HPE aruba networking

Five design principles for a smarter data center

Reduce data center complexity and accelerate automation, integration, and security with distributed services architectures.

HPE GreenLake

Introduction

As the demand to deliver value grows at a rapid pace, IT leaders have turned to application modernization and cloud operating models as keys to successful digital initiatives. To support this, traditional infrastructure must be modernized, but that can create as many challenges as it solves. Many organizations are unprepared, facing pockets of siloed compute and storage, disjointed network and security architectures, and operations that hinder centralized IT management, orchestration, security, policy, and visibility.

While data center networking technology has evolved to provide higher-performing 100/400G leaf-spine fabrics, security and services architectures have not advanced accordingly. Network modernization requires automation and API-based programmability to integrate with cloud orchestration and management platforms. This shift means infrastructure and operations must also be modernized to align with the cloud-centric, microservices-based app architectures and agile IT service provisioning operations that hyperscale data centers have been using for years.

It is a paradox that while the network is more important than ever, it must also work invisibly—without hindering application developers or business processes.

To compensate, many organizations use a stateless data center fabric that inefficiently bolts on network services and applies complex service chaining. But the sheer volume, velocity, and variety of data traffic requires a transition from manual, reactive, and siloed processes for managing network connections and data flows to ML/AI-powered platforms that establish, scale and secure connections and manage the infrastructure.

There is a better way—a network architecture that is simpler to provision, deploy, and manage and is transparently responsive to the needs of application developers, IT Ops, DevOps, and the business.

Next-generation data center fabrics allow organizations to move away from legacy architectures and compete on the same level as hyperscalers by consolidating stateful functions to the entire fabric, providing a wide variety of infrastructure services in a new and integrated manner. This requires a shift from thinking of the fabric as just a segmentation and connectivity solution, to one that supports all infrastructure services allowed for workload scale.

This paper examines five critical data center design principles to consider when designing a future data center:

- Modernize with hardware-accelerated DPU-enabled switches
- Transition to a fourth-generation distributed services architecture
- Extend Zero Trust closer to applications
- Blend network and security AIOps
- Leverage the edge, colocation, and IaaS

1 – Modernize with DPU-enabled switches

Traditionally, CPUs alone provided the processing power for hyperscale and enterprise data centers. More recently, GPUs (graphics processing units), have taken a significant role. Originally used to deliver rich, real-time graphics, their parallel processing capabilities make them ideal for accelerated computing tasks, including artificial intelligence, deep learning, and big data analytics applications.

A new type of processor has emerged on the scene called a DPU (data processing unit), quickly becoming an important part of the data-centric accelerated computing family. DPUs are purpose-built silicon used to offload data traffic so compute-intensive tasks can be optimized on CPU and GPU resources.

Using its own hardware-based processor, DPUs are often deployed in hyperscale data center servers and perform a large set of accelerated compute, cloud, network, security, and storage functions including encryption, firewall, load balancing, NAT, telemetry and more. These capabilities enable the isolated, bare-metal, cloud-native computing platforms that define the next generation of cloud-scale computing.

DPU technology has evolved from being server only to now being available in top-of-rack switches. This new category of distributed services switch marries standards-based Ethernet/IP based switching with embedded, hardware accelerated, fully programmable DPU technology in a unified, high performance secure networking solution for the data center and cloud.

”HPE Aruba Networking and AMD Pensando have enabled the industry’s first distributed services architecture that allows enterprises to create and operate network infrastructures that perform and scale just like the hyperscale infrastructure giants,”

Alan Weckel, founder and technology analyst, 650 Group.”

This allows operators to extend industry-standard leaf-spine networking with stateful distributed micro-segmentation, east-west firewalling, NAT, encryption, and telemetry services—delivered closer to where critical compute and storage workloads are being processed at the compute-network edge.

