



Xilinx Data Center Spring Launch

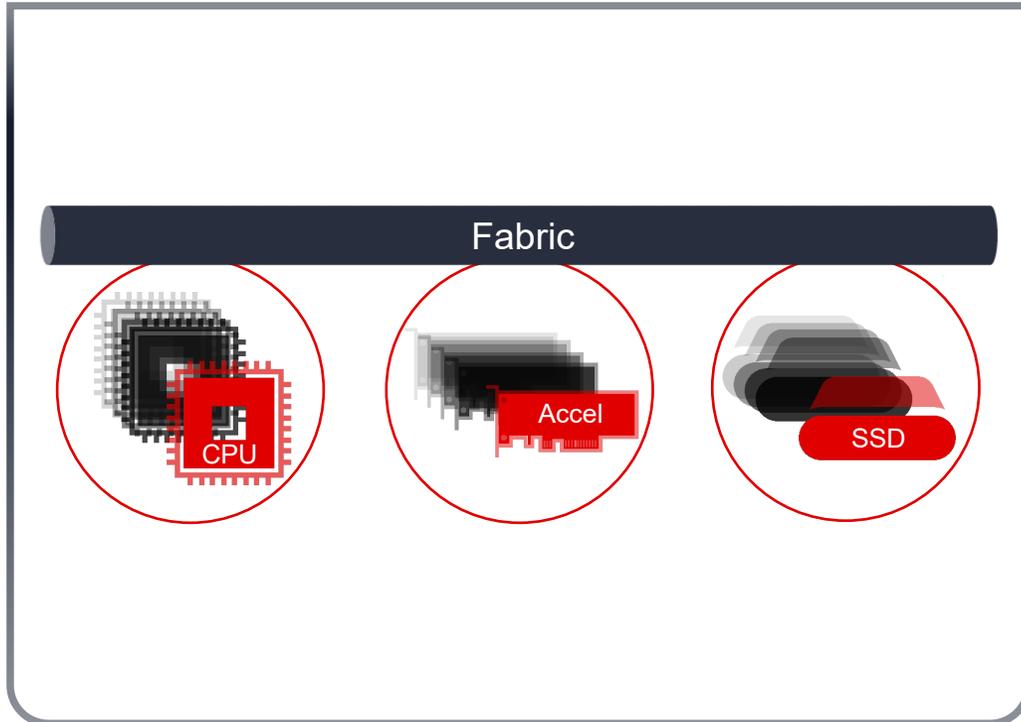
February 2021



The Composable Datacenter

There Is No “Typical” Datacenter

Requires not just disaggregated compute, network, and storage...



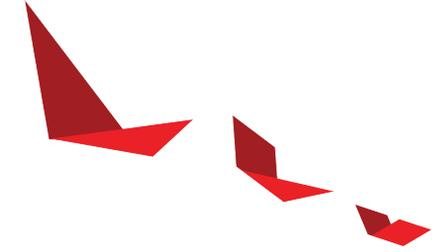
...but composability of the *device* itself

Software Defined



Hardware Accelerated





- ▶ **The Alveo SN1000 Smart NIC** - hardware performance and software adaptability - *Kartik Srinivasan*
- ▶ **Smart World Video Analytics** - An AI video analytics platform and solutions ecosystem built for the most critical applications – *Ed Wright and Guruprasad MP*
- ▶ **Accelerated Algorithmic Trading** - Built on Alveo, enabling a broad spectrum of traders to be newly competitive in high-frequency trading - *Alastair Richardson*

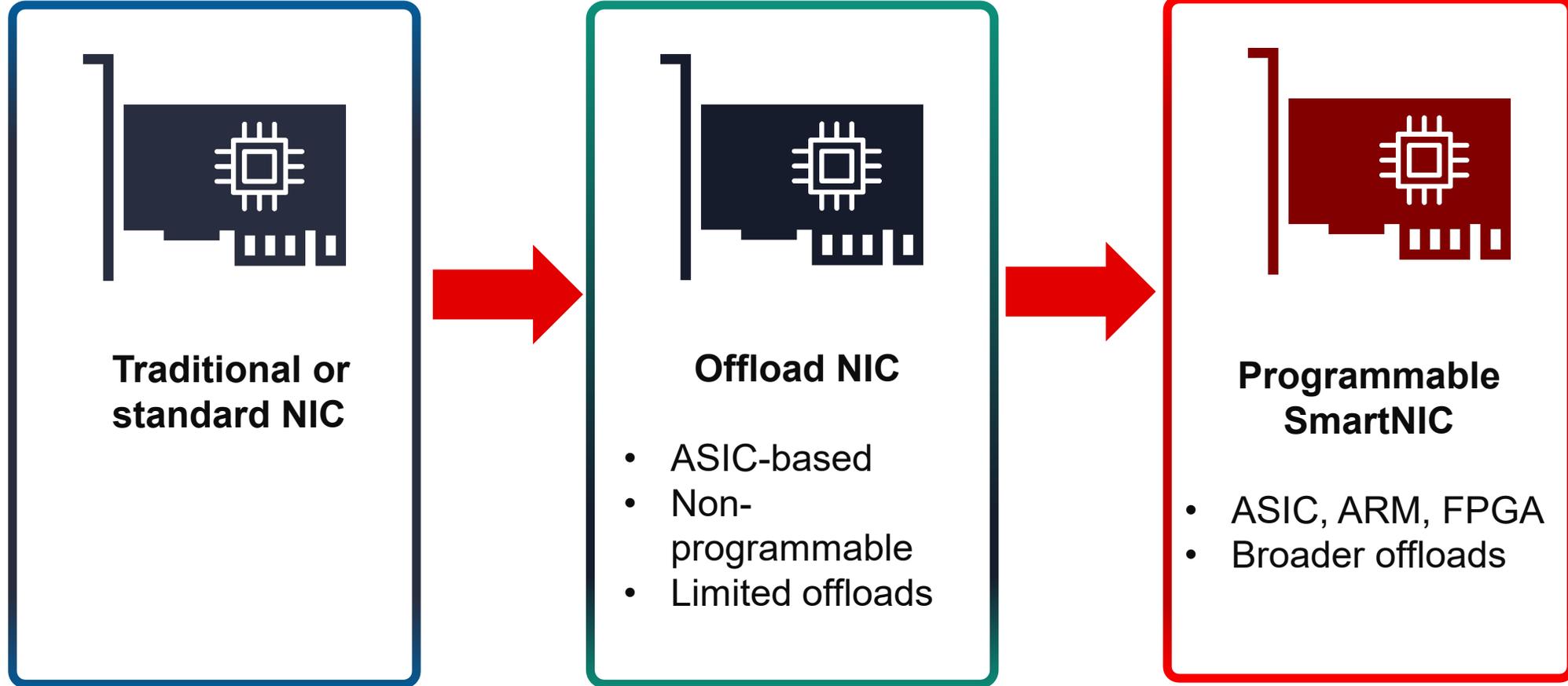


SN1000 Smart NIC

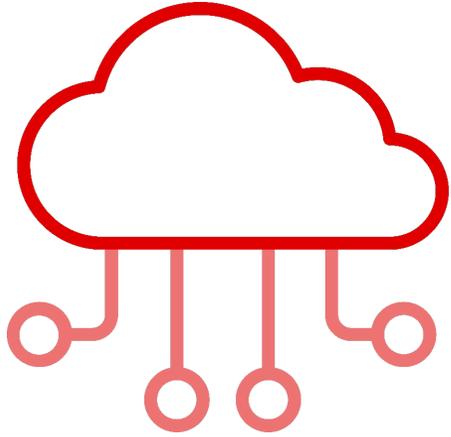
The *Composable* Smart NIC



Evolution of the Smart NIC



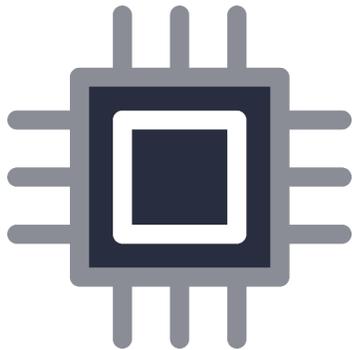
SmartNICs: Emerging Limitations



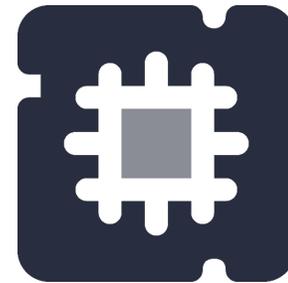
Cloud providers need both performance *and* adaptability

- ▶ Fast pace of change
- ▶ Wide variety of network functions
- ▶ Every hyperscaler and CSP has different needs

BUT...



