Power Usage in Data-Intensive Applications using MapReduce over MPI

Presenter: Paula Olaya

It is not all about runtimes

As we build workflows for data analytics, execution times are still traditional metrics of success

- Measuring and modeling performance is our bread and butter

But about power usage?

- Measuring performance versus power usage

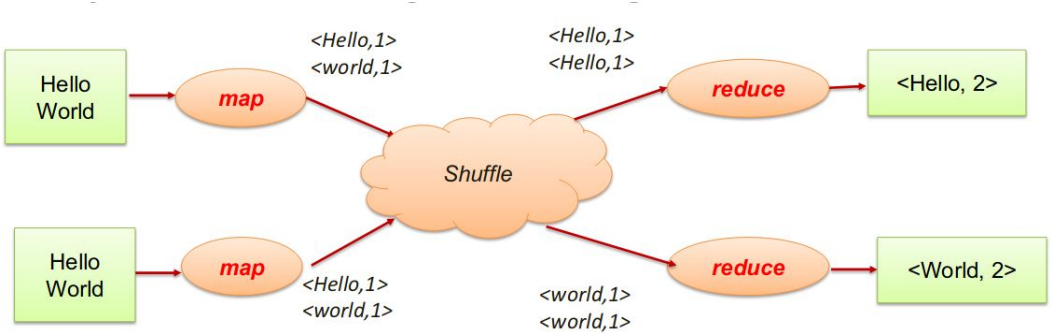
Measuring Power Usage

Building workflows for data analytics

- Measuring performance versus power usage
- Collaborators: P. Balaji (ANL), H. Jagode (ICL), Anthony Danalis (ICL)

Data Analytics in HPC

- Data analytics and data-intensive workloads are gaining representation at peta- and exascale
- **MapReduce** has gained the most traction in the HPC community



Mimir: a novel MapReduce over MPI framework

Mimir tackles:

- Skewed data
- Imbalance in memory usage
- Loss in data scalability

by implementing three optimizations:

- Combiner optimizations
- Dynamic repartitions
- Split method to handle datasets with superkeys

Mimir: Optimizations and Benchmarks

Optimizations	Benchmarks
• Combiner Optimizations • Dynamic Partitions • Superkeys and Splitting	• **WordCount (WC):** Number of each unique word. • **Octree Clustering (OC):** clustering algorithm for points in a three-dimensional space • **Breadth-First Search (BFS):** Traversal algorithm that generates a tree rooted at a source vertex • **Join:** Combines data from two or more tables by certain conditions

Mimir: Optimizations and Benchmarks

Benchmarks:

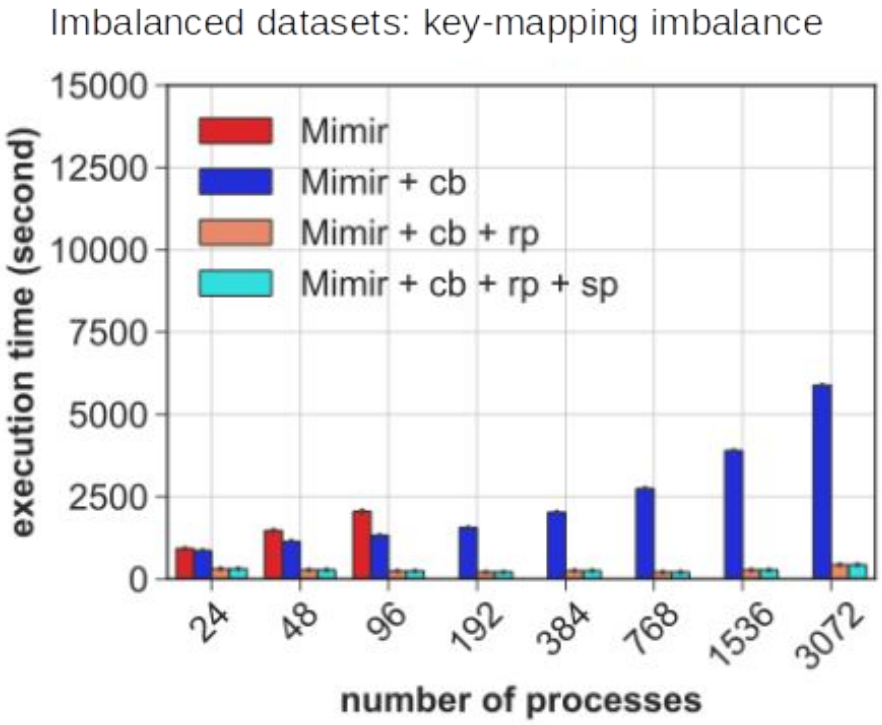
- **Octree clustering (OC)**

Optimizations:

