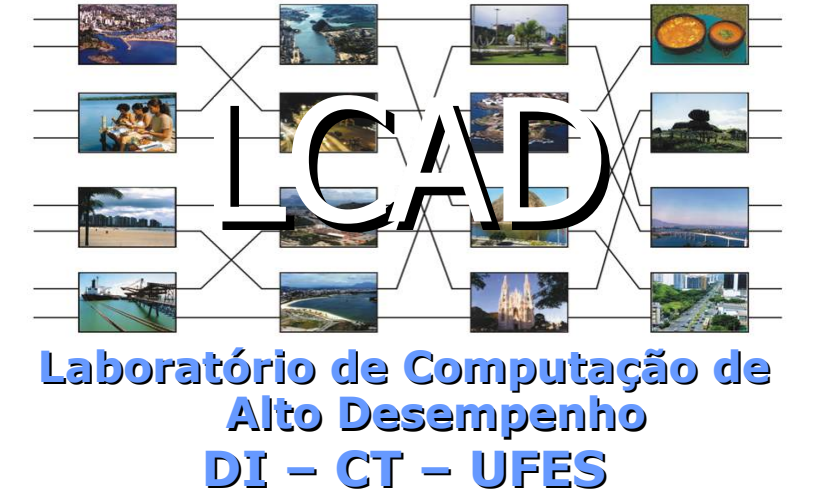


Integração do Cluster Enterprise ao Grid Sinergia



Laboratório de Computação de Alto Desempenho
Departamento de Informática – Centro Tecnológico
Universidade Federal do Espírito Santo
www.lcad.inf.ufes.br



Alberto Ferreira De Souza, Renato Luiz de Freitas Cunha

Cluster Enterprise

O Cluster Enterprise é o cluster computacional do Laboratório de Computação de Alto Desempenho (LCAD) da Universidade Federal do Espírito Santo (UFES) e é composto por 64 nós de processamento Athlon XP 1800+, contando com desempenho teórico máximo de 195,8 GFLOPS, e um nó de administração (master).

Com 256MB de memória e 20GB de armazenamento por nó, o cluster totaliza 16GB de memória RAM e 1.2 TB de armazenamento.

O Cluster Enterprise conta com o *Sun Grid Engine (SGE)* para o escalonamento dos processos executados no cluster, cuja tarefa é aceitar *jobs* submetidos para o cluster, colocá-los em uma fila de espera até que possam ser executados, enviá-los da fila de espera para o dispositivo de execução, gerenciá-los durante a execução e manter um registro da execução quando ele for excluído.

Grid Sinergia

O Grid Sinergia é uma iniciativa multi-institucional para projetar, desenvolver e disponibilizar uma infraestrutura de grid (recursos e *middleware*), explorando a rede óptica de alta performance da Rede Nacional de Pesquisa (RNP), Rede GIGA, para execução eficiente de aplicações científicas paralelas ou distribuídas. Mas especificamente, esse grid é composto pelas seguintes instituições: IC/UFF, CAT/CBPF, LNCC, DI/PUC-Rio, IC/UNICAMP e DI/UFES.

Seu objetivo é prover aos pesquisadores um ambiente realístico e prático para o desenvolvimento de pesquisas de computação distribuída e oferecer a administradores de sistemas experiência prática na administração e operação de recursos computacionais distribuídos em grids.

