Distributing Common Data to Worker Nodes

February 15, 2012 NuComp Meeting
Marko Slyz, with help from the DOCS group
and Fermigrid

Transfer methods:

CVMFS: Has a persistent cache on the worker node.
   Pulls down an entire file at a time.
   No throttling of bandwidth.

NFS: No throttling so server and network can bog
   down. At least NFS 4 can have replication.

Hadoop, Lustre, xrootd: No throttling of bandwidth, but
   Can keep data in several places. Hadoop can have
   "rack awareness". xrootd can get closest copy.

SRM: Queues up requests so server and network won't
   bog down.   No local caching.

Distribution Strategies

These can be used with any(?) of the transfer methods.

Spread-out accesses to data:
Stagger job starts -- maybe with *next_job_start_delay.*
Or maybe jobs can do I/O throughout their run-time.
Avoid peaks in I/O.

Pulling down data in pilot job:
Can use custom scripts to bring one copy of the
data files, and then run several user jobs with that
pilot.
Or just run longer jobs if that's convenient.