

# Midsize Datacenter Design

Presented by:

Jason Bomar, CCIE #9316



# Introduction

- \* Network engineer for almost 20 years
- \* Focus is on R/S and DC infrastructure solutions
- \* Cisco VAR
- \* More Enterprise experience than midsize business, but plenty of each

# Midsized? WTF is “midsized”

- \* Your DC is big enough that you are willing to invest in switches dedicated to JUST the DC.
- \* No sharing L2 domains between users and servers.
- \* There is an expectation of growth in the DC
- \* You are interested in, and possibly implementing, some DC-only features:
  - \* VPC
  - \* FabricPath
  - \* FCoE
  - \* 10Gb/40Gb/100Gb

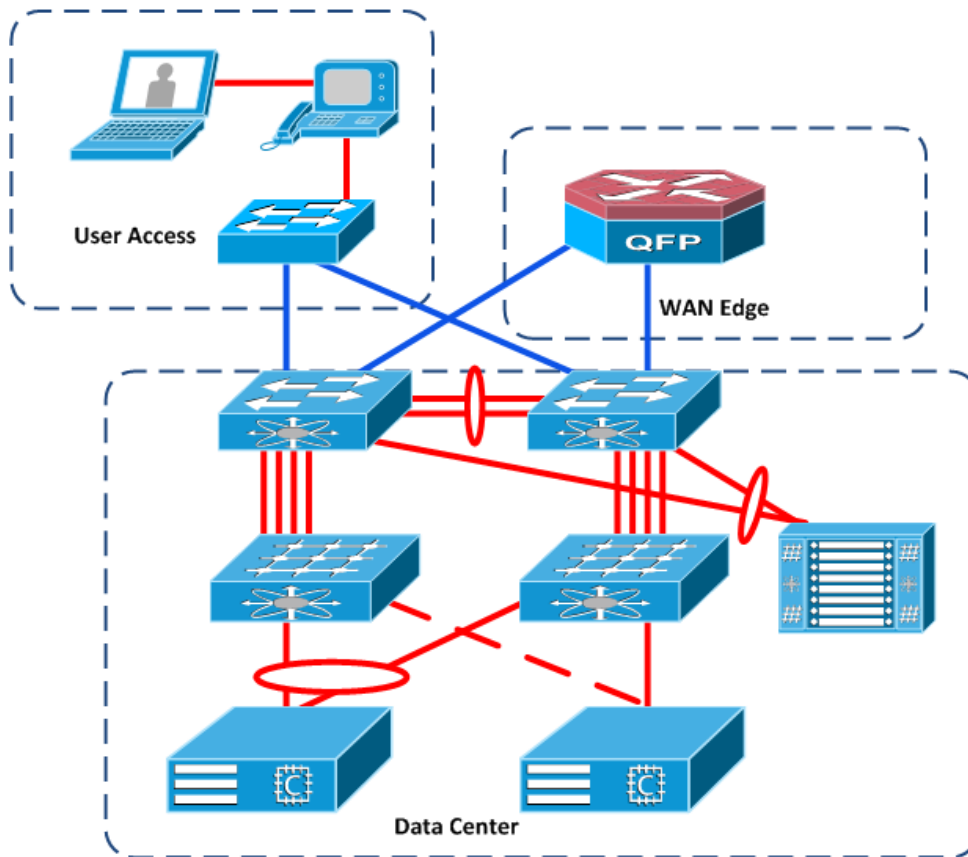
# Housecleaning

- \* Let's have a frank discussion on design:
  - \* You need to invest in scalability
  - \* You need to understand YOUR apps and YOUR business
  - \* You need to understand where to set L2/L3 boundaries and why
  - \* IP addressing
  - \* Advanced features
  - \* What are YOUR goals for the network?
    - \* Uptime
    - \* Growth
    - \* Redundancy