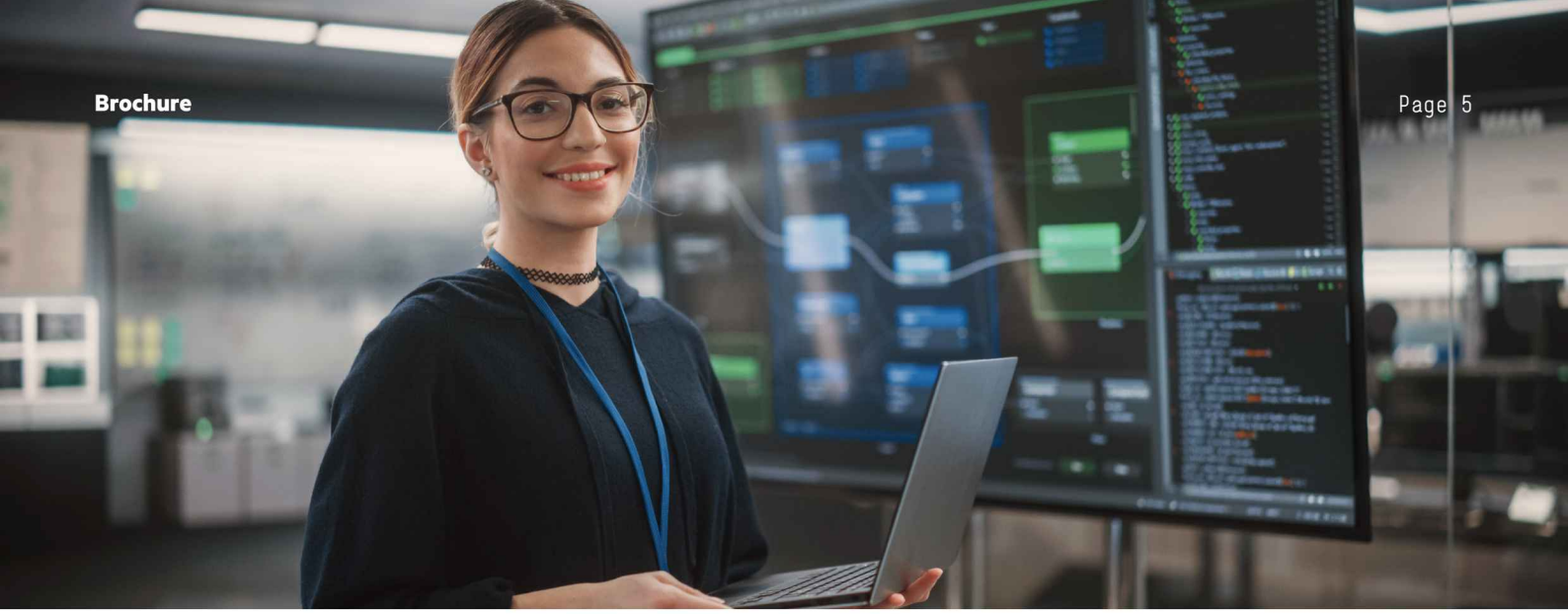
Unlike dedicated SmartNICs, which are installed in server(s), a distributed services switch can be deployed at the server top-of-rack and provides distributed services to all servers and hosts in the rack.

Distributed services switches do not require changes to server hardware or software or make assumptions on server operating systems, require any driver or agent to be installed on servers to provide distributed services at scale and wire-rate, and can be integrated into new or existing brownfield enterprise data centers and private clouds.

2 – Transition to a fourth-generation distributed services architecture

Today, silicon and software allow data center fabrics to provide the infrastructure services required to support workloads at scale, expanding them beyond just a segmentation and connectivity solution.

Hyperscalers discovered a decade ago that to scale their fabrics, they needed to eliminate the complexity associated with needing a different appliance for each OS (operating system) and service in the data center. Instead of bringing in a new vendor for each infrastructure service, they built each function into a single OS managed through a single automated controller. This integration and simplification allowed them to support millions of workloads.



