Future of Adaptive Computing:
The Composable Data Center

Salil Raje, EVP & GM, Data Center Group

March 2021

Data Center Trends

- Explosion of unstructured data
- Diversity of workloads
- Latency-sensitive applications
- Evolving standards

Online Gaming

Video Conferencing

Ecommerce

Streaming

TIME

XILINX

Key Challenges for Data Centers

Scale infrastructure to meet demand

- Meet performance and latency requirements
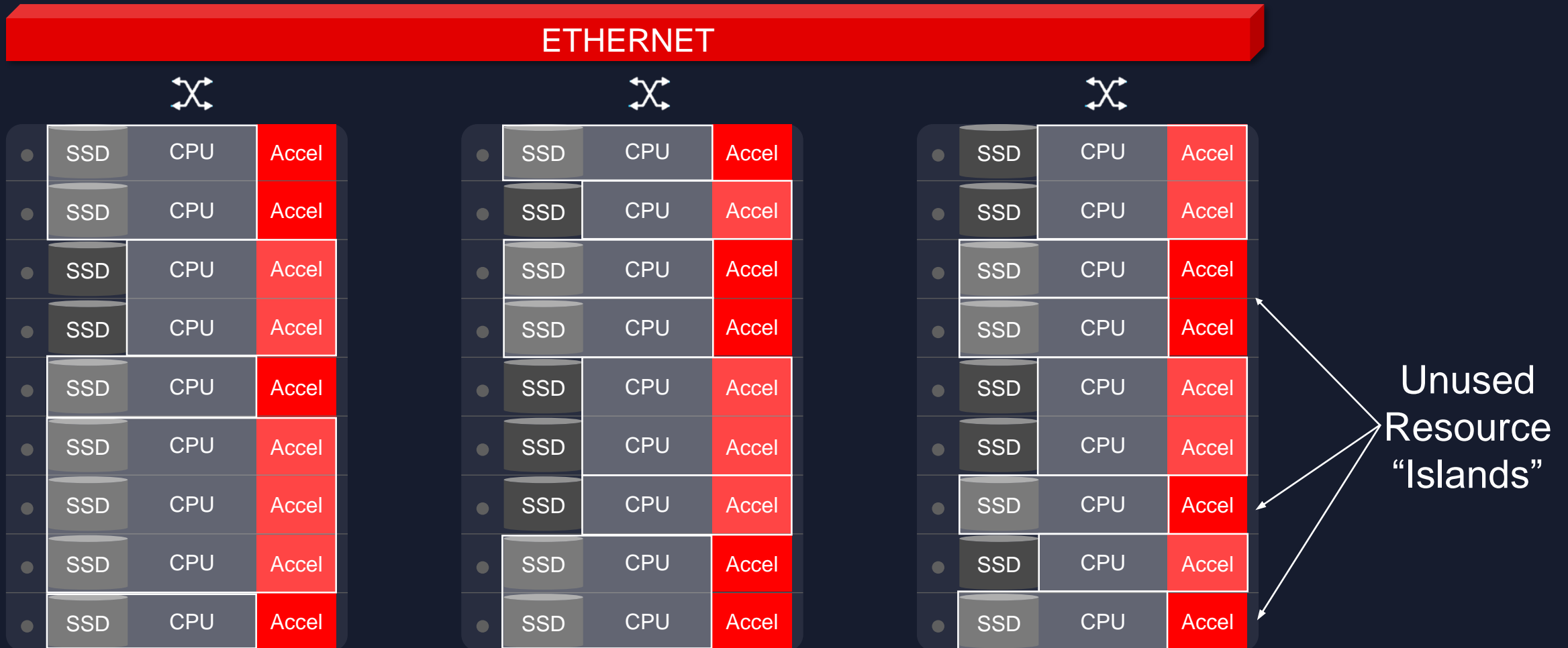- Quickly adapt to changing workloads
- Minimize TCO and power
- Evolve to more efficient architectures
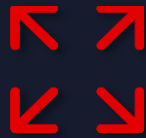
Limitations of Current Data Center Architecture:

Fixed Configuration + Sub-optimal Utilization

Unused Resource "Islands"

