Router Virtualization in Service Providers

Executive Summary

Several factors are contributing to a growing need for virtualization in the point of presence (POP) and data center: a desire to reduce capital costs by buying fewer chassis, a desire to reduce operational costs by deploying fewer chassis and simplifying topologies, and a strong push to decrease the environmental impact by using less power. These goals must coincide with existing requirements of stability, resiliency, and service isolation. This paper discusses how the needs of POP deployments compare and contrast with those of the data center and which virtualization architectures are available to address those needs. It then covers the architectural benefits of virtualization in Cisco® IOS® XR Software that contribute to these goals.

Deployment Considerations

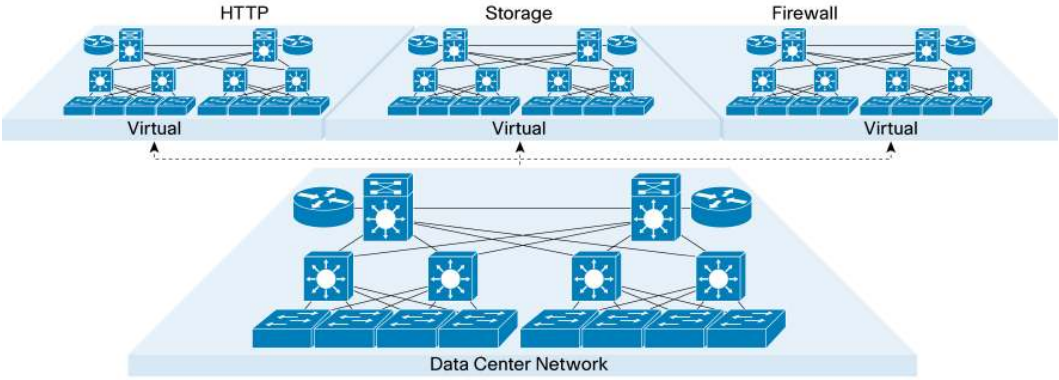
Historically, virtualization occurs much more frequently in the data center than the POP because of cost savings and overlapping administrative domains. However, environmental concerns and recent architectural advances have expanded the possible range of virtualization opportunities outside the data center. High-end routers require substantial power and cooling just for the basic chassis components, so any virtualization can provide a significant power savings by simply adding hardware components to existing equipment. The following sections will analyze two key deployment scenarios – POP and data center – and evaluate the requirements of each.

Data Center

Virtualization in the data center traditionally refers to applications and application servers, such as HTTP and SAP. Because rack space is such a precious resource, adding extra servers and routing instances without an increase in rack space is highly desirable. Space and power are more expensive because typically a third party, not the company deploying the equipment, owns the facilities. Figure 1 provides a typical virtualization topology in a data center environment.

© 2008 Cisco Systems, Inc. All rights reserved. This document is Cisco Public Information.

Figure 1. Data Center Virtualization Technology



Multiple groups within the same organization (servers, firewall, storage, network architecture) are often accessing the same network device and thus employ virtualization to prevent infighting about who “owns” it. These routers typically will not carry full Internet routes, but just a handful of local routes for other destinations in the data center and a default route toward the network core for everything else. Furthermore, there are few routing adjacencies because most of the connections to the data center router are hosts. The flows tend to be lower speed (less than 1 Gbps) because of bandwidth constraints on the servers. Features and services, if configured, are consistent for all virtualized routing instances; it is unlikely that one instance will be running a particular feature set and another instance will run an entirely different feature set.