- **Combiner Optimizations**
- **Dynamic Partition**
- **Superkeys and Splitting**

System:

- **Tianhe-2 128 24-core nodes**



Mimir: Optimizations and Benchmarks

Benchmarks:

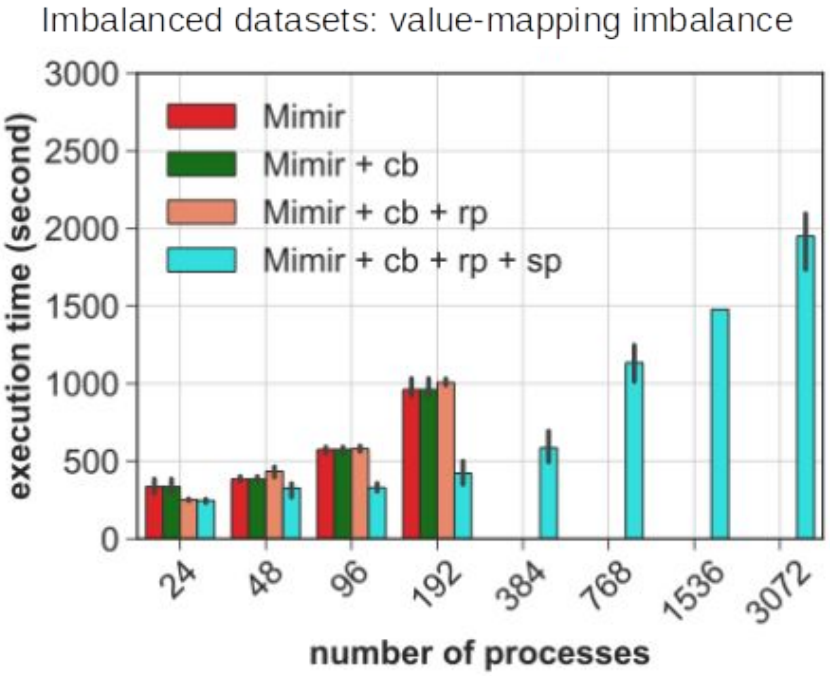
- **Join**

Optimizations:

- **Combiner Optimizations**
- **Dynamic Partition**
- **Superkeys and Splitting**

System:

- **Tianhe-2 128 24-core nodes**



Mimir: Power vs Performance

Optimizations:

- W/o local combiner
- With local combiner

Benchmarks:

- Word count (WC)

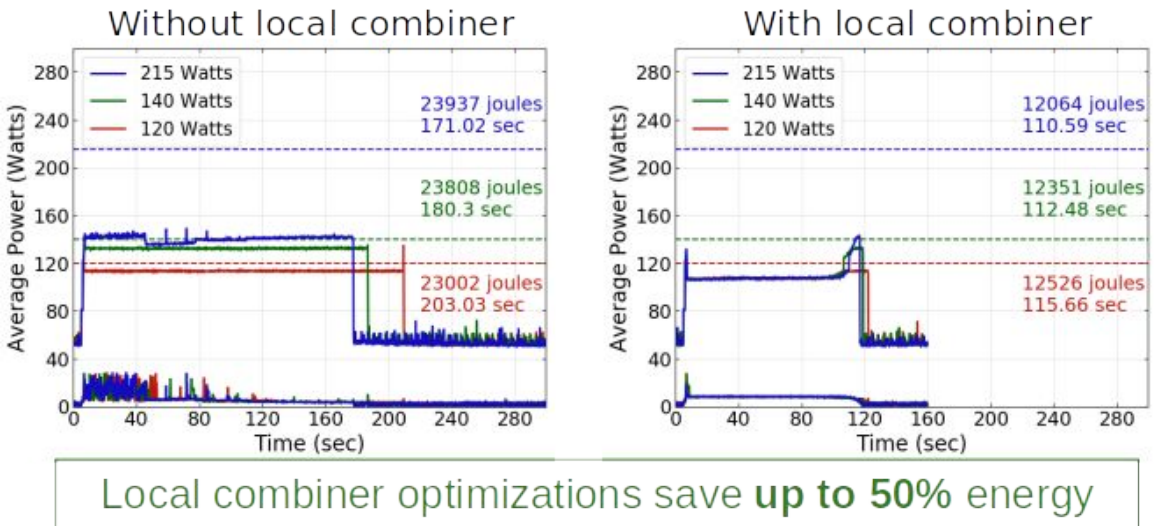
System:

