EVS Network Design Guide

LAN infrastructure recommendations for EVS product line

5 January 2016/version 1.0

Corporate
+32 4 361 7000

North & Latin America
+1 973 575 7811

Asia & Pacific
+852 2914 2501

Other regional offices
www.evs.com/contact

Contents

INTRODUCTION

As media workflows in the broadcast industry are gradually shifting from legacy technologies toward an Internet Protocol architecture it is of increasing importance for broadcasters and media facilities to carefully design the underlying architecture supporting the different models of media applications. Providing low latency, high reliability and overall highest performance of the network allows for optimal delivery of media contents within a production or post-production environment.

WHAT THIS DOCUMENT IS ABOUT

The purpose of this document is to describe the current state of the EVS products line in terms of Ethernet connectivity and give valuable guidelines and best practices for their integration within an IP network infrastructure. It starts with outlining the EVS recommendations for network segmentation, laying the foundations for media segregation and a performance oriented design for EVS applications. Specific requirements in terms of Ethernet and IP connectivity for common EVS products are then discussed. Special emphasis will be placed on the ever-increasing importance of IP multicast traffic. Integration within an existing infrastructure at the customer's premises is another important point we will focus on as well as integration with 3rd-party video editing appliances. Finally an optimal network design for the EVS ecosystem based on a 2-tier leaf and spine architecture will be depicted and thoroughly discussed.

This design guide is aimed at Broadcast/IT professionals and should serve as a reference for anyone willing to integrate EVS products within an Ethernet/IP infrastructure.

Hereinafter described guidelines and recommendations are the result of a long-standing experience and testing within our labs and large-scale facilities in order to insure their stability, as well as many successful practical implementations at our customers' premises.

When mentioned throughout this document, the term "server" must be understood in the most general sense and refers to any computer running software capable of accepting requests from clients. EVS XT/XS servers will be referred to specifically.

WHAT THIS DOCUMENT IS NOT ABOUT

As one may have worked out from the preceding paragraph this document focuses primarily on considerations related to the Layer 2 (Data Link i.e. Ethernet) and Layer 3 (Network i.e. Internet Protocol) of the OSI model. Specifics of Layer 1 (Physical) or Layer 4 and above (Host Layers), nonetheless important matters, won't be discussed hereinafter.

For hardware recommendations such as Ethernet switches, transceivers, cables, Network Interface Cards and compatibility with EVS products please refer to the individual hardware datasheets.

For software specifics such as used TCP/UDP ports, API's or application protocols details please refer the individual software user guides.

This document was not written with the aim to be a comprehensive network course for Broadcast/IT engineers. Readers are assumed to be familiar with basic IP and Ethernet concepts and terminology. We won't get into details for each technology or protocol discussed hereinafter as this guide is primarily meant to be a practical guide for EVS products integration.

Some EVS software rely on, or can be integrated with, cloud platforms and web services. This topic won't be discussed in this document as it currently focuses only on LAN requirements. WAN requirements may be included in future releases.

Finally general network security aspects such as access control, encryption, access-lists or firewalls won't be considered at the moment.

GENERAL ETHERNET DESIGN

The main aspects to consider while designing a network infrastructure for the EVS ecosystem are: performance, reliability and scalability.

In order to achieve optimal performance EVS recommends using **Virtual Local Area Networks** (or VLANs). A VLAN is a logical grouping of two or more devices. This logical grouping may extend across many switches thus allowing to group hosts regardless of their physical location. This grouping results in segmentation and isolation between groups of hosts thus allowing for broadcast control, security, Layer-3 address management, and traffic-flow management between VLANs.

BROADCAST CONTROL

As the number of devices within a broadcast domain increases, so does the broadcast rate within that broadcast domain. Moreover switches that cannot handle multicast packets or for which multicast management is poorly configured will treat multicast packets as broadcast and will flood them within their originating VLAN.
The broadcast rate is significant since each device must process each broadcast packet to determine whether the content of the packet should be pushed up the protocol stack. For each broadcast that is received, the receiving device must interrupt the CPU to evaluate the content of the broadcast frame. Apart from wasting CPU cycles broadcast traffic wastes bandwidth available on each host. Excessive broadcast and multicast traffic can seriously degrade network performance and will eventually induce latencies that can severely impact time-sensitive applications and servers. An important aspect of VLANs is that broadcasts transmitted in one VLAN are not propagated to other VLANs. Within an EVS workflow it is thus crucial to limit chatty protocols, most of the time used for control and management purposes and relying on broadcast and multicast packets, to a single VLAN in order to spare servers and interfaces transferring heavy and time-sensitive media content.

NETWORK SEGMENTATION

Network segmentation allows for traffic segregation thus limiting the impact of traffic patterns within one VLAN for other VLANs. By separating types of traffic, quality of service can also be applied in order to reduce delay, jitter and packet loss for specific traffic classes. This aspect will become increasingly important as Video Production facilities are transitioning to a converged IP-based infrastructure.

EVS recommends segmenting the network into the following four types of traffic:

	Control[1]	File Based[2]	NLE[3]	Real Time[4]
MTU	1500	9000	9000	1500
VLAN ID[5]	201	202	203	204
IP Subnet[5]	172.29.201.0/24	172.29.202.0/24	172.29.203.0/24	172.29.204.0/24

Unicast routing should be enabled between these VLANs. Multicast routing will be discussed further below.

This network segmentation applies to the EVS production "island" only. EVS products will often also exist in the customer's house network within another defined network address space and with specific constraints imposed by the customer's network equipment. Pre-requisites and limitations for these workstations and servers should be carefully studied on a per-product basis before deployment.

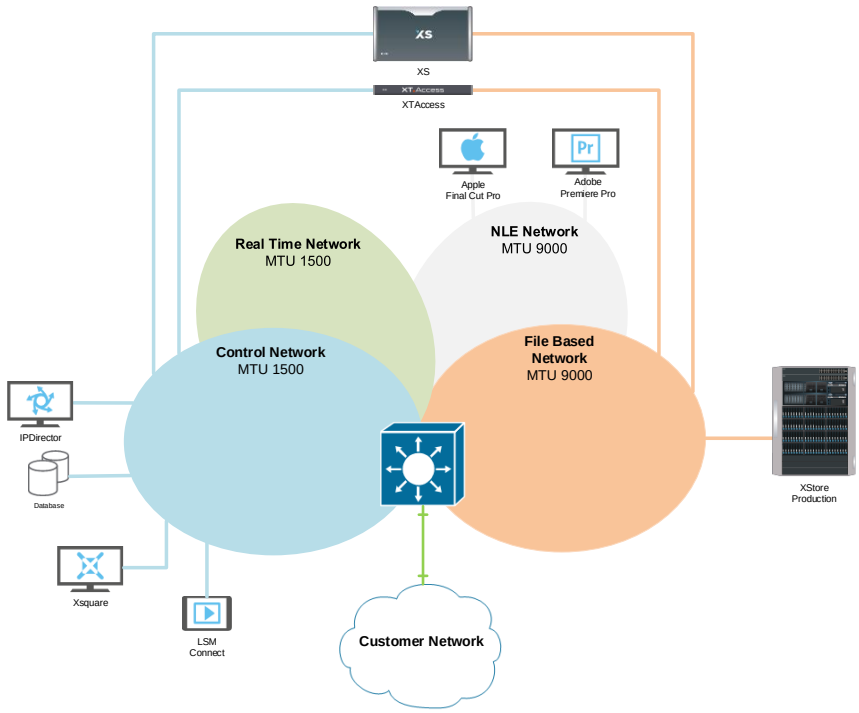


Figure 1 - Recommended VLAN segmentation

[1] Database transactions, jobs control, orchestration, …
[2] File transfers (mostly between encoding/playout servers and storage using XTAccess software)
[3] Common Non-Linear Editing Systems such as Adobe Premiere Pro, Apple Final Cut Pro, Xedio CleanEdit. Avid Technology systems are handled differently and discussed later on in this document.
[4] Live Video Production with real-time constraints. Requirements are subject to change as standards are currently being defined.
[5] Suggested value

ETHERNET MTU

The MTU is the maximum payload length for a particular transmission media. The MTU for Ethernet is typically 1500 bytes. The maximum packet length for Ethernet is typically 1518 bytes with 14 bytes of Ethernet header and 4 bytes of CRC included, leaving 1500 bytes of payload. In order to provide maximum network throughput several EVS products are configured to use, whenever possible, Ethernet frames with a payload of 9000 bytes. Those non-standard Ethernet frames are more commonly referred to as Jumbo Frames. Consequently in order for these packets to be correctly forwarded, the data path between EVS servers configured with Jumbo Frames must accept Ethernet frames with a payload of 9000 bytes. Jumbo Frames usage is required within the “File Based” and “NLE” VLANs and routing of such frames should be possible between those two VLANs. Jumbo Frames will never be routed to other VLANs and will be concealed from non-Jumbo hosts.
Although frequently misunderstood, TCP communications between hosts on a non-Jumbo VLAN and hosts whose NIC is configured to accept Jumbo Frames will remain possible. The TCP handshake will determine the Maximum Segment Size (or MSS) between two hosts and will not exceed the smallest MTU of both hosts.

It is recommended to configure a maximum Ethernet Payload of 9000 bytes for each network interface cards connected to a Jumbo-capable VLAN and to set on each corresponding switch port the maximum MTU size allowed. This maximum MTU size allowed on a switch port is platform dependent and should be greater than 9000 bytes.
If the requirement to have Jumbo Frames compliant VLANs is not feasible for technical reasons it is strongly advisable to inform EVS staff promptly as effective bandwidth available from each server within the “File Based” and “NLE” VLANs may be lessen.

