Hydra Cluster Audit Report

Comprehensive Discovery, Cleanup & Hardening

Infrastructure Team
SUNY New Paltz Computer Science

February 4, 2026

Contents

1 Executive Summary

This document records the full audit, cleanup, and hardening of the 3-node Hydra RKE2 Kubernetes cluster performed on February 4, 2026. Key actions taken:

Completed Actions:

- Disabled host-level MariaDB systemd service (duplicate of Docker container)
- Disabled and masked Samba (smbd/nmbd) — no configured shares, ports 139/445 closed
- Deleted 4 orphaned ZFS storage classes (hydra-hot, hydra-warm, hydra-cold, hydra-gpu)
- Backed up all Docker Compose files, Apache configs, .env files, Traefik dynamic configs
- CS Lab backend + MariaDB fully deployed to K8s (15 tables, all data migrated)
- Created and applied 8 K8s IngressRoutes replacing all Apache proxy rules
- Audited Chimera and Cerberus (both clean — no Samba, RKE2 agents healthy)
- Archived and removed `/var/www` static content (∼436 MB)
- Archived Apache configuration files
- Stopped and removed 4 redundant Docker containers (cs-lab-backend, cs-lab-db, hydra-saml-auth, traefik)
- Cleaned dangling Docker images and unused volumes
- Removed stale ExternalName service `cs-lab-website`
- Fixed GPU operator: removed NVIDIA labels from Hydra (no GPU), GPU pods now only run on Chimera/Cerberus
- Fixed n8n and git-learning 502 errors (changed localhost bindings to 0.0.0.0)
- Fixed admin API unreachable (changed `app.js` listen to 0.0.0.0)
- Created shell aliases and test scripts (`~/.hydra-aliases`, `scripts/test-routes.sh`)
- Verified /etc/hosts consistency on all 3 nodes

Remaining:

- UFW firewall setup on all 3 nodes
- Cerberus kernel update reboot
- Backup automation (etcd snapshots, DB backups)
- Metrics agents on Chimera/Cerberus
- Migrate remaining Docker services to K8s (hackathons, n8n, java-executor, git-learning, sshpiper)

2 Phase 0: Discovery & Snapshot

2.1 Cluster Topology — Verified

Node	IP	Role	OS	Resources
Hydra	192.168.1.160	Control plane, etcd, master	Ubuntu 22.04.5	20 CPU, 256 GB RAM, 21 TB RAID-10
Chimera	192.168.1.150	Worker (GPU inference)	Ubuntu 24.04.2	48 CPU, 256 GB RAM, 3× RTX 3090
Cerberus	192.168.1.233	Worker (GPU training)	Ubuntu 24.04.3	48 CPU, 64 GB RAM, 2× RTX 5090

Table 1: Verified cluster topology. Cerberus confirmed at .233 (not .242 as in old Ansible inventory).

RKE2 Cluster: All 3 nodes Ready. RKE2 v1.28.4+rke2r1. Cluster age: 16 days. Containerd 1.7.7.

2.2 Chimera Status

- CPU: AMD Ryzen Threadripper 3960X (48 cores), 251 GB RAM
- GPU: 3× NVIDIA RTX 3090 (24 GB each) — all healthy, low utilization
- Disk: 3.5 TB Samsung NVMe, 54% used
- Services: OpenWebUI (:3000), Ollama (:11434), metrics agent (:9100)
- Uptime: 5 weeks, 6 days
- No Docker, no Samba — clean RKE2 agent with containerd

2.3 Cerberus Status

- CPU: AMD Ryzen Threadripper PRO 7965WX (48 cores), 62 GB RAM
- GPU: 2× NVIDIA RTX 5090 (32 GB each) — healthy, idle
- Disk: 3.6 TB MSI NVMe, 6% used
- Services: metrics agent (:9100), SSH
- Uptime: 3 days, 19 hours
- **Kernel update pending — system restart required**
- No Docker, no Samba — clean RKE2 agent