ASIC implementations
lack customization
capabilities



CPU/SOC
implementations suffer
performance hits at scale

Common Offload Types



NETWORKING



SECURITY



STORAGE

Common Offload Types



OVS | SR-IOV |
VIRTIO.NET | LOAD
BALANCING | NAT |
OVERLAYS | CONNec-
TION TRACKING | VXLAN
| PACKET GENERATION |
TELEMETRY | NVGRE |
GENEVE | VROUTER

NETWORKING



BULK CRYPTO |
IPSEC | SSL/TLS |
KTLS | STATEFUL FIRE-
WALL | MULTIPLE CI-
PHERS | HARDWARE
ROOT OF TRUST |
IDS/IPS

SECURITY



COMPRES-
SION/DECOMPRES-
SION | HASH ACCELE-
RATION | NVME ACCELE-
RATION | NVMEOF | DEDUPLI-
CATION | ERASURE
CODING | FLASH CON-
TROLLER | VIR-
TIO.BLK

STORAGE

Introducing The Xilinx Alveo SN1000



The Adaptability of Xilinx
FPGAs



Isolated Control and Data
Planes



Vitis Networking:
P4, C, C++ Programming

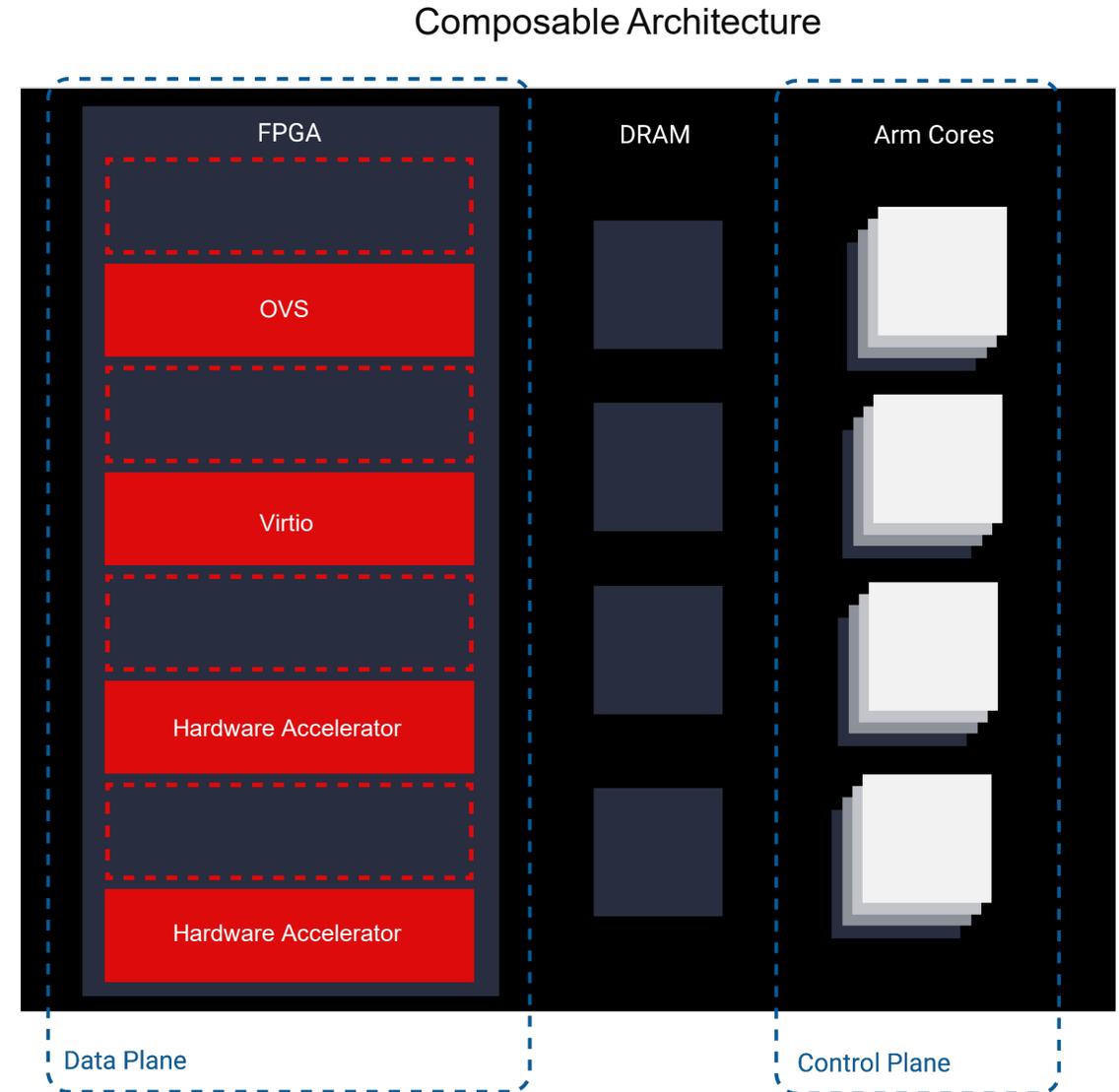
**The Industry's First SmartNIC
with Composable Hardware**



2x 100 Gb SmartNIC	
Hardware	
FPGA	XCU26 Xilinx UltraScale+
Network Interfaces	2x QSFP28
PCI Express	PCIe x16 lanes (Gen 3 x 16, Gen 4 x 8)
On-board CPU	16-core NXP Arm SoC
Performance	
Packet Rate (64Byte)	100Mpps
TDP	70W
Look-up Tables (LUTs)	1M LUT FPGA fabric
Physical Dimensions	
Form Factor	FHHL PCIe

Introducing The Xilinx Alveo SN1000

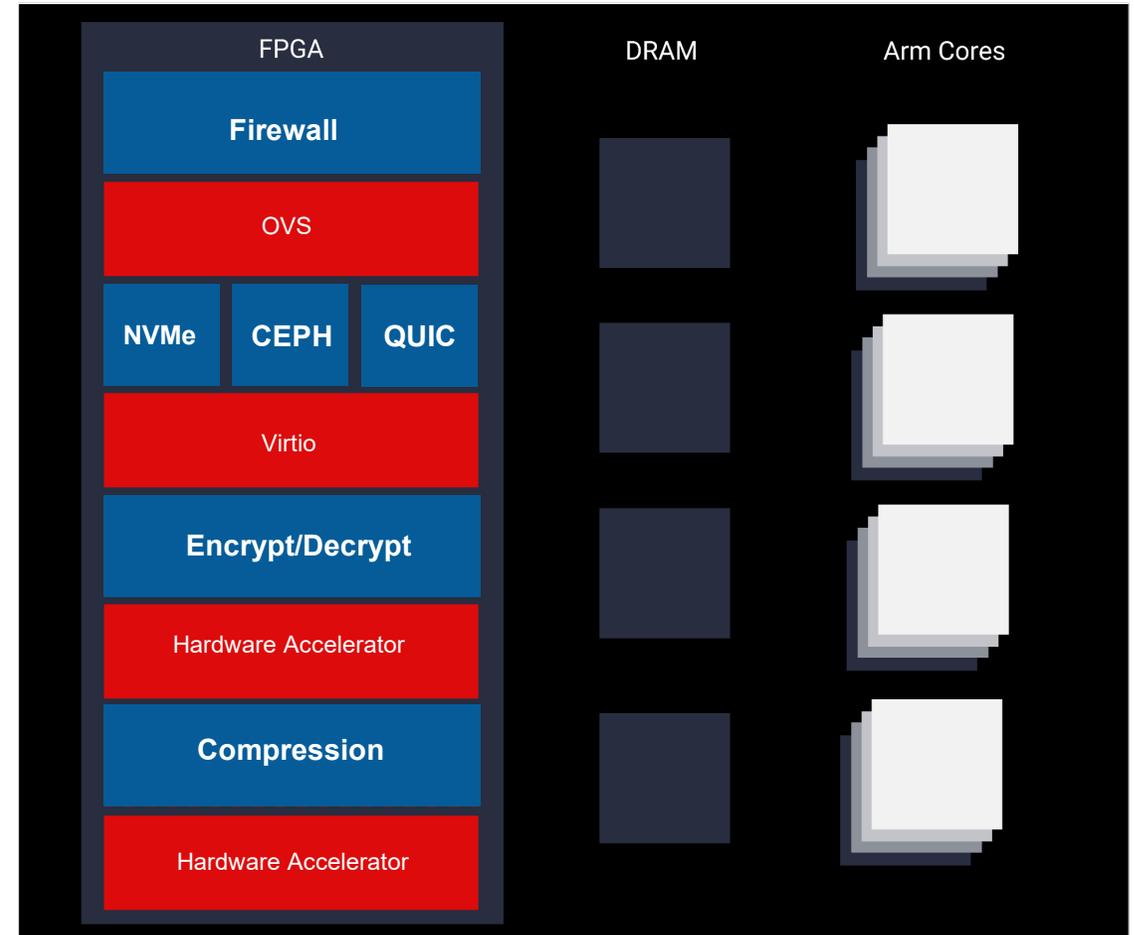
- ▶ **Software-defined hardware acceleration**
- ▶ Application specific data paths
- ▶ Build **custom offloads** or **extend existing offloads** to handle new protocols and applications



Introducing The Xilinx Alveo SN1000

- ▶ **Software-defined hardware acceleration**
- ▶ Application specific data paths
- ▶ Build **custom offloads** or **extend existing offloads** to handle new protocols and applications

