BOISE STATE UNIVERSITY
College of Engineering

Massively Parallel Incompressible Navier-Stokes Computations on the NCSA Lincoln Tesla Cluster

Dana Jacobsen and İnanç Şenocak

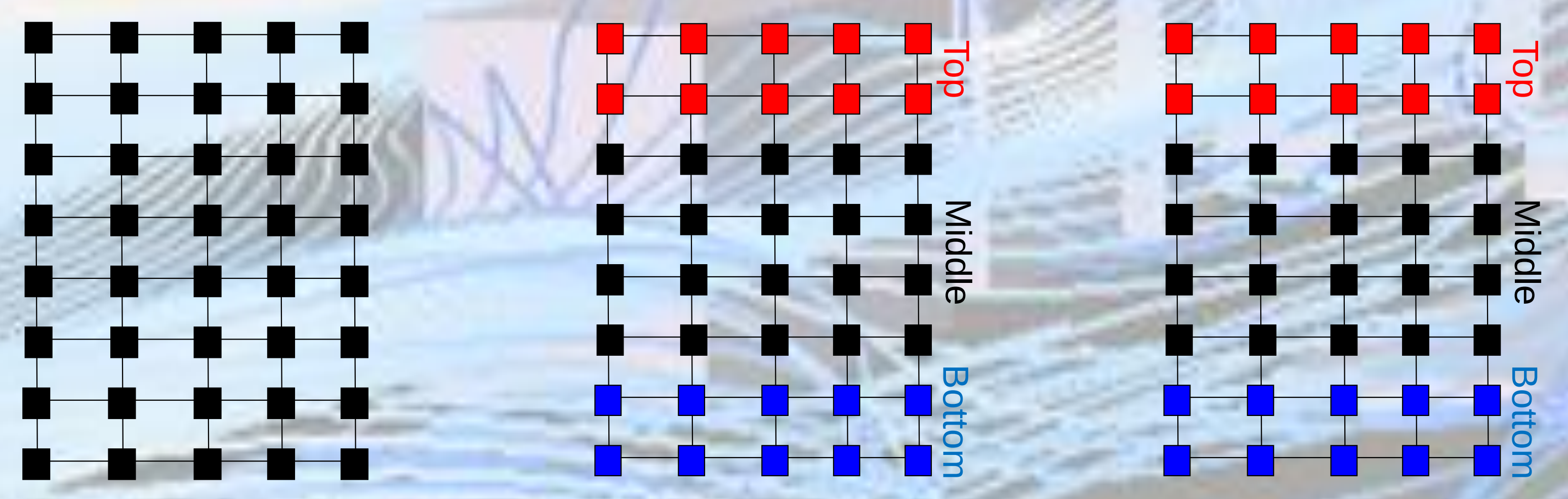
NVIDIA CUDA

Objective: Develop a *scalable* and *efficient* incompressible Navier-Stokes solver for emerging massively parallel *multi-GPU* clusters

Pseudocode MPI+CUDA Implementation

```
for (timestep = 1 .. ntimesteps) {
   temperature<<< u,v,w,phi >>>
   Copy_Exchange_Ghost_Cells(phi)
   momentum<<< u,v,w,phi >>>
   Obstacle_Logic
   Copy_Exchange_Ghost_Cells(u,v,w,p)
   divergence<<< u,v,w,div >>>
   for (jacobisteps =  1 .. Iterations) {
      pressure<<< div, p >>>
      Copy_Exchange_Ghost_Cells(p)
   }
   velocity_correct<<< u,v,w,p >>>
   Obstacle_Logic
   Copy_Exchange_Ghost_Cells(u,v,w)
}
```

Copy/Exchange Strategies



No Overlap

- Compute full domain
- Copy and send top/bottom using MPI async calls
- Finish MPI receives, copy top/bottom to device

- Does not overlap computation
- MPI and host/device copies are interleaved

Overlap

- Compute top
- Compute bottom
- Copy top/bottom to host
- Send top/bottom using MPI async calls
- Compute middle
- Finish MPI receives
- Copy top/bottom to device
- ThreadSync to ensure middle is done

