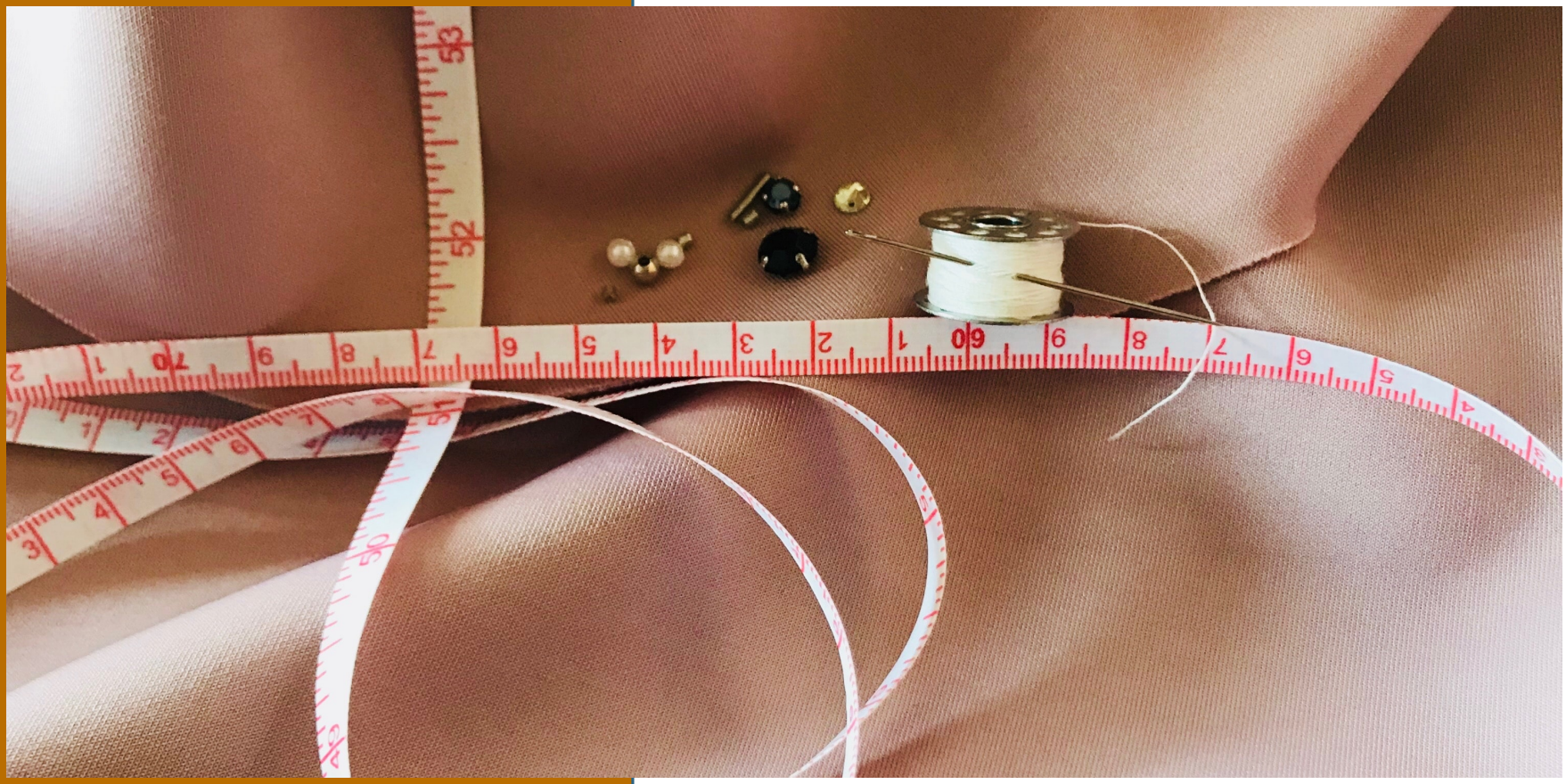
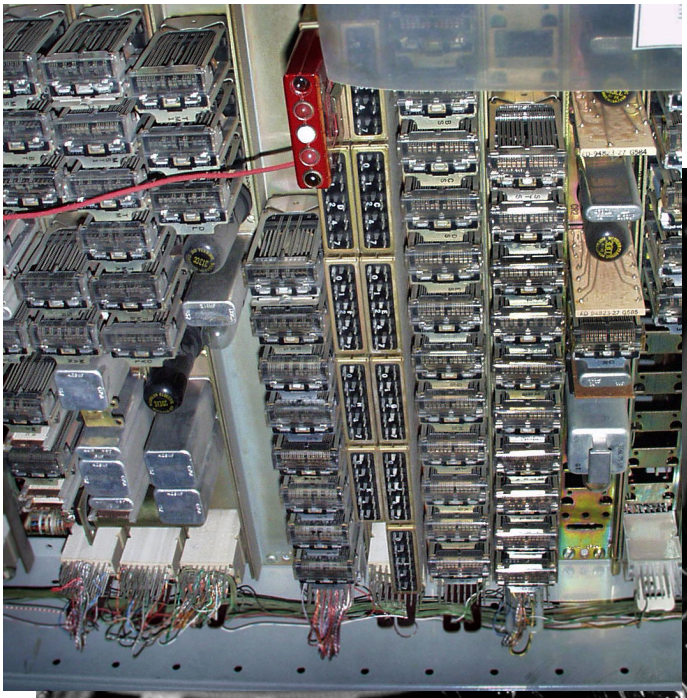
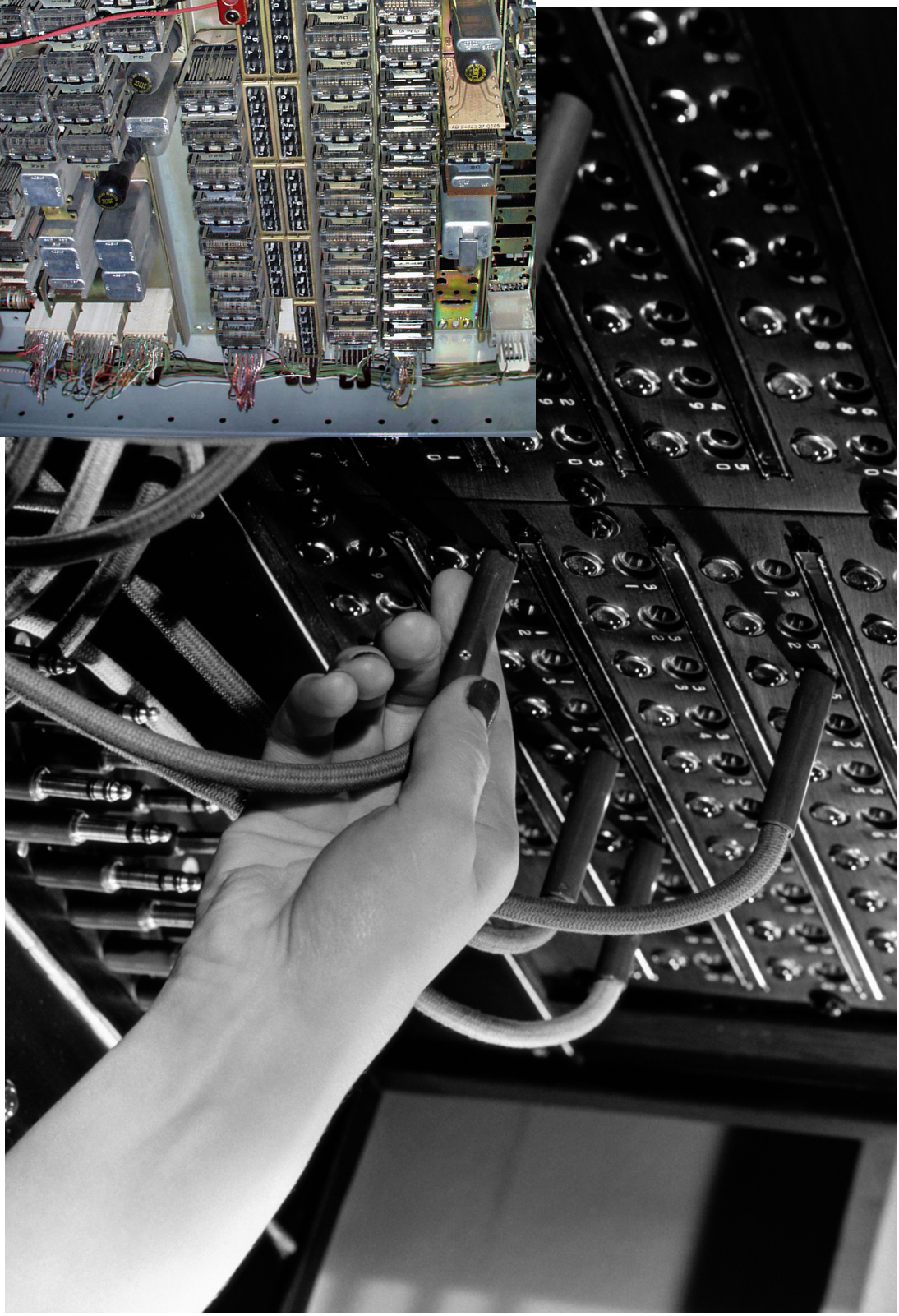
A short history of the spine and leaf



