

GeekPie_HPC

ShanghaiTech University



TEAM

GeekPie_HPC



hpc
is now.

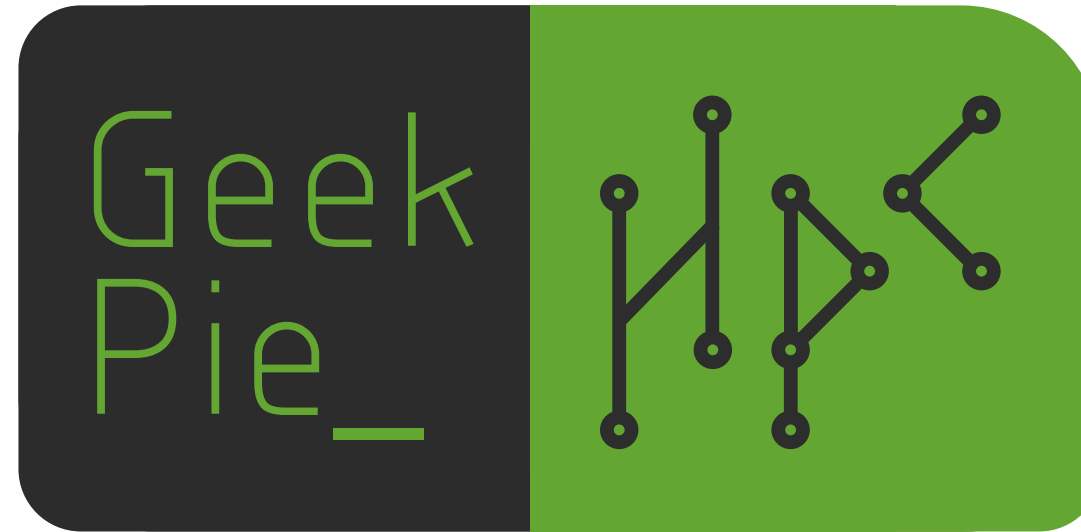
A photograph of the ShanghaiTech University campus at dusk. The image shows a modern building with a curved, overhanging roof supported by a series of white columns. A long, narrow pool of water in the foreground reflects the building and the sky. In the background, a tall, lattice-structured tower is visible against the sunset sky. The overall atmosphere is serene and modern.

*Small-scale
Cutting-edge*

ShanghaiTech University

TEAM

Team Info



GeekPie

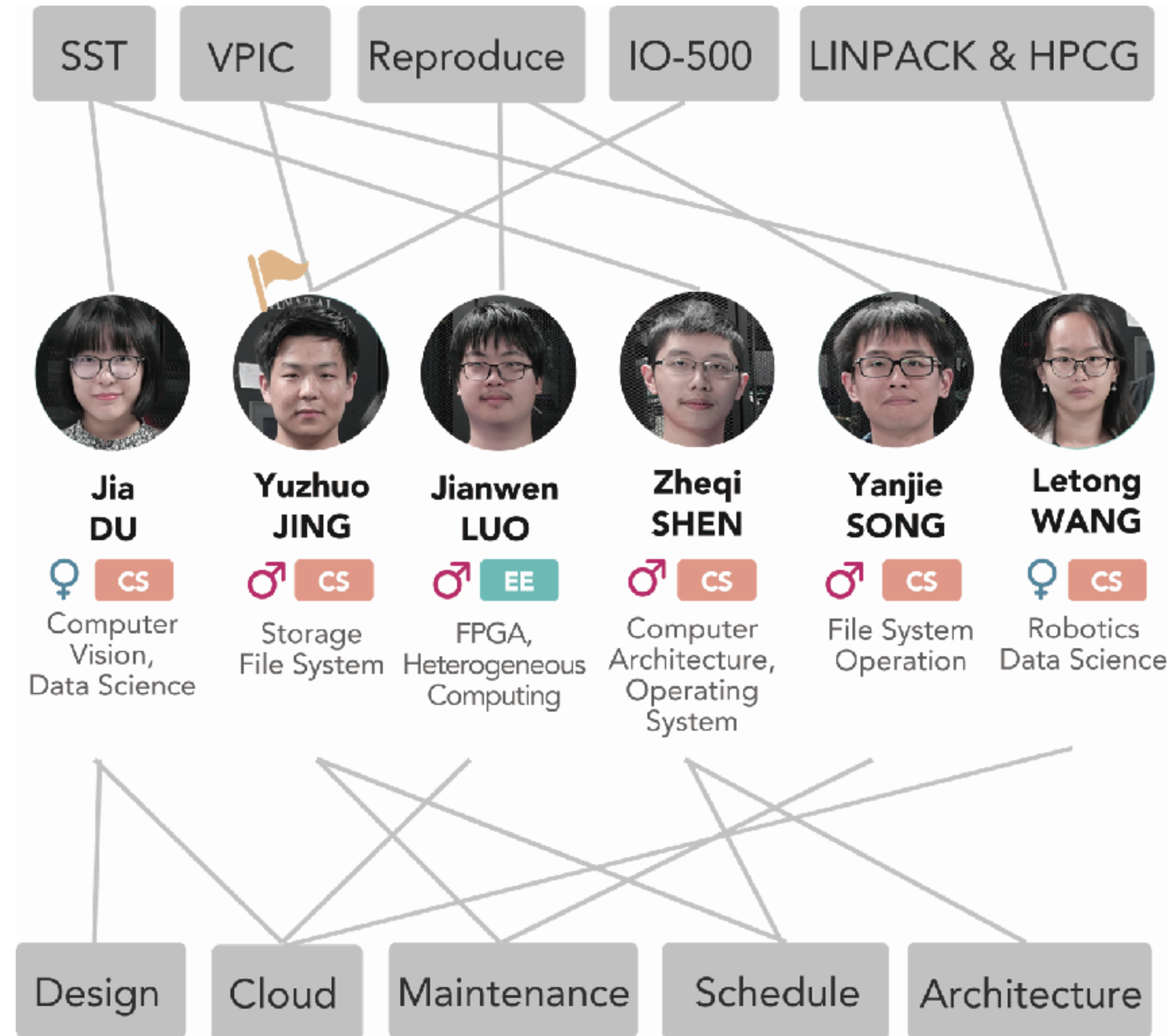
- The largest student association in our school, attracting science enthusiasts of different backgrounds.
- “Pie (派)” means “group” in Chinese;
- The association has various branches including
 - GeekPie_Robotics
 - GeekPie_Security
 - GeekPie_DevOps
 - ...