XILINX

The Solution: Composable Data Center

Decoupling of resources

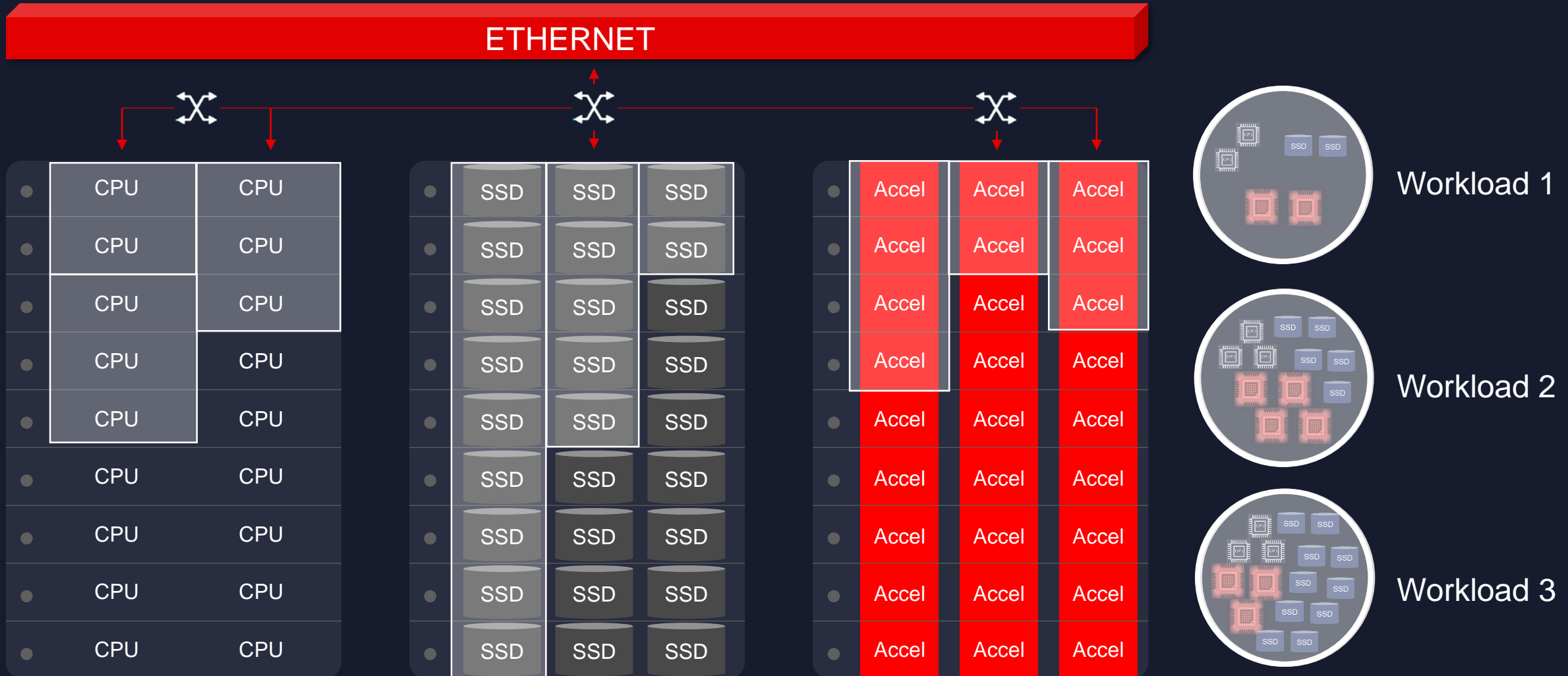**Pooling them + accessible from anywhere**

Provisioning workloads with exact resources needed

Reconfigurable via software

XILINX®

Composable Data Center Utilizes Resources Efficiently

Standard based framework for provisioning and composing resources

ETHERNET

CPU	CPU
CPU	CPU
CPU	CPU
CPU	CPU
CPU	CPU
CPU	CPU
CPU	CPU
CPU	CPU
CPU	CPU

SSD	SSD	SSD
SSD	SSD	SSD
SSD	SSD	SSD
SSD	SSD	SSD
SSD	SSD	SSD
SSD	SSD	SSD
SSD	SSD	SSD
SSD	SSD	SSD
SSD	SSD	SSD

Accel	Accel	Accel
Accel	Accel	Accel
Accel	Accel	Accel
Accel	Accel	Accel
Accel	Accel	Accel
Accel	Accel	Accel
Accel	Accel	Accel
Accel	Accel	Accel

Workload 1

Workload 2

Workload 3

XILINX®

Main Challenge of Composable Data Center: Latency

Standard based framework for provisioning and composing resources
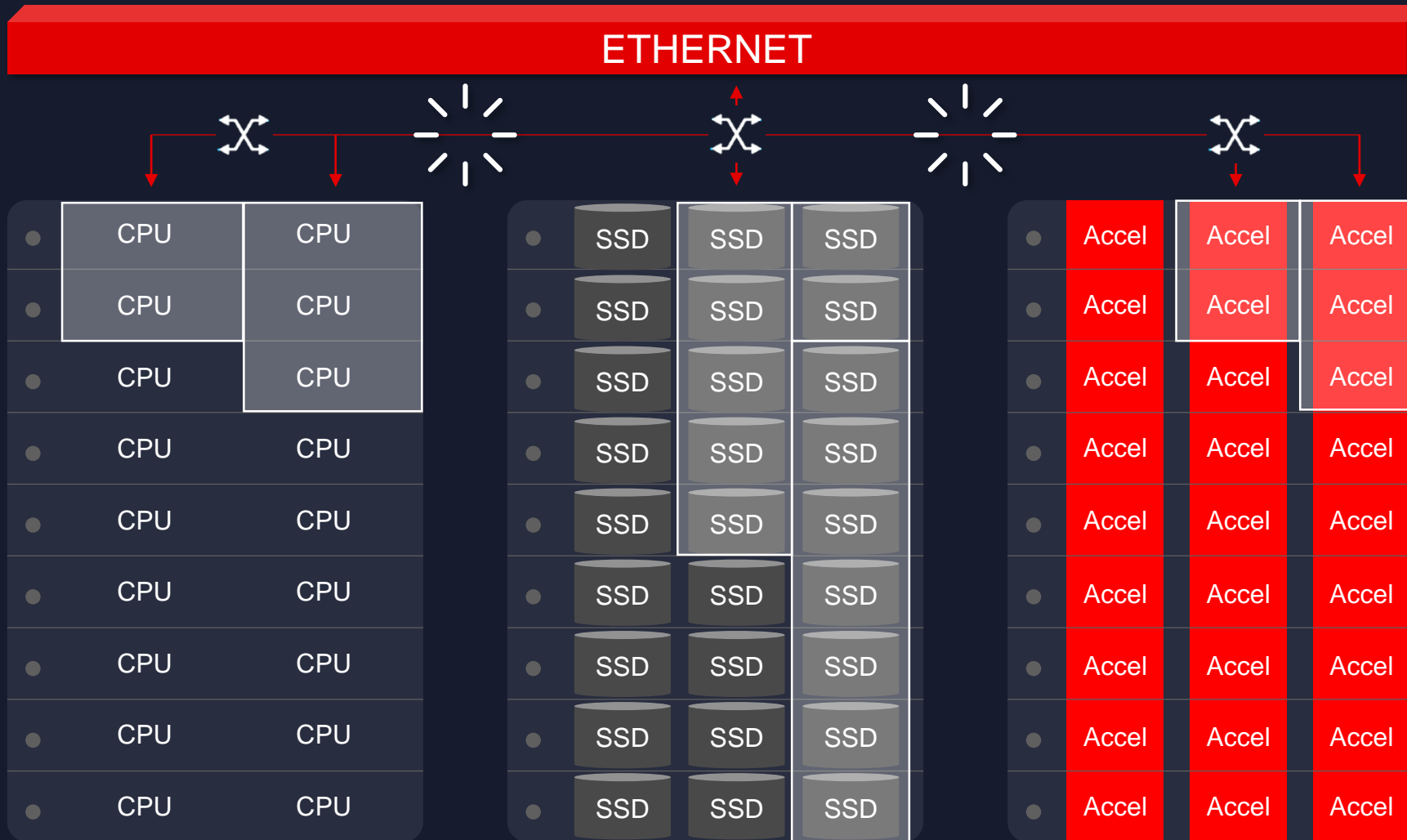
Increased Latency & Reduced Bandwidth

