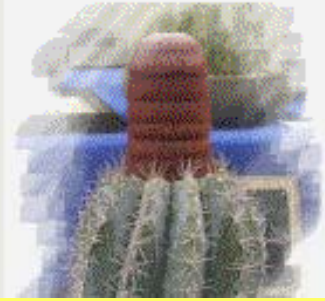
Scalable Methods for Performance and Power Data Collection and Analysis

Karen L. Karavanic
Portland State University

Andrés Márquez
Pacific Northwest National Laboratory
Rashawn Knapp, Agniv Adhikari, Mike Smith
Portland State University

Los Angeles Cactus and Succulent Society

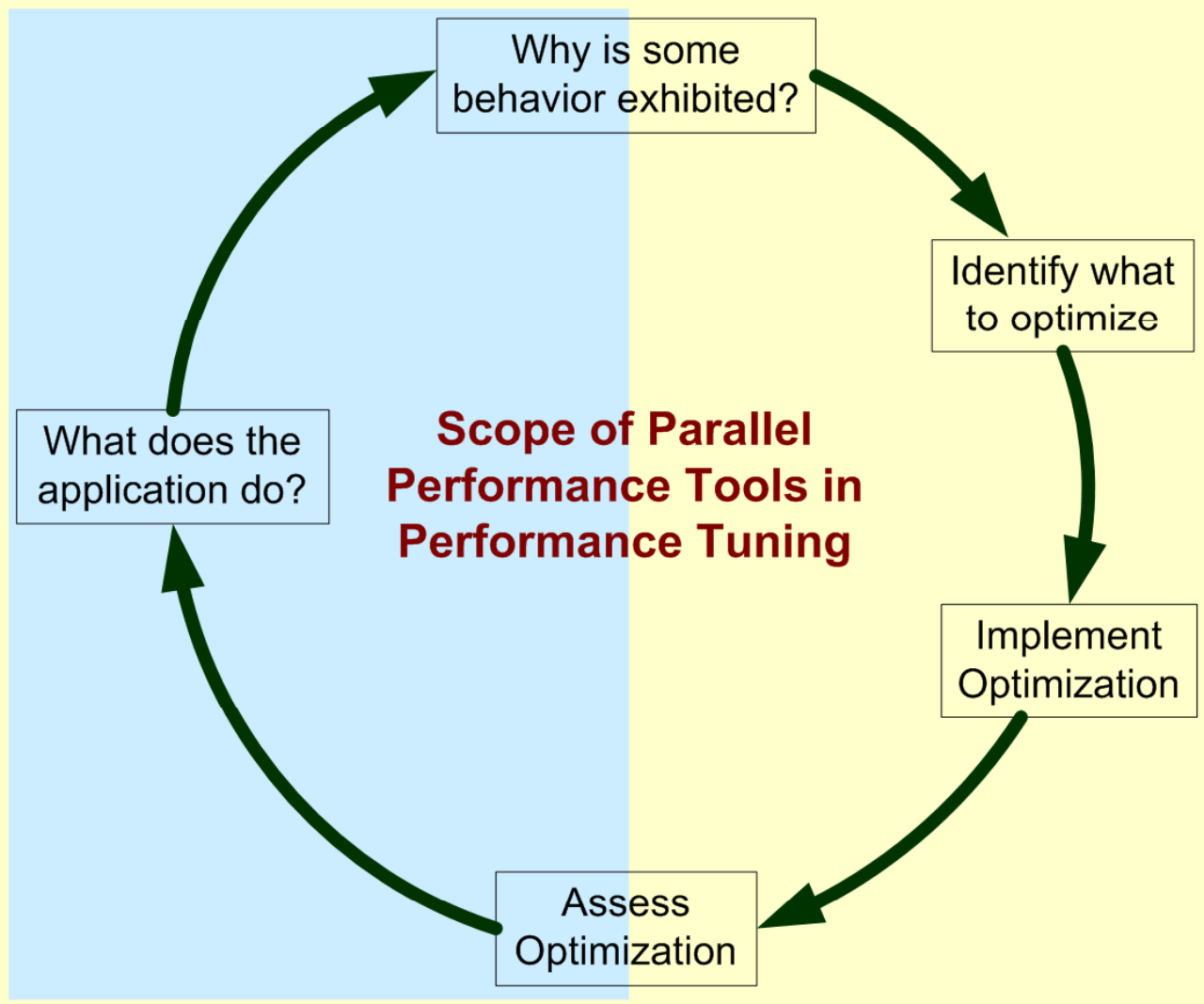
Serving the Los Angeles Cactus and Succulent Community since 1935



July 2009 Program
"The Art of Staging"
by Artie Chavez & Steven Frieze

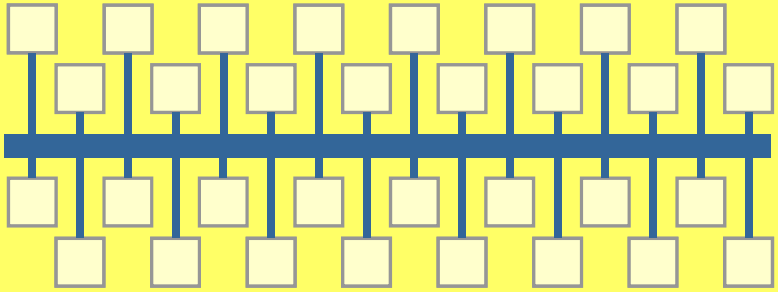
Our featured speakers for our July program will be Artie Chavez & Steven Frieze, two of the clubs most long-standing members. "The Art of Staging" will be a workshop in the art of staging your plants. Artie and Steven have a combined 50 years of experience in growing and staging succulents and cacti. Each has won numerous ribbons, rosettes, and trophies in Southern

Performance Tuning and the Scope of Parallel Performance Tools

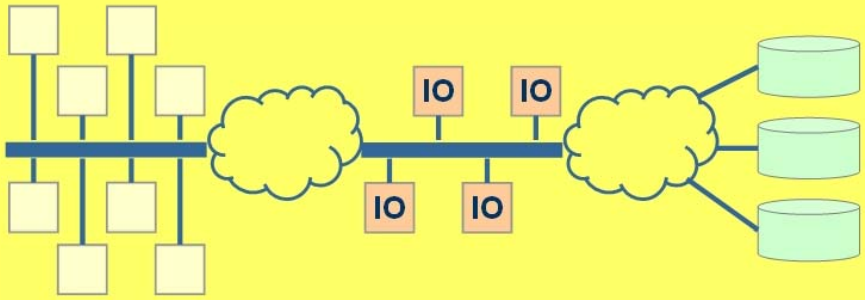


Limitations of Application Focused Performance Analysis

Faulty Hardware



OS Interference



File System

Traditional methods misdiagnose or do not detect an important class of performance problems