GeekPie_HPC

- As the youngest branch, GeekPie_HPC focuses on HPC-related training and competitions.

TEAM

Role



TEAM

Family Photo!!

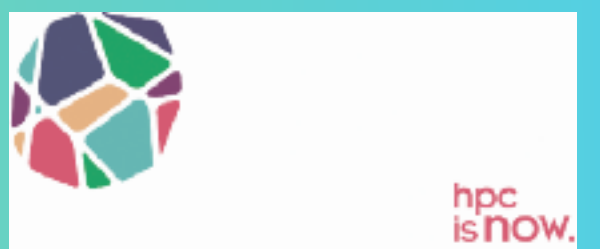


HARDWARE

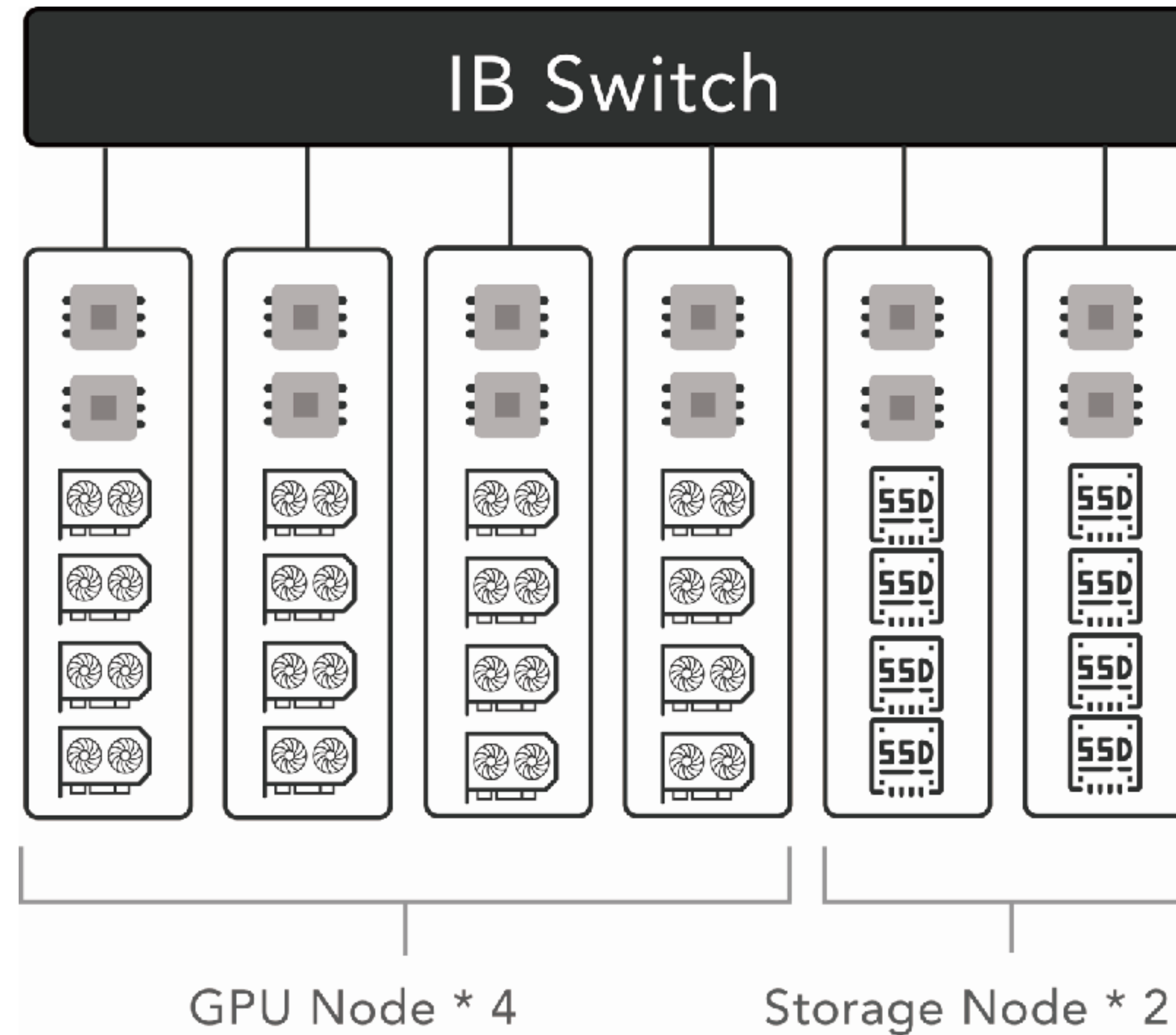
SOFTWARE

CLOUD

GeekPie_HPC

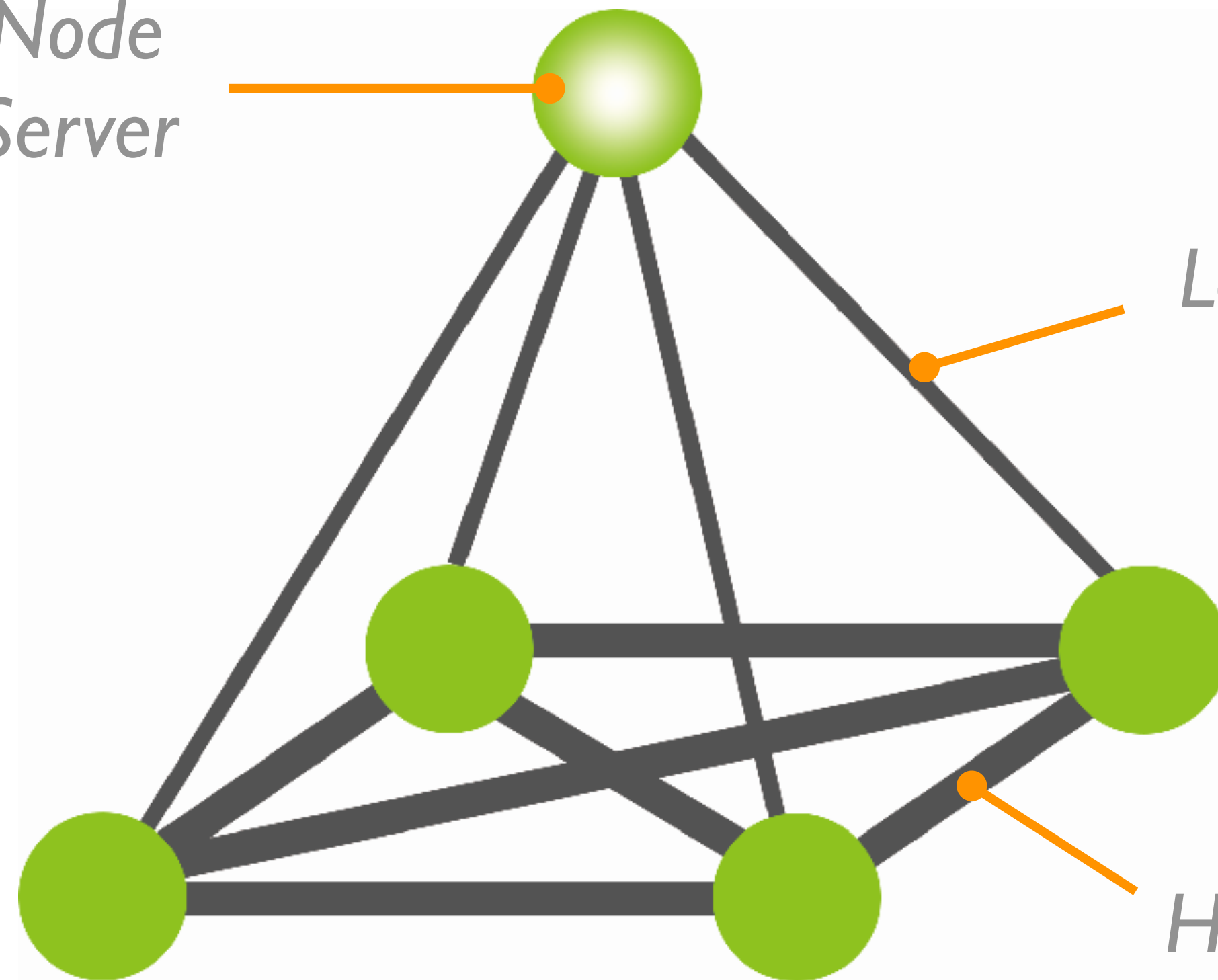


ARCHITECTURE



CLOUD

