



Distributed Virtual Routing

Distributed Virtual Routing (DVR) is a powerful capability available with Extreme Fabric Connect. It enables a single tier routing architecture for distributed Fabric-based networks. It distributes the routing function (virtual default gateway) so that it is always available at first network hop. As a result, any wired, wireless or virtual host can physically move to any point in the network fabric while benefiting from shortcut switching and routing.

With this technology, tromboning is avoided, performance is enhanced and traffic flows are optimized. It is applicable for Data Center deployments, Data Center Interconnect deployments and Campus networks, especially wireless deployments to optimize client roaming.

The Networking Challenge With Roaming Devices

When devices roam around in a network infrastructure, whether those devices are virtual machines or whether they are wireless mobile clients in a campus infrastructure, they face the same issue. In order to keep continuous network connectivity, the IP address cannot change irrespective of the physical location and network attachment.

In network environments this is achieved by stretching the IP subnet to the location where the device resides. The result is that in a campus or data center the IP subnets could be extended quite far across the physical topology.

Fabric technologies allow stretching of L2 domains (and thus IP subnets) quite elegantly without the need of building Spanning Tree domains.

The challenge that remains is the exit point out of the IP subnet i.e., default router. If a subnet only has one (or two) default routers, typically the router is most of the time at the wrong location and thus traffic will “trombone” across the network, latency is increased, bandwidth is wasted and debugging connectivity failures becomes much harder.

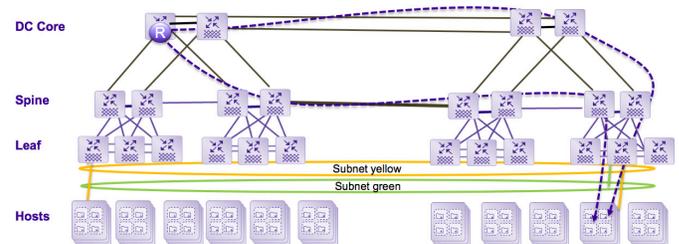


Figure 1: Traffic Tromboning

The way to address this problem is to distribute the routing (Default Gateway Function) to the network access points, where the roaming devices connect to.

An additional challenge is to ensure that the connectivity service (IP subnet) is always dynamically extended to the roaming device. Extreme Fabric Connect addresses this elegantly; however this topic is not part of this document.

Introduction to Extreme Fabric Connect Based Network Virtualization

Extreme Fabric Connect, based on the IEEE 802.1aq Shortest Path Bridging (SPB) standard, leverages a service instance concept (ISID), which is part of the Ethernet MacInMac (IEEE 802.1ah) data plane. These ISIDs virtualize any type of connectivity service. Those connectivity services can be point-to-point or any-to-any L2 LAN extensions. However at the same point in time these L2 extensions can support virtualized L3 routing and L3 virtualization (VRF) capabilities, enabling flexible bridged and routed tenant (user or application) virtualization.

One key element of an SPB based Ethernet fabric is the separation of network infrastructure and network services. End-to-end Network services are only provisioned at the edge of the SPB based Ethernet network infrastructure. The fabric automatically connects the services between the service access points; the fabric itself is zero-touch.

Introducing the DVR Concept

Stretching IP subnets across multiple locations, racks in a data center or floors or buildings in a wireless deployment, enables hosts (virtual machines or wireless users) to move around freely without having to change their IP addresses. Typically however, when IP subnets are stretched, the problem of inefficient routing (traffic tromboning) occurs, because the default gateway router in most cases is not local to the roaming hosts. Distributed Virtual Routing (DVR) addresses this issue by distributing a Virtual IP (VIP) routing instance to all access switches that have a presence in the IP subnet/VLAN. In the Data Center use case, virtual machines use their first Top of Rack switch as their default gateway. A DVR-enabled TOR switch can route locally without having to forward traffic first to a routing point somewhere deep in the network. This means that all L2 and L3 traffic is efficiently "east-west" short cut switched between any host, irrespective of its location, avoiding inefficient double paths and traffic tromboning.

