

NSF/IUCRC CAC PROJECT

Monitoring Power Usage of Jobs Running on Quanah Cluster

Jie Li

PhD Student, TTU

09/18/2019

Advisors:

Mr. Jon Hass, SW Architect, Dell Inc.

Dr. Alan Sill, Managing Director, HPCC, TTU

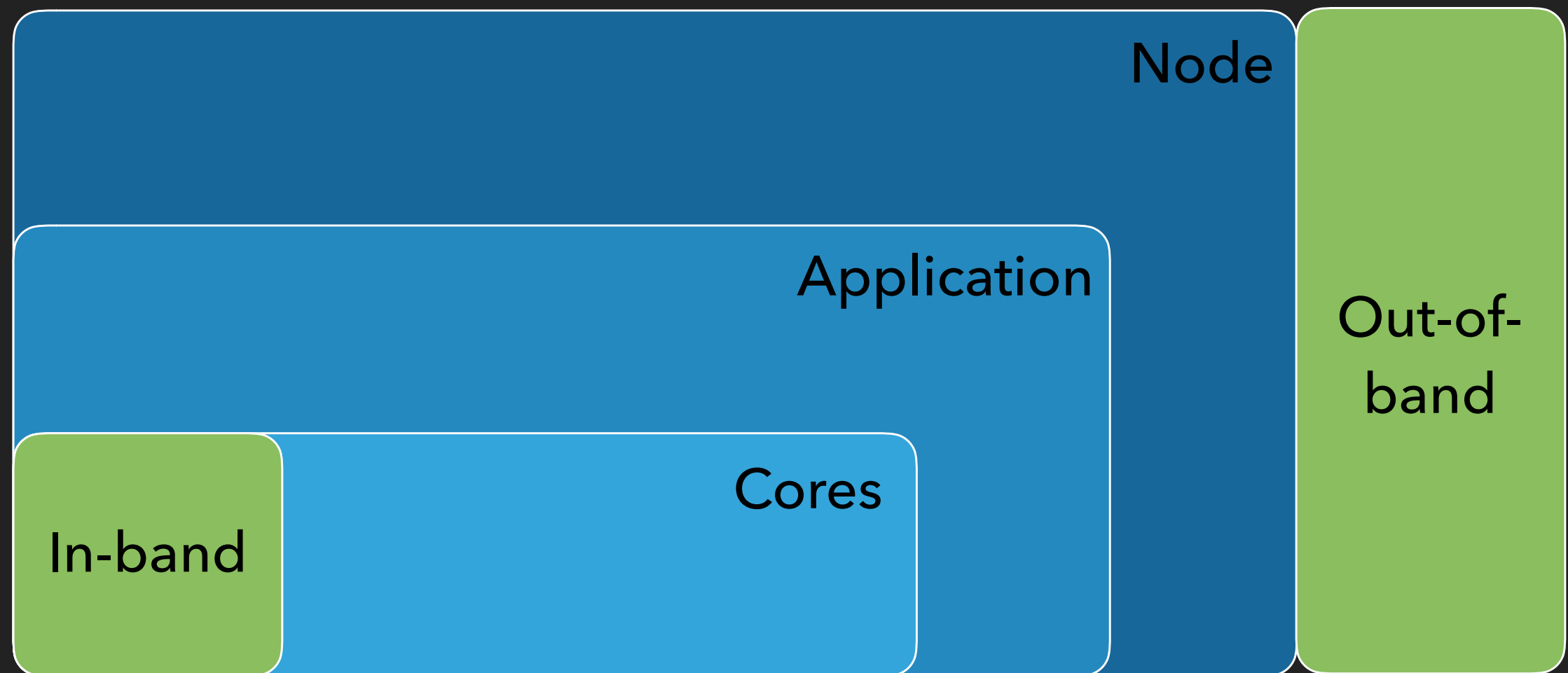
Dr. Yong Chen, Associate Professor, CS Dept, TTU

AGENDA

- ▶ Background
- ▶ Power Measurement Techniques
- ▶ Implementation
- ▶ Results and Future Work

Out-of-band Measurement	In-band Measurement
Use external equipment	Use device-level integrated measurement capabilities
Not provide easily accessible information to the running processes	Provide both CPU and memory subsystem energy measurement
Avoid perturbing the ongoing computation	Require active participation of the compute cores
Capture whole node energy profiles more easily	Not possible to capture whole node energy profiles

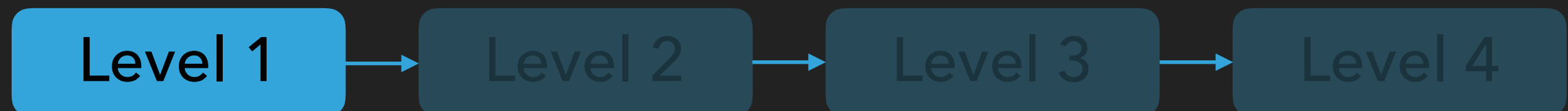
BACKGROUND



- ▶ Job-wide Aggregate Information
- ▶ Periodic Sampling
- ▶ Application Instrumentation
- ▶ Multi-Level Correlation

