2021 中国智能网卡研讨会

CHINA SMARTNIC WORKSHOP

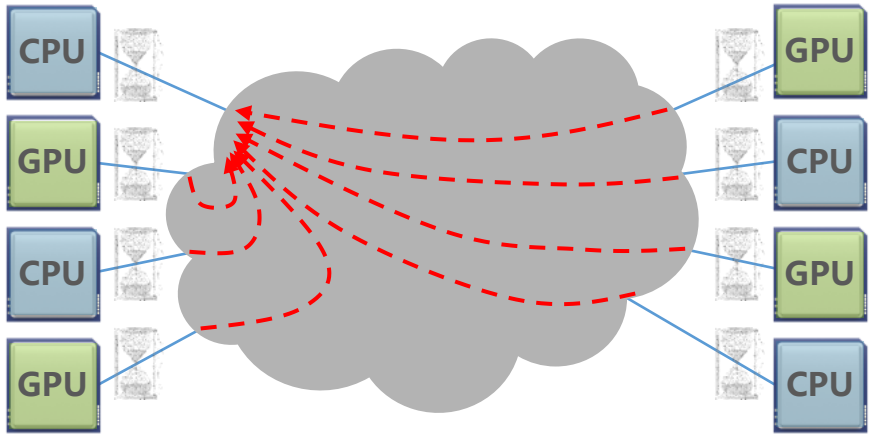
DPU驱动数据中心成为计算单元

宋庆春，NVIDIA

以数据为中心的计算架构成为了趋势

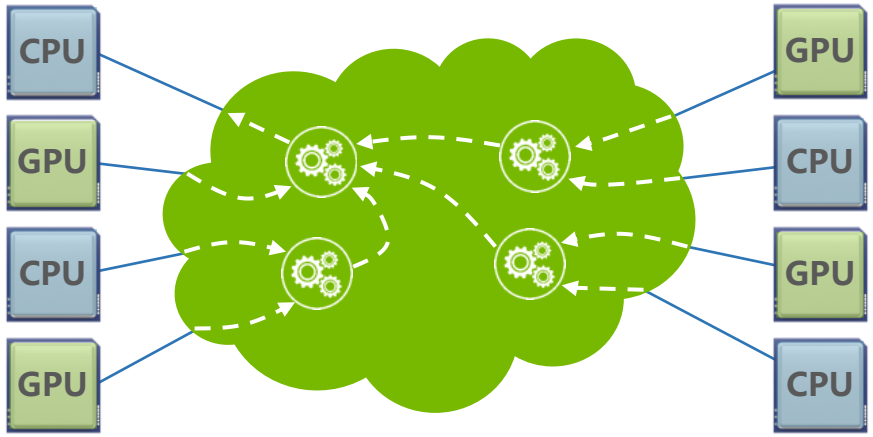
网络计算和DPU成为以数据为中心计算架构的核心

以 CPU 为中心 (Onload)



