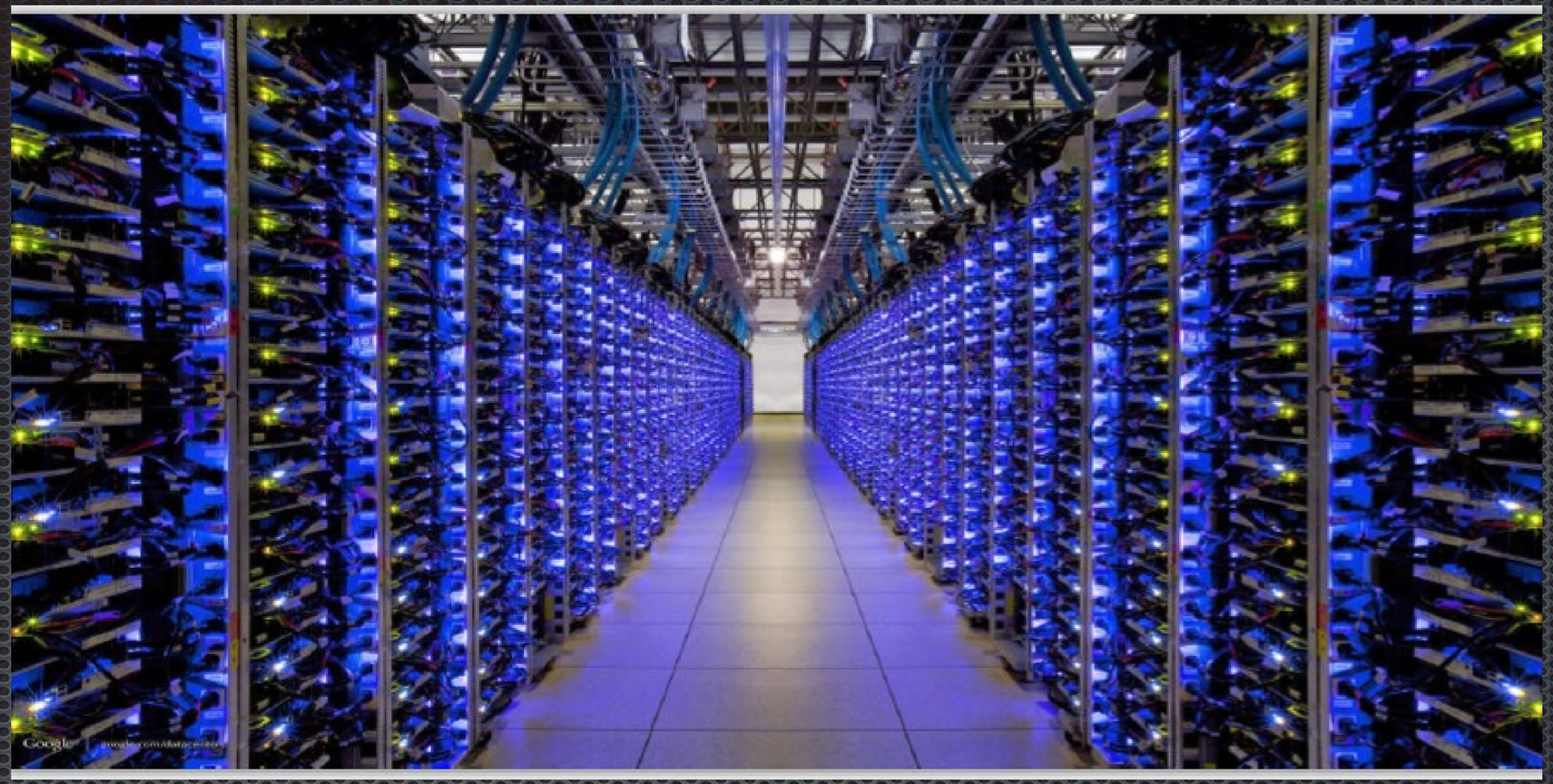


Self Service Infrastructure Environment for Next Generation High Performance Test and Evaluation

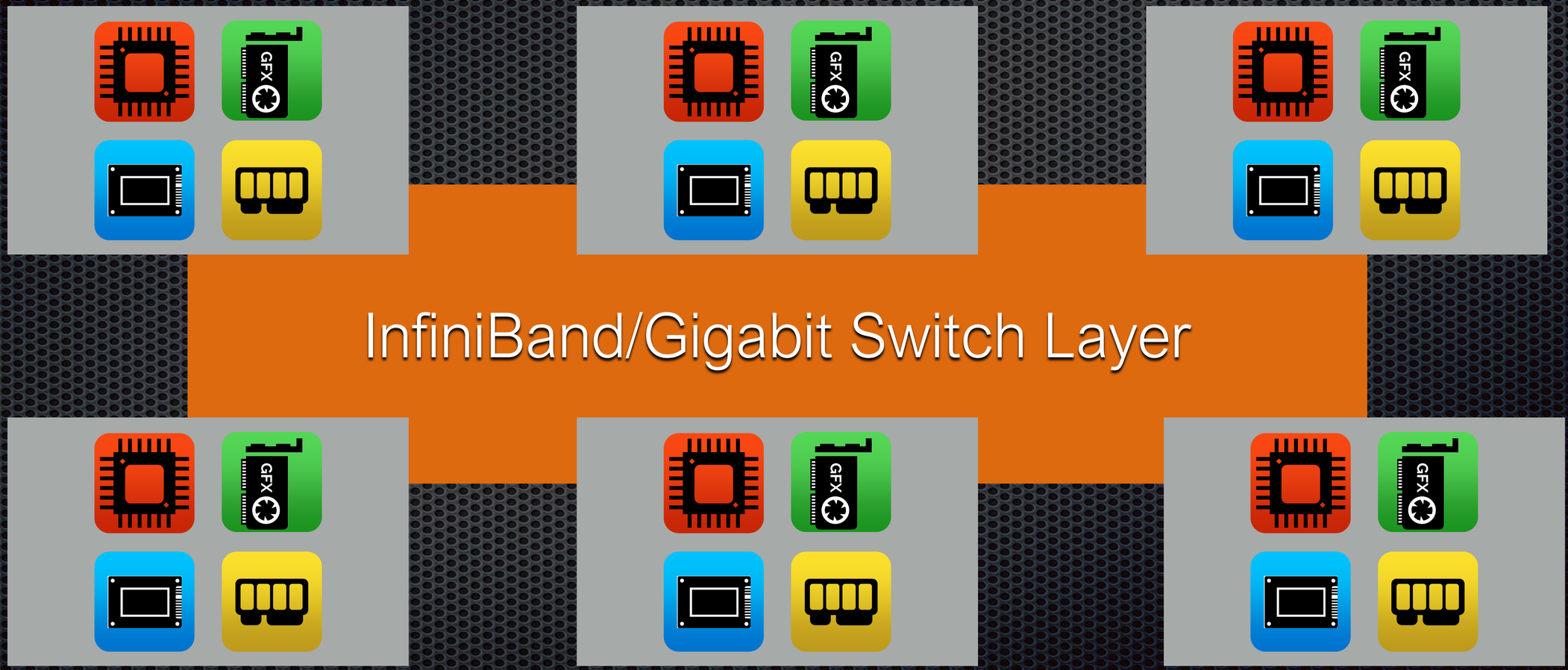
Agenda

Introduction
Charles (Chuck) T. Reynolds CTO TSI

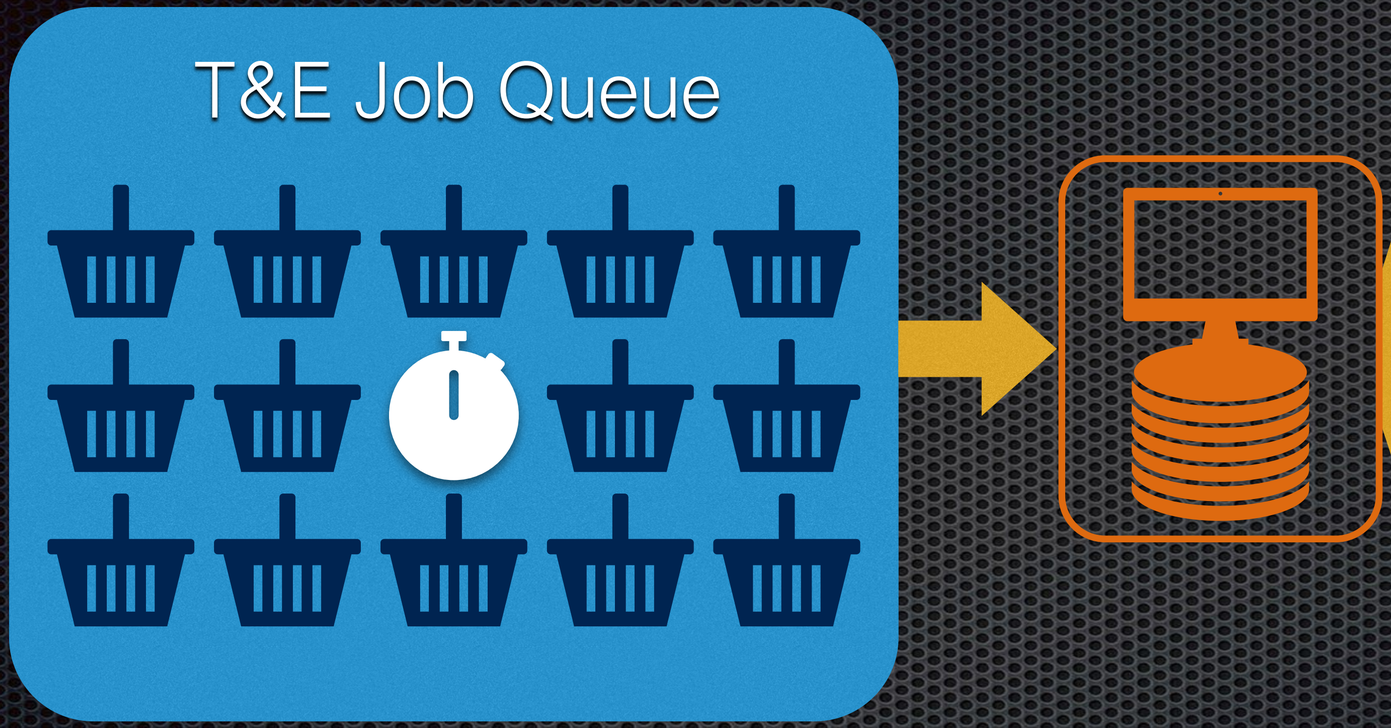
Resource Allocation Challenges
Network Fabric Performance
Self Service Portal for Infrastructure
Solution: TECaaS

