ARCNet: Building the National Nervous System for Industrial AI

A federated, exascale architecture for autonomous discovery and manufacturing.

ARCNet

The ARCNet Vision: A Closed-Loop Infrastructure for Sensing, Reasoning, and Action

The Brain: Exascale HPC

ORNL Frontier & Lux

The cognitive core for training massive-scale AI models, running high-fidelity simulations, and developing global optimization strategies.

The Nervous System: Federated AI Network

~1000 x 1MW Nodes

A distributed network of modular AI data centers for high-performance inference, localized training, and regionaal data processing.

Senses & Actuators: Edge Devices

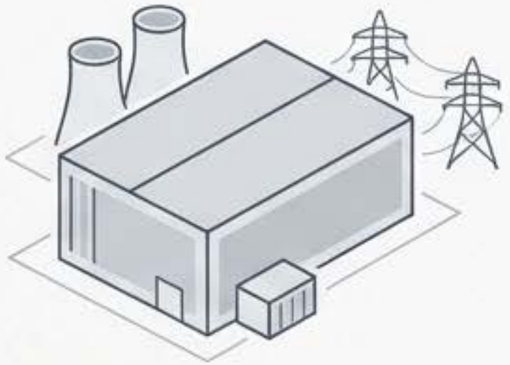
ADAM Platform

Autonomous machines that execute tasks, conduct experiments, and feed real-world data back into the system.

Like Ethernet standardized data exchange, **ARCNet** provides the universal protocol linking HPC, distributed compute, and edge devices for autonomous operations.

Our Strategic Pivot: From Centralized Hyperscale to a More Resilient, Agile Federated Model

To accelerate time-to-market and mitigate critical deployment risks, ARCNet has pivoted from an initial plan for a 1 GW hyperscale build-out to a more agile federated architecture. This decision is based on a rigorous analysis of technical, economic, and logistical realities facing large-scale infrastructure projects today.

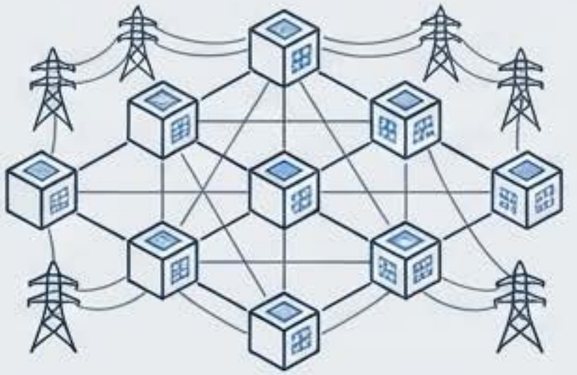


The Original Plan (Hyperscale)

10 x 100 MW hyperscale data centers.

Key Challenges

- Long deployment timelines (18-36 months per site)
- Severe grid interconnection bottlenecks
- Concentrated points of failure



The Strategic Approach (Federated)

~1000 x 1 MW modular data centers.

Key Advantages

- Rapid, parallel deployment (6-12 months per node)
- Flexibility in power sourcing and grid connection
- Inherent resilience and fault tolerance

The Federated Model De-Risks and Accelerates Our Path to 1 GW of Capacity

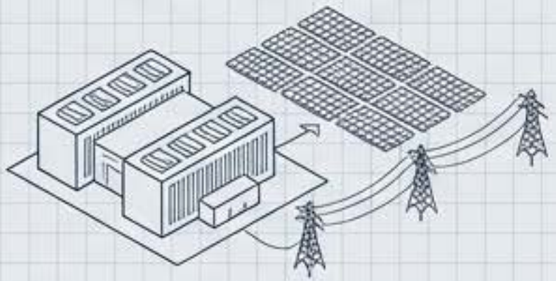
Aspect	10×100 MW Hyperscale Model	1000×1 MW Federated Model
Grid Interconnection	**Major Bottleneck.** Requires new substations and high-voltage lines. **Multi-year queue waits are common** (2-5 years in Northern Virginia). Risk of completed data centers sitting idle for power, as seen in Santa Clara (~100 MW unenergized).	**Highly Flexible.** Can tie into existing distribution grids. Shorter interconnect process (months, not years). Exploits pockets of available power, avoiding regional grid strain.
Deployment Speed	**Slow & Sequential.** Each facility takes **18-24+ months** to build and energize. Total build-out could take years.	**Fast & Parallel.** Modular 1 MW units can be installed in **6-12 months each.** Hundreds can be deployed concurrently, delivering capacity incrementally.
Redundancy & Resilience	**Low.** A single site outage removes **100 MW (10% of total capacity)**. Vulnerable to regional disasters or grid failures.	**Extremely High.** A single node failure removes only **0.1% of capacity**. Geographic diversity protects against regional outages, creating a self-healing network.
Performance Trade-off	Excellent for tightly-coupled HPC workloads within a single campus.	WAN latency impacts cross-site jobs. Mitigated by intelligent scheduling that co-locates tasks and data, making it ideal for distributed AI inference and asynchronous workloads.