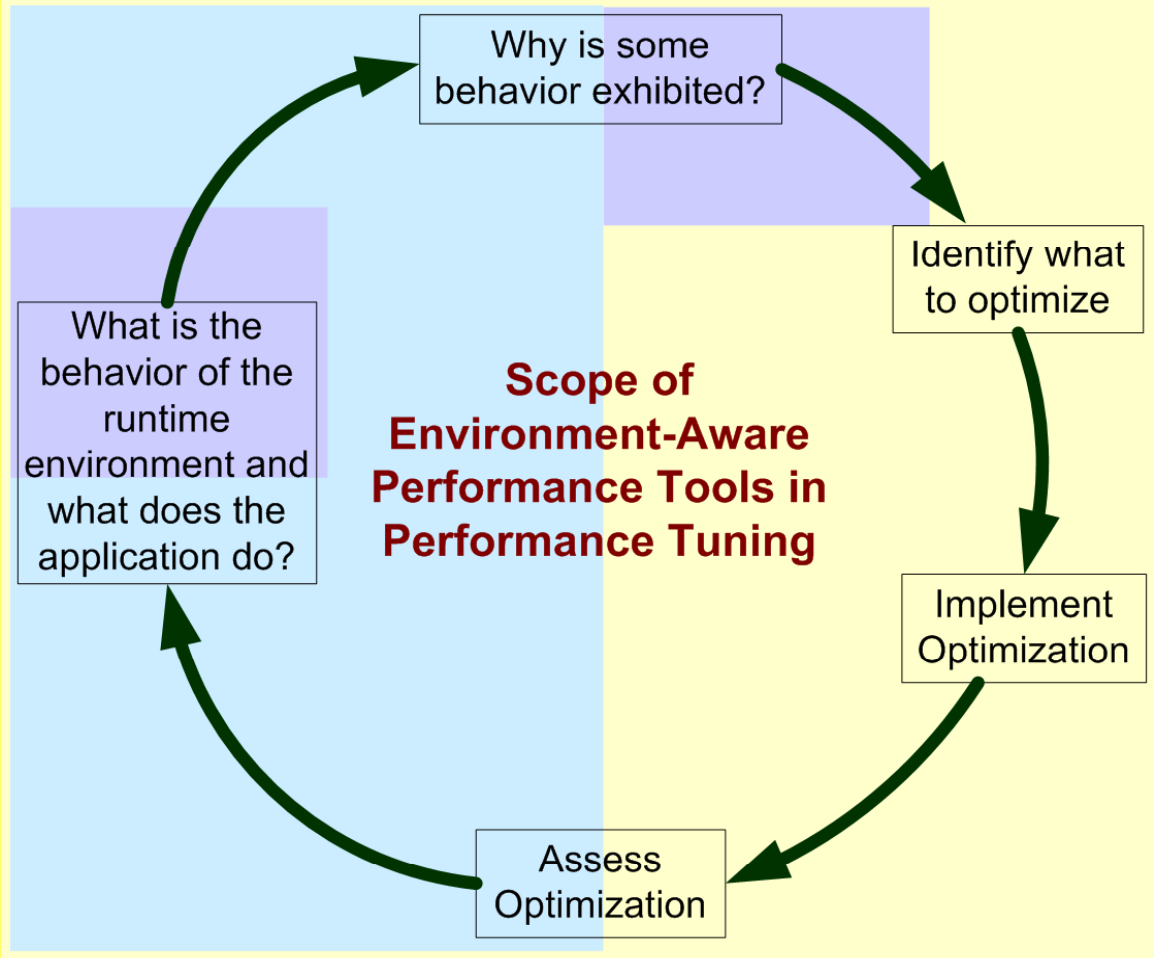
Characterized by:

- Difficult diagnosis
- Application metrics are not sufficient
- Requires performance data about entire system

Our Approach

Environment Aware Performance Diagnosis (EAPD)

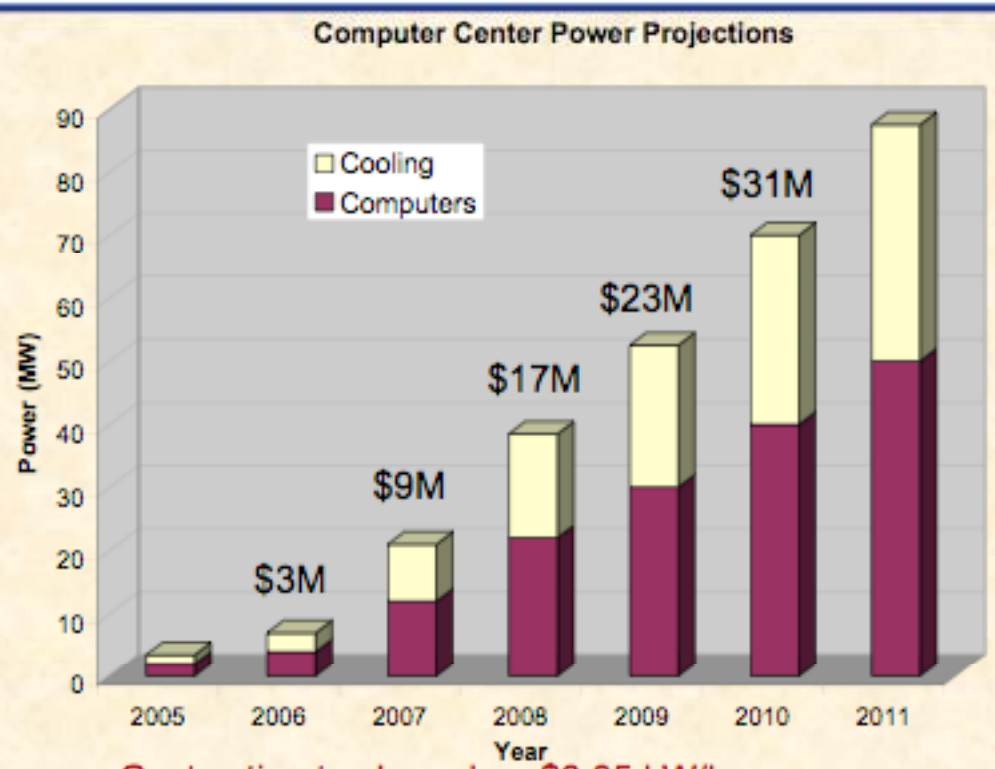
Application and *system* level metrics



Online and *automated* methods

ORNL Computing Power and Cooling 2006 - 2011

- Immediate need to add 8 MW to prepare for 2007 installs of new systems
- NLCF petascale system could require an additional 10 MW by 2008
- Need total of 40-50 MW for projected systems by 2011
- Numbers just for computers: add 75% for cooling
- Cooling will require 12,000 – 15,000 tons of chiller capacity



Cost estimates based on $0.05 kW/hr

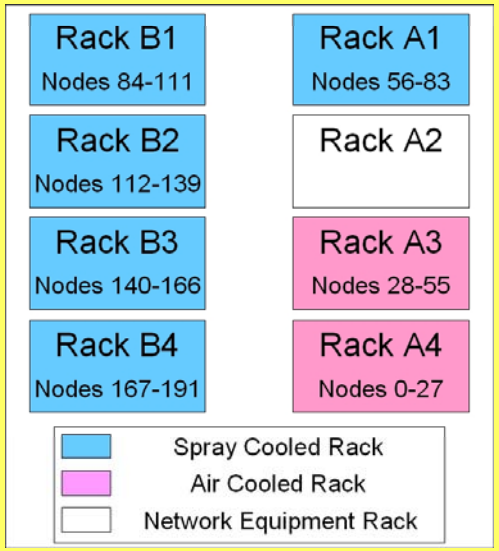
Annual Average Electrical Power Rates $/MWh

Site	FY 2005	FY 2006	FY 2007	FY 2008	FY 2009	FY 2010
LBNL	43.70	50.23	53.43	57.51	58.20	56.40 *
ANL	44.92	53.01				
ORNL	46.34	51.33				
PNNL	49.82	N/A				

Data taken from Energy Management System-4 (EMS4). EMS4 is the DOE corporate system for collecting energy information from the sites. EMS4 is a web-based system that collects energy consumption and cost information for all energy sources used at each DOE site. Information is entered into EMS4 by the site and reviewed at Headquarters for accuracy.

