Carlos Eckert

HPC & AI Infrastructure Engineer
Philadelphia, PA | 267-850-8461 | carlos.eckert.dev@gmail.com | linkedin.com/in/carlos-eckert/ | github.com/toucheLos

SUMMARY

HPC engineer running production GPU infrastructure at scale. Sole administrator of three production clusters at Temple University - 34 GPU nodes, 100+ CPU nodes, 78+ researchers - and spent a summer tuning GPU allocation across 356 H100 and A100 nodes on Mount Sinai's 11-petaflop Minerva supercomputer. Deep Slurm internals and cluster security experience. Author of VIEB, a GPU-accelerated behavioral analysis pipeline processing 17M+ frames, in active use in live neuroscience experiments.

PROFESSIONAL EXPERIENCE

High-Performance Computing Engineer, Temple University *September 2024 – Present*
• Sole administrator responsible for end-to-end operation of three HPC clusters (compute, storage, networking), supporting 78+ researchers across multi-partition SLURM environments
• Provisioned and operated NVIDIA A100 GPU fleets on AlmaLinux 9 with Slurm, Cobbler, Ansible, and Ganglia; own full node lifecycle from bare-metal PXE provisioning through production scheduling
• Conducted internal security assessment of Slurm attack surface - prolog/epilog injection, munge key handling, module system and SUID escalation paths - and hardened cluster configuration accordingly
• Diagnosed and resolved fleet-wide NVIDIA driver generation mismatch (570 vs 595), iDRAC event storms, and recurring GPU node scheduling failures across heterogeneous hardware
• Instructed HPC module for undergraduate REU cohort spanning biology, applied math, bioengineering, chemistry, and computer science

Computational Neuroscience Assistant, Temple University *December 2025 – Present*
• Built GPU-accelerated ML pipeline extracting 49-99 kinematic features per frame across 3,846 videos and 22M+ frames of behavioral recording
• Debugged JAX/CUDA and cuML/RAPIDS at scale - implemented chunked UMAP transforms in 200K-row batches to fix GPU memory constraints
• Performed IHC, Jess automated western blot (GFAP, GluA1, GluA2, NMDA1), and histological sectioning in a transgenic Alzheimer's mouse model

HPC Minerva Intern, Mount Sinai Hospital, *June 2025 – August 2025*
• Optimized and supported operations on Minerva, Mount Sinai's 11+ petaflop supercomputer supporting over 2,000 research workflows and $142M in NIH funding
• Developed utilization tracking scripts across 356 GPUs including H100 and A100 nodes to dynamically rebalance allocations
• Reduced idle compute through usage-aware scheduling strategies, improving resource efficiency across the cluster

PROJECTS

VIEB (**Video Interpreter Excluding Bias**), github.com/toucheLos/VIEB
• Unsupervised behavioral state discovery pipeline in active use in live neuroscience experiments
• Identified a fear discrimination state invisible to traditional human scoring
• DeepLabCut pose estimation -> kinematic feature extraction -> UMAP and HDBSCAN clustering without predefined labels
• Built six-axis evaluation framework measuring temporal coherence, usage balance, DBCV separability, ARI/NMI stability, cross-validated predictive decoding, and cross-method convergent validity

Research Assistant - NeuroVISOR Software, Temple University *September 2025 – May 2026*
• Designed extensible, object-oriented scientific visualization software for neuronal modeling, enabling dynamic addition and removal of ion channel components
• Implemented an Euler method solver for action potential dynamics, cross-validated against Yale NEURON

EDUCATION

Temple University, B.S. Computer Science and Mathematics *May 2026*
Coursework: Numerical Analysis, Linear Algebra, Operating Systems, Stochastic Processes, Math Modeling
Organizations: Game Theory Club (Founder)

SKILLS & INTERESTS

HPC & Infrastructure: Slurm (multi-partition, GPU scheduling, prolog/epilog, fairshare), AlmaLinux / RHEL, Ansible, Cobbler, Ganglia, munge, Buildbot, bare-metal provisoning, cluster security, Git, Linux
GPU & Parallel Computing: CUDA, NVIDIA A100 / H100, JAX, cuML / RAPIDS, LAMMPS, Kubernetes
Languages: Python, C, C#, C++, Java, Shell, CUDA, SQL
Spoken: Spanish (Advanced), Japanese (Intermediate)