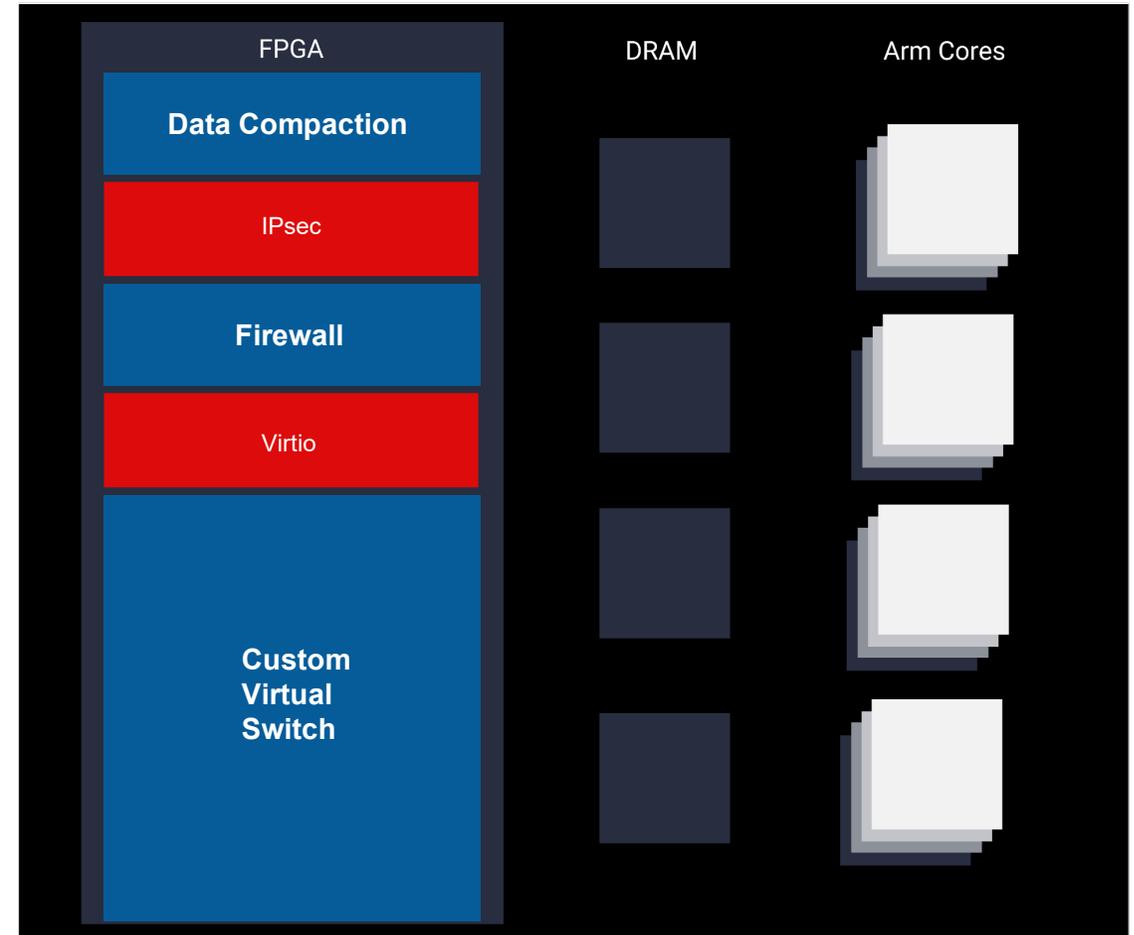
Composability Example 1



Introducing The Xilinx Alveo SN1000

- ▶ **Software-defined hardware acceleration**
- ▶ Application specific data paths
- ▶ Build **custom offloads** or **extend existing offloads** to handle new protocols and applications

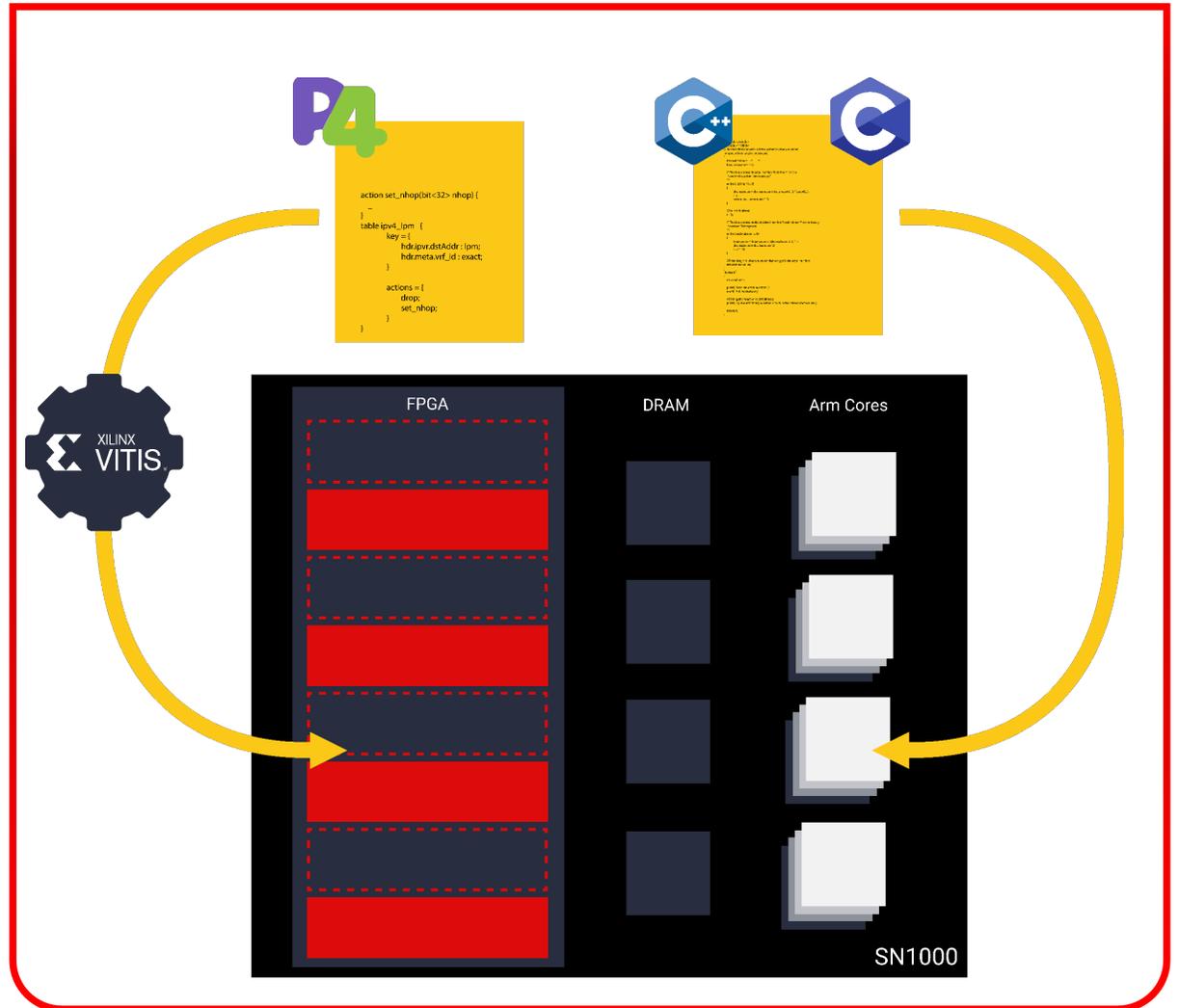
Composability Example 2



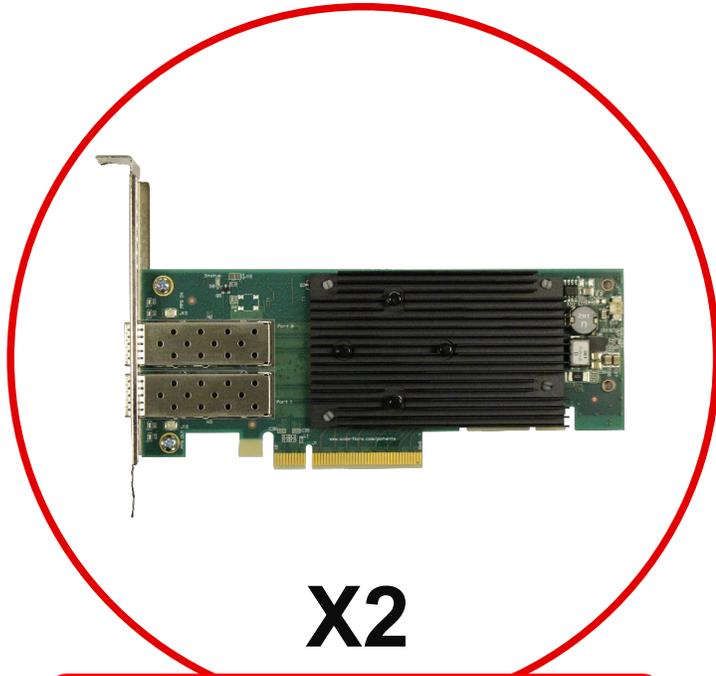
Vitis Networking

- ▶ **Customize with ease, without sacrificing performance**
- ▶ **P4: the perfect match for “Match-Action” processing**
 - Tailored for high-performance networking
 - Includes high performance algorithmic CAM technologies
- ▶ **Vitis RTL/HLS-** Mature developer tools for any compute or storage offloads at HW speeds with powerful high level language support
- ▶ **Xilinx SmartNIC Plug-In Framework**
 - Customizations can be easily embedded into the powerful SN1000 SmartNIC flow

Software-Defined Hardware Acceleration



Xilinx NIC Family



X2

10/25/100Gb Offload NIC



U25

25Gb SmartNIC



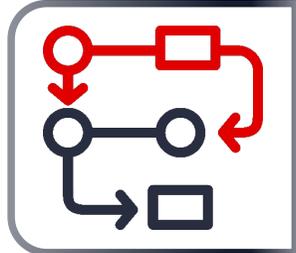
SN1000

100Gb SmartNIC

The SN1000 Difference



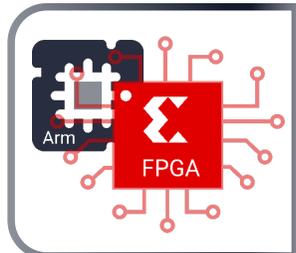
Software-defined hardware acceleration for all offloads



