Data Center Design Considerations

Unit 4

Outline

Data Center Design

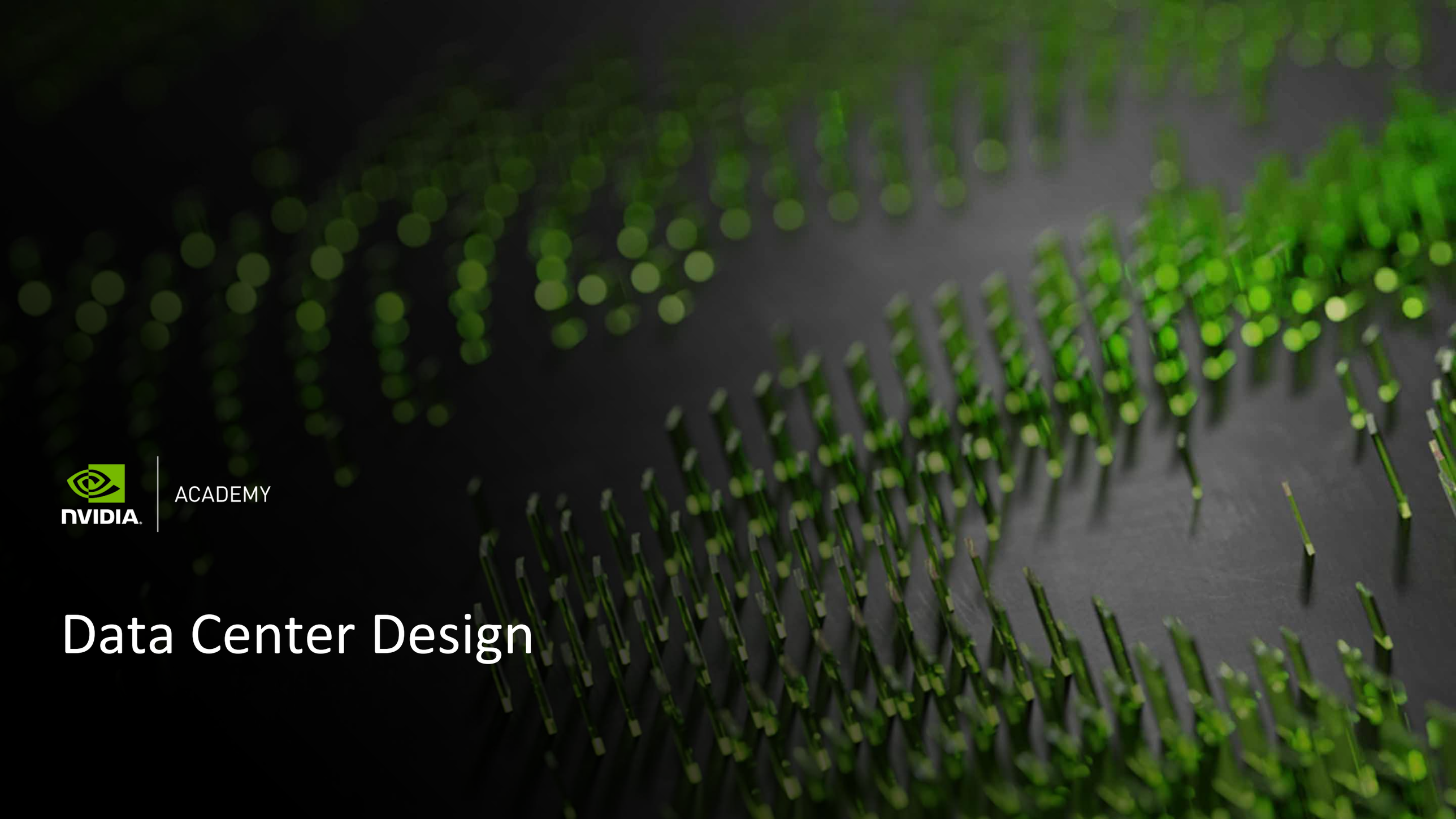
- Network Topology
- Hierarchical Network Design
- Leaf-Spine Design

Layer 2 Networks

- Layer 2 Protocols
- Layer 2 Network Design

Layer 3 Networks

- Layer 3 Protocols
- Layer 3 Network Design



Data Center Design

Data Center

A facility used to house computer systems and associated components

- A data center is a facility used to house computer systems and associated components, such as compute, network, and storage systems.
  - From the size of a small server room to a few buildings geographically distributed
  - A critical business asset where companies run their workloads
- The modern data center has evolved from on-premises to cloud infrastructures requiring:
  - Scalability
  - Resiliency
  - Security
  - Efficiency

Data Center Design

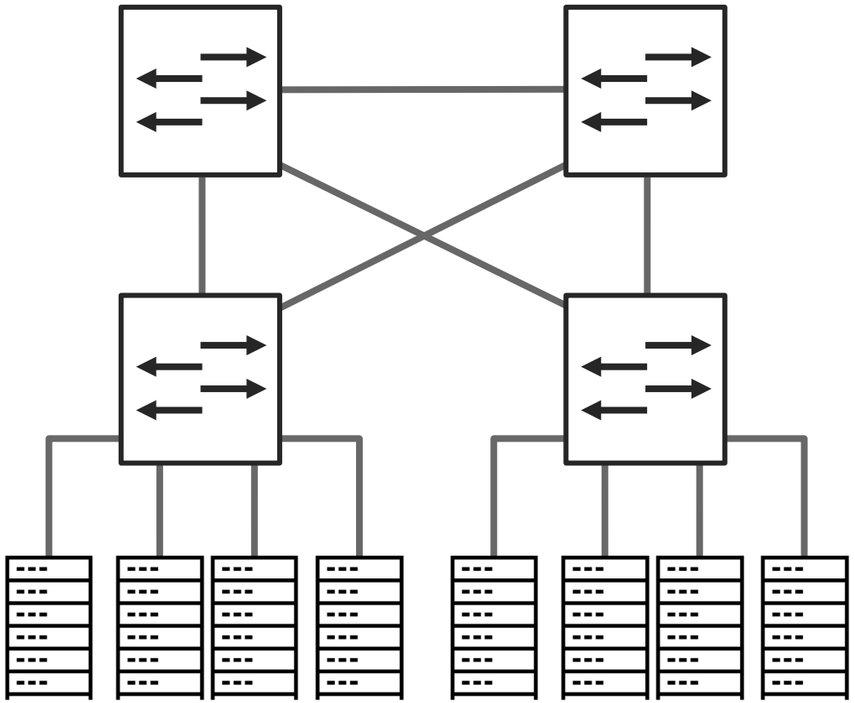
Planning essential data center parameters

- Data center design is the process of planning all of a data center's essential computational and non-computational parameters:
  - Number and type of required servers
  - Network layout
  - Power, cooling, and ventilation systems
  - Physical data center security
  - Disaster recovery
  - Business continuity planning
- Modern data center design will incorporate many different aspects that will make it easy to manage and more efficient.