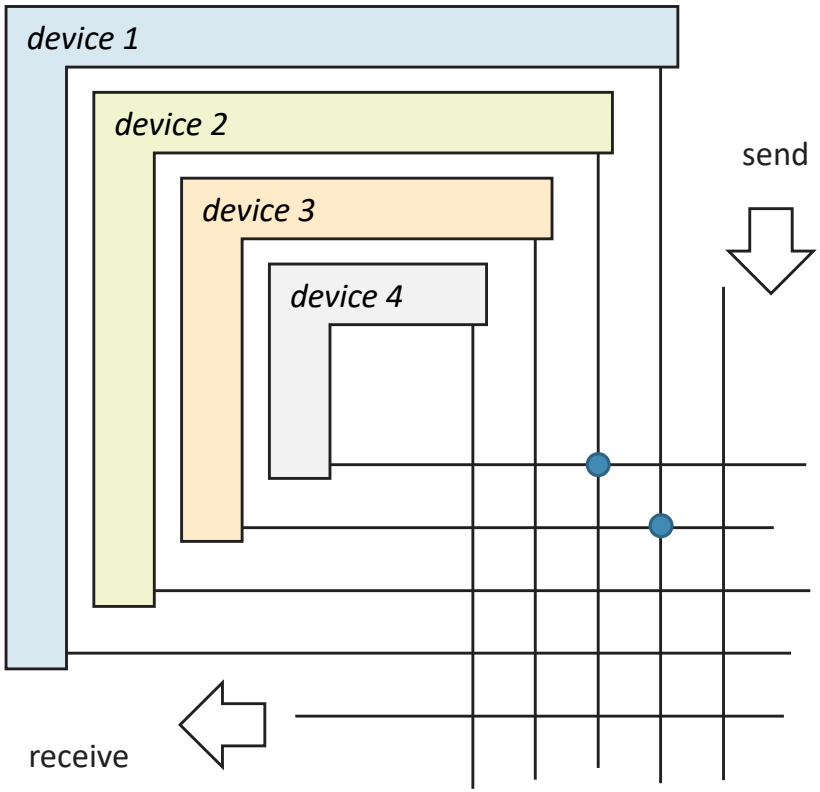
JUNIPER NETWORKS | Engineering Simplicity

