International Collaborations with TSUBAME

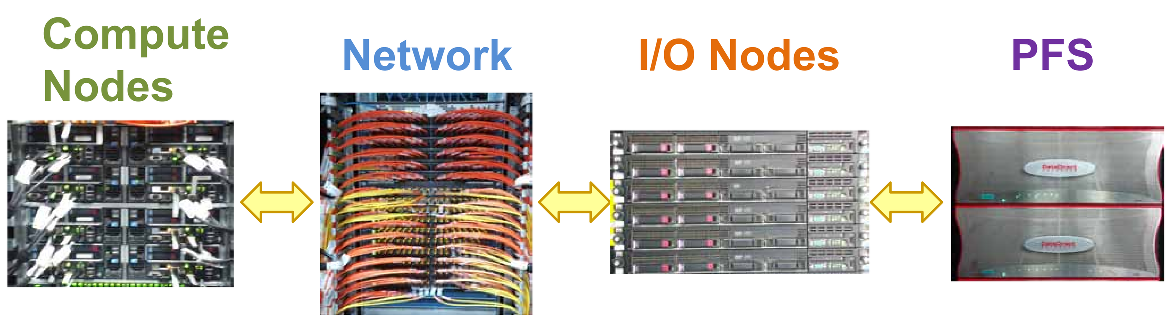
TOKYO TECH
Pursuing Excellence

FTI: Fault Tolerant Interface

FP3C

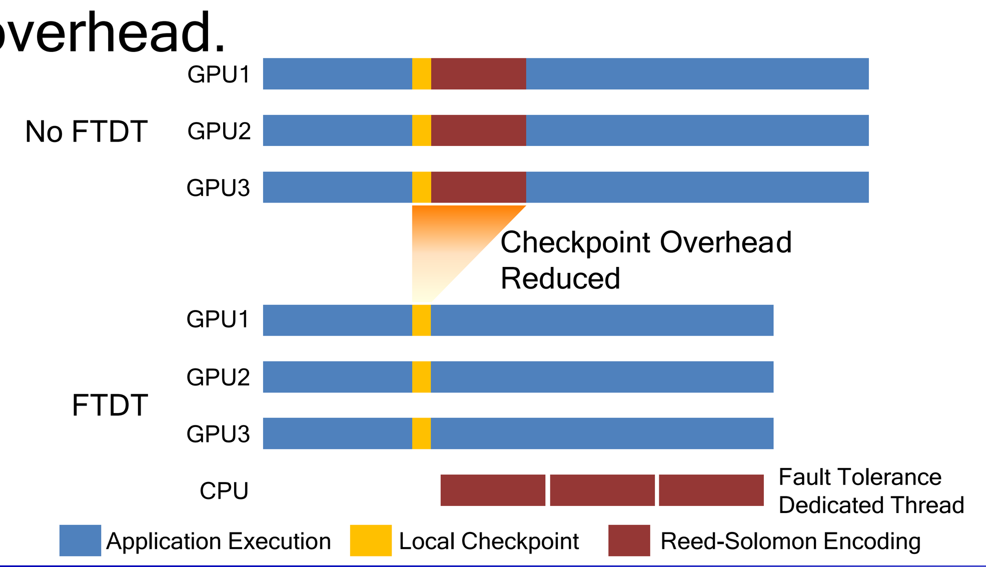
Background

Scalable fault-tolerance techniques are necessary for post-petascale systems.
Parallel file system based checkpoint restart cannot scale because of the I/O bottleneck.
Local disks such as solid-state-drives can provide the scalability needed to checkpoint large applications.



