

Using MuMMI to Model and Optimize Energy and Performance

Xingfu Wu and Valerie Taylor
Texas A&M University

Scalable Tools Workshop 2015, Lake Tahoe, CA
August 3, 2015

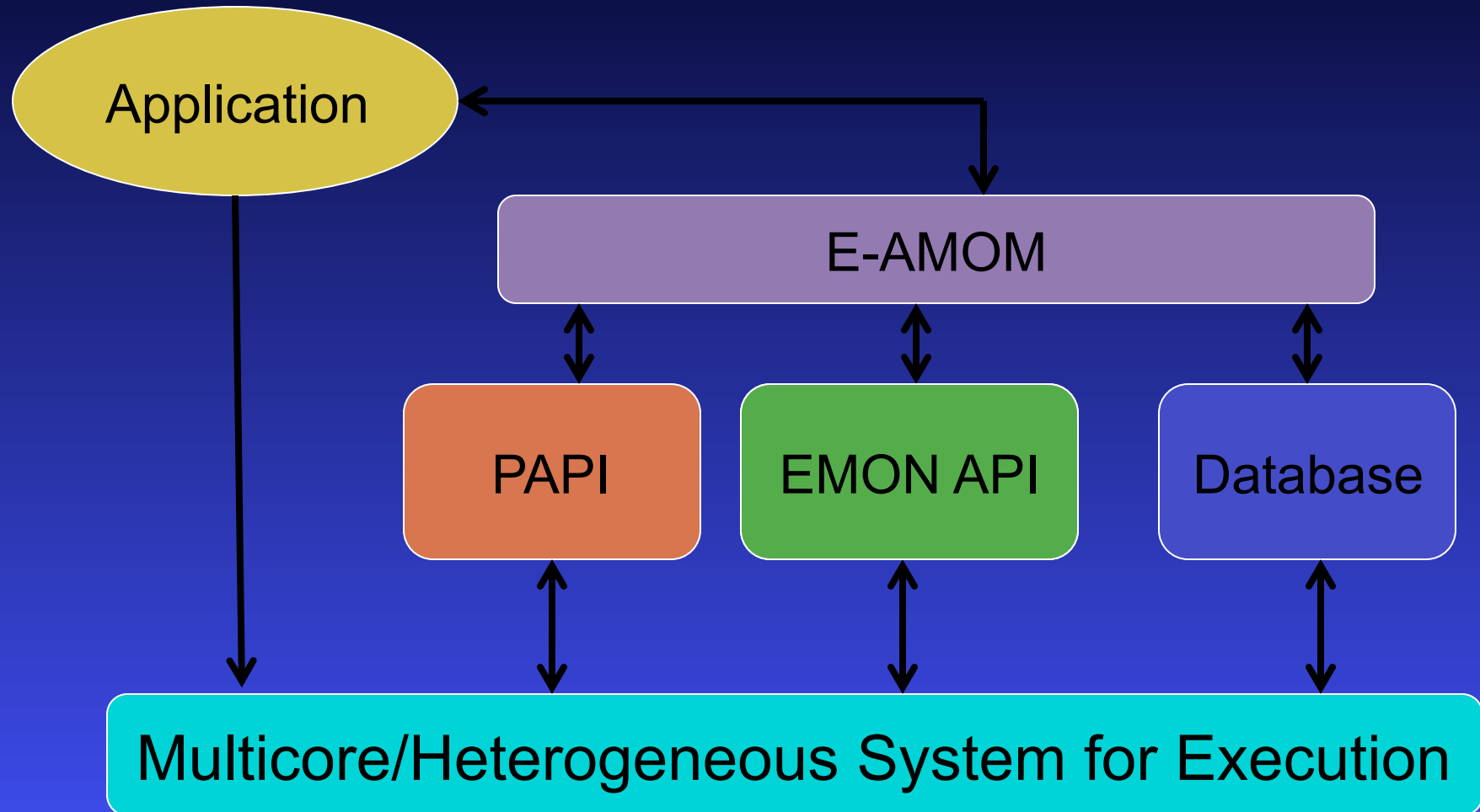


<http://www.mummi.org>

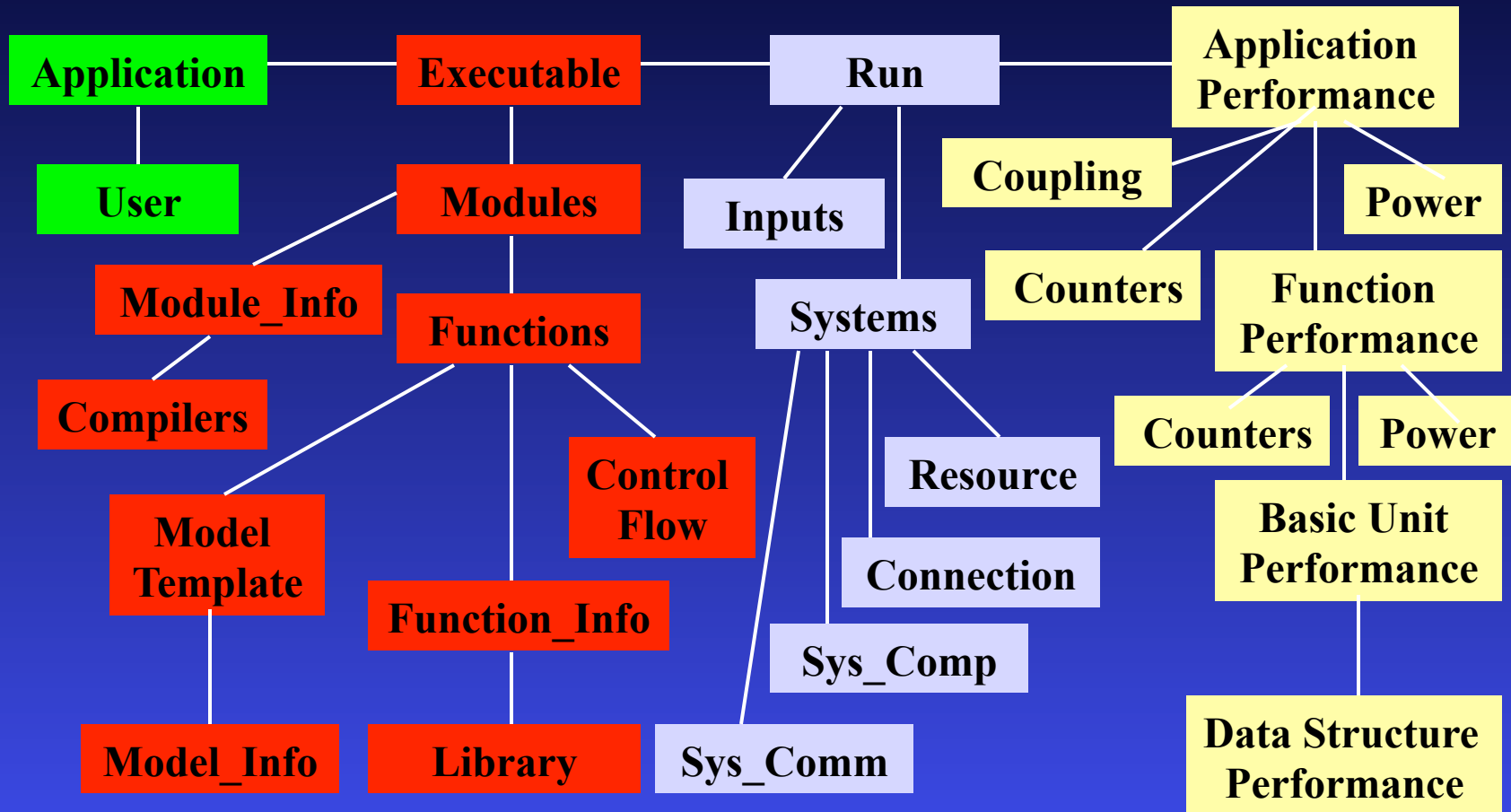
Outline

- Recent Development of MuMMI
- Power Measurement Tools
- Hardware Performance Counter Tools
- Performance Counter-based Modeling
- Performance Counter-guided Optimization

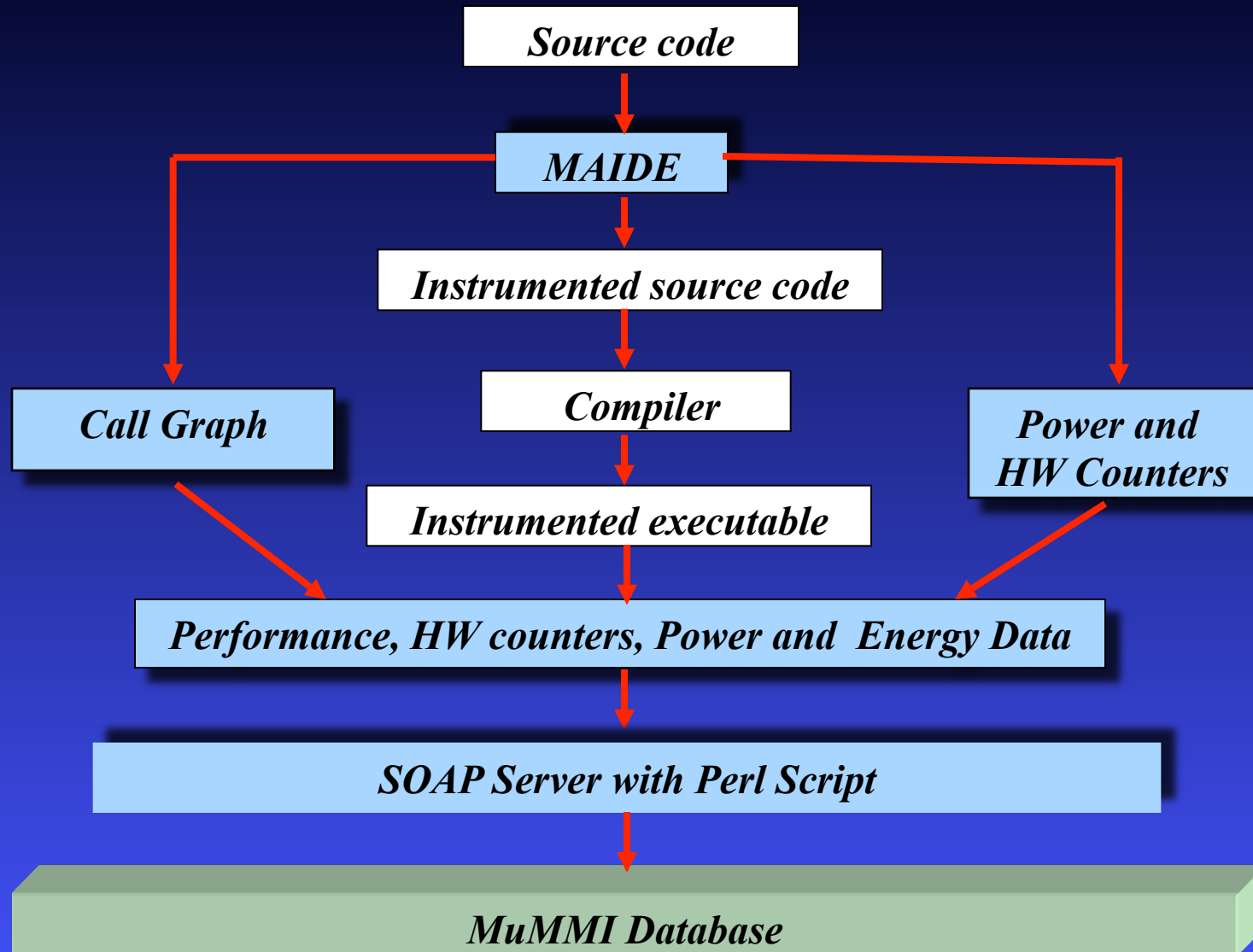
MuMMI (Multiple Metrics Modeling Infrastructure) Project