Russ White
https://rule11.tech





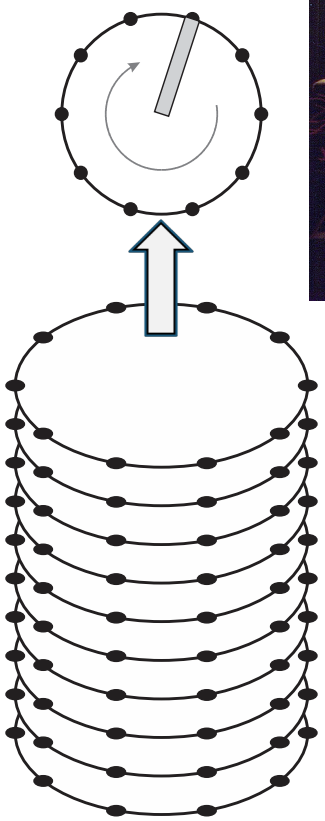
crossbar

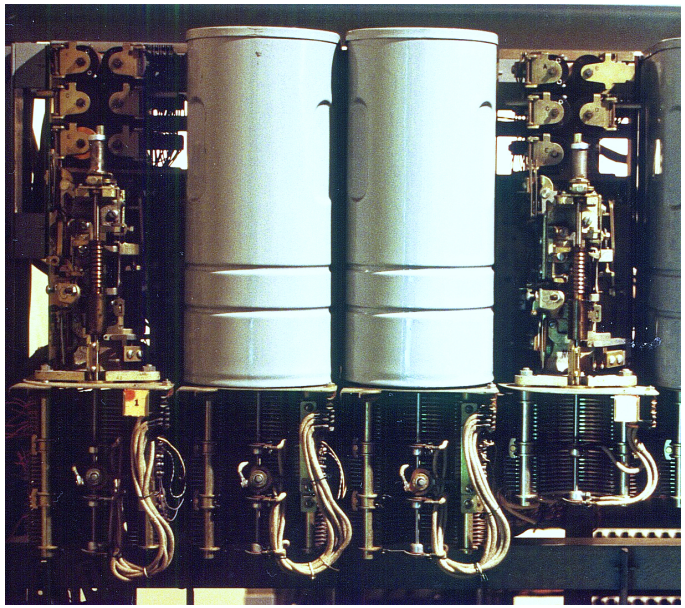


Device 2 made to device 4

Device 1 made to device 3

crossbar





Almon Brown Strowger, 1891

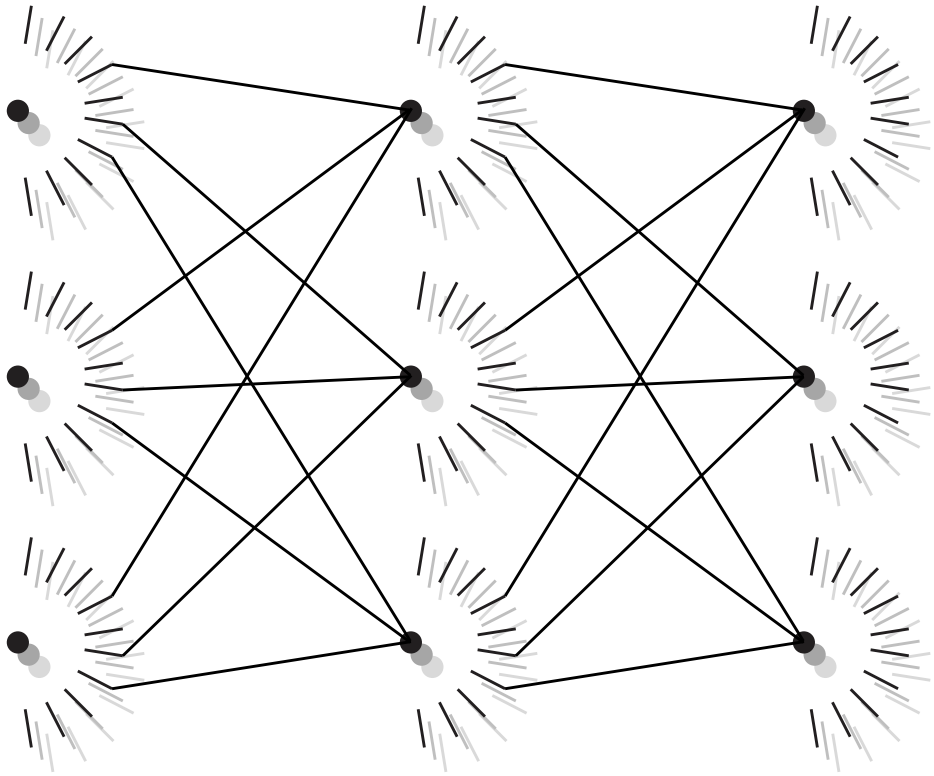
"Step by step" switch

Designed to take the operator out of the circuit make process

stowger

- Trunk lines to connect to other frames
  - Not many, expensive to drive
  - So “long distance” was expensive
- Tens/Hundreds of thousands of lines
- Often paired with relay crossbars at BDFs

Stowger frame

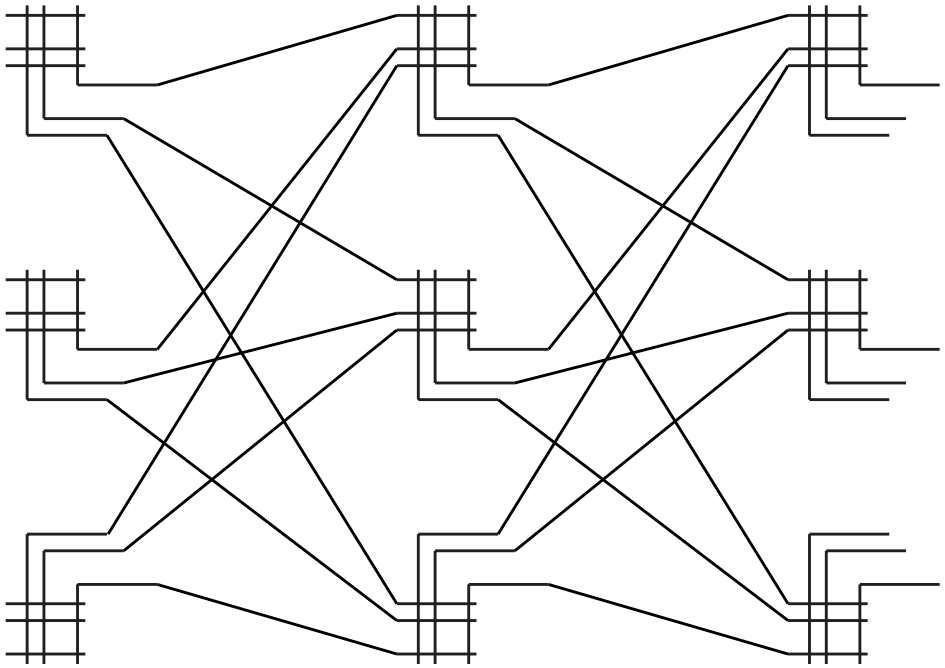


Each switch is one port in, one hundred ports out

Wired as a fabric

Three stages gives you six digits

Stowger frame



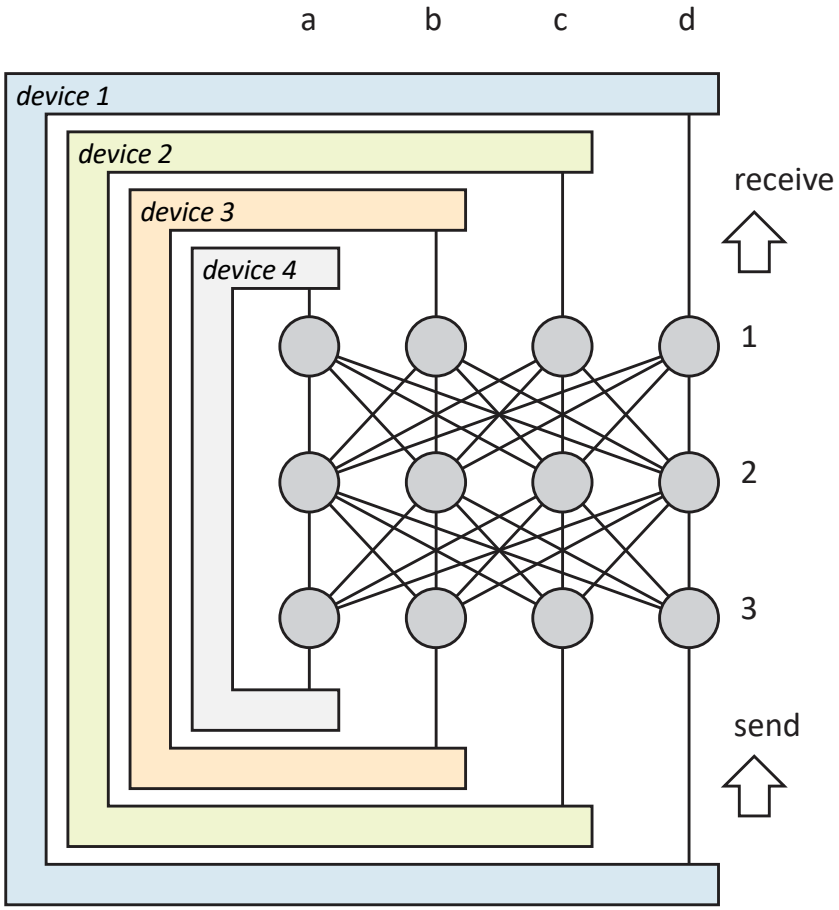
Charles Clos in 1952-3

Replace the Stowger switches with small, fixed size crossbars

Multiple inputs allowed a more compact fabric

3 stages
(or arys)