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



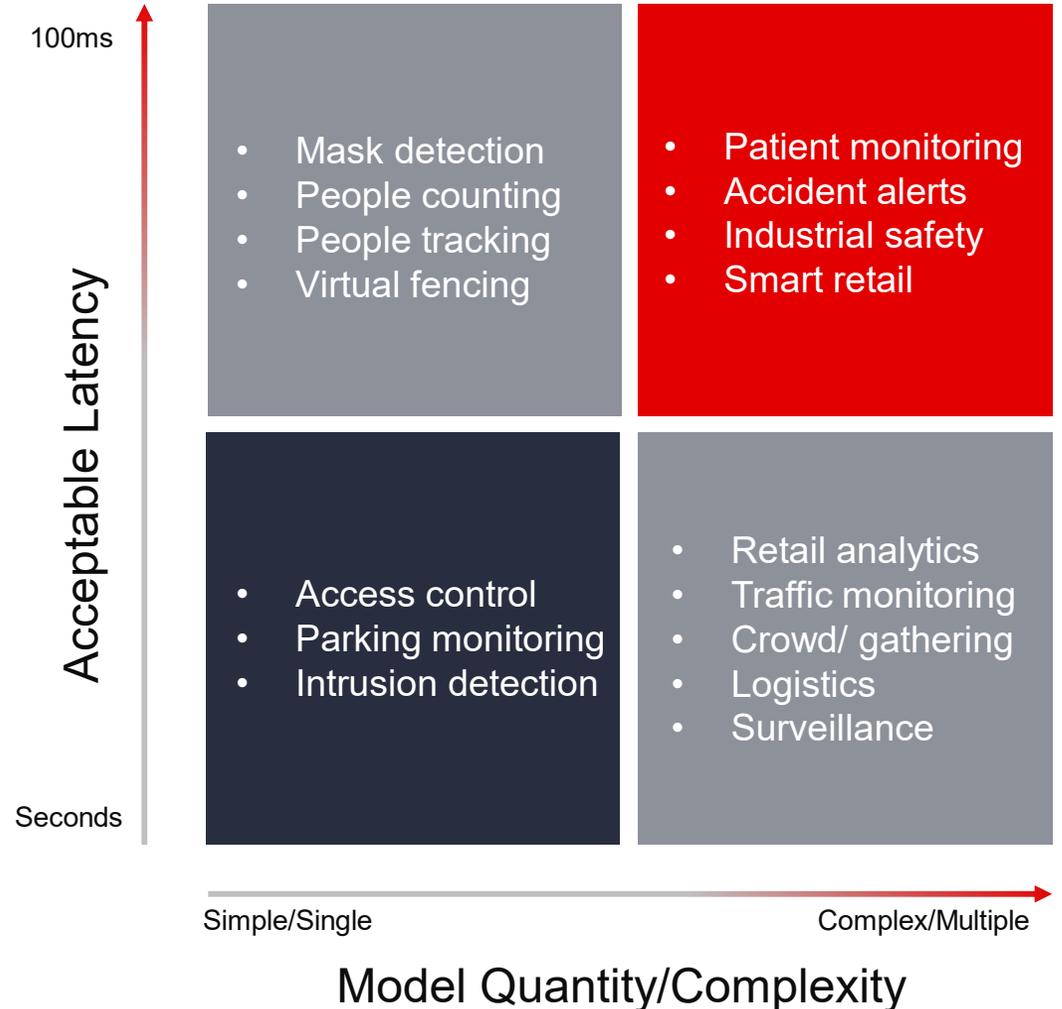
Xilinx Smart World

The Platform for Critical AI Video Analytics



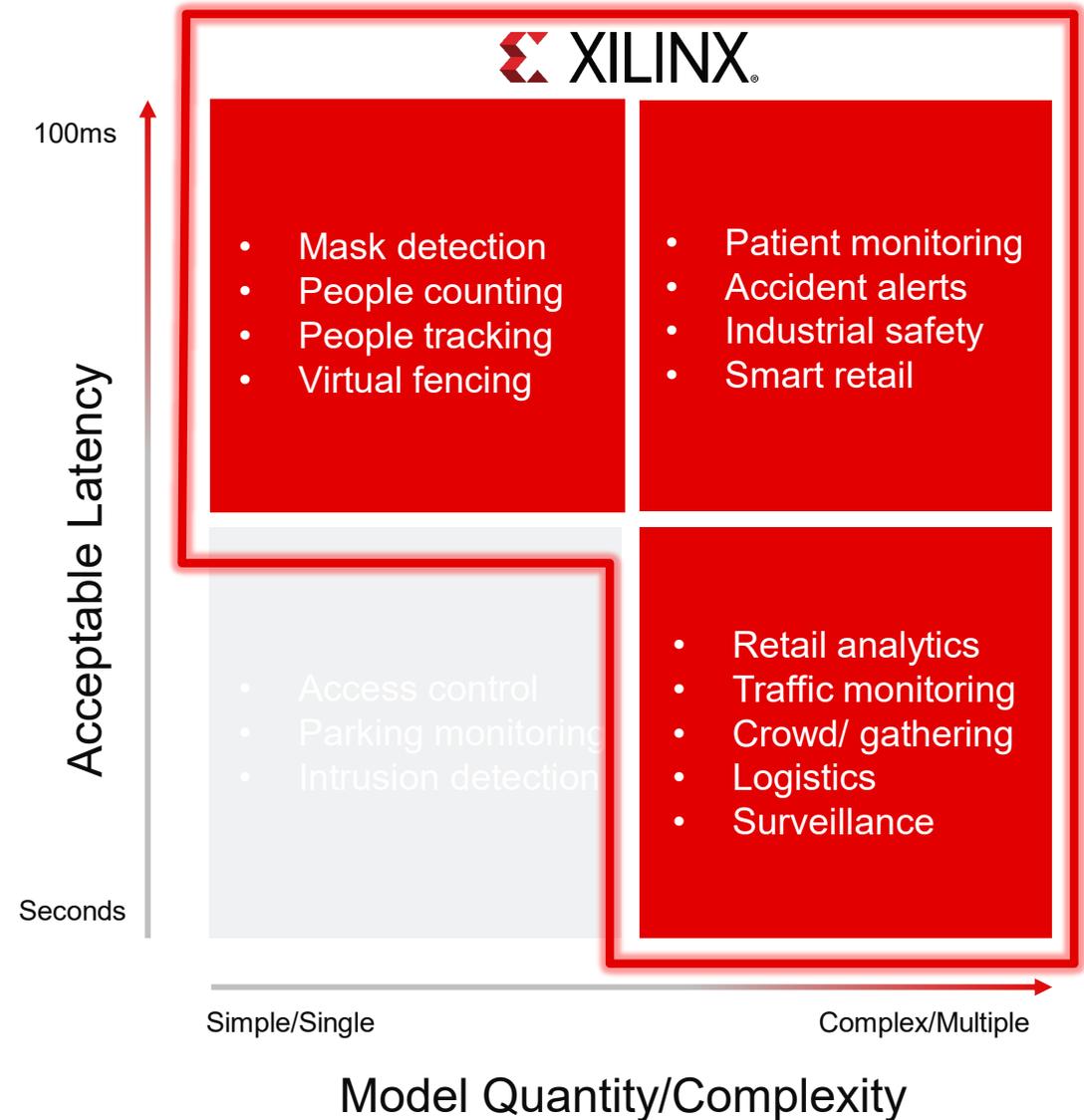
Critical Video Analytics Apps

- ▶ The most critical AI video analytics applications are those that protect human life, health and property
- ▶ These applications are increasingly complex, and complexity puts pressure on architecture
 - Deterministic low latency becomes more difficult to achieve
 - OPEX (space, power) and CAPEX (hardware costs) skyrocket with increased complexity

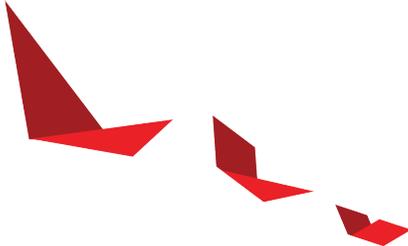


Introducing Xilinx Smart World Video Analytics

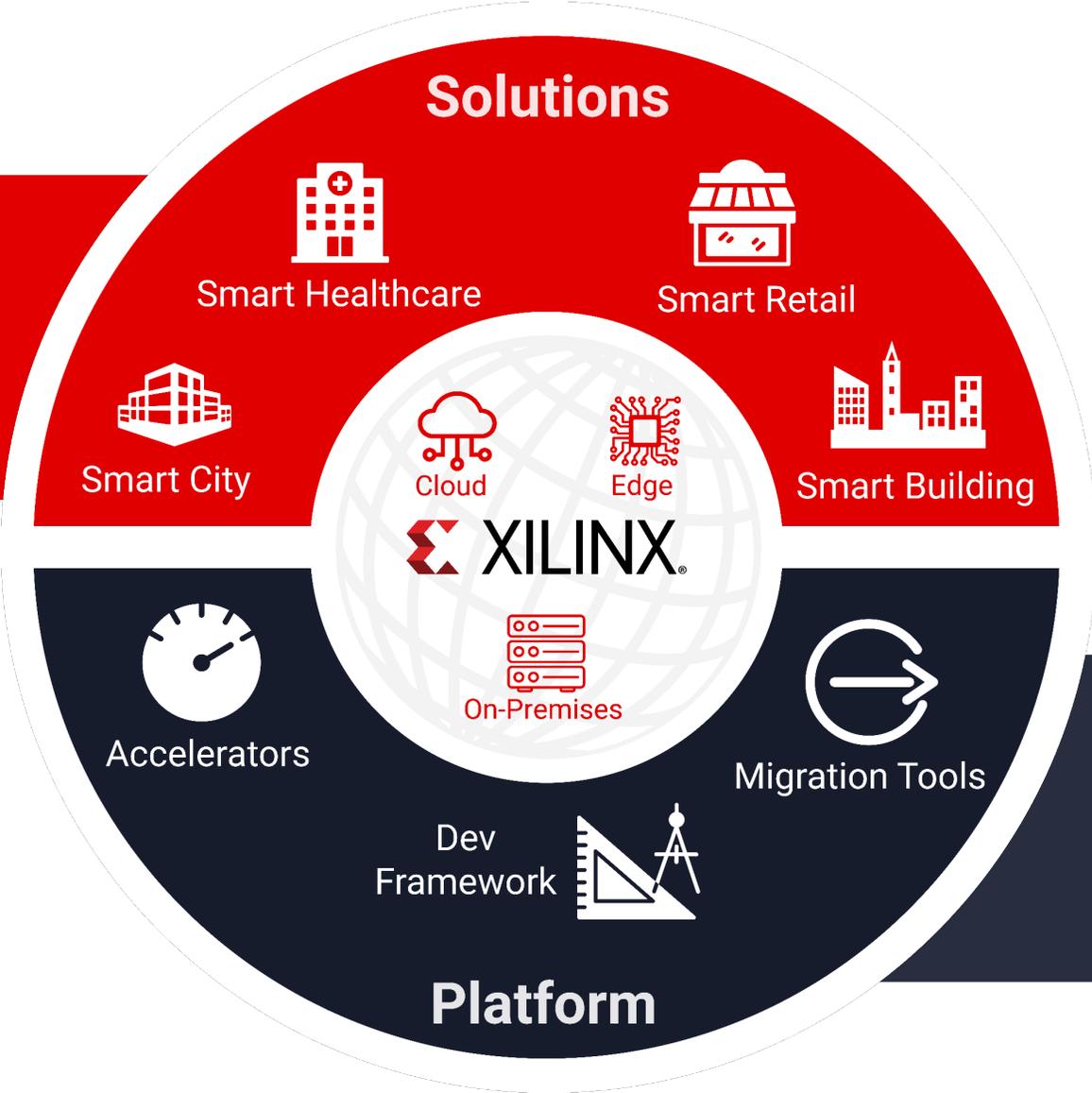
- ▶ Designed to optimize performance for the most complex AI video challenges
- ▶ Massively parallel to handle multiple models elegantly on minimal hardware
- ▶ Deterministic sub-100ms pipeline latency
- ▶ Built on Xilinx's proven Alveo accelerator cards
- ▶ The industry's lowest TCO



