FUNGIBLE

The Fungible DPU™:

A New Category of Microprocessor for the Data-Centric Era

Hot Chips 2020

Our Mission

Revolutionize the **performance, economics, reliability** and **security** of all scale-out data centers

Core Technology: a new category of microprocessor called the **Fungible DPU™**, associated software, and systems

Context

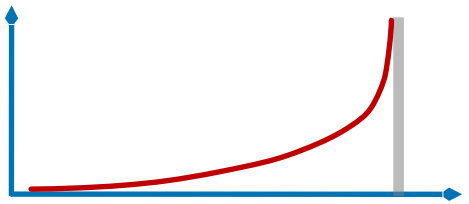
Problems Facing Data Centers

Large footprint (power and space)

- Inability to pool expensive resources
- Inefficient execution of *Data-Centric Computations*

Scaling challenges (very small & very large)

Increasing complexity (technology limits)



Security vulnerabilities

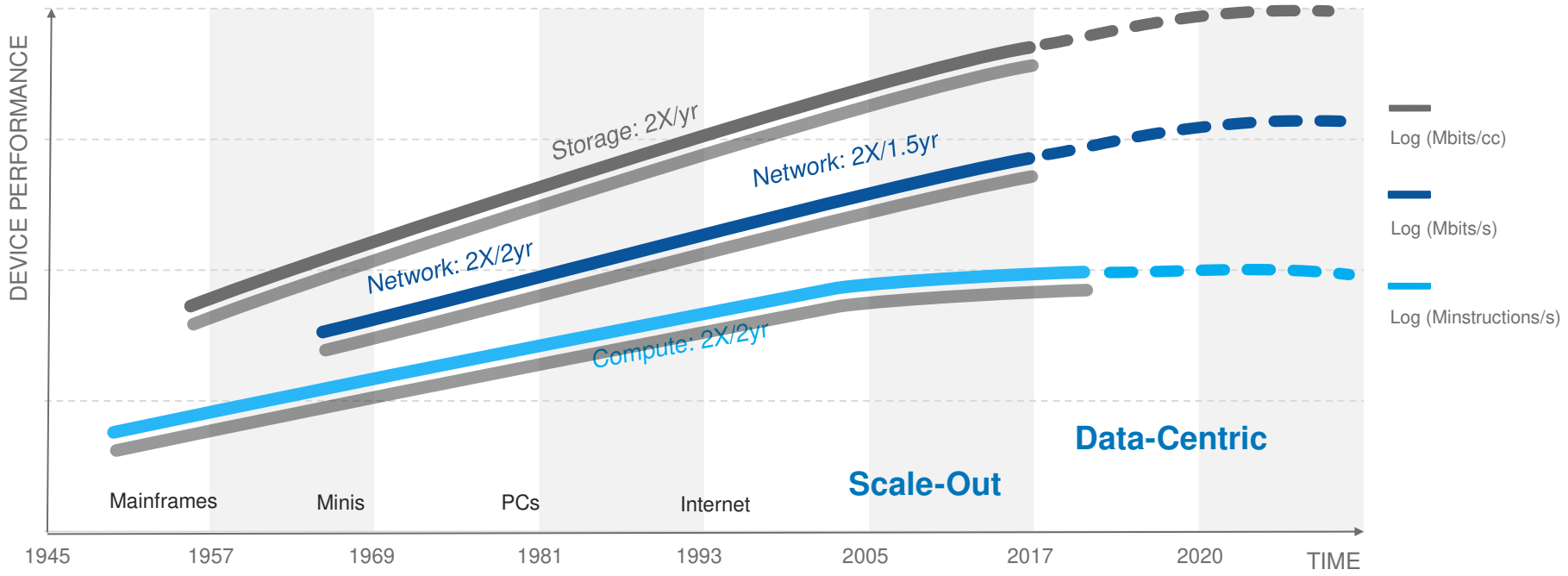
Headwinds

Network and storage speeds are increasing faster than compute

