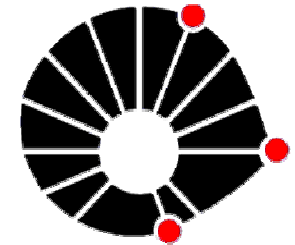
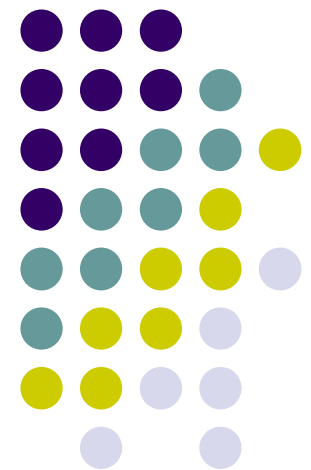


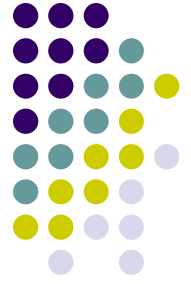
BlueGene/L



UNICAMP

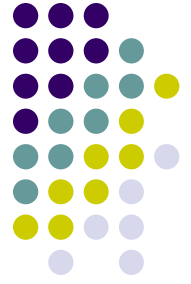
Douglas Gielo Quinellato





Introdução

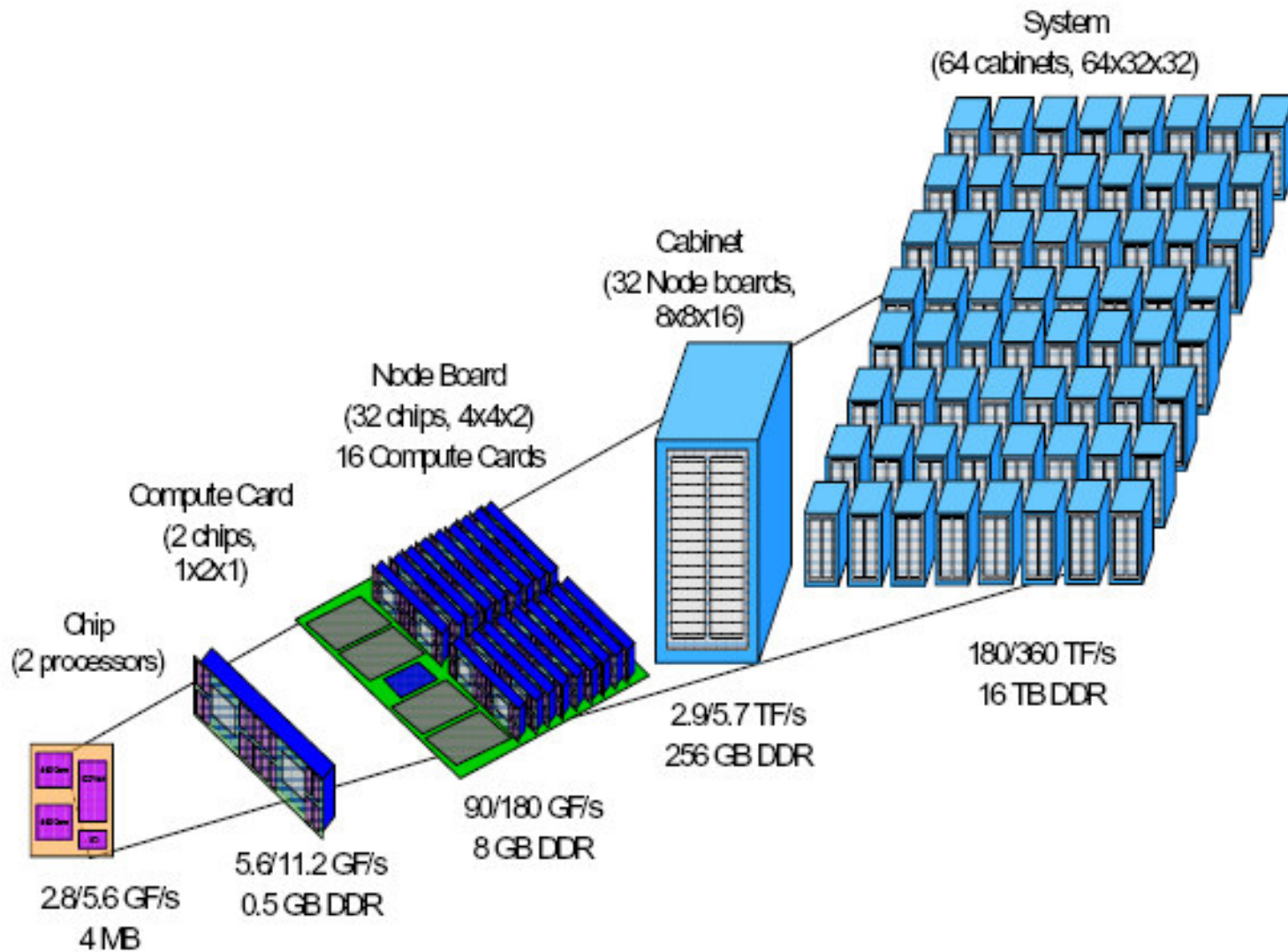
- Supercomputador massivamente paralelo
 - Desenvolvido pela IBM, em parceria com o Lawrence Livermore National Laboratory
- Pico de performance: 360 teraFLOPS
- Até 65.536 nós
 - Processadores de baixo custo/consumo
- Modelo de passagem de mensagem
- Primeiro e segundo lugares na lista top500
 - Primeiro: 131072 processadores (LLNI)
 - Segundo: 40960 processadores (IBM Thomas J. Watson)



