HUSSH TECHNOLOGIES CORPORATION · ENGINEERING

🤫

# 🤫 One — Engineering Deep Dive

## Protocol · Product · System, on the NVIDIA stack

Prepared for Jensen Huang. Systems architecture to design to the engineered product — one consent protocol, one placement engine, one agent, vertically integrated across NVIDIA's three planes: **personal supercomputing** above all, **edge**, and the **data center**.

THE INTEGRATION THESIS

Most software meets NVIDIA at an API. This plan integrates at four depths: **silicon features** (unified memory, FP4, confidential computing), **runtime** (DGX OS, NemoClaw, OpenShell, NIM, Dynamo, NeMo RL), **the open agent-security stack** (Alliance, NOOA, SPIFFE), and **economics** (receipts as meters; intelligence per dollar, down to N = 1). Protocol, product, and system — not a wrapper.

Companion to the open letter (hushh.ai/one/nvidia) and "The Fourth Utility." Every NVIDIA capability cited from their published materials; every hussh claim graded **IMPLEMENTED** · **IN DEVELOPMENT** · **ROADMAP**. Measured results shown are from our test suite; TARGET figures are budgeted design goals, never presented as measured.

WHAT'S INSIDE

# Contents



### How to read this

This is the builder's document behind the strategy. It assumes the reader has shipped silicon, runtimes, and rack-scale systems, and wants the actual design: message shapes, decision functions, lifecycle sequences, failure behavior, and the acceptance demo for every milestone. Where our published technical specification already fixed a design (control/data-plane split, PCHP types, the tail-tolerant executor, the 60-second budget), this document extends it onto the NVIDIA stack rather than restating it. Where a schema or flow is **proposed joint work**, it is labeled exactly that — nothing unimplemented is dressed as shipped.

01

# System Overview

One protocol, one engine, one agent — deployed across three NVIDIA planes. The architecture separates a **data plane** (personal data read, transformed, stored — always on hardware the owner controls) from a **control plane** (grants, receipts, the transparency log, placement — small, auditable, carrying capabilities and proofs, never content).

| LAYER | PERSONAL PLANE · SPARK / STATION / RTX | EDGE PLANE · GARAGE GRID / JETSON THOR | DATA CENTER PLANE · NVL72 / VR200 |
|---|---|---|---|
| AGENT | Resident Agent One — fetch, organize, guard, act; always-on | Fleet workers + physical actuators under scoped grants | No resident agent — jobs only, no custody |
| CONTROL | Grant store · receipt signer · transparency log (device-rooted) | spaceID node identity · metering agent · matched-book scheduler | Attestation verifier · settlement · teardown proofs |
| DATA | Vault + models on unified memory; nothing leaves un-granted | Envelope-encrypted job payloads only; no plaintext personal corpus | Ephemeral: decrypt inside TEE, compute, zeroize |

**The invariants (fixed before any integration)**

— **Locality.** Default site of computation is the owned device; moving data is the expensive, risky operation. Remote compute is the exception, in a tenant the owner governs — never a new custodian.

— **Legible consent.** Every access emits a signed, content-addressed receipt and an entry in an append-only, hash-chained log the owner can read in full. Revocation is a first-class, low-latency operation, not a support ticket.

— **Single-party sovereignty.** One principal, no second reader — not the vendor, the advertiser, the model provider, or us.

**WHY THE SPLIT MAPS CLEANLY ONTO NVIDIA**

NVIDIA's own factory architecture separates infrastructure processing (BlueField DPU domain) from tenant compute; NemoClaw separates the secure assistant substrate from the application; the Alliance separates harness from model. PCHP is the same instinct applied to personhood: the control plane that decides and proves is physically and cryptographically distinct from the plane that touches the data. Integration is therefore seam-matching, not surgery — §04–§08 walk each seam.

