Data Management & Storage for NGS

2009 Pre-Conference Workshop

Chris Dagdigian

BIOTEAM
Enabling Science

BioTeam Inc.

- **Independent Consulting Shop:**
  Vendor/technology agnostic

  Staffed by:
  - Scientists forced to learn High Performance IT to conduct research

- **Our specialty:**
  Bridging the gap between Science & IT

BIOTEAM
Enabling Science

Setting the stage

- Data Awareness
- Data Movement
- Data Management
- Storage & Storage Planning
- Storage Requirements for NGS
- Putting it all together …

BIOTEAM
Enabling Science

The Stakes …

180+ TB stored on lab bench

The life science “data tsunami” is no joke.

BIOTEAM
Enabling Science

Data Awareness

- First principals:
  - Understand chemistry changes faster than IT
  - Understand the data you will *produce*
  - Understand the data you will *keep*
  - Understand how the data will *move*
- Second principals:
  - One instrument or many?
  - One vendor or many?
  - One lab/core or many?

BIOTEAM
Enabling Science

Data You Produce

- Important to understand data sizes and types on an instrument-by-instrument basis
  - How many instrument runs per day/week?
  - What IT resources required for each basecall made?
- Will have a significant effect on storage performance, efficiency & utilization
- Where it matters:
  - Big files or small files?
  - Hundreds, thousands or millions of files?
  - Does it compress well?
  - Does it deduplicate well?

BIOTEAM
Enabling Science

Data You Produce

- Cliché NGS example
  - *Raw* instrument data
    - Massive image file(s)
  - *Intermediate* pipeline data
    - Raw data processed into more usable form(s)
  - *Derived* data
    - Results (basecalls & alignments)
    - Wiki's, LIMS & other downstream tools

BIOTEAM
Enabling Science

Data You Will Keep

- Instruments producing terabytes/run are the norm, not the exception
- Data triage is real and here to stay
  - Triage is the norm, not the exception in 2009
  - Sometimes it is cheaper to repeat experiment than store all digital data forever
- Must decide what data types are kept
  - And for how long …
- Raw data $\Rightarrow$ Result data
  - Can involve 100x reduction in data size

BIOTEAM
Enabling Science

General Example - Data Triage

- Raw Instrument Data
  - Keep only long enough to verify that the experiment worked (7-10 days for QC)
- Intermediate Data
  - Medium to long term storage (1year to forever)
  - Tracked via Wiki or simple LIMS
  - Can be used for re-analysis
    - Especially if vendor updates algorithms
- Result Data
  - Keep forever

BIOTEAM
Enabling Science

Applying the example …

- Raw Instrument Data
  - Instrument-attached local RAID
  - Cheap NAS device
  - Probably not backed up or replicated
- Intermediate Data
  - Almost certainly network attached
  - Big, fast & safe storage
    - Big for flexibility & multiple instruments
    - Fast for data analysis & re-analysis
    - Safe because it is important data & expensive to recreate
- Result Data
  - Very safe & secure
  - Often enterprise SAN or RDBMS
  - Enterprise backup methods

BIOTEAM
Enabling Science

NGS Vendors don't give great advice

- Skepticism is appropriate when dealing with NGS sales organizations
  - Essential to perform your own diligence
- Common issues:
  - Vendors often assume that you will use only their products; interoperability & shared IT solutions are not their concern
  - May lowball the true cost of IT and storage required if it will help make a sale

BIOTEAM
Enabling Science

Data Movement

- Facts
  - Data captured does not stay with the instrument
  - Often moving to multiple locations
  - Terabyte volumes of data could be involved
  - Multi-terabyte data transit across networks is rarely trivial no matter how advanced the IT organization
  - Campus network upgrade efforts may or may not extend all the way to the benchtop …

BIOTEAM
Enabling Science

Data Movement - Personal Story

- One of my favorite ‘09 consulting projects …
  - Move 20TB scientific data out of Amazon S3 storage cloud
- What we experienced:
  - Significant human effort to swap/transport disks
  - Wrote custom DB and scripts to verify all files each time they moved
    - Avg. 22MB/sec download from internet
    - Avg. 60MB/sec server to portable SATA array
    - Avg. 11MB/sec portable SATA to portable NAS array
  - At 11MB/sec, moving 20TB is a matter of *weeks*
  - *Forgot to account for MD5 checksum calculation times*
- Result:
  - Lesson Learned: data movement & handling took 5x longer than data acquisition