*Master Node
as NFS Server*



Low Speed Link

High Speed Link

APPLICATION

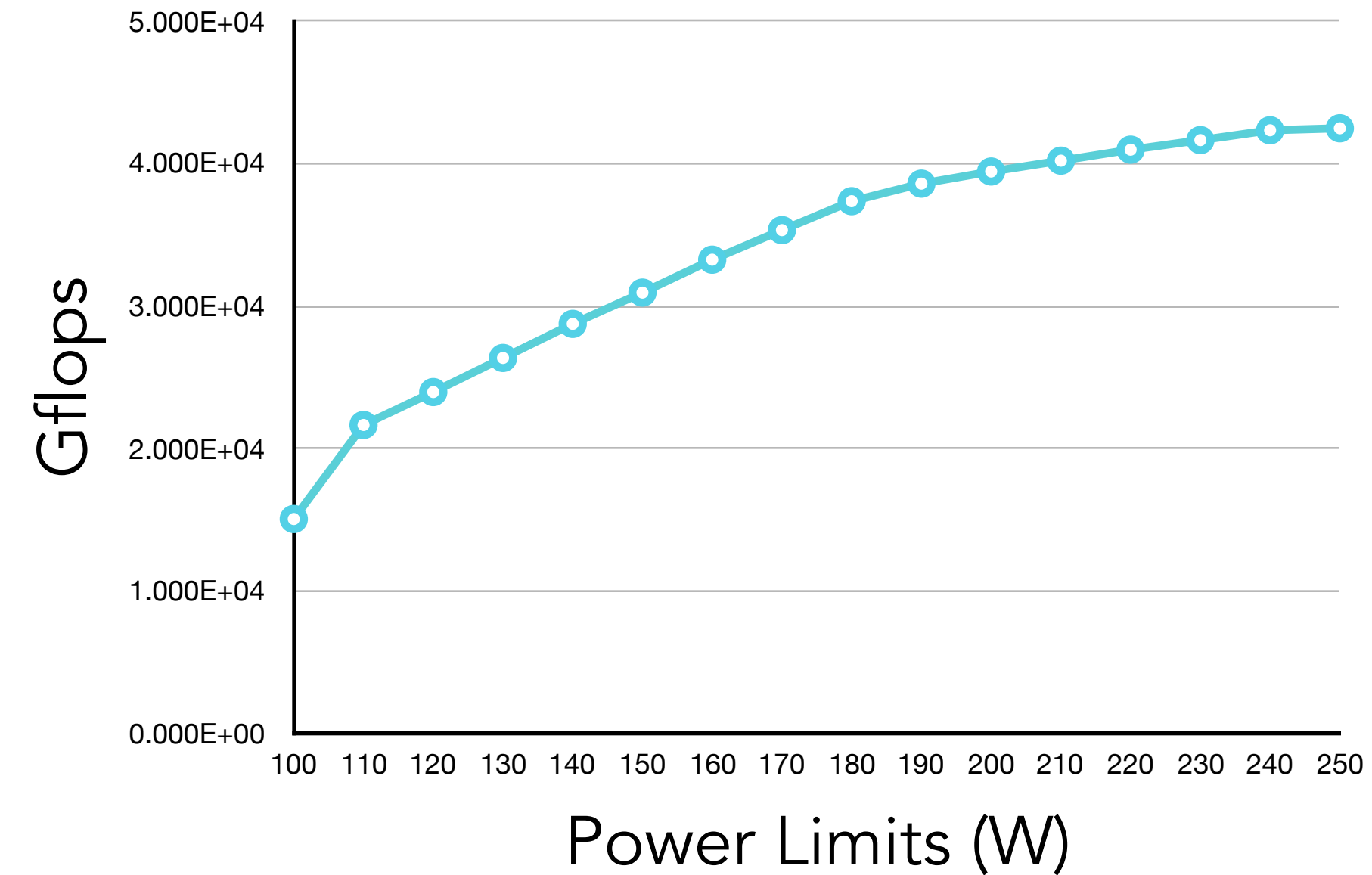
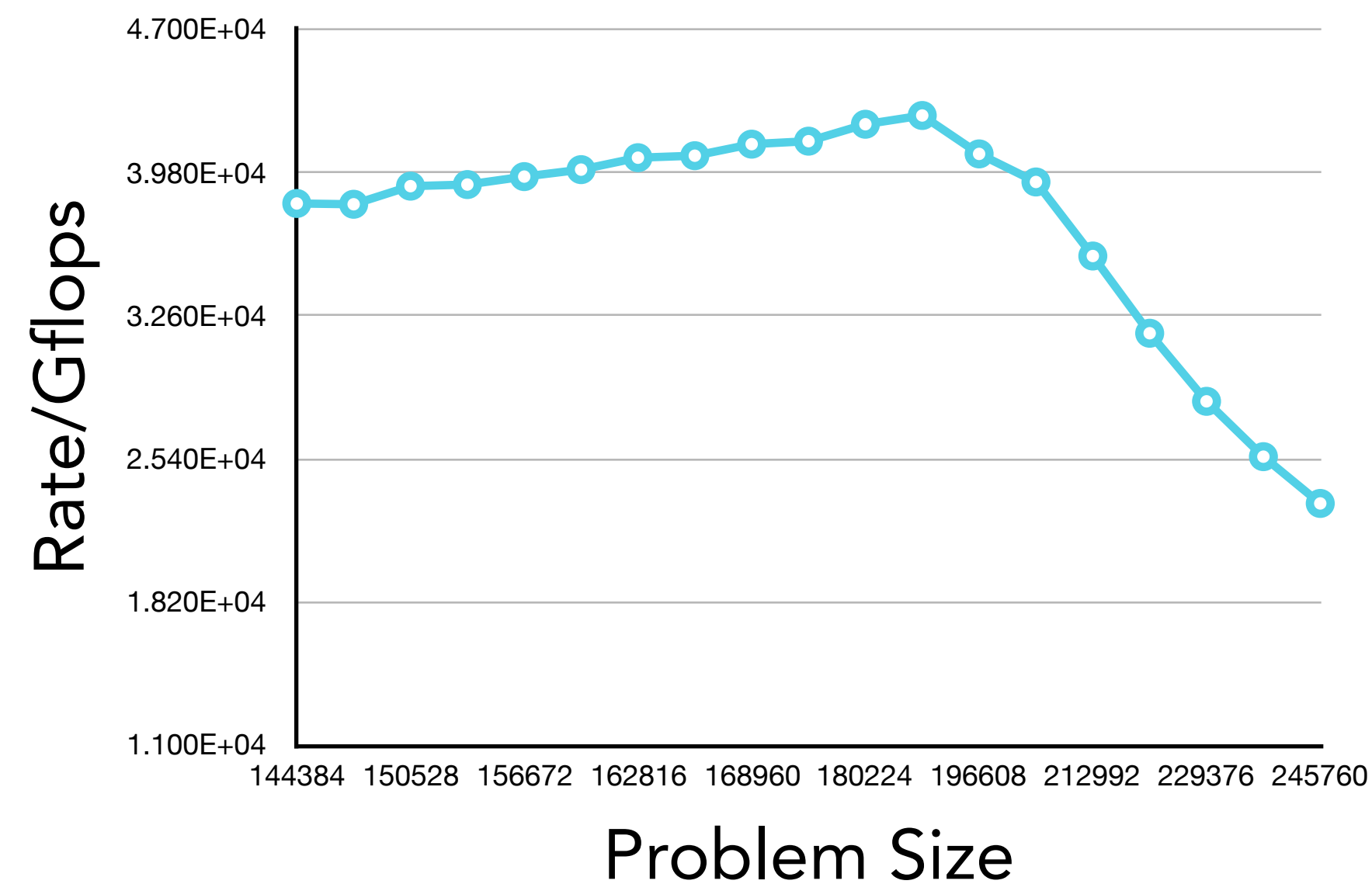
STRATEGY



