

Network Design Principles

Network Design Principles

Carrier grade

IP Address Management



Content

- Dynamic Routing Protocols
 - BGP
 - OSPF
- Logical Network Design
 - Switched
 - Statically Routed
 - Dynamically Routed

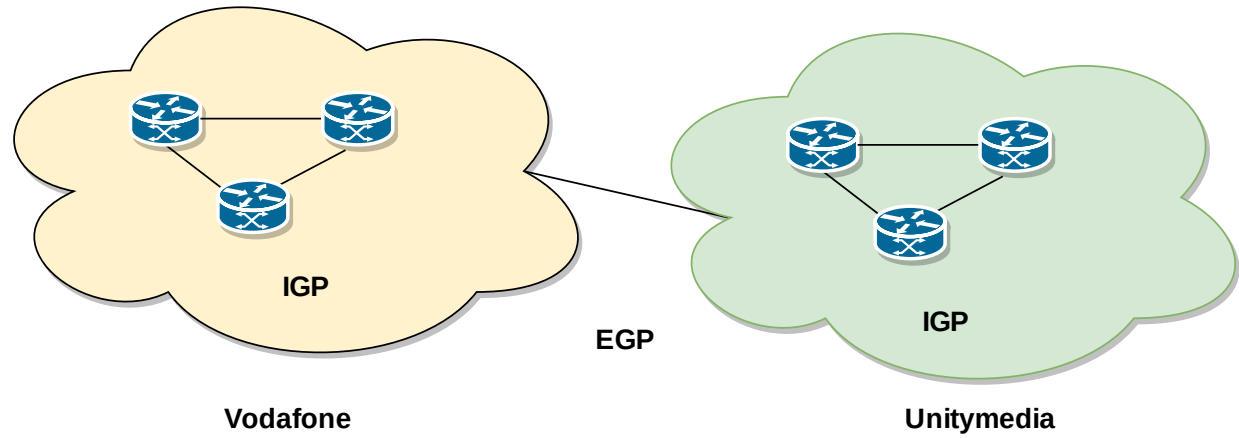
Dynamic Routing

- Why?

Dynamic Routing

- Handle failover
- Discover remote networks
- Choose the best path to the destination
- Lower configuration costs
- Lower failure points?

Dynamic Routing



- IGP = Interior Gateway Protocol
- EGP = Exterior Gateway Protocol

Dynamic Routing

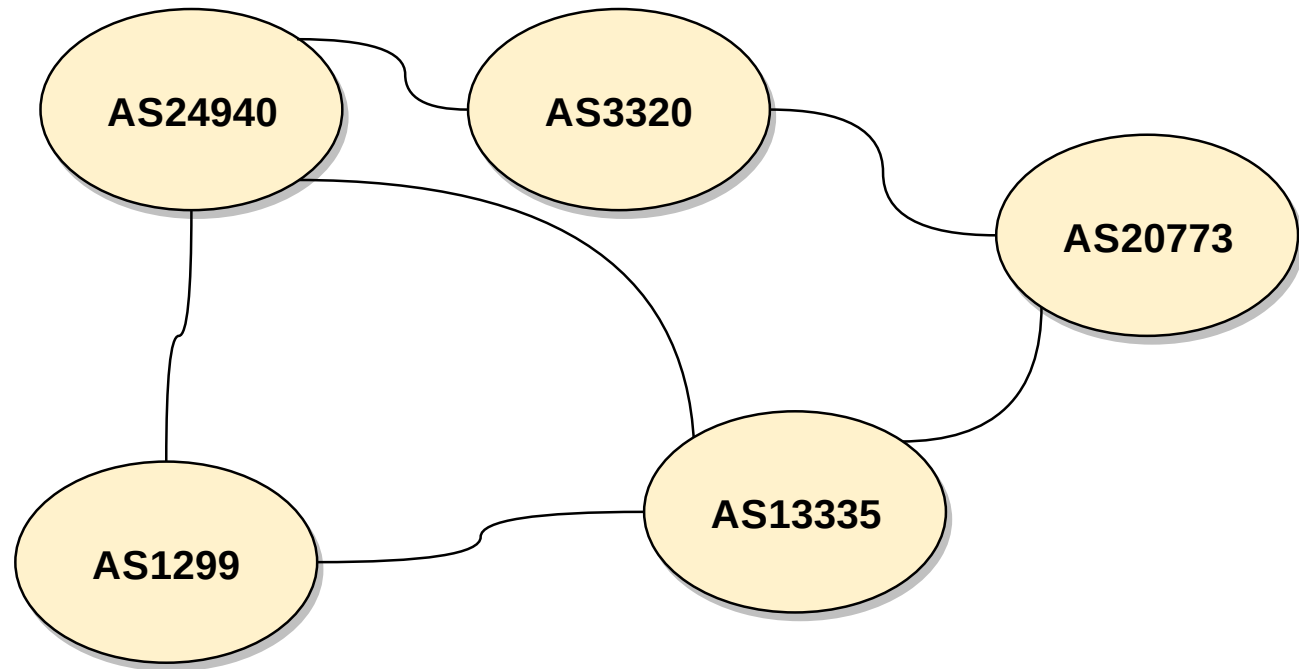
BGP

- Border Gateway Protocol
- Distance Vector Routing Protocol
- EGP
- Aggregates all vectors to the most simple one
- Pushes updates based on events
- Only talks to direct peers
- Prevents routing loops by checking the AS-Path
- Announces the configured routing table

