A tale of 3 storage requirements

Archive, Science Operations, and Admin systems

Brennan Hay

Definitions, Functions & Caveat

- Administrative
  - Business
  - headquarters data
- Mission
  - Weather observation
  - Archive systems
  - Science data
  - Supercomputing / HPCC
- May not cover unknown storage capabilities

CLASS for archive

- An archive
- Not a general storage pool
- Extremely high latency
  - Request responses in Days
- No backend (direct) interfaces?
  - tape storage
  - FC
- Incredibly Slow performance
- Structured datasets (exclusively?)

Science Ops Requirements

- High performance
  - SSD or Spinning disk
- Direct attached, SAN, or Networked
  - 10Gb, Infiniband, or 4/8GbFC
- Format
  - Blocks and Files
- Low latency
- Sized appropriately to data sets
  - Depends on size of data set
  - Most likely in terabytes to a Pb.
- Data can be unstructured or structured

Admin System Storage Requirements

- General storage pool
- May include PII (SECURITY)
- Low performance for users
  - Slow reads not a big issue for individuals
    - May be an issue in the aggregate
- High performance for servers
  - IOPS is an issue
- Extremely duplicative
  - NOAA stores copies of the same thing.
- Format
  - Blocks and Files

Enterprise Storage solutions Visualized by similarities

Archive type storage in NOAA

High performance storage

CLASS

HPCS

(Nothing enterprise)

General unstructured data

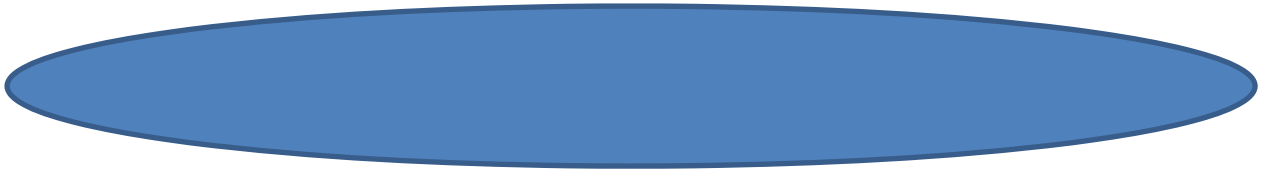
(Nothing enterprise)

Asking the right question

- Why aren't CLASS and HPCS sharing resources?
- Do we need a general storage, or high performance storage capability for science operations and admin systems?
- Can one capability meet all of above?
  - Archive
  - High performance
  - Low performance / PII

A service oriented capability?

Staff



Storage

Performance	**Slow**	**Fast**	**Ludicrous**
Interfaces	**Direct | SAN | Network | Infiniband**		
Availability	**Best effort | 99% | 99.9% | 99.99% | 99.999% | 99.9999%**		
Bandwidth	**10Mb | 100Mb | 1000Mb**		