HPCMP Enhanced User Environment

Workflow Demonstration

Presented by

Dave Cronk

Date: 01/27/2012

Outline

- Overview
- Utility Server
- Center-Wide Job Management
- Center-Wide File System
- Storage Lifecycle Management
- Secure Remote Desktop
- Demo

Overview

- Greater Horsepower leads to larger data
- Larger input decks
- More complex post processing
- More demanding data management needs
- Users don’t just need more horsepower, but need an infrastructure to improve their entire workflow

Utility Server

- **The Utility Server provides a single point of access**
  - **Use the Utility Server to submit and manage HPC jobs (Center-Wide Job Management)**
  - **Use the Utility Server to manage movement of data (Center-Wide File System)**
  - **Use the Utility Server as a pre/post processing engine**
  - **Use the Utility Server to remotely visualize your data (Secure Remote Desktop)**

Center-Wide Job Management

- **Center Wide Job Management (CWJM) allows users to manage their batch jobs on any unclassified HPC resource within a center from a single point of access, the Utility Server**
- **Supports both batch and interactive processing**
- **Uses standard PBS commands qsub, qstat, and qdel**
- **stdout/stderr are automatically returned to the Utility Server, in the directory from which the qsub command was executed**
- **Since the Utility Server is itself an HPC resource, you can also submit jobs directly to the Utility Server compute nodes**
  - Use normal PBS commands and scripts to run directly on the Utility Server
  - Add “@Remote_host” to the job script or command line for remote execution

Center-Wide File System (CWFS)

- Medium-term (30 day TTL) near HPC storage
  - More time to decide what data is important and needs to be archived
  - More time to post-process data
- Mounts to the HPC login nodes and the Utility Server.
- Provides, quick, easy way to move files between machines and file systems
- The CWFS at each center will contain approximately 1 Petabyte (PB) of usable online storage capacity
- NOT a replacement for scratch on HPC machines

Storage Lifecycle Management (SLM)

- Archive growth is outpacing technology advancements
- Most users use convoluted directory and file names to identify content
- Archive and forget
- SLM allows users to add metadata tags to their data objects
- Improved query functionality
- Set retention time
  - Data deleted from archive after no longer needed

What is Secure Remote Desktop?

- A client-server model integrated with the HPCMP security stack to deliver pixels to the desktop over a wide-area network and is be supported from the Utility Server

- Public Key Infrastructure Virtual Network Computing (PKI-VNC)
  - Allows a remote user to securely access a Linux desktop on the Utility Server. In practice, this allows a user to run ParaView, EnSight, Matlab, or any other software program that can run on Linux.

How Does it All Fit Together?

- Use the Utility Server's interactive, large memory capabilities to pre-process data and create an input deck
- Create a batch script which extracts the input deck and uses Center-Wide Job Management to execute the job on a large HPC system
- Copy the results back to the Center-Wide File System for medium term storage and post processing
  - CWFS provides sharing between HPC systems and Utility Server
- Using a Secure Remote Desktop (SRD) session on the Utility Server, conduct remote data processing and visualization without having to copy data to a local workstation
  - SRD gives you an actual desktop on the Utility Server for visualization on your system

DEMO