2.4 Docker Containers on Hydra (Pre-Audit)

Container	Port(s)	Image	Status
hydra-saml-auth	host network	hydra-saml-auth	Up 3d
traefik (Docker)	8081, 8082	traefik:v3.6	Up 3d
sshpiper	2222	farmer1992/sshpiperd	Up 3d
cs-lab-backend	5001	custom build	Up 2d
cs-lab-db	3306 (int)	mariadb:10.11	Up 3d
hackathons-app	45821	hackaton-voting-app	Up 3d
java-executor-service	55392	docker-java-executor	Up 6d
traefik-n8n-traefik-1	8080 (lo)	traefik:v3.3	Up 6d
traefik-n8n-n8n-1	5678 (lo)	n8n	Up 6d
gg-git-learning-app-1	38765 (int)	gg-git-learning-app	Up 6d
n8n-user-manager (x2)	3000 (int)	n8n-user-manager	Up 6d
traefik-n8n-postgres-1	5432 (int)	postgres:16-alpine	Up 6d

Table 2: 13 Docker containers running across 7 Compose projects.

2.5 Storage Architecture

- **Boot disk:** `/dev/mapper/ubuntu-vg-ubuntu-lv` — 1 TB LVM, 27% used
- **Data array:** `/dev/md0` — 21 TB mdadm RAID-10 (6 SSDs), mounted at `/data`, 0.3% used
- **No ZFS pools** — `zpool status` returns "no pools available"
- **SSD alignment verified:** Physical block size 4096, filesystem block size 4096, scheduler `mq-deadline`

2.6 RAID Details

```
/dev/md0: RAID-10, 6 active devices (sda-sdf)
Chunk Size: 512K, Layout: near=2
State: clean, 0 failed devices
Total: 20.96 TiB (23.04 TB)
Filesystem: ext4, 4096-byte blocks
```

3 Phase 1: Services Disabled

3.1 MariaDB Host Service — DISABLED

`systemctl stop mariadb && systemctl disable mariadb`
MariaDB 10.6.23 was running as a systemd service *and* as Docker container `cs-lab-db` (MariaDB 10.11). The Docker container is the active instance used by CS Lab Backend. The host service was a duplicate.

3.2 Samba — DISABLED & MASKED

`systemctl stop smbd nmbd && systemctl disable smbd nmbd && systemctl mask smbd nmbd`
Default Ubuntu Samba config with **no custom shares defined**. Only the stock `[printers]`

and [print$] shares existed. Neither Chimera nor Cerberus have Samba installed. No CIFS mounts anywhere in the cluster. Ports 139/445 are now closed.

3.3 Hydra Backend (Port 5002) — KEPT

`/srv/hydra-backend/app.js` is **not legacy**. It is the admin API that:

- Creates user accounts via `/opt/hydra-scripts/create_user.sh`
- Queries the `ServerDatabaseForm` table in MariaDB
- Retrieves PM2 and journal logs
- Authenticated via token in `/srv/hydra-backend/.token`

This service needs to be kept and eventually migrated to K8s.

4 Phase 2: Storage Class Cleanup

Deleted 4 orphaned ZFS storage classes:

```
kubectl delete storageclass hydra-hot hydra-warm hydra-cold hydra-
    gpu
```

These referenced `zfs.csi.openebs.io` but no ZFS pools exist. Storage is a single 21 TB mdadm RAID-10 SSD array.

Remaining storage classes:

Name	Provisioner	In Use
hydra-local	rancher.io/local-path	Yes (30 student PVCs)
hydra-nfs	nfs.csi.k8s.io	Yes (1 PVC)
local-path	rancher.io/local-path	No (duplicate)

5 Phase 3: CS Lab Migration to Kubernetes

5.1 Docker → K8s Migration

The CS Lab website (React frontend + Express backend + MariaDB 10.11) was migrated from Docker containers to Kubernetes deployments with full data preservation.