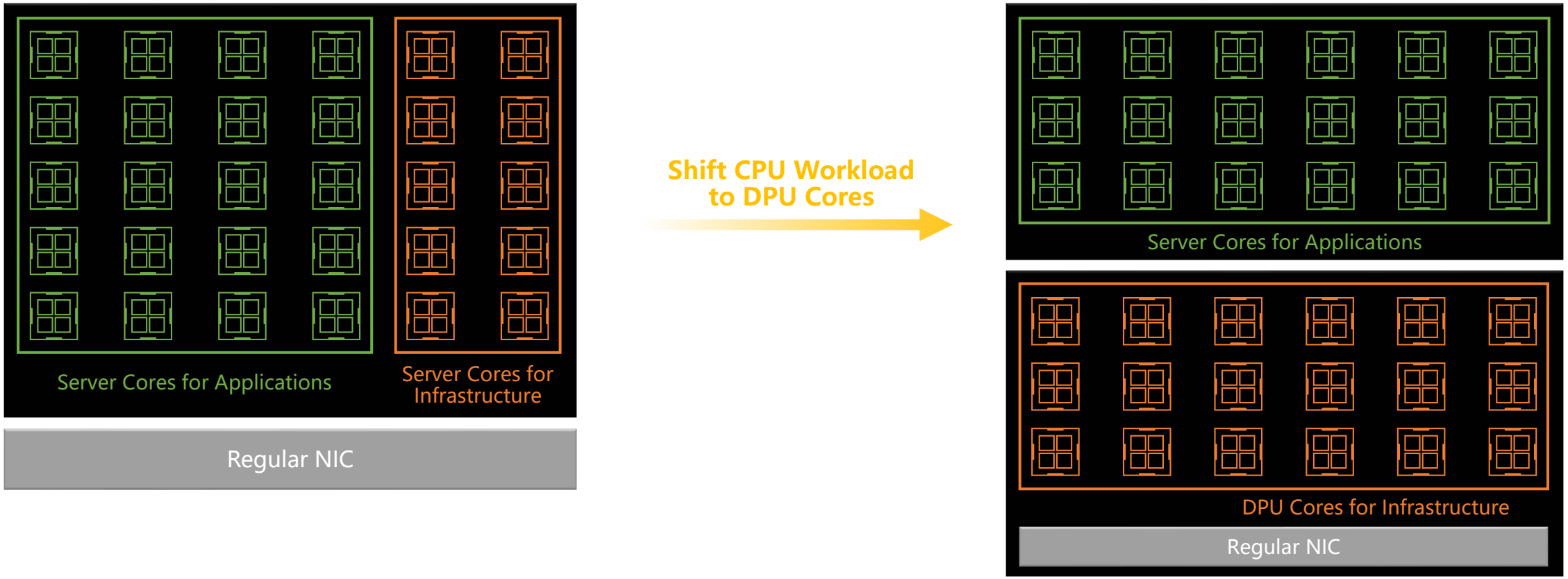
典型通信延时: 30-40 微妙

以数据为中心 (Offload)

