1. Astrophysics Imaging/Spectroscopy-Experiments/Modeling

- Palomar Transient Factory- Kulkarni - Caltech
  La Silla SN Search - Baltay
- BOSS - Schlegel - LBNL
- DES - Marriner/Nichols - FNAL/Portsmouth
- BigBOSS - Schlegel/Levi - LBNL
- JDEM - DoE/NASA
- SASIR - Mexico/UC - Bloom

Each of these groups will or are currently obtaining optical/NIR imaging and/or spectroscopic datasets of a large fraction of the sky. The goals of these programs are often multi-fold, spanning cosmological measurements from a variety of methods, to pure astrophysics and involve the creation of large catalogs of data both of a temporal and cumulative static nature. These observation are then typically confronted directly with simulation.

The projects acquire on order of 50 - 100's GBs per night and require real-time analysis to asses both the quality of the current data as wel las to discover transients such as supernovae.

2. Current HPC Requirements

(see slide notes)

- Range from large clusters (like PDSF) to running on Franklin
- For many 2GB/node is often enough, though for several applications having machines like davinci with 100 GB seen by several processors is a much more effective computational route.
- 100 - 1000 GB / day I/O
- Access to a system like NGF where a variety of computational resources can seen both the experimental and simulations is crucial from the infrastructure standpoint.
- Span most everything: FFT's, linear solvers, MC's,
- Known limitations: having access to db's on the compute nodes

3. HPC Usage and Methods for the Next 3-5 Years

(see slide notes)

- Upcoming changes to codes/methods/approaches: None right now
- Changes to Compute/memory load: real-time computing will become a priority
- Will grow to LSST proportions and the data becomes more valuable along the way
- Anticipated limitations/obstacles/bottlenecks on 10K-1000K PE system.
- Strategy for dealing with multi-core/many-core architectures is already in place, though small memory per node will be an obstacle as the image sizes grow in the future.

4. Summary

- Access to 100's of TB's of disk on NGF.
- The ability for users to work on/visualize the resultant data sets through the science gateway nodes
- Real-time computing capabilities (perhaps through cloud-computing)
- Access to the db's which contain information on both the observations and past simulations from the compute nodes.