Organização do *middleware* Globus Toolkit

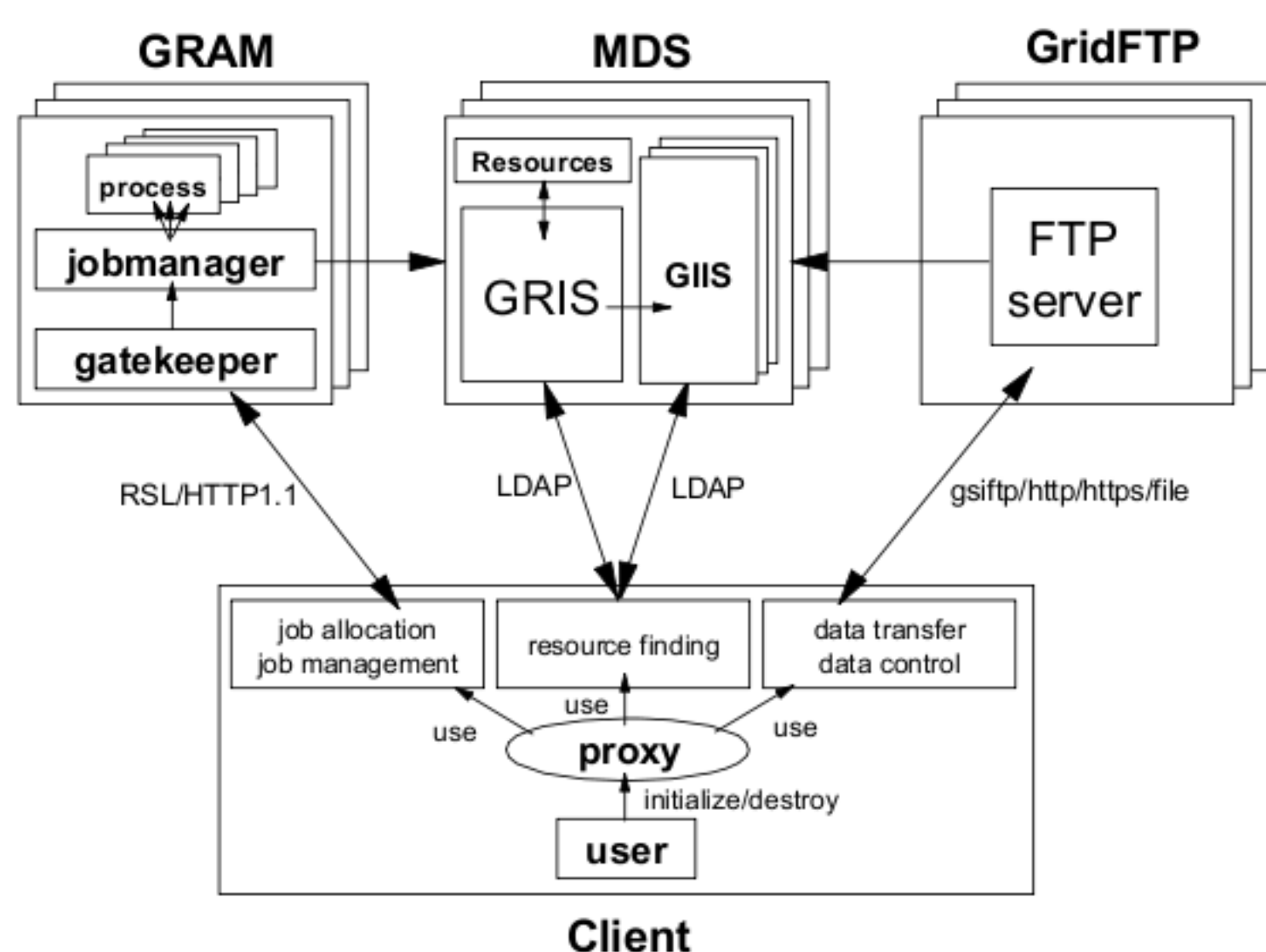
O Globus Toolkit disponibiliza um conjunto de serviços que facilitam a construção de infraestruturas para computação em grid. Ele possui três componentes principais, que são responsáveis pela Gerência de Recursos, Gerência de Serviços e Informação e Gerência de Dados.

Todos esses componentes são têm sua segurança garantida pela *Grid Security Infrastructure (GSI)*, que provê funções de segurança para autenticação, comunicação confidencial, autorização de acesso a recursos, dentre outras.

- **Gerência de Recursos:** Provê a execução remota e a gerência do estado de uma aplicação através do *Grid Resource Allocator Manager (GRAM)* e *Global Access to Secondary Storage (GASS)*.