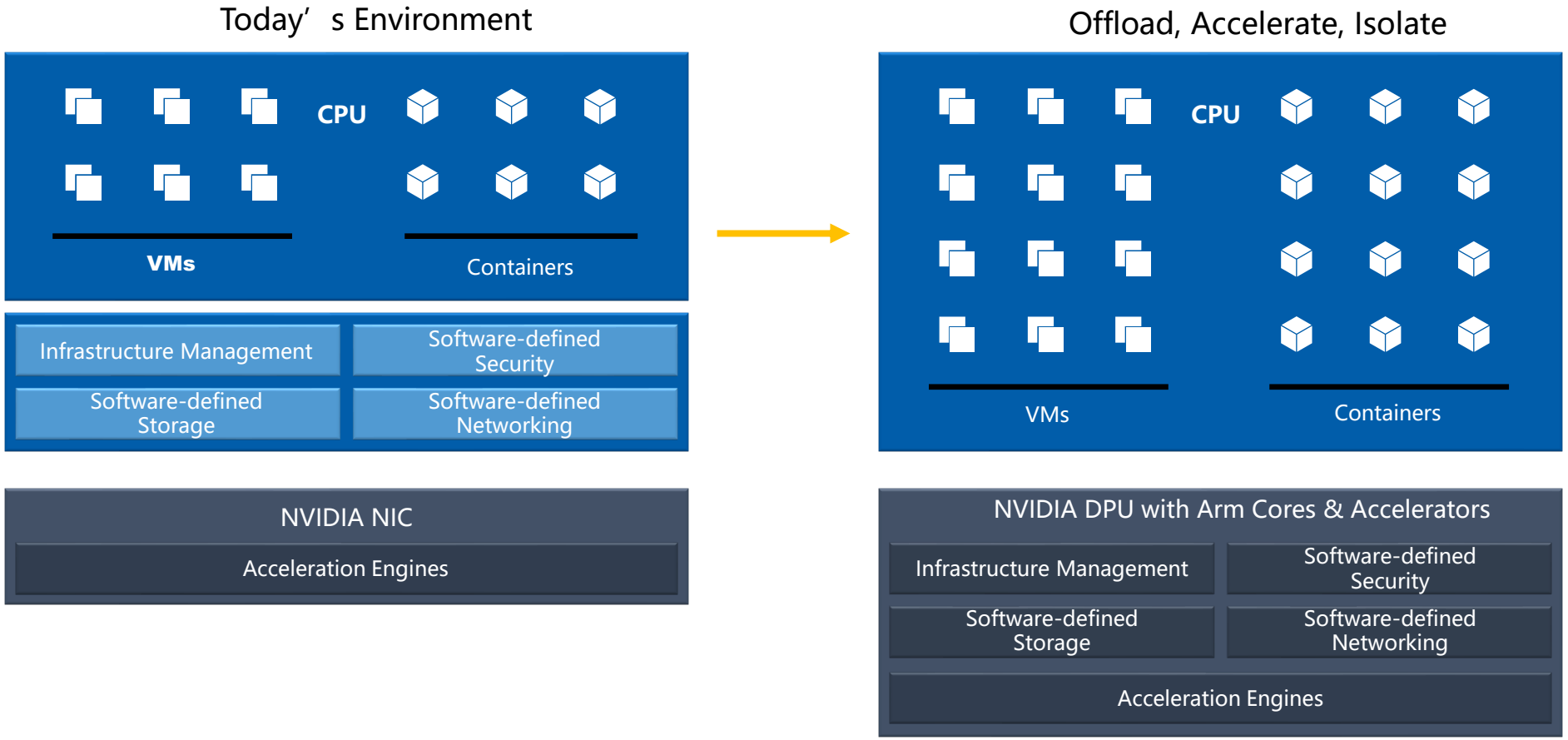


典型通信延时: 3-4 微妙

DPU 是单纯接替 CPU 的工作负载吗?



DPU 硬件卸载和加速才是关键



NVIDIA BlueField-2 DPU

Data Center Infrastructure on a Chip

集数据中心技术设施于芯片

- Up to 200Gb/s Ethernet and InfiniBand, PAM4/NRZ
- ConnectX-6 Dx inside
- 8 Arm A72 CPUs subsystem
- Integrated PCIe switch, 16x Gen4.0
- Single DDR4 channel
- 1GbE Out-of-Band management port
- Accelerated security, storage, networking

NVIDIA DOCA简介

轻松构建 BlueField 生态系统

- Software Development Framework for BlueField DPUs
- Software Compatibility for Generations of BlueField DPUs
- Offload, Accelerate, and Isolate Infrastructure Processing
- Support for Hyperscale, Enterprise, Supercomputing and Hyperconverged Infrastructure
- DOCA is for DPUs what CUDA is for GPUs

