



The bridge to possible

Trends & Topology Options for SP Networks

Adapting Webscale Technologies in the Provider
Universe

LJ Wobker, Principal TME
@ljwobker

Cisco Webex App

Questions?

Use Cisco Webex App to chat with the speaker after the session

How

- 1 Find this session in the Cisco Live Mobile App
- 2 Click “Join the Discussion”
- 3 Install the Webex App or go directly to the Webex space
- 4 Enter messages/questions in the Webex space

Webex spaces will be moderated until February 24, 2023.



Abstract – or your chance to drop the class...

- The rise of web, or hyperscale networks, has fundamentally changed the networking space. Combinations of applications, system designs, and network topologies have been deployed at sizes and scale unfathomable just a few years back. Should providers be building big leaf-spine topologies? Do we still need large linecard based modular systems? This session discusses the advantages and disadvantages of different device designs and topologies, and how they apply (or sometimes don't apply!) to the provider space. There will be some quantitative analysis (“is thing X or thing Y cheaper or lower power?”) as well as some less objective metrics and even some outright opinions begging to be challenged.

*Can I use web scale
technologies and/or
topologies in my provider
network???*



“Yes!”

and...

“No!”





The bridge to possible

Thank you

CISCO *Live!*

Agenda

- Introduction
- How does web run networks?
- Topology discussion
- Routing systems – abstracted
- “Service scale” vs. “bandwidth scale”
- Conclusion

A decorative graphic in the top right corner of the slide, consisting of a dense cluster of circles in various sizes and colors, including shades of blue, green, orange, and red. The circles are arranged in a way that they appear to be floating or expanding from the right edge of the frame.

What's web-scale,
anyway?

Web-scale networks

- Massive
- Virtualized
- Divide and Conquer
- Layered
- Parallel

Provider networks

- Massive [Not nearly as, if at all]
- Virtualized [Less so, if at all]
- Divide and Conquer [Often: “combine and suffer”]
- Layered [Flattened]
- Parallel [Collapsed and path-y]

Webifying provider networks

what sorts of web-ish things...

- Massive
 - Virtualized
 - Divide and Conquer
 - Layered
 - Parallel
- (Not as) massive
 - Virtualize where you can
 - underutilized!
 - also: underutilized...
 - similar but at smaller scale

Topologies

Clos was the Datacenter Choice

Topology choice: Clos

Multiple definitions exist...

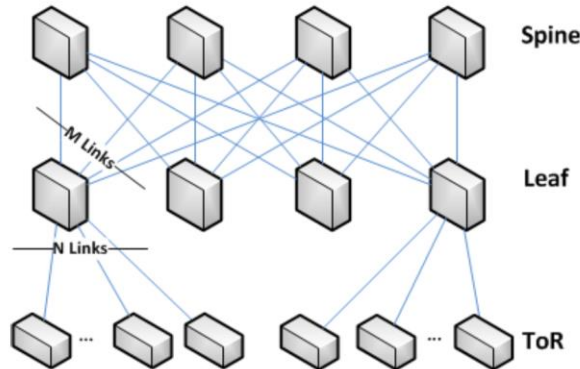
Has N stages
($N=3,5,7..$)

Folded on diagram

Full bisection
bandwidth if $M \geq N$

Natural link load-
balancing

ECMP Based



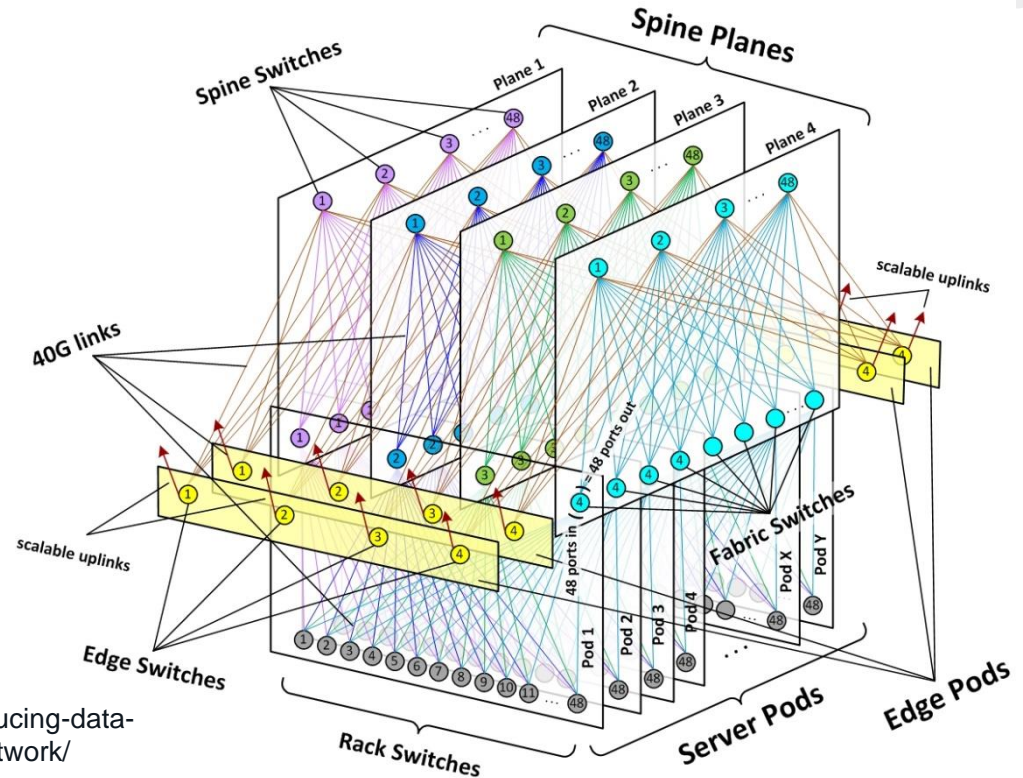
3-Stage Folded Clos Topology

<https://www.nanog.org/meetings/nanog55/presentations/Monday/Lapukhov.pdf>

Potential for Massive Scale

The Data Center Fabric

- “**hundreds of thousands** of 10G-connected servers”
- “**multi-petabit** bisection bandwidth”



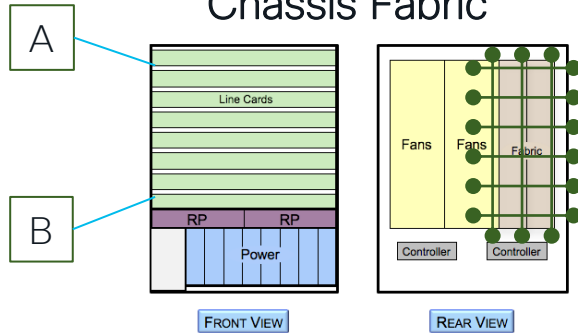
<https://code.facebook.com/posts/360346274145943/introducing-data-center-fabric-the-next-generation-facebook-data-center-network/>

