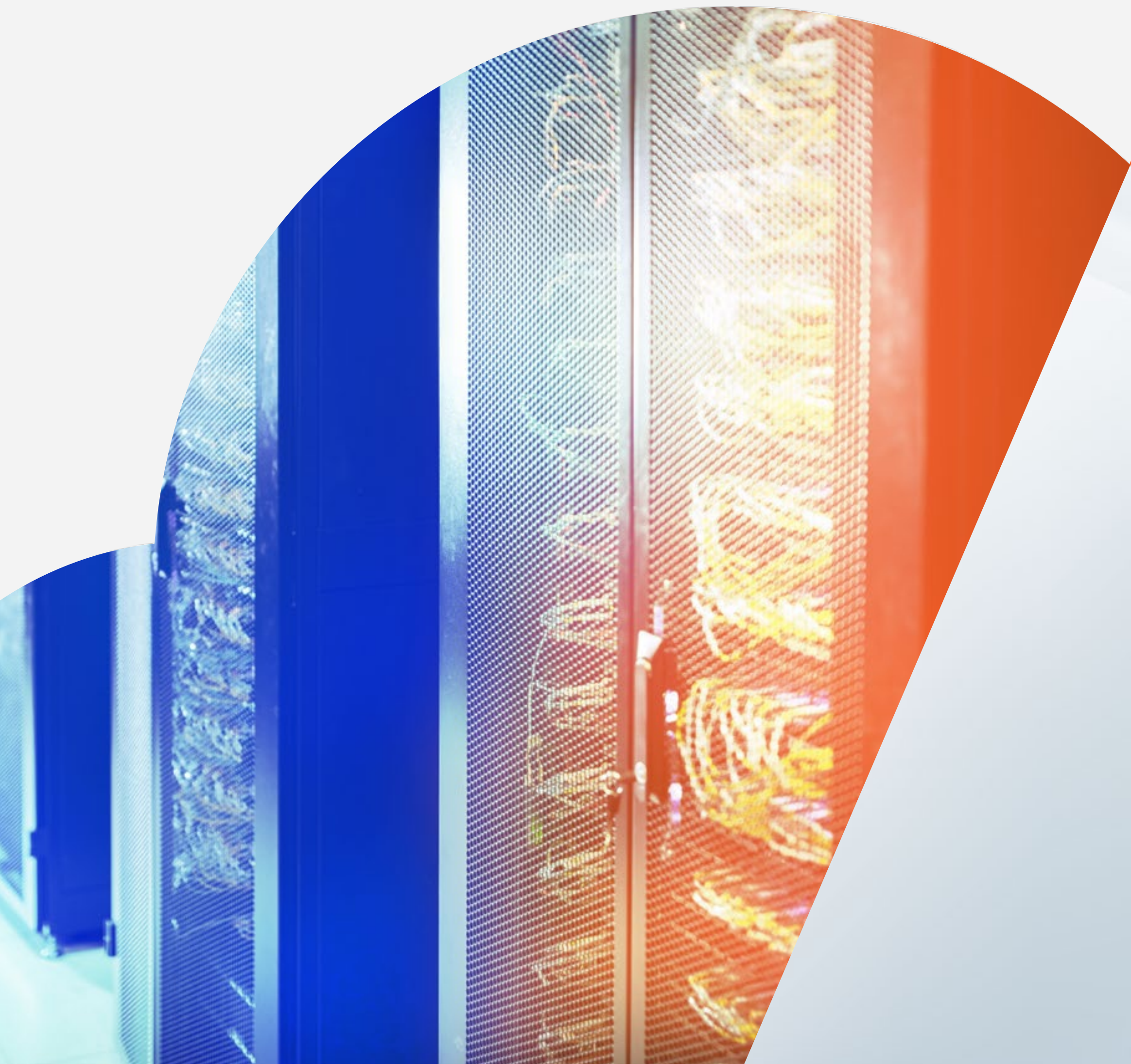
Which Network Architecture Is Right for You?

