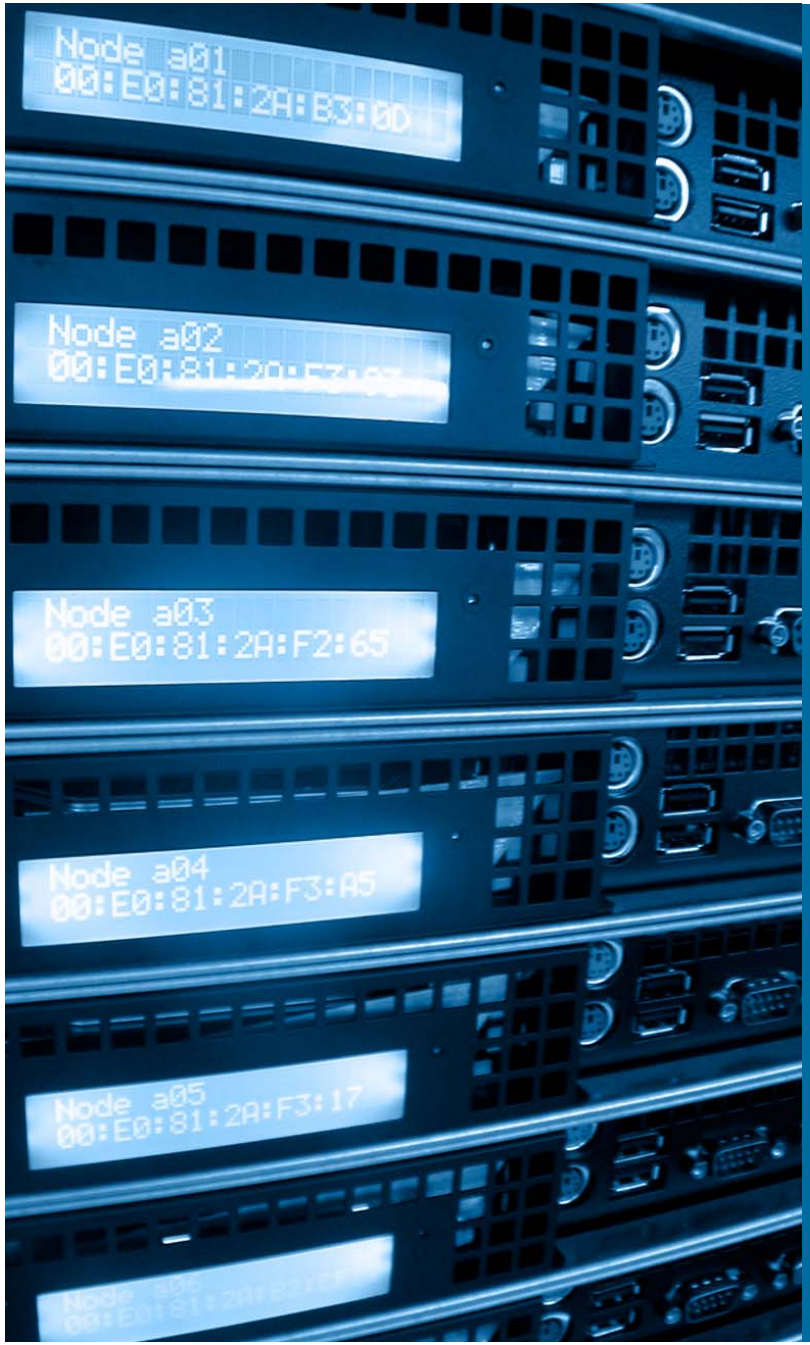
BIOTEAM
Enabling Science



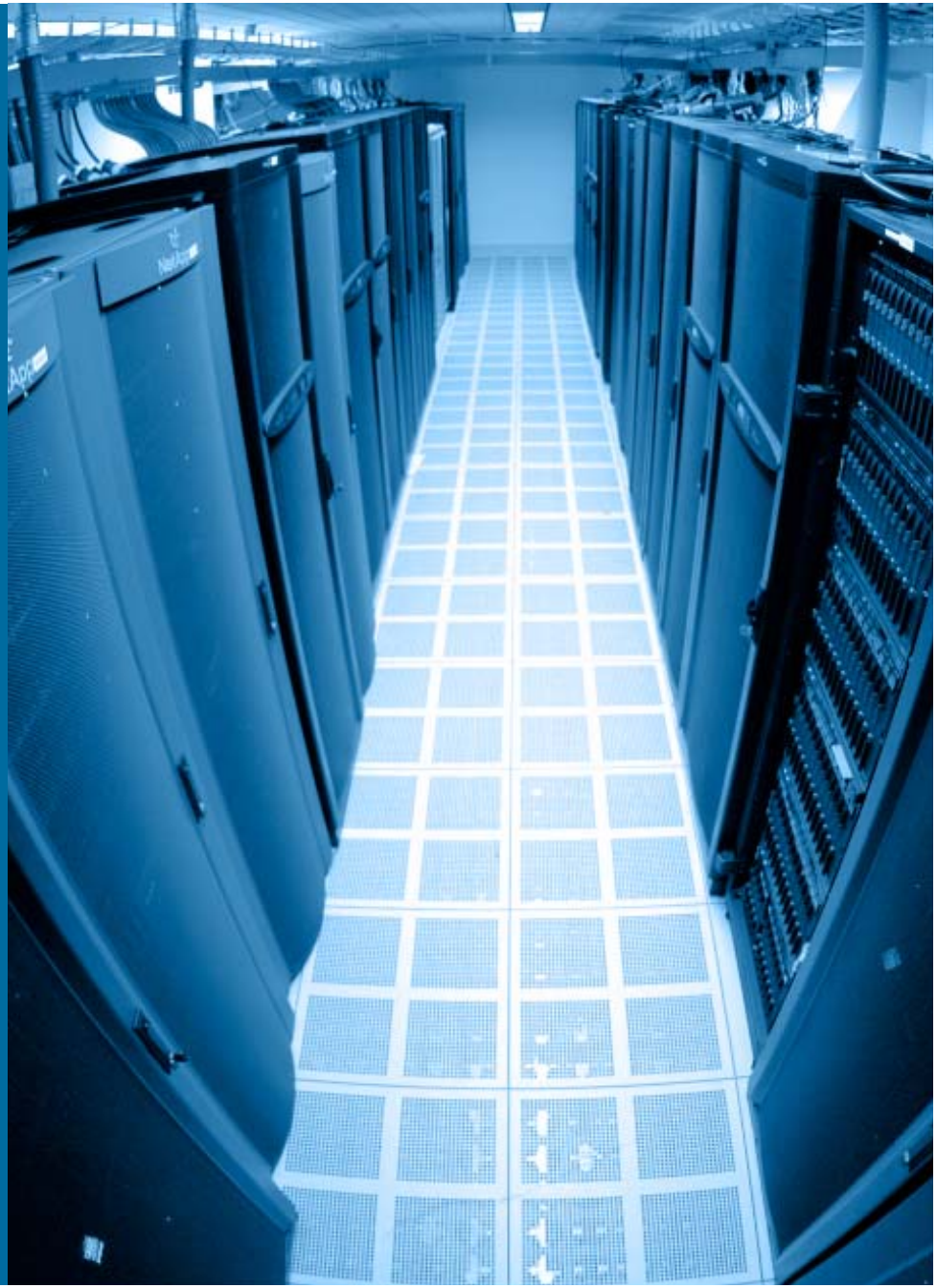
Plenary Keynote

Chris Dagdigian

Bio·IT World
CONFERENCE & EXPO '09

Kicking Things Off

- Honored to be here!
- Backstory:
  - *Consider this an evolved form of the "Trends from the Trenches" talks from prior years*
- Conference tracks look great
  - Maps well to what we've seen out in the field
- Housekeeping note:
  - Check http://blog.bioteam.net for the latest version of talk slides



BIOTEAM
Enabling Science

BioTeam Inc.

