Supercomputing and Opensource after the K-Computer

Christian Külker

Debian Edu/ Skolelinux

v0.1

2012-03-18

Contents

What is a supercomputer?

What is a Supercomputer?

- A Supercomputer a big computer
- »Super« stands for something extraordinary in terms of performance
- A unambiguous definition do not exist, because the method of measuring the performance (speed of calculation) is not possible on all high performace computers in the same manner
- A popular definition is, at least all computers out of the Top500 and GreenTop500 list are Supercomputers.

Different Types of Supercomputers

- Scalar processors <70th
- Vector processors >70th - mid 80th
- Parallel processing mid 80th - 90th
- Custom made processors (APE) and commodity processors (Intel, AMD, Alpha, ...)
- modern supercomputers highly-tuned computer clusters using commodity processors combined with custom interconnects
- CPU/GPU and other accelerators (FPGA, ...)
- Different coupling (Strong: APEmille, Loose: Cluster)
- Different Networks (Ethernet, Infiniband, Torus, Mesh, ...)

Top500/Green Top500

Top500 - World fasted Computers

- The Top500 provides rankings of the most powerful supercomputers in the world.
- GFLOPS, TFLOPS, PFLOPS
- http://www.top500.org

The top 10 Hit Parade - of Top500, Nov. 2011

#	Name	OEM	Country	Cores	Rmax	Eff.	Processor	OS	Accel.
1	K computer	Fujitsu	Japan	705024	10.510	93.17	SPARC64	**Linux**	None
2	Tianhe-1A	NUDT	China	186368	2.566	54.58	Xeon	**Linux**	NVIDIA
3	Jaguar	Cray Inc.	USA	224162	1.759	75.46	Opteron	**Linux**	None
4	Nebulae	Dawning	China	120640	1.271	42.59	Xeon	**Linux**	NVIDIA
5	Tsubame 2.0	NEC/HP	Japan	73278	1.192	52.11	Xeon	**Linux**	NVIDIA
6	Cielo	Cray Inc.	USA	142272	1.110	81.27	Opteron	**Linux**	None
7	Pleiades	SGI	USA	111104	1.088	82.72	Xeon	**Linux**	None
8	Hopper	Cray Inc.	USA	153408	1.054	81.79	Opteron	**Linux**	None
9	Tera-100	Bull SA	France	138368	1.050	83.7	Xeon	**Linux**	None
10	Roadrunner	IBM	USA	122400	1.042	75.74	PowerXCell	**Linux**	IBM

Rmax [PFLOPS], Eff. [%]

Green Top500 - Supercomputing goes Green

- Green 500 (http://www.green500.org)
- In the context of The Green500 List, a supercomputer is a computing system that is fast enough to appear of the latest Top500 List.
- In the context of The Little Green500 List, a supercomputer is a computing system that achieves performance on the HPL benchmark at a high-enough level to have secured entry into the oldest Top500 list released within 19 months.
- http://www.green500.org/docs/pubs/RunRules_Ver0.9.pdf

Green Top500 - Environmentally Responsible Supercomputing

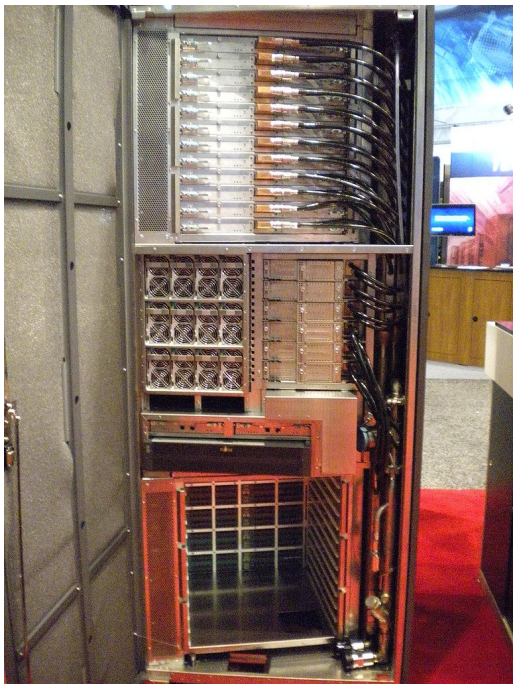
- The Green500 provides rankings of the most energy-efficient supercomputers in the world.
- raise awareness about power consumption,
- alternative performance metrics
- MFLOPS/W (Example IBM Blue Gene/Q 2026 MFLOPS/W)

Green Top500 - trends of the 10 greenest supercomputers in the world

- (1) aggregating many low-power processors like IBM BlueGene/Q mostly from the embedded world
- (2) using energy-efficient accelerators, typically from the gaming/graphics market, e.g., AMD Radeon GPU, NVIDIA Tesla Fermi GPU, Cell, and Intel Knights Corner
- => http://www.green500.org/

Hardware: The K computer

The 京Computer



- named for the Japanese word "kei" (京), meaning 10 quadrillion (German: 10 Billiarden, 10 Peta).
- installed at the RIKEN Advanced Institute for Computational Science (AICS) campus in Kobe, Japan.