Tip
It’s easy to test Jumbo Frames capability between two hosts using the following command on of them in order to ping the other one’s IP address:
ping -l 8972 -f *172.29.202.50*
The “-f” option set the “Don’t fragment” flag in the packet
The “-l” option set the payload of the ICMP echo request. 8972 is the maximum value that you can test for 9000 bytes frames. But why? Because 8972 is the ICMP payload size, to which you must add 8 bytes for the ICMP header size and 20 bytes for the IP header size making 9000 bytes of Ethernet payload in total.
The command above will work on Microsoft Windows hosts. From a Linux host the matching command would be:
ping -s 8972 –M do *172.29.202.50*

LINK AGGREGATION

Each EVS product can be connected to one or two VLANs using for each connection one or two network interfaces depending on the redundancy needs.

For each product one of the following teaming modes for network interfaces will be used:

Adapter Fault Tolerance (AFT)
Used when the network interface cards are connected onto the same network switch (or stack of switches). A failed primary adapter will pass its MAC and Layer 3 address to the failover (secondary) adapter. Spanning Tree Protocol (STP) must be disabled on the corresponding switch ports.

Switch Fault Tolerance (SFT)
Used when two adapters connect to two separate switches to provide a fault tolerant network connection in the event that the first adapter, its cabling or one of the switch fails.

Link Aggregation Control Protocol (LACP)
Must be used with an 802.3ad capable switch. The Link Aggregation Control Protocol (LACP) allows the exchange of information with regard to the link aggregation between the two members of said aggregation. This information will be packetized in Link Aggregation Control Protocol Data Units (LACDUs).

Each end of the link must be configured as an active or passive LACP port.

Passive LACP: the port prefers not transmitting LACPDUs. The port will only transmit LACPDUs when its counterpart uses active LACP (preference not to speak unless spoken to). This is the preferred mode for switch ports.

Active LACP: the port prefers to transmit LACPDUs and thereby to speak the protocol, regardless of whether its counterpart uses passive LACP or not (preference to speak regardless). This is the preferred mode for servers NICs.

On most hardware platforms LACP allows for the aggregation of more than 2 links (often up to 8 aggregated links but this limit is platform-dependent). A standard LACP setup will then protect a host from (at least) a single NIC or link failure. LACP however requires every interface of a single host to be connected to the very same switch thus not protecting the host from a switch failure. To overcome this limitation two approaches are possible. The first one consists in using stackable switches. Stackable switches are most of the time connected through a proprietary stacking backplane and operate as one single logical switch allowing LACP to run across links connected to different members of the stack. The other approach when connecting a single host across a pair of redundant

switches consists in using a multi-chassis link aggregation protocol like e.g. Cisco's Virtual Port Channel (vPC) or Arista's Multi-Chassis Link Aggregation Groups (MLAG) which allows physically independent switches to share LACP status information.

LACP is not only a fault-tolerant protocol but also allows for the distribution of Ethernet frames to all physical links available within the aggregated link thus presenting to the host a network bandwidth that is equal to the sum of all active interfaces. This aggregated bandwidth available to the host often leads to misunderstanding from the user as LACP suffers from significant shortcomings in terms of load-balancing that are worth discussing here.

How well the individual Ethernet frames will be distributed and by how much the practically possible data throughput will be increased depend heavily on the specific implementation of the link aggregation in a given switch or NIC driver.

The LACP standard does not define a specific algorithm for frame distribution but specify that the algorithm cannot cause out-of-order delivery or duplication of frames that are part of any given conversation (or flow). The consequence of these requirements is that all frames that are part of a given flow are transmitted on a single link in the order that they are generated by the host.

This means that, for example, when transferring a file between two servers, each with 4 LACP-enabled Gigabit links (thus 4 Gbps of bandwidth available), the transfer will be limited to a single link within the link aggregation on each end resulting in only 1 Gbps of actual bandwidth available between the two hosts.

Moreover the load-balancing algorithm acting on a per-flow basis (instead of per-packet basis) may often result in unevenly distributed load across the links within the bundle, some flows being much bandwidth-heavier than others. This will eventually lead to network latencies appearing when the network is under stress but way before reaching its theoretical bandwidth limits.

It is however possible to tweak the load-balancing efficiency of a switch or host by carefully choosing the fields used by the LACP hashing algorithm when determining which link to select, effectively defining how the algorithm distinguish flows. A load-balancing method based on **Source and Destination IP address** often provides suitable outcome and should be the preferred configuration for most cases.

In order to illustrate the impact of the LACP hashing algorithm let's consider the following example where two XTAccess are writing to a central storage equipped with two Gigabit interfaces configured with LACP. In the first scenario, the switch hashing algorithm is based on the "Destination IP" field alone within the IP packets. As both XTAccess are writing to the same destination IP address this will result in the switch choosing the exact same egress interface for all traffic going out of the switch to the storage thus wasting in this case half of the available bandwidth.

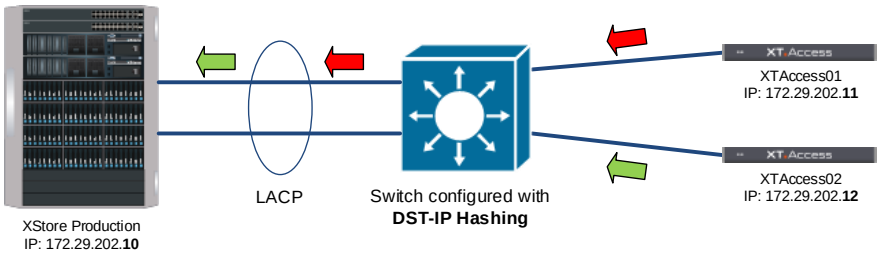


Figure 2 - Destination IP hashing

Carefully configuring the Ethernet switch algorithm to use both the “Source IP” and “Destination IP” fields will result in two different hash outcomes for packets coming either from XTAccess01 or XTAccess02 thus effectively sharing the load between both available interfaces to the storage.

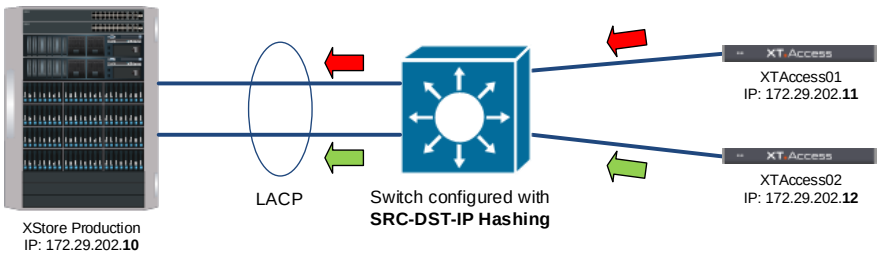


Figure 3 - Source/Destination IP hashing

Finally the following table sums up the main characteristics of the different link aggregation methods:

	SINGLE SWITCH	STACKED SWITCHES	PAIR OF SWITCHES	FAULT TOLERANCE	BANDWIDTH AGGREGATION	SERVER SIDE ONLY	BOTH ENDS CONFIG
AFT	✓	✓		✓		✓	
SFT			✓	✓		✓	
LACP	✓	✓		✓	✓		✓
LACP + MC-LAG			✓	✓	✓		✓

EVS PRODUCTS CONNECTIVITY

The following table describes how common EVS products should be connected to the Ethernet network:

	VLAN	Network Interfaces	Recommended Teaming Mode	NIC Speed	MTU
C-Cast Agent	Control	1	-	1 Gbps	1500
	File Based	1	-	1 Gbps	9000
	WAN	1	-	1 Gbps	1500
DB Server	Control	2	SFT/AFT[6]	1 Gbps	1500
Epsio	Control	2	SFT/AFT[6]	1 Gbps	1500
IPDirector	Control	2	SFT/AFT[6]	1 Gbps	1500
IPDirector with HD sw player	Control	1	-	1 Gbps	1500
	File Based	1	-	1 Gbps	9000
LSM Connect	Control	1	-	100 Mbps	1500
Multireview	Control	1	-	1 Gbps	1500
	File Based	1	-	1 Gbps	9000
XFile3	Control	2	SFT/AFT[6]	2 x 1Gbps	1500
	File Based	2	LACP	2 Gbps	9000
XIP	Control	2	-	2 x 1Gbps	1500
	Real-Time	2	-	2 x 10 Gbps	1500
XT3/XS[7]	Control	1	-	100 Mbps	1500
	File Based	2	LACP	2 / 20 Gbps	9000
XStore A/P[8]	File Based	2 / 4	LACP	20 / 40 Gbps	9000
XTAccess	Control	2	LACP	2 Gbps	1500
	File Based	2	LACP	2 / 20 Gbps	9000
Xedio CleanEdit	NLE	2	LACP	2 / 20 Gbps	9000

[6] Depending on the network topology
[7] Equipped with TGE network card
[8] Please refer to the “EVS XSTORE 10G – Network requirements” document