Independent Consulting Shop

- Vendor/technology agnostic

Staffed by:

- Scientists forced to learn High Performance IT
- Many years of industry & academic experience

Our specialty:

- Bridging the gap between Science & IT

How I Got Here

Recipe for current career arc:

- Find people smarter than myself
- Watch what they do
- Try to understand the '*why*'
- *Shamelessly copy them*

Partial list of Bio-IT role models:

- James Cuff, Harvard FAS
- Tim Cutts, Welcome Trust Sanger Inst.
- Matthew Trunnel, Broad Institute
- Joe Landman, Scalable Informatics
- Chris Smith, Platform Computing
- Dan Templeton, Sun Microsystems
- Stu Jackson, Rosetta/Merck
- … and many more …

cluster building in '02

BIOTEAM
Enabling Science

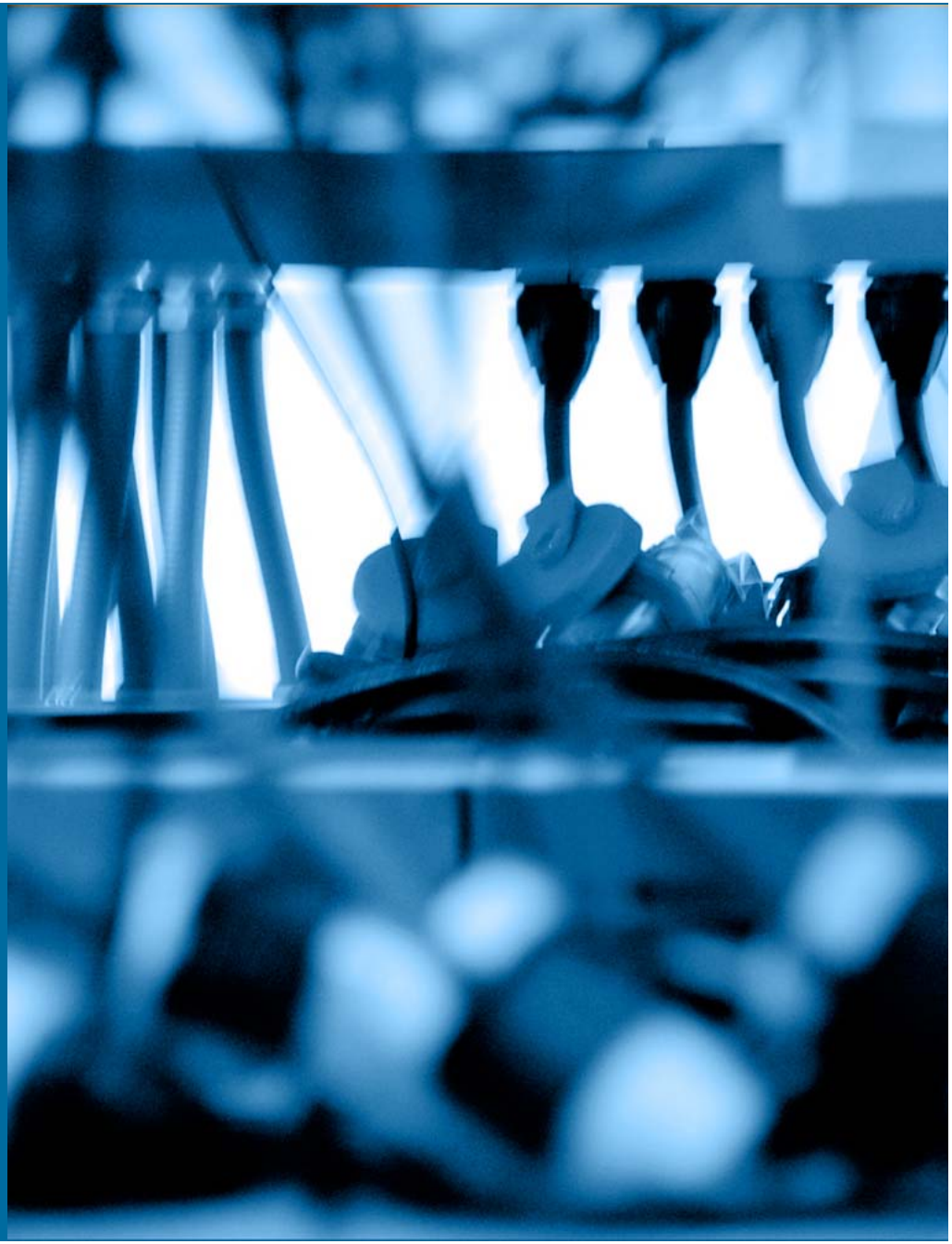
Some Reservations …

I'm uncomfortable giving this particular address:

- My role models did not get to where they are today via punditry & plenary addresses
- Most don't have the time to talk in public about their latest efforts

Trying really hard to avoid being seen as an industry figure or talking head

- Results more important than words



BIOTEAM
Enabling Science

Topics for Today

Old News

- *Once trendy, now mainstream*

Currently Exciting

- *Overhyped but still useful*

Soon To Be Exciting

- *What we are watching*

Already Mainstream

Virtualization &
Bio-IT Storage “Tsunami”

Virtualization In Research IT

Still the lowest hanging fruit in most shops

- Tremendous benefits for
  - Operators, end-users, environment & budgets
- Tipping point for me was
  - Live migration of running VMs without requiring a proprietary file system underneath

Virtualization In Research IT

Seen in 2009:

- Campus “Virtual Colocation Service”
  - Deployed when HVAC/power hit facility limits
  - Available campus-wide to all researchers & groups
  - Built with VMWare & NetApp on Sun hardware
  - Aggressive thin-provisioning & content optimization
- Truly significant payoff:
  - ~400 servers currently virtualized
  - Large # of physical servers retired & shut down
  - Storage savings from de-dup, compression & thin provisioning
  - Significant electrical & HVAC savings
  - Full delegation of administrative control to owners

Virtualization In Research IT

Another benefit for Research IT shops:

- Lets scientists design, deploy and manage apps and services that are not part of the “enterprise” portfolio
- Solves a common problem in large IT environments:
  - Scientists routinely building web apps and services to satisfy individual or workgroup level requirements
    - Often need administrative control over the web server and elevated access permissions on the base OS
  - Apps and services do not meet enterprise standards for support, security, documentation & lifecycle management

Virtualization In Research IT

Research Services & Web Applications:

- Virtualization makes the support line clear:
  - IT provides a ‘known good’ guest OS image
  - End user handles everything else
  - If end-user breaks the system, IT reverts to prior snapshot

And coming soon:

- Virtualized cluster head nodes
- Custom VM sandboxes for all!