Fabrics Come In All Shapes and Sizes

NPU Fabric

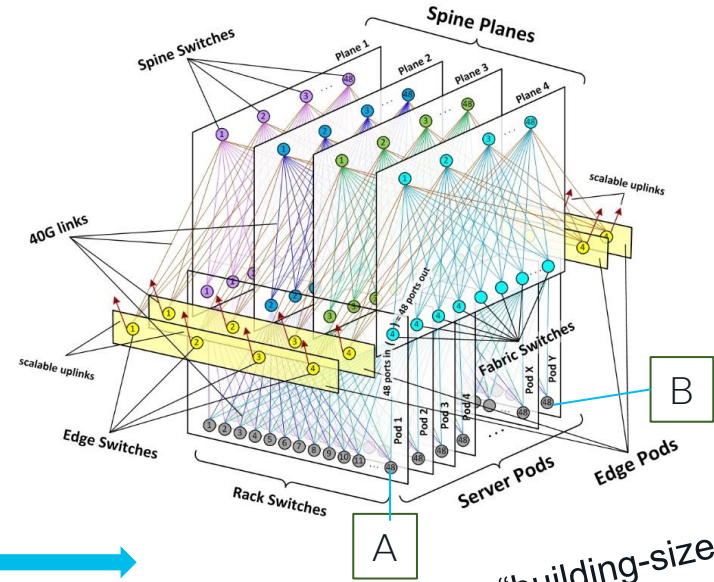


Chassis Fabric



Internal

Data Center Fabric

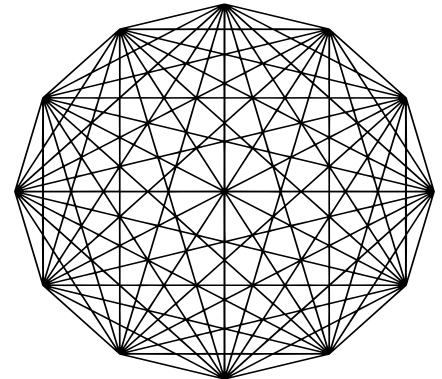
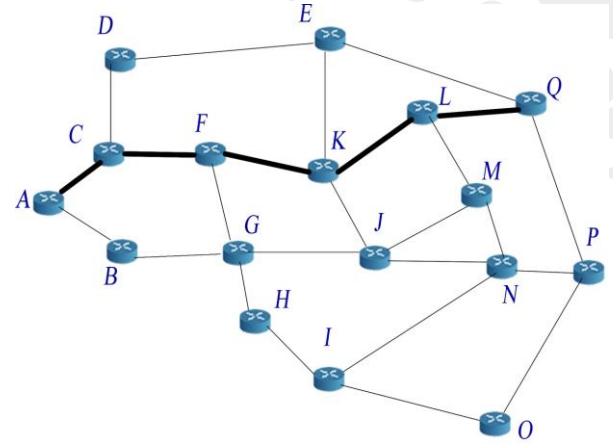


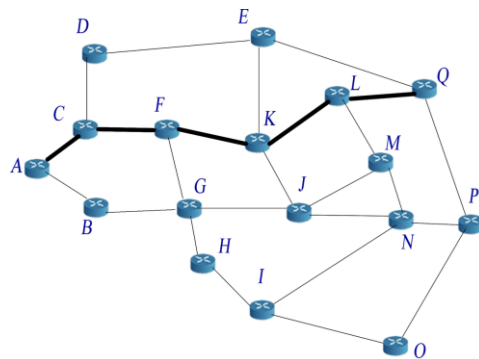
External

“building-sized router”

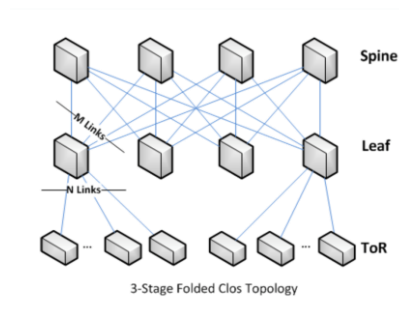
Topologies : meshes

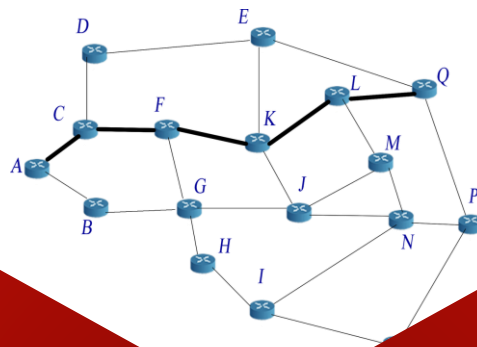
- Meshes can be damn near anything....
 - Often (much?) less regular
- Can be full or partial
 - Sparseness is incredibly important to how we use these things!
- Often follow “human” topologies
 - Geography
 - Population
- Our meshes often grow more “organically”





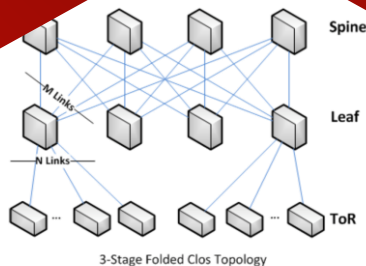
Web networks are all fabrics.
SP networks are all meshes.





Web networks are not fabrics.
 SP networks are not meshes.

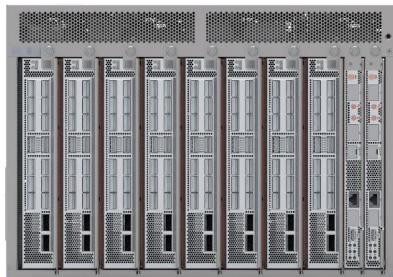
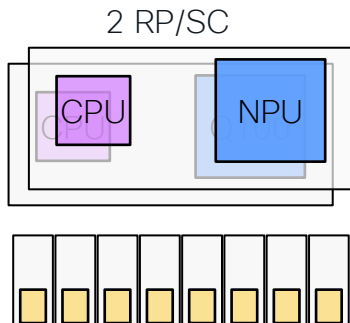
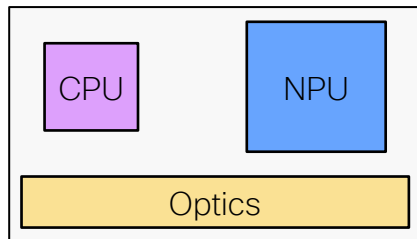
NO



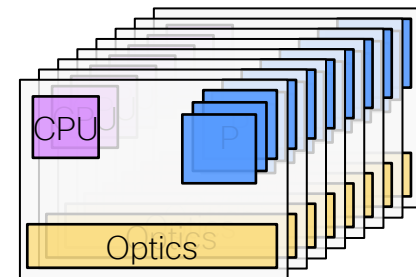
Systems

3 Chassis Types


very broadly speaking, routers almost always look like one of these...



Fabric (8)



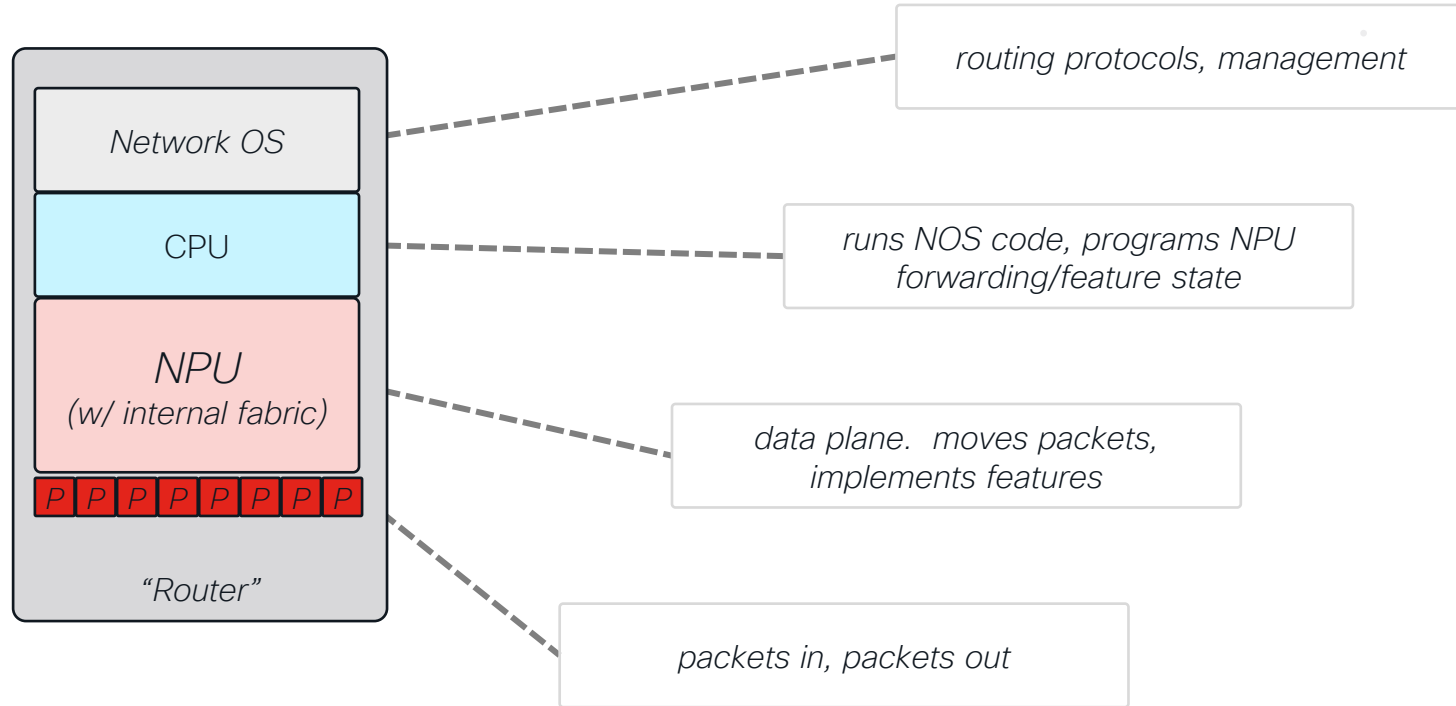
LC
(4-18)