Applications need to access ever larger data sets

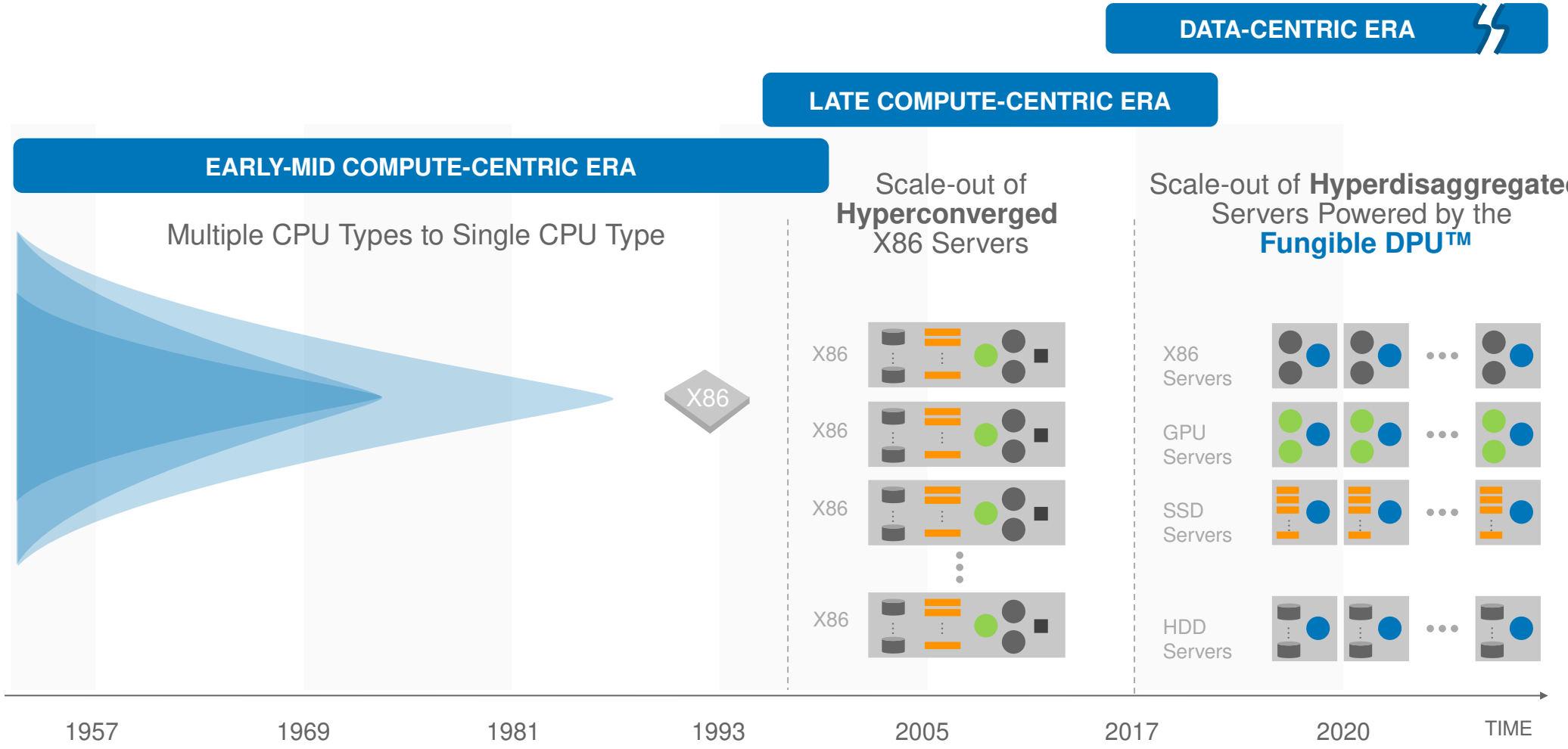
Moore's Law slowing and will likely plateau

Security attacks increasing in frequency and sophistication



FUNGIBLE

Data-Centricity Will Drive the Architecture

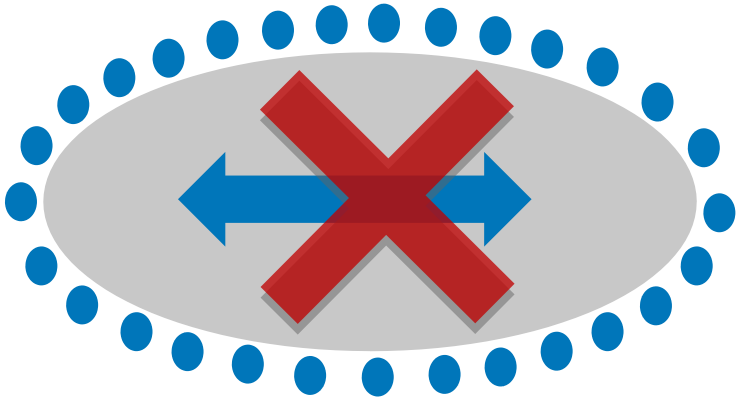


Our Approach:

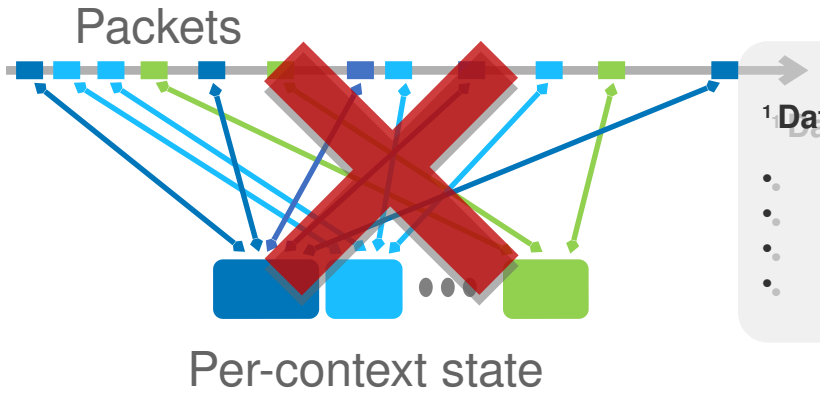
Clean Sheet, Fundamentals Based Design

Confront the Root Causes

Inefficient data interchange between nodes



Inefficient execution of *data-centric*[1] computations inside nodes



[1] **Data-Centric Computations:**

- All work arrives as **packets**
- Require frequent **context switching**
- Involve modification of **state**
- **I/O dominates** arithmetic and logic

