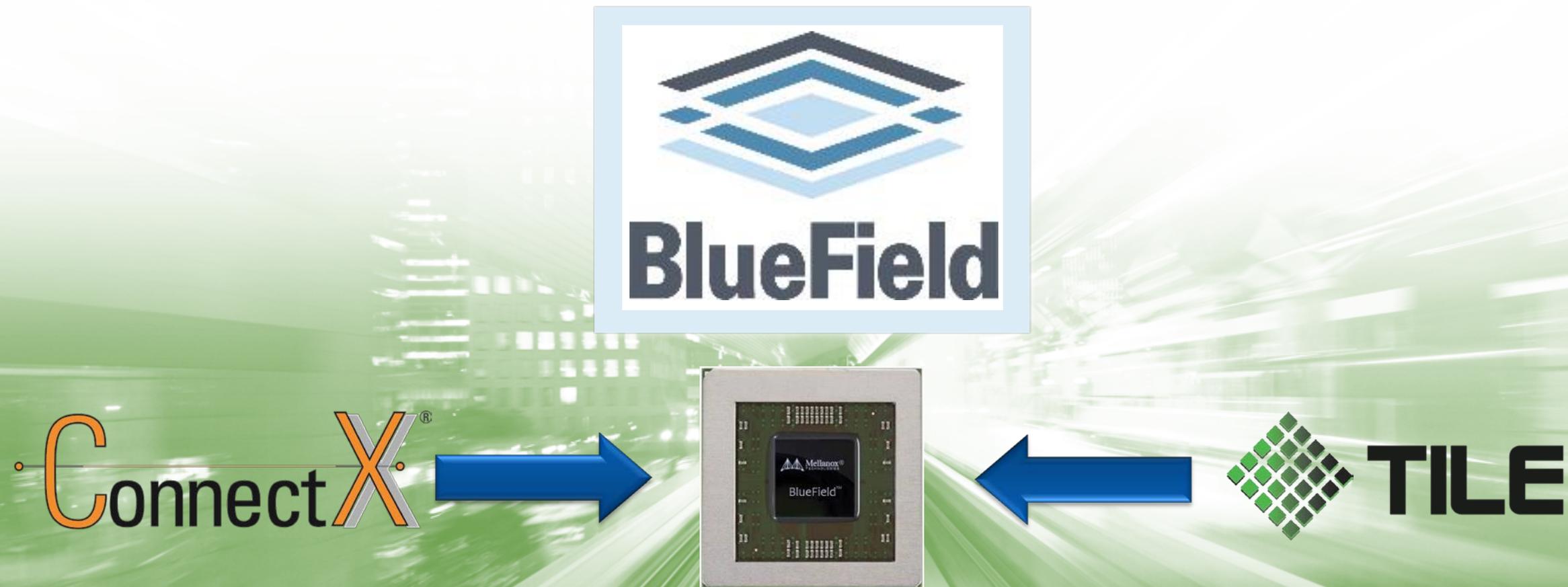




推出 BlueField™ - 智能网络的创新

2016年11月

 **Mellanox**
TECHNOLOGIES
Connect. Accelerate. Outperform.™



**行业领先的 ConnectX 智能卸载
与 Tile 多核 ARM 体系架构相结合**

存储

NVMe 闪存存储阵列
横向扩展存储 (NVMe over Fabric)
灵活的 HBA



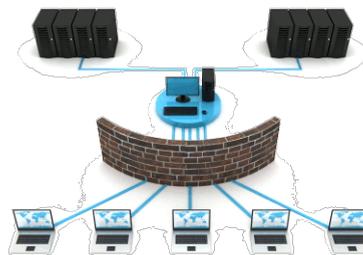
NFV

加速 VNF
开放虚拟交换机 (OVS)、SDN
叠加网络卸载



安全性

入侵检测/预防 (IDS/IPS)
防 DDoS、防火墙
加密协议 (IPsec、SSL/TLS、DTLS...)