Clos fabric

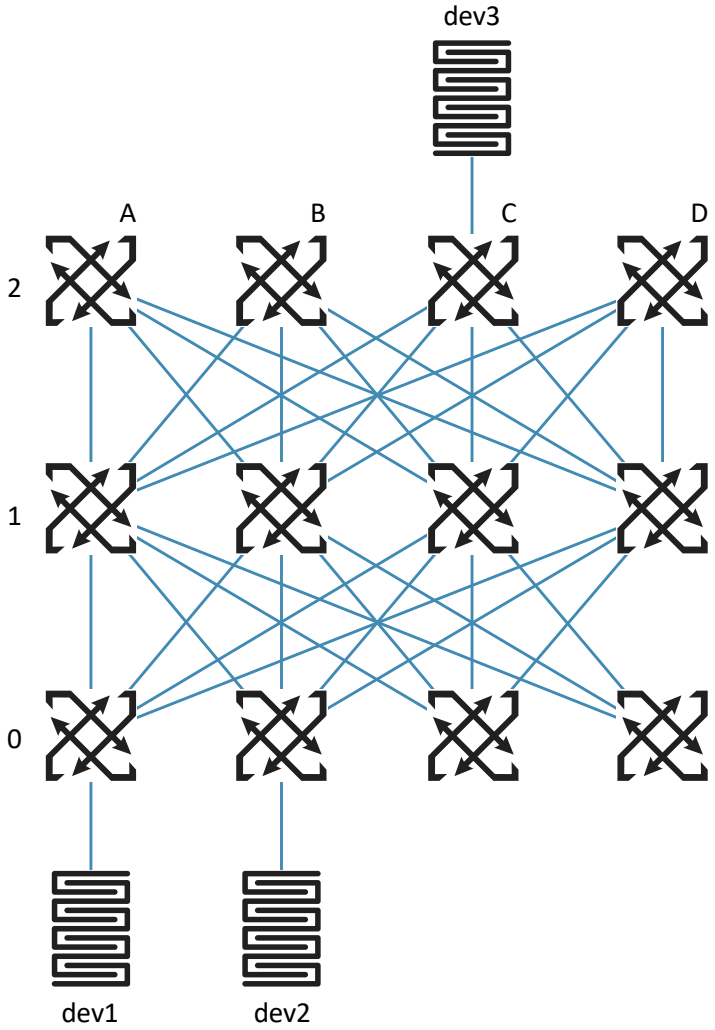


Unidirectional circuit setup

Nonblocking

- Circuits fail to make, rather than the network dropping traffic

Clos fabric

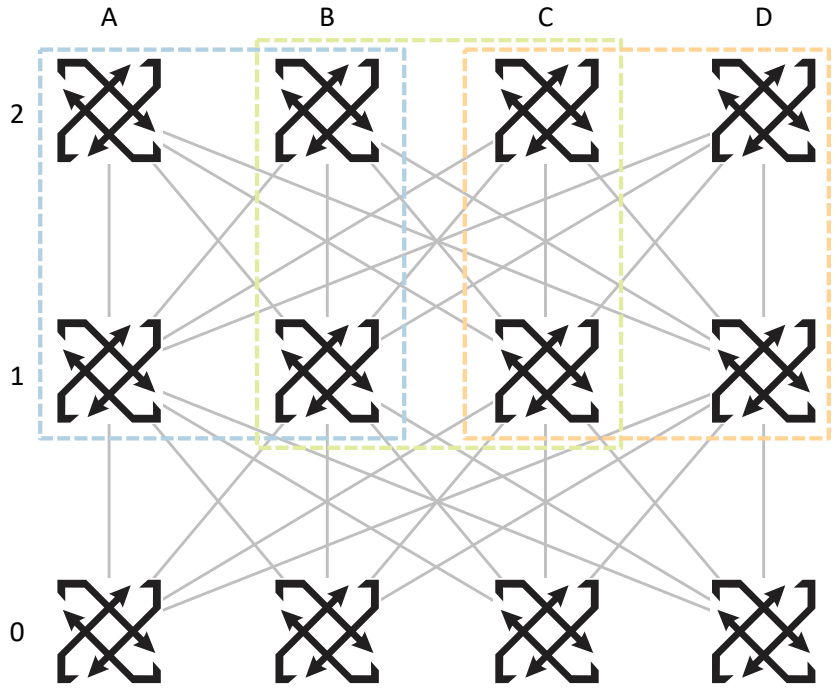


Packet switching ==
Bidirectional Forwarding
==
Folded Clos

Packets can be dropped in the network

Non-contending rather than non-blocking

Clos fabric (folded)



Regular & repeating

Nonplanar

Scale out

- Scale limited by device port count

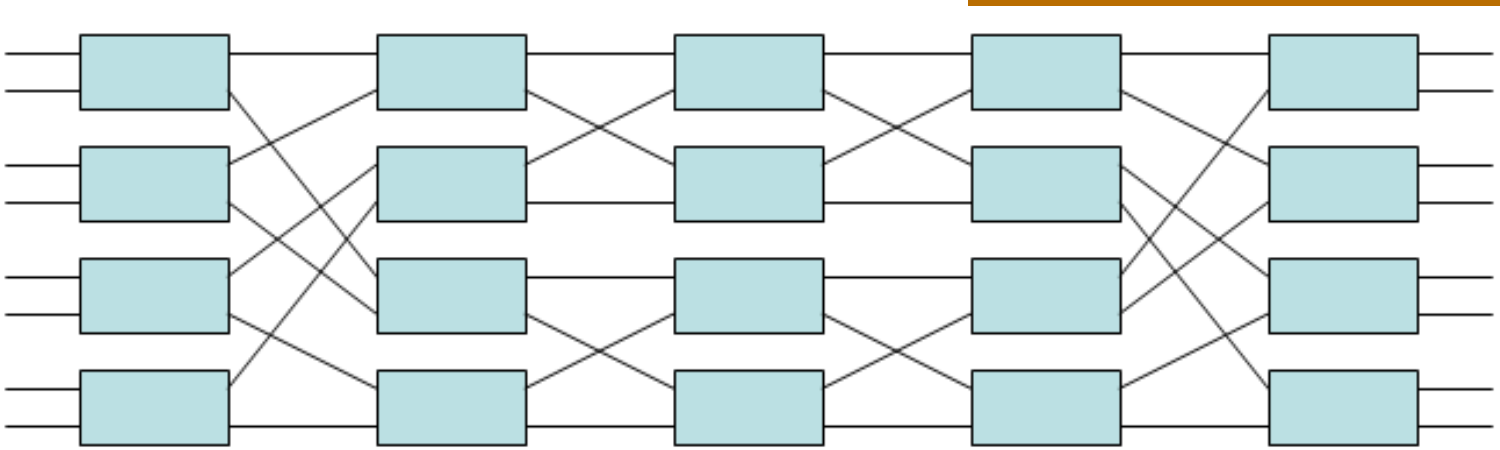
No “cross links”