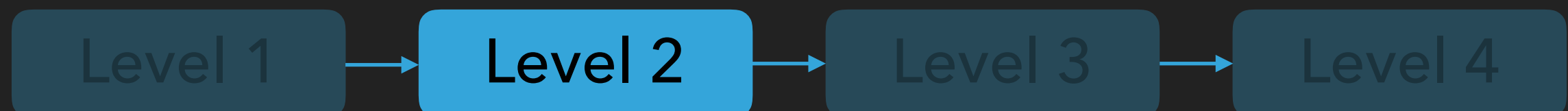
JOB-WIDE AGGREGATE INFORMATION

- ▶ **Coarse-grained** power and energy, total energy usage by each application.
- ▶ Information from each component is **aggregated** over an entire application run rather than point-in-time samples or per-node information.
- ▶ Help to understand how **energy-to-solution changes** for different application optimizations.
- ▶ It only provides insight into energy usage behavior in aggregate, **not the varying rates** of energy consumption throughout an application's execution.

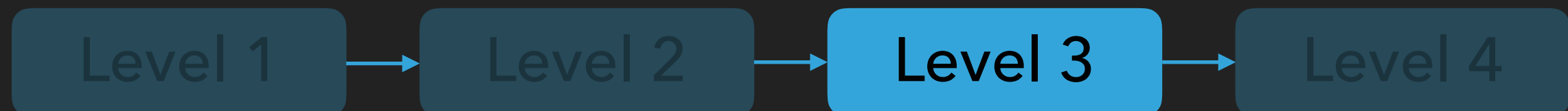


PERIODIC SAMPLING

- ▶ **Fine-grained** detail is provided by periodically sampling power levels and energy usage over time.
- ▶ Sampling may be implemented **in-band** or **out-of-band**.
- ▶ Information can be used to **plot power usage versus time**.
- ▶ **Retain** potentially **large volume** of point-in-time sample information

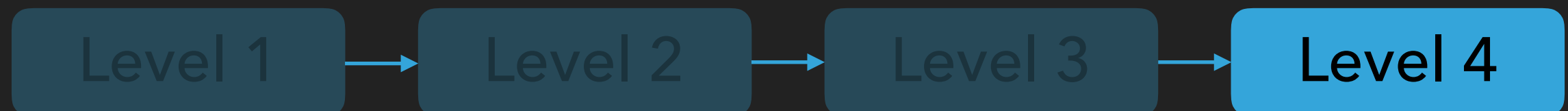


- ▶ A **finer level**, use application knowledge to interpret the recorded information
- ▶ Modify an application to **instrument code regions of interest**.
- ▶ Information can be analyzed to **characterize each instrumented region's power and energy usage behavior**.
- ▶ Require **application modifications** and may **reduce performance** due to instrumentation overhead.



MULTI-LEVEL CORRELATION

- ▶ **Cross-correlated** information gathered from previous levels.
- ▶ **Difficult to Align** in-band and out-of-band periodic sampling measurements can be difficult.