XILINX.

FPGAs Enabling the Composable Data Center

01.

Adaptable & scalable acceleration improves workload performance

02.

Bringing compute closer to data to reduces latency & bandwidth

03.

Adaptable Intelligent fabric enables efficient pooling of resources
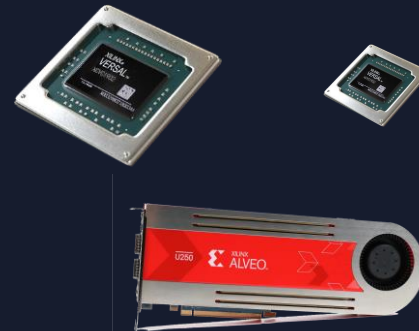
XILINX®

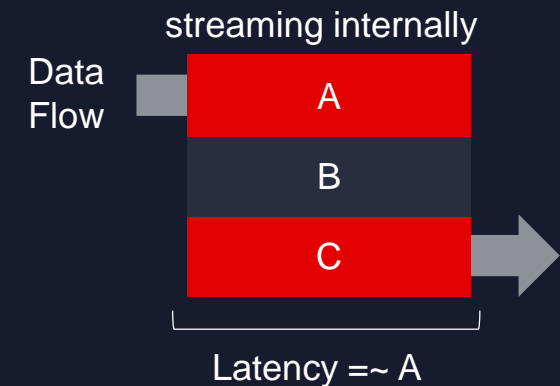# Adaptable & Scalable Acceleration Improves Performance

FPGA advantages in data center

Customizable architecture adapts to specific applications to maximize workload performance

Encryption

Analytics

Decryption
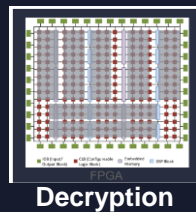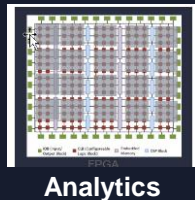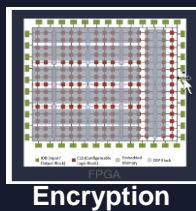
Same architecture can be scaled and algorithms ported to fit different place in the data center

Whole application acceleration enables handling high-speed real-time streaming data flows
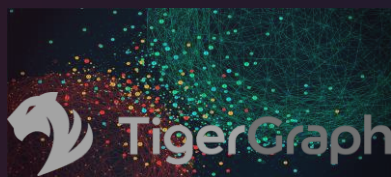
streaming internally

Data Flow

A

B

C

Latency =~ A

XILINX®

FPGAs Enabling Workload Acceleration

Xilinx Smart World
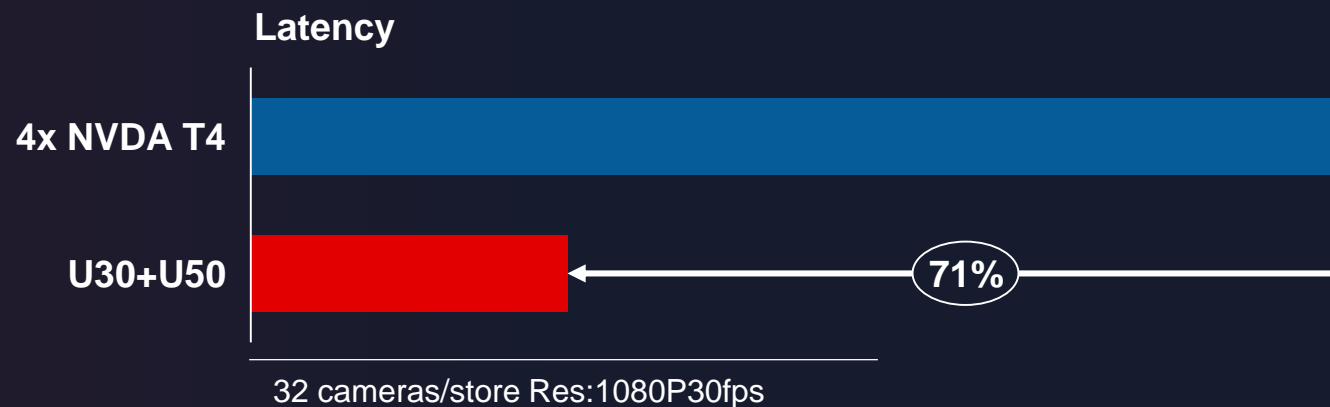
Graph Database Analytics