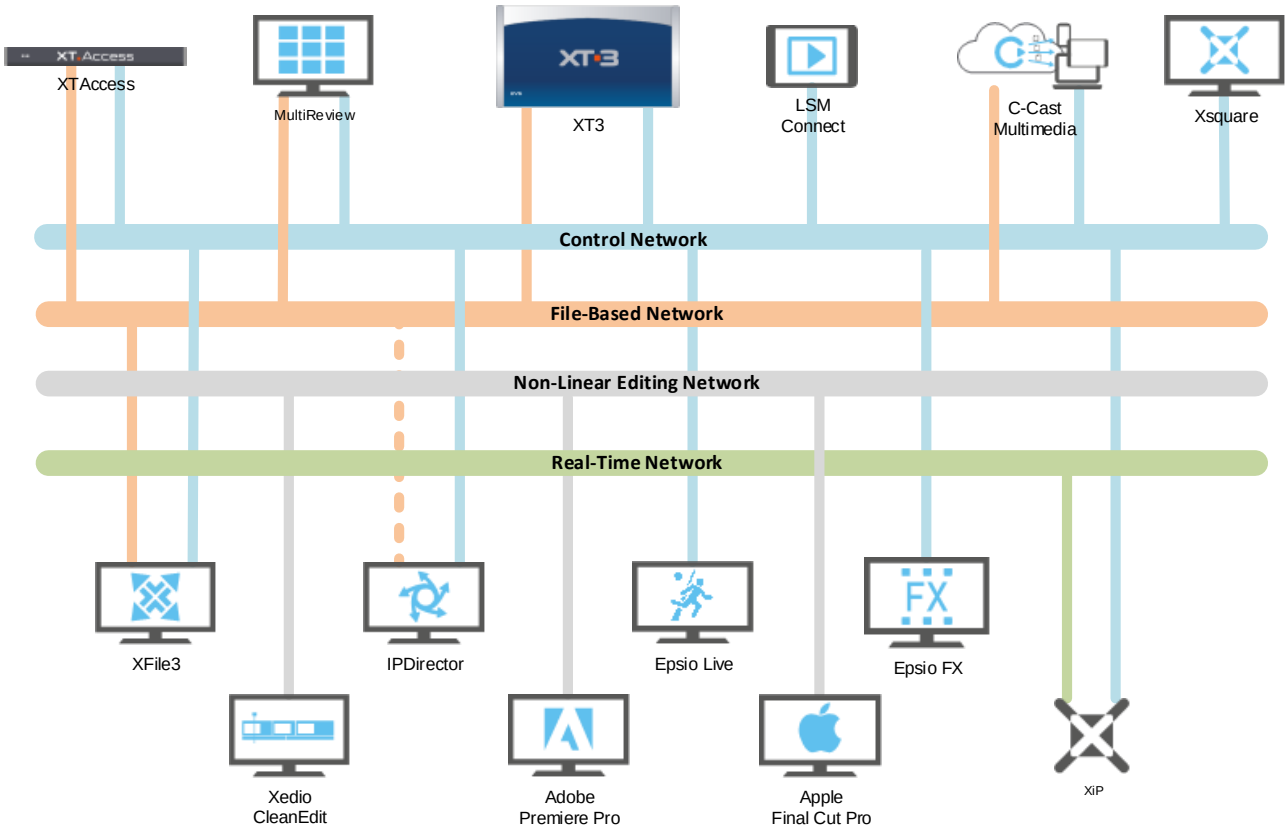


Figure 4 - Mapping products to VLANs

In order to avoid **duplex mismatch** issues that can lead to severe performance drops it is strongly discouraged to manually set interfaces' duplex mode or speed on any side of the link. Auto negotiation should be enabled at any time, both on the host interfaces and on the switch ports.

On the Ethernet switch, disabling **spanning-tree protocol** on a per port basis for EVS servers is recommended as it allow those ports to enter the forwarding state immediately. However extreme care needs to be taken whenever disabling spanning tree features as this can create Layer 2 loops within the network. Do not disable spanning tree on ports connecting to other switches or routers and for optimal security please consider using BPDU filtering / guarding techniques whenever you disable spanning-tree protocol on switch ports.

Similarly disabling **dynamic trunking protocol** on switch ports connecting to hosts and statically configuring them as access ports will speed up link initialization.

IP REQUIREMENTS

SERVICE DISCOVERY AND CONTROL PROTOCOLS

EVS software and servers use a collection of service discovery protocols which allow automatic detection of devices and services offered by these devices on the network. These protocols dictate in some way how the network must be designed and configured. Amongst the protocols used by EVS software let's discuss the most noteworthy:

LinX Protocol: LinX is a proprietary protocol developed by EVS for discovery and control. In its current state LinX relies on broadcast messages to discover XT/XS Servers available on the network. As broadcast packets are never routed outside of a layer 2 domain (or broadcast domain) this discovery is therefore limited to the local Ethernet segment. Any EVS products using LinX to communicate with XT/XS Servers must thus have at least one network interface card located on the same VLAN as the management interface of the XT/XS Servers.

We give here below the list of EVS software that rely on LinX for some of their operations:

> C-Cast Agent	> Ingest Funnel	> XFile3[9]
> C-Cast IPconnect	> LSM Connect	> XFileLite
> Dyvi	> Multicam	> XFileStreamer
> Epsio FX	> Multireview	> Xplore
> GX Server	> Nano Air	> Xsquare Suite[9]
	> TruckManager	> Xtract

When one of these software is installed on a server it is mandatory to have at least one Ethernet port of the machine in the same VLAN as the management port of the XT/XS servers for the specific discovery operation.

For other operations LinX uses the multicast address 225.0.0.64. Technically this multicast group could be routed across VLANs, however as per the previously mentioned requirement to have all network interfaces running the LinX protocol within the same subnet it is most of the time unnecessary and greatly ease the network configuration.

[9] These software provide a workaround to the broadcast-based discovery by allowing the user to manually add "non-discoverable"/remote XT/XS servers.

Bonjour Discovery: Along with the LinX protocol, Xsquare software relies on Apple's Bonjour Protocol in order to discover XTAccess servers available on the network. This protocol uses the well-known multicast IP address 224.0.0.251 which is part of the multicast 224.0.0.0/24 range of addresses which is reserved for local traffic only. Usually Bonjour packets have a TTL (Time To Live) of 1 and therefore cannot be easily routed between VLANs. It is thus recommended to have at least one interface of each XTAccess reachable in the **Control Network** by the Xsquare software. Although several methods allow for the routing of the Bonjour Protocol across different subnets (i.e. Bonjour Gateway, mDNS, Wide-Area Bonjour, …) none of them is currently qualified or supported in an EVS environment.

IPDirector Discovery: the IPDirector suite makes extensive use of multicast packets for various administration tasks, devices discovery and events management. This protocol is proprietary but the multicast addresses that are used are readily configurable by the user. In order to avoid complex multicast routing across the network it is recommended to have all IPDirector workstations connected into the **Control Network**. In many cases instances of IPDirector (e.g. IPBrowse, IPEdit, …) will exist in the Customer's house network as well and will require multicast groups used by the IPDirector software to be routed through the network. IP Multicast routing will be discussed further on in this document. In some limited cases when only a handful of IPDirector stations are located remotely and routing IP Multicast is not feasible it is possible to use the LAN/WAN parameters of the Remote Installer software. In that case unicast packets will be used for remote IPDirector stations instead of multicast.

EVS Truck manager: in mobile units, the Truck Manager software can help engineers to speed up the IP configuration of the various EVS devices. In order to do so the software must be installed on a machine that is connected to the Control Network.

MULTICAST IP REQUIREMENTS

As discussed previously EVS software heavily rely on multicast traffic for control and management protocols. In the near future the transition of Live Video Production from SDI-based connectivity to IP networks will lead to a considerable increase of the amount of multicast traffic conveyed in the network. Multicast traffic uses network infrastructure efficiently by requiring the source to send a packet only once, even if it needs to be delivered to a large number of receivers. The nodes in the network take care of replicating the packet to reach multiple receivers only when necessary. It is thus crucial in a Broadcast oriented IP network to configure those nodes carefully to manage the multicast traffic in the best way possible. Indeed poorly configured multicast management can ultimately lead to wasted bandwidth, misbehaving applications and performance drop.

Sources and receivers use a common IP multicast group address taken from the reserved multicast address space (224.0.0.0/8 in IPv4) to send and receive multicast messages. Sources use the group address as the IP destination address in their data packets. Receivers use this group address to inform the network that they are interested in receiving packets sent to that group.

IGMP

In IPv4 networks, in order for a host connected to a network switch to receive multicast packets for a specific group, it has to register to this group through IGMP (Internet Group Management Protocol). EVS products currently rely on IGMP version 2.
The IGMP membership process occurs on a local network segment where the host requests membership to a group through its local router while the router listens for these requests and periodically sends out subscription queries.

IGMP SNOOPING

IGMP Snooping is a function which examines IGMP protocol messages within a VLAN in order to avoid all ports of the switches to be flooded with multicast packets. With IGMP Snooping accurately configured on a switch, multicast traffic will be sent only to hosts interested in receiving this traffic. If IGMP snooping is disabled within a VLAN or if IGMP snooping is not available on the switch then incoming multicast packets will be treated as broadcast packets and will be flooded out to every switch ports. This behaviour is not suitable as it implies lot of wasted bandwidth and can eventually oversubscribe some links.

EVS recommends to make sure that **IGMP Snooping** is **enabled** in each VLAN that was defined for EVS servers. Most Layer 3 switches available on the market right now are

IGMP Snooping capable but one has to make sure that the service is effectively running, more specifically by configuring an **IGMP snooping querier** that will send out membership queries. When an IGMP snooping querier is enabled, it sends out periodic IGMP queries that trigger IGMP report messages from hosts that want to receive IP multicast traffic. IGMP snooping listens to these IGMP reports to establish appropriate forwarding.

Although this process usually works effortlessly behind the scene two particular cases have to be briefly discussed:

- Multicast addresses in the 224.0.0.0/24 range are considered link-local multicast addresses. They are used for discovery protocols and will not be constrained by IGMP snooping but flooded to every port thus wasting network bandwidth and hosts' CPU cycles. These addresses are reserved and should not be used for user application.