POWER MEASUREMENT TECHNIQUES

L1: Aggregate Information

Pros: Easily obtainable, quick summary info

Cons: Coarse-grained information

L2: Periodic Sampling

Pros: In-band or out-of-band sampling

Cons: Retain large volume of information

L3: Application Instrumentation

Pros: Users mark app regions, more info

Cons: Degrade application performance

L4: Multi-level Correlation

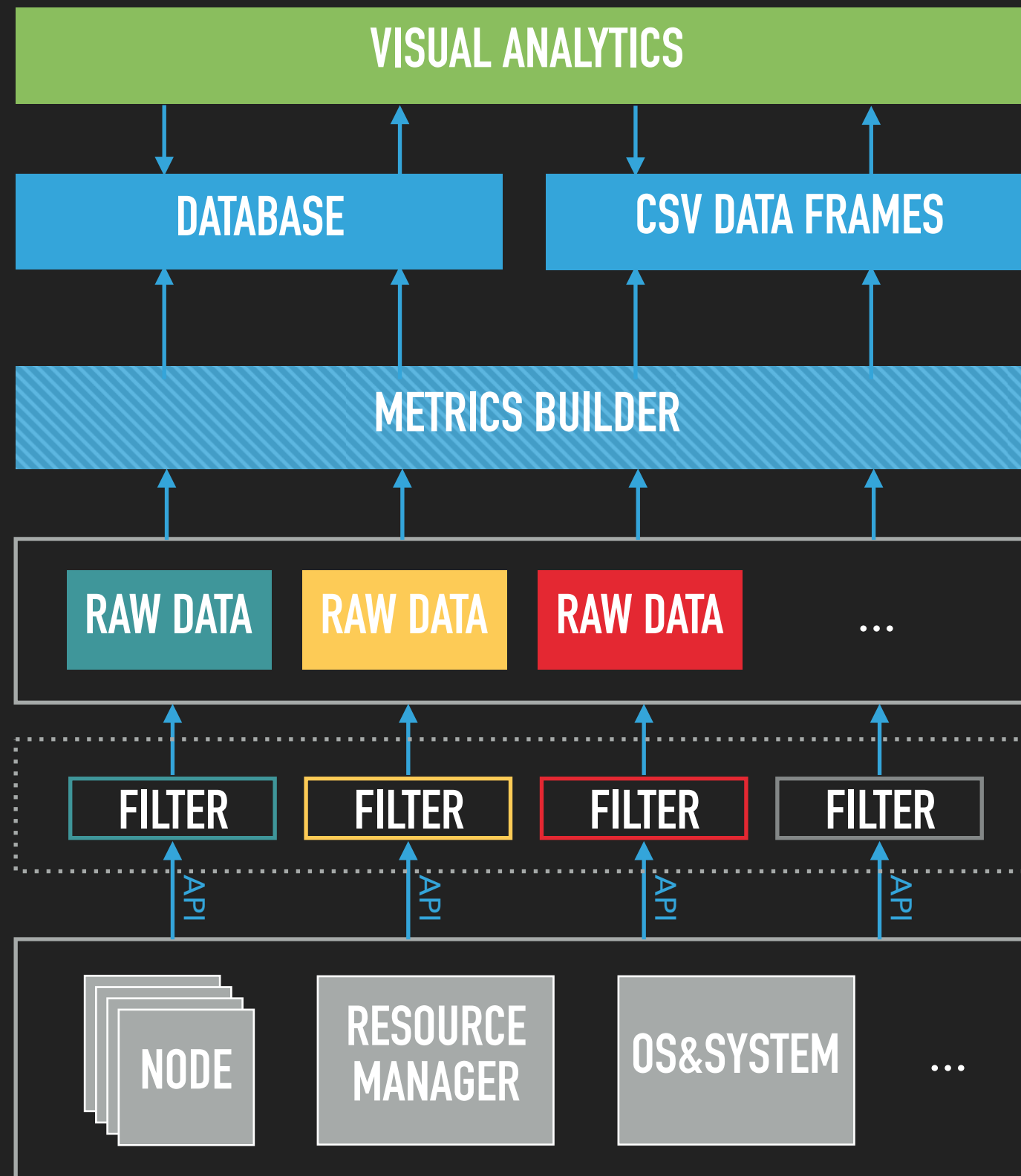
Pros: Information Fusion, intra-region insight