A Comparison of Chassis-based Router, Clos and DDC

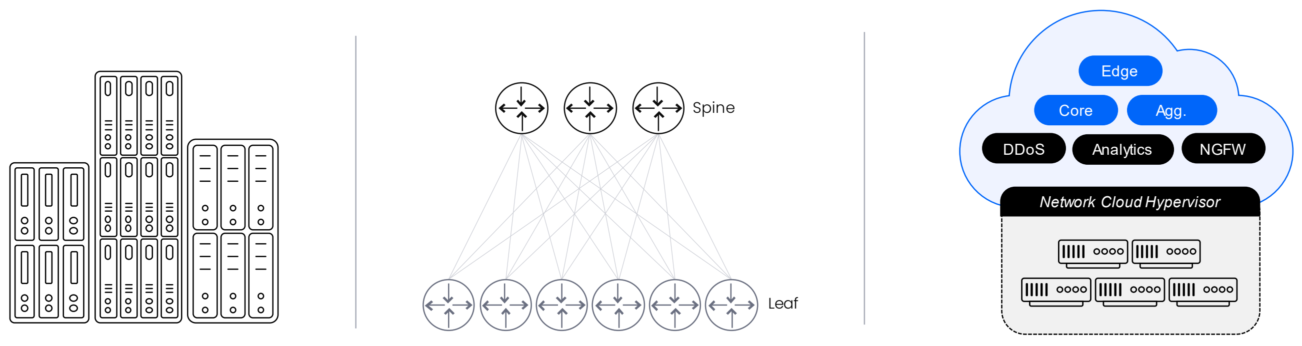
WHITE PAPER



Introduction

After more than 25 years of deploying chassis-based routing platforms, service providers (SPs) across the globe have recognized the challenges imposed by such platforms in terms of flexibility, scale and cost. Over the years, options for expanding the capacity of a chassis-based router beyond its physical slots (which happens quite often) ranged from deploying additional chassis, through upgrading to a larger underutilized chassis with more slots and line cards, to adopting a multi-chassis solution. None of these options is easy or ideal. Adding a router is expensive and adds routing and trunk complexity. Upgrading requires more power and cooling, not to mention forklifting. That cycle, which repeats every few years, is painful, to say the least.

Network Architecture



Chassis-Based Router | Clos (Spine & Leaf) | Distributed Disaggregated Chassis (DDC)

With the emergence of data center (DC) server applications and associated spine-and-leaf deployments (3-stage Clos) in large cloud providers, SPs also began looking at Clos architecture as an option for scaling out a large core or edge router. This was the case even though SP applications may not align 100% with the traffic patterns and applications in the cloud provider's DC. Notably, top-of-rack (ToR) leaf/spine switches and associated chips are less applicable to carrier-class routing needs of SPs. Moreover, traffic flow patterns are less east-west dominated at SP co-locations. Lastly, the consistent and deterministic quality of service (QoS) treatments plus convergence needs are quite different. All this without mentioning the burden of managing and automating tens, if not hundreds, of leaf devices by the SP instead of a single routing entity.

The point here is that SPs need to carefully weigh the different considerations before deciding on Clos for handling large routing fabric scale. It is a good practice not to simply assume that you need to use a Clos design because massive cloud scale providers are doing so. It is also important to ask yourself the question "Are you ready to all of a sudden automate and manage the life cycle of massive amounts of switches/mini-routers?"

Fortunately, a modern hybrid solution – Distributed Disaggregated Chassis (DDC) – is now available. This DDC routing system architecture was specified and published by the Open Compute Platform (OCP) and extended by the Telecom Infra Project (TIP) and its Disaggregated Distributed Backbone Router (DDBR). DDC solutions already are offered by a number of players (including DriveNets) and the supported vendor ecosystem is growing.

On the one hand, DDC architecture addresses the scaling issue encountered in chassis design to avoid the need for the multi-chassis approach and offers unlimited data plane and control plane scaling out. On the other hand, DDC's physical connectivity is very similar to Clos architecture, while maintaining the management simplicity and carrier-class router performance of a traditional chassis-based router. Even though a DDC router is also made of multiple pieces of spine-and-leaf components, it is logically enclosed via software to have a single control plane and management plane.