Unlocking Significant Economic Advantages Through a Distributed Footprint

30% Federal Investment Tax Credit (ITC) on Renewables

The federated model is designed to maximize economic benefits. By keeping each site's solar and storage project under the 1 MW AC threshold, we automatically qualify for the full 30% ITC without needing to meet complex prevailing wage and apprenticeship requirements that apply to larger projects.

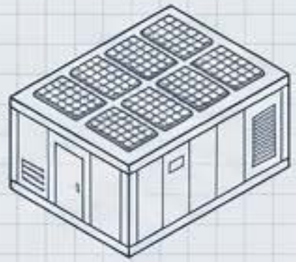
Hyperscale



Projects >1 MW must meet stringent labor rules for 30% ITC, otherwise they receive only a **6% base credit.**

Across 1000 sites, this translates to a potential ~$600M in tax savings, significantly offsetting the base capital expense and improving the ROI for each node.

Federated



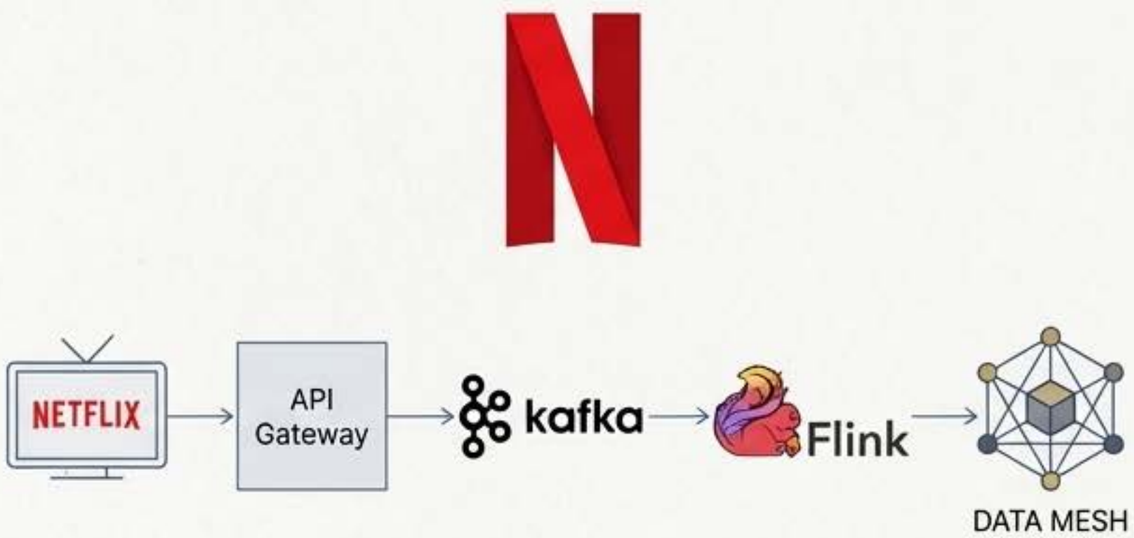
Projects <1 MW **automatically qualify for the 30% ITC.**

