napa:tech;

SMARTNIC+IPU SOLUTIONS

SOLUTION DESCRIPTION

Storage Offload

FOR INTEL-BASED INFRASTRUCTURE
PROCESSING UNITS (IPUs)

napa:tech;

SMARTNIC+IPU SOLUTIONS

SOLUTIONS
Telecom
Cloud and Edge
Cybersecurity
Monitoring
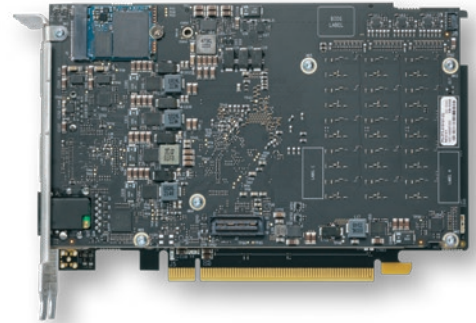Infrastructure
Financial

PLATFORMS
Link-Capture™ Software
Link-Inline™ Software
Link-Virtualization™ Software
Link-Storage™ Software
Link-Security™ Software
Link-Programmable™
Smart Network Interface Cards (SmartNICs)
Infrastructure Processing Units (IPUs)

APPLICATIONS
5G UPF Offload
vSwitch Offload
Network Offload
Security Offload
Storage Offload
Data Capture
+ More

SERVICES
Professional Services
Custom Development

NAPATECH F2070X IPU

Napatech Infrastructure Processing Unit (IPU) solution maximizes the performance of data center storage based on NVME over TCP (NVME/TCP)

Enterprise and cloud data centers are increasingly adopting the Non-Volatile Memory Express over Transmission Control Protocol (NVMe/TCP) storage technology because of the advantages it offers in terms of performance, latency, scalability, management and resource utilization. However, implementing the required storage initiator workloads on the server's host CPU imposes significant computational overheads and limits the number of CPU cores available for running services and applications.

This solution brief explains how an integrated hardware-plus-software solution from Napatech addresses this problem by offloading the storage workloads from the host CPU to an Infrastructure Processing Unit (IPU) while maintaining full software compatibility at the application level.

The solution not only frees up host CPU cores which would otherwise be consumed by storage functions but also delivers significantly higher performance than a software-based implementation. This significantly reduces data center CAPEX, OPEX and energy consumption. It also introduces security isolation into the system, increasing protection against cyber-attacks, which reduces the likelihood of the data center suffering security breaches and high-value customer data being compromised.

Napatech F2070X IPU
The Napatech F2070X Infrastructure Processing Unit (IPU) is a 2x100G PCI Express (PCIe) card with an Intel® Agilex® F-Series FPGA and an Intel® Xeon® D processor, in a Full Height, Half Length (FHHL), dual-slot form factor.
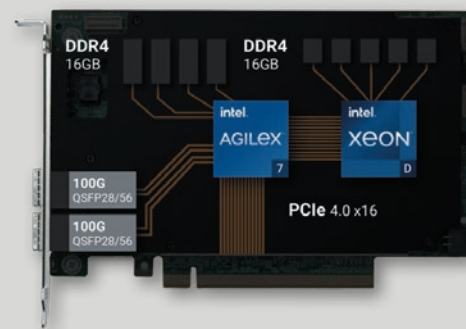
The standard configuration of the F2070X IPU comprises an Agilex AGF023 FPGA with 4 banks of 4GB DDR4 memory, together with a 2.3GHz Xeon® D-1736 SoC with 2 banks of 8GB DDR4 memory, while other configuration options can be delivered to support specific workloads.

The F2070X IPU connects to the host via a PCIe 4.0 x16 (16 GT/s) interface, with an identical PCIe 4.0 x16 (16 GT/s) interface between the FPGA and the processor.

Two front-panel QSFP28/56 network interfaces support network configurations of
• 2x 100G;
• 8x 10G or 8x 25G (using breakout cables).

Optional time synchronization is provided by a dedicated PTP RJ45 port, with an external SMA-F and internal MCX-F connector. IEEE 1588v2 time-stamping is supported.