---

02

## PCHP at the Wire · SSH for humans, literally

SSH's contribution was a handshake so trustworthy that strangers' machines could speak safely. PCHP applies the pattern to personhood: **no agent touches data, compute, or an actuator without a handshake that authenticates, authorizes, and leaves a proof.** Six steps, every time:

```
1 IDENTIFY   caller presents workload identity (SVID / device key) + HusshID
principal
2 GRANT      control plane resolves an unexpired, sufficiently-scoped, un-
tombstoned Grant
3 ATTEST     execution environment presents evidence (Enclave / TEE / device) —
§03
4 EXECUTE    data plane performs the action; withheld fields never reach the
output
5 RECEIPT    ed25519-signed, content-addressed Receipt — action, fields, purpose,
hash, ts
6 LOG        Receipt appended to the hash-chained transparency log; chain verifies
O(n)
```

**The implemented types IMPLEMENTED · TESTED**

```
Grant   { grant_id, principal, source, scope[], expiry, revoked, sig }
Receipt { grant_id, action, fields[], purpose, content_hash, ts, sig }
LogEntry{ prev_hash, receipt, sig, hash = sha256(prev || canon(body)) }
# revocation writes a tombstone; the scheduler refuses any future action
# whose grant chain is tombstoned — and the refusal is itself logged.
```

Measured properties from the passing suite: tampering any receipt field breaks the signature; mutating any past log entry fails `verify_chain()`; withheld fields never reach the normalized picture; revoked or expired grants raise `ConsentError` before any fetch. A receipt proves a specific action over specific fields occurred — without disclosing the field values. Defaults are closed: a fresh device consents to nothing until the owner grants, item by item, in the Day-0 ceremony.

### Transport & surface

PCHP rides the agent ecosystem's lingua franca: the platform exposes a Model Context Protocol surface, so any MCP-capable agent — the owner's, or a trusted partner's — reads the catalog, places consent-gated actions, and checks status. Our OpenClaw MCP server (14 tools) is shipped and in the registry today; agent-to-agent flows and human-authorized payment mandates carry the same receipt discipline. **MCP SURFACE SHIPPED** — and the spec stays open: we monetize the network, never the protocol.

---

0 3

## Attestation-Bound Receipts · consent × confidential computing

The novel joint engineering: extend the receipt so it proves not only *what* was authorized, but *where* it ran — cryptographically, per job, across all three planes.

```
Receipt v1.1                        # PROPOSED JOINT EXTENSION — not yet
implemented
{ grant_id, action, fields[], purpose, content_hash, ts, sig,
  attest: { platform,               # secure_enclave | nv_cc | device_key
            measurement,            # enclave/TEE build measurement
            report_hash,            # hash of the raw attestation report
            verifier_sig } }        # control-plane verifier countersignature
```

### Flow on a burst job

— Before any ciphertext leaves the household, the control plane requests attestation evidence from the target — NVIDIA confidential computing report on Hopper/Blackwell-class GPUs in the data center; Secure Enclave assertion on-device; device-key assertion on edge nodes.

— The verifier checks measurement against an allow-list of known-good runtimes (the NemoClaw/One Engine build actually deployed), then — and only then — releases the job's ephemeral data key (§08).

— The completed job's receipt embeds the report hash and verifier countersignature: one record now answers the compliance question in full — *this job ran inside the enclave it claimed, on the fields it was granted, under this owner's consent, at this position in an unbroken chain.*

**WHAT NVIDIA'S STACK PROVIDES; WHAT WE ADD**

**Theirs:** hardware-rooted TEE and attestation on the GPU (models and data encrypted in use); BlueField-isolated infrastructure domain; OpenShell guardrails and the open NOOA harness for governing agent behavior. **Ours:** the ownership semantics above it — grants, receipts, revocation, and a log the *person* holds. **Joint:** the v1.1 binding plus NOOA-hook enforcement, so a harness physically cannot execute an un-granted tool call — grant checks move from convention to the harness layer. **PROPOSED — WORKSTREAM W2, §12**

Why this matters commercially as well as technically: it converts a stack of compliance PDFs into a query the owner can run — the enterprise version (§05 of "The Fourth Utility") of the household promise, and a capability no cloud-custody architecture can honestly offer.

---

04

## One Engine — Placement · where a job runs, and why

Placement is a scheduling decision, not dogma. The engine scores each task against five inputs and selects the lowest rung that satisfies all constraints — because every rung climbed is data moved, and moving data is the risk.

```
place(task):
  m  = model_footprint(task)          # weights+KV vs unified memory on each rung
```

```
l  = latency_class(task)            # interactive | human-time | batch
c  = consent_ceiling(task.grants)   # highest rung the grants permit
e  = energy_budget(device)          # thermal/battery state, always-on envelope
$  = unit_cost(rung)                # metered rate, matched-book only
for rung in [device, mac, spark_station, edge_batch, dc_surge]:
    if rung > c: break                        # consent is a hard ceiling
    if fits(m, rung) and meets(l, rung): return rung
return degrade(task)                # partial result beats silent escalation
```