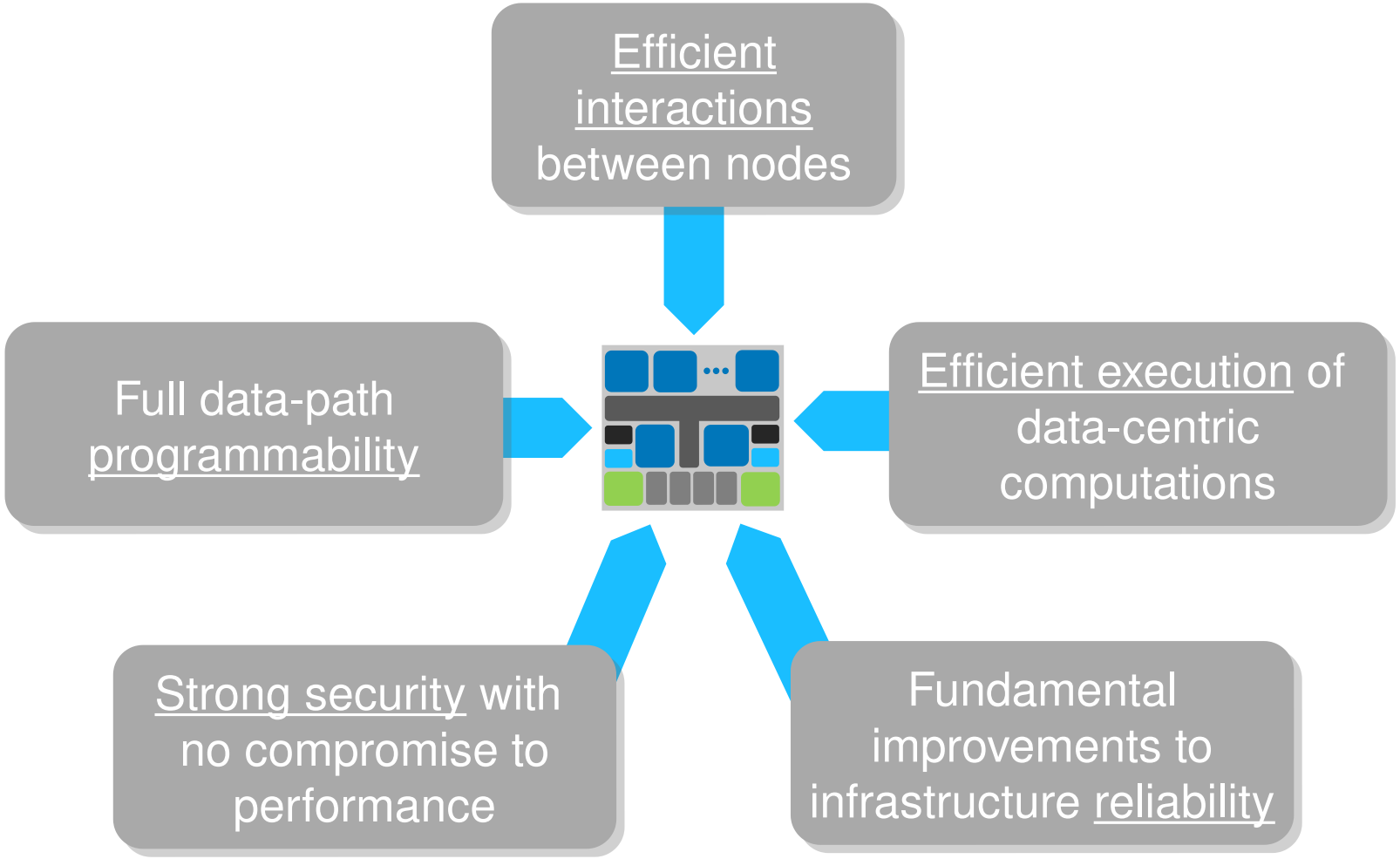
Unreliability

Inflexibility

Insecurity

FUNGIBLE

Fungible DPU™ Addresses all Five Root Causes



Where does the Fungible DPU™ sit?

X86 Compute Node

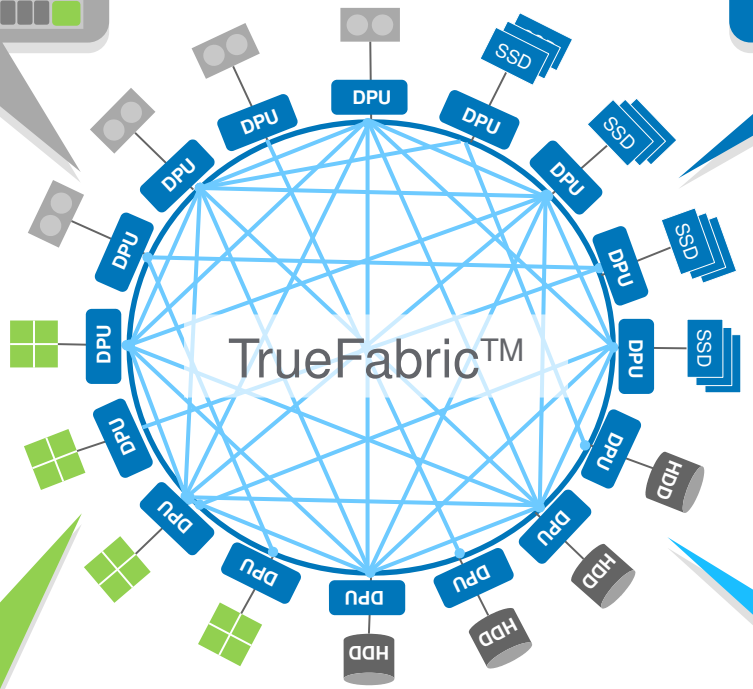
Fungible DPU™ in each node

- Connects node to network
- Functions as intra-node hub

Programmable data and control planes

Handles data-centric computations

End point of TrueFabric™



SSD Storage Node

Provides:

- ✓ Full flexibility with high performance
- ✓ Strong end-to-end security
- ✓ Pooling of resources
- ✓ Reliable low-jitter, low latency fabric
- ✓ Observability and telemetry