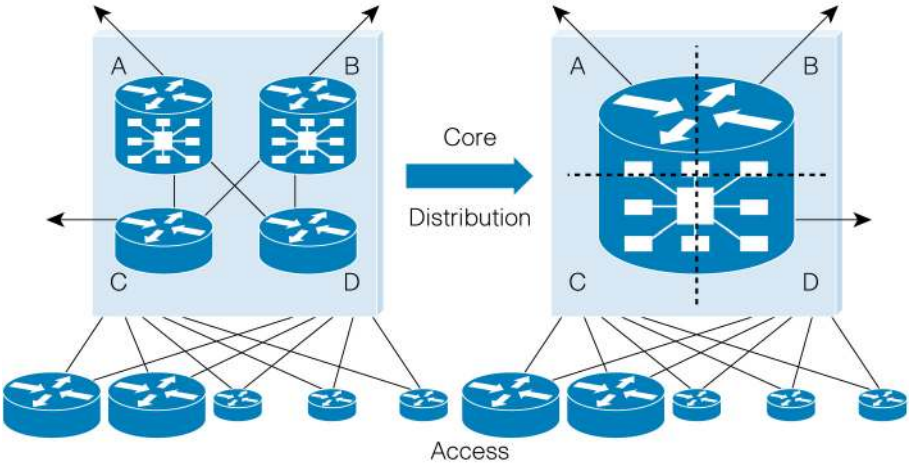
Thus, we can classify the data center environment as relatively nonintensive for a router; the router is effectively acting as a conduit to connect the data center services.

POP

Virtualization at the POP level typically results from one of two main goals:

1. Consolidation of multiple layers in the routing topology, as demonstrated by Figure 2

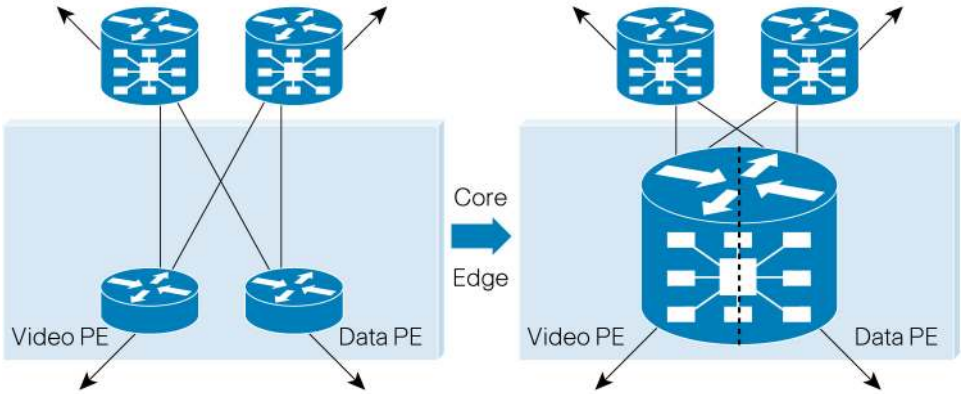
Figure 2. POP Layer Consolidation



2. Consolidation of multiple networks (such as data and video) within the same provider network, as demonstrated by Figure 3

© 2008 Cisco Systems, Inc. All rights reserved. This document is Cisco Public Information.

Figure 3. POP Network Consolidation



In either case, the feature set can be very different between the two virtual routing instances; a provider and a provider edge (PE) device have very different roles in a network, and a router in a data network is likely to have different service requirements than one in a video network. Many POPs are located on premises owned by the deploying provider, thus they own rack space and power instead of leasing it. Because the POP is the aggregation point of all customers in a particular geographical region, the number of routing adjacencies is high, and full Internet routes are expected to be exchanged with most routing peers. Additionally, these devices aggregate traffic from lower layers of the network, so the bandwidth demands are extremely high (greater than 10 Gbps).

Thus, we can classify the POP environment as quite resource-intensive both from a control plane and data plane perspective.

Tables 1 and 2 summarize deployment considerations for the data center and POP, respectively.

Data Center

Table 1. Data Center Deployment Considerations

Category	Deployment Consideration	Characteristic	Notes
Cost	Power and cooling	Expensive	Typically leased space, thus higher cost
	Rack space	Expensive	
Scale	Prefix scalability	Very low	Default route to core
	Routing adjacencies	Few	Connected to many hosts, not other routers
	Bandwidth	Low	Constrained by server bandwidth
Manageability	Feature sets between virtualized instances	Common	Minimal feature disparity
	Administrative domains	Multiple	Multiple groups share router

© 2008 Cisco Systems, Inc. All rights reserved. This document is Cisco Public Information.

