

Downtime

Downtime is a system for detecting failures within a cluster, tracking their resolution, and correlating them with service outages. It integrates with the Fault Tolerant Backplane to track managed/averted failures, and Cobalt to calculate the affect of failed systems.

There are some more concerns in the ImplementationPath, but some work has been done on a more formal SystemDesign.

collaborations

- Fault Tolerant Backplane
- Cobalt

questions

- How many times has a given system failed?
- How much compute time was lost?
- How many reservations were interrupted?

scope

- node availability
 - ◆ compute nodes
 - ◆ management nodes
- shared filesystem
 - ◆ PVFS
 - ◆ NFS
- network
 - ◆ internal network (inter-node)
 - ◆ external network

components

- issue tracker
- reporting scripts
- correlation engine
- report generation