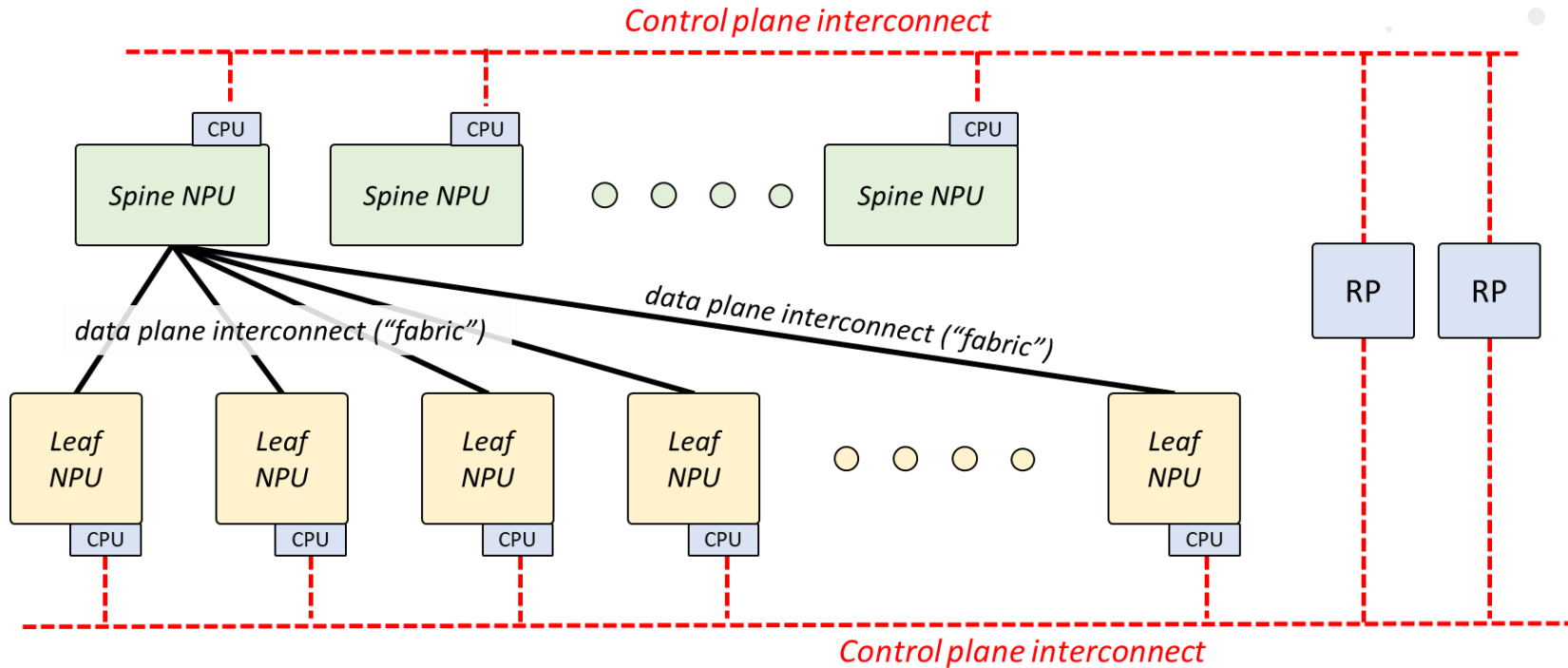
“All single-NPU systems are alike. All distributed systems are distributed in their own way...”

-- LJ Wobker, with apologies to Leo Tolstoy

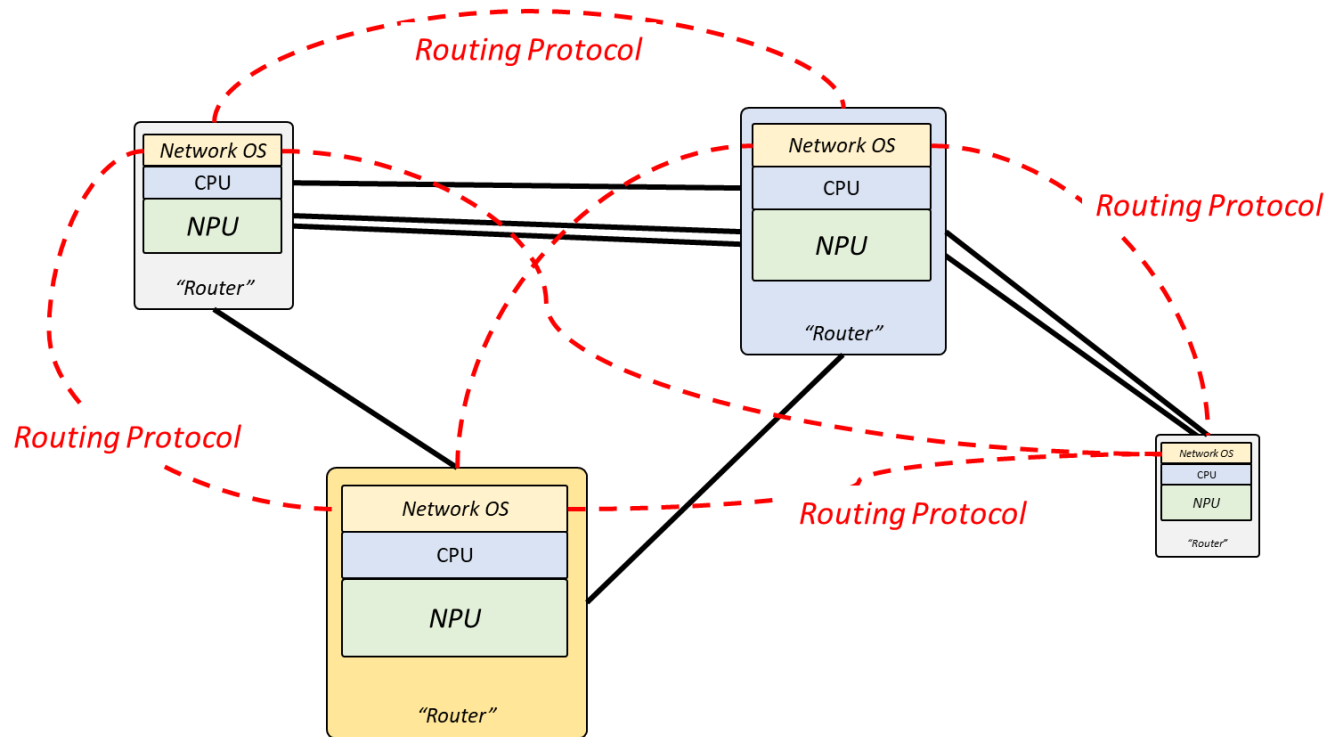
A simple, fixed routing system.



