Converge! NETWORK DIGEST | AvidThink®

2020 Next-Gen Infrastructure Acceleration and SmartNICs Report

RESEARCH BRIEF

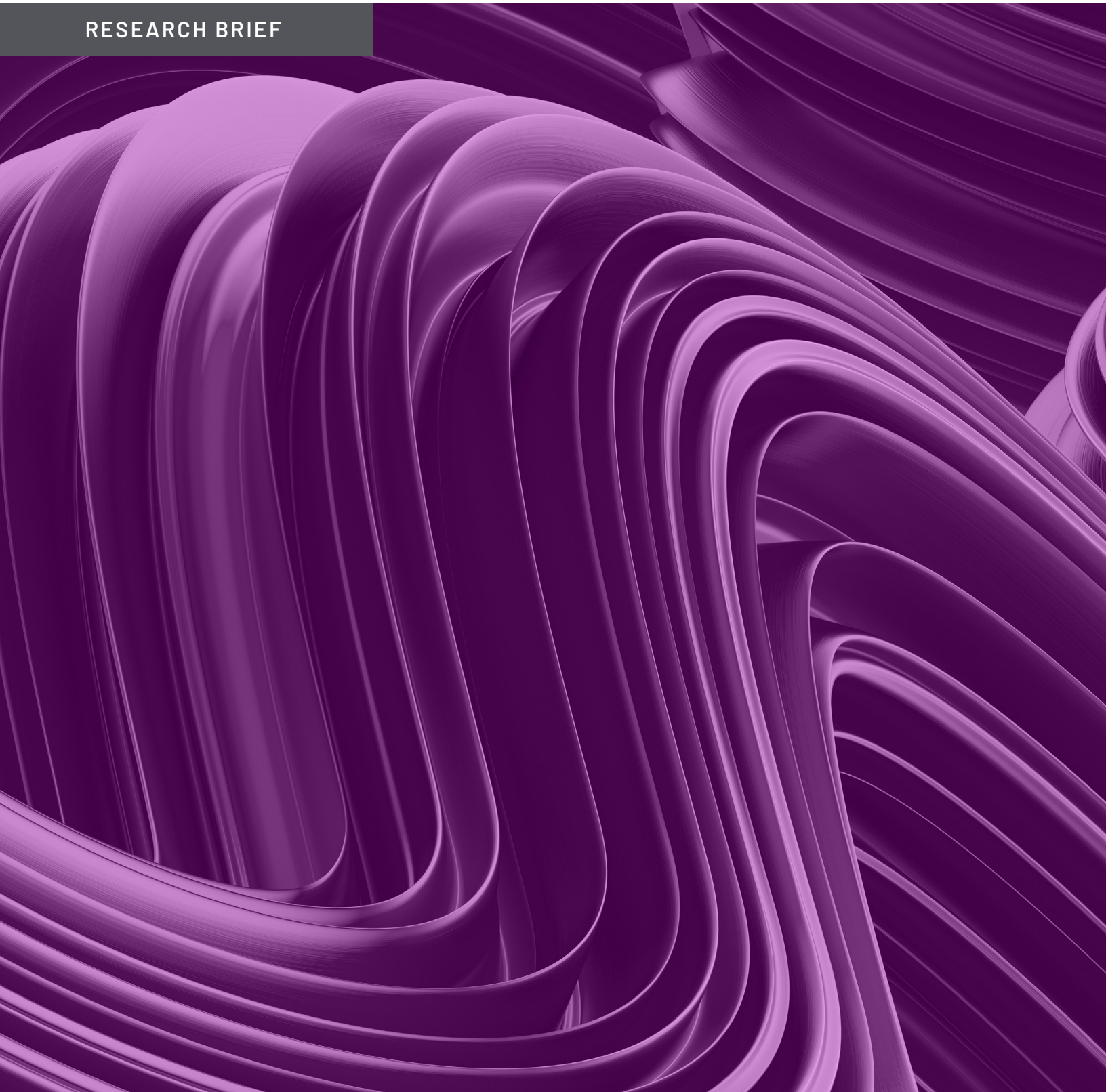


Table of Contents

© Copyright 2020 AvidThink LLC, All Rights Reserved

Table of Contents

Research Briefs are independent content created by analysts working for AvidThink LLC. These reports are made possible through the sponsorship of our commercial supporters. Sponsors do not have any editorial control over the report content, and the views represented herein are solely those of AvidThink LLC. For more information about report sponsorships, please reach out to us at **research@avidthink.com**.

About AvidThink™

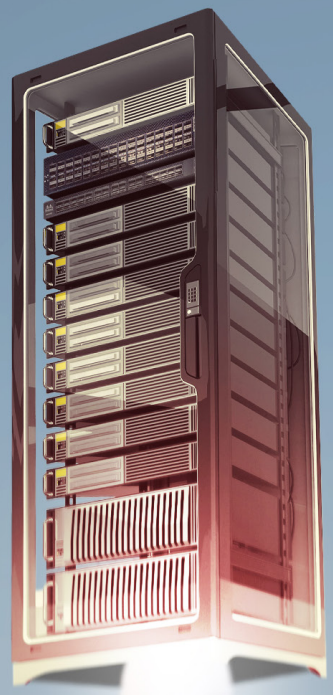
AvidThink is a research and analysis firm focused on providing cutting edge insights into the latest in infrastructure technologies. Formerly SDxCentral's research group, AvidThink launched as an independent company in October 2018. Over the last five years, over 110,000 copies of AvidThink's research reports (under the SDxCentral brand) have been downloaded by 40,000 technology buyers and industry thought leaders. AvidThink's expertise covers Edge and IoT, SD-WAN, cloud and containers, SDN, NFV, hyper-convergence and infrastructure applications for AI/ML and security. Visit AvidThink at **www.avidthink.com**.

