

Building a Terabit Internet Exchange

Arnold Nipper, CTO

MENOG6

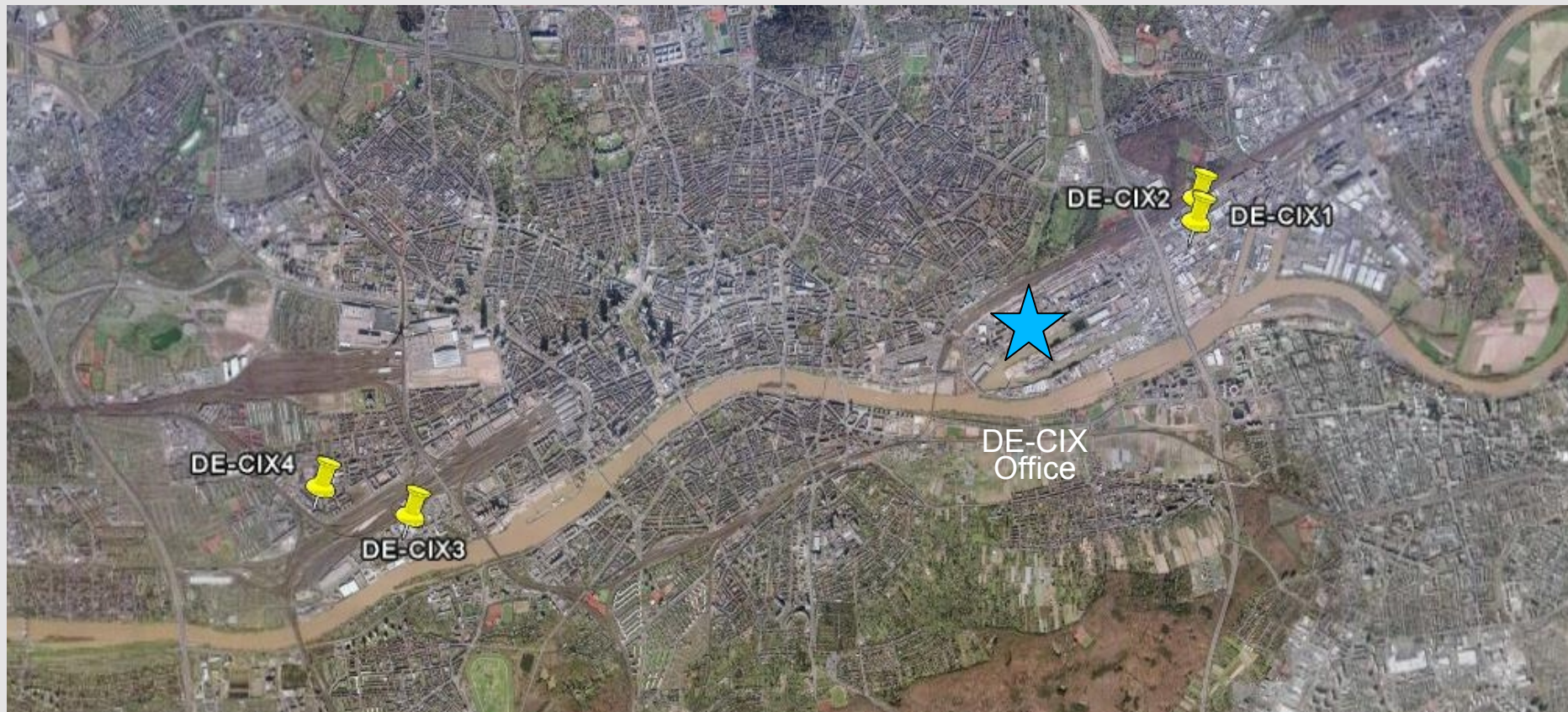
**Sheraton Riyadh Hotel,
Riyadh, 2010-04-14**



Agenda

- **Motivation**
- Design of the infrastructure
- Scaling the infrastructure
- Summary

