

APPRO XTREME-X SERVER

Papers and Abstracts

Nelsi Warken

¹Programa de Pós-Graduação em Informática
Mestrado em Ciência da Computação
Universidade Católica de Pelotas, RS, Brasil

nelsi.warken@gmail.com

Resumo. *Este artigo mostra uma visão geral a respeito do Supercomputador Appro Xtreme2, fabricado pela High-Performance and Enterprise Computing. Em novembro de 2007 estava em vigésimo nono colocado no ranking dos Top500 super computers. Appro Xtreme-X Supercomputing é baseado no conceito de cluster escalável (Scalable Unit), gerenciados como um sistema stand alone de alta capacidade e desempenho. Possui como funcionalidades o Directed Airflow Cooling e Appro Cluster Engine Management Software. Utilizado em várias áreas de pesquisa científica, defesa e segurança nacional. O maior cliente é o Lawrence Livermore National Laboratory, Laboratório de Segurança Nacional, localizado em Livermore, CA, USA.*

1. Histórico

A empresa APPRO High-Performance and Enterprise Computing foi fundada em 1991 por Daniel Kim. Sua sede localiza-se nos Estados Unidos da América, em Milpitas, na Califórnia. Possui também um escritório regional em Houston no Texas. A empresa tem penetração mundial via revendedores internacionais [APPRO 2007].

Em 2003, a APPRO expandiu sua equipe de fabricação, engenharia e desenvolvimento de produtos com a criação de uma subsidiária na Ásia. Neste ano, também, lança no mercado sua primeira solução de arquitetura de supercomputador, o APPRO Hyper-Blade Cluster Server Solution, um cluster altamente flexível, modular e escalável, com característica de gestão remota e alta velocidade.

Em 2006 lança no mercado o Appro Xtreme-X Supercomputing, baseado em processadores AMD Opteron. Este lançamento se baseia no conceito de cluster escalável, próxima geração de clusters com desempenho de supercomputadores. Sua arquitetura flexível reúne grupos de clusters integrados em uma unidade escalável (*Scalable Unit - SU*), gerenciados como um sistema *stand alone* de alta capacidade e desempenho.

2. Motivação

A Appro aplica novas reflexões e idéias para criar simples, valiosas e confiáveis experiências com foco na tecnologia de computação de alto desempenho (*High Performance Computing - HPC*). Ao fornecer as melhores soluções, a Appro ajuda a cientistas, engenheiros e analistas financeiros a utilizarem os seus recursos de supercomputação em uma grande quantidade de dados e tarefas, na investigação da descoberta científica, modelagem de dados, em simulações de engenharia e visualização sísmica.

Seu desafio está no desenvolvimento de soluções cuja relação custo/desempenho seja a melhor possível, utilizando uma arquitetura equilibrada, padrões abertos e conhecimentos de engenharia.

3. Objetivos

A Appro tem como objetivos: desenvolver *Workstations*, servidores, clusters, supercomputadores e soluções, através do compromisso com a inovação do produto, qualidade e serviço ao cliente. Seu nicho de mercado são as empresas de médio a grande porte.

O objetivo da criação do XTREME-X2 é oferecer a mais alta performance por dollar e por watt.

A figura 1 mostra o supercomputador XTREME-X2 da Appro High-Performance and Enterprise Computing.



Figure 1. Supercomputador XTREME-X2

4. Onde está instalado

1) Administração Nacional de Segurança Nuclear (NNSA): Lawrence Livermore National Laboratory, Los Alamos National Laboratory e Sandia National Laboratories. A APPRO vendeu supercomputadores para atender ao projeto TLCC07 da NNSA [TLCC 2007]. Lawrence Livermore National Laboratory (LLNL) é um laboratório de segurança nacional com a responsabilidade de garantir que as armas nucleares dos Estados Unidos permaneçam salvas, seguras e confiáveis. Fundado em 1972 em Livermore, CA, o laboratório desenvolve atividades de design e pesquisa em armas nucleares. Além disso, aplica seus conhecimentos especializados e suas capacidades multidisciplinares em programas de pesquisa em meio ambiente e pesquisa em biociência para desenvolvimento de biodefesas, gerenciamento do estoque de armas nucleares, e em esforços para evitar a disseminação de energia nuclear e outras armas de destruição de massa, além de reforçar a segurança interna do país. Muitos projetos são elaborados em colaboração com universidades, indústrias e outros laboratórios [LLNL 2007]. Em julho de 2007,