Another key difference between DDC and Clos architectures is that DDC does not need to run routing protocols among the hundreds of links between spine-and-leaf boxes, which drastically simplifies operations in terms of getting a core/edge router instantiated. Fewer protocols translate to easier troubleshooting and monitoring.

In the sections below, we will dive into the pros and cons of each approach to further illustrate the main points above.

Data Plane and Control Plane Scalability

Both DDC and Clos can provide data plane scale by adding more spine-and-leaf components plus the associated links between them. This horizontal scale-out capability stands out favorably compared to vertical scale-up in chassis-based solutions.

Regarding control plane scale, currently DDC scales out the control plane on a pair of powerful x86 servers. However, in today's Clos architecture, without major innovations and architectural changes, the control plane can only be scaled out per each leaf node as the control plane is fundamentally distributed in nature. No easy method exists for extracting or peeling off the control plane function out of each leaf-and-spine device and consolidating them into a single control plane virtual machine (VM) or container running on a pair of centralized x86 appliances. This is a big consideration for control plane scaling out.

Logical Instances and Grouping

DDC with multi-router architecture allows data plane resources to be separated and grouped into several individual service instances (SI) while providing total independence and isolation among the instances. Each SI has its own routing and management, and database containers and resources are allocated individually so one SI does not infringe on the resources dedicated to another SI, providing cloud-like multi-tenant resource sharing. This is a very important factor to consider since this ultimately reduces network building costs and fully leverages the massive forwarding capacity provided by modern white boxes at tens of terabits per 2RU device. Such a network design lowers infrastructure costs, accelerates innovation, and enables the running of multiple applications over a shared pool of resources. This leads to both CapEx and OpEx savings through improved resource utilization, as well as software automation simplicity.

In contrast, as far as we know today, Clos architecture has no logical grouping and aggregation of resources into isolated logical instances out of the whole pool of spine/leaf devices. This is the case unless one views each device as a logical system by itself (which is true). But there is no innovation beyond that – for instance there is no current way to group three spine devices and use them as single entity with unified management and single control/data planes.

Deterministic Forwarding and QoS Treatment

DDC architecture uses a chip's cell, switching between line cards and fabric cards. This is based on the design philosophy that DDC is targeted for carrier-class routers in large SP networks, as such deterministic forwarding behavior is absolutely required in par with field-proven chassis-based routers in production. This is done to achieve predictable latency and jitter among packets. This is by no means proprietary to DriveNets; moreover, since DriveNets is a software company, other fabric switching technology (IP based) can be adopted.

DriveNets is continuously investigating, monitoring, and evaluating the whole ecosystem for newer chipset suppliers and white box ODMs to make the DDC solution more diverse. Needless to say, there is no benefit to DriveNets in sticking with one supplier or another on the hardware front since DriveNets is an open software company focusing on SW innovation. Having said that, DriveNets takes responsibility for the total solution (HW and SW combined).