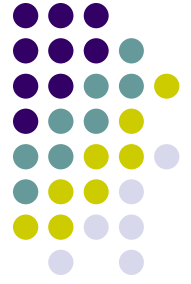
Arquitetura

- Nós formados por 2 processadores PowerPC
 - 700MHz
 - 2 unidades de ponto flutuante
 - Tecnologia System on a chip
- Rede de interconexão em toro, dimensões 64x32x32
 - Comunicação ponto a ponto entre os nós
- Mais 4 redes adicionais, com finalidades específicas

Packing



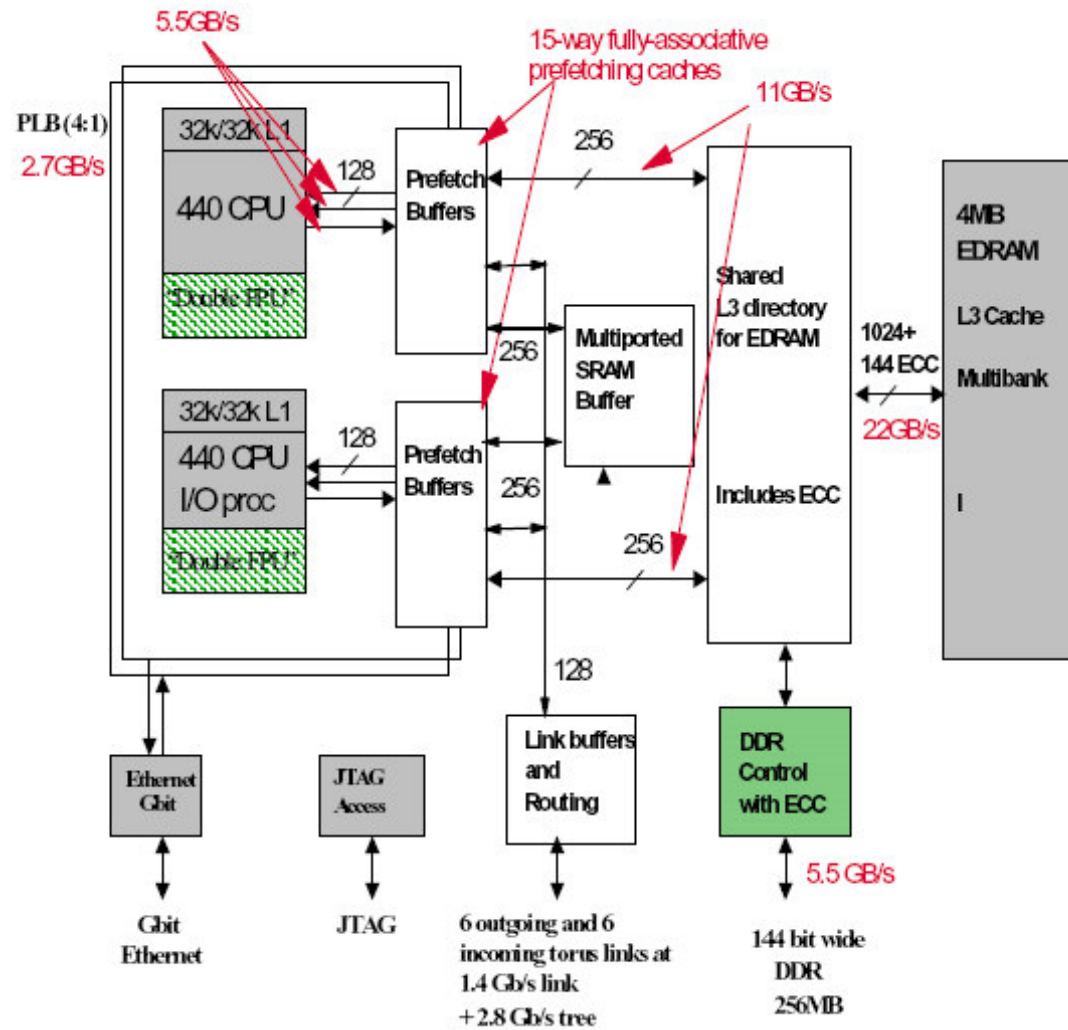
Nós

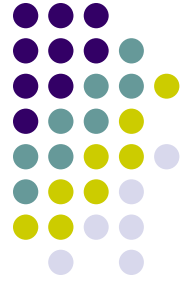


- 2 processadores PowerPC 440
 - Sem coerência de cache L1
- Co-processador FP2, com duas unidades de ponto flutuante, capazes de executar até 8 instruções de ponto flutuante por ciclo
- Cache L2 2KB com prefetch de dados
- Cache L3 4MB
- Interfaces JTAG, ethernet, memória externa
- Geralmente um dos processadores é dedicado ao envio/recepção de mensagens