Cons: Synchronize metrics

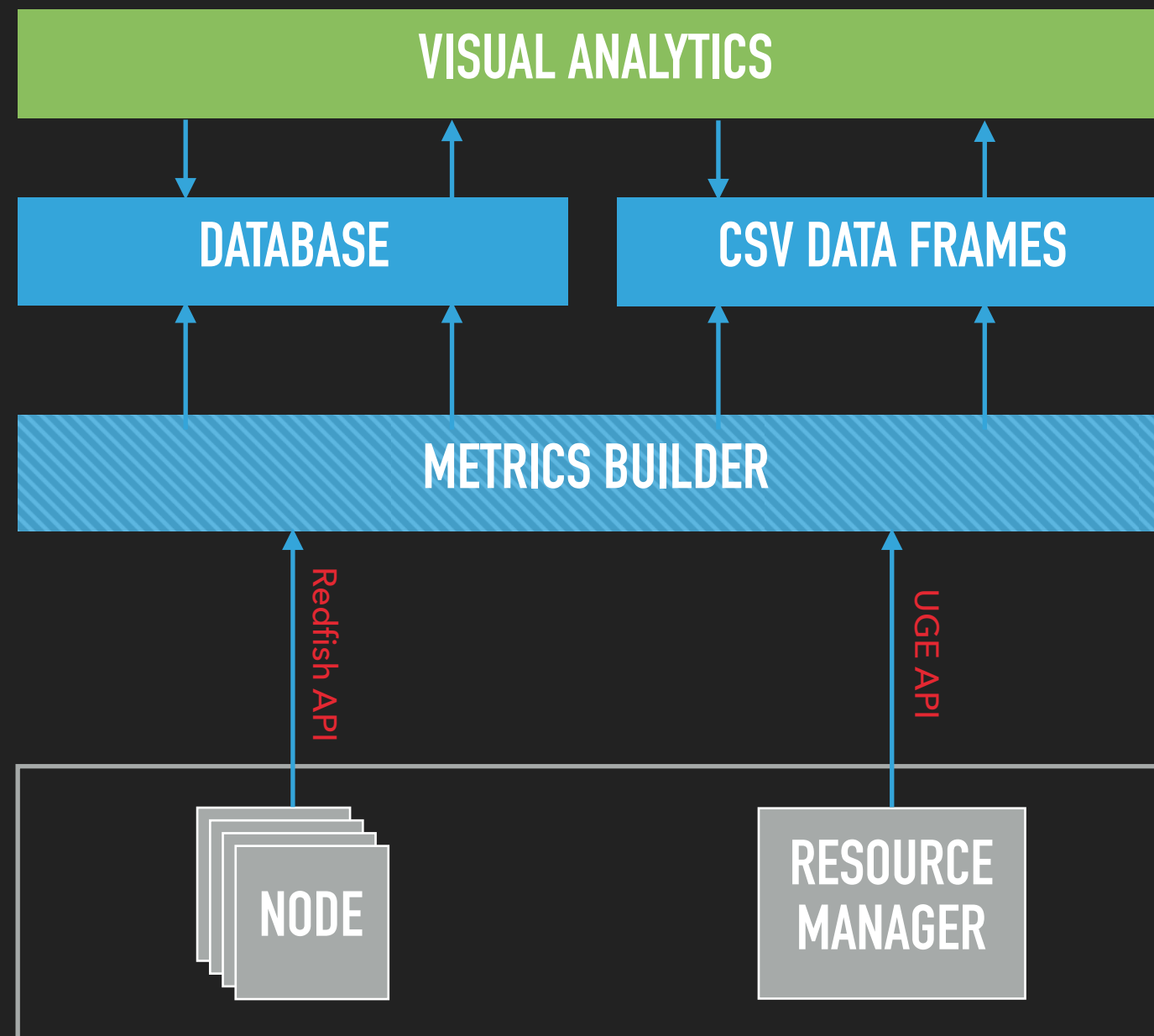
Effort



MONITORING FRAMEWORK

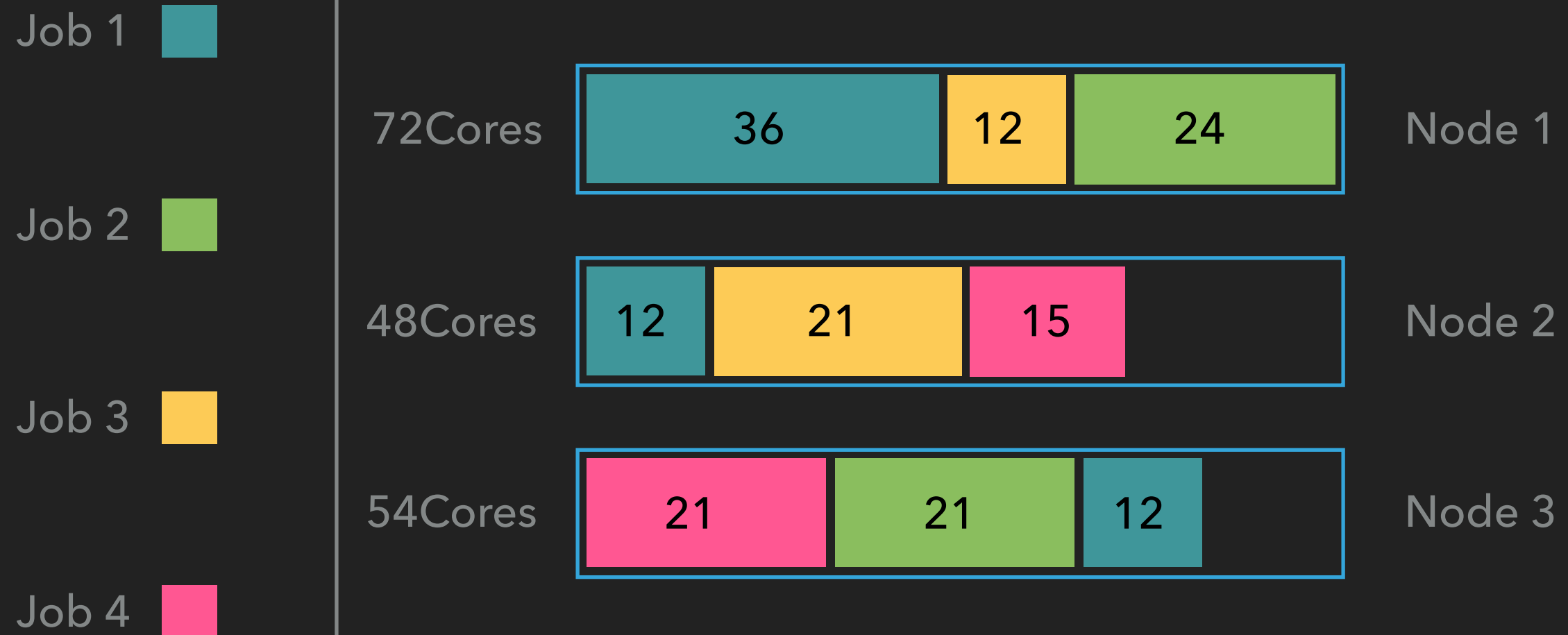


IMPLEMENTATION



How to get/estimate the power usage of each job?

METHODOLOGY



Assume power usage is proportional to the core usage!