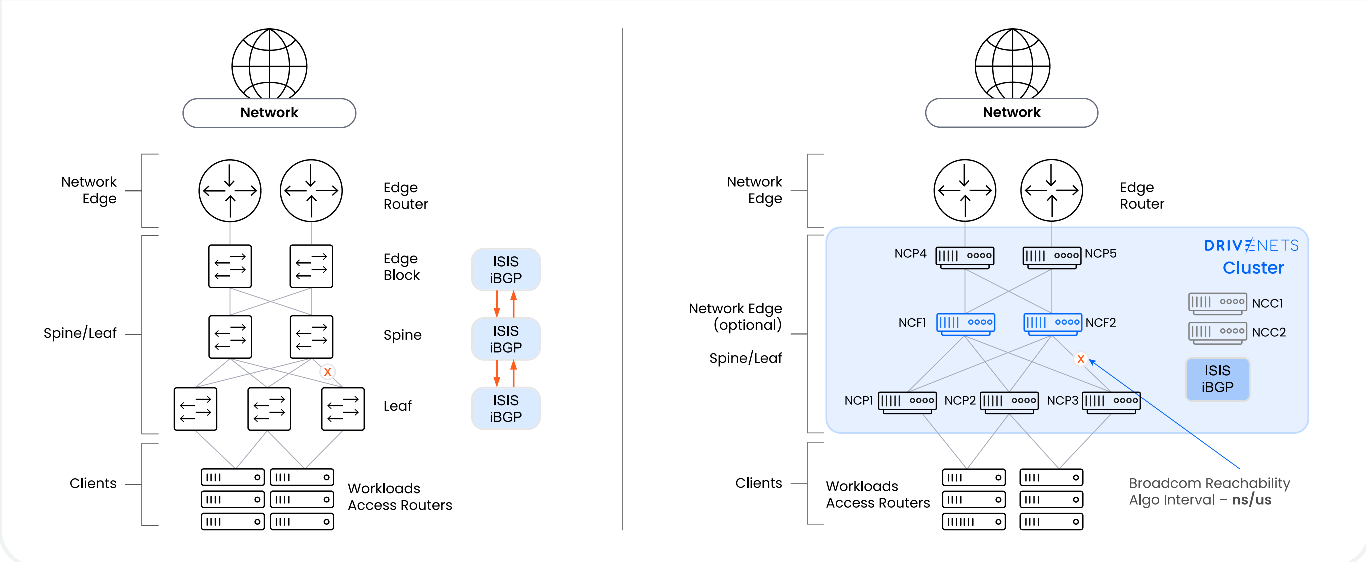
In the case of Clos, it is less clear whether it matches the carrier-class router behavior of chassis or DDC architectures. With Clos, packets travel normally through three stages (leaf to spine to leaf), and each stage can individually add buffer/latency.

Another area of concern is consistent end-to-end QoS treatment and behavior in Clos architecture as QoS treatment is highly ASIC-dependent. With multiple devices conducting ingress and egress QoS handling, there are always variations in virtual output queueing (VOQ) implementation and delay buffering (some are done at ingress, while other devices may perform just regular egress buffering). This poses a concern regarding predictable and consistent QoS behavior of Clos. With DDC, this is not a concern since multiple devices are treated as one single node and QoS policy enforcement is consistent across all white boxes.

Router Convergence in the Event of Failures

Regarding cluster component failure, including link problems, the DDC solution can quickly reconverge due to its single routing/control plane design and built-in high availability. With Clos, convergence depends on many protocols and many devices performing convergence at the same time. This is no longer router convergence, but mini-network convergence. Both underlay and overlay can add and prolong the convergence time due to the many devices and protocols involved, and the varied timing for each component to kick in. This is a serious performance metric that SPs should evaluate before adopting Clos architecture.

DDC improved Convergence



Manageability and Operational Impact

Both chassis-based and DDC-based architectures have a single management plane for operators to use for managing the router (irrespective of slots/capacity/scale). This familiar and easy management plane for access and control is an industry norm. In addition, DDC further improved the network operating system (NOS) to be more granular and container based. This provides flexibility in operational support and day-to-day management, allowing more granular hot patches on certain software-components rather than having to upgrade the whole NOS.

Compared to chassis and DDC-based, Clos-based management poses several challenges:

- Each leaf-and-spine device has its own management plane; with large Clos deployments, the device count can be daunting when it comes to log-in, management and troubleshooting.
- Network provisioning tasks become more complex as SPs do not have the same automation and scripting skillsets as cloud providers.
- Code upgrades and maintenance window counts skyrocket, placing a significant challenge on operations teams dealing with the many devices in the network.

