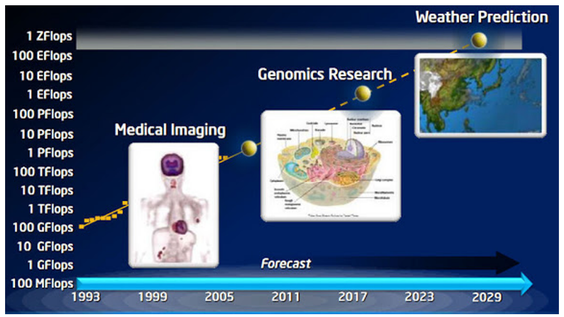
Optimizing Power for Tomorrow's Supercomputer

Stephanie Labasan, Hank Childs

Department of Computer and Information Science, University of Oregon

High-Performance Computing (HPC) is an essential component in understanding and advancing scientific fields ranging from climate to fuel efficiency. These fields are requiring greater and greater computational power to achieve the next set of breakthroughs, which has consequently led to the study of Exascale systems: machines capable of 10^{18} operations per second. The transition to Exascale introduces many new challenges with the most significant centering on power and energy consumption. This work investigates techniques to increase energy efficiency for visualization and data analysis routines.

Motivation

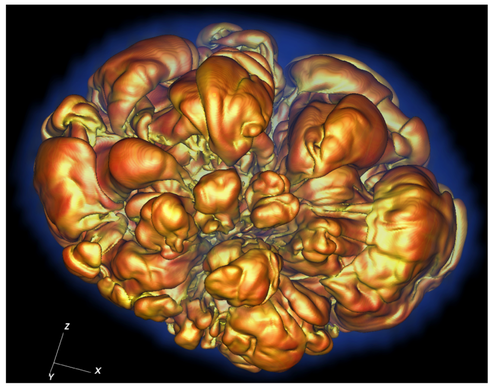


More computationally powerful supercomputers are necessary to continue advancing scientific fields, such as climate change, fluid dynamics, and nuclear fusion. We need to save **100X** in power bills in order to meet the **$20M per year** goal. [1]

"Power costs for the largest petaflop systems are in the range of $5-10M annually…So, to achieve an Exascale system using current technology, the annual power cost to operate the system would be above **$2.5B per year**."

- *ASCAC Summary Report [2]*

Isosurfacing Algorithm



Isosurface of temperature from star explosion simulation (ignore blue haze).

- Surface represents points of a constant value within a volume
- Common visualization technique
- Acts as case study because of its power intensive routine characterized by:
  - cache misses
  - big data processing

Conclusion

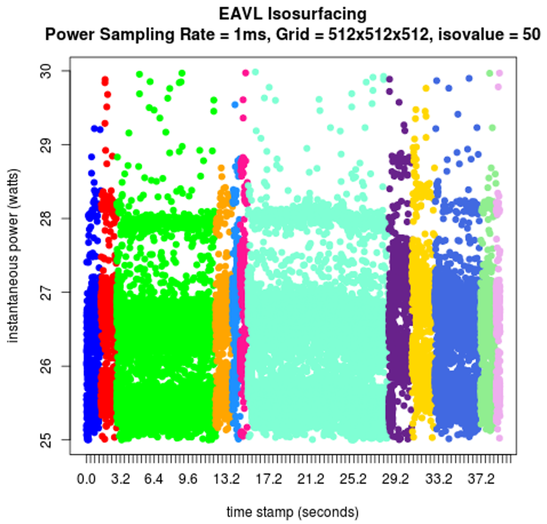
- High density between 25W and 30W
- Bi-modal and tri-modal trends in watt values
- Baseline of ~27W, visualization application contributes ~5W

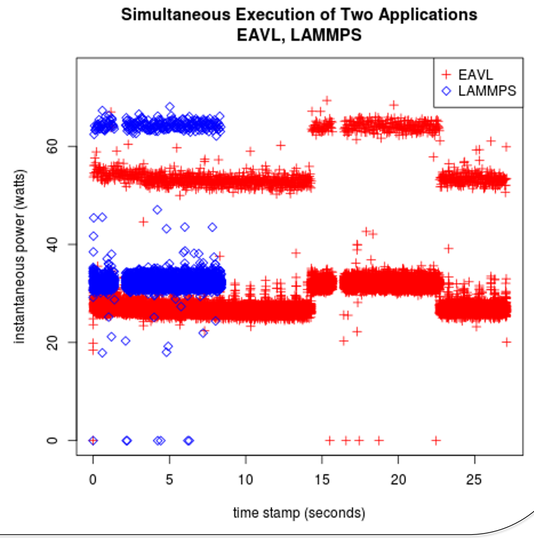
Exascale Computing

Systems	2009	2011	2018 (not optimized)	2018 (optimized)
System Peak Flop/s	2 Peta	20 Peta	**1 Exa**	**1 Exa**
System Memory	0.3 PB	1 PB	50 PB	10 PB
Node Performance	125 GF	200 GF	10 TF	1-10 TF
Node Memory BW	25 GB/s	40 GB/s	2,000 GB/s	200-300 GB/s
Interconnect BW	1.5 GB/s	10 GB/s	500 GB/s	50 GB/s
System Size (Nodes)	18,700	100,000	5,000,000	Millions
Storage	15 PB	30 PB	1,500 PB	300 PB
I/O	0.2 TB/s	2 TB/s	100 TB/s	20 TB/s
Power	6 MW	~10 MW	~500 MW	~20 MW

[3]

Results





Future Work

- Detailed attribution of peaks and valleys to performance counters
- Power optimization of ~5W in visualization application

References

[1] http://www.33rdsquare.com/2012/02/exascale-computing-race-heats-up.html

[2] http://science.energy.gov/~/media/ascr/ascac/pdf/reports/exascale_subcommittee_report.pdf

[3] http://www.scidacreview.org/1001/html/hardware.html