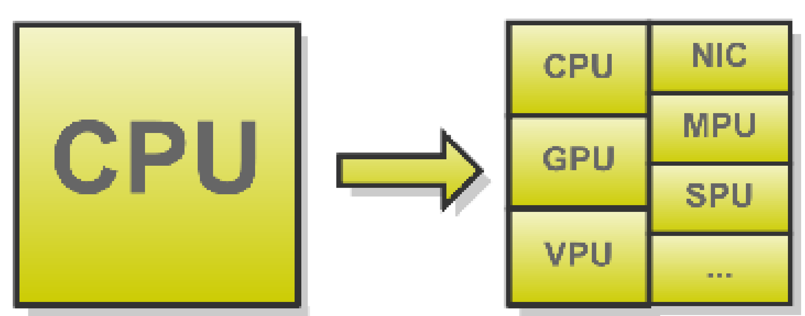
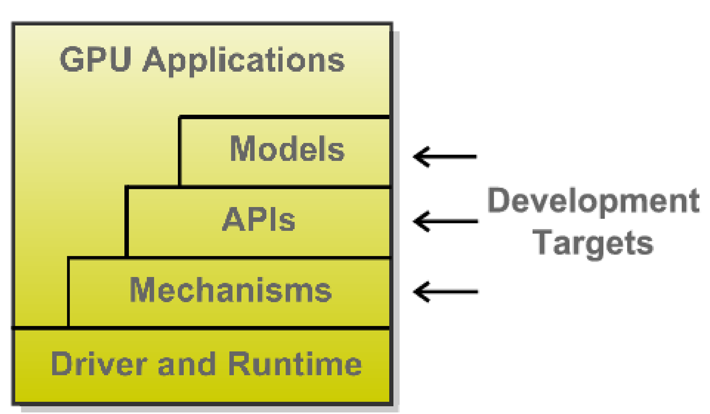
Cluster-Wide Multi-GPU Computing

Jeff Stuart — stuart@cs.ucdavis.edu
John D. Owens — jowens@ece.ucdavis.edu
University of California, Davis

Motivation & Approach

- Moving towards System on a Chip (SoC)
- GPU performance over CPU is substantial.
- GPU power consumption per watt is much better than that of CPU.





- The biggest impact of software systems comes in building GPU-application middleware.
- We've experimented with work in all three areas.
- Still many more opportunities.