BIOTEAM
Enabling Science

Data Movement: Recommendations

- Network & network design matters
- Gigabit Ethernet has been a commodity for years
  - Don't settle for anything less
- 10 Gigabit Ethernet is reasonably priced in 2009
  - We still mostly use this for connecting storage devices to network switches
  - Also for datacenter to lab or remote building links
  - 10GbE to desktop or bench top not necessary
    - *10GbE to nearby network closet may be*
- Portable disk enclosures might be appropriate
  - Remember to account for time needed for copying and checksum activities
  - Safe & secure storage is important

BIOTEAM
Enabling Science

Data Movement: Recommendations

- Don’t bet your experiment on a 100% perfect network
  - Instruments writing to remote storage can be risky
  - Some may crash if access is interrupted for any reason
  - Stage to local disk, then copy across the network

- Network focus areas:
  1. Instrument to local capture storage
  2. Capture device to shared storage
  3. Shared storage to HPC resource(s)
  4. Shared storage to desktop
  5. Shared storage to backup/replication

BIOTEAM
Enabling Science

Data Management

- Very important
  - Especially if multiple IT & storage systems involved
- There is no universal solution
  - Too many variables in protocol, data & research flows
- We have seen many different methods adopted
  - LIMS, Wiki, Spreadsheets, etc.
  - All have pros and cons
- Choosing
  - Flexibility is key
    - Chemistry or SOP might change faster than a typical LIMS design lifecycle can handle
    - All solutions are useless if unused

BIOTEAM
Enabling Science

Storage Requirements for NGS

What features do we actually need?

BIOTEAM
Enabling Science

“Must Have”

- High capacity & scaling headroom
- Variable file types & access patterns
- Multi-protocol access options
- Concurrent read/write access

BIOTEAM
Enabling Science

“Nice to have”

- Single-namespace scaling
  - No more “/data1”, “/data2” buckets …
  - Horrible cross mounts, bad efficiency
- Low Operational Burden
- Appropriate Pricing*
- “A la cart” feature and upgrade options

BIOTEAM
Enabling Science

Capacity

- Chemistry/instruments improving faster than our IT infrastructure
  - Flexibility is essential to deal with this
- If we don’t address capacity needs:
  - Expect to see commodity NAS boxes crammed into lab benches and telco closets
  - Expect hassles induced by island of data
  - Backup issues (if they get backed up at all)
  - … and lots of USB drives on office shelves …

BIOTEAM
Enabling Science

Remember The Stakes …

180+ TB stored on lab bench

The life science “data tsunami” is no joke.

BIOTEAM
Enabling Science

File Types & Access Patterns

- Many storage products are optimized for particular use cases and file types
- Problem
  - Life Science & NGS can require them all:
    - Many small files vs. fewer large files
    - Text vs. Binary data
    - Sequential access vs. random access
    - Concurrent reads against large files

BIOTEAM
Enabling Science

Multi-Protocol Is Essential

- The *overwhelming* researcher requirement is for *shared* access to *common* filesystems
  - Especially true for next-gen sequencing
  - Lab instrument, cluster nodes & desktop workstations all need access the same data
  - This enables automation and frees up human time
- Shared storage in a SAN world is non-trivial
- Storage Area Networks (SANs) are not the best storage platform for discovery research environments

BIOTEAM
Enabling Science

Storage Protocol Requirements

- NFS
  - Standard method for file sharing between Unix hosts
- CIFS/SMB
  - Desktop access
  - Ideally with authentication and ACLs coming from Active Directory or LDAP
- FTP/HTTP
  - Sharing data among collaborators

Concurrent Storage Access

- Ideally we want read/write access to files from
  - Lab instruments
  - HPC / Cluster systems
  - Researcher desktops
- If we don't have this
  - Lots of time & core network bandwidth consumed by data movement
  - Large & possibly redundant data across multiple islands
  - Duplicated data over islands of storage
  - Harder to secure, harder to back up (if at all …)
  - Large NAS arrays start showing up under desks and in nearby telco closets