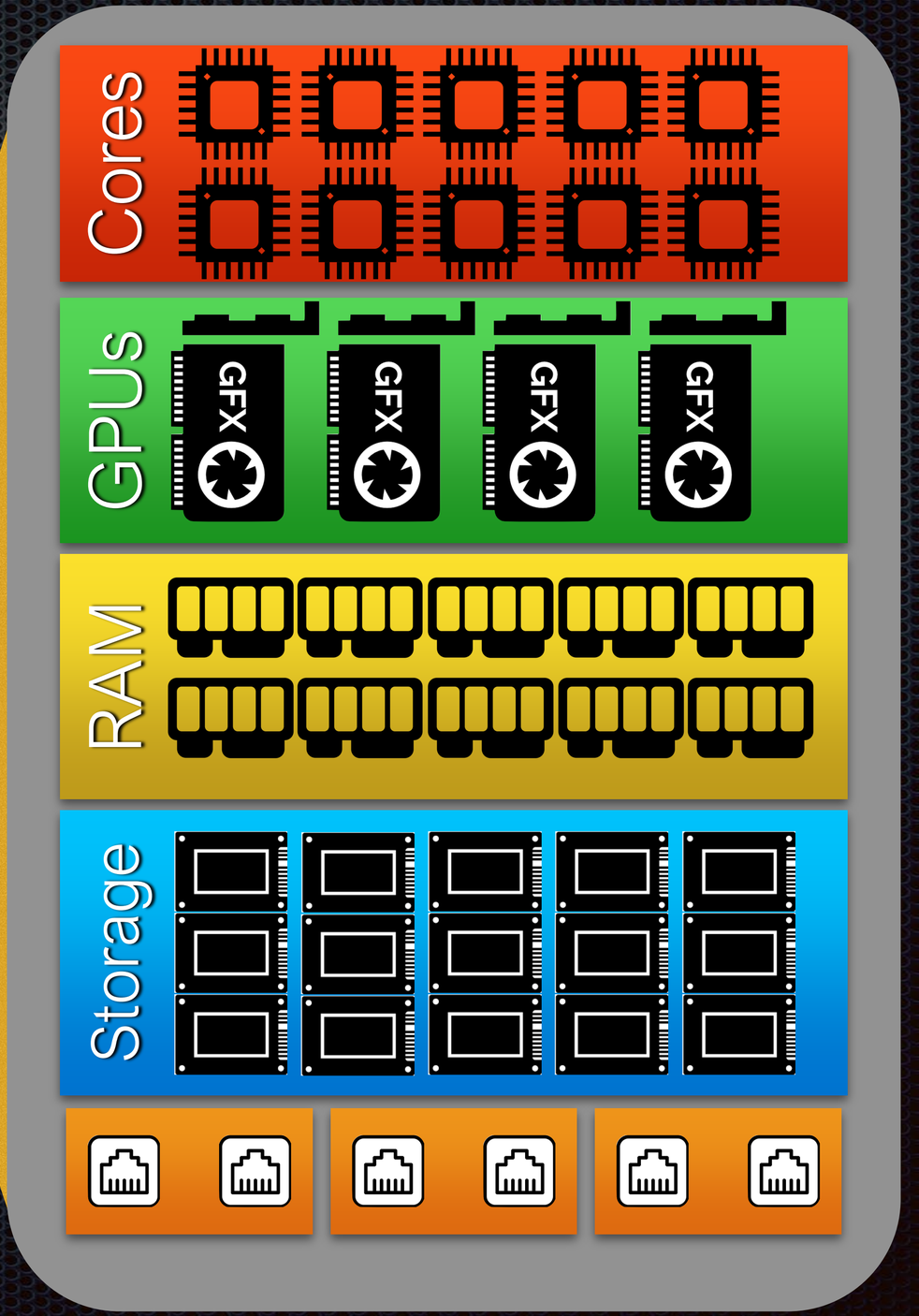
Challenge One: Resource Allocation

Connecting HPC Today

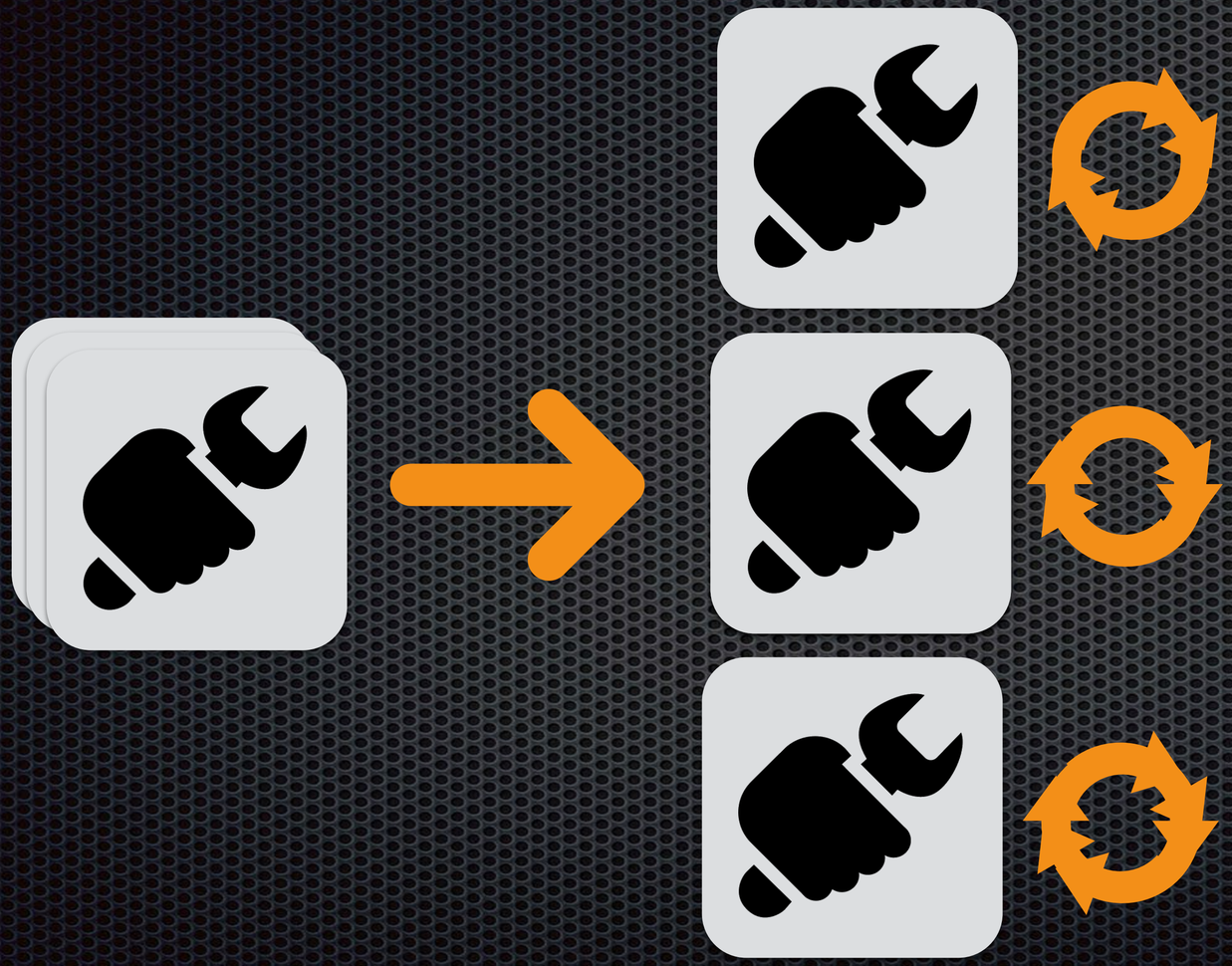


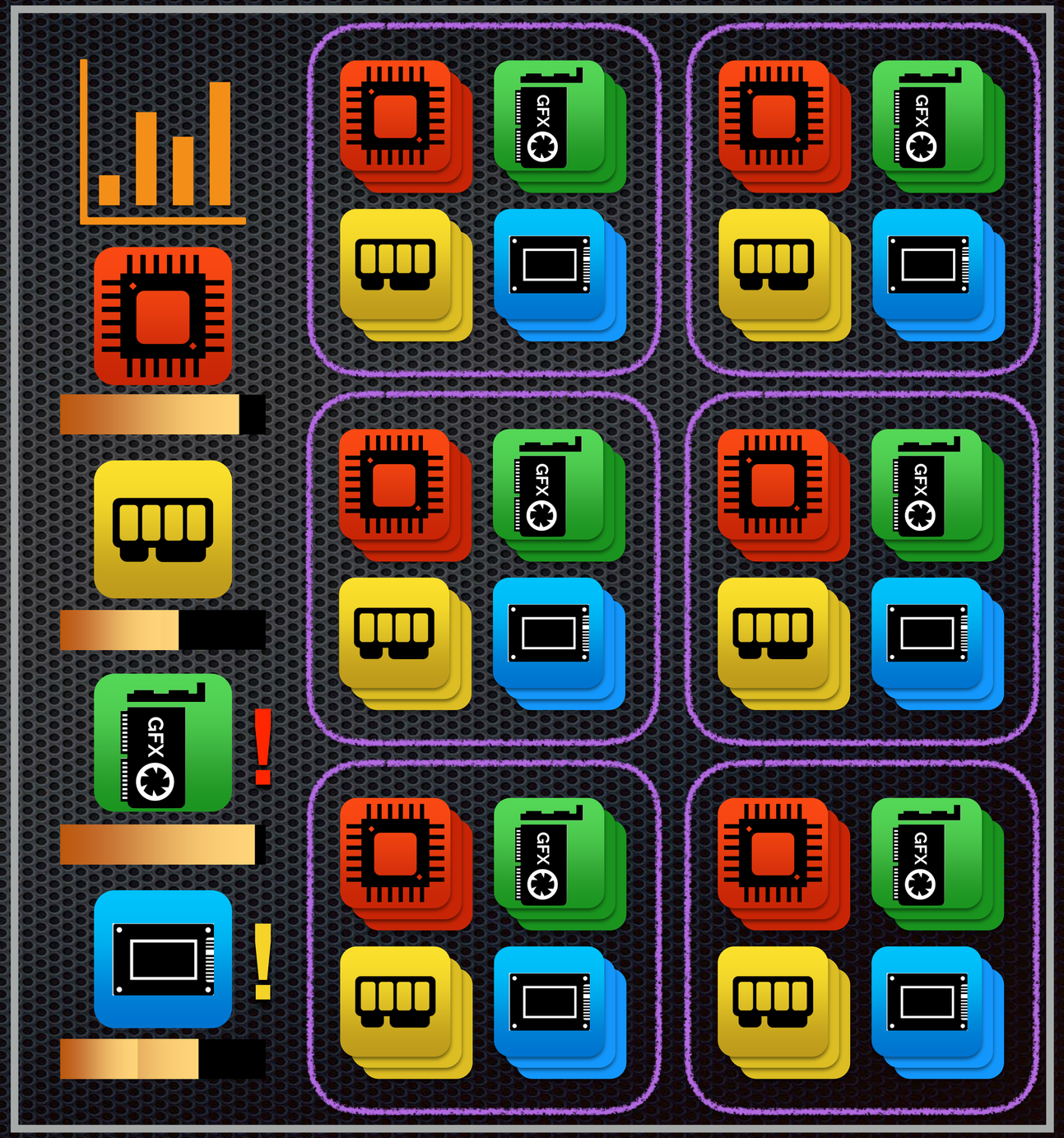
Infrastructures Are Currently One-Size-Fits All



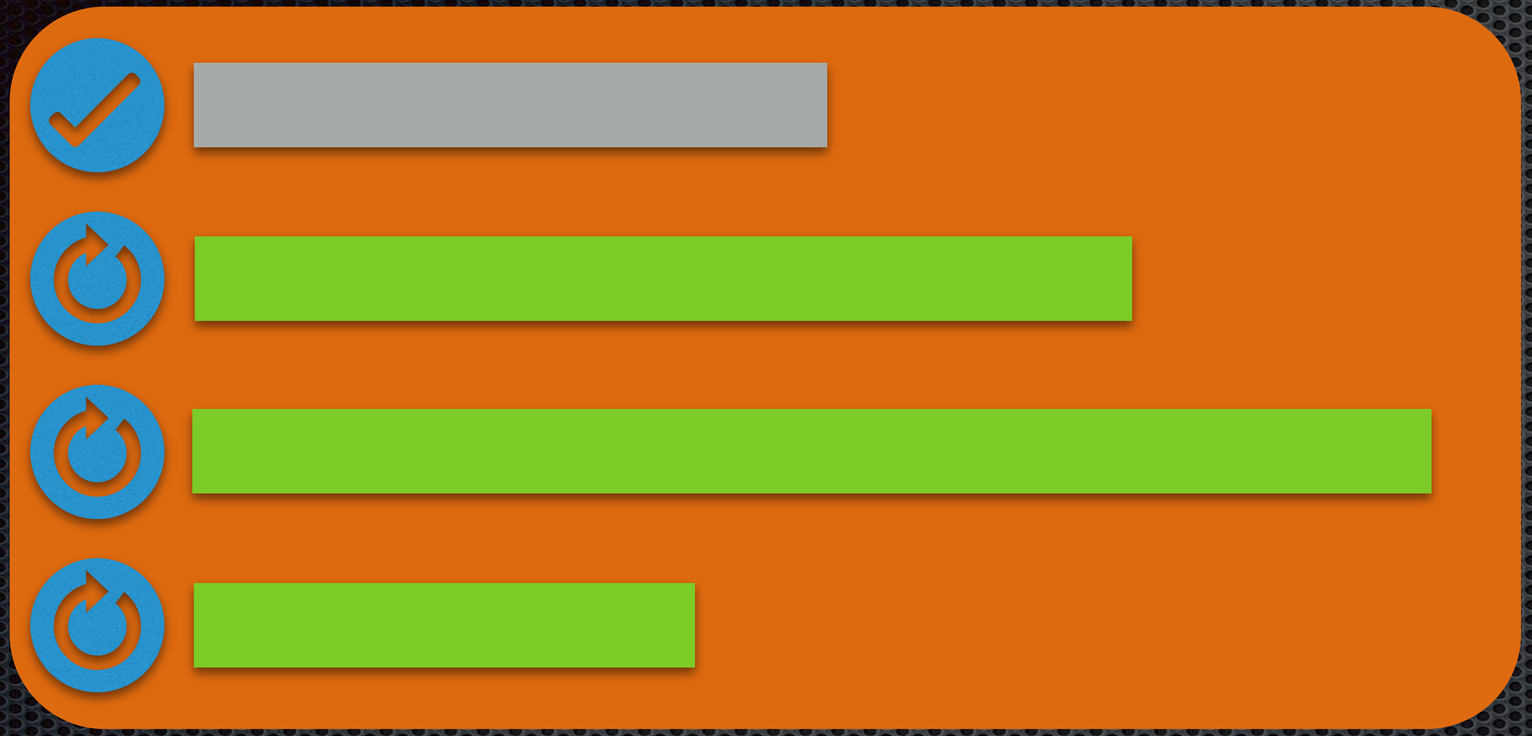


Other times the jobs aren't balanced, resulting in constraints & idle resources





Different jobs - same hardware
Little or no optimization for different types of jobs



- Jobs compete for constrained resource
- Resources left idle
- Inconsistent job times
- Single job queue
- Infrastructure not optimized for the job

16 Workflows in the queue