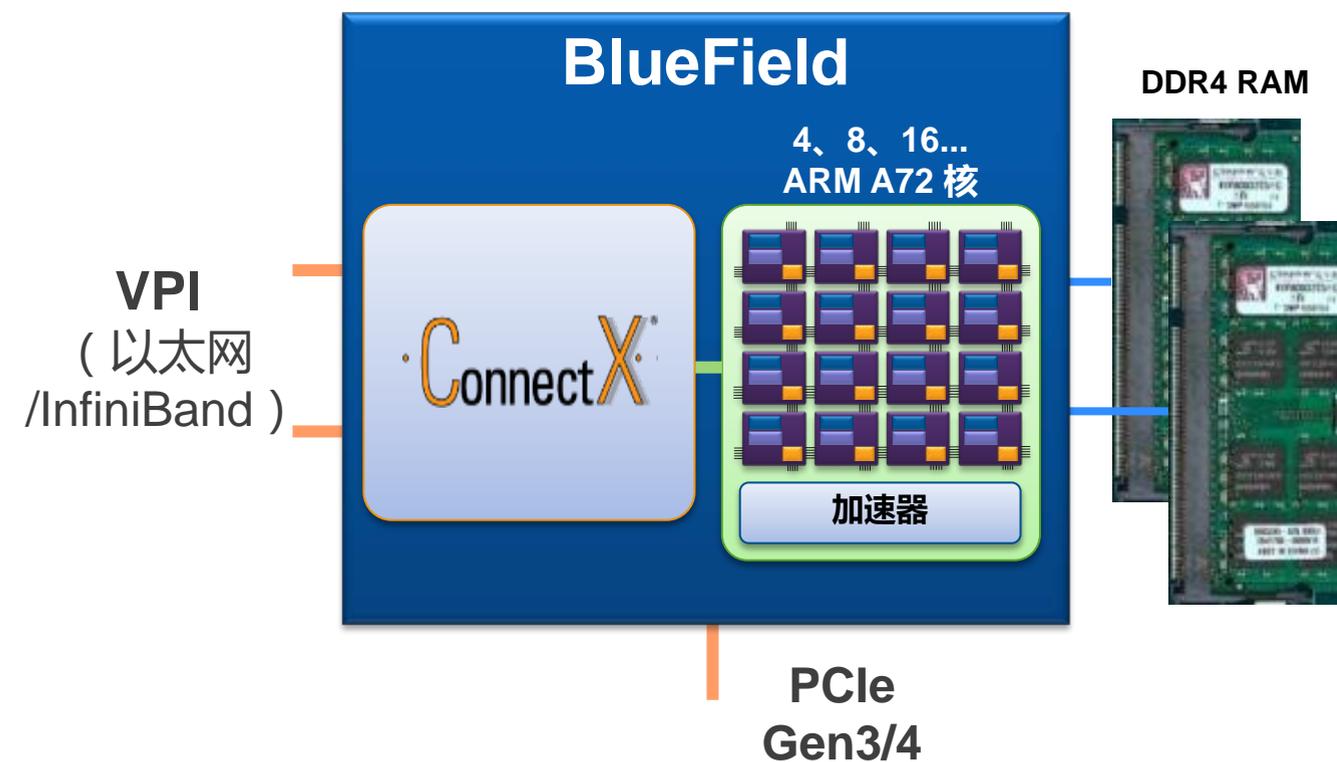
■ ConnectX® 与基于 Tile 的多核 ARM 的集成

■ 最先进功能

- 10/25/40/50/100G 以太网和 InfiniBand
- PCIe Gen3/Gen4 , 与网络带宽相匹配
- 64 位 ARM 多核阵列 - 完全可编程
- 硬件加速卸载
 - 加密、数据包处理、OVS 卸载
 - RDMA、RoCE、NVMeF、RAID

■ 全系列软件兼容的产品/芯片

- 包含不同 ARM 核数量, I/O 端口数量和 I/O 接口速率产品
- 适用不同的性价比场景



■ 具有 16 个 ARM[®] A72 CPU 的 Tile 体系架构

- SkyMesh[™] 完全一致性低延迟互连
- 8MB L2 高速缓存 (每个 Tile 1MB, 共 8 个 Tile)
- 每个 A72 核具有 48KB I-Cache、32KB D-Cache

■ 2 个共享 L3 高速缓存列

- 总计 12MB L3 最后一级高速缓存

■ 集成 ConnectX[®]-5 子系统

- 支持双接口 100Gb/s 以太网/IB, 与 ConnectX-5 兼容

■ PCIe Gen 4

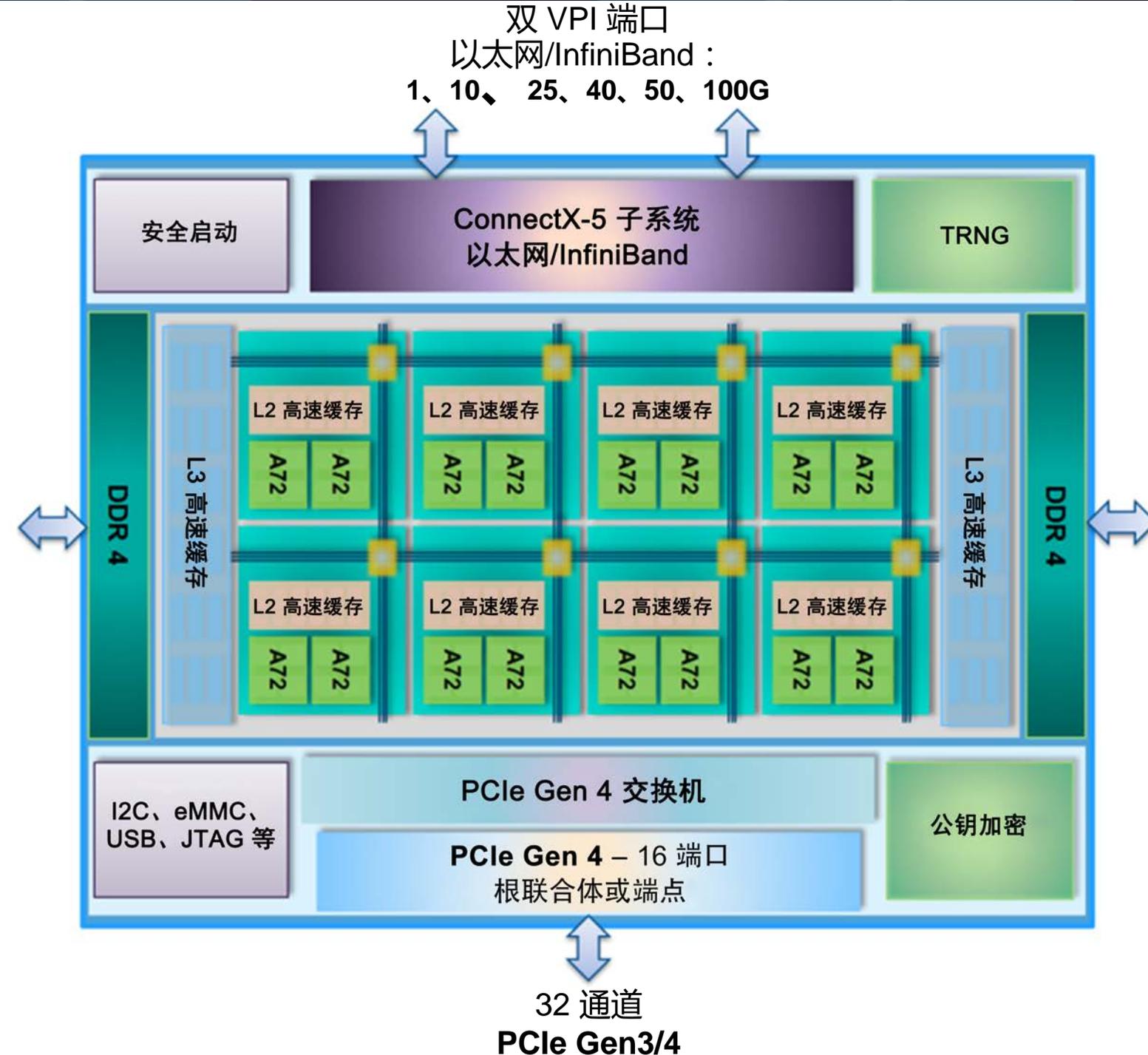
- 32X PCI Gen4 通道 (最高 200Gb/s)
- 根联合体或端点模式
- 2x16、4x8、8x4 或 16x2 配置

■ 加密

- 公钥加速、AES、SHA-1/2
- 真随机数生成器 (True RNG)