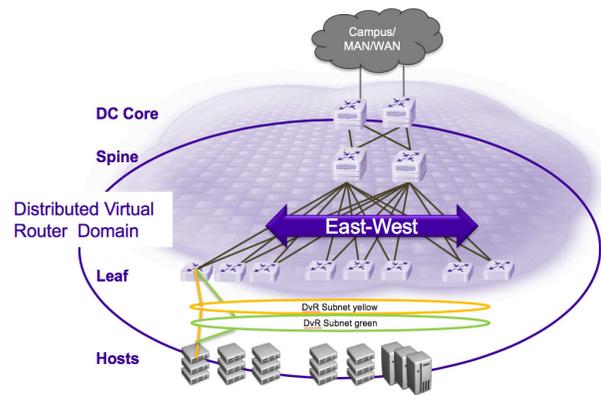


Figure 2: East-West Short Cuts

Similarly, for campus-based tunnel-less, AP switched wireless deployments, users can easily roam between buildings using DvR. Each building provides default gateway routing capabilities for the users at the point of entry to the network. With short-cut switching and routing enabled, all the L2 and L3 traffic from mobile users is efficiently short-cut switched north south to the campus core and Data Center, avoiding tromboning and improving performance.

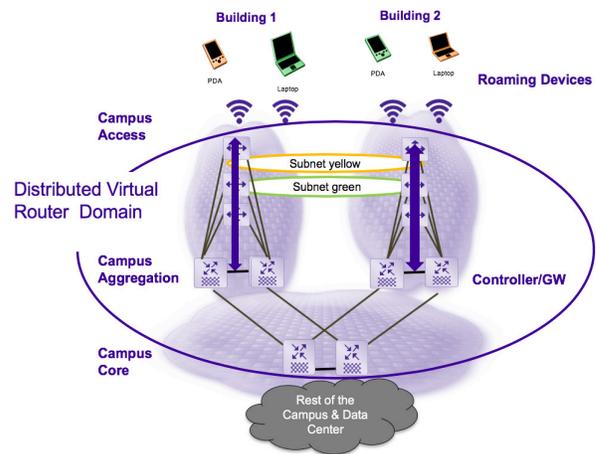


Figure 3: VIP Distribution

Simplifying Data Center Deployments

To ensure scale and efficiency, DVR leverages domains within the Fabric Connect network. Typically, a DVR domain represents a single data center or a single building in a campus. A DVR domain consists of redundant controllers and leaf nodes. The controller function, typically deployed on a DVR compatible VSP switch, distributes all default gateway VIP configuration to DVR leaf nodes. It also pushes all the L3 unicast, multicast and VRF configuration to the leaf nodes - requiring only the fabric infrastructure configuration as well as the L2 service configuration to be implemented on the leaf nodes.

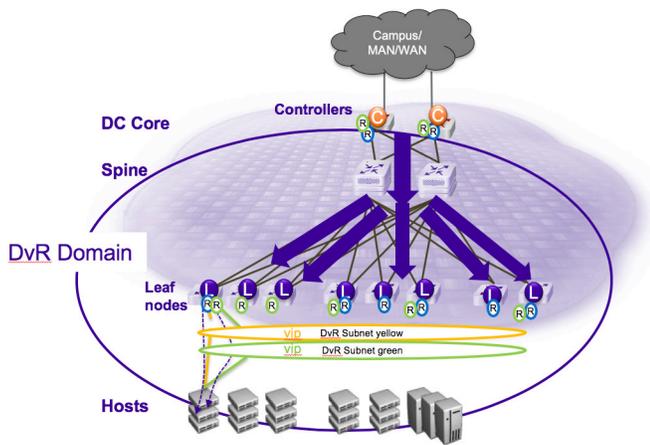


Figure 4: DvR Data Center Configuration

Short-Cut Routing

DvR not only simplifies VIP provisioning, but also introduces an optimized short-cut routing concept that ensures that all host-to-host traffic is shortest path switched and routed. This optimized routing is achieved by maintaining a shared and distributed host route information table in the DvR domain, such that every host location is known to all DvR members and changes are updated among the members dynamically.

This results in a well-balanced and optimized traffic flow in the data center without any wasted bandwidth. Traffic latency is brought to a minimum because of the most efficient traffic path is always chosen. Link based IS-IS metrics also allow engineering of the forwarding paths to some degree.

Multi-Pathing

DvR leverages SPB Equal Cost Trees, IP Equal Cost Multi-Pathing and Link Aggregation and thus spreads traffic over multiple links or paths if the metrics and the switch capabilities allow for it.

Data Center Interconnect

For stretched Data Center deployments or Data Center migrations, DvR supports the multi-domain concept. DvR controllers, that share the same fabric connect network, are automatically linked together through a DvR Backbone domain. This concept, if required, allows stretching IP subnets between multiple DvR domains enabling seamless roaming while still keeping the individual Data Center DvR domains autonomous.

As a result virtual machines can be freely moved between the DvR domains while still maintaining the same IP address and default gateway.

The use of DvR Domains is flexible allowing the use of a single domain across many smaller Data Centers or the more common use of one DvR domain per Data Center, as well as the ability to have several DvR domains within each Data Center in order to scale to even the largest possible Data Center. The solution supports up to 16 individual DvR domains per DvR Backbone.

