Teaching Storage Systems with Interactive Courselets on Chameleon

Chameleon Cloud Webinar

Erez Zadok and Tyler Estro
Stony Brook University
File Systems and Storage Lab (FSL)

FOUNT
Scaffolded, Hands-On Learning for a Data-Centric Future

Stony Brook University

Overview

- Challenges in storage education
- Course design and infrastructure
- Courselet development and publication
- Lessons learned

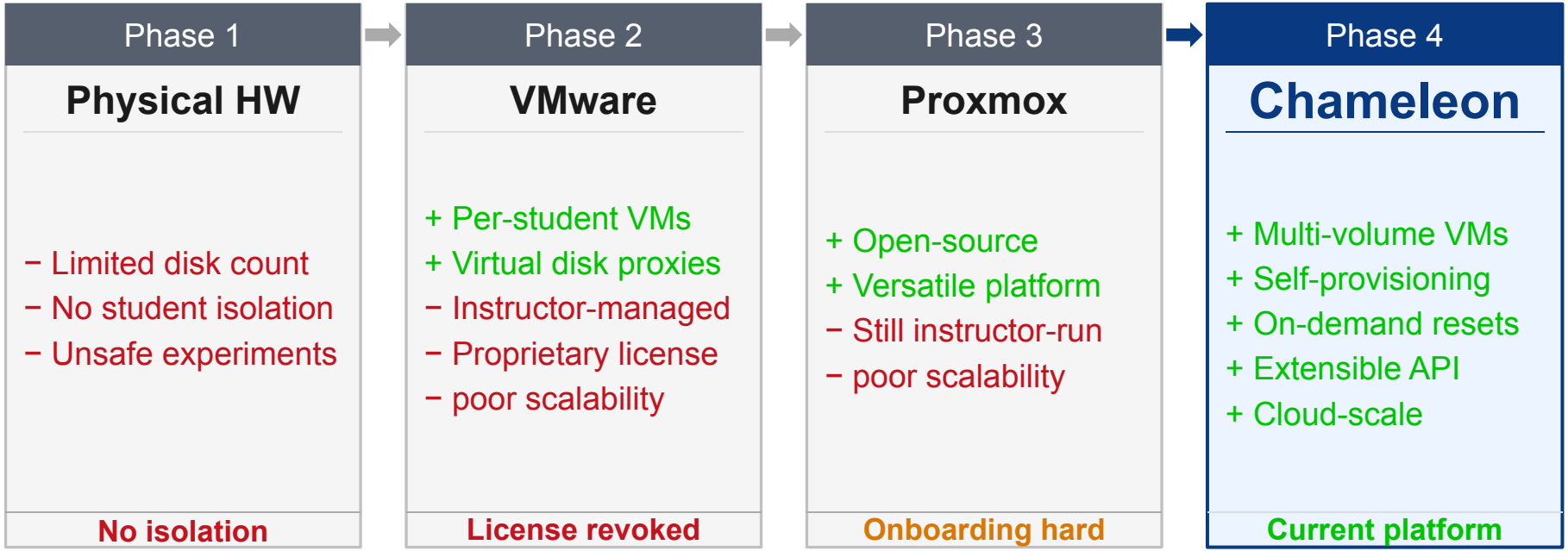
Storage Is Hard to Teach

- Storage experiments are difficult to deploy at classroom scale
  - Stateful and slow
  - Sometimes destructive
  - Require per-student isolation
- Multi-device setups and resettable environments are hard to provide
- Physical machines too slow/limited to use
- Storage labs need more than a single default VM

Course Design

- CSE 590: Special Topics – Storage Systems
  - 20 graduate students
  - Hardware, file systems, RAID, VFS, and more
- Two phases
  - Phase 1: Run few existing courselets as onboarding, fill survey
  - Phase 2: Author their own courselets on a storage topic, demo
- Guest lectures (senior PhD student)
  - One on using courselets, one on developing them

The Infrastructure Journey



Stony Brook University

Teaching Storage on Chameleon

Safe Failure Scenarios

- RAID degradation demos
- No production data at risk
- No shared system ever at risk

Repeatable Runs

- Reset environment per experiment
- Mistakes become learning, not disasters
- Iterate without hardware delays
- Can run performance benchmarks

Flexible Topology

- Complex multi-disk topologies
- Floating IPs and networking
- Per-experiment environment layout
- Realistic (root /bin/sh commands)

Reusable Artifacts

- Courselets published on Trovi
- Used for lab onboarding
- Reusable across semesters

The Instructor Baseline

- 5 courselets sequenced from simple to complex
  - VM Initialization at KVM@TACC
  - Disk Partitioning
  - Benchmarking Disk I/O
  - RAID 0, RAID 1
- Dual purpose
  - Onboard students
  - Establish a template for courselet development

Students as Authors

- Shifted students from users to creators
  - Teams of 1-4 developed 16 courselets on storage topics
  - Each team worked for 3 weeks
  - Built on the instructor courselet model
- Student work was held to publication standards
  - Reviewed for correctness, clarity, and reproducibility
  - Revised through iterative feedback
- Best 12 submissions curated and published on Trovi

Anatomy of a Courselet

- Every courselet follows the same structure
  - Learning objectives declared upfront
  - Hardware requirements
  - Step-by-step cells alternating explanation and executable code
  - Document command purpose, arguments, and expected output
- Verification is mandatory
  - Every code block's effect is confirmed before moving on
    - Partition a disk, then confirm the layout with lsblk
- Frees all Chameleon resources
  - Students own the full lifecycle, including cleanup

Published Courselets

Instructor-Authored (5)

- VM Init at KVM@TACC
- Disk Partitioning Demo
- Benchmarking Disk I/O
- RAID 0 (Disk Striping)
- RAID 1 (Disk Mirroring)