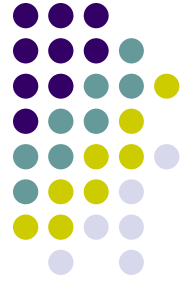
Nós - Diagrama





Redes de Interconexão

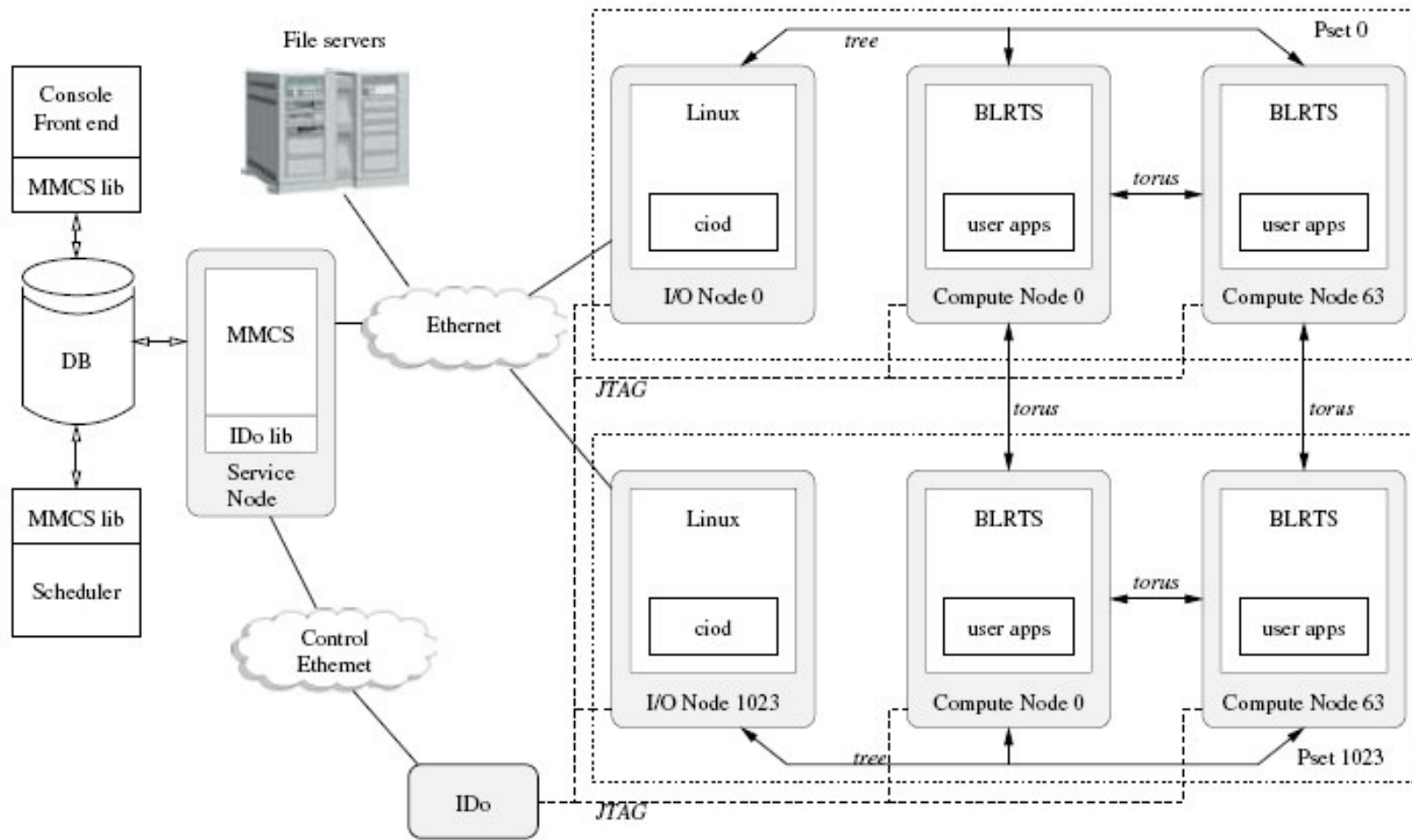
- Toro
 - Links seriais ponto a ponto entre os nós
 - Cada nó ligado com 6 vizinhos
 - Garantia de entrega e checagem de erro
- uma árvore de combinação/broadcast, para otimização de operações do tipo AllReduce();
- uma rede para barrier e interrupt;
- uma rede gigabit para conexão JTAG, para transmitir informações de controle e testes;
- uma rede gigabit para conexão com outros hosts.

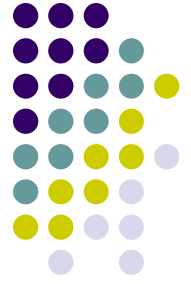


Organização

- Nós podem ser de processamento ou I/O
 - Processamento – utilizado para execução dos processos do usuário
 - Kernel monotarefa (BlueGene/L Run Time Supervisor - BLRTS)
 - I/O
 - Execução de serviços relativos a sistemas operacionais
 - Sockets, sistemas de arquivo...
 - Kernel Linux multitarefa
- Em geral 1 nó de I/O para cada 64 de processamento

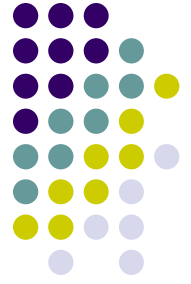
Organização





Sistema Operacional

- Linux
 - Kernel minimalista nos nós de computação
 - Realizam o mínimo necessário, deixando o processamento o maior tempo possível com o processo
 - Kernel normal nos nós de I/O
 - Realizam tarefas relativas a chamadas de sistema
- Sistema Operacional Externo
 - Cria partições, carregam o programa nos nós, etc



Modelo de Programação

- Ferramentas GNU
 - Compilação cruzada
- Padrão MPI para passagens de mensagens
- 3 camadas para envio de mensagens
 - Pacotes
 - Mecanismos básicos para envio e recepção de mensagens
 - API simples (inicializar, enviar, receber)
 - Operações assíncronas
 - Mensagens
 - Envio de mensagens de tamanho arbitrário
 - Divisão da mensagem em pacotes
 - Reordenamento
 - MPI
 - API no nível do usuário



Desempenho

- <http://www.newscientisttech.com/article/dn9394-atomic-simulation-most-intensive-computer-program-ever.html>
- 23 Junho 2006
- Performance mantida (sustained) de 207.03 teraFLOPS
- Qbox
 - Simula iterações entre 1000 átomos em alta pressão utilizando modelo quântico