- KNM (72 cores) + PAPI

Data:

- 4 billion words
- 72 unique words (4B72)

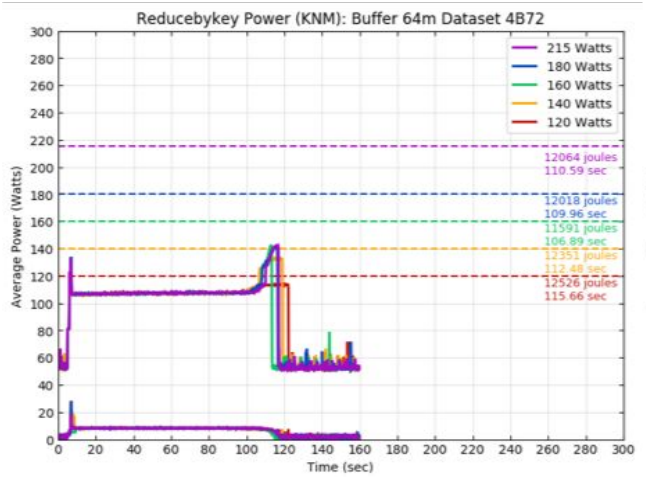
Data exhibits up to a 99% combinability rate before the shuffling



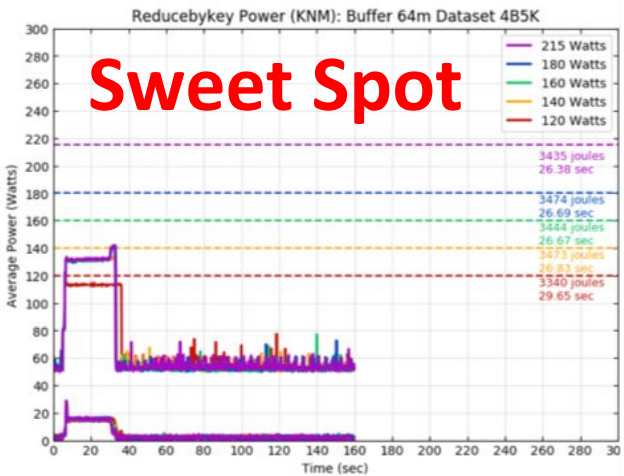
Local combiner optimizations save **up to 50%** energy

Mimir: Power vs Performance

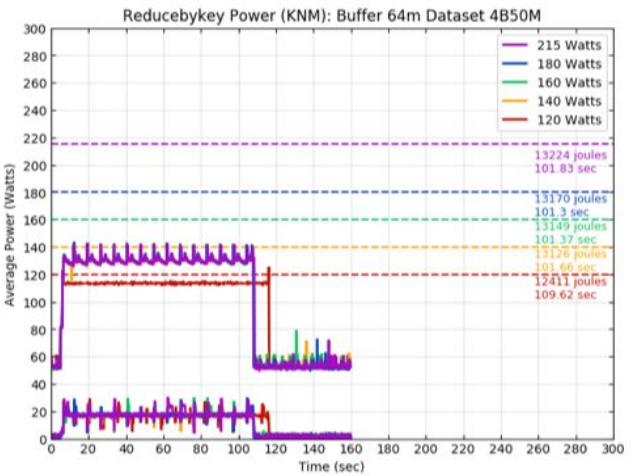
Data combinability ranges from 99% (4B72) to 10% (4B50M)



4B words,
72 unique words



4B words,
5K unique words



4B words,
28M unique words

Summary

- Optimizations on Mimir offer significant time and power performance benefits
- Power usage gives insight into data management patterns indistinguishable by runtime performance
  - Data analytics workflows are different than compute intensive workflows
  - The challenge is how to capture (and model) what factors that result in power usage