© Copyright 2020 AvidThink LLC, All Rights Reserved

NVIDIA

ACCELERATE YOUR HYPER-CLOUD INFRASTRUCTURE WITH NVIDIA SMART NETWORKING SOLUTIONS

10G to 400G end-to-end Smart Networking solutions | ConnectX SmartNICs and BlueField DPUs accelerate Cloud and Telco networks | Open Ethernet Switching Fabric - Spectrum Ethernet Switches with the Cumulus Linux NOS make cloud fabrics faster, smarter and flexible | Smart-everywhere security protects applications, data, and infrastructure | GPU Acceleration provides performance that scales



www.nvidia.com

2020 Next-Gen Infrastructure Acceleration Report

Introduction

When Communications Service Providers (CSPs) teamed up in 2012 to launch Network Functions Virtualization (NFV), they outlined a vision whereby virtualized networking workloads would run on the same type of standard servers as deployed in their IT data centers. As NFV progressed, however, it became clear that the workloads imposed too much CPU overhead for this approach to be cost-effective. Recognizing that accelerated computing needs accelerated networking, CSPs started considering acceleration technologies such as fast-path software, Smart Network Interface Cards (SmartNICs), Application-Specific Integrated Circuits (ASICs), Field-Programmable Gate Arrays (FPGAs) and Graphics Processing Units (GPUs).

These strategies mirror the approaches adopted by cloud service providers such as Amazon, Google, and Microsoft to maximize the efficiency of their cloud and edge data centers: offloading acceleration functions frees up processor cores to run monetizable workloads.

The research brief primarily discusses infrastructure acceleration technologies adopted by CSPs and explores the products available from leading vendors. However, the same techniques and observations apply to data center use cases for both enterprises and cloud providers as well. Fundamentally, the workloads across all these different verticals will likely be similar, with the exception that hyperscale cloud providers will need to run these workloads at scale.

Next-Generation Workloads for Next-Generation Use Cases

Across all use cases, one of the key drivers behind the use of infrastructure acceleration technologies is the steep increase in bandwidth utilization. Globally, the average 5G connection will generate 2.6X more traffic than the average 4G connection by 2022[1]. Meanwhile, average broadband speeds will increase from 45.9 Mbps in 2018 to 110.4 Mbps in 2023[2]. All that traffic terminates in an application living on a data center somewhere and has to be processed.

To enable servers to process that data efficiently requires the use of infrastructure acceleration technology. Beyond general drivers, there are next-generation use cases on the horizon that will further drive new workloads with high-performance requirements.

5G Services

5G continues to be rolled out worldwide despite the recent pandemic. While raw data rates anticipated for 5G mobile broadband connections (up to 10 Gbps) will require unprecedented network processing bandwidth, the nature of 5G traffic and applications present additional challenges.

For example, the 5G Control and User Plane Separation (CUPS) architecture offloads data plane processing to the User Plane Function (UPF), enabling a programmable data path that can evolve to match new use cases. Similarly, 5G will bring a massive increase in the number of supported devices, both human and IoT, increasing the number of unique flows both at the edge and

[1] https://blogs.cisco.com/sp/5g-gets-top-billing-but-dont-forget-wi-fi

[2] https://www.cisco.com/c/en/us/solutions/executive-perspectives/annual-internet-report/index.html

© Copyright 2020 AvidThink LLC, All Rights Reserved

in the core. This will require secure, ultra-low latency classification, steering, and processing of network traffic.

In addition, the recent trends around disaggregation of the radio access network (RAN) in the Open RAN initiative from the Telecom Infra Project (TIP), the ORAN Alliance, and the Open Networking Foundation, point to commercial-off-the-shelf (COTS) servers being used to process traffic from the 5G RAN. This will likely require hardware acceleration beyond what general-purpose CPUs can provide.

Edge Compute

One of the critical elements of a 5G network is edge compute (labeled multi-access edge compute or MEC in the telecoms space). Edge compute existed prior to 5G, but it has become prominent because many of 5G's capabilities like ultra-low latency require the use of edge computing. CSPs worldwide are exploring edge compute applications that represent new business opportunities with new categories of customers. Many of these edge applications are not only data-intensive but also impose high-performance requirements on the networking infrastructure.

Edge compute existed prior to 5G, but it has become prominent because many of 5G's capabilities like ultra-low latency require the use of edge compute.

Artificial Intelligence (AI) and Machine Learning (ML) are critical to edge use cases such as assisted driving, Augmented Reality (AR), video surveillance, and voice control. Both AI and ML workloads are executed more efficiently on accelerated networking platforms than on standard servers.

In the IoT market, analytics, big data filtering, and image processing are examples of applications where edge-hosted acceleration is key to cost-effective deployments.

Storage-Defined Storage

In CSP-centric discussions, we sometimes forget that most of our new workloads will require significant storage. While not a vertical use case, nor new technology, storage is increasingly moving towards a disaggregated software-defined model both in the data center and even at the edge. A quick look at the Open Compute Project (OCP) portfolio of designs, and it's clear that storage servers make up an essential element of the next-gen data center. Flash and especially Non-Volatile Memory Express (NVMe) are critical storage technologies. Even more so are the protocols to access them over a network, such as NVMe over Fiber (NVMe-oF), NVMe over TCP (NVMe/TCP), and Remote Direct Memory Access over Converged Ethernet (ROCEv2). Accelerating these protocols can improve application performance significantly.