MuMMI Database Schema



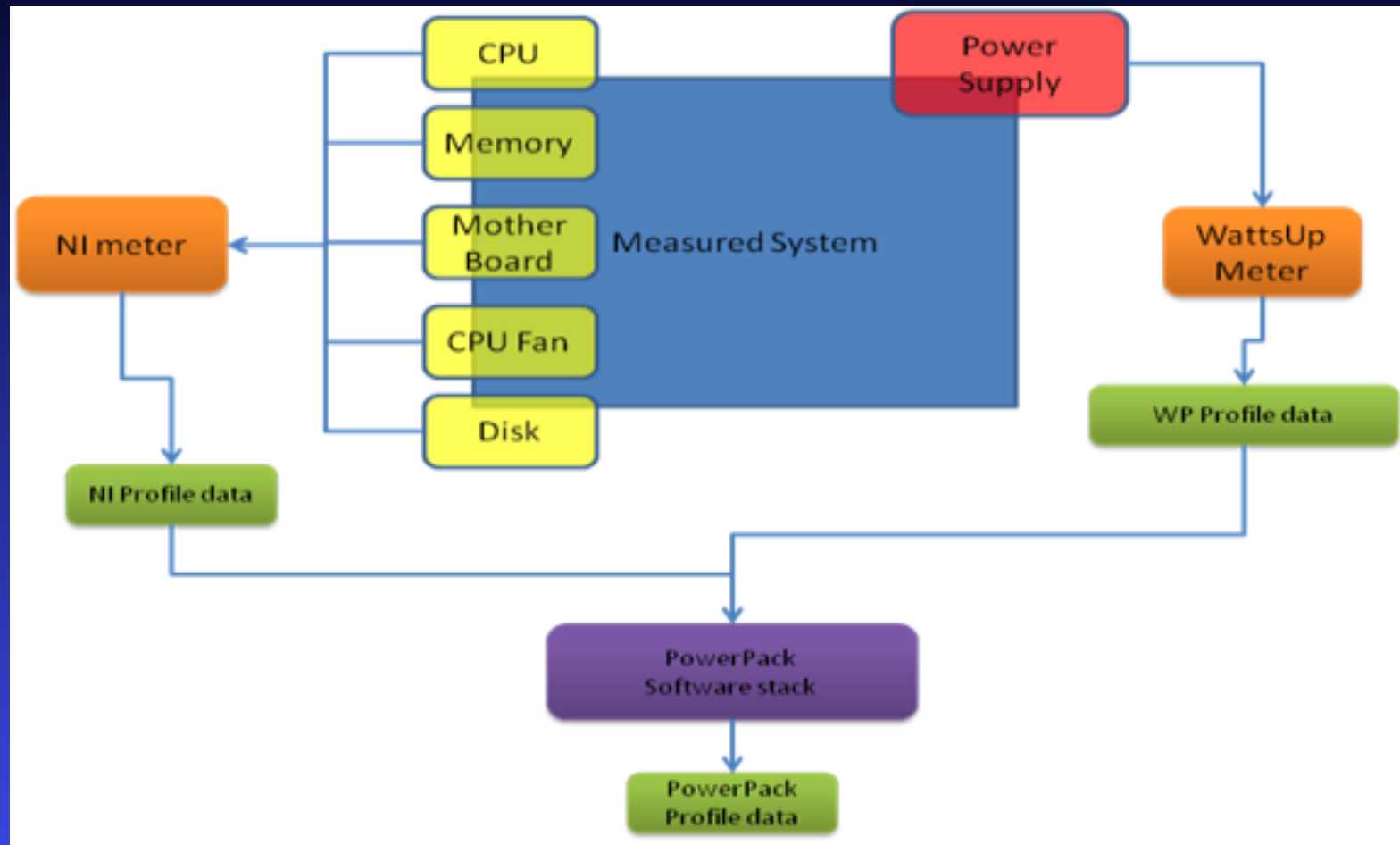
Data Collection: MAIDE System



Power Measurement Tools

- IBM EMON API on BlueGene P/Q (MonEQ)
- Intel RAPL
- NVIDIA MLPM
- PowerMon2 (RENCI)
- PowerPack (VT)

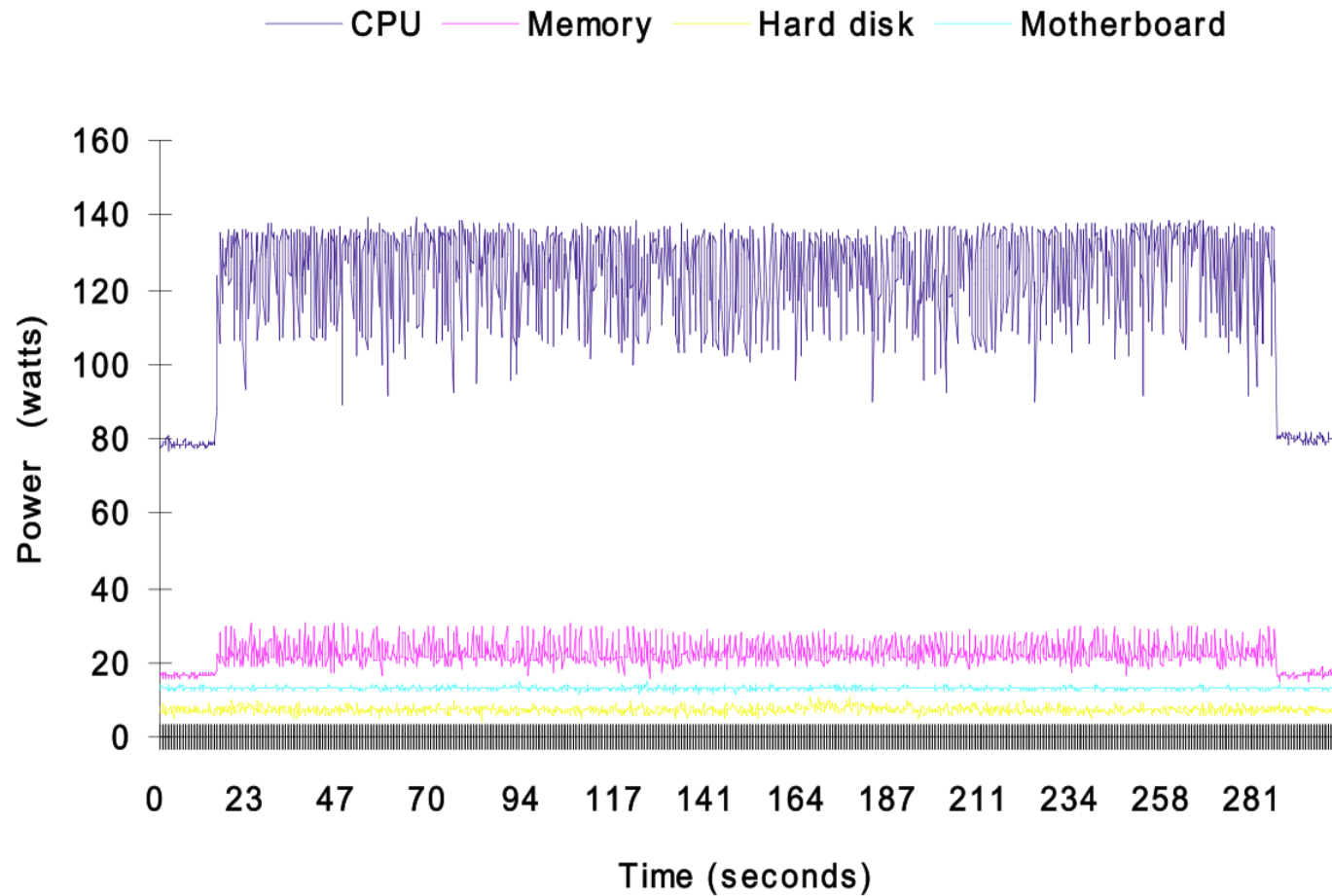
PowerPack Schema (Virginia Tech)



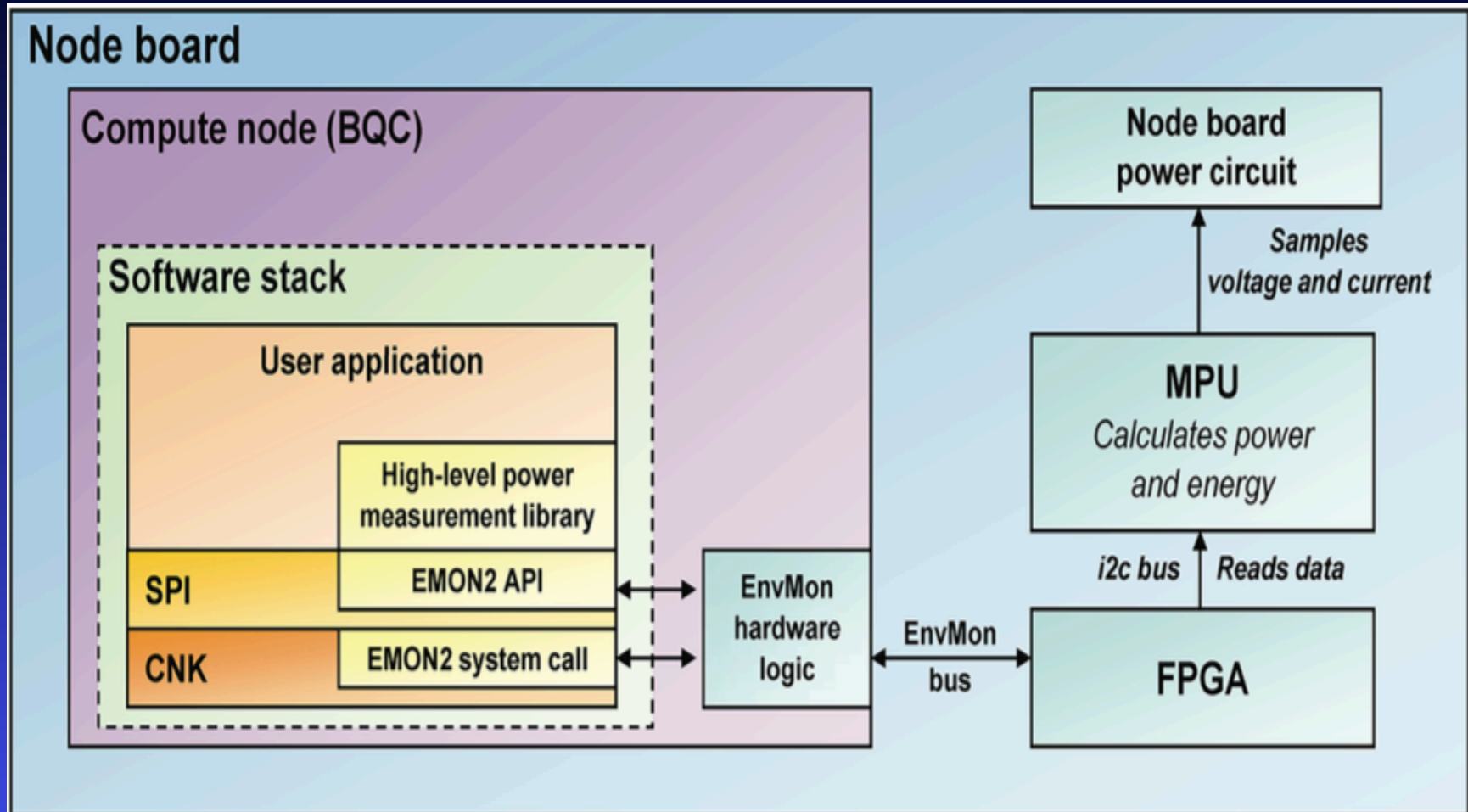
Power sampling frequency: 1 sample per second

