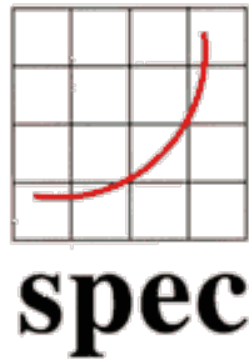# Implications of SPECPower and the Green500 for HPC

Kirk W. Cameron

Virginia Tech
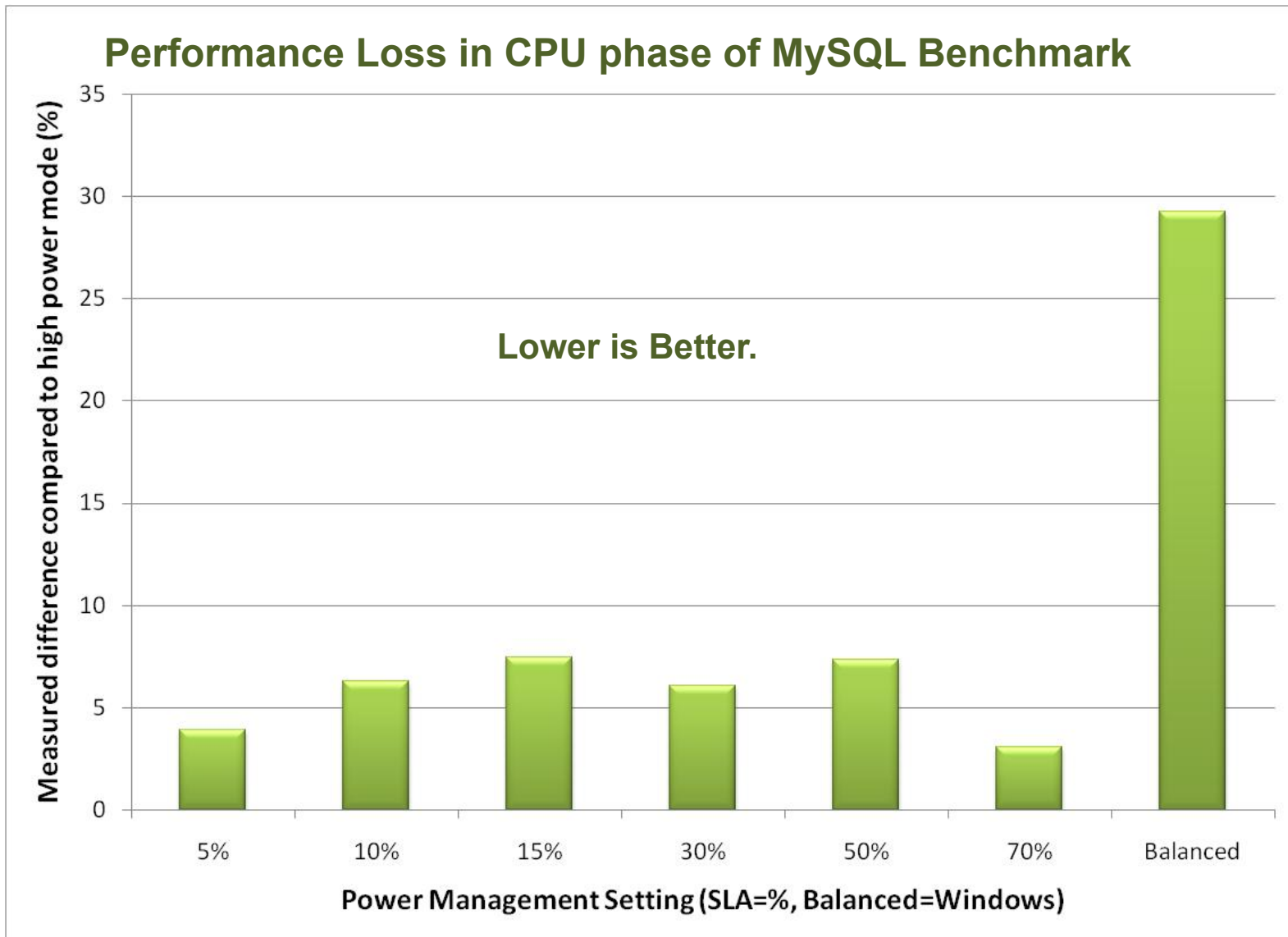
The Problem SCAPE@VT Addresses...

Amount and cost of power continues to increase.