The Energy Smart Data Center at PNNL

- NWICE - 192 nodes, 2.3 GHz Intel (quad-core) Clovertown,
- Five racks spray-cooled
- Two racks air-cooled
- Sensors to monitor temperatures, power usage, cooling capacity
- FRED measurement and analysis tool



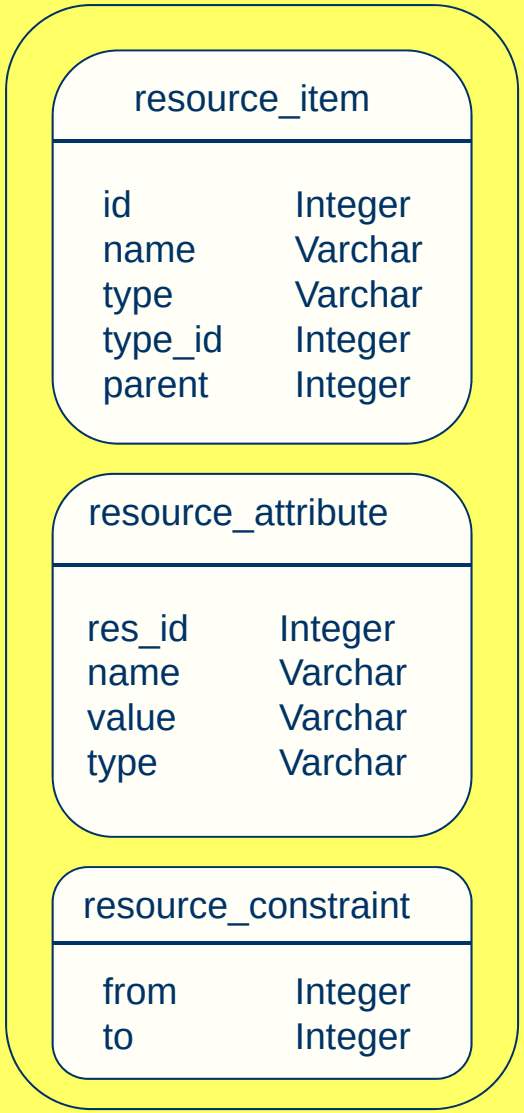
The PerfTrack Project

PerfTrack is a tool for storing, exploring, and analyzing application performance data

- Collect and store description of each platform
- Collect and store description of each build and run of an application
- Integrate DBMS into a performance analysis tool
- Store a wide variety of performance data
  - Data from different measurement tools:

    Tracing, DPCL, Paradyn, TAU, Vampir, Speedshop, HW counters, native application performance measurements, etc.
- GUI for data navigation and querying
- Shield tool user from DBMS internals

PerfTrack Design



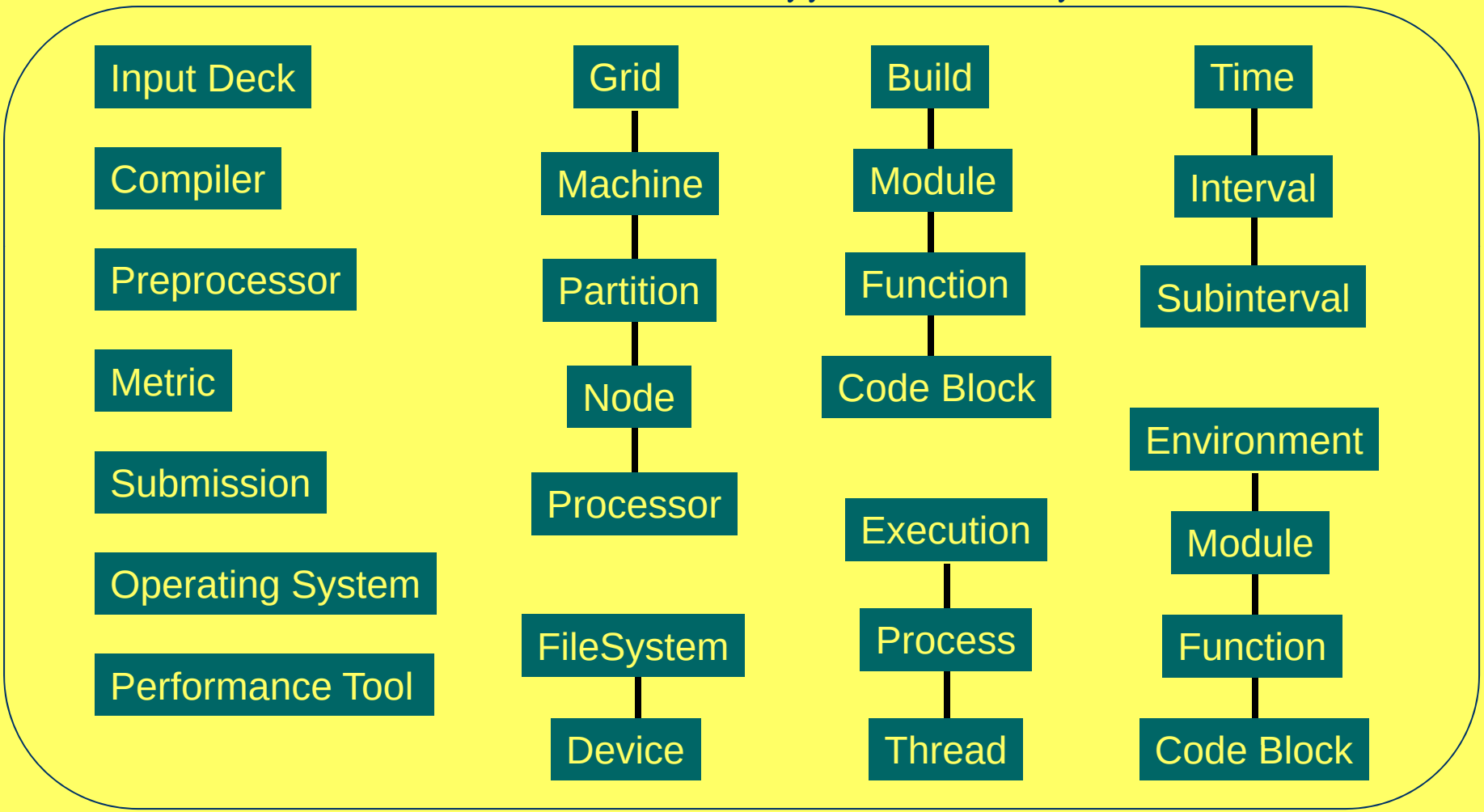
Generic Database Schema

PerfTrack Data Format (PTDF)

```
ResourceType resourceTypeName

Application appName

Execution execName appName

Resource resourceName resourceTypeName
   execName

Resource resourceName resourceTypeName

ResourceAttribute resourceName attributeName
   attributeValue attributeType

ResourceConstraint resourceName1
   resourceName2

PerfResult execName resourceSet perfToolName
   metricName value units startTime endTime
```