Latency

4x NVDA T4	
U30+U50	71%

32 cameras/store Res:1080P30fps

Search per sec

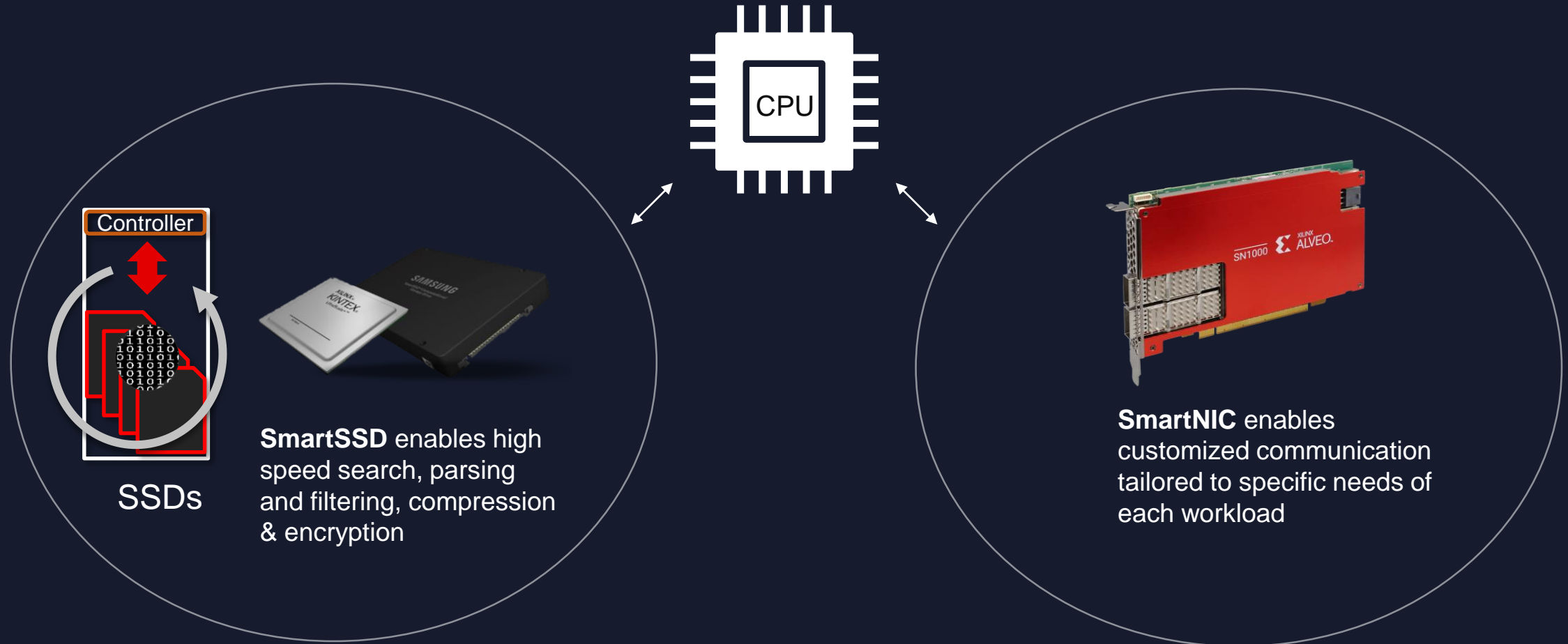|	
CPU*	400x
5x U50	

* CPU: 2x Xeon(R) Silver 4116 @ 2.10GHz, 48 cores, 528 GB RAM

XILINX

Bringing Compute Closer to Data at Rest & In Motion

Reduces Bandwidth Need & Improves Latency
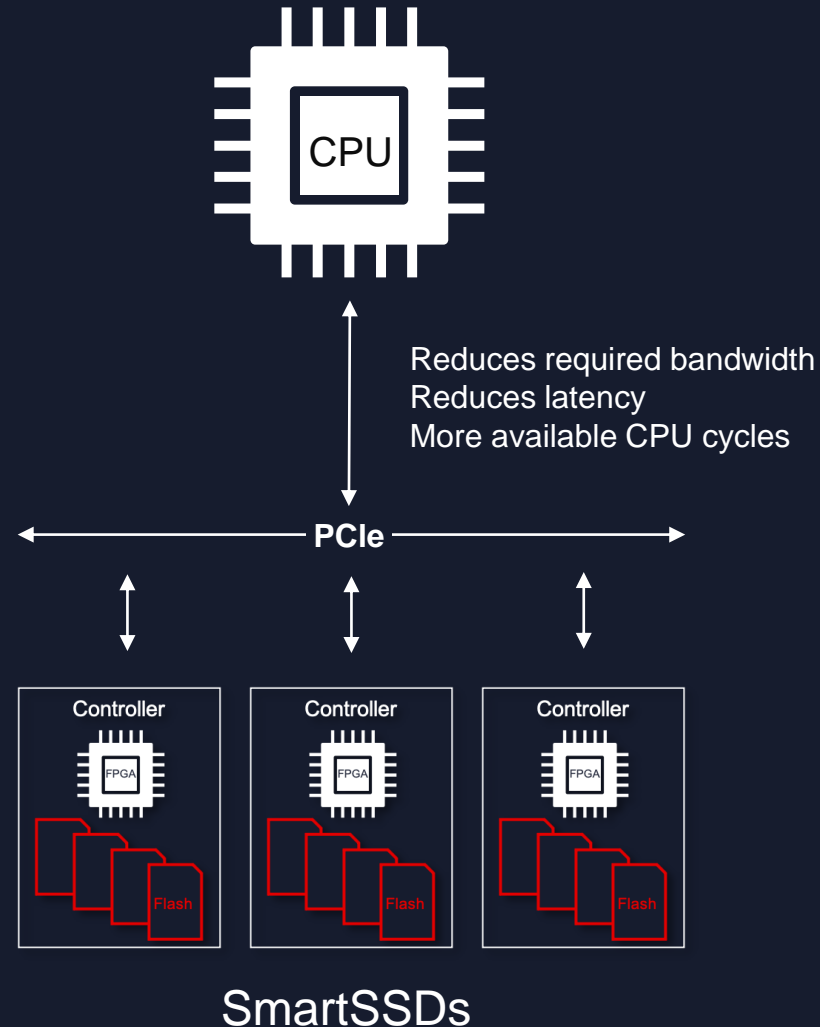