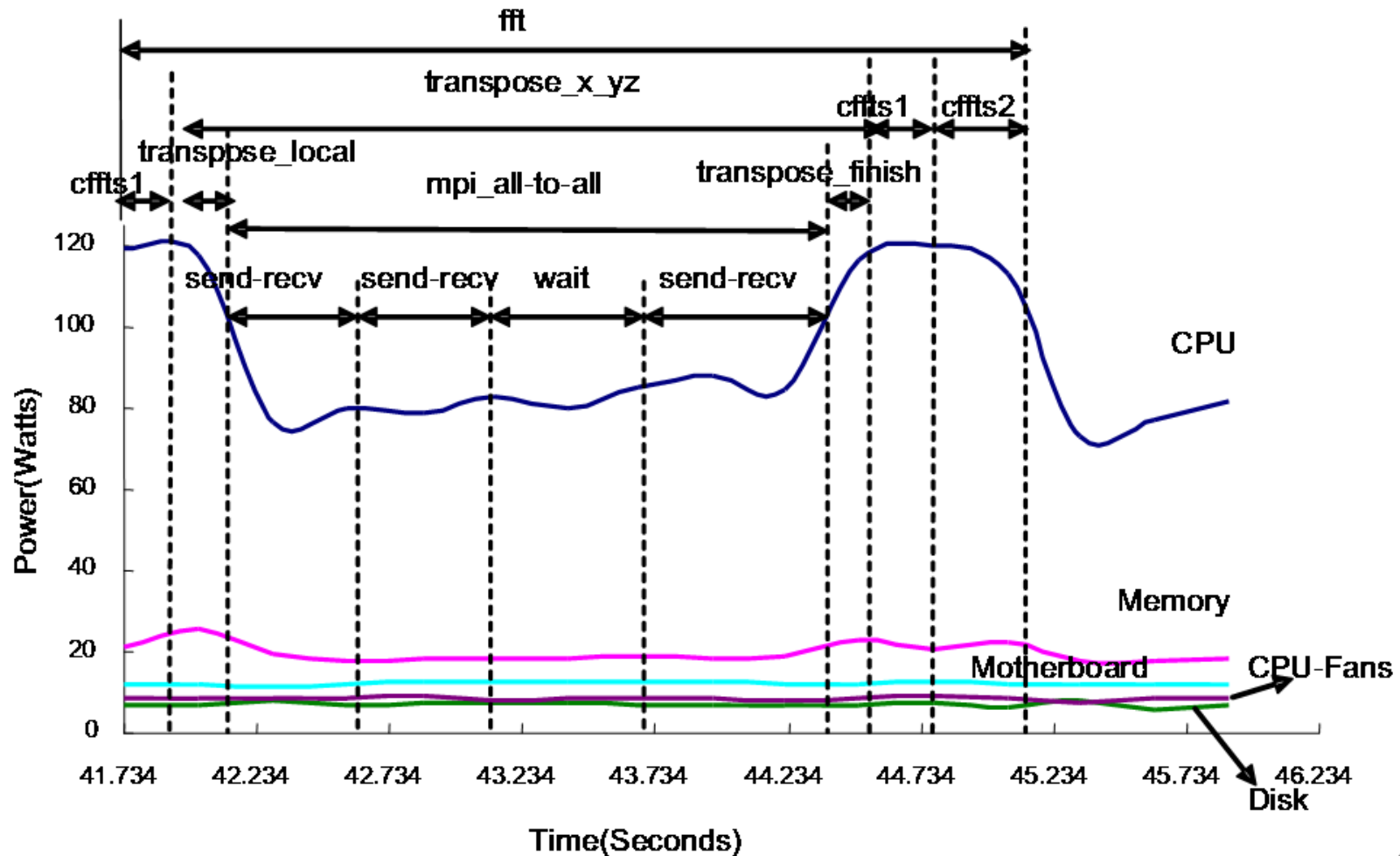**Power management features disabled by default.**

Why is PM turned off?

Performance Loss in CPU phase of MySQL Benchmark

Lower is Better.

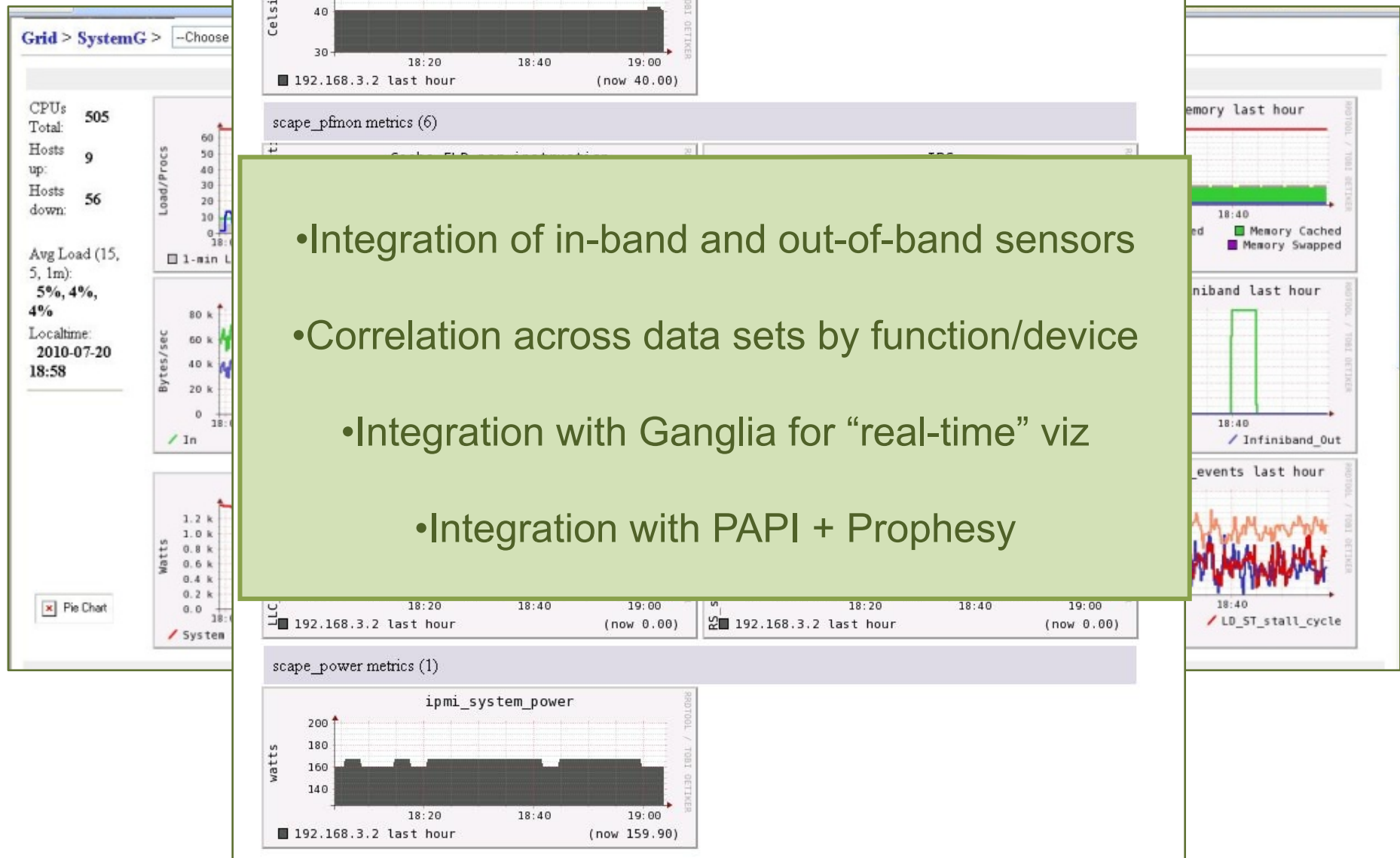Power Management Setting (SLA=%, Balanced=Windows)