Data Analytics and AI

Related to storage, many of the workloads today, whether at the edge or in central data centers, involve moving around and processing large amounts of data. Whether part of big data analytics with software platforms based on Apache Spark, Hadoop, Kafka, or machine and deep-learning platforms based around Tensorflow, PyTorch, or Keras, there's a significant amount of data shuffling and need to move large data sets between servers rapidly.

© Copyright 2020 AvidThink LLC, All Rights Reserved

Key Infrastructure Acceleration Technologies

There are two main classes of infrastructure acceleration technology. The first is software-based, focused on providing a fast-path for packets within a server while taking advantage of unique CPU and server architecture features. The second class is hardware-based, utilizing different processing architectures more suited to parse and dispatch network packets than general-purpose CPUs. In many deployments, data center operators will use a combination of both software and hardware technologies to provide acceleration. For instance, using DPDK with SmartNICs based on various technology types, or combining VPP with DPDK and a SmartNIC. We'll now cover the major frameworks and solution types for both classes and then wrap up the section with observations from our conversations with various carriers, cloud providers, and vendors.

Software Accelerators

The main idea behind software-based acceleration is to provide the means to accelerate packet handling and processing without necessitating the use of proprietary and non-standard hardware on servers. Often, these software approaches involve finding best practices in processing network traffic and codifying these practices in the forms of reusable libraries. Because these techniques may be operating system (OS) or instruction-set-architecture (ISA) dependent, software application developers who rely on these libraries need to be aware of the target deployment systems (e.g., x86 versus Arm ISA). Even on the same ISA, there may be variations in processor families — some will support acceleration instruction sets, and some won't. Developers will need a graceful fallback strategy for these circumstances.

In some situations, because of the coupling between software drivers and hardware NIC versions, an application that has taken advantage of some of these techniques have to be recompiled for different targets. In the end, software techniques are generally more portable across the different underlying infrastructure.

For more information on the various software acceleration techniques, we would encourage the reader to check out our recent report on "**Myth-busting DPDK in 2020**," at our **nextgeninfra.io** resource site. That report, sponsored by the Linux Foundation, contains a lot more detail about DPDK and the various software technologies for network data path acceleration.

Packet Processing Acceleration

Generically, several companies provide software that accelerates network packet processing by, for example, a fast path architecture in which the data plane is split into two layers.

The lower layer, the fast path, processes the majority of incoming packets outside the OS environment and without incurring any of the OS overheads that degrade overall performance. Only those packets that require complex processing are forwarded to the OS networking stack (the upper layer of the data plane), which performs the necessary management, signaling, and control functions.

Some of the more open frameworks in this space include BPF/eBPF (extended Berkeley Packet Filter), including eXpress Data Path (XDP), that allow user applications to compile packet handling code to run within the kernel and take appropriate actions, which might involve invoking the OS networking stack if needed.

DPDK

The Data Plane Development Kit (DPDK) software is a set of Linux user-space libraries and drivers that accelerate packet processing workloads running on all major CPU architectures. DPDK is probably the most popular software-based acceleration solution for CSP workloads in NFV. It works through bypassing the OS kernel and handling the packets in application user space.

An open-source project run by the Linux Foundation, DPDK doesn't replace the standard kernel network stack but is used for specific networking functions where high throughput and low latency are critical, such as wireless core, wireless access, wireline infrastructure, routers, load balancers, firewalls, video streaming, voice over IP and more. Several standard Linux distributions such as Fedora and Ubuntu include DPDK support in their packaging systems.

© Copyright 2020 AvidThink LLC, All Rights Reserved

VPP

Vector Packet Processing (VPP) is an extensible packet processing framework for network-intensive applications. It comes with a set of layer 2/3 switching and routing functionality built-in. Initially developed by Cisco Systems, it is now part of the open-source project FD.io hosted by the Linux Foundation. VPP is used in many layer 4-7 network functions to achieve higher throughput and better CPU efficiency.

VPP can run on DPDK, so that instead of using kernel drivers to get packets from the hardware, it takes direct hardware control to speed up the packet path, resulting in faster processing. VPP ensures that the fewest cycles possible are spent on packet processing by processing multiple packets in batches, so the CPU's caches remain hot, and cache misses are avoided.

Semiconductor Platforms

When NFV was first adopted, most CSPs were resistant to the idea of using hardware acceleration, preferring standard server motherboards with no dependence on specialized hardware. However, after the significant cloud providers like Microsoft Azure and Amazon Web Services unveiled platforms showing their use of SmartNICs to improve overall performance, CSPs changed their strategy and started adopting SmartNICs.

Regardless of how SmartNICs are built, most of these SmartNICs today interact with server OSes via relatively standard interfaces, often supporting Single-Root Input/Output Virtualization (SR-IOV), which virtualizes a physical NIC into multiple virtual NIC interfaces. We'll start our examination of hardware-based acceleration by taking a look at SmartNICs.

SmartNIC Cards

A SmartNIC, otherwise known as an "intelligent NIC," is a network adapter that offloads processing tasks from the main CPU. Using either an onboard multicore processor, NPU, FPGA, ASIC, or a combination thereof, the SmartNIC performs functions such as encryption, decryption, compression, decompression, firewall, filtering, TCP/IP processing, HTTP processing, virtual switching, storage acceleration, AI and ML.

SmartNICs are ideally suited to applications such as network acceleration and high-throughput web servers, enabling customers to maximize the consolidation of their workloads on virtualized environments while also accelerating their execution. We'll discuss the evolution of SmartNICs and their role in the next-gen data center later in this research brief. For now, we'll visit each of the critical component technologies that are used to power these SmartNICs.

FPGAs

FPGAs have traditionally been used as a preliminary step in the design of ASICs, accelerating time-to-market while volumes are low. FPGAs are developed using the same tools as those used to design ASICs but can be rewritten or reconfigured on the fly. The downside to FPGAs is their unit cost and their tendency to require larger power budgets.