Disaggregated Composable Infrastructure

"Intel® Rack Scale Design is an industry-wide architecture for disaggregated, composable infrastructure that fundamentally changes the way a data center is built, managed, and expanded over time, which enables service providers to cost-effectively accelerate data center innovations,"

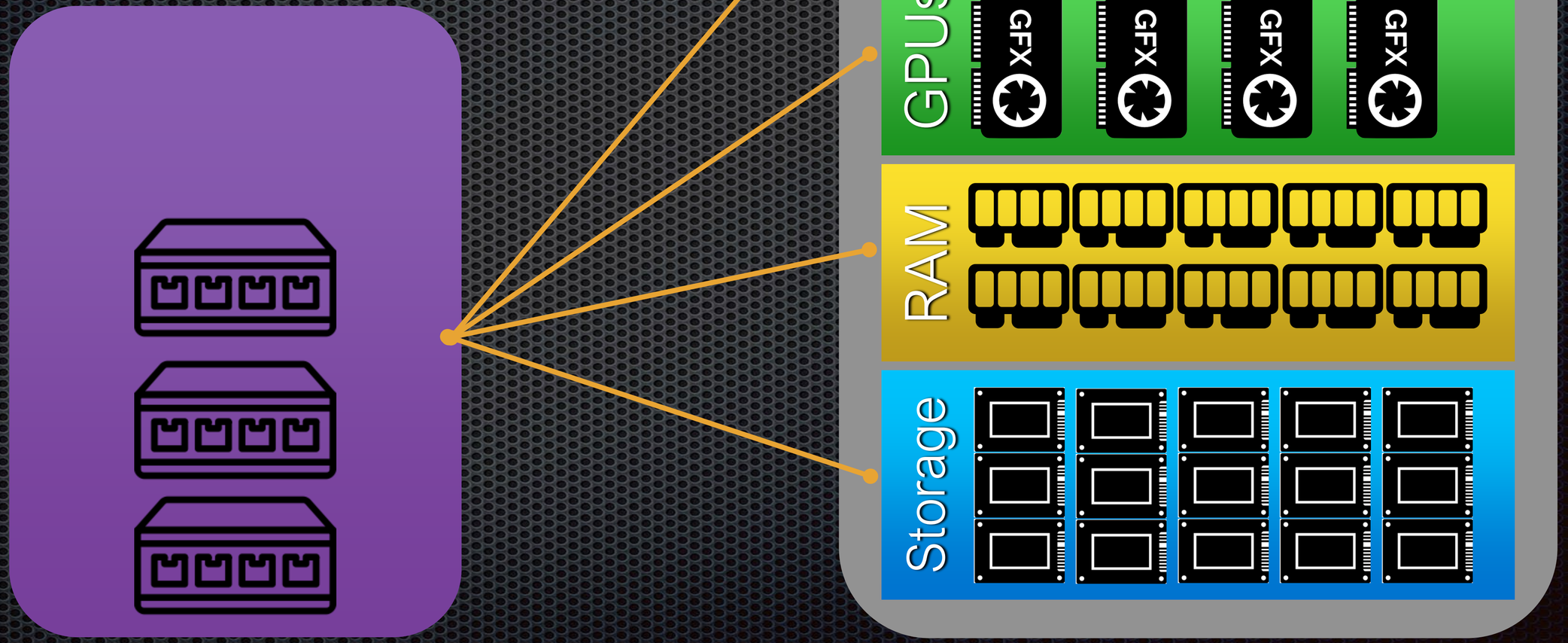
--Figen Ulgen, GM, Intel® Rack Scale Design

"As good as disaggregation sounds, we must not forget using discrete resource pools means developing interconnections that were previously contained on the server motherboard. Is that even possible?"

