Colorado State University

Low-cost BYO Mass Storage Project

James Cizek

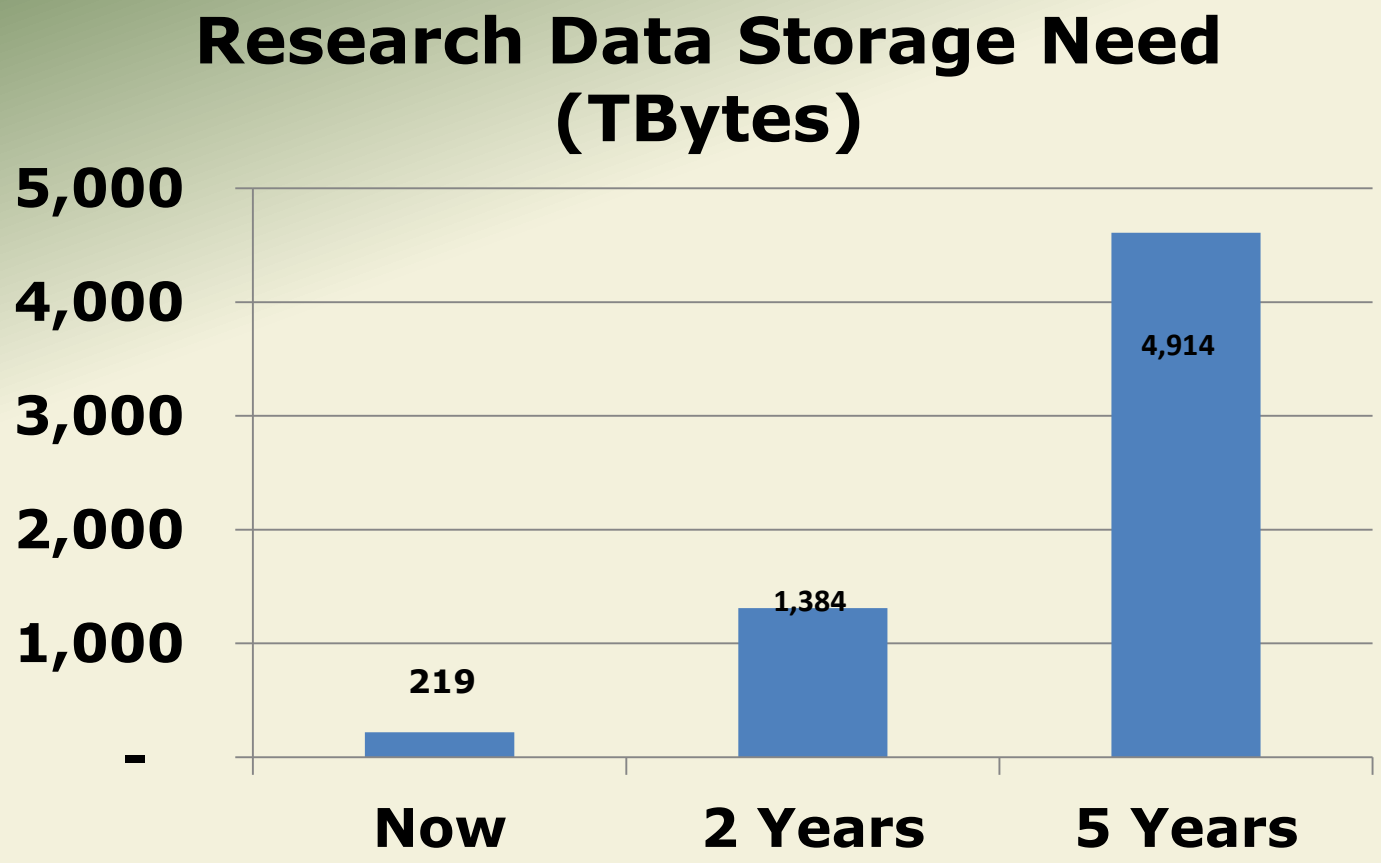
Unix Systems Manager

Academic Computing and Networking Services

The Problem

- Reduced Budget
- Storage needs growing
- Storage needs changing (Tiered Storage)
- Long term (Archiving) storage needs growing