When used for network processing, FPGAs can process up to about 200 Gbps without exceeding the power constraints of PCI Express (PCIe) slots. For edge applications like AI and ML, FPGAs combine dedicated hardware acceleration with software-like flexibility.

NPUs

NPUs or Network Processing Units were more prevalent in network acceleration in the early 2000s, with Intel's IXP family and EZChip (acquired by Tilera, in turn, acquired by Mellanox, finally acquired by NVIDIA) as leading vendors. Netronome subsequently licensed the IXP NPU design from Intel and became the most prominent NPU solution provider in the field. In recent years, NPU-based SmartNICs have yielded the market to a combination of ASIC and FPGA-based SmartNICs.

© Copyright 2020 AvidThink LLC, All Rights Reserved

GPUs

As their name implies, GPUs were originally designed to render computer graphics. However, their architectures, comprising multiple processor cores working in tandem, also make them valuable as coprocessors for workloads that are inherently parallelizable. GPUs can be used in network equipment to boost network applications like radio-access network (RAN) processing in 5G networks, replacing functions previously performed by dedicated digital signal processors (DSPs).

GPUs are often used for training in AI and ML applications, too, since they can process more functions in a given period of time than standard CPUs, making them well-suited to building and executing neural networks that must process high volumes of data.

Specialized ASICs

Once the functionality of an acceleration subsystem has been verified, often using an FPGA, an ASIC implementation is always more cost-effective, power-efficient, and performant than an FPGA. However, the upfront development and fabrication costs can run into millions of dollars for even a moderately-sized ASIC, so ASICs are typically used only for high-volume applications where the design is stable.

A key advantage of ASICs for network infrastructure applications is that, unlike FPGAs, they can include analog circuits such as transceivers on the same die as CPU cores. Some of the ASICs that can be used for infrastructure acceleration (whether as SmartNICs or in a different form factor) include those designed for network switching. These merchant-silicon switching ASICs are becoming increasingly programmable, supporting higher-level languages like the P4 programming language, which allow them to be used to support network acceleration functions.

Other Hardware Approaches — SOCs and DPUs

One of the areas of confusion around FPGA, ASICs, and System-on-Chip (SOCs) is how different vendors classify them. Generally speaking, a piece of silicon is classified as an SoC if it has a processor or microcontroller core on it. So, if an ASIC implementation bundles an Arm core inside, it's considered an SoC. This can get confusing because some of the more recent FPGAs also include processor cores embedded, for increased programming flexibility. At this point, the FPGA technically becomes an SoC.

More recently, the field has seen the introduction of a new category of processors, the Data Processing Unit (DPU). While the generic term has likely been used in the past, Fungible (which we cover later in this report) filed a patent in July of 2017, describing a system architecture centered around a DPU. NVIDIA's networking unit, Mellanox, has recently been using the same DPU term, which they've defined as a SOC consisting of a multicore CPU, NIC functions, and programmable acceleration engines. Both view the DPU as primarily the 3rd socket in the data center, with the CPU and GPU being the other two sockets.

Regardless of nomenclature, end-users will likely be more interested in the ultimate manifestation of these components: SmartNICs, programmable network switches, or network appliances. And even more, the real-world performance, manageability, programmability, and compatibility of these hardware accelerators with their data center hardware and software architectures.

Infrastructure Acceleration in Public Clouds

The real push behind the adoption of SmartNICs as crucial elements of infrastructure acceleration came from the public cloud providers. Microsoft Azure's Catapult project helped establish the use of FPGA-based SmartNICs for data center workloads. Starting around 2012, Microsoft developed the first incarnation of Catapult for its cloud storage platform, followed by its second generation for Bing and Azure workloads in 2015. Microsoft's Catapult, along with Amazon's use of their Nitro SmartNIC after 2015, helped legitimize the role that SmartNICs played in the data center, prompting CSPs around the world to start their own investigations using commercially available SmartNIC solutions.

© Copyright 2020 AvidThink LLC, All Rights Reserved

Amazon AWS

Amazon's Nitro system is the underlying platform for their Elastic Compute Cloud (EC2) web service that offloads networking, storage, and management services from the host servers to dedicated hardware. The initial version of Nitro was based on commercial-off-the-shelf (COTS) silicon, but in 2015 AWS acquired Annapurna Labs to develop custom ASICs that improve hardware performance, move Elastic Block Store (EBS) processing from the main CPU and offload all I/O virtualization to separate cards.

Google's analysis shows that these capabilities have allowed them to seamlessly upgrade their network infrastructure across five generations of virtual networking, increasing VM-to-VM bandwidth by nearly 18x while reducing latency by 8x.

The Amazon family of Nitro cards includes cards for Virtual Private Cloud (VPC), Elastic Block Store (EBS), instance storage, control plane, and security. Nitro now controls all AWS's compute infrastructure. The Nitro Hypervisor is a lightweight hypervisor that manages memory and CPU allocation, delivering a performance that is indistinguishable from bare metal.

Google

Within the Google Cloud platform, the Andromeda virtual network stack provides up to 100 Gbps bandwidth by leveraging NVIDIA GPU accelerator platforms. Google positions their "100 Gbps Accelerator VMs" as being ideal for high-throughput applications like scientific modeling, high-performance web servers, virtual network appliances, highly-scalable multiplayer gaming, video encoding services, distributed analytics, machine learning, and deep learning.