# Terms and Definitions

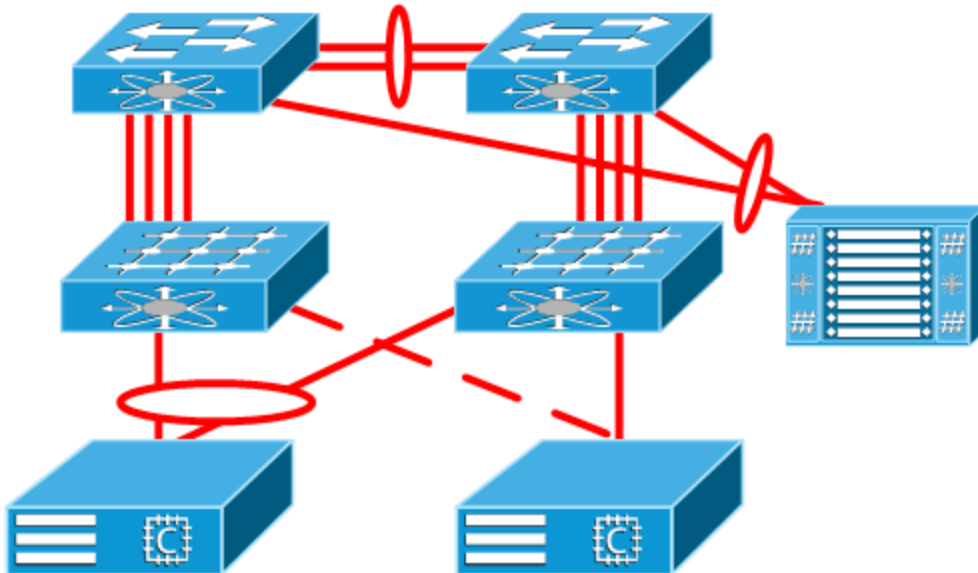
- \* FEX – Fabric Extender, this is a Nexus 2000. It is NOT a switch, but more akin to a line card in a chassis ... just minus the chassis.
- \* VPC – Multichassis Etherchannel within the DC.
- \* FabricPath – Removal of Spanning tree from the L2 DC network, introduction of the leaf and spine.
- \* Storage – This refers to Fibre Channel, FCoE, iSCSI, NFS, etc.

# The POD concept



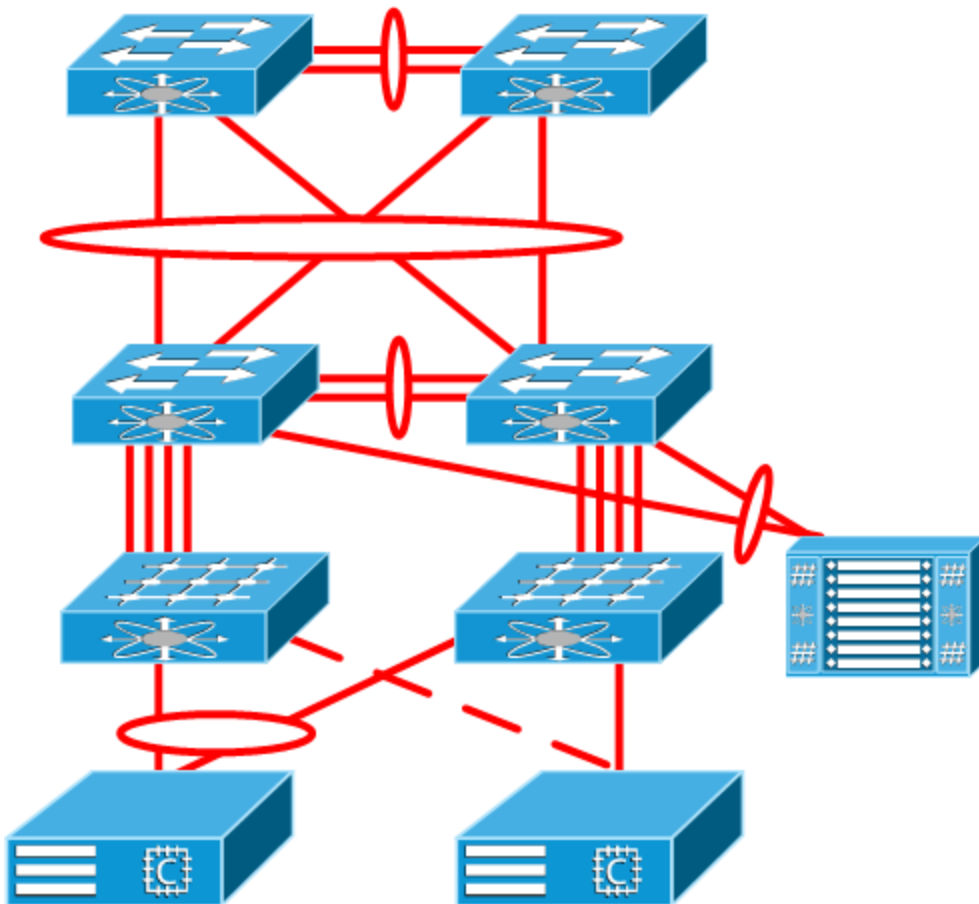
- The entire network is broken into “modules” or “blocks”.
- Traffic between these blocks is typically routed
- The DC will typically behave as a Core to the non-DC blocks
- The DC will behave as a collapsed core/agg/access within the DC block (in a single tier POD).