METHODOLOGY

Job 1



Job 2



Job 3



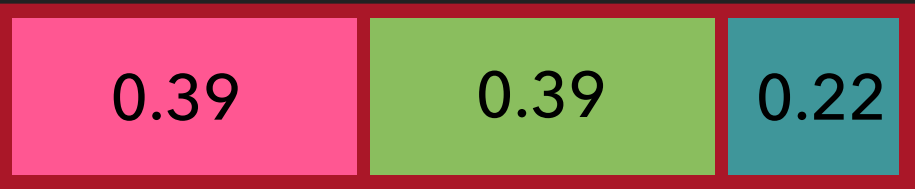
Job 4



Node 1



Node 2



Node 3

METHODOLOGY

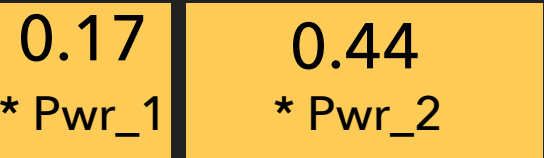
Job 1



Job 2



Job 3

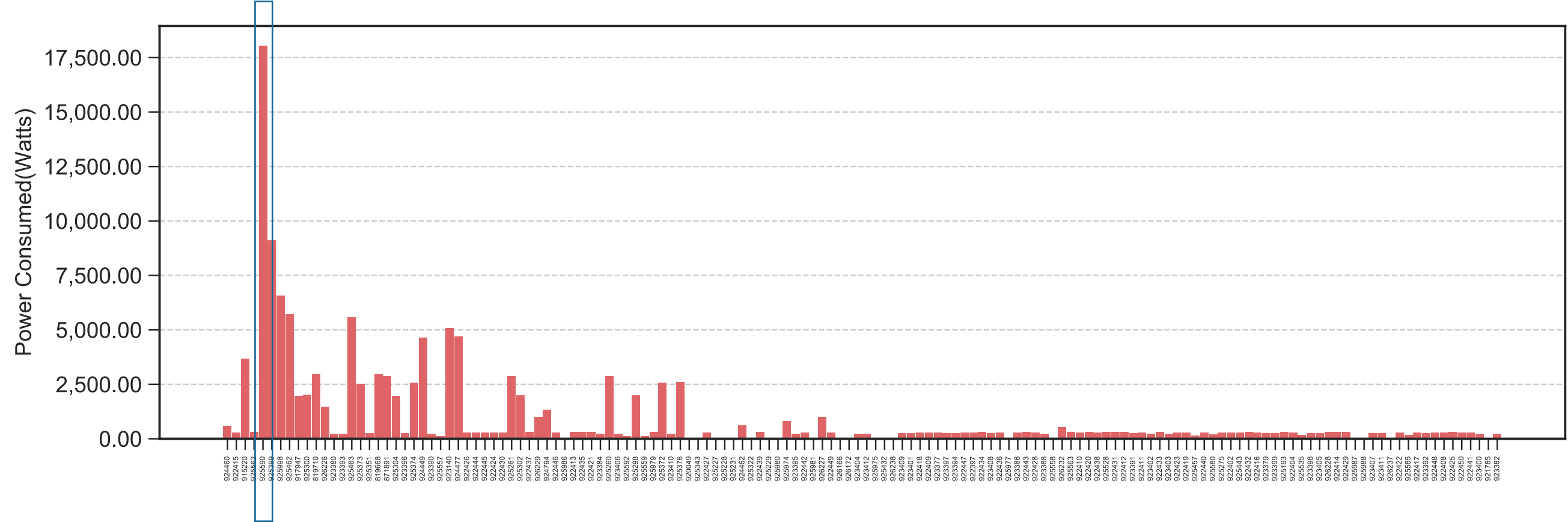


Job 4



Demo

RESULTS