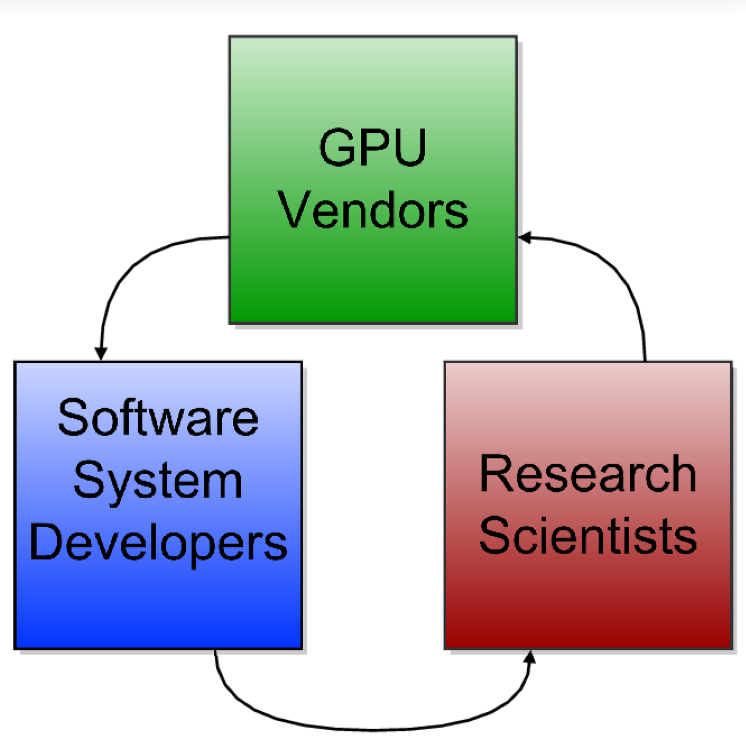
Development Cycle

- Wish list

  More autonomy

  GPU ←→ NIC

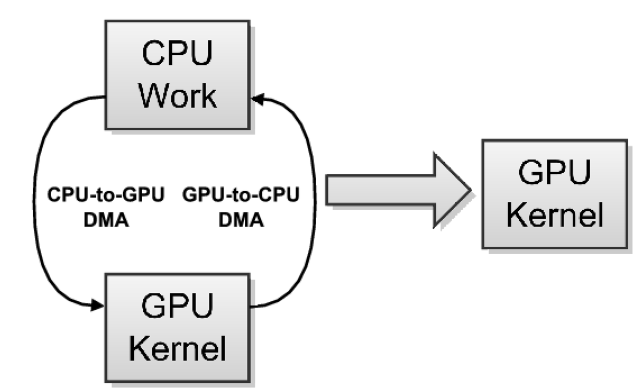
  Native TCP/IP, MPI



- As we implement support software, scientists use it and come up with new demands, spurring changes from GPU vendors, on top of which we implement new software systems.

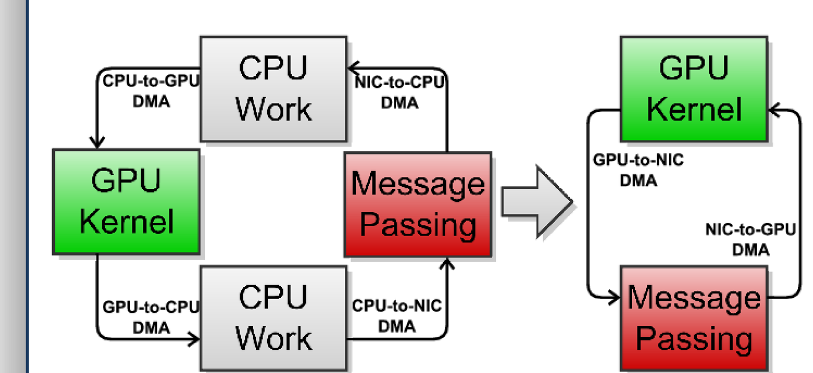
Projects

Callbacks [1]

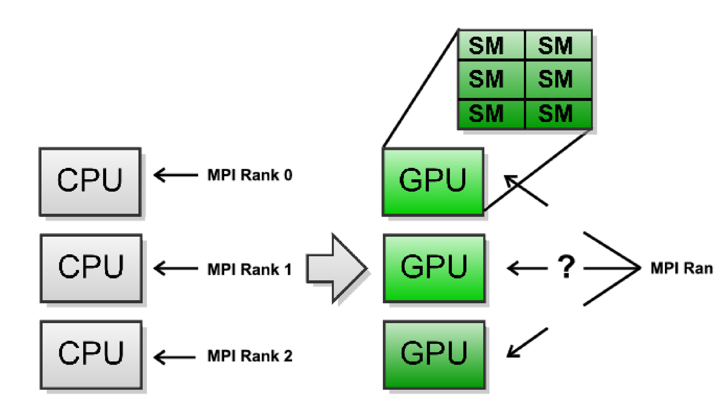


- Allows system calls
- GPU can DMA
- Arbitrary code execution
- Grants GPU autonomy

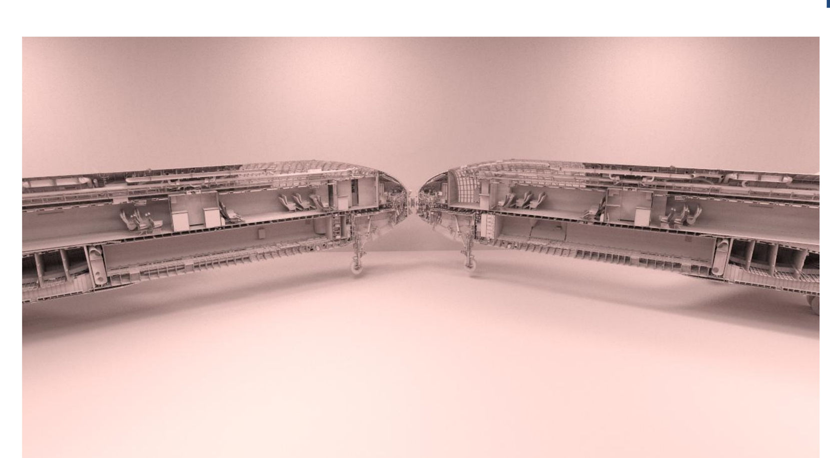
DCGN [2]