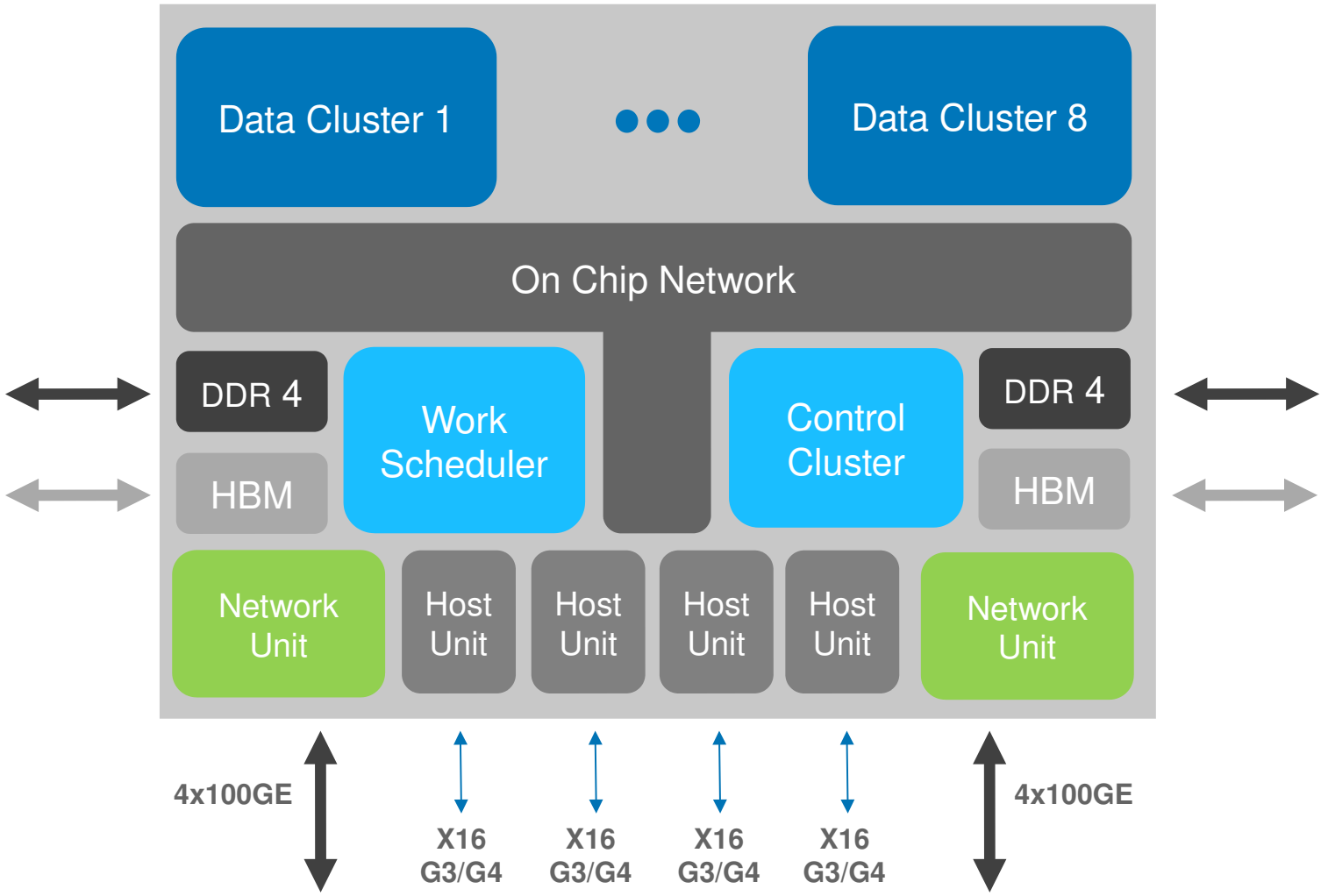
GPU ML/Analytics Node

HDD Storage Node

FUNGIBLE

The Fungible F1 DPU™

Fungible F1 DPU™ Architecture



8 Data Clusters

- 192 processor threads
- Full cache coherency
- Tightly integrated accelerators

Control Cluster

- 8 processor threads
- Secure complex
  - HSM
  - Root-of-trust
- Work scheduler