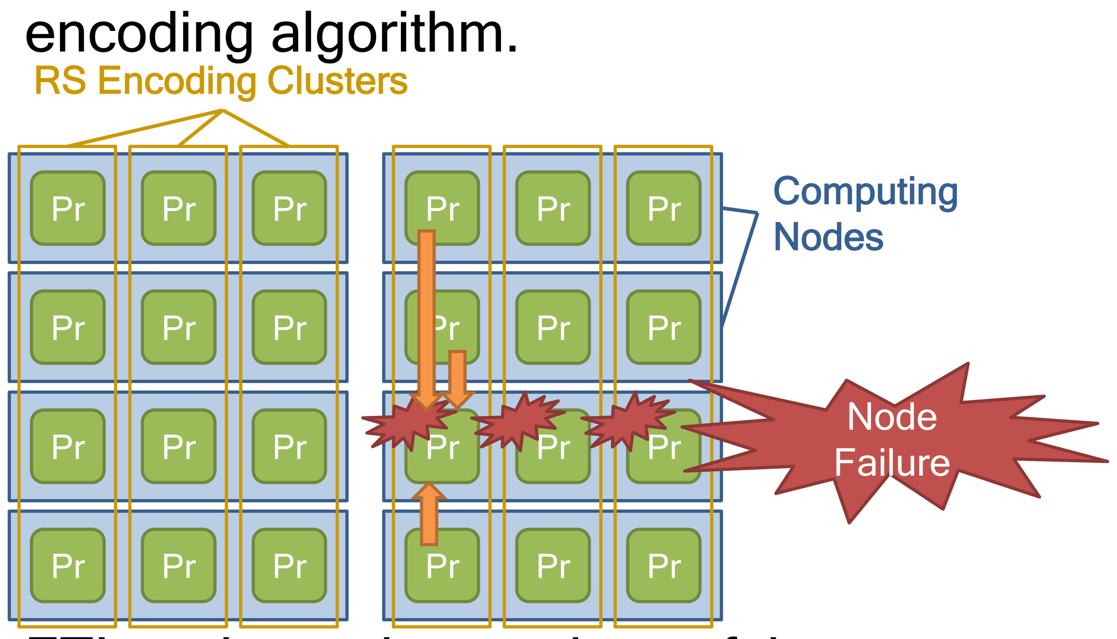
FT dedicated thread

In heterogeneous systems many applications do most of the computational work on GPUs leaving some CPU resources available.
We leverage those resources to encode the checkpoints, reducing the checkpoint overhead.



Topology-aware RS Encoding

Checkpoint replication or erasure codes are needed to tolerate hard failures.
FTI implements a scalable Reed-Solomon encoding algorithm.