PowerPack

MPI BT running on 1 node with 4 cores/node

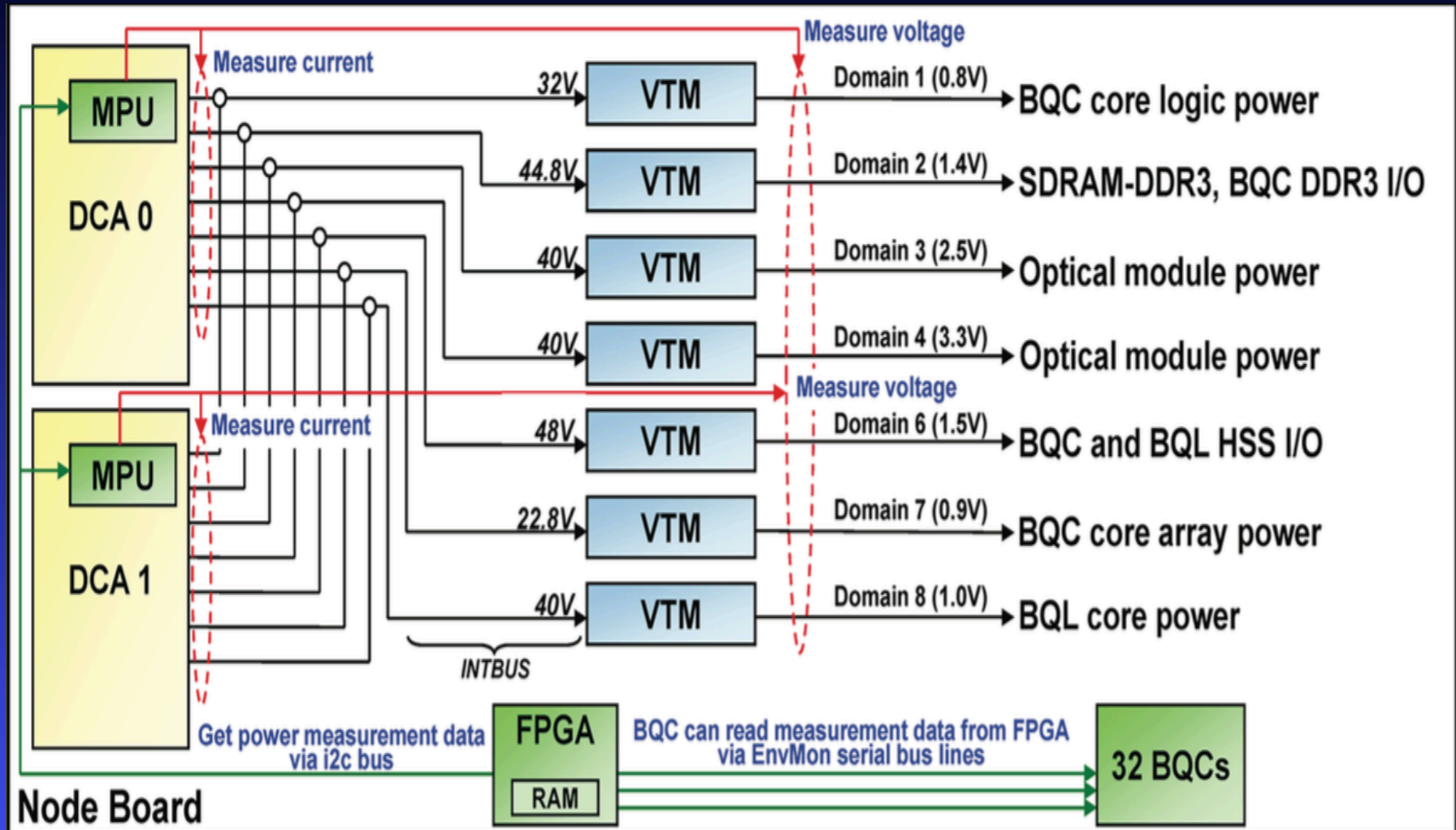


IBM EMON API



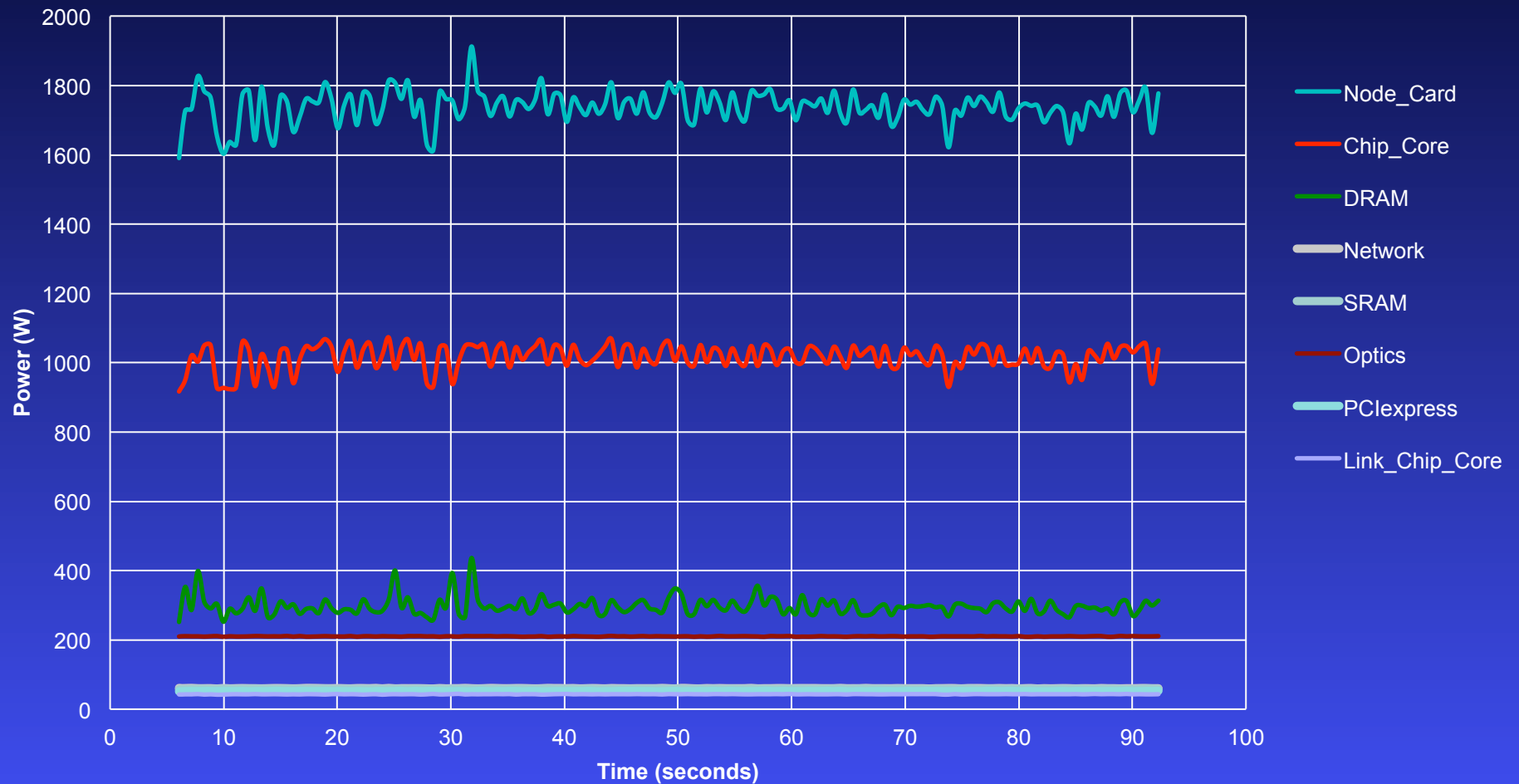
Power sampling frequency: ~2 samples per second

IBM EMON API



IBM EMON API

Power per node card for GTC on 16384 nodes of ANL BGQ Mira



IBM EMON API

Average Power per node for GTC on 16384 nodes of ANL BGQ Mira



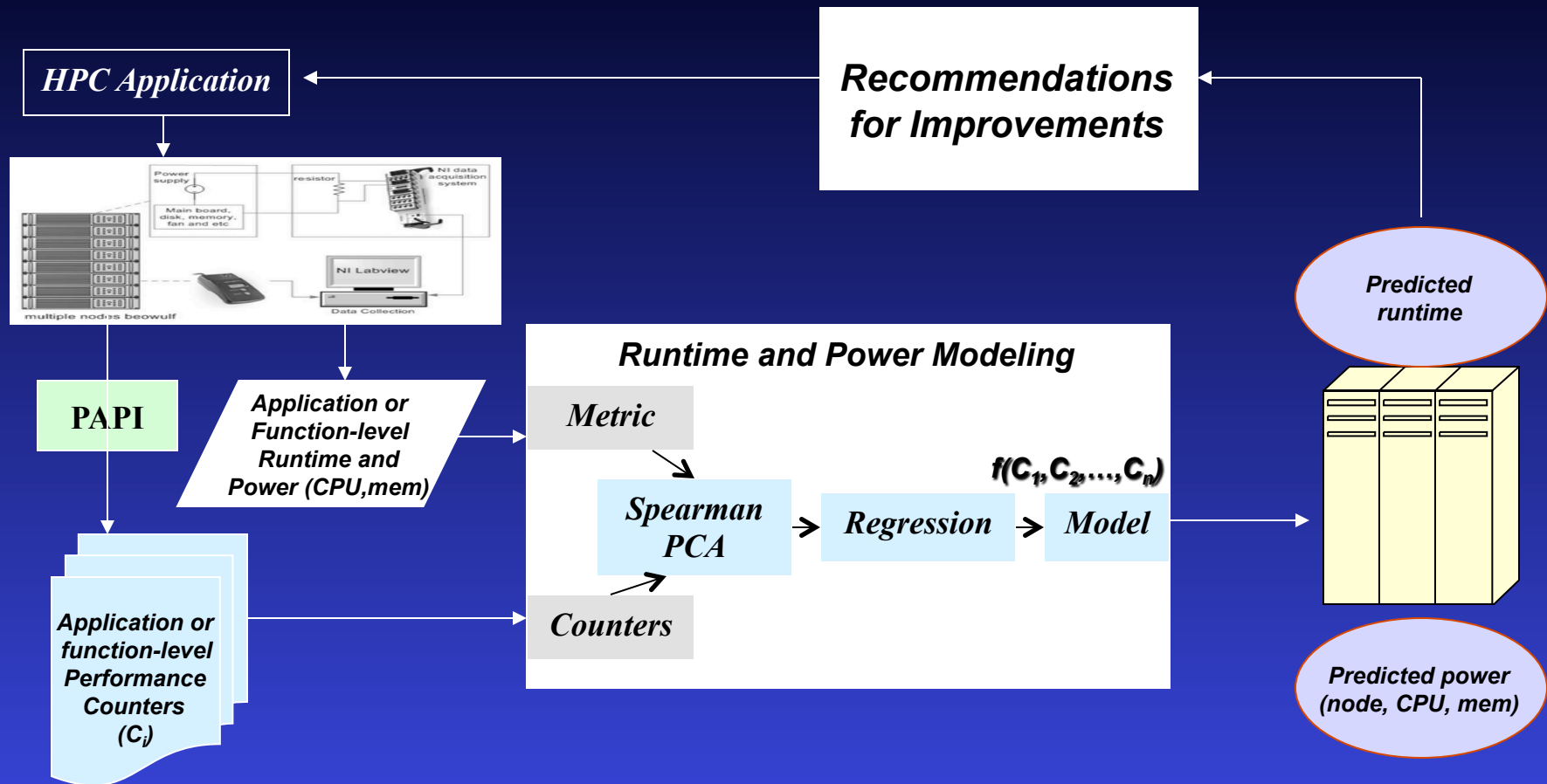
Performance Counter Tools

- Perf_events (Linux)
- HPM (IBM)
- perfmon (Linux)
- PAPI (UTK)

Outline

- Recent Development of MuMMI
- Power Measurement Tools
- Hardware Performance Counter Tools
- Performance Counter-based Modeling
- Performance Counter-guided Optimization

Overview: Performance Counter-based Modeling



Four metrics: runtime, node power, CPU power, memory power

Four Models of Parallel eq3dyna on SystemG

The screenshot shows the MuMMI website interface. At the top, there's a navigation bar with 'MuMMI', 'Home', 'Browse', and 'About'. Below this, there's a section for 'Counter' and 'Percentage Change' with a 'What-If' button. A row of tabs includes 'Modeling', 'Ranking', 'Predictions', 'Prediction Error Rate', 'Suggestions', 'Stacked Predictions', 'Aligned Predictions', 'Graph', and 'Prediction Graphs'. Below these are 'Giant Scatterplot' and 'Relevant Scatterplot' buttons, with a 'Saved' status indicator. The main heading is 'eq3dyna 1.0; eqdyna3drun-hybrid-original Model Coefficients'. Below this, there's a 'Show' dropdown set to '25' and a search bar. The main content is a table with 6 columns: Counter, runtime, power_sys, power_cpu, power_mem, and Description. The table lists various performance metrics and their values. At the bottom, it says 'Showing 1 to 15 of 15 entries' and has navigation buttons for 'Previous', '1', and 'Next'.

Counter	runtime	power_sys	power_cpu	power_mem	Description
Frequency	3.9416004e-10	41.56653	26.452081	4.3977528	Frequency
PAPI_BR_NTK	7.6851559e-11				Conditional branch instructions not taken
PAPI_FDV_INS				2470.4353	Floating point divide instructions
PAPI_FML_INS			10.692587		Floating point multiply instructions
PAPI_FP_INS	4.5223155e-10	104.25188	13.946254		Floating point instructions
PAPI_L1_ICM	6.5852623e-06	1317108.0	563801.12	333958.76	Level 1 instruction cache misses
PAPI_L1_STM		3347.641		3210.2844	Level 1 store misses
PAPI_L1_TCA		57.274734		24.390351	Level 1 total cache accesses
PAPI_L2_DCW	3.7509496e-07	20716.209	4464.7716	3030.2913	Level 2 data cache writes
PAPI_L2_ICA	3.5201417e-06				Level 2 instruction cache accesses
PAPI_L2_STM				13629.208	Level 2 store misses
PAPI_L2_TCW		3347.641			Level 2 total cache writes
PAPI_RES_STL		104.57014	52.800714	35.363755	Cycles stalled on any resource
PAPI_SR_INS				20.396028	Store instructions
PAPI_TLB_DM		1016.2855	7921.7689	327.73056	Data translation lookaside buffer misses

Four Models of Parallel eq3dyna on ANL Mira

The screenshot shows the MuMMI web application interface. At the top, there's a navigation bar with 'Home', 'Browse', and 'About' links. Below this, a 'Counter' section allows selecting an option and viewing percentage change. A 'What-If' button is also present. A row of tabs includes 'Modeling' (active), 'Ranking', 'Predictions', 'Prediction Error Rate', 'Suggestions', and 'Stacked Predictions'. Below these are more specific tabs: 'Aligned Predictions', 'Graph', 'Prediction Graphs', 'Giant Scatterplot', 'Relevant Scatterplot', and a 'Saved' button. The main heading is 'eq3dyna 2.3; ANL BGQ Mira' followed by 'Model Coefficients'. A 'Show 10 entries' dropdown and a search bar are located above a table. The table has columns for Counter, runtime, power_sys, power_cpu, power_mem, and Description. It lists 9 entries related to various hardware and instruction metrics. At the bottom, it shows 'Showing 1 to 9 of 9 entries' and navigation buttons for 'Previous', '1', and 'Next'.