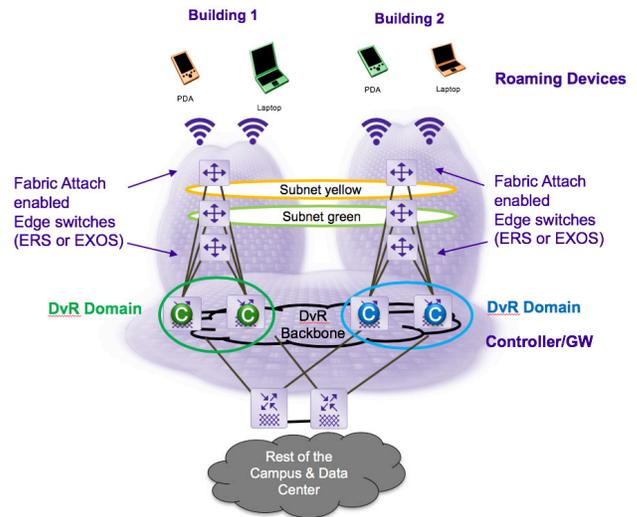


Figure 5: DvR Deployment Data Center Interconnect

Simplifying Campus-Based Wireless Deployments

In campus-based wireless deployments, because DvR is compatible with Extreme Fabric Attach, DvR controllers typically communicate not with leaf nodes, but interoperate with Fabric Attach enabled switches. Fabric Attach enables non-Fabric Connect enabled devices to participate in the Fabric, providing dynamic auto-attach capabilities of users and devices into Fabric-based services at the entry point to the network. With Fabric Attach support on EXOS based switches, DvR based campus networks can be designed with a mix of ERS and EXOS based switches at the edge of the network.

Like in the Data Center interconnect scenario, campus implementations would be multi-domain - with each separate building, encompassing redundant controllers and numerous Fabric Attached closet switches, being its own DvR domain.

DVR controllers in each building, that share the same Fabric Connect network, are automatically linked together through a DVR Backbone domain. This allows IP subnets to be stretched between multiple DVR domains. The result is clients can move freely within and between buildings while still maintaining the same IP address and default gateway.

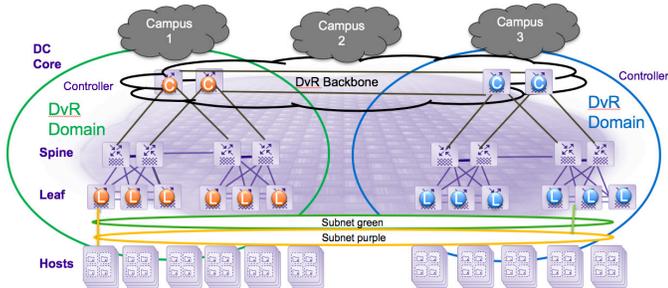


Figure 6: DVR Deployment Scenario in the Campus Network

The Enabling Technology of DVR

DVR is an extension to SPB and enhances the capability of Extreme Fabric Connect. It leverages a domain ISID as a communication channel to share a distributed IP host route data base, thus not overloading the IS-IS LSDB with host route information. The host routes themselves are exchanged through a separate instance of ISIS reserved for DVR exchanges. The shared domain communication channel is also used to exchange configuration information among the DVR domain nodes. Similarly the DVR backbone domain communication channel is also based on a pre-defined backbone domain ISID. Traffic forwarding is solely based on SPB forwarding mechanisms.

Local host ARP requests are handled by each DVR node locally ensuring highest possible data center scaling. DVR recognizes host movements across a DVR domain by monitoring gratuitous or reverse ARPs triggering a host move event in case of host location changes.

Resilient Architecture

DVR offers a hugely scalable architecture with an active-active resiliency architecture. Not only can DVR be deployed with one or more DVR Domains, but also within each domain one or more Controllers can be defined. For a resilient solution at least two Controllers must be deployed within each domain with up to a maximum of 8 Controllers to spread the load in large leaf spine designs. All controllers are active and a DVR Leaf can fully operate as long as at least one Controller is online within the domain.

Conclusion

DVR provides a highly scalable and robust network infrastructure for environments with dynamic roaming end points - whether they are virtual machines or wireless users or devices.

With DVR in place, traffic tromboning is avoided, thus reducing network latency. Traffic is optimized either east west within the data center or north south within the campus. DVR is a seamless extension of the Extreme Fabric Connect technology and is compatible with our VSP, ERS and EXOS based switches.

For more information on DVR and or Fabric Connect please contact an Extreme representative.