| TASK CLASS | TYPICAL PLACEMENT | RATIONALE |
|---|---|---|
| Triage, wake, approvals | `iPhone (ANE + Secure Enclave)` | Identity root; consent surface; instant |
| Interactive inference, daily picture | `Spark / Station resident` | ≤200B-class local models in 128GB unified; sub-second turns |
| Overnight embed, index, fine-tune prep | `Edge batch (grid)` | Latency-tolerant; envelope-encrypted; cheapest honest watt |
| Model-scale train / trillion-class context | `DC surge (NVL-class)` | Provably exceeds device; TEE + teardown; §08 |

**Failure behavior — measured, not asserted** **IMPLEMENTED · TESTED**

The fetch stage talks to sources we do not control, so it runs a deadline-aware, tail-tolerant executor: deadline propagation, bounded concurrency, full-jitter backoff, per-source circuit breakers, graceful partials. Measured, from the suite: with one source deliberately hung 5.000s against a 1.000s stage deadline, the pipeline returned in **1.004s** with a correct picture from the 2 healthy sources, the hung source cleanly marked degraded, and the transparency log holding **exactly 2 entries — no receipt for data never accessed.** Sixteen tests across consent, log, pipeline, and executor: all passing. Bounded latency with honest degradation beats unbounded waiting, on every plane.

05

## The Personal Plane · the resident agent on Spark and Station

Above anything else, this. NVIDIA's own materials position Spark as a complete platform for local autonomous agents built for always-on

workloads, and ship NemoClaw in DGX OS for secure private assistants. The personal plane makes that hardware *somebody's*.

**Resident architecture — the Ultra runtime port** ROADMAP · WORKSTREAM W1

- **Substrate.** DGX OS + NemoClaw as the secure always-on base; One Engine and Agent One as a resident service above it — the same layered split as our shipping Apple-native build, one codepath apart (RHEL variant for enterprise racks).
- **Models.** FP4-quantized open weights within 128GB unified memory on Spark (NVIDIA rates the class to ≤200B parameters); Station's 748GB coherent memory admits trillion-parameter-class residents. BYOA holds: the owner brings any model; the consent fabric doesn't care whose weights think.
- **Serving.** TensorRT-LLM / NIM microservices for the local endpoints; every tool call passes the §02 handshake; withheld fields never enter a prompt.
- **Envelope.** Always-on means power honesty: the engine schedules heavy passes against the device's thermal and energy state (the `e` term in `place()`), and the PWS-1 meter (§11) makes the cost visible instead of vibes-based.

**Day 0, unchanged by heavier iron**

Every Ultra ships the White-Glove Human-First install: a person delivers it; the consent ceremony sets every permission to the owner's choice, item by item from closed defaults; the first fetch runs live in about a minute (§11's budget); and the install isn't done until the owner can drive it alone. Until the Linux runtime port ships, Ultra machines run under hussh-managed infrastructure — labeled honestly, with the full experience delivered the moment the port lands. The defining demo — one connected account to a full financial picture with a dollar figure attached, on-device — is **IN DEVELOPMENT · THE CRITICAL PATH**; we show it before we sell on it.

THE GTC DEMO, DEFINED PRECISELY

A family's Agent One, resident on a Spark on stage: Day-0 ceremony → live 60-second first fetch → a question answered from the assembled picture → the transparency log

read aloud, receipt by receipt → one burst job to an NVL-class rack with the v1.1 attested receipt shown on return. Ten minutes, zero slides, every claim executable.

---

06

## Personal Post-Training · intelligence per dollar, N = 1

NVIDIA's thesis: post-training is continuous, the central workload of the agentic era, and the driver of intelligence per dollar. The household is that loop at its smallest and most private — the model that serves one family, refined by that family's corrections, on silicon that family owns.

### The loop, engineered ROADMAP · WORKSTREAM W5

```
signal   owner corrections, approvals/refusals, task outcomes   (all receipted)
rollout  NeMo Gym environments replaying household task classes on-box
reward   verifiable outcomes first (did the renewal get caught? did the
         draft get accepted?) — never engagement, never attention
update   NeMo RL — LoRA-class adapters on an open base (Nemotron 3 Ultra
         class: open weights, post-trainable on proprietary data, local)
custody  adapters live in the Private Vault; base + adapters never leave
         hardware the family owns; every rollout emits a Receipt
```

— **Why adapters, not full fine-tunes:** Station-class memory admits large residents, but the loop's economics want small, frequent, reversible updates — an adapter per domain (money, health, home), each revocable like any grant, each auditable in the log.