- *This is not a hierarchical design*

fabric properties

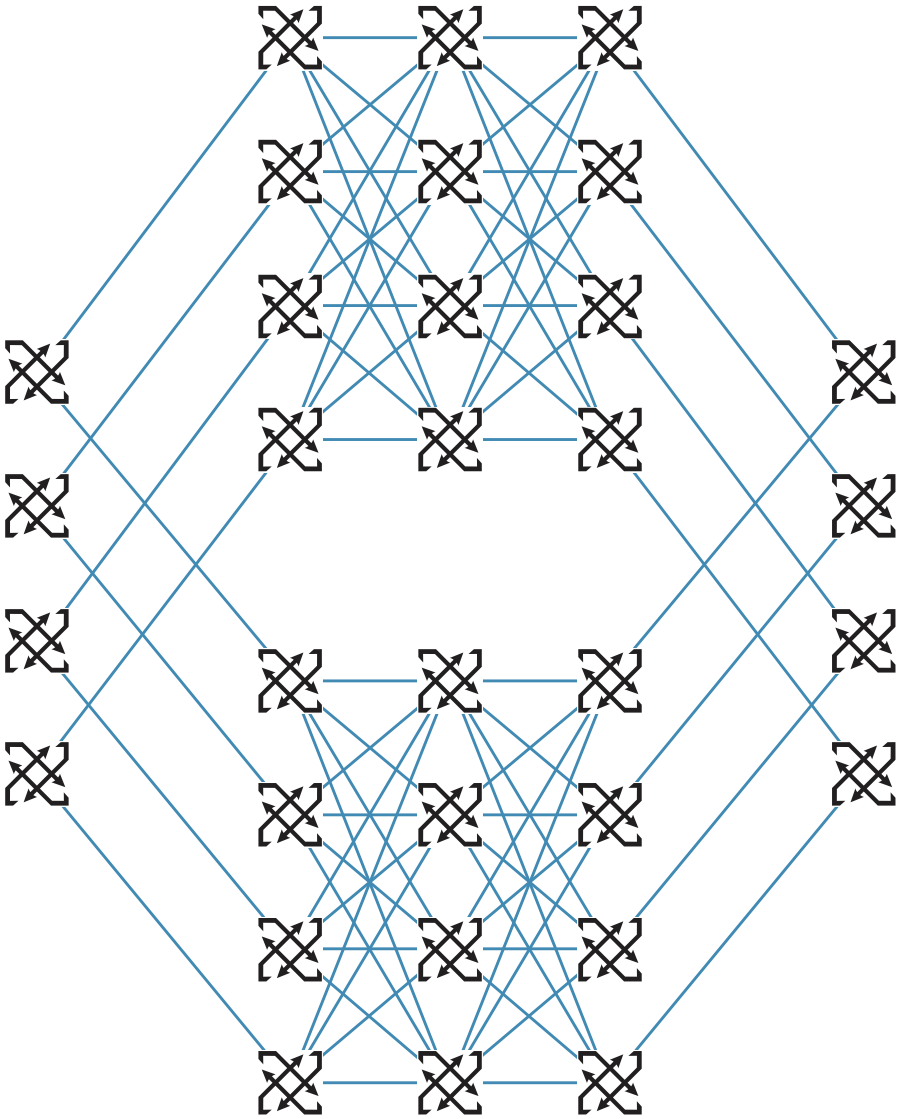
Václav E. Beneš
(although described by others before)