- Some hosts may fail to answer to an IGMP membership query and consequently will not receive the multicast packets they may be expecting. In that case it is mandatory to configure a 'static-join' IGMP command for the specific multicast groups on the switch ports where those hosts are connected.
  For example, due to a shortcoming of Android 4.1, the **LSM Connect** tablet does not send IGMP join packets and therefore the switch port to which it is connected should be configured with a static join on IP multicast group **225.0.0.64**.

MULTICAST ROUTING

When a specific multicast group has receivers not only in the subnet where the sender resides but also in other remote networks then multicast packets have to be routed.

In a typical 'closed' EVS ecosystem, the network segmentation presented earlier make it unnecessary to route any multicast group by grouping senders and receivers for a single multicast group into the same VLAN.

However when EVS software or servers exist in remote networks (e.g. in the customer's house network or at remote locations) it may be necessary to then route traffic for a selected subset of multicast groups.

Multicast routing is a feature that is commonly found on high-end network switches and routers but is less likely to be found on middle or low end appliances. Sometimes those features require a licence update which often comes with a high cost. It is thus of the most importance to carefully select the appropriate hardware that will make up your network infrastructure in order to meet your needs. EVS has qualified a list of network equipment suited for most use cases. Your EVS sales or pre-sales representatives can help you determine the most appropriate model of switch for your setup.

We will not discuss multicast routing here in great lengths as this topic alone could be the subject of a very thick book. As previously stated this document is not meant to serve as a network course but it is nonetheless interesting to present the basic concepts and requirements of an IP network where multicast routing is unavoidable.

First of all it is worth noting that for most protocols designed to send and receive multicast routing updates, proper operation depends on knowing the unicast paths towards sources and nodes. Therefore it is crucial to first have a properly configured unicast routing IP infrastructure before adding any multicast routing on top of it.

Amongst the common delivery and routing protocols used for multicast distribution, Protocol Independent Multicast (or PIM) is widely available on many platforms and most of the time the best suited for a typical EVS environment.

Several flavours of PIM exist amongst which **PIM Sparse Mode** (PIM-SM) is the most widely deployed and generally scales well when relatively few routers are involved in each multicast group. It is well suited for groupware applications where many senders and many receivers interact with each other using a common multicast group.

With PIM-SM, routers do not forward multicast packets for a group, unless there is an explicit request for traffic. Requests are accomplished using PIM join messages which are sent hop by hop toward the **rendezvous point** (RP). The RP keeps track of multicast groups, and the sources that send multicast packets are registered with the RP by the first-hop router of the source.

The rendezvous point can be specified by multicast group. It can be set either statically on each node (Static RP) or dynamically learned (BSR, Auto-RP, Anycast RP, …). As a typical EVS production environment has usually a limited number of routers involved Static RP is usually the preferred and most straight-forward method used. However multicast routing just as much as unicast routing is very project specific and those topics should be discussed between the local IT team and the EVS project/integration team in order to reach agreement about the best technologies to use for a particular environment.

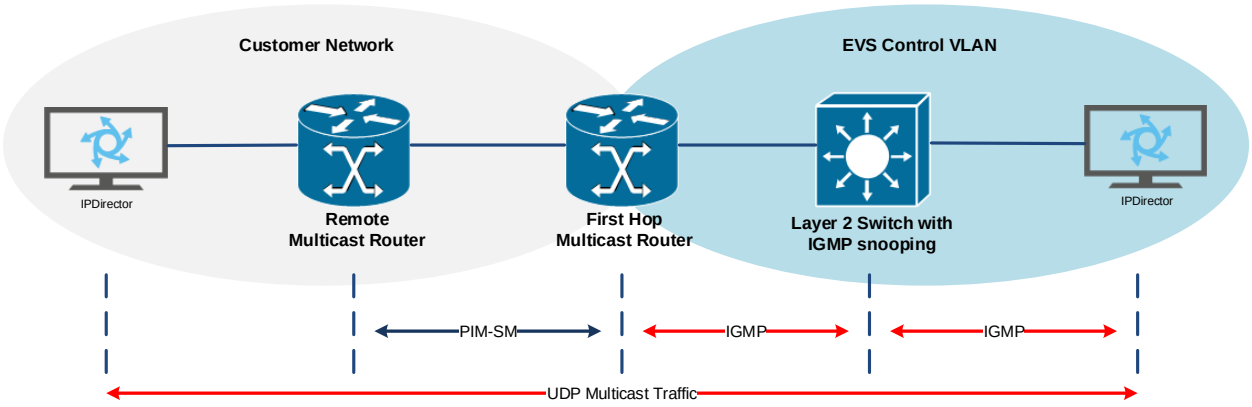


Figure 5 - End-to-end Multicast routing

LIST OF MULTICAST GROUPS

We give here below the list of multicast groups used by EVS software.

Multicast group	Purpose	Configurable
224.0.0.251	Bonjour Protocol	NO
224.14.0.69	IPDirector – Routing Protocol	YES
224.14.0.79	IPDirector – Remote Installer Discovery	YES
224.14.0.89	IPDirector – Remote Installer Package Distribution	YES
225.0.0.64	LinX Protocol	NO
239.193.0.1	Multireview Media Streaming	NO

RECOMMENDED DESIGN FOR STUDIOS

LEAF/SPINE MODEL

To provide a **resilient** and reliable Ethernet network in order to keep outages to a minimum while maximizing the available bandwidth between hosts as well as ensuring an optimum and linear cost model for scaling host densities up, EVS recommends using a two tiers Data Center topology often referred to as the Leaf/Spine model.

A typical implementation of this model consists of a pair of redundant **Spine switches** running in a resilient configuration. The Spines in turn connect to **Leaf** (Top of Rack) **switches** which terminate connections for hosts, storage, and other service nodes. All default gateway and routing functions are served from the Spine.

In this optimal model EVS recommends dual homing each host to a pair of interconnected switches (either 1G or 10G switches). Using a Multi-Chassis Link Aggregation Group (MC-LAG) protocol, these two switches will advertise themselves as a single LACP peer to all connected hosts allowing all network card interfaces of the host to be active.

Similarly these pair of Leaf switches will connect into the Spine tier in a bowtie fashion. The **bowtie MC-LAG** is simply a large MC-LAG between two pairs of MC-LAG peers making all links active and forwarding data unlike in a traditional Spanning Tree Protocol environment where redundant links would be blocked and waiting for a failure to happen before becoming active. In a MC-LAG deployment, the spanning-tree protocol remains in place as a fallback protection in the unlikely event of a MC-LAG failure or misconfiguration.

In order to provide first hop redundancy, EVS recommends using a "**L3 Anycast Gateway**" Protocol at the Spine level. Besides to providing first-hop (i.e. default gateway) redundancy, this kind of protocol will allow active/active router redundancy by sharing a virtual IP and MAC address (unlike well-known VRRP and HRSP protocols). Both Spine switches will respond to ARP requests for the virtual IP address effectively balancing the load between all active paths.

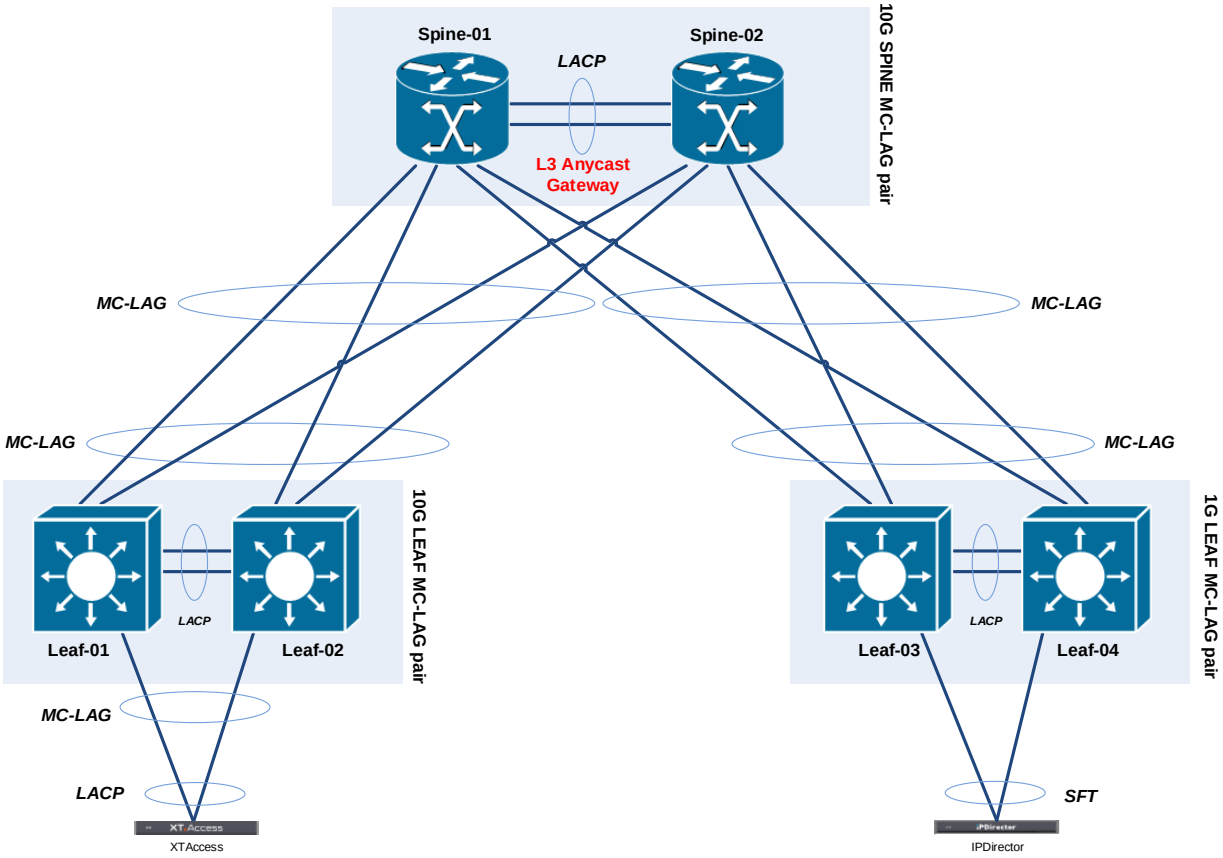


Figure 6 - Optimal active/active redundant Leaf/Spine topology

SCALED-DOWN TOPOLOGIES

While optimal, this active/active redundant Leaf/Spine topology does not always make sense for smaller setups. Designing an Ethernet network is always a trade-off between key factors such as capacity, scalability, flexibility, redundancy, performance and last but not least cost. Hereafter we will present scaled-down topologies for various less demanding use cases.