-- *DataCenterDynamics*

PCIe Fabric

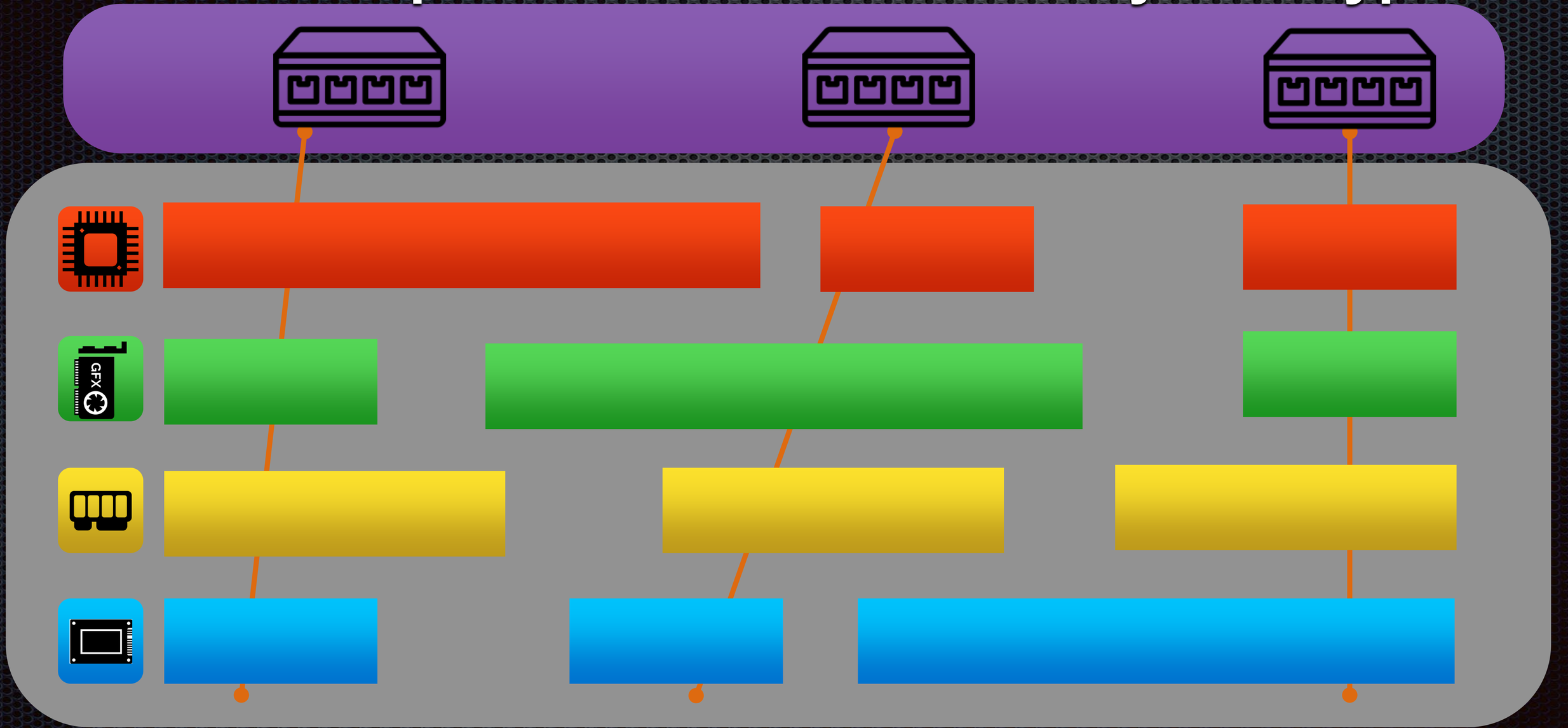
Provides Disaggregated Composable Connections



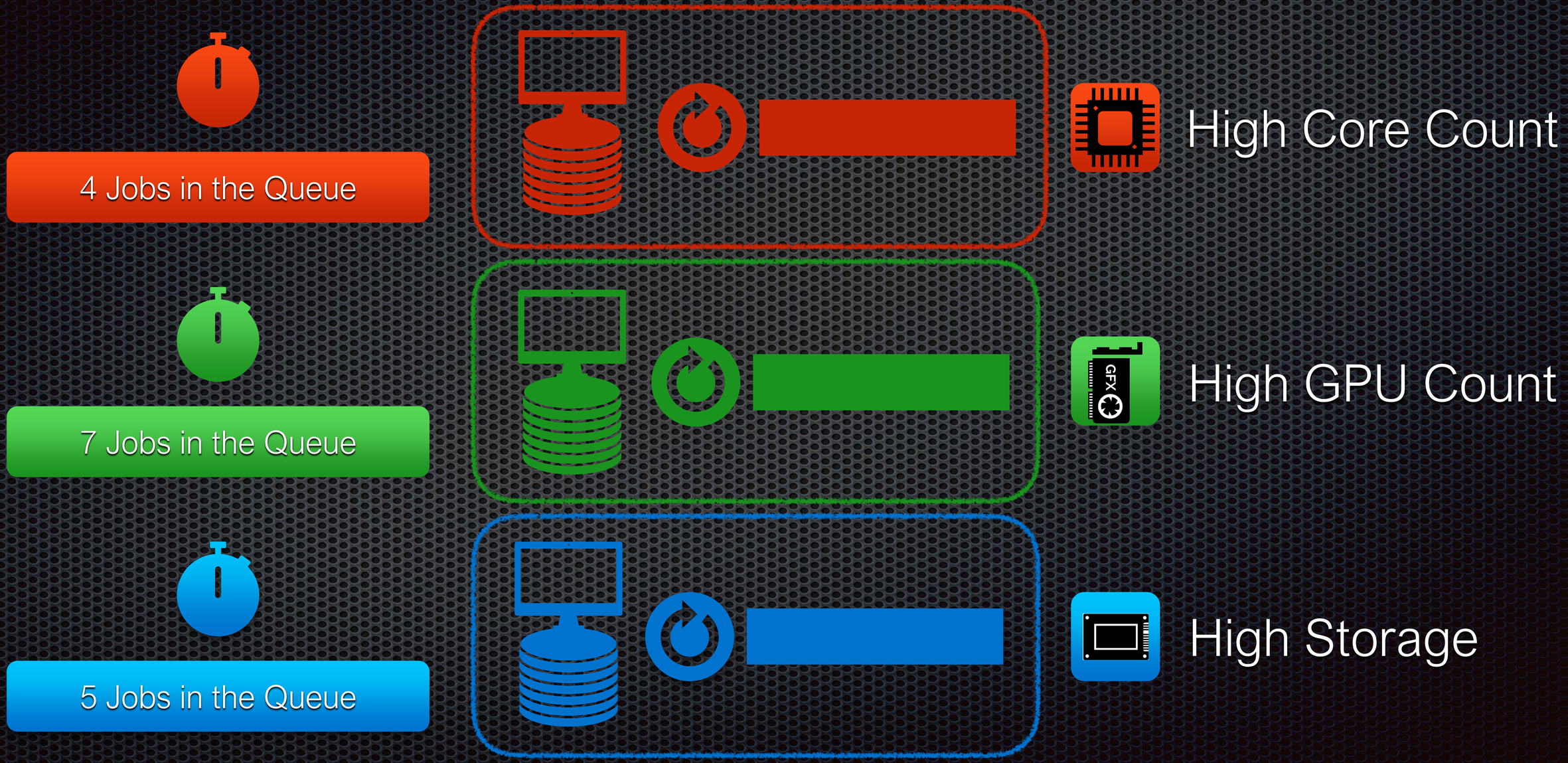
PCIe Fabric
Provides Pools of Composable Resources

PCIe Fabric

Provide Optimized Infrastructure by Job Type



Optimized Infrastructure for Different Jobs



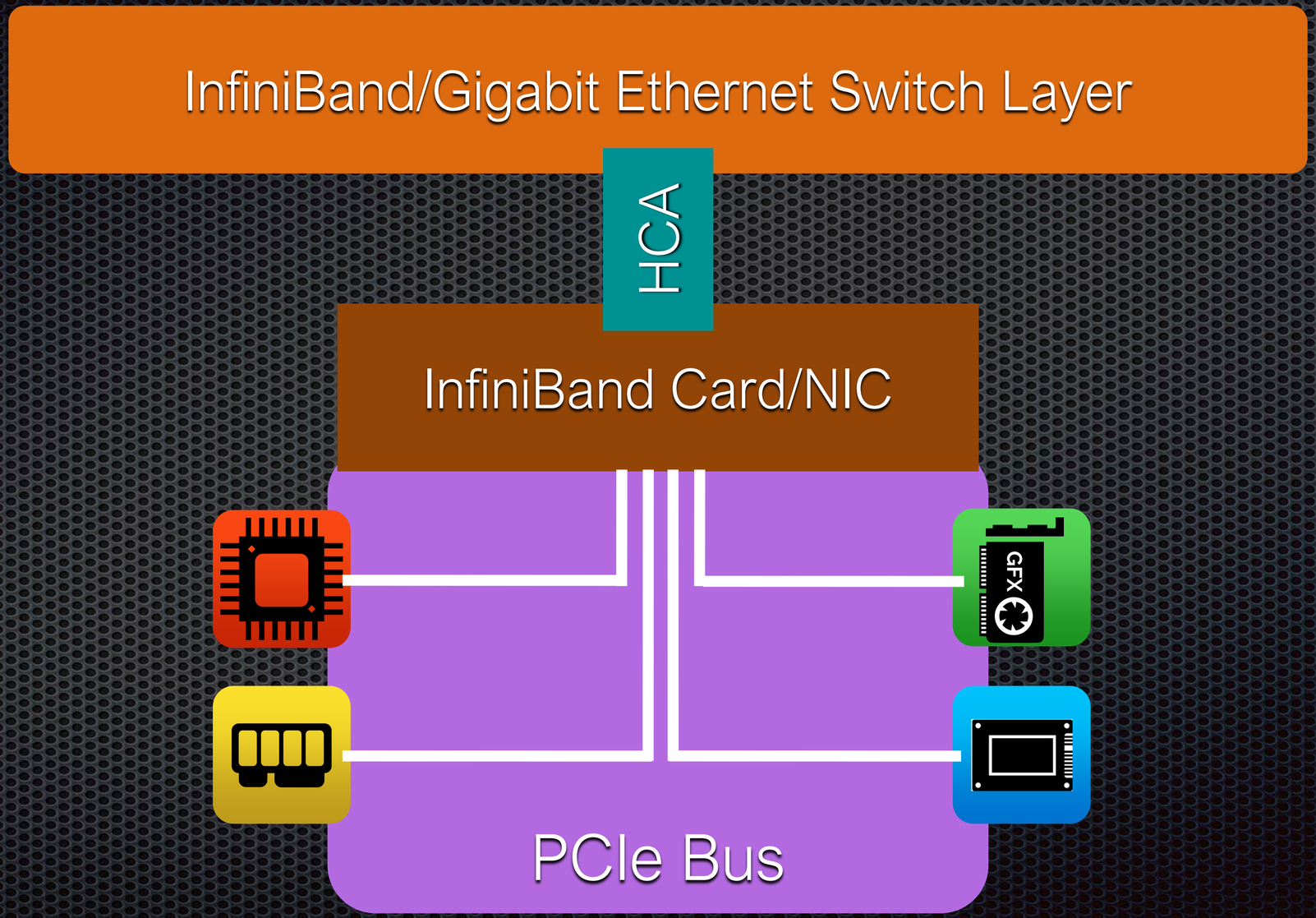
Composable ***PCIe Fabric***

PCIe Fabric brings a fundamentally new architecture to the T&E Compute Center

- PCIe based connection layer
- Direct access to components
- Eliminates translation layers (green)
- Provides a unified network
- Software-defined & composable
- API integrations to automate control
- Infrastructure matches the job