Software Agility & Innovation Impact on OAM&P

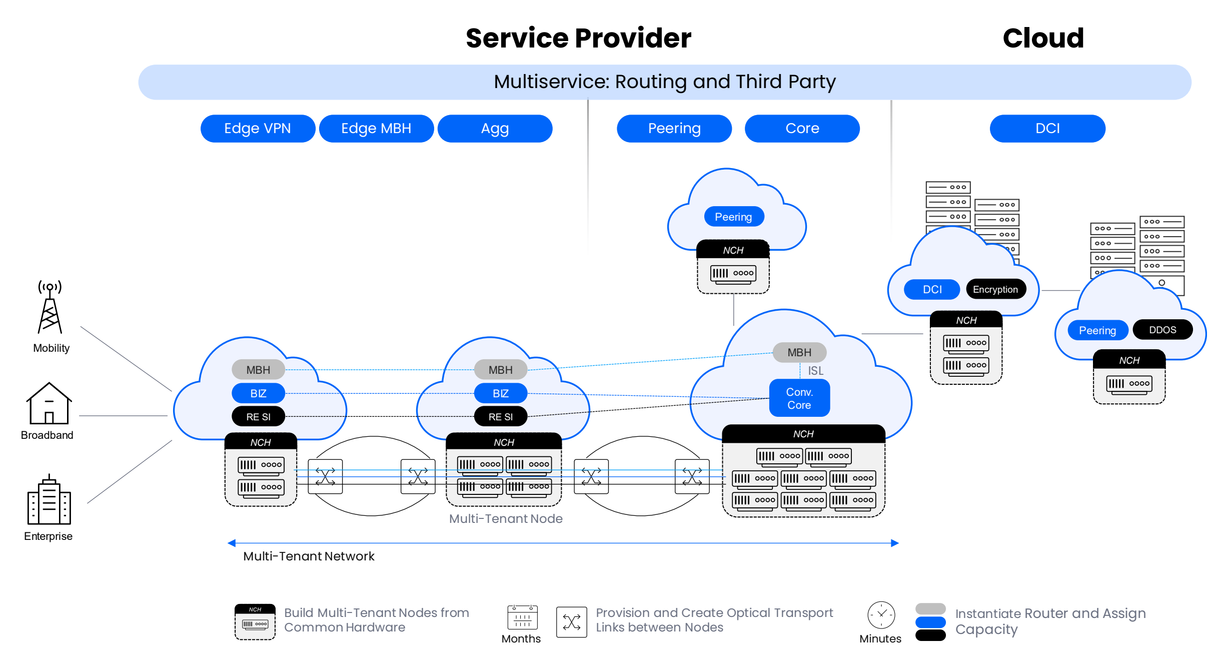
SPs should not overlook how DDC's software agility and innovation can impact the operations, administration, maintenance and provisioning (OAM&P) paradigm in the future.

DriveNets' DDC multi-router architecture supports SI's-port-level granularity, allowing another level of software-driven OAM&P not fully explored. Plenty of unforeseen opportunities exist for changing how the network is provisioned, turned up and tested, and for simplifying tasks among fiber/transport and IP multi-router devices.

Take, for example, the Clos routing instance (RI) application, which has the most traffic demand and growth. If the allocated data plane resources run out, the RI needs to be shut down and reconfigured. Additional line cards must be added into the RI resource pool and respun up the expanded RI. Even if the RI needs only 2 additional 100G or 400G ports, this can be a long and involved process; factors include whether a chassis has additional slots, whether spare line cards are available, and the levels of provisioning and IT work required for enabling the new change. With DriveNets DDC multi-router solution, this can be done by dynamically allowing the extra 100/400G ports (free from the forwarding pool) to be added into the service instance (SI) within hours if not minutes; all this can be carried out via DriveNets Network Orchestrator (DNOR), without truck roll or manual intervention.

Another example is cutting down the time required for end-to-end service readiness. Today, in order to roll out and pretest a single chassis-based edge service, fiber connectivity and transport gear testing and turn-up are required. Sometimes this process has to be conducted in a serial manner; namely transport cannot test circuits until the terminating (chassis-based edge) equipment has the right port and transceiver in place to check out end-to-end service readiness (light test). Unfortunately, the chassis line card and port cannot be nailed ahead of time, which prolongs the service turn-up procedure and involves multiple teams, calls and manual technician on-site verifications. With DriveNets' multi-router port granularity, the transport, access and fiber patch teams can preselect a white box port for use (transceiver and port enabled) and use it to map and test end-to-end readiness to the edge node. Later on, when the edge device port is ready to be assigned to an RI and a customer, the last step can all be done via SW without any further truck roll or manual work. This, in our estimate, can reduce service readiness time by at least 60%. It is not difficult to imagine that many other potential applications and use cases may be improved due to DDC network agility and innovation.