Migration completed:

1. Dumped Docker MariaDB data (132K lines, all tables)
2. Deployed K8s MariaDB pod with PVC for persistent storage
3. Imported initial 4-table dump into K8s MariaDB
4. Discovered missing tables (Events, Admins, Faculty, etc.) from legacy `db.txt`
5. Imported full schema: 15 tables with all data from `server/src/db.txt`
6. Deployed K8s CS Lab backend connecting to K8s MariaDB
7. Verified all API endpoints working: `/api/events`, `/api/faculty`, `/api/courses`, etc.
8. Stopped and removed Docker containers (cs-lab-backend, cs-lab-db)

5.2 Database Schema — 15 Tables

Table	Rows	Purpose
Admins	14	Admin users with bcrypt password hashes
Events	5	Department events with admin FK
Faculty	13	Faculty directory
FacultySemesters	6	Faculty-semester assignments
Faqs	6	FAQ entries
NoSchoolDays	2	Holiday calendar entries
SchoolCalendar	1	Academic calendar config
ServerDatabaseForm	2	Server access request forms
StudentHighlightBlog	7	Student project showcase
StudentResources	15	External learning resources
TechBlog	12	Tech blog articles
profiles	1	Admin profile details
Courses	40	CS course catalog
CourseResources	18	Course materials/links
CompExamSettings	1	Comprehensive exam config

Table 3: All 15 tables migrated from Docker MariaDB + legacy dump.

5.3 K8s Manifests

Manifests in `k8s/components/cs-lab/`:

- `deployment.yaml` — Backend (port 5001) + MariaDB (port 3306) deployments
- `service.yaml` — ClusterIP services for backend and DB
- `secret.yaml` — Database credentials
- `pvc.yaml` — 5 Gi PVC for MariaDB data
- `external-services.yaml` — ExternalName services for Java executor, Git learning, hydra-backend

6 Phase 4: Traefik Consolidation

6.1 Pre-Audit State: 3 Traefik Instances

Three separate Traefik instances were running:

1. **K8s Traefik** (v2.11) — hostPort 80/443, Let's Encrypt ACME, CRD provider
2. **Docker Traefik** (v3.6) — port 8082, file provider for student routes
3. **n8n Traefik** (v3.3) — port 8080 localhost, n8n-specific routing

6.2 Apache Configuration (Inactive)

Apache was **inactive/disabled** but had a 370-line config at `/etc/apache2/sites-enabled/hydra.newpaltz.` containing proxy rules for all services. This config served as the historical reference for what routes need to exist.

Routes extracted from Apache config:

- `/dashboard`, `/login`, `/logout`, `/auth`, `/token`, `/check`, `/servers` → hydra-saml-auth (:6969)
- `/api/courses`, `/api/faculty`, `/api/faq`, `...` → CS Lab Backend (:5001)
- `/students/*` → Docker Traefik (:8082) → student containers
- `/hackathons/` → hackathons-app (:45821)
- `/java/` → java-executor (:55392)

- `/git/` → git-learning (:8080)
- `/admin-api/` → hydra-backend (:5002)
- `gpt.hydra.newpaltz.edu` → chimera:3000 (OpenWebUI)
- `n8n.hydra.newpaltz.edu` → n8n (:5678)
- `/placeframe/` → :6721 (NOT RUNNING)
- `/studentmvp/` → :5175 (NOT RUNNING)
- `/minecraftdashboard/` → 192.168.1.145:3000 (stale external IP)

6.3 New K8s IngressRoutes Created

All Apache proxy rules have been converted to K8s IngressRoute CRDs:

File: `ingressroute-production.yaml`

- `hydra-main` — All hydra-saml-auth routes (dashboard, login, auth, servers, API)
- `cs-lab-website` — All CS Lab API and frontend routes (20+ path rules)
- `hackathons` — /hackathons/ path (updated from existing)
- `java-executor` — /java/ path
- `git-learning` — /git/ path
- `hydra-default` — Catch-all at priority 1 (CS Lab frontend)