Data Center Design

A network topology is the schematic arrangement of network elements

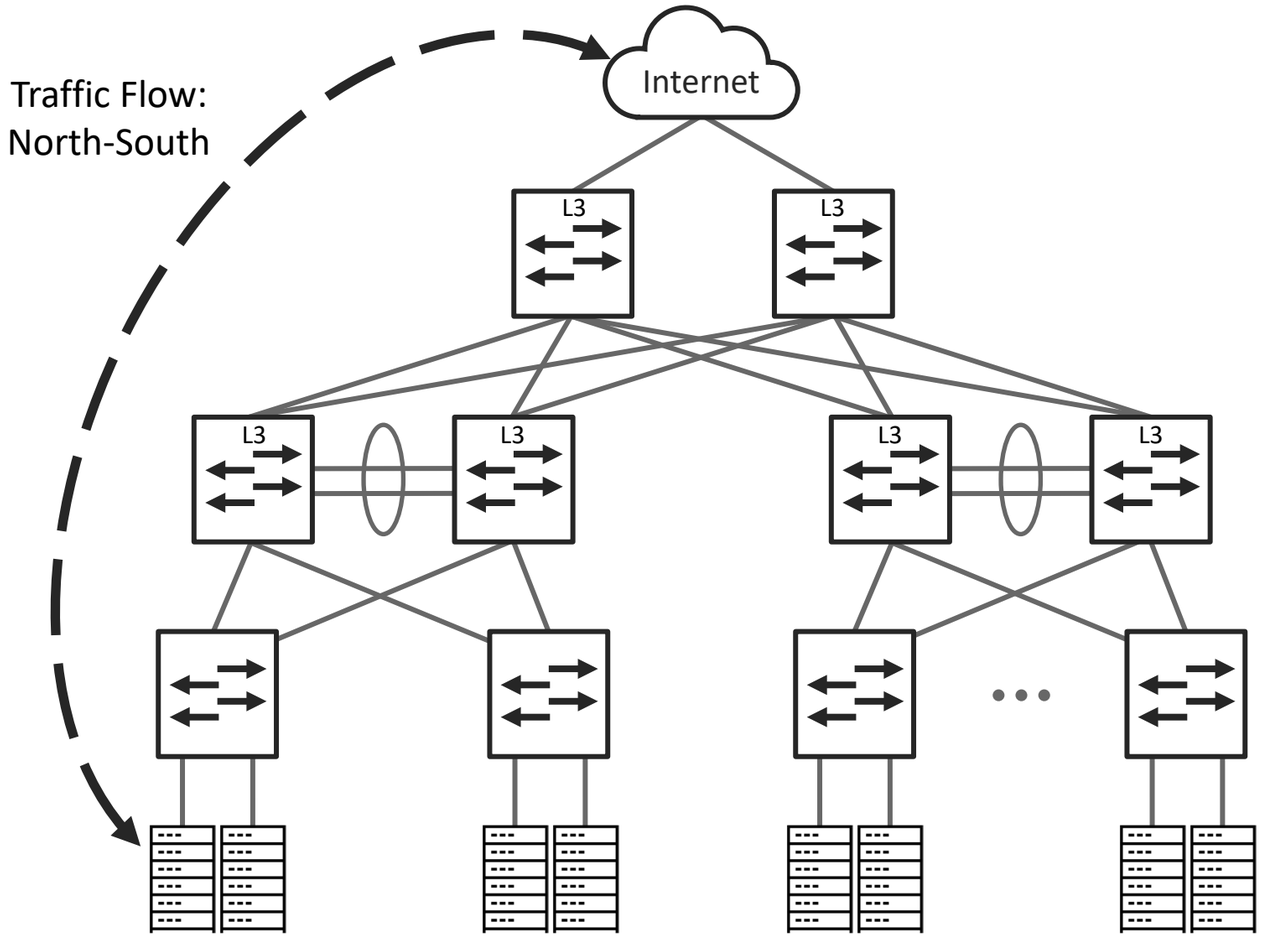
- A network topology is the schematic arrangement of network elements, such as links and nodes..
- The physical topology describes how devices on the network are connected
- The logical topology describes how data moves from one node to another.
- Considerations when choosing a topology:
  - Availability - redundancy and fault tolerance
  - Reliability - downtime and delays are unacceptable
  - Performance - locate faults, troubleshoot errors, allocate resources
  - Future growth - add new nodes without negatively affecting performance or the user experience
  - Budget – effective and affordable



Hierarchical Network Design

Traditional modular data center design

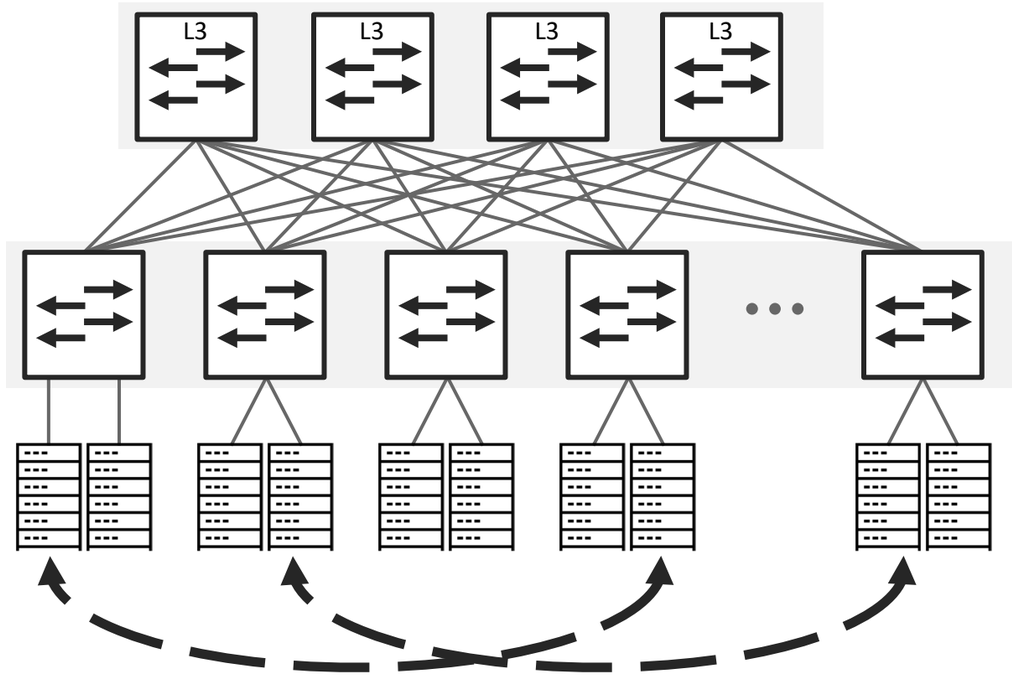
- A hierarchical design separates a network into distinct layers, where each layer has its role in the network.
- Most effective for north-south traffic
  - Traffic in and out of the data center
  - Client-server applications
- The model scales somewhat well, because of its modularity.
- May be subject to bottlenecks if uplinks between layers are oversubscribed.



Modern Data Centers

Modern data center environments are becoming much more dependent on east-west traffic

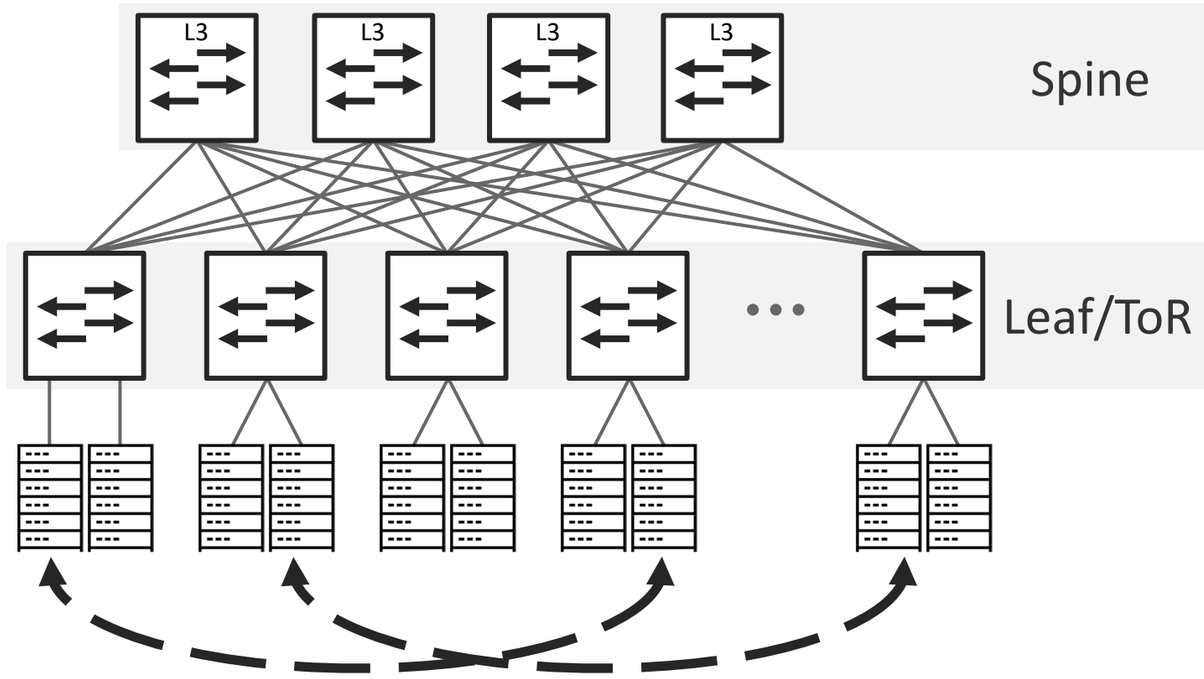
- Data center transformation is driven by new trends: cloud computing, virtualization, big data.
- Traffic patterns have changed to east-west communication:
  - More data moves between servers and storage nodes
- Modern data center design requirements:
  - Increased server-to-server communication
  - Scalability
  - Resiliency



