

Sanil Rao*, **Mohammad Alaul Haque Monil**[†], Het Mankad[†], Narasinga Rao Miniskart[†], Keita Teranishi[†], Jeffrey Vetter[†], Franz Franchetti*
*Carnegie Mellon University, Pittsburgh, PA, USA. [†]Oak Ridge National Laboratory, Oak Ridge, TN, USA. (monilm@ornl.gov)

IRISX: Introduction and Motivation

- A performance portable solution for scientific computing
 - Combines code generation and runtime orchestration
 - Targets multi-accelerator platforms
 - Capable of adjusting concurrency based on hardware
 - Kernel and graph level transformation using SPIRAL code generation engine and IRIS Runtime
 - Finds the best configuration suited for a platform
- Outperforms state-of-the-art scientific library

IRIS Runtime and SPIRAL Code Generation Engine

IRIS Runtime

- A portable task-based runtime for heterogeneity
- Automated and optimized data movement
- Multi-vendor, multi-device and multi-stream asynchronous execution and a rich set of schedulers
- Supports: CPUs, GPUs, FPGAs and DSPs
- Automatic orchestration of heterogeneous processors



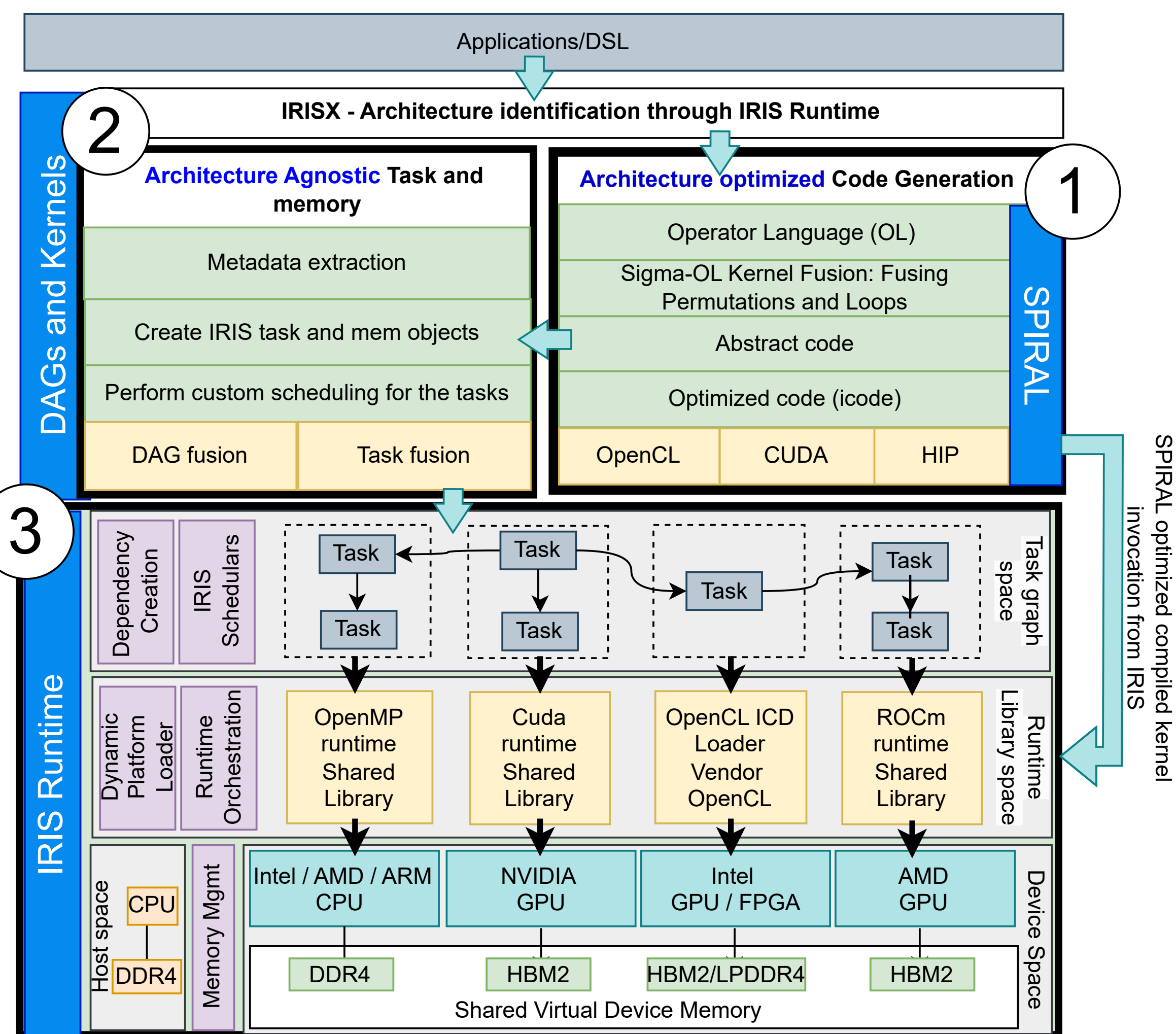
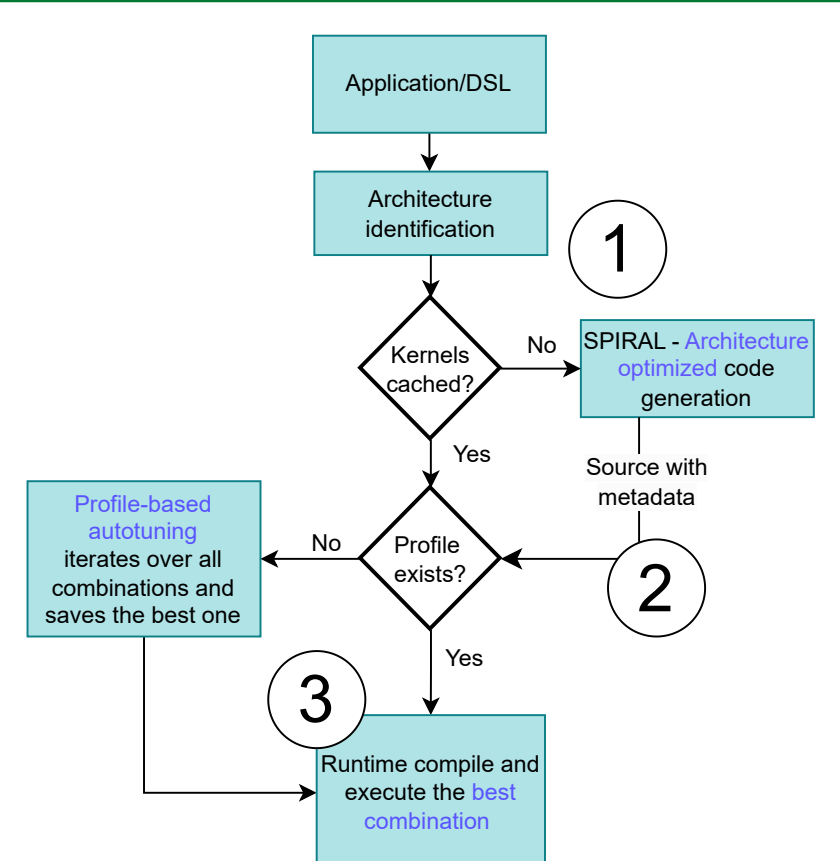
SPIRAL Code Generation Engine

- Generates architecture-optimized code
- Expresses the problem in high level Operator Language
- Capable of performing various levels of kernel fusion
- Provides ready-to-use optimized scientific kernels for various graph algorithms, PDEs, and cryptography

IRISX Design and Flow

IRISX performs three main steps

- Generates optimized kernels with metadata for the available hardware. Applies kernel fusion
- Tasks, and graph is created and invokes IRIS Runtime
- Performs data movement, orchestration, DAG and task fusion and stores the profile