■ 内存控制器

- 2 通道 DDR4 内存控制器, 带 ECC



网络和安全市场机遇

■ 主要的技术转型

- 将网络服务分离到 NFV 机架式服务器
 - “软件定义网络”的基本促成因素 - 根据需要添加容量
- 多家供应商提供虚拟路由器 -证实了这一趋势

■ 但性能密度滞后

- 服务器的设计并未针对低延迟、高吞吐量的数据层面
- 基本数据面报文处理消耗了太多 x86 计算资源
- 加速数据层面
 - 为 VNF 腾出更多 x86 CPU 计算资源
 - 更低的延迟和可预测的吞吐量
 - 更出色的小数据包性能

专用
网络设备



数据中心级 NFV

■ 集成 I/O

- 具有与 I/O 相匹配带宽的 PCI Express Gen3/4
- 支持 10、25、40、50、100 Gb/s 的以太网接口速率

■ 具备数据路径的网络感知能力

- 在超高流量负载（如 DDoS）下具有可预测的性能
- 在虚拟化主机环境下支持多 VNF
 - 每端口具有 256 个虚拟功能的 SR-IOV
- 第 2 层/第 3 层网络卸载
- 安全性卸载

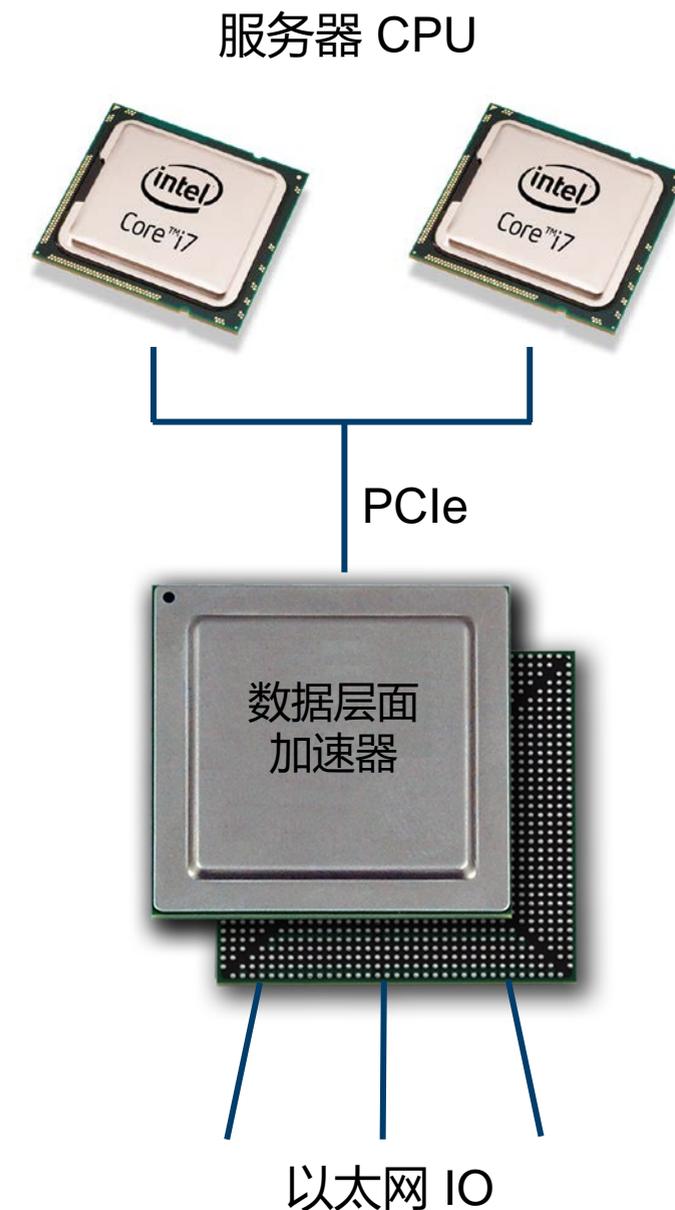
■ 可编程能力

- 支持不断发展的功能和网络协议

■ 安全性

- 安全执行环境，包括安全启动
- 独立的软件域，与服务器分离

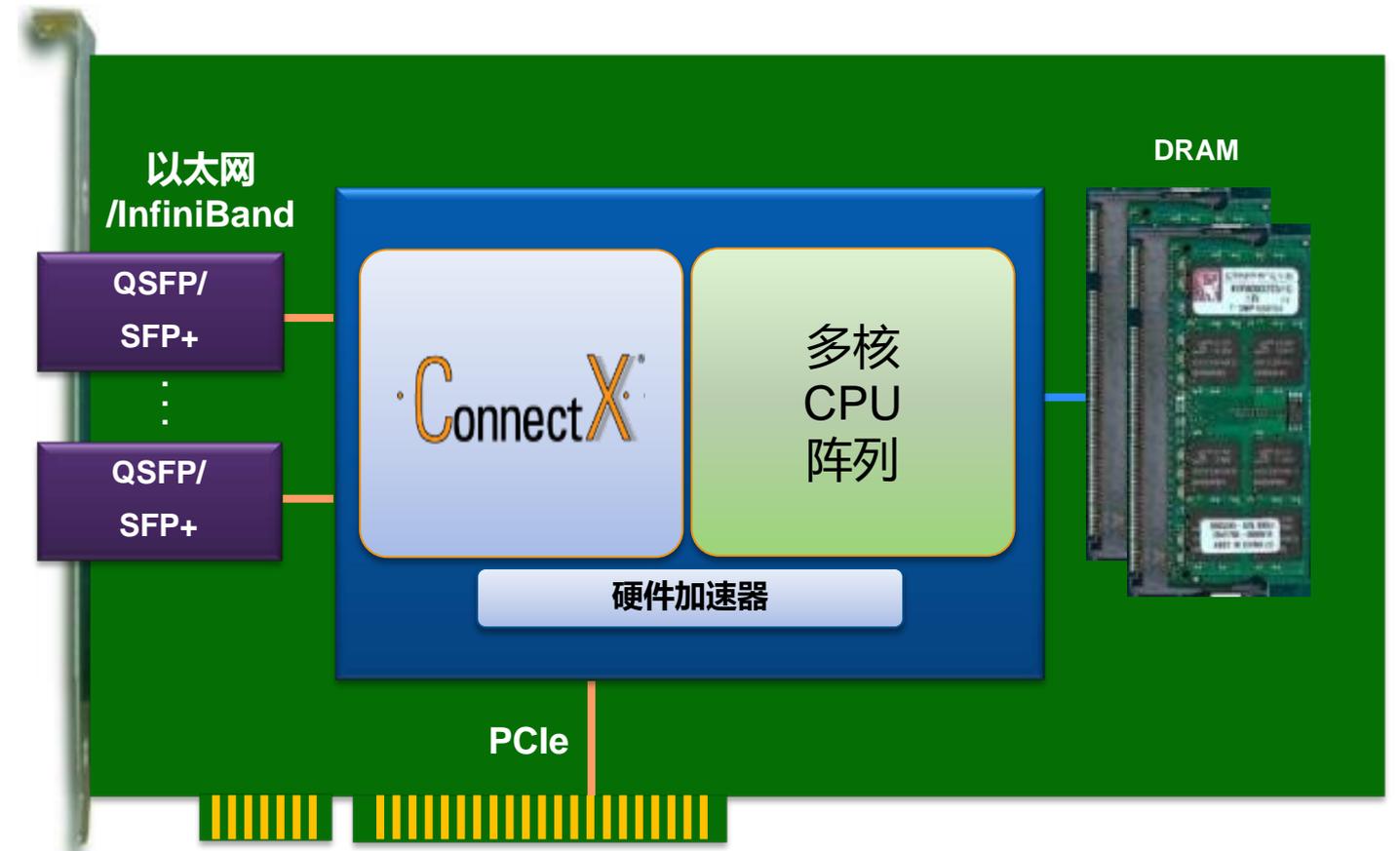
■ 可扩展性



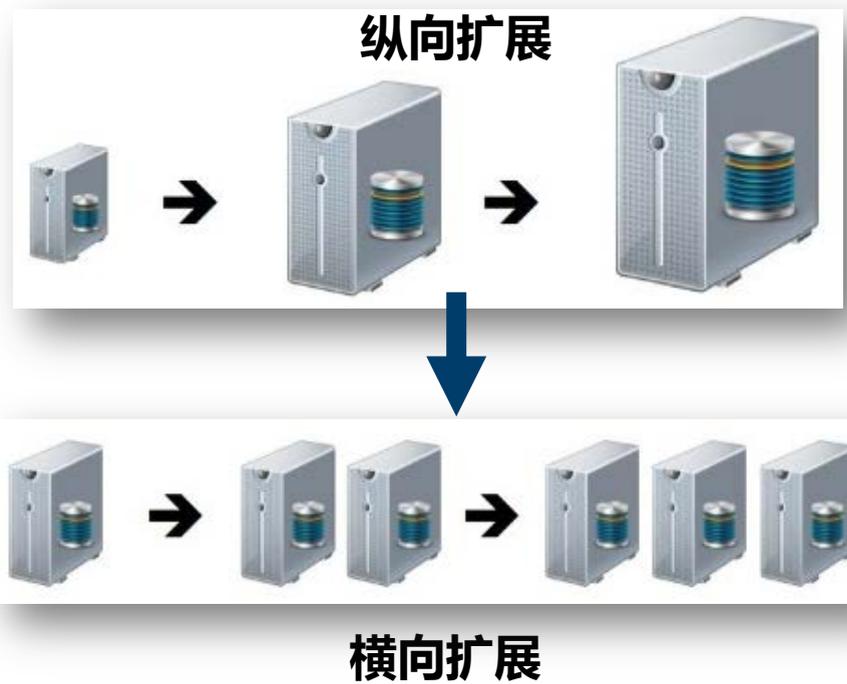
- ConnectX® 支持以太网协议加速
 - 以太网卸载 (RSS、TSS、LRO、LSOv2)
 - 基于嵌入式流表的 “eSwitch”
 - OVS 加速
 - RDMA、RoCE
 - 报文头重写
 - 针对 NVGRE、VXLAN、GENEVE、MPLS 的无状态卸载
