Multi-System Data Collection With

Power Data Aggregation Monitor (PowerDAM)

EnviroInfo and ICT4S 2015

06.09 -09.09 2015

Dr. Hayk Shoukourian
Outline

• Why do we need a multi-system data collection?
• How do we tackle that with PowerDAM?
• What can PowerDAM currently do?
• What is the collected data currently used for?
  ▪ Operational Data Provision to MYNTS & Verification of Simulation Results
  ▪ Basis for modeling and predicting the energy and power consumption of large-scale applications executed on different HPC systems
Why do we need multi-system data collection?

- Collect data from all aspects of the data center
- Calculate the energy consumption of applications, evaluate system heat reuse, etc.
- Reports PUE, ERE, WUE, etc.
Why do we need multi-system data collection?

Data Collection
- Collect data from all aspects of the data center

Data Processing & Analysis
- Calculate EtS of applications

Reports
- KPIs (EtS, PUE, ERE, WUE, etc.)

PowerDAM

- HPC Systems
- Infrastructure
- Reuse Technology

© Hayk Shoukourian
Why do we need multi-system data collection?

Who is currently using PowerDAM?

PowerDAM

- Collect data from all aspects of the data center
- Calculate EET of applications...
- KPIs (ET, PUE, ERE, WUE, etc.)...

Infrastructure

HPC Systems

Reuse Technology

Agents

Plug-In

Data Collection

Data Processing & Analysis

Reports

References:


Cooling Networking

Infrastructure

Consumed Energy for user: [redacted] on system: mpp1

PUE of SuperMUC

overall average PUE  average PUE per timestamp

Created by PowerDAM
Publish/subscribe communication model via MQTT protocol

**Agents**
- Infrastructure
- HPC Systems
- Reuse Technology

**IClient Publisher**
- Publishing Topic
- Control Topic
- General Topic
- Registration Topic

**IClient Subscriber**
- Publishing Topic
- Control Topic
- General Topic
- Registration Topic

**BROKER**

**PowerDAM**
- Data Collection
- Data Processing & Analysis
- Reports

© Hayk Shoukourian
What else do we use the data for?

State of the Art and Perspectives

\[ TtS(n, f) = (1 - p) \left( 1 - \alpha_s + \alpha_s \frac{f_0}{f} \right) + \frac{p}{n} \left( 1 - \alpha_p + \alpha_p \frac{f_0}{f} \right) + O_{n,f} \]

\[ TtS(n, f) = \frac{t_1}{f} + \frac{t_2}{n} + \frac{t_3}{nf} + t_4n + t_5 \frac{n}{f} + t_6 \]

\[ APC(1, f) = P_{CPU,dynamic} + P_{CPU,idle} + P_{memory,dynamic} + P_{memory,idle} + P_{other} \]

\[ P_{CPU,dynamic} = P_{switching} + P_{short-circuit} + P_{static} \]

\[ P_{CPU,idle} \in O(1) \]

\[ APC(n, f) = k_1nf^3 + k_2n + k_3 \]

\[ EtS(n, f) = \int_{t_1}^{t_2} P(t)dt = APC(n, f) \cdot TtS(n, f) \]
Why do we need a predictor?

- Introduced average power consumption constraint during the maintenance of cooling towers.

- Can the job J be scheduled and the power consumption constraint preserved?
- What is the energy-optimal execution configuration?

```bash
#!/bin/bash
#@
job_type=parallel
#@
node = 270
...

  echo -n "Starting job J"
  mpiexec -n 270 ./myJobJ
  echo -n "Job J finished"
```

© Hayk Shoukourian
Some LACP prediction results

LACP \textbf{EtS} prediction results for \textbf{EPOCH} when executed using 287 compute nodes. Available data points are for 74, 112, and 210 compute nodes.

LACP \textbf{APC} prediction results for \textbf{EPOCH} when executed using 96 compute nodes. Available data points are for 74, 112, and 210 compute nodes.

\textbf{3} EtS, \textbf{1} APC, \textbf{5} TtS out of \textbf{68} HYDRO (available history of \textbf{4} runs) and \textbf{1} EtS, \textbf{3} APC, \textbf{8} TtS out of \textbf{106} EPOCH (available history of \textbf{3} runs) measurements had \textgreater= \textbf{10\%} error.
Differences In Node Power Draw

CoolMUC @ MPRIME

≈ 8.4 %
Differences In Node Power Draw

SuperMUC @ MPRIME

≈ 20 %
Upgrading the LACP model ...

\[ ||APC^j||_{min} = APC(j)^i - \sum_{u \text{ utilized node of } j}(P_u - P_{min}) \]

\[ ||APC^j||_{max} = APC(j)^i + \sum_{u \text{ utilized node of } j}(P_{max} - P_u) \]

System-vendor provided approximation for 311 nodes is: 118,108.47 W

LACP predicted maximum power consumption for node count 311 is: 55,993.13 W

Data center specified power cap
THANK YOU!