Not coming soon:

- Grids & Clusters distributing entire VMs for task execution
  - *Technically possible but still behind the“hype vs. practicality” curve*

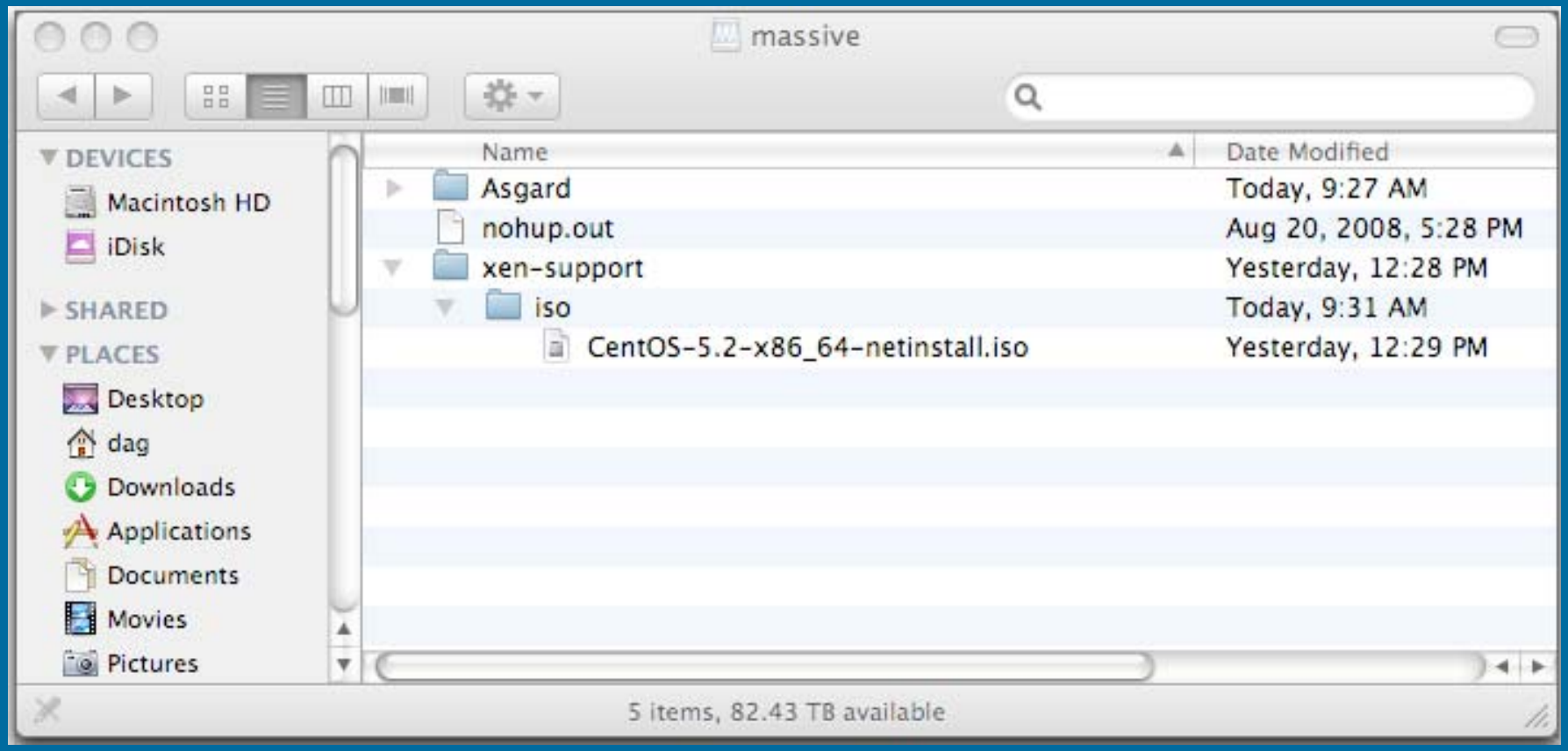
Storage

Storage Trends

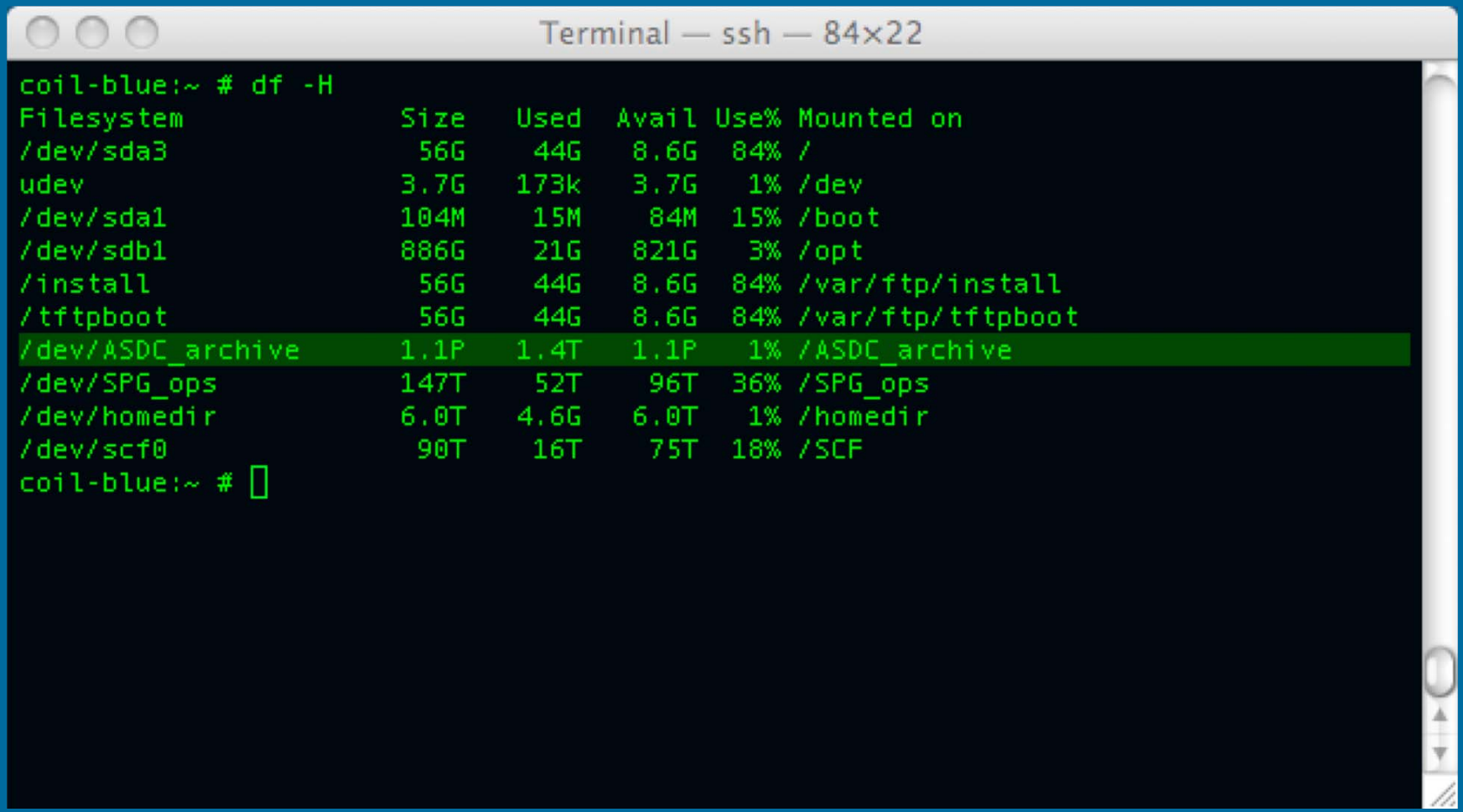
Seen in 2008

- First 100TB single-namespace project
- First Petabyte+ storage project
- 4x increase in “technical storage audit” work
- First time witnessing 10+TB catastrophic data loss
- First time witnessing job dismissals due to data loss
- Data Triage discussions are spreading well beyond cost-sensitive industry organizations