Projected Needs (2009 Survey)



The Goal

- Find a mass storage solution that won't break the bank
- Vendors sell high-speed, costly systems (suitable for Amazon, Google, etc.), but we want slower, low-cost
  - Looking at vendor offerings, we decided to "roll our own"
- Maximize TB/$$ with reasonable assurance that data are redundant and safe

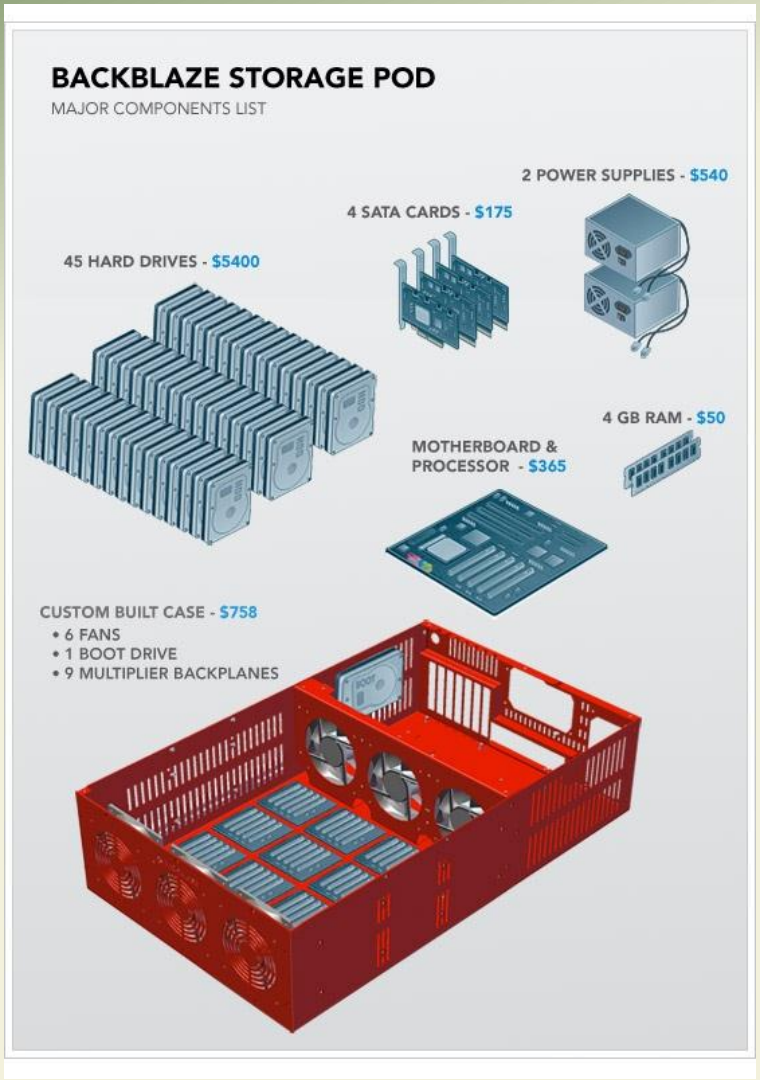
Some Understandings

- Approached this project as “Secondary” or “Tier 2” type storage, not intended to replace extremely fast, ultra-reliable, expensive disk systems
- Realized that device management, support, and component failure need to be addressed

A starting point

- Online backup company "Backblaze" open-sourced their storage pod design, see https://www.backblaze.com/petabytes-on-a-budget-how-to-build-cheap-cloud-storage.html
- Thought that starting with a proven design would eliminate many unknowns and speed up our design process
- Turned out to be helpful, but ran into many of our own headaches

The BackBlaze design



BackBlaze vs. CSU design goals

- Realized that the BackBlaze design didn't exactly meet our requirements
- No redundant power supplies
- Cheap SATA cards didn't take advantage of performance available by having large number of spinning hard drives
- Case too small to accommodate server-class motherboard
- Single "system" hard drive is single point of failure.
- Realized the need to over-engineer cooling and vibration reduction (2 major contributors to drive failure)
- Chassis was red instead of CSU green!

