altran
Part of Capgemini

MULTI-UNIT STACKING

GOING BEYOND THE STANDARDS

TABLE OF CONTENTS

INTRODUCTION

Switch stacking is a technology that connects multiple physical switches into a single logical switch. Network administrators can add or remove switches to increase or decrease port count without getting a multi-card chassis switch and dealing with the configuration of multiple switches. The network can evolve over a period with a pay as you grow model and therefore benefit from both the switch vendors and their customers.

Leading switch vendors like Cisco, Juniper, HPE, Dell, Extreme Networks, Huawei, etc., support stacking in their enterprise and campus switch models. Leading silicon vendors like Broadcom and Marvell also support feature extensions that enable switch stacking. Performance is best if the underlying hardware also supports stacking, but switches can be stacked even without hardware support for the feature. Though the goals and objectives of stacking are the same, vendor implementations are completely proprietary.

In this white paper, we will discuss the objectives, topologies, operations and maintenance of stacked switches, and also look at some specific vendor implementations to compare them.

STACKING TERMINOLOGY

Before we dive deeper into stacking, it's topologies and its benefits, it is important to have clarity on to some of the basic stacking terminologies outlined in the below table:

Table 1 Basic Stacking Terminology

Terminology	Description
Master Switch	The switch chosen to control the stack
Slave Switch	A switch that participates in the stack and accepts being controlled by the master
Standby Switch	One of the slave switches which will control the stack if the master switch fails. We also refer to it as Backup master
Stack link	A link between 2 switches in the stack. This link is not exposed to other switches in the network. There may be multiple stacking links between adjacent switches
Stack Port	The port, on a switch, that the stack link is connected to

OBJECTIVES OF SWITCH STACKING

The Ethernet switch market is estimated at around $ 3.5 billion currently and is expected to move upwards to about $ 6 billion in the next five years. Stackable switches provide network administrators the ability and flexibility to move switches from areas where WiFi is replacing Ethernet and reuse them in areas where more port count is needed, without having to discard them.

Network administrators achieve the following by stacking multiple switches

Flexibility – Stacking technology provides the ability to scale the networks up and down as required. Switches can be added or removed from the stack on the fly to increase or decrease port count at any location without affecting the functionality or the networking capabilities of the existing switches.

Simplified configuration and manageability – Stacking architecture ensures that multiple switches can be managed as a single switch,i.e., through the master, thereby reducing the amount of configuration for an administrator and ensuring ease of configuration and management.

Ethernet Switch Market Size, Share, Growth, Industry Trends & Analysis
https://www.marketresearchfuture.com/reports/ethernet-switch-market-4576

STACKING TOPOLOGIES

A stacked switch can be created by interconnecting stacking capable switches in a variety of ways such as Daisy Chain, Ring, Mesh, etc. The Daisy Chain and Mesh are the most commonly employed. Some vendor designs use the regular switch ports for stacking while others need dedicated special port types.

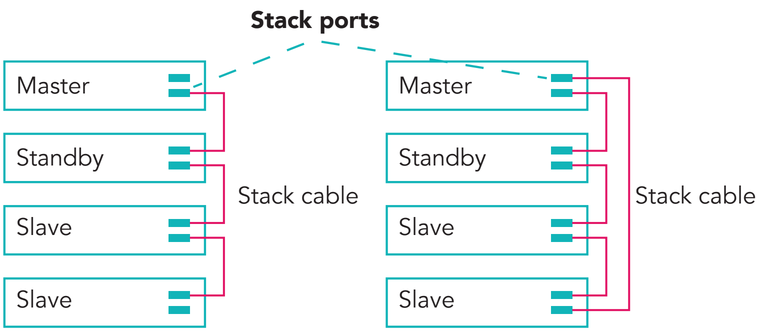


Figure 1: Daisy Chain and Ring Topology

The daisy chain is a linear connection of switches connected via stacking ports. This is very simple to put together but leads to stack truncation when any one of the stack ports interconnecting switches fails. So, this is usually used only when there is a large distance between the first and last switches. The ring is similar to the daisy chain except that the final switch in the stack connects back to the first switch. This is more robust as it supports redundancy and the topology can converge even if a stacking link fails. A mesh is much more robust and tolerant to even multiple failures, but the number of interconnections makes it more expensive and complex.