Memory & I/O Interfaces

- 800 Gbps network unit
- 4x16 G3/G4 PCIe unit
- DDR4
- HBM

On Chip Network

- High performance
- Low latency
- Any unit to any unit

FUNGIBLE

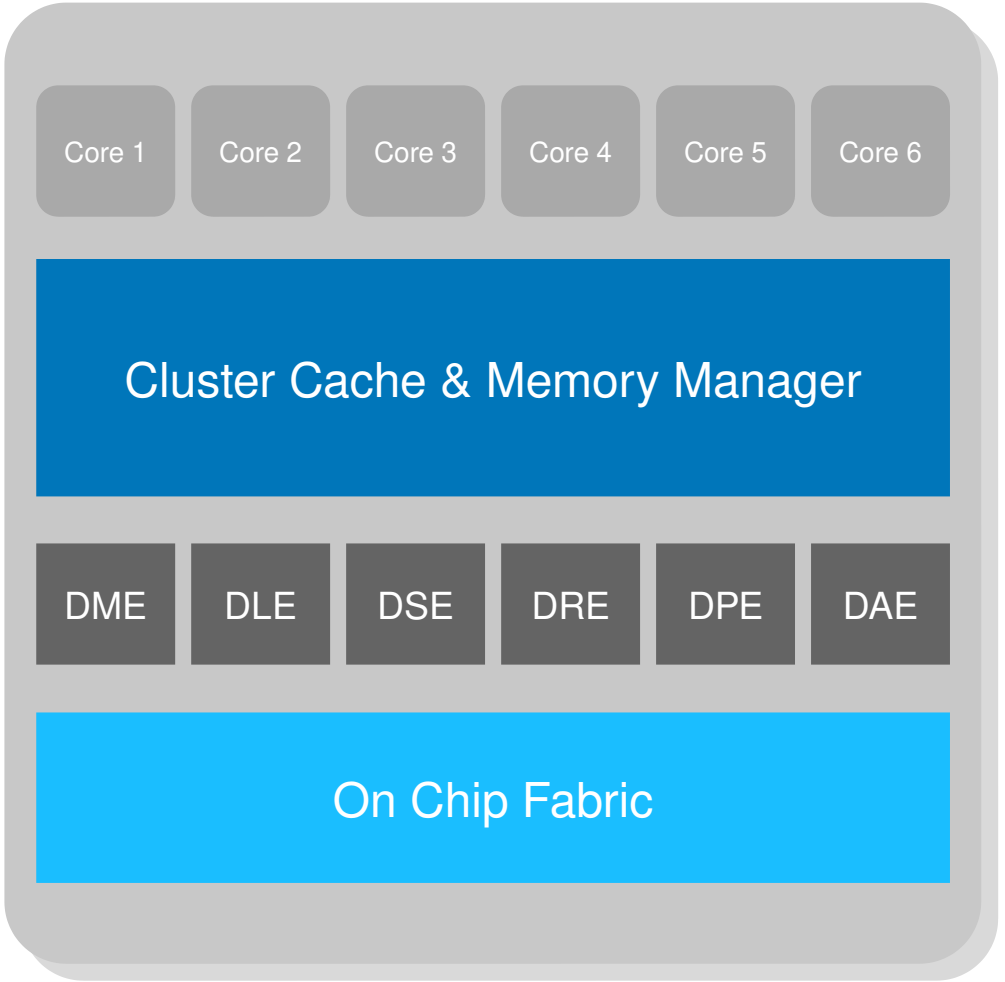
Data Cluster

Runs the Data Plane

6 Cores * 4 Threads

Multi-Threaded Accelerators

- Data movement
- Data lookup
- Data security
- Data reduction
- Data protection
- Data analytics



Control Cluster

Runs the Control Plane on Linux

4 cores * 2 threads

Secure Enclave

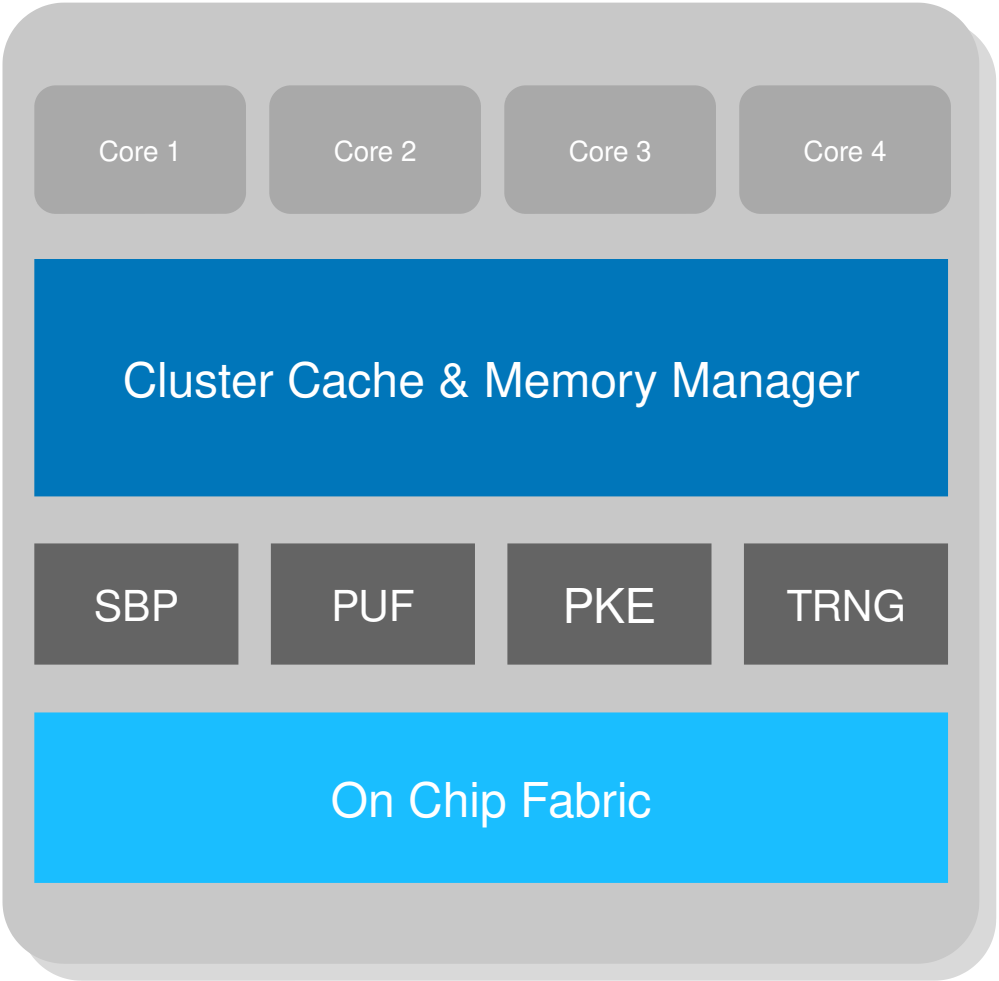
- Secure Boot
- Secure Key Vault
- Binary Signing and Authentication

Public Key Crypto Engines

- RSA
- Elliptic Curve

True Random Number Generator

Physically Unclonable Function



CPUs, Caches, External Memory

CPUs

- MIPS-64, 9-stage, dual-issue, 4xSMT, FPU/SIMD unit
- IPC on data-centric workload close to CPU-max
- Full hardware virtualization
- Large I+D L1$, shared L2$, full system-wide coherency

High Bandwidth HBM2 Memory

- 8GB, 4Tbits/sec
- Integrated in the package

High Capacity DDR4 Memory

- 2xDDR4 controllers, ECC enabled, up to 2666 MHz
- Up to 512GB
- Support of RDIMM, NVDIMM-N

- **Fully general programmability**
- **All code in ANSI-C**
- **Fast thread switching**
- **Tight coupling with accelerators**
- **No performance compromises**

High-Performance (800G) Flexible Network Engine

Implements TrueFabric™ end point

Low latency Ethernet MAC with FEC

Integrated L2/L3/L4 forwarding

Low latency transit switching

Support of general virtualization protocols

Tight integration with data clusters

P4-like language controls

- Parsing, encapsulation, decapsulation
- Rx/Tx acceleration
- Lookup acceleration