Board management is provided by a dedicated RJ45 Ethernet connector. Secure FPGA image updates enable new functions to be added, or existing features updated, after the IPU has been deployed.

The processor runs Fedora Linux, with a UEFI BIOS, PXE boot support and full shell access via SSH and a UART.

NVMe over TCP: the optimum storage technology for modern data centers
NVMe over Transmission Control Protocol (NVMe/TCP) is a storage technology that allows Non-Volatile Memory Express (NVMe) storage devices to be accessed over a network using standard data center fabrics.

NVMe-oF   NVMe-oF

Application Servers   Switch Fabric   NVMe Storage Array
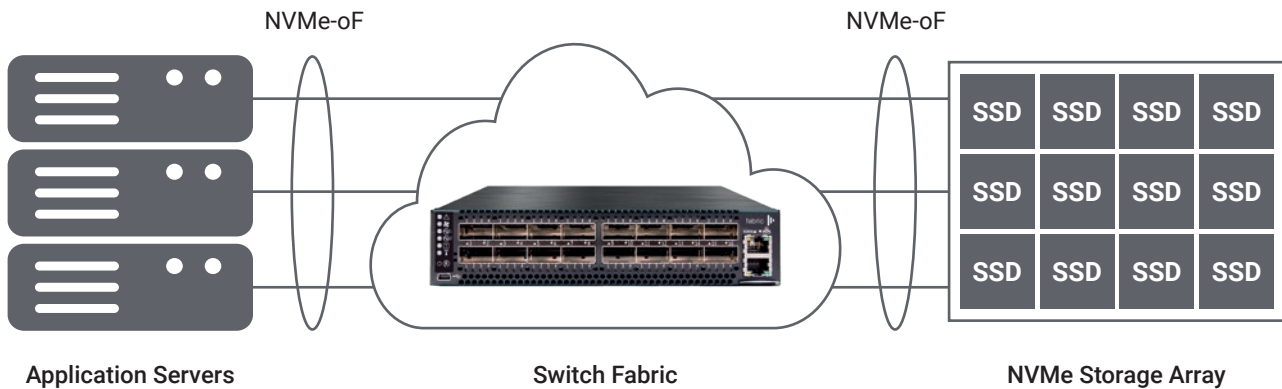
Modern cloud and enterprise data centers are increasingly adopting NVME/TCP due to the compelling advantages that it offers over older storage protocols like Internet Small Computer System Interface (iSCSI) Fibre Channel:

- **Higher Performance:** NVMe is designed to take full advantage of modern high-speed NAND-based Solid-State Drives (SSDs) and offers significantly faster data transfer rates compared to traditional storage protocols. NVMe/TCP extends these benefits to a networked storage environment, allowing data centers to achieve high-performance storage access over the fabric.
- **Reduced Latency:** The low-latency nature of NVMe/TCP is critical for data-intensive applications and real-time workloads. By minimizing the communication overhead and eliminating the need for protocol conversions, NVMe/TCP can help reduce storage access latencies and improve overall application performance.
- **Scalability:** Data centers often deal with large-scale storage deployments and NVMe/TCP allows for seamless scalability by providing a flexible and efficient storage access solution over a network. As the number of NVMe devices grows, data centers can maintain high levels of performance without significant bottlenecks.
- **Shared Storage Pool:** NVMe/TCP enables the creation of shared storage pools accessible to multiple servers and applications simultaneously. This shared storage architecture improves resource utilization and simplifies storage management, leading to significant cost savings.
- **Legacy Infrastructure Compatibility:** Data centers often have existing infrastructure built on Ethernet, InfiniBand or Fibre Channel networks. NVMe/TCP allows them to leverage their current fabric investments while integrating newer NVMe-based storage technology without having to overhaul the entire network infrastructure.
- **Efficient Resource Utilization:** NVMe/TCP enables better utilization of resources by reducing the need for dedicated storage resources on each server. Multiple servers can access shared NVMe storage devices over the network, optimizing the use of expensive NVMe storage resources.
- **Future-proofing:** As data centers continue to evolve and adopt faster storage technologies, NVMe/TCP provides a forward-looking approach to storage access, ensuring that storage networks can keep up with the growing demands of modern applications and workloads.