STACK OPERATIONS

Connecting multiple switches to form a stack involves several operations, all of which need to happen in parallel. These are outlined below

Role selection – In a stack, one switch must be master, one a Backup master, and the rest, Slaves. An election mechanism usually achieves this role selection that the switches execute between themselves, complexities include dealing with new switches introduced into the stack and switches removed from the stack.

Topology Discovery – The switches must discover how they are connected, and each switch must know the optimal route to reach every other switch in the stack, both for stack maintenance as well as for switching user traffic.

Loop Detection – The switches must ensure that in-ring and mesh topologies, traffic entering the stack links does not cause loops.

Fail-over – The Backup master and slaves must know when the master fails and must re-orient to keep the stack functional. This is usually achieved using the periodic Heart Beat messages.

The above functions are necessary to ensure that the stack will adapt dynamically and allow switches to be added or removed from the stack, and to support known problems like link failures. Vendors have implemented proprietary algorithms and methods to take care of resolving these operational issues.

STACKING ATTRIBUTES

There are several attributes of stacking that enable smooth operation of the stack. They are outlined in this section:

MAC Address – The logical stacked switch must use a single MAC address consistently irrespective of which switch unit in the stack is actually sending or receiving PDUs. One way of achieving this is to use a shared MAC address called Persistent MAC or Switch MAC, for all the switches in the stack.

Switch Programming – All the switches in the stack must have identical forwarding tables in the hardware, including MAC addresses, IP route entries, access control lists, etc. This is usually achieved by the switches synchronizing the information between themselves or by having the master alone setup the hardware tables in all the elements of the stack.

Firmware Upgrade – All switches in a stack must use the same firmware version. A common way to achieve this is by having the master overwriting the firmware on all the slaves with the version on the master.

File Synchronization – All switches in a stack must also use the same startup and configuration files. A common way to achieve this is by having the master overwriting these files on all the slaves with the files available on the master.

Switch vendors use different, sometimes proprietary, IPC mechanisms between the switches in the stack. These must support both unicast communication for a specific switch to switch messages and multicast communication from master to all other switches in the stack.

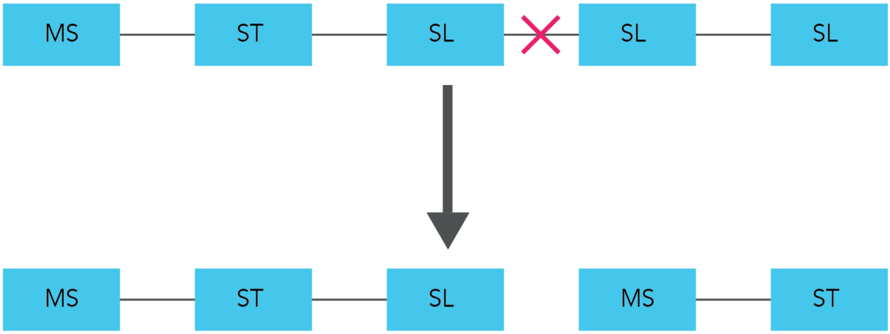
STACK INITIALIZATION AND MAINTENANCE

There are several operations and procedures involved in initializing and maintaining a stack. This includes handling stack split and stack merge scenarios:

Stack Initialization – Switches in a stack must ensure that there is only one master for the stack. They must also ensure that all switches in a stack are initialized with the same configuration.

Adding a switch to the stack – A new switch, introduced into a working stack, may become a slave or the master in the stack, depending upon the specific vendor implementation of stacking. The stacking solution must ensure that there is always one and only one master in the stack. Disturbance to the other units in the stack and the network must be kept to a minimum. This is usually achieved by avoiding unnecessary changes in master/slave roles.

Stack Split – A working stack may split into two either because of switch unit failures or because of stacking link failures. This happens with a single failure in Daisy Chains and with more than one failure in ring topologies. Stack split leads to dual master scenarios and different switch vendors handle the stack split in different ways. Each resulting stack must have one and only one master. Also, the two stacks resulting from the split must not end up using the same MAC or IP address.



MS: Master, SL: Slave, ST: Standby

Figure 2: Stack Split Scenario

Stack Merge – Two stacks may merge into one. This could happen when a failed switch, causing a stack split, recovers and starts up again. The resulting stack must have only one Master.

