Discovering Energy Resource Usage Patterns on Scientific Clusters

Matthew Bae

Harvey Mudd College

Introduction

- Power consumption has been a large concern in supercomputing within the past decade
- Currently addressed by techniques such as dynamic voltage and frequency scaling
- Identify patterns of HPC jobs that consume different amounts of electricity in the context of system resources

Motivation

- Discovering resource usage patterns can be conducted by monitoring performance from scientific clusters.
- Observations done on SLURM, a widely used, open-source job scheduler.
- Data from Cori, LBNL's Cray XC40 supercomputer

Data

- Cori had 1630 compute nodes, each with 32 cores.
- Generated from 5951 jobs from May 2016
- Preprocessed through Python and Apache Spark

Results

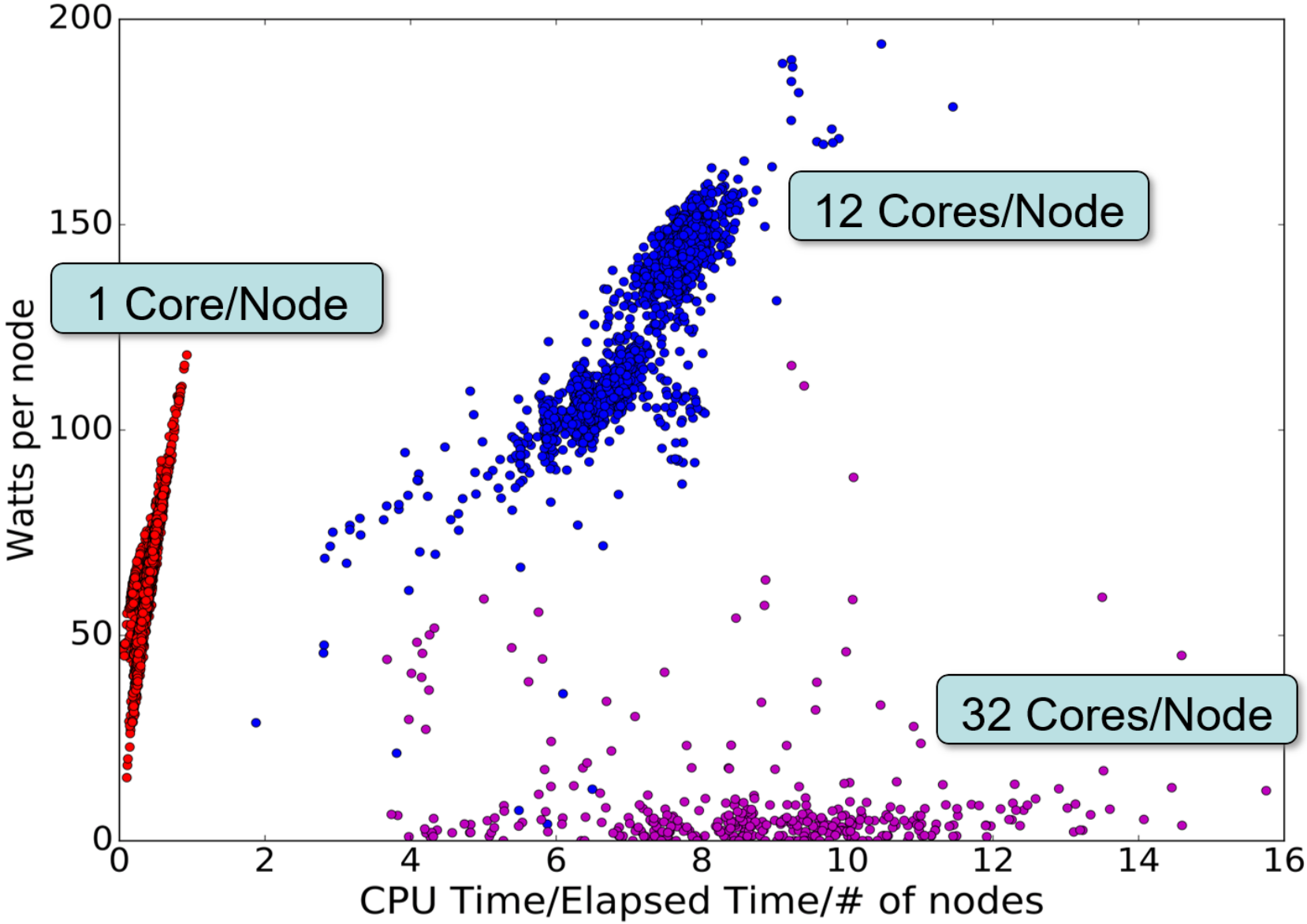

Figure 1: Jobs that use 80 nodes from 1 user

Tables and Figures

Cluster	Elapsed Time (s)	AveCPU Time (s)	Energy/Node (J)
Red	8.79e0	4.65e-2	5.47e2
Blue	2.44e1	1.37e1	3.222e3
Purple	1.23e3	3.35e2	6.74e3