# The POD Concept



- When we need to expand our POD design, our first step is to acquire a true distribution.
- I would recommend at this point that the 6K be strongly considered at this time.

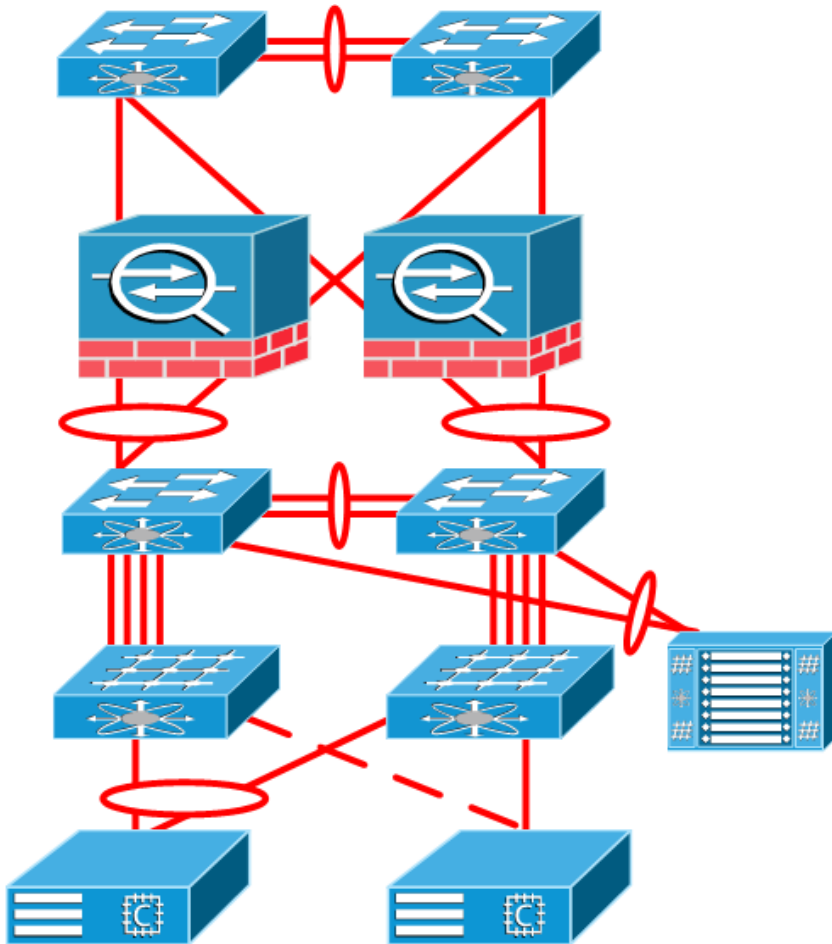
# The POD Concept



- At this point you will MOVE the L3 boundary up to the new distribution.
- This should involve REMOVING the L3 features on the original collapsed distribution.
- You want to do this to allow the ability to perform ISSU on the (now) Access switch.
- You can use VPC to increase reliability to between Access and Distribution