PerfTrack Design

Base Resource Type Hierarchy



Integrating ESDC Data with Parallel Performance Data

Extended Resource Type Hierarchy

```
DataCenter
├─ Chiller
├─ CoolingTower
├─ HeatRecovery
├─ AirConditioner
├─ Pump
└─ Rack
      PowerUnit
      TMU
Grid
└─ Machine
      Partition
         Node
            Processor
            DIMM
            ASIC
Build
└─ Module
      Function
         CodeBlock
Environment
└─ Module
   └─ Function
      └─ CodeBlock
Execution
└─ Process
   └─ Thread
Time
└─ Interval
   └─ Subinterval
FileSystem
└─ Device
Compiler
OperatingSystem
InputDeck
Preprocessor
Metric
Submission
PerformanceTool
```

Modified PTDF

```
PerfResult execName focus perfTool
   metricName value units sTime eTime

Result resourceName focus perfTool
   metricName value units sTime eTime
```

New Metrics

Cooling Tower: Fan KW, Fan Speed, etc

Nodes: CPU Temp, Dimm Temp

Racks: Supply temp, Return temp

Selecting Data in PerfTrack

Select Data (on ruby.cat.pdx.edu)

Choose resource names and attributes to search for in the left pane. Add them to the Selection Parameters, then press Get Data to retrieve results.

Resources

metric

Show resource information

Name	Type
TMU VCU Press Status	metric
TMU VCU Vent Status St...	metric
TMU Water Inlet Temp	metric
TMU Water Inlet Temp St...	metric
TMU Water Outlet Temp	metric
TMU Water Outlet Temp ...	metric
TMU Water Valve1 Statu...	metric
TMU Water Valve2 Statu...	metric
Total Interrupts	metric
Tower EFan KW	metric
Tower EFan Speed	metric

Attribute	Value

Add to Selection Parameters

Add Resource Type

Selection Parameters

Rela	Type	Value	Count
D	datacenter|rack	B1	68990
D	datacenter|rack	B2	68865
D	grid|machine|partition|node	B1Node111	14770
D	grid|machine|partition|node	B2Node112	14780
D	metric	CPU0 Temp	21780
D	metric	CPU1 Temp	21780
D	metric	DIMM1 Temp	7560
D	metric	DIMM2 Temp	7560
D	metric	Node Air In Temp	68604
D	metric	Rack Return Temp	600
D	metric	Rack Supply Temp	600
D	metric	Total Interrupts	34

Items matching all parameters: 23403

Clear Highlighted Parameters

Performance Result Label

TI_4n4ppn
TI_8n2ppn
t01_8n2ppn

Clear All Entries

Combine Data | Cancel | Get Data

Selecting Data in PerfTrack

Select Data (on ruby.cat.pdx.edu)

Choose resource names and attributes to search for in the left panel. Add them to the Selection Parameters, then press Get Data to retrieve results.

Resources

metric

Show resource information

Name	Type
TMU VCU Press Status	metric
TMU VCU Vent Status St...	metric
TMU Water Inlet Temp	metric
TMU Water Inlet Temp St...	metric
TMU Water Outlet Temp	metric
TMU Water Outlet Temp ...	metric
TMU Water Valve1 Statu...	metric
TMU Water Valve2 Statu...	metric
Total Interrupts	metric
Tower EFan KW	metric
Tower EFan Speed	metric
Tower EFan Status	metric