Counter: Select an Option, Percentage Change: %, What-If

Modeling, Ranking, Predictions, Prediction Error Rate, Suggestions, Stacked Predictions

Aligned Predictions, Graph, Prediction Graphs, Giant Scatterplot, Relevant Scatterplot, Saved

eq3dyna 2.3; ANL BGQ Mira

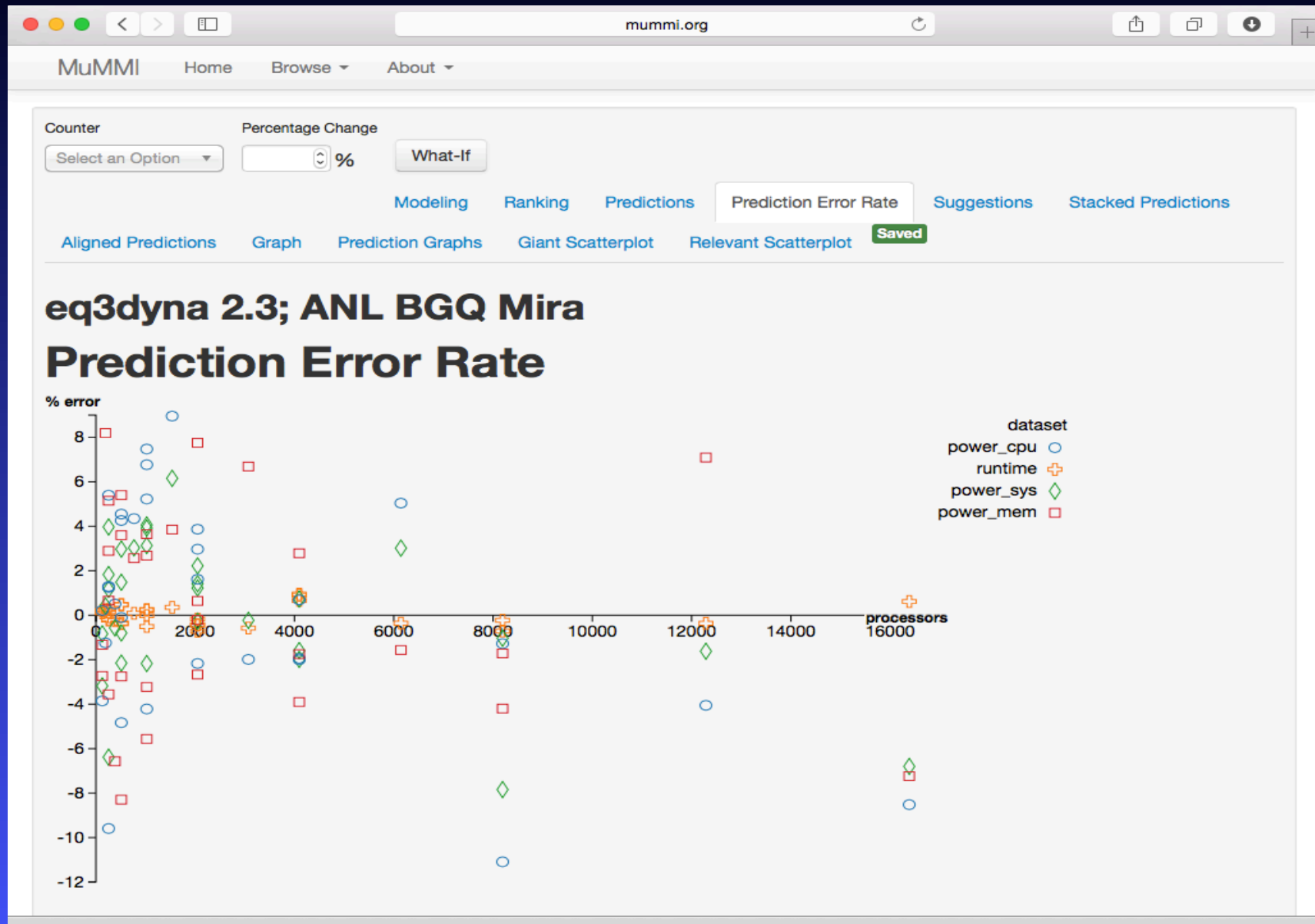
Model Coefficients

Show 10 entries, Search:

Counter	runtime	power_sys	power_cpu	power_mem	Description
Frequency	9.1191088e-11	25.35867	13.126316	0.02118797	Frequency
PAPI_BR_MSP	5.0554256e-09			65.966871	Conditional branch instructions mispredicted
PAPI_BR_NTK		45.484849	21.748956		Conditional branch instructions not taken
PAPI_L1_DCM				0.58197173	Level 1 data cache misses
PAPI_L1_STM	2.1072027e-10				Level 1 store misses
PAPI_LD_INS	2.9397223e-09	101.24107	96.410533		Load instructions
PAPI_RES_STL	5.4960219e-10			6.506956	Cycles stalled on any resource
PAPI_SR_INS				71.098626	Store instructions
PAPI_VEC_INS	7.5915981e-08	72921.882	68085.045		Vector/SIMD instructions (could include integer)

Showing 1 to 9 of 9 entries, Previous, 1, Next

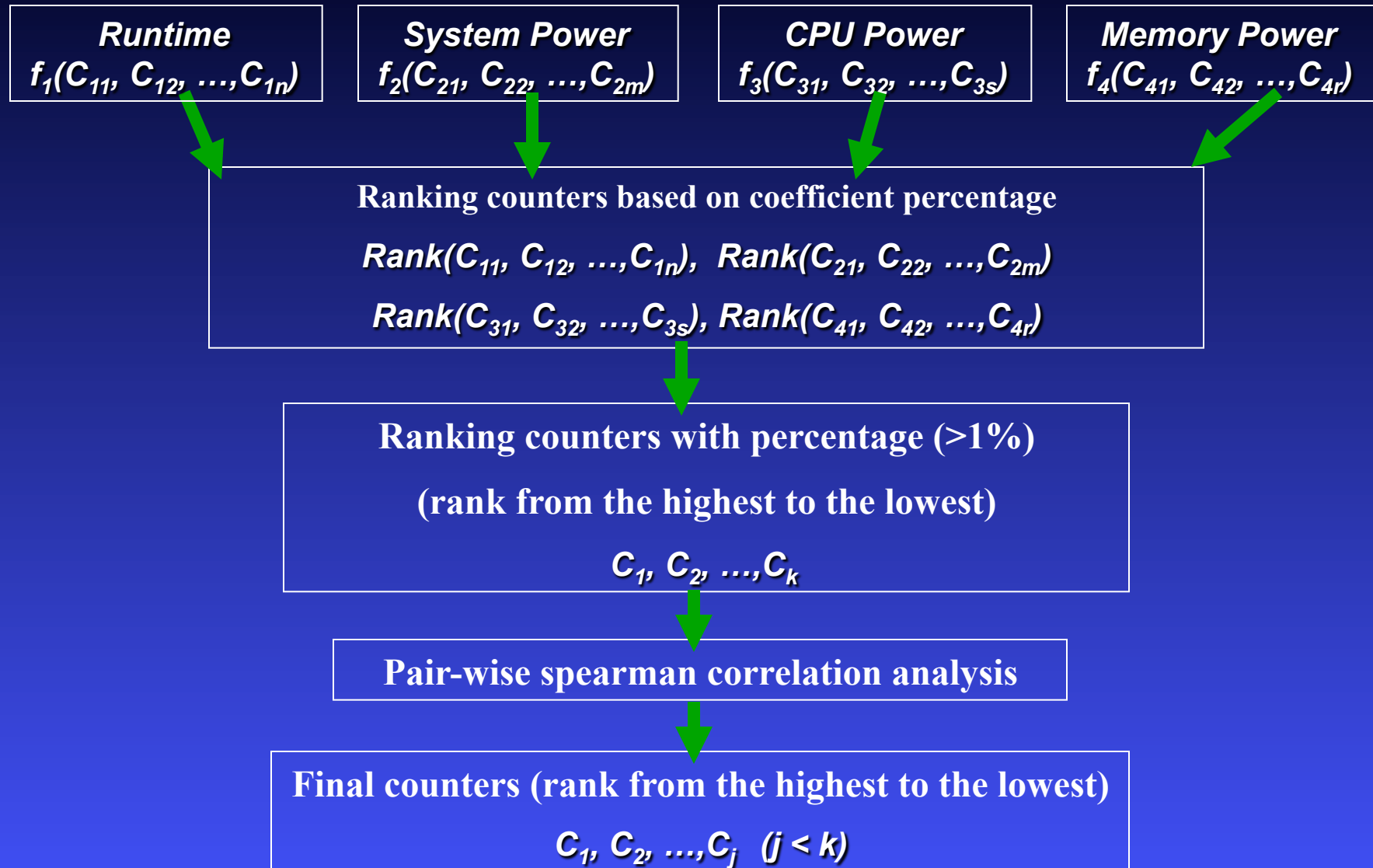
Prediction Error Rates on ANL Mira



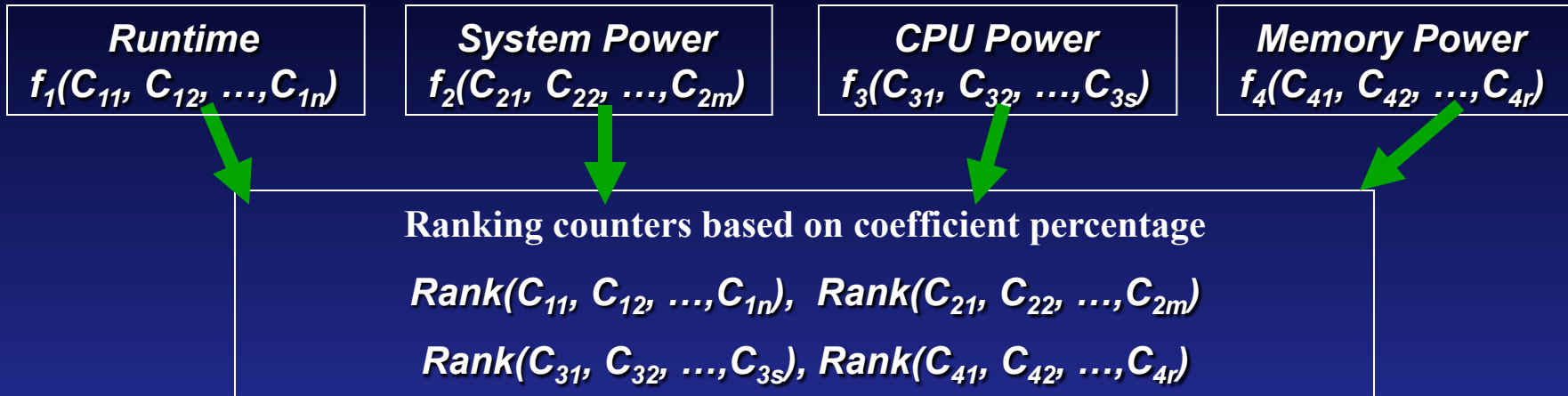
Outline

- Recent Development of MuMMI
- Power Measurement Tools
- Hardware Performance Counter Tools
- Performance Counter-based Modeling
- Performance Counter-guided Optimization

Overview: Counter-Guided Optimizations



Counter Ranking



For example, given a parallel aerospace simulation PMLB:

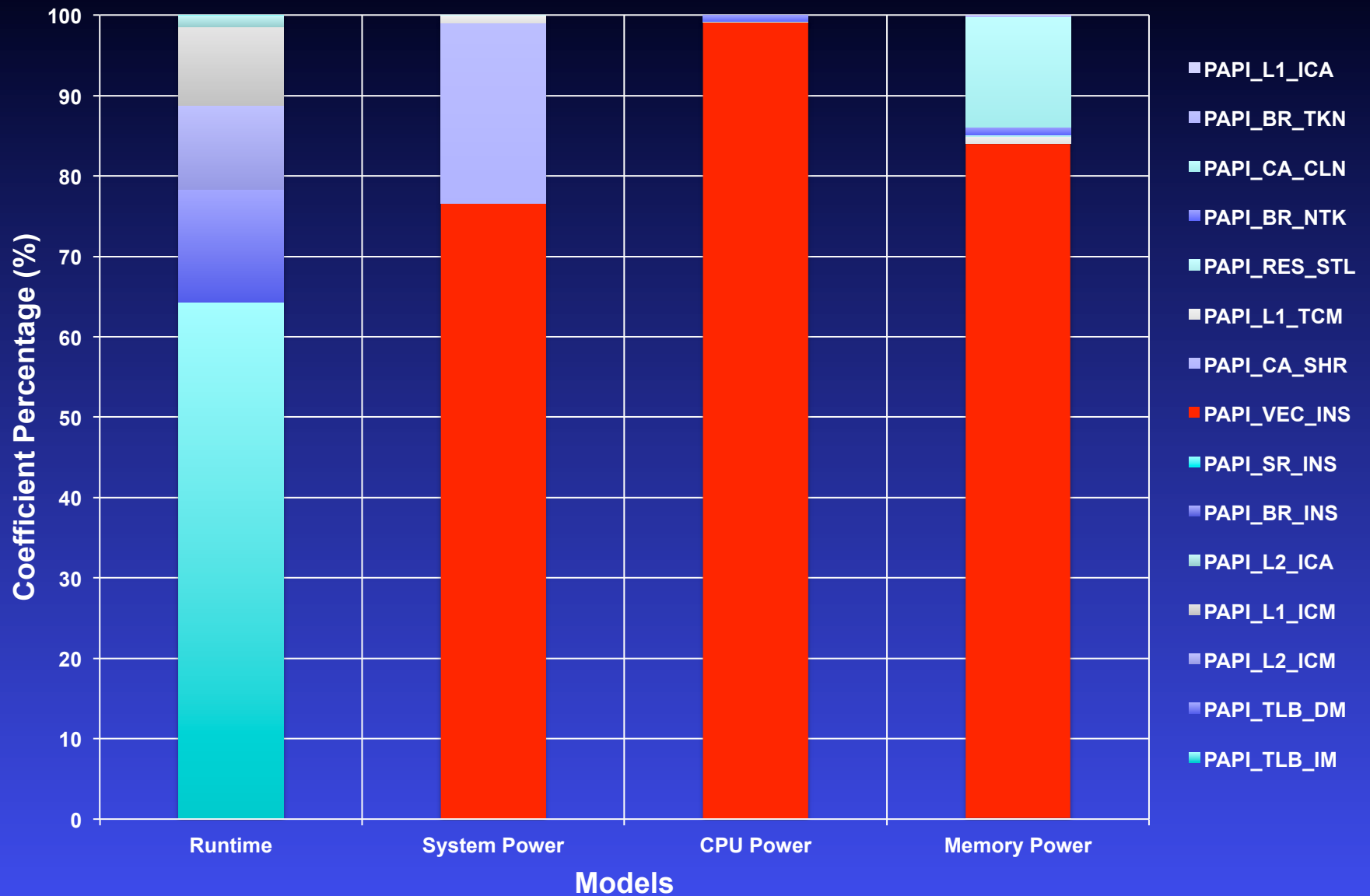
Runtime
TLB_IM: 64.29%
TLB_DM: 14.03%
L2_ICM: 10.49%
L1_ICM: 9.75%
L2_ICA: 1.40%
BR_INS: 0.03%
SR_INS: 0.01%

Node Power
VEC_INS: 76.64%
CA_SHR: 22.45%
L1_TCM: 0.89%
RES_STL: 0.02%

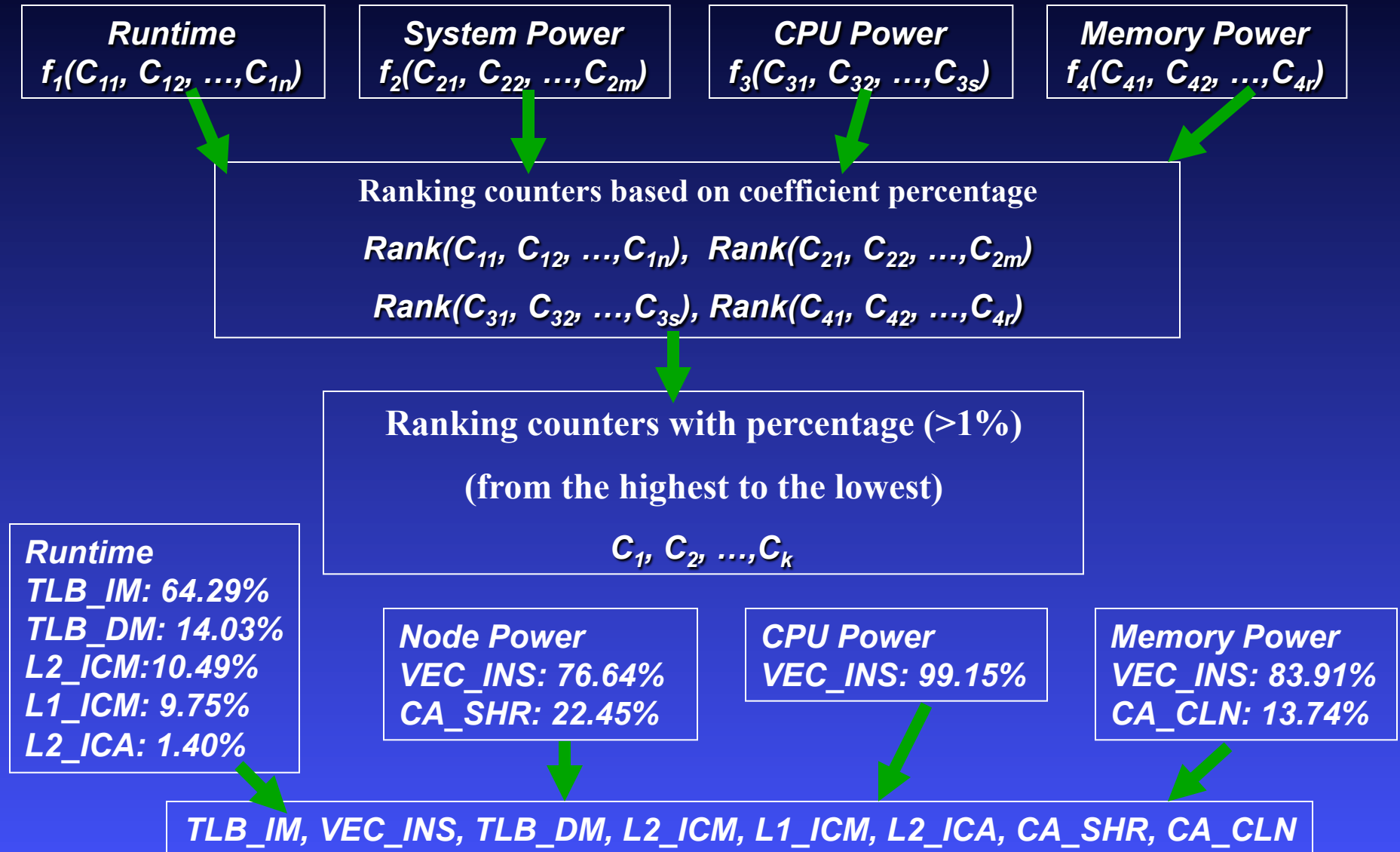
CPU Power
VEC_INS: 99.15%
BR_NTK: 0.81%
RES_STL: 0.04%

