

# *NVSHMEM: OPENSHELL FOR GPU-CENTRIC COMMUNICATION*

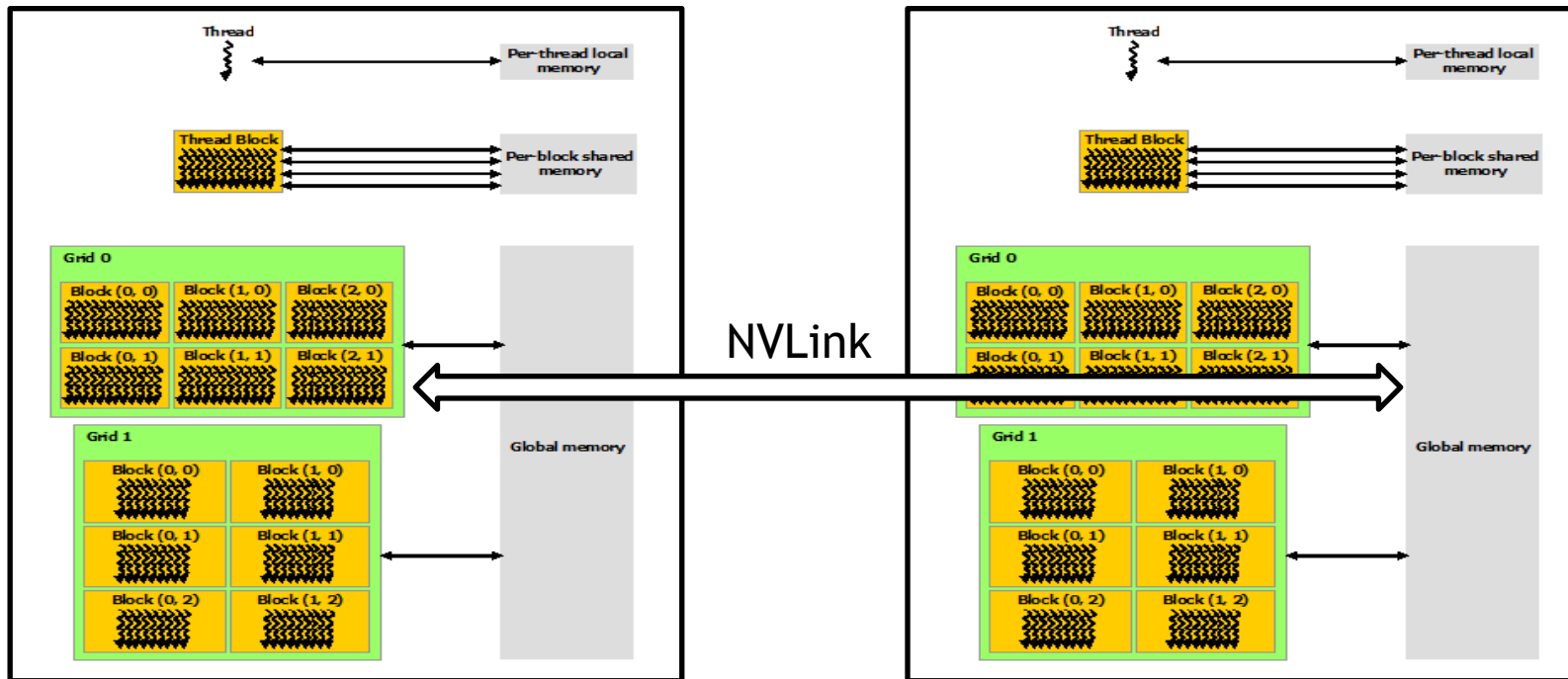
Sreeram Potluri

Senior Software Engineer



# BULK SYNCHRONY, THE BIG HAMMER

- More threads per GPU, more GPUs per system
- NVLINK connects GPUs in a node - LD/ST access to peer-GPU memory
- Reliance on CPU - limits scaling, limits utilization, harder to program