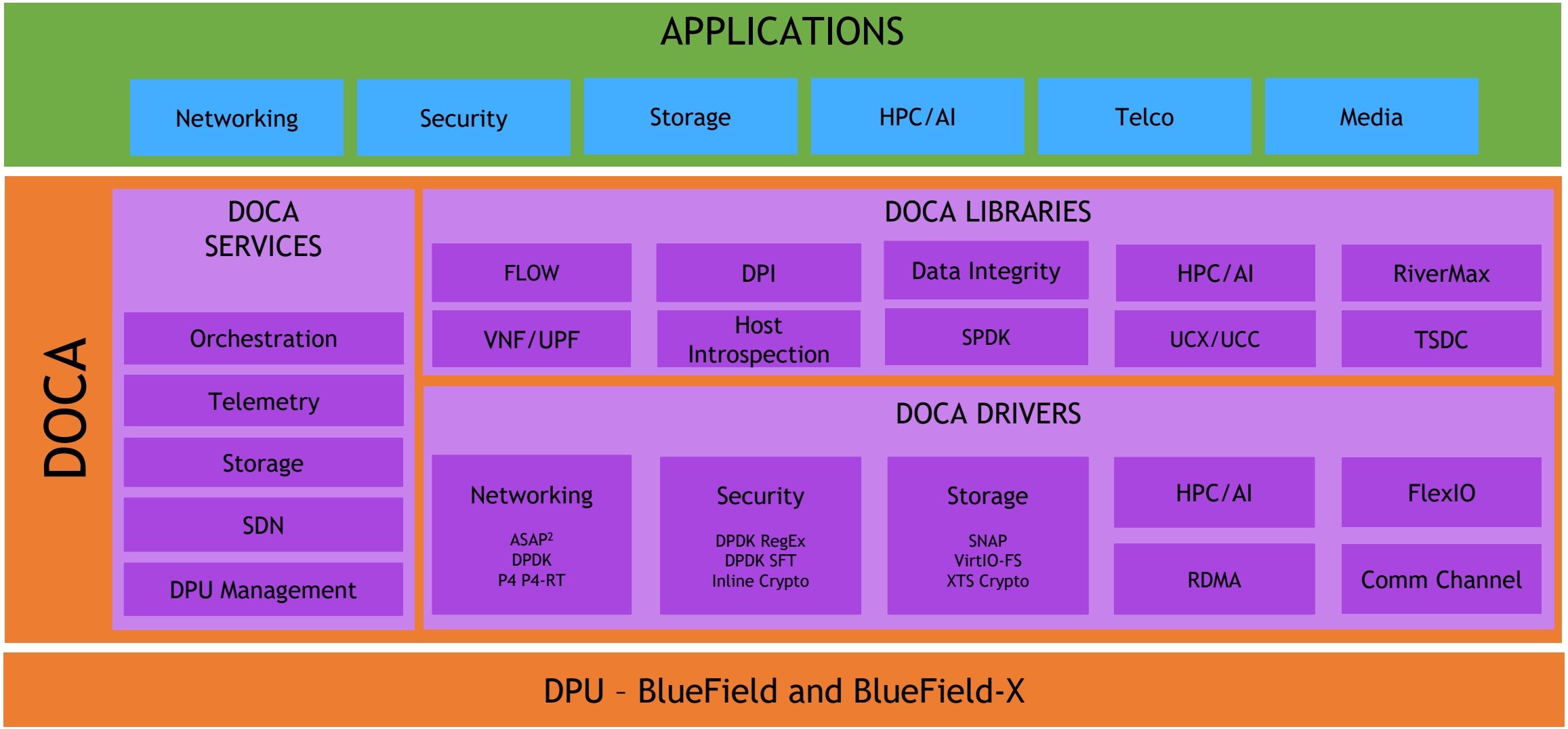
软件定义且硬件加速的基础设施

软件定义网络	vRouter vSwitch	Video Streaming	NAT/ Load Balancer	VMs and Containers	Telemetry/ PTP
软件定义存储	NVMe-oF	Encrypt	Elastic	Compress	Dedupe
软件定义安全	Distributed Next-Gen Firewall	IDS/IPS	Root of Trust	Micro Segmentation	DDOS Prevention

DOCA
Open and Programmable API Framework
Easy, Flexible Programming of Infrastructure / Acceleration and Security