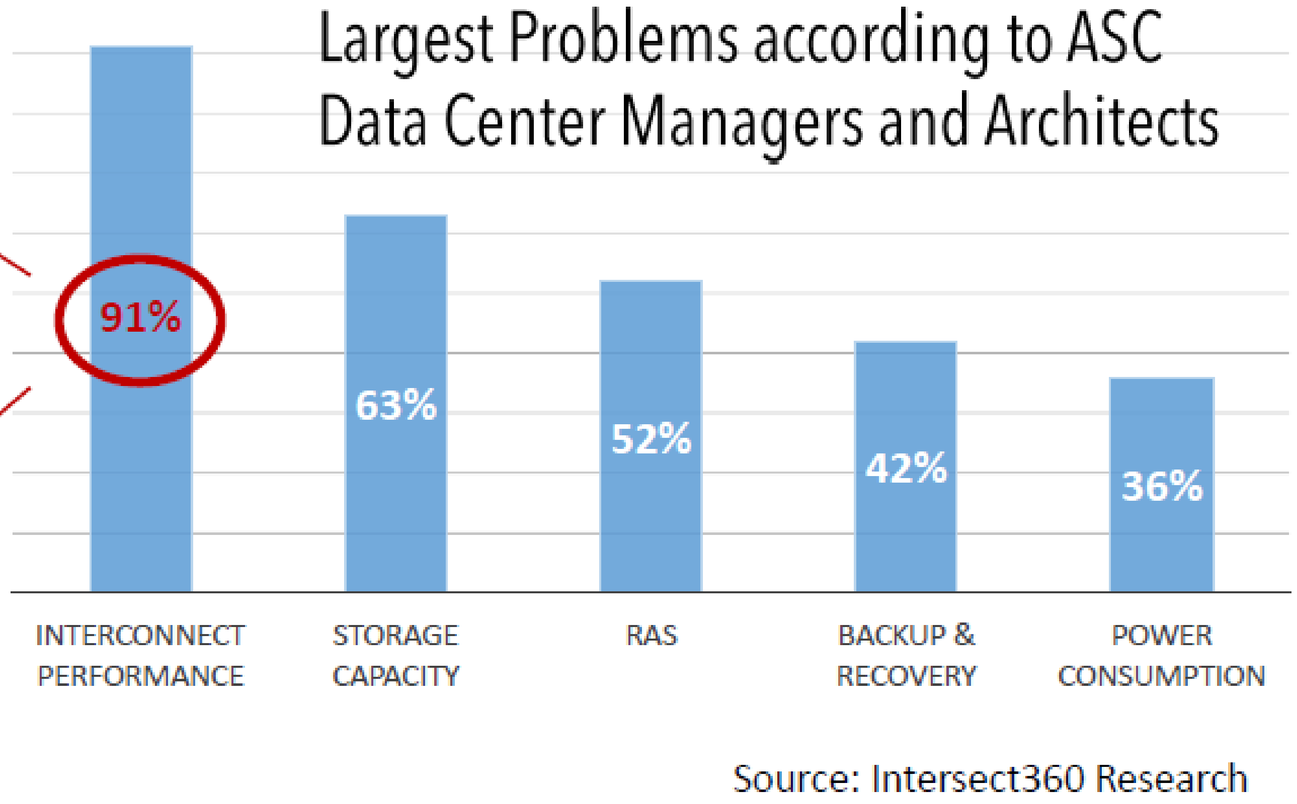
Challenge Two:
Network Fabric Performance

