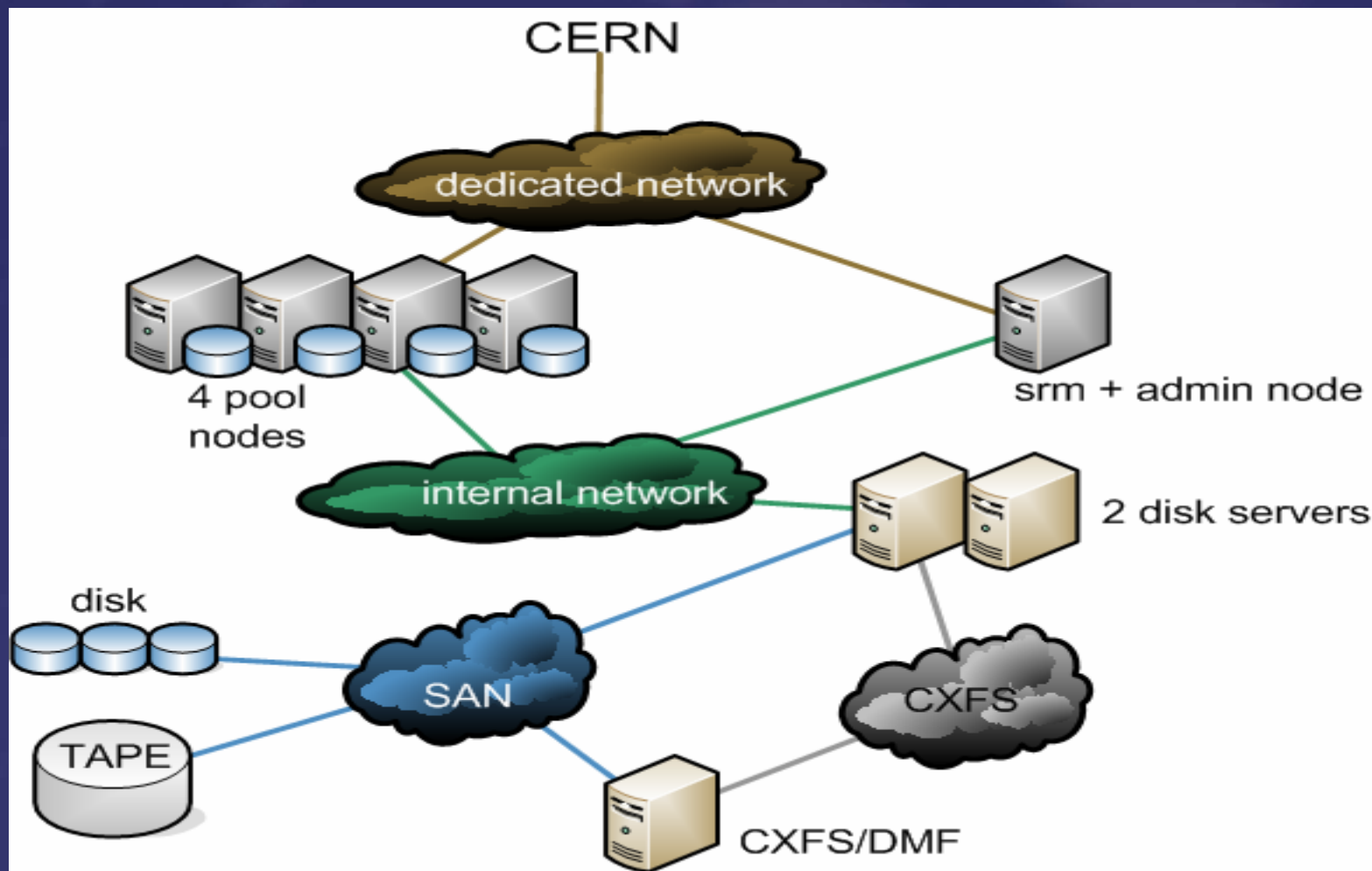


Status & results SC3 throughput phase SARA/Nikhef

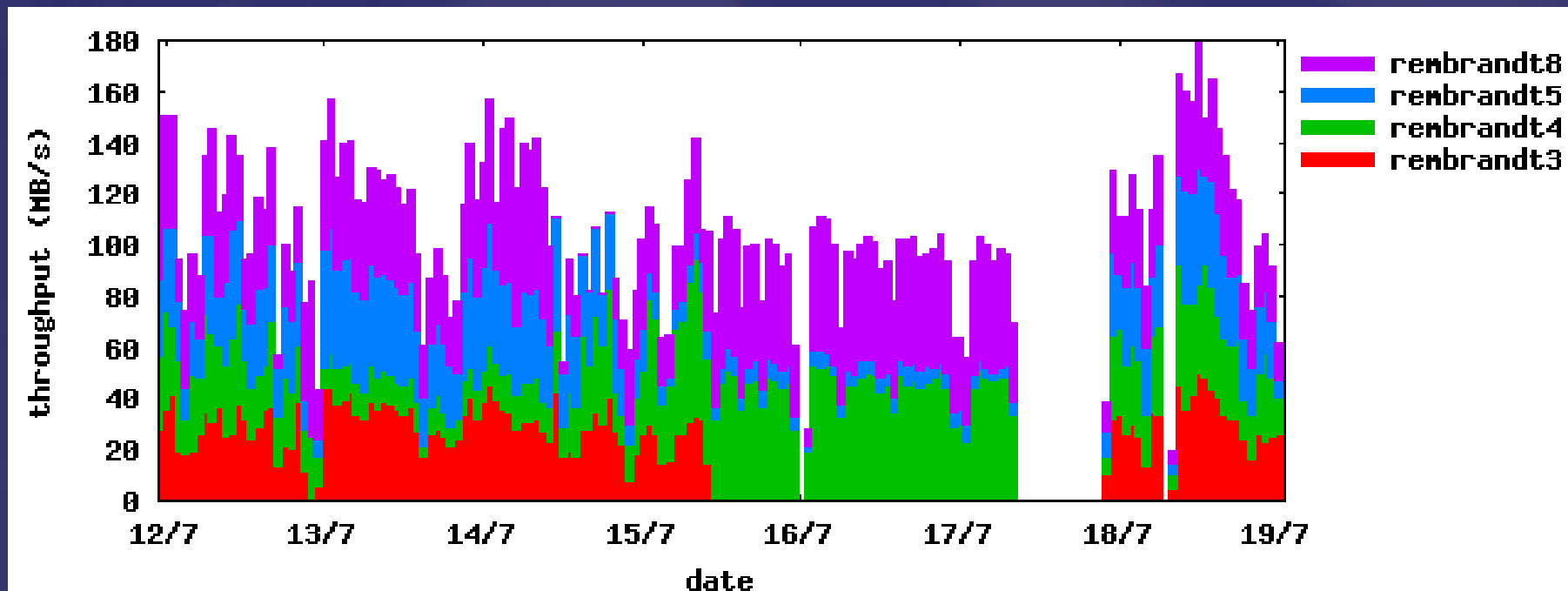
Mark van de Sanden



- CERN intake nodes (pool nodes)
 - 4x dual Opteron's, 4GB memory, 2x 1GE
 - 2TB disk cache, 12x 250GB SATA, 3ware RAID controller, disk I/O 200MB/s RAID0 (used during SC3) and 100MB/s RAID5
 - every node runs gridftp-door
- SRM node
 - dual Xeon, 4GB memory, 2x 73GB internal disk, 2x 1GE
 - runs also dCache admin, pnfs server and gridftp-door
- MSS gateway nodes (disk servers)
 - 2x dual Xeon, 4GB memory, 2x 73GB internal disk, 2x 1GE, dual HBA FC, 1.6 TB CXFS filesystem (SAN shared filesystem)
 - runs CXFS client, read/write data directly to/from CXFS filesystem
 - and rfio daemon to put/get data to/from pool nodes
- MSS server (CXFS/DMF)
 - 4 cpu R16K MIPS, 4GB memory, 12x FC, 4x GE, 2x 36GB internal disk, 1.6 TB CXFS filesystem (SAN shared filesystem), 3x STK 9940B tape drives
 - CXFS MDS server, regulates access to CXFS filesystem
 - DMF (Data Migration Facility = HSM system), migrates data from disk to tape and back
- Network
 - dedicated 10GE network between CERN – Amsterdam
 - GE internal network between pool nodes and MSS gateway nodes