File: `ingressroute-subdomains.yaml`

- `openwebui` — `gpt.hydra.newpaltz.edu` → chimera:3000
- `n8n` — `n8n.hydra.newpaltz.edu` → n8n:5678

Stale routes removed (services no longer running):

- `/placeframe/` — port 6721 not listening
- `/studentmvp/` — port 5175 not listening
- `/minecraftdashboard/` — 192.168.1.145 is not a cluster node

7 Phase 5: Cleanup Actions

7.1 Static Content Archived & Removed

`/var/www/` (∼436 MB) archived to `var-www-backup.tar.gz` then removed:

- `/var/www/FLAPJS-WebApp/` — JFLAP automata simulator
- `/var/www/interview-coach/` — Interview practice app
- `/var/www/lccjs/` — LC-3 JavaScript simulator
- `/var/www/LccWebUI/` — LC-3 Web UI
- `/var/www/gpt/` — Old GPT static files
- `/var/www/SUNYCAT.png` — 772 KB image
- `/var/www/html/` contents (kept `.well-known/` for ACME)

Apache was already inactive. None of this content was being served.

7.2 Apache Config Archived

Full Apache config archived to `apache-full-backup.tar.gz`. Stale files removed:

- `hydra.newpaltz.edu.conf.bak` — removed
- `lccjs.conf` symlink — removed

Main config retained as reference in `sites-available/`.

7.3 Docker Cleanup

Containers stopped and removed (now running in K8s):
- `cs-lab-backend` — replaced by K8s deployment
- `cs-lab-db` — replaced by K8s MariaDB pod
- `hydra-saml-auth` — replaced by K8s `hydra-auth` deployment
- `traefik` (standalone, ports 8081/8082) — replaced by K8s Traefik pod

Docker images removed:
- `newpaltz-cs-lab-website-backend`, `mariadb:10.11`, `traefik:v2.11`
- `hydra-saml-auth` (multiple tags), `portainer/portainer-ce`
- `nginx:stable-perl`, `busybox`, duplicate `traefik-n8n-user-manager`
- 100+ dangling image layers pruned

35 unused Docker volumes removed.

Remaining Docker containers (8, still needed):

Container	Port	Why Still Docker
hackathons-app	45821	ExternalName svc, not yet K8s native
traefik-n8n-traefik-1	8080	Routes git-learning via ExternalName
traefik-n8n-n8n-1	5678	n8n workflow engine
traefik-n8n-n8n-user-manager	3000	n8n user management API
traefik-n8n-postgres-1	5432	PostgreSQL for n8n
sshpiper	2222	Student SSH proxy
gg-git-learning-app-1	38765	Git learning app
java-executor-service	55392	Java code executor

8 Phase 6: Backups Created

All backups stored in `/home/infra/hydra-audit-20260204/backups/`:

File	Contents
`cslab-docker-compose.yml`	CS Lab Docker Compose
`hydra-saml-auth-docker-compose.yaml`	Main auth app + Docker Traefik
`sshpiper-docker-compose.yaml`	SSHPiper config
`hackathons-docker-compose.yml`	Hackathon voting app
`java-executor-docker-compose.yml`	Java code executor
`traefik-n8n-docker-compose.yaml`	n8n + Traefik + Postgres
`gg-git-learning-docker-compose.yml`	Git learning app
`apache-hydra.conf`	Apache main vhost config
`apache-lccjs.conf`	Apache lccjs vhost
`traefik-dynamic/`	All Traefik file-provider configs
`cslab.env`	CS Lab environment variables
`hydra-saml-auth.env`	Auth app environment
`hydra-backend.env`	Admin API environment
`cslab-mariadb-full-dump.sql`	Full MariaDB dump (5 MB)
`var-www-backup.tar.gz`	Archived /var/www content
`apache-full-backup.tar.gz`	Archived Apache config

Table 4: Complete backup manifest.

