ENERGY-EFFICIENT VISUALIZATION PIPELINES

A CASE STUDY IN CLIMATE SIMULATION

Vignesh Adhinarayanan
Ph.D. (CS) Student
Synergy Lab, Virginia Tech

INTRODUCTION

"Supercomputers are constrained by power"

- Power budget for Los Alamos county = 66 MW
- Power budget for Trinity supercomputer alone = 15 MW

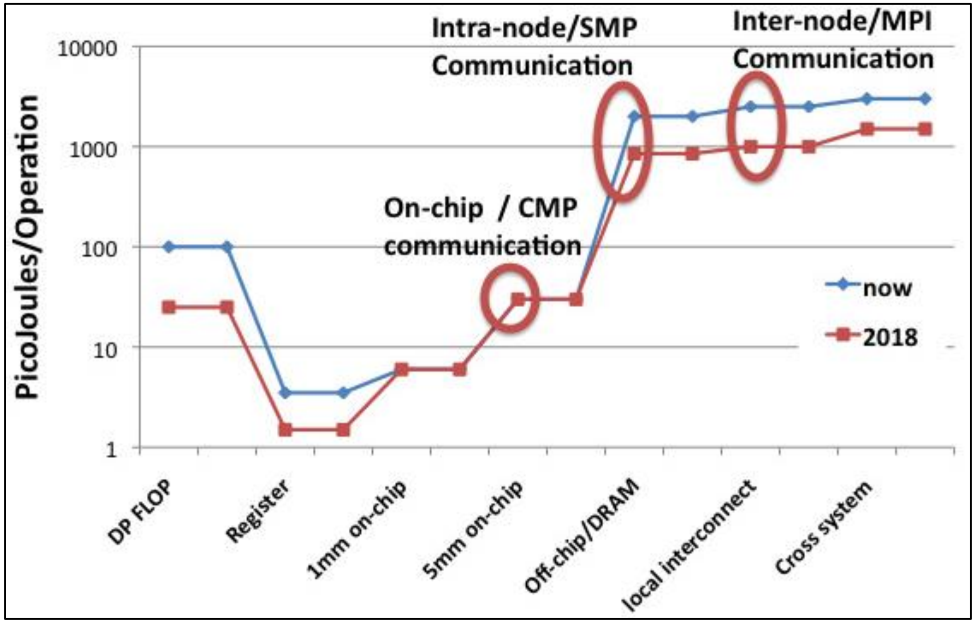
- Exceeding power budget ➔ Brownouts in Los Alamos
  - *Installing and starting ASCI White believed to play a part in the rolling California brownouts in 2001*

INTRODUCTION

"Supercomputers are constrained by energy"

- 1 MW power consumption → 1 million dollars per year
  - Operating cost of supercomputers is comparable to the acquisition cost
    - The gap is expected to narrow down in the future