Traditional/legacy Third generation	Next generation Fourth generation
• Switching and connectivity centric • Network/security services are bolted on per VM • Highly centralized L4-7 switches, limited scale • High complexity and cost (appliances, agents) • Limited ability to automate due to complexity	• Cloud-centric operational model • Fabric and infrastructure service unification • Fully distributed services across all data center workloads • Simplified service insertion at the data center fabric • Comprehensive automation, visibility and telemetry

Figure 1. Transition to a next-generation data center fabric

The **fourth generation** of data center architectures brings this same consolidation of stateful functions to the entire fabric. Rather than a stateless interconnect vehicle stitching workloads and services together, the fabric can now provide a wide variety of simplified and integrated infrastructure services, helping reduce design and provisioning complexity while ensuring that stateful services are available at the fabric edge.

This begins with two of the most crucial functions in the data center: east-west security, required for any Zero Trust deployment, and complete (not sampled) network telemetry. Both are essential for decomposed workloads. Security within the data center is critical to prevent breaches, while telemetry opens opportunities for emerging machine learning-based solutions that can automate security and network operations in ways not possible without high fidelity data center telemetry.

This new architecture also overcomes a sub-optimal design choice that many organizations have deployed—installing software agents on their servers to provide microsegmentation. By embedding these services into the network fabric, operators now have a superior design option that frees up valuable server CPU cycles—which would otherwise be stolen by having to process compute-intensive network services.

While creating such an architecture may be feasible when building out new hyperscale cloud environments from the ground up, how can existing data centers take advantage of this leading-edge technology?

The logical place to start would be to deploy these services in the top-of-rack (ToR) leaf switch to take advantage of a distributed services architecture without a costly, time-consuming forklift upgrade on the entire data center. This deployment strategy is very appealing as it allows organizations to migrate individual server rack(s) or data center PODs in a non-disruptive manner.

A fourth-generation distributed services architecture can:

- Reduce latency and improve security by applying services as close as possible to applications
- Eliminate appliance sprawl, lowering infrastructure and maintenance costs
- Reduce or eliminate the need to deploy costly software server agents (both in licensing and CPU processing)
- Optimize network performance and bandwidth by reducing latency via deploying services at the fabric edge
- Help drive operational and policy efficiencies for network and security operations teams

3 - Extend Zero Trust closer to your applications

Cybersecurity threats have changed dramatically in recent years. Zero Trust is an essential enterprise security practice that prevents data breaches and limits internal lateral movement by assuming an attacker is present in the environment. In the data center, this means distrusting any entity and all traffic on the network unless a security policy explicitly allows them. Segmentation statefully inspects all east-west traffic in the data center and applies policies to stop bad actors from moving laterally through the internal network.

Network services need to support disaggregated application scale. Historically, these services have all been deployed as discrete appliances or VMs connecting to the network but have not been part of the fabric. This leads to complexity, with different vendors to manage, traffic tromboning across the fabric, and complexity between networking and services teams.

“Our collaboration with AMD and HPE Aruba Networking in using the CX 10000 as the foundational element for our ‘DXC Secure Network Fabric’ for global data centers, has revolutionized our zero-trust security architecture for the data center and edge. Previously what required hundreds of virtual and physical firewalls to meet our segmentation and compliance requirements, is now being delivered natively inline on the platform resulting in projected TCO savings of 83%.”

- Nitin Jain, Global Network Lead, DXC Technology

By integrating stateful service capabilities within a data center fabric, security and visibility are pushed closer to where applications and workloads are processed without impacting the existing network architecture or software configurations. This improves data center security posture and visibility, while reducing the cost of acquisition and simplifying operations. The architecture inspects traffic right at top-of-rack (ToR) switches—removing the need to hairpin traffic through traditional centralized appliances and reducing network congestion and complexity.

A distributed services architecture expands Zero Trust deeper into the data center, to the network-server edge, delivering fine grain micro-segmentation, dramatically scaling and strengthening the security of mission-critical workloads—delivering 100X the scale, 10X the performance at a third of the TCO of traditional solutions.