atendendo um contrato, a Appro forneceu para a LLNL o Xtreme-X High Performance Computing, clusters Linux escaláveis de alto desempenho com até 438 teraflop/s em 21 Unidades Escaláveis (SU), 3024 nodos, 12096 processadores com 48384 *cores* e 96,8TB de memória. O projeto ainda tem uma opção adicional para 10 SU, 200 teraflops/s de poder de computação para ser implantado em 8 clusters Linux nos três laboratórios da NNSA.

2) A Renault F1 Team Computation Fluid Dynamics (CFD) realizou um contrato com a Appro, anunciado em 04 março de 2008. CFD utiliza moderna tecnologia de condução de pesquisa realizada para colocar a Renault na *pole position* na competição de F1. O Appro Xtreme-X2 Supercomputer será transferido e instalado na Renault F1 CFD Center, no Reino Unido, até o final de Junho de 2008. Este supercomputador é baseado em processador Quad-Core AMD Opteron, tendo uma capacidade máxima de 38 TF com gerenciamento de um sistema unificado de supercomputação. O sistema tem 4,4 TB de memória local de alto desempenho. Cada *dual processor node* tem capacidade pico de mais de 70 gigaflops. Os nodos no sistema são interligados por um Double DDR de Infiniband (4Gbits/s), reduzindo a latência inter-processador e melhorando o sistema de confiabilidade.

3) A Universidade de Tsukuba do Japão contratou um 95TF Appro Xtreme-X Supercomputer para o Centro de Ciências Computacionais, localizado no centro da cidade de Tsukuba Science. A região é conhecida por sua grande concentração de instituições de pesquisa avançada em diversas áreas da ciência e da tecnologia. Os supercomputadores Xtreme-X3 serão utilizados, principalmente, para fornecer suporte computacional em várias áreas de pesquisa e educação, incluindo inteligência artificial, computação paralela, modelagem matemática, sistemas de software, arquitetura de computadores e engenharia média, além de outras áreas de vanguarda da ciência e da tecnologia. O Appro Xtreme-X3 Supercomputer será transferido e instalado até maio de 2008. Sua arquitetura de 674 processadores Quad-Core AMD Opteron e 95TF de capacidade de computação, podendo ser escalado até 120TF. O sistema fornece 21,5TB de memória local de alto desempenho.

A figura 2 mostra as Unidades Escaláveis Xtreme-X instaladas no Lawrence Livermore National Laboratory.

5. Aplicações

Appro Xtreme-X Supercomputer é utilizado nas seguintes áreas:

- Computacional Fluid Dynamics;
- simulações de engenharia aeroespacial e automotiva;
- exploração e produção petrolífera;
- visualização científica;
- descoberta e refinamento de petróleo;
- pesquisa sísmica;
- meteorologia e ciências ambientais;
- pesquisa em energia, meio ambiente, biociência e biodefesas;
- simulações meteorológicas;
- pesquisa em armas nucleares (defesa e segurança);

- inteligência artificial;
- modelagem matemática.