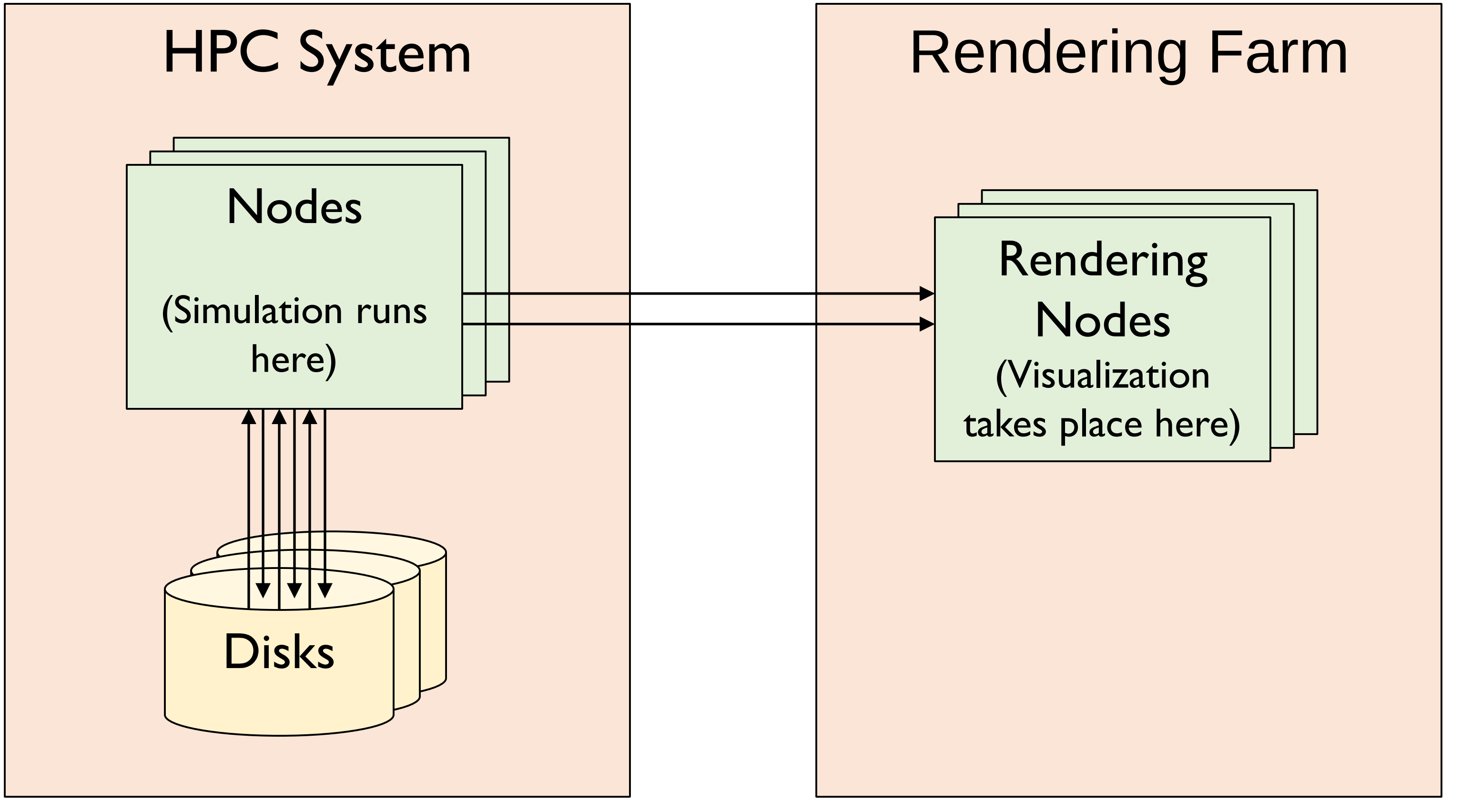
THE ENERGY CHALLENGE



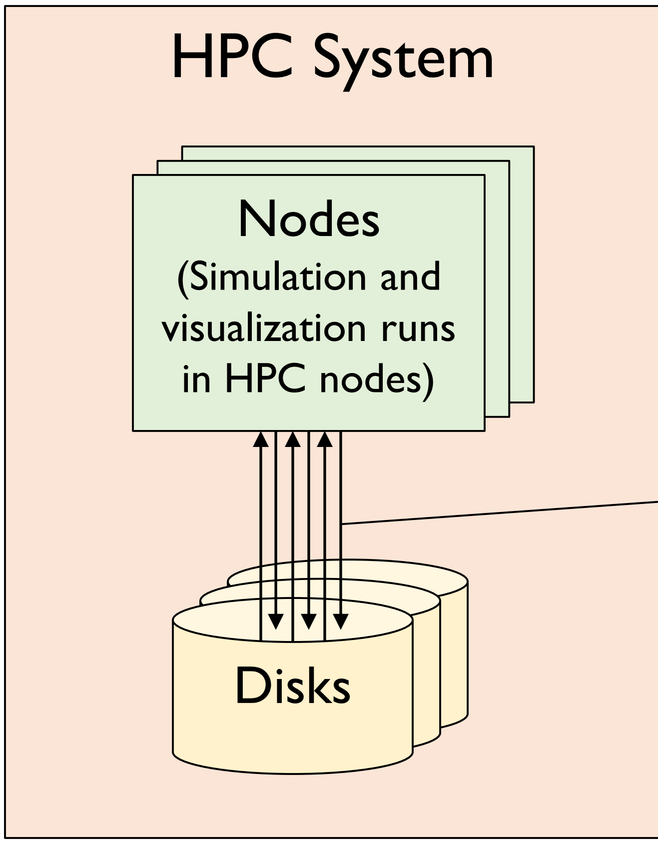
- Off-chip data movement cost nearly *hundred* times as much energy as on-chip data movement

Image source: J. Shalf et al., "Exascale Computing Technology Challenges", VECPAR 2010

Traditional "Post-Processing" Visualization



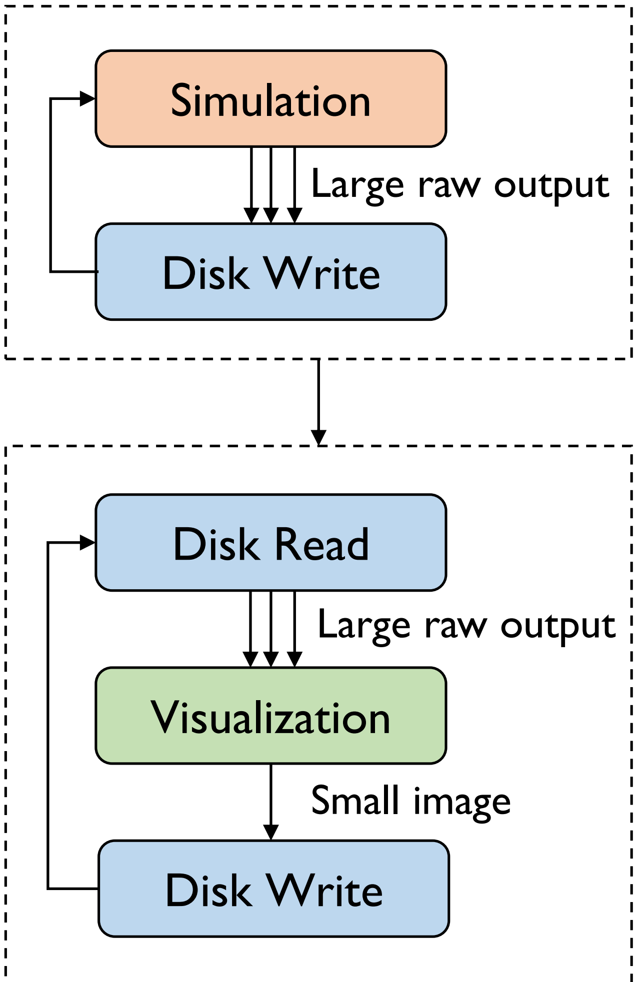
Modern "Post-Processing" Visualization



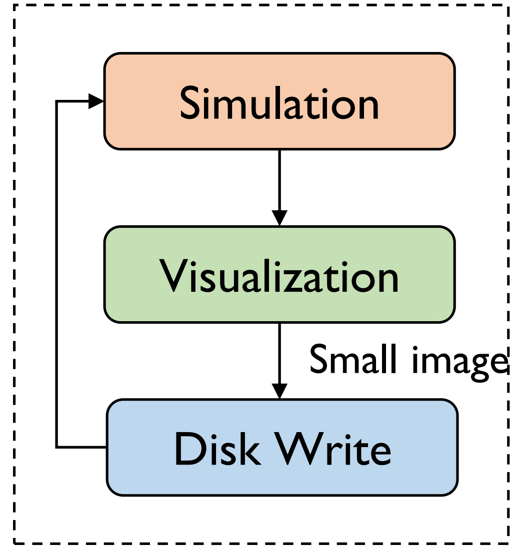
Also write raw output only every few iterations (i.e., temporal sampling technique is used)