Memory Power
VEC_INS: 83.91%
CA_CLN: 13.74%
BR_NTK: 0.98%
L1_TCM: 0.92%
RES_STL: 0.18%
BR_TKN: 0.16%
L1_ICA: 0.11%

Counter Ranking for Original PMLB on SystemG



Counter Ranking

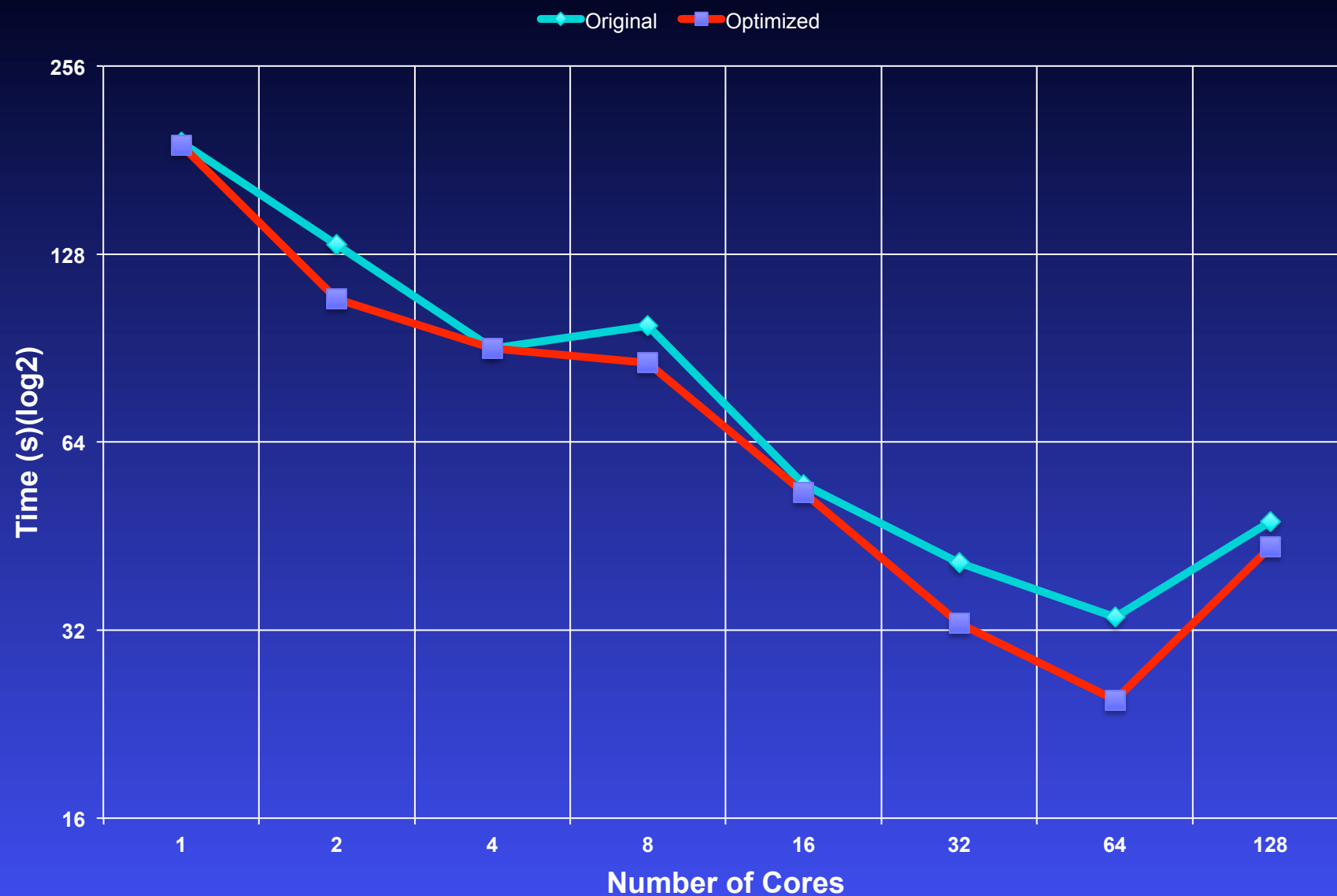


Correlation Analysis Using Pair-wise Spearman

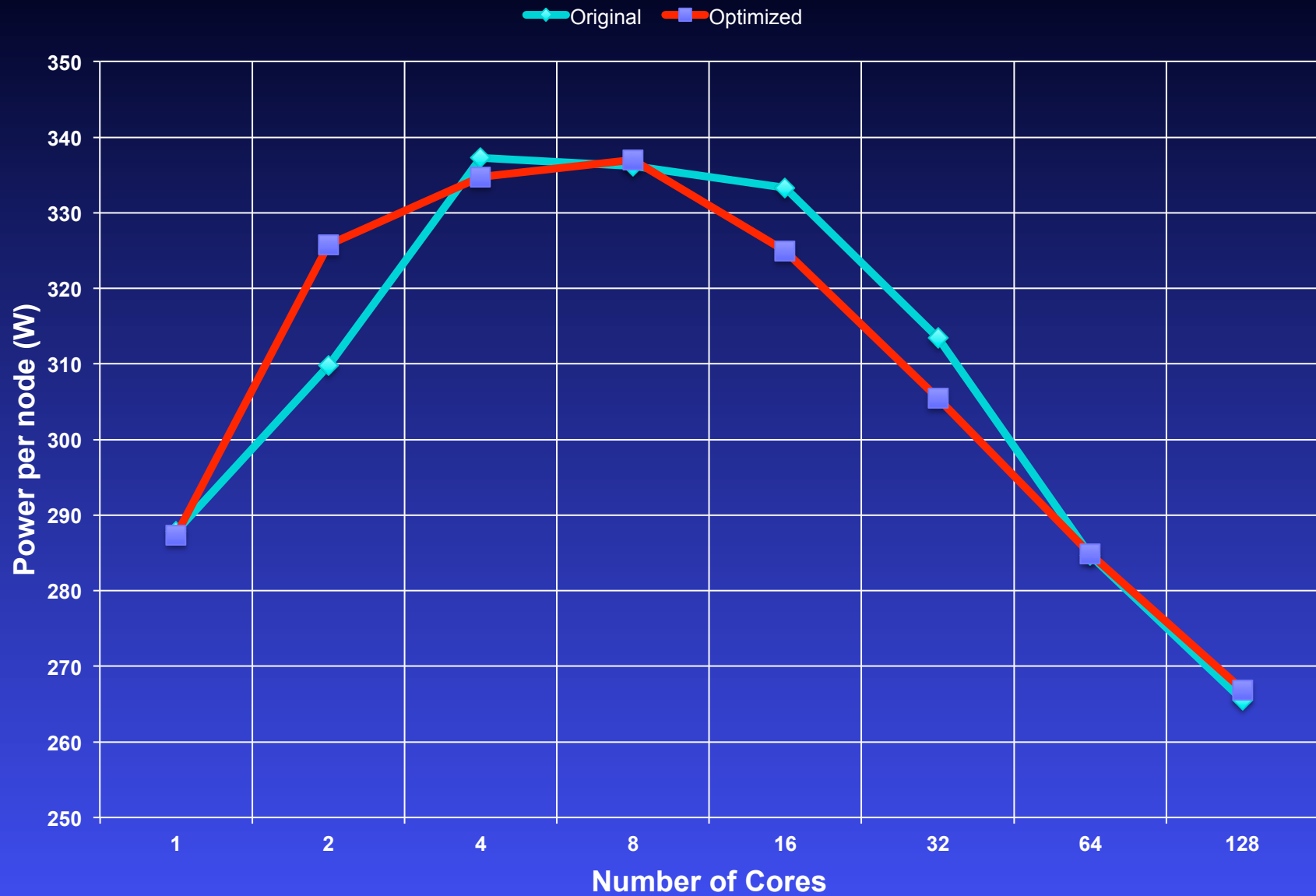
- **TLB_IM: Occurred in**
TLB_DM occurred in Runtime
BR_N occurred in CPU, Memory
L2 occurred in Runtime
L occurred in Runtime
L2 occurred in Runtime
BR occurred in Memory
BR occurred in Runtime
- **VEC_INS: Occurred in CPU, Memory**

*Final counters: TLB_IM
and VEC_INS
for optimization focus*

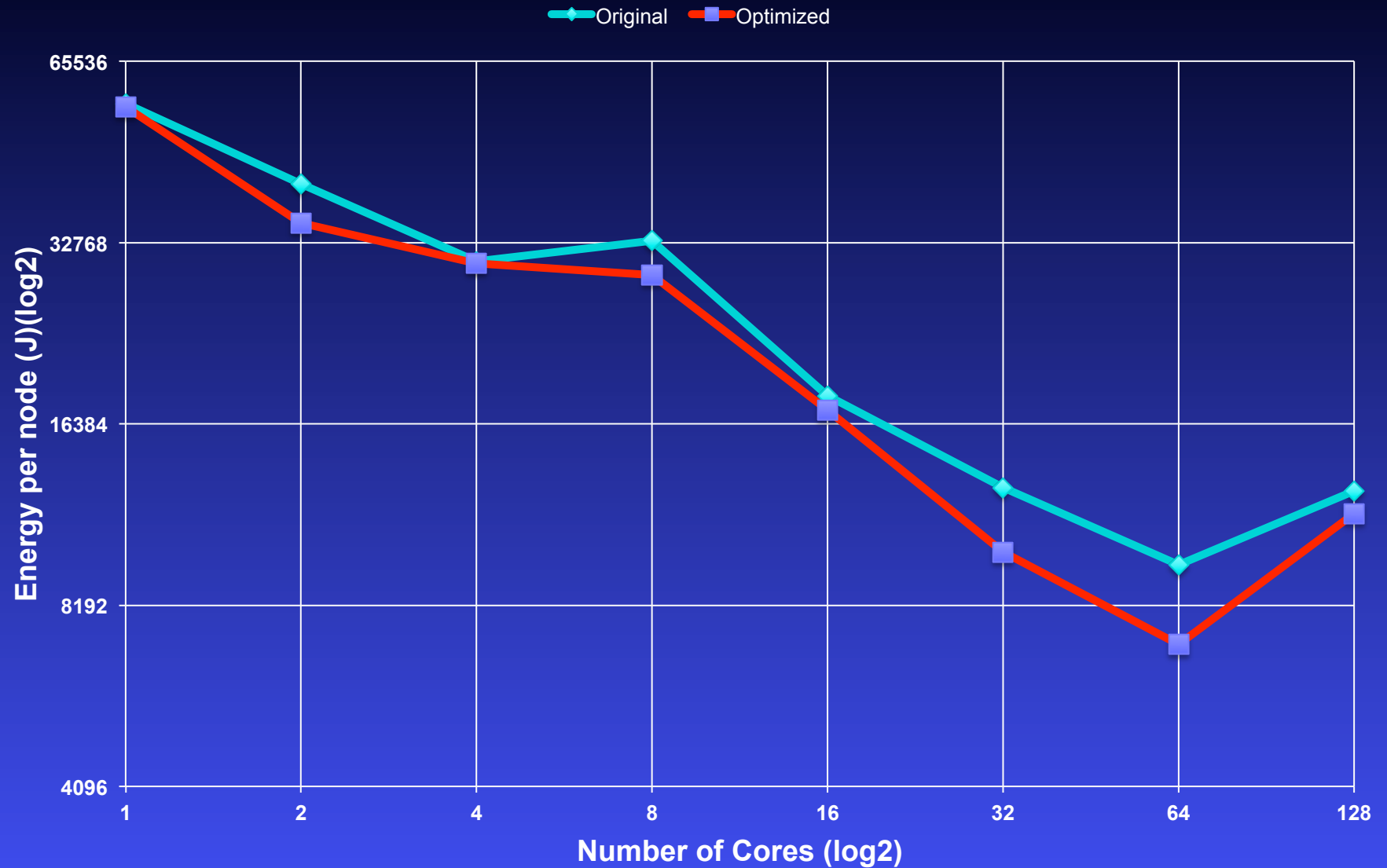
Performance for PMLB with 128x128x128 on SystemG