NVIDIA BlueField DPU

DOCA 软件栈

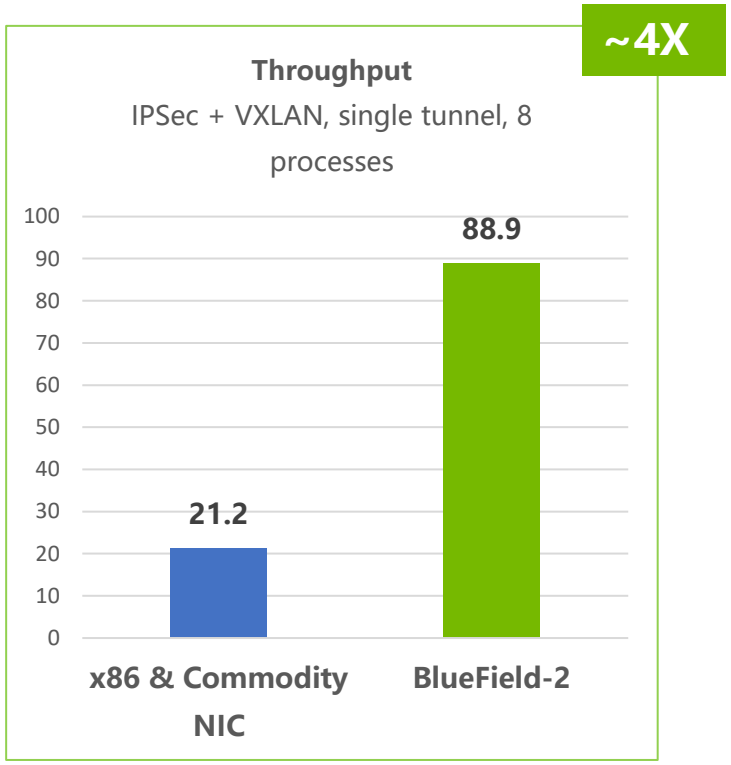

APPLICATIONS
Networking
Security
Storage
HPC/AI
Telco
Media
DOCA
DOCA SERVICES
Orchestration
Telemetry
Storage
SDN
DPU Management
DOCA LIBRARIES
FLOW
DPI
Data Integrity
HPC/AI
RiverMax
VNF/UPF
Host Introspection
SPDK
UCX/UCC
TSDC
DOCA DRIVERS
Networking
ASAP²
DPDK
P4 P4-RT
Security
DPDK RegEx
DPDK SFT
Inline Crypto
Storage
SNAP
VirtIO-FS
XTS Crypto
HPC/AI
FlexIO
RDMA
Comm Channel
DPU – BlueField and BlueField-X

DOCA开发区

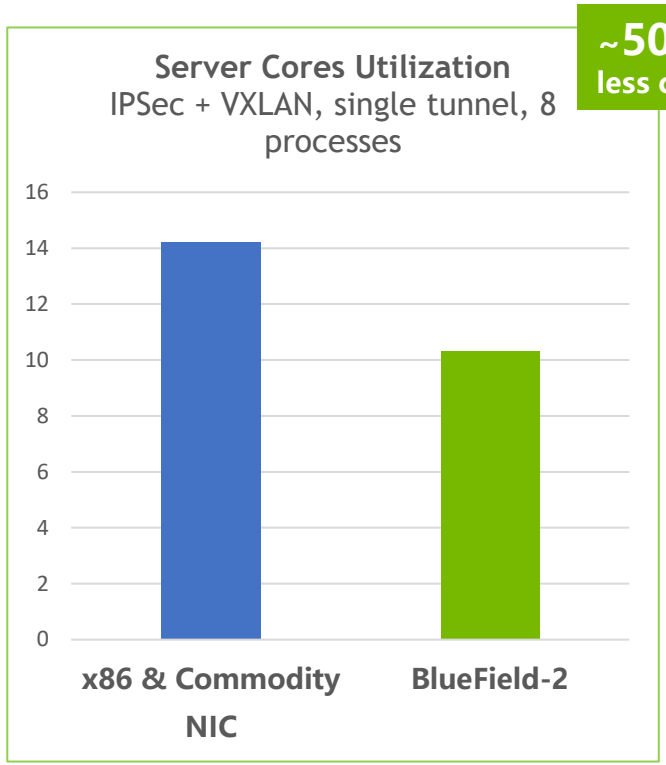
- **Installation guide**
  - SDK manager
  - Manual Installation
- **Programmer's guide**
  - DPI
  - Netflow
  - DPDK – Upstream
- **DOCA and DPDK APIs**
- **Reference application**
  - Application Recognition
  - East-West Overlay Encryption
  - L4 OVS Firewall
  - URL Filter
  - Netflow Exporter
  - Simple forward VNF
- **Tools**
  - DPI Compiler
  - Regex Compiler
  - RXP Bench
- **DOCA Run-Time guides**