CSU design changes

- Lengthened case by 3 inches to accommodate dual CPU server-class motherboard
- Added more RAM for file system buffering (6 GB compared to BackBlaze 4GB)
- Added larger, redundant power supplies - individual supply can run entire case
- Used "Enterprise" grade drives instead of consumer grade, after much research
  - Drives selected have vibration sense / damping
- Replaced cheap SATA cards with high-performance PCI-e cards

CSU chassis nearing completion

CSU chassis nearing completion

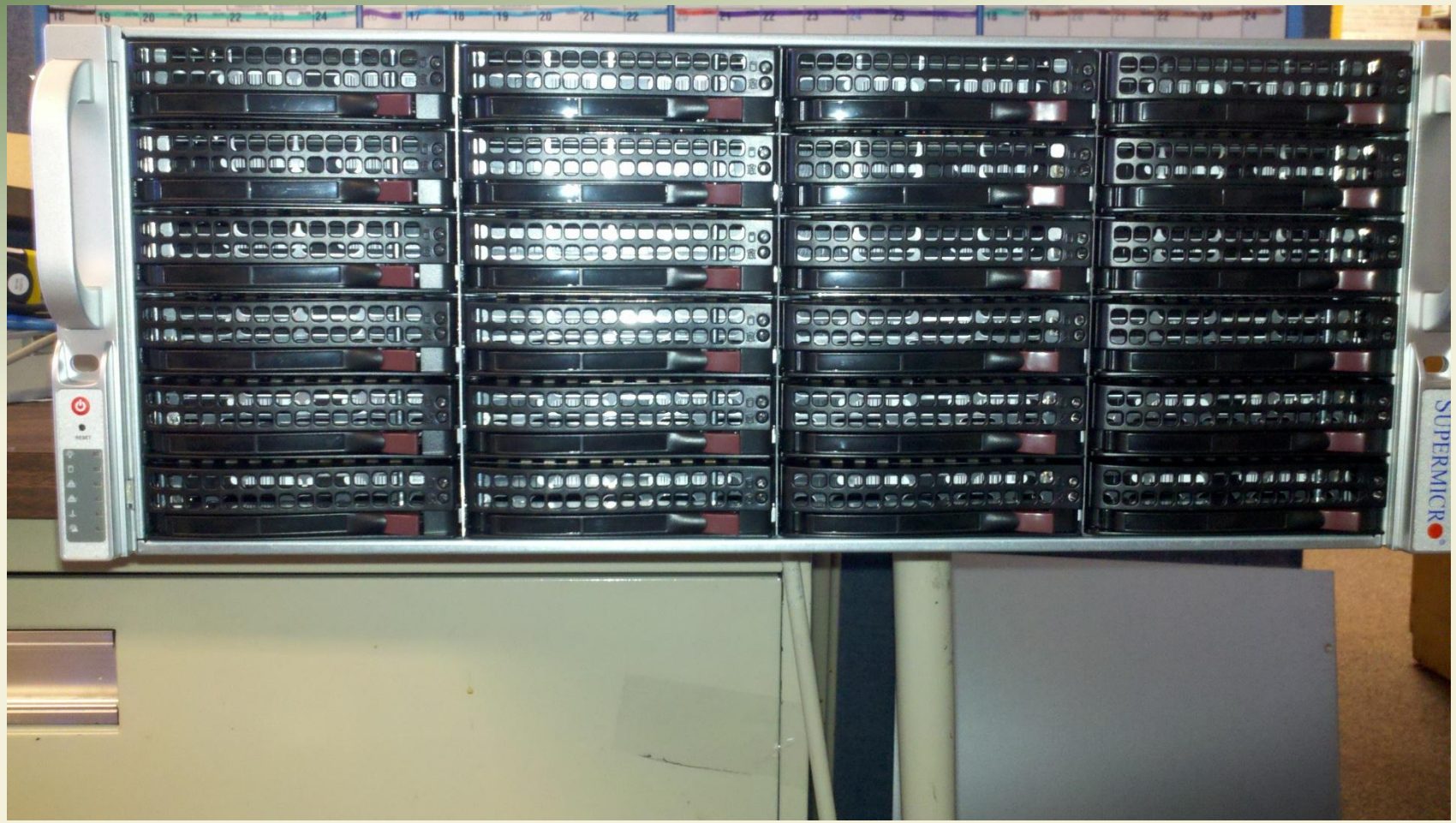
DIY storage- V2.0

- As time progressed, CSU found a commercial offering that nearly replicates our "Green Box"