- Overlaps computation with exchange

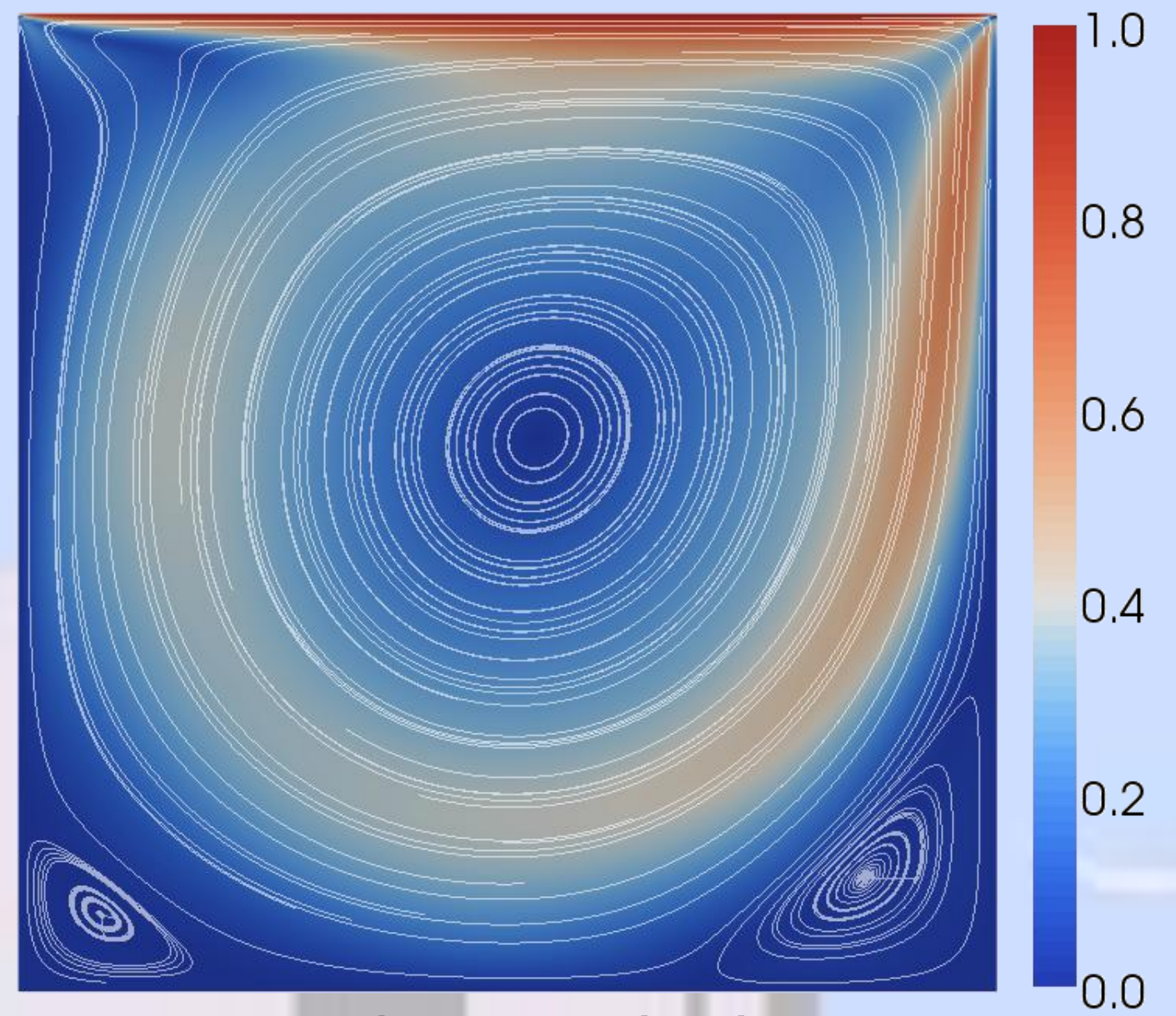
CUDA Streams

- Compute top
- Compute bottom
- ThreadSync to ensure computations complete
- Use CUDA Streams to simultaneously copy top/bottom from device while computing middle.
- Exchange top/bottom using MPI async calls
- Async copy top/bottom to device
- ThreadSync to ensure all work is done