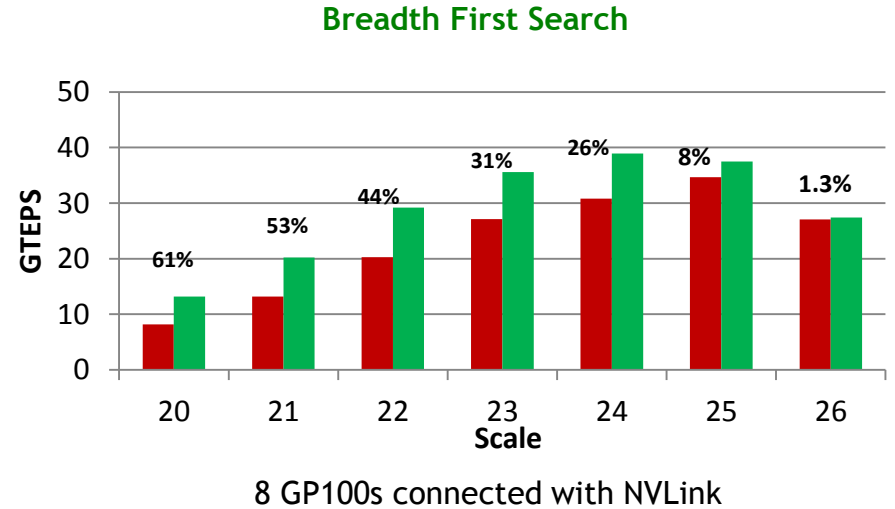
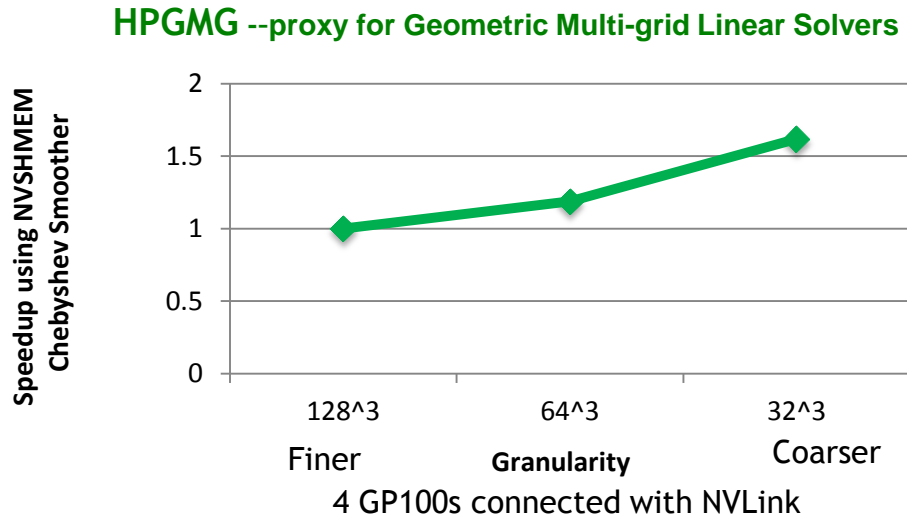
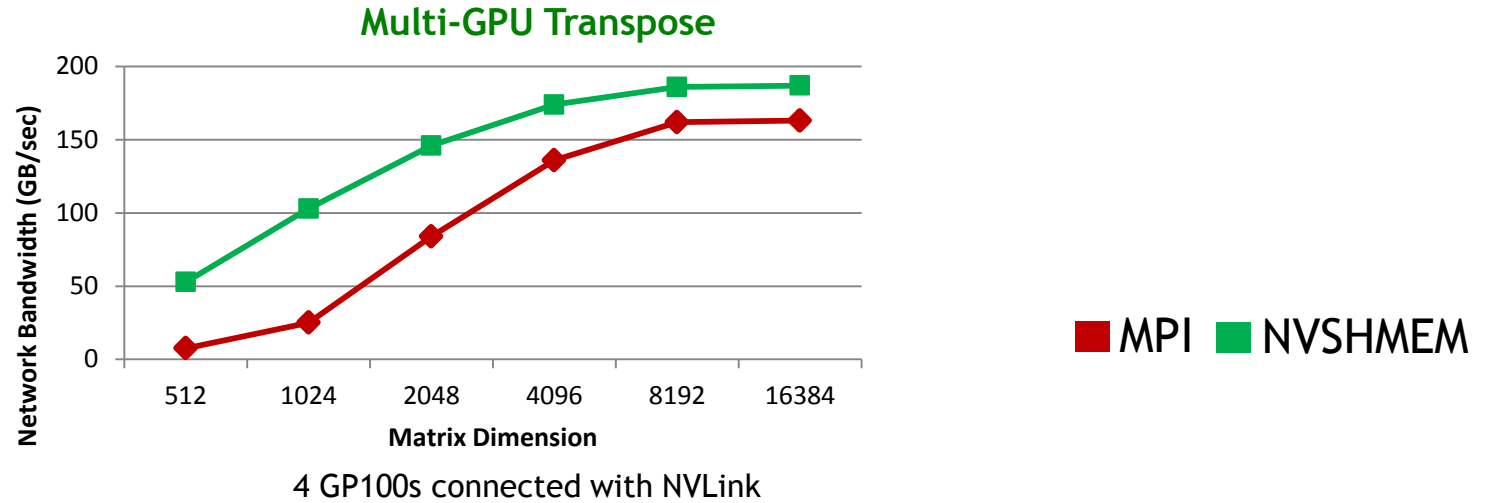


The need for a solution is increasing

# OPENSHMEM FOR GPU-INITIATED COMMUNICATION

- NVSHMEM is based on OpenSHMEM 1.3 (+ thread safety)
- GPU-side API for remote data access, collectives and synchronization
- Extensions for GPUs:
  - Stream-based API - CPU initiated communication but offloaded to GPU
  - Threadgroup-based API - GPU-side API to take advantage of concurrency on the GPU
- Early Access planned for Jan 2018 - GPUs within a single node - (NVLink, PCIe, QPI)
- Implementation over InfiniBand is under investigation
- Get in touch if you are interested in co-designing your application

# PERFORMANCE WITH MINI-APPS



Acknowledgements: Effort supported by funding from ORNL and DoD