Panasas Update

HEPiX at DESY
Spring 2007

Robert Petkus

RHIC/USATLAS Computing Facility
Brookhaven National Laboratory

Panasas – Object Based Storage

- Panasas is a hardware/software object-based storage appliance intended for use in high performance computing clusters
- Features include:
  - Direct and parallel data access
  - Global namespace (ala AFS)
  - Load balancing

Two Hardware components

- Storage Blade (OSD)
  - Store and retrieve data objects
  - Handle client I/O
- Director Blade (File and metadata server)
  - DirectFlow or NFS file server
  - Coordinates between client and storage blades
  - (re)Distributes data across storage blades

Panasas – Object Based Storage

Two Software components

- Active Scale (Director Blade)
  - Divides data into objects for distribution across storage blades
  - Distributes load across storage blades
- Direct Flow
  - Linux kernel module for client node
  - Offers direct data path from client to storage blade
  - Data cache and prefetch
  - File reconstruction

Panasas at BNL

- Panasas is used as a centralized, POSIX-like, unix-browsible file system.
- It was intended as a replacement for traditional NFS storage
- Current deployment consists of 100 TB on 20 shelves divided into several realms (think AFS cell)
- 2Gb/s connectivity to each shelf
- DirectFlow client deployed on >1700 nodes

Future use of Panasas at BNL

BNL will discontinue use of Panasas when its warranty expires this summer

Why?

- Escalating disk and blade hardware failures
- Diminished quality of support
  - Client-side crashes related to Nessus security scanning not resolved
- Administrative hassles maintaining DirectFlow on the Linux farm
- Upgrades of server ActiveScale requires concurrent client DirectFlow upgrade (for major releases)
- Expensive

Panasas at JLAB

- Panasas was purchased to provide relief from the I/O bottlenecks inherent in a traditional NFS deployment
- Used primarily as a data store for experiment work areas
- 41 TB usable storage across 5x5TB and 3x8TB shelves
- DirectFlow is deployed on ~200 nodes with NFS access available to other systems

Future Use

- Panasas will likely be replaced in the future in favor of a solution like the SunFire x4500 (Thumper) or another system that will undergo evaluation
- Client hang-ups and freezes have been slow to resolve
- Costs are too expensive

Panasas at DESY - DV

- A 9.5 TB test system of (2) shelves containing (3) Director Blades and (8) 1Gb channel-bonded links performed well during a 6 month evaluation phase and was purchased.
- The system has been in use since May 2006
- 120 node farm accesses the storage via Direct Flow with no latency or performance issues.
- The experience with Panasas has been mostly positive
  - Recent ActiveScale upgrades have gone smoothly
  - A storage blade failure during the test phase only resulted in a single file that was unavailable for some time without corruption
- Some issues include
  - The cumbersome process of requesting a new DirectFlow client each and every time the kernel is updated
  - The prohibitively high cost

Questions / Comments ?