Leaf-Spine Design

Modern data centers have transitioned to a leaf-spine design

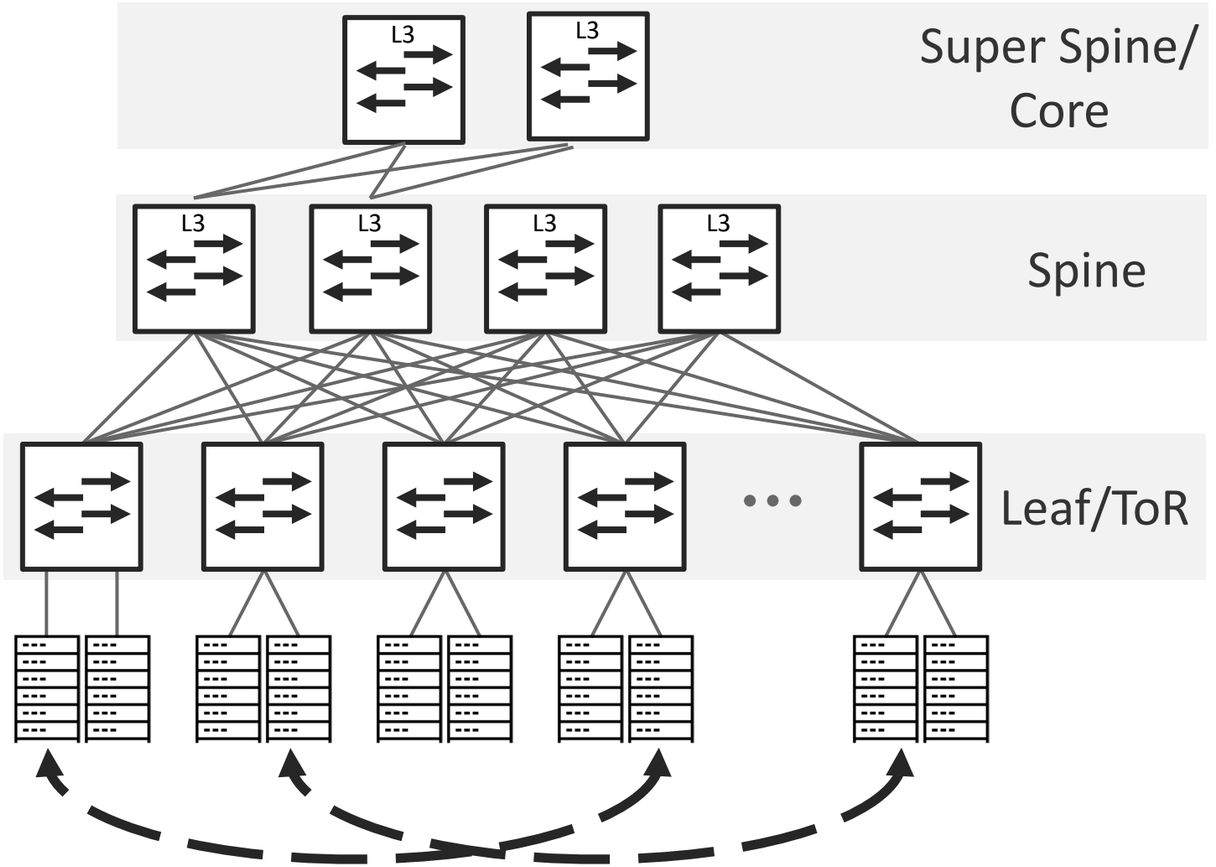
- In the past few years, many data centers have transitioned to a leaf-spine design.
- Leaf-spine architecture is a multilayer, full-mesh topology composed of a leaf layer and a spine layer.
- Leaf-spine design advantages:
  - Provides predictable and deterministic latency
  - Improves scalability
  - Improves redundancy
  - Increases bandwidth
  - Offers congestion avoidance



3-level Leaf-Spine Design

Modern data centers have transitioned to a leaf-spine design

- When the network reaches a certain scale, a third level of switches should be considered.
- The upmost level, called the super spine or core, is used to interconnect the spine switches.