But you may miss out on important simulation events

Post-Processing vs Insitu Pipelines



Post-processing



In-situ

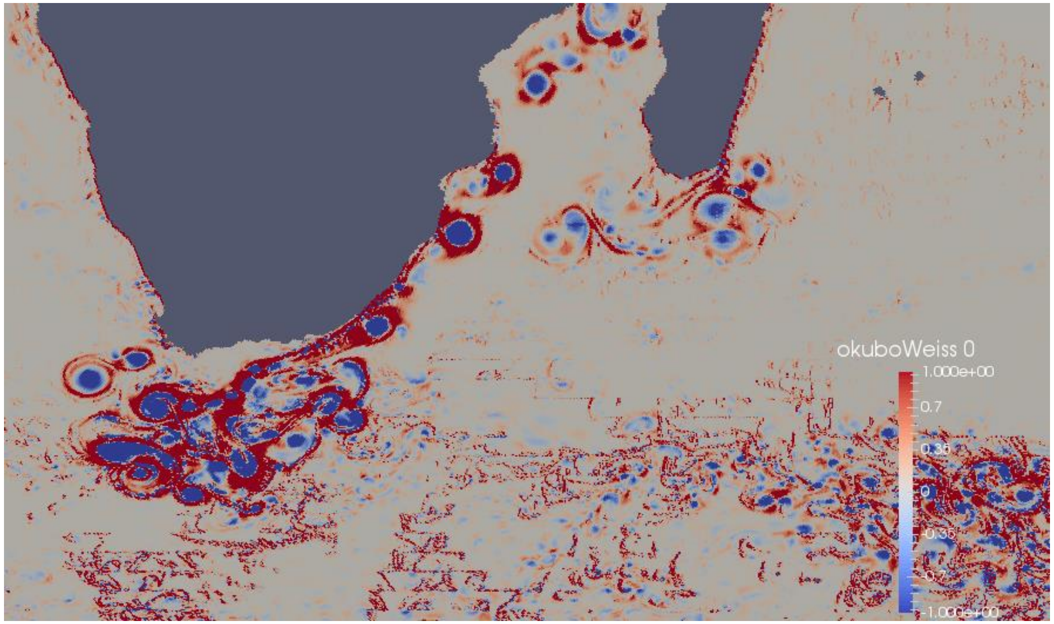
Traditional Post-Processing: Post-processing without any sampling
Modern Post-Processing: Post-processing with temporal sampling (write output every few iterations – here every 24 iterations)
In-situ: Produce images on the fly *and* do so only every few iterations

GOAL

"Study the performance, power, and energy trade-offs among traditional post-processing, modern post-processing, and in-situ visualization pipelines"

- Detailed sub-component level power measurements within a node to gain detailed insights
  - i.e., measure power consumption of CPU, memory, and disk
- Measurements at scale to understand problems unique to big supercomputers

APPLICATION



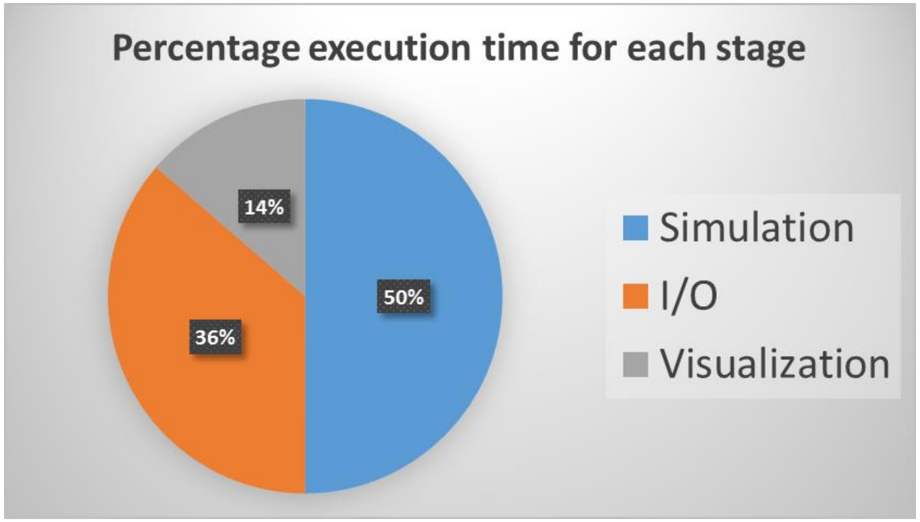
Eddies near Southern Africa

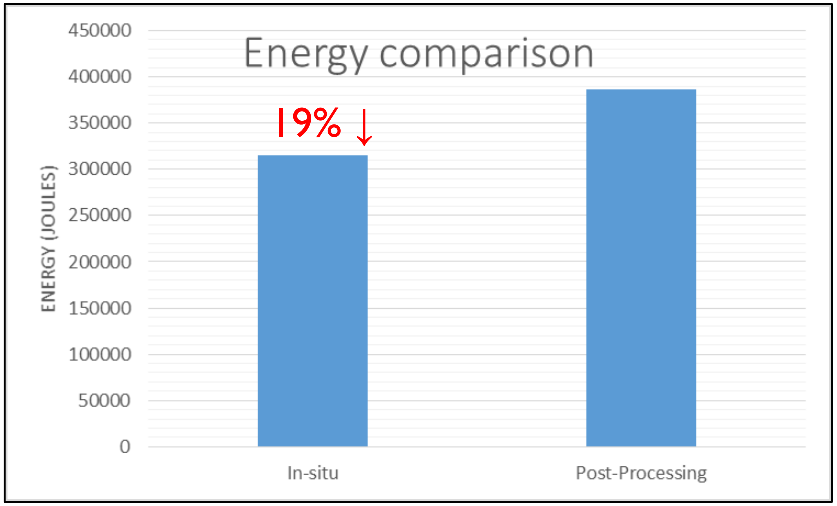
- *Modeling and Prediction Across Scale (MPAS) Ocean Simulation*
  - Solves an unstructured mesh problem
  - End goal: Identify eddies in the ocean