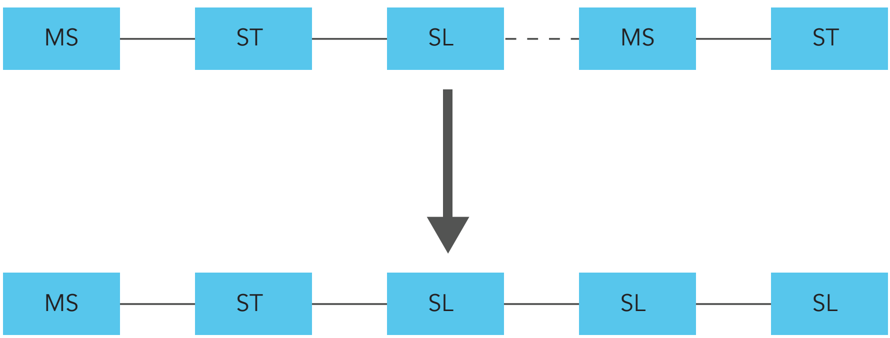


Figure 3: Stack Merge Scenario

HARDWARE PROGRAMMING FOR SWITCH STACKING

There are several aspects that need to be considered when programming the hardware when the switch is used in a stacked environment. The stacking link is used for sending stacking related control messages, protocol packets and data packets. The control messages include heart messages between master and standby switch, election and role assignment messages, messages from the master to program other switches in the stack, etc. These messages should be provided a higher priority over the protocol packets, which in-turn need to have higher priority over data packets. The QoS programming in the chipset shall be used to provide the priority assignment to the different types of packets.

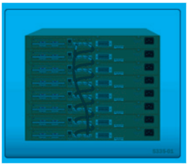
All the protocol packets should be handled by the master switch and for this purpose, each switch in the stack should program its hardware to re-direct the packets received on its front panel ports to the master switch via the stacking port. Different policing limits need to be applied for each of the protocol packets to ensure that the master switch CPU is not overwhelmed. The master switch should be able to transmit a packet on any of its ports or the port of any of the switches in the stack by adding metadata to the transmitted packet. This metadata should be added in the format supported by the specific chipset.

VENDOR IMPLEMENTATIONS

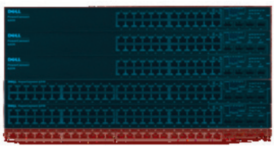
Vendors like Cisco, Dell, Extreme, etc. support stacking in access and aggregation switches. These are deployed in Campus LAN networks, Enterprise and branch office networks, and in Datacenters. We provide here an overview of some of the leading vendor implementations.

MAC Address – The logical stacked switch must use a single MAC address consistently irrespective of which switch unit in the stack is actually sending or receiving PDUs. One way of achieving this is to use a shared MAC address called Persistent MAC or Switch MAC, for all the switches in the stack.

Cisco Catalyst



Extreme Summit



Dell PowerConnect

Huawei iStack

Juniper EX

Figure 5: Vendors stacks for Comparison

Cisco switch – www.cisco.com Huawei switch – www.huawei.com

Though the switches offer similar capabilities of ports, topologies supported, resiliency and protocols supported, there are multiple variations of capacity, master election procedure and methods adopted for specific feature realization. We summarize this below.

Table 2: Vendor Stacks Comparison

Switch Stack	Number of switches in stack	Topology supported	Remarks
Cisco Catalyst	9	Daisy Chain, Ring	Support for election based on MAC address and priority configured in each node. Uses Cisco StackPower technology to create a pool of power resources that can be shared across switches.
Extreme Summit	8	Daisy Chain, Ring	Support election based on master Capability configuration and priority, which enables switches to be included or excluded from the master election.
Dell PowerConnect	12	Daisy Chain, Ring	Supports the master election based on priority and MAC address. The standby switch can be configured or auto-selected.
Huawei iStack	9	Daisy Chain, Ring	Supports the master election based on priority, MAC and startup time. The switch which completes in startup procedure fully in the shortest time is elected as standby. Support for Multi-Active detection techniques during stack split.
Juniper EX Series	10	Daisy Chain, Ring, Mesh	Supports the master election based on priority, system uptime, MAC address. Support for stacking with other Juniper switches models in the EX series.