Rearrangeable Clos

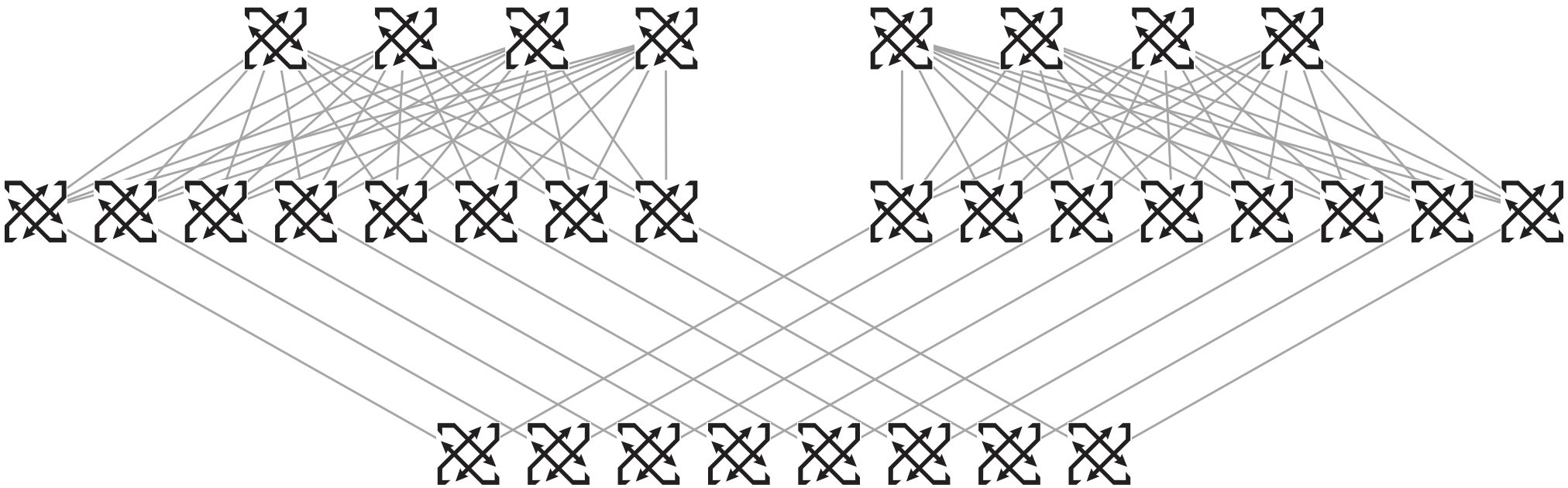


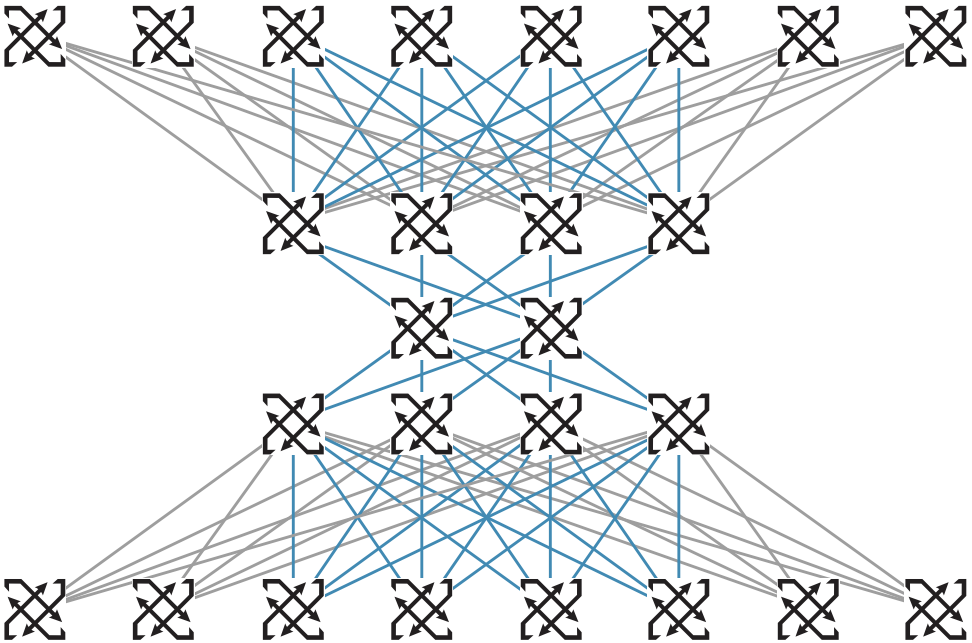
Introduces idea of planes

Benes fabric



Benes fabric





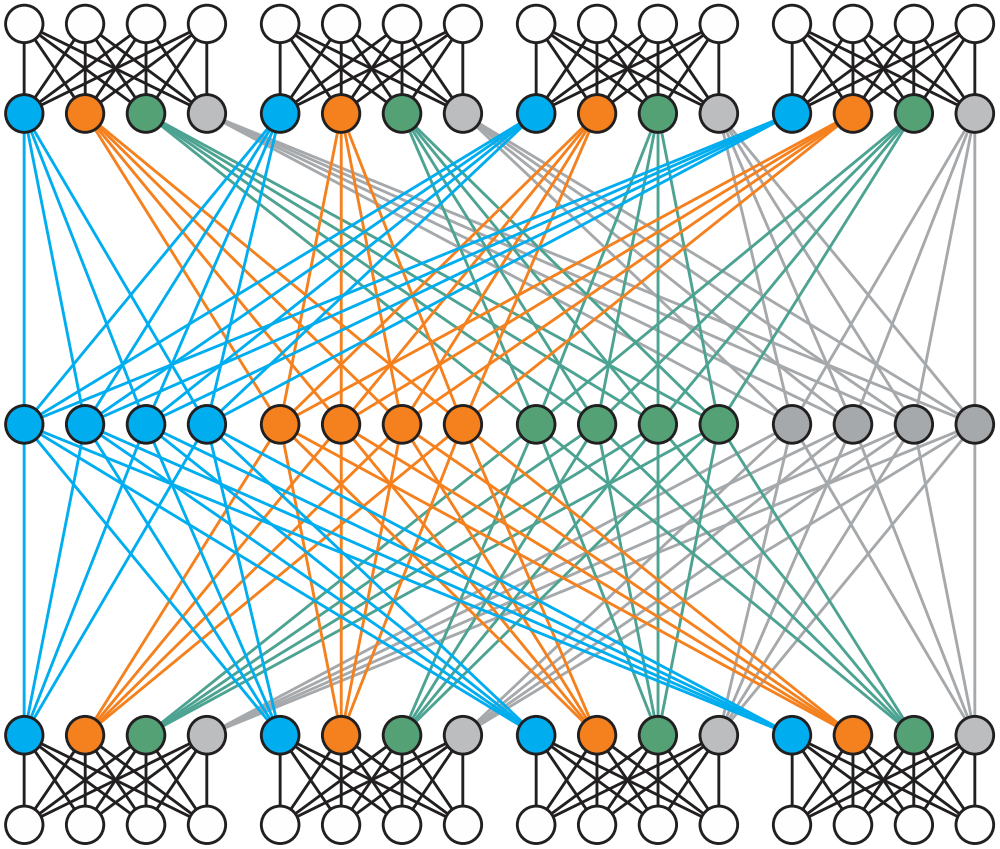
Charles E. Leiserson, 1985

Each stage has twice the bandwidth

Universal topology

Almost anything with increasing speed towards the top of fabric is now called a "fat tree"