HARDWARE PLATFORM

- Compute nodes
  - 64 nodes
    - Each node contains 2x Intel Xeon E5-2670 and 64 GB of RAM
  - Nominal power consumption
    - 6000 W (idle) to 20000 W (workload such as MPAS)
- Storage nodes
  - Lustre file system
  - 5 nodes configured as 1 master + 2 MDS + 2 OSS
    - 1 RAID storage per MDS and OSS
  - Nominal power consumption
    - 2500W (idle) to 2800W (active)

ENERGY COMPARISON





Real measurements

Partial measurement and estimation

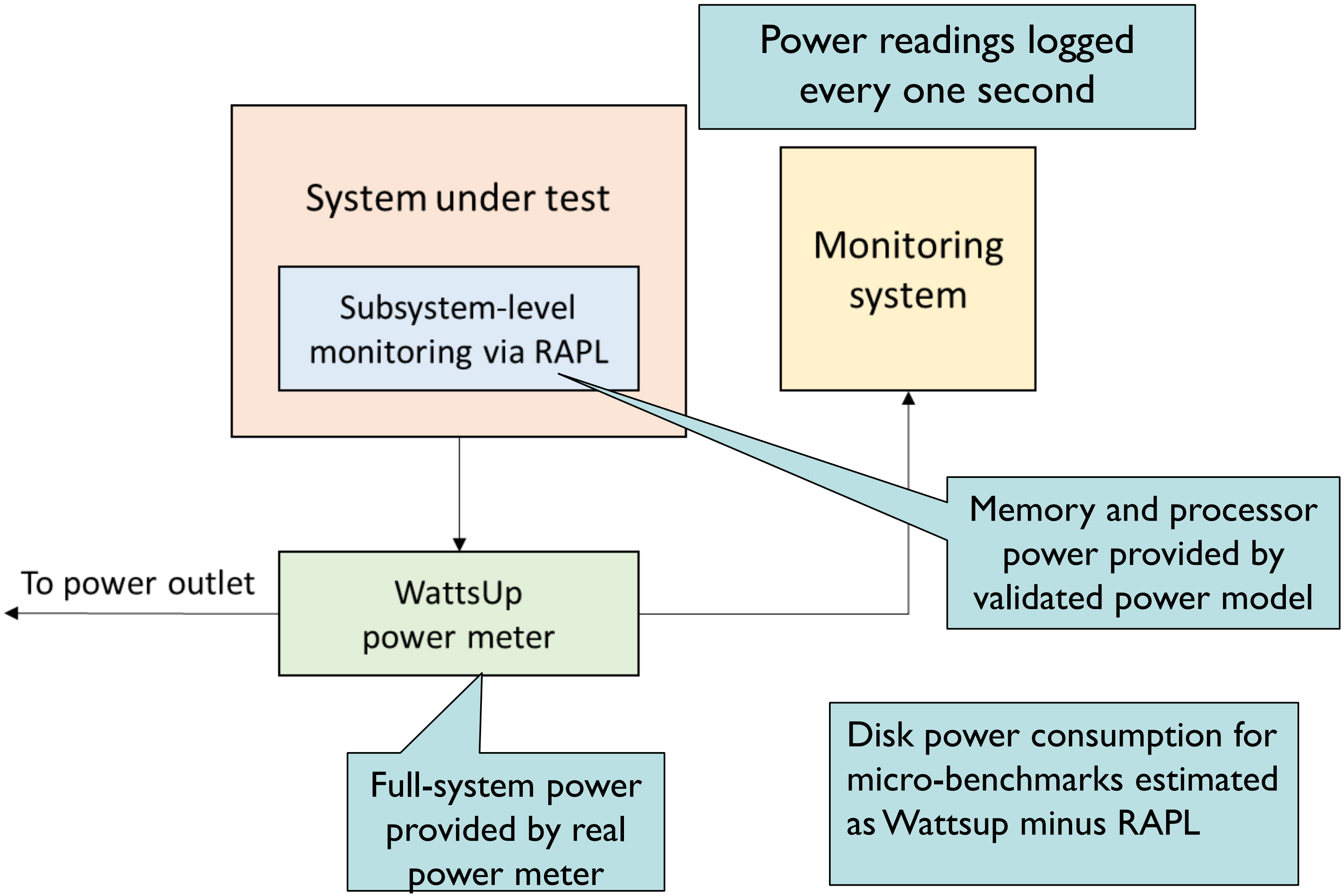
In-situ consumes 19% lower energy than post-processing