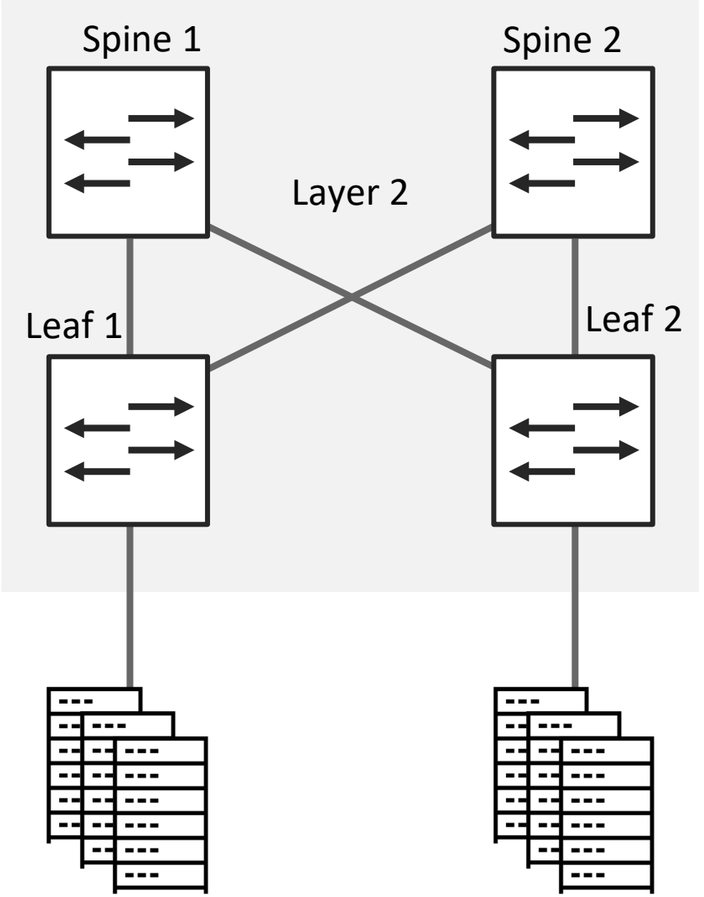


Layer 2 Networks

Layer 2 Networks

The networking devices are configured for switching

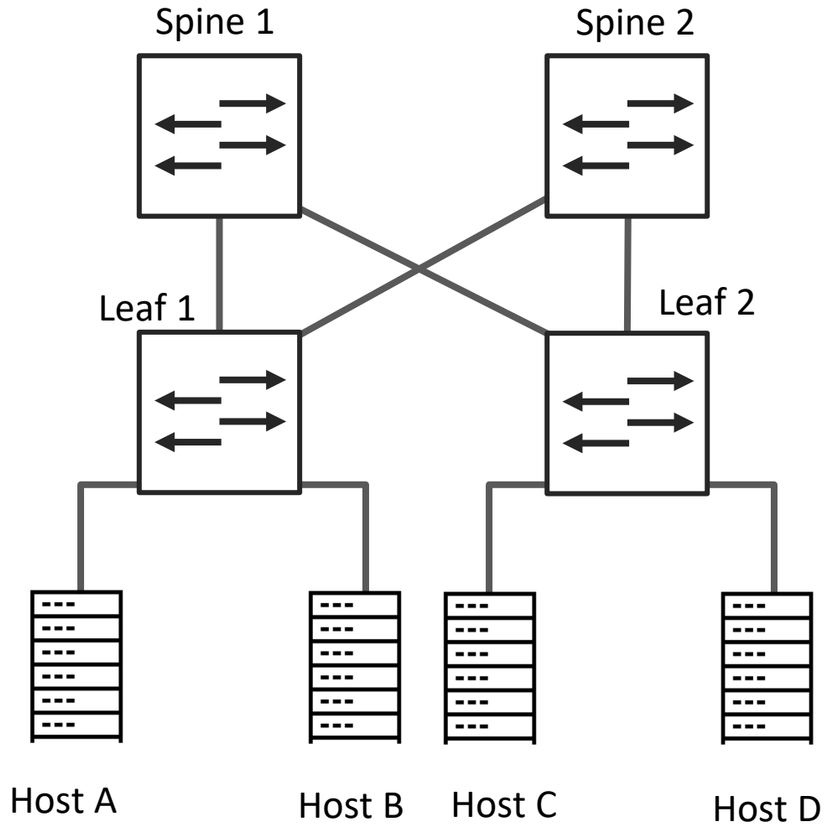
- In a layer 2 network, the networking devices are configured for switching
  - Build and maintain a MAC address table
  - Forward Ethernet frames based on the destination MAC address
- A layer 2 network compared to a layer 3 network provides:
  - Lower scalability
  - Lower bandwidth
  - Lower performance
  - Difficult to achieve redundancy and multipath support



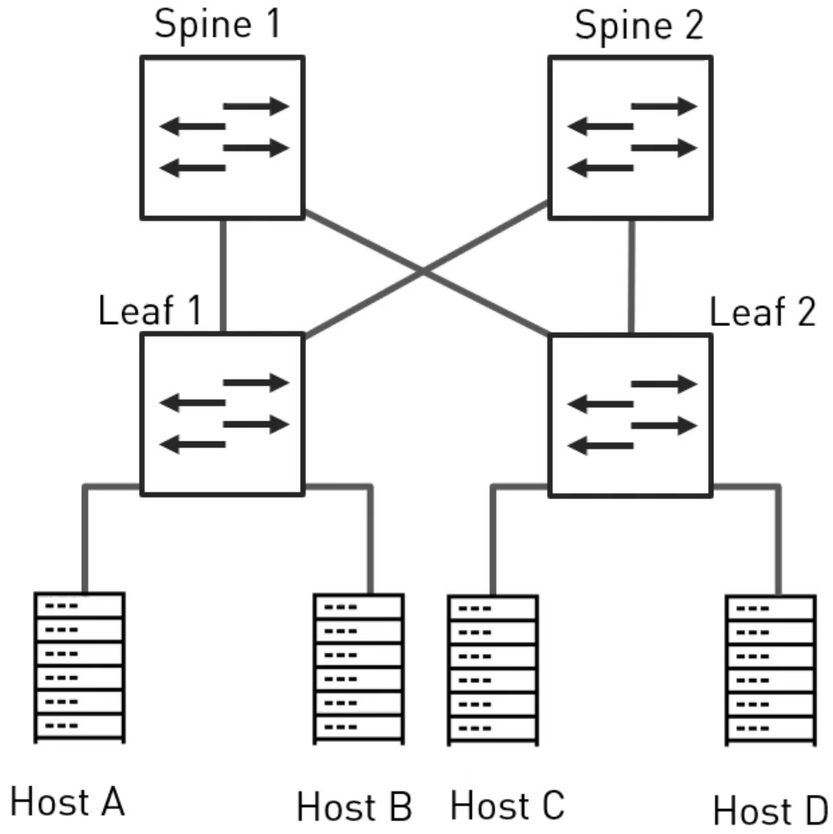
Redundancy and Multipath Support

Layer 2 Loops

- Redundant links are required to provide a backup path in case of a link or switch failure
- Redundant links allow multiple paths between a pair of nodes
- Redundant links result in layer 2 loops as there are multiple paths between a pair of nodes



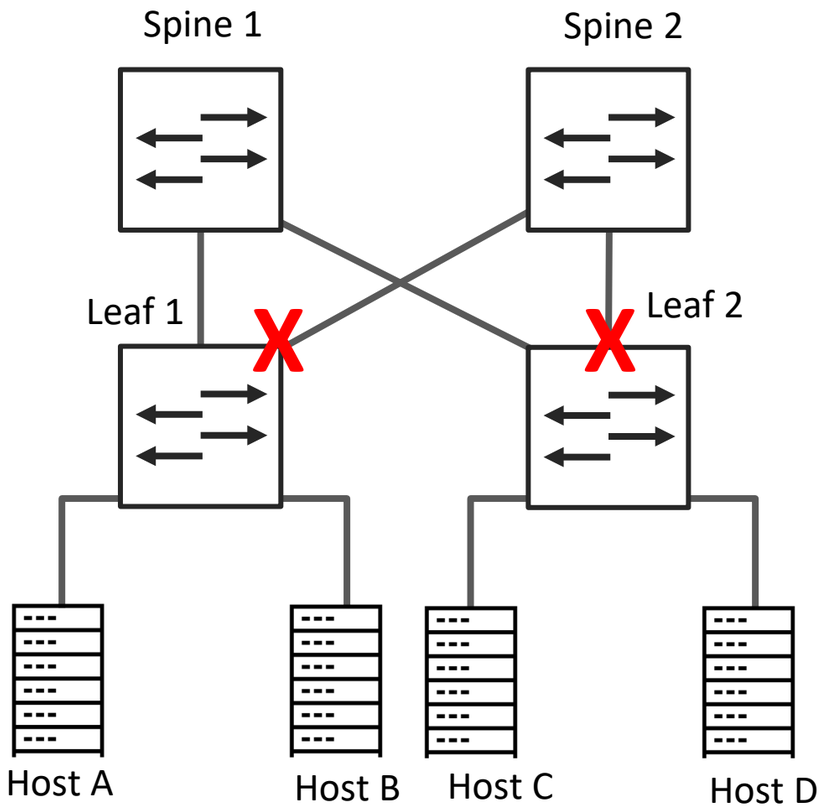
Broadcast Storms



Spanning Tree Protocol (STP)

STP ensures a loop-free logical topology for Ethernet networks

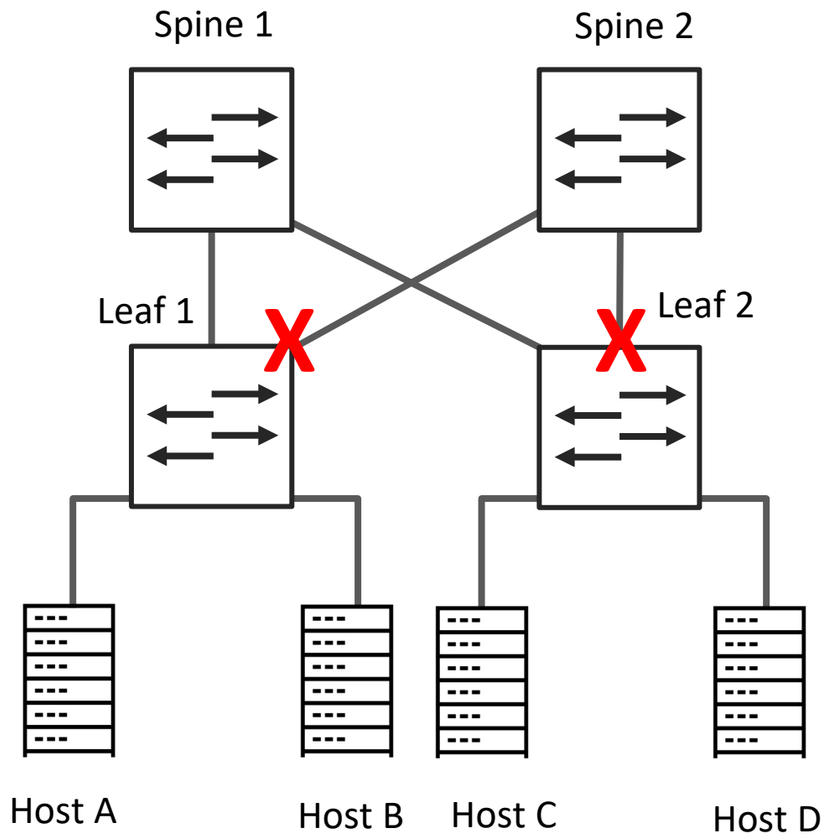
- STP allows a network design to include redundant links and to provide automatic backup paths, if an active link fails
- STP identifies redundant links and puts redundant ports in blocking state
- When a topology change occurs, STP reacts and moves blocked ports to forwarding state
- Convergence time is 30-50 seconds