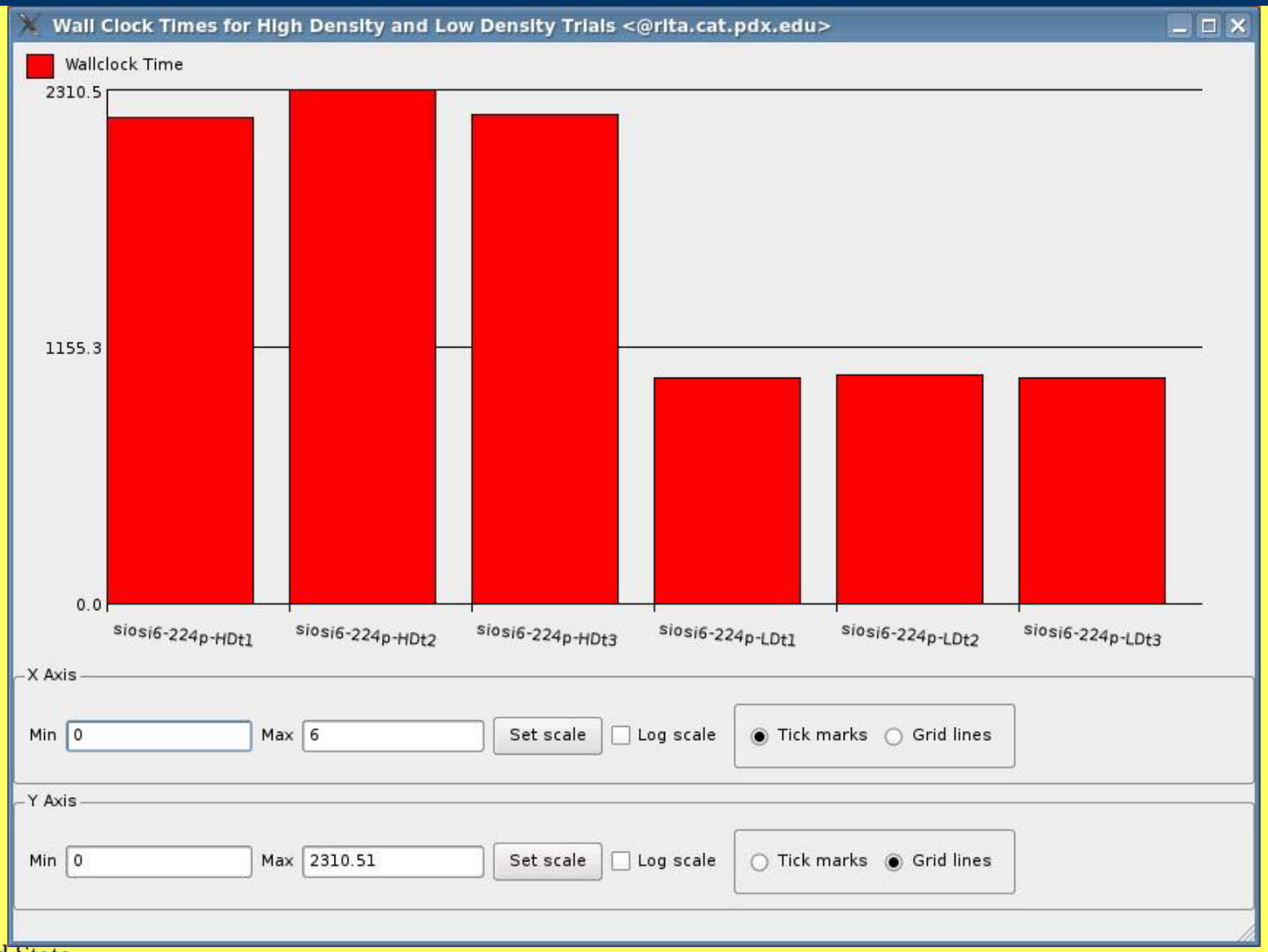
Attribute	Value

Add to Selection Parameters

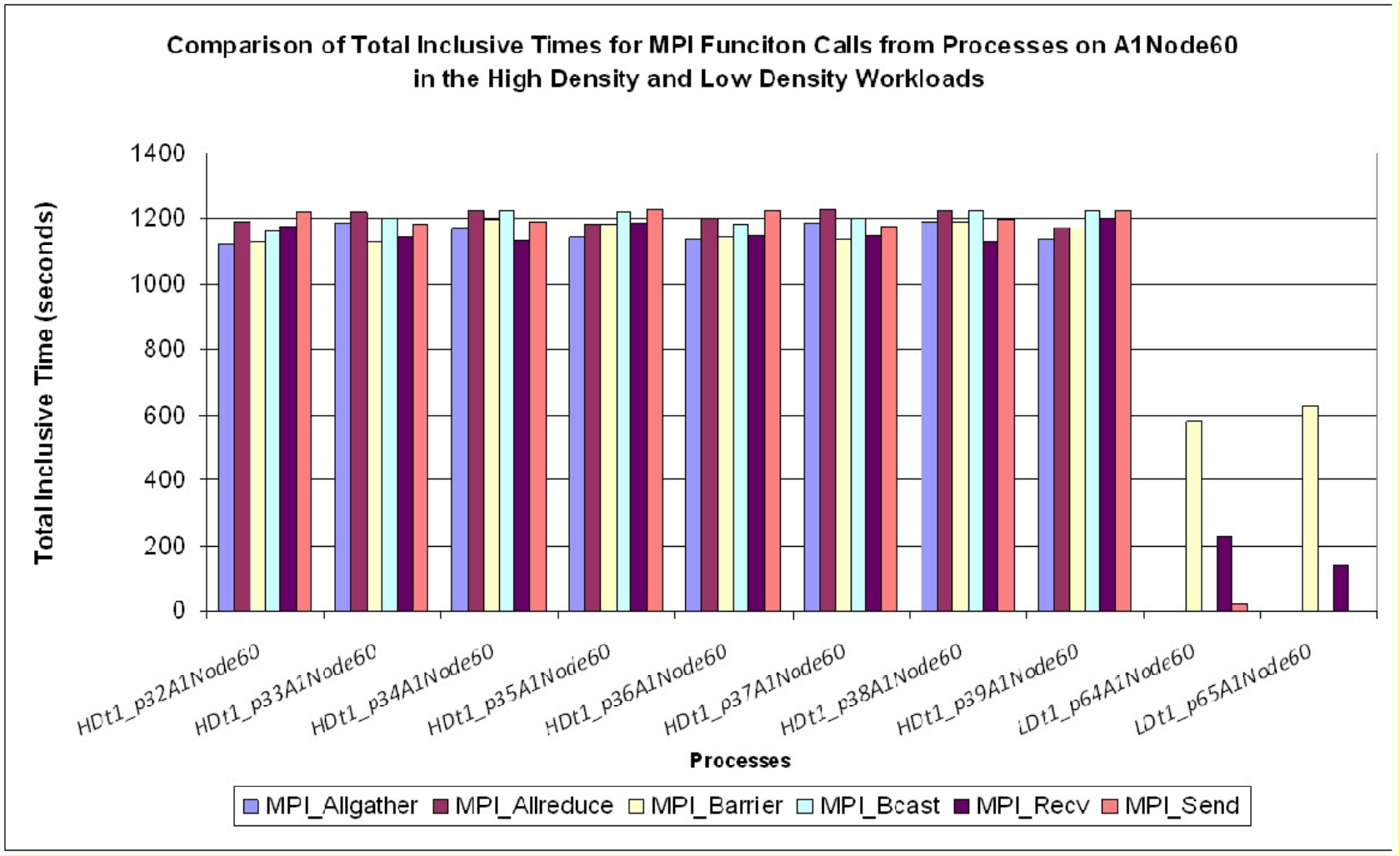
Selection Parameters

Rela	Type	Value	Count
D	datacenter|rack	B1	68990
D	datacenter|rack	B2	68865
D	grid|machine|partition|node	B1Node111	14770
D	grid|machine|partition|node	B2Node112	14780
D	metric	CPU0 Temp	21780
D	metric	CPU1 Temp	21780
D	metric	DIMM1 Temp	7560
D	metric	DIMM2 Temp	7560
D	metric	Node Air In Temp	68604
D	metric	Rack Return Temp	600
D	metric	Rack Supply Temp	600
D	metric	Total Interrupts	34

Using Power, Cooling, and Application Performance Data in a Workload Placement Study