HARDWARE PLATFORM

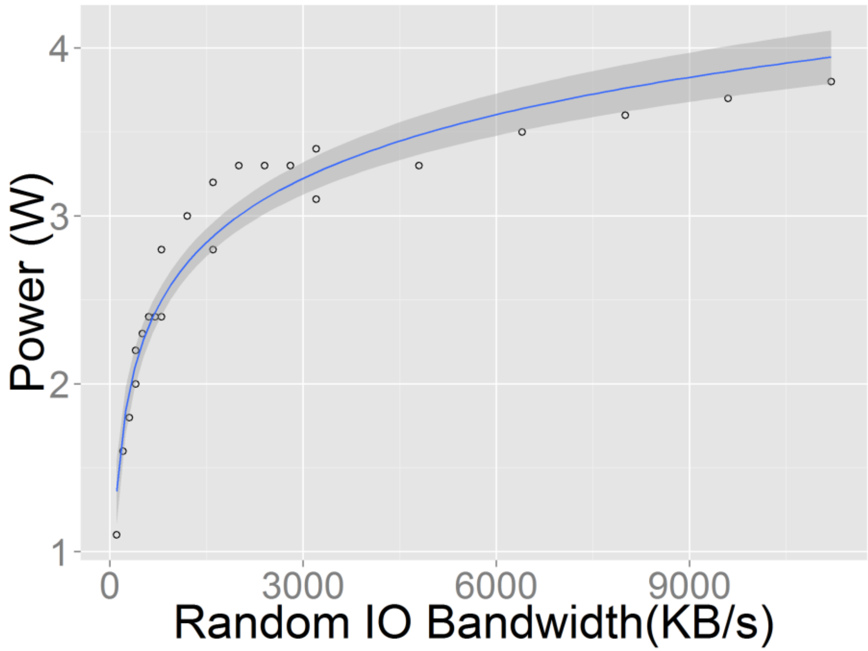
CPU	2x Intel Xeon E5-2665
CPU frequency	2.4 GHz
Last-level cache	20 MB
Memory	4x 16GB DDR3-1333
Memory size	64 GB
Hard disk	Seagate 7200rpm disk
Storage size	500GB
Disk bandwidth	6.0 Gbps