Rapid Spanning Tree Protocol (RSTP)

RSTP provides rapid recovery after a topology change

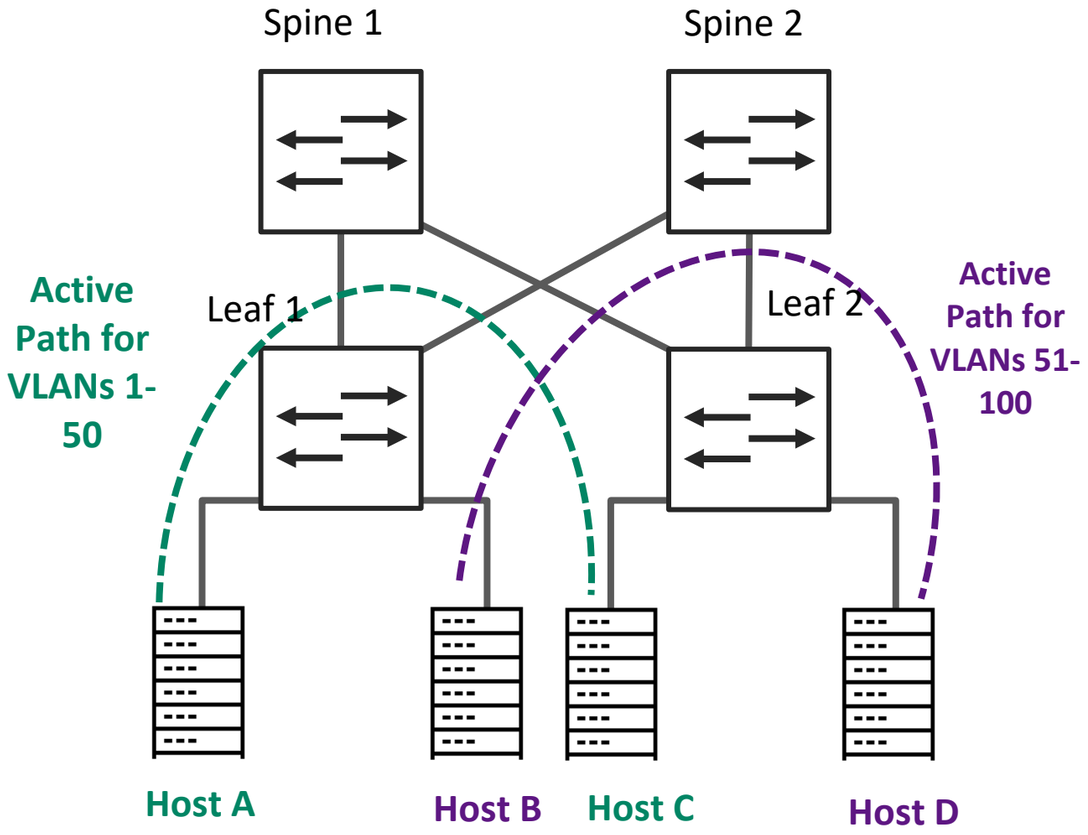
- RSTP switches react immediately to a topology change and start negotiating port states
- RSTP provides an alternate port that immediately moves to forwarding state, if the port becomes unavailable
- RSTP is typically able to respond to changes within a few seconds or as fast as a few milliseconds on a physical link failure



Multiple Spanning Tree (MST)

MST enables to configure multiple STP instances

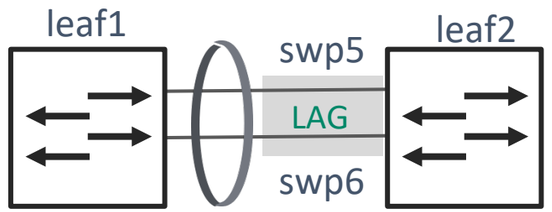
- MST enables configuring multiple STP instances, with a group of VLANs mapped to each instance
- Each instance may have a different active path
- MST provides redundancy, load-balancing, and multipath support, hence better overall network utilization
- Multiple MST instances are difficult to manage at scale



Link Aggregation Group (LAG)

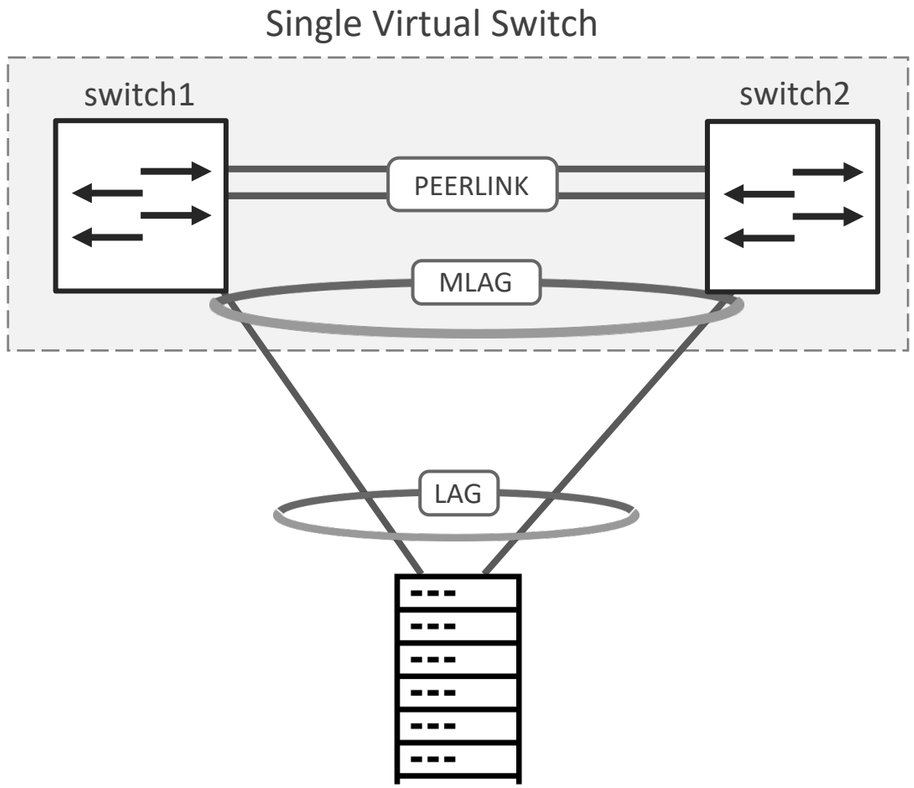
Link aggregation, LAG, port channel, bonding

- LAG enables the aggregation of multiple physical ports into one logical port
- The benefits of link aggregation include the following:
  - **Higher bandwidth** - LAG provides bandwidth that is the sum of the physical links' bandwidth
  - **Load balancing** - Traffic is distributed on all physical links, which are part of the LAG
  - **High availability** - When a physical link fails, other links of the LAG continue to serve the traffic



Multi Chassis LAG (MLAG)

- Physical ports of two separate switches are aggregated into one logical port
- MLAG switches appear as a single layer 2 switch
- The dual connected device (host or switch) runs a standard LAG and is unaware of the fact that its LAG is connected to two separate switches
- MLAG provides:
  - High bandwidth and load balancing
  - High availability in case of link failure
  - High availability in case of switch failure or switch software upgrade