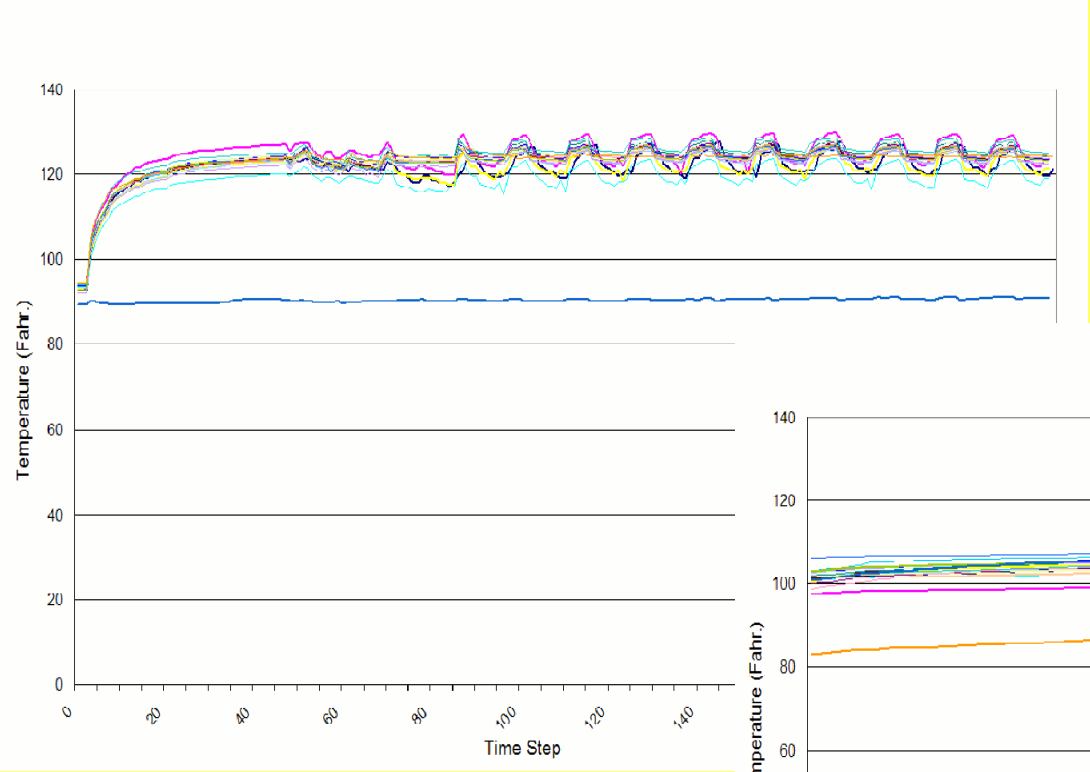


Workload Placement Study: Application Level

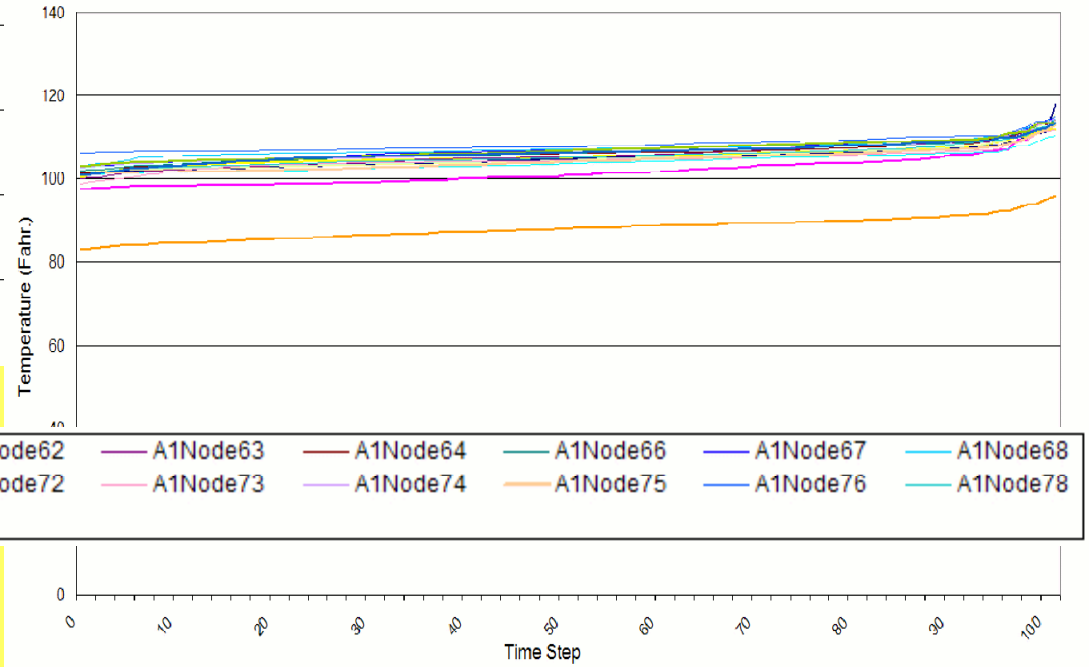


Workload Placement Study: CPU Temperatures

High Density Workload

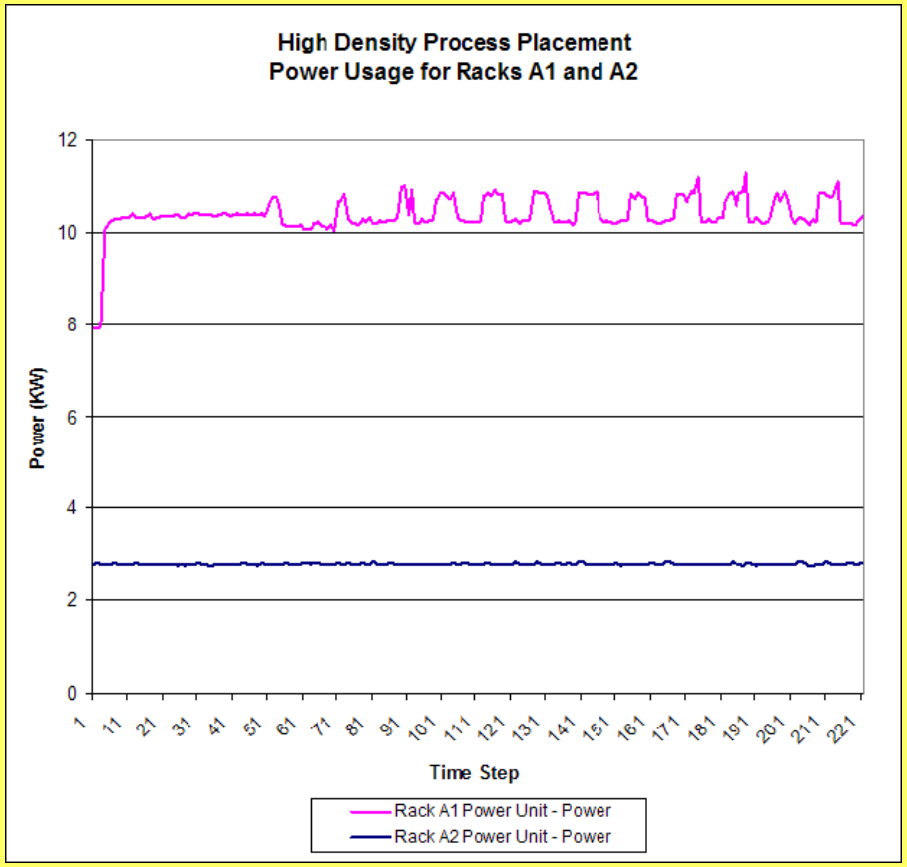


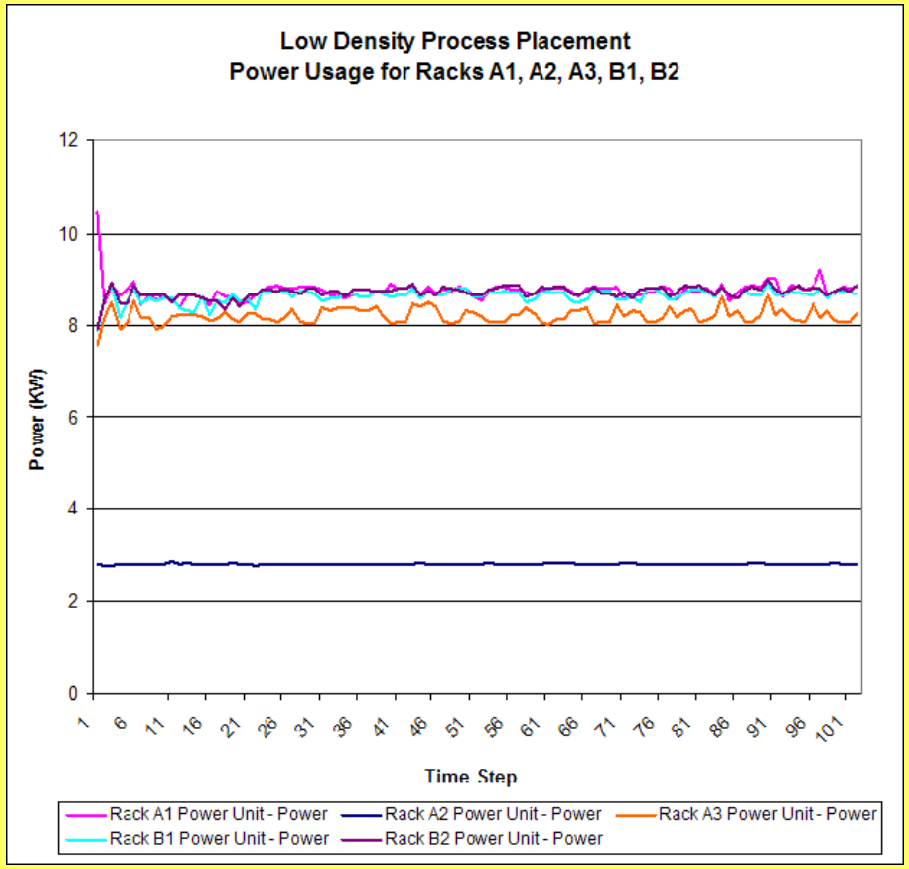
Low Density Workload



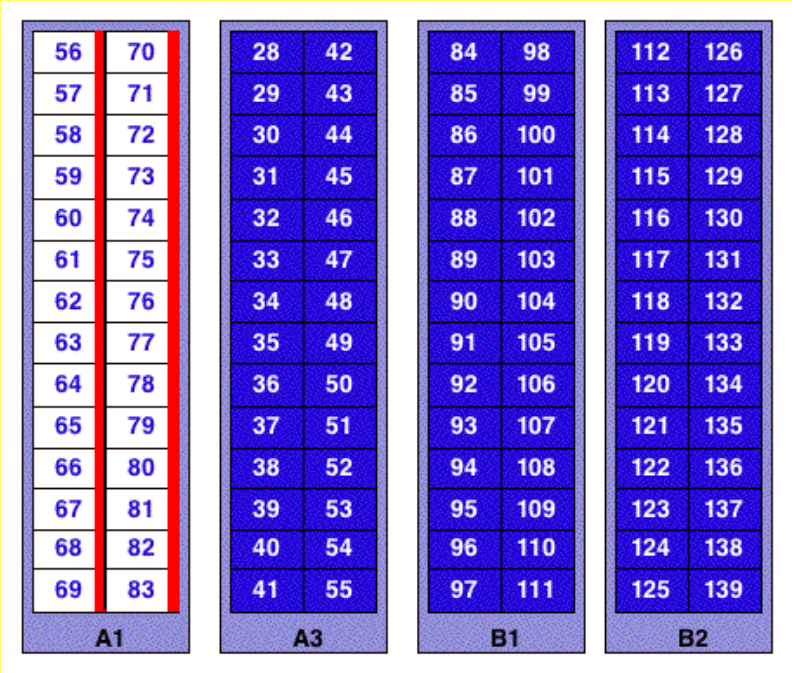
A1Node56 A1Node60 A1Node61 A1Node62 A1Node63 A1Node64 A1Node66 A1Node67 A1Node68