SCAPE@VT Related Work

- Metrics and Outreach
  - Research on SPECPower, Green500
  - NSF CCF 0848670 (VT)
- Power Measurement/Analysis/Prediction
  - PowerPack v1.0 → 3.0
  - Used by about a dozen universities
  - On Clusters: NSF CNS 0910784 (Texas A&M, UTK)
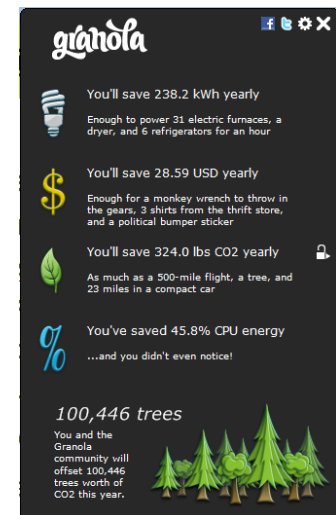 - On Grids: NSF CNS 0905187 (UTSA, UF, UTK)

PowerPack 1.0→2.0

PowerPack 3.0

- Integration of in-band and out-of-band sensors

- Correlation across data sets by function/device

- Integration with Ganglia for "real-time" viz

- Integration with PAPI + Prophesy

SCAPE@VT Related Work

- Power Management Research
  - Multi-core Clusters: NSF CNS 0720750
  - SystemG (28 Tflops): NSF CNS 0709025
  - HPPAC (2004-2009): NSF CCF 0614705
  - Up to 35% energy savings with performance SLA

- VT IP Commercialization
  - Venture-backed Startup: MiserWare
  - Software for PM + SLAs
  - Free Version: http://grano.la
    - 100K Downloads in first 100 days

granola

You'll save 238.2 kWh yearly
Enough to power 31 electric furnaces, a dryer, and 6 refrigerators for an hour

You'll save 28.59 USD yearly
Enough for a monkey wrench to throw in the gears, 3 shirts from the thrift store, and a political bumper sticker

You'll save 324.0 lbs CO2 yearly
As much as a 500-mile flight, a tree, and 23 miles in a compact car

You've saved 45.8% CPU energy
...and you didn't even notice!

100,446 trees
You and the Granola community will offset 100,446 trees worth of CO2 this year.
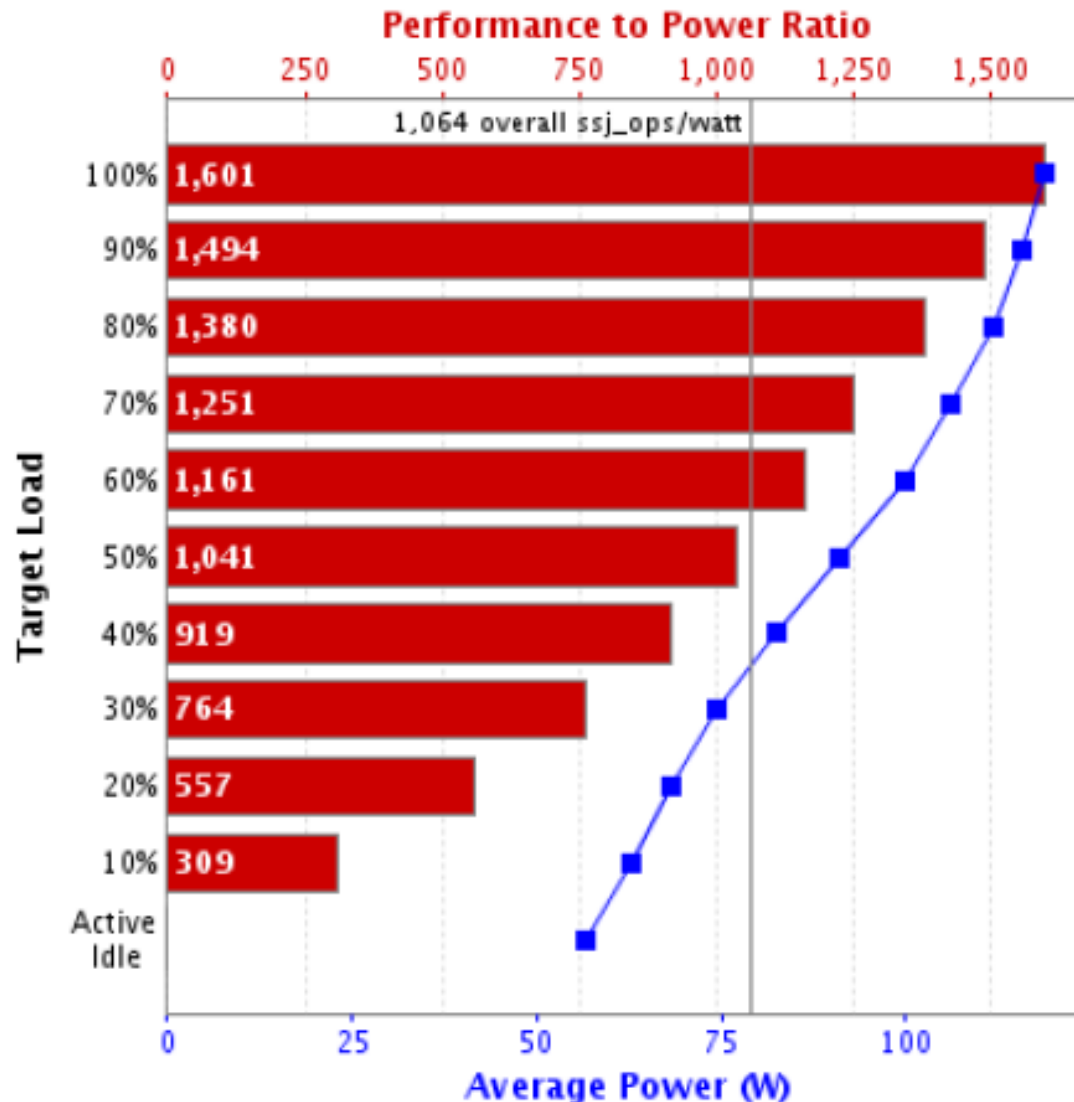
Focus of this talk

- SPECPower (~3 years later)
  - How it works
  - Trends and analysis
- Green500 (~3 years later)
  - How it works
  - Trends and analysis
- Implications for HPC

