HYPHA

Distributed Compute Engine

Benchmark Report & Technical Overview

Jack Mays, Founder

Danville, Virginia

April 2026

Executive Summary

Hypha is a distributed compute engine that dispatches workloads as micro-operations across autonomous agents. It processes data entirely in memory and writes only final structured output to disk. Agents pull work independently with no orchestration layer, no scheduler overhead, and no dependency on cloud infrastructure.

This document presents the results of a sustained production benchmark: the full HCP Young Adult connectome dataset processed by a single Hypha agent running on a 13-year-old Dell Precision T7610 workstation. The run completed without intervention, without crashes, and without cloud infrastructure of any kind.

The results demonstrate that Hypha delivers production-grade distributed compute on commodity hardware that organizations already own.

Benchmark: HCP Young Adult Connectome

Dataset

The Human Connectome Project (HCP) Young Adult dataset is one of the most complex neuroimaging datasets in existence. It contains high-resolution multi-modal MRI data from

approximately 1,200 subjects and is used by neuroscience researchers worldwide to study brain connectivity, neurological conditions, and cognitive function.

For this benchmark, the full synapse export was processed: 166 JSONL source files representing the complete synapse-level connectivity data of the human connectome.

Hardware Configuration

Component	Specification
System	Dell Precision T7610
CPU	2x Intel Xeon E5-2637 v2 @ 3.50GHz
CPU Age	12 years (released 2013)
Cores / Threads	8 cores / 16 threads (dual socket)
RAM	251 GB DDR3 ECC
Storage	915 GB SATA, single drive
Disk Read Speed	198 MB/sec (buffered)
Operating System	Ubuntu 24.04, kernel 6.17
Python	3.x, pandas 3.0.1, pyarrow 23.0.1
Hypha Agents	1 (single agent)
Cloud Infrastructure	None

Benchmark Results

Metric	Value
Input files	166 JSONL files
Input size	~136 GB
Output files	166 Parquet files
Output size	7.1 GB
Total synapse records	166,216,068
Rows per file (avg)	~1,001,964
Compression ratio	95% (136 GB → 7.1 GB)
Wall clock time	58 hours 59 minutes
Throughput	~2.8 million records/hour

Metric	Value
CPU utilization	92-94% user, sustained
I/O wait	~0% (compute-bound)
Memory used	~85 GB of 251 GB available
Swap activity	Minimal (cold process memory only)
Crashes	0
Manual interventions	0
Cloud infrastructure used	None

Data Quality

Post-run validation confirmed the output data is clean and correctly typed:

Column	Type	Null Count	Status
pre_cell_id	Int64	13,074	Expected — orphan synapses
post_cell_id	Int64	22,502	Expected — orphan synapses
pre_site_id	Int64	0	Clean
post_site_id	Int64	0	Clean
x_nm	Int64	0	Clean — all synapses located
y_nm	Int64	0	Clean — all synapses located
z_nm	Int64	0	Clean — all synapses located
synapse_type	Int64	0	Clean
contact_area	Float64	All	Absent in source dataset
confidence	Float64	0	Clean
source_file	string	0	Clean

Every synapse record contains valid 3D coordinates (x_nm, y_nm, z_nm). Null values on pre_cell_id and post_cell_id reflect orphan synapses in the source data — a known characteristic of connectome datasets, not a pipeline error. contact_area is absent across all 166 source files, confirming this field was not captured in this export of the H01 dataset.

How Hypha Works

Core Philosophy

Most distributed compute systems are orchestrated top-down: a central controller assigns work to workers, monitors them, and retries on failure. This creates coordination overhead, scheduler bottlenecks, and systems that are only as reliable as their orchestration layer.

Hypha inverts this model. Agents are autonomous. They pull work from a job queue independently, process it entirely in memory, write only final structured output to disk, and return to a waiting state. The controller dispatches jobs and tracks state — it does not manage agents directly.

This architecture is inspired by mycelium biology: distributed, self-organizing, and resilient. No single node is critical. No orchestration layer can fail. Work finds the compute that can handle it.

