STORM employs several innovative mechanisms to enable extremely fast job launching, even on thousands of nodes:
- use of hardware collective communications to multicast binary and data files.
- LOO-bypass mechanism to transmit the files directly by the NICs, without CPU intervention.

Figure 1 shows measured results for job launching on a 236-processor Amdahl K840 cluster at LANL, (154th in the top500 list). Results are shown for three binary image sizes and are split into the time to send the binary file from the file server to all the compute nodes, and the time to actually execute it.

Figure 2 shows the measured and predicted launch times of a 12MB executable for up to 16,384 nodes. We used a very detailed model to predict STORM's launch time, which remains well under a second even for 16K nodes.

Figure 3 shows the effect of using different time-quantum values when running two copies of a job concurrently. We consider a synthetic compute-bound job and SWEEP3D, a real LANL kernel. The overhead of context-switching only appears when using time quantum of less than 2ms. In fact, Figure 3 shows that STORM can perform a global context-switch with similar frequency and efficiency of that of a local-node, enabling interactive responsiveness even in a large cluster.

### Goals

- Scalable, lightweight and fast resource-management:
  - Resource Allocation
  - Job-launching
  - Cluster-wide synchronization and context-switching

Increase the usability of a cluster:
- Checkpointing and fault-tolerance
- Improved system utilization
- Improved system responsiveness

Testbed for current and new scheduling algorithms.

Orders of magnitude faster than existing production systems (ASCI Q at LANL, C-Plant at Sandia) and the best published results.

For more information:
"STORM: Lightning-Fast Resource Management"
In Proceedings of the IEEE/ACM SC2002
http://www.c3.lanl.gov/~fabrizio