Locations



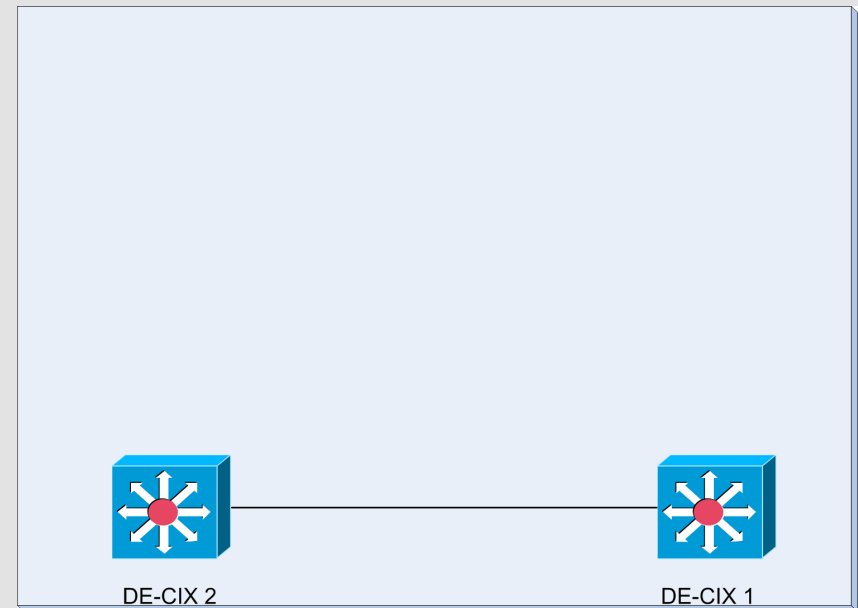
Evolution of an IXP infrastructure

- Started with single switch in May 1995
- Moved to better colocation in 1998
- Single switch is ideal
 - No need for STP
 - No need for backbone ports
 - Backplane has high bandwidth



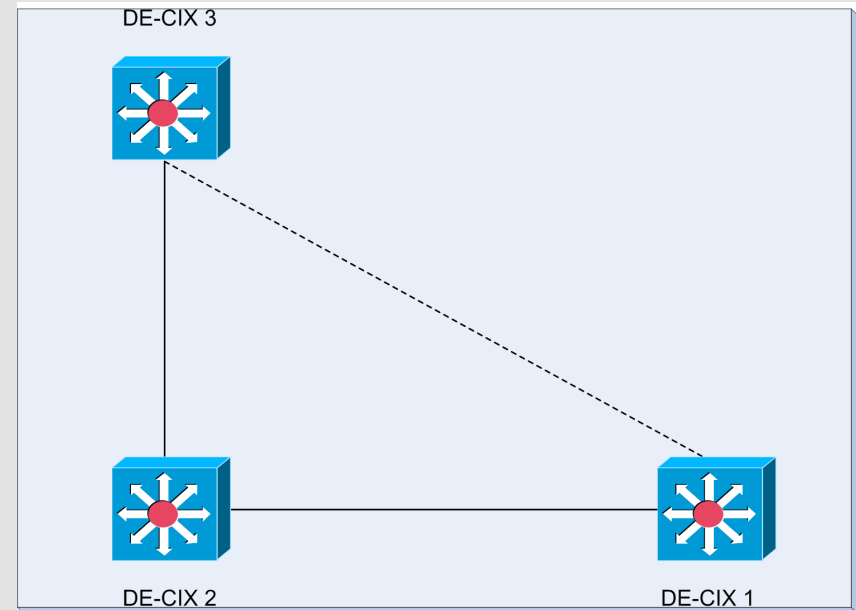
Evolution of an IXP infrastructure

- Added second switch in 2001
- Same operator but different building
- Still almost ideal