SPECPower Example (specjbb2008)

Source: SPECPower Benchmark Methodology v2.0 (http://www.spec.org/power)

SPECPower Example (specjbb2008)

Performance			Power	Performance to Power Ratio
Target Load	**Actual Load**	**ssj_ops**	**Average Power (W)**	
100%	99.8%	190,234	119	1,601
90%	90.7%	172,967	116	1,494
80%	80.8%	154,130	112	1,380
70%	69.7%	132,811	106	1,251
60%	60.8%	115,866	99.8	1,161
50%	49.6%	94,582	90.9	1,041
40%	39.7%	75,792	82.5	919
30%	29.8%	56,857	74.4	764
20%	19.9%	37,980	68.2	557
10%	10.2%	19,410	62.8	309
	Active Idle	0	56.7	0
			$\sum$ssj_ops / $\sum$power =	1,064

Source: SPECPower Benchmark Methodology v2.0 (http://www.spec.org/power)

SPECPower Load-level Comparison

Source: Technet (http://blogs.technet.com)
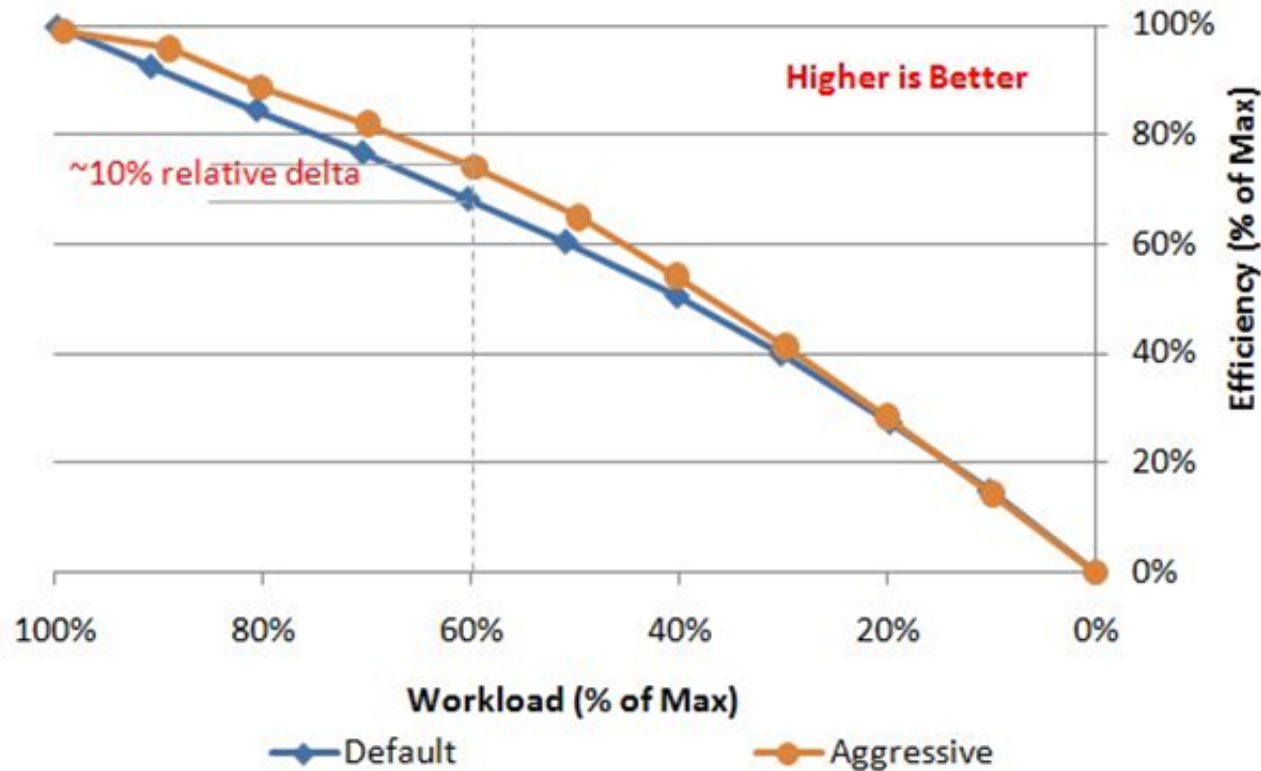
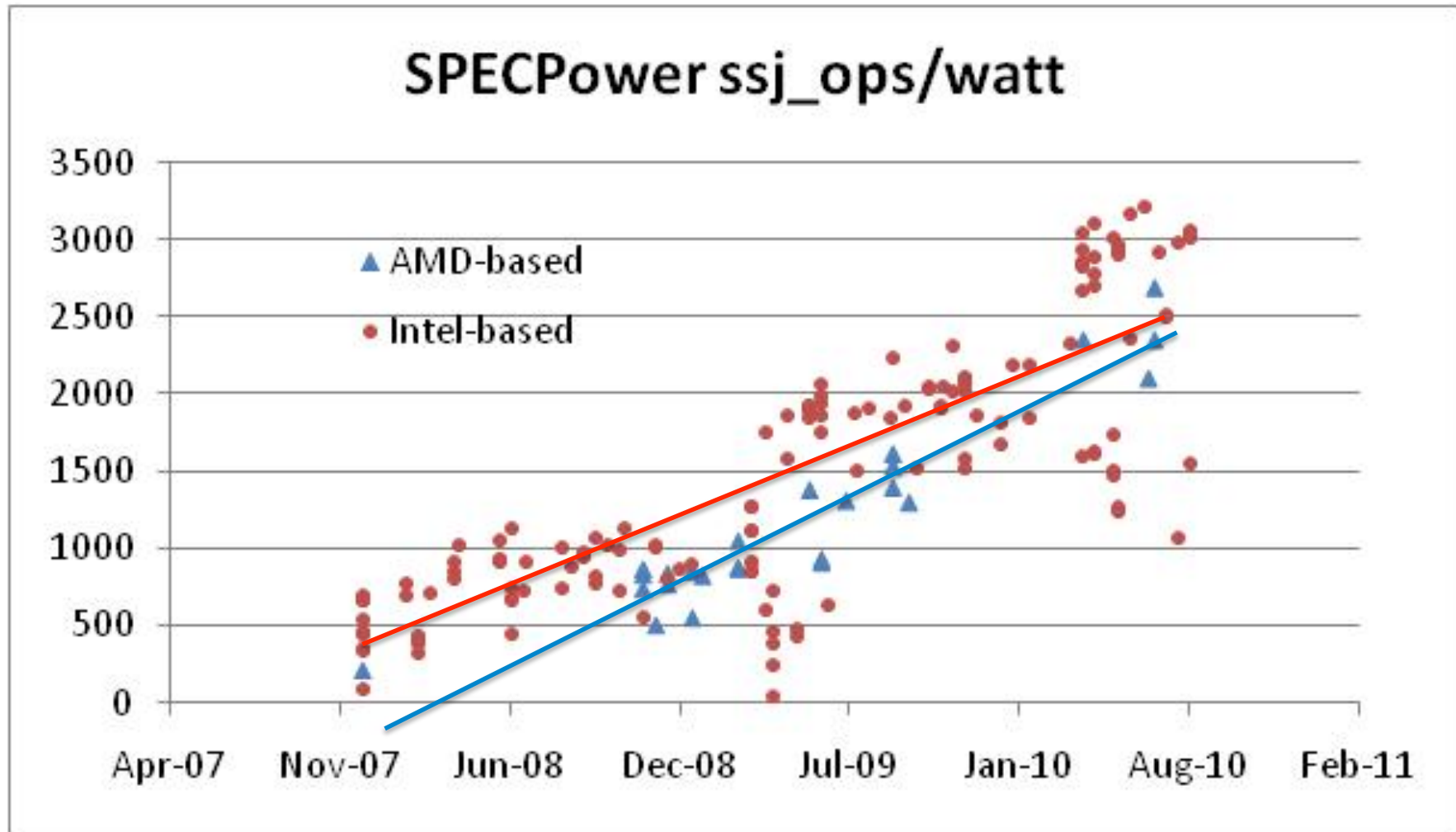# Some data about the SPECPower data