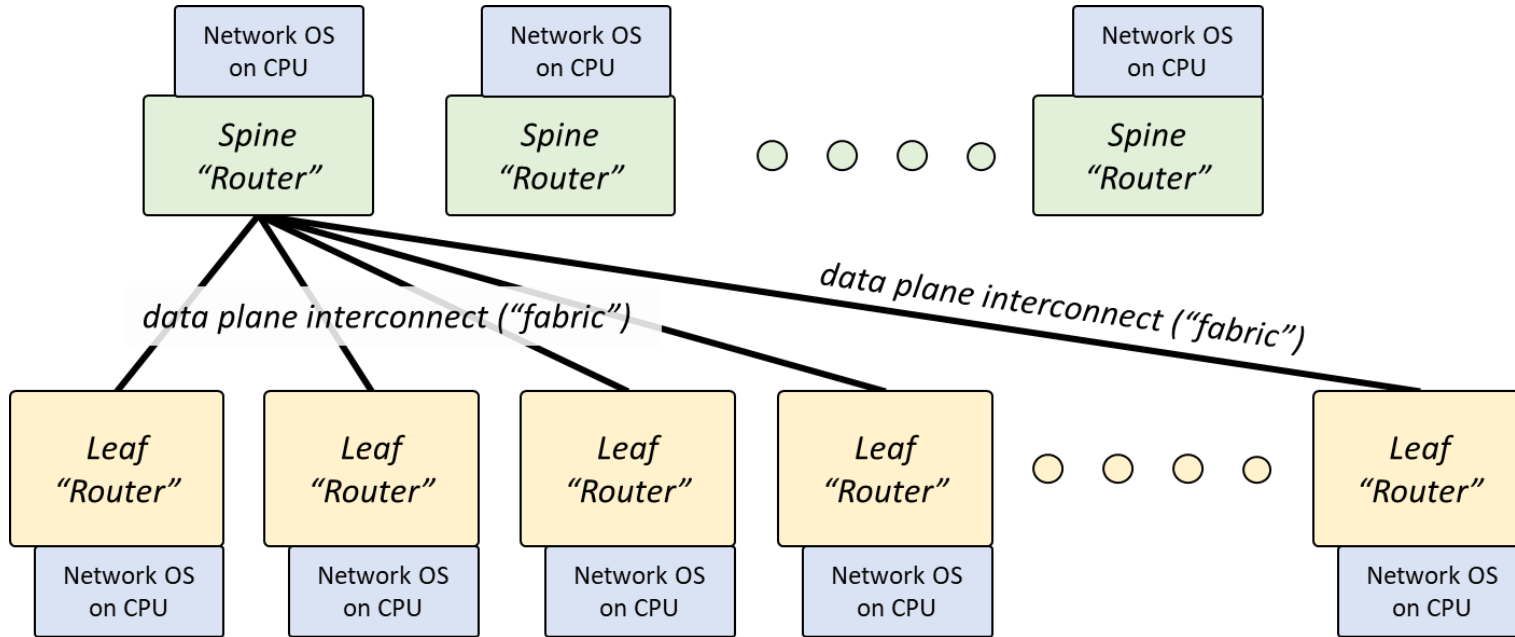
The parts of a “big router” ... abstracted.