MID-SIZED/MIXED SPEEDS SETUP

Although the best practice would be to keep all hosts connected into the Leaf tier, this constraint would quickly drive costs up in a mixed 1G/10G mid-sized environment.

If the number of 10G-capable hosts is limited or if the total number of Leaf switches is small, plenty of 10G ports will remain available at the Spine tier and it makes more sense to connect 10G hosts directly into the Spine while keeping only 1G Leaf switches.

If 10G needs become more significant the topology could easily transition to a standard Leaf/Spine model with 10G Leaf switches.

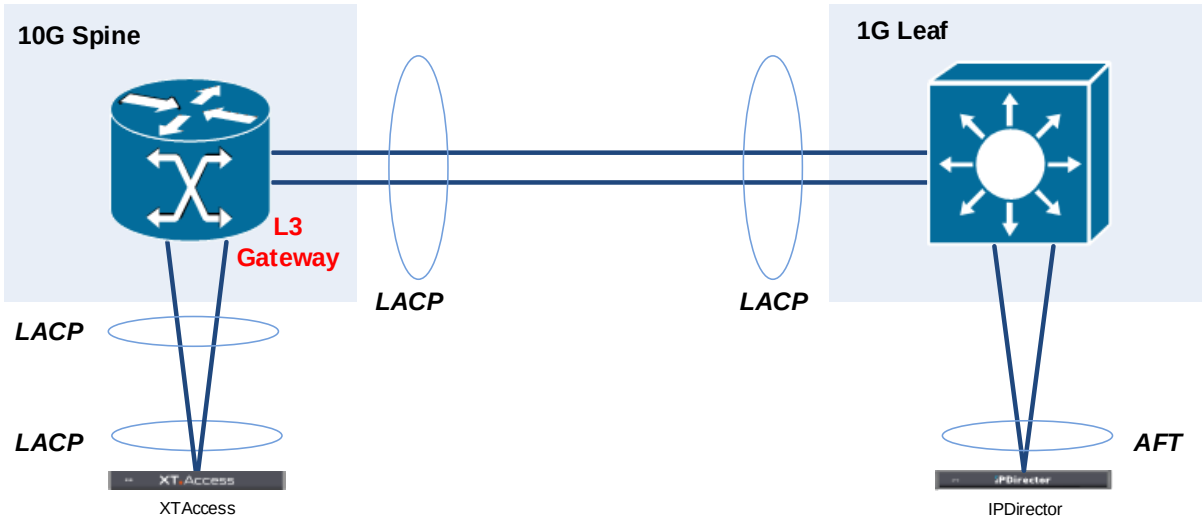


Figure 7 - Mixed Speed Leaf/Spine topology without switch redundancy

SMALL SETUP

If the longer-term requirements for number of ports can be fulfilled in a single switch (or pair of switches in a highly available design), then there's no reason why a single tier collapsed design could not be used. Single tier/Collapsed designs will always offer the lowest CapEx and OpEx as there are a minimal number of devices and no ports used for interconnecting tiers of switches.

SMALL/MIXED SPEEDS SETUP

There is unfortunately no ideal solution for small setup with mixed speeds hosts. Most 1G switches currently available on the market offer four 10G interfaces often used in order to interconnect them with other switches. With a collapsed design these available 10G ports can be used for connecting 10G hosts though their low number remains an important constraint.

Another alternative would be to opt for a 10G switch while using 10G-to-1G transceivers to allow for 1G connectivity. This choice also allows for future painless transition of 1G to 10G connectivity.

SWITCH SELECTION

As previously stated, it is of the most importance to carefully select the appropriate hardware that will make up your network infrastructure in order to meet your needs. EVS has qualified a list of network equipment suited for most use cases. Your EVS sales or pre-sales representatives can help you determine the most appropriate model of switch for your setup.

SPINE SWITCH SELECTION

Spine switch will typically be high-density 10G switches. The number of 10G interfaces needed will depend on the number of 1G Leaf switches to be connected. In a fully redundant topology we will need two 10G ports on each Spine switch for each pair or 1G Leaf switches. If 10G connectivity is required for a small number of hosts they need to be accounted for in the total number of 10G ports at the Spine tier. If 10G connectivity is required for a larger number of hosts then 10G Leaf switches will probably be necessary. In this case 40G links are recommended for interconnecting the Leaf switches to the Spine and proper model selection for the required interfaces is needed.

As previously discussed Spine switches will provide Layer 3 functionalities and thus should be capable of routing. Dynamic routing protocols for unicast (i.e. OSPF) and multicast (i.e. PIM-SM, IGMP querier, RP) are in most cases desired for integration with the existing customer's network. 'L3 Anycast Gateway' and 'L2 Multi-Chassis Aggregation' capabilities are mandatory for smooth integration with Leaf switches.
Finally the ability to route Jumbo Frames between VLANs is required by some EVS workflows.

LEAF SWITCH SELECTION

Leaf/Top of Rack switch will typically be high-density 1G switches where hosts will connect. Each switch will need two 10G ports to connect into the redundant Spine tier as well as two 10G ports to connect to a Peer Leaf switch when redundancy is needed for hosts.
When a large amount of 10G ports is needed the Leaf could be a high-density 10G switch with 40G ports for Spine and Peer interconnection.
As Layer 3 functionalities will be provided by the Spine tier it is then not mandatory to have Layer 3 capable switches at the Leaf tier. It is however mandatory that these switches are capable of Layer 2 Multi-Chassis Link Aggregation in order to provide active/active links to the hosts and into the Spine.
For optimal bandwidth management these switches should of course be capable of VLAN segmentation, 802.1q trunking and IGMP snooping.

Typically Leaf switches will have interfaces with different speeds (1G/10G or 10G/40G) and this speed mismatch may lead to buffering of traffic at egress ports. In-cast traffic patterns where multiple hosts discuss with another single host (e.g. database or storage) will also often lead to packet buffering. If the switch is not capable of coping with this buffering, packets will be dropped and latencies will unfavourably impact applications. It is highly advisable to select Leaf switch with a substantial buffer size in order to avoid such buffer overflow that can have a dramatic effect on the performance of applications and speed of data transfers.
Finally Leaf switches must be able to handle Jumbo Frames at the Layer 2 level.

SINGLE-TIER SWITCH SELECTION

As single-tier/collapsed design switches essentially combine functions of Spine and Leaf switches recommendations formulated above for both of them should be taken into consideration.

Required Minimum Functionalities for Ethernet Switches			
	H/W	**LAYER 2**	**LAYER 3**
SPINE	> High density 10Gb SFP+ ports > Optional: 40Gb ports (for 10Gb Leaf switches) > Redundant & hot-swappable power supply	> 802.3ad Link Aggregation/LACP > Multi-Chassis Link Aggregation > 802.1Q VLANs/Trunking > Jumbo Frames > IGMP snooping	> Equal Cost Multipath Routing for load balancing and redundancy > PIM-SM > Anycast Gateway Protocol > Jumbo Frames routing
LEAF	> High density 100BaseT with 10Gb uplinks or High density 10Gb SFP+ with 40Gb uplinks > Substantial buffering capabilities > Redundant & hot-swappable power supply	> 802.3ad Link Aggregation/LACP > Multi-Chassis Link Aggregation > 802.1Q VLANs/Trunking > Jumbo Frames > IGMP snooping	

INTEGRATION WITH CUSTOMER'S NETWORK

In the previous section we have discussed an optimal network topology for the EVS ecosystem in the studio. Most of the time this 'closed' EVS network will have to be interconnected with others such as the customer's house network.

The recommended model of interconnection between the EVS Spine switches and the customer's switches is a full-mesh topology using Layer 3 point-to-point links as presented below. This topology gives full redundancy and optimal load-balancing (thus maximum available bandwidth) to the EVS network.