A1Node69 A1Node70 A1Node71 A1Node72 A1Node73 A1Node74 A1Node75 A1Node76 A1Node78
A1Node79 A1Node80 A1Node83

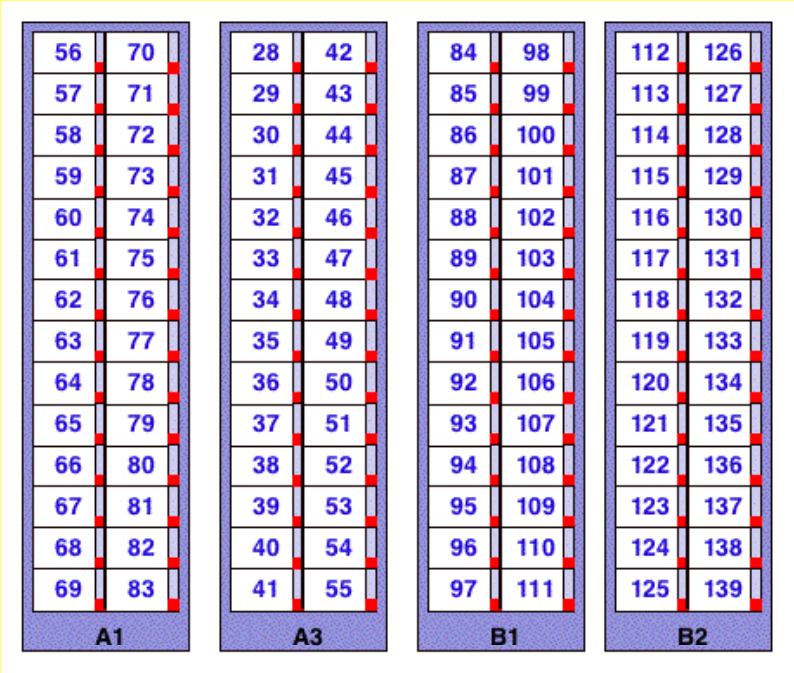
Workload Placement Study: Rack Power Usage





New Visualizations for Integrated Data

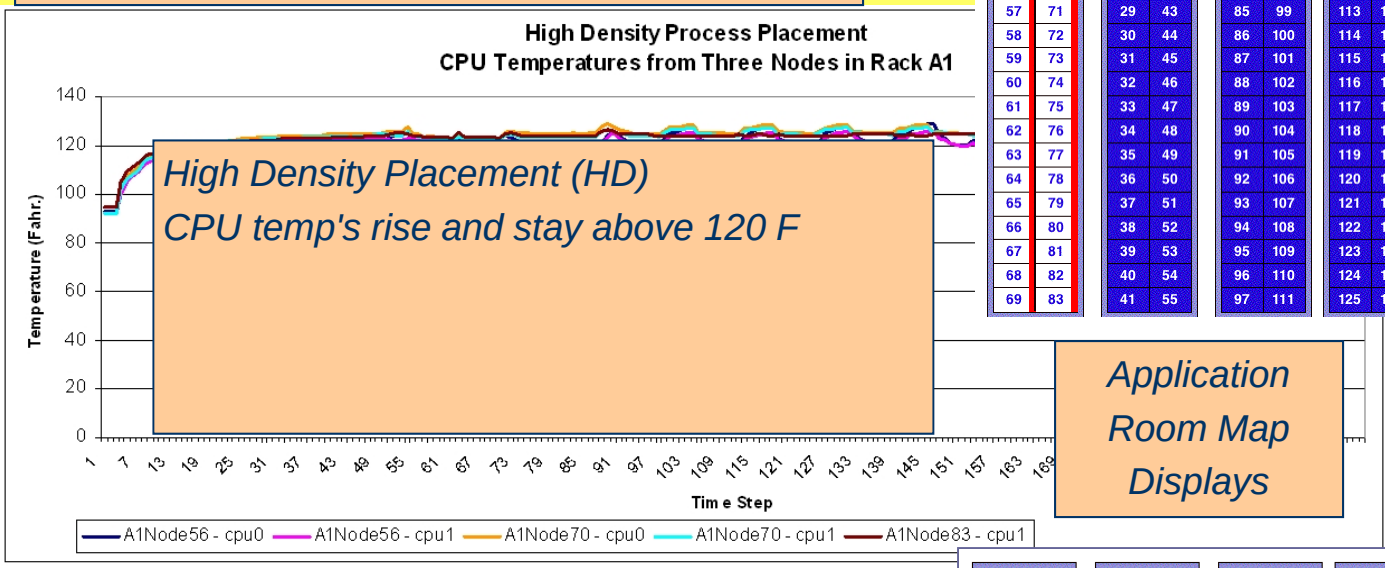




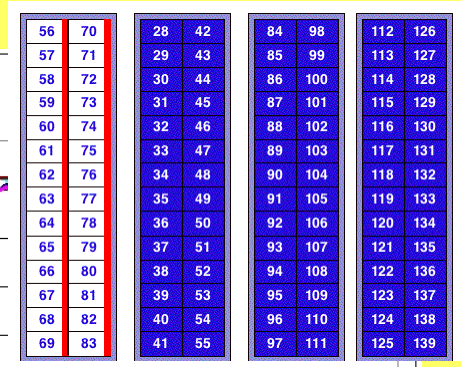
RoomPrint

Integrating Power/Cooling Into Application Performance Analysis: Example

224 processes on 1 rack (8 proc.s per node)