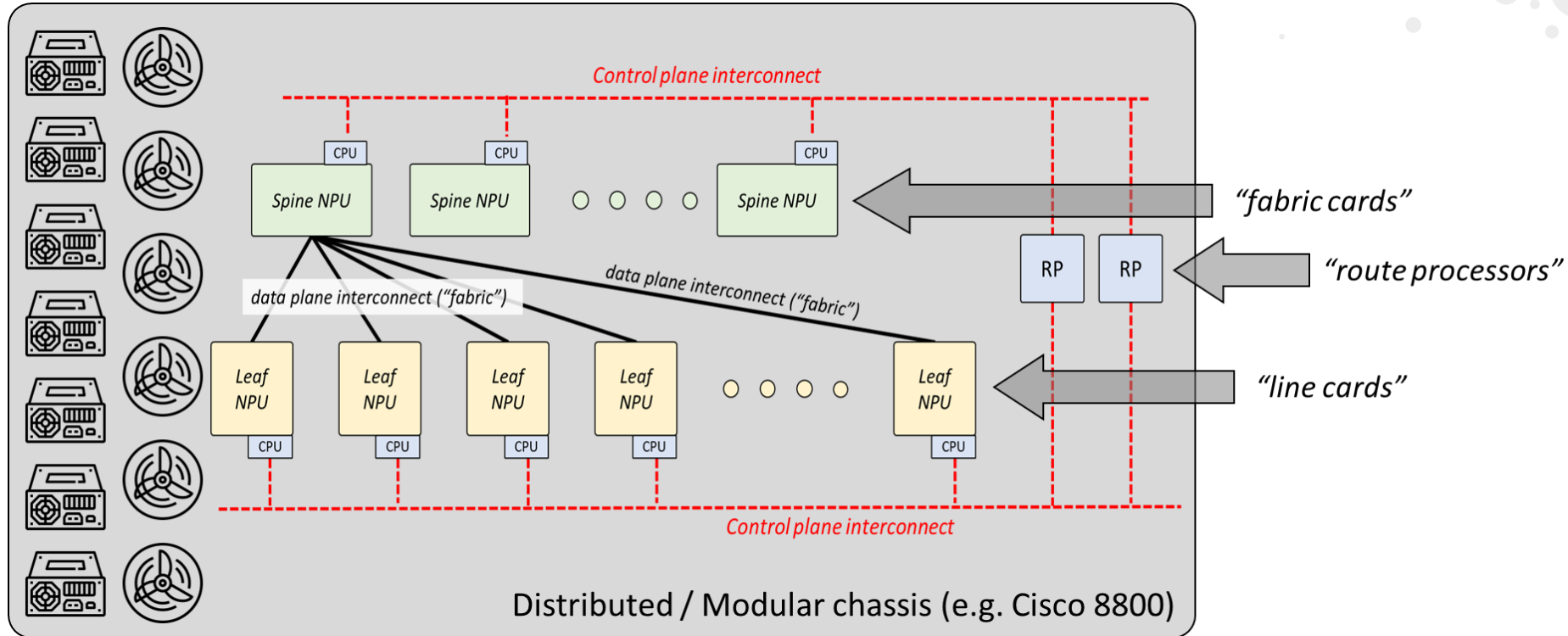
A “network” of routers



A bunch of routers, in a fabric

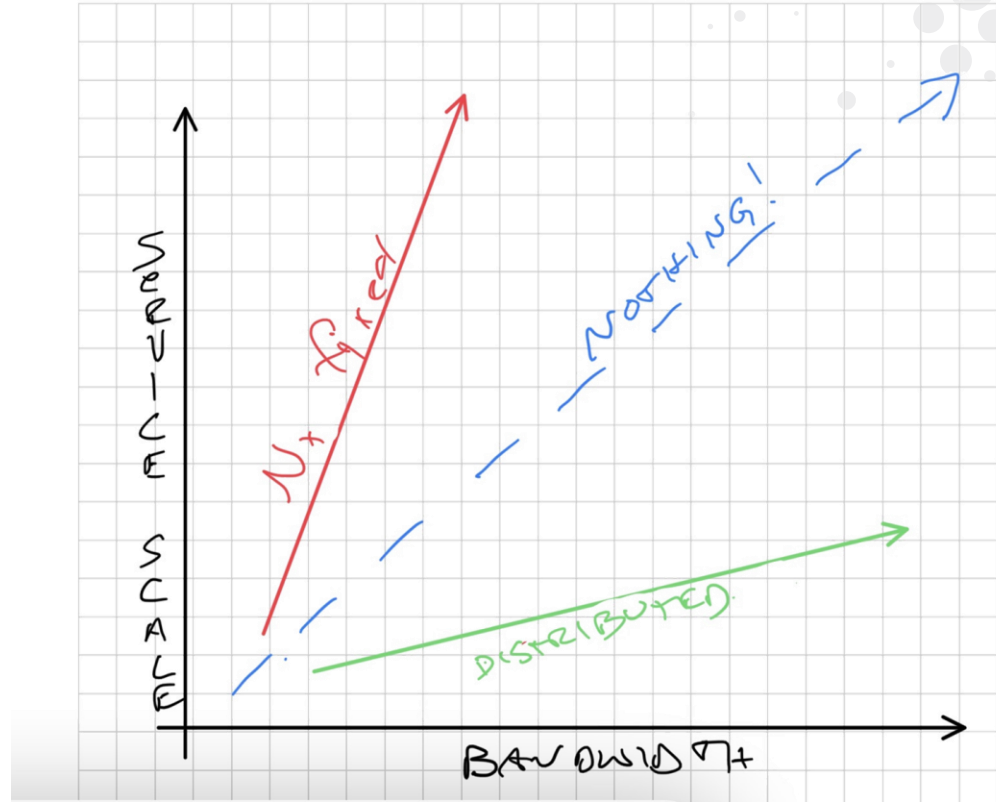


A distributed router.

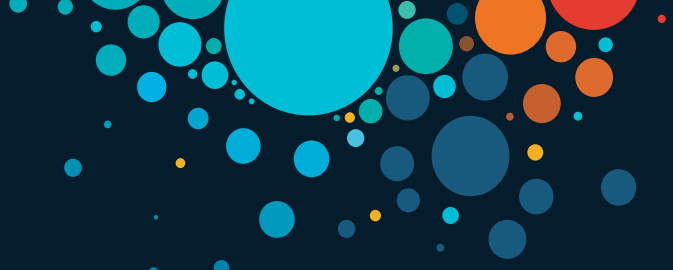


Takeaway: Scaling

- $N \times$ fixed scales services
- $N \times$ linecards scale bandwidth
- nothing scales both
 - (without substantial tradeoffs...)



Distributions



I want you to figure out a way to run the network, delivering more bandwidth and more services, but at a lower total cost. This should be easy, so just let me know when you're done. . .

-- every executive, at every network operator, everywhere
(and, the bane of our collective existences)

Simple math... cost calculator

*What I think most of us **have been using** as the formula:*

Simple math... cost calculator

*What I think most of us **have been using** as the formula:*

*$cost = bandwidth * some_loaded_cost_per_gigabit$*

Simple math... cost calculator

What I think most of us **have been using** as the formula:

*cost = bandwidth * some_loaded_cost_per_gigabit*

What I think most of us **should be using** as the formula:

Simple math... cost calculator

What I think most of us **have been using** as the formula:

$$\text{cost} = \text{bandwidth} * \text{some_loaded_cost_per_gigabit}$$

What I think most of us **should be using** as the formula:

$$\begin{aligned} \text{cost} = & \text{bandwidth} * \text{some_loaded_cost_per_gigabit} \\ & \text{PLUS} \\ & \text{number_of_services} * \text{some_loaded_cost_per_service} \end{aligned}$$

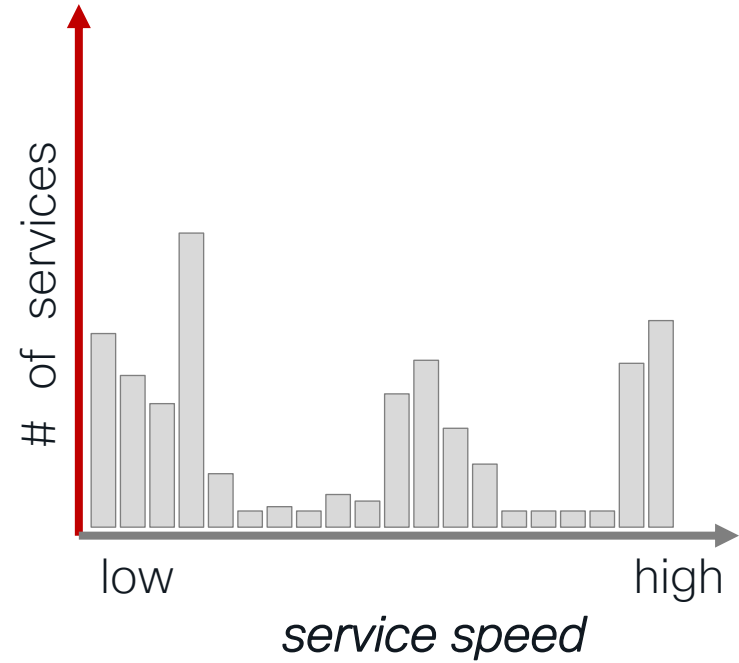
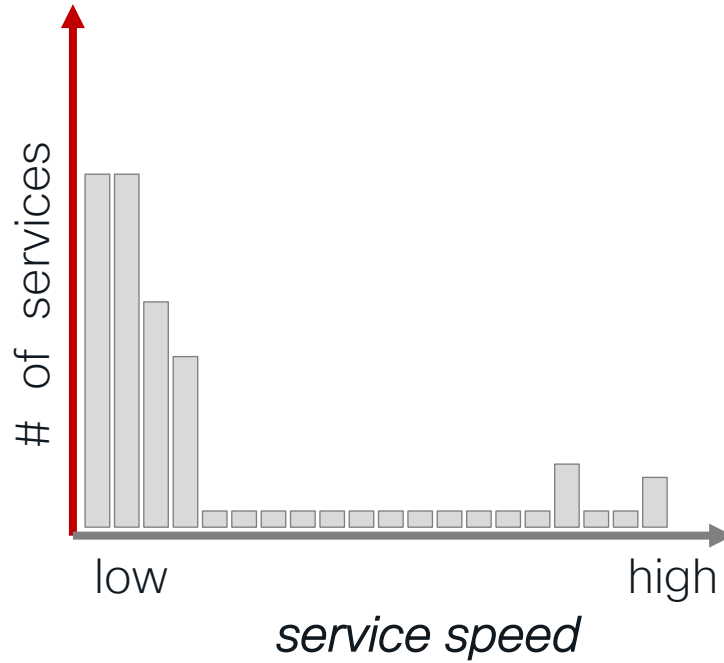
Why?

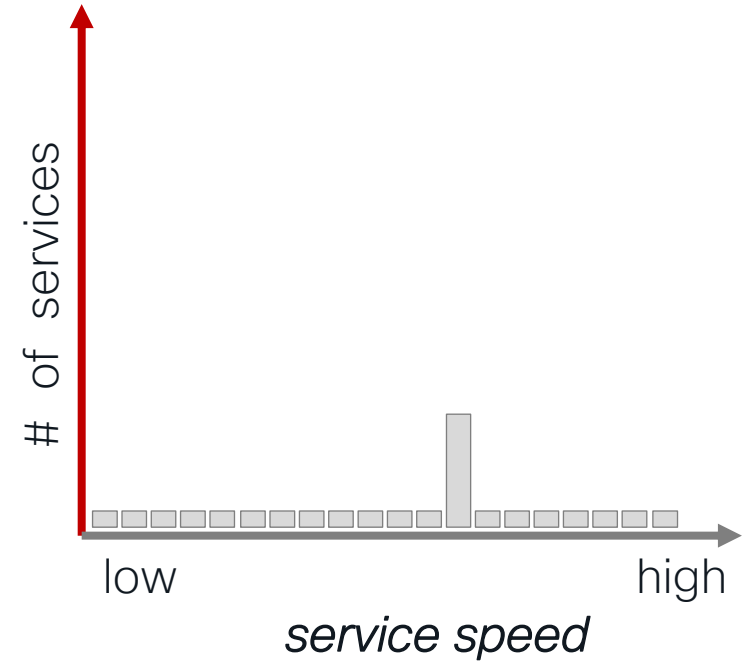
1. Bandwidth was a reasonable proxy for cost.
2. The distributions of service vs. bandwidth were not as wide.