BIOTEAM
Enabling Science

Data Drift: Real Example

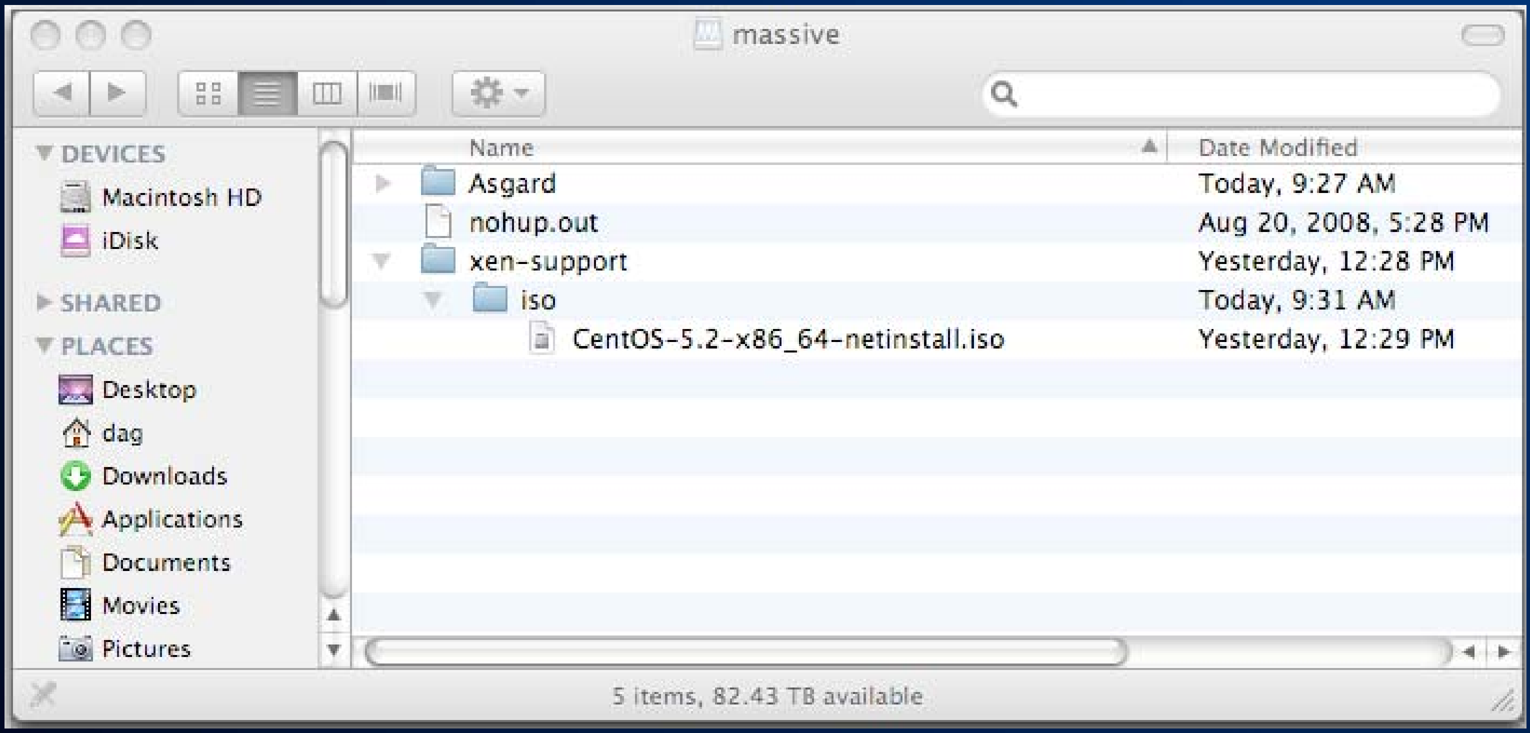
- Non-scalable storage islands add complexity
- Example:
  - Volume “Caspian” hosted on server “Odin”
  - “Odin” replaced by “Thor”
  - “Caspian” migrated to “Asgard”
  - Relocated to “/massive/”

- Resulted in file paths that look like this:

```
/massive/Asgard/Caspian/blastdb
/massive/Asgard/old_stuff/Caspian/blastdb
/massive/Asgard/can-be-deleted/do-not-delete…
```