82TB Folder. Still satisfying.

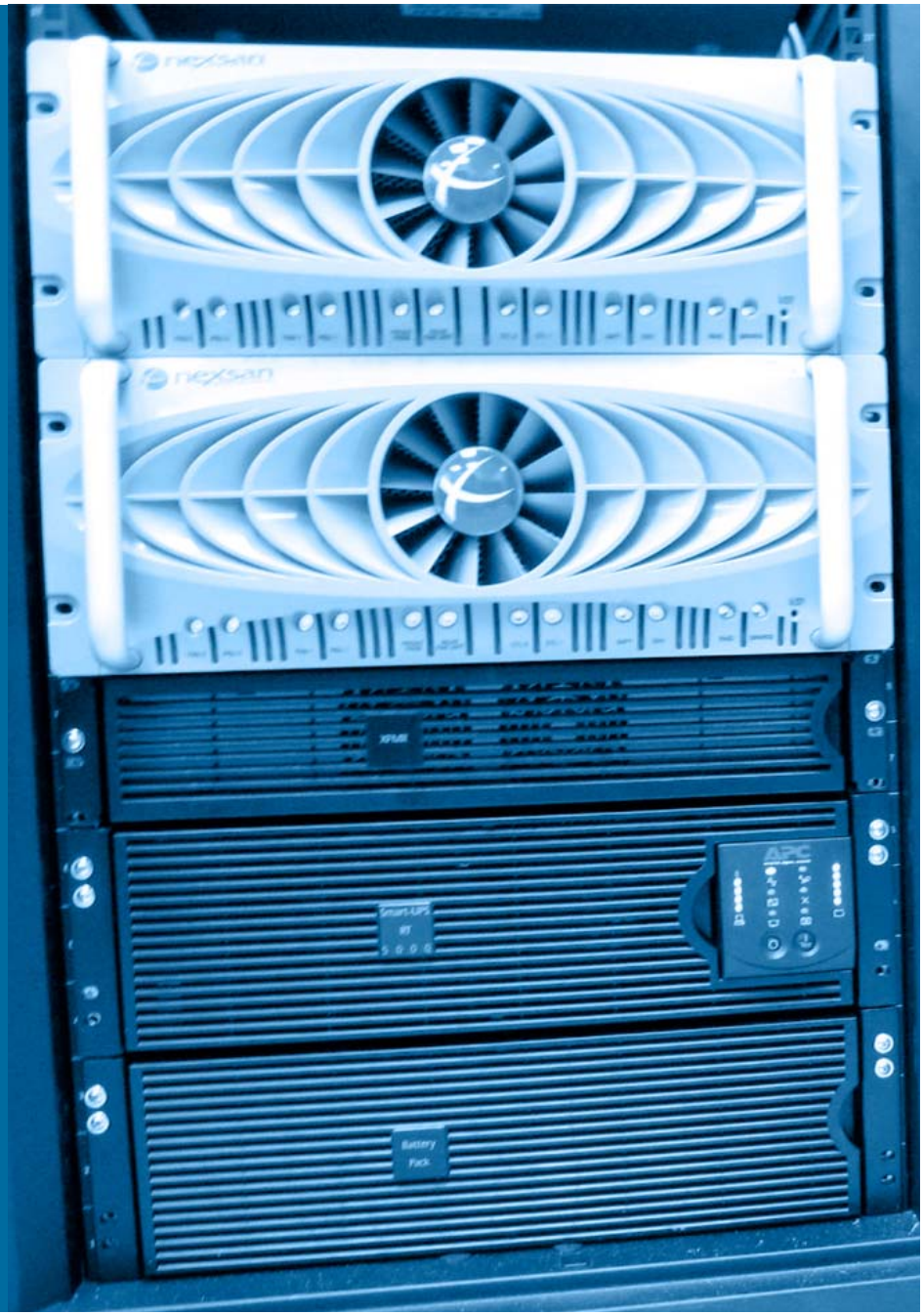


1PB volume - Even more satisfying



General Observations

- Storage is still “cheap” & getting cheaper
- Operational costs remaining the same
  - *But there is hope …*
- Backup & data continuity costs are exploding
  - *Still in awe of the backup experts who are still staying afloat in the age of 1TB SATA disks …*