This Architectural Pattern is Proven at the Highest Scale: The Netflix Precedent

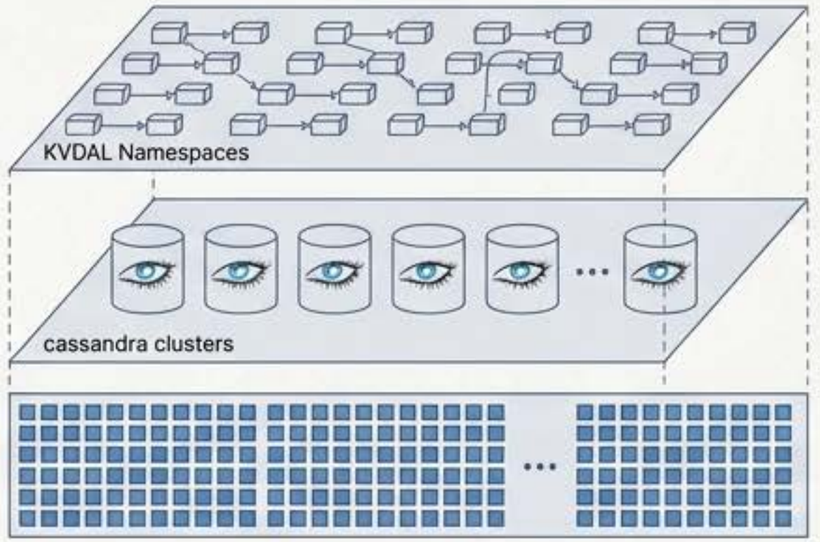
Netflix faced a similar challenge: unifying data from hundreds of siloed microservices to understand user behavior in real-time. Instead of a monolithic database, they built a Real-Time Distributed Graph (RDG) on a flexible, scalable key-value storage system (Cassandra). This approach allowed them to scale to billions of entities and handle millions of events per second.

Key Parallel

ARCNet's federated model follows the same core principle: leveraging a distributed, scalable architecture to achieve performance and flexibility that a monolithic system cannot. We are building a "distributed supercomputer" just as Netflix built a "distributed graph."