Overall, NVMe/TCP offers a powerful and flexible storage solution for data centers, enabling high performance, low latency and efficient resource utilization in a shared and scalable storage environment.

Limitations of software-only storage architectures

Despite the compelling benefits of NVMe/TCP for storage, data center operators need to be aware of significant limitations associated with an implementation in which all the required storage initiator services run in software on the host server CPU.

First, a system-level security risk is presented if either the storage virtualization software, the hypervisor or the virtual switch (vSwitch) is compromised in a cyber-attack.
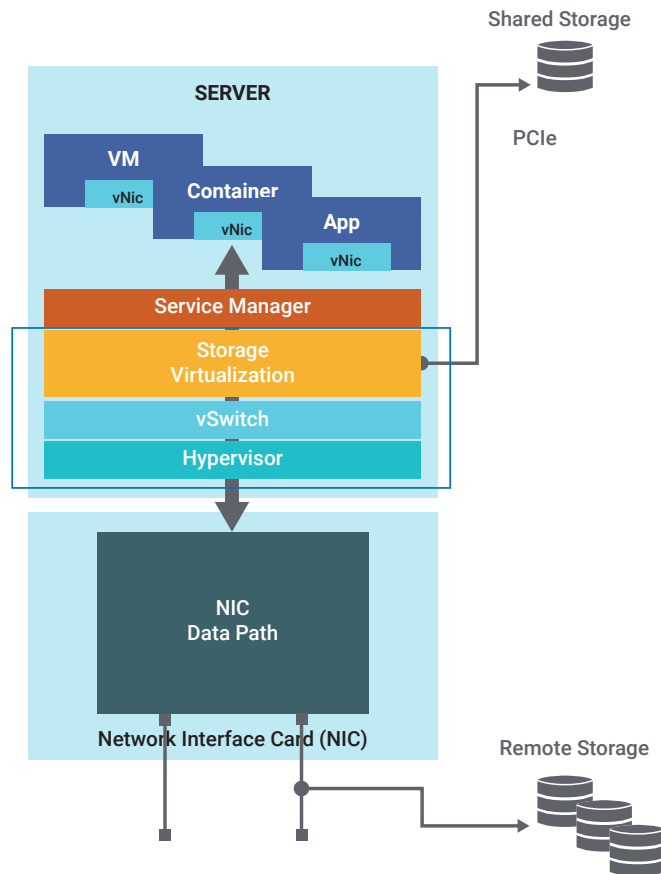
Second, there is also no way to ensure full isolation between tenant workloads. In a multi-tenant environment, a single architecture hosts multiple customers' applications and data. The "noisy neighbor" effect occurs when an application or VM uses the majority of available resources and degrades system performance for other tenants on the shared infrastructure.

Finally, a significant fraction of the host CPU cores are required for running infrastructure services such as the storage virtualization software, the hypervisor and the vSwitch. This limits the number of CPU cores that can be monetized for Virtual Machines (VMs), containers and applications. Reports indicate that between 30% and 50% of data center CPU resources are typically consumed by infrastructure services.

In a high-performance storage subsystem, the host CPU might be required to run a number of protocols such as Transmission Control Protocol (TCP), Remote Direct Memory Access over Converged Ethernet (RoCEv2), InfiniBand and Fibre Channel. When the host CPU is heavily utilized running these storage protocols as well as other infrastructure services, the number of CPU cores available for tenant applications is significantly reduced so that, for example, a 16-core CPU might only deliver the performance of a 10-core.

For these reasons and more, a software-only architecture presents significant business and technical challenges for data center storage.