BlueField-2 100G DPU 的 IPsec TCP 性能

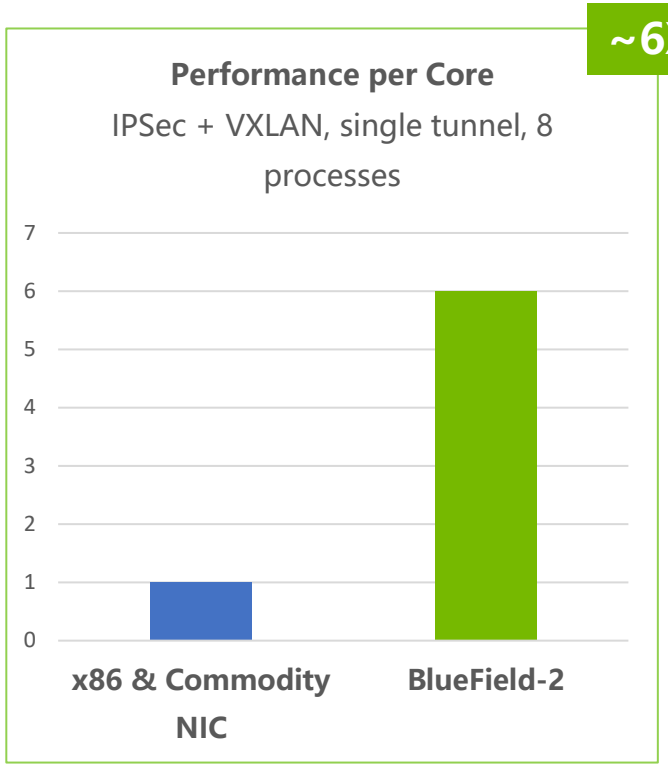
更高的性能| 更优的投资



- BlueField-2 P-series 100GbE Single port
- Intel(R) Xeon(R) CPU E5-2687W v4 @ 3.00GHz
- Host OS : RHEL 8.3 (Ootpa)



- Each SW core does : 21.2 / 14.2 = ~1.5G
- To reach ~90G in SW ~60 cores
- To reach ~90G w/ DPU ~10 cores

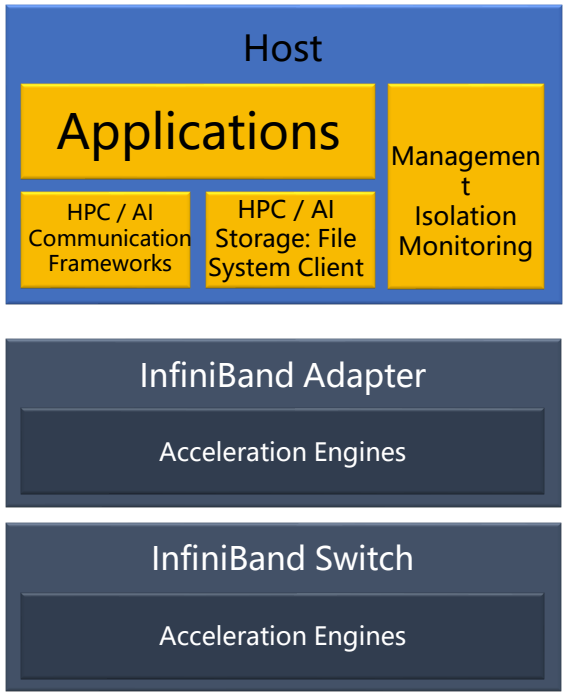


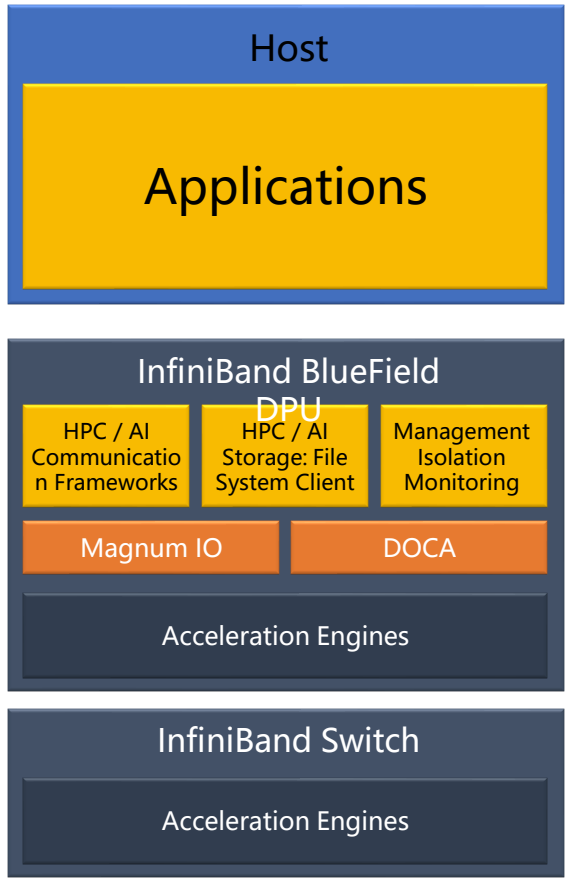
- SW cores do : 21.2 / 14.2 = ~1.5G
- With DPU SW cores do : 88.9 / 10.3 = ~9G
- Core performance increased by magnitude 6X

