Disk Cleanup

Introduction

Due to budget issues at SLAC, we are planning to clean off existing disk to make ready for launch. The plan is to clean off 100 TB of our existing 200 and buy 150 more. A concern is that we may have passed the time where ordering the disk will get it installed in time for launch.

The current usage so far is something like:

- I&T - 40 TB
- BT - 9 TB
- 40 TB - free
- 15 TB - various earmarks (systests, RM builds, user disk etc)
- MC - ~100 TB (by subtraction from above) - usage by task as of 1/20.

Note that we still need to run the Big Run/OpsSim2 datasets which should eat up most of the remaining free space.

There are complications with the I&T and BT data in that they are tied to the old eLog. And not all products linked to by the eLog are registered datasets in PipelineI. However we will need to do this sooner or later, since the disks will be retired at some point and the output on them must be saved for posterity,

Possible Strategy for MC Datasets

- identify obsolete simulations and simply expunge them
- identify key datasets and remove the MC, Recon and Digi files, leaving only the merit (etc) tuples
  - if not already in xrootd and the data catalogue, copy these into xrootd and the DC.