BIOTEAM
Enabling Science

Single-namespace is valuable



BIOTEAM
Enabling Science

Things To Think About

An attempt at some practical advice …

BIOTEAM
Enabling Science

Storage Landscape

- Storage is a commodity in 2009
- Cheap storage is easy
- Big storage getting easier every day
- Big, cheap & *SAFE* is much harder …
- Traditional backup methods may no longer apply
  - Or even be possible …

BIOTEAM
Enabling Science

Storage Landscape

- Still see extreme price ranges
  - Raw cost of 1,000 Terabytes (1PB):
    - $125,000 to $4,000,000 USD
- Poor product choices exist in all price ranges

BIOTEAM
Enabling Science

Poor Choice Examples

- On the low end:
  - Use of RAID5 (unacceptable in 2009)
  - Too many hardware shortcuts result in unacceptable reliability trade-offs

BIOTEAM
Enabling Science

Poor Choice Examples

- And with high end products:
  - Feature bias towards corporate computing, not research computing - pay for many things you won't be using
  - Unacceptable hidden limitations (size or speed)
  - Personal example:
    - $800,000 70TB (raw) Enterprise NAS Product
    - *… can't create a NFS volume larger than 10TB*
    - *… can't dedupe volumes larger than 3-4 TB*

BIOTEAM
Enabling Science

One slide on RAID 5

- I was a RAID 5 bigot for many years
  - Perfect for life science due to our heavy read bias
  - Small write penalty for parity operation no big deal
- RAID 5 is no longer acceptable
  - Mostly due to drive sizes (1TB+), array sizes and rebuild time
  - *In the time it takes to rebuild an array after a disk failure there is a non-trivial chance that a 2nd failure will occur, resulting in total data loss*
- In 2009
  - Only consider products that offer RAID 6 or other "double parity" protection methods
  - Even RAID 6 is a stopgap measure …

BIOTEAM
Enabling Science

Research vs. Enterprise Storage

- Many organizations have invested heavily in centralized enterprise storage platforms
  - Natural question: *Why don't we just add disk to our existing enterprise solution?*
  - This may or may not be a good idea
    - NGS capacity needs can easily exceed existing scaling limits on installed systems
    - Expensive to grow/expand these systems
    - Potential to overwhelm existing backup solution
    - NGS pipelines hammering storage can affect other production users and applications

BIOTEAM
Enabling Science

Research vs. Enterprise Storage

- Monolithic central storage is not the answer
- There are valid reasons for distinguishing between enterprise storage and research storage
- Most organizations we see do not attempt to integrate NGS process data into the core enterprise storage platform
  - Separate out by required features and scaling needs

BIOTEAM
Enabling Science

Putting it all together …

Remember this slide ?

- First principal:
  - Understand the data you will *produce*
  - Understand the data you will *keep*
  - Understand how the data will *move*
- Second principal:
  - One instrument or many?
  - One vendor or many?
  - One lab/core or many?

BIOTEAM
Enabling Science

Putting it all together

- Data Awareness
  - What data will you produce, keep & move?
  - Size, frequency & data types involved
- Scope Awareness
  - Are you supporting one, few or many instruments?
  - Single lab, small core or entire campus?
- Flow Awareness
  - Understand how the data moves through full lifecycle
    - Capture, QC, Processing, Analysis, Archive, etc.
    - What people & systems need to access data?
    - Can my networks handle terabyte transit issues?

BIOTEAM
Enabling Science

Putting it all together, cont.

- Data Management
  - What files & data need to be tracked?
  - … and by what method
    - LIMS, WIKI, Excel, Dir Structure, etc.
- Data Integrity
  - Backup, replicate or recreate?

BIOTEAM
Enabling Science

Example: Point solution for NGS

Self-contained lab-local cluster & storage for Illumina

BIOTEAM
Enabling Science

Example: Small core shared IT

100 Terabyte storage system and 10 node / 40 CPU core Linux Cluster supporting multiple NGS instruments

BIOTEAM
Enabling Science

Example: Large Core Facility

Matthew will discuss this in detail during the third talk…

BIOTEAM
Enabling Science

End;

- Thanks!
- Lots more detail coming in next presentations

- Comments/feedback:
  - chris@bioteam.net

BIOTEAM
Enabling Science