Dynamic Routing

BGP

Distance Vector Routing

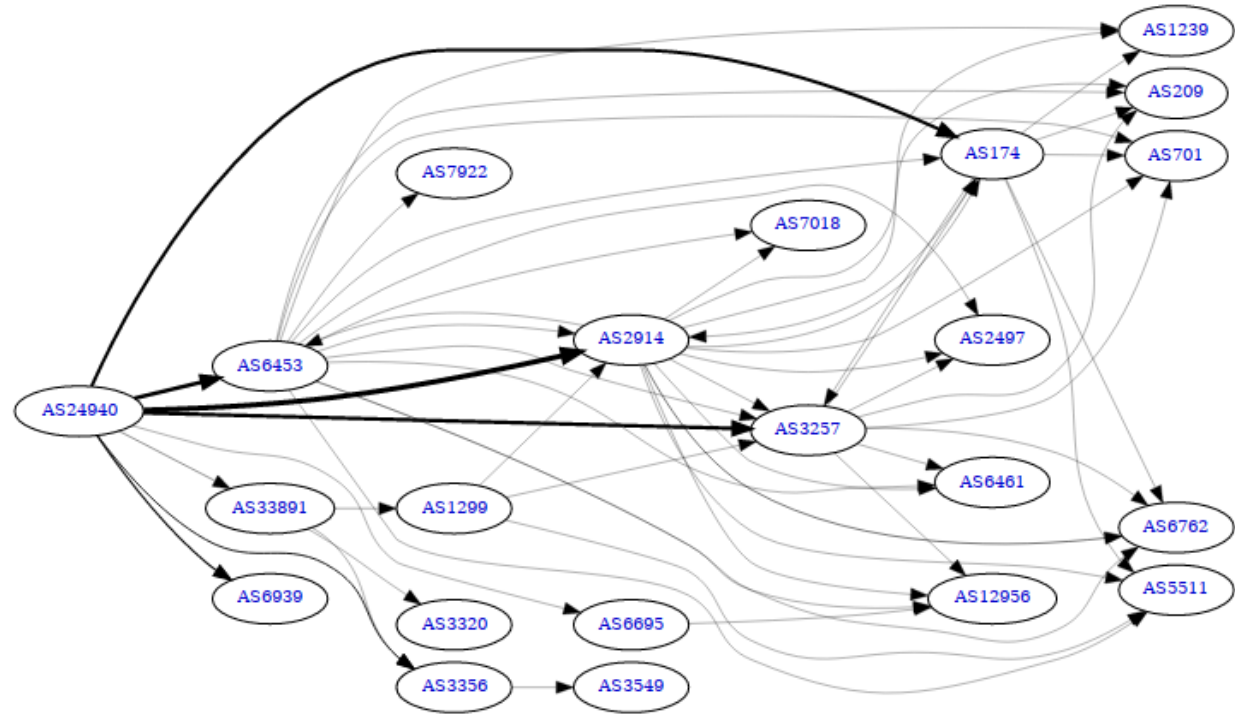


Element	Vector
-----	-----
AS24940	AS3320
AS24940	AS3320 -> AS20773 -> AS13335
AS24940	AS3320 -> AS20773 -> AS1299
AS24940	AS1299
AS24940	AS1299 -> AS13335