SmartSSD enables high speed search, parsing and filtering, compression & encryption

SSDs

SmartNIC enables customized communication tailored to specific needs of each workload
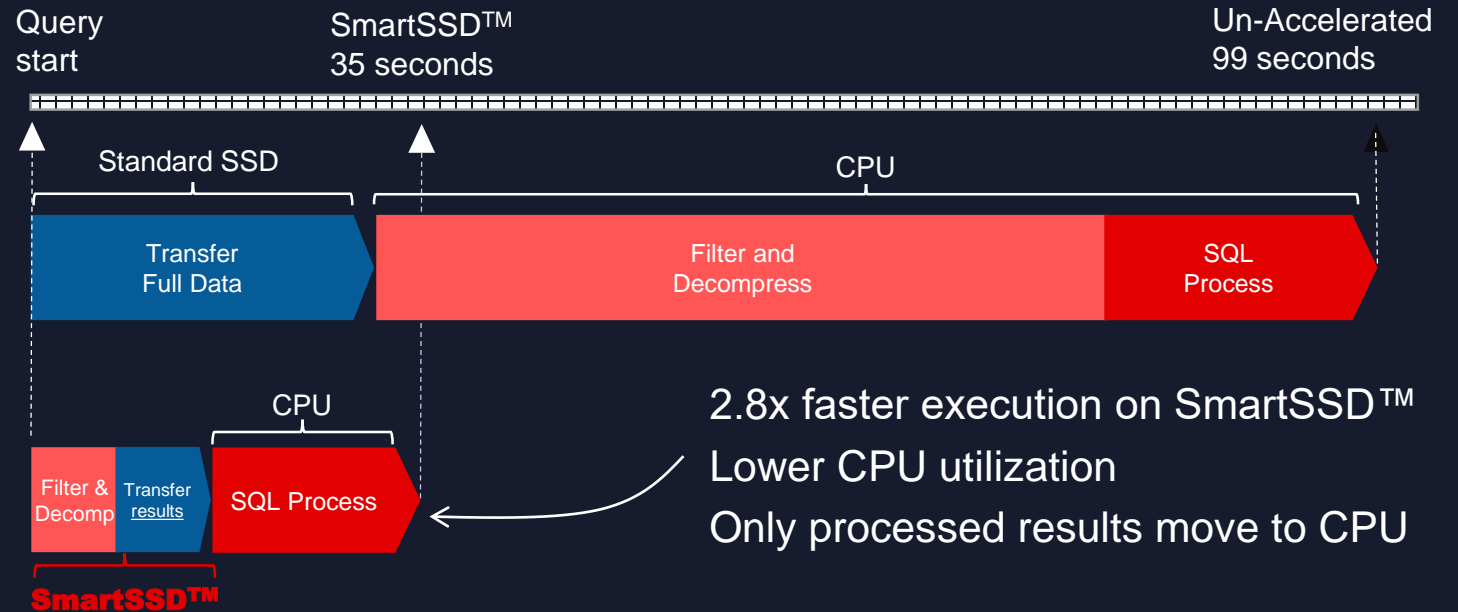
XILINX

Moving Compute Closer to Storage

Storage functions accelerated by FPGAs

- High speed NPU search
- Parsing and filtering
- Compression/Decompression
- Encryption/Decryption