4 - Blend network and security AIOps

Telemetry is relied on as a source of truth for what is happening in the data center. But to ensure that it is true, telemetry must be accurate and ubiquitous across the data center.

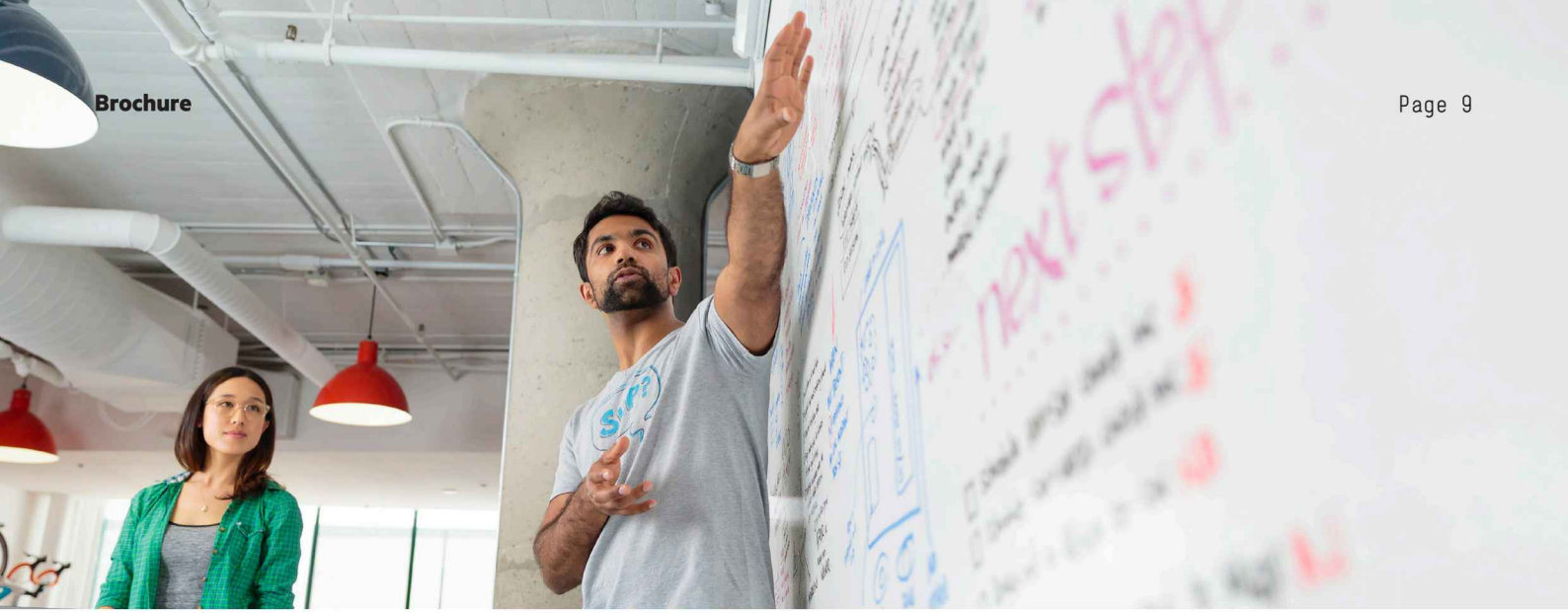
Network operations teams often struggle with not having enough telemetry to help address automation, simply visualizing the limited telemetry they have. Because today's data center fabrics are not capable of full telemetry, network probes and software agents must be deployed to know what is happening. Probes or agents can deliver very detailed telemetry, but only in the locations where they reside, thereby requiring sampled traffic flows across the data center to increase visibility. This method only captures snapshots of data traffic and falls short of the fidelity required by today's ML-based automation solutions.

This is a legacy third-generation problem based on a fragmented approach.

A distributed services architecture solves these challenges by providing accurate and ubiquitous telemetry across the entire data center, natively, with no traffic impact or bump in the wire insertion of devices (failure points). Probes and agents are no longer necessary, nor is the need for TAP-aggregation networks to gather telemetry. And with telemetry now part of the data center fabric, data center operators can reduce the extent of their telemetry "blind spot."