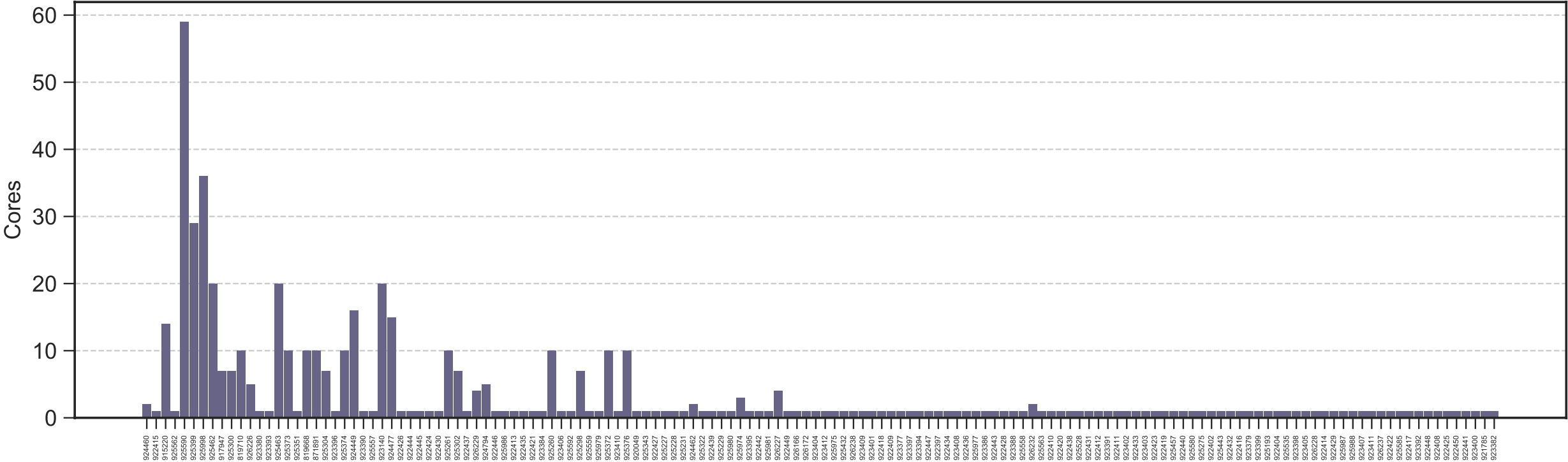


Job 925590

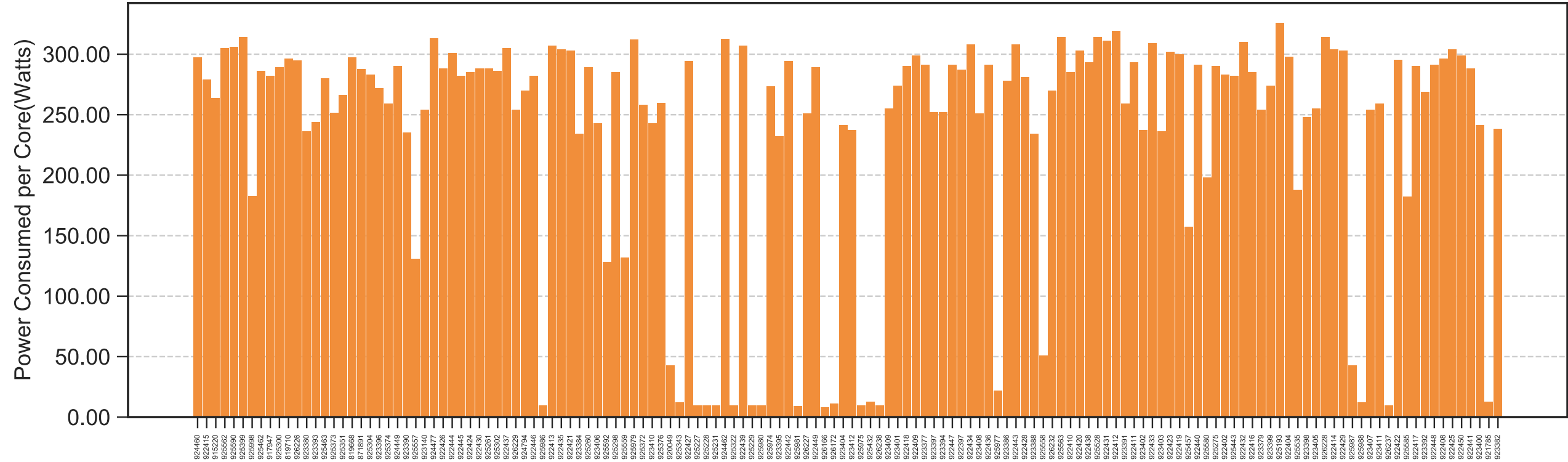
Power Consumed Per Job

Total Jobs: 144

RESULTS



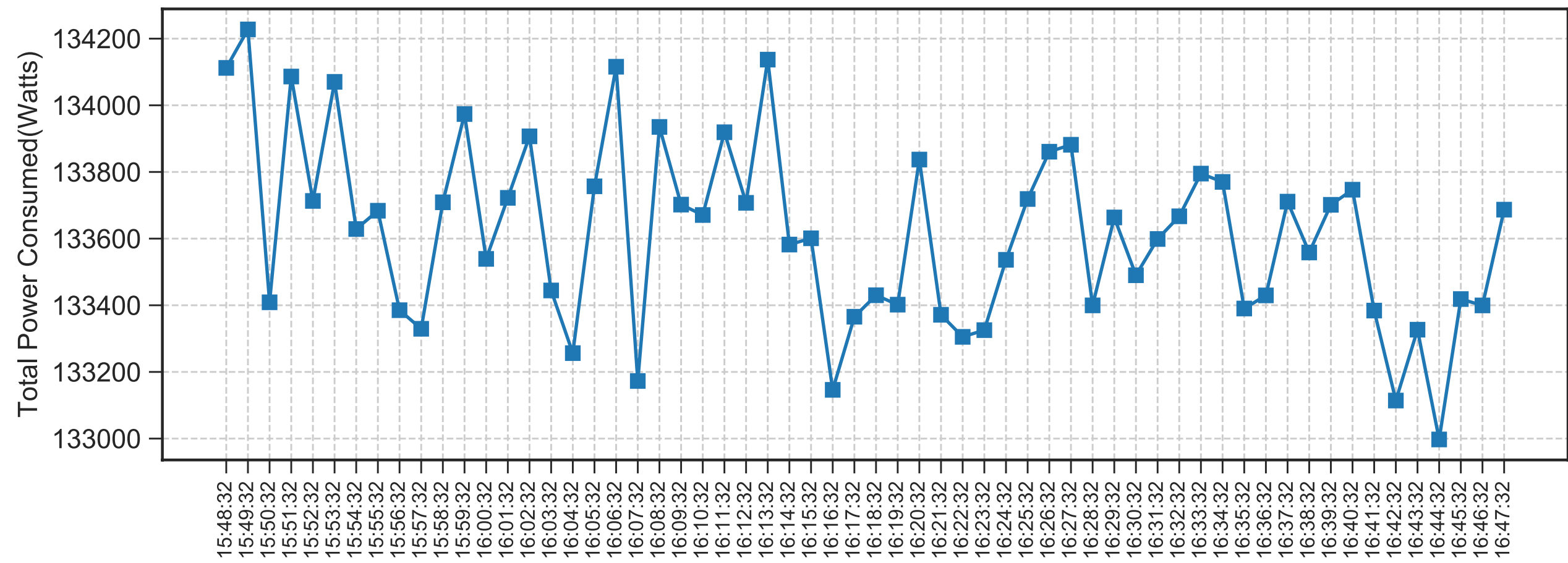
RESULTS



Average Power Used Per Core of Each Job

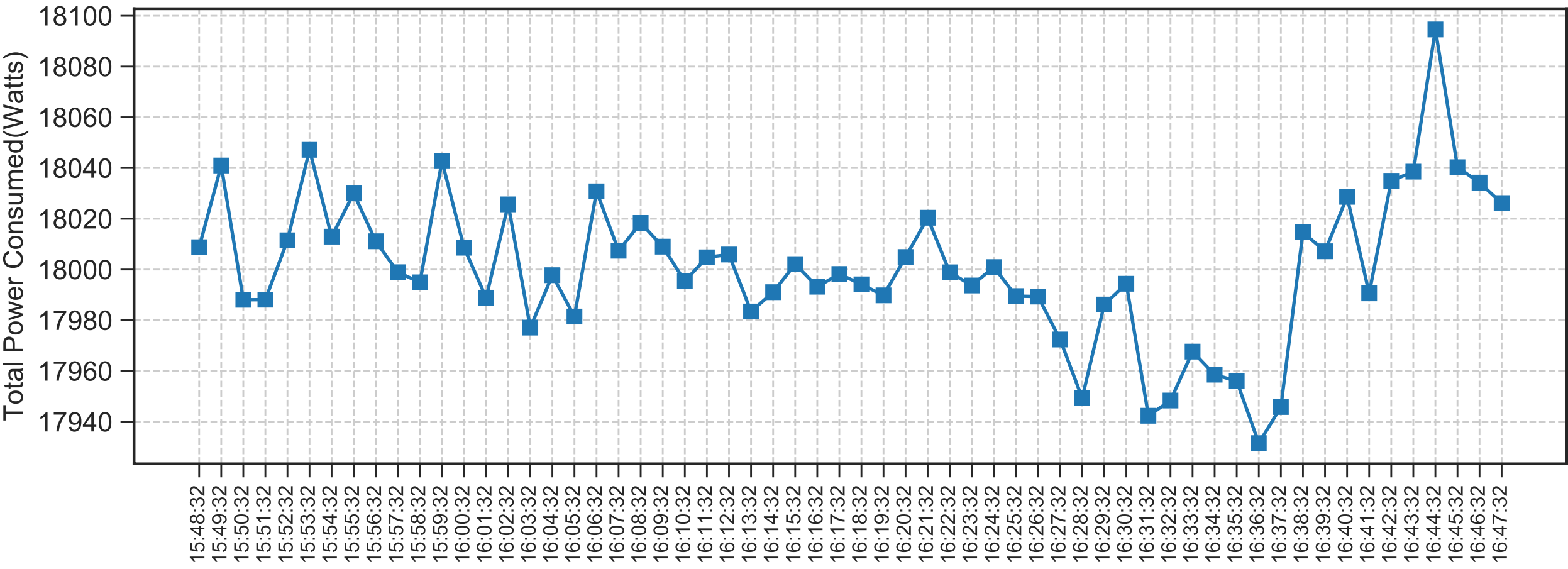