GeekPie_HPC

Benchmark - HPL

Application

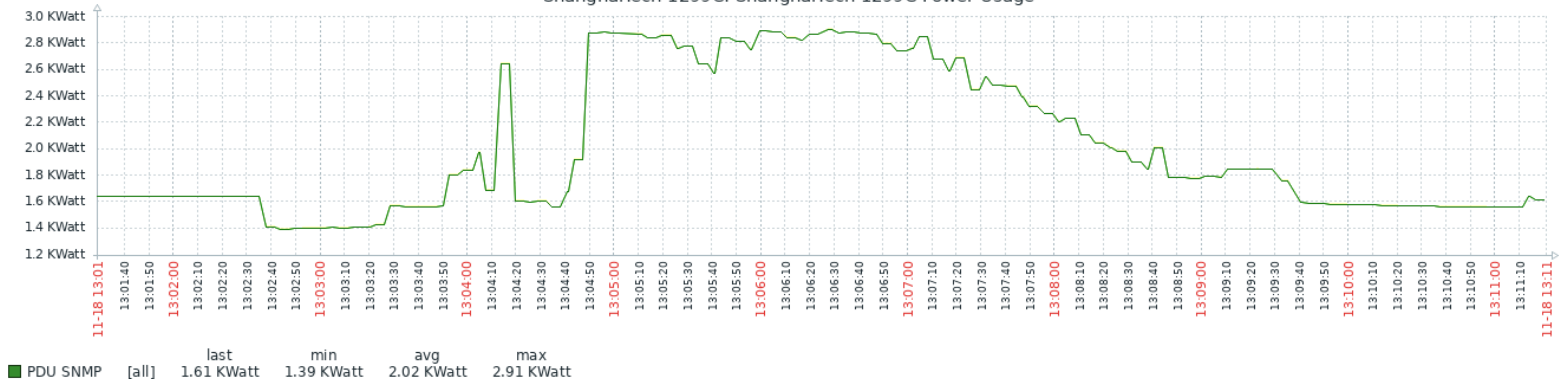


- Solve dense linear system
- Some parameters (e.g. problem size), are related to performance because of hardware configuration (e.g. memory usage)
- Performance is affected by the limited power supply