Andromeda also takes advantage of Intel QuickData DMA Engines to offload payload copies of larger packets. Driving the DMA hardware directly from the Andromeda stack, which bypasses the OS, enables the stack to spend more time processing packets rather than moving data around. The Andromeda architecture allows Google to offload other virtual network processing to hardware opportunistically, improving performance and efficiency without requiring the use of Single Root Input / Output Virtualization (SR-IOV) or different approaches that tie a VM to a physical machine for its lifetime.

Google's analysis shows that these capabilities have allowed them to seamlessly upgrade their network infrastructure across five generations of virtual networking, increasing VM-to-VM bandwidth by nearly 18x while reducing latency by 8x.

Microsoft

Microsoft continued work on its Catapult platform, launching the third generation around 2017 to both accelerate AI workloads and increase network speed for their Bing services. Microsoft also adopted the name "Azure Accelerated Networking" (AccelNet) for the solution that offloads host networking to their Catapult FPGA-based SmartNICs. Microsoft selected an FPGA approach because they believed that ASICs would not provide sufficient programmability, limiting their adaptability over time, while embedded CPU cores in an ASIC would not provide scalable performance, especially on single network flows. Multicore, System-on-Chip (SoC)-based SmartNICs were also evaluated, but considered to be inferior in terms of latency, price, and power, especially at high bandwidths.

Azure SmartNICs implementing AccelNet have been deployed on all new Azure servers since late 2015, with AccelNet service available to customers since 2016, providing consistent sub-15µs VM-to-VM TCP latencies and 32 Gbps throughput. AccelNet today supports primary Microsoft services such as Bing and Microsoft365.

© Copyright 2020 AvidThink LLC, All Rights Reserved

Architecturally, AccelNet enables SR-IOV to a VM as a high-performance data path bypassing the host and thereby reducing latency, jitter, and CPU utilization. All the AccelNet SmartNICs in Azure are upgradeable, with the rules tables and configuration typically being upgraded quarterly with new functionality.

The above examples of hyperscalers include only those based in the US. However, hyperscale cloud providers like Baidu, Alibaba and Tencent in China have also made extensive investments in SmartNICs. For the sake of brevity, we will not delve into their use cases in this research brief.

Communications Service Providers

Many CSPs are known to have deployments of infrastructure acceleration platforms in their networks. With 5G and the edge, CSPs are either using SmartNICs in their telco clouds or contemplating doing so.

Examples of CSPs who have spoken publicly about their use of SmartNICs include:

- AT&T is using SmartNICs to provide real-time measurement of traffic flows in the AT&T cloud;
- AT&T is also using 6WIND's packet processing software for accelerating networking functions;
- China Mobile, China Telecom, and China Unicom leverage Ethernity's acceleration technology (throughOEM partners);
- CenturyLink is using SmartNICs as part of its telco cloud infrastructure.

In addition to telcos and the major hyperscale cloud providers, other cloud hosting platforms, like Equinix's Packet bare metal services, have also indicated that they provide platforms with SmarNICs to accelerate network-centric workloads.

Vendor Solutions

Select vendors are discussed in this section, and this is not an exhaustive list of vendors. Instead, these are vendors that were mentioned by network operators in both the cloud and telecommunications space during AvidThink's research.

This space includes a large number of proprietary vendors who focus on developing unique IP, sometimes for niche applications — such as Algo-Logic, which provides the financial services with low-latency, high-performance solutions for trading and other use cases.

Some of these companies have been acquired by larger companies. For instance, Titan IC was acquired by Mellanox in March of this year, before NVIDIA closed on its acquisition of Mellanox. And in December last year, Cisco Systems Inc. grabbed Australian firm Exablaze, which developed packet processing cards for the algorithmic stock trading sector. Prior to this, Arista Networks bought field-programmable gate array maker Metamako Inc.

6WIND

6WIND provides server software for networking, deployed globally for 20 years at tier-1 customers such as AT&T, Cisco, NEC, and Nokia.

6WIND's vRouter is a software router that replaces expensive hardware for Border Router, CG-NAT, VPN Concentrator, Site-to-Site IPsec VPN, Mobile Security Gateway and hypervisor acceleration solutions. It is deployed on bare metal on commercial-off-the-shelf x86 servers such as Advantech, Dell PowerEdge, HPE ProLiant, Lanner, and Supermicro, or in virtual machines such as VMware ESXi, KVM or AWS, with licenses from 100 Mbps to 200 Gbps throughput. Management includes CLI and YANG-based NETCONF APIs for integration with third-party management tools and orchestrators.

6WINDGate is packet processing software delivered in source code form for line-rate Linux networking. It helps OEMs build 5G, TCP, telecom infrastructure, and network appliance solutions on x86 and Arm processors in months versus years for bare metal, virtual machine, and container applications. Management includes CLI and YANG-based NETCONF APIs for integration into OEM systems.

© Copyright 2020 AvidThink LLC, All Rights Reserved

AMD

AMD provides solutions for high-performance computing, graphics, and visualization technologies: building blocks for gaming, immersive platforms, and the data center.

AMD's discrete graphics processors are capable of accelerating computing workloads. The Radeon Instinct compute cards deliver high levels of performance for deep learning, high-performance computing (HPC) and cloud computing. This new accelerator is designed with optimized deep learning operations, double-precision performance, and hyper-fast HBM2 (High Bandwidth Memory).

To provide AI and machine learning acceleration, AMD combines their GPU solutions with the ROCm open ecosystem that includes Radeon Instinct optimized MIOpen libraries supporting AI frameworks like TensorFlow PyTorch and Caffe 2.