POP

Table 2. POP Deployment Considerations

Category	Deployment Consideration	Characteristic	Notes
Cost	Power and cooling	Less expensive	Premises typically owned by provider
	Rack space	Less expensive	
Scale	Prefix scalability	High	Full Internet routes, possibly full VPN routes
	Routing adjacencies	Many	Connected to many routers
	Bandwidth	High	Aggregates traffic from lower layers of network
Manageability	Feature sets between virtualized instances	Different	Provider or PE or different types of networks
	Administrative domains	Single/few	Provider and PE same owner, possibly different groups with multiple PEs

Virtualization Concepts and Considerations

Though varying degrees of routing virtualization exist, this paper will focus on the two main techniques for creating virtualized router entities as defined by their physical and operational characteristics. A Hardware-Isolated Virtual Router (HVR) has hardware-based resource isolation between routing entities, whereas a Software-Isolated Virtual Router (SVR) comprises software-based resource isolation between routing entities.

Within SVRs, there are several models for achieving virtualization. One model allows for multiple guest operating systems to overlay on a host operating system. This approach tends to have a detrimental impact on scale because it introduces significant contention of resources. Rather than attempting to manage this contention with algorithms and enhancements to divide resources, a common approach overprovisions resources on all SVRs so that no individual SVR is likely to affect the others. Unfortunately, this approach wastes resources while decreasing overall scale.

Another model integrates the virtualization into the kernel itself, a design decision in contrast to the overlay model. Though this technology area is relatively nascent and will surely mature, it still suffers the same fundamental limitations of resource contention. Furthermore, though kernel virtualization could improve the processing performance relative to the overlay model, it introduces extra complexity and instability into the kernel.

A third model does not incorporate multiple operating systems, instead providing virtualization in the individual applications. Though such a model provides better scale from lower overhead, it also complicates the design, testing, and especially management of the SVRs. In addition, it requires the applications to understand some level of virtualization, which requires substantial work on each component individually. This development effort can take a long time before achieving consistency. Thus, there could be incongruence between features supported on standalone systems as opposed to SVRs.

Regardless of the software virtualization model, a common characteristic of SVRs is sharing hardware resources in the data plane. While rudimentary arbitration features like memory partitioning may be employable, high levels of virtualization in the hardware are expensive and thus rarely integrated. As a result, vigilant resource monitoring must extend beyond the control plane.

© 2008 Cisco Systems, Inc. All rights reserved. This document is Cisco Public Information.

In contrast, the HVR approach dedicates both control plane and data plane resources on a per-module boundary to individual virtual entities, so there is no sharing of either control plane or data plane resources. It is sometimes said that the only resource HVRs share is sheet metal. A lightweight shim layer provides low-level communication between HVRs, who otherwise believe they are independent router entities. Because of dedicated control plane and data plane resources, software applications and forwarding hardware need not implement virtualization. This separation effectively eliminates arbitration for resources between virtual routing entities.

Table 3 summarizes the defining characteristics of HVRs and SVRs.

Table 3. Comparison of Virtualized Routing Architectures

Category	Hardware-Isolated Virtual Router	Software-Isolated Virtual Router
Control plane resources (CPU, memory)	Dedicated	Shared
Data plane resources (forwarding engine, queues)	Dedicated	Shared
Chassis resources (power supplies, blowers, fabric)	Shared	Shared
Management, configuration	Dedicated	Typically shared, but varies depending on degree of virtualization
Connections between virtualized routing entities	Typically external	Typically internal, but possibly external
Per-chassis scalability (routing adjacencies, prefixes)	Increased with additional logical routers	Unaffected by additional virtual routers

Discussion

The two approaches have different impacts on considerations that are important to providers. These considerations are key to selecting the proper technology for the virtualization application.