Key Properties

- Agents pull work autonomously — no top-down assignment
- Processing happens entirely in memory — no intermediate disk writes
- Only final structured output is written to disk
- Hardware-agnostic — CPU, GPU, NPU, TPU agents in the same fleet
- No cloud dependency — runs on hardware organizations already own
- No orchestration overhead — no Kubernetes, no scheduler, no coordinator
- Scales horizontally by adding agents — no reconfiguration required
- Heterogeneous fleet support — fast and slow agents coexist naturally

Workload Dispatch

Jobs are submitted to the controller as structured payloads specifying an operation name and input parameters. The controller assigns leases to agents that claim jobs. Agents execute the operation, return a result, and the controller records completion. Failed jobs are requeued automatically.

The path contract — standardized input and output paths shared across all agents — means the controller requires no knowledge of individual agent hardware or capability. The right workload finds the right machine naturally.

Why This Matters

The Problem with Existing Solutions

Enterprises processing large datasets today face a difficult choice: pay for cloud compute that sends sensitive data outside their walls, wait in queue for shared HPC resources, or invest in expensive dedicated infrastructure that sits idle between workloads.

For regulated industries — healthcare, legal, pharmaceutical, financial services, government — the cloud is often not an option. Data residency requirements, compliance obligations, and security policies mean that sensitive data must stay on-premise.

What Hypha Changes

Hypha turns the hardware an organization already owns into a production compute fleet. The Dell Precision T7610 used in this benchmark is a 13-year-old workstation available on the secondary market for under $500. It processed 166 million synapse records from one of the most complex neuroscience datasets in the world, without cloud infrastructure, without a dedicated HPC allocation, and without a single crash.

The same architecture that processed the HCP connectome can process any structured workload: financial transactions, legal documents, genomic data, IoT sensor streams, insurance actuarial models, manufacturing quality data.

The engine does not care what the data is. It cares that the work gets done correctly.

Observable Compute

Because every job knows which agent ran it, how long it took, and what it produced, Hypha generates a natural audit trail of compute activity. Which agents worked together on a task. Which machines were idle. Which workloads saturate specific hardware.

This transforms infrastructure monitoring from passive observation into active intelligence — a correlation map of your own compute fleet, showing the relationships between machines during a workload in the same way a functional connectivity matrix shows relationships between brain regions.

Deployment Model

On-Premise. One Time. Yours Forever.

Hypha is sold as an on-premise deployment. No subscription. No usage metering. No data leaving the building. Organizations purchase the system, it runs on their hardware, and they own it.

Packages

Package	Includes	Price
Starter	Controller + 20 agents, 5-day onsite setup, 30-day hypercare, 1-year support	$50,000
Professional	Controller + 100 agents, custom pipeline, extended onsite, 1-year support	$150,000
Enterprise	Full deployment, custom ops, multi-site, white glove	Negotiated

Annual maintenance at 15-20% of purchase price covers updates, new ops development, and ongoing support. Additional ops — custom pipelines for specific workloads — are scoped and priced separately.

Current Status & Next Steps

Completed

- Core distributed engine: agent lifecycle, epoch fencing, lease TTL, requeue
- Full HCP Young Adult synapse pipeline: 166M records, 59 hours, zero crashes
- GPU acceleration: FAISS similarity search across 4x NVIDIA M10 GPUs in parallel
- Dual-path CPU/GPU validation with reconciliation op
- Provisional patent filed (Womble Bond Dickinson)
- Benchmark suite: HCP connectome, NYC Taxi, Gutenberg corpus

In Progress

- Multi-agent scaling validation: 10-subject calibration across heterogeneous fleet
- NFS shared storage layer for multi-node agent coordination
- HCP Young Adult full pipeline: FC dense, FC parcel, QC, package
- NYC Taxi benchmark: Hypha vs pandas vs Dask comparison

Target Markets

- Neuroscience and neuroimaging research laboratories
- Regulated industries: healthcare, pharmaceutical, legal, financial services
- Commercial research organizations requiring on-premise AI compute
- Enterprises with existing hardware seeking compute efficiency

Contact

Jack Mays

Founder, Hypha LLC

Danville, Virginia

HYPHA | Patent Pending

© 2026 TKT, Inc. All Rights Reserved.

HYPHA technology is the subject of pending patent applications. The contents of this white paper, including its text, illustrations, diagrams, and technical descriptions, are protected by applicable copyright laws. No portion may be reproduced, distributed, or transmitted without prior written permission from TKT, Inc.

This document is confidential and intended for the recipient only.