The networking world's most important ratio

- Features vs bandwidth
- Or: "control plane" vs "data plane" -- (arguable context?)
- "control plane" includes stuff like:
 - Peers, routes, labels, interface... But ALSO:
 - Counters
 - Policies
 - Layer/encap combinations
- Maybe: "anything that is not just moving bits is control-plane" ?


“Provider-like” distributions





Distribution of service speeds

- The distributions of service bandwidth matter.
- They matter MUCH more than the “average bandwidth”
- Do any of these look like yours?
- Do you have a Really Good Idea what yours looks like?



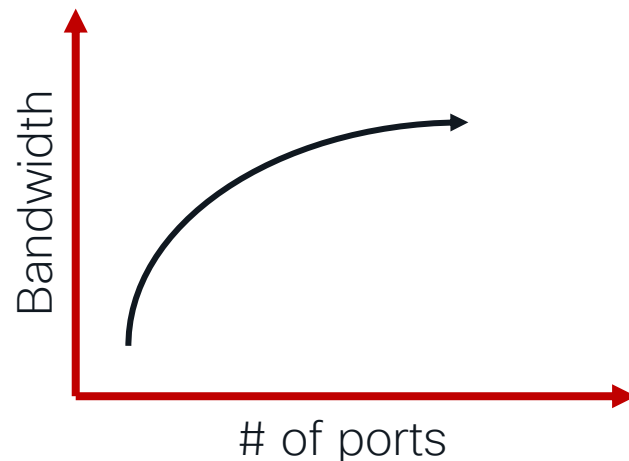
*Even more important: do you have a
Really Good Idea what your distribution is
next year ... in 3 years? ... 5 years?*

-- something I really think you should be thinking about

Contentions

I shall contend... (operator side)

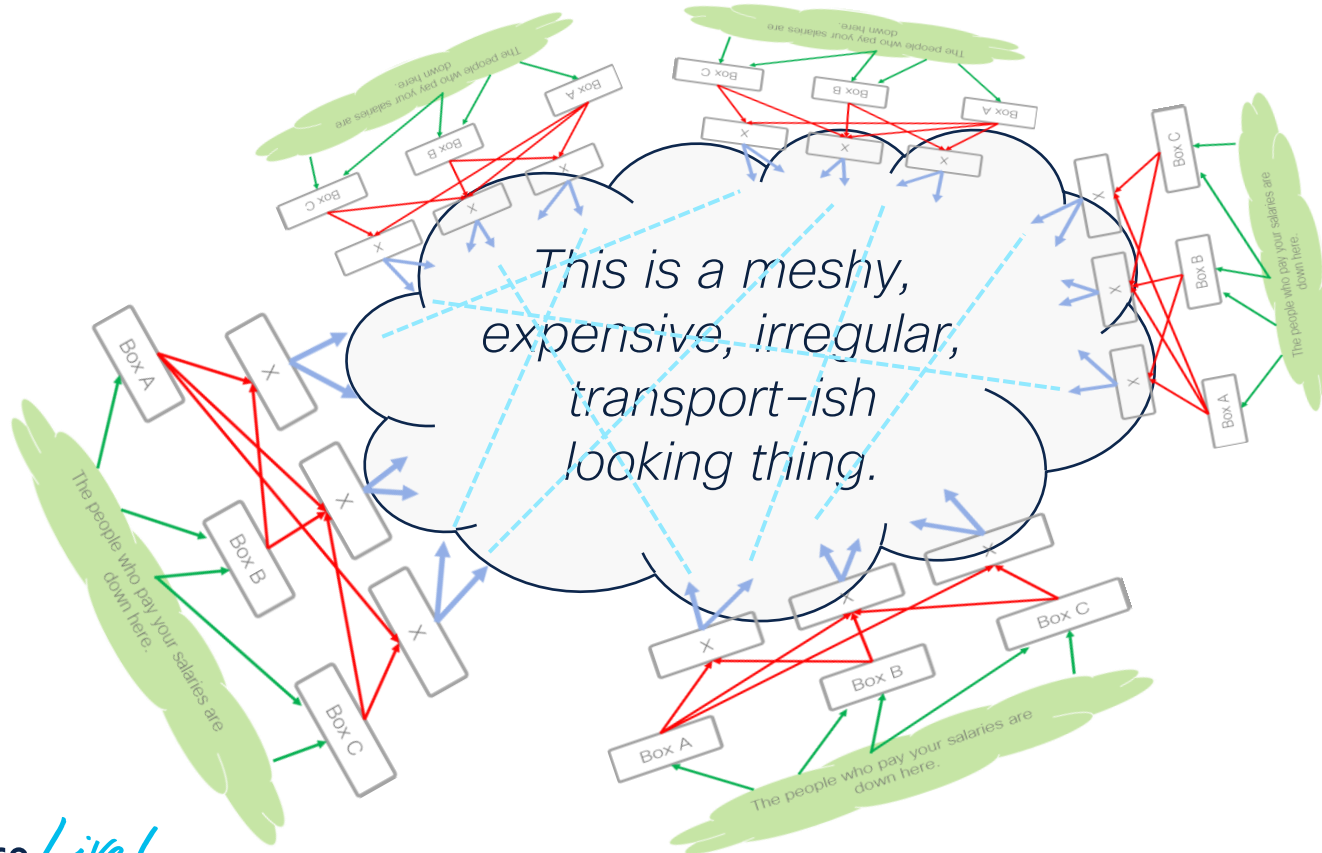
- Bandwidth continues to grow, but slowing.
- BUT - bandwidth growth is unevenly distributed
- The gap between the "fast" and "slow" continues to widen
 - Now "multi gigabit" and "multi kilobit"
- Zipf / power law like distributions
 - continued stretch between fast and slow



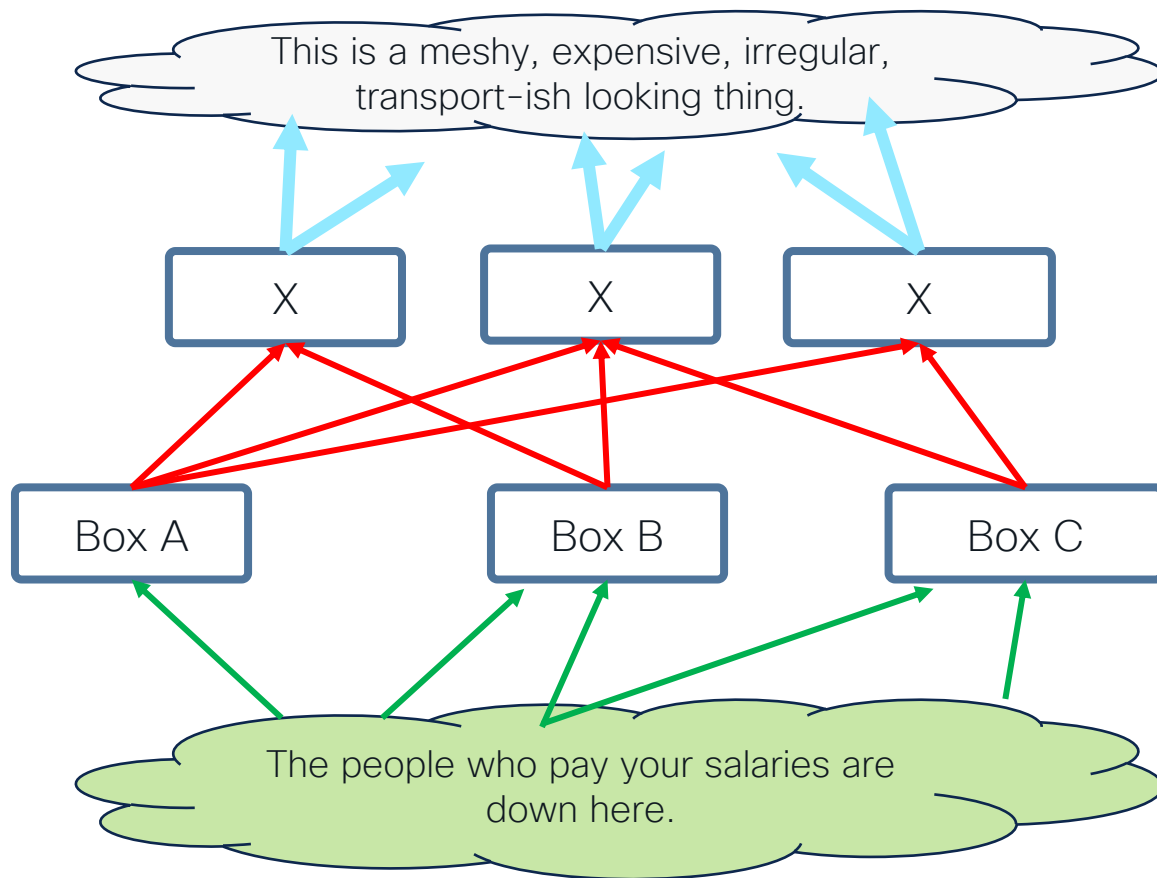
I shall further contend... (vendor side)