IPU-based storage offload

Offloading the NVMe/TCP workload to an Infrastructure Processing Unit (IPU), in addition to other infrastructure services such as the hypervisor and vSwitch, addresses the limitations of a software-only implementation and delivers significant benefits to data center operators:
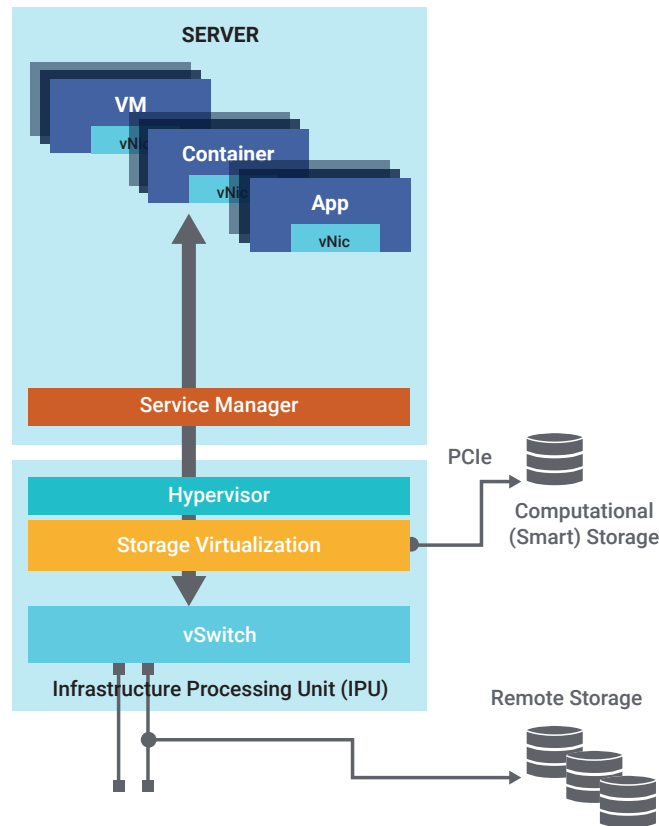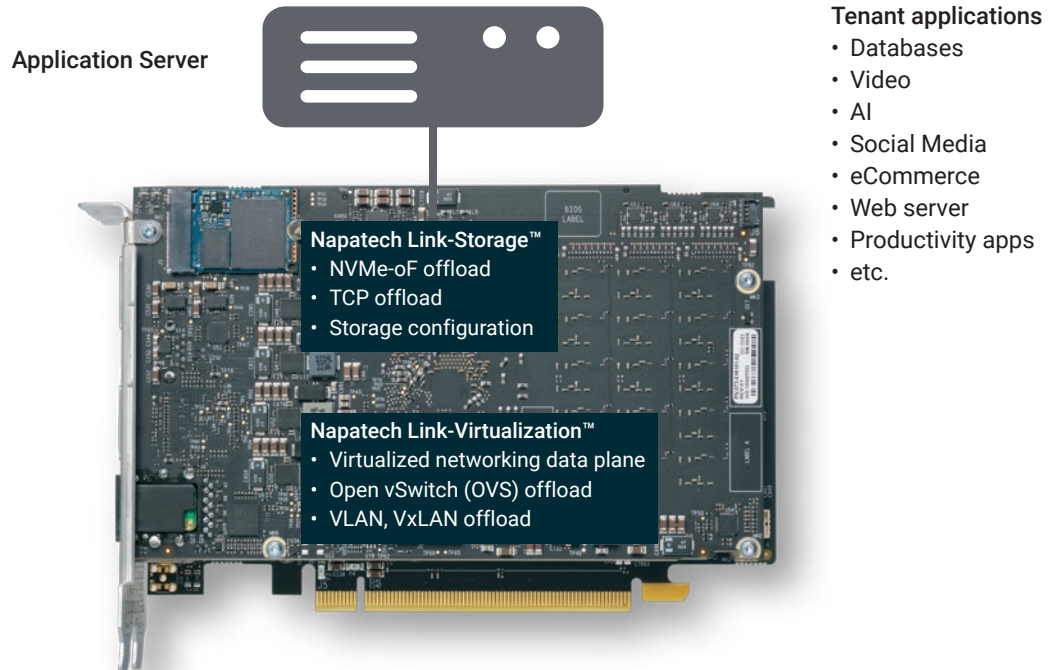