- Overlaps computation with copies and exchanges

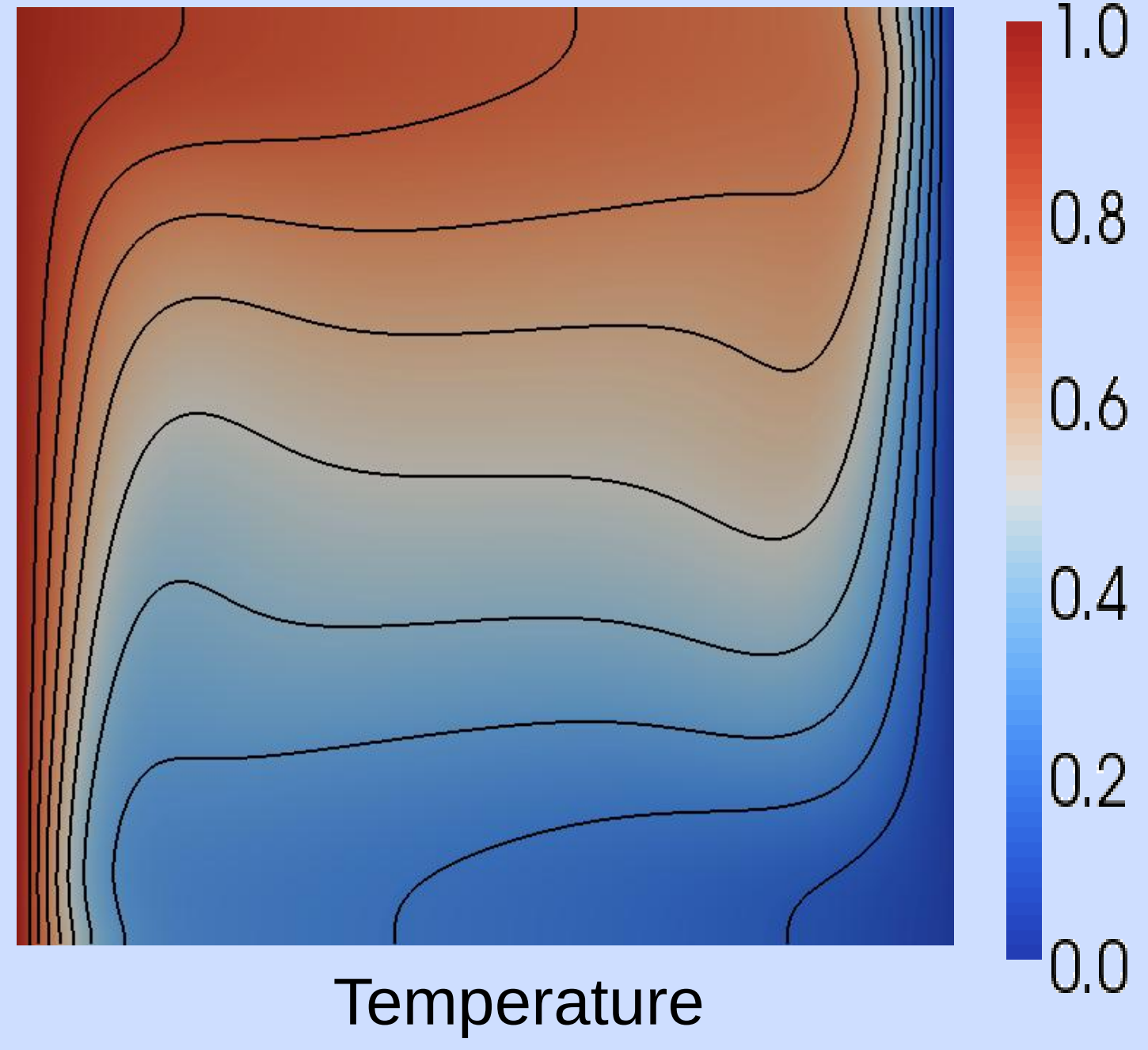
GIN3D Simulation Capabilities

Lid Driven Cavity



Velocity Magnitude

Heated Cavity

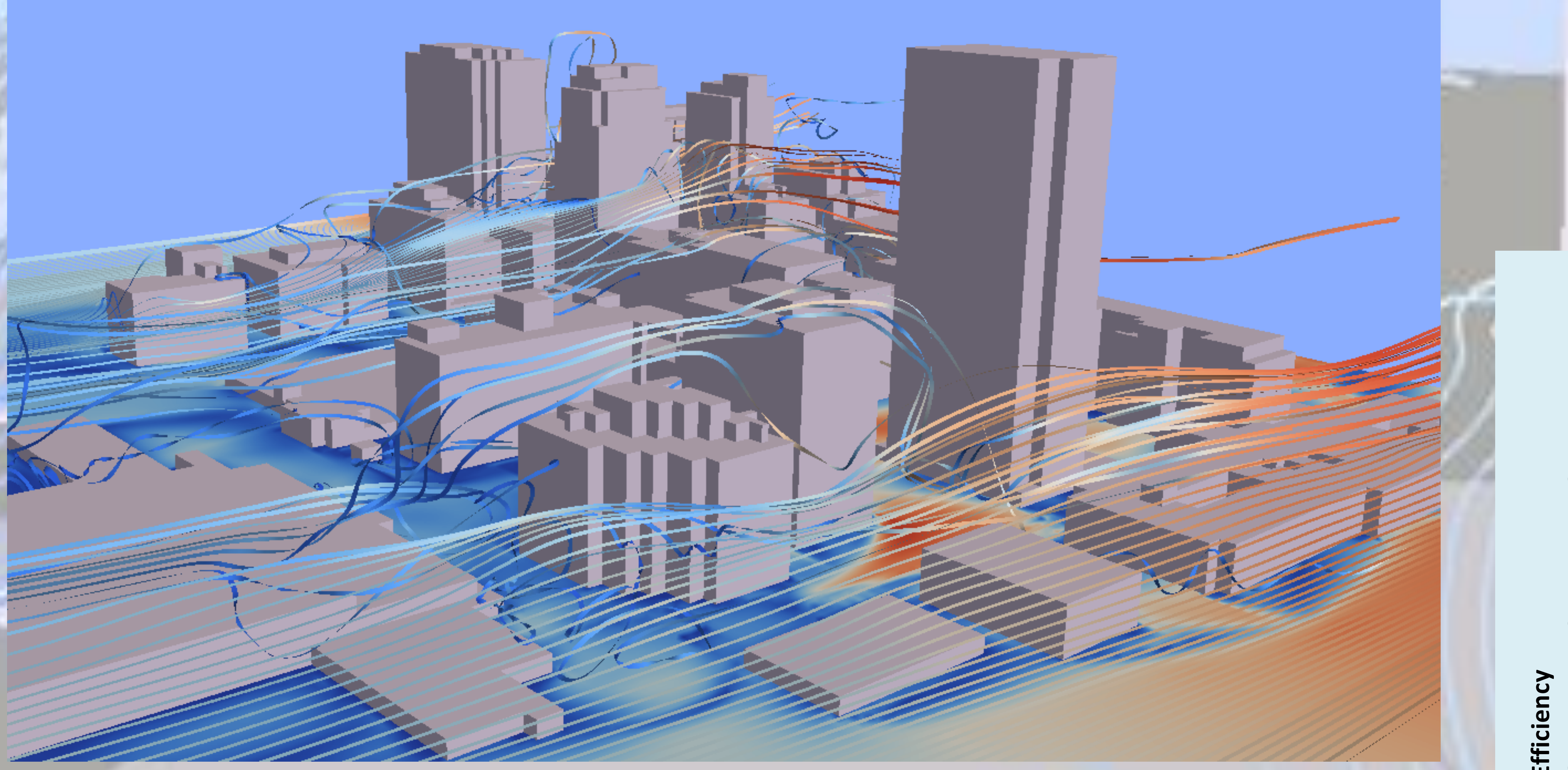