All packets are AES-GCM encrypted

Precision time protocol

- **Low, deterministic latency**
- **Full cross-section bandwidth**
- **End-to-end congestion control**
- **End-to-end error control**
- **End-to-end encryption**
- **Network virtualization**
- **Granular QoS**
- **Enables disaggregation at scale**

FUNGIBLE

High-Performance (512G) Flexible Host Engine

Includes 16 independent dual-mode controllers

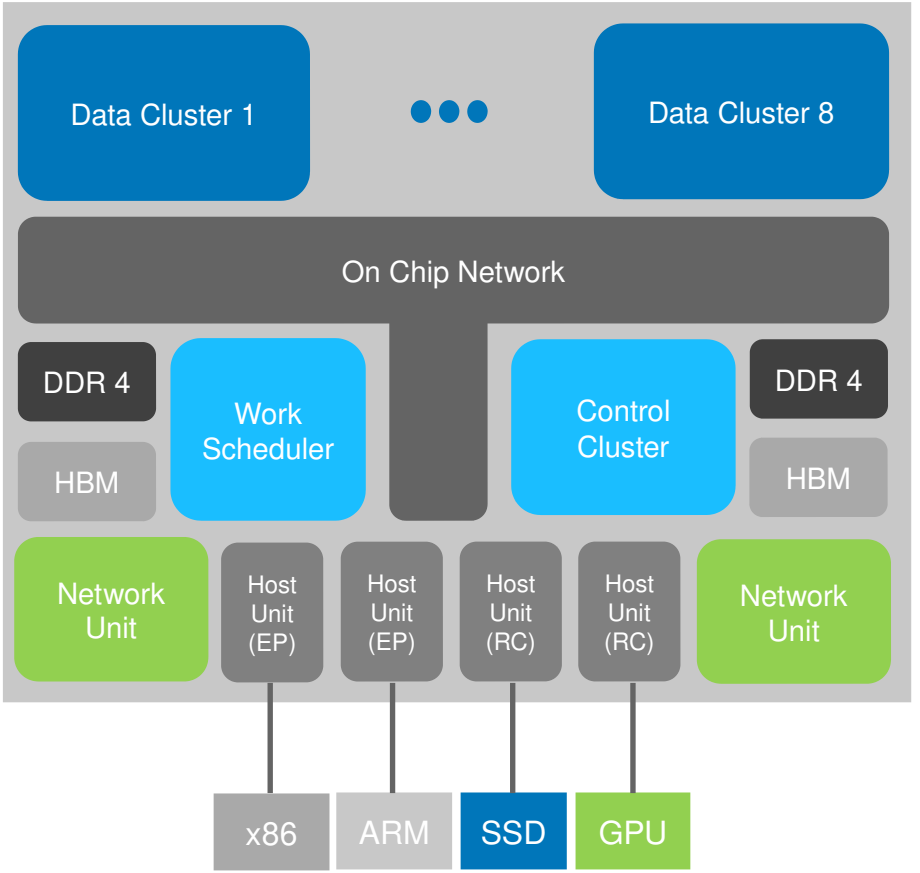
- •EP/RC in any combination (4x16 to 16x4)

End point for X86 or ARM CPUs

- •Hardware virtualization
  - •SR-IOV with 64 PFs, 1024 VFs
  - •Fine-grain QoS support
- •Software flexibility
  - •Full network, storage, and security virtualization

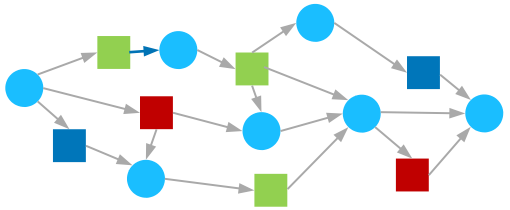
Root complex

- •High performance abstraction layer
- •Fully flexible data and control planes
- •Connects to and abstracts SSDs, GPUs, FPGAs