FTI analyzes the topology of the system and creates encoding clusters that increase the resilience.

Multi-level Checkpointing

FTI has three levels of resiliency to tolerate different kinds of failures.
The three levels have a good resiliency overhead ratio.

L1: Local Storage (SSDs, PCRAM, Mem)
Transient Faulures

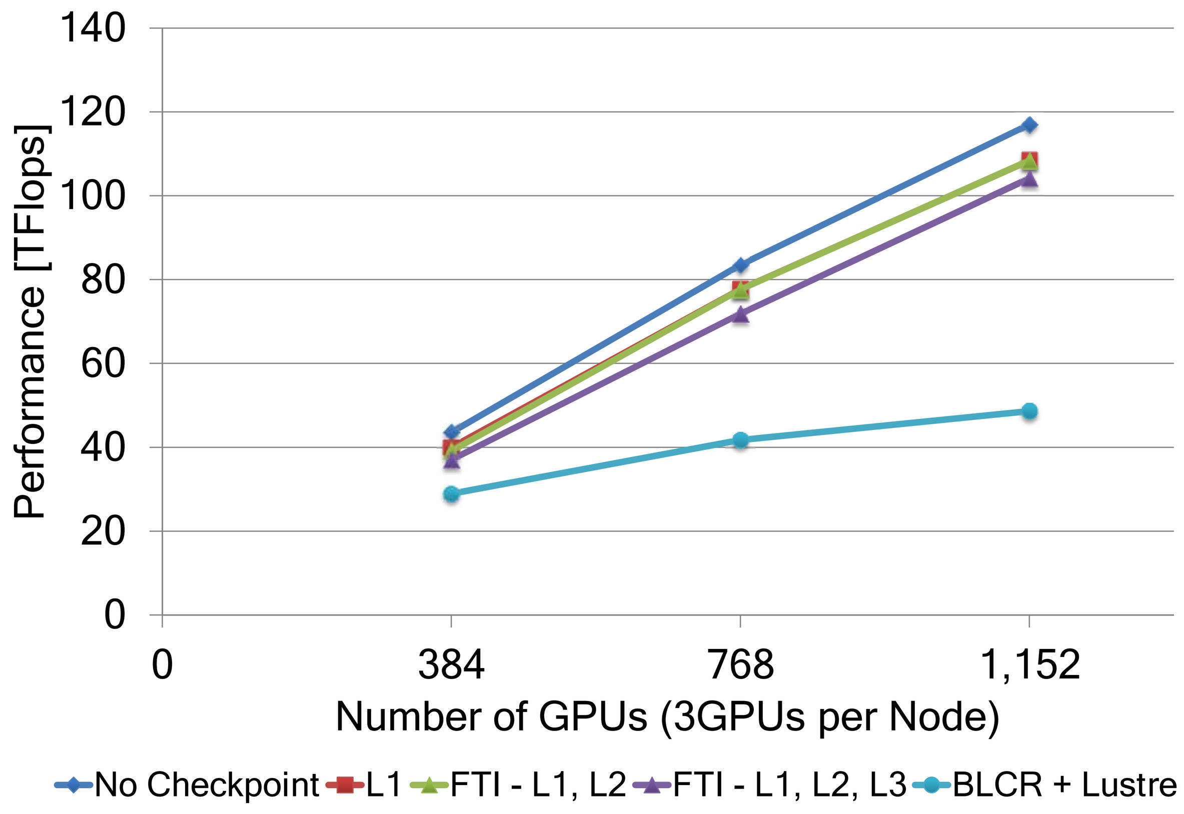
L2: Local Storage + TA-RS Encoding
Single or Multiple Node Crashes