- MSDC will drive astonishing bandwidth/density improvements
- Silicon will get "leaner" (at least comparatively...)
- Fixed systems will become increasingly relevant
- Optics-to-other price ratio will continue to widen
- Rate of new systems to consume will increase

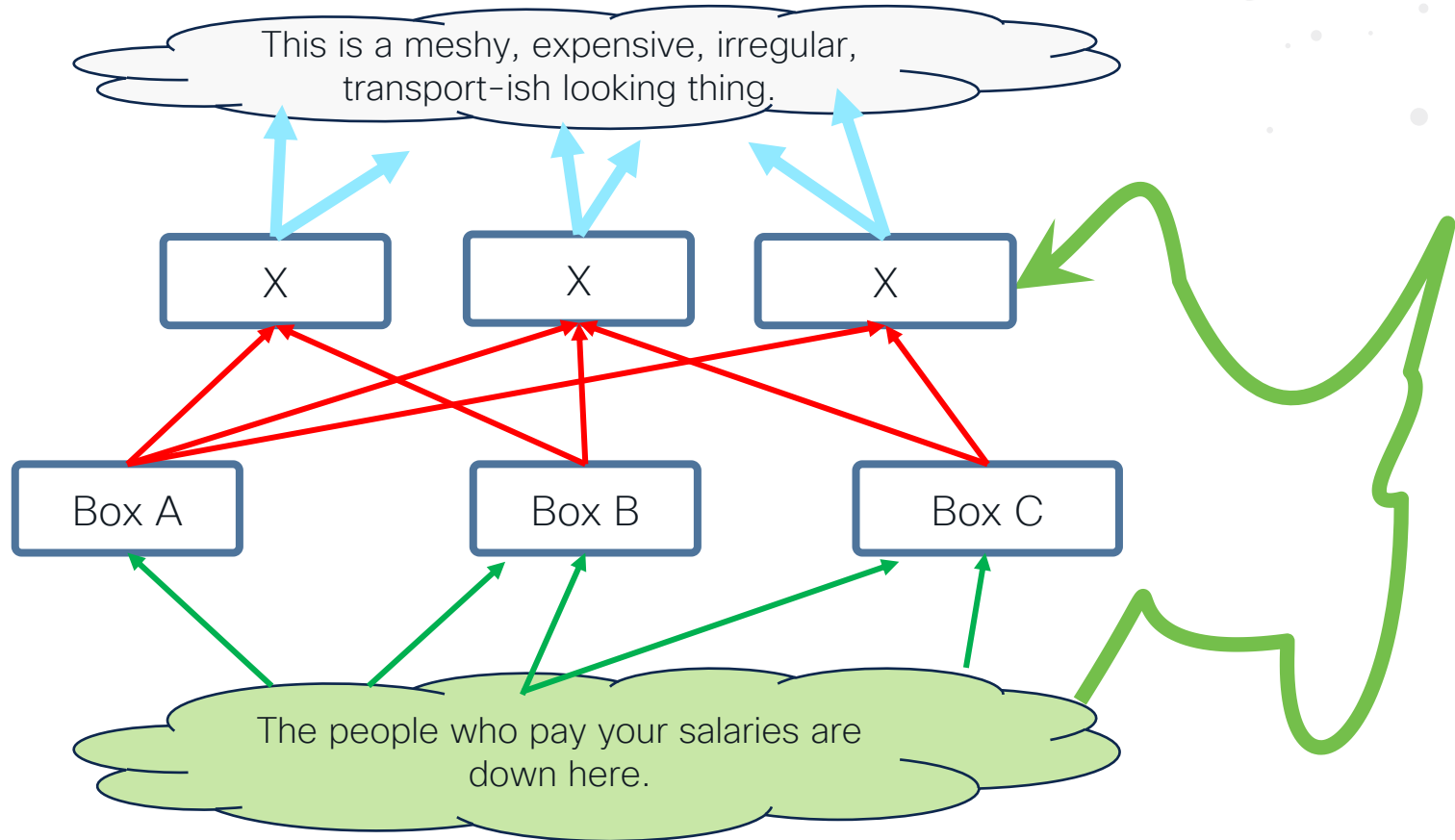
The "generic SP" network diagram



Zoom in some...

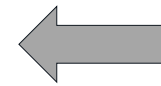
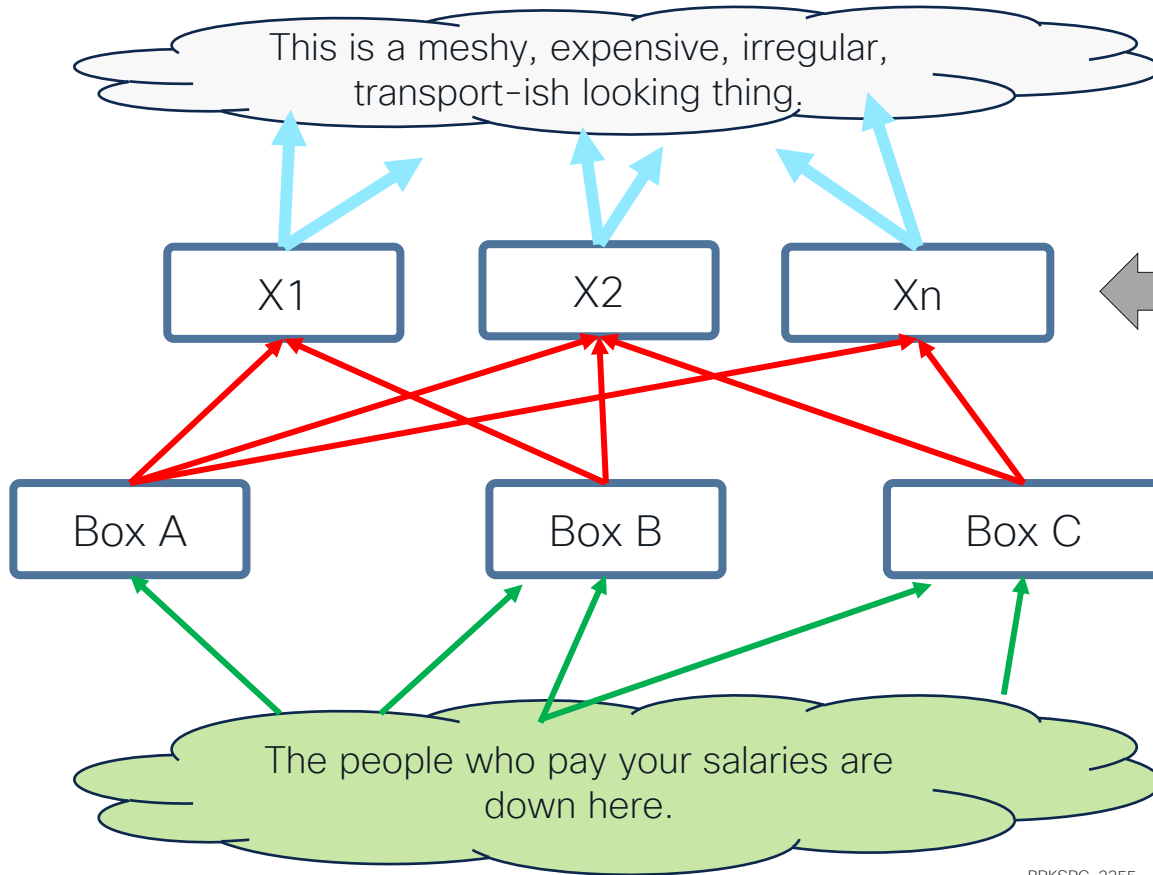