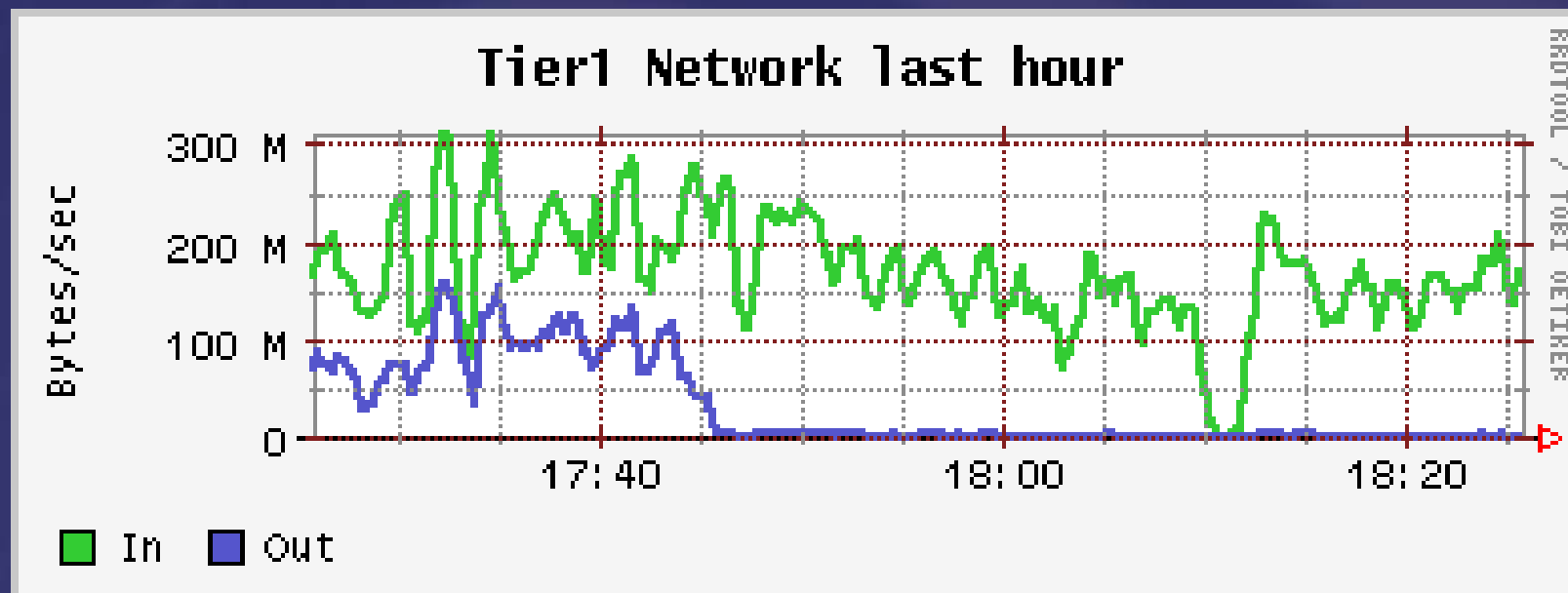
Gridftp performance

- ▣ 30-50MB/s average performance per pool node, much lower than hardware I/O performance
- ▣ 50-150MB/s aggregate performance pool nodes, not consistent and lower then SC3 goal



- First use of dCache/SRM
 - Not easy to manage, tune and to debug, little documentation
- Tuned IO movers down, default is 100, set it between 5-8
- Tuned heartbeat for better load balancing across the pools, default 120 sec, set it to 10 sec
- Gridftp server from dCache has low single stream performance (~1,7MB/s) and low single transfer performance (15MB/s with 10 streams)
 - dCache gridftp server has different behavior as Globus gridftp server (different tuning on host level)
- Maximum ~8 transfers per pool node, nodes failing due high load and memory usage
 - high cpu load due to large number gridftp servers with 8 concurrent transfers and 10 streams (~100 java server threads)
- Time outs on TURL returns from dCache
 - Cleaned up Postgress tables and restarted helped
- transfers between Gridftp-door and pool nodes (example rembrandt4 gridftp server writes to rembrandt3 pool) due to FTS implementation
 - srm PutRequests versus srmcp's

- Less cpu intensive, better manageable
- No network traffic between pool nodes



- Few down time periods due to CERN/Castor
- Most sites does not met SC3 disk-2-disk throughput goal
- Gridview
 - Have had some unstable moments
 - Too slow with updates, always trails behind a few hours, cannot be used for active monitoring (we use ganglia for host monitoring), is good for statistics.
 - Not all functions are implemented
 - Hour overview is only from the last 24 hours, wants to see longer periods