- **Gerência de Serviços de Informação:** Formada pelos componentes *Grid Resource Information Service (GRIS)* e *Grid Index Information Service (GIIS)*, é organizada de maneira hierárquica de modo a adquirir informação a respeito do grid e distribuí-la. A esses dois serviços é dado o nome *Monitoring and Discovery Service (MDS)*.

- **Gerência de Dados:** Implementada pelo *Grid File Transfer Protocol (GridFTP)*, um protocolo de dados seguro e de alta performance baseado no *File Transfer Protocol (FTP)*.

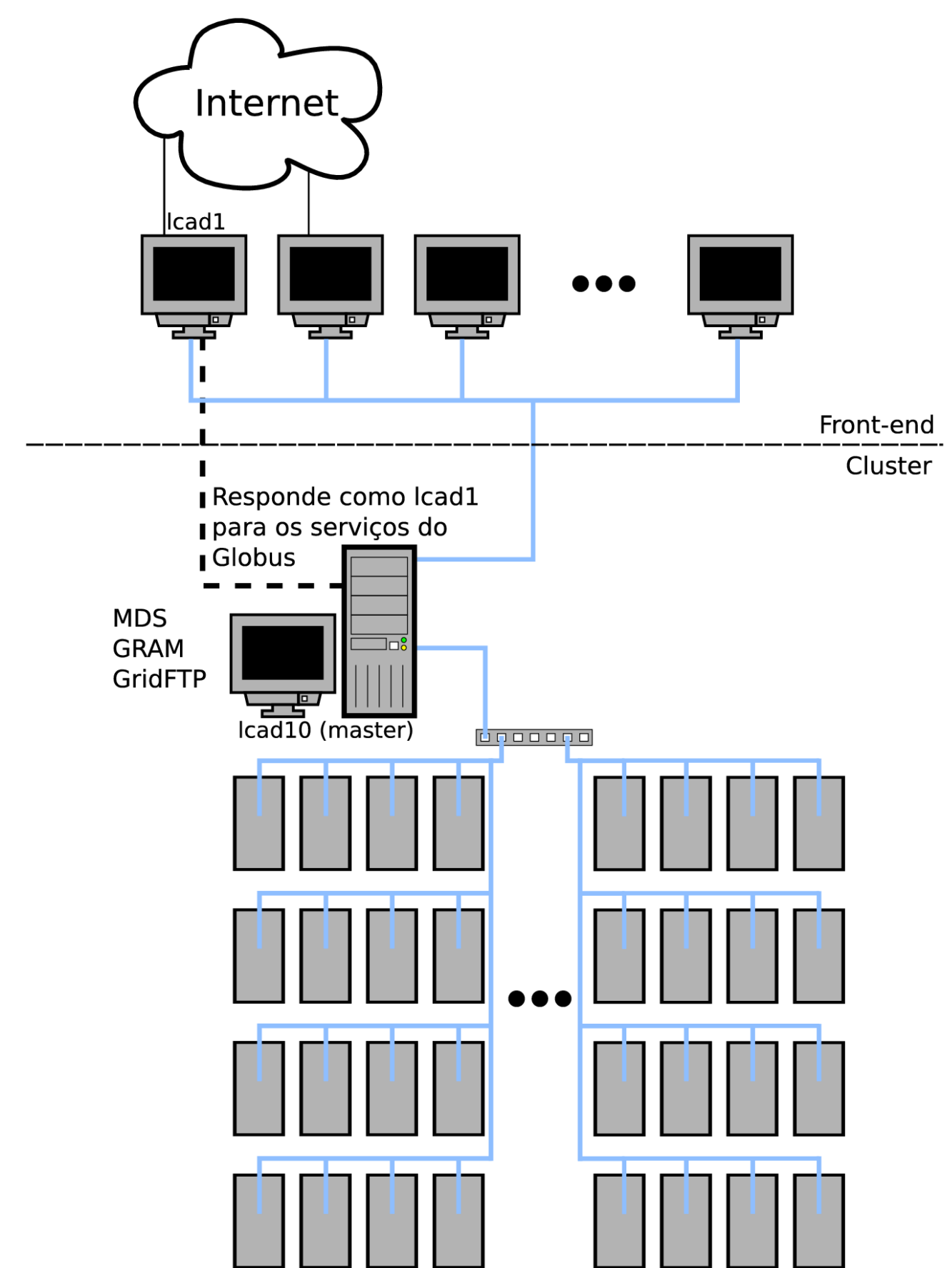


Considerações sobre a Integração

A configuração de rede do Cluster Enterprise não permite que usuários externos à rede do LCAD submetam jobs para o cluster e nem que alguma máquina pertencente ao cluster seja acessível através da internet, sendo essa uma restrição de segurança imposta pela administração da rede do Departamento de Informática. No entanto, para que a integração com o Grid Sinergia fosse realizada com sucesso, seria necessário que disponibilizar um outro método de submissão de jobs para usuários do Grid Sinergia sem obrigá-los a inicialmente entrar na rede do LCAD para então submeter seus jobs. Da mesma forma, seria inconveniente para o administrador do sistema ser obrigado a incluir em seu sistema usuários de outras instituições associadas ao Grid.

A solução encontrada foi a de configurar uma máquina front-end do LCAD como gateway da *master* do cluster utilizando *Network Address Translation (NAT)* para as portas utilizadas pelos serviços do Globus. Para evitar o problema de validação do *host* a executar os jobs realizada pelo código das aplicações clientes do Globus (devido à GSI), a implementação do Globus Toolkit no LCAD segue um esquema de instalação semelhante à de hosts com múltiplas interfaces de rede. Onde é forçado um valor para o *Fully Qualified Domain Name (FQDN)* da *master* no Globus.

Configuração do Cluster Enterprise para o Grid Sinergia



Monitoramento do Cluster

Visando simplificar o processo de obtenção de informação sobre nós pertencentes ao cluster, removendo a necessidade de se utilizar exclusivamente ferramentas de linha de comando e esperando que essa informação pudesse estar disponível ao longo do tempo, optou-se por instalar o software Ganglia nos nós de processamento, ao invés de utilizar diretamente o MDS do Globus.