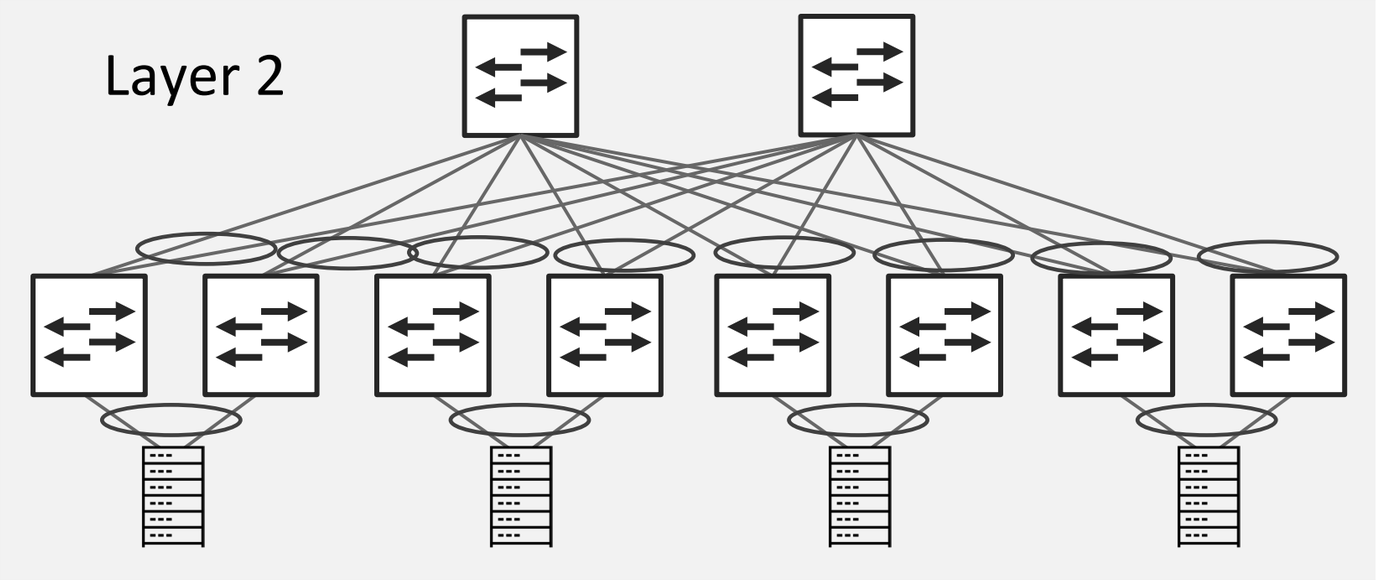


Layer 2 Network Design

Full layer 2 network design:

- MSTP for multipath
- MLAG for redundancy

- Pros:
  - Simple design, well understood, and easy to configure at a small scale
- Cons:
  - Layer 2 protocols are difficult to manage at scale
- Typical deployments:
  - Traditional enterprises with few networking needs

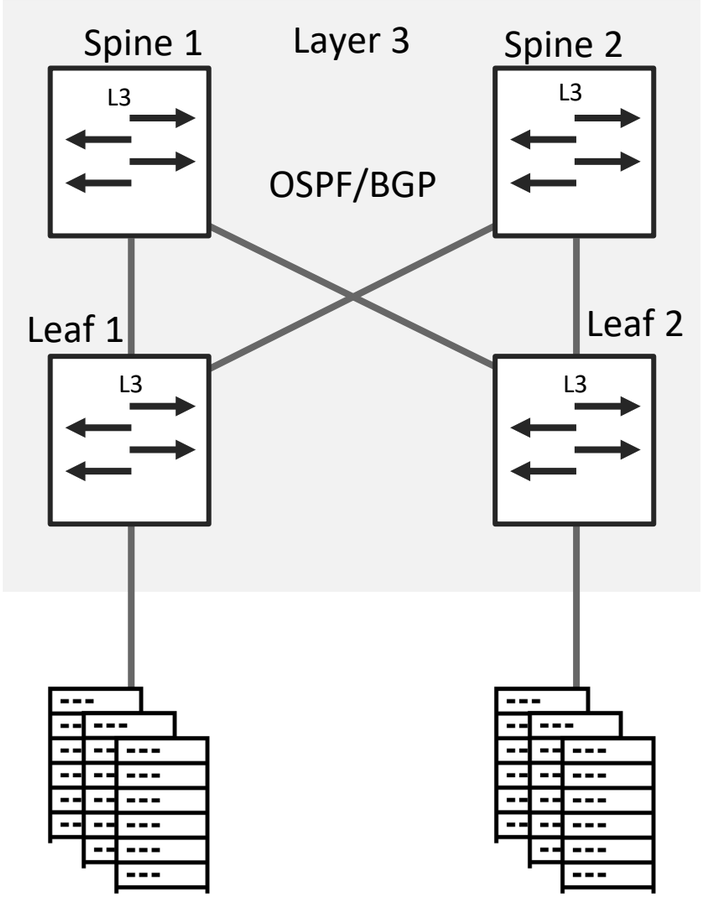


Layer 3 Networks

Layer 3 Networks

Routers or layer 3 switches configured for routing

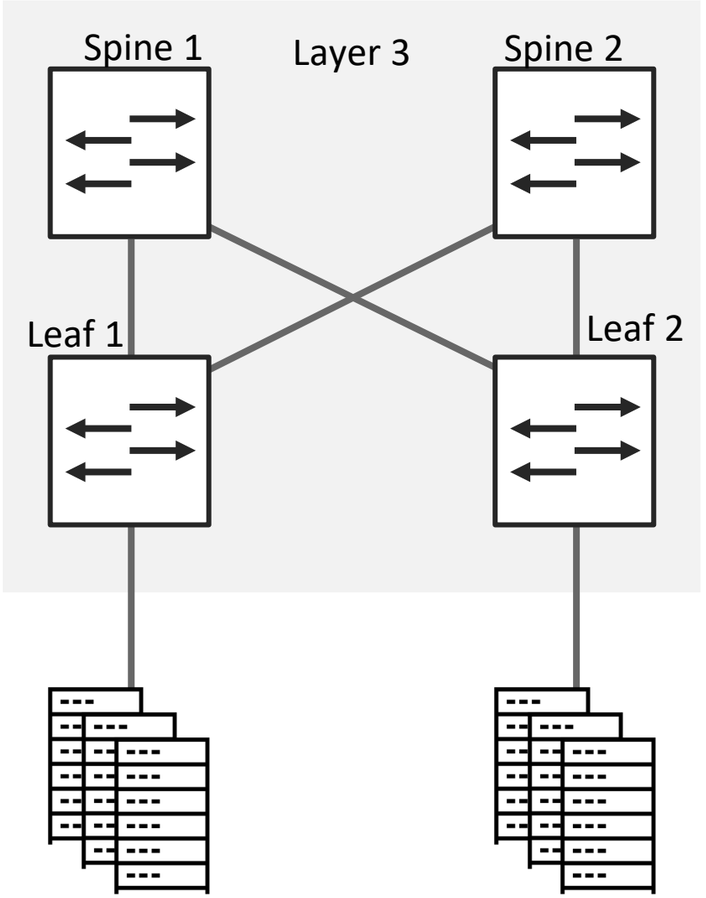
- Routers or layer 3 switches configured for routing
- Routing functionalities:
  - Build and maintain a forwarding database called a routing table
  - Include entries that map remote IP networks to next hop routers
  - Configured with static routes or a dynamic routing protocol, such as OSPF/BGP
  - Forward IP packets to the destination based on the routing table entries