Xilinx Smart World Video Analytics



Customer Focus
An ecosystem of solutions, ready to deploy for critical video AI analytics applications



Developer Focus
The VMSS platform enables partners and developers to deliver low latency solutions and plugins for complex AI inferencing

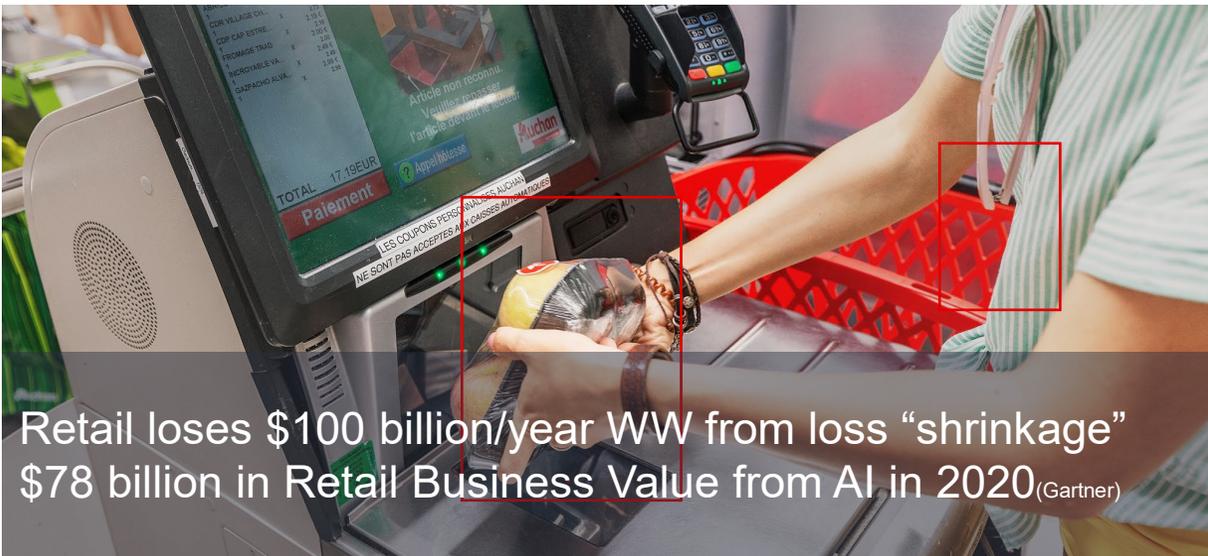
Low Latency Inference Market Drivers



US work-related injuries cost \$171 billion in 2019
105 million days of work were lost



The global incident and emergency management market is expected to grow from \$117.2 billion in 2020 to \$156.1 billion by 2025



Retail loses \$100 billion/year WW from loss “shrinkage”
\$78 billion in Retail Business Value from AI in 2020 (Gartner)



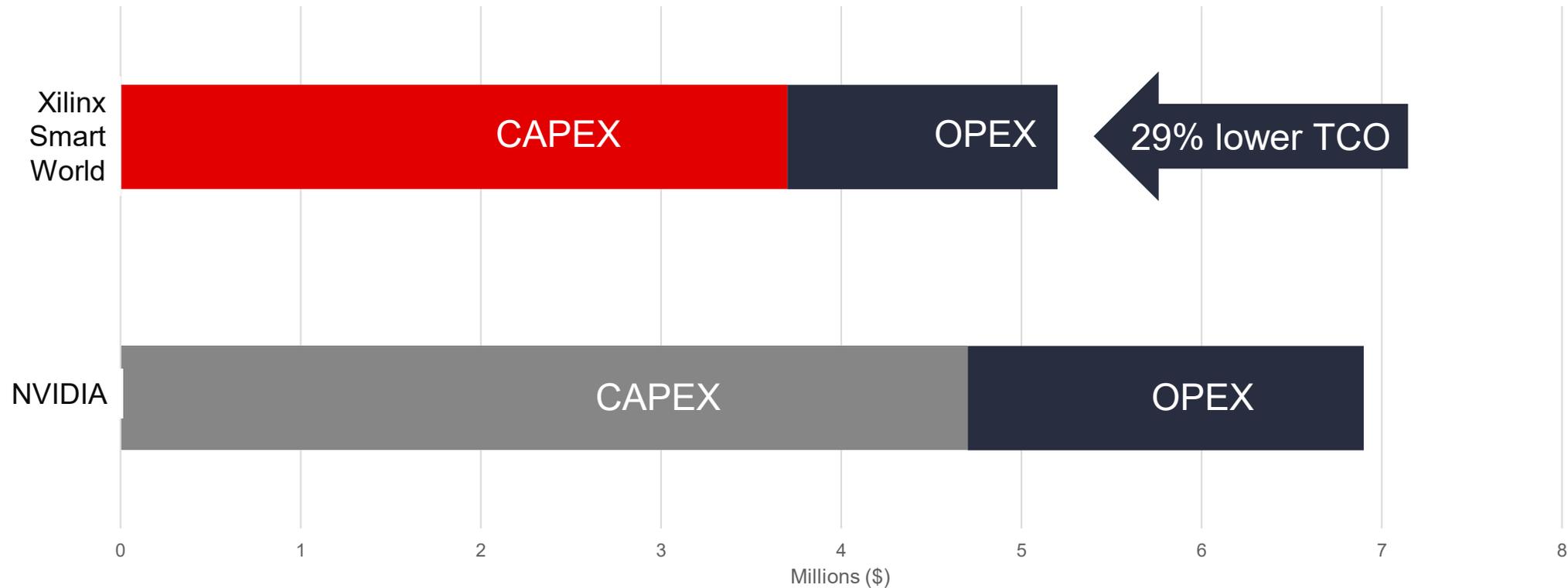
Economic burden for global critical care is \$270 billion
Nursing staff represents 30% of the total cost
ICU mortality in developed countries is still close to 20%

Xilinx Smart World TCO Advantage

Two Xilinx Alveo cards vs 4 Nvidia T4s for 32 cameras

One server per location vs two for Nvidia

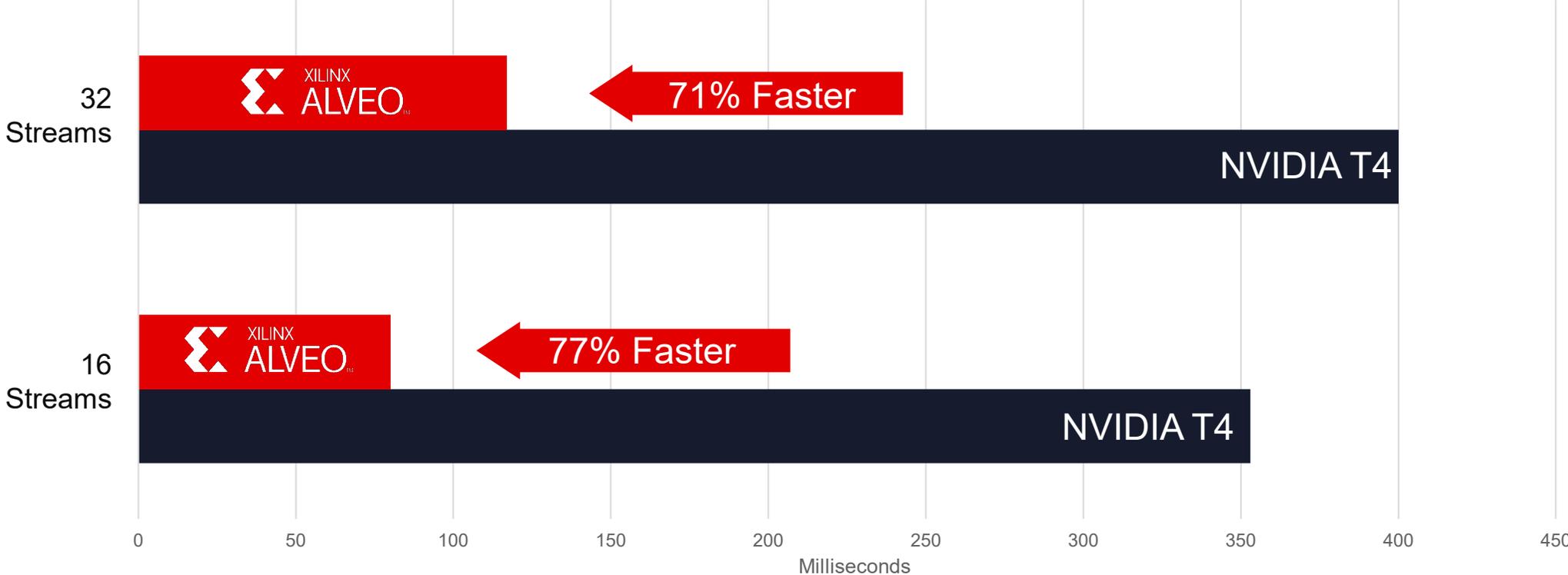
Lower maintenance, power, and cooling costs