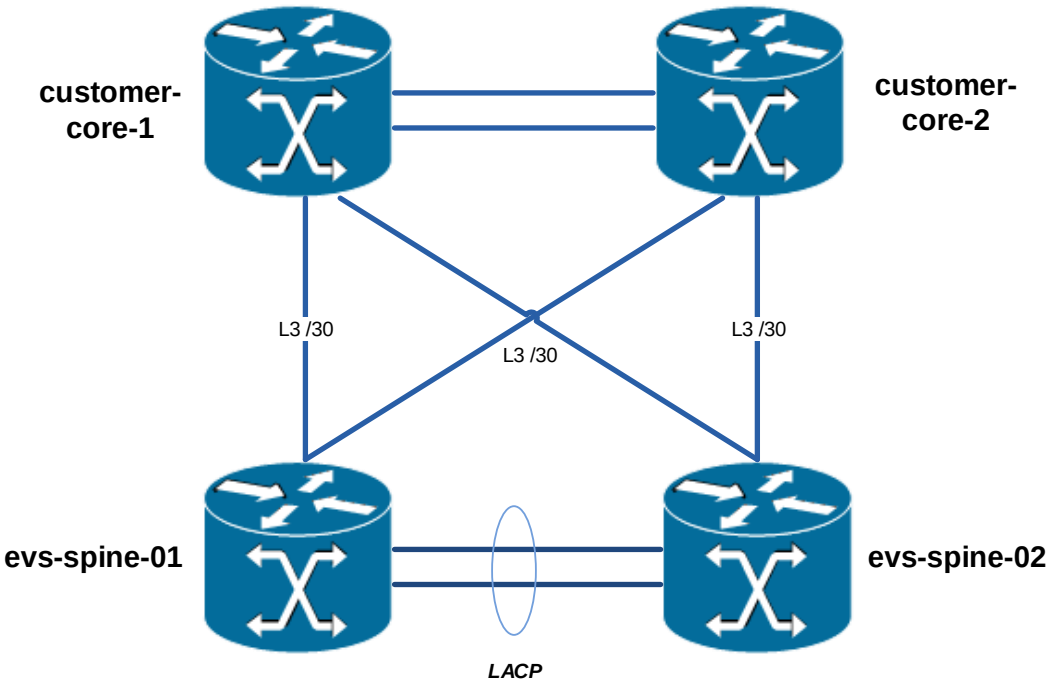


Figure 8 - Full-mesh interconnection with customer's network

In order to cut costs or if the number of available interfaces for uplinks is limited **and** if the combined bandwidth need from the house network to the EVS network is not greater than 20 Gbps in nominal mode or 10 Gbps in degraded mode then only one 10 Gbps link is sufficient to each Spine switch. In this case the topology will be as follows:

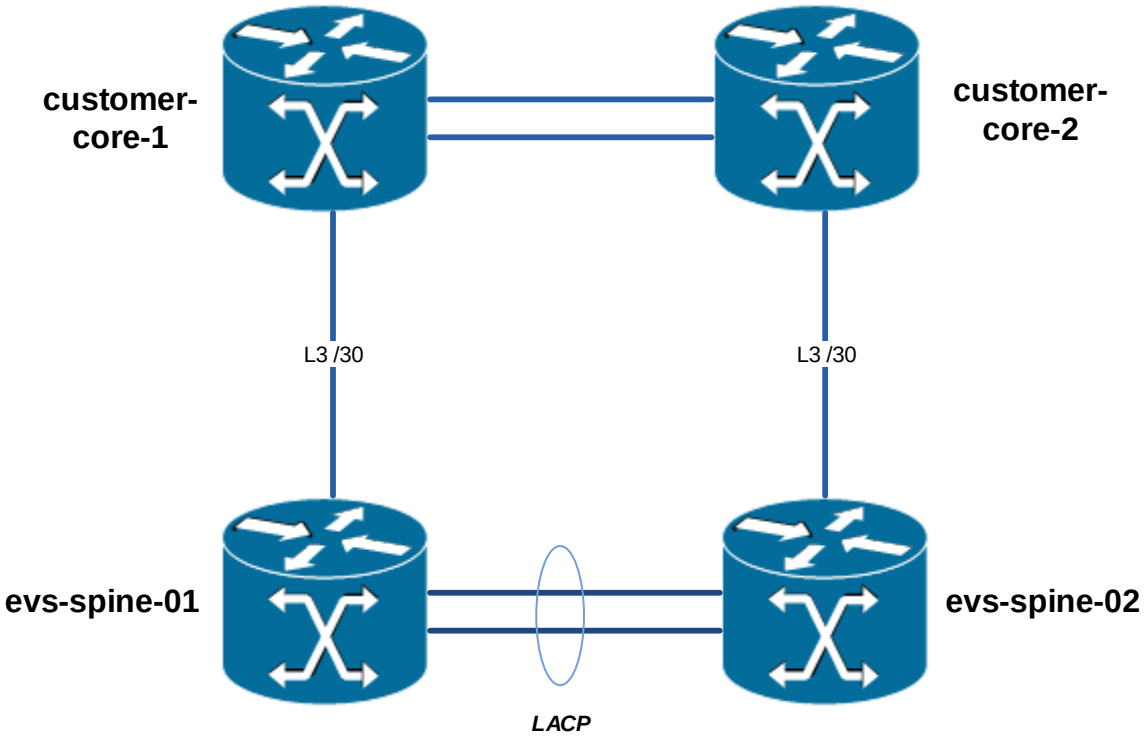


Figure 9 - Squared interconnection with customer's network

It is important to note that in such a case, in order to keep upstream traffic to the EVS network evenly load-balanced across the two Spine switches, the incoming flows must be evenly distributed across both available paths. In this situation the efficiency of the load-balancing will thus depend on a proper configuration on the customer's side of the network.

EVS recommends using an open dynamic IP routing protocol between customer's and EVS network, namely Open Shortest Path First (OSPF). This will allow load-balancing across equal-cost multi-pathing (ECMP) which results in an optimal use of the available network bandwidth between both networks.

If OSPF is not an option because of hardware constraints, then static routing might be an acceptable fall-back solution. In that case special care is needed to make sure that there is no bandwidth bottleneck created by a possible lack of equal-cost multi-pathing for static routes.

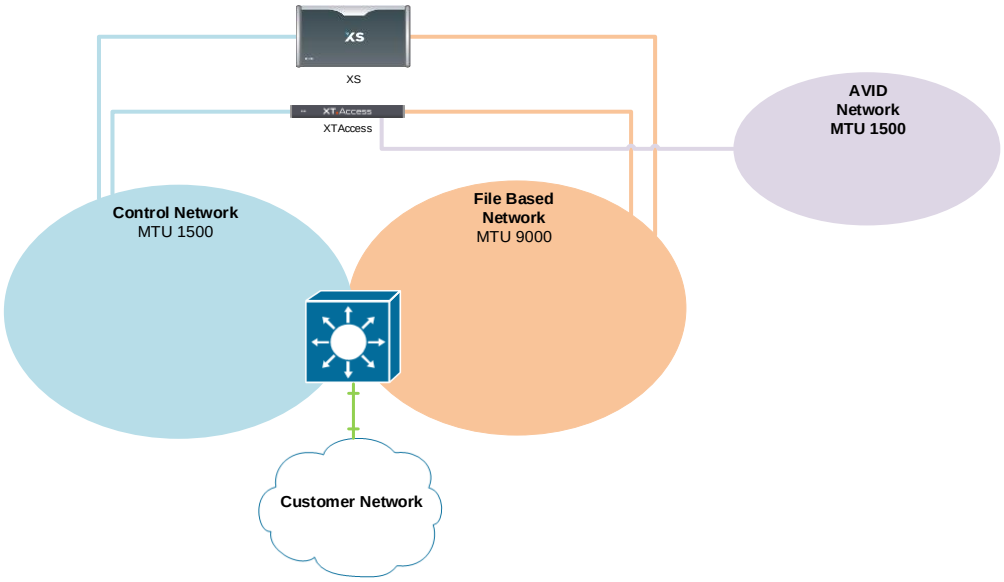
INTEGRATION WITH AVID ISIS

Avid Technology has some very strict rules for the Ethernet network supporting their ISIS storage solution. One of them is that hosts writing to their storage should never be configured with a MTU greater than 1500 bytes. Jumbo Frames enabled EVS servers are thus not allowed to write as such on an Avid system.

EVS suggests two solutions in order to allow for the integration of Jumbo Frames enabled EVS servers with the Avid storage.

EXTRA NIC ON XTACCESS SERVERS

The first option is to add one extra NIC into one or several XTAccess servers. This interface would then be directly connected onto the Avid IP network with a standard MTU size. This way those XTAccess will still be able to write using Jumbo Frames to EVS servers and storage while using standard frames to the Avid storage.



STATIC ROUTES ON XTACCESS

Another possibility not requiring extra connectivity is to keep XTAccess servers connected to the Control and File-Based VLANs only, as recommended previously, but to use their Control interface for outgoing traffic to Avid storage. The Control interface having a standard MTU of 1500 the Avid requirement is fulfilled. This method requires however to set up persistent static routes on the XTAccess servers in order to force them to use their Control network card when trying to reach the Avid subnet address range.

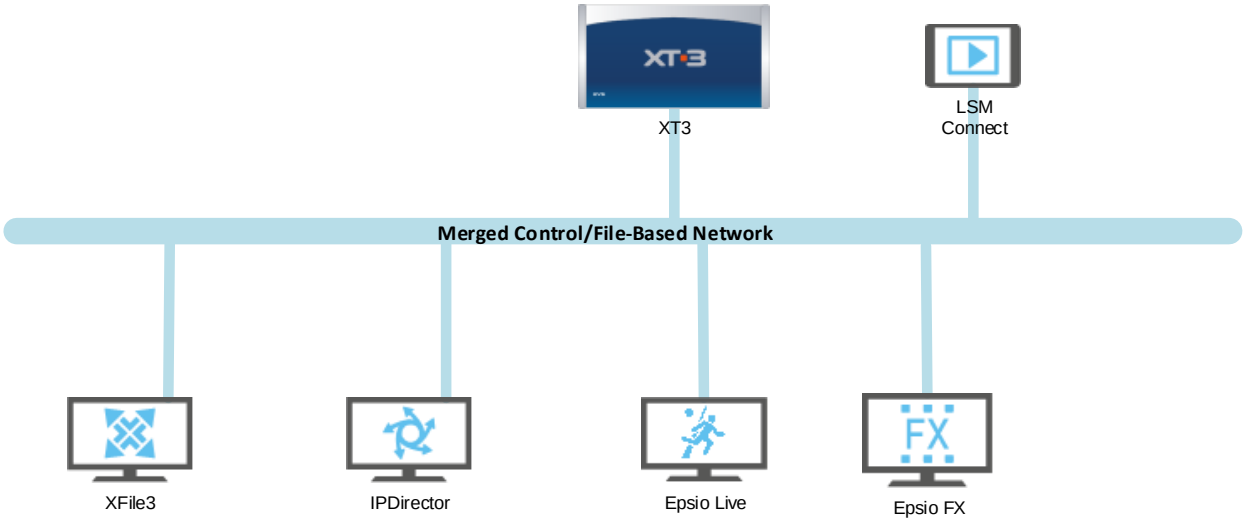
RECOMMENDED DESIGN FOR OB VANS

Although usually less demanding in terms of resiliency and performance, the Ethernet/IP network in Outside Broadcast facilities should be nonetheless carefully designed and managed.