Redundancy and Multipath Support

A layer 3 design provides redundant active-active links with no blocked ports

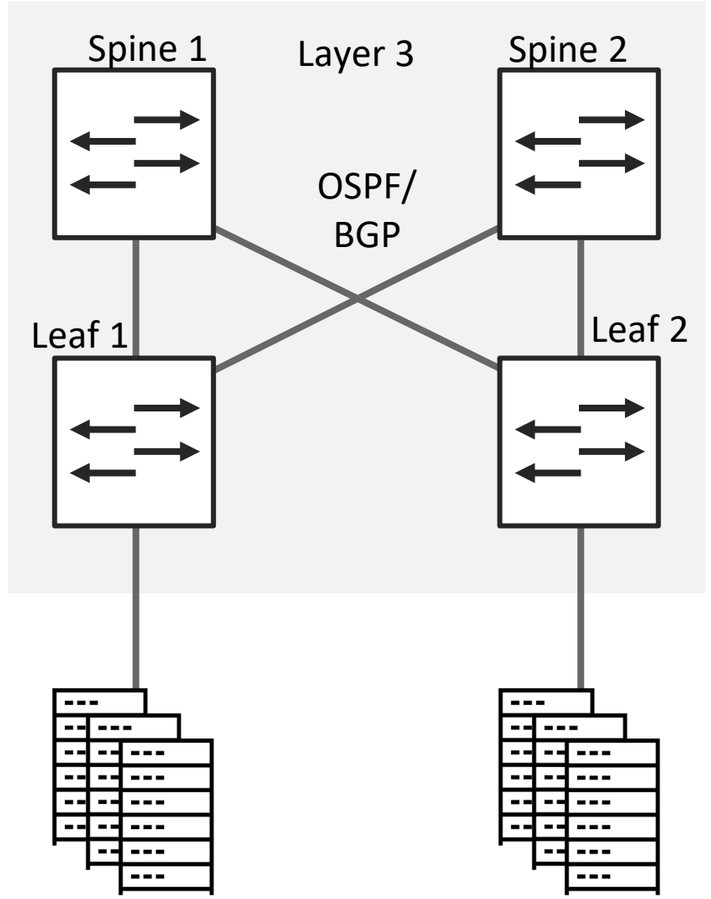
- Routing protocols identify and eliminate routing loops
  - Provide redundant active-active links with no blocked ports
- Equal Cost Multi Path (ECMP) is a routing strategy that allows packet forwarding to a single destination over multiple best paths
  - Multipath support
  - Load balancing
- IP packets carry a Time to Live (TTL) field
  - The number of hops (routers) the packet is allowed to traverse
  - Avoids packets endlessly circulating in a loop



Routing Protocols

An open-standard, scalable, and robust protocol should be chosen

- Dynamic routing protocols:
  - Open Shortest Path First (OSPF)
  - Border Gateway Protocol (BGP)
- BGP is the routing protocol of choice in modem data centers



BGP in the Data Center

BGP is an increasingly popular protocol for use in the data center

- BGP is the routing protocol that makes the Internet work
  - Allows service providers to exchange routing and reachability information
  - Designed for scalability and stability in very large networks
- BGP is an increasingly popular protocol for use in the data center as it lends itself well to the rich interconnections in a leaf-spine topology
- The following are some notable BGP properties:
  - Does not require routing state to be periodically refreshed (unlike OSPF)
  - Has many robust vendor implementations
  - Thoroughly tried and tested protocol that comes with years of operational experience

Layer 2/3 Network Design

Layer2/3 network design:

- Layer 2 from hosts to ToRs with MLAG for redundancy
- Layer 3 network with ECMP for redundancy and load balancing

Pros:

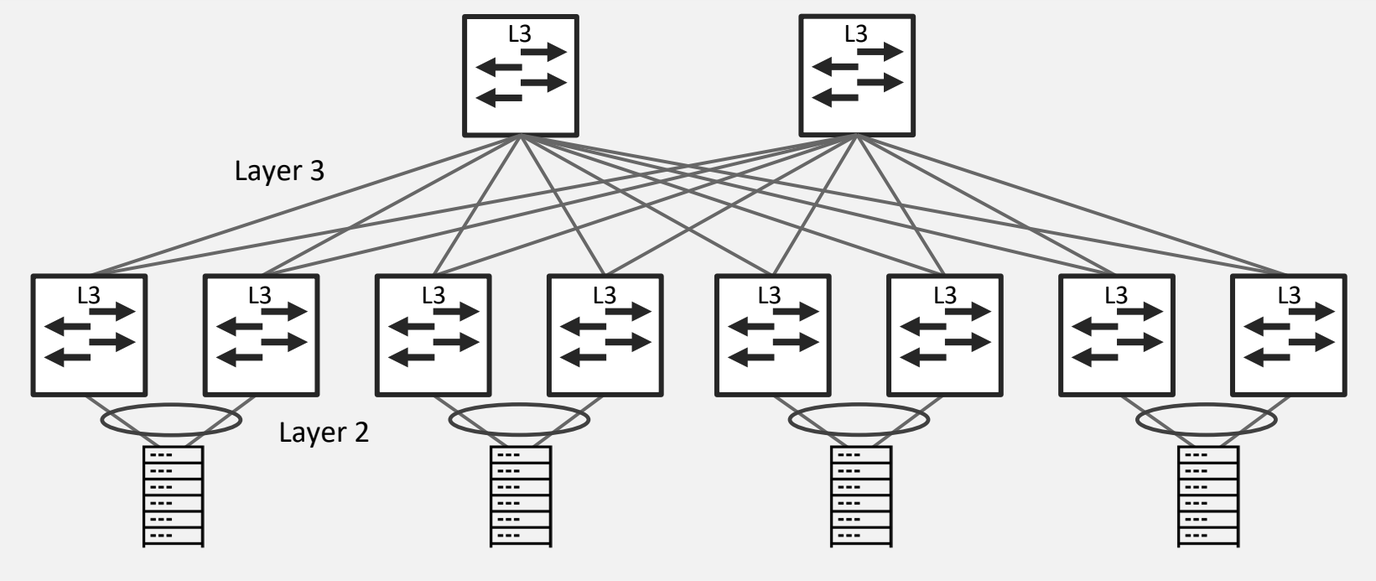
- The designs are simple, well understood, and easy to configure

Cons:

- Still uses layer 2

Typical deployments:

- Data centers and big enterprises
- High-scale bare metal hosts
- High-scale virtualized hosts



Full Layer 3 Network Design

Network design:

- Full layer 3 from network to hosts with ECMP for redundancy even at the host level

Pros:

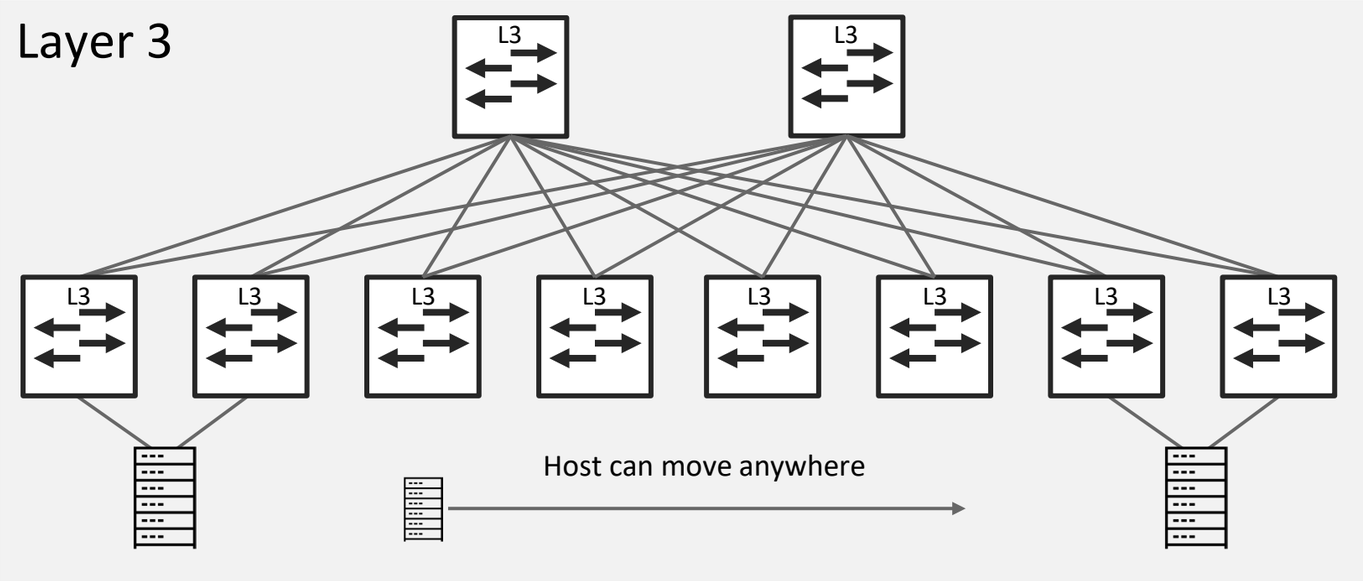
- Full layer 3 – highly redundant
- Designed for high scale and virtualization
- Fewer protocols – no MLAG, no STP