Dynamic Routing

BGP

AS24940 IPv4 Route Propagation

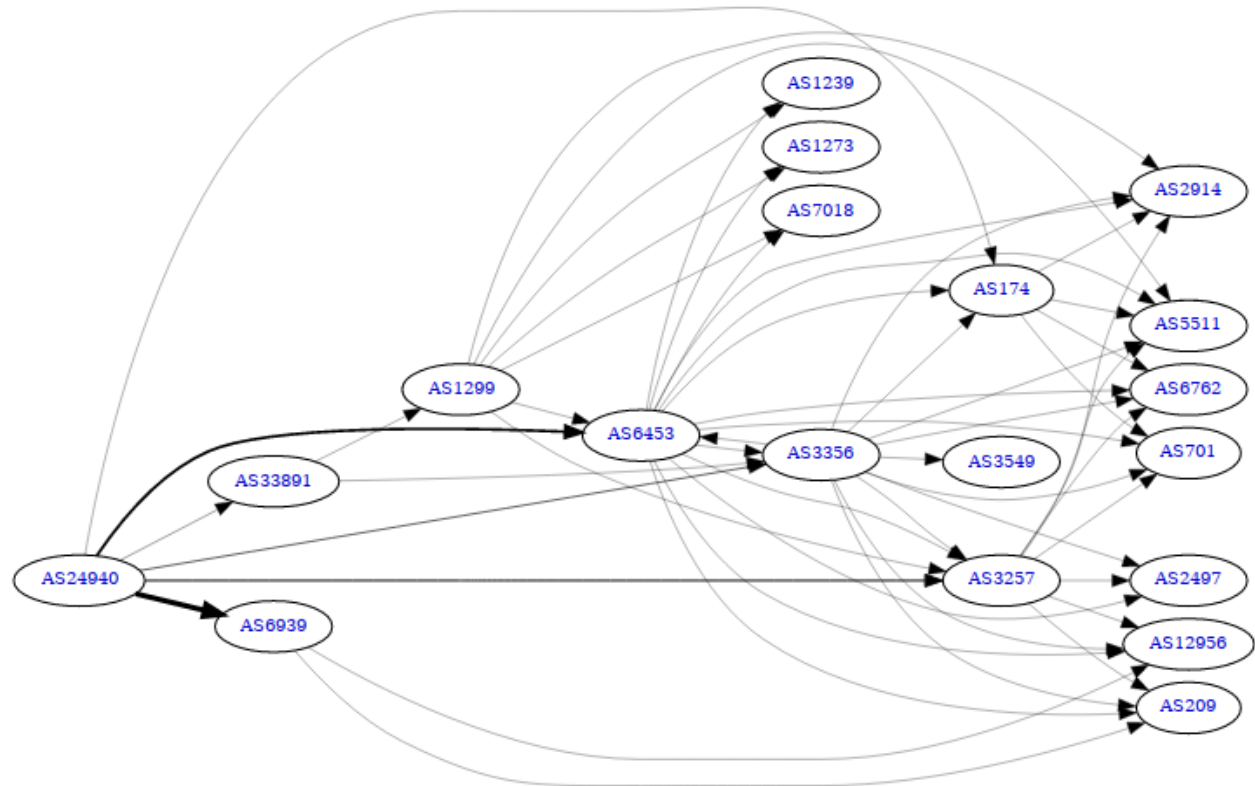


© http://bgp.he.net/AS24941#_asinfo

Dynamic Routing

BGP

AS24940 IPv6 Route Propagation



© http://bgp.he.net/AS24940#_asinfo

Dynamic Routing

BGP

Route Selection

- Ignores Link-State
- Ignores Hop Count
- Honors AS-Path
- Honors proprietary values

Dynamic Routing

BGP

OSPF

- Open Shortest Path First
- Link-state routing protocol
- IGP
- Interval based

Dynamic Routing

BGP

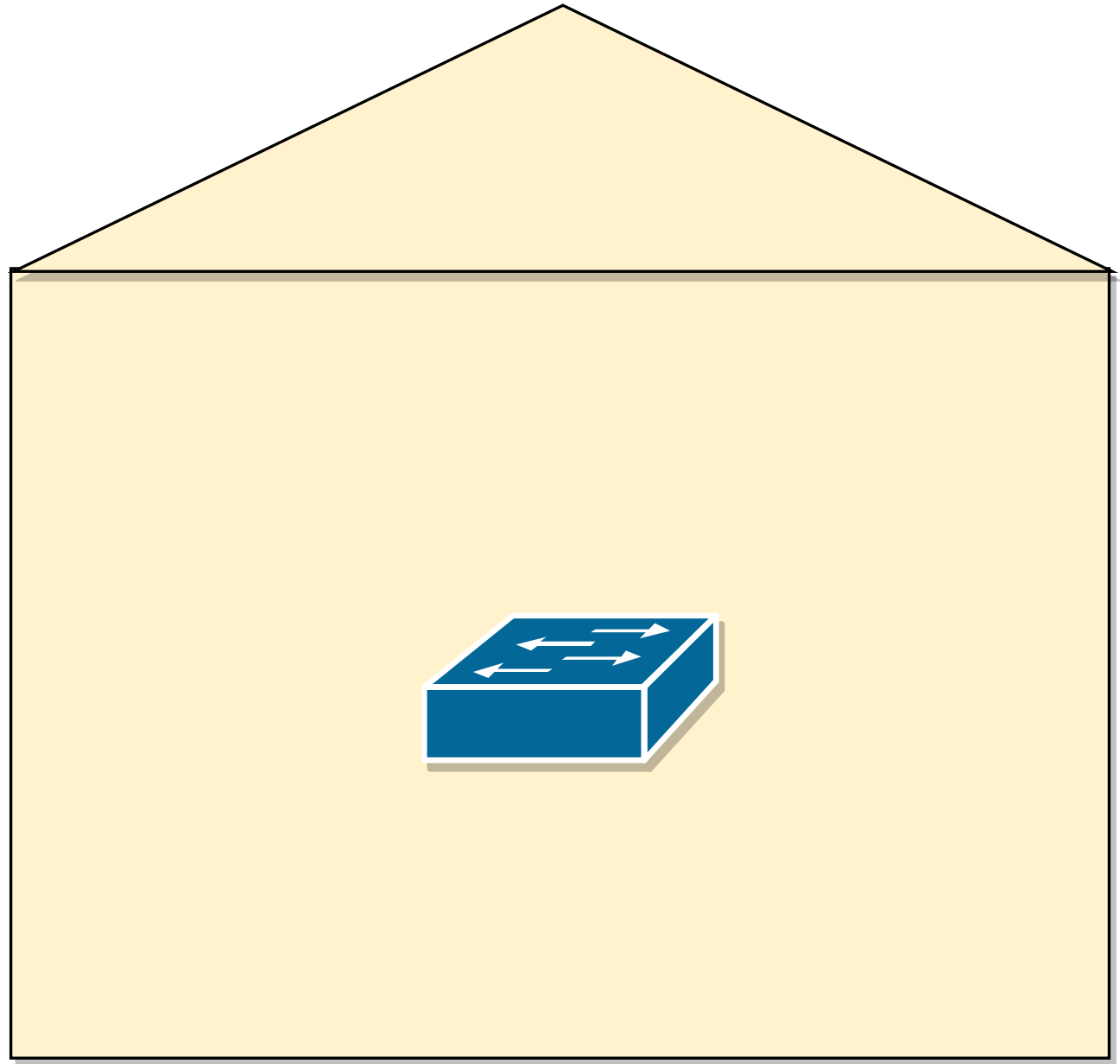
OSPF

Link-state routing

- Every router in an AS talks to every other
- Uses SPF to create a tree
- Periodically sends a flood
 - Causes a huge amount of traffic