HPC Node Level View

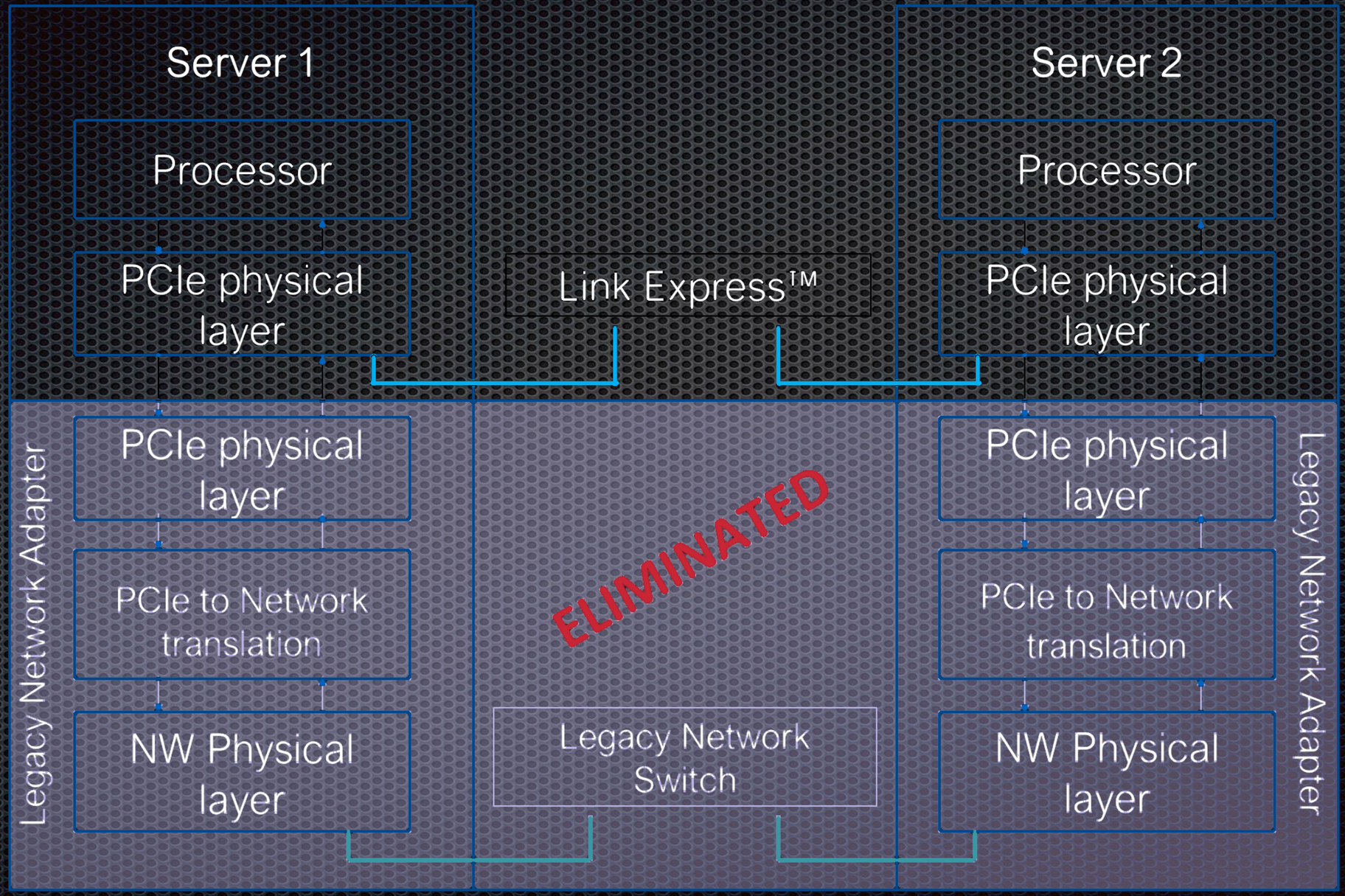


The Network is the Bottleneck

- **Too Slow**
- **Too Rigid**
- **Too Expensive**



Use PCIe to Eliminate Conversions – Key to Performance

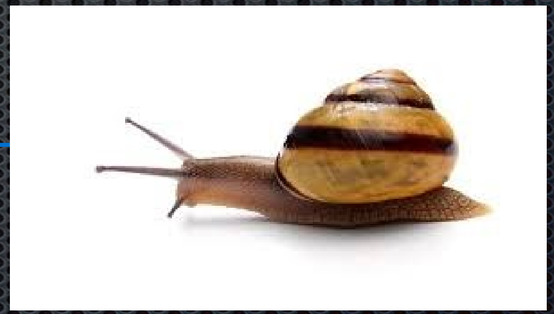


Breathtaking Performance

Typical latency (full one way) – *Less is better* *(source: OSU and GigaIO)*

PCIe
167 nanoseconds

InfiniBand
3,000 nanoseconds



Ethernet
20,000 nanoseconds

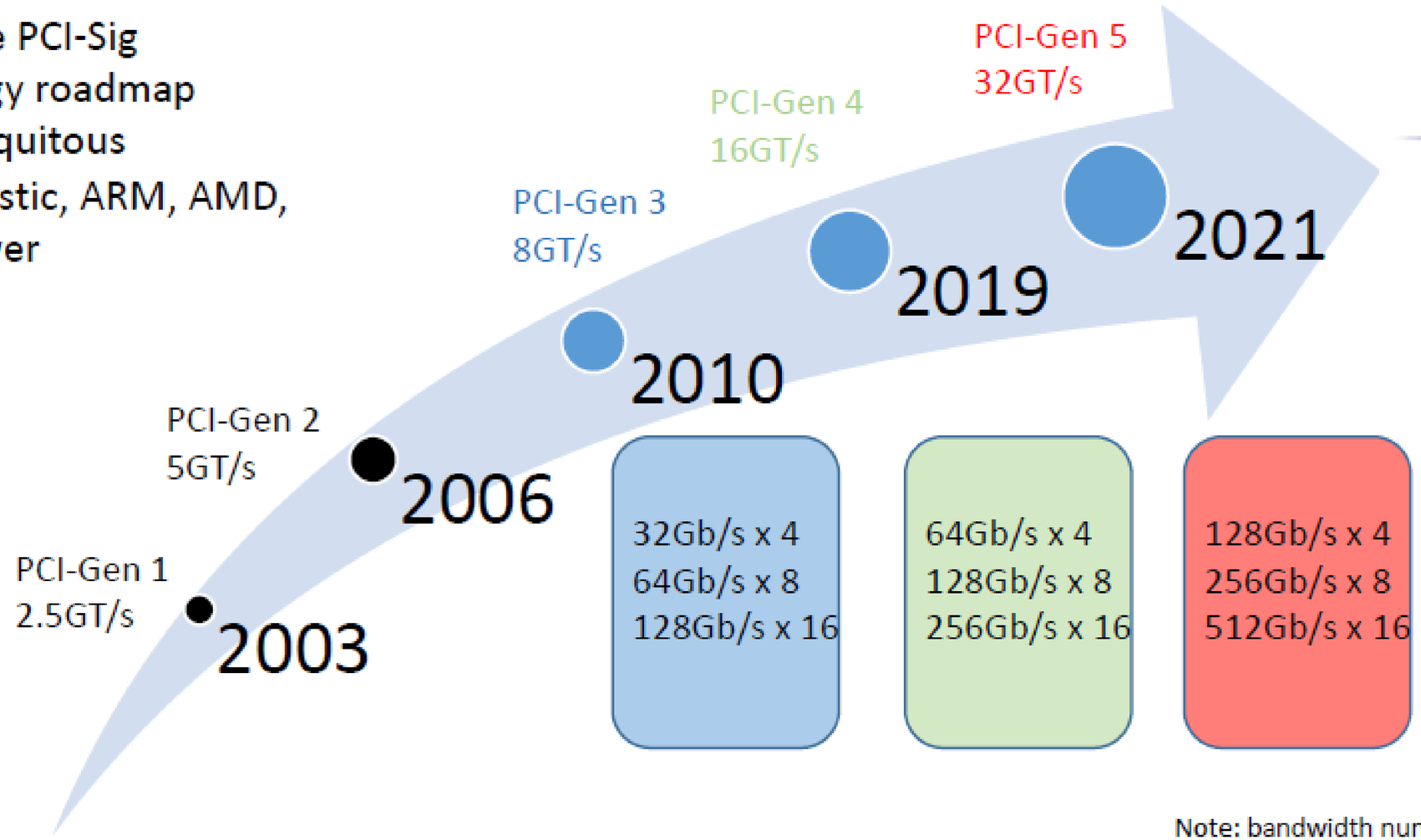
Performance of ***PCIe Fabric***

PCIe Fabric brings a new levels of performance to the HPC

- Near theoretical bandwidth
- Ultra low latency (10x better!)
- Greener (less power)
- Legacy and Future support (PCIe 1,2,3,4,5…..) - 256 Gb/s, .5 Tb/s

PCIe Future Leading Performance

- Riding the PCI-Sig technology roadmap
- PCI is Ubiquitous
- CPU agnostic, ARM, AMD, Intel, Power



Note: bandwidth numbers are Half duplex

The Advantage of ***PCIe***

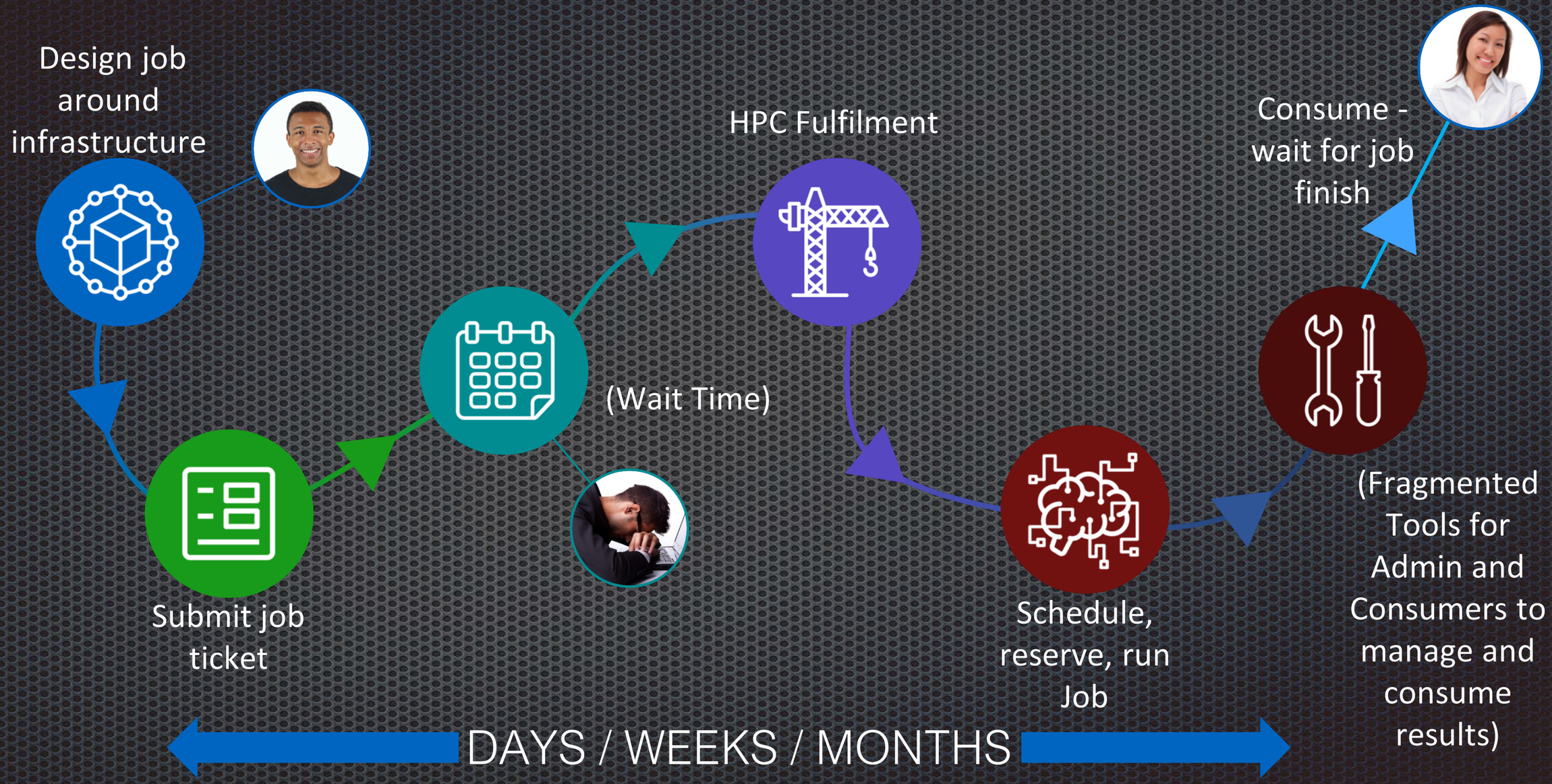
Other Network Fabrics	vs	PCIe
Various ~ 100 ? Gb/s	Performance	128 Gb/s (PCIe 3)
> 3000 ns	Latency	< 1000 ns
Proprietary Bus	Connects To	Directly to PCIe BUS
Only propriety fabric	Supports	All PCIe (1,2,3,…)
Limited – Replacement	Future	PCIe 4 Ready - 256 Gb/s PCIe 5 - .5 Tb/s 2-3 yrs

SDSC Performance Testing Results

- 25% Higher performance for 1/3 the cost of hardware (latency less by almost 10X)
- PCIe drives much higher levels of resource utilization due to aggregation and scaling effectiveness
- PCIe bandwidth superior to competition (IB) (115Gb/s vs 82 Gb/s)

Challenge Three:
Lack of Self Service

Traditional HPC Job Process Challenged



Lack of Self Service

- 1970’s methodology
- Submit a job against standard infrastructure
- Many different tools to support (reservation, scheduling, job submittal, monitoring, etc.)
- Different administrative environments and tools
- User has little control over defined infrastructure
- Set configurations of deployed apps (Openstack, Linux, Containers - typically one flavor)
- Little user customization of application layer or automation for application deployments
- Admins required to deploy new application environments

Middleware Infrastructure Framework

Workflows

- Automated Test Suites
- CI/CD Pipeline
- Big Data Jobs
- Service & Support
- Technical Escalation
- Training - Internal / Partner
- AI/DL/ML Jobs
- Cyber Exercise/Training