Appro Cluster Engine Management Software possui características com solução de gerenciamento de toda configuração global do sistema e vem pré-instalado em todos os supercomputadores XTREME-X. Este software oferece a facilidade de interface web, permitindo controlar o Appro Xtreme-X2 a partir de qualquer local. Os módulos do software incluem gerenciamento de Rede, gerenciamento de servidores, gerenciamento de cluster e gerenciamento de armazenamento. Além disso, o ACEMS suporta múltiplas topologias de rede, configurações *diskless*, *network failover* para atingir máxima confiabilidade, desempenho e alta disponibilidade. Ele também suporta múltiplos *root file systems*, oferecendo instantâneo provisionamento para rápida instalação de padrão Linux em grandes sistemas *diskless*, permitindo dar boot de 100 a 1000 *blades* quase ao mesmo tempo. O Appro Cluster Engine Management Software oferece uma completa, confiável e alta disponibilidade em arquiteturas que suportam escalabilidade com multi-configurações de rede.

6. Sistema Operacional

O sistema operacional utilizado nas máquinas XTREME-X é *Operating System RedHat AS Version 5*.

7. Configuração da Arquitetura

Características da arquitetura do XTREME-X:

- Solução de pré-pacote totalmente integrado;
- Oferece uma variedade de configurações de sistema;
- Suporte para Quad-Core AMD Opteron Processors;
- 128 sockets com 512 núcleos (*cores*);
- 4TB de memória principal por rack;
- Cache 2MB L2 cache per processor;
- Eficiência energia X refrigeração
- Arquitetura *Cooling*: Appro Directed Airflow;
- Redundant (Dual-Rail) Infiniband Interconnect Fabric;
- Redundant (Dual-Rail) Ethernet Management Fabric de 10 GB;
- Três controles independentes de refrigeração por nodo;
- Fontes redundantes N+1 de alimentação, 90% de eficiência;
- Completo sistema de gerenciamento;
- Sistema de produção de acordo com a solicitação dos requisitos;
- Interconexões Intel fiber optic entre racks e a tecnologia de multi-core.

Arquiteturas de Referência do XTREME-X2 Rack Cabinet:

- 1 Rack escala até 4.5TF com 512 cores;
- 2 Racks escala até 9TF com 1024 cores;
- 4 Racks escala até 18TF com 2048 cores;
- 8 Racks escala até 36TF com 4096 cores;

- 18 Racks escala até 81TF com 9216 cores;
- 36 Racks escala até 162TF com 18432 cores;
- 72 Racks escala até 324TF com 36864 cores;

Permite um ambiente de supercomputação até 1000TF.

Um único equipamento rack pode ter 64 nodos com 128 processadores e 512 núcleos. O pico de capacidade de um único rack é 4,5TF por segundo. A capacidade máxima de memória de um único rack é 4TB e a largura de banda de memória agregada é mais 3TB por segundo. A configuração modular de um rack de 4,5TF é uma unidade escalável (SU) que pode ser replicada até formar sistemas com até 72 racks, através de três níveis de rede fat-tree.

8. Processadores

Quad-Core AMD Opteron Processor

A figura 3 mostra a estrutura do processador AMD Opteron.

9. Estrutura Física

Outra facilidade disponível nas máquinas Appro XTREME-X é o *Directed Airflow Cooling*, influenciando a estrutura física do local de instalação:

- reduz espaço de piso - 30%;
- maior eficiência e refrigeração no consumo de energia;
- acesso frontal para manutenção;
- montagem em racks numa Unidade Escalável (SU) uma ao lado da outra, beneficiando-se do sub-chão frio disponível, fornecendo ar frio diretamente para o equipamento, eliminando a necessidade de ar entre os racks;
- fluxo de ar entre o menor numero de componentes;
- pode ser ejetado, através de acesso no painel frontal;
- mais da metade do painel frontal é perfurado para minimizar turbulência e reduzir a pressão.

10. Rede de Interconexão

A rede de interconexão é DDR InfiniBand com ConnectX host channel adapters, com desempenho escalável de Interconexão.