— **Why verifiable rewards:** the reward channel is the value channel; tying it to outcomes the owner can verify keeps the loop aligned with the household, not with usage metrics. The transparency log doubles as the training ledger — what the model learned from is a query, not a mystery.

— **Where NVIDIA compounds it:** NeMo RL/Gym turn the loop into repeatable infrastructure rather than bespoke code; Vera-class CPUs' RL-sandbox throughput and Rubin's training efficiency lower the cost of every point of intelligence a family builds into their own agent.

> **Cloud AI asks the family to fund a model that serves everyone. Personal post-training funds a model that serves exactly one household — and proves it, receipt by receipt.**

07

## The Edge Plane · garage-grid node anatomy

The letter's north star, engineered: Puppy machines in garages and warehouses, linked over LEO transport, sited where power is cheap and the sun is abundant — owned by their hosts, metered by the handshake.

### Anatomy of a grid node

| COMPONENT | DESIGN |
|---|---|
| IDENTITY | One spaceID per node, keypair device-rooted; SPIFFE SVID for workload identity (§09) so any Alliance-grade tooling can verify the machine. |
| WORKLOADS | Envelope-encrypted job payloads only — **no plaintext personal corpus ever resides on a brokered node.** Job classes: batch embedding, indexing, overnight adapter training, render/sim. Interactive low-latency inference stays on the personal plane by design — LEO transport serves async and human-time work honestly; we route around what it can't do rather than pretend. |
| METERING | Each job's receipt carries units (tokens, joules, wall-seconds). Receipts are the usage records; usage records are the settlement input. One event, three functions. |
| SCHEDULING | Matched-book only: a node accepts work solely against committed demand the broker already holds. No naked inventory — the iron rule that keeps a grid from becoming a prop desk. |
| TELEMETRY | Thermal, power, and duty-cycle feed the PWS-1 meter and the siting model — a PhysicsNeMo/Omniverse-class thermal-solar twin of the garage is the siting tool (roadmap), which turns "where power is cheap and sun is abundant" into an optimization target. |

| | |
|---|---|
| ACTUATORS | Jetson Thor-class robotics compute joins the household fleet as an *actuator class*: a robot that touches the physical world acts only through scoped, revocable grants, every action receipted like any data access. Physical AI inherits the consent fabric; it does not get a bypass. |

Grid liquidity runs in closed loops we control both sides of — the first federated garage node, then one campus loop (idle cluster nights as supply, researcher burst as demand), then One power users, then the open market. Watt Score certification stays firewalled from the resale P&L (church/state); the catalog's honest 0/100 starting position fills from live routed jobs, never vendor claims. **GRID: ROADMAP · SEQUENCED AFTER THE CRITICAL PATH**

08

## The Data Center Plane · governed surge, keys never persisted

The rack is rented capability, never a custodian. A burst exists for exactly as long as its job, inside a tenant the owner governs, and leaves nothing behind but a receipt.

### Burst lifecycle — the full sequence DESIGNED · NOT YET WIRED (W4)

```
1 DECIDE     place() proves the task exceeds every lower rung          (§04)
2 GRANT      burst-scope grant resolved; crossing the boundary is itself
             a logged consent event
3 PROVISION  One Engine requests capacity via Dynamo-orchestrated pool;
             BlueField-isolated infra domain; CC-mode GPUs
4 ATTEST     TEE evidence verified against runtime allow-list             (§03)
5 KEYS       ephemeral job keypair minted; data key wrapped to the TEE's
             attested public key — decryptable only inside the enclave
6 EXECUTE    compute runs; partial telemetry only (units, health) egresses
7 RECEIPT    v1.1 receipt: fields, units, report_hash, verifier sig
8 TEARDOWN   session keys destroyed, KV/activations zeroized, capacity
             released; the teardown itself emits a final receipt.
             Nothing persists but ciphertext the owner holds + the log.
```

### What the household actually sends

Honest job shapes, not hypotheticals: corpus-scale re-embedding after a big grant change; a long-context synthesis exceeding resident memory; the quarterly adapter consolidation; a practice's multi-client batch under each client's own grants. Rack-scale exists because these

are real — and NVIDIA's delivered-FLOPs discipline (world-record MoE pre-training on GB300 NVL72; NVLink-5's 1.8 TB/s per GPU; 97%+ scaling efficiency to 1,024 GPUs; throughput still compounding from software alone) is why the surge tier's unit economics keep improving on silicon already in the fleet.