Resiliency and Security

A key requirement for deployment is that fundamental aspects of router operation should get no worse as a result of implementing virtualization, and resiliency is among the most important. An outage in one virtual routing entity, whether planned or unplanned, should have no impact on the others. With HVRs, this requirement proves relatively easy to meet because entire physical modules are associated with one and only one virtual routing entity. In contrast, SVR resiliency depends on the virtualization method (overlay, kernel, or application), as well as the quality of the implementation. In most scenarios, however, the lack of physical isolation inherently means that extra care needs to be taken with SVRs to preserve resiliency. Malicious attacks become amplified with the number of SVRs and DoS protection becomes vital. Software defects require additional scrutiny, because multiple SVRs could be affected simultaneously instead of just a single router.

Management

The hardware isolation of HVRs makes most aspects of planning straightforward. Because each HVR has dedicated control plane and data plane resources, careful planning and resource management are not necessary. SVRs, on the other hand, require meticulous planning and policy construction to ensure that one SVR does not negatively impact others in the chassis. Constant competition for resources, both in the control plane and the data plane, typically requires constant monitoring and retooling for optimal performance. Because SVRs share so many resources, software isolation can be difficult. Thus, software upgrades and downgrades affect all SVRs and do not allow for individual control. In contrast, the separation of resources between HVRs allows HVRs to be upgraded or downgraded individually without affecting the operation of the others.

© 2008 Cisco Systems, Inc. All rights reserved. This document is Cisco Public Information.

Other aspects of management become problematic with SVRs because ownership of the physical nodes is unclear, leading to inconsistencies in configuration maintenance and resource monitoring. A lack of total visibility could prevent the operator from noticing critical trends in overall router functionality (such as CPU utilization or memory usage), with negative impact on resiliency.

One additional issue with management is the connections between virtual routing entities. HVRs typically require external connections, which can increase the total capital expenditure of the solution, but also provide a clean inter-HVR architecture that works exactly the same as a connection to any other router. In contrast, inter-SVR connections typically exist as a kind of a virtual interface. However, these virtual connections must be managed by the software as a special case and do not follow the same forwarding path for normal traffic, instead using one of two approaches:

1. Modifying the existing forwarding path, such that a route lookup may point to a forwarding table for another SVR. This adds performance impact to the forwarding and can have a drastic impact on the scale from the hardware holding prefixes for all SVRs.
2. Placing the virtual interface on dedicated hardware and offloading the extra processing through packet diversion. This option has additional challenges, though, as SVRs must share the internal bandwidth to the virtual interface hardware. Without a very robust fabric quality-of-service (QoS) implementation, it is difficult to ensure that inter-SVR traffic will be handled properly. It must be able not only to prioritize within flows of the same SVR, but also arbitrate between flows from different SVRs, which again adds more complexity and overhead into the forwarding path.

Using dedicated hardware for virtual interfaces consumes space on the router – typically an entry slot that could use a physical interface. This dedicated slot provides some additional forwarding capacity, but often adds nothing to the control plane processing power. HVRs require use of an additional route module to provide control plane and router ownership services, but that extra hardware also adds to the control plane processing power and allows for more routing adjacencies. An external connection is required for HVR, but this paradigm allows for true isolation of HVR resources and precludes the use of complex changes to the forwarding path to allow for inter-HVR forwarding.

Scale

Perhaps the most important distinction between HVRs and SVRs is the direction of scale: does the virtualization technology scale upward or downward? In the case of HVRs, each additional routing entity adds to the number of prefixes, routing adjacencies, and logical interfaces that the chassis can support. In the case of SVRs, each routing entity leaves the number of these scaling elements unchanged or slightly lower from the extra overhead. In essence, an SVR implementation divides the available resources whereas an HVR implementation multiplies the available resources. This important distinction is key to understanding optimal virtualization in various network topologies, as it affects deployment models and solution development.