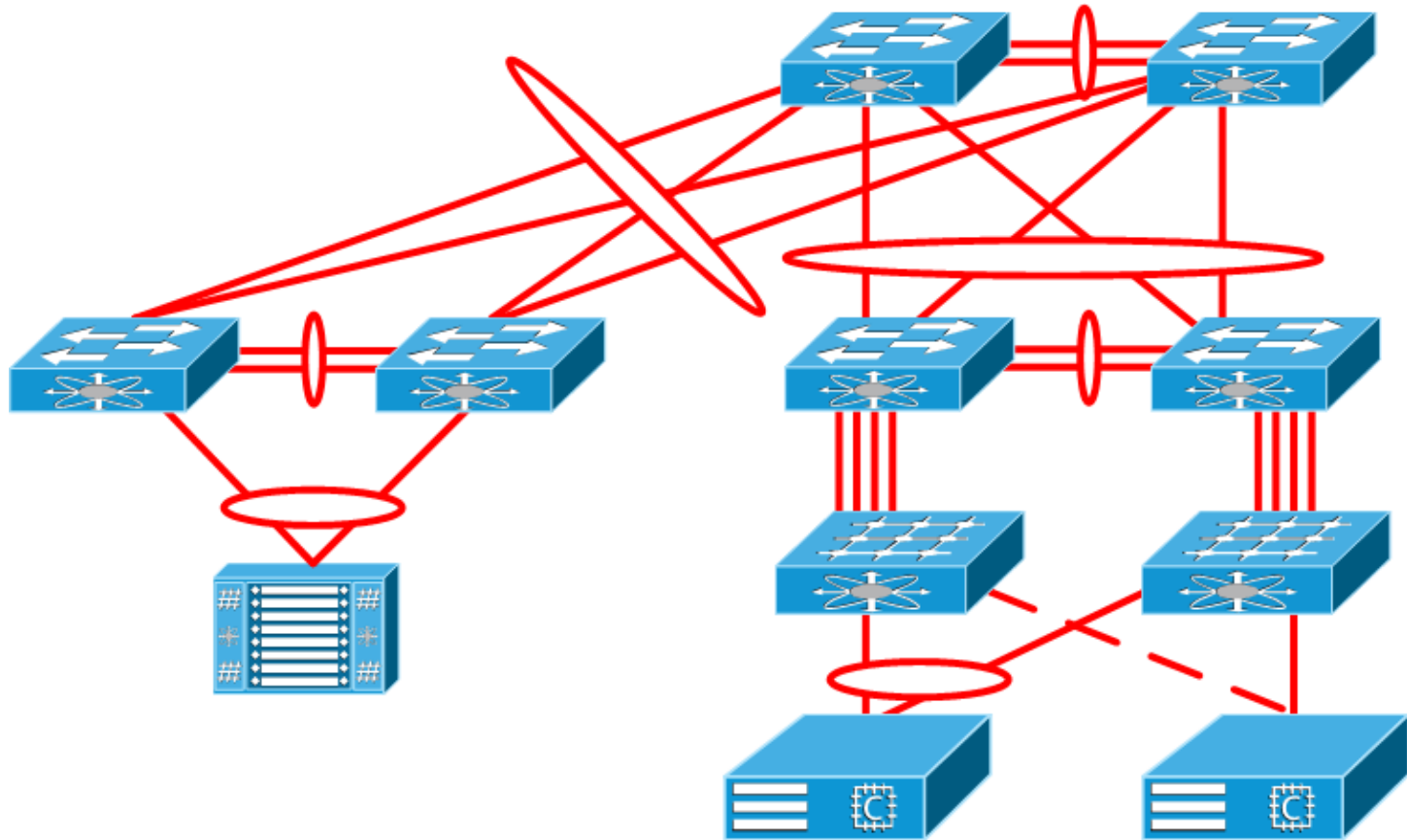


# The POD Concept



- Another option is to insert a Services Layer between the distribution and the access.
- Here we show a pair of firewalls, but there could be other services as well.
- Depending on the service device in question you may or may not be able to run VPC between the layers.