- 可编程网卡提供了更丰富的功能：
 - L2-L7 协议处理
 - L3/4 加密/解密 (IPsec、SSL、SRTP、DTLS 等)
 - 安全处理 (防火墙、IDS/IPS、防 DDoS 等)
 - 应用识别、DPI
 - 数据压缩/解压缩



- 可编程软件定义网卡
 - 加速和卸载商用服务器
 - 在针对任务高度优化的处理器上执行网络工作负载
- ConnectX® 前端
 - 驱动程序支持广泛的服务器硬件和操作系统
 - 灵活的以太网/InfiniBand I/O 能力
 - 1/10/25/40/50/56/100 Gb/s
- 多核 CPU 阵列
 - CPU 卸载数据层面功能
 - 用于网络协议卸载的加速器引擎
 - 标准 Linux/BSD/KVM 编程环境
- 本地 DRAM 内存子系统
 - 支持大型流表和深度数据包缓冲区



存储市场机遇



■ 主要技术转型

- 高级服务与存储容量分离
 - “软件定义存储”的基本促成因素 - 根据需要添加计算或容量
- SSD/NVMe 取代机械硬盘
 - NAND 目前占据主导地位，新技术即将来临 - 3D-XPoint

■ 全都需要高性能网络

- 集群存储要求低延迟和高带宽
- 更快的存储需要更快的网络（一个 NVMe 驱动器大约需要 20Gbps 带宽）

NVMe over Fabric

SanDisk、Micron、HGST、Seagate 和多家初创公司

■ 集成 I/O

- PCI Express , Gen3/4
- 以太网接口 : 10、25、40、50、100G
- InfiniBand 接口 : 40、56、100G

■ 数据路径具备存储感知的能力

- NVMeF 协议转换 (PCIe <> NVMeF)
- RDMA、RoCE (NVMeF 所需)
- RAID/纠删码卸载
- 数据完整性卸载 (即 T10-DIF)