- Offered in 36 drive "smart" and 45 drive "dumb" models
- Price point equivalent to in-house unit
- New technology (SAS expander vs. SATA multiplier) offers twice the speed and far more flexibility (More OS choices, drive choices)

DYI storage- V2.0 (cont)

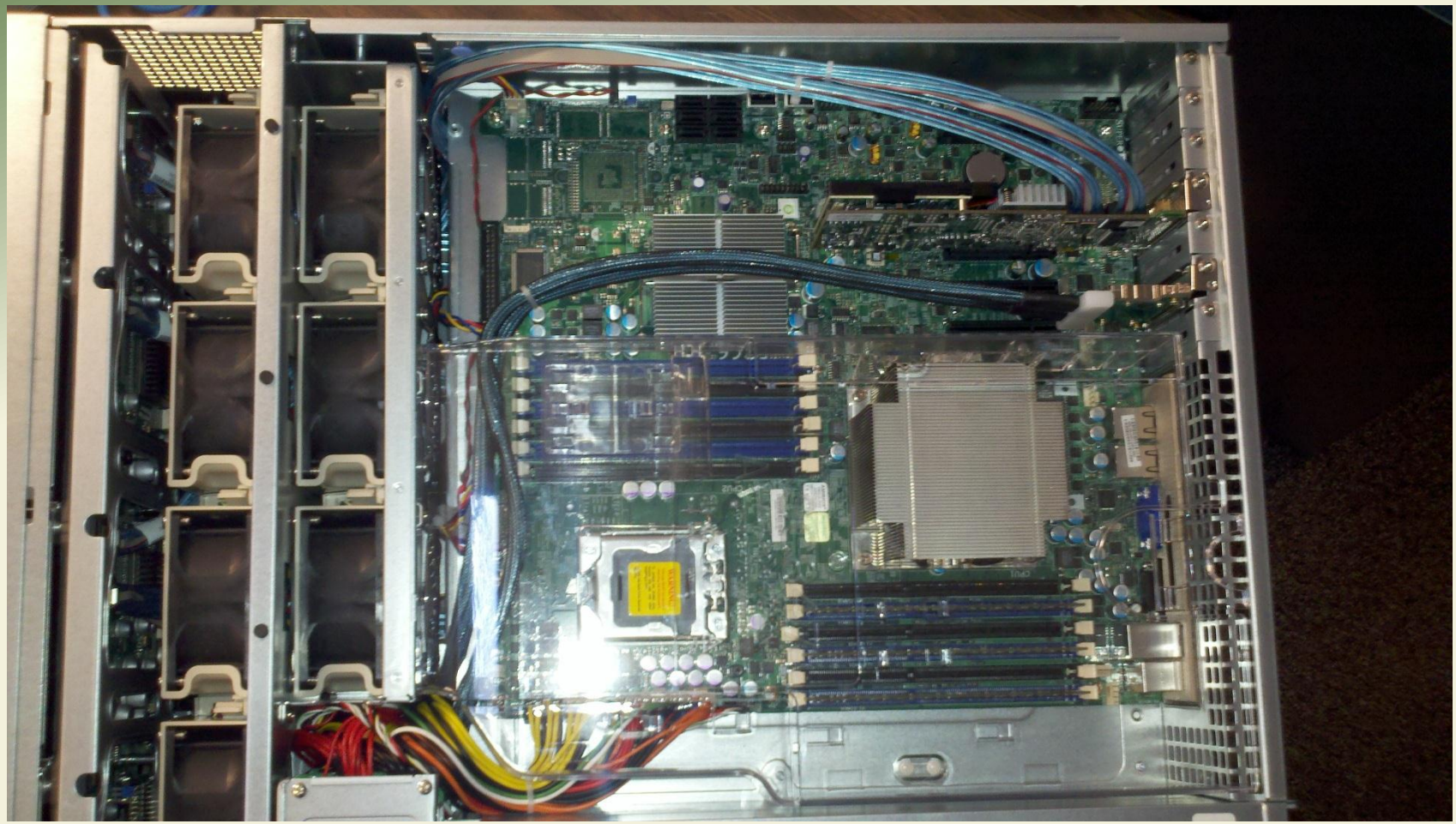
- Although dampens the "DIY" spirit of the project, the new chassis offers more reliability, expansion and speed than "Green box"
- Still functionally identical, but can be built from "opening the box" to "installed in rack" in 2 hours!
- Allows multiple hosts to access simultaneously (i.e. Windows, Unix)