- All data available at http://spec.org/power
- ~175 systems "officially" reported (Q107-Q310)
  - Data dominated by Intel systems (~125)
  - AMD-based Systems: ~50% HP systems
  - Intel-based Systems: one of everything

  Dell, HP, Fujitsu, Toshiba, IBM, NEC, Apple, etc.
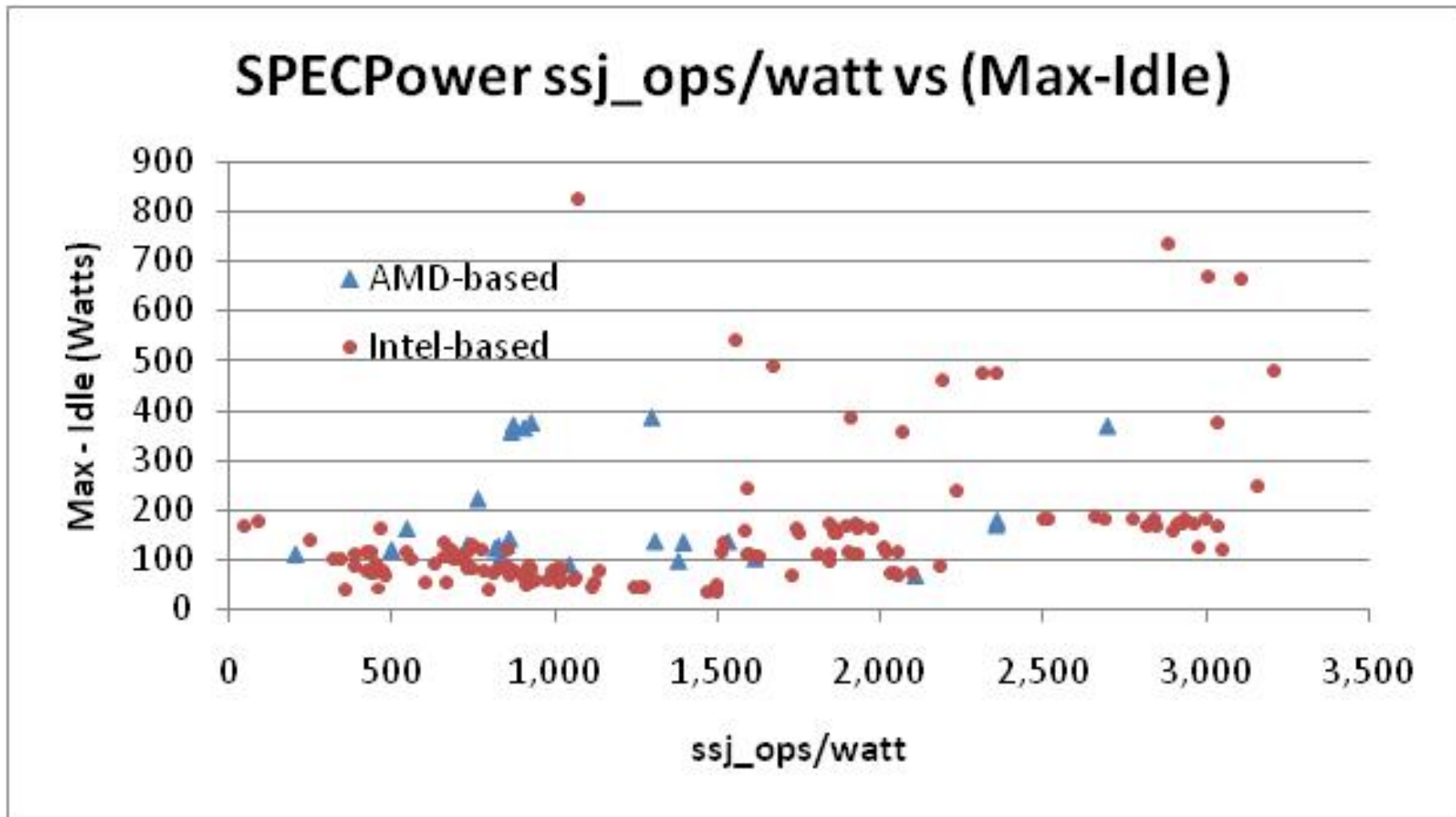- Probably not enough data for conclusions

