Cross-Sector Applications & Solutions

Wed. Nov 11, 2026 4:00 PM - 5:30 PM JST | Wed. Nov 11, 2026 7:00 AM - 8:30 AM UTC | Room D (Icho) (3F)

Oral D1 Cross-Sector Applications & Solutions

Chair: Takashi Enomoto, Arne Jørgen Berre

5:00 PM - 5:15 PM JST | 8:00 AM - 8:15 AM UTC

[BSG7SMHd-05] Accelerating a Fully Coupled Earth System Forecast Model through an Efficient CPU-to-GPU Development Approach

*Yun-Chuan Shao[1] (1. National Taiwan University (Taiwan))

Over the past decade, limitations in GPU hardware and development technologies have made it difficult to migrate global numerical weather prediction models to GPU architectures. Although these models run efficiently on CPU-based systems, porting them to GPU typically requires substantial development effort to rewrite low-level code for specific GPU architectures. In addition, the high double-precision requirements of global numerical weather prediction models introduce challenges related to rounding errors and numerical accuracy during GPU implementation. Similar issues have been reported in GPU-based weather and climate model development, particularly when different numerical precisions and CPU–GPU implementations are compared.
To address these issues, our team developed a CPU-to-GPU migration framework that combines OpenACC with the NVIDIA Collective Communications Library (NCCL). Using this framework, we upgraded the Central Weather Administration's Global coupled Earth system Prediction System, GEPSv3, from its original CPU-based high-performance computing platform to a GPU-based architecture. The proposed approach aims to reduce the need for extensive low-level code rewriting while addressing the inter-GPU communication and numerical-accuracy requirements of a fully coupled global ocean–atmosphere forecasting system.

A performance comparison was conducted between the original configuration using 496 CPU cores on the FX1000 system and the GPU-based configuration using 16 CPU cores and eight GPUs on an A100 node. For a 45-day integration of GEPSv3, the GPU-based configuration reduced the computational time by 86%, while the node-hour consumption was reduced to approximately 1.4% of that of the original CPU-based configuration. These results demonstrate that the proposed NCCL-OpenACC framework can substantially improve computational efficiency and reduce resource consumption, while providing an effective approach for migrating a fully coupled global ocean–atmosphere forecasting system from a CPU-based to a GPU-based architecture.

© 2026 DITTO Summit Local Organising Committee. All rights reserved.