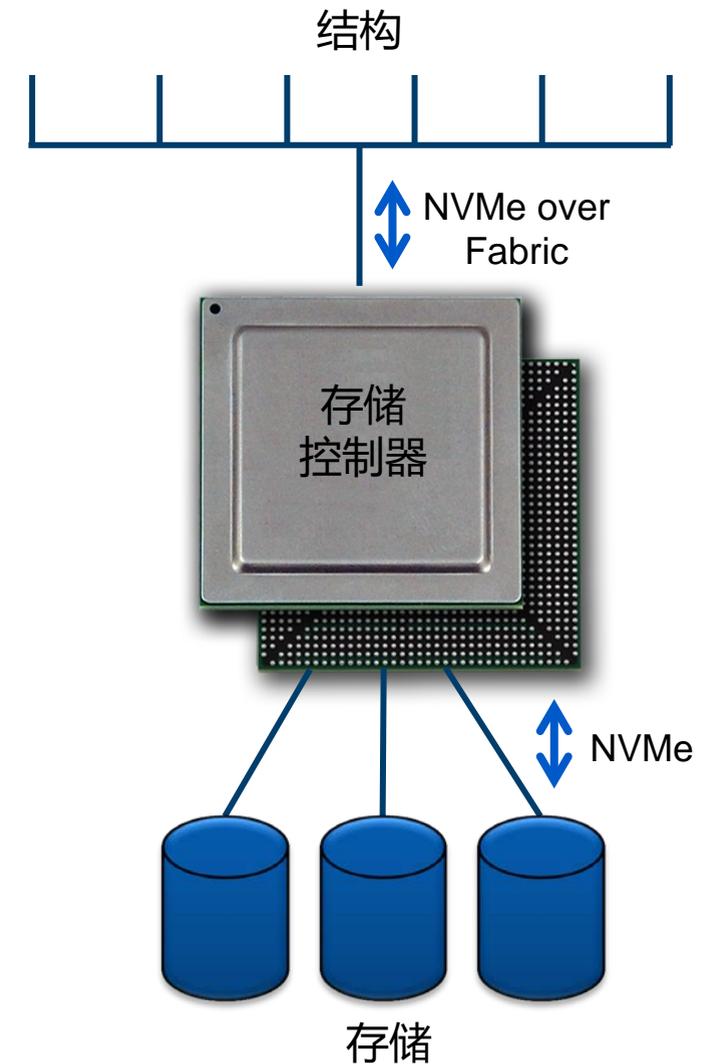
■ 可编程能力

- 为了支持存储内 (in-storage) 功能和计算

■ 可靠性和高可用性

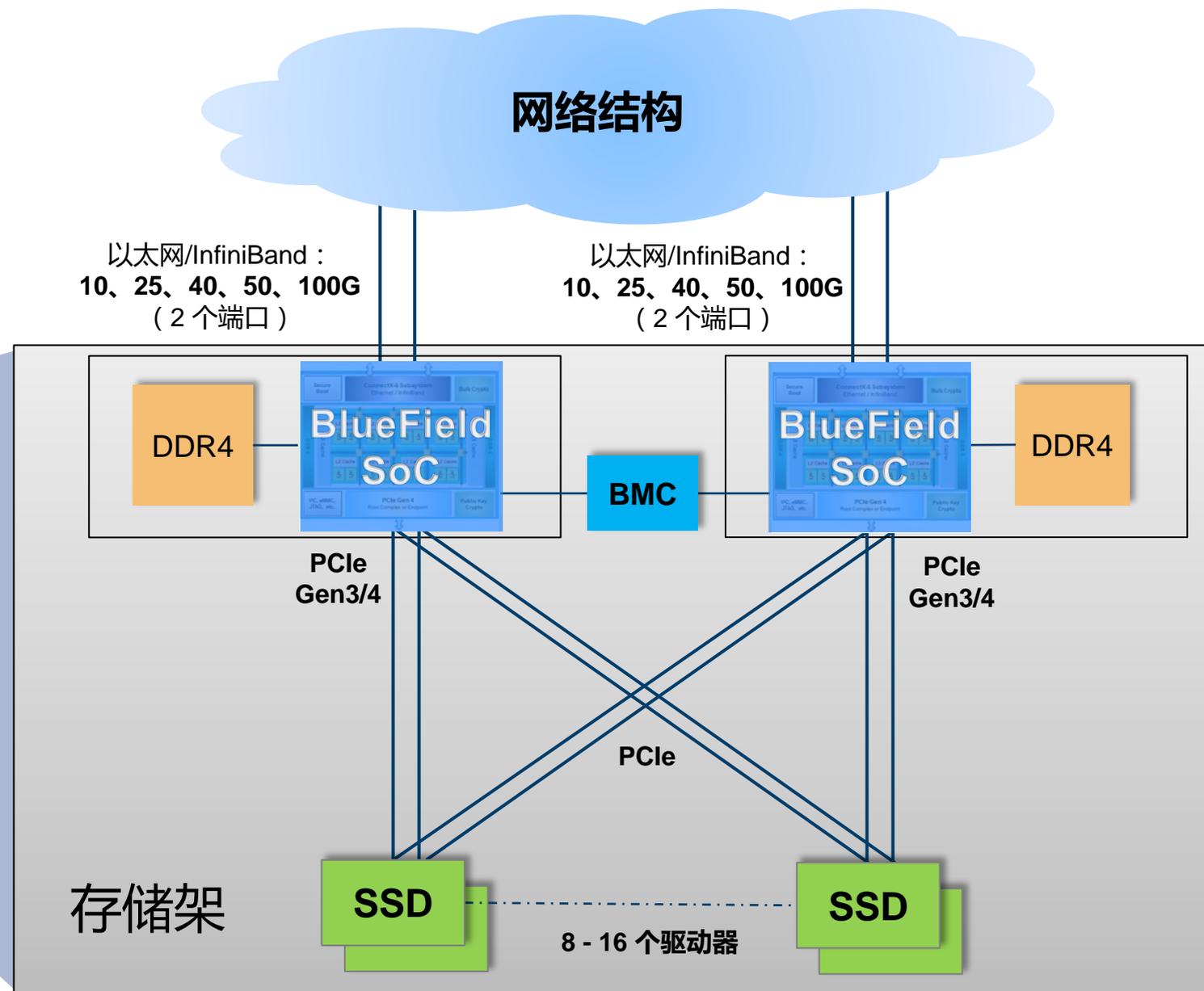