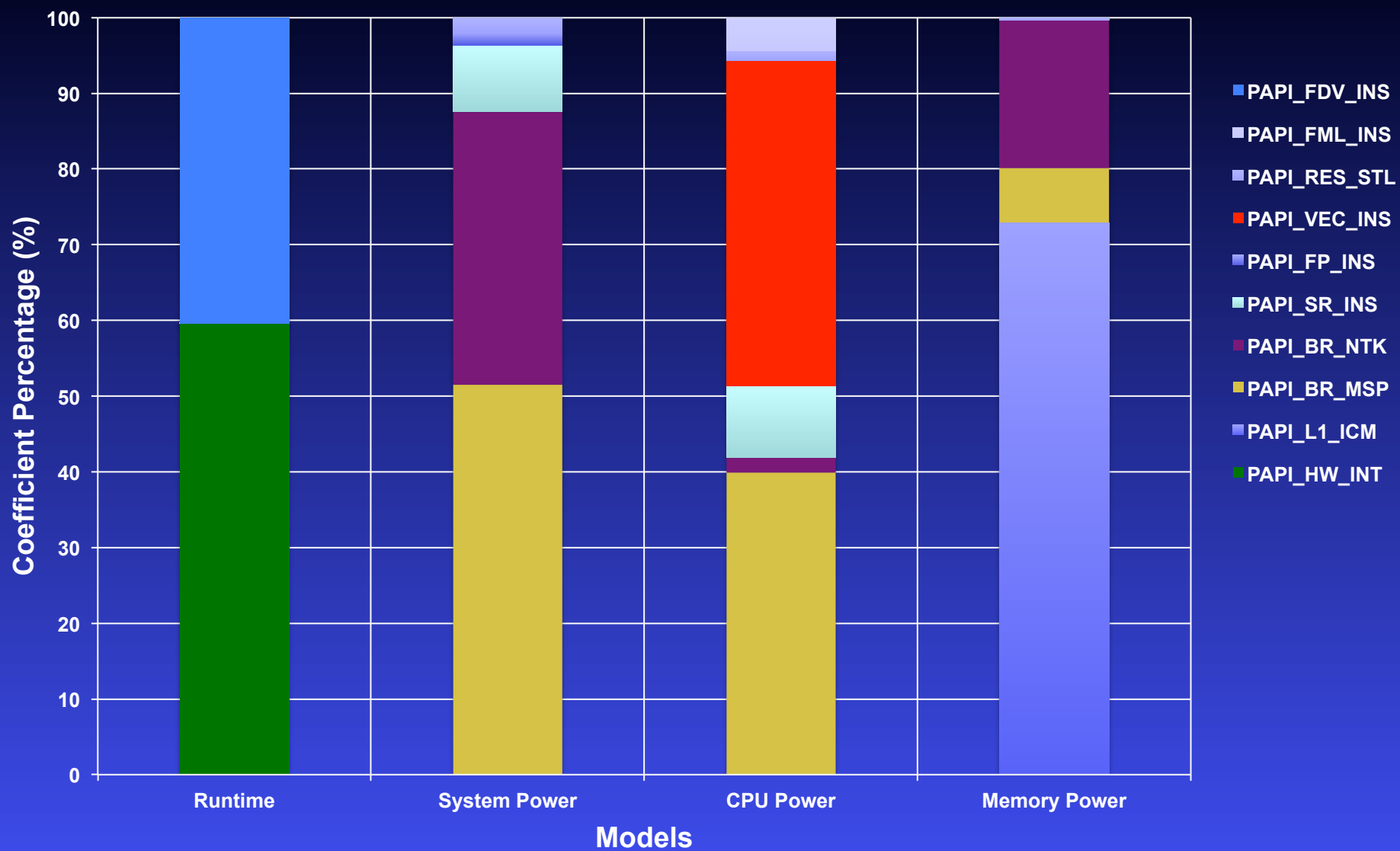
Node Power Comparison on SystemG



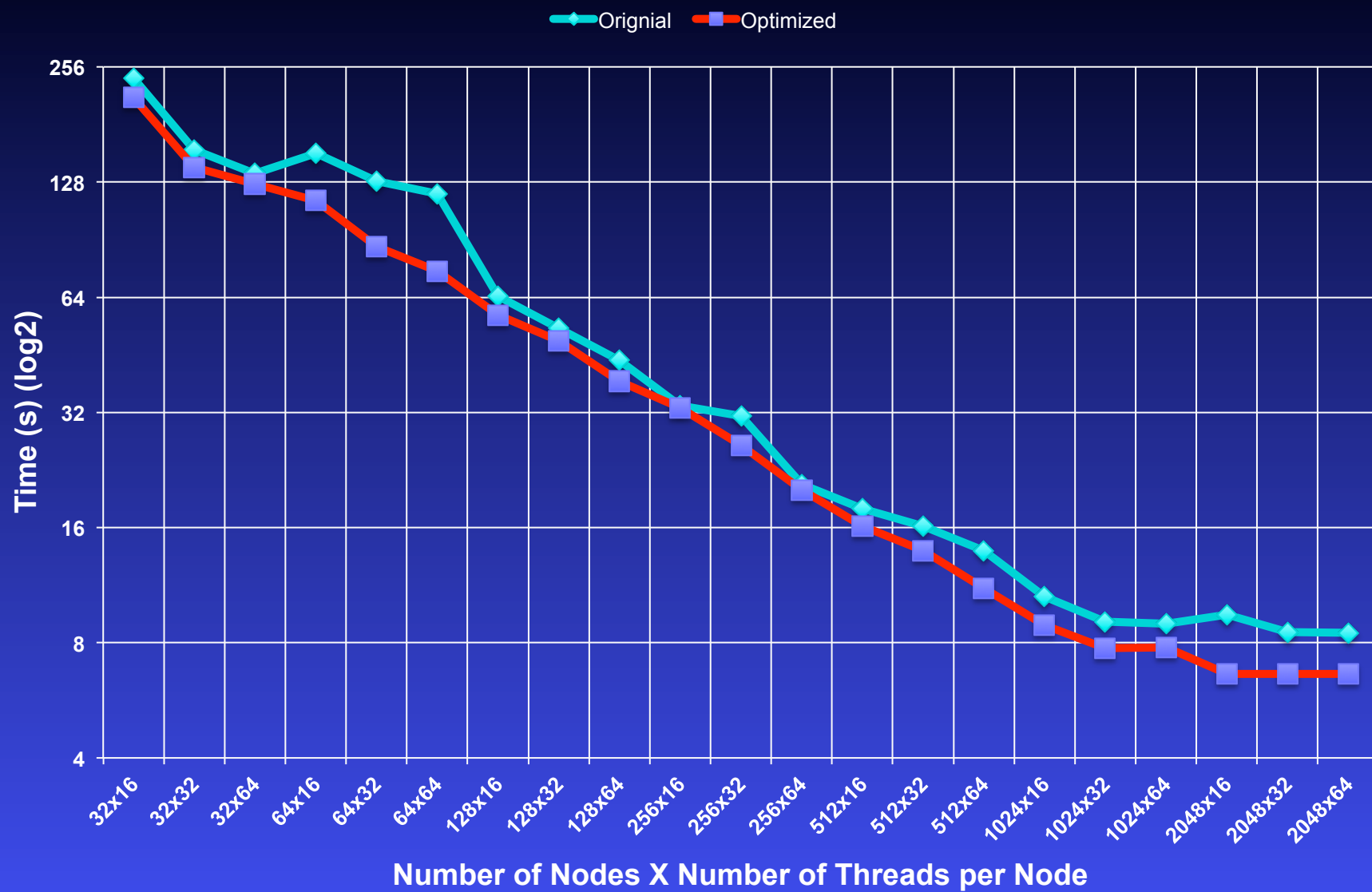
Energy Comparison for PMLB on SystemG



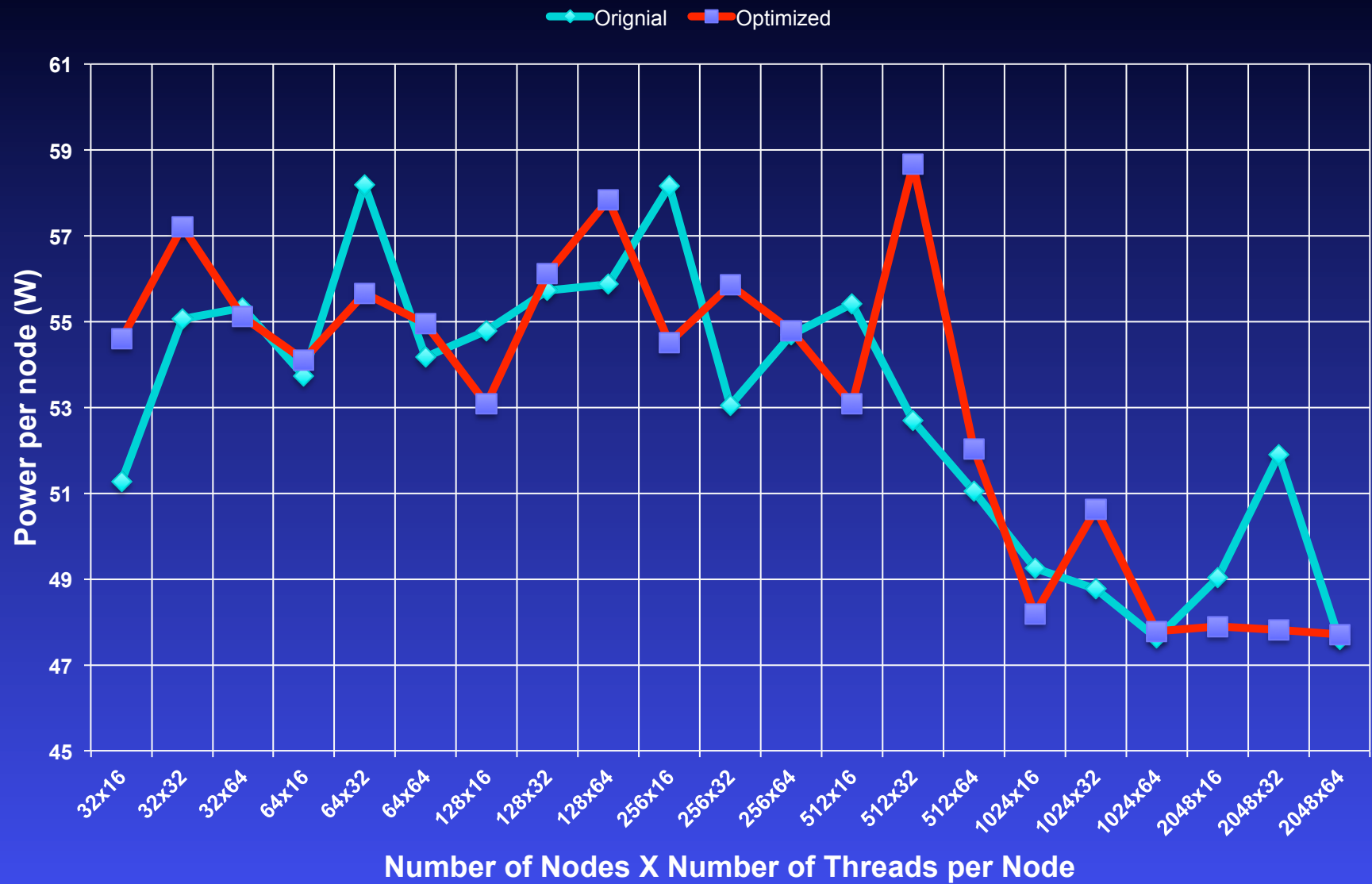
Counter Ranking for Original PMLB on Mira