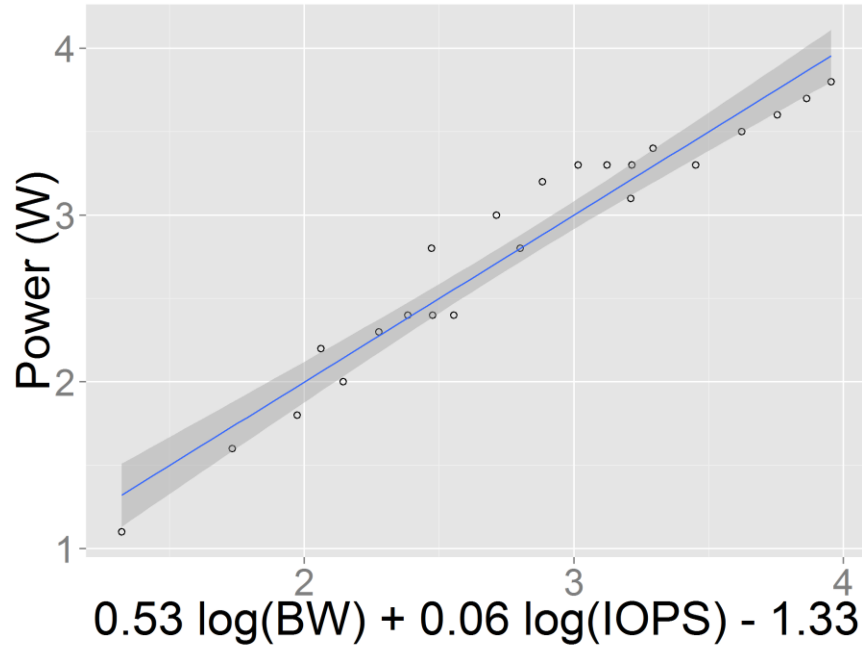
Hardware configuration

DATA COLLECTION



DISK POWER MODEL

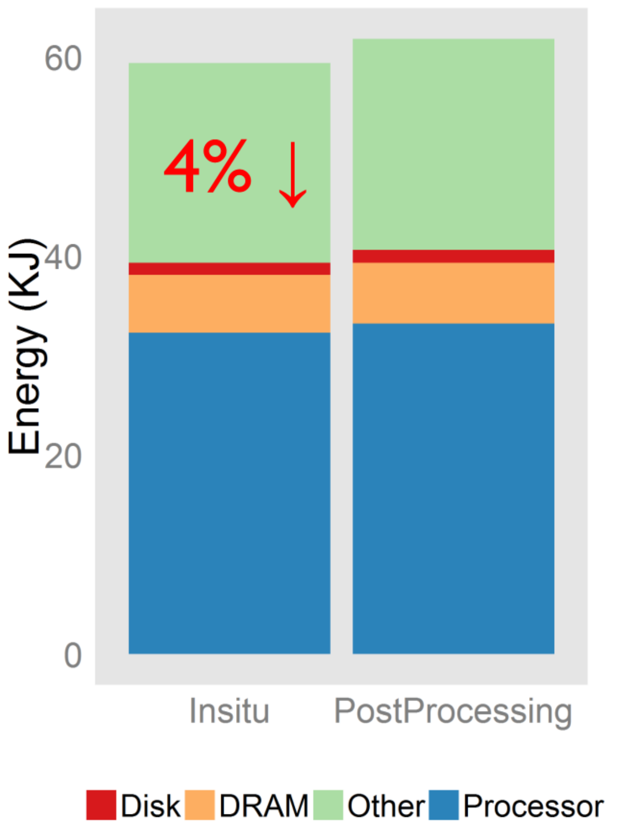


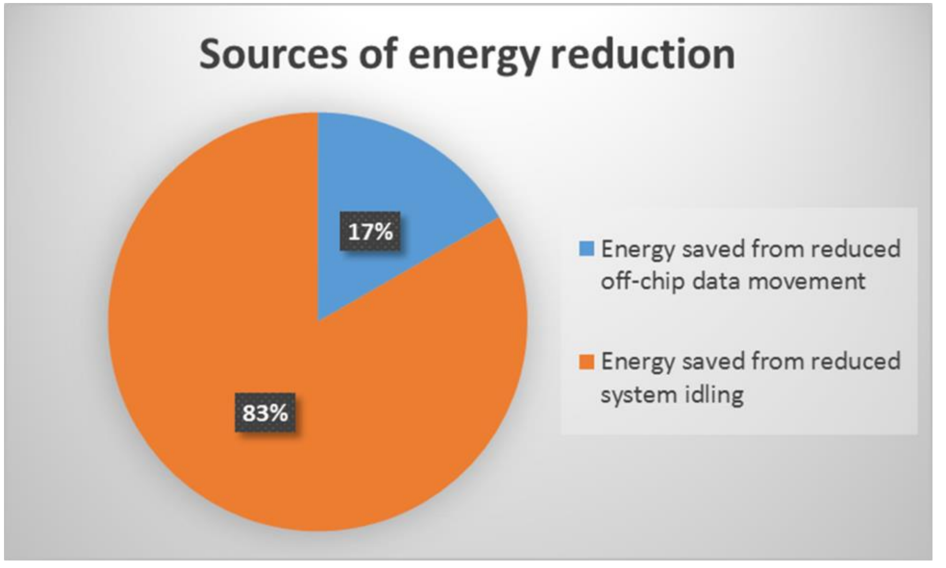


- Constant power from the *spinning* of disk
- Power consumption of read/write head dependent on *number of I/O operations*
- Power consumption of actual reads and writes dependent on *volume of data*

Single-Node Experiments

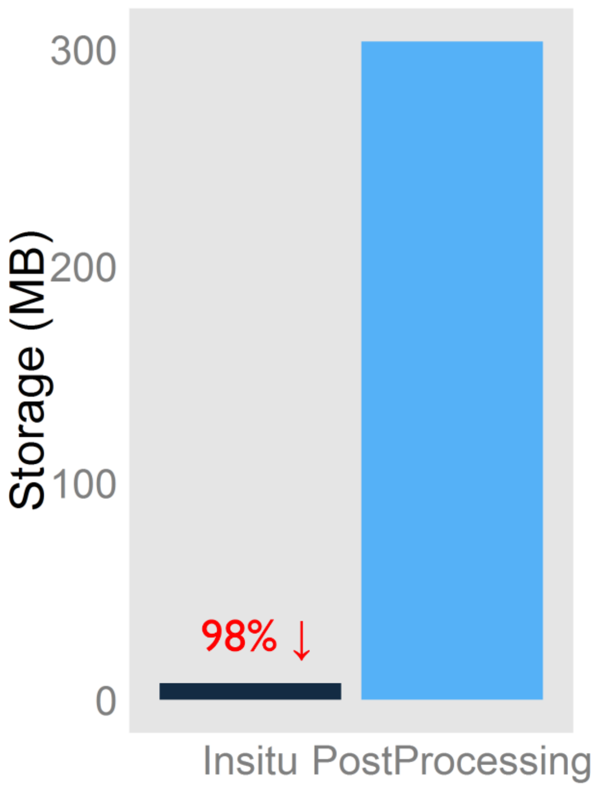
Energy Comparison





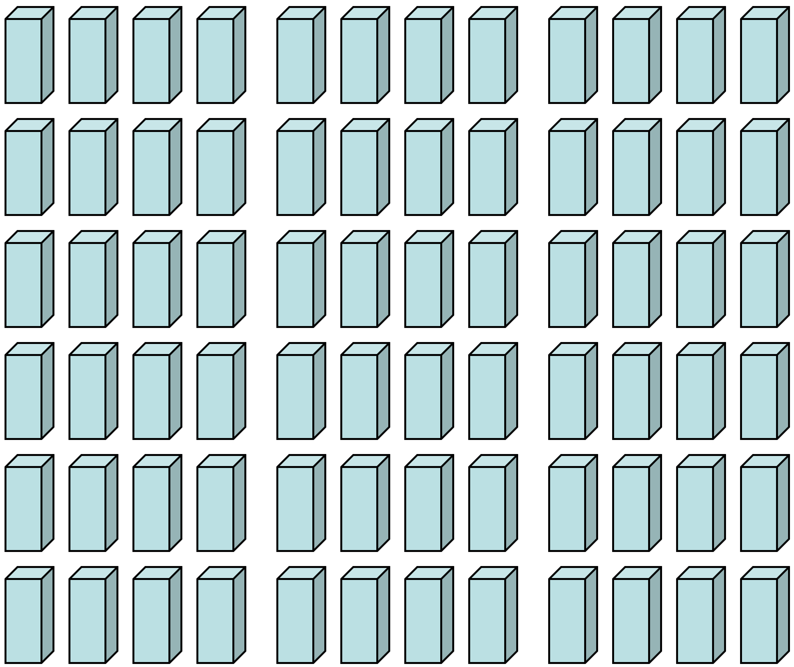
- Processor and memory consume lot of energy while waiting for I/O
- Worthwhile to minimize energy consumption while idling

Single-Node Experiments
Storage Requirements

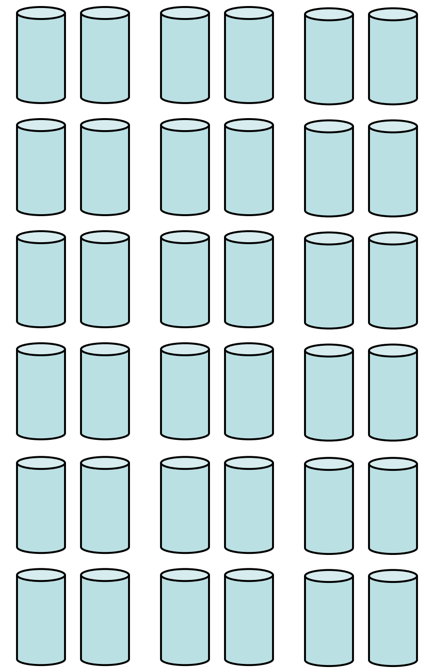


- ~97.5% lower storage requirement for the in-situ pipeline
  - Implies smaller storage cluster
  - Implies lower power consumption