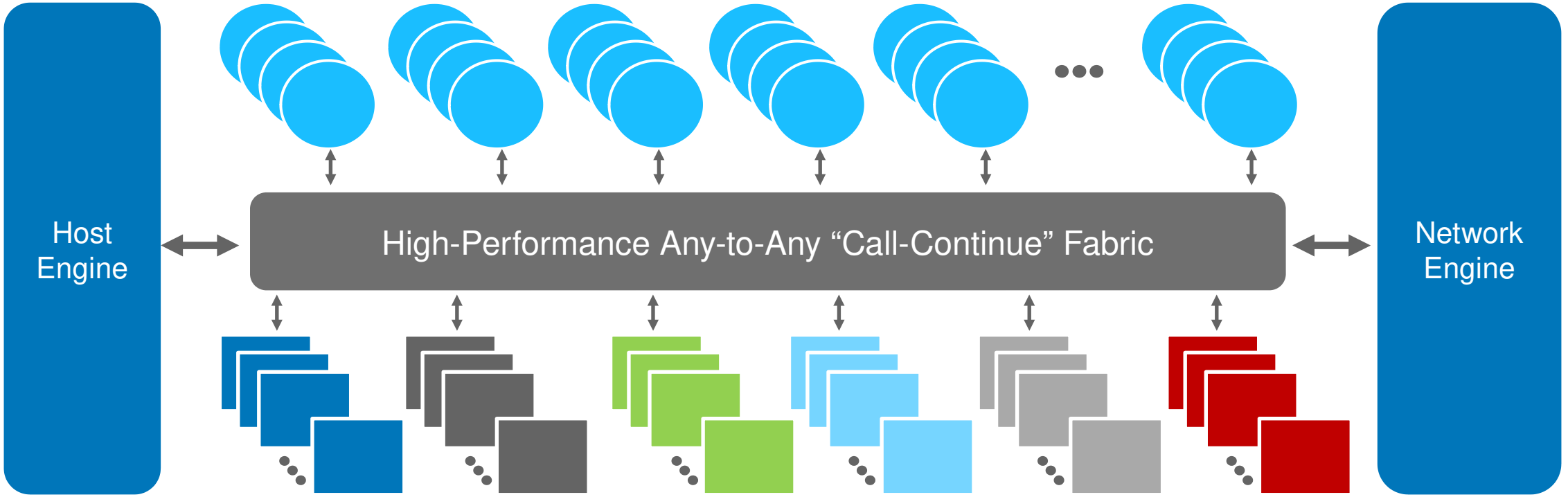


FUNGIBLE

Data Path Programming Model

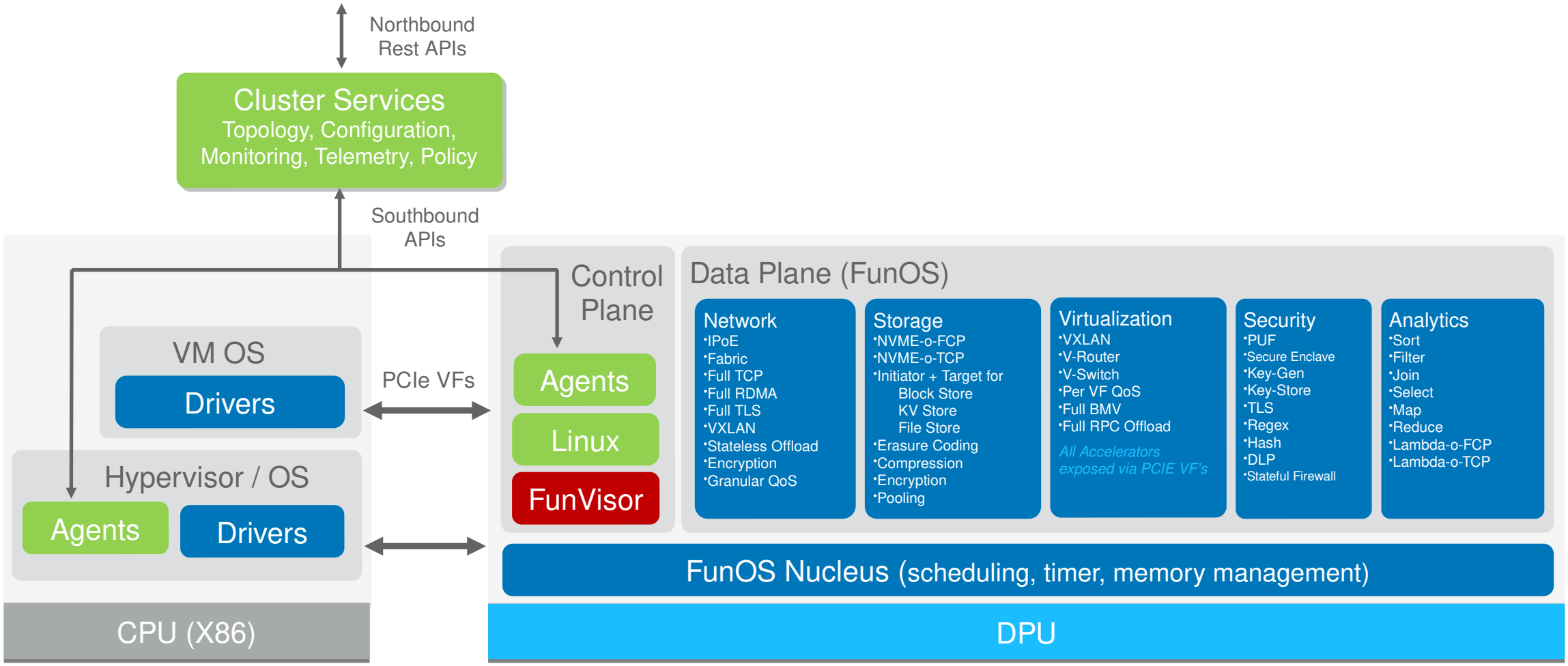


MIPS-64 Hardware Threads Execute Run-To-Completion C-Code



Heterogeneous Accelerator Threads

Fungible DPU™ Software



Multiple Levels of Programmability

Software running on the DPU

- DPU control plane
- DPU data plane

Software running on a PCIe connected X86 Host

- OS drivers and Agents
- Data path code execution via eBPF

Cluster Services for management and control of multiple DPU systems

- Northbound APIs for orchestration systems

Infrastructure Services Performance

Service	Measured Performance	Estimated[1] Performance
TCP[2] (Single Flow, Multi Flow)	50Gbps, 250Gbps	70Gbps, 400Gbps
TLS[2] Session Setup Rate	32,000/sec	100,000/sec
IPSEC[2] (Single Flow, Multi Flow)	-	10Gbps, 250Gbps
Stateful Firewall[2]	-	370Gbps
OVS	-	400Gbps
Load Balancer	256Gbps	300Gbps
Block Store (4K IOPS)	8M	10M
Video Streaming	256Gbps	300Gbps
TPC-H Benchmark (relative to X86)	3X-100X	-

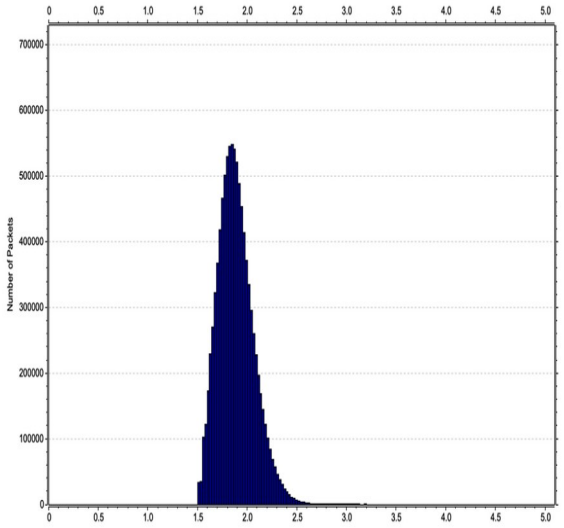
[1] Full chip dedicated to service
[2] All measurements are full-duplex