Deployment Evaluation

Now that we've established the characteristics of data centers and POPs, as well as the qualities of SVRs and HVRs, let's evaluate the compatibility of these virtualization technologies in the two environments.

© 2008 Cisco Systems, Inc. All rights reserved. This document is Cisco Public Information.

Data Center

Table 4. Data Center Evaluation

Category	Deployment Consideration	Characteristic	Optimal Virtualization Approach (SVR/HVR)
Cost	Power and cooling	Expensive	SVR
	Rack space	Expensive	
Scale	Prefix scalability	Very low	SVR
	Routing adjacencies	Few	SVR
	Bandwidth	Low	SVR
Manageability	Feature sets between virtualized instances	Common	Either
	Administrative domains	Multiple	Either

In Table 4, it is clear that SVRs are a better match for the data center because of the low scale and price premium for rack space.

POP

Table 5. POP Evaluation

Category	Deployment Consideration	Characteristic	Optimal Virtualization Approach (SVR/HVR)
Cost	Power and cooling	Less expensive	Either
	Rack space	Less expensive	
Scale	Prefix scalability	High	HVR
	Routing adjacencies	Many	HVR
	Bandwidth	High	HVR
Manageability	Feature sets between virtualized instances	Different	HVR
	Administrative domains	Single/few	HVR

In Table 5, it is clear that HVRs are much better suited for POP applications. The amount of scale required is prohibitive for SVRs, as simply running full Internet routes on multiple SVRs could significantly drain the available resources. The multiplicative nature of HVR scalability ensures sufficient resources and provides effectively infinite scale. More importantly, the requirement for a different feature set and the ability to independently manage software on the routing instances make SVRs inherently incompatible.

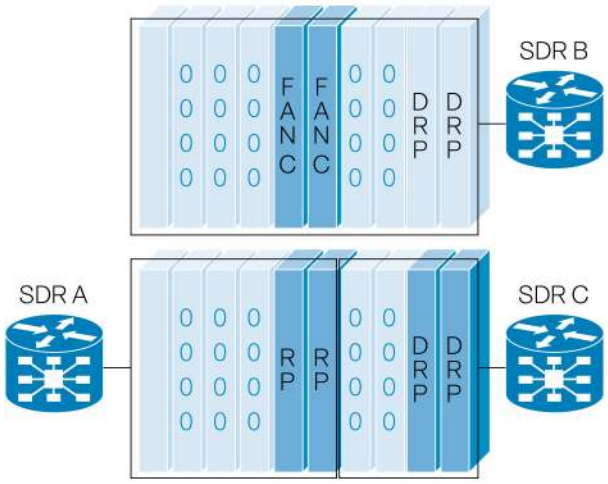
Note that in data center scenarios where scale and bandwidth requirements increase, such as a hybrid data center router that aggregates greater amounts of routing or traffic, HVRs may also be a better option. In such a case, a routing platform would benefit from offering both SVR and HVR services.

Virtualization in Cisco IOS XR Software: Secure Domain Routers

To accommodate the high bandwidth and control plane needs in provider networks, especially POPs, Cisco IOS XR Software includes support for an HVR technology known as Secure Domain Routers (SDRs). SDRs provide full isolation between virtualized routing instances through the use of Distributed Route Processors (DRPs) for extra control plane resources. SDRs are defined on per-slot boundaries, with entire Route Processors (RPs) and Modular Services Cards (MSCs) dedicated to an SDR. Figure 4 depicts the deployment of SDRs on a Cisco CRS-1 Carrier Routing System running Cisco IOS XR Software.

© 2008 Cisco Systems, Inc. All rights reserved. This document is Cisco Public Information.

Figure 4. SDR Example Deployment on Cisco CRS-1/16



In Figure 4, the RPs provide ownership of SDR A, whereas DRPs provide ownership of SDRs B and C. Cisco IOS XR Software allows for partitioning the router in this way to maximize chassis scalability both in terms of control plane and forwarding plane. Each SDR maintains a separate configuration and manages its own interfaces independently of the other SDRs in the chassis. They also maintain separate routing, forwarding, and adjacency tables. In fact, the only parts of the chassis that are shared are the fabric, the fans, and the power supplies, objects that require a marginal amount of environmental monitoring provided by one of the RPs.