…and now for some conclusions…

Max Power Consumption Per Node

Max Power Per Node (100% Load)
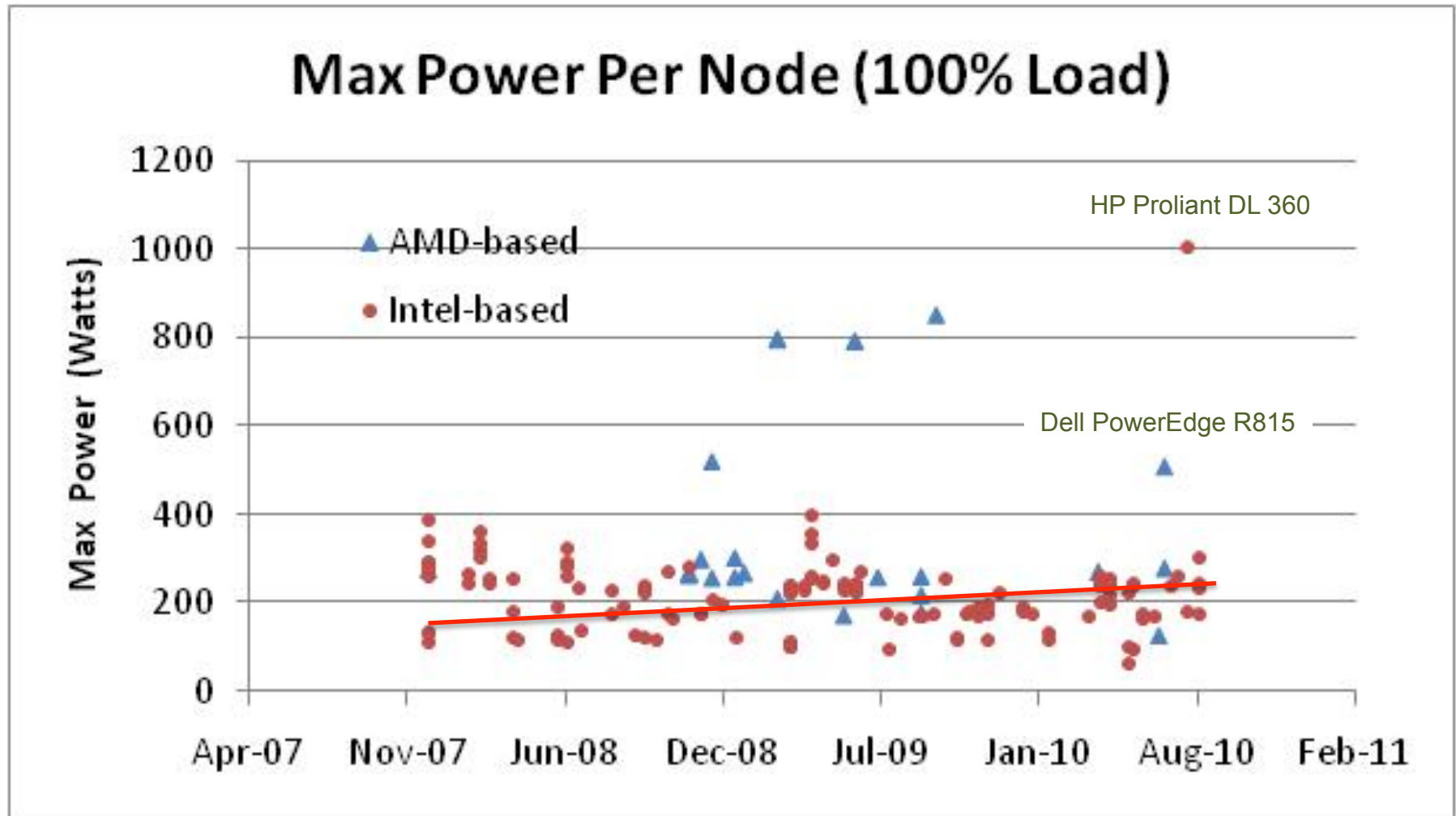
Are things getter better?

Technology trends (Max – Active Idle)

SPECPower ssj_ops/watt vs (Max-Idle)

Are more cores more energy efficient?

