



CzechAI Server Inventura — Kom- pletní audit infrastruktury plat- formy

Master of Science HES SO in
Engineering

Profil/Orientation Orientation

Date of publication: 10.05.2026

Produced by:

Author Bohemia Shield · CzechAI Operations

Under the supervision of Teacher Name

In collaboration with Company Name



Table of Contents

Declaration of authenticity	3
Acknowledgments	4
Resume (English version)	5
Résumé (version française)	5
1. Úvod	6
1.1. Cíle auditu	6
2. Inventura serverů	6
2.1. Přehled – 6 serverů	6
2.2. Klíčové metriky	6
3. Detail serverů	7
3.1. Server 179 – czechai (orchestrace)	7
3.2. Server 117 – Docker host	7
3.3. Server 134 – Reality scrapery (kandidát na vyřazení)	7
3.4. Server 40 – Cenová mapa V24 (nově doplněn)	7
3.5. Server dedi9544 – Hetzner Robot	7
3.6. Server GPU 170 – LLM inference	7
4. Doporučení	8
5. Závěr	8





Declaration of authenticity

I hereby affirm that this assignment is my own written work and that I have used no other sources or aids other than those indicated. All passages that have been quoted from publications or paraphrased from these sources are indicated as such. Additionally, any assistance from artificial intelligence tools has been appropriately acknowledged and documented.

Date : 10.05.2026

Name : Bohemia Shield · CzechAI Operations



structured-mse-thesis.typ	To be finalized	
HES-SO / Domaine I&A / MSE / Travail de Master	10-05-2026 - Bohemia Shield · CzechAI Operations	- 3 / 8 -



Acknowledgments

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aequaleam animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut.



structured-mse-thesis.typ	To be finalized	
HES-SO / Domaine I&A / MSE / Travail de Master	10-05-2026 - Bohemia Shield · CzechAI Operations	- 4 / 8 -



Resume (English version)

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aequi doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut.

Résumé (version française)

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aequi doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut.



structured-mse-thesis.typ	To be finalized	
HES-SO / Domaine I&A / MSE / Travail de Master	10-05-2026 - Bohemia Shield · CzechAI Operations	- 5 / 8 -



1. Úvod

CzechAI je technologická platforma kombinující AI agenty, multi-modelové LLM inference, real-estate data pipeline a workflow automatizaci. Provozována je na rozsáhlé infrastruktuře 6 serverů u poskytovatele Hetzner (kombinace Cloud + Robot Dedicated) a externí GPU provider.

Tento dokument představuje kompletní audit infrastruktury k datu 10. 5. 2026 a navrhuje konsolidační kroky vedoucí ke snížení provozních nákladů a zlepšení výkonu.

1.1. Cíle auditu

- Identifikovat všechny servery a jejich aktuální využití.
- Zmapovat běžící služby (PM2 procesy, Docker kontejnery).
- Detekovat přetížené nebo nevyužité servery.
- Navrhnout konsolidační strategii.
- Aktualizovat dokumentaci (CLAUDE.md, MKDocs).

2. Inventura serverů

2.1. Přehled — 6 serverů

Tabulka níže sumarizuje všechny servery včetně provider, hardware, OS, využití disku a primárního účelu.

Server	IP	OS	RAM	Disk	Účel
179 czechai	46.224.121.179	Ubuntu	30 GB	150 GB · 85 %	Orchestrace, ng-inx, PM2
117	94.130.218.117	Ubuntu	251 GB	875 GB · 78 %	Docker host
134 czechaiscraper	46.224.232.134	Ubuntu 24	7.6 GB	150 GB · 72 %	Reality scrapery (k vyřazení)
40 (NEW)	94.130.223.40	AlmaLinux 8.10	251 GB	7.4 TB	Cenová mapa V24, reality
dedi9544	188.40.28.226	Debian 13	252 GB	1.7 TB · 3 %	Idle (migrace připravena)
GPU 170	217.198.112.170	—	96 GB VRAM	—	LLM inference

2.2. Klíčové metriky

- **Servery v ekosystému:** 6 (5 Hetzner + 1 externí GPU)
- **Celková RAM:** 796 GB
- **Celkový disk:** 10 TB
- **GPU VRAM:** 96 GB
- **PM2 procesů:** 135
- **Docker kontejnerů:** 74
- **Nginx subdomén pod czechai.io:** 150+



3. Detail serverů

3.1. Server 179 — czechai (orchestrace)

Hlavní orchestrační server. Ubuntu, 8 vCPU (sdílený), 30 GB RAM, 150 GB SSD (85 % využito), load average 4.84.