Benchmark - HPL

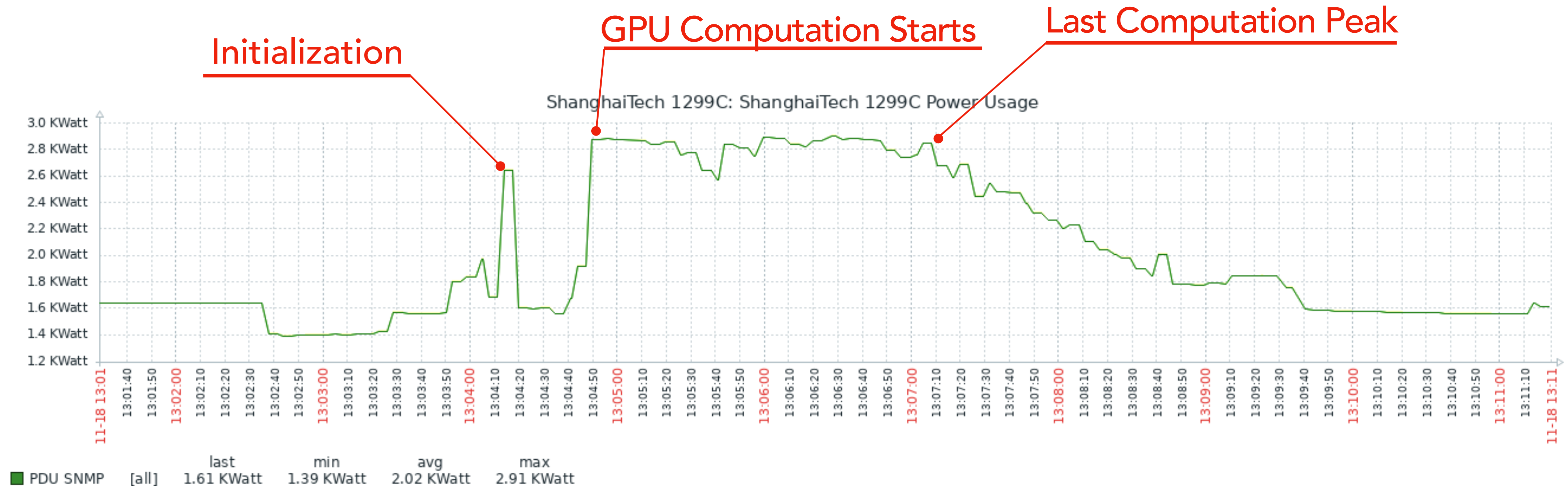
Application

ShanghaiTech 1299C: ShanghaiTech 1299C Power Usage



Benchmark - HPL

Application



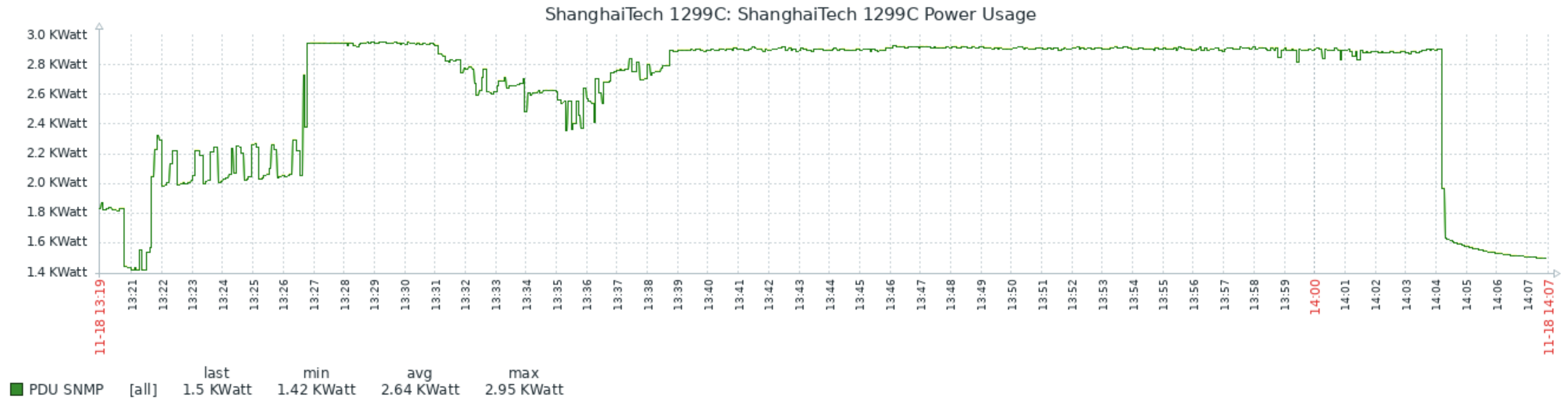
Benchmark - HPCG

Application

- Solve sparse linear system
- Tune the local grid dimension to fit our cluster's architecture
- Dynamically balance fan speed and GPU power