Temperature

Rayleigh-Bénard Convection

Temperature Field, Pr = 12.5, Ra = 0.35

Flow in Urban Environments

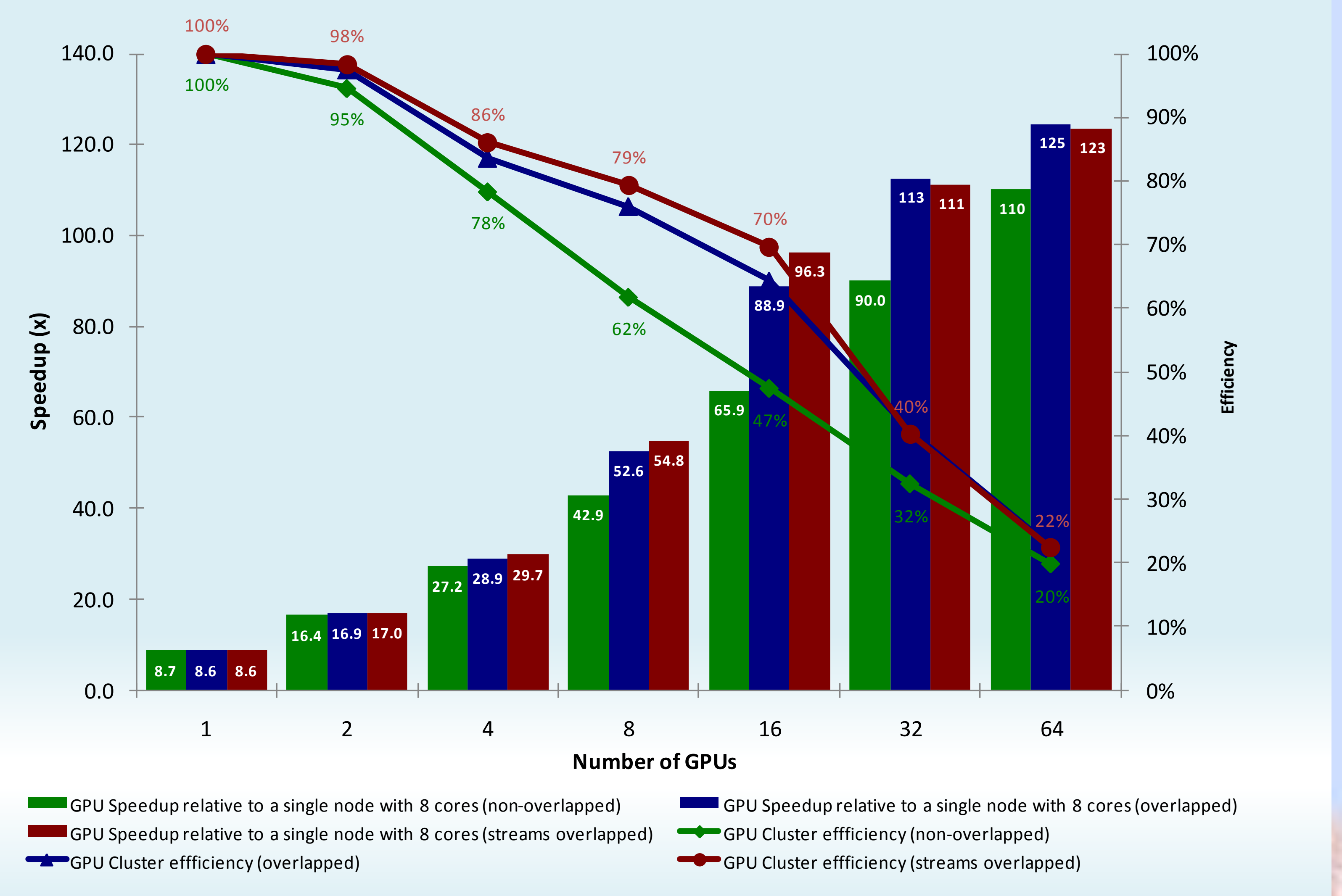


Conclusions

- Overlapping MPI with device copy and computation is essential for scalability
- For fixed problem size, GPU interexchange quickly becomes a bottleneck
- Speedups relative to 8 CPU cores: 17x for 1 GPU node, 123x for 32 GPU nodes, and 453x for projected CPU performance

