Science-Centric Storage

Chris Dagdigian

2010 Bio-IT World Expo Europe

Hannover, Germany

BIOTEAM
Enabling Science

Word of warning

- Known for speaking quickly and having a large slide deck
  - 60 slides in 30 minutes is pretty normal for me, sorry!
- Also an unrepentant PowerPoint fiddler
  - Final version will be posted on http://blog.bioteam.net/

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

Science Driven Storage

BIOTEAM
Enabling Science

First the good news …

BIOTEAM
Enabling Science

Late 2010 – The Good News

- Far less scary today than it was in 2004-2006
- Petascale storage is no longer a risky technological gamble
  - And hasn't been for years
- Now it's just an IT design & budget exercise
  - Proven solutions available from multiple vendors
  - Existing petascale customers are speaking about and sharing their experiences
  - Choices for nearline & tiered storage increasing every day

BIOTEAM
Enabling Science

Late 2010 – The Good News

- Instrument-driven NGS “data deluge” becoming far more manageable (and eventually going away)
- Why?
  - Data triage is an accepted common practice
    - 1st talk today (Sanger): “BAM files only”
      - No images, no srf, no fastq
  - Current generation instruments doing more internal processing & data reduction
  - Next generation instruments moving entirely away from image-based sensor platforms
    - *Especially true for 4th-gen sequencing platforms*

BIOTEAM
Enabling Science

Now for the bad news …

Late 2010 – The Bad News

- NGS instruments are not the only "terabyte scale" tools showing up in wet labs
  - Many analytic tools coming out in HT form
  - Many imaging platforms switching going 2D to 3D/4D
    - Confocal microscopy, ultrasound, etc.
  - Cheaper lab instruments often mean that more are purchased
- End result still seems to be a need for large scale storage in discovery research environments

BIOTEAM
Enabling Science

Late 2010 - More bad news

- Managing User Expectations Still Difficult
  - End users still have no idea about true costs of keeping data accessible & available during it's lifecycle
    - *"I can get a terabyte from Costco for $220!"*
    - *"I can get a terabyte from Costco for $160!" (Oct 08)*
    - *"I can get a terabyte from Costco for $124!" (April 09)*
    - *"I can get a terabyte from NewEgg for $84!" (Feb 10)*
    - *2TB SATA for $109, 1TB SATA for $69 (Oct 10)*
- IT needs to be involved in setting expectations and educating on true cost of keeping data online & accessible
  - Everyone benefits when this happens

BIOTEAM
Enabling Science

Late 2010 – The Bad News

- Most worrisome trend:
  - As NGS instrument data rate declines, rate at which researchers are consuming downstream storage is increasing
    - *First heard this 1.5 years ago, now confirmed at multiple sites*
  - Why is this worrisome?
    - Storage requirements of researchers are far less predictable than instruments
    - Data mashups & widespread collaboration breeding tremendous data duplication

BIOTEAM
Enabling Science

Infrastructure Tour

What does this stuff look like in the real world?

BIOTEAM
Enabling Science

Infrastructure Tour

- Research protocols are changing faster than the underlying IT infrastructures that support them
- Vendors, products & strategies will differ depending on size, scope & services
- Are you:
  - Individual lab or PI?
  - Workgroup or department?
  - Small core facility?
  - Large core facility?
  - Other

BIOTEAM
Enabling Science

Example: Point solution for NGS

Self-contained lab-local cluster & storage for Illumina

Example: Point solution for NGS

Datacenter-resident infrastructure for 1-2 NGS systems

BIOTEAM
Enabling Science

Example: Shared IT for Midsized Core

BIOTEAM
Enabling Science

Example: Large Genome Center

BIOTEAM
Enabling Science

Example: Large Core Facility

BIOTEAM
Enabling Science

Example: Data Transfer Station

More & more human-driven large scale data movement is
being seen out in the field, for various reasons. It must be planned for.

BIOTEAM
Enabling Science

Example: Data Transfer Station

External e-SATA ‘toasters’ & portable RAID units

Example: Data Transfer Station

Data transfer station using hot-swap SATA disk bays

BIOTEAM
Enabling Science

Example: ‘naked’ data transfer

eSATA / USB 2.0 “toaster”

BIOTEAM
Enabling Science

Example: ‘naked’ data archive @ scale

Far cheaper than a true offline archive/tape solution. Not unreasonable in many use cases, especially for non-unique data..

BIOTEAM
Enabling Science

Example: ‘naked’ data archive @ scale

… prettier picture

BIOTEAM
Enabling Science

And we are all trying to avoid this …

BIOTEAM
Enabling Science

The Stakes …

180+ TB stored on lab bench. Primary data. No RAID, no backup.

