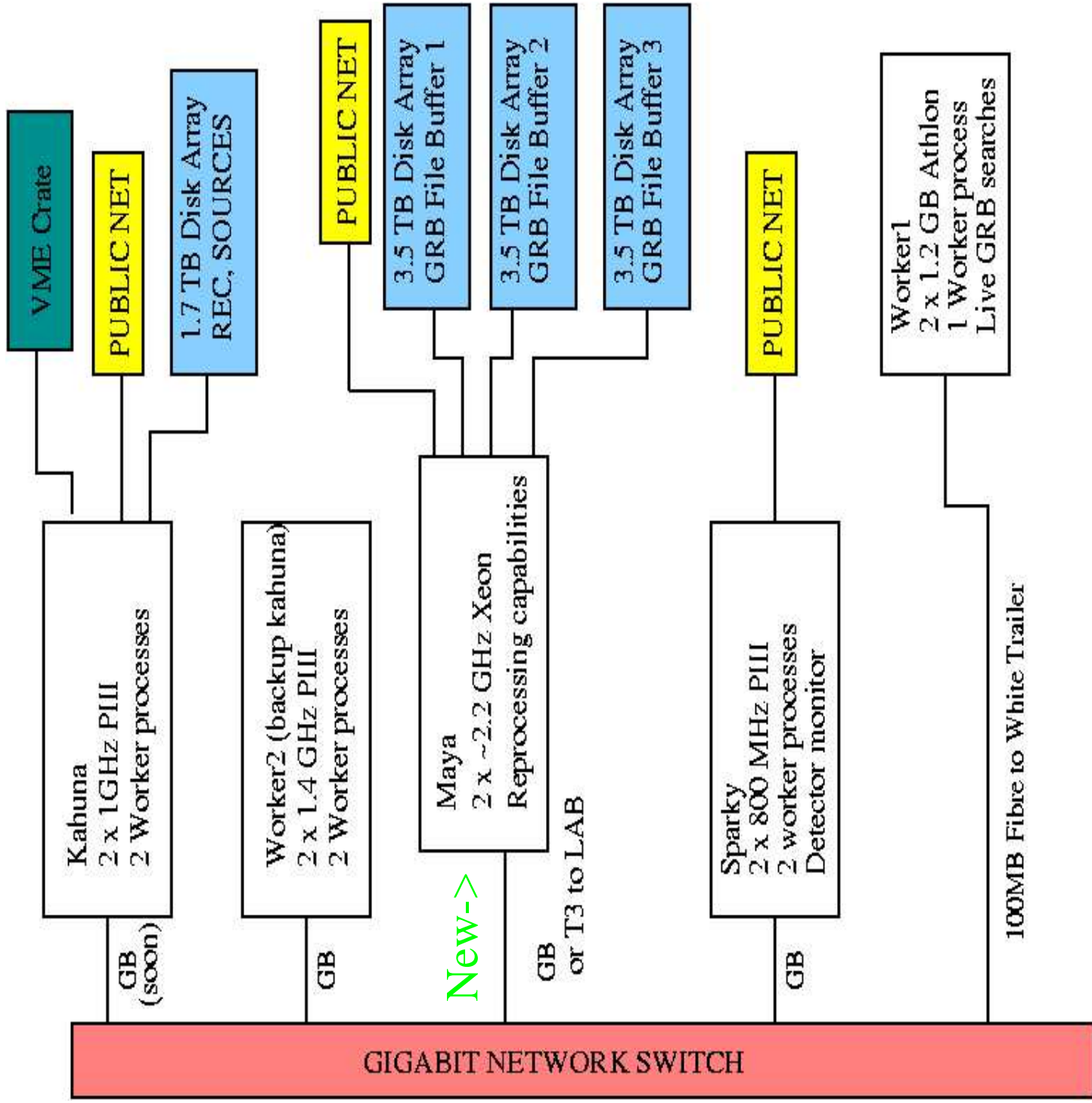


Archiving in the Tapeless Age

Erik Blaufuss-UMD

- Whats new at the site?
 - ~10 TB of disk and a server installed at site in mid-June.
- Current plan for archiving
- Options for moving data from the site
- The possibility that changes everything.

What's at the site now....



Current Plan

- Buffer of file with all RAW data (GRB files) saved at the site
 - ~10 TB of disk space gives ~70 days of data.
- Crab, GRB, SH, REC buffered to current array
 - REC- transferred over the network to lab/UMD
 - SH – transferred over network to UCI
 - Crab/GRB burst data- transferred via portable media to lab, and then on to UMD once at the lab.
- We are no longer writing the raw data streams:
 - Sun, Moon, MrkXXX, Save* (get them from REC)
 - * Save data is being written during the transition

Current status

- Disk arrays and server installed at the site
 - Ready to serve as RAW data buffer
- Julie is modifying the saveGRB and Archiver scripts to use these disk buffers instead of tape
 - Shift duties will include monitoring an “archiver status” page that monitors transfer of data to arrays.
- Testing and development of all parts of the new system will continue for the rest of the month.

Data transport method

- With limited network bandwidth to the site, we need a method to transfer Crab and GRB burst data to the lab.
- Data needs to be transferred and verified before array at site fills.
 - Need dedicated data manager onsite to oversee.
- 250 GB firewire portable disks are about \$300/ea
 - Fully reusable
 - Hold roughly 14 days of Crab data.
 - Tested with linux, works reasonably well.
- Tapes for Crab/burst data are another alternative.

Everything changes if...

- Gus recently found out about the possibility of getting a T3 (45 Mbps, 28 T1s) line to the site.
- If we can get one, we can reduce the amount of equipment at the site for archiving...
 - New disk arrays and server located at the lab
 - Kahuna, workers and modest disk buffer at site
 - All data (Raw, Rec, Crab) from the site is transferred over the network to the lab, where it is buffered and synced to UMD, as needed.
 - Disk arrays will be much happier in a clean, AC'd lab
 - No need for portable disks/tapes for moving data

What else is needed?

- Work needed to get the “2nd” site (the lab) for data storage up to speed.
 - Need a “data manager” to setup new equipment and babysit when they are cranky.
 - Still some equipment needed here
 - UPS, racks, servers
- Software to transfer and check data after transfer is needed
 - Needs to be automated.

Schedule

- Until T3 issue is decided, continue as planned
 - Should know within a week or so...
- ~July 2003
 - Software for archiving and syncing data to disk arrays finished and tested
 - Portable disk option/data manager “installed”
 - Resolve kahuna/VME/GB issue.
- August-September 2003
 - Save data stream stopped as GRB buffer size grows.