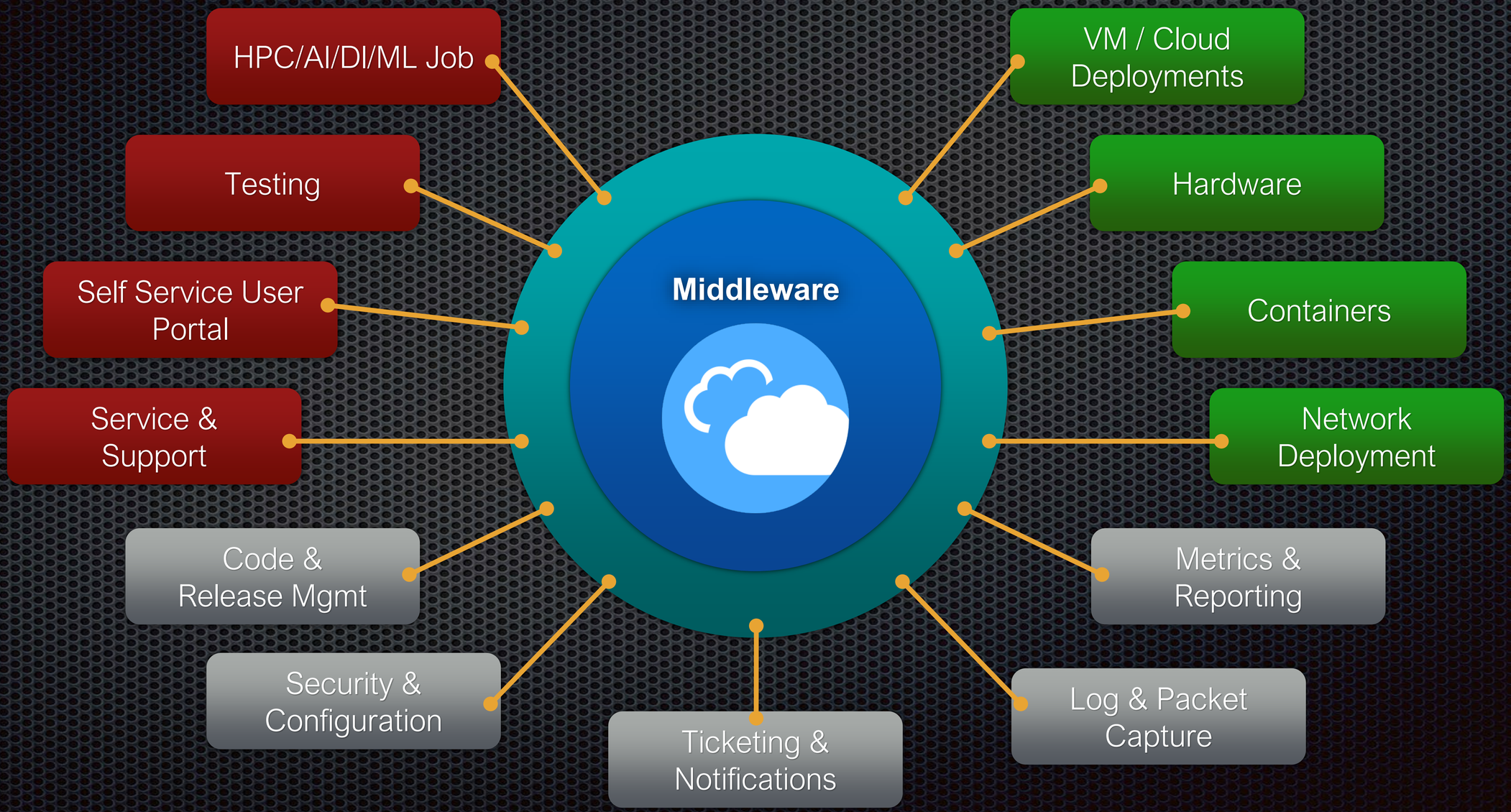
Middleware

Middleware able to link various infrastructure to various use types

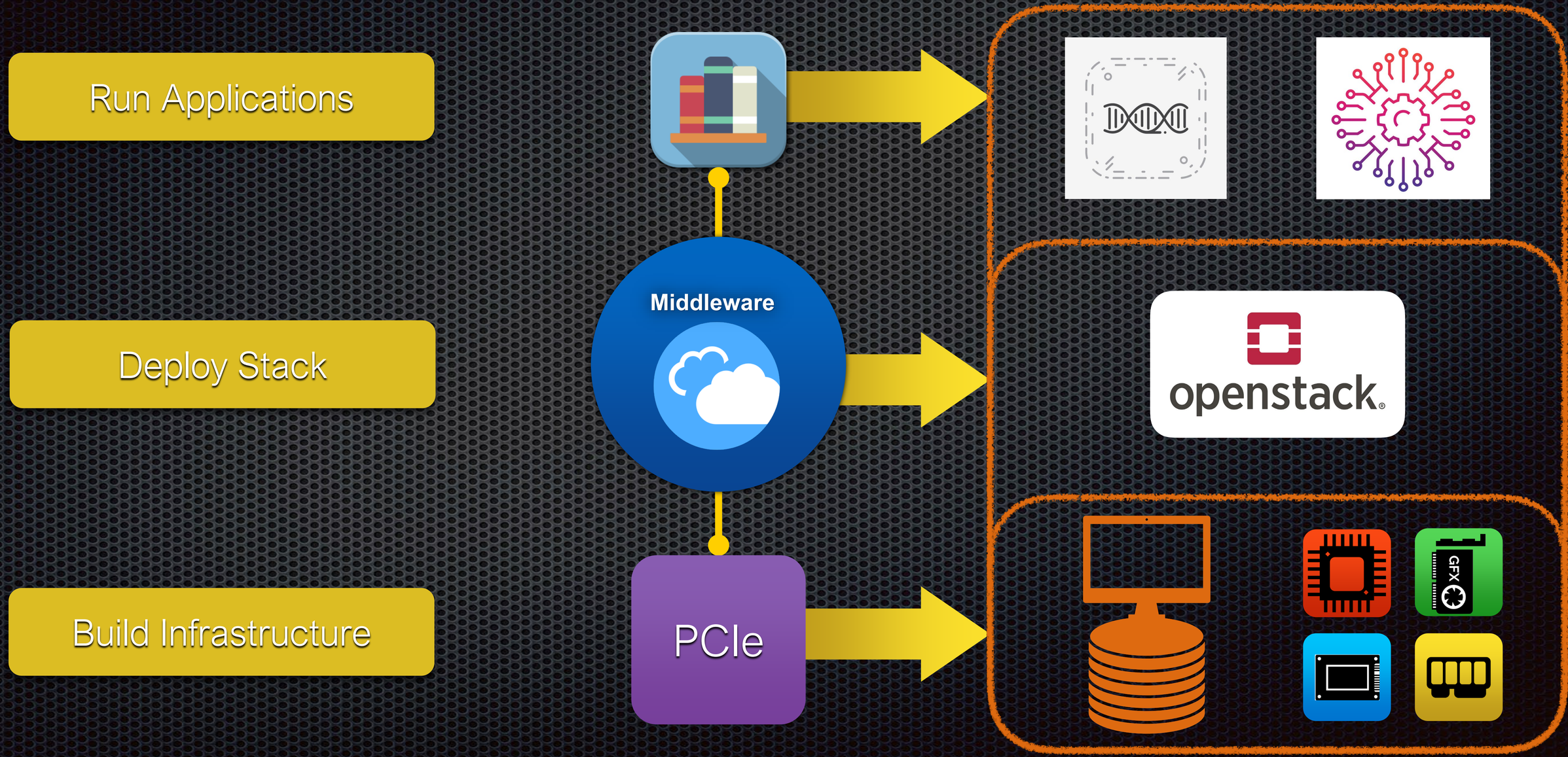
Infrastructure

- Off-Prem VM Services
- On-Prem VM Services
- Physical Network Equipment
- Firewall / Security / Load Bal
- Data / Storage Services
- Device Farm
- HPC
- Cyber Range

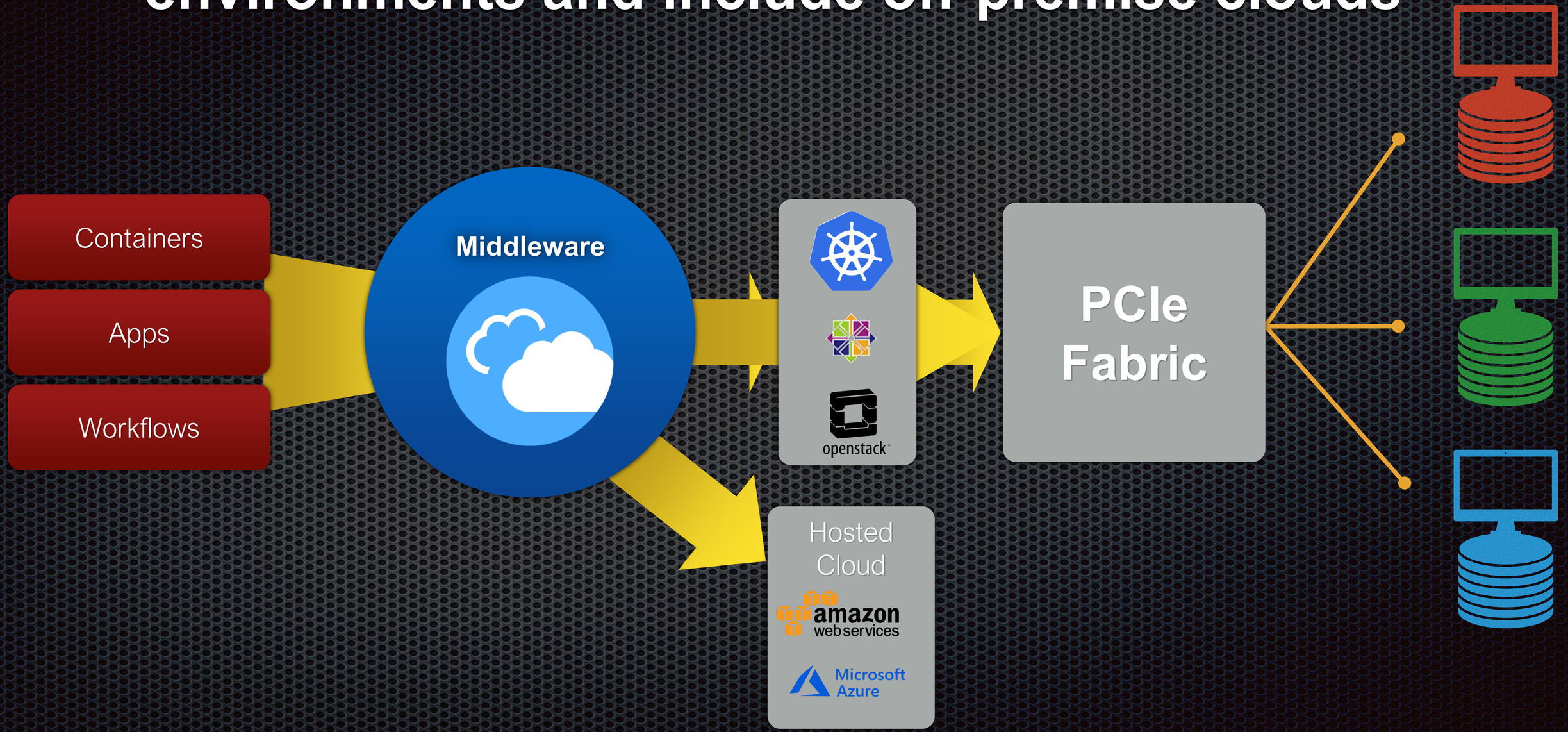
Tying it all together so self service can happen