CPU

Reduces required bandwidth
Reduces latency
More available CPU cycles

PCIe

Controller

FPGA

Flash

Controller

FPGA

Flash

Controller

FPGA

Flash

SmartSSDs

XILINX®

SmartSSD Computational Storage Device

Query start

SmartSSD™ 35 seconds

Un-Accelerated 99 seconds

Standard SSD

CPU

| Transfer Full Data | Filter and Decompress | SQL Process |

CPU

| Filter & Decomp | Transfer results | SQL Process |

SmartSSD™

2.8x faster execution on SmartSSD™

Lower CPU utilization

Only processed results move to CPU
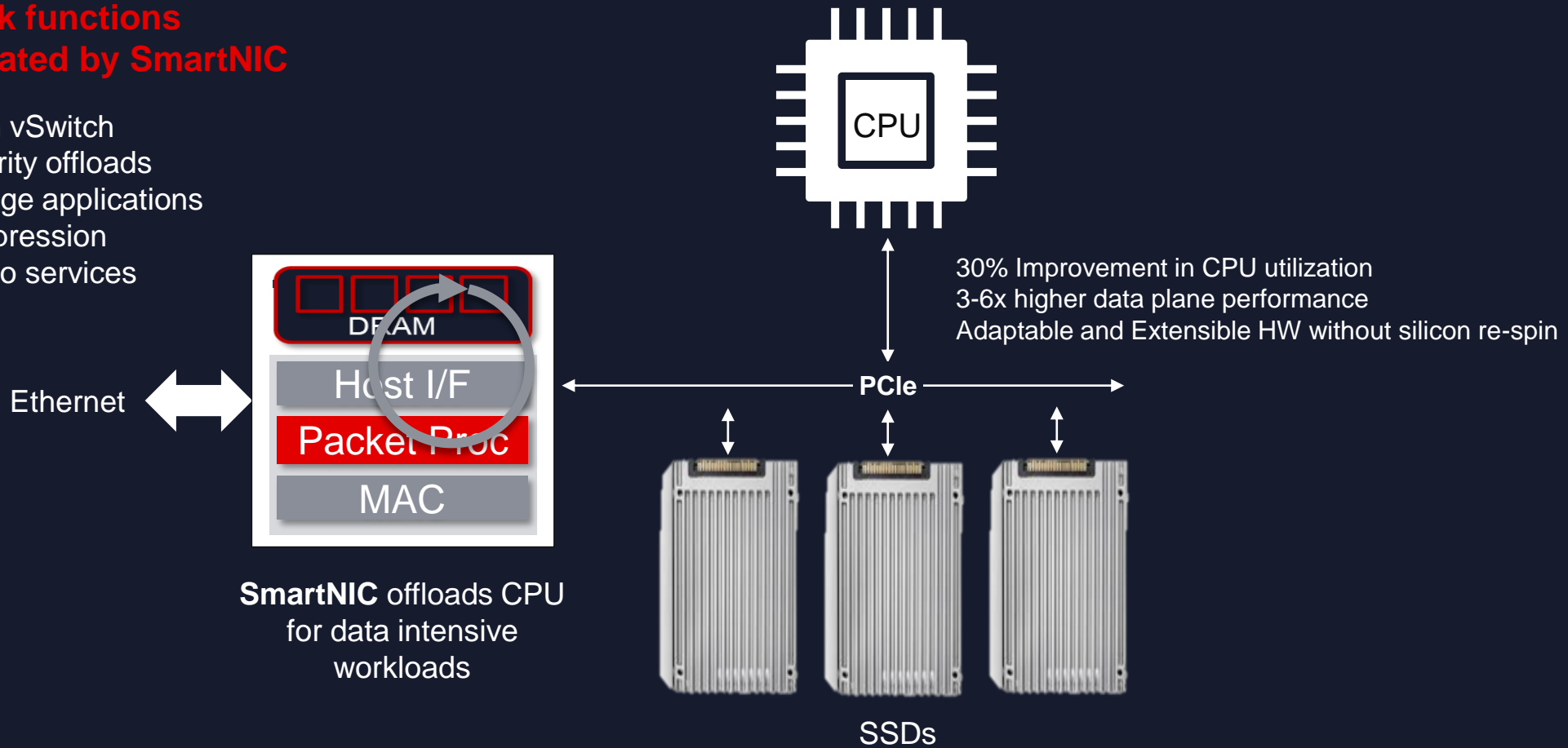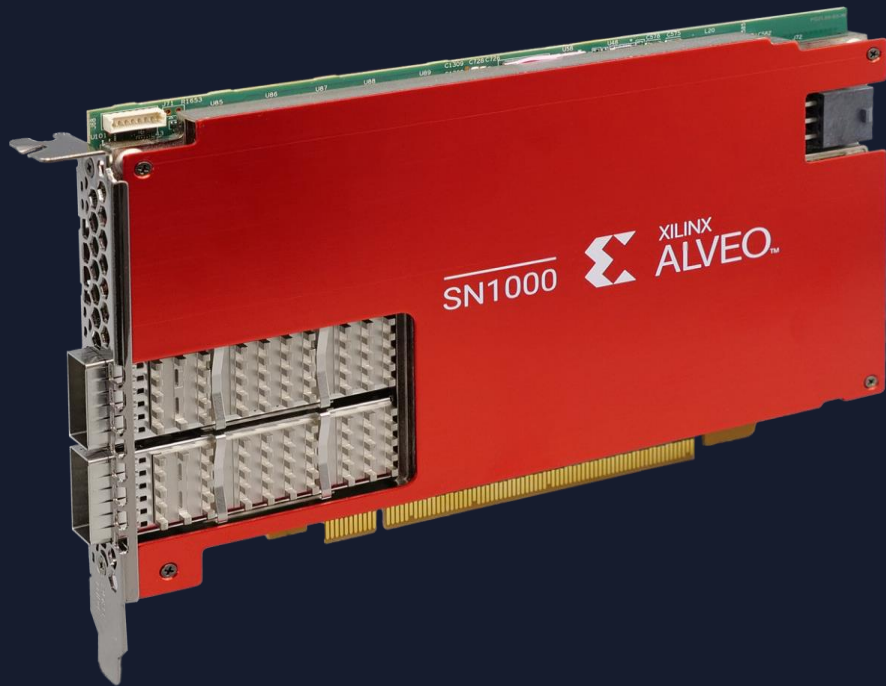
XILINX.