fat tree

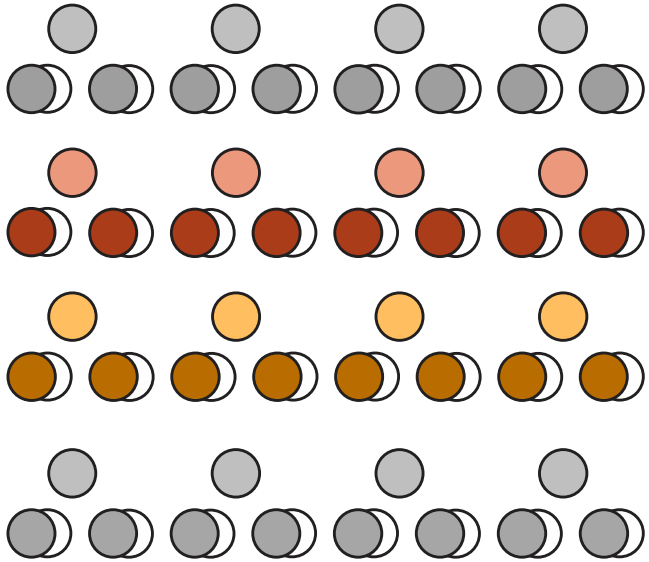


Pod size == device port count/2

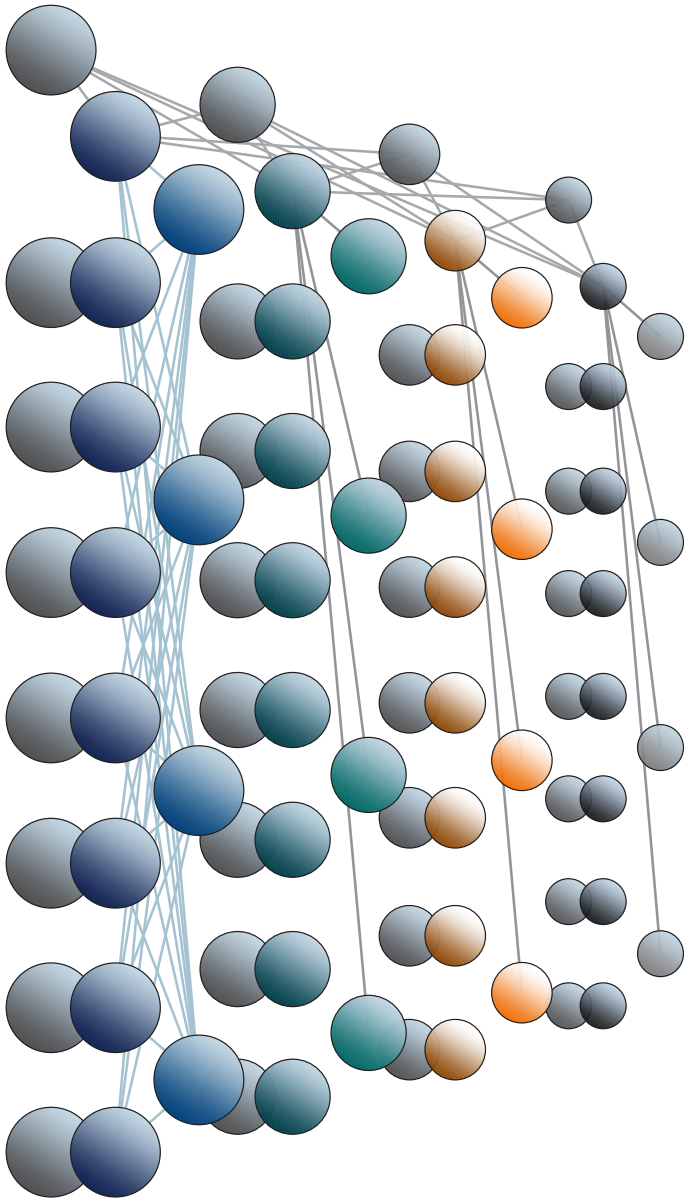
Plane count == device port count/2

"butterfly" fabric

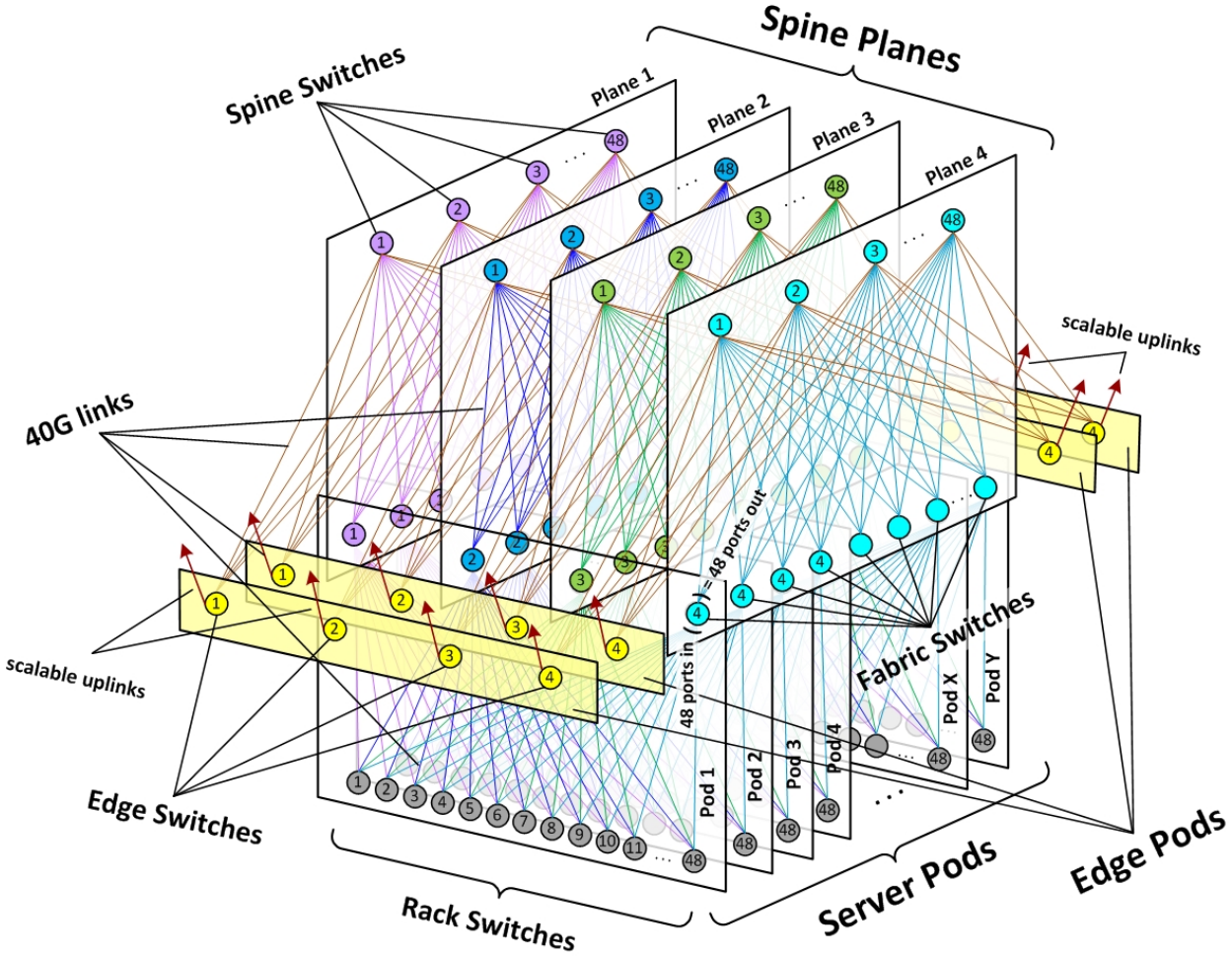
“butterfly” fabric (from the top)



also call “folded”

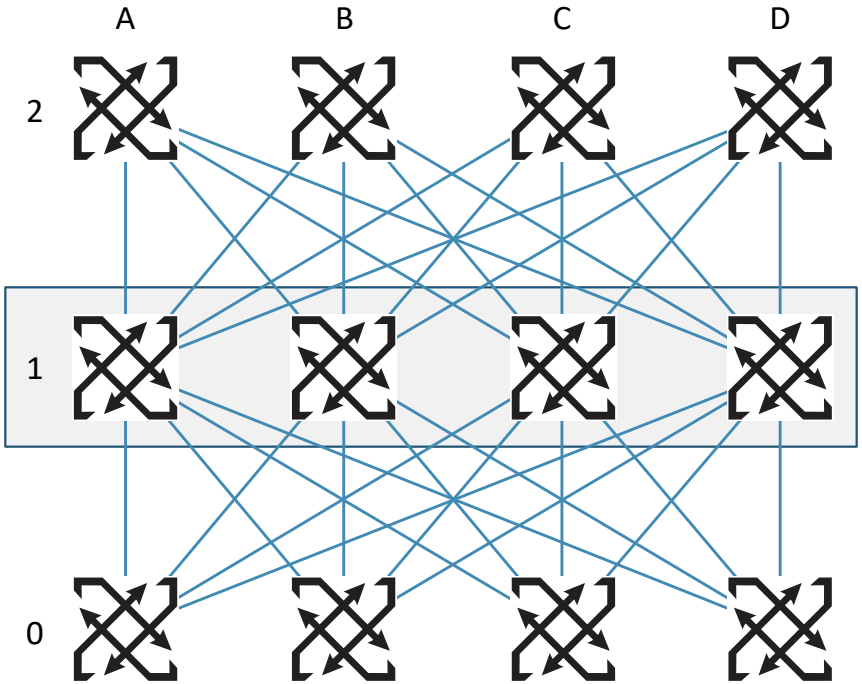


“butterfly” fabric (3 dimensional)



https://code.fb.com/production-engineering/introducing-data-center-fabric-the-next-generation-facebook-data-center-network/