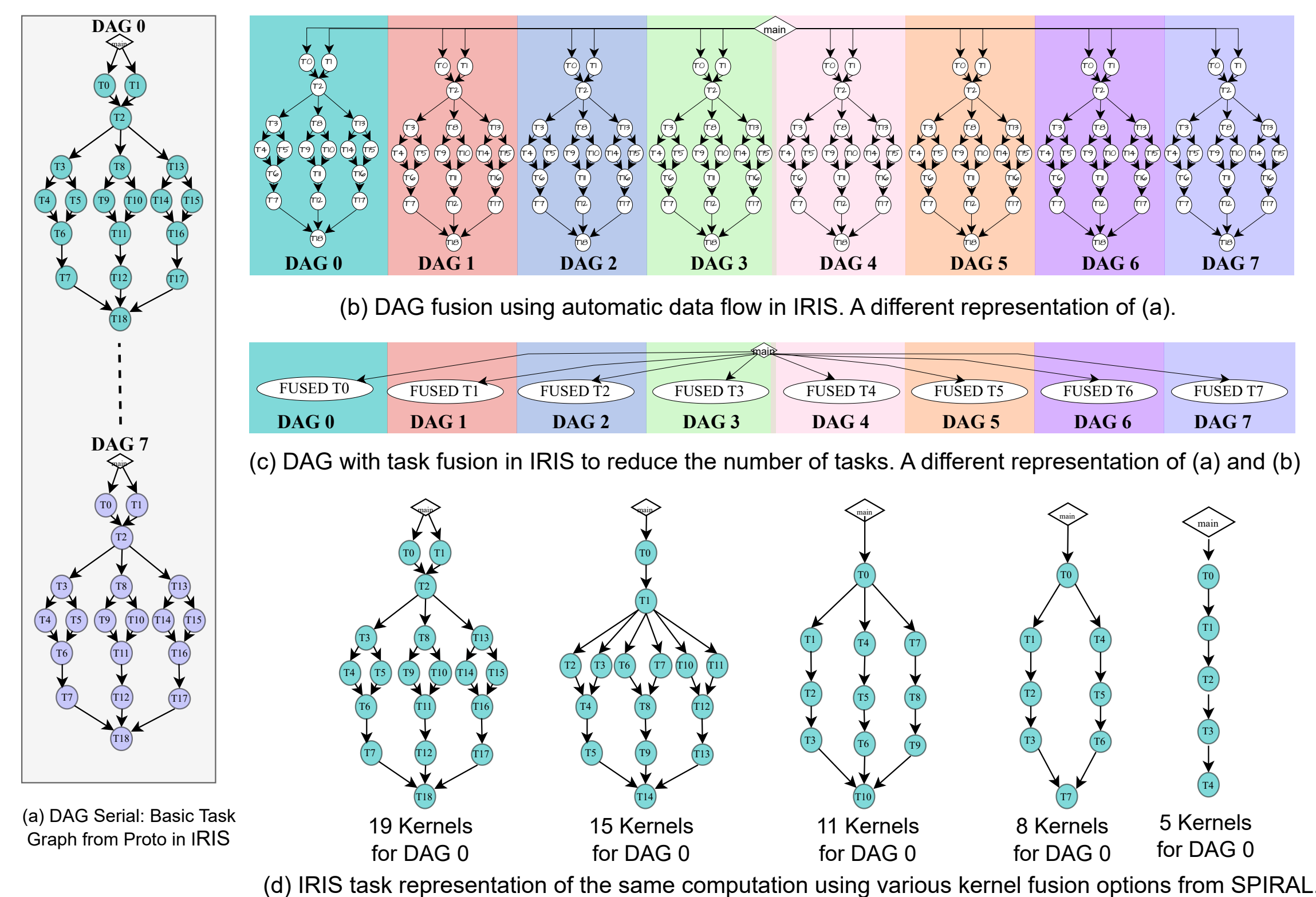


Various Representations of Solving 3D Euler equation