Here below we will present three common scenarios and the resulting recommended network design.

SMALL SETUP

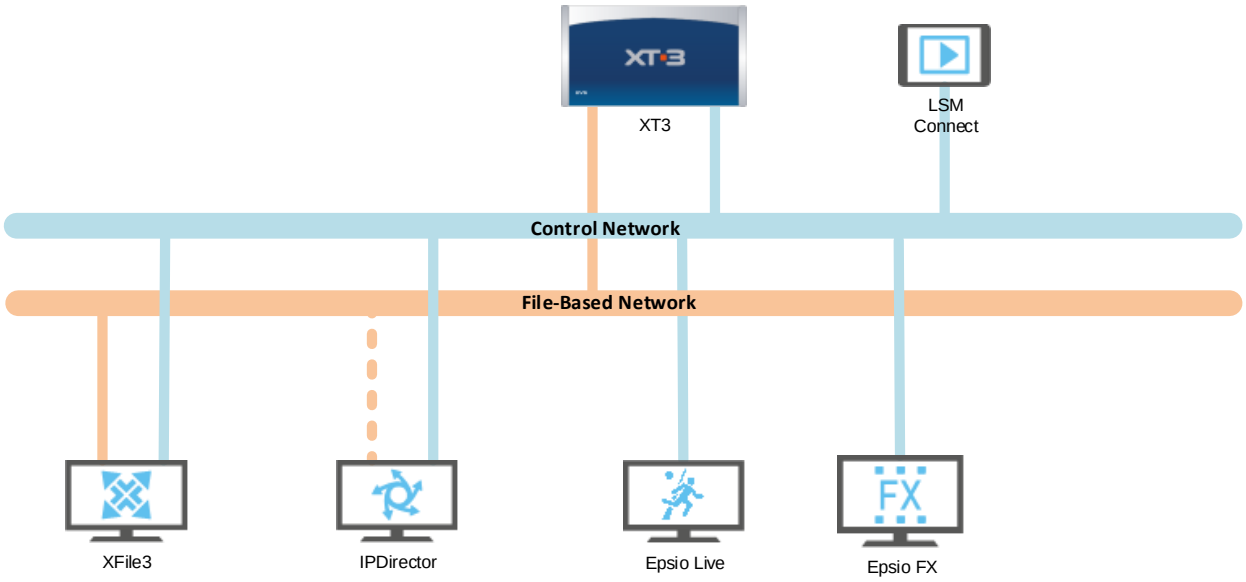
For small setups with just a few machines and a single Ethernet switch, a flat network design (i.e. without VLAN segmentation) can certainly meet the reasonable needs of end-users. The low complexity of such setup allows for very little or no specific configuration of the network equipment. Simplicity however comes with drawbacks including the lack of redundancy and scalability. Control and File-Based Media traffics coexisting on a same VLAN can interfere and increase latencies as the number of servers and workstations as well as the load on the network increase. For these reasons it is always recommended when possible to prefer a multiple-VLANs design as depicted later over the simple but limiting single-VLAN option.



MEDIUM SETUP

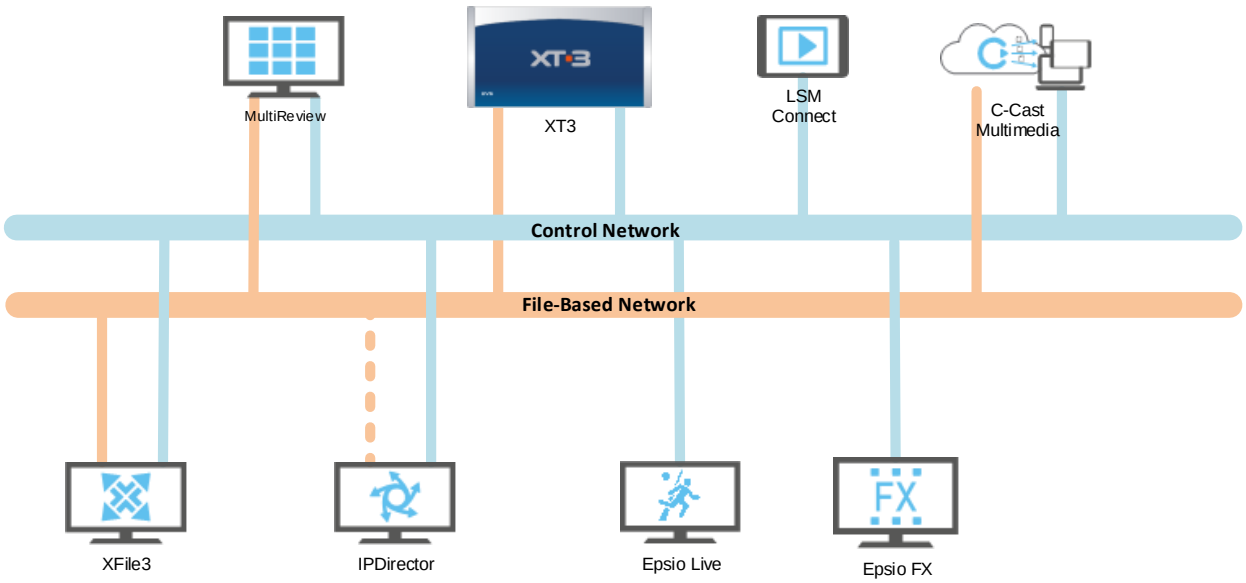
A safer though slightly more complex design requires to implement network segmentation as detailed earlier on in this document. In this case Control and File-Based media traffic are isolated within their own VLAN. This of course requires specific configuration of the network equipment as well as multiple Ethernet links for each server that needs dual

connectivity. Although slightly more complex to set up, this is the recommended design that will allow for painless addition of new servers to the setup when the need arises.



SETUP WITH MULTIREVIEW OR C-CAST

When the setup includes either a MultiReview or C-Cast product then a two-VLAN design is not only recommended but is a pre-requisite.



SETUP WITH MULTIPLE TRUCKS

For bigger setups where multiple trucks must be interconnected, depending upon the capabilities of the underlying network equipment, two scenarios are possible:

- Layer 2 interconnection with stretched VLANs across trucks
- Layer 3 interconnection with specific VLANs for each truck

The preferred solution for interconnecting a set of trucks is to stretch VLANs across each truck's network equipment either using well-known Layer 2 trunks (IEEE 802.1q) or some kind of network virtualization overlay such as VXLAN. This way all requirements related to multicast and broadcast IP traffic previously established in this document can bet met.

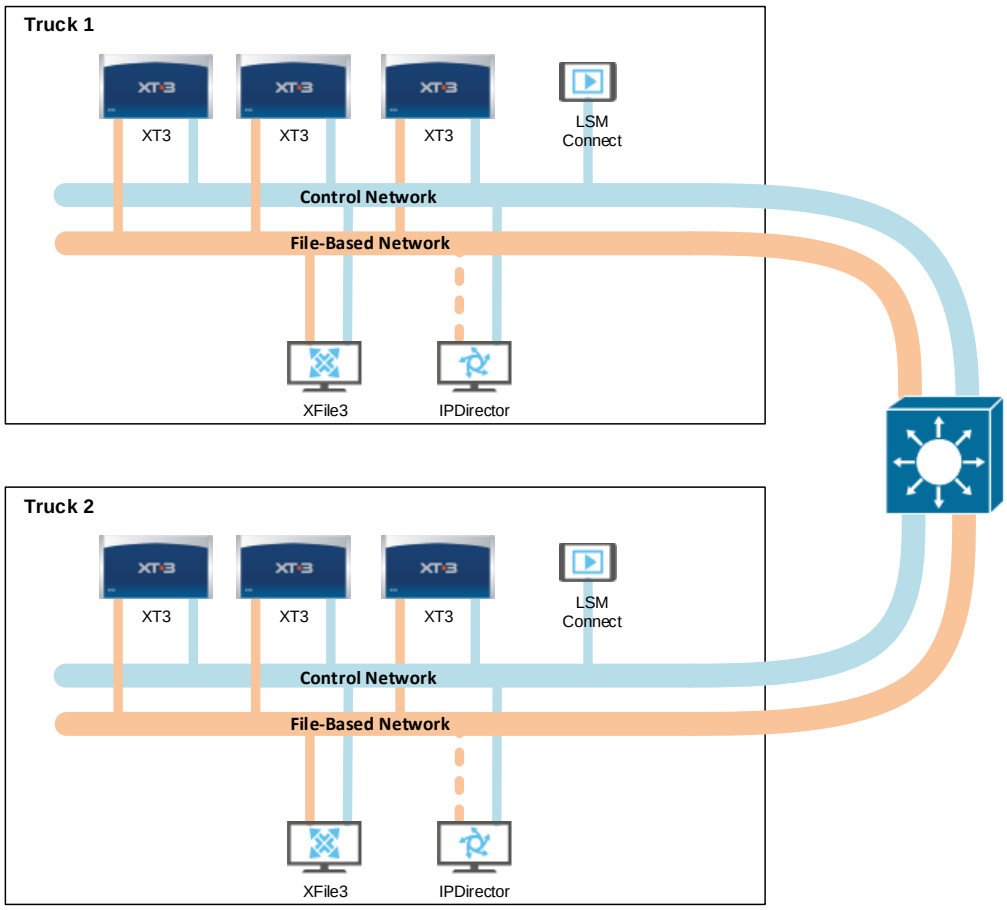


Figure 10 - Truck interconnection in Layer 2

If stretching VLANs across several trucks is not an option for technical reasons then traffic must be routed between each network segment. However this comes with limitations and constraints regarding the various EVS discovery and control protocols.