Proof Points: Impressive Scale Metrics



~27 KVDAL Namespaces

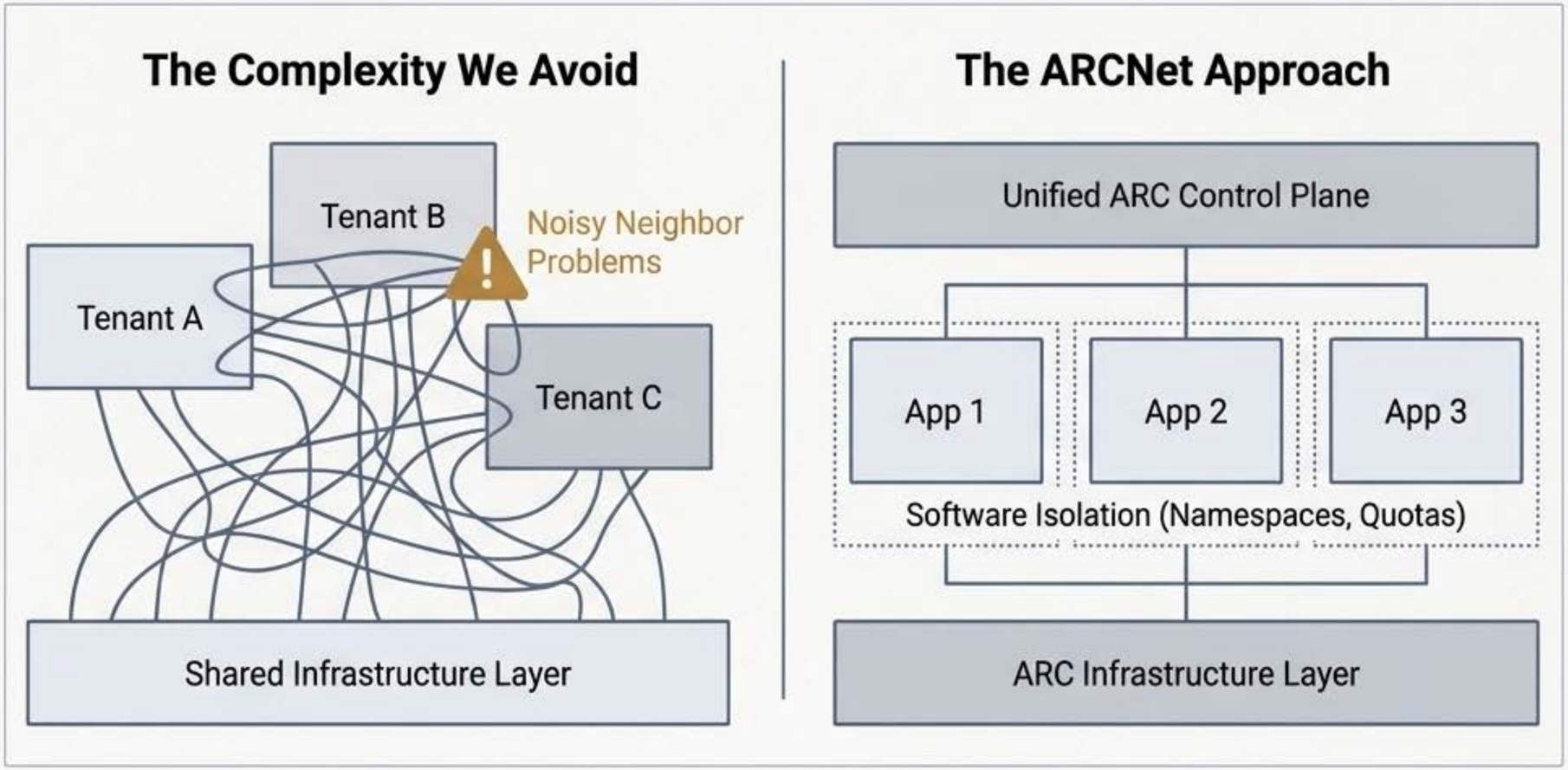
~12 Cassandra Clusters

~2,400 EC2 Instances

Supporting 8B+ nodes and 150B+ edges.

A Simpler, More Powerful Operating Model: Single-Tenant, Multi-Application

We eliminate an entire class of operational complexity. **ARCNet** is not a multi-tenant IaaS platform like AWS. Instead, ARC operates the entire infrastructure as a **single, cohesive unit**. Users bring their applications and workflows; ARC provides the optimized execution environment, model selection, and placement.



Key Benefits:

- **No "noisy neighbor"** resource contention.
- **Global optimization** of model placement, caching, and data replication.
- **Simplified** security, governance, and auditability.
- ARC manages **model selection**, routing users to the best model for the task (cost/latency/quality).

The ARCNet Node: A Replicable 1 MW ‘AI Factory’ Blueprint

Each node is a standardized, self-contained AI/HPC cluster designed for rapid, repeatable deployment and high efficiency.

Compute:
1x ~1MW rack-scale NVIDIA Rubin system.