O InfiniBand ainda está em desenvolvimento e promete várias novidades para o futuro. O principal objetivo deste novo barramento de dados é interligar servidores e dispositivos de armazenamento localizados a curtas distâncias, servindo como uma opção mais rápida às redes Ethernet. Usando o InfiniBand um servidor de bancos de dados poderia acessar um dispositivo de armazenamento externo, sem nenhum gargalo, como se fosse um dispositivo local, o que abre muitas possibilidades nos clusters e servidores de alto desempenho.

Infiniband - inclui qualidade de serviço, *failover* e pode ser escalável. A especificação da arquitetura Infiniband define uma conexão entre nodos processadores e nodos de I/O de alta performance, tais como dispositivos de armazenamento.



Figure 2. Lawrence Livermore National Laboratory

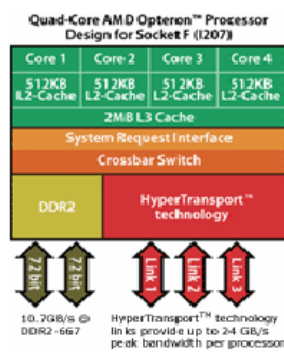


Figure 3. Processor Quad-Core AMD Opteron

Failover - capacidade de trocar automaticamente para um redundante ou *standby* servidor, sistema ou rede, quando ocorre uma falha [Infini 2007].

No Appro XTREME-X2 os nodos são conectados com um Infiniband Dual-Rail de 20Gb/s, 4X *Double Data Rate fabric (DDR)*. Tem duas portas de gerenciamento Ethernet Gigabit e duas conexões DDR para interligar o InfiniBand fabric.

11. TOP 500

- Site: Lawrence Livermore National Laboratory
- System Family: Appro HyperBlade
- System Model: Appro XtremeServers
- Computer: Appro Xtreme Server - Quad Opteron Dual Core 2.4GHz Infiniband
- Vendor: Appro International
- Main Memory: 18432 GB
- Installation Year: 2007

Performance Linpack Data:

Processors	Rmax(GFlops)	Rpeak(GFlops)	Nmax
9216	36620	44240	1021440

Informações gerais e conceituais:

- Processadores Core 2 Quad: são processadores formados por duas pastilhas de silício, integradas no mesmo chip, onde cada uma possui dois núcleos (*cores*), formando um total de quatro núcleos.
- Gflop/s é a taxa de execução, bilhões de operações de ponto flutuante por segundo. Sempre que este termo é utilizado ele irá referir-se a operações de ponto flutuante 64 bits (adição ou multiplicação). Também um Tflop/s é utilizado para o mesmo fim, mas refere-se a trilhões de operações de ponto flutuante por segundo.
- Rmax - Maxima Performance alcançada pelo Linpack.
- Rpeak - performance teórica de pico.
- Nmax - tamanho máximo do problema conseguido no Rmax: numero de equações lineares utilizadas (métrica).
- Linpack - é um benchmark que utiliza a melhor performance como medida. Está disponível na maioria dos sistemas relevantes. Usado para solucionar um denso sistema de equações lineares.
- A ordem de classificação no Top500: maior Rmax. Se for igual ao de outra máquina, é também considerado o maior Rpeak. E para sites que tem a mesma máquina, a ordem é dada pelo tamanho da memória e depois por ordem alfabética.
- História do *Ranking* do computador Appro Xtreme Server - Quad Opteron Dual Core 2.4GHz Infiniband:

List	Rank	Processors	Rmax(GFlops)	Rpeak(GFlops)
11/2007	29	9216	36620	44240
06/2007	19	9216	36580	44240

[Top 2007]

References

APPRO (2007). Appro high-performance and enterprise computing. <<http://www.appro.com>>.

Infini (2007). Infiniband. <<http://http://www.linuxhpc.org>>.

LLNL (2007). Lawrence livermore national laboratory. <<http://www.llnl.gov>>.

TLCC (2007). Projeto tlcc07. <<https://asc.llnl.gov/publications/sc2007-tlcc.pdf>>.

Top (2007). Top500 sites list. <<http://www.top500.org/>>.