Benchmark - HPCG

Application

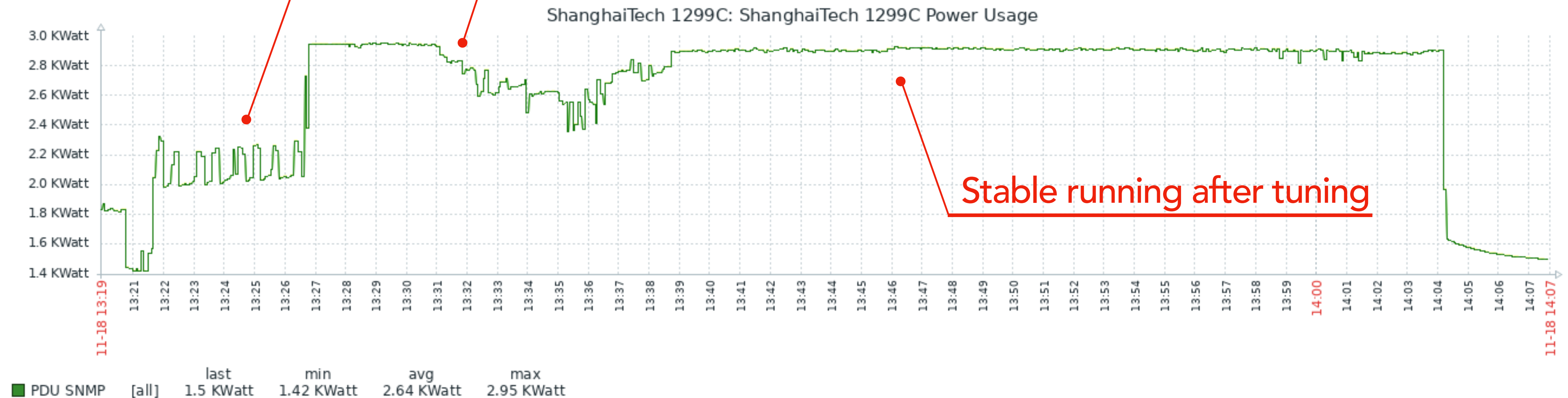


Benchmark - HPCG

Application

Initialization on CPU

Throttling due to temperature wall



- Lesson learned: careful of temperature wall for longer running time

Benchmark - IO-500

Application

- IO is becoming a dominate composition of many HPC workload
- Require balanced bandwidth and IOPS
- No PFS due to small cluster size
- RAID-0 for best performance, data crash is negligible in 48 hrs
- Using Intel® SSD DC P4510 Series (U.2)

What is?

- A parallel, componentized simulation framework for scalability and flexibility

How to?

- Adjust compile configurations of both core and involved elements
- Reconstruct the division of MPI ranks manually to optimize/balance the workload

What is?

- Planetary Normal Modes Simulation

How to?

- Understand mathematical details
- Summarize the logic flow of the code and algorithms
- Profile the program to find optimization hotspots
- Automate the collection of system characteristic parameters

What is?