DPU 让云原生计算成为可能

在零信任的前提下实现多租户的隔离

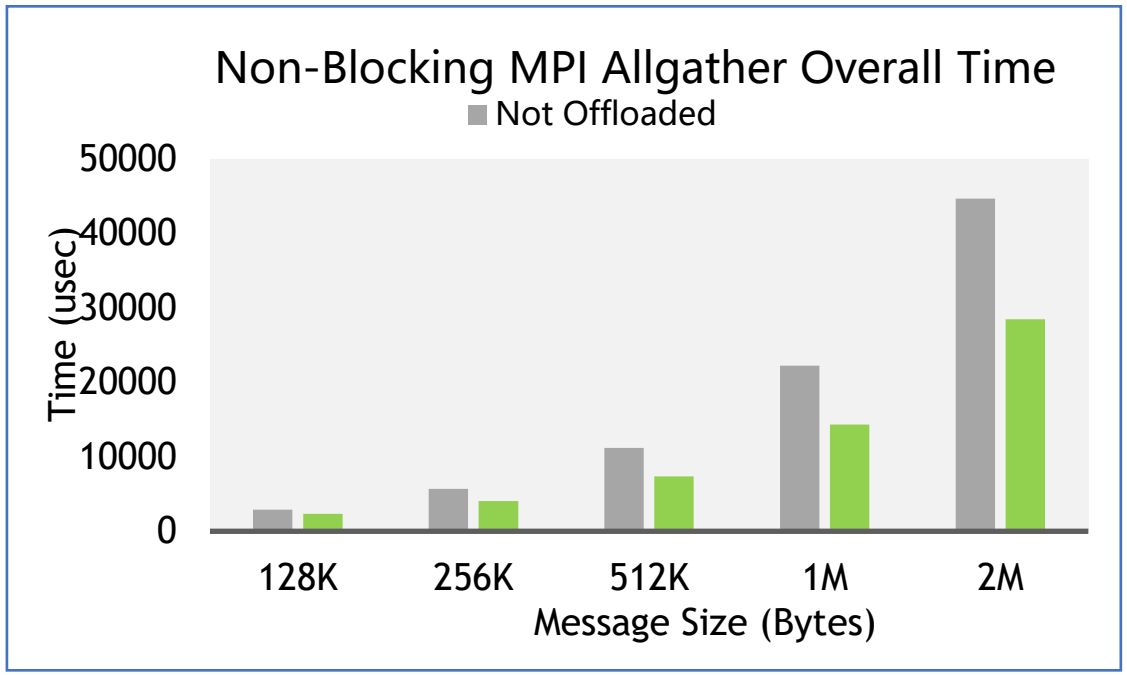
- Collective offload with UCC accelerator
- Smart MPI progression
- User-defined algorithms
- 1.4X higher application performance