BIOTEAM
Enabling Science

Science Driven Storage

Sizing, selecting & purchasing

BIOTEAM
Enabling Science

Data Awareness

- First principals:
  1. Understand science changes faster than IT
  2. Understand the data you will *produce*
  3. Understand the data you will *keep*
  4. Understand how the data will *move*

- Second principals:
  1. One instrument or many?
  2. One vendor or many?
  3. 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 Keep

- Terabyte scale instruments are the new norm
- No longer possible to give “unlimited” storage to researchers
- Data triage has not been controversial for years
  - Giant facilities: BAM files only, toss everything else
  - Mainstream: toss images, keep some intermediate data
- Key Questions
  - What do you keep? What do you throw away?
  - Online, nearline or offline?
  - What does “forever” mean?
  - What crazy things are your researchers going to do with all that data?

BIOTEAM
Enabling Science

Data You Move

- 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
  - Carrying/shipping physical media can be a solution
  - New in 2010:
    - Terabyte volumes of data arriving from outsourced providers

BIOTEAM
Enabling Science

Flops, Failures & Freakouts

Learning from past mistakes.

BIOTEAM
Enabling Science

#1 - Unchecked Enterprise Architects

- Scientist: "*My work is priceless, I must be able to access it at all times*"
- Storage Guru: "*Hmmm…you want high availibility, huh?*"
- *System delivered:*
  - 40TB Enterprise SAN
  - Asynchronous replication to remote site
  - Can't scale, can't do NFS easily
  - $~500K/year in maintenance costs

BIOTEAM
Enabling Science

#2 - Unchecked User Requirements

- Scientist: "*I do bioinformatics, I am rate limited by the speed of file IO operations. Faster disk means faster science.* "
- System delivered:
  - Budget blown on top tier 'Cadillac' system
  - Fast *everything*
- Outcome:
  - System fills to capacity in 9 months

BIOTEAM
Enabling Science

#3 - D.I.Y Cluster/Parallel File systems

- Common source of storage unhappiness

- Root cause:
  - Not enough pre-sales time spent on design and engineering
- System as built:
  - Not enough metadata controllers
  - Poor configuration of key components
- End result:
  - Poor performance or availability

BIOTEAM
Enabling Science

Lessons Learned

- End-users are not precise with storage terms
  - "Extremely reliable" means *no data loss*, not millions spent on 99.999% high availability
- When true costs are explained:
  - Many research users will trade a small amount of uptime or availability for more capacity or capabilities
  - … will also often trade some level of performance in exchange for a huge win in capacity or capability

BIOTEAM
Enabling Science

Lessons Learned

- End-users demand the world but are willing to compromise
  - Necessary to *really* talk to them and understand work, needs and priorities
  - Also necessary to explain true costs involved
- People demanding the “fastest” storage often don’t have actual metrics to present

BIOTEAM
Enabling Science

Lessons Learned

- Software-based parallel or clustered file systems are non-trivial to *correctly* implement
- Essential to involve experts in the initial design phase
  - *Even if using 'open source' version …*
- Commercial support is essential
  - *And I say this as an open source zealot …*

BIOTEAM
Enabling Science

Informatics Storage Requirements

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

- Don’t forget “lessons learned” ...

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

- Our science is expanding 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 …

Motivated researchers will solve their own problems with or without help from IT …

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

BIOTEAM
Enabling Science

Concurrent Storage Access

- Ideally we want read/write access to files from
  - Lab instruments & instrument control workstations
  - 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:
  1. Volume “Caspian” hosted on server “Odin”
  2. “Odin” replaced by “Thor”
  3. “Caspian” migrated to “Asgard”
  4. Relocated to “/massive/”