**DYNAMO SEAM, PRECISELY**

One Engine is a Dynamo *client*, not a rival scheduler: Dynamo owns intra-pool orchestration; we own the consent ceiling, the attestation gate, and settlement. The integration surface is deliberately thin — request shape, attestation exchange, usage telemetry — so the seam survives both roadmaps. **PROPOSED · WORKSTREAM W4**

---

09

## The Identity Fabric · HusshID × spaceID × SPIFFE

Three identities, three jobs, one chain of custody for every action in the system.

| IDENTITY | ANCHORS | ANSWERS |
|---|---|---|
| `HusshID` | The human principal — phone number as the natural key; Secure Enclave device key as the cryptographic root | *Whose* data, *whose* grant, *whose* log. The owner of record, 24/7/365. |
| `spaceID` | The machine — one per node (a phone, a Spark, a garage rack), device-rooted keypair | *Where* a job ran; the grid's unit of certification, metering, and reputation. |
| `SPIFFE SVID` | The workload — short-lived verifiable identity for the process actually executing | *What code* acted — verifiable by any zero-trust tooling, not just ours. |

Every receipt binds all three: principal (HusshID), node (spaceID), workload (SVID), plus the attestation evidence of §03. That quadruple is the full answer to

the only question that matters in agentic security — **who authorized what code to do what, where, to whose data** — as a single signed record.

THE ALLIANCE ALIGNMENT, MADE CONCRETE

The Open Secure AI Alliance's stated frame — identity, permissions, harnesses, guardrails, logs, evaluation — maps term-for-term: identity = this fabric (with spaceID→SPIFFE interop so our nodes verify under anyone's tooling); permissions = grants; harness enforcement = NOOA hooks (W2); logs = the transparency chain; evaluation = the measured suite and PWS-1. Our contribution application proposes PCHP as the open consent-and-receipt layer of that stack — the piece every member needs and none has shipped. **APPLICATION — THIS QUARTER (W3)**

Design note: the phone number is the *natural* key because it is the key the world already banks personal data against — the point is to invert who controls it, not to invent a new namespace nobody indexes on.

---

10

## Threat Model, Three Planes

| THREAT | PLANE | POSTURE |
|---|---|---|
| Central breach of hussh | ALL | **Mitigated by construction.** No central corpus exists; compromising the platform yields capabilities and hashes, not anyone's personal data. |
| Silent access | ALL | **Detectable.** Every read is grant-gated and hash-chain logged; tampering any past entry invalidates every subsequent hash. |
| Custody creep | ALL | **Prevented.** No second reader; partner and model agents act through scoped, revocable grants and never receive a durable copy. |
| Device loss | PERSONAL | **Contained.** Keys in the Secure Enclave; data encrypted at rest; the lost node is revoked and its grants tombstoned from the control plane. |

| | | |
|---|---|---|
| Rogue or curious edge host | EDGE | **Bounded by design.** Brokered nodes receive envelope-encrypted job payloads only — never a personal corpus; identity and attestation gate acceptance; misbehavior burns the node's spaceID reputation and its certification. |
| Metering / billing fraud | EDGE · DC | **Signed at the source.** The usage record *is* the signed receipt; settlement inputs inherit receipt integrity, and disputes replay the chain. |
| Exfiltration during burst | DC | **TEE-bounded.** Data keys wrapped to attested enclaves only; zeroize-on-teardown; the teardown receipt is the proof of erasure the owner keeps. |
| Supply-chain compromise of our own binaries | ALL | **In development.** Signed, attestable provenance for agent components, aligned with the Alliance's signed-patch and safe-format direction; the runtime allow-list (§03) refuses unrecognized builds. |

**Not claimed — stated so trust is earnable**

— An attacker with persistent code execution on an *unlocked* owner device is outside the model; we reduce blast radius (Enclave keys, least-scope grants) but claim no immunity.

— FedRAMP High authorization is in pursuit and never stated as held.

— The 60-second live first-fetch against real sources — the demo that defines the product — is in active development; we show it before we sell on it. Honesty about the boundary is part of the specification.