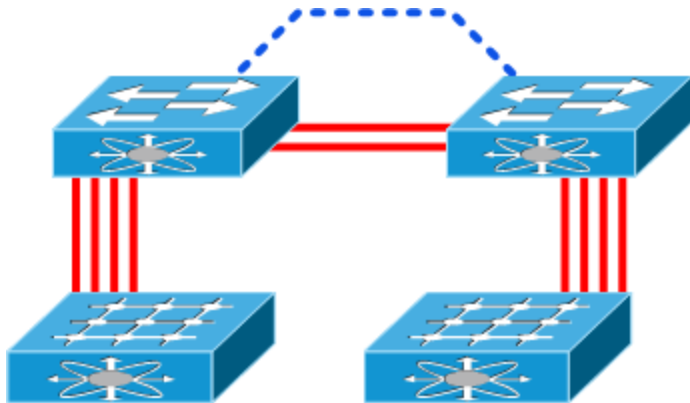
# The POD Concept



# Goals of the POD

- \* The block infrastructure allows a scalable growth pattern while limiting the failure domains.
- \* The goal is that as the POD grows from Single Layer to Dual Tier to Multi Tier you can use a common look/feel/operational procedure to maintain it.
- \* Components will be able to bolt on due to the modularized design.

# 6K/5K/2K Components



A Year ago, there was one choice up here, and it was the Nexus 5500. Now there is a very strong contender, the Nexus 6000. They both have similar features – the 6K is newer and faster, but a little more expensive. It does NOT support Fibre Channel at this time however.

The FEX's (FEXes? FEXen?) that you choose will allow you to connect different devices within your DC fabric. Do you need lots of 10G? Is it copper or fiber? How about storage – FC? iSCSI? NFS? FCoE?

There are FEX options to cover all manner of sins, and new ones come out regularly. You can mix and match, you can also deploy them connected to a single upstream switch OR you can cross connect them to two different upstream switches (more redundancy, less scalability).

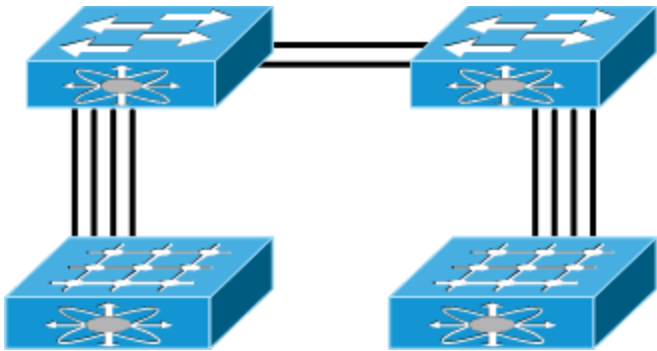
# 6K vs. 5K

Feature	Nexus 55xx	Nexus 6000
Latency	~1.8us	~1us
MAC table	32K	<b>256K MAC/ARP (flexible)</b>
LPM Routes	16K	<b>32K</b>
Hosts	16k	<b>128K</b>
Multicast route	8K	<b>32K*</b>
Bridge Domains	4K	<b>16K</b>
ACLs	4K flexible	<b>4K flexible</b>
IGMP Snooping groups	4K	<b>32K</b>
ECMP	64 way	<b>1K*</b>
VRFs	1K	<b>4K</b>
SPAN	4	<b>31, 16 can be ERSPAN*</b>
Buffer	640K per port dedicated	<b>25MB per 3 QSFP</b>
Number of FEX Support	24 L2 / 16 L3	<b>24 for Both L2 &amp; L3</b>