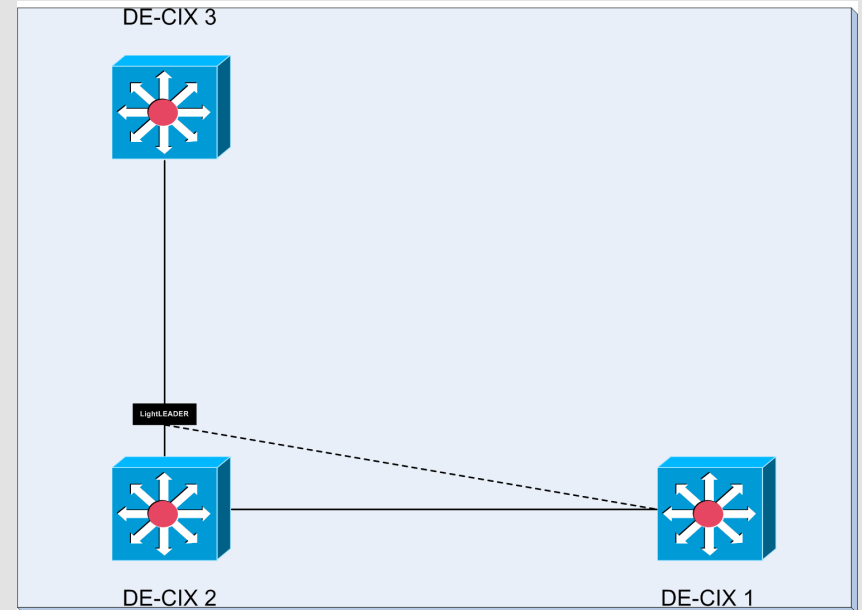
Evolution of an IXP infrastructure

- Added third switch in 2003
- Different colo operator to get better coverage
- Problems
 - Need for STP
 - Need for (expensive) DF
 - Have to deal with multiple operators



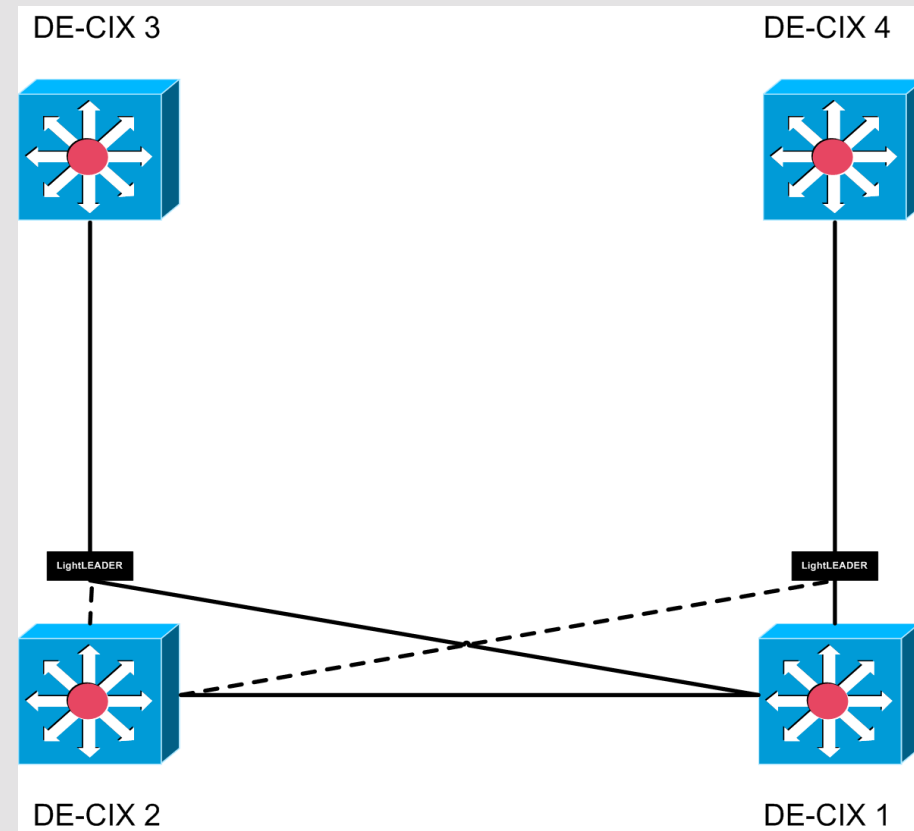
Evolution of an IXP infrastructure

- Optimise infrastructure
- Reduce complexity