Co tam běží: 79 PM2 procesů, 30 Docker kontejnerů, 150+ nginx subdomén. Klíčové porty: 22 (SSH), 80 / 443 (Nginx), 8090 (Router API), 3002 (Chat frontend), 8093 (AML platform), 8095 (RAG API).

3.2. Server 117 — Docker host

Hetzner CCX, dedikovaný CPU, 251 GB RAM, 875 GB disk (78 %).

Hostované služby: Twenty CRM, Outline Wiki, SiYuan, Blinko, Affine, Supabase, Chatwoot, WAHA, Bolt.diy, Big AGI, CrewAI Studio. Záložní Ollama na portu 11434.

3.3. Server 134 — Reality scrapery (kandidát na vyřazení)

Hetzner CPX32, 8 vCPU, 7.6 GB RAM. **Server je dlouhodobě přetížený:** swap thrashing 8.7 GB / 11 GB, load average 5–7.

Co tam běží: 18 web scraperů (Bezrealitky, Bazoš, iDNES, RemAX, Drazby), reality pipeline, PostgreSQL databáze reality_scraper (1.2 GB).

3.4. Server 40 — Cenová mapa V24 (nově doplněn)

AlmaLinux 8.10 (RHEL-based), 12 cores, 251 GB RAM (245 GB volné), celkem 7.4 TB disk. Load 0.79 (idle).

Důležitý kontext: Tento server byl chybějící v původní inventuře z 9. 5. 2026 a doplněn 2026-05-10 na základě upozornění.

Co tam běží: Produkce shadow.czechai.io (cenová mapa V24, port 5163), reality pipeline (paralelní s 134), shadow-url-estimator. Identifikován crash loop procesu merge-guardian s 30 056 restarty.

3.5. Server dedi9544 — Hetzner Robot

Debian 13, 252 GB RAM (idle), 1.7 TB disk (3 % využito), load 0.23.

Stav migrate: 80 % migrate ze 134 je již připraveno (/scrapers/reality-pipeline/ s 6.9 GB kódu, /data/reality-pipeline/db/ s 1.6 GB MASTER_LIVE + ARCHIVE), ale 0 procesů aktuálně běží.

3.6. Server GPU 170 — LLM inference

RTX PRO 6000 Blackwell, 96 GB VRAM. Ollama na portu 11434.

Modely (4, celkem 79.4 GB):

- huihui_ai/Qwen3.6-abliterated:35b — chat + vision (256K context, 22 GB)
- huihui_ai/deepseek-r1-abliterated:70b — reasoning (40 GB)
- huihui_ai/qwen3-coder-abliterated:30b — code (17 GB)
- nomic-embed-text — embeddings (0.3 GB)

4. Doporučení

1. **Migrate 134 → 40 + dedi9544.** Server 40 (compute) přebere reality pipeline a browser scrapery. dedi9544 (data) převezme PostgreSQL databázi a backupy. Po ověření funkčnosti server 134 vyřadit. Očekávaná úspora: 10.49 €/měs.
2. **Hetzner Cloud Private Network.** Servery 179, 117, 40 v jednom privátním síťovém segmentu 10.10.0.0/16 (zdarma). dedi9544 připojen přes vSwitch (2 €/měs). Veškerý mezi-server traffic interní = rychlejší
 1. bezpečnější + bez billingu.
3. **Disk cleanup 179.** /opt/router-api/venv obsahuje 2.2 GB CUDA knihoven, ale server nemá GPU. Reinstall PyTorch jako CPU-only. Statické GIS shapefiles (861 MB) přesunout na dedi9544.
4. **Fix merge-guardian crash loop na 40.** Proces zaznamenal 30 056 restartů. Před migrací nutná analýza příčiny a oprava.
5. **Aktualizovat CLAUDE.md.** Sekce „Cenová mapa V24 – port 5163 na 179“ je chybná. Skutečnost: shadow.czechai.io produkce běží na serveru 40.

5. Závěr

Audit identifikoval 6 serverů s celkovou kapacitou 796 GB RAM, 10 TB diskového prostoru a 96 GB VRAM. Hlavním navrženým krokem je konsolidace serverové infrastruktury přes migraci 134 → 40 + dedi9544, která sníží měsíční náklady o 10.49 € a vyřeší dlouhodobý problém swap thrashingu na serveru 134.

Současně doporučujeme aktualizaci dokumentace CLAUDE.md, kde je nesprávně lokalizována produkční služba shadow.czechai.io (běží na serveru 40, ne na 179).