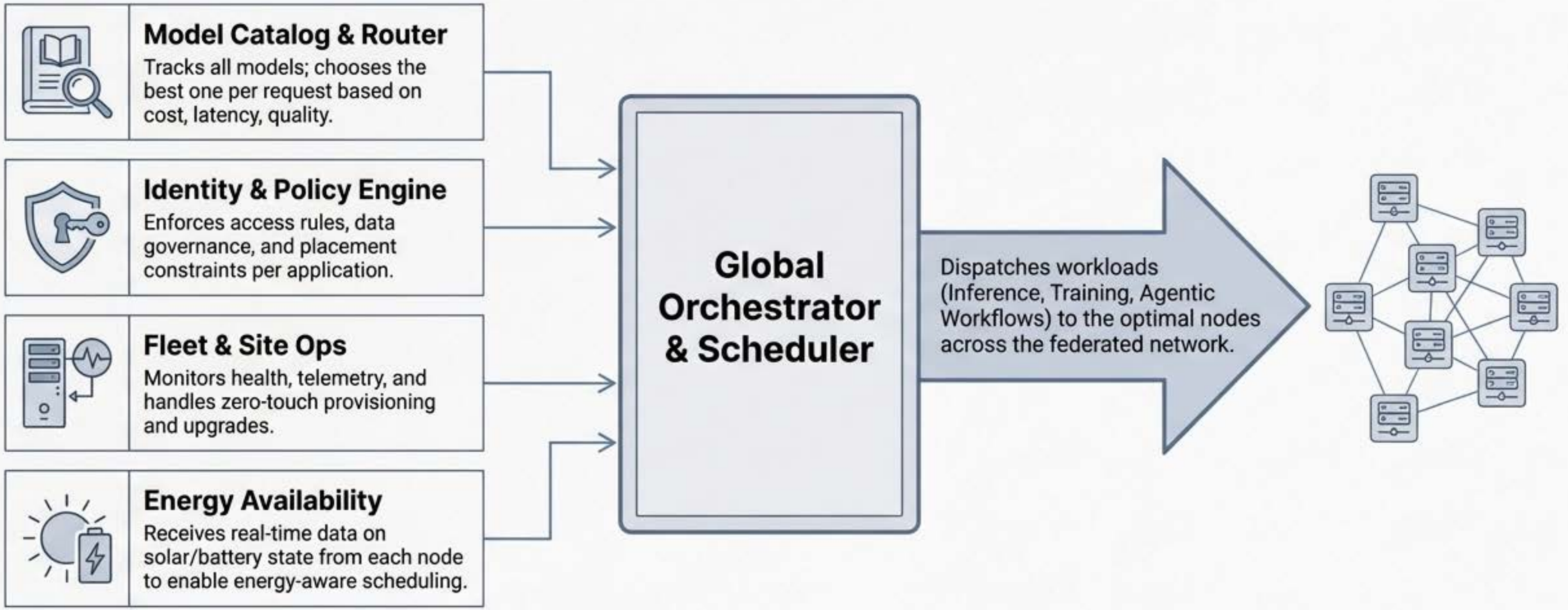
Energy Storage:
Multi-layer buffering. Rack-level ultracaps handle millisecond spikes, while site-level BESS (e.g., Yotta Blocks) provides ride-through and smoothing.

Power:
End-to-end 800V DC backbone. Facility 13.8kV AC is rectified to 800V DC and fed directly to the rack, minimizing conversion losses.

Cooling:
High-efficiency liquid cooling. Designed for 45°C warm-water loops, reducing energy overhead.

The Global Control Plane: ARCNet's 'Operating System'

The control plane is the global intelligence layer that **unifies** the federated nodes. It decides what runs where, on which model, with which data, under what policy, and logs everything for a complete audit trail.



A Phased Rollout to Mitigate Risk and Scale Intelligently

Phase 0: Reference Stack

Node Count
2 Lab Nodes

Goal
Harden the software platform.

Key Activities
- Implement core control plane services (model catalog, orchestration, audit logging).
- Validate reproducible workflows and remote fleet management.

Phase 1: MVP Federation

Node Count
10-20 Nodes

Goal
Prove distributed viability across 2-3 geographic regions.

Key Activities
- Launch inference-first services.
- Offer fine-tuning as a managed, single-site service.
- Deploy initial agentic workflows.
- Achieve "N-1 site loss" tolerance within each region.

Phase 2: Regional Clusters

Node Count
50-100 Nodes

Goal
Densify regions and enhance automation.

Key Activities
- Introduce automated dataset replication policies.
- A learned model router.
- Canary deployments for model updates.

Phase 3: Network Effects

Node Count
200+ Nodes

Goal
Operate as a national-scale utility.