Multi-Tenant DDC Goals



Mix & Match Flexibility for Different Vendor Devices

Regarding vendor diversity, the flexibility to mix and match vendor boxes in today's DDC solution is relatively limited. Although OCP-compliant/certified white boxes from different ODMs can be mixed inside the same cluster, this often requires integration and code work. Mixing white boxes made with different chip technologies is one of the tasks set forth by the Ultra Ethernet Consortium (UEC), yet it is still not straightforward today. The OCP architecture does not prevent this, and it should become a reality as DDC further matures.
Clos architecture undoubtedly holds an edge in this capability since vendor A's spine plus vendor B's and C's leaves can participate inside a single Clos cluster. Of course, alongside that positive flexibility are the negative planning, sourcing, testing, and certification overheads associated with multiple ASIC technologies and the more boxes/vendors to manage.

Third Party Application and Beyond

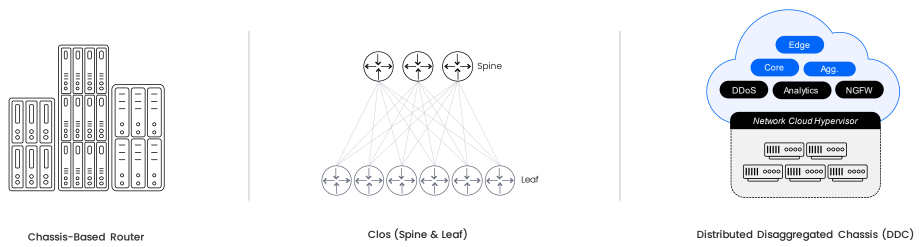
Today, x86-based network appliances beyond basic routing (such as DDoS, NGFW, CGNAT and DPI) are at odds with traditional routers. Some vendors design server-based line cards to address this issue, and Clos attaches a server appliance to an end leaf node. These solutions can bind two distinct application domains together but cannot truly converge the two into the same forwarding network infrastructure. DriveNets has conducted proof of concept (POC) trials and demos beyond DDC multi-router by integrating multiple services from third-party non-routing applications inside the same DDC cluster complex. The solution leverages the powerful data plane ASIC (aka NPU – network processing unit) to take ~75% of the heavy-duty lifting away from the x86 computing;

as a result, data path acceleration occurs at the white box NPU level. With further development of DDC multi-service a true and efficient convergence between the routing data path and the x86 compute data path is inevitable. This will make service chaining straightforward without packets "jumping hoops" back and forth with layers of orchestration tools in between.

Conclusion

Choosing the "right" next-generation architecture for multiservice routing applications always has been a compromise, driven by the dominance of various parameters and practical considerations. For service providers, where operational simplicity is a stronger consideration than hyperscale, there is a tendency to choose a chassis-based architecture. When it comes to data centers (DC), which are built to scale, Clos (spine and leaf) usually is chosen.

Network architecture comparison



	Chassis-Based Router	Clos (Spine & Leaf)	Distributed Disaggregated Chassis (DDC)
Capacity & Scale	Fixed and limited capacity Fix size granularity Forklift scaling	High-scale solution IGP routing scale No service providers features	Seamless scale-out No capacity limits Same white box across the entire network Service providers features
Management & Operation	Simple: The chassis is a single entity	Complex: Each box is managed separately	Simple: The entire cluster is a single entity
Cost of Ownership (TCO)	Proprietary hardware Multiple router families	Higher OpEx Underutilized hardware	Low-cost white boxes Common hardware across the network Single software package Optimal resource utilization

Service providers must evaluate carefully whether a traditional Clos solution meets their needs regarding carrier-class performance, unique traffic flow patterns with multiservice routers, and costs related to complicated and distributed control and data plane features.

Fortunately for SPs, there is another option available that combines the best of both chassis and Clos architectures – DDC.

DRIVENETS

DriveNets is a leader in cloud-native networking software and network disaggregation solutions. Founded in 2015 and based in Israel, DriveNets offers service providers and cloud providers a radical new way to build networks, substantially growing their profitability by changing their technological and economic models. DriveNets' solution – Network Cloud – adapts the architectural model of cloud to telco-grade networking. Network Cloud is a cloud-native software that runs over a shared physical infrastructure of standard white-boxes, radically simplifying the network's operations, offering telco-scale performance and elasticity at a much lower cost.

For more information, visit us at **www.drivenets.com**