Next-generation architectures provide benefits that go beyond telemetry by:

- Providing networking teams MTTI (mean-time-to-innocence), with a time machine to look back at traffic, on a per app basis, for any flow in question and determine root cases to poor application performance
- Integrating via a Rest API and providing flow data to a wide variety of security and network performance tools including Advanced Security ML (XDR), Application Dependency Mapping (ADM), AI/Operations (AIOps), SIEM/SOAR, firewall compliance rules, and identity group mapping tools
- Automatically detecting anomalies, grouping them into related incident roots and notifying operations consoles, ticketing systems, and automation systems with real-time streaming data analysis, freeing up operational teams from having to visualize raw telemetry
- Moving away from having to deploy high cost, complex TAP aggregation networks that only offer limited insight, to fully leveraging AI/ML tools with complete high-fidelity telemetry that a fourth-generation data center fabric provides



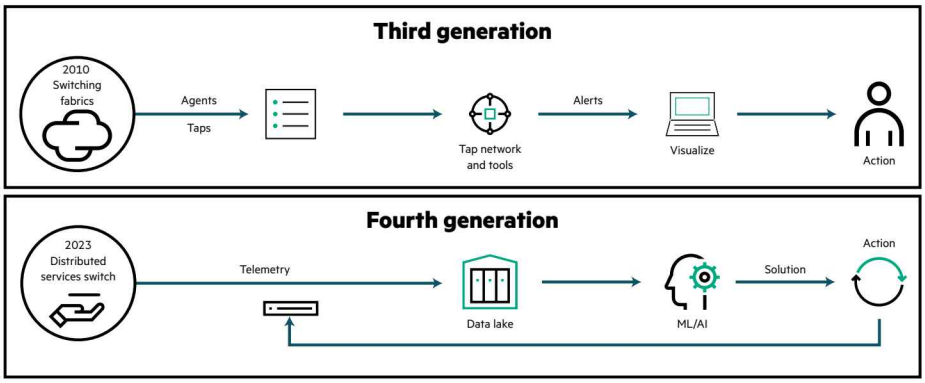


Figure 2. A new era of network operations

5 - Leverage the edge, colocation, and IaaS

Two powerful IT trends are converging: colocation, and edge-to-cloud infrastructure as a service. Most existing implementations are based on centralized architectures that collect and process data at core data centers, colocation centers, or in the cloud. But in today's world, an abundance of data is generated at the edge—in remote locations like factory floors, retail sites, healthcare facilities, smart buildings and cities.