9 Phase 7: Network & Security Findings

9.1 Externally Exposed Ports (Pre-Hardening)

Port	Service	Status	Action
22	SSH	Required	Keep
80	HTTP (Traefik)	Required	Keep
443	HTTPS (Traefik)	Required	Keep
111	rpcbind (NFS)	Exposed	Restrict to cluster
139	Samba	**Closed**	Samba disabled
445	Samba	**Closed**	Samba disabled
2049	NFS	Exposed	Restrict to cluster
2222	SSHPiper	Required	Keep
5001	CS Lab	Exposed	Move behind Traefik
6443	K8s API	Exposed	Restrict to cluster
6969	hydra-auth	Exposed	Move behind Traefik
8081	Traefik Dash	Exposed	Restrict to localhost
8082	Docker Traefik	Exposed	Consolidate to K8s
9345	RKE2 Reg	Exposed	Restrict to cluster
45821	Hackathons	Exposed	Move behind Traefik
55392	Java Executor	Exposed	Move behind Traefik

Table 5: UFW firewall not yet enabled — pending final validation.

9.2 /etc/hosts Verification

All three nodes have consistent `/etc/hosts` entries:

```
192.168.1.160 hydra
192.168.1.150 chimera
192.168.1.233 cerberus
```

10 Phase 8: GPU Operator Fix

Problem: GPU operator DaemonSet pods stuck in `Init:0/1` on Hydra (no GPU). The error was: `no runtime for "nvidia" is configured`.

Fix: Set all `nvidia.com/gpu.deploy.*` labels to `false` on Hydra node:

```
kubectl label node hydra \
  nvidia.com/gpu.deploy.container-toolkit=false \
  nvidia.com/gpu.deploy.device-plugin=false \
  nvidia.com/gpu.deploy.gpu-feature-discovery=false \
  nvidia.com/gpu.deploy.operator-validator=false \
  nvidia.com/gpu.present=false --overwrite
```

Result: Only the NFD worker (expected on all nodes) remains on Hydra. All GPU pods properly running on Chimera and Cerberus only.

11 Phase 9: Route Testing & IngressRoutes

All 8 IngressRoutes applied and verified:

IngressRoute	Match	Backend	Status
hydra-main	/dashboard, /login, /auth, etc.	hydra-auth:6969	200/302
cs-lab-website	/api/*, /courses, /events, etc.	cs-lab-backend:5001	200
hackathons	/hackathons/	hackathons:45821	200
java-executor	/java/	java-executor:55392	Routed
git-learning	/git/	git-learning:8080	Routed
hydra-default	catch-all (priority 1)	cs-lab-backend:5001	200
openwebui	gpt.hydra.newpaltz.edu	chimera:3000	200
n8n	n8n.hydra.newpaltz.edu	n8n:5678	200

Fixes applied during testing:

- n8n Docker port changed from `127.0.0.1:5678` to `0.0.0.0:5678`
- n8n-traefik Docker port changed from `127.0.0.1:8080` to `0.0.0.0:8080`
- `/srv/hydra-backend/app.js` listen changed from `127.0.0.1` to `0.0.0.0`
- IngressRoute API group fixed: `traefik.io/v1alpha1` → `traefik.containo.us/v1alpha1`

12 Remaining Work

1. **UFW firewall** — Enable firewall on all 3 nodes (do last, keep SSH fallback)
2. **Cerberus reboot** — Pending kernel update requires system restart
3. **Backup automation** — Daily etcd snapshots, K8s MariaDB dumps, config exports
4. **Metrics agents** — Deploy Prometheus node exporter on Chimera and Cerberus
5. **Migrate remaining Docker → K8s** — hackathons, n8n stack, java-executor, git-learning, sshpiper
6. **n8n auth typo** — Fix `mplotkin@jewpaltz.com` → `mplotkin@newpaltz.edu` in PostgreSQL