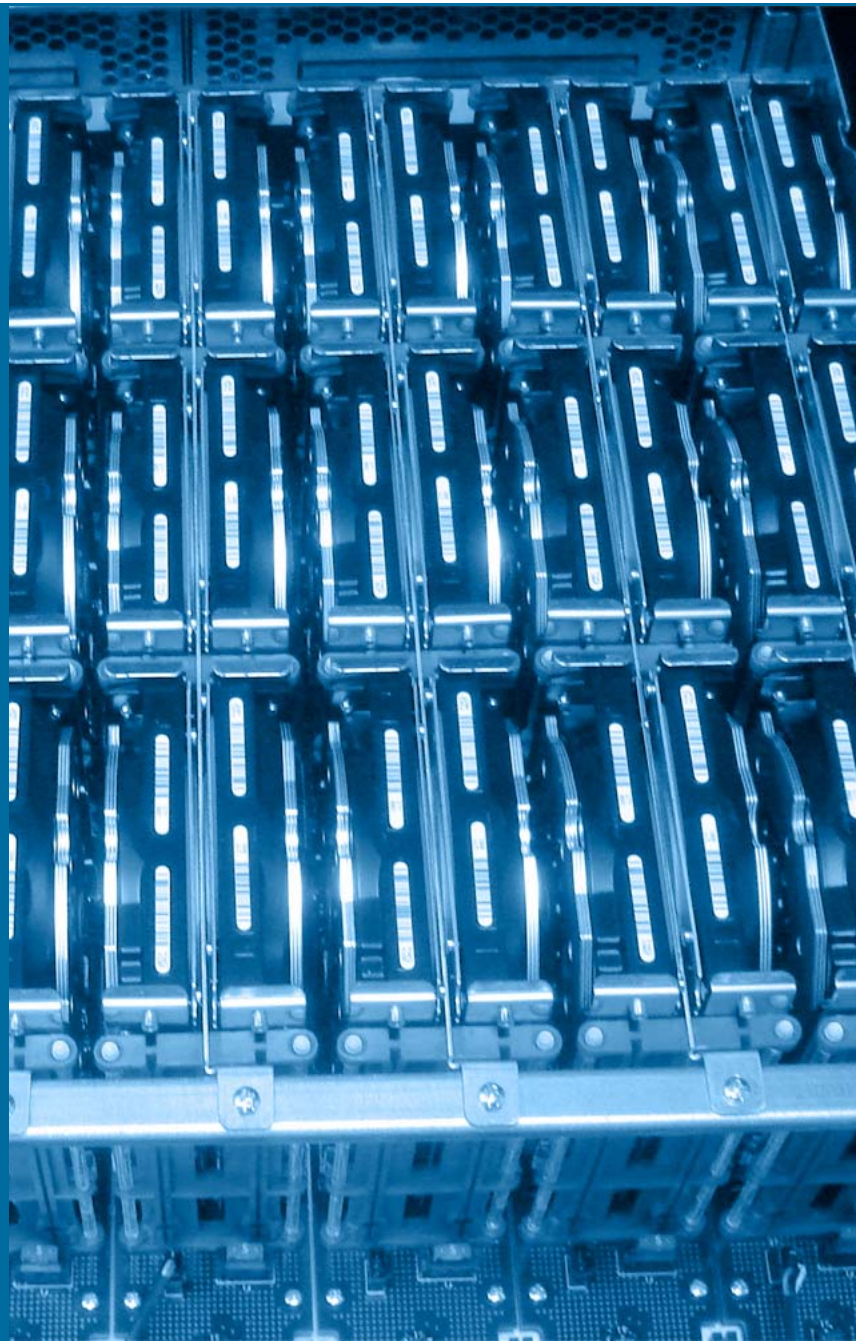


Observations cont.

- End users still have no clue about the true costs of keeping data accessible & available
  - *"I can get a terabyte from Costco for $220!" (Aug 08)*
  - *"I can get a terabyte from Costco for $160!" (Oct 08)*
  - *"I can get a terabyte from Costco for $124!" (April 09)*
- IT needs to be involved in setting expectations and educating on true cost of keeping data online & accessible
- Organizations need forward looking research storage roadmaps

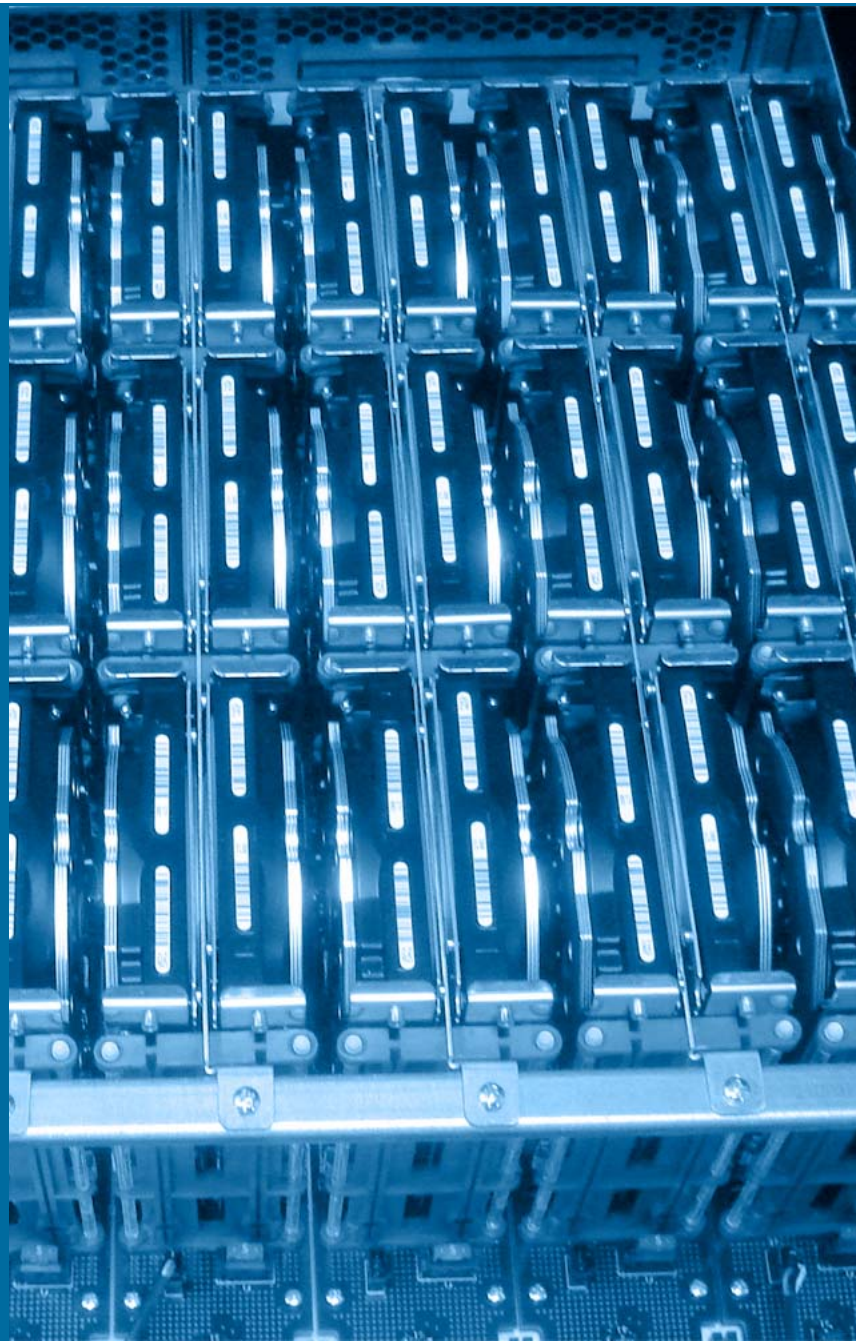
Observations cont.

- The rise of “terabyte instruments” is already having a major disruptive influence on existing environments
  - We see individual labs deploying 100TB+ systems
- I was wrong when I said
  - *“petabyte scale storage needs will appear within the decade …”*
  - That time is now for *some* large organizations



Homework Exercise

- Select three vendors
- Build quotes for 100TB single-namespace NAS solution
- My results:
  - *$160K to $1.5M range*
- *Repeat every six months*



Capacity Dilemma: Data Triage

Data Triage

- The days of unlimited storage for research are over
- Rate of consumption increasing unsustainably
- First saw triage acts in 2007 (industry client)
- Becoming acceptable practice in 2008
- Absolutely a given in 2009 for most projects we see

Why delete data?

- Given full lifecycle cost of data, sometimes repeating the experiment is cheaper than storing the results forever

Data Loss: Lessons learned

The 2008 event:

- Metadata for clustered SAN file system irrevocably corrupted
- 10+ TB of scientific data lost forever
- 3 people “fired”

Vastly simplified root cause:

- Double-disk failure on RAID5 volume holding SAN FS metadata
- Significant operational errors

Data Loss: Lessons learned

Observing this event taught us some lessons:

- We no longer use RAID5 on large filesystems
- Everything on RAID6 or other double-parity system
- Mandatory use of SNMP & email reporting
- Aggressive disk handling: Replace drive on warning
- Background scrub & verify operations
- We reject storage/controller products that are not proactive about disk scrubbing and consistency checking

Thoughts for ‘09 and beyond

Double-parity or bust!