Logical Network Design

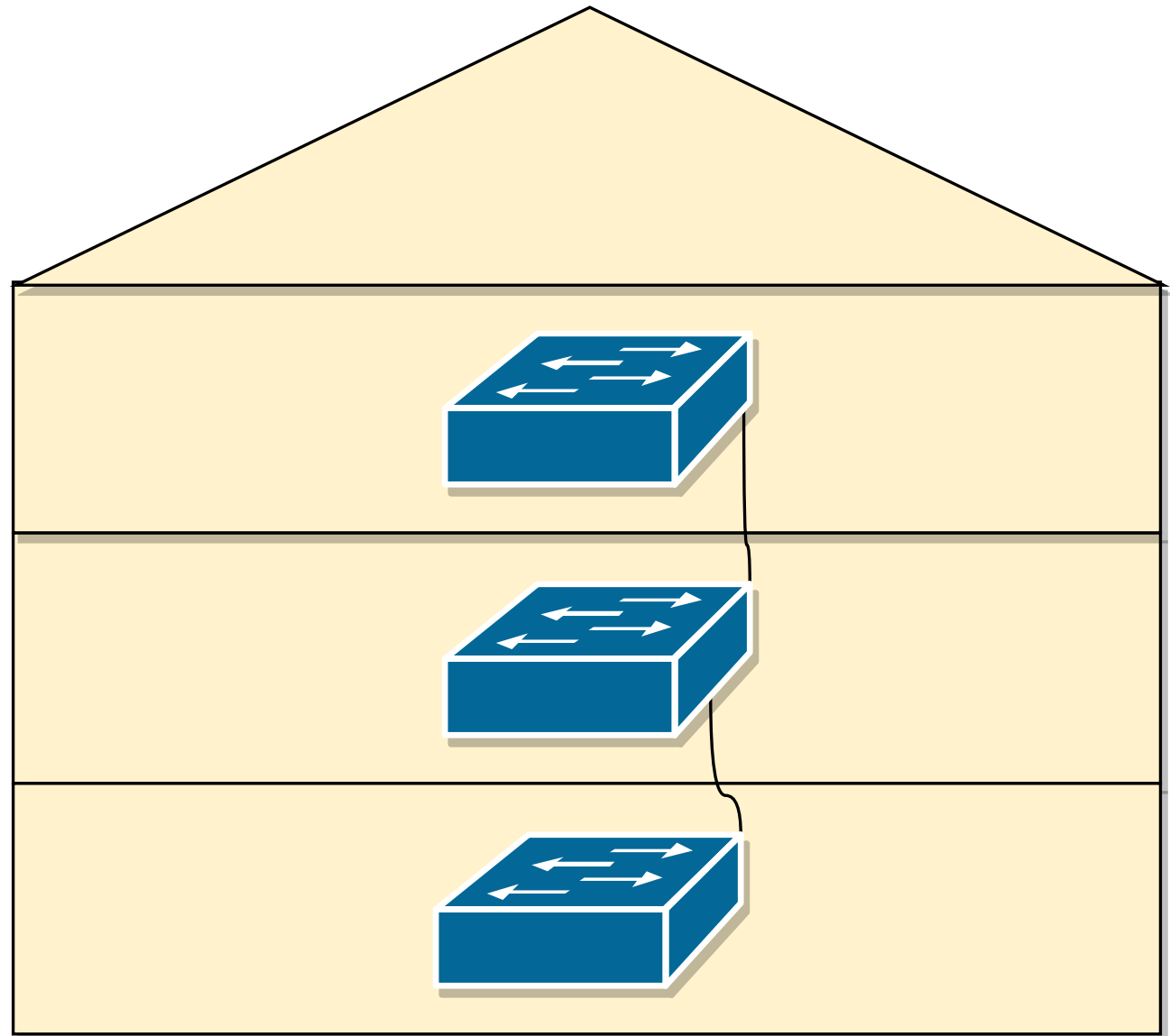
Switched



One Classroom in one house

Logical Network Design

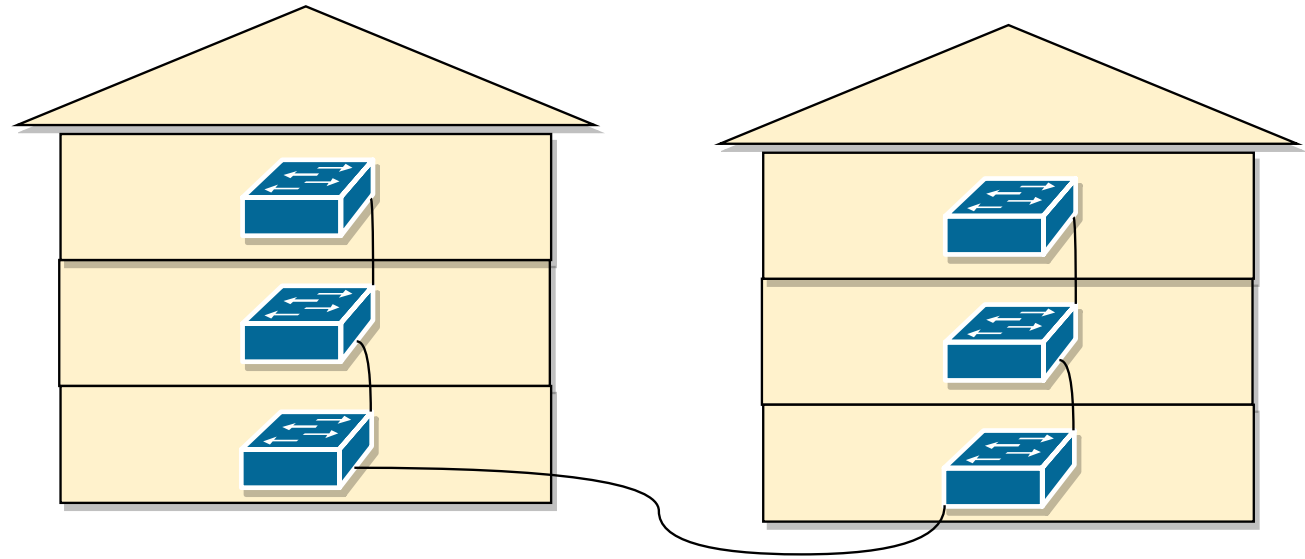
Switched



Three classrooms in one house

Logical Network Design

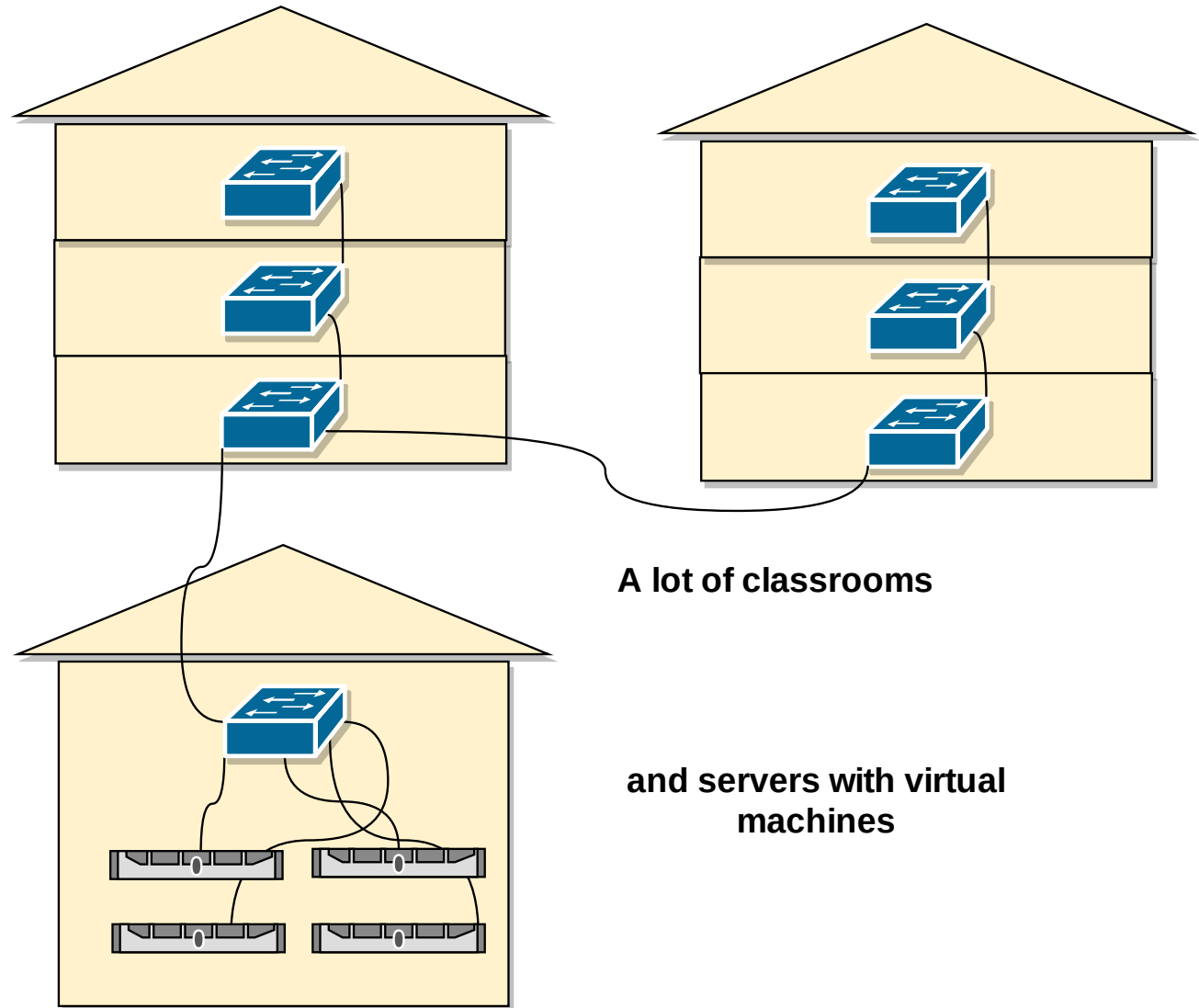
Switched



A lot of classrooms

Logical Network Design

Switched



Logical Network Design

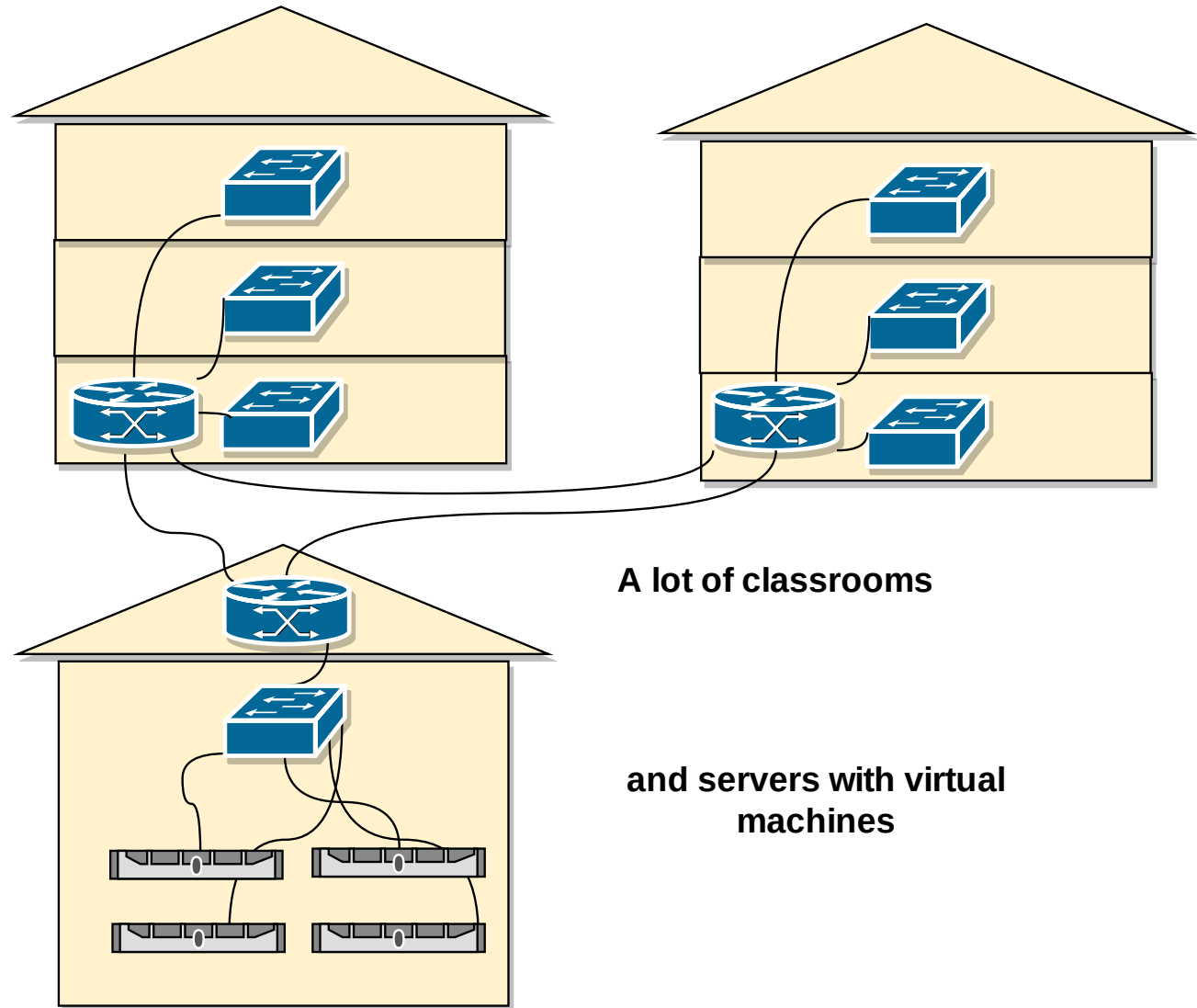
Switched

- Many SPOF, unable to mesh
 - would create loops
- CAM table size is limited
 - Content Addressable Memory = MAC table
- All switches hold up all MAC addresses in the same Layer2/VLAN
 - limit is around ~512-4096 addresses per device
 - Juniper EX2200-48T: 16.000 entries, 864€
 - Juniper EX3200-48T (EOL): 32.000 entries, 1975€
 - TP-Link T3700G-28TQ: 32.000 entries, 1698€
 - Netgear GS724Tv4: 16.000 entries, 150€

Logical Network Design

Switched

Statically Routed



- One Subnet for each switch

Logical Network Design

Switched

Statically Routed

- Static routes configured on all switches
- Works fine on Layer 2 until there are VLANs across routers
 - CAM tables will again be too big
- Requires manual configuration for routes
- Redundant ring

Logical Network Design

- Less stuff to configure by hand
- Better failover for new links
- TOR switch for hypervisors is still a bottleneck
- We could use OSPF or RIP