# 6K vs. 5K

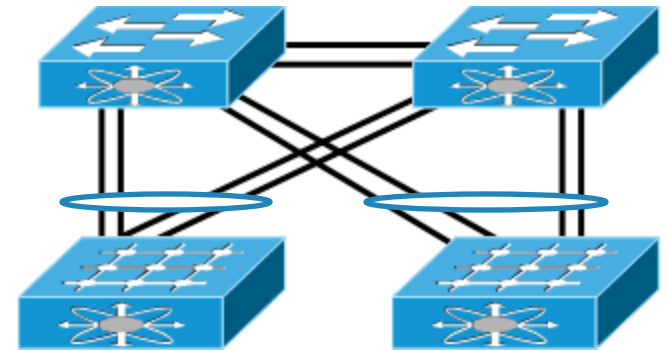
- \* The 6K does NOT support FC
- \* The 5K with an L3 module does NOT support ISSU
- \* The 6K costs more, but provides better performance/scalability
- \* The 6K is brand new, the 5K has been around for a while
- \* The 6K has a much richer routing feature set

# FEX



## Straight Through FEX Design

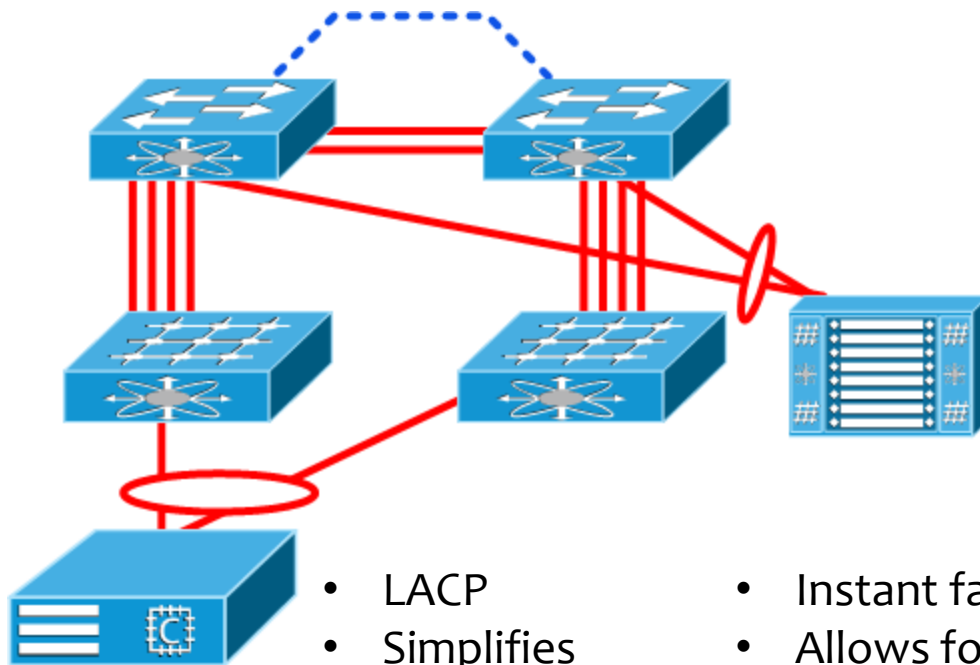
- Up to 24 FEX per 5K
- Prefer if your systems are dual attached (LACP or active/passive)



## Cross Through FEX Design

- Up to 12 FEX per 5K
- Prefer if your systems are single attached.