Resources for Post-Processing

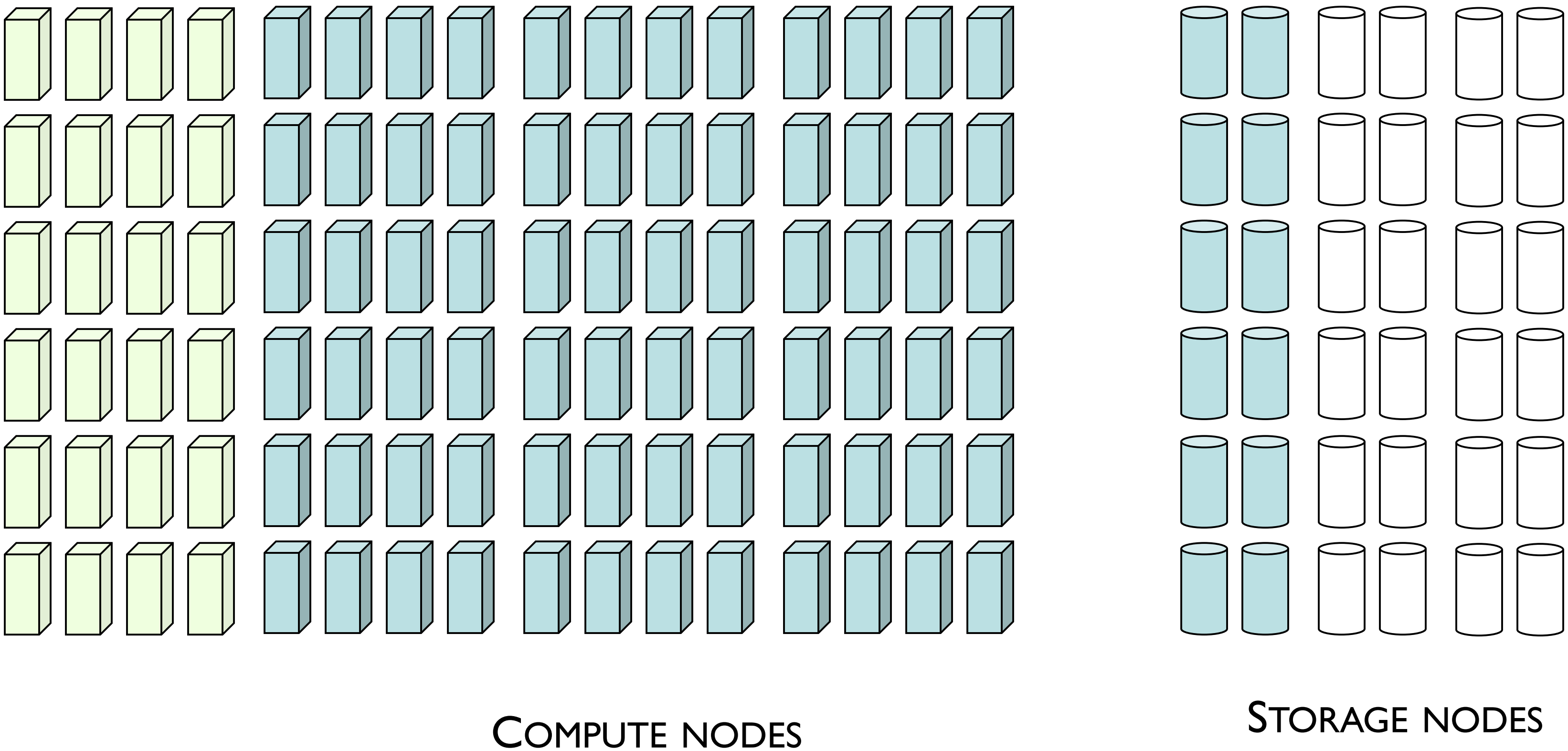


Compute nodes

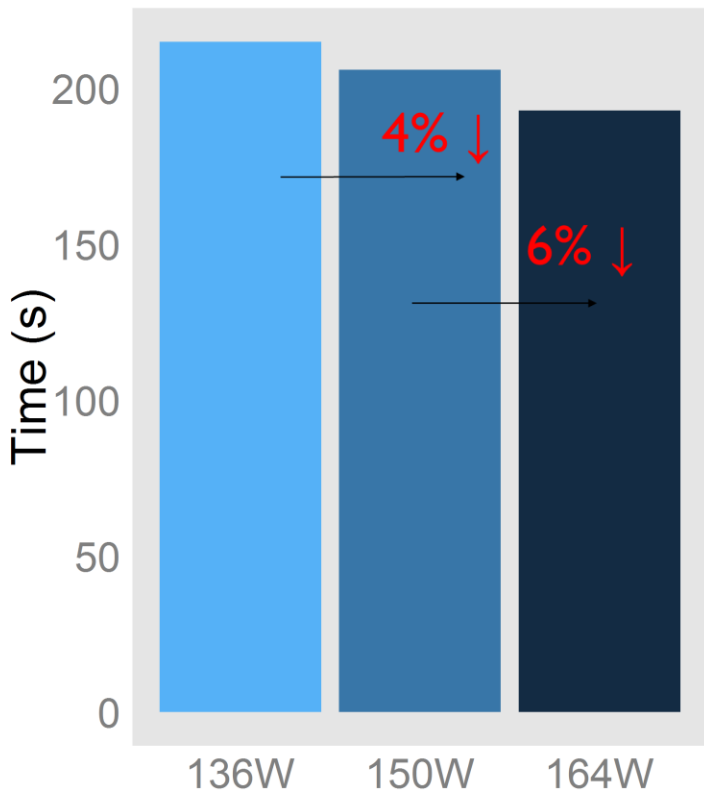


Storage nodes

Resources for Insitu



Redistributing Storage Power to Compute Nodes: Impact on Performance



Assuming reduced storage nodes results in 10% of total power redirected to compute nodes