L3: Parallel File System
Large Power Outage

Fast ↕ ***Reliable***

Evaluation Result

We evaluated FTI with the seismic wave simulation application: SPECFEM3D.



FTI scales up to more than one thousand GPUs on TSUBAME2.0 while checkpointing every 6 minutes.
The checkpoint overhead is as low as 8% in comparison with a non checkpointed execution.

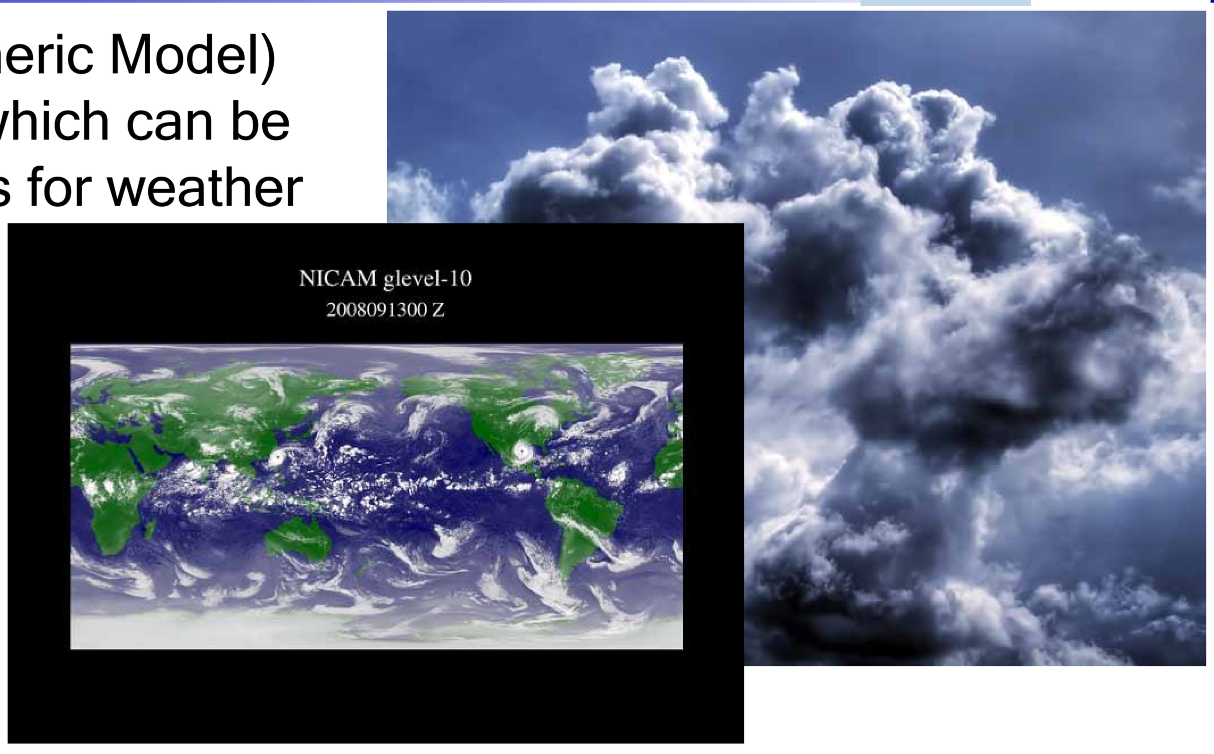
References

L. **Bautista-Gomez et al.** - Hierarchical Clustering Strategies for Fault Tolerance in Large Scale HPC Systems, Cluster2012, Beijing CHINA, 2012.
L. **Bautista-Gomez et al.** - Scalable Reed-Solomon-based Reliable Local Storage for HPC Applications in IaaS Clouds, EuroPar2012, Rhode Island, GREECE, 2012.
L. **Bautista-Gomez et al.** - FTI: high performance Fault Tolerance Interface for hybrid systems, SC11, Seattle, USA, 2011.

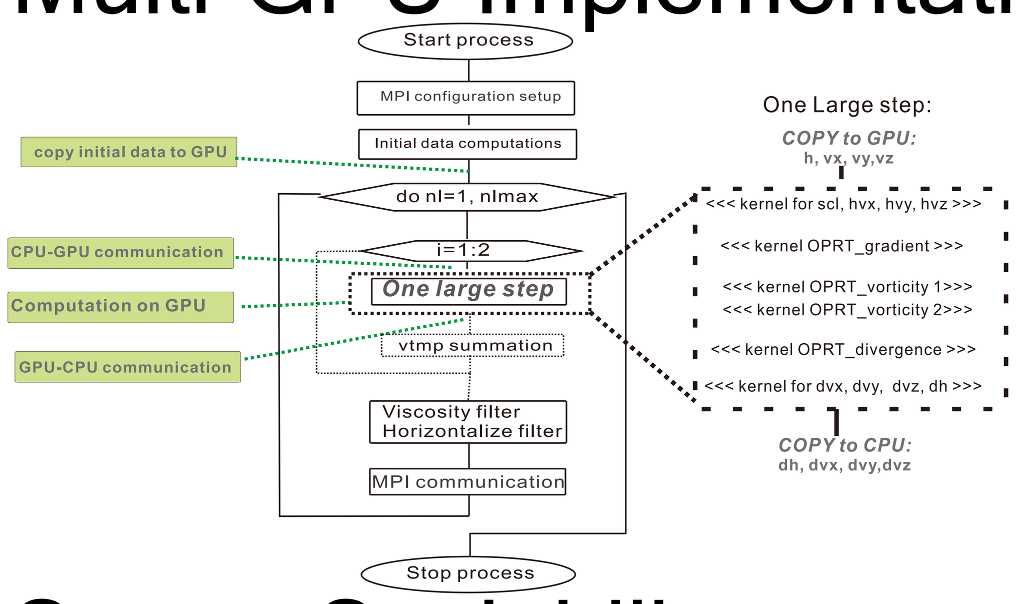
NICAM Atmospheric Model on GPUs

G8 RCI ECS

NICAM (Nonhydrostatic Icosahedral Atmospheric Model) is a Global Cloud Resolving Model (GCRM) which can be used for both short term numerical predictions for weather systems and long term climate simulations.