High Density Placement (HD)
CPU temp's rise and stay above 120 F

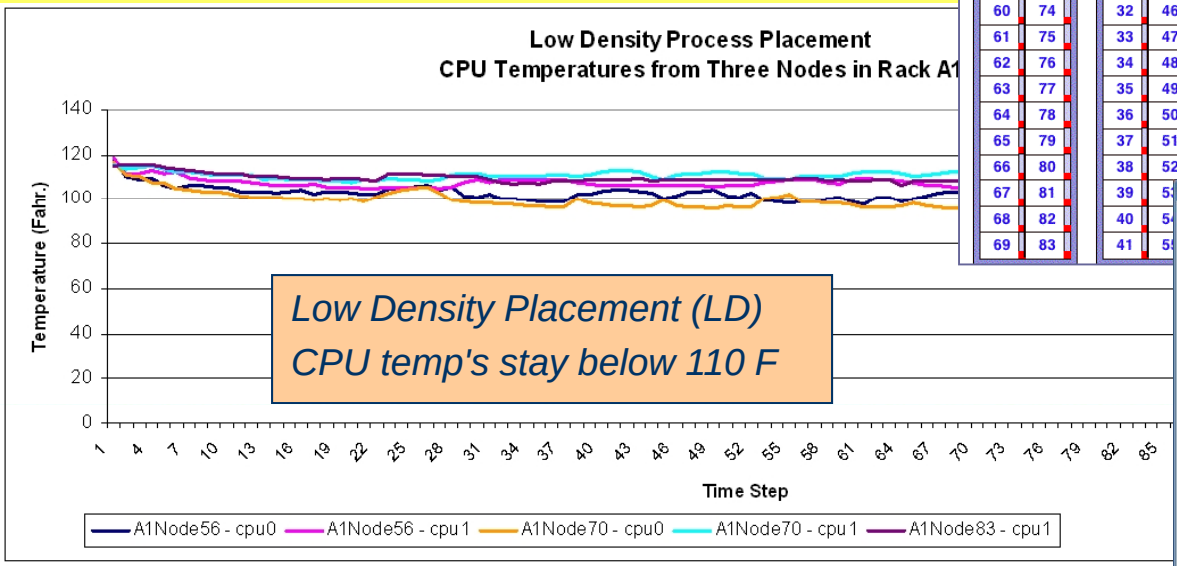


Application Room Map Displays

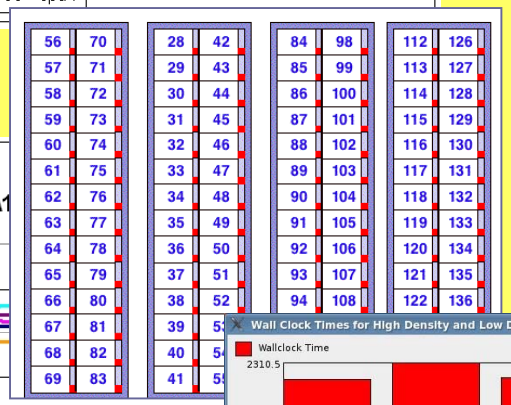
Total KWH

Rack	HD	LD
A1	6.397	2.448
A2	1.721	0.789
A3		2.299
B1		2.420
B2		2.443
ALL	8.119	10.401

224 processes on 4 racks (2 proc.s per node)

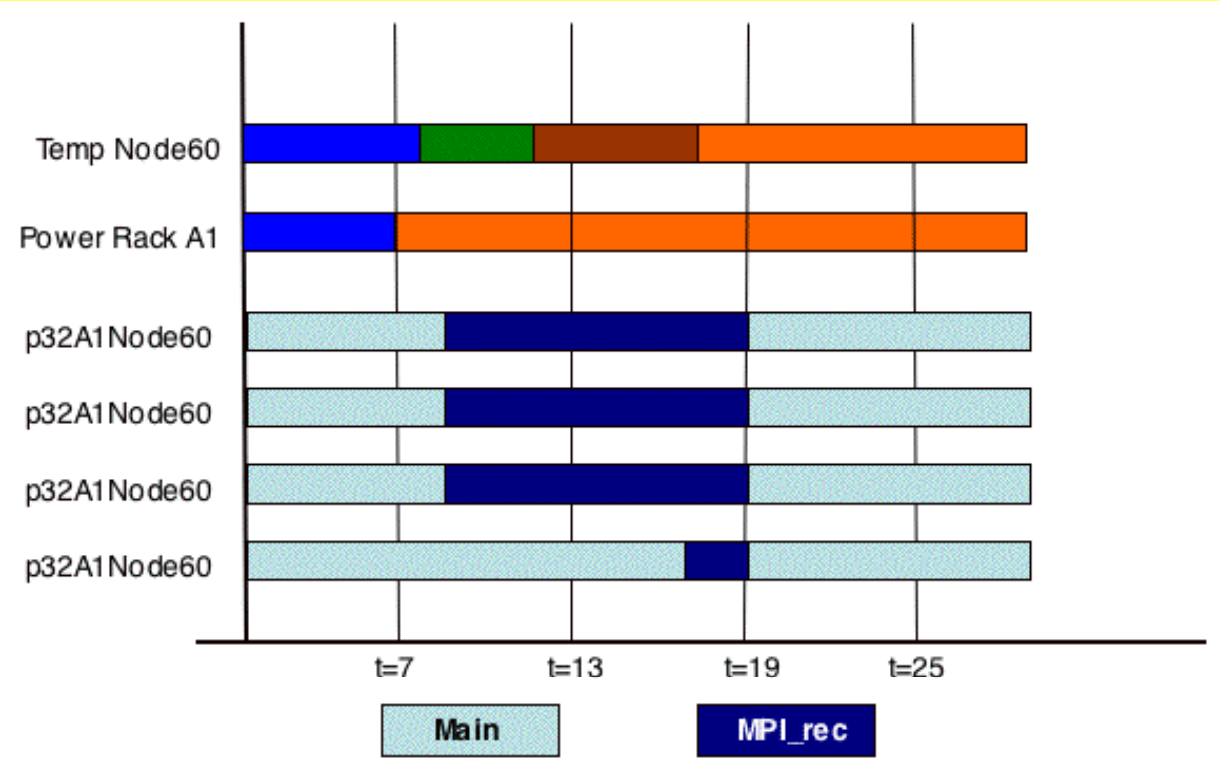


Low Density Placement (LD)
CPU temp's stay below 110 F



PerfTrack visi shows HD times significantly longer

New Visualizations for Integrated Data



Spectrum bars

Contact Information

Karen L. Karavanic

karavan@cs.pdx.edu

http://www.cs.pdx.edu/~karavan

***Of Interest* SC09: Paper presentation on Scalable Event Tracing by Kathryn Mohror**

Acknowledgments

This work was supported in part by funding from PNNL/Battelle, the Portland State University Center for Sustainable Processes and Practices, the James F. and Marion L. Miller Foundation, *UC/LLNL and the DOE Office of Science.*

- PerfTrack
  - Collaborator: Dr. John May (Lawrence Livermore National Laboratory)
  - *Contributing PSU students: Kathryn Mohror, Rashawn Knapp, Brian Pugh, Dylan Enloe, Aaron Amauba, Mike Smith, Agniv Adhikari*
  - *Contributing H.S. students: Travis Chapman (OSU), Thomas Conerly (CMU), Abraham Neben (Northwestern)*