- GPU Kernels now have pseudo autonomy on the network.
- Introduces "slots"
  Intelligently multiplexes ranks from MPI
- Allows comm. between any two resources (CPUs and GPUs)



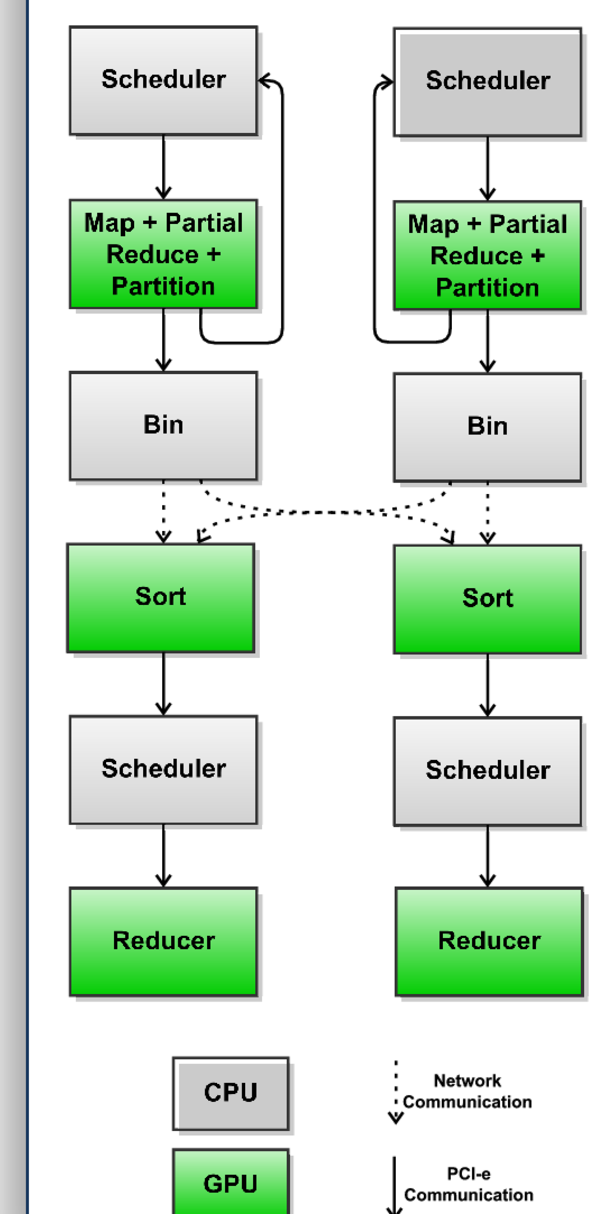
MGI [3]



- Resource manager for heterogeneous clusters.
- Run kernels on CPUs and GPUs.
- Move computation to data.
- Decide best resource based on hints from kernel descriptions, and on data locations. We tested using a path tracer.

Projects Continued

GPMR [4, 5]



- Cluster-wide implementation of MapReduce for GPUs
- Modifies MapReduce to better match GPU
  - Partial Reductions
  - Map/Reduce data merged into chunks
  - All modifications meant to reduce PCI-e and/or network traffic.

Future Work

- Explore heterogeneous scheduling
- Push for more GPU autonomy
- Develop infrastructure for exascale machines

[1] Jeff A. Stuart, Michael Cox, and John D. Owens. GPU-to-CPU Callbacks. In UnConventional High-Performance Computing 2010 as part of the EuroPar 2010 Workshop Series. September 2010.

[2] Jeff A. Stuart and John D. Owens. Message Passing on Data-Parallel Architectures. In Proceedings of the 23rd IEEE International Parallel and Distributed Processing Symposium, May 2009.

[3] Brian Budge, Tony Bernardin, Jeff A. Stuart, Shubhabrata Sengupta, Kenneth I. Joy, and John D. Owens. Out-of-core Data Management for Path Tracing on Hybrid Resources. In Computer Graphics Forum (Proceedings of Eurographics 2009), April 2009.

[4] Jeff A. Stuart, Cheng-Kai Chen, Kwan-Liu Ma, and John D. Owens. Multi-GPU Volume Rendering using MapReduce. In MAPREDUCE '10, The First International Workshop on MapReduce and its Applications. as part of the High-Performance and Distributed Computing 2010 Workshop Series. June 2010.

[5] Jeff A. Stuart and John D. Owens. Multi-GPU MapReduce on GPU Clusters. To Appear International Parallel and Distributed Processing Symposium (IPDPS), May 2011.