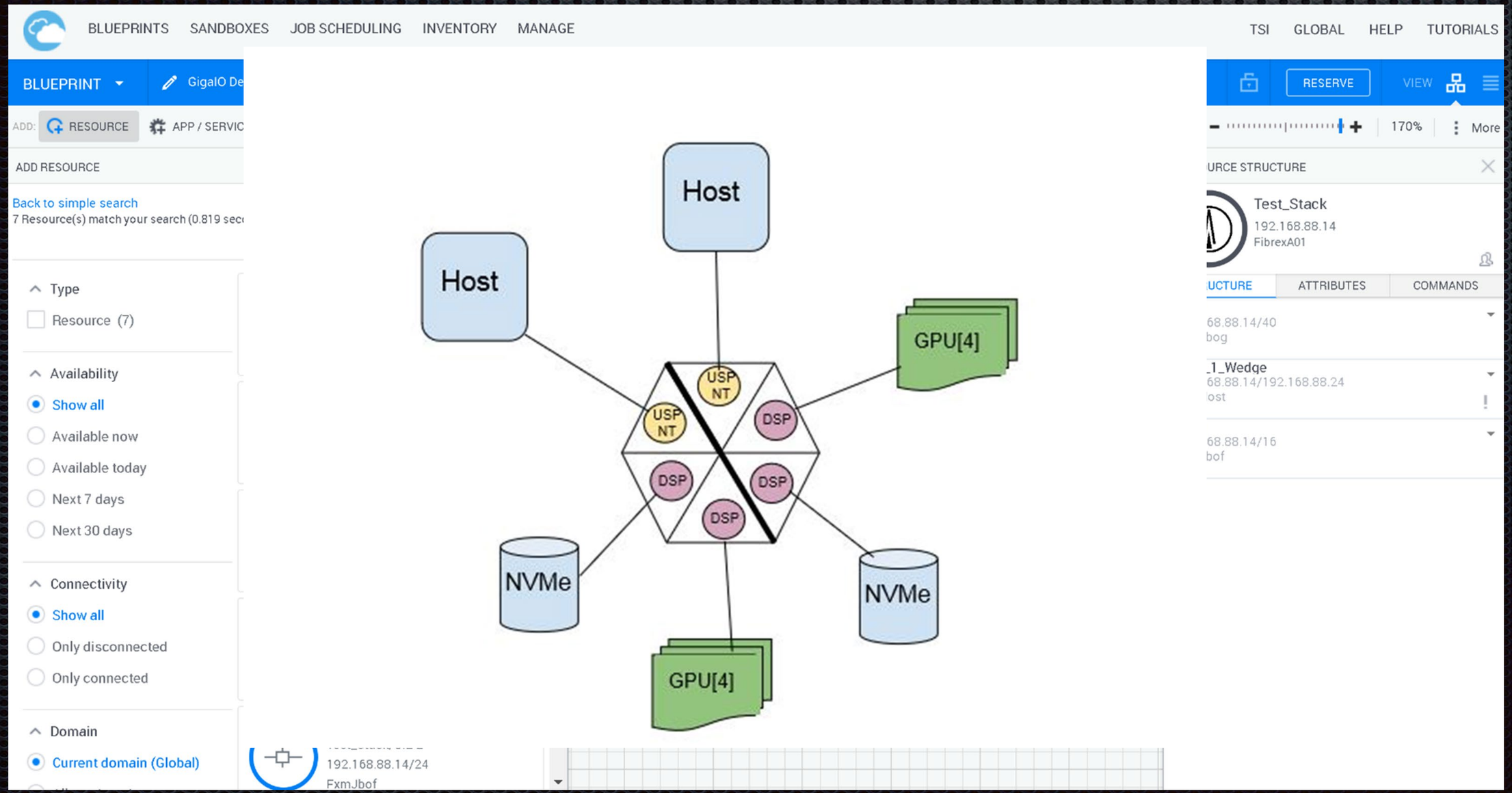
Middleware Framework + PCIe Fabric (TECaaS)
Can Orchestrate Infrastructure from Bottom to Top



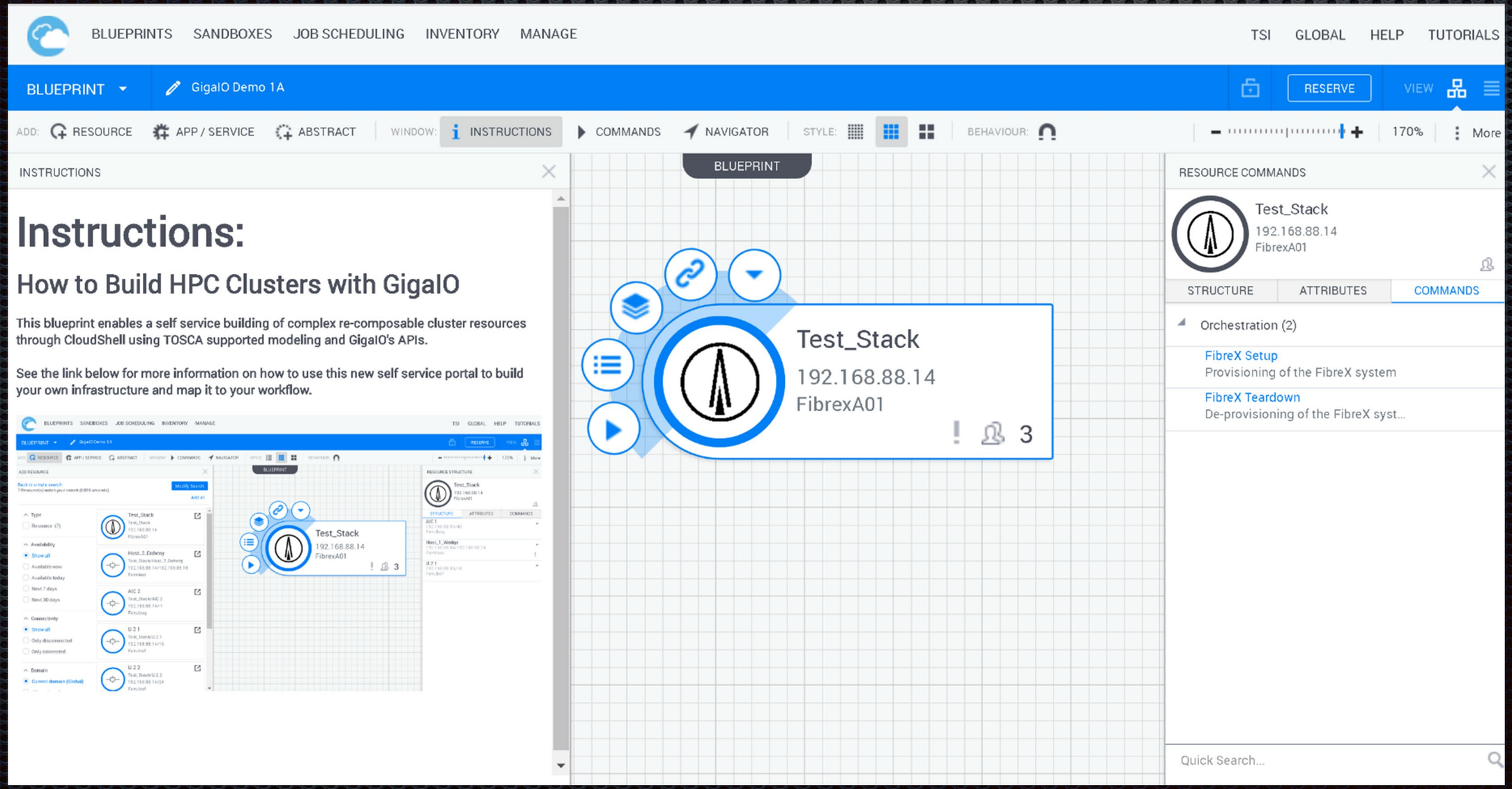
TECaaS can orchestrate infrastructure environments and include off-premise clouds



Build blueprints of full stacks with Composable PCIe Resources



Deliver Blueprint with Self-service Instructions and Automation



TECaaS Advantage

Test and Evaluation Compute Center with PCIe Network Fabric

	Traditional Approach		HPC-as-a-Self-Service
Design Environment Manual: Visio, Powerpoint	Hours	Minutes	Model & Publish Blueprint Drag n' drop from asset inventory, connectivity, abstract blueprint
Request I.T. Dedicated P/V Infrastructure	Hours	Minutes	Self-Service On-demand equipment reservation and scheduling
I.T. Fulfillment Rack and Stack, Configure, Validation, Approval	Days/ Months	Minutes	Automate & Orchestrate Simplified configuration and provisioning, save and restore.
Fragmented Access Telnet, SSH, RDP, API, CLI	Complex	Simple	Unified Access Embedded Web-portal
Fragmented Users Single user access, conflicts, hoarding	Siloed	Shared/ Reused	Multi-Tenant + Scalable Consolidated labs, global shared user base, Community

VS

Thank You for attending!

For more information visit us at:

www.tsieda.com

Or email us @ info@tsieda.com

White Paper Available