Cons:

- A routing engine is required on hosts - RoH

Typical deployments:

- Highly virtualized environments with a high degree of mobility



Summary

Data Center Design

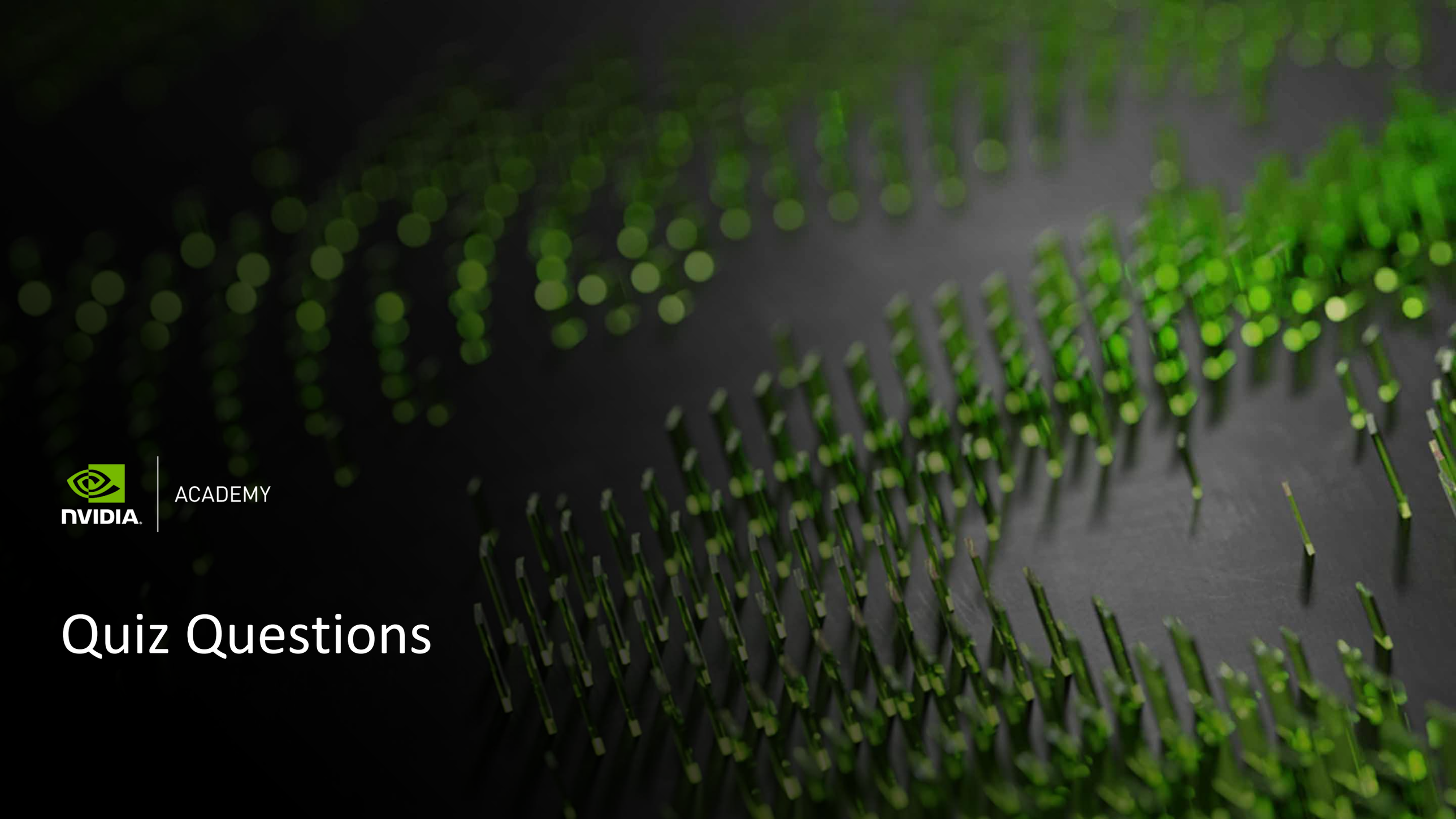
- Network Topology
- Hierarchical Network Design
- Leaf-Spine Design

Layer 2 Networks

- Layer 2 Protocols
- Layer 2 Network Design

Layer 3 Networks

- Layer 3 Protocols
- Layer 3 Network Design

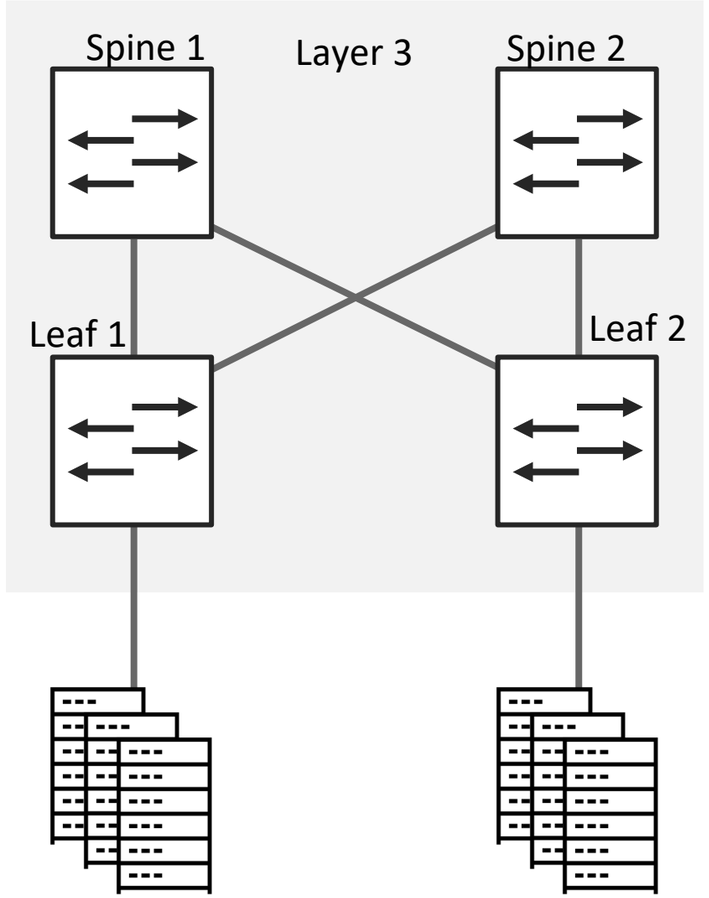


Quiz Questions

Examine Your Knowledge 1

In a **layer-3 data center** design how can you achieve multipath support?

A. BGP with ECMP can be configured
B. A layer-3 design cannot support multipath
C. STP with ECMP can be configured
D. An MSTP instance per VLAN group can be configured



Examine Your Knowledge 3

BGP is an increasingly popular protocol for use in the data center thanks to the following properties: (Select two)

A. Periodic updates that allow fast convergence
B. Scalable and stable
C. Multipath support
D. BGP was specifically designed for leaf-spine topologies.

Examine Your Knowledge 4

Which two of these are used to prevent loops in a layer 2 network? (Select two)

A. STP
B. BGP
C. MLAG
D. RoH
E. ECMP

Examine Your Knowledge 5

What is the purpose of the Time To Live (TTL) field carried in an IP packet?

A. Do determine the number of routers the packet is allowed to traverse
B. Do determine what is the trip time allowed for a packet till it gets to the destination node
C. To avoid packets from endlessly circulating in a loop
D. Do record the time it takes for a packet to get to the destination node

Summary Exam

Examine Your Knowledge 2

In a **layer-2 data center design** how can you achieve multipath support?

A. BGP with ECMP can be configured
B. A layer-3 design cannot support multipath
C. STP with ECMP can be configured
D. An MSTP instance per VLANs group can be configured