Switched

Statically Routed

Dynamically
Routed



Logical
Network
Design

Switched

Statically Routed

Dynamically
Routed

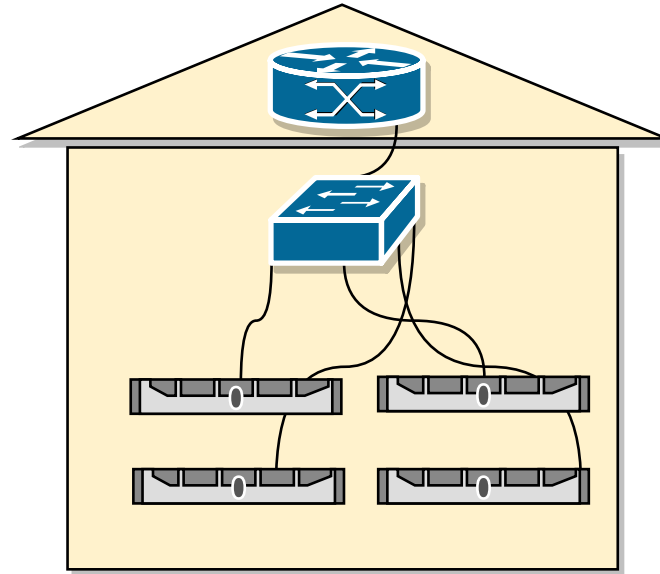


Logical Network Design

Switched

Statically Routed

Dynamically
Routed



**and servers with virtual
machines**

Logical Network Design

Switched

Statically Routed

Dynamically
Routed

Figure 1 Topology of Cisco MSDC Design Evolution—Phase 1

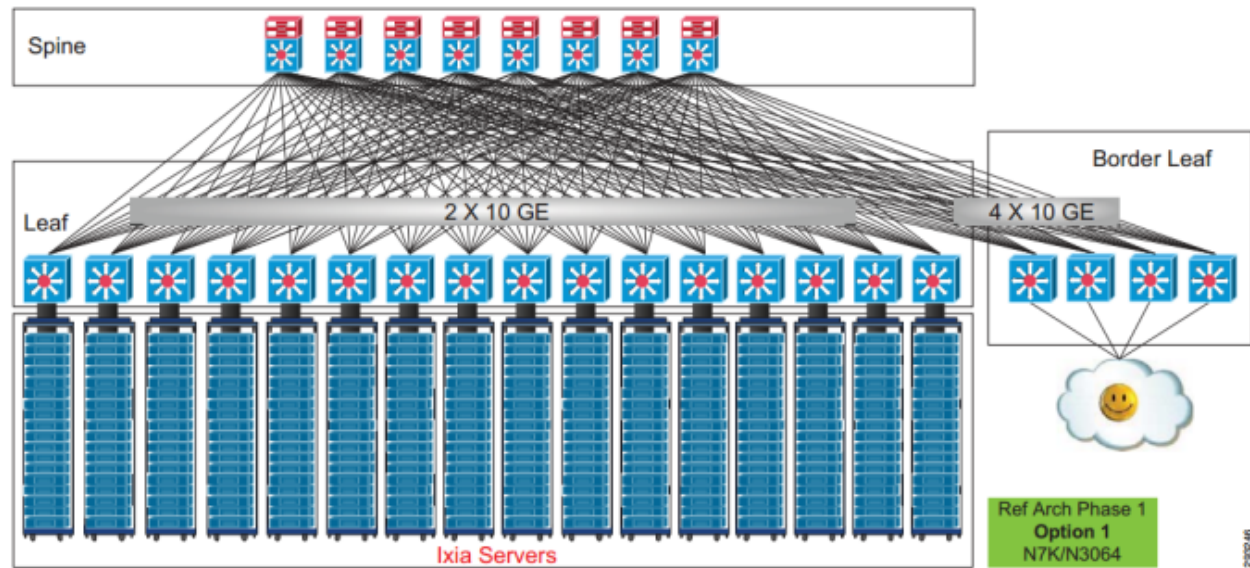


Figure 2 Topology of Cisco MSDC Design Evolution—Phase 2

- 16 Racks
- 16 Switches
- 8 Routers
- $16 \times 40 = 640$ Servers
- $640 \times 100 = 64,000$ Virtual Machines

Logical Network Design

- Impossible to have 64.000 Virtual Machines + Hypervisors in the same Layer 2
 - Would require switches with 64,000 MAC entries
 - Juniper EX92008, 8RU, ~45.000€, per rack

Switched

- Dynamic routing protocols route prefixes to a next hop
- Normally used to route nets to a router

Statically Routed

- You can route to any IP address?