# Virtual Port Channel



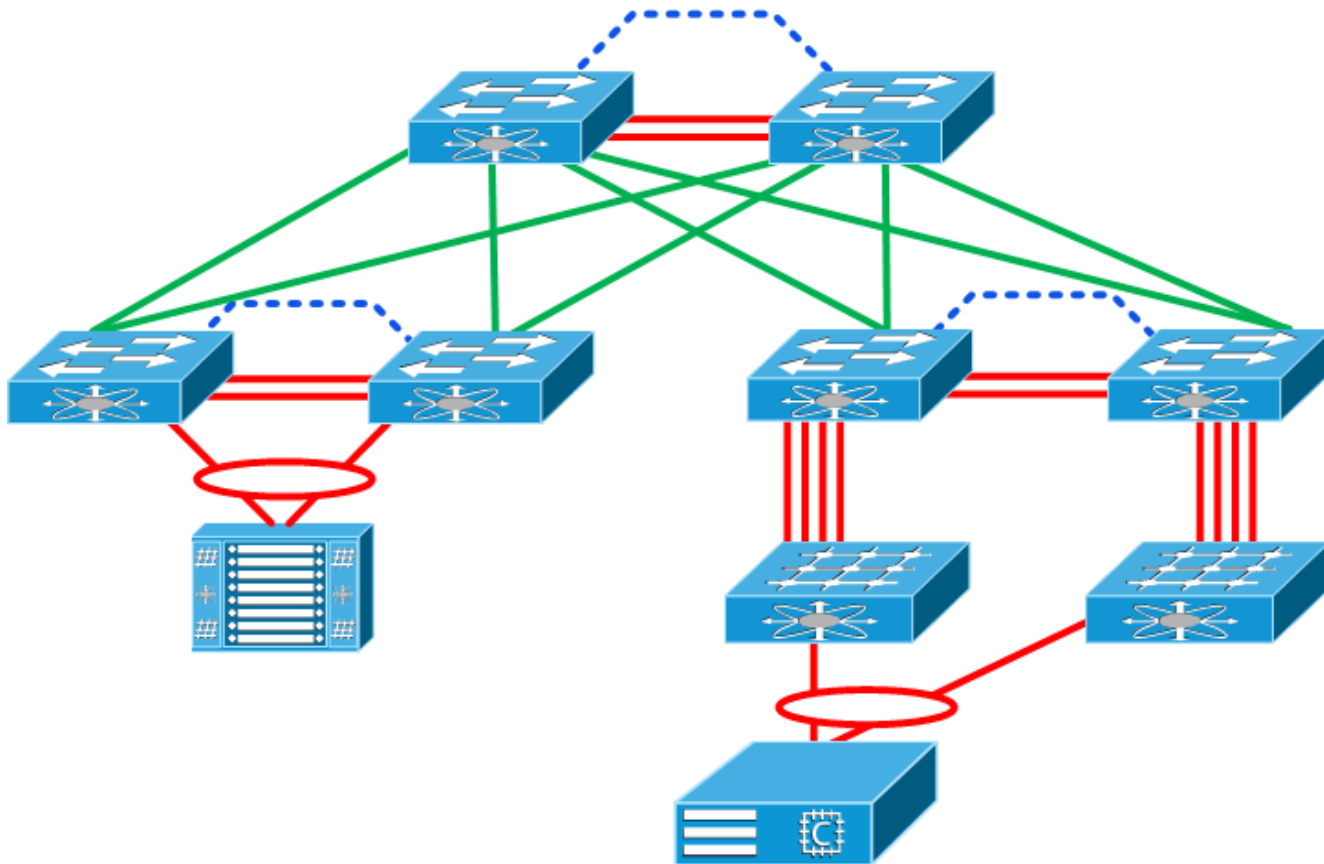
- LACP
- Simplifies Spanning Tree
- Forwarding on multiple links