STACKING AND ALTERNATE TECHNOLOGIES

Multi-card chassis design is another way to build switches with higher port counts. These typically have two control cards working in 1:1 redundant mode and multiple line cards to provide the increased port counts. Line cards can be added on the fly up to the maximum number of slots available in the chassis. These are used when the port count requirement is very high. These have reduced the switching latency due to a fabric-based architecture. Since the initial investment and scale are high, these are used for core and aggregation deployments. Stacking, due to its flexibility to grow and shrink support model is deployed in access and aggregation deployments. The following figure shows a typical deployment of a chassis along with stacking in a network.

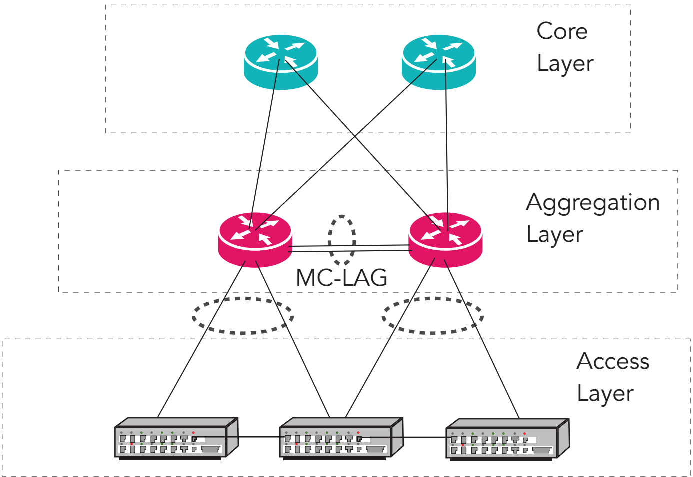


Figure 6: Stacking Deployment Scenario

ALTRAN FRAMEWORKS FOR MULTI-UNIT STACKING

Stacking is a complex technology providing resiliency, scalability and involving close interaction between switch software and hardware. Coordination between the stacking related protocols operating independently in each switch in the stack is required to realize stacking. Therefore, developing the stacking framework involves considerable investment and time for equipment vendors compared to selecting a framework that already supports stacking. Selecting an appropriate stacking framework allows vendors to add their value additions and thereby develop a differentiated product with time to market advantage.

Altran offers a licensable software framework termed ISS (Intelligent Switch Solution), which is widely deployed in switches used in Industrial. Enterprise, Datacenter and Metro Ethernet networks across the world. The framework supports all standard switch technologies, including Layer2, Layer3, MPLS, Telemetry and application software, along with stacking extensions. The stacking software has been developed using standard networking protocols with proprietary extensions and is referenced on standard merchant silicon. The framework supports centralized firmware upgrade, logging and management. This allows vendors to quickly adapt the framework to their hardware platform and develop a switch stacking solution reducing both development time and cost & addressing their customer need in an accelerated fashion.

CONCLUSION

Stacking allows switches to be deployed initially with lower port density and with increasing demand for services allows more switches to be added to the network element without disturbing the existing network. The flexible growth model combined with resiliency, simpler management model and scalability has led to stacking becoming the de-facto standard in access and aggregation switches. Therefore, all silicon vendors and software vendors support stacking in the frameworks. In summary, stacking reduces the cost of ownership, supports an agile model for switch deployment, making it an ideal option for the evergrowing enterprise and campus networks.

Contact us

marketing@altran.com

Author

Varadharajan Parthsarathy
Assistant Vice President, Technology

About Altran

Altran is the world leader in engineering and R&D services. Altran offers its clients a unique value proposition to meet their transformation and innovation challenges. Altran supports its clients, from concept through industrialization, to develop the products and services of tomorrow and has been working for more than 35 years with major players in many sectors: Automotive, Aeronautics, Space, Defense & Naval, Rail, Infrastructure & Transport, Energy, Industrial & Consumer, Life Sciences, Communications, Semiconductor & Electronics, Software & Internet, Finance & Public Sector. Altran has more than 50,000 employees operating in over 30 countries.

Altran is an integral part of Capgemini, a global leader in consulting, digital transformation, technology and engineering services. The Group is at the forefront of innovation to address the entire breadth of clients' opportunities in the evolving world of cloud, digital and platforms. Building on its strong 50-year + heritage and deep industry-specific expertise, Capgemini enables organizations to realize their business ambitions through an array of services from strategy to operations. Capgemini is driven by the conviction that the business value of technology comes from and through people. Today, it is a multicultural company of 270,000 team members in almost 50 countries. With Altran, the Group reported 2019 combined revenues of €17billion.

www.altran.com


altran
Part of Capgemini