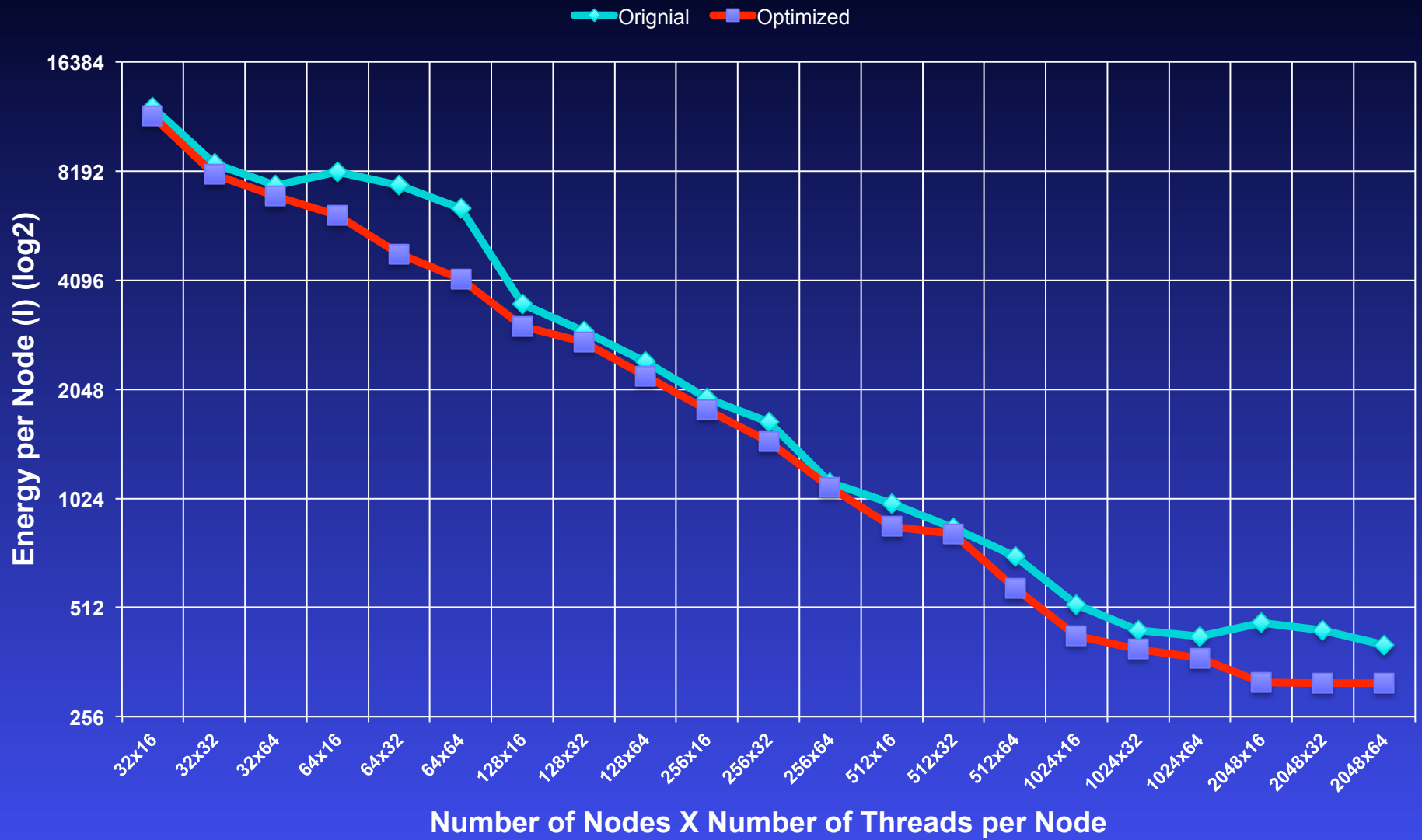
Performance Comparison on Mira



System Power Comparison on Mira

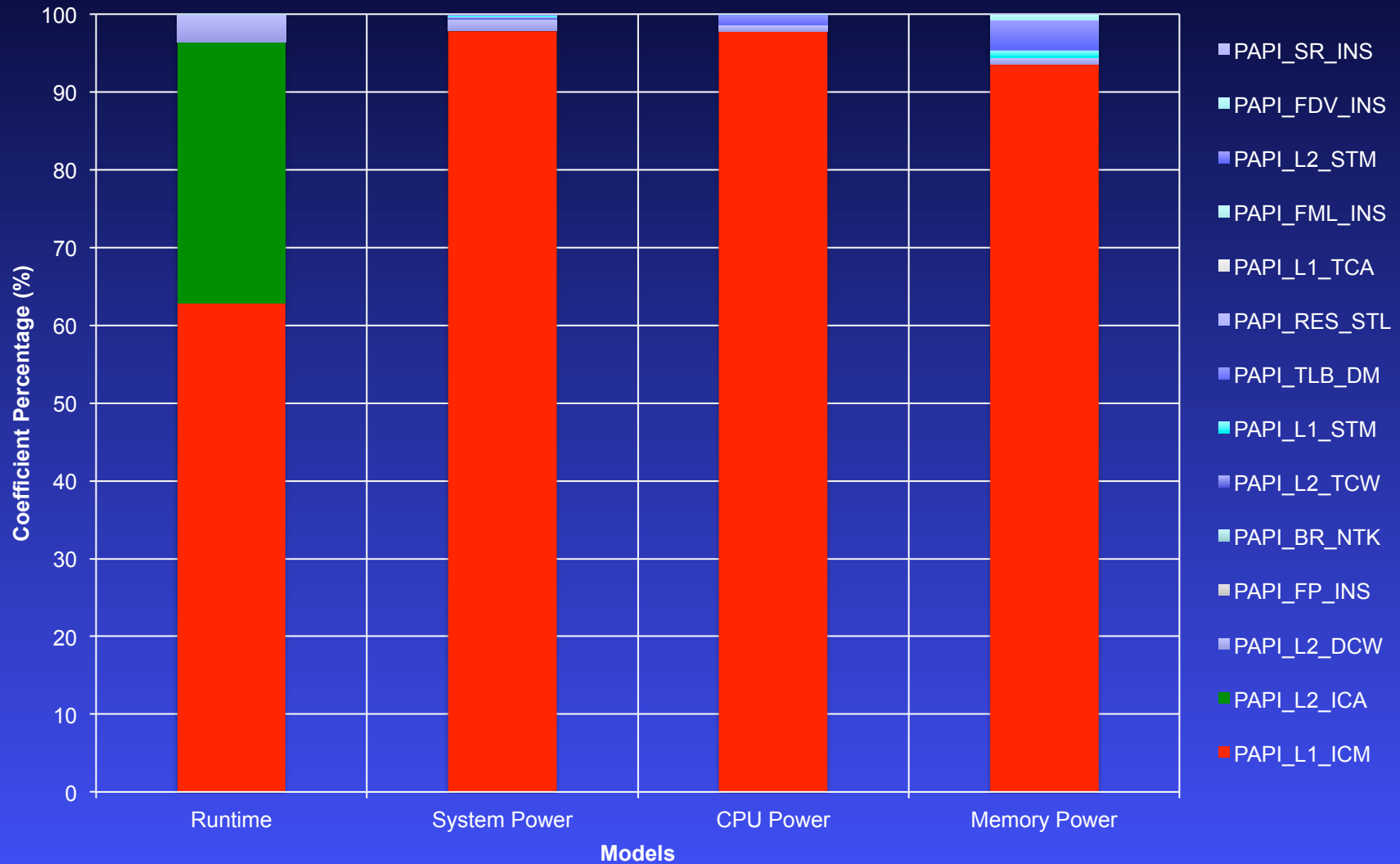


Energy Comparison for PMLB with 512x512x512 on Mira

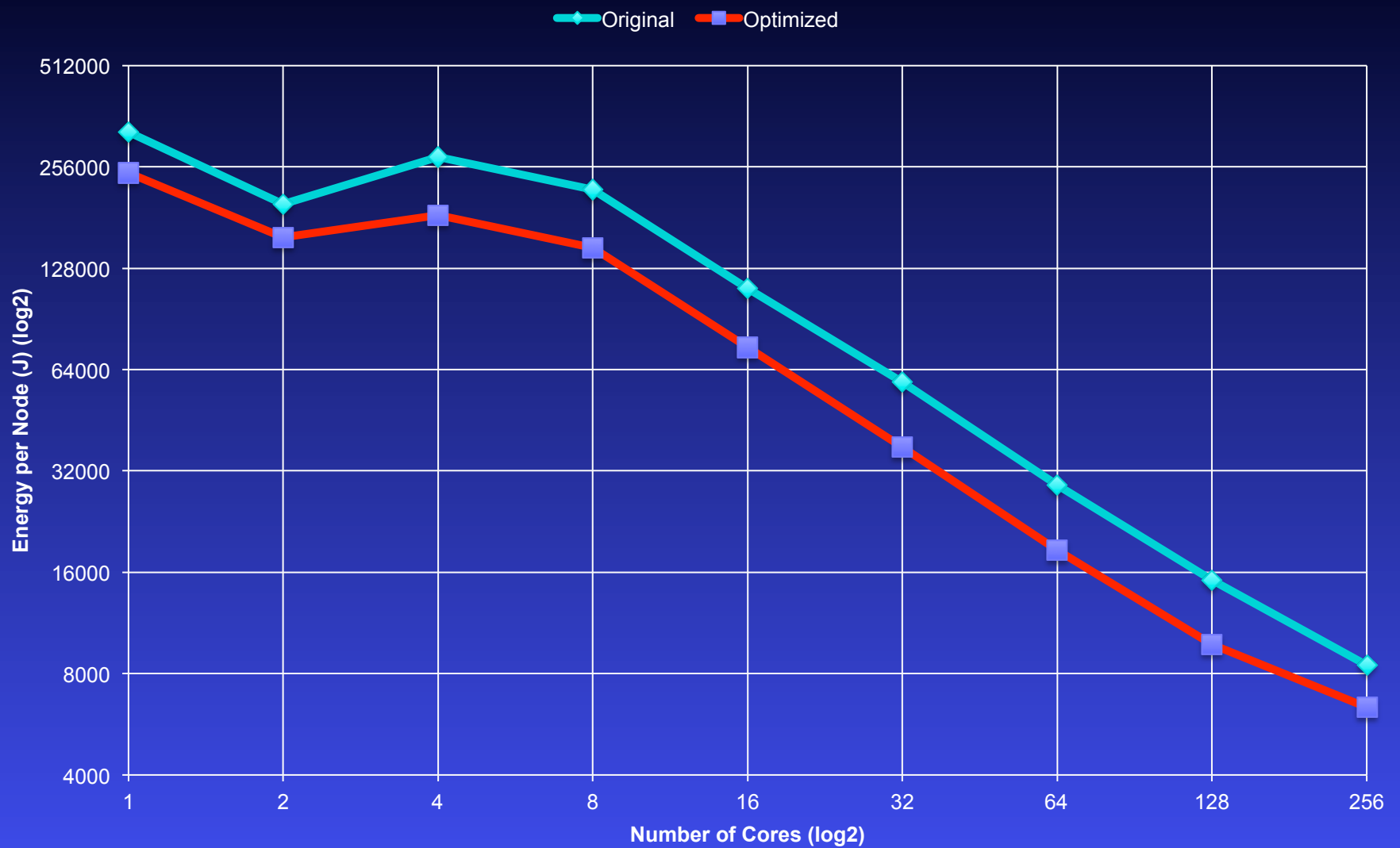


Counter Ranking on SystemG

Counter Ranking for Original eq3dyna on SystemG

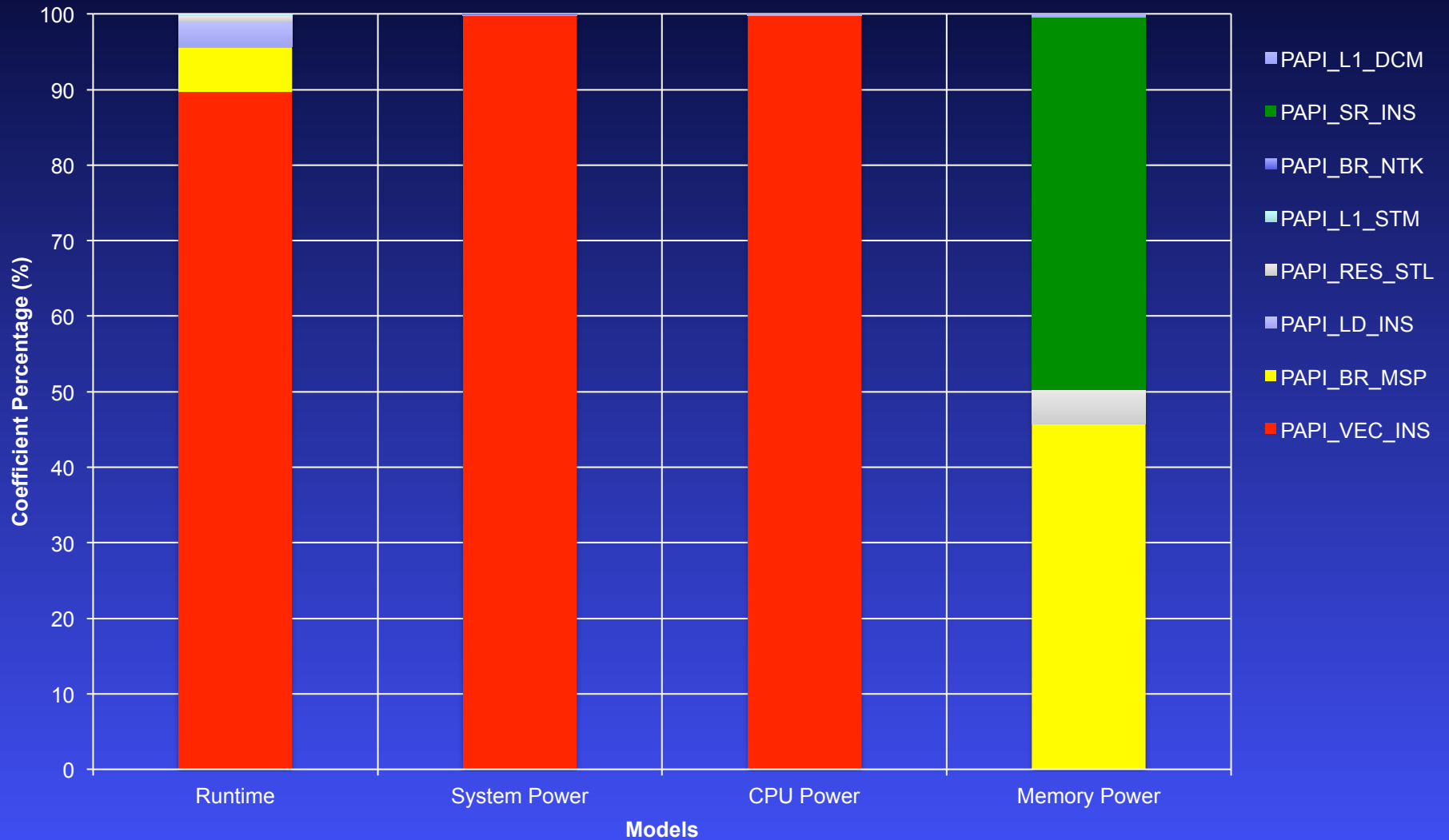


Energy Comparison of eq3dyna on SystemG

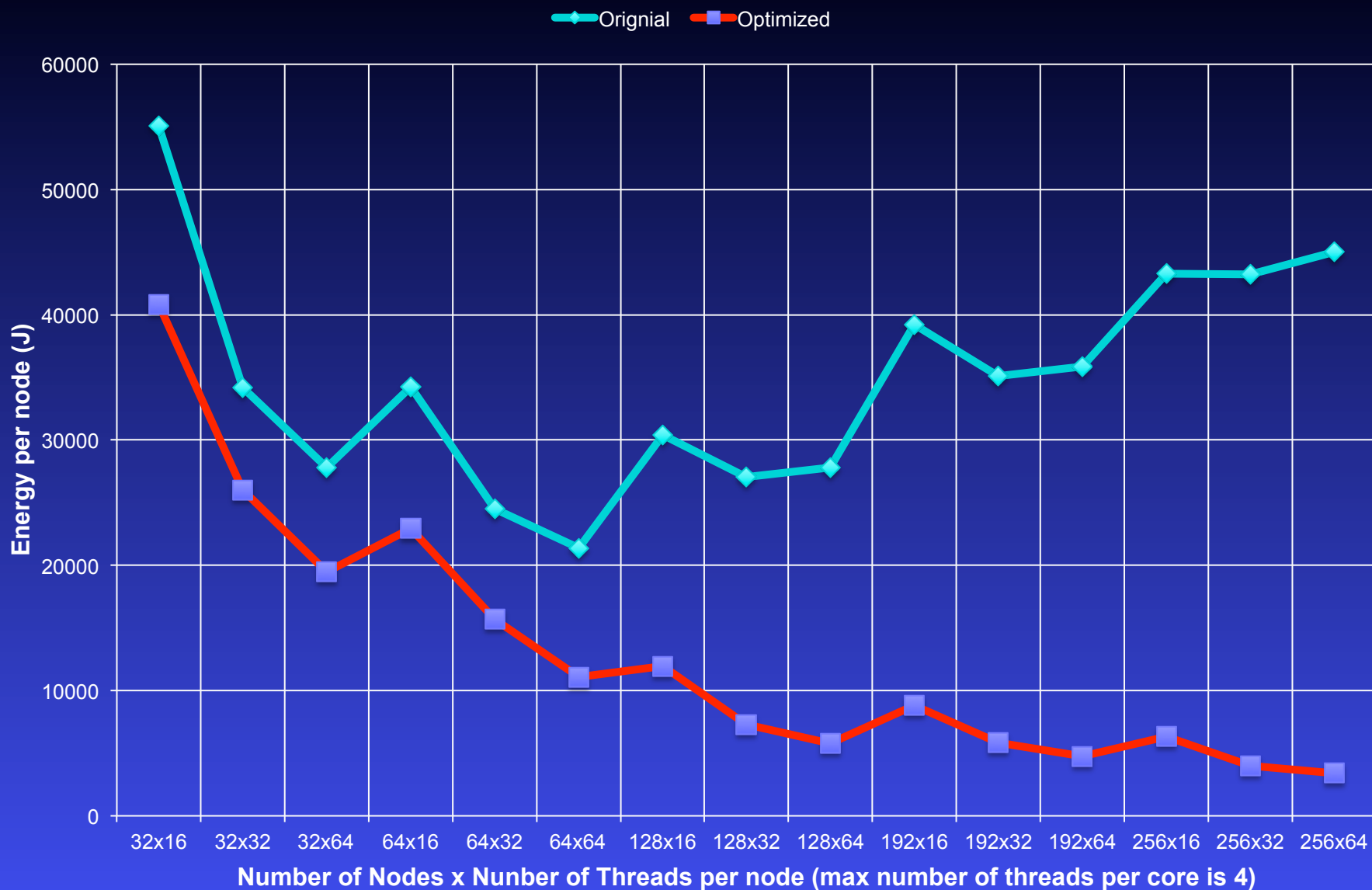


Counter Ranking on ANL BGQ Mira

Counter Ranking for Original eq3dyna on ANL BGQ Mira



Energy Comparison for eq3dyna with 100m on ANL BG/Q Mira



Summary and Future Work

- We used runtime and power models to identify the most important counters for optimization focus
- Our counters-guided optimizations save energy by
 - ◆ Eq3dyna:
Average 48.65% on Mira; 30.67% on SystemG
 - ◆ PMLB:
Average 18.28% on Mira; 11.28% on SystemG
- We believe that the modeling and optimization methodology can be applied to large-scale scientific applications on other architectures such as GPU and Xeon Phi by utilizing counters for GPU and Xeon Phi.