Wayne Hendricks

twh@waynehendricks.com · waynehendricks.com · github.com/twh · ORCID 0000-0001-9111-8968

Summary

Paige.AI's founding HPC engineer, from 2018 through its $81M acquisition by Tempus. I built the training infrastructure behind Virchow (*Nature Medicine*, 2024), PRISM, and the company's FDA-cleared diagnostic products, scaling it from 5 to 40+ ML researchers. Before that, HPC at Memorial Sloan Kettering and Caltech.

Career History

Senior HPC Engineer

Tempus AI (formerly Paige.AI), New York City *2018–Present*

- **Founding HPC engineer at Paige.AI, from 2018 through the September 2025 Tempus acquisition.** Owned end-to-end GPU, storage, scheduling, and multi-cloud infrastructure for clinical foundation model training.
- Built the training infrastructure behind the Virchow, Virchow2, PRISM, and PRISM2 pathology foundation models and the OmniScreen biomarker screening system.
- Architected "Azul", a dynamic Azure CycleCloud and Slurm cluster scaling on-demand to hundreds of GPUs with 7 PB of whole-slide imaging on Azure Managed Lustre; ran 1.5M jobs in 2025.
- Grew the environment from 5 to 40+ ML researchers across four NVIDIA generations (Pascal, Volta, Ampere, Hopper) and from on-prem to Azure and AWS.
- Helped hire the early management and engineering staff, including ML, infrastructure, and platform leads.
- Now migrating Paige's training and inference infrastructure into Tempus's GCP environment as its pathology models fold into Tempus's oncology foundation-model effort.

HPC Engineer

Memorial Sloan Kettering Cancer Center, New York City *2018–2022 (concurrent with Paige.AI)*

- Designed and built a 100-GPU DGX-1 cluster with Cisco RoCE networking and Pure Storage FlashBlade for high-throughput AI/ML workloads.
- Expanded storage to 4 PB of object (S3/Qumulo) and 500 TB of Pure FlashBlade by migrating from a flat to a tiered model based on data temperature.
- Evaluated and helped acquire a turn-key scientific datacenter facility.

Computing and Software Systems Research Engineer

California Institute of Technology High Energy Physics – CMS, Pasadena, California *2015–2018*

- Principal administrator of the Caltech Tier-2 cluster for the Compact Muon Solenoid experiment at CERN's Large Hadron Collider, one of the largest production scientific computing environments in physics.
- Operated a 7,300-slot HTCondor scheduling system with 4.5 PB of storage in continuous collaboration with Fermilab, CERN, and the Open Science Grid.

- Redesigned the core network topology into a tiered model to minimize latency and increase redundancy for distributed scientific workflows.
- Developed a state-of-the-art SDN testbed for testing scientific workflows via contributions to the ESnet project.
- Presenter and technical exercise contributor to ESnet network demonstrations at Supercomputing; also presented at HEPiX and OSG/HTCondor Week.

Earlier Career

Duke University, Durham, NC — *2011–2015* — Principal Unix/Linux and network analyst on the IT Operations Management team. Owned monitoring, alerting, and change management across all academic and enterprise infrastructure; principal administrator for application performance and 24/7 NOC alerting systems.

FedEx, Memphis, TN — *2005–2011* — Senior Technical Analyst and Technical Analyst supporting Unix/Linux production systems for global logistics, SCADA infrastructure for sorting facilities, and the www.fedex.com distributed infrastructure. Started as Intern II in 2005 with the Systems Administration and Consulting group.

Contractor and university work, 2000–2005 — Building network and wifi upgrades, departmental websites, and a business backup solution for critical data.

Technical Expertise

NVIDIA DGX and GPU cluster architecture across Pascal, Volta, Ampere, Hopper, and Blackwell. Slurm, Azure CycleCloud, HTCondor, Kueue, and Kubernetes for scheduling and orchestration. InfiniBand, RoCE, NVLink/NVSwitch, GPUDirect RDMA, NCCL, and DOCA for high-performance fabrics. Lustre, Weka, GPFS, NetApp, Pure, and S3/Blob for high performance storage and data archival at petabyte scale. Azure, AWS, and GCP with Terraform, Packer, and Ansible. Network performance and systems debugging on Linux, and automation with mostly shell and Python.