More specifically:

- Discovery of XT/XS servers by LinX-based EVS software will be limited by each truck's network boundaries. As previously explained a number of EVS software are able to overcome this limitation with manual static configuration.
- Discovery of XTAccess by Xsquare will be limited by each truck's network boundaries.
- Multicast protocols such as IPDirector's groupware protocol or LinX will be confined within each truck's network unless multicast routing is explicitly configured.
- Multicast streams such as Multireview will be restricted to one truck unless multicast routing is explicitly configured.

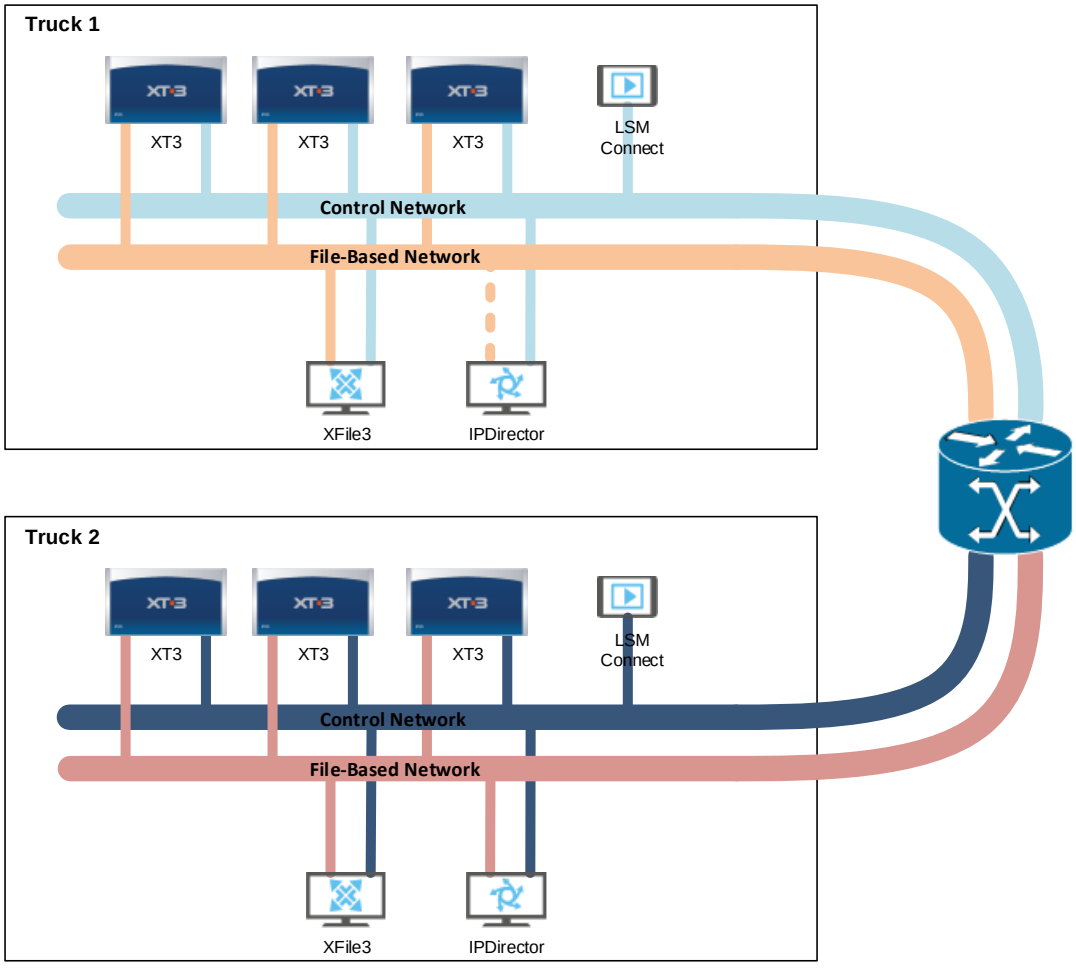


Figure 11 - Truck interconnection in Layer 3

QUALITY OF SERVICE

To date EVS has no peremptory request for Quality of Service (QoS) rules to be implemented at the network nodes level.

Instead of complex QoS control mechanisms the conventional policy consists in provisioning a network so that bandwidth capacity is based on peak traffic load estimates. This approach is simple for networks with predictable peak loads and the performance achieved is satisfactory for most applications.

GLOSSARY

AFT: Adapter Fault Tolerance. A type of link aggregation group (LAG).

ARP: Address Resolution Protocol. Protocol used for resolution of network layer addresses into link layer addresses.

BPDU: Bridge Protocol Data Units. Frames sent by the Spanning tree protocol (STP).

CapEx: Capital Expenditure. Fixed asset of a company such as acquired equipment.

CPU: Central Processing Unit. The electronic circuitry within a computer that carries out the instructions of a computer program by performing the basic arithmetic, logical, control and input/output (I/O) operations specified by the instructions.

CRC: Cyclic Redundancy Check. Error-detecting code used in digital networks and storage devices to detect accidental changes to raw data.

DTP: Dynamic Trunking Protocol. A proprietary networking protocol developed by Cisco Systems for the purpose of negotiating trunking on a link between two VLAN-aware switches, and for negotiating the type of trunking encapsulation to be used.

ECMP: Equal-cost Multi-Path routing. A routing strategy where next-hop packet forwarding to a single destination can occur over multiple "best paths" which tie for top place in routing metric calculations.

HSRP: Hot Standby Router Protocol (HSRP). A Cisco proprietary redundancy protocol for establishing a fault-tolerant default gateway.

IEEE 802.1q: The networking standard that supports virtual LANs (VLANs) on an Ethernet network. The standard defines a system of VLAN tagging for Ethernet frames and the accompanying procedures to be used by bridges and switches in handling such frames.

IGMP: Internet Group Management Protocol. A communications protocol used by hosts and adjacent routers on IPv4 networks to establish multicast group memberships.

IPv4: Internet Protocol version 4 (IPv4). The fourth version in the development of the Internet Protocol (IP). IPv4 is a connectionless protocol for use on packet-switched networks.

Jumbo Frames: Ethernet frames with more than 1500 bytes of payload.

LACP: Link Aggregation Control Protocol. A type of link aggregation group (LAG). Also known as IEEE 802.3ad

MC-LAG: Multi-Chassis Link Aggregation Group. A type of link aggregation group (LAG) with constituent ports that terminate on separate chassis, primarily for the purpose of providing redundancy in the event one of the chassis fails.

MSS: Maximum Segment Size. A parameter of the Options field of the TCP header that specifies the largest amount of data, specified in octets, that a computer or communications device can receive in a single TCP segment.

MTU: Maximum Transmission Unit. The largest size packet or frame, specified in octets (eight-bit bytes), that can be sent in a packet- or frame-based network.

NIC: Network Interface Controller. A computer hardware component that connects a computer to a computer network.

NLE: Non-Linear Editing system is a video or audio editing workstation that performs non-destructive editing on source material.

OpEx: Operating Expense. The ongoing cost for running a product, business, or system.

OSI model: Open Systems Interconnection model. A conceptual model that characterizes and standardizes the communication functions of a telecommunication or computing system without regard of their underlying internal structure and technology.

OSPF: Open Shortest Path First. A routing protocol for Internet Protocol (IP) networks.

PIM: Protocol-Independent Multicast. A family of multicast routing protocols for IP networks.

QoS: Quality of Service. The capability of a network to provide better service to selected network traffic.

RP: Rendezvous Point. A router in a multicast network domain that acts as a shared root for a multicast shared tree.

SDI: Serial digital interface. A family of digital video interfaces.

SFT: Switch Fault Tolerance. A type of link aggregation group (LAG).

STP: Spanning Tree Protocol. A network protocol that ensures a loop-free topology for any bridged Ethernet local area network.

TCP: Transmission Control Protocol. A standard that defines how to establish and maintain a network conversation via which application programs can exchange data.

TTL: Time to live. A mechanism that limits the lifespan or lifetime of data in a computer or network.

UDP: User Datagram Protocol. One of the core members of the Internet protocol suite. UDP uses a simple connectionless transmission model with a minimum of protocol mechanism.

VLAN: Virtual Local Area Network. A group of devices on one or more LANs that are configured to communicate as if they were attached to the same wire, when in fact they are located on a number of different LAN segments.

VoIP: Video over IP.

vPC: virtual PortChannel. A Cisco proprietary protocol that allows links that are physically connected to two different switches to appear as a single link to a third device.

VRRP: Virtual Router Redundancy Protocol. A networking protocol that provides for automatic assignment of available Internet Protocol (IP) routers to participating hosts. This increases the availability and reliability of routing paths via automatic default gateway selections on an IP subnetwork.

WAN: Wide Area Network. A telecommunications network or computer network that extends over a large geographical distance.