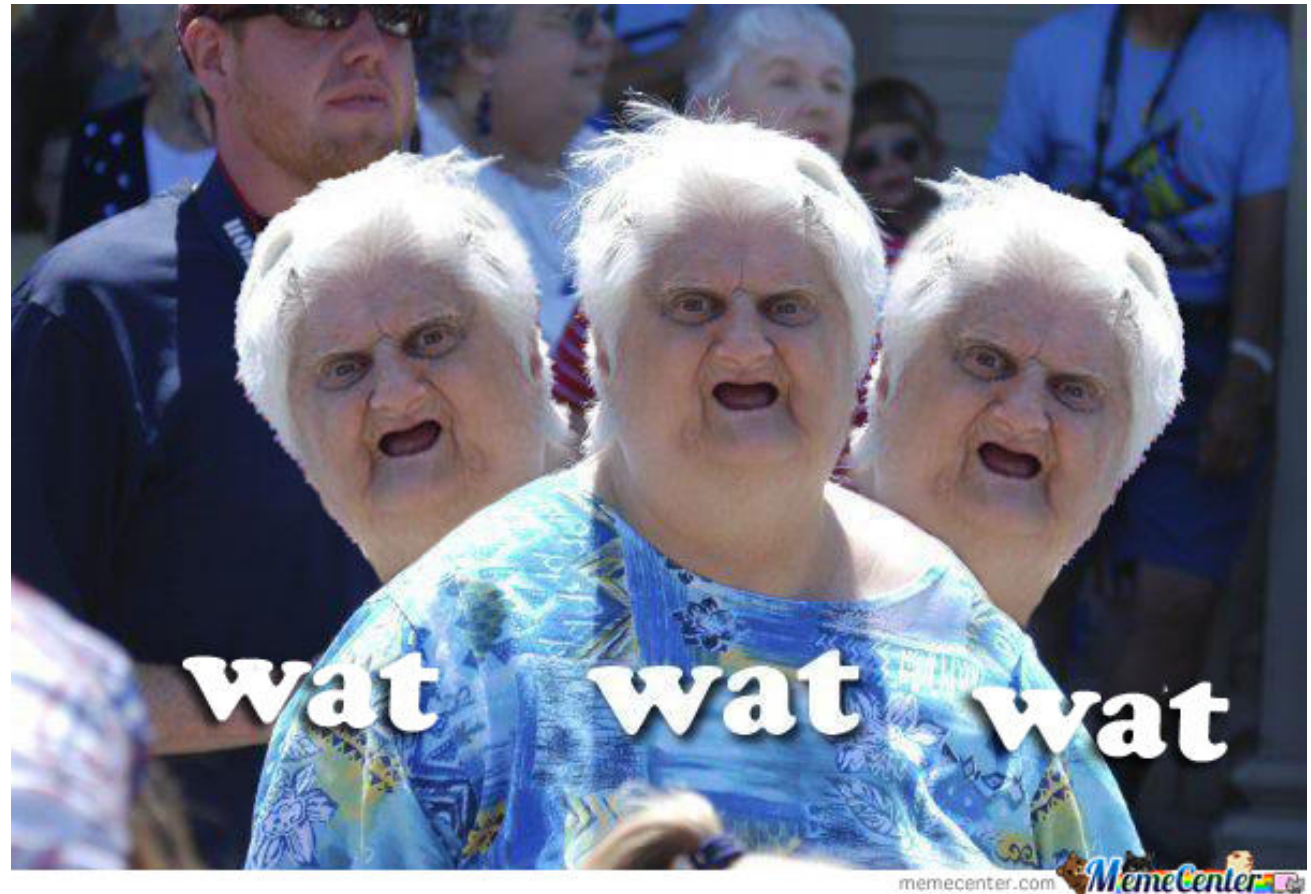
Dynamically Routed

Logical
Network
Design

Switched

Statically Routed

Dynamically
Routed



Logical Network Design

Switched

Statically Routed

Dynamically Routed

BGP as IGP

- Bad: Every router would talk to every other router (fully meshed)
 - Fine for small networks, doesn't work in this size
 - Each hypervisor will act as a router
- Bad: Needs something to get the routing table
 - "Announces the configured routing table" - Slide 7
- Nice: Simple Table of: Sourcenet -> Next Hop

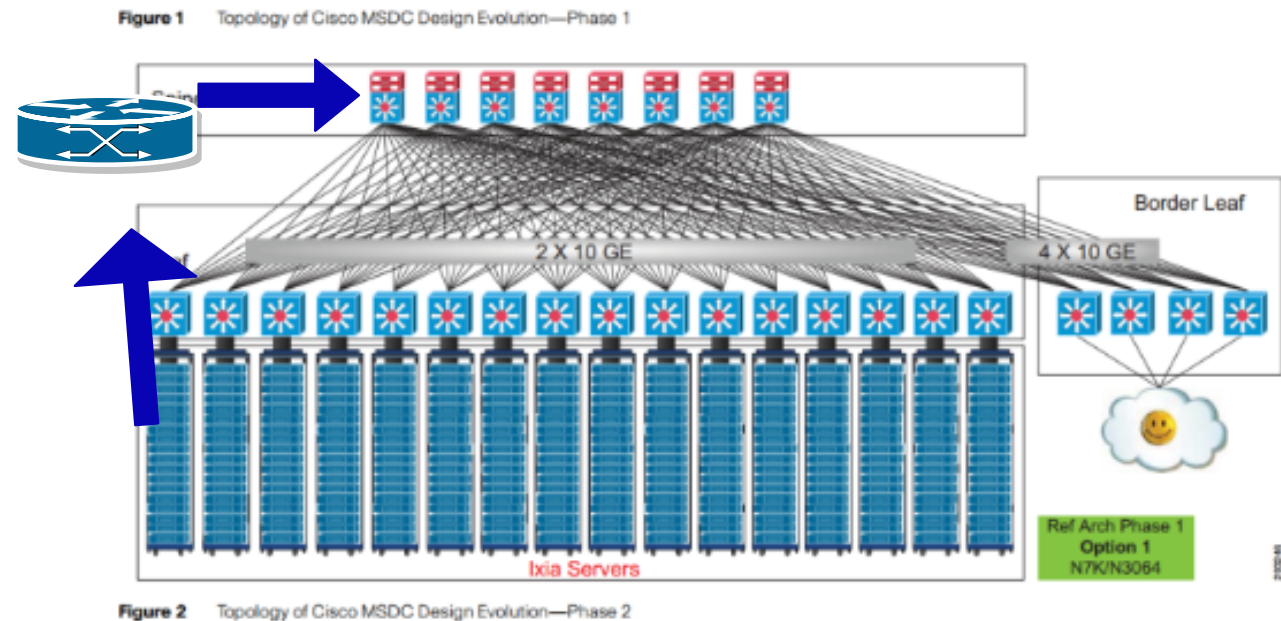
Logical Network Design

Switched

Statically Routed

Dynamically
Routed

BGP route reflector



- Everybody sends updates to a central router
- Router pushes updates to all routers that need this information

Conclusion

- Dynamic routing is fun
- OSPF is fine for announcing hypervisor networks
- BGP is perfect for cloud
- Don't trust vendors for CAM table size