- RAID5 declines in use as drive counts increase

Data Triage is here to stay; acceptance less of an issue

Exponential capacity growth hopefully slowing …

- Better “in-instrument” data reduction
- Customer trust in vendor & instrument increases
- Explicit data triage within organizations
- “Only” 100-200 TB/year new data generation within the biggest organization we’ve visited this year*

Software RAID on hosts & arrays gaining popularity

Storage devices running more 3rd party software

- Example: Ocarina reader running inside Isilon nodes
- Example: Ocarina Optimizer on BlueArc chassis

Thoughts for ‘09 and beyond (cont.)

Backup

- 2008: Became something of a sick joke
- Storage products leave backup products in the dust
  Almost too far ahead to even attempt to keep up

Starting to see in 2009:

- Complete re-think of backup paradigm
- New expectations, new procedures in an age where “nightly full” may never happen again
- Observed recently:
  50TB of Isilon storage for HPC cluster
  Users told it is scratch space; not backed up at all

Thoughts for ‘09 and beyond (cont.)

Cloud Storage

- James Hamilton has blogged some interesting numbers:
- *Cold storage geographically replicated 4x can be achieved at scale for $.80 GB/year (and falling)*
- Can you match that in your datacenter?
- Some potentially interesting use cases for cloud stores

More on this later …

Thoughts for ‘09 and beyond (cont.)

- IT organizations may need to relinquish some control over research data administration
- IT groups are should not be the sole decision maker …
  - For data triage efforts
  - For aggressive storage optimization efforts
- Science staff must be active in these decisions

Final thoughts on storage for 2009

- Yes the problem is real
  - More and more “terabyte instruments” are coming to market
  - Many of us have peta-scale storage issues today
- “Data Deluge” & “Tsunami” are apt terms
- But:
  - The problem does not feel as scary as it once did
  - Chemistry, reagent cost & human factors are natural bottlenecks
  - Data Triage is an accepted practice, no longer heresy
  - Data-reduction starting to happen within instruments
  - Customers starting to trust instrument vendor software more
  - We see large & small labs dealing successfully with these issues
  - Many ways to tackle IT requirements

    Mix and match solutions to fit local need …

BIOTEAM
Enabling Science

Becoming Mainstream

Green IT for cynics &
Utility Computing

Green IT for Cynics

What we have seen in 2008 & 2009:

- Production deployment of storage arrays with AutoMAID
- Virtual “colocation facilities”
- High density datacenter redesign efforts
  - Blade deployments driving other facility efficiency wins
- Auto-scaling HPC cluster in big Pharma
  - Multi-thousand core Linux HPC cluster:
    - Server power state(s) harmonized with Platform LSF queue(s)
    - Unattended auto-sizing of cluster to match scientific workflow
    - Real-time savings metrics kept & tracked
  - *First time in enterprise environment, old news in some academic and research settings*

BIOTEAM
Enabling Science

Green IT for Cynics

Hyped beyond all reasonable measure

- Reminds me of dotcom-era WAN-scale grid computing PR
- Promising more than can be realistically delivered

Marketing aside, still worth pursuing:

- Shrinking physical footprint, reducing power consumption and better management of HVAC is better for the planet
- … and has tangible results on the bottom line

BIOTEAM
Enabling Science

Green IT for Cynics

Where Green IT matters to me:

- Putting more capability into a smaller space
- Putting capability into spaces that previously could not support it
  - Telco closets, underneath wetlab benches, etc.
- Working within power or HVAC-starved machine rooms
- Reducing power draw & increasing power use efficiency
- Reducing cooling costs & increasing air handling efficiency

Green IT for Cynics

My 2008 “aha!” moment:

- NexSan SATABeast Storage Arrays
  - 48x 1TB SATA disks in 4U enclosure w/ FC interconnects
  - “AutoMAID” options built into management console
- What we had:
  - Three 48 terabyte SATABeasts in two racks
  - APC PDUs with per-outlet monitoring features
- What we saw:
  - 30% reduction in power draw
  - No appreciable impact on cluster throughput

Green IT for Cynics

Best 2009 moment:

- Linux HPC cluster on west coast
  - IBM BladeCenter based w/ thousands of CPU cores
  - Chilled water to the racks
  - Thousands of cores managed 24x7 by ~2 FTEs
  - Management interface tuned for VT100 text display
    - Remote cluster admin is *reasonable* with a phone-based SSH client
- Internally developed tools
  - Scripts pull state & event log data from enclosures
  - All provisioning 100% automated
  - Probe scripts catch majority of system errors before they affect cluster jobs or are noticed by end-users
  - Deep integration with Platform LSF queues
    - Automatic ‘bclose’ when problems detected
    - Impressively “science-aware”

Green IT for Cynics

Best 2009 moment:

- HPC cluster continued
  - Automatic workflow-aware server power-on and power-off
  - Email status alerts to management:
    - `"Hello, I've saved $80,000 in facility costs this year!"`

Green IT for Cynics

- Recommendation:
  - Use Green IT for political cover on facility efficiency efforts
- Example:
  - You: "*I want to implement automatic server power on|off*"
  - Network Ops: *"Our monitoring infrastructure is based on ping testing of IP addresses, you will flood our system with false downtime alerts if you shut down functional systems!"*
  - You: "*Green IT.*"
  - Boss: "*Make it happen.*"

Utility Computing

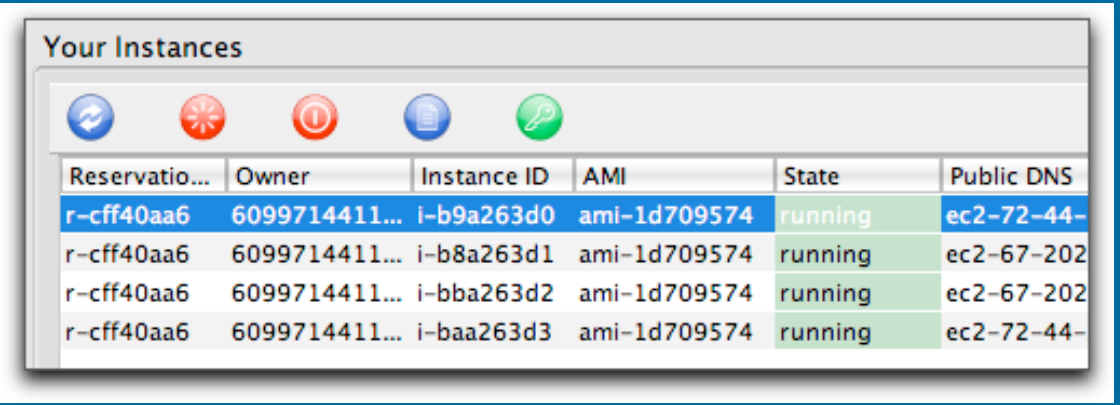
In 2009 I will try hard never to use the word “cloud” in any serious technical conversation …