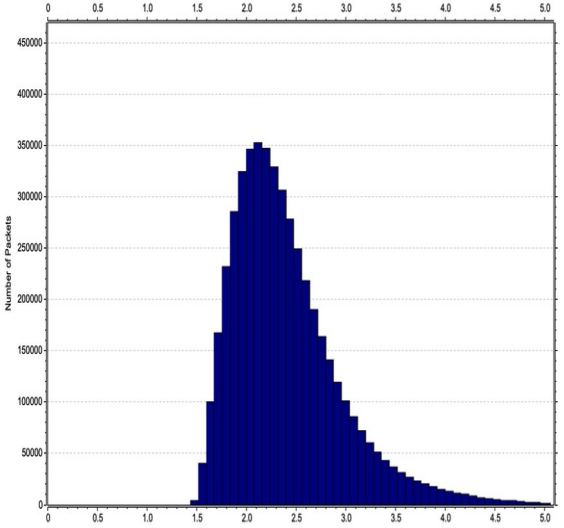
TrueFabric™ Performance

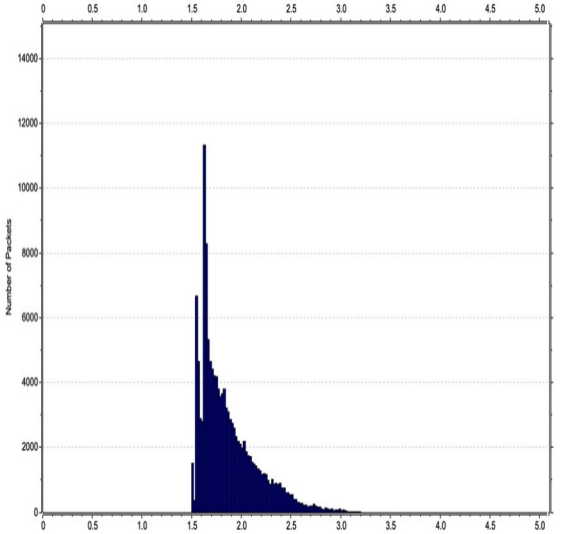
Traffic Pattern	1024 * (Node to Node)	1024 Node to 1024 Node	1024 Nodes to 1 Node
Fabric Utilization	90.7%	93%	90%
Latency Mean	1.84µs	2.10µs	1.71µs
Latency Variance	0.13µs	0.32µs	0.12µs
Latency P99	2.14µs	3.30µs	1.75µs

Network Configuration:

- 1024 Nodes
- 200Gbps/Node
- Two-tier leaf-spine
- Leaf ZLL: 500ns
- Spine ZLL: 500ns
- iMix packet profile





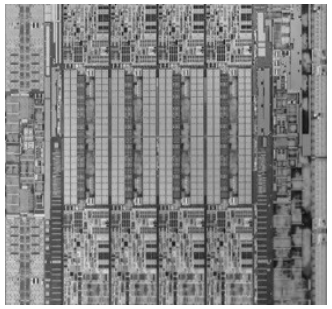


FUNGIBLE

A New Category of Microprocessor

Purpose-built for the data-centric era

CPU



General-purpose

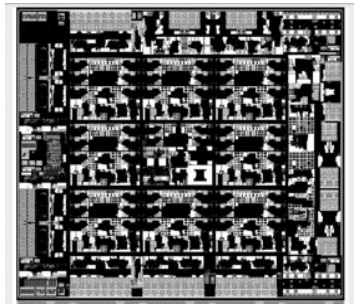
Multi-core, MIMD
High IPC for single threads
Fine-grain memory sharing
Classical cache coherency
Based on locality of reference
Ideal for low to medium I/O

GPU

Vector floating point

Multi-core, SIMD
High throughput for vector processing
Coarse-grain memory sharing
Relaxed coherency
Based on data >> instructions
Ideal for graphics, ML training

Fungible DPU™

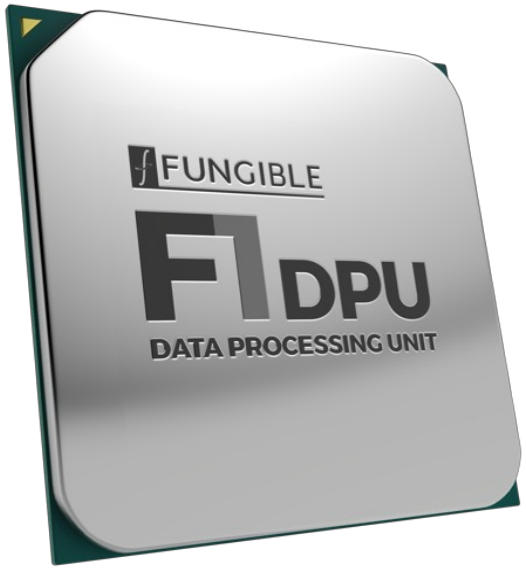


Data-centric

Multi-core, MIMD + tightly-coupled accelerators
High throughput for multiplexed workloads
TrueFabric™ enables disaggregation and pooling
Specialized memory system and on-chip fabric
Ideal for network, storage, security, virtualization
Data-centric computations run >10X more efficiently

Announcing the Fungible F1 and S1 DPUs

Common Architecture and Programming Model



- Storage target
- AI Server
- Security appliance
- Analytics

- Bare-metal virtualization
- Storage initiator, local instance storage
- NFV applications
- Node security

FUNGIBLE

THANK YOU