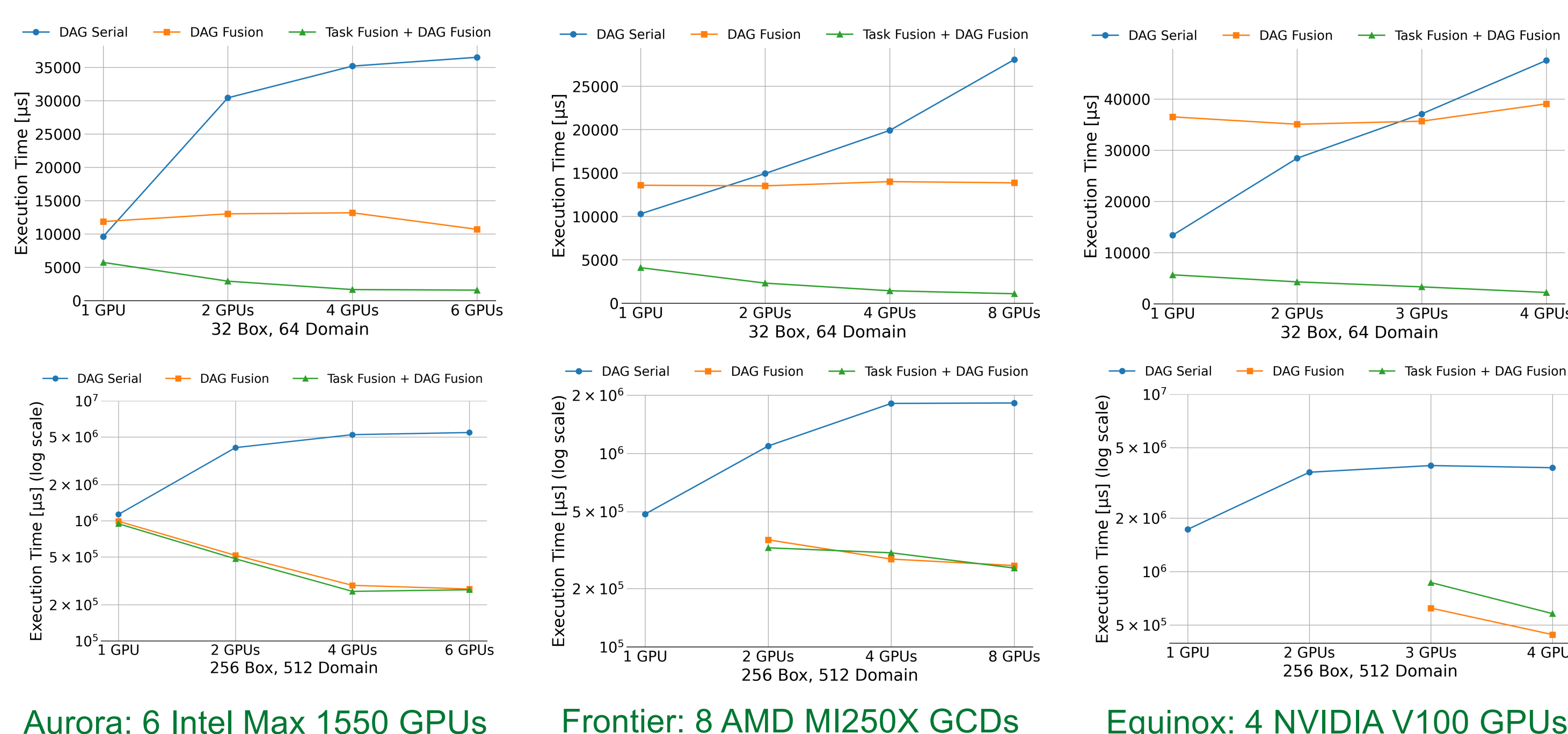
Kernel Fusion: Fuses multiple kernels using SPIRAL