32 cameras/store Res:1080P30fps
AI model: Resnet 50 & TinyYoloV3
HW: (1 x U30 + 1 x U50) vs 4 X Nvidia T4

Xilinx Smart World Latency Advantage

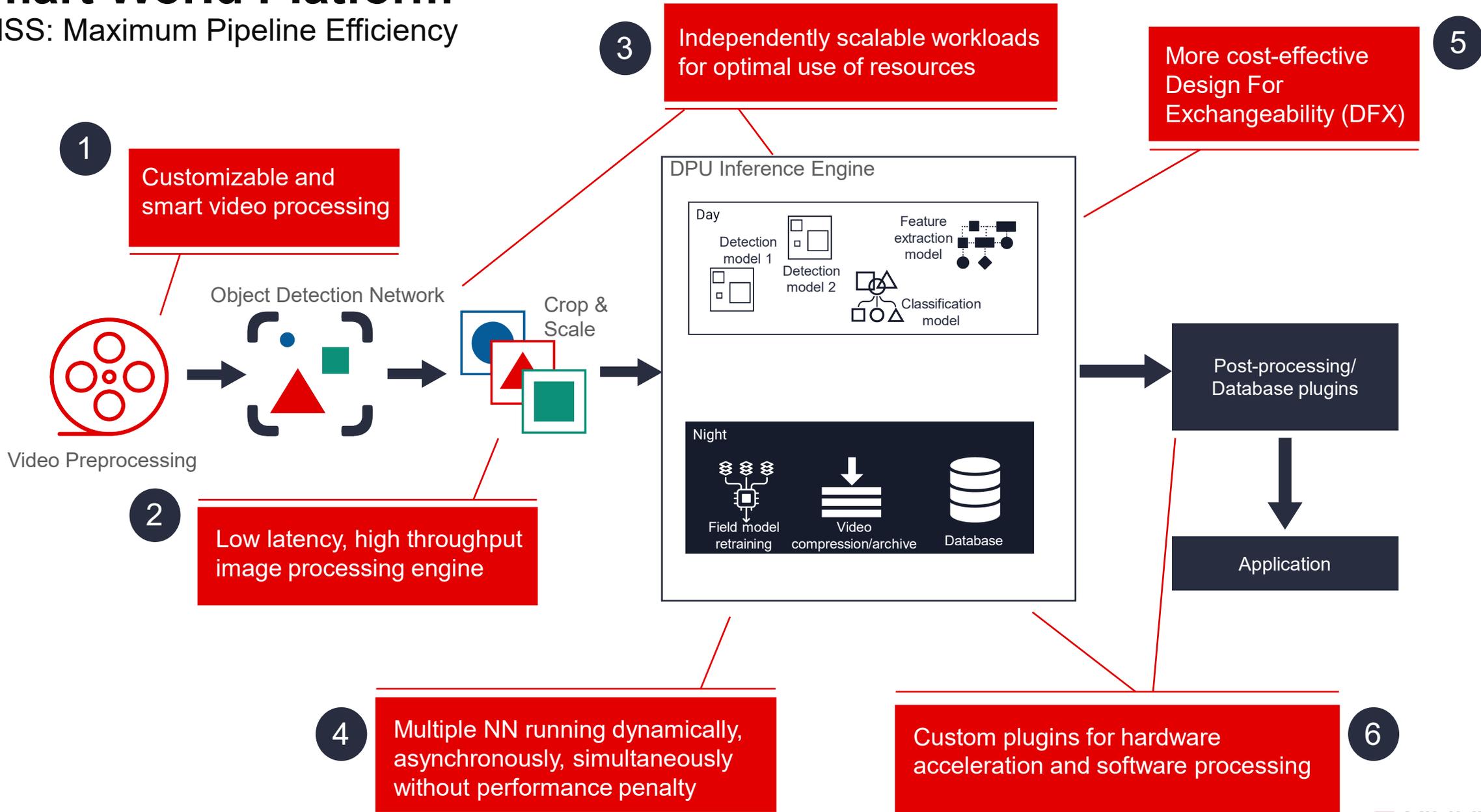
Superior End-To-End Latency



32 cameras/store Res:1080P30fps
AI model: Resnet 50 & TinyYoloV3
HW: (1 x U30 + 1 x U50) vs 4 X Nvidia T4

Smart World Platform

VMSS: Maximum Pipeline Efficiency



Featured Smart World Solutions

Migration and Acceleration

Mipsology

- Toolset delivering easy migration of existing AI applications
- High-performance plug-and-play AI inference accelerator

Edge AI Training

dEEPAI

- AI training at the edge on FPGA with a 10x performance/cost advantage vs GPUs
- Support for Tensorflow, PyTorch and Keras.

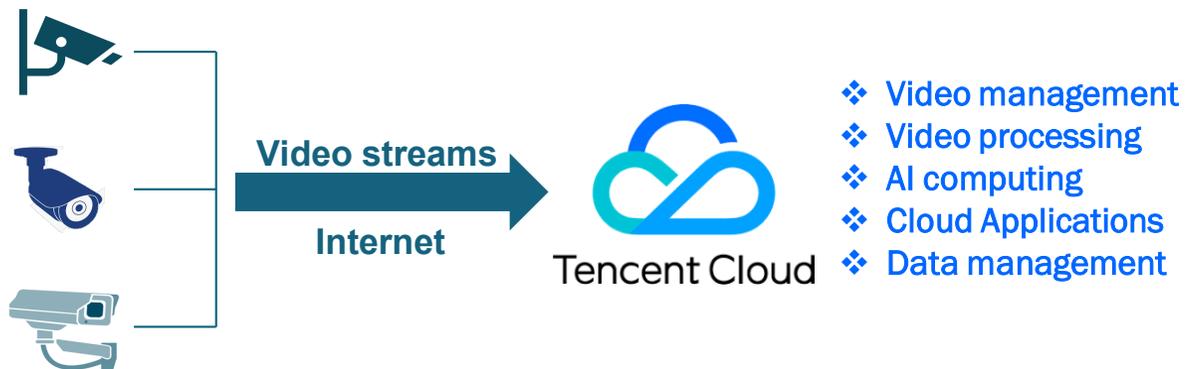
Smart Retail and Smart City

 **Aupera**
Making Video Alive

- Full video AI solutions at the edge
- Smart city, smart building, and smart retail
- High efficiency, low latency, and scalability

Tencent WeLink is an IoT operating system that **monitors, controls and manages** all connected devices within the building.

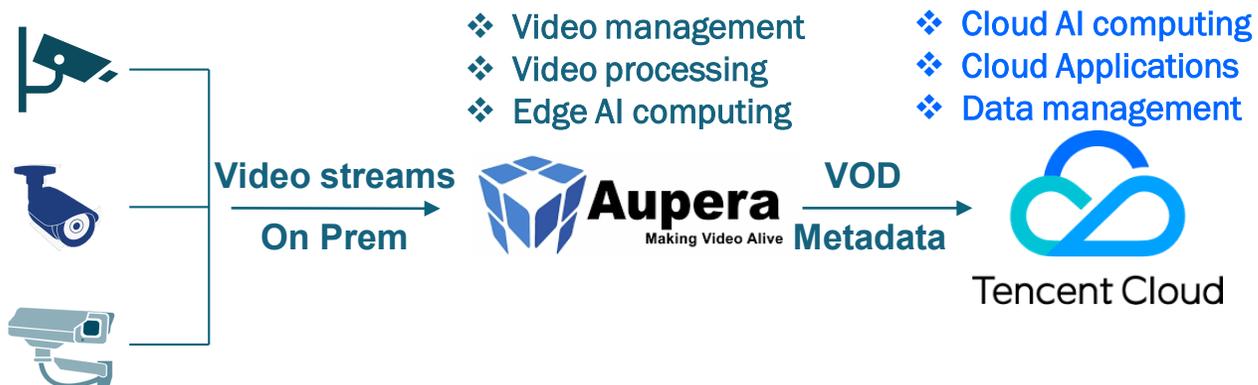
Before



Challenge

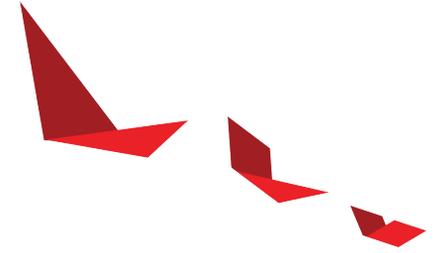
- Expensive bandwidth cost with all cameras streaming to cloud processing
- Non-unified streaming protocols, complicated integration
- Overconsumption of cloud computing resources
- Slow scale deployment

With Aupera - Currently in Deployment to Enable 5,000+ Cameras



- 90%+ bandwidth cost savings on VOD to cloud and AI local processing
- Video gateway to unify different protocols to Tencent Cloud, with seamless integration
- Video pre-processing & Edge AI offloading central cloud computing
- Ultra-low, deterministic latency with instant response
- 100X camera management capacity vs traditional solution
- Remote upgrade with single click

Xilinx Smart World Video Analytics



The lowest TCO and lowest deterministic latency for demanding AI video applications



Xilinx App Store

The One-Stop Shop to Eval and Buy Xilinx Solutions



Introducing The Xilinx App Store

Pre-built, containerized apps that deliver an easy way to evaluate, purchase, and deploy accelerated applications in minutes.