- 内存错误保护
- 软错误恢复能力

■ 低功耗



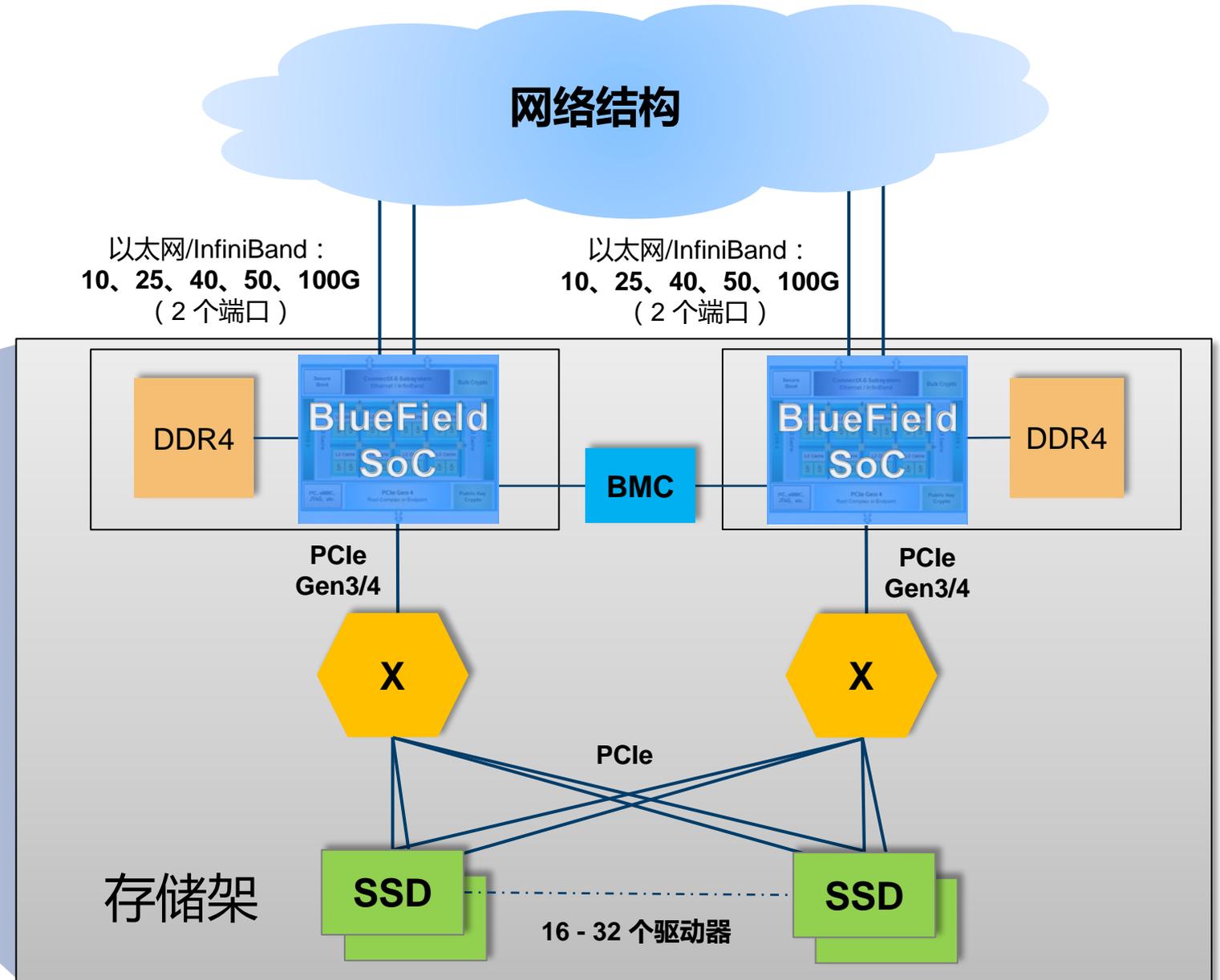
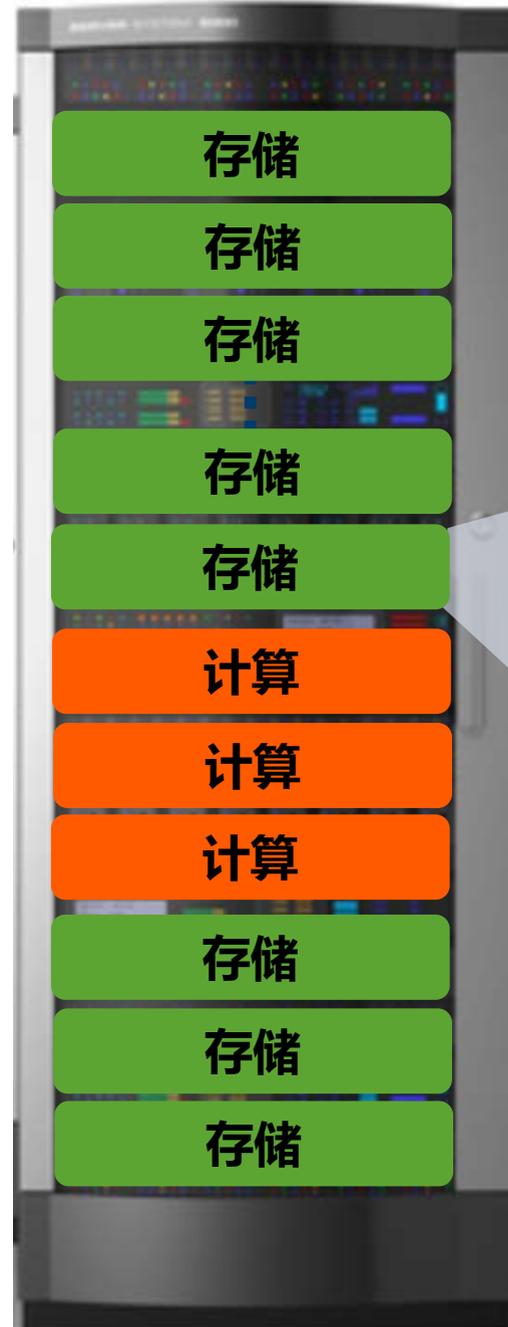
使用 BlueField 的横向扩展 NVMe 存储系统