"butterfly" fabric (as commonly drawn)



64x100g

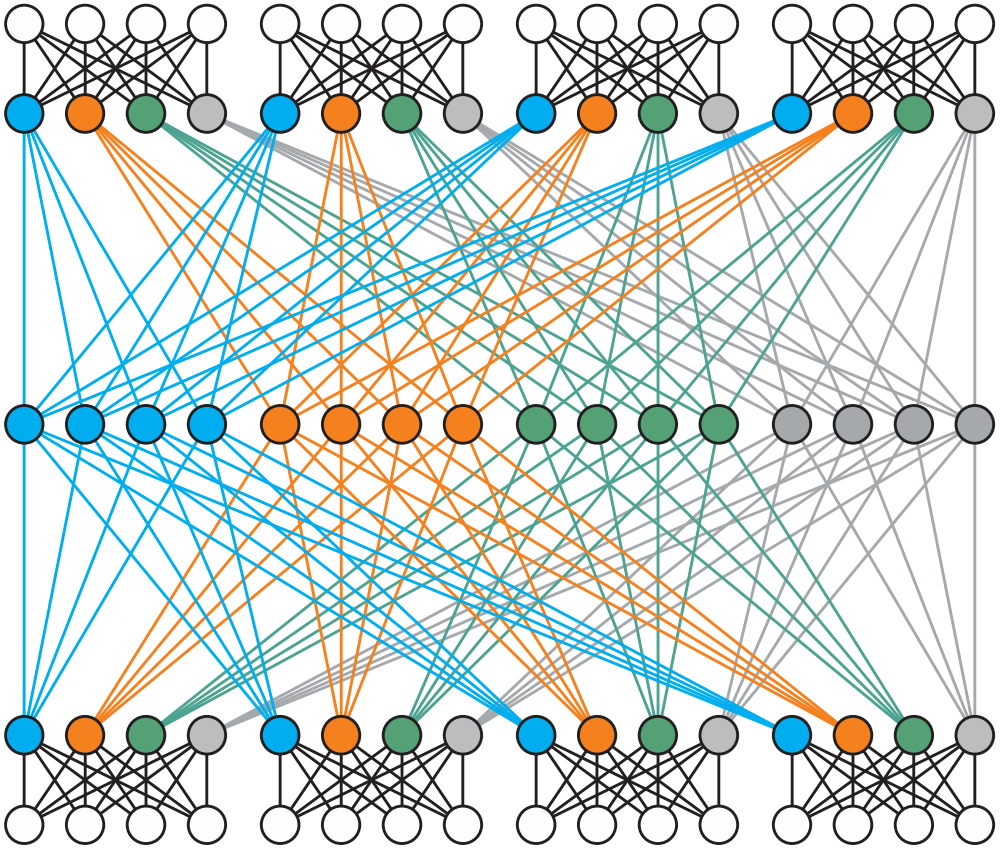
- 32x3 fabric
- 2,048x100g facing out
- 8,192x25g facing out
- 96x2ru switches per fabric

32x100g

- 16x3 fabric
- 512x100g facing out
- 2,048x25g facing out
- 48x1ru switches per fabric

Number of ports on spine == number of outfacing ports

scaling



32x100g

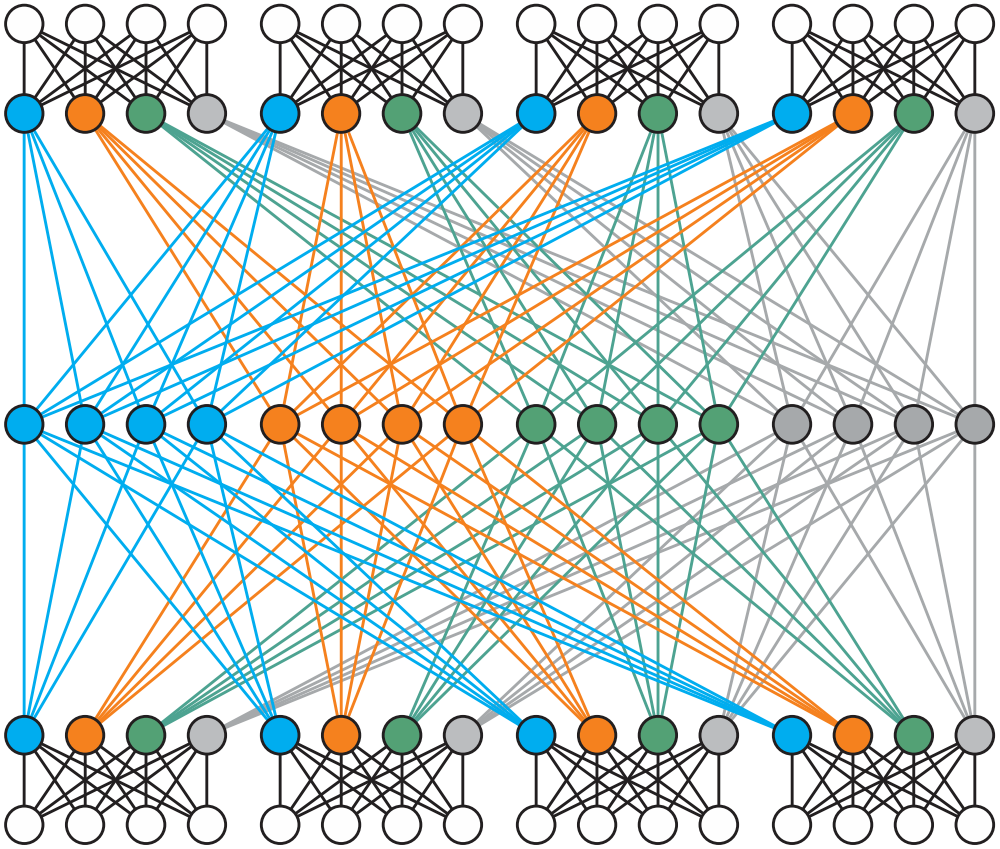
4 fabrics

8,192x25g facing out

16 fabrics

32,768x25g facing out

scaling



64x100g

4 fabrics

32,768x25g facing out

16 fabrics

131,072x25g facing out

32 fabrics

262,144x25g facing out

scaling