CSU – Commercial Chassis



CSU Commercial Chassis

CSU Commercial Chassis



Costs (Greenbox V1.0)

- Case: $700
- 1 TB Drives: $99 x 45 ($4,455)
- Motherboard / Processors / Memory: $900
- Power Supplies: $200
- SATA cards: $300
- Ethernet card with iSCSI offload: $350
- SATA Multipliers: $45 x 9 ($405)
- Fans/Cables/Hardware/DVD/Mounts/etc.: $1,000

Total: 45 Raw TB for $8,310!

Costs (Silverbox V2.0)

- Case: $3600 (includes MB, memory, etc..)
- 2 TB Drives: $189 x 36 ($6,804)

- Total: 72 Raw TB for $10,404!

Initial Performance

- ✓ Indicates that these units will be more than adequate as secondary and long term storage
- ✓ May be fast enough to replace expensive primary storage in the future (Lot of spindles)
- ✓ Has proven expandability is easy and cost effective

Cost / Performance Comparison

- As primary storage: Current models ~$25K for 1 Terabyte
- As secondary storage: Current models ~$8K for 16 Terabytes
- Our system: 72 Terabytes for ~$10K
  - In initial performance testing, our system beats the best published performance numbers of both our primary and secondary storage systems

Challenges ahead

- Support management (What happens when a disk fails?)
- Backup and protection of stored data
  - Mirroring units
  - Avoid backing up to enterprise backup system
- Data storage and protection policies
- Parallel file system

Where will this be useful?

- Library digital repository
- Research computing
- HPC, tier 2
- Campus wide "Cloud" storage
- Second or Third Tier storage for your Enterprise backups
- Email/File archiving
- Database "snapshots" kept for long term (LMS)

In Summary

- At $139 / Terabyte, this solution provides mass storage cheaper than anything found to date
- Low price allows building 2 and mirroring (Much less expensive than tape backup!)
- Space, power, and cooling savings are substantial over other offerings
- Provides a simple solution to allow *all* research data to be kept from a project instead of discarding portions
- Flexible enough to fit many applications where large data storage is a necessity
- Reusable! After project completion, can be reconfigured to fit needs “on the fly”

Questions?

- james@colostate.edu