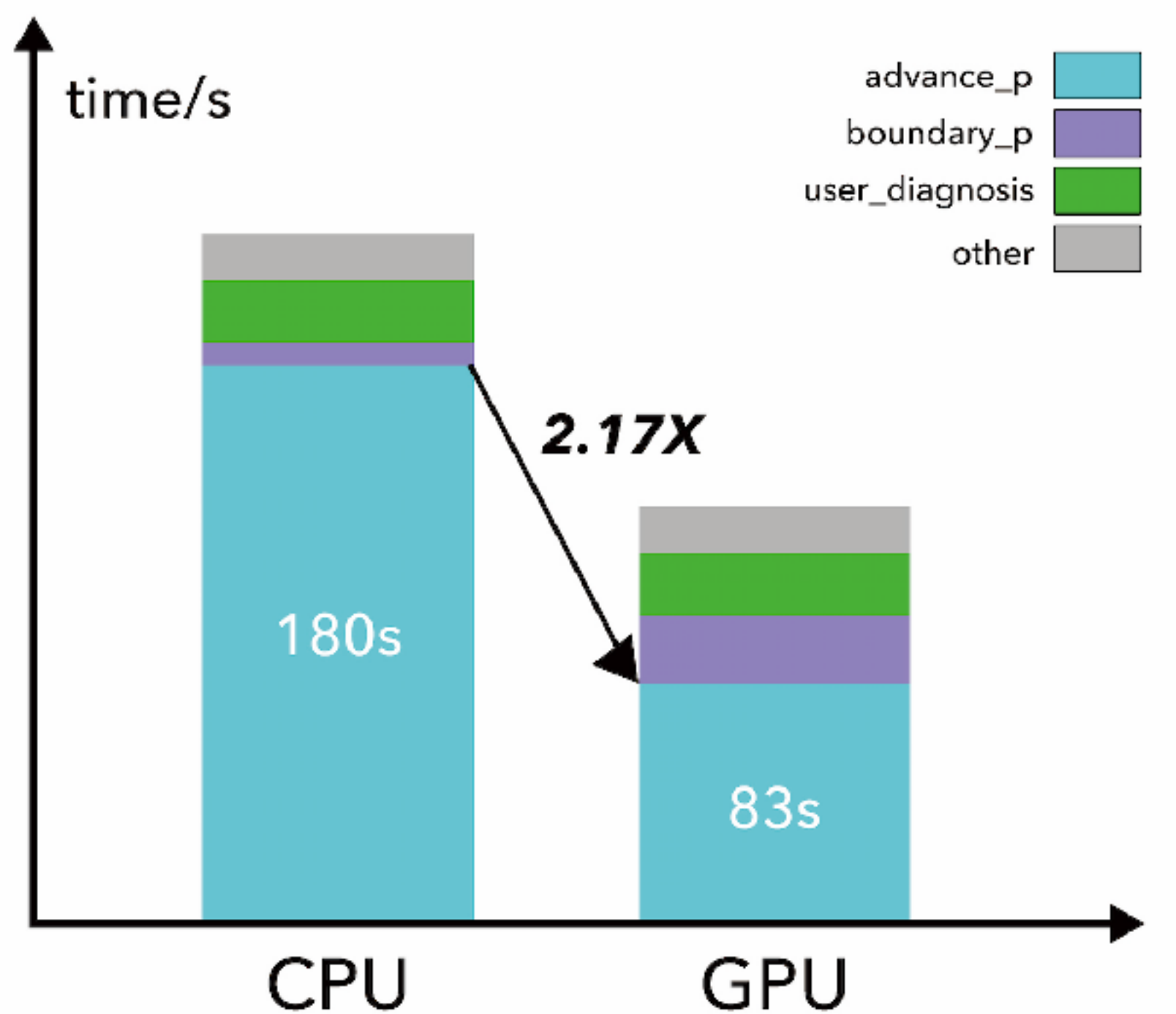
- Particle-In-Cell simulation software

How to?

- Analyzed application hotspot
 - advance_p: computation intensive ~65%
 - boundary_p: communication intensive ~25%
- Ported VPIC onto GPU, and optimized IO bottleneck
- Compared the performance of the two implementations

VPIC

Application



CPU and GPU Performance Comparison

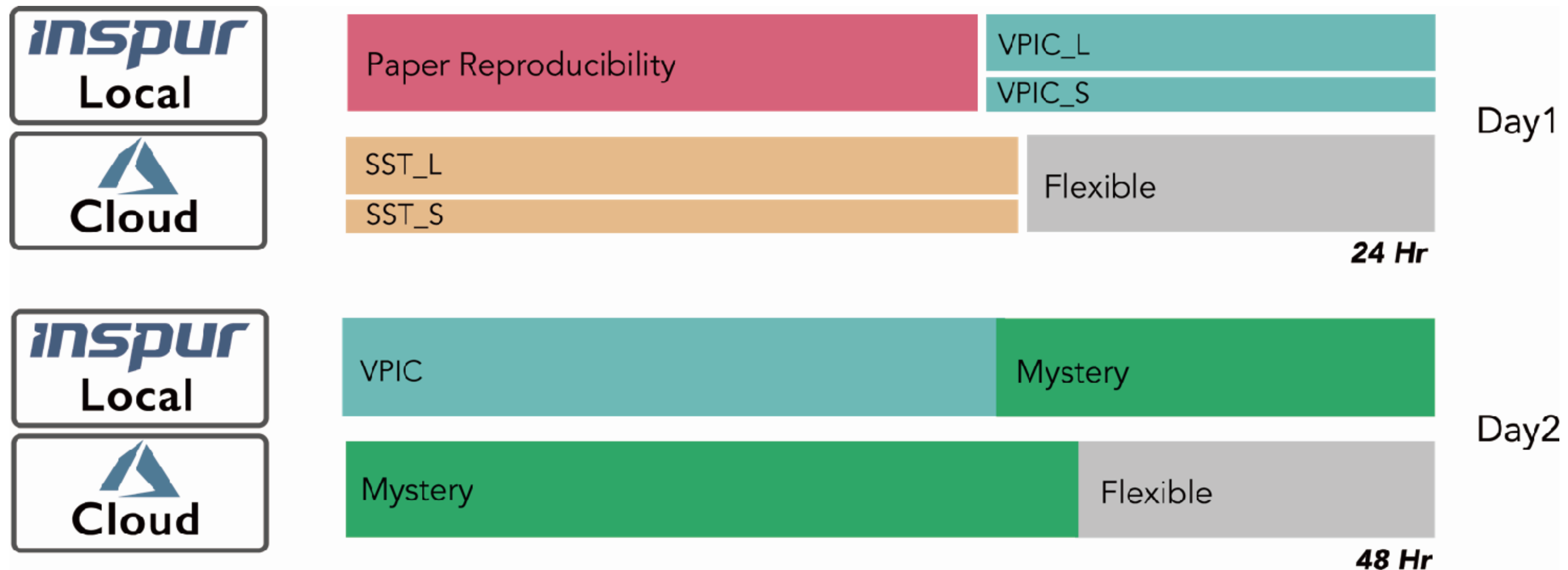
Advantage

- Leverage the massive parallelism of GPU
- Lazy-load data via CUDA Unified Memory (UM)
- Overlap computation & IO via *cudaMemPrefetchAsync*

Disadvantage

- Poor scalability
- Checkpoint will introduce extra overhead
- Hard to adopt user-customized code

Scheduling Strategy



Power Outage

Strategy

Data Consistency

- Separate read-only OS from writable data storage
- Replace OS disk on failure
- Recover data storage from replication
 - Use Data Center level SSD for better RAID robustness

Data Recovery

- Checkpointing
- Time-consuming tasks are postponed after the power outage / uploaded to the Cloud

Human Resource Strategy



- Three-shifts system:
keep at least ***half of the team*** awake at any time

WE WILL WIN!

Diverse & Dynamic



Well-Designed



Clear-Cut



Team

Architecture

Strategy

THANK YOU!!!

GeekPie_HPC



上海科技大学
ShanghaiTech University



inspur