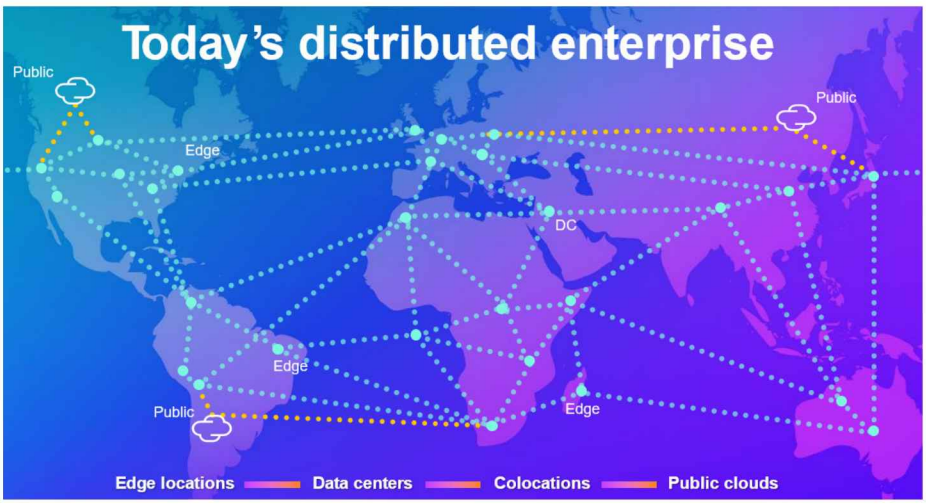
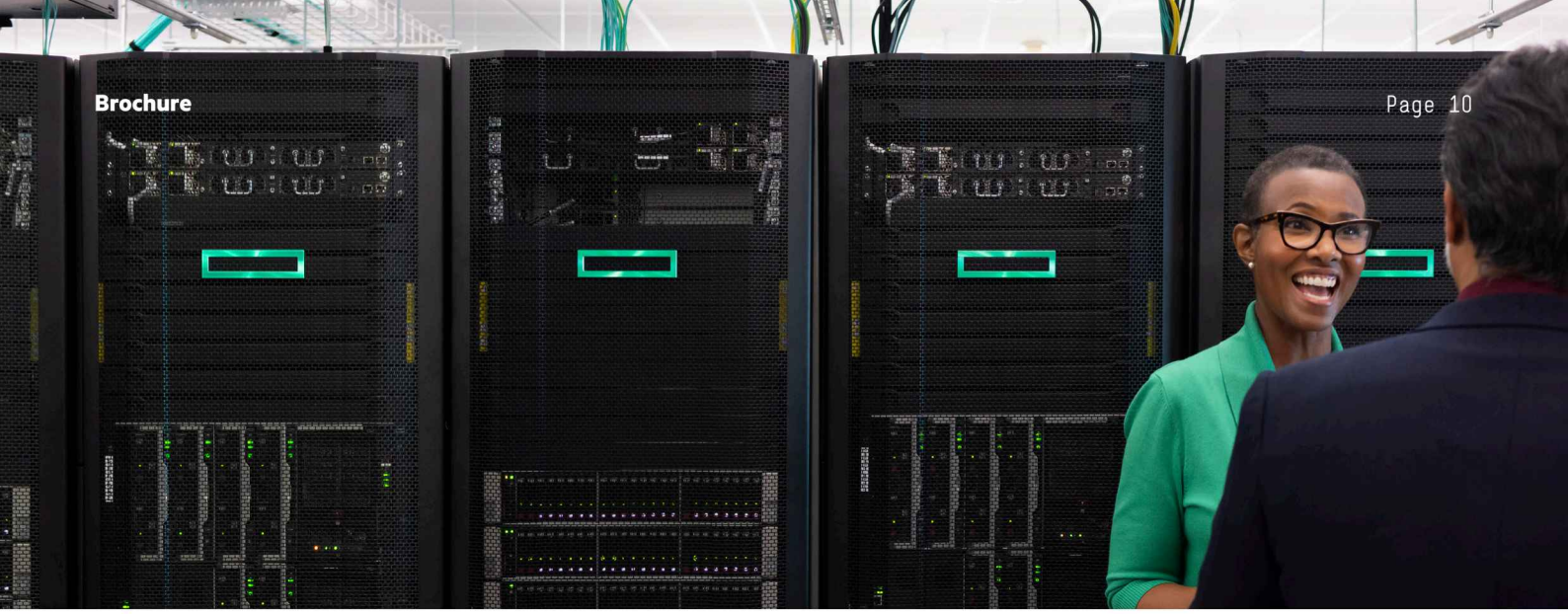