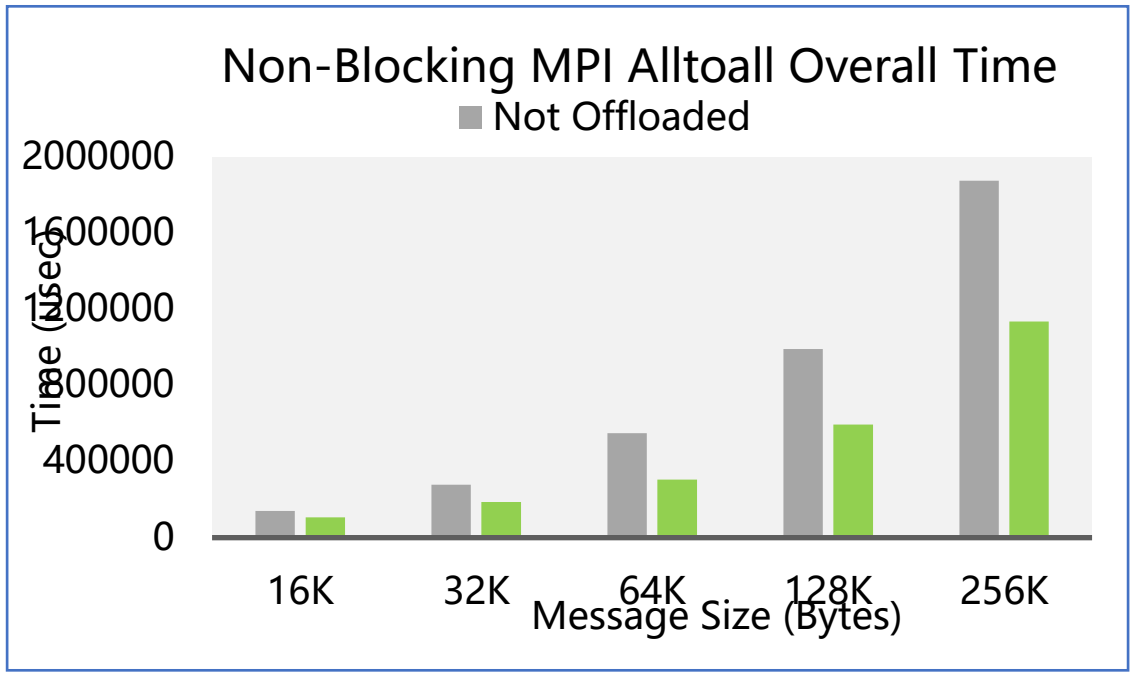




BlueField DPU 大幅提升无阻塞MPI 性能

MPI iAlltoall 性能提升44% | MPI iAllgather 性能提升36%

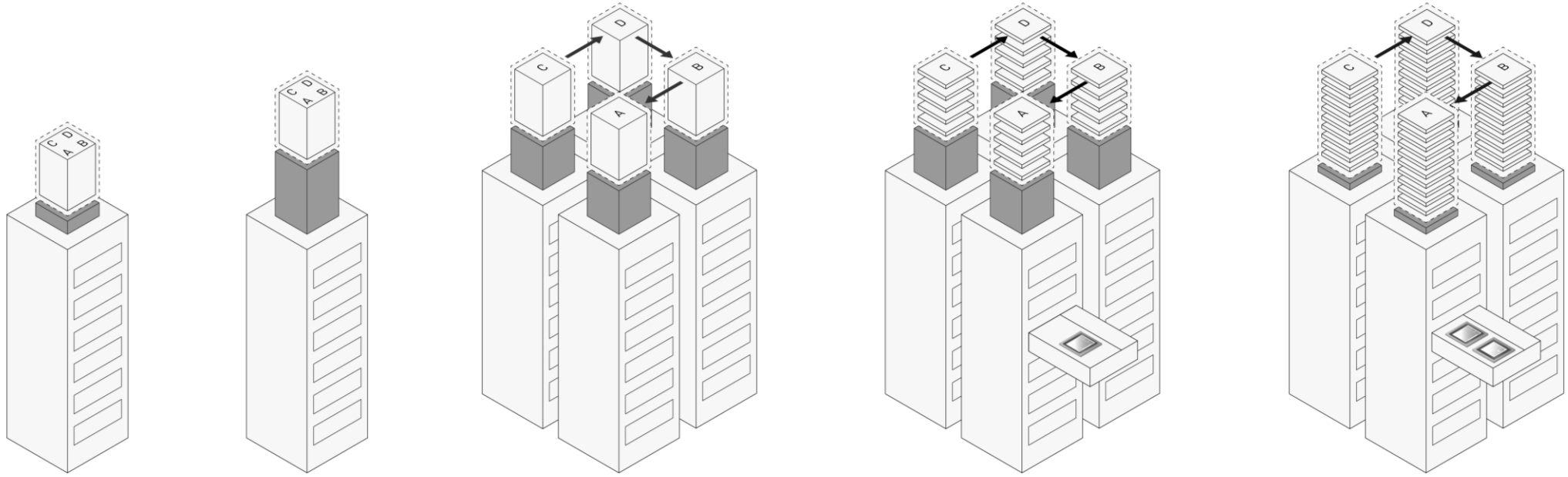




Source: Courtesy of Ohio State University MVAPICH Team and X-ScaleSolutions

MVAPICH X-ScaleSolutions

总结: 数据中心成为了新一代计算单元



THANKS