Broadcom

To complement its well-known merchant switching chipset product line, Broadcom has a family of networking acceleration products. Broadcom's Stingray SOC family combines a full-featured 100GbE NIC, a powerful 8-core CPU running at 3 GHz and hardware engines for packet processing, encryption, RAID, and de-duplication. Broadcom also ships SmartNIC solutions based on its SOC technology, such as the Stingray PS225, a SmartNIC with 2×25GbE ports. In addition to its onboard Arm CPU, it supports 16GB of DDR4 memory and comprehensive security features, including silicon root of trust and a 100Gbps hardware encryption engine. The PS 225 is designed for data plane acceleration and software-defined storage (SDS), including NVMe-over-Fabrics (NVMe-oF). Typical examples of offloaded services for the Stingray include virtual switching (OVS), software-defined storage, and data and network encryption (IPsec/SSL).

Broadcom has secured large customers, such as Baidu, who use its products in their cloud infrastructure. According to Dell'Oro Group's 4Q19 Controller and Adapter report, Broadcom has the highest market share of the SmartNIC market among Ethernet adapter vendors at the end of 2019.

Cisco (VPP)

Cisco's Ultra Cloud Core User Plane Network Function (UPF) uses Vector Packet Processing (VPP) technology with added support for mobility functions, running in Linux user-space, and leveraging DPDK to provide a high-performance User Plane using commodity hardware.

The VPP software uses fast packet processing technology that provides lower cost-per-bit to the customer compared with software based on either packet forwarding technology or on open virtual switch technology. Cisco has successfully demonstrated 400 Gbps throughput per server using VPP technology.

The Cisco UPF software architecture enables operators to deploy a User Plane either at a remote edge or at central sites without sacrificing any benefits. The majority of standard or custom features are offloaded to software-based packet forwarding processes with added plug-ins to support mobility-related functions. This software-based traffic handling technology is designed to support future hardware-based traffic offloading technologies in order to achieve even higher throughput, scaling, and performance.

© Copyright 2020 AvidThink LLC, All Rights Reserved

Ethernity Networks

Ethernity Networks provides networking and security solutions on programmable hardware for accelerating telco/cloud edge networks. Ethernity's FPGA logic offers complete data plane processing with a rich set of networking features, security, and a wide range of virtual function accelerations for optimizing edge networks.

Ethernity's advantage comes from its Router-on-NIC offering, which provides switch/router functionality and an NFVI gateway, delivered on their own FPGA-based SmartNIC while also available as an SoC. With a programmable data plane and features such as load balancing, monitoring, and SLA, Ethernity's Router-on-NIC enables acceleration of SD-WAN, vRouter, vBNG, and numerous other edge applications.

With operators seeking to build open 5G networks based on network virtualization, Ethernity's Router-on-NIC provides 5G User Plane Function (UPF) acceleration, which requires complete offload of the data plane to achieve 5G performance targets. The solution offers advanced features such as traffic management, deep packet buffers, classification, tunneling, flow counters, IPsec, and more. The platform is programmable and integrates with any VNF running over DPDK, making it adaptable to evolving architectural requirements as the 5G market develops.

Fungible

Fungible was founded by Pradeep Sindhu (formerly founder and CEO at Juniper Networks) and Bertrand Serlet (formerly SVP of Software Engineering at Apple).

Fungible's Data Processing Unit (Fungible DPU) has been positioned as the "third socket" in data centers, complementing the CPU and GPU, and promising to deliver benefits not just in performance per unit power and space, but also strengthening reliability and security. Fungible views its DPU as key to powering next-generation, high-performance, efficient, and cost-optimized scale-out data centers.

Fungible's Data Processing Unit (Fungible DPU) has been positioned as the "third socket" in data centers, complementing the CPU and GPU, and promising to deliver benefits not just in performance per unit power and space, but also strengthening reliability and security.

The Fungible DPU is purpose-built to address two of the biggest challenges in scale-out data centers — inefficient data interchange between nodes and inefficient execution of data-centric computations. Data-centric computations are increasingly prevalent in data centers, with notable examples being the computations performed in the network, storage, security, and virtualization data-paths. However, existing processor architectures cannot process these computations efficiently, leading to over-provisioning and underutilization of resources. This drives up the costs to build and operate data centers. Fungible believes that eliminating these inefficiencies will also accelerate the proliferation of modern applications, such as AI and analytics.

Intel

The Intel Agilex family combines an FPGA fabric built on Intel's 10nm process with heterogeneous 3D System-in-Package (SiP) technology. This enables the integration of analog, memory, custom computing, custom I/O and Intel eASIC device tiles into a single package with the FPGA fabric.

Diamond Mesa, Intel's latest eASIC device, allows reuse of IP from FPGA-based designs and can cut power consumption by as much as 50% while lowering unit costs, in only half of the time needed to develop a similar ASIC.

© Copyright 2020 AvidThink LLC, All Rights Reserved

Intel offers PCIe-based FPGA Programmable Acceleration Cards (PACs). The N3000 is a PAC that features an integrated network interface card (NIC) in a small form factor. It enables high throughput, low latency, and low power-per-bit performance for custom networking pipelines.

The Intel® Acceleration Stack for Intel Xeon® CPU with FPGAs provides optimized and simplified hardware interfaces as well as software APIs. It supports virtualization platforms such as vSphere from VMware. Intel FPGAs also support the Intel oneAPI unified programming model.

Marvell

Marvell's SmartNICs solutions are primarily powered by its assets from Cavium, which they acquired in 2018. Cavium has been one of the industry leaders in building multicore infrastructure processors built around the Arm architecture, but with additional acceleration functions for compression/decompression, pattern-matching and security. Their LiquidIO Smart NIC family of intelligent adapters benefits from Cavium's long history in infrastructure acceleration and provides programmable server adapter solutions for various data center deployments. Meanwhile, their OCTEON family of multicore infrastructure processors target a wide variety of networking and wireless equipment, including switches, routers, secure gateways, firewalls, network monitoring, 5G Macro Base station for Transport.