机架视图



添加 PCIe 交换机以获得更大的磁盘驱动能力

机架视图



- ConnectX® 支持以太网和存储协议加速
 - RDMA、RoCE
 - NVMe over Fabric 协议处理
 - 用于 RAID 的纠删码卸载
 - 签名移交卸载
 - 用于数据完整性的 T10-DIF
- 可编程内核支持：
 - RAID 卸载（校正子生成和重建）
 - 增强的闪存转换层 (FTL)
 - NVMe→SAS 转换
 - 磁盘扇区加密/解密
 - 数据压缩/解压缩
 - 重复数据消除





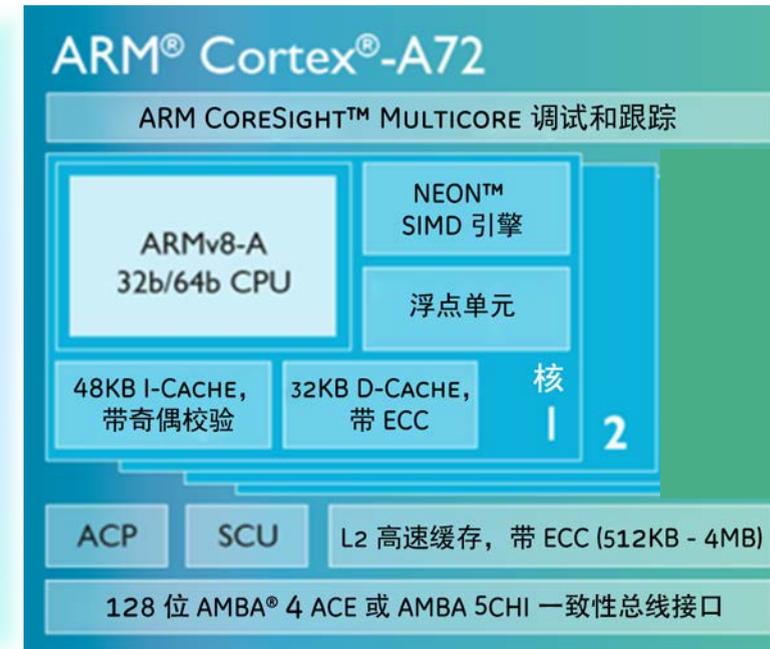
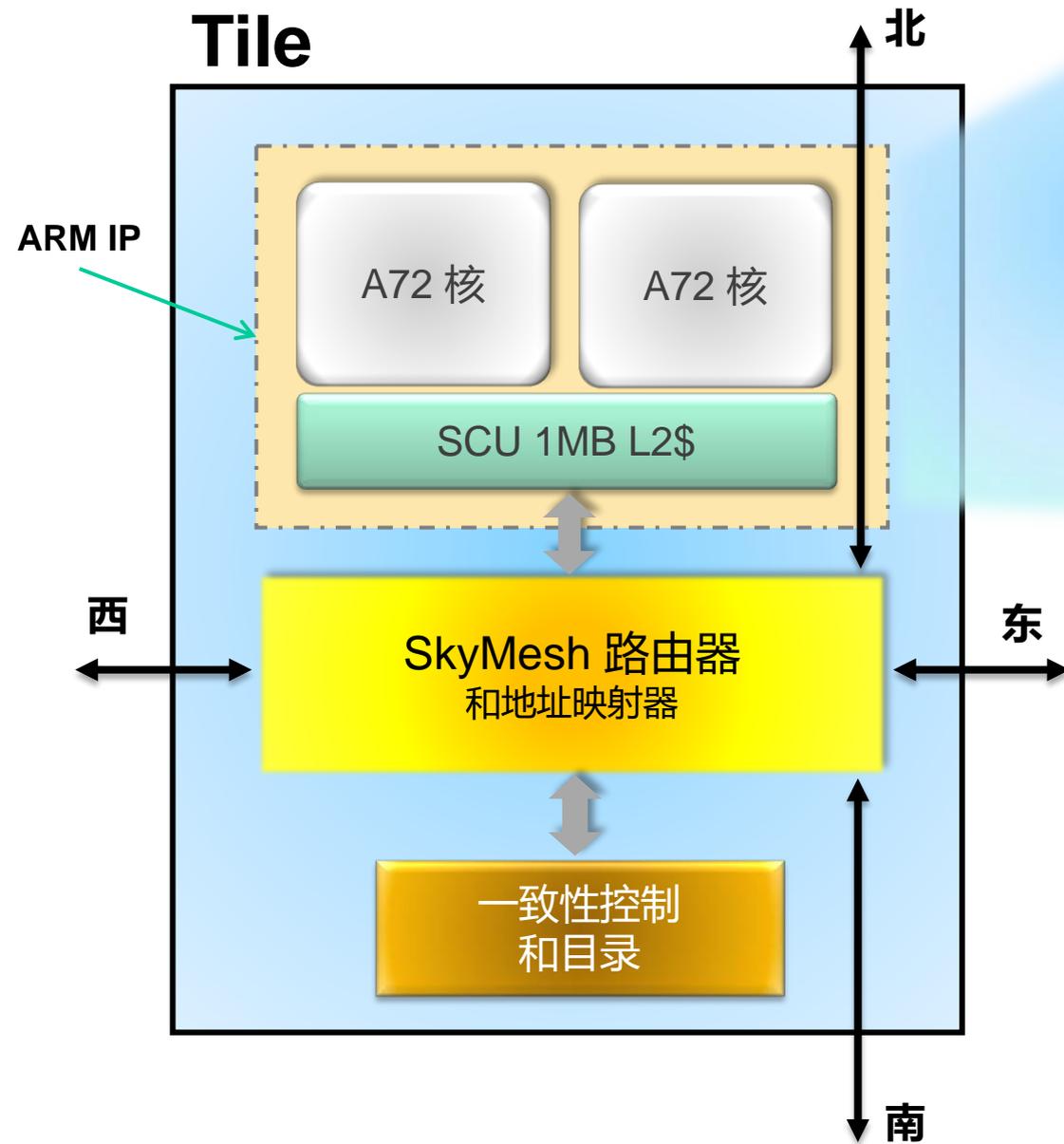
针对嵌入式网络、计算和存储的
最高性能 SoC