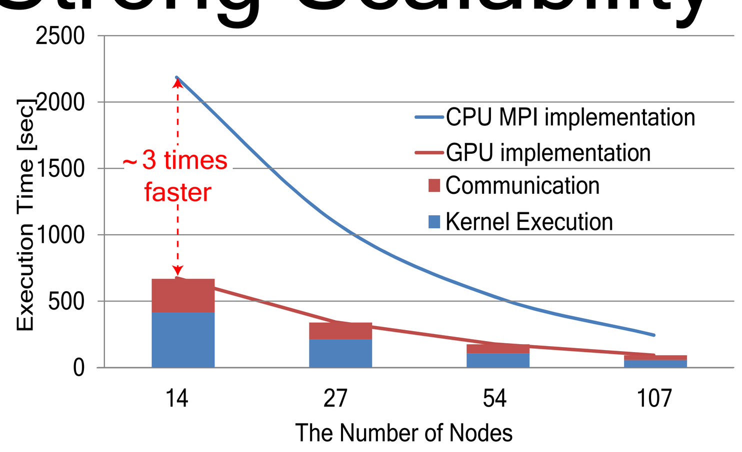


Multi-GPU Implementation



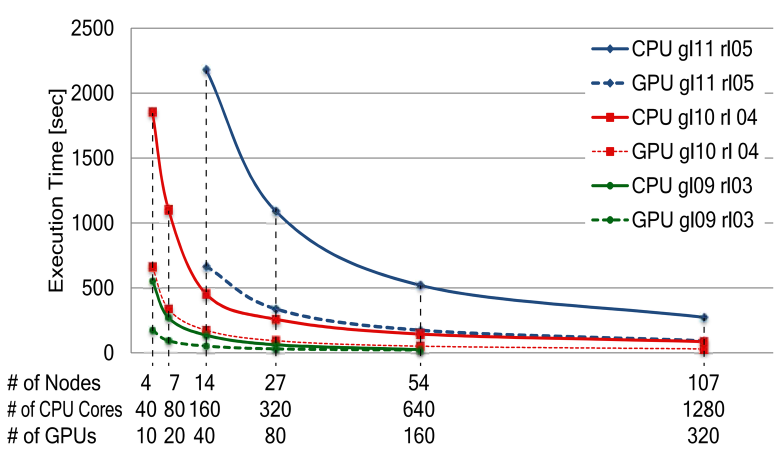
Multi-GPU implementation is based on moving the most computational intensive part ("One large step" module) to GPUs by using PGI CUDA Fortran.
The implementation was evaluated on TSUBAME2.0.

Strong Scalability



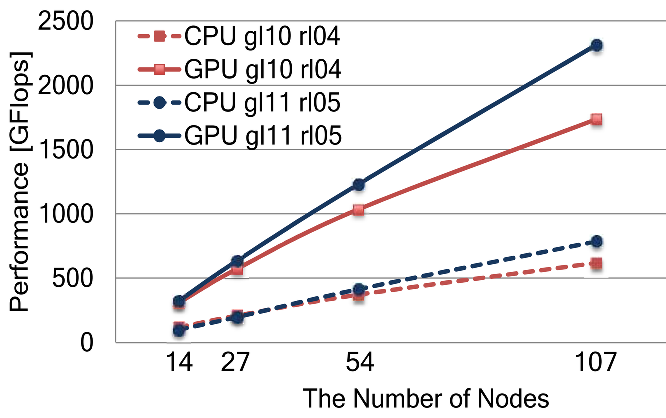
Multi-GPU implementation gives 3 times higher performance comparing with the maximum performance of the MPI-parallel version.

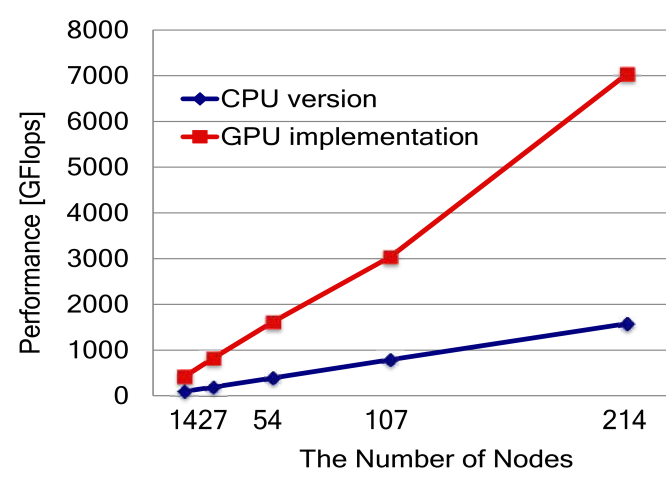
Performance of the proposed GPU implementation is almost optimal for "reasonable-sized" problems.