- Performance improves by up to 6% for MPAS-O

FINDINGS

- Most energy savings come from reducing system idling (i.e., from reducing the I/O wait time)
- Further savings possible if we can reduced size of the storage nodes

CONCLUSION

- In-situ visualization offers the following advantages:
  - Reduced energy consumption (by reducing system idling or I/O wait time)
  - Reduced power (by using fewer storage nodes)
  - Improved performance (by reducing I/O wait time and by making more power available for compute nodes)

APPENDIX

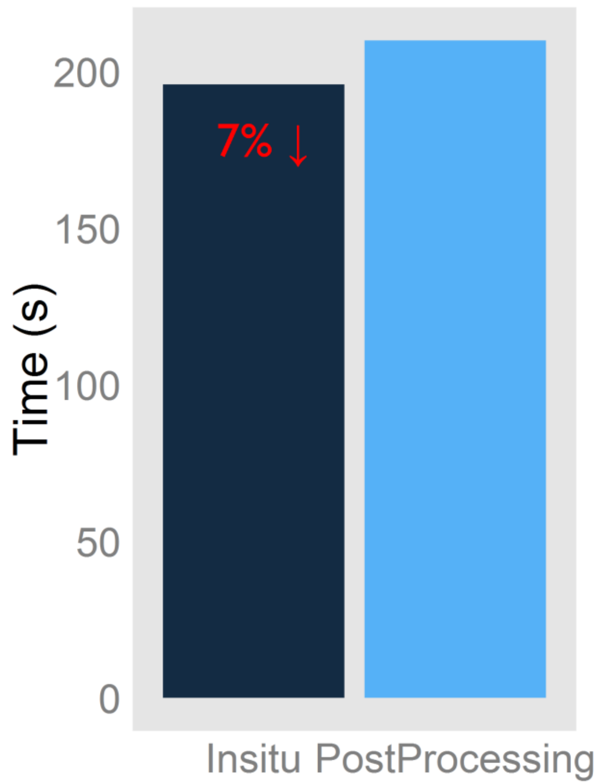
Expectations for a supercomputer

- Increased I/O wait time
  - Storage separated from compute by network
  - Longer execution time and corresponding increase in energy
- Additional energy consumption from data movement through the network
  - No data transfer via network cables in single-node
- Power/energy overhead for storage higher
  - Separate cluster for storage → additional CPUs, memory, cooling etc.
  - Storage sub-system shared with compute sub-system in single-node

Future Directions

- Enhancing HPC systems
  - Flash buffers and SSDs can reduce I/O wait time
    - Downside: Introducing more components can increase power consumption
- HPC system design changes
  - Bringing storage nodes and compute nodes together
    - Similar to Memory in Processor or Processor in Memory concepts in the computer architecture community
- Runtime system changes
  - Energy proportional computing and storage
    - Putting compute nodes to sleep states during I/O
    - Putting some storage nodes to deep sleep state when bandwidth and storage requirements are lower

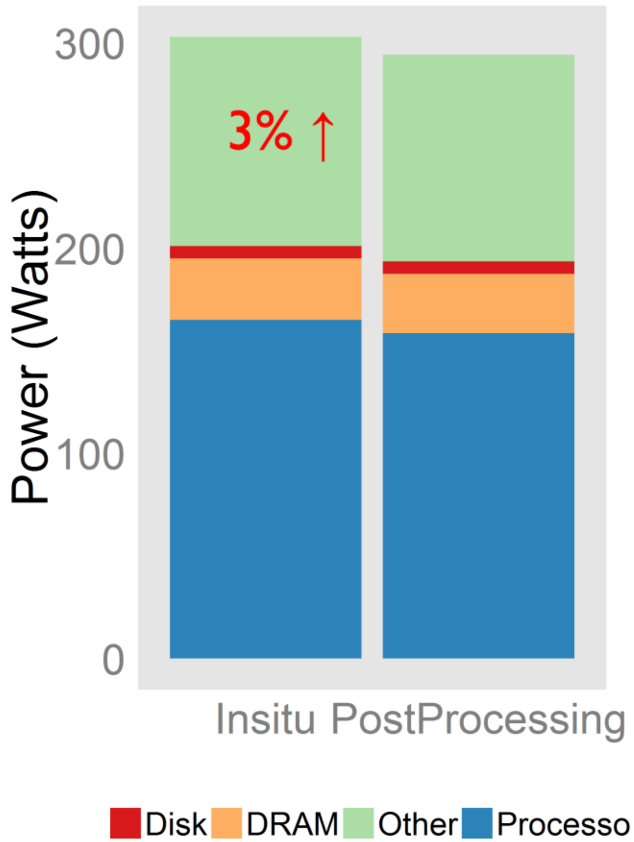
EXECUTION-TIME COMPARISON



- In-situ consumes 7% lower execution time than modern post-processing
  - Reduced I/O wait time
- The difference will be significant for an HPC system
  - Details later

Single-Node Experiments

Power Comparison



- In-situ consumes 3% more power than *modern* post-processing
  - Difficult trade-off choice

- Might not be the same for a supercomputer
  - Details later