- Instant failover
- Allows for highly redundant fabrics

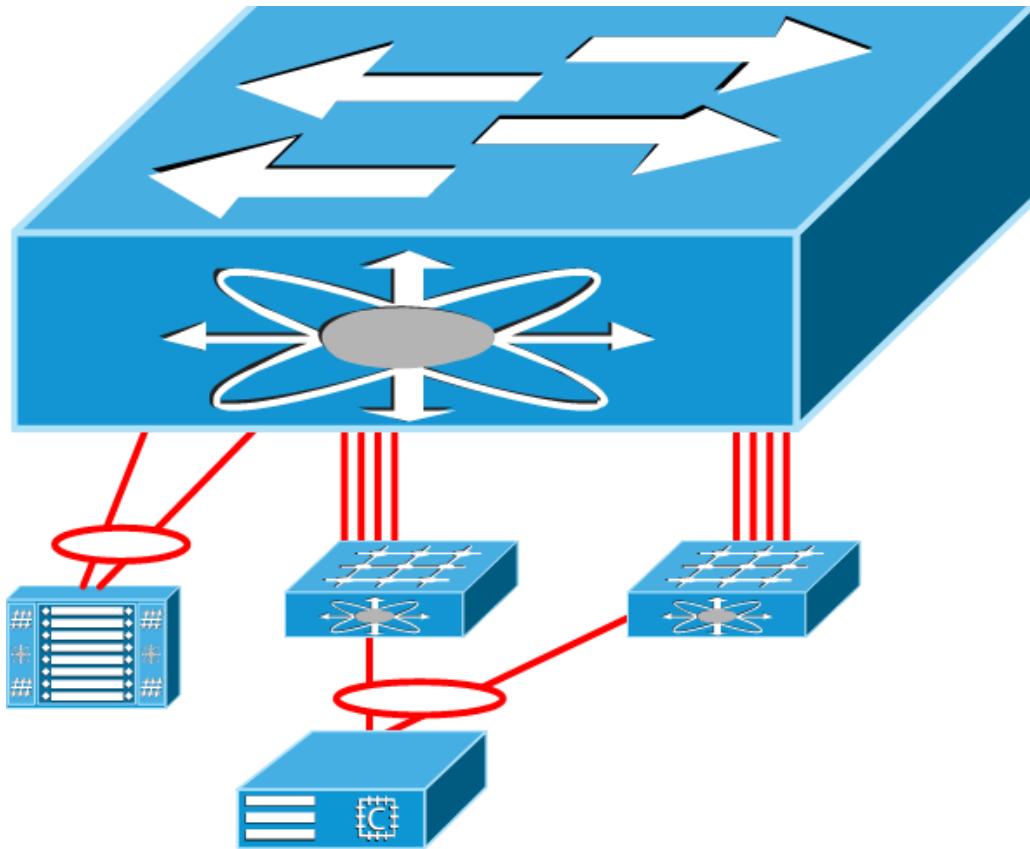
- Peer link
- Peer keepalive
- Orphan ports
- Integrates with FP
- HIGHLY recommended
- You need to be careful with routing adjacencies and peer-links.



# Fabric Path



# Fabric Path



- In a FabricPath domain there are two domains – Classical Ethernet and Fabric Path.
- Logically, the FP domain looks like one large switch from the standpoint of the CE domain.
- The FP domain is effectively routing for the various FP switches in the domain.
- FP is probably NOT a technology you turn on in the single POD situation, but as you expand into multi-tier POD's then it is a very useful technology.

# Links, documents, etc.

- \* <http://bradhedlund.com/2010/12/16/routing-over-nexus-7000-vpc-peer-link-yes-and-no/>
- \* [https://www.ciscolive365.com/connect/sessionDetail.ww?SESSION\\_ID=7730&backBtn=true](https://www.ciscolive365.com/connect/sessionDetail.ww?SESSION_ID=7730&backBtn=true) (need login)
- \* [http://www.cisco.com/en/US/docs/switches/datacenter/sw/design/vpc\\_design/vpc\\_best\\_practices\\_design\\_guide.pdf](http://www.cisco.com/en/US/docs/switches/datacenter/sw/design/vpc_design/vpc_best_practices_design_guide.pdf)
- \* [http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9670/guide\\_c07-690079.html](http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9670/guide_c07-690079.html)
- \* [http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps12806/ps12869/data\\_sheet\\_c78-726128.html](http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps12806/ps12869/data_sheet_c78-726128.html)
- \* [http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9670/data\\_sheet\\_c78-618603.html](http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9670/data_sheet_c78-618603.html)
- \* [http://docwiki.cisco.com/wiki/Cisco\\_NX-OS/IOS\\_Port-Channel\\_Comparison](http://docwiki.cisco.com/wiki/Cisco_NX-OS/IOS_Port-Channel_Comparison)