Total Jobs: 144

RESULTS



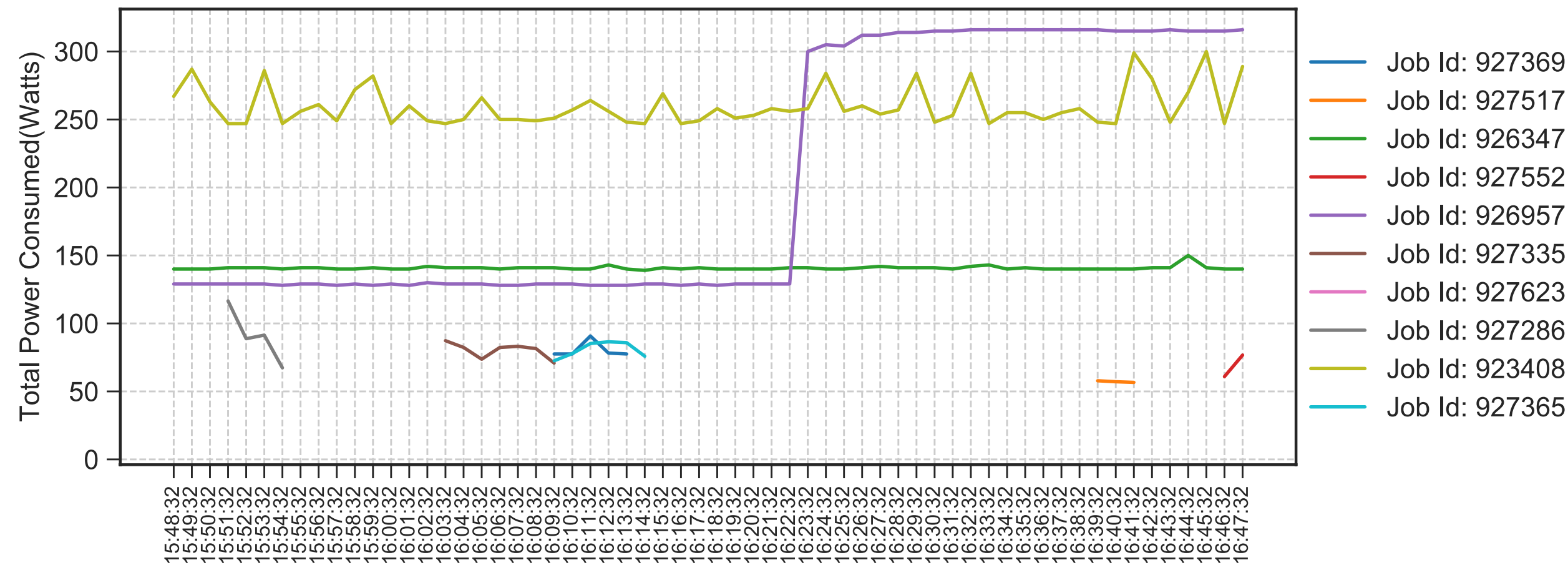
Total Power Consumed on Quannah Cluster

RESULTS



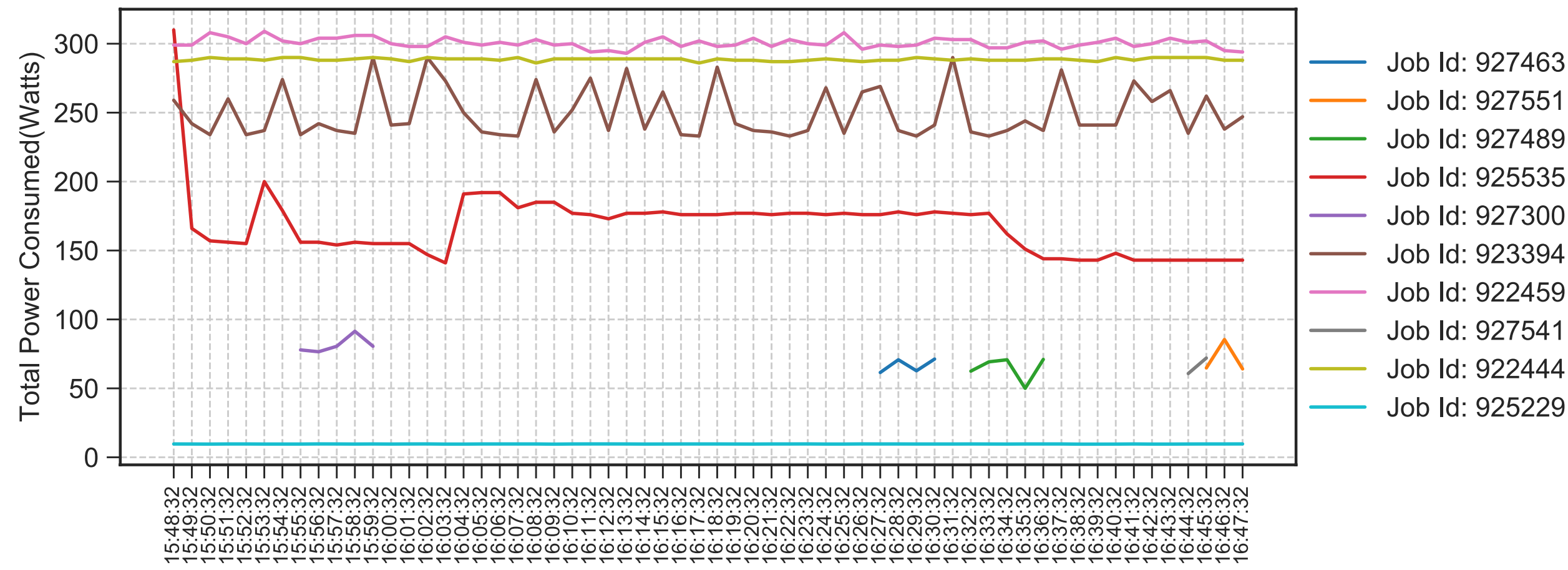
Total Power Consumed of Job 925590

RESULTS



Total Power Consumed of Some Random SelectedJobs

RESULTS



Total Power Consumed of Some Random SelectedJobs

FUTURE WORK

- ▶ Work with data visualization team to nail down the data frame structure
- ▶ Implement the filters
- ▶ Research on reducing the latency of fetching BMC metrics



QUESTIONS?/COMMENTS?