功能特性	优势
ARMv8 多核处理器阵列	<ul style="list-style-type: none">• 易于编程和定制功能集• 庞大的软件生态系统• 软件可以与 x86 相互移植
ConnectX [®] 网络子系统	<ul style="list-style-type: none">• 行业领先的网络加速和卸载• 最低的端到端延迟• 跨硬件和操作系统的全面驱动程序
可扩展的 SkyMesh [™] 片上网络	<ul style="list-style-type: none">• 硬件一致性可简化编程• 性能随内核数线性扩展
SoC 集成 I/O 和 DDR 控制器	<ul style="list-style-type: none">• 节省 PCB 面积和功耗• 最低延迟和最小尺寸数据包吞吐量
硬件加密加速器	<ul style="list-style-type: none">• 不影响安全性的高性能• 受保护的执行环境



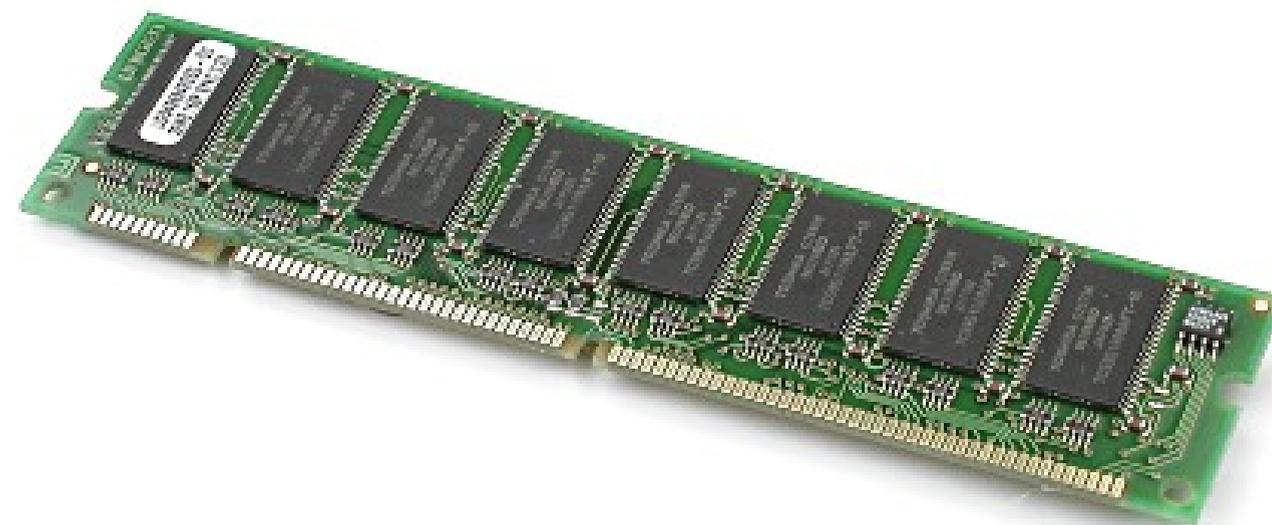
谢谢

备用



- Tile 集成双 ARM A72 核，每个核具有：
 - 48KB L1 iCache、32KB dCache
 - 1MB L2 高速缓存 - 共享
 - NEON SIMD 引擎
 - 浮点单元
 - 核内加密指令
 - AES、SHA-1、SHA-224、SHA-256、Galois Field
- 一致性控制器提供：
 - 分布式目录（监听过滤器）
 - 保证内存访问的序列化和一致性

- 2 通道，72 位 DDR3/4，带 ECC 错误保护
- DDR4 DIMM 支持：
 - 每通道最高 256GB，总计 512GB
 - 每通道最多 2 个 DIMM（相同）
 - 最多 16 个 DDR Rank：
 - 每通道最多 4 个 Package Rank
 - 每 Package Rank 最多 4 个逻辑 Rank (3DS)
 - 性能：1600 - 2667MT/s
 - 支持 DDR4 UDIMM、RDIMM 和 LRDIMM（带或不带 ECC）
 - NVDIMM-N 支持
- 支持的请求类型：
 - 带/不带 ECC 的 64B 读取
 - 带/不带 ECC 的 64B 写入
- 性能：
 - 多 Rank 下的目标是 90% 总线利用率



■ 加密应用

- 网络加密和身份验证协议
 - IPsec
 - SSL/TLS/DTLS

■ 加密引擎

- 对称加密功能
 - AES-GCM/CTR/CBC/XTS , 采用 128/192/256 位密钥
 - SHA-1、SHA-2-224/256/384/512
- 公钥加速器
 - RSA、Diffie-Hellman、DSA、椭圆曲线加密 (ECC)、ECDSA、ECDH - 最高 4096 位模数
- 真随机数生成

■ 性能

- 对称加速 - 100Gb/s 加密卸载 (AES-GCM、SHA-1、SHA-2)
- 公钥引擎 - 20K 次运算/秒 RSA 2K
- 真随机数生成器 - 500Mbps



- GPIO 引脚 (48 个引脚)
 - 输入、输出、开漏
 - 可配置中断 I/O
 - 可通过外部 CPLD 扩展 I/O
- I²C/SMBus (9 个端口)
 - 每个端口可配置为主和/或从
- MDC/MDIO (6 个端口)
 - 由 I²C 叠加
- eMMC (2 个端口)
 - 一个用于启动, 另一个用于 NV 存储
- SPI
 - 用于外部闪存的四路 SPI
 - 用于固件和配置启动
- USB 2.0 端口 (ULPI)
 - 端点模式
 - 可配置为启动端口
- DMTF 管理接口
 - 支持 MCTP over SMBus 的 SMBus 接口
 - RBT (RMII) 管理端口
 - 支持 MCTP over PCIe VDM 的 PCIe 接口
- 用于内部 eSwitch 的 SDN 管理接口
- UART (2 个端口)
 - 可配置为启动端口
- 1PPS 精度时间戳 I/O
 - 1PPS 输入和 1PPS 再驱动输出
- JTAG 边界扫描端口
 - 支持 IEEE 1149.1 和 1149.6
- CoreSight™ 调试端口
- LED 驱动接口
 - 4 个用于网络链路指示的输出 (与 GPIO 多路复用)
- 风扇控制/过热关机接口



支持社区工作

客户相信什么

我们加速的方面

Mellanox 交付项目

运行时软件

开发工具

社区
所有其他...

社区
所有其他...

关键软件
关键市场/应用

增值
性能和可见性

增值
框架

基础
* 必须拥有

- Yocto ARM 发行版
- BlueField BSP
- Mellanox OFED 驱动程序/API 套件

基础
必须拥有