Task Fusion: Fuses multiple Tasks using IRIS Runtime

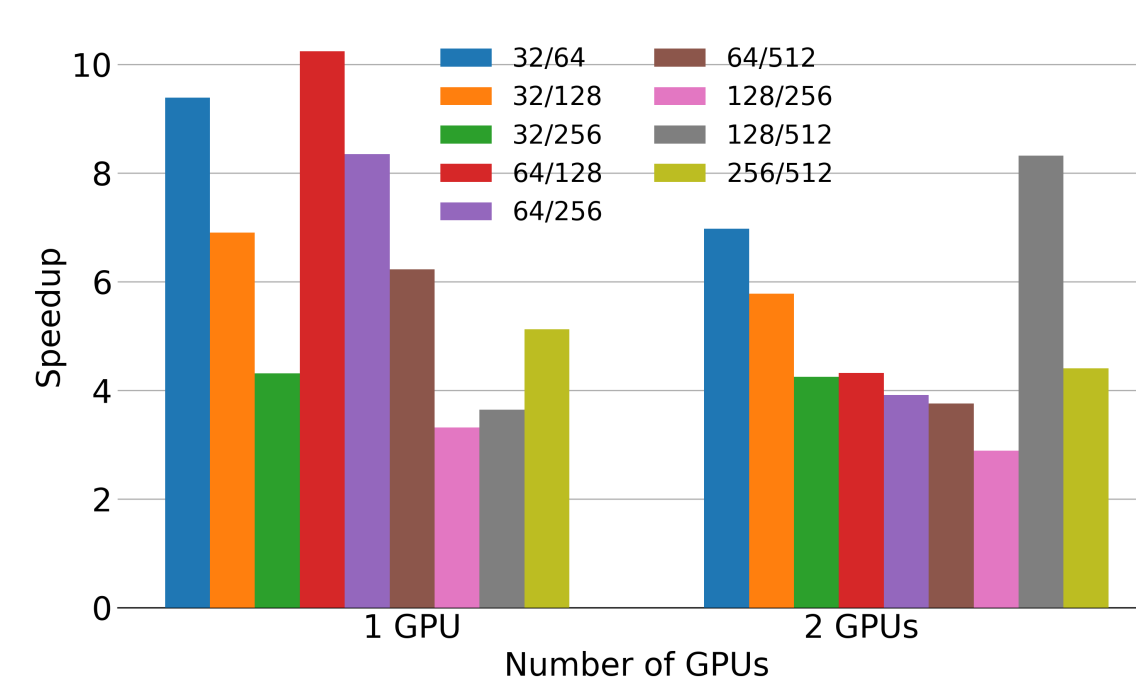
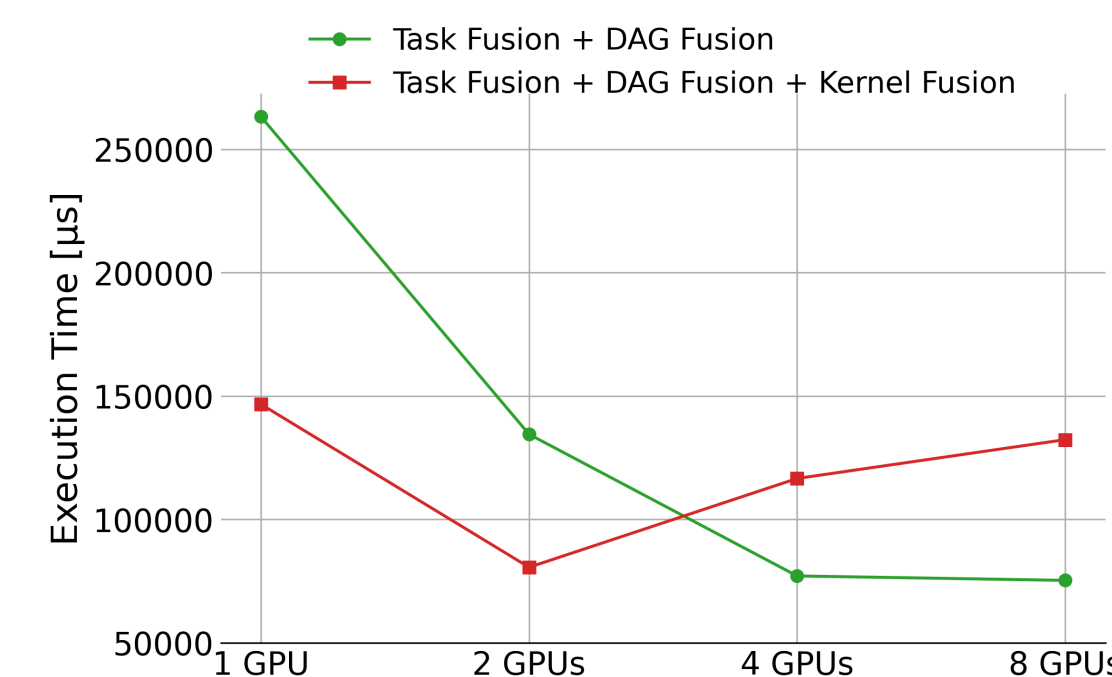
DAG Fusion: Fuses multiple DAGs using Data flow analysis in IRIS Runtime



IRISX Mitigates Performance Variability and Finds the Best Configuration



- The Left charts show the performance differences due to Task and DAG fusion.
- The right chart shows performance difference due to kernel fusion



- IRISX provided up to 10x performance improvement when compared with Proto library on 2 NVIDIA A100 GPUs