FPGAs as SmartNIC Network Accelerators

Network functions accelerated by SmartNIC

- Open vSwitch
- Security offloads
- Storage applications
- Compression
- Crypto services

Ethernet

DRAM
Host I/F
Packet Proc
MAC

SmartNIC offloads CPU for data intensive workloads

CPU

30% Improvement in CPU utilization
3-6x higher data plane performance
Adaptable and Extensible HW without silicon re-spin

PCIe

SSDs

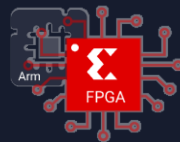XILINX®

Alveo SN1000 SmartNIC

Software-defined hardware acceleration for all offloads

Application specific data paths at line-rate performance

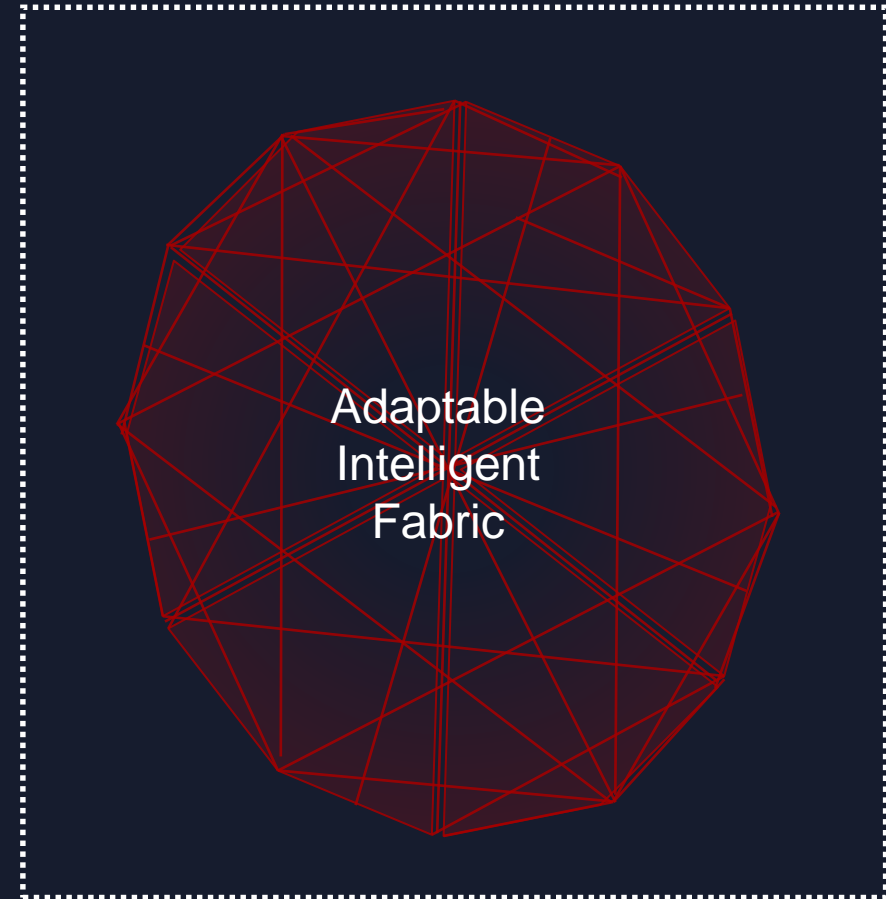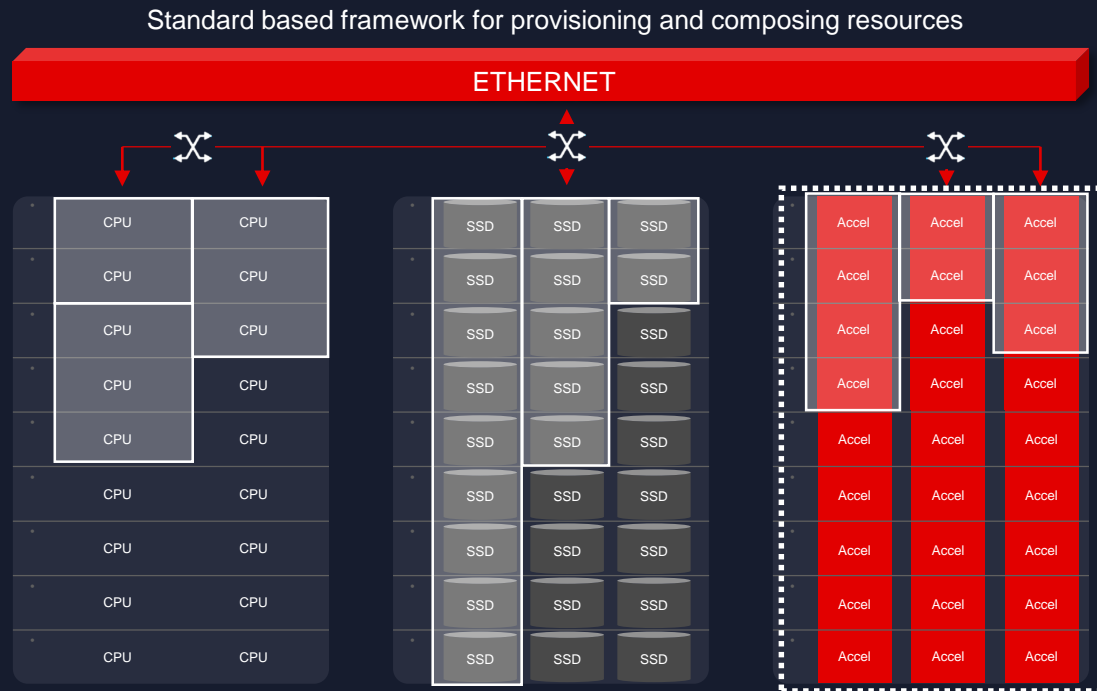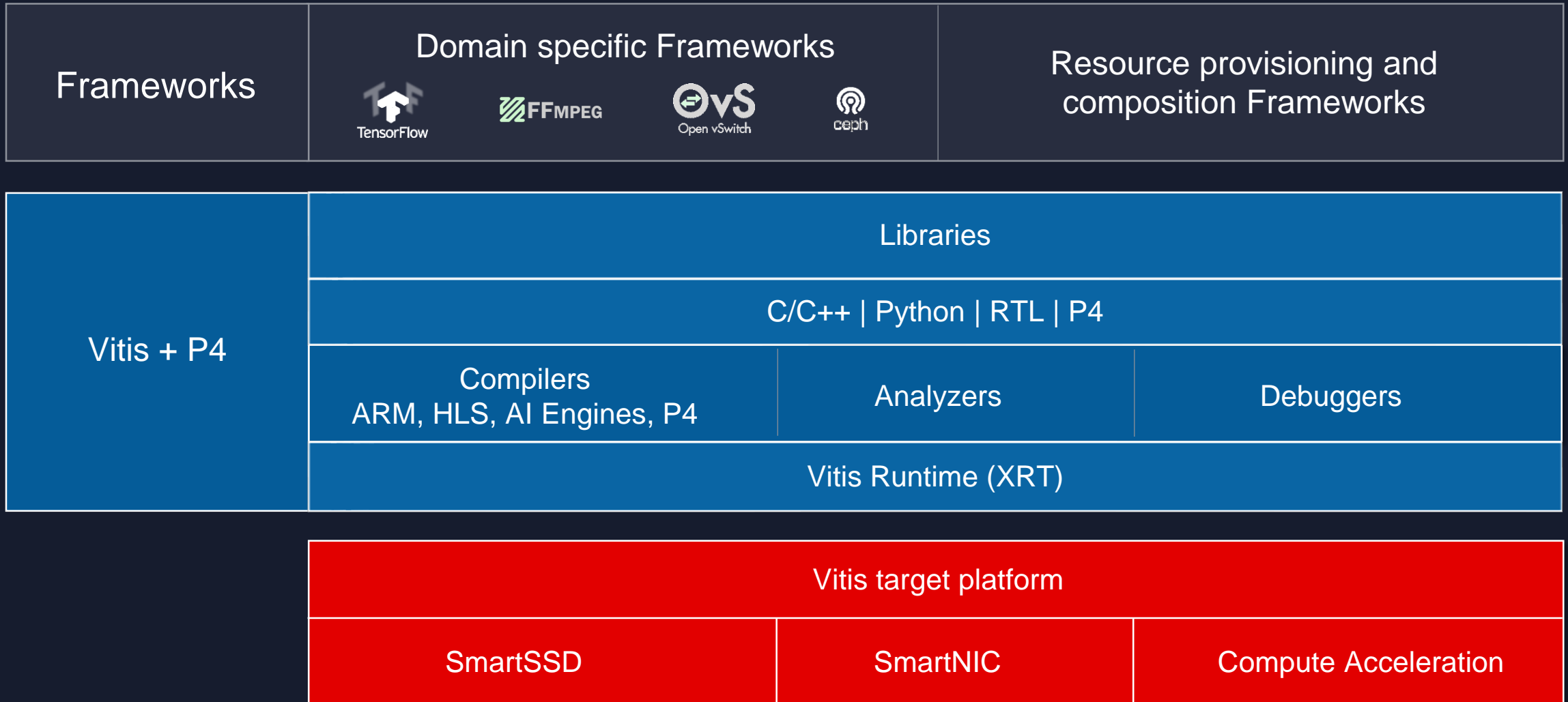<P4> P4, C, C++ programming for fast, adaptable hardware acceleration

Heterogeneous architecture with control and data plane isolation

XILINX

Next Gen FPGAs Enable Adaptable Intelligent Fabric

Local communication among accelerators reduces latency & traffic on the main network

Standard based framework for provisioning and composing resources

Adaptable Intelligent Fabric

XILINX®

Integrated XILINX VITIS + P4 to Program Adaptable Elements

Frameworks	Domain specific Frameworks TensorFlow · FFMPEG · Open vSwitch · ceph	Resource provisioning and composition Frameworks

Vitis + P4	Libraries		
	C/C++ | Python | RTL | P4		
	Compilers ARM, HLS, AI Engines, P4	Analyzers	Debuggers
	Vitis Runtime (XRT)		

Vitis target platform		
SmartSSD	SmartNIC	Compute Acceleration

XILINX

XILINX® Paving the Way to the Composable Data Center of the Future

Rapid Reconfigurability | Reduced Latency | Flexible Architecture

XILINX®