Hardware: The K Computer

Active	Operational June 2011
Sponsors	MEXT, Japan Japan
Operators	Fujitsu
Location	RIKEN Advanced Institute for Computational Science
Architecture	88,128 SPARC64 VIIIfx processors,
	Tofu interconnect
	Linux-based enhanced operating system
Speed	10.51 petaflops (Rmax)
Ranking	TOP500: 1, November 2011

=> http://en.wikipedia.org/wiki/K_computer

Hardware: The K Computer

- uses 88,128 2.0GHz 8-core SPARC64 VIIIfx processors
- packed in 864 cabinets
- a total of 705,024 cores
- 45 nm CMOS technology
- Each cabinet contains 96 computing nodes
- 6 I/O nodes.
- a computing node contains a single processor and 16 GB of memory

Software: How it is normally done?

Software: How it is normally done? (exaggerated)

- build computer
- find users
- develop or port software
- run software
- improve software

Software: How was it done for the K computer?

Software: How was it done for the K computer?

- find problems
- find users
- develop or port software
- build small subset of computer
- run software on a small set, evaluate
- build real computer
- improve software

K Computer Software Goals

- Analyzing the behavior of nanomaterials through simulations and contributing to the early development of such next-generation semiconductor materials, particularly nanowires and carbon nanotubes
- Predicting which compounds, from among a massive number of drug candidate molecules, will prevent illnesses by binding with active regions on the proteins that cause illnesses
- Simulating the actions of atoms and electrons in dye-sensitized solar cells (higher energy efficiency)
- Simulating seismic wave propagation and tsunamis to predict the effects they will have on human-made structures (design of quake-resistant structures)
- Conducting high-resolution (400-m) simulations of atmospheric circulation models to provide detailed predictions of weather phenomena

K Computer - Scaling Software

Program	Fields	Application Overview	Techniques
NICAM	Earth Science	Nonhydrostatic icosahedral atmospheric model	FDM (atmosphere)
Seism3D	Earth Science	Seismic wave propagation / strong seismic motion simulation	FDM (wave motion)
PHASE	Nanoscience	Plane wave expansion first-principle molecular dynamics analysis	plane wave DFT
FrontFlow/Blue	Engineering	Unsteady flow analysis based on large eddy simulation (LSE)	FEM (fluid)
RSDFT	Nanoscience	Real-space first-principle molecular dynamics calculation	real-space DFT
LatticeQCD	Physics	Lattice QCD simulation based elementary particle and nucleus research	QCD

Hardware: Interconnects

Hardware: Network Inter-Connects

- Link technologies: ethernet, infiniband, custom
- Links topologies: star, ring, fat tree, ..., torus
- 3D/6D Torus
- Challanges: fault/error tollerance

Your Laptop is a Supercomputer

Compiling a Program

- Compiling and Linking MPI Programs
- Linux with Ethernet or Infiniband (OpenMPI):

```
mpicc hello.c -o hello
```

- Running MPI Programs
- for OpenMPI create machine file (hosts to be used)
- define your resources CPU, Cores
- load your environment (mpi-selector)
- Example run (Default network, for example infinband):

MPI Communication pattern

- point2point
- Collective: broadcast
- one2all
- all2one
- all2all n+1 -> n-1

MPI C Example

```
/* C Example */
#include <stdio.h>
#include <mpi.h>

int main (argc, argv)
     int argc;
     char *argv[];
{
  int rank, size;

  MPI_Init (&argc, &argv); /* starts MPI */
  MPI_Comm_rank (MPI_COMM_WORLD, &rank); /* get current process id
  MPI_Comm_size (MPI_COMM_WORLD, &size); /* get number of processe
  printf( "Hello world from process %d of %d\n", rank, size );
  MPI_Finalize();
  return 0;
}
```

mpirun

```
ckuelker@hpc> mpirun -np 4 -machinefile machinefile.hpc hello
    Process 0 on host01 out of 4
    Process 1 on host01 out of 4
    Process 3 on host02 out of 4
    Process 2 on host02 out of 4
```

- run via scheduler/ queue

```
qsub -pe nc0 16 ~/bin/hello-world.sh
```

SC11

Supercomputing 11 - Seattle

- Exascale
- Knights Corner
- K Computer
- AMD Interlagos
- Cooling
- Power saving
- Cloud

Thank you for listening

Christian Külker

Open Source Projects:
- Debian Edu
- CipUX

christian.kuelker@cipworx.org

Occupation:

HPC Project Manager

Partnership Program Coordinator

Eurotech - ETH Lab - Business Unit HPC