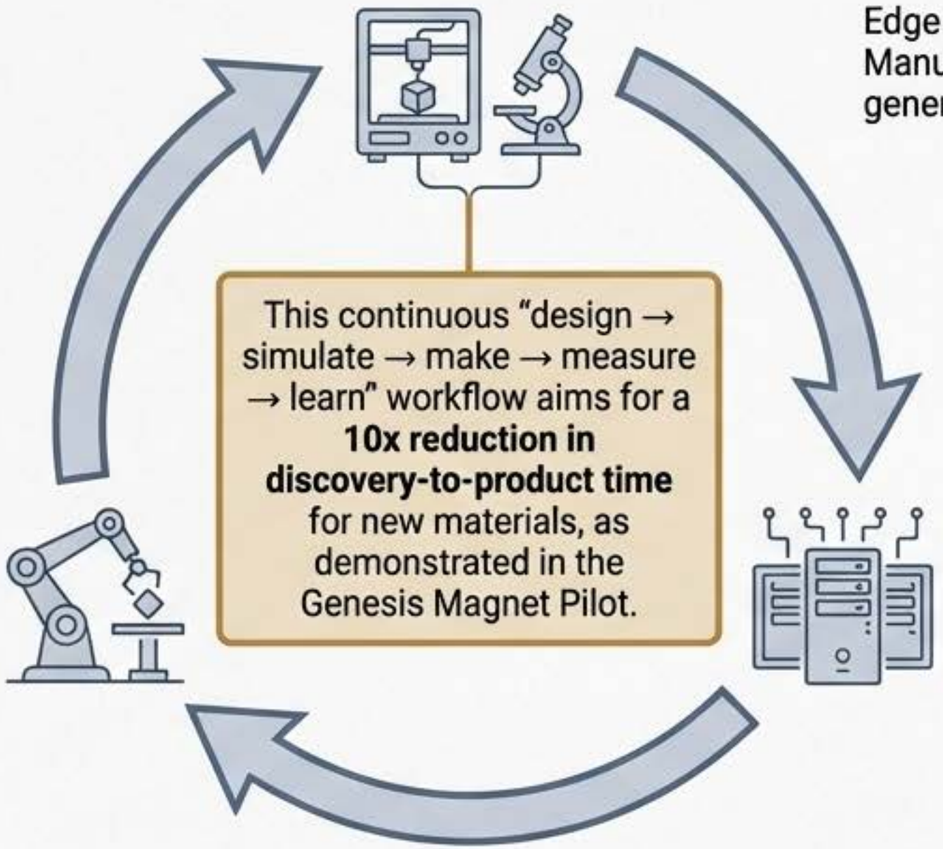
Key Activities
- Treat the network as an internal compute marketplace, routing jobs to optimal locations for energy, cost, and data locality.
- Introduce specialist nodes (e.g., long-context inference clusters).

The Decisive Advantage: Enabling Closed-Loop Autonomy at Scale

ARCNet is more than a compute provider; it is an engine for accelerating discovery. By tightly integrating sensing, reasoning, and acting, our platform creates a virtuous cycle where every operation makes the entire system smarter.

SENSE

Edge instruments at ORNL's Manufacturing Demonstration Facility generate data on a new printed alloy.



This continuous "design → simulate → make → measure → learn" workflow aims for a **10x reduction in discovery-to-product time** for new materials, as demonstrated in the Genesis Magnet Pilot.

REASON

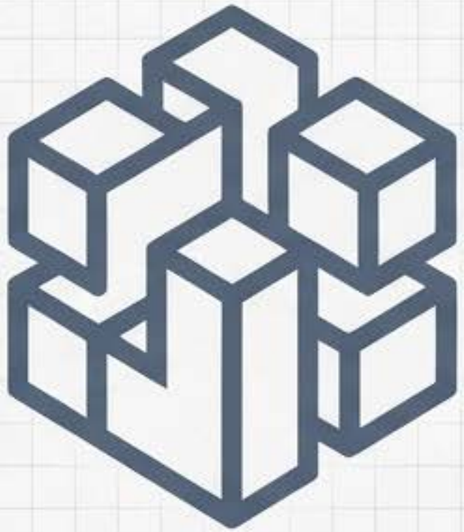
Frontier/Lux trains an evolving materials model on the aggregated data from all experiments, discovering optimal new designs.

ACT

The ADAM orchestrator dispatches the next experiment back to the the edge device, instructing it to fabricate and test the new design.

The Foundation of Trust: An Open, Universal, and Auditable Platform

We are building ARCNet as a neutral, public-benefit utility for AI-driven autonomy. Our commitment to an open architecture and a complete, unalterable audit trail ensures transparency and control for our partners.



Open & Model-Agnostic

Supports any AI model or agent framework. We define a universal protocol to connect intelligence to the physical world, not a walled garden.

Always-On Provenance

The 'trust layer' is a core primitive. Every workflow execution—from inference requests to multi-step agentic tasks—produces a complete, reproducible audit trail (datasets, model versions, parameters, outputs).

Secure by Design

Built from Day 1 for export control compliance and governed data flow, enabling secure collaboration between government, national labs, and industry on sensitive projects.

ARCNet is the Foundational Protocol for the Physical AI Economy