ssj_ops/watt vs # cores

Number of cores

△ AMD-based

• Intel-based

Dell PowerEdge R815   HP Proliant DL 170

ssj_ops/watt

Conclusions (for SPECPower)
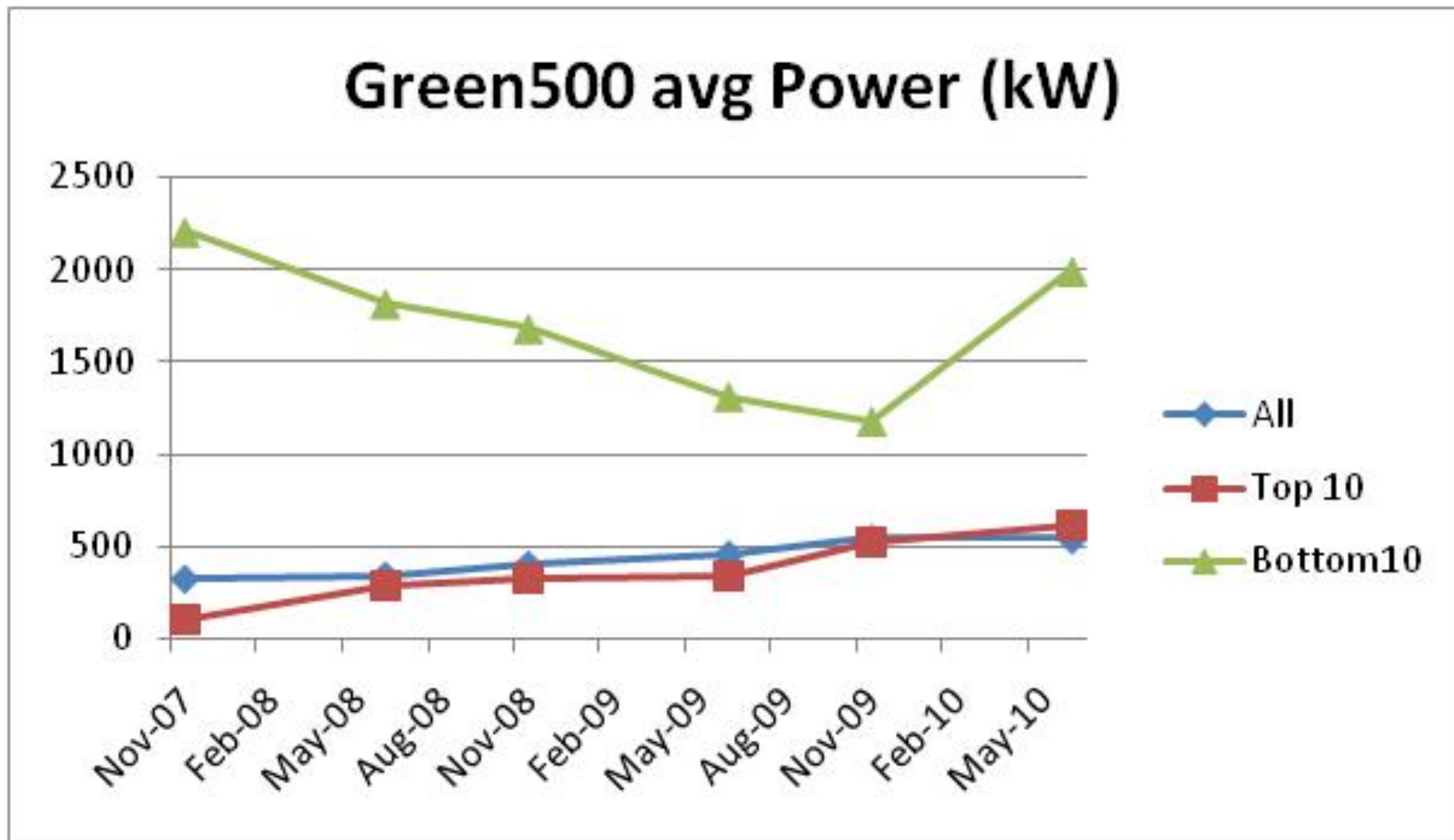
- Server efficiency is getting better
  - 5.4X increase in ssj_ops/watt since Q107
- Power consumption per node increasing
- More cores = more energy efficient (mostly)
- AMD- vs Intel-based systems
  - Intel fares better on raw numbers
  - AMD trend looks better
 - Key may be in (max – idle) + PM

The Green500

- Goal: Raise energy efficiency awareness in HPC
- Co-founded with Wu Feng in 2007
  - 6 lists so far (Nov+June since 11/07)
- Methodology
  - Linpack + simplified PowerPack for measurement
  - Focus on steady-state, nodal power and extrapolate
- This data available at http://green500.org
- Probably not enough data for conclusions

   …and now for some conclusions...