Publications

Papers where I am credited for infrastructure contributions.

Virchow: A Million-Slide Digital Pathology Foundation Model

E. Vorontsov, et al. · arXiv preprint, 2023

A 632M-parameter self-supervised vision transformer for computational pathology, trained on a million-slide whole-slide image archive from Memorial Sloan Kettering Cancer Center. Later published in *Nature Medicine* (2024).

Contribution: Built and operated the GPU compute infrastructure and high-throughput storage environment used to train and validate the model.

arXiv · PDF

Virchow2: Scaling Self-Supervised Mixed Magnification Models in Pathology

E. Zimmermann, S. Liu, et al. · arXiv preprint, 2024

Three vision transformer foundation models trained on 3.1 million histopathology whole-slide images: Virchow2 (632M), Virchow2G (1.9B), and Virchow2G Mini (22M distilled). State-of-the-art on 12 tile-level tasks.

Contribution: Designed and operated the GPU compute infrastructure and high-throughput storage environment used to train and validate all three models.

arXiv · PDF

PRISM: A Multi-Modal Generative Foundation Model for Slide-Level Histopathology

S. Liu, et al. · arXiv preprint, 2024

A multi-modal generative foundation model operating at the slide level for computational pathology, jointly modeling histology and clinical text.

Contribution: Built and maintained the HPC infrastructure supporting large-scale whole-slide image preprocessing, model training, and validation.

arXiv · PDF

PRISM2: Unlocking Multi-Modal General Pathology AI with Clinical Dialogue

E. Vorontsov, G. Shaikovski, A. Casson, et al. · arXiv preprint, 2025

A multi-modal slide-level pathology foundation model trained on 700,000 diagnostic specimen-report pairs and 14M question-answer pairs, the largest vision-and-language histopathology dataset to date. Aligns histomorphologic features with diagnostic language via clinical-dialogue supervision.

Contribution: Built and maintained the HPC infrastructure supporting large-scale whole-slide image preprocessing, model training, and validation.

arXiv · PDF

Screen Them All: High-Throughput Pan-Cancer Genetic and Phenotypic Biomarker Screening from H&E Whole Slide Images

Y. K. Wang, L. Tydlitatova, J. D. Kunz, et al. · arXiv preprint, 2024

OmniScreen, a high-throughput system predicting a broad range of clinically relevant molecular biomarkers across cancers from H&E whole-slide images, built on Virchow2 embeddings from 60,529 patients with paired MSK-IMPACT panels.

Contribution: Built and maintained the HPC infrastructure supporting large-scale whole-slide image preprocessing, embedding generation, model training, and validation.

arXiv · PDF

High Speed Scientific Data Transfers Using Software Defined Networking

H. Newman, et al. · INDIS ‘15 (SC15), Austin, Texas

Second Workshop on Innovating the Network for Data-Intensive Science, co-located with SC15: The International Conference for High Performance Computing, Networking, Storage and Analysis.

Contribution: Built and operated the SDN testbed and high-speed data-transfer infrastructure used in the demonstrations.

DOI · PDF

SDN-NGenIA: A Software Defined Next Generation Integrated Architecture for HEP and Data Intensive Science

J. Balcas, T. W. Hendricks, D. Kcira, A. Mughal, H. Newman, M. Spiropulu, J. R. Vlimant · Journal of Physics: Conference Series, Vol. 898 (CHEP 2016)

A software-defined next-generation integrated architecture supporting high-energy physics and data-intensive science. Presented at the 22nd International Conference on Computing in High Energy and Nuclear Physics (CHEP 2016), San Francisco.

Contribution: Built and operated the SDN testbed and high-speed data-transfer infrastructure used in the demonstrations.

DOI · PDF

HTTP as a Data Access Protocol: Trials with XrootD in CMS's AAA Project

J. Balcas, B. P. Bockelman, D. Kcira, H. Newman, J. Vlimant, T. W. Hendricks · Journal of Physics: Conference Series, Vol. 898 (CHEP 2016)

Evaluation of HTTP as a data access protocol for the CMS Any-data Anytime Anywhere (AAA) project, comparing performance and operational characteristics against XrootD's native protocol.

Contribution: Built and operated the SDN testbed and high-speed data-transfer infrastructure used in the demonstrations.

DOI · PDF