11

## Metrics & SLOs · delivered, not peak

PWS-1 — THE PUPPY WATT SCORE, FORMALLY

`PWS-1 = geomean( tokens/joule , tokens/TCO-$-hr )` · TCO over 60-month depreciation at $0.169/kWh reference power · measured on live production jobs routed through the broker, never vendor-claimed · certification firewalled from resale (church/state). It is cost-per-token and intelligence-per-dollar's sibling, read at the

household meter — the same delivered-not-peak philosophy NVIDIA's own record posts are measured by.

### The 60-second first-fetch budget (TARGET — instrumented, so the number is measured, not asserted)

| STAGE | BUDGET | ENGINEERING NOTE |
|---|---|---|
| OAuth consent + token | ≤ 8 s | Human-paced; one tap, scoped read grant |
| Source enumeration | ≤ 6 s | Identify financial senders/domains, 12 mo |
| Fetch + parse | ≤ 25 s | Parallel connectors; incremental parse (dominant) |
| Normalize → picture | ≤ 8 s | Accounts, balances, recurring, renewals |
| Local inference | ≤ 10 s | On-device model; dollar-denominated insight |
| Render + receipts | ≤ 3 s | Picture view + transparency-log writes |
| **TOTAL** | **≤ 60 s** | Tail target; p50 materially below; bursts a step only if over budget |

### Handshake overhead (design targets)

The consent fabric must be invisible at human timescales: grant resolution and receipt signing are budgeted to sub-perceptual overhead per action, log appends are asynchronous off the hot path, and chain verification is an owner-initiated O(n) audit, not a per-action cost. These are TARGET figures under instrumentation — published as measurements only when they are measurements, in the same discipline as everything else on this page. What is already measured: the executor's bounded-latency behavior (§04) and the integrity properties of the receipt/log suite (§02) — sixteen tests, passing.

---

12

## The Engineering Program · every milestone is a demo

| # | WORKSTREAM | ACCEPTANCE DEMO — WATCHABLE, NOT SLIDEWARE |
|---|---|---|

| | | |
|---|---|---|
| W1 | Ultra runtime port — Agent One on DGX OS / NemoClaw | Day-0 ceremony + live first-fetch, resident on a Spark |
| W2 | PCHP × OpenShell/NOOA + attestation binding (v1.1) | Un-granted tool call physically refused at the harness; burst receipt verifies against TEE report |
| W3 | Alliance application + spaceID→SPIFFE interop | hussh node verified end-to-end by third-party SPIFFE tooling |
| W4 | Burst boundary on Dynamo | Full 8-step lifecycle (§08) with teardown receipt, live |
| W5 | Personal post-training recipe — NeMo RL on Station | Adapter trained from receipted household rollouts; log replayed as the training ledger |
| W6 | Subscription-rail pilot | One metered burst SKU on one partner statement, end to end |

Sequencing is governed by the critical path — the live 60-second first-fetch — and the operating gates (G0 first-dollar unblock; G1 matched-book until brokerage GMV proven; G2 runtime attach before channel scale; G3 external PCHP behind its review gate; G4 market-data only after liquidity). The ask is unchanged from the letter, now with an engineering shape: **a joint pod on W1–W2, a GTC demo slot for the §05 demonstration, Alliance sponsorship for W3 — one conversation, one pilot, one garage.**

### Sources

**NVIDIA primary materials (July 2026; retrieved Aug 3, 2026):** Open Secure AI Alliance (blogs, Jul 27) · Vera Rubin intelligence-per-dollar post-training (blogs, Jul 17) · Agent Toolkit + PhysicsNeMo/CUDA-X (newsroom, Jul 26; Omniverse libraries Jul 20; Jetson Thor Jul 15) · GB300 NVL72 MoE world record (developer blog, Jul 21) · DGX Spark product page (NemoClaw, always-on local agents; retrieved Aug 3).

**hussh canon:** Personal Supercomputing Infrastructure — Technical Specifications (rev 2026-06-12, including all measured results and implemented types cited herein) · One Puppy Partner Master Pack · the open letter (hushh.ai/one/nvidia) · "The Fourth Utility" (rev 2026-08-03) · compute flywheel & Puppy 100 canon (wiki, Jul 2026).

owners. Hushh Technologies Corporation (brand: "hussh") is independent and not affiliated with, endorsed by, sponsored by, or partnered with any company named here; we name a partner only once an agreement is executed.  Revision 2026-08-03.