Table 6 revisits our previous chart, comparing virtualization approaches with SDRs.

Table 6. Comparison of Virtualization Technologies with Cisco IOS XR Software-Supported Secure Domain Routers

Category	Hardware-Isolated Virtual Router	Software-Isolated Virtual Router	Secure Domain Routers
Control plane resources (CPU, memory)	Dedicated	Shared	Dedicated
Data plane resources (forwarding engine, queues)	Dedicated	Shared	Dedicated
Chassis resources (power supplies, blowers, fabric)	Shared	Shared	Shared
Management, configuration	Dedicated	Typically shared, but varies depending on degree of virtualization	Dedicated
Connections between virtualized routing entities	Typically external	Typically internal, but possibly external	External
Per-chassis scalability (routing adjacencies, prefixes)	Increased with additional logical routers	Unaffected by additional virtual routers	Scales linearly with number of SDRs

Summary

Both SVRs and HVRs can be part of a provider's IP Next-Generation Network (NGN) deployment strategy, but care must be taken to ensure the proper application for the particular technologies. If not, a defect, operator error, or malicious attack could have drastic implications on the resiliency of the network. In low-scale and less intensive environments, Software-Isolated Virtual Routers provide the ideal architecture for virtualization. However, control-plane-intensive and high-bandwidth environments require the extra scale offered by Hardware-Isolated Virtual Routers for highly resilient and predictable performance and service. Cisco IOS XR Software's Secure Domain

© 2008 Cisco Systems, Inc. All rights reserved. This document is Cisco Public Information.

Routers provide this HVR functionality and are ideally suited to help providers scale upward as bandwidth and processing needs increase.

For More Information

For more information about Cisco IOS XR Software Secure Domain Router (SDR), visit http://www.cisco.com/en/US/docs/ios_xr_sw/iosxr_r3.7/system_management/configuration/guide/yc37sdr.html or contact your local account representative.

For more information about virtualization in service provider POPs, visit: http://www.cisco.com/en/US/prod/collateral/routers/ps5763/prod_white_paper0900aecd8036355e.pdf

For more information about the Virtual Devices Contexts (VDCs) for data centers, visit: http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9402/ps9512/White_Paper_Tech_Overview_Virtual_Device_Contexts.html

For more information about the virtualization architecture using the Cisco Nexus 7000 Series, go to: http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9402/ps9512/brochure_cisco_nexus_7000_series_virtualization_arch.pdf

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV
Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

CCDE, CCENT, Cisco Eos, Cisco Lumin, Cisco Nexus, Cisco StadiumVision, Cisco TelePresence, Cisco WebEx, the Cisco logo, DCE, and Welcome to the Human Network are trademarks; Changing the Way We Work, Live, Play, and Learn and Cisco Store are service marks; and Access Registrar, Aironet, AsyncOS, Bringing the Meeting To You, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, CCVP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Collaboration Without Limitation, EtherFast, EtherSwitch, Event Center, Fast Step, Follow Me Browsing, FormShare, GigaDrive, HomeLink, Internet Quotient, IOS, iPhone, iQuick Study, IronPort, the IronPort logo, LightStream, Linksys, MediaTone, MeetingPlace, MeetingPlace Chime Sound, MGX, Networkers, Networking Academy, Network Registrar, PCNow, PIX, PowerPanels, ProConnect, ScriptShare, SenderBase, SMARTnet, Spectrum Expert, StackWise, The Fastest Way to Increase Your Internet Quotient, TransPath, WebEx, and the WebEx logo are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0809R)

Printed in USA C11-512753-00 12/08

© 2008 Cisco Systems, Inc. All rights reserved. This document is Cisco Public Information.