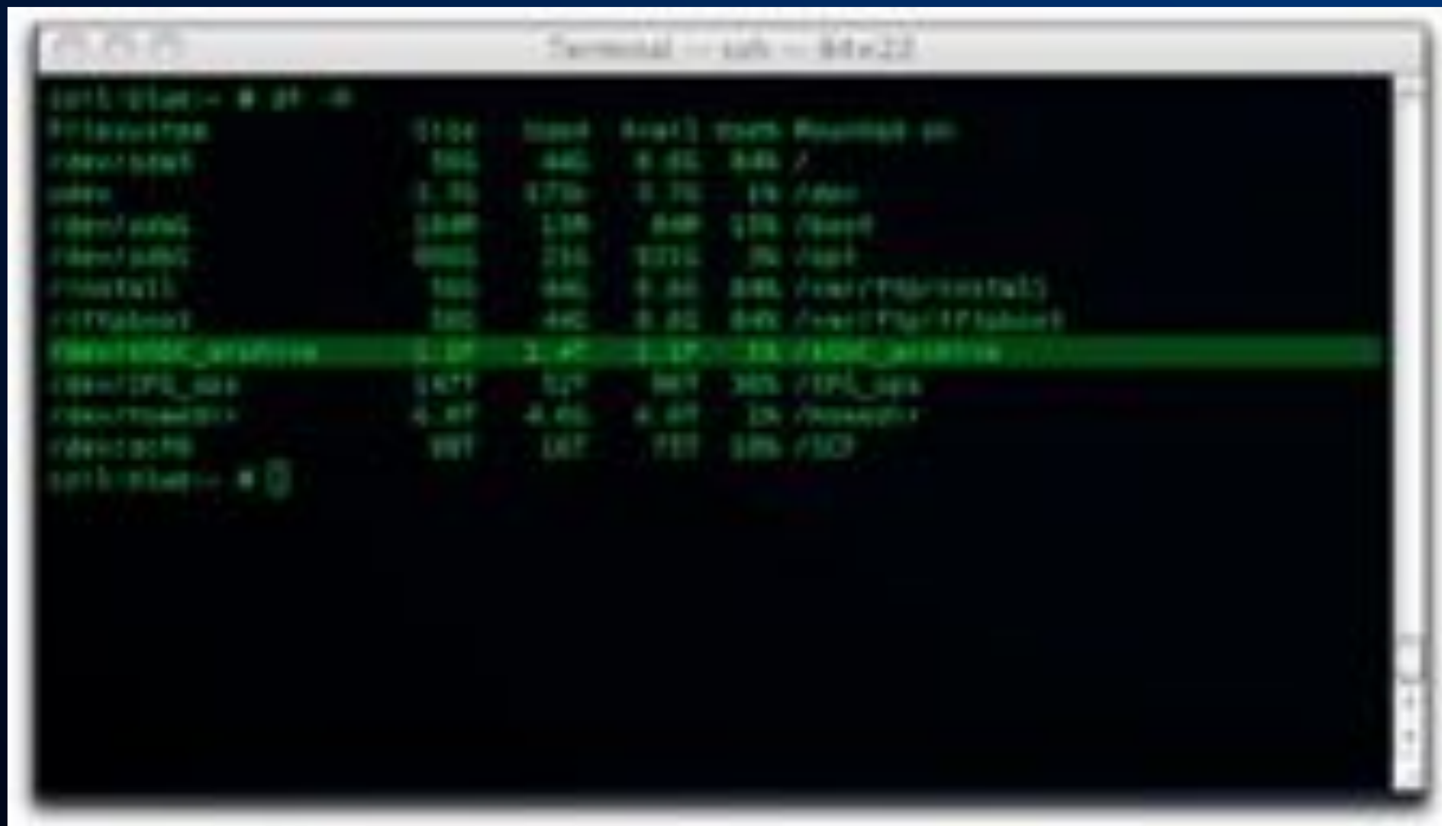
- Resulted in file paths that look like this:

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

BIOTEAM
Enabling Science

Nerdvana

- 1 petabyte available in single folder:



BIOTEAM
Enabling Science

Things To Think About

An attempt at some practical advice …

BIOTEAM
Enabling Science

Storage Landscape

- Storage is a commodity in 2010
- 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 for years now)
  - 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)
  - 2009 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*

And Finally …

Can’t have an IT talk without the “C-word”

BIOTEAM
Enabling Science

Few Slides on Cloud Storage

- I drank the kool-aide
- I think cloud storage is the future
- Economics alone make this inescapable

BIOTEAM
Enabling Science

Cloud Storage Inevitable

- I don’t care where you work
  - You will never come close to the at-scale operational efficiencies seen by Amazon, Google & Microsoft
- Sheer obscene scale allows sale of well-engineered services cheaper than we can deploy ourselves
  - With healthy profit margins too

BIOTEAM
Enabling Science

Amazon Example

- Amazon S3
  - Most expensive tier is $0.15 per GB/month
  - Design goal: 99.999999999% durability
  - Automatic replication, designed to survive simultaneous loss of two datacenters
- Amazon “Reduced Redundancy” S3
  - “*Only*” 99.99% durability at ~30% less cost
  - Still get replication
    - Designed to survive loss of one datacenter

BIOTEAM
Enabling Science

Amazon Example …

- If we are truly honest about the actual fully-loaded cost of keeping our storage online, safe & accessible than cloud storage economics start to become compelling
- The only thing holding many of us back is the speed of our internet connections
  - Amazon physical media ingest/export can solve this for some use cases

BIOTEAM
Enabling Science

My thoughts …

- Cloud storage is a good fit for archive and several other use cases
- Great way to share data among collaborators or between service provider & client
- Low-volume lab instruments will soon routinely “write directly into the cloud”

BIOTEAM
Enabling Science

My thoughts …

- Expect to see this in 2011:
  - Storage vendors building cloud storage support into native controllers & arrays
    - *It's just a REST or SOAP API call away …*
  - Likely as an inexpensive archive or backup option to primary & nearline stuff running locally
    - Or as front-end cache to primary data residing in the cloud

BIOTEAM
Enabling Science

End;

- Thanks!

- Comments/feedback welcome:
  - chris@bioteam.net

BIOTEAM
Enabling Science