Figure 2. A new era of network operations



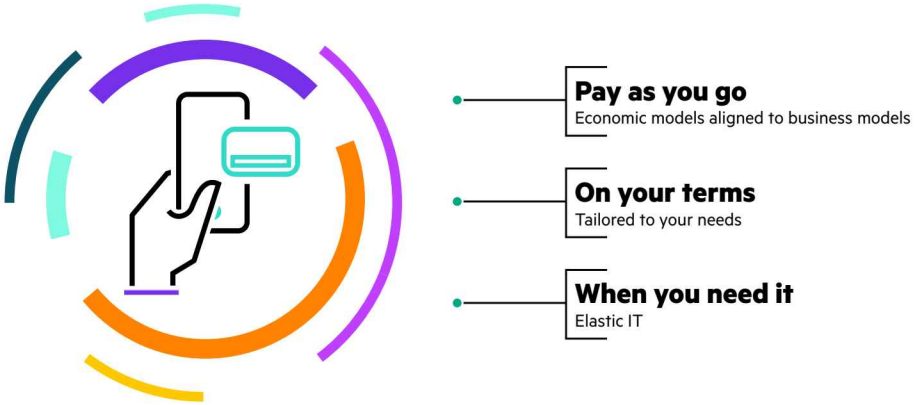
Placement of application workloads drives infrastructure decisions rather than infrastructure determining application location. Hybrid cloud, defined as a combination of public cloud, edge, colocation and on-premises locations has become the new reality for enterprise mission-critical workloads and a wide variety of on-demand, as-a-service offerings.

Colocation paired with an as-a-service platform is a great solution. Colocation can offer:

- Simplified single-tenant, cloud experience, while keeping control of applications and data
- Low latency, high bandwidth connection to other major cloud providers and network providers
- Improved transaction speed with direct connection to a broad ecosystem of adjacent businesses
- Support for your sustainability goals with elimination of over-provisioning and energy-efficient facilities
- Optimized IT spend with no upfront payments, pay only for what you use, and no egress charges

Combining these benefits with a pay-as-you-go approach for your infrastructure and operational needs, brings the best of both worlds—public cloud-like agility and collocated infrastructure—with a single contract, single invoice and single point of contact, and the opportunity to pivot staff from running a data center to working on other high-value activities.

Modern, fourth-generation architectures take full advantage of the deployment flexibility and consumption models afforded by as-a-service and colocation data center service offerings.



According to a Forrester study, customers who deployed HPE GreenLake have reported up to 80% faster time to market while deploying complex global IT projects.

Conclusion

As the shift from centralized to distributed modern edge-to-cloud data centers continues, new architectures are needed to provide secure connectivity that delivers an exceptional experience for users and applications. This next wave of data center connectivity requires higher performing fabrics, distributed services, and flexible consumption options.

The new era of data centers collapses infrastructure services into a fourth-generation fabric, eliminating the need to rely on discrete hardware appliances and software agents deployed in highly centralized, sub-optimal architectures.

Access to the simplicity and scale that was previously reserved for hyperscalers is now widely available, with a simplified consumption model, accelerated network services, and the choice of how and where to locate workloads.

The software and silicon are now in place to allow for all data center fabric services—stateless and stateful—to be delivered from a common platform. Finally, all customers have the choice and the scale they require for their business.

HPE Aruba Networking

In 2022, HPE Aruba Networking and AMD Pensando™ partnered to deliver the industry's first distributed services switch. The HPE Aruba Networking CX 10000 Series Switch represents a new category of data center switch that combines best-of-breed L2/3 Ethernet switching with embedded AMD Pensando DPU technology.

Now, data center operators can seamlessly insert stateful services into their networks in a distributed fashion—simplifying data center network design and increasing security posture.

Next-generation data center switching from HPE Aruba Networking allows enterprises to deliver great digital experiences to customers and employees with unprecedented scalability, performance, and operational efficiency.

HPE GreenLake is a portfolio of cloud and as-a-service solutions that helps simplify and accelerate your business. It delivers a cloud experience wherever your apps and data live – edge, data center, colocations, and public clouds. Available on a pay-as-you-go basis, HPE GreenLake runs on an open and more secure edge-to-cloud platform with the flexibility you need to open new opportunities.

How can you learn more?

IDC examines how to meet the need for high performance and security through a modern datacenter fabric

Learn more about HPE Aruba Networking data center modernization

Visit **ArubaNetworks.com**

Make the right purchase decision.
Contact our presales specialists.

Contact us

Hewlett Packard Enterprise

BR_SmartDataCenter_RVK_100523