For buyers

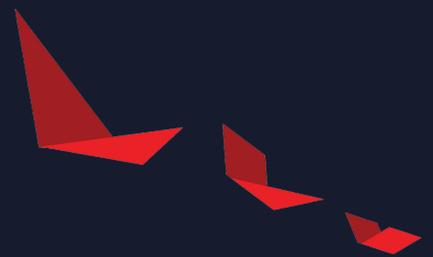
- ▶ 10 minutes to running application on your on-prem Alveo or cloud instances
- ▶ No hardware expertise needed – fully containerized applications with software APIs
- ▶ Flexible deployment options from pay-per-use to perpetual licensing

For sellers

- ▶ Proven secure platform incorporating Digital Rights Management (DRM) IP
- ▶ Global sales / marketing outreach through self-service platform
- ▶ Real-time analytics and lead generation for maximum business intelligence



App Store



Xilinx Accelerated Algorithmic Trading

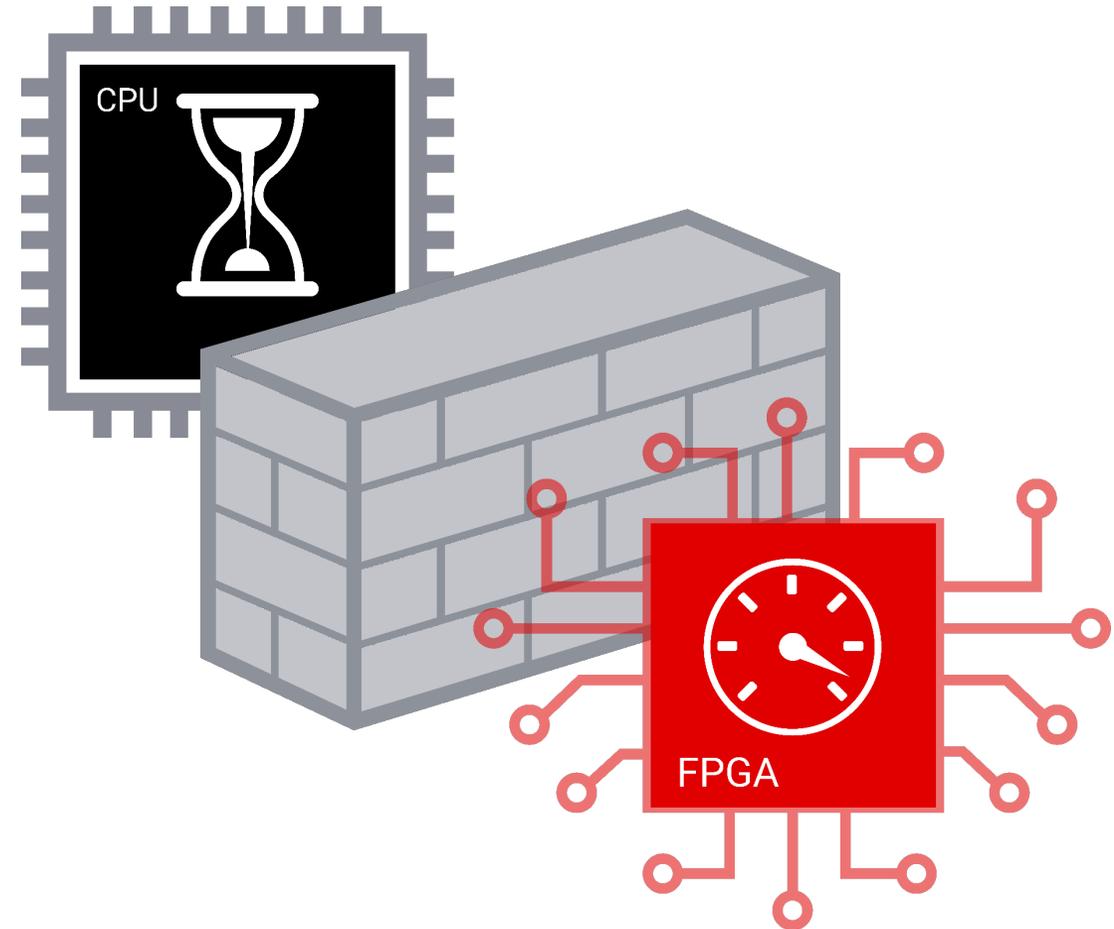
Hardware Accelerated Algorithmic Trading Made Easy

Algorithmic Trading Today

Split into those using hardware and those using software, creating a huge gap in capabilities and performance

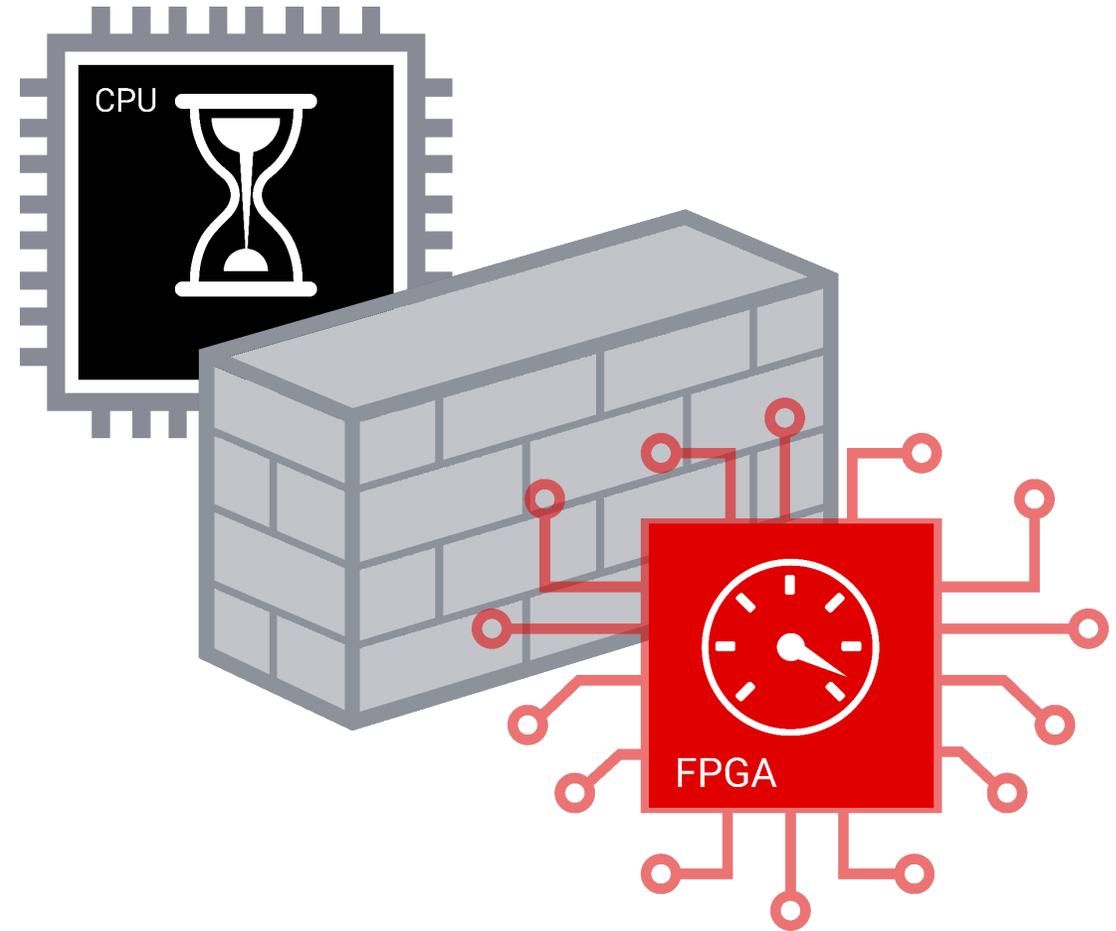
High barriers to entry:

- ▶ **Hardware developers**
- ▶ **High costs**
- ▶ **Long lead times**
- ▶ **High risks**