Cluster	AveRSS/Node (b)	AvePages/Node	AveCPUFreq
Red	5.16e4	4.56e4	1.42e9
Blue	6.11e3	1.41e6	2.05e9
Purple	1.14e6	2.26e6	5.84e8

Table 1: Partial table of data for Figure 1

Results

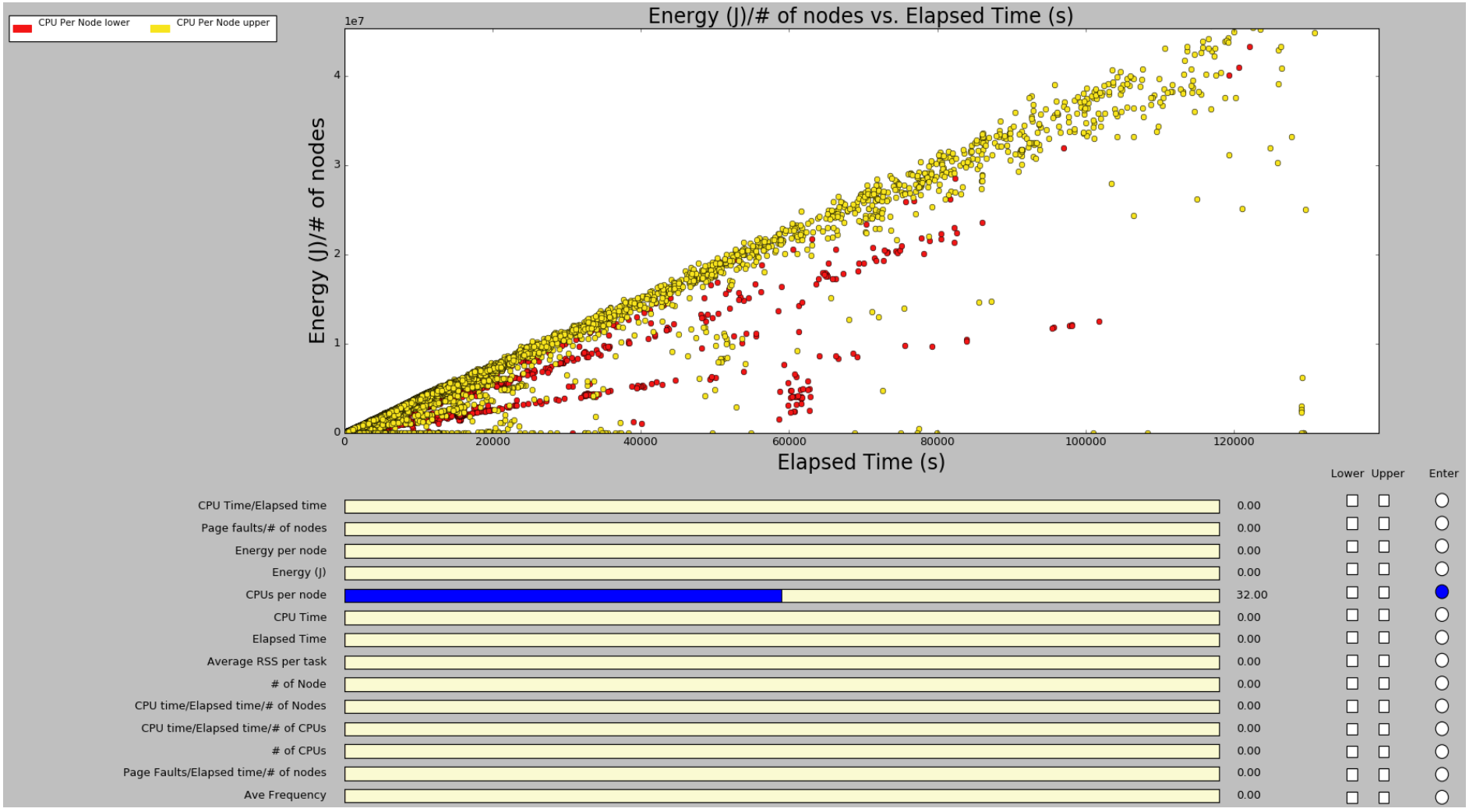

Figure 2: Use of interactive plot on Energy/Node vs. Elapsed Time

Results

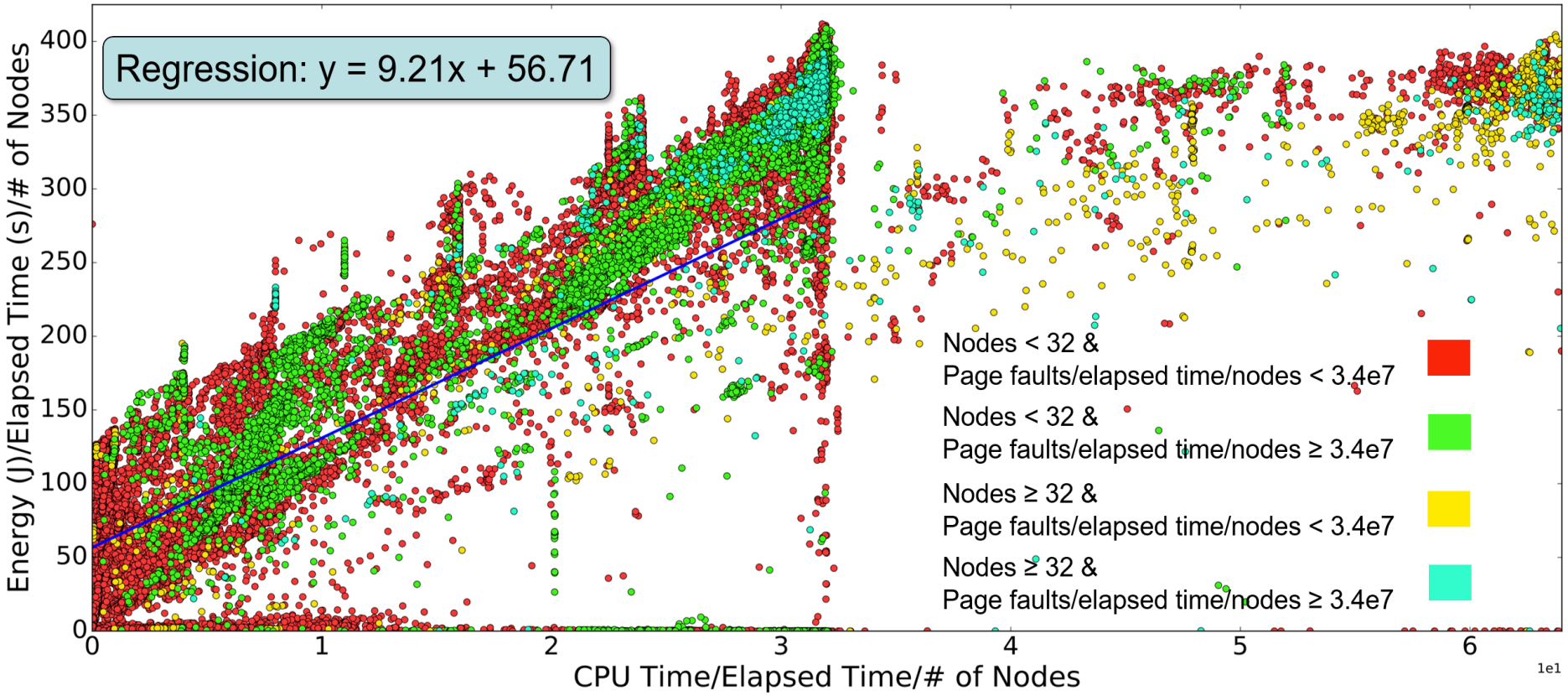

Figure 3: Regression on an interactive plot of WPN vs. CPU workload

Conclusion

- Monitoring energy performance in relation to other resources can lead to discovering energy usage patterns.
- Energy consumption patterns arise based on different variables such as CPU load and CPU utilization.
- Differences in WPN shows potential in energy savings.

Future Work

- Analyzing spread of WPN values and understand causes of low WPN values
- Optimizing the number of CPUs per node depending on the resource usage patterns of job executions
- Develop suggestions to allow users and those maintaining the supercomputers to conserve energy

Thank you