- **CPU Utilization:** NVMe/TCP communication involves encapsulating NVMe commands and data within the TCP transport protocol. Without offloading, the host CPU is responsible for processing these encapsulation and de-encapsulation tasks. Offloading these operations to dedicated hardware allows the CPU to focus on other critical tasks, leading to improved overall system performance and improved CPU utilization.
- **Lower Latency:** Offloading the NVMe/TCP communication tasks to specialized hardware can significantly reduce the latency associated with processing storage commands. As a result, applications can experience faster response times and better performance when accessing remote NVMe storage devices.
- **Efficient Data Movement:** Offloading non-CPU application tasks to discrete hardware accelerators enables data movement operations to be performed more efficiently than using a general-purpose CPU. It can handle large data transfers and buffer management effectively, further reducing latencies and improving overall throughput.
- **Improved Scalability:** Offloading NVMe/TCP tasks allows for better scalability in large-scale storage deployments. By relieving the CPU from handling the network communication, the system can support a higher number of concurrent connections and storage devices without becoming CPU-bound.
- **Energy Efficiency:** By offloading certain tasks to dedicated hardware, power consumption on the host CPU can be reduced. This energy efficiency can be especially important in large data center environments where power consumption is a significant consideration.

In addition to the above benefits that apply to the NVMe/TCP storage workload, the IPU-based system architecture provides incremental security isolation options, whereby the infrastructure services are isolated from tenant applications. This ensures that the storage, hypervisor and vSwitch services cannot be compromised by a cyber-attack launched by a tenant application. The infrastructure services themselves are secured since the boot process of the IPU itself is secure, while the IPU then acts as the root of trust for the host server.

Napatech storage offload solution

Napatech provides an integrated, system-level solution for data center storage offload, comprising the high-performance Link-Storage™ software stack running on the F2070X IPU.

See the sidebar for details of the F2070X IPU hardware.

Application Server

Tenant applications
- Databases
- Video
- AI
- Social Media
- eCommerce
- Web server
- Productivity apps
- etc.