The performance model was developed and validated for a full-GPU implementation of the NICAM. Results show 4.5x potential acceleration over parallel CPU and 7 TFLOPS potential performance on 217 nodes for a multi-GPU implementation.

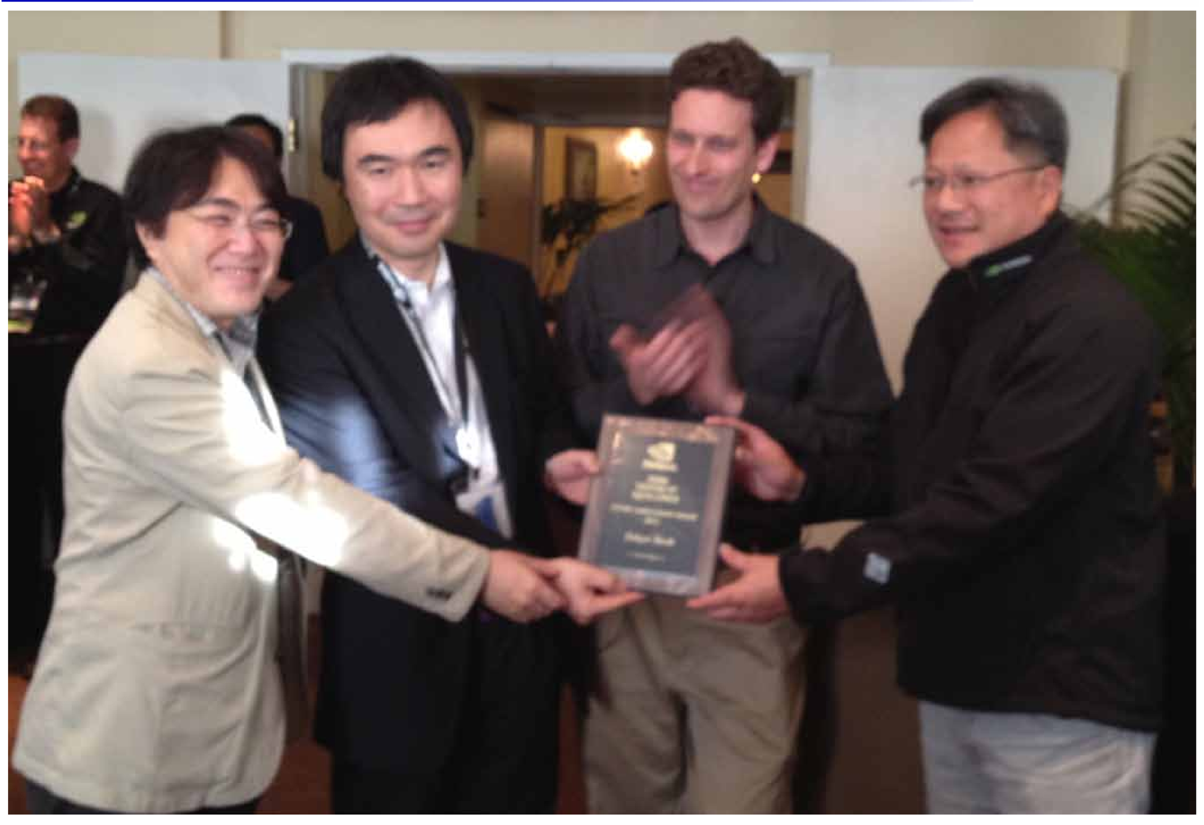
Performance





CUDA CoE

NVIDIA.



Tokyo Institute of Technology is selected as the first CUDA Center of Excellence(CCOE) in Japan.
We snagged the first-ever Achievement Award for CUDA Centers of Excellence (CCOE), for our research with TSUBAME 2.0.
We are using NVIDIA GPUs for various fields of research including

- Physis: Large-scale stencil programming framework
- Petascale phase-field simulation (2011 Gordon Bell Award)
- Accelaration of meta-genome analysis
- AMR for large-scale applications
- OpenACC performance