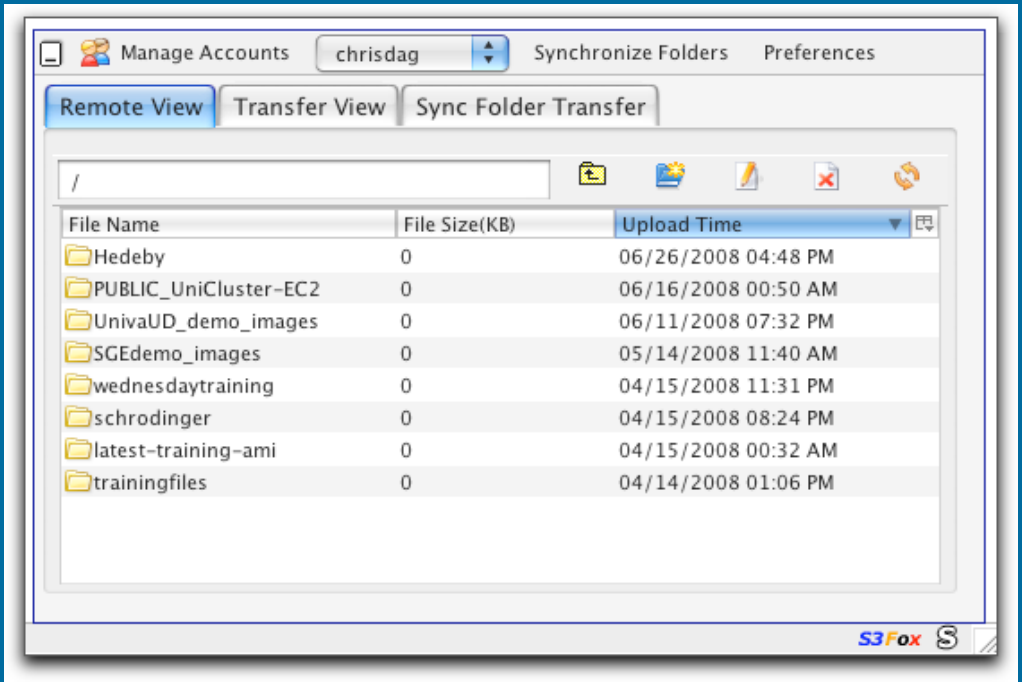
Utility Computing: Setting the Stage

- Speaking of utility computing as it resonates with *infrastructure* people
- My building blocks are servers or groups of systems, not software stacks, developer APIs or commercial products
  - Goal: Replicate, duplicate or relocate complex systems

Lets Be Honest

- This stuff is not rocket science
- Fast becoming accepted and mainstream
- Easy to understand the pros and cons





Utility Computing: Economics

- It is expensive to design for peak demand in-house
- Pay-as-you-go can be compelling for some workflows

Consider: 100 CPU hour research problem

- EC2: 10 large servers @ .40/hr for 10 hours
  - Work done in 10 hours at cost of $40 USD
- EC2: 100 large servers @ .40/hr for 1 hour
  - Work done in 1 hour at a cost of $40 USD

Amazon Web Services - Vocabulary Primer

EC2 - “Elastic Compute Cloud”
S3 - “Simple Storage Service”
EBS - “Elastic Block Store”
SQS - “Simple Queue Service”

While I'm being honest …

My own $.02:

- Today Amazon AWS ***is*** the cloud
  - Simple, practical, understandable and usable today by just about anyone
  - Rollout of features and capabilities continues to be impressive
    - Latest example: Elastic MapReduce
      - » Run Hadoop jobs from your browser
      - » Available now: CloudBurst short read mapping
  - VARs are springing up (Cycle Computing, RightScale, etc.)
- Competitors are pretty far behind
  - … and tend to believe too much of their own marketing materials

BioTeam & Amazon AWS

Why I drank the EC2 Kool-Aid

- Saw it, used it, solved actual customer problems with it

Timeline

- Late 2007
  - Initial experimentation & test cases
- Early 2008
  - By March, every single BioTeam consultant had independently used EC2 to solve a customer facing problem
- Late 2008
  - Commercial and OSS application EC2 integration requests are coming in almost weekly
- 2009
  - Industry pressuring more and more ISVs to support EC2 model
  - Real workflows moving into AWS

Recent Example: Docking on EC2

BioTeam & Pfizer, using Rosetta++ software (*Baker Lab, UWash*)

- Implemented by Adam Kraut
- Addressed in detail today at the HPC pre-conference workshop

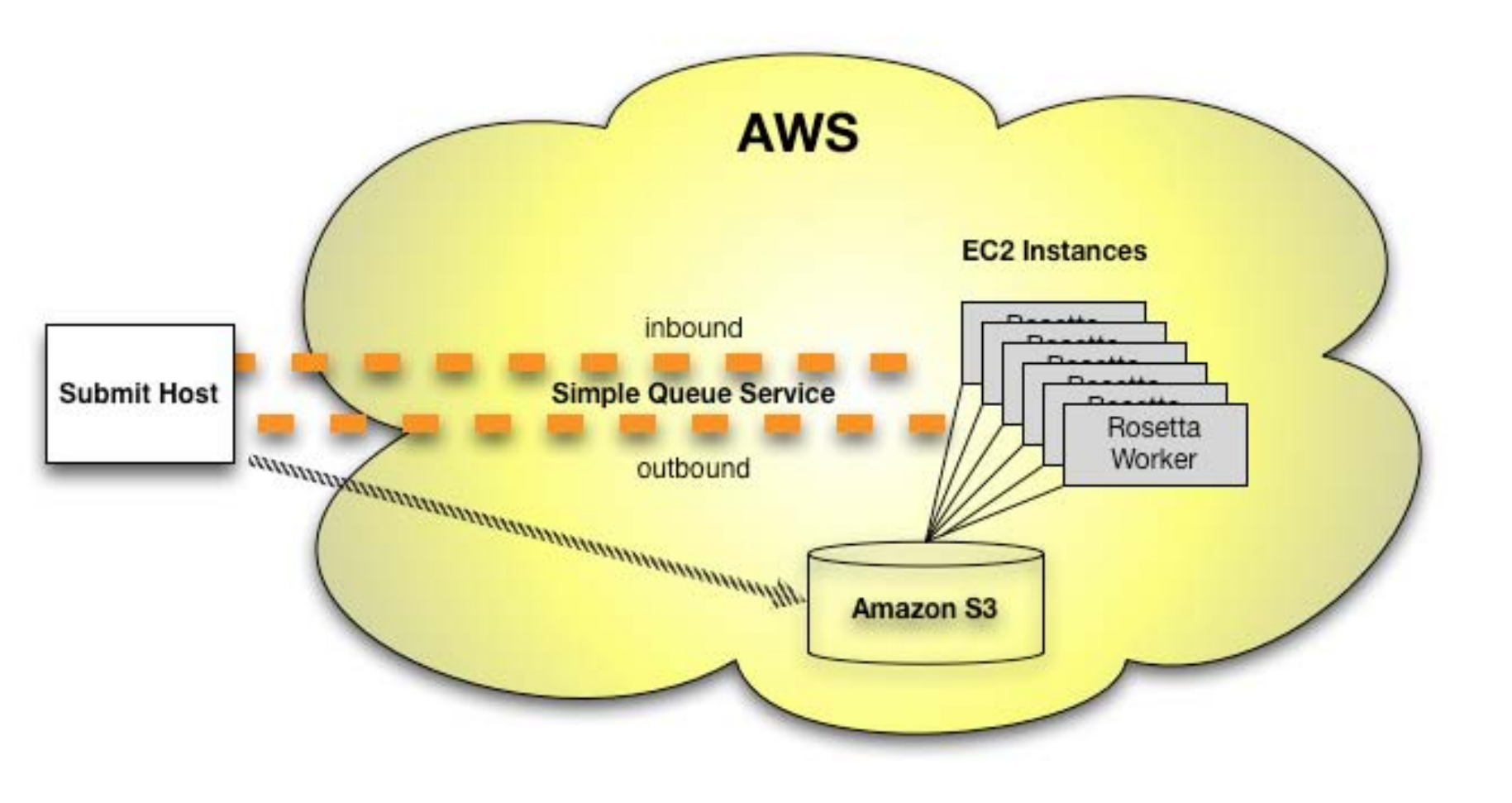
Chosen Workflows:

- Global docking of Antibody-Antigen complex by RosettaAntibody and ZDOCK
  - ~1 hour CPU time per target
- Full-atom refinement of predicted complexes by Rosetta++
  - ~1000 hours CPU time per target

EC2 Project Architecture: Docking

- Borrows heavily from Rightscale Grid Edition
- Inbound, outbound Amazon SQS queues
- Job specification in JSON format
- Data for each work unit pushed into Amazon S3 buckets
- Workers pull from S3, push back when finished
- Job provenance/metadata stored in Amazon SimpleDB
- Independent work units allow dynamic allocation of Worker instances

EC2 Architecture: Docking



EC2 Project Architecture continued: Docking

- Structures are scored in EC2
- ALL results are in S3
  - EC2 <-> S3 bandwidth is free
- Download the top scoring docked complexes
- Launch post-processing EC2 instances to score, rank, filter, and cluster the results in S3
  - “Bringing the compute to the data”

Utility Computing Pet Peeve

My take:

- Amazon, Google & Microsoft probably have better internal operating controls than you do
- All of them are happy to talk as deeply as you like about all issues relating to security
- Do your own due diligence & don’t let politics or IT empire issues cloud decision making

Utility Computing Pet Peeve

Don’t want to belittle security concerns:

- But whiff of hypocrisy is in the air …
  - Is your staff *really* concerned or just protecting turf?
  - It is funny to see people demanding security measures that they don’t practice internally across their own infrastructure

State of Amazon AWS

New features are being rolled out fast and furious

- Almost monthly these days

But …

- Still no solution for the data ingestion problem
  - How to push 1TB/day into Amazon S3 or EBS?
- EC2 nodes still poor on disk IO operations
- EBS service can use some enhancements
  - Many readers, one-writer on EBS volumes would be fantastic
- Poor support for latency-sensitive things and workflows that prefer tight network topologies

This matters because:

- Compute power is easy to acquire
- Life science tends to be IO bound
- Life science is currently being buried in data

If ingest problem can be solved …

- I think there may be petabytes of life science data that would flock to utility storage services
  - Public and private data stores
  - Mass amount of grant funded study data
  - Archive store, HSM target and DR store
- Why?

  Providers operating at such extreme scale can offer efficiencies unmatchable by traditional IT groups
  - *Remember Hamilton's figure: 4x geographically replicated storage for $.80 GB/year and falling …*
  - Amazon already has a "requestor-pays" download model

If ingest problem can be solved …

Next-gen sequencing example:

- Instruments in wet-lab
  - Connected to reasonable compute and storage resources
  - Possibly small enough to be lab-local
- Data analysis & reduction done onsite
- Derived data (basecalls, etc.) kept onsite
- Run data *ingested* into utility storage service
  - Cheap, scalable long-term storage of scientific data
  - Probably a one-way trip

Next-Gen & Utility storage

What this would mean:

- Primary analysis onsite; data moved into remote utility storage service after passing QC tests
- Data would rarely (if ever) move back
- Need to reprocess or rerun?
  - Spin up “cloud” servers and re-analyze in situ
  - Terabyte data transit not required

Summary:

- Lifesci data; 1-way transit into the cloud
- Archive store or public/private repository
- Any re-study or reanalysis primarily done in situ
- Downside: replicating pipelines & workflows remotely

Cloud Sobriety

McKinsey presentation "Clearing the Air on Cloud Computing" is a must-read

- Tries to deflate the hype a bit
- James Hamilton has a nice reaction:
  - http://perspectives.mvdirona.com/

Both conclude:

- IT staff needs to understand "the cloud"
- Critical to quantify your own internal costs
- Do your own due diligence

BIOTEAM
Enabling Science

Worth Watching

Scale-out best practices &
Federated storage

“Trickle-Down” Scale-out Tips

- Google, Microsoft & Amazon all operate at extreme scales that few of us come close to matching
  - Datacenter scaling & efficiency measures taken by these companies are tightly held trade secrets for competitive reasons
- This is starting to change and ***we will all benefit***
  - Both from “trickle-down” best practices & hard data
  - And with vendor products improved to meet these demanding customers

“Trickle-Down” Scale-out Tips

April 2009:

- Google Datacenter Efficiency Summit
  - Presentations now online
- Google video on Youtube showing ‘04-era 780W/sq ft containerized datacenter
- If Google was doing this stuff in 2004, what are they up to now?

And from server vendors:

- Dell warranting servers at 95F input temperature
- Rackable Systems warranting it’s C2 rack at 104F inlet temperature

“Trickle-Down” Scale-out Tips

James Hamilton’s blog is required reading:

- http://perspectives.mvdirona.com/
- Former MS employee, now Amazon VP & Distinguished Engineer on AWS team
- Focused on infrastructure efficiency, reliability, and scaling
- Well written, really useful information backed by real data

Moving forward

- Realize significant benefit from best practices trickling out of these sites
- Also benefit from vendor products engineered to satisfy the extreme scale folks

And finally …

Ongoing research in 2009

Federated Storage Experiments in 2009

- Storage “islands” have always been an issue in our field
  - Made worse by lab-local large storage and (*future*) cloud storage
  - Bad for scientists (lots of rsync & wasted productivity)
  - Bad for IT (duplicated content multiplies backup & operational hassles)

Federated Storage Experiments in 2009

- Interesting ongoing efforts regarding global namespaces and federated data stores
- Got into this via Jacob Farmer
- Using technology from General Atomics
  - “Nirvana” Storage Resource Broker
  - Has some interesting capabilities for life science, especially in collaborative environments
- Can’t speak much yet; pilots ongoing
- Look for more talks/papers on this in 2009 …

Wrapping things up

Once hyped, now mainstream:

- *Virtualization in research IT*
- *Storage Deluge Problems*

Currently hyped, but worthwhile:

- *Green IT*
- *Utility Computing*

Future Focus:

- *"Trickle-down" Best Practices*
- *Federated Storage*

End;

- Thanks!
- BioTeam people will be at Booth #113
- Presentation slides will appear here:
  - http://blog.bioteam.net
- Comments/feedback:
  - "chris@bioteam.net"