Napatech Link-Storage™
- NVMe-oF offload
- TCP offload
- Storage configuration

Napatech Link-Virtualization™
- Virtualized networking data plane
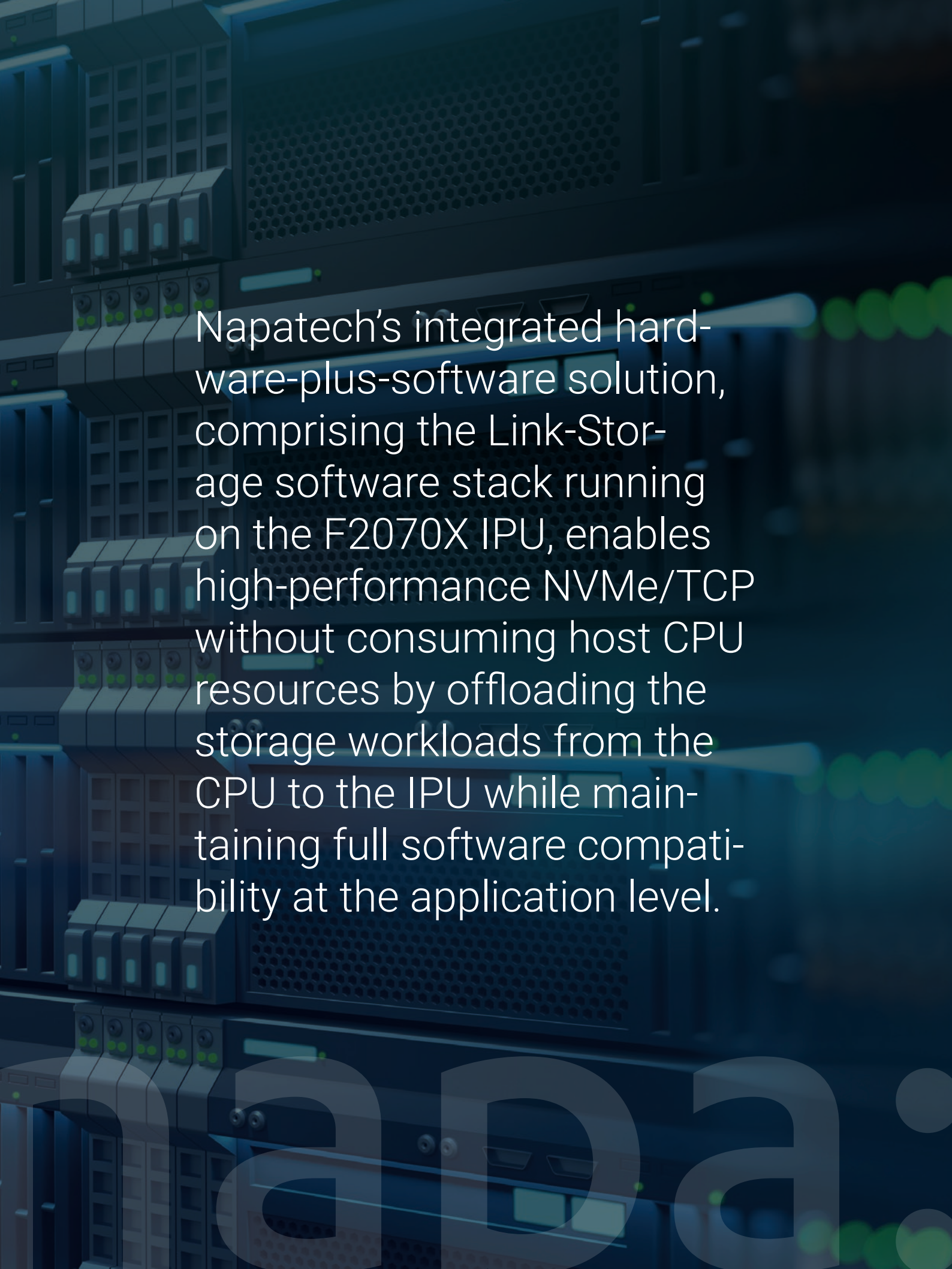- Open vSwitch (OVS) offload
- VLAN, VxLAN offload

The Link-Storage software incorporates a rich set of functions, including:

- Full offload of NVMe/TCP workload from the host to the IPU;
- Full offload of TCP workload from the host to the IPU;
- NVMe-TCP initiator;
- Storage configuration over Storage Performance Development Kit Remote Procedure Call (SPDK RPC) interface;
- Multipath NVMe support;
- Presents 16 block devices to the host via virtio-blk interface;
- Compatible with standard virtio-blk drivers in common Linux distributions;
- Security isolation between the host CPU and the IPU, with no network interfaces exposed to the host.

In addition to Link-Storage, the F2070X also supports the Link-Virtualization™ software which provides an offloaded and accelerated virtualized data plane including functions such as Open vSwitch (OVS), live migration, VM-to-VM mirroring, VLAN/VxLAN encapsulation/decapsulation, Q-in-Q, RSS load balancing, link aggregation and Quality of Service (QoS).

Since the F2070X is based on an FPGA and CPU rather than ASICs, the complete functionality of the platform can be updated after deployment, whether to modify an existing service, to add new functions or to fine-tune specific performance parameters. This reprogramming can be performed purely as a software upgrade within the existing server environment, with no need to disconnect, remove or replace any hardware.

Napatech's integrated hardware-plus-software solution, comprising the Link-Storage software stack running on the F2070X IPU, enables high-performance NVMe/TCP without consuming host CPU resources by offloading the storage workloads from the CPU to the IPU while maintaining full software compatibility at the application level.