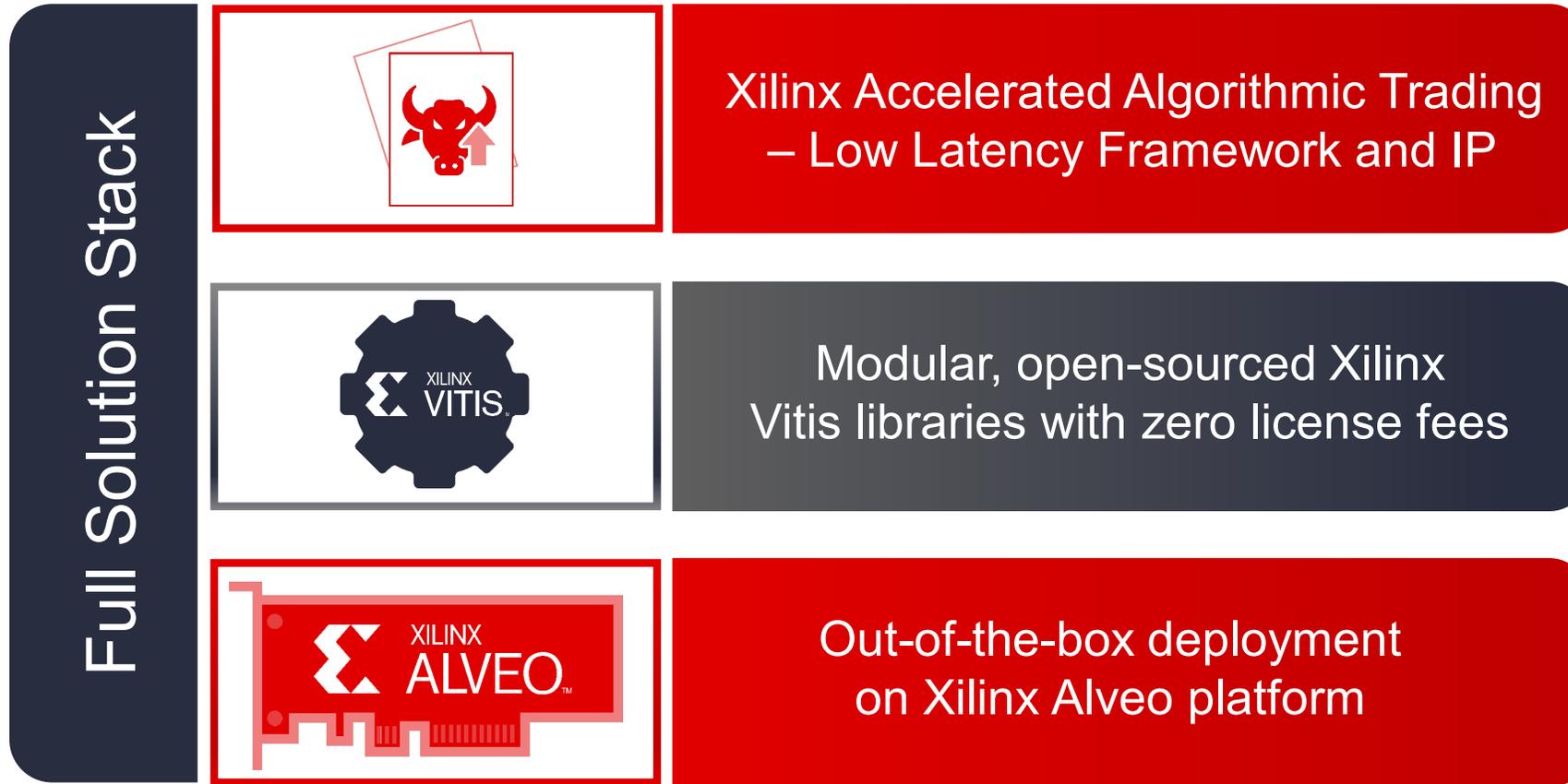
The Need for Lower Latency Trading

- ▶ Failure to compete on latency is costly and negatively impacts Transaction Cost Analysis (TCA)
- ▶ Breaking the microsecond latency barrier gives traders a significant advantage and minimizes losses to High Frequency Trading (HFT)
- ▶ CPU's have hit their limit
 - No longer getting faster
 - Not network connected, PCI slows down trading
- ▶ Entry to HFT market is expensive and not affordable to the wider market



Introducing Xilinx Accelerated Algorithmic Trading

A composable, open-source trading system that enables traders to implement sophisticated strategies with sub-microsecond latency

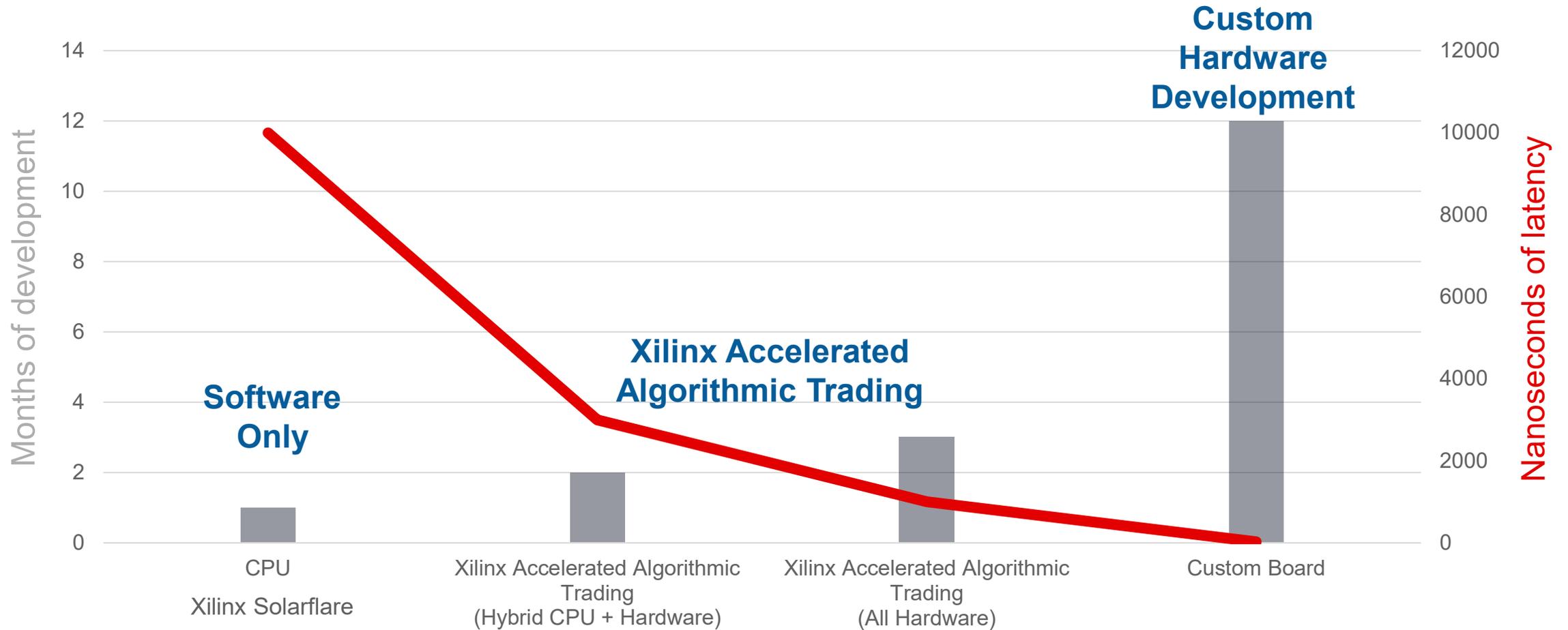


Xilinx Accelerated Algorithmic Trading Features

- ▶ Modular, open-sourced Xilinx Vitis® library with zero license fees
- ▶ Example design for tick to trade on the Chicago Mercantile Exchange (CME)
- ▶ Flexibly compose your trading architecture with a suite of libraries
- ▶ Vitis programmability with C/C++
- ▶ Flexible easy integration with “in-house” or 3rd party apps
- ▶ Accelerates time-to-market – weeks, not years



A Pathway To Lower Latency



Xilinx Accelerated Algorithmic Trading Use Cases

- ▶ Brokers
- ▶ Exchanges
- ▶ Market Data Vendors
- ▶ Sell Side Vendors
- ▶ Proprietary Traders

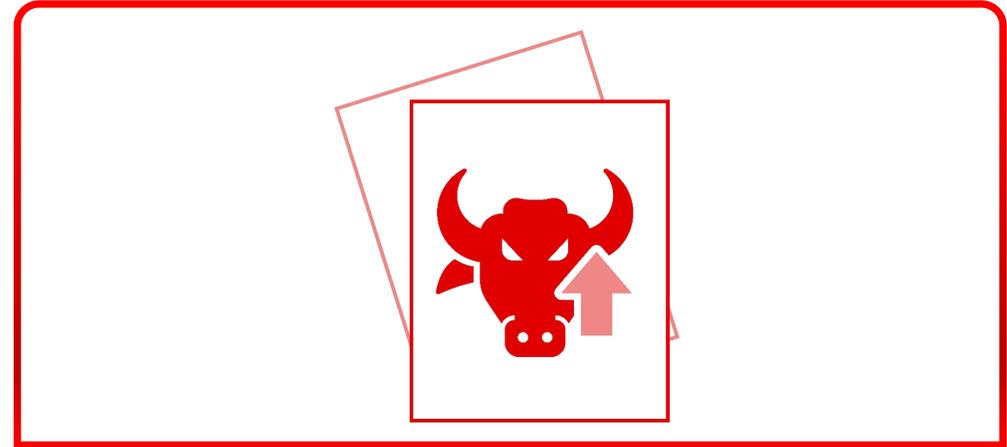


Get Started With Xilinx Accelerated Algorithmic Trading

Regain The Latency Edge



+

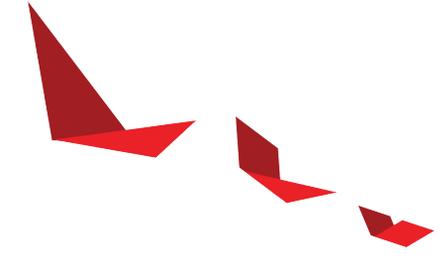


Alveo U50 & U250 - available through
Xilinx VARs and VADs

Free open-source download from
Xilinx.com, no license fees

Xilinx Data Center Group Spring 2021

The Composable Datacenter: Software-defined, hardware accelerated

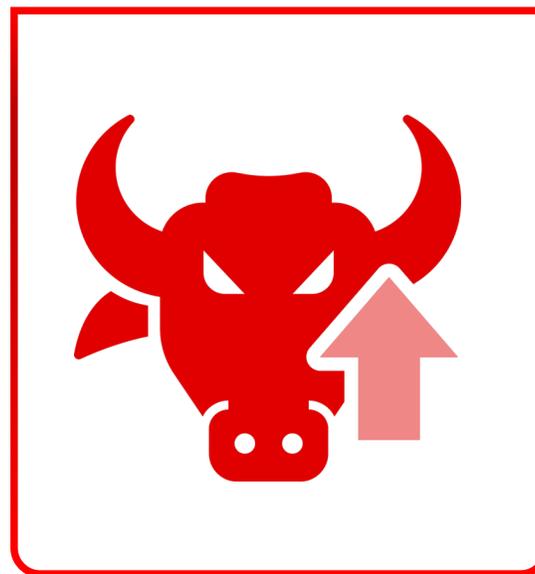


The Alveo SN1000
Smart NIC - hardware
performance and
software adaptability

An AI video analytics
platform and solutions
ecosystem built for the
most critical applications

Accelerated Algorithmic
Trading – Enabling a
broad spectrum of
traders to be competitive
in HFT

Xilinx App Store – A one-
stop-shop to evaluate
and purchase Xilinx
solutions





Thank You