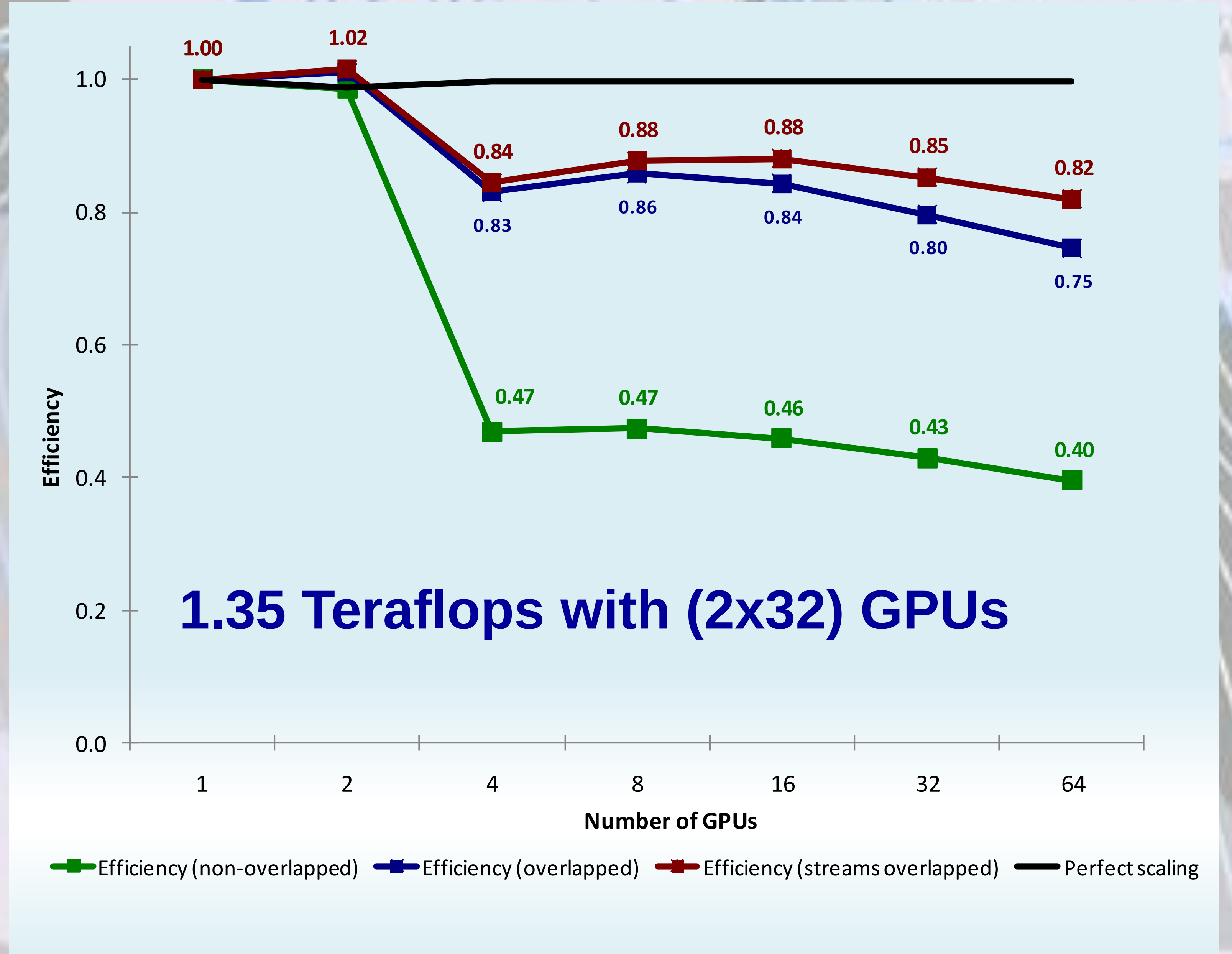
Strong Scalability

Problem size fixed at 1024x64x1024



Weak Scalability

Problem size per GPU fixed at ~3.8 GB



1.35 Teraflops with (2x32) GPUs

Acknowledgments: ■ ***NASA Idaho EPSCoR Research Initiation Grant*** ■ ***NVIDIA Professor Partnership Program*** ■ ***NCSA grant number ASC090054***