Student-Authored (12)

File Systems

- Symbolic Links
- Hard Links
- Btrfs File System
- XFS File System

Space Efficiency

- Compression
- Dedup + Compression

Block Management

- LVM
- Storage Caching

Data Security

- dm-integrity
- dm-crypt

Network Storage

- NFSv3
- NFSv4

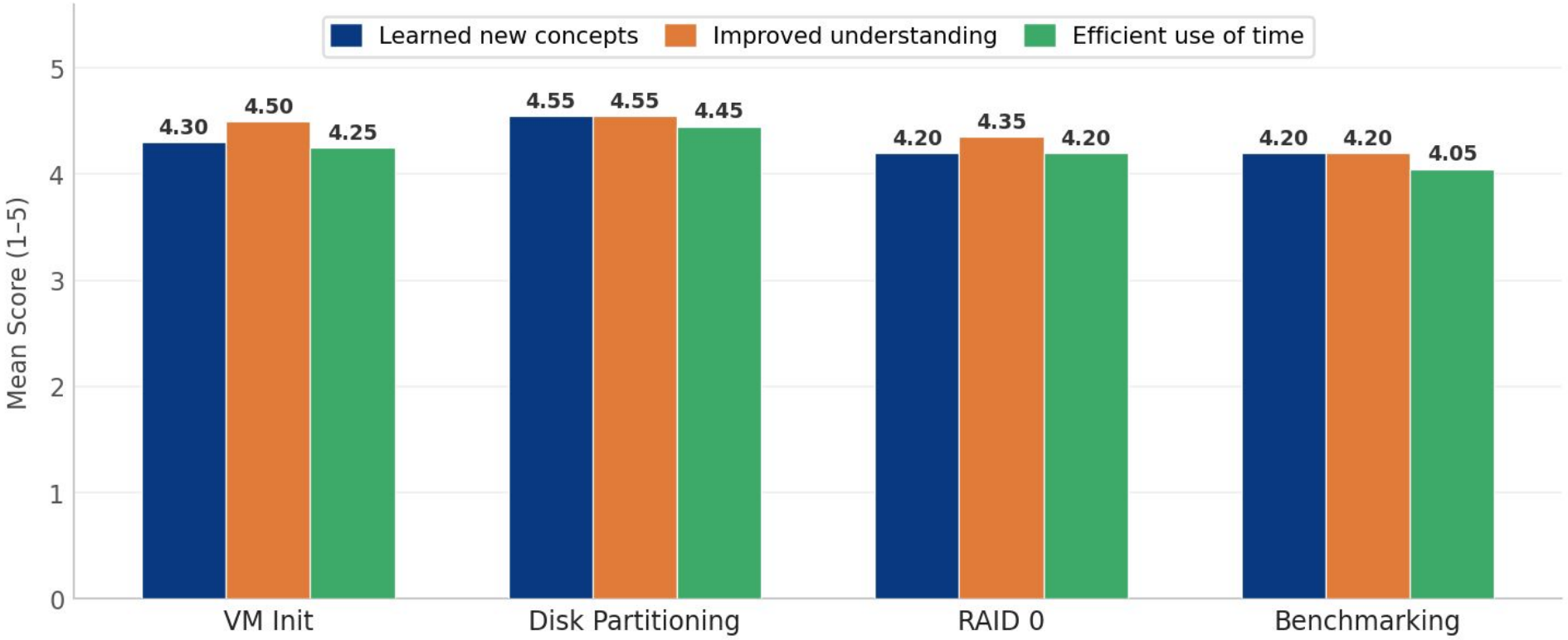
Live Demo

Walkthrough of a student-authored courselets on Trovi

Trovi: Topics in Storage Systems

Student Survey Results

n = 20 students · Scale 1–5 · 3 questions per courselet



Stony Brook University

Student Feedback: What Worked

Self-paced format

> *I like the Jupyter notebook style — I can go at my own pace and see what each command does as I go.*

RAID was the standout

> *I didn't know anything about RAID. This courselet was really informative — the best out of all.*

CLI unlocked new skills

> *I only ever used a GUI for disk management. Now I can do this from the command line on any device.*

Student Feedback: Constructive

Missing output context

> *I ran the benchmarks but had no clue what the results meant. Guidance on interpreting the output would help.*

Needs more background

> *More background on basic file and storage concepts would better benefit beginning learners.*

Supplement dependency

> *This was greatly helped by the Zoom walkthrough. It might have been too dense or cumbersome otherwise.*

Student Feedback: Critical

Passive by design

> *It felt like a walkthrough, not an assignment. No real incentive to experiment through the interface.*

AI prose concerns

> *When I realize text might be AI-generated, my confidence in its correctness instantly drops.*

Format skepticism

> *Courselets is not the way to go about teaching this. It just feels very disconnected.*

Learning Through Friction

- Chameleon provided pedagogically useful friction, not its elimination
- Provisioning and recovering from mistakes is the curriculum
  - Students should not be shielded from the mechanics being tested
- Failures were recoverable
  - Break a configuration, reset, and rerun with no shared system at risk
  - Mistakes became experiences rather than incidents

Conclusion

- Chameleon made storage labs practical
  - Safe, resettable multi-disk environments
  - Classroom-scale deployment
- Students learned through courselet development
  - 17 courselets curated and published
- The right infrastructure turns one semester into lasting contributions
  - Reusable for onboarding and future teaching

Teaching Storage Systems with Interactive Courselets on Chameleon

Thank You

Q & A

Erez Zadok and Tyler Estro
Stony Brook University
File Systems and Storage Lab (FSL)

FOUNT

Scaffolded, Hands-On Learning for a Data-Centric Future

Stony Brook University