Supercomputer Power Consumption

Supercomputer Power Efficiency

Green500 avg mflops/watt
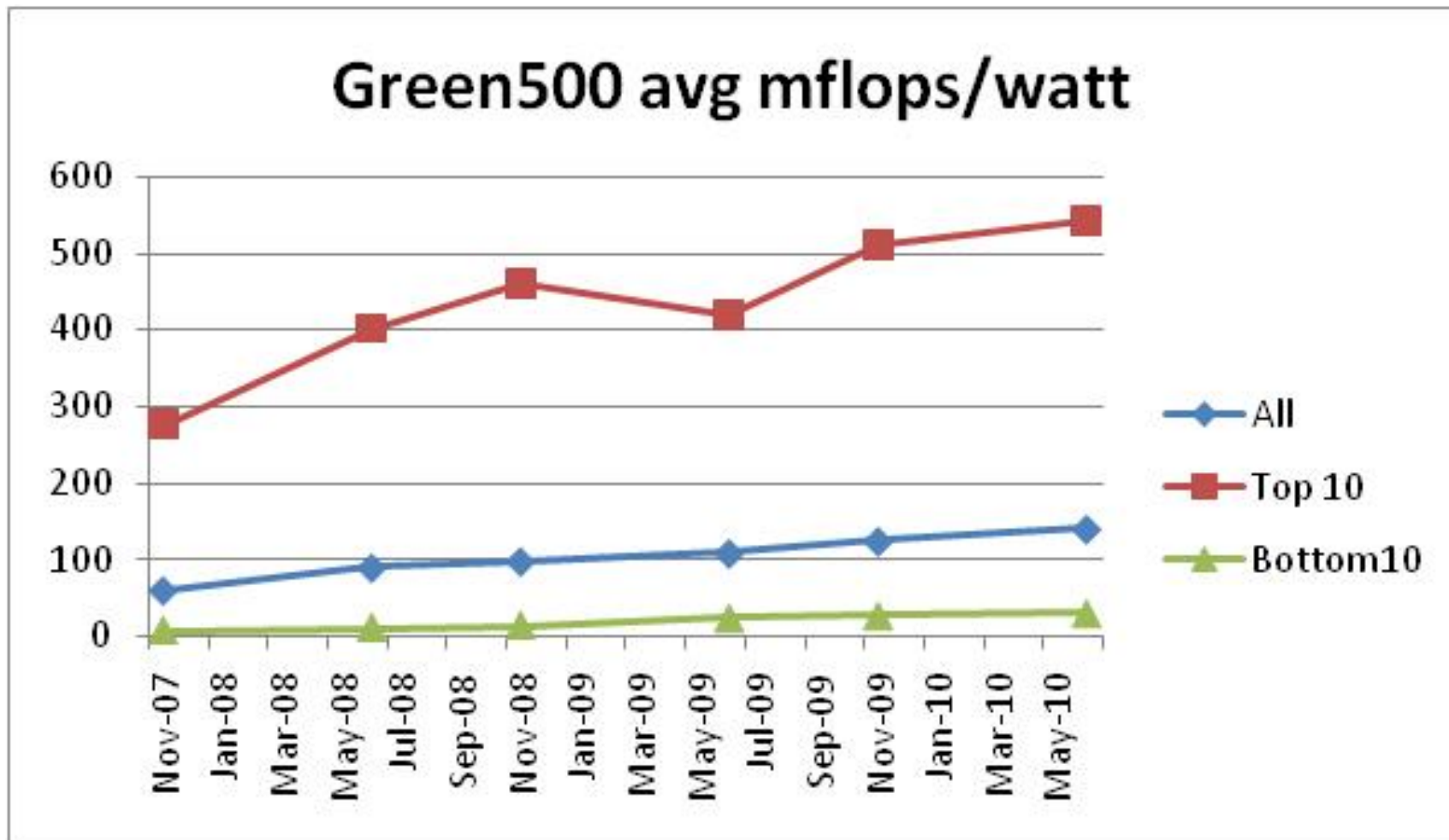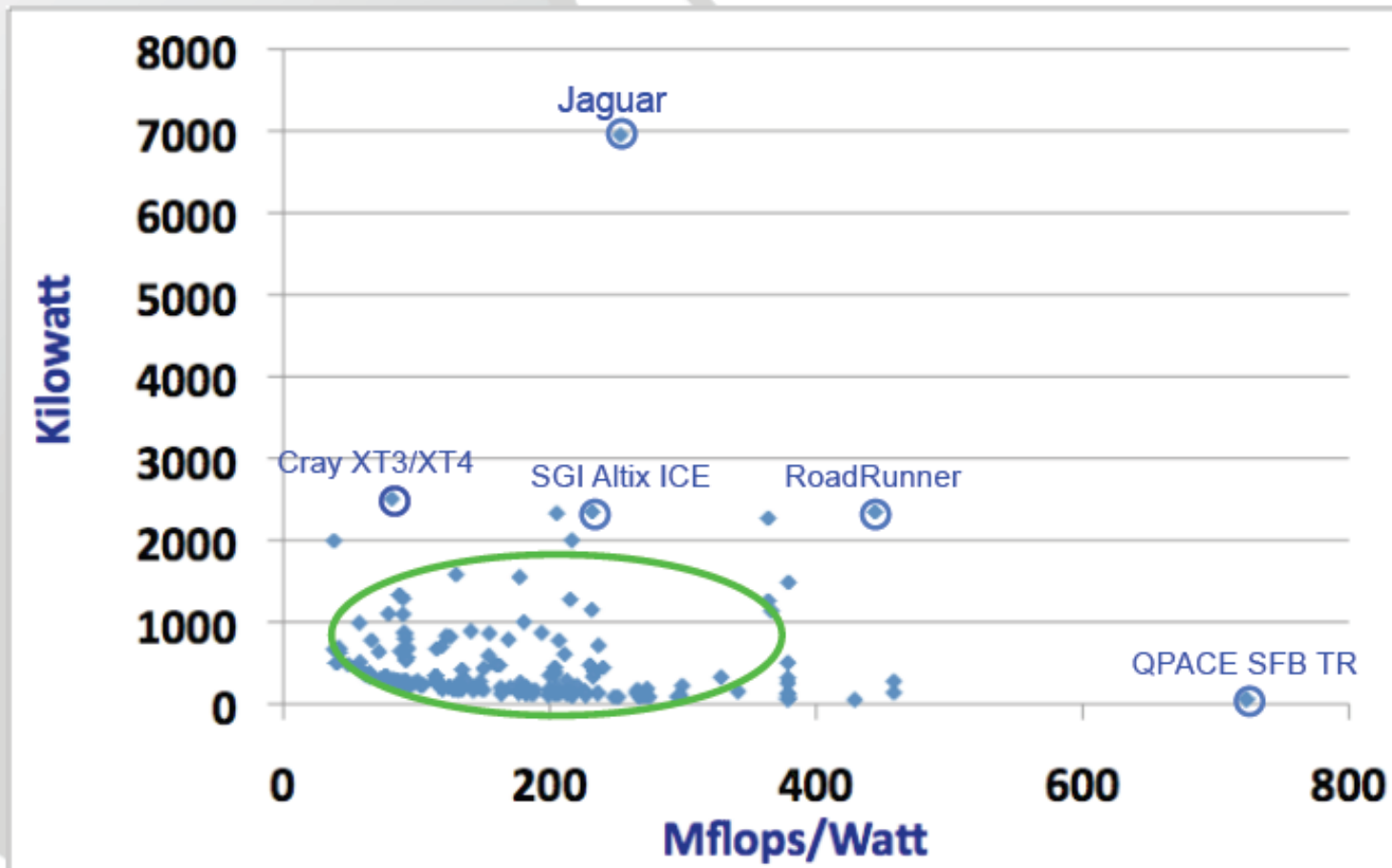
Total Power vs. Efficiency

Power Consumption vs. Mflops/Watt

Source: Wu Feng, HPPAC 2010

Conclusions (for Green500)

- Supercomputer efficiency is getting better
  - 3X increase in Flops/watt for Top10 since Nov 2007
  - #1 → 400 Mflops/watt (2007), 700 Mflops/watt (2010)
  - 1.7X increase in avg power consumption
- Accelerators = more energy efficient (mostly)
  - E.g. IBM Qpace emerges in Top 10
- Commodity system efficiency catches up with custom designs in about 18 months
  - Blue Gene/P 380 Mflops/watt in 2007
  - Intel Xeon E5540-based 380 Mflops/watt in 2009

Implications for HPC (SPECPower)

- If SPECPower trends continue
  - Implications for clusters from commodity parts
  - 28X improvement in efficiency by Q313
  - 1.6X increase in power per node from Q107
  - Avg Nodes: 320W for Q310 → 406W in Q313
  - 1,000 node cluster (~1kW more power)

Implications for HPC (Green500)

- If Green500 trends continue
  - Implications for custom designs + cluster designs
  - Top 10: 36X increased power by 2013 since 2007
  - Avg: 2.8X increased power by 2013 since 2007
  - Top 10: 3.8X increased efficiency by 2013 since 2007
  - Avg: 5.4X increased efficiency by 2013 since 2007
  - Green500 Top10 Supercomputer in 2013
    - 3.6 MW power consumption
    - 1450 Mflops/watt

Future Work

- SPECPower
  - Extension to SERT for Energy Star Ratings
  - Extension of methodology throughout SPEC
- Green500
  - Little Green 500 (smaller supercomputers)
  - Open Green500
  - HPCC Green500
  - New benchmarks

Thanks.

Kirk W. Cameron

cameron@vt.edu

http://scape.cs.vt.edu
http://grano.la
http://green500.org
http://spec.org/power