Industry-leading performance

The Napatech F2070X-based storage offload solution delivers industry-leading performance on benchmarks relevant to data center use cases, including:

- **Latency:** less than 10µs latency added for remote NvME access compared to local access.
- **Throughput:** 2x 100G minimum throughput for both read and write operations (constrained by the bandwidth of the PCIe bus and the target storage array).
- **Input/Output Operations per Second (IOPS):** 6M IOPS at 4KB I/O block size.
- **CPU core utilization:** no host CPU cores utilized for NVMe/TCP or supporting networking operations (assuming that Napatech's network offload solution is also deployed).

These benchmarks are measured as follows:

- The Flexible IO Tester (fio) performance benchmarking tool is used.
- **Latency:** measured based the delay of an I/O send/complete.
- **Throughput:** bandwidth of a transfer in the case of a block size of 128KB under sequential reads and writes.
- **IOPS:** measurements based on a block size of 4KB for random reads and writes.
- **CPU core utilization:** total number of CPU cores consumed by I/O and networking operations.

Note that the above benchmarks are preliminary pending general availability of hardware and software.
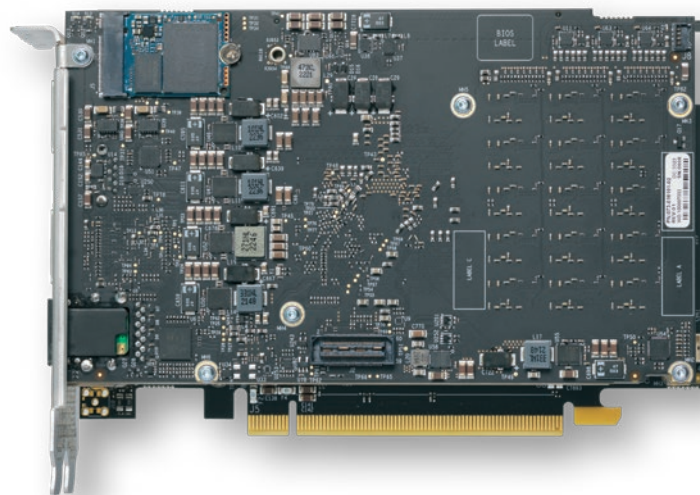
Summary

Enterprise and cloud data centers are increasingly adopting the NVMe/TCP storage technology because of the advantages it offers in terms of performance, latency, scalability, management and resource utilization. However, implementing the required storage initiator workloads on the server's host CPU imposes significant compute overheads and limits the number of CPU cores available for running services and applications.

Napatech's integrated hardware-plus-software solution, comprising the Link-Storage software stack running on the F2070X IPU, addresses this problem by offloading the storage workloads from the host CPU to the IPU while maintaining full software compatibility at the application level.

Napatech's storage offload solution not only frees up host CPU cores which would otherwise be consumed by storage functions but also delivers significantly higher performance than a software-based implementation. This significantly reduces data center CAPEX, OPEX and energy consumption.

The Napatech solution also introduces security isolation into the system, increasing protection against cyber-attacks, which reduces the likelihood of the data center suffering security breaches and high-value customer data being compromised.

For more information, visit: Napatech at https://www.napatech.com/support/resources/solution-descriptions/storage-offload-solution/

napa:tech;

SMARTNIC+IPU SOLUTIONS

Napatech is the leading supplier of Smart Network Interface Card (SmartNIC) and Infrastructure Processing Unit (IPU) solutions used in cloud, enterprise, and telecom datacenters.

Through commercial-grade software suites integrated with high-performance hardware, Napatech accelerates network infrastructure and security workloads to deliver best-in-class system-level performance while maximizing the availability of server compute resources for applications and services.

Additional information is available at:
www.napatech.com

NAPATECH RECONFIGURABLE COMPUTING

EUROPE, MIDDLE EAST AND AFRICA
Napatech A/S
Copenhagen, Denmark
T +45 4596 1500
info@napatech.com
www.napatech.com

NORTH AMERICA
Napatech Inc.
Portsmouth, NH, USA
T +1 888 318 8288
info@napatech.com
www.napatech.com

ASIA-PACIFIC
ntapacsales@napatech.com
www.napatech.com