Do you do this????



Business Trends, Design Thoughts

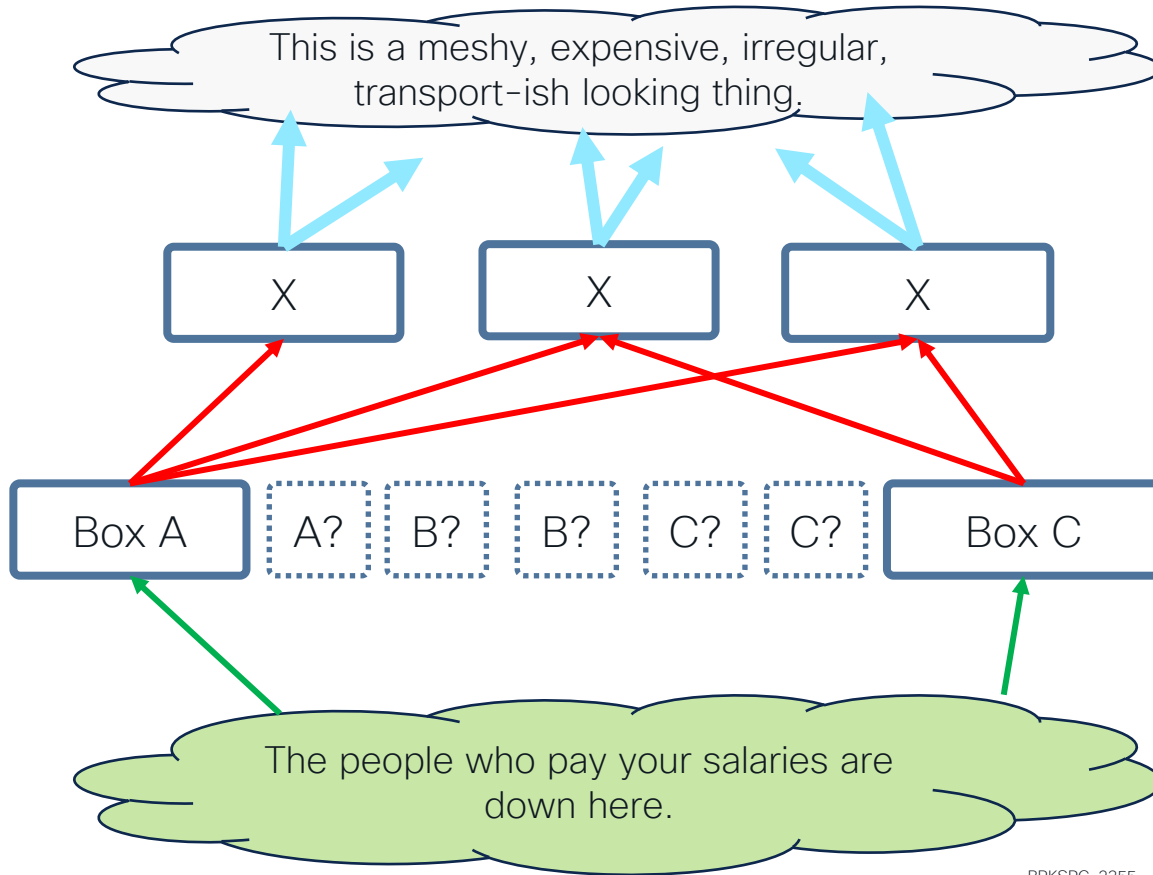
“Web as SP” or “SP as Web”



Size this day-one.
Don't run out!

*(This is much less painful than
this used to be, due to a
combination of business and
technology reasons)*

Expanding the network part over time...



Super important:
You do NOT have to know how many "A" vs "C" boxes you need on day-one.

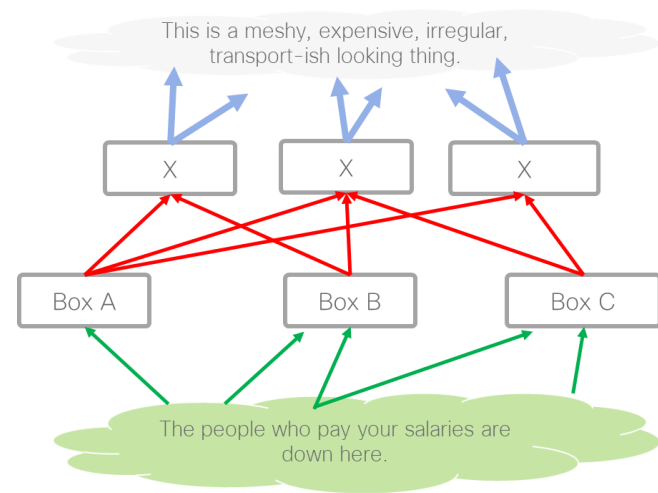
If you have free space at the top ("spines"), it's trivial to scale A/B/C later ("leafs").

In *whatever ratios* you end up needing (!!!)

And - it's easy/cheap to *change* them later.

A possible summary?

- Transport is *expensive*, so optimize here.
- Red links are NOT: don't worry as much
- Pick how much “box” complexity you need.
- NOT A PRESCRIPTION FOR:
 - A specific number of levels/layers (agg below is fine)
 - A specific boxes: fixed/modular, (non)-redundant
 - A specific vendor (though we have a preference...)



One last thing before we go...

I'm grateful for the opportunity to be here and to talk about things (that at least to me!) are fascinating. It's important to me (personally) that attendees find these presentations helpful and/or valuable. Not everyone will enjoy every presentation – and that's OK. You folks sacrifice time and money to come here; Cisco and the speakers do a ton of work to put on the event and the presentations. If the material is excellent, we've done our jobs. If there's something we can do better, please let us know... I know you're not supposed to on the internet, but I really do read the comments. ; -)

Thank you for your time today.

Lawrence J Wobker, Cisco

February 2023

Complete your Session Survey

- Please complete your session survey after each session. Your feedback is important.
- Complete a minimum of 4 session surveys and the Overall Conference survey (open from Thursday) to receive your Cisco Live t-shirt.
- All surveys can be taken in the Cisco Events Mobile App or by logging in to the Session Catalog and clicking the "Attendee Dashboard" at <https://www.ciscolive.com/emea/learn/sessions/session-catalog.html>



Continue Your Education



Visit the Cisco Showcase for related demos.



Book your one-on-one Meet the Engineer meeting.



Attend any of the related sessions at the DevNet, Capture the Flag, and Walk-in Labs zones.



Visit the On-Demand Library for more sessions at ciscolive.com/on-demand.

Q & A Time...



The bridge to possible

Thank you

CISCO *Live!*

CISCO *Live!*

ALL IN