Ganglia é um sistema de monitoramento para sistemas computacionais de alta performance, como clusters e grids, desenvolvido buscando baixo overhead por nó e alta concorrência e que consiste em dois tipos de processos *daemons*, um programa de linha de comando e uma biblioteca para implementação de clientes.

As maiores facilidades providas pelo Ganglia para esse projeto foram a interface de visualização web com informações completas e a possibilidade de extensão das informações coletadas através do programa de linha de comando *gmetric* aliados ao fato de ser um projeto de Software Livre gratuito, de qualidade e utilizado em várias organizações.

Juntamente com a página web que agrega informações coletadas pelo Ganglia, o site web do LCAD foi reestruturado, teve seu endereço mudado para dentro do domínio do LCAD e conta agora, também, com páginas de documentação do laboratório armazenadas em um sistema *wiki*.

Para evitar o *overhead* gerado por dois sistemas de monitoramento executando simultaneamente no Cluster Enterprise, foi implantado no cluster um provedor de informações do Ganglia para o MDS do Globus. Dessa forma, as informações geradas pelo Ganglia são reaproveitadas pelo MDS, executado apenas na *master* do cluster, que devido à estrutura hierárquica do Globus, disponibiliza as informações do cluster para o MDS do Grid Sinergia.

Proteção do Cluster

Instabilidades da rede elétrica do Departamento de Informática da UFES eventualmente levavam a falhas no sistema de refrigeração do cluster, ocasionando no superaquecimento de alguns nós. Para contornar esse problema foi criado um script que consulta a temperatura dos nós e, se ela ultrapassar um limite pré-estabelecido, então um email é enviado aos administradores do LCAD notificando a situação de superaquecimento e o nó é desligado. As temperaturas coletadas são disponibilizadas pelo Ganglia.

Procedimento semelhante é realizado quando há falta de energia elétrica para o cluster, onde todos os nós são desligados e um email de notificação é enviado aos administradores.