With applications in hyperscale data centers, hybrid public/private clouds, security VPN appliances, and telco NFV cloud environments, LiquidIO Smart NICs offload server-based networking functions from the server CPUs directly onto the PCIe adapter, maximizing application line-rate performance, increasing the number of VM instances per core. This reduces the total cost of ownership (TCO) is reduced for data center deployments. These adapters are offered with software packages supporting OVS offload, IPSec security/DDoS mitigation, and compression/decompression acceleration for virtualized servers while enabling customers to customize the product for specific use cases.

NVIDIA BlueField DPU accelerates control and data plane application performance and functionality for cloud and edge, such as NVMe SNAP storage virtualization, bare metal networking, or cloud-native and security workloads including analytics, micro-segmentation, IDS/IPS, and firewalls.

NVIDIA Networking (Mellanox)

The NVIDIA Mellanox ConnectX and BlueField Data Processing Unit (DPU) SmartNIC product families deliver efficient 10G-200G Ethernet and InfiniBand to accelerate cloud, storage, broadcasting, AI, edge and telco workloads.

Industry analysts indicate NVIDIA SmartNICs have over 65 percent market share in the 25G and above Ethernet adapter market. The ConnectX SmartNICs feature built-in acceleration engines for RDMA over Converged Ethernet (RoCE), TLS/IPsec crypto offloads, Accelerated Switch and Packet Processing (ASAP^2) for virtual switching/routing, and NVMe over Fabrics (NVMe-oF) for storage, in addition to traditional networking offloads. Selected ConnectX models also deliver accurate time synchronization capabilities to enhance both 5G wireless infrastructure and IP video streaming.

NVIDIA BlueField DPUs integrate advanced ConnectX networking with an array of 64-bit Arm cores into a single system on a chip (SoC), delivering a blend of software-defined, hardware-accelerated application offload engines at speeds up to 200Gb/s. BlueField accelerates control and data plane application performance and functionality for cloud and edge, such as NVMe SNAP storage virtualization, bare metal networking, or cloud-native and security workloads including analytics, micro-segmentation, IDS/IPS, and firewalls.

© Copyright 2020 AvidThink LLC, All Rights Reserved

Napatech

Napatech is a provider of reconfigurable computing platforms that helps mobile operators by bringing hyper-scale computing benefits to their networks. Napatech enhances open and standard virtualized servers to boost innovation and release valuable computing resources that improve services and increase revenue. Their Reconfigurable Computing Platform is based on a broad set of FPGA software for leading IT compute, network, and security applications that are supported on a wide array of FPGA hardware designs.

Napatech's solution includes a product family of SmartNICs spanning the price-performance spectrum, including 1/10/25/40/50/100/200/400 GbE. The SmartNICs feature application-centric, production-grade software packages that improve application performance on any standard Linux- or Windows-based server. Combined, the software and hardware deliver line-rate networking performance, reduced server CPU utilization, and user programmability that delivers hardware performance at the speed of software innovation required in mobile, cloud and, edge networking.

Netronome

Netronome's Agilio SmartNICs offer a comprehensive network data plane acceleration platform, with support for accelerating Open vSwitch (used by Nokia Nuage, Ericsson, and Red Hat Enterprise Linux), uniquely Contrail vRouter / Tungsten Fabric (used by Juniper) and uniquely also eBPF based data planes, all at speeds up to 100 Gbps.

Netronome's Express VirtIO (XVIO) supports VM migration while offloading the hypervisor's networking functions. Acceleration also extends to security features like connection tracking, DoS prevention, and IPsec / TLS encryption, as well as QoS features like metering, scheduling, and shaping. Transparent inline TLS / SSH decryption with built-in load balancing enables scaling of content-aware security and visibility applications.

Agilio FX SmartNICs accommodate Linux-based control planes to support bare metal server configurations. Custom applications (e.g., VNFs) can be supported using server-based offload APIs and programmability in C, P4, and eBPF.

Pensando

Pensando Systems was founded by the famous "MPLS" (Mario Mazzola, Prem Jain, Luca Cafiero, and Soni Jiandani) executive team from Cisco, and is also supported by John Chambers (Cisco's former CEO), who is both its chairman and an investor. With investment from Lightspeed Venture Capital, Goldman Sachs, and HPE, Pensando has designed and built its Distributed Services Platform (DSP) — essentially a SmartNIC, which includes a custom, programmable P4 processor called Capri built on an Arm core. Pensando's Capri is optimized to provide acceleration for networking, storage, and security stacks with minimal latency, jitter, and low power requirements (30W @ 100 Gbps). Pensando ships both a 25 Gbps and a 100 Gbps version of their Pensando Distributed Services Card (DSC).

The DSC provides software-defined services at the server edge, eliminating an assortment of discrete appliances throughout the data center and simplifying IT operations. The Pensando DSC additionally enables pervasive network visibility using its hardware bi-directional flow streaming and traffic mirroring capabilities.

Pensando also provides a software orchestration solution in the form of its Policy and Services Manager (PSM), which delivers centralized lifecycle management, enterprise-grade security, and visibility across the data center.

© Copyright 2020 AvidThink LLC, All Rights Reserved

VMware

In 2016, VMware acquired PLUMgrid, a Software-Defined Networking (SDN) startup that had been one of the initial contributors to the IO Visor open-source project.

IO Visor is an I/O hypervisor engine that resides between the Linux OS and the hardware, along with a set of development tools. It is a kernel-based Virtual Machine (VM) for I/O instructions, rather than a replacement for a hypervisor like ESX since it only performs I/O functions. The IO Visor project was announced in 2015 as a project hosted by the Linux Foundation, comprising the IO Visor engine along with a set of development, management and operation tools as well as applications and IO modules. Besides PLUMgrid, other founding members included Broadcom, Cavium, Cisco, Huawei, and Intel.

IO Visor is available as a VMware vSphere Installation Bundle (VIB), which is a software package installed on a vSphere ESXi Host. Products such as the Cisco HyperFlex platforms leverage IO Visor for networking acceleration, as one of the installed VIBs.

Notably, the IO Visor project is also the progenitor of the XDP or eXpress Data Path framework. By leveraging eBPF, XDP provides a high performance, programmable network data path in the Linux kernel. XDP provides bare metal packet processing at the lowest point in the software stack, reducing latency, and the XDP framework also allows the addition of new processing functions without the need to modify the kernel.

Xilinx

Xilinx provides both standalone FPGAs and SmartNICs that accelerate networking infrastructure.

The Virtex UltraScale+ FPGAs provide high performance and integration at the 16nm process node, including serial I/O bandwidth and logic capacity. This family is used for applications like 400G networking.

The Alveo U50 family of SmartNICs, built on the UltraScale+ architecture, accelerates a range of workloads, including high-performance computing, networking (security, TOE), computational storage acceleration, data analytics, and video processing. Packaged in a 75-Watt, small form factor and equipped with 100 Gbps networking, PCIe Gen4, and HBM2, the Alveo U50 family can be deployed in a wide variety of servers.

Through the 2019 acquisition of Solarflare, Xilinx offers the XtremeScale X2 series of SmartNIC Ethernet adapters with Onload kernel bypass technology. X2 SmartNICs can provide real-time packet and flow information across thousands of virtual NICs. The combination of ultra-high bandwidth, ultra-low latency, ultra-scale connectivity, and packet telemetry allows these adapters to scale with the addition of new servers, Virtual Machines (VMs), or containers.

The upcoming Xilinx Radio Frequency System-on-Chip (RFSoC)-based NIC will improve 5G base station performance while reducing latency. PCIe cards based on the Versal Adaptive Compute Acceleration Platform (ACAP) will accommodate telecom acceleration workloads with further improvements in performance.

© Copyright 2020 AvidThink LLC, All Rights Reserved

What's Next for Infrastructure Acceleration

As discussed earlier, with the increasing importance of 5G and edge deployments, infrastructure acceleration technologies will be a critical part of any edge data center (or micro-data center) installation. SmartNICs will likely no longer be the exception. Just as Microsoft Azure has deployed SmartNICs on all new Azure servers since 2015 (over a million deployed), the majority of edge servers will have some hardware acceleration component — probably a SmartNIC. It's relatively well understood at this point that CPU scaling limits are close to being reached and that combining general-purpose CPUs with specialized hardware tuned for specific workloads is the path forward.

SmartNICs will get smarter and morph over time. Some of the innovation areas that AvidThink sees the most traction in include:

- **Security**: More recent incarnations of SmartNICs include a silicon root-of-trust (RoT) for secure bootup, as well as hardware enforcement of security zones, allowing the NIC to enforce security policy and protect the server against both internal and external attacks. In addition, SmartNICs are located in an ideal part of the network topology to inspect and provide visibility into all traffic entering and exiting a server (given that any encryption or decryption happens on the SmartNIC).
- **Bootstrap Platform**: Related to the silicon RoT capabilities, SmartNICs could serve as server bootup platforms, able to validate the overall system to ensure that the server has not been compromised.
- **Storage Applications:** SmartNICs are increasingly more storage protocol-aware, with some of these SmartNICs able to create a scalable fabric for distributed storage across a data center.
- **Composable Systems:** Fungible views its DPU-based SmartNIC as essentially enabling a data-center-wide composable system, building a fabric across multiple server racks that can compose banks of CPUs, FPGAs, GPUs, and NVMe storage devices into dynamically-partitioned systems. This is similar in concept to other composable data center solutions from companies like Liqid, which use PCIe switching fabrics to do the same within a cluster.

Further, alternate hardware-accelerated topologies are also being explored, such as using programmable switching chipsets on top-of-rack switches as hardware-assist for 5G workloads. For instance, Kaloom, a networking startup, in combination with Red Hat, have demonstrated a novel way to use switches with the Intel/Barefoot Tofino P4-programmable switch chip to accelerate 5G user-plane function workloads.

Summary and Recommendations

Through product evaluations as well as commercial deployments, CSPs have confirmed that running virtualized networking workloads on the same type of server configurations deployed in their IT data centers is not cost-effective, because of the CPU overhead imposed by these workloads. When implemented along with standard server platforms, however, acceleration technologies such as fast-path software, SmartNICs, FPGAs, DPUs and GPUs enable CSPs to meet their performance and cost goals. This mirrors approaches already successfully adopted by cloud service providers to maximize the efficiency of their data centers.

CSPs, System Integrators, and Network Equipment Providers should explore both hardware and software technologies for infrastructure acceleration in order to realize the business objectives of network virtualization. With the ongoing innovation in infrastructure acceleration, especially around SmartNICs, AvidThink expects that the telecommunications and enterprise markets will see a substantial return on their investment in this area.

© Copyright 2020 AvidThink LLC, All Rights Reserved

AvidThink, LLC
1900 Camden Ave
San Jose, California 95124 USA
avidthink.com

© Copyright 2020 AvidThink, LLC, All Rights Reserved
This material may not be copied, reproduced, or modified in whole or in part for any purpose except with express written permission from an authorized representative of AvidThink, LLC. In addition to such written permission to copy, reproduce, or modify this document in whole or part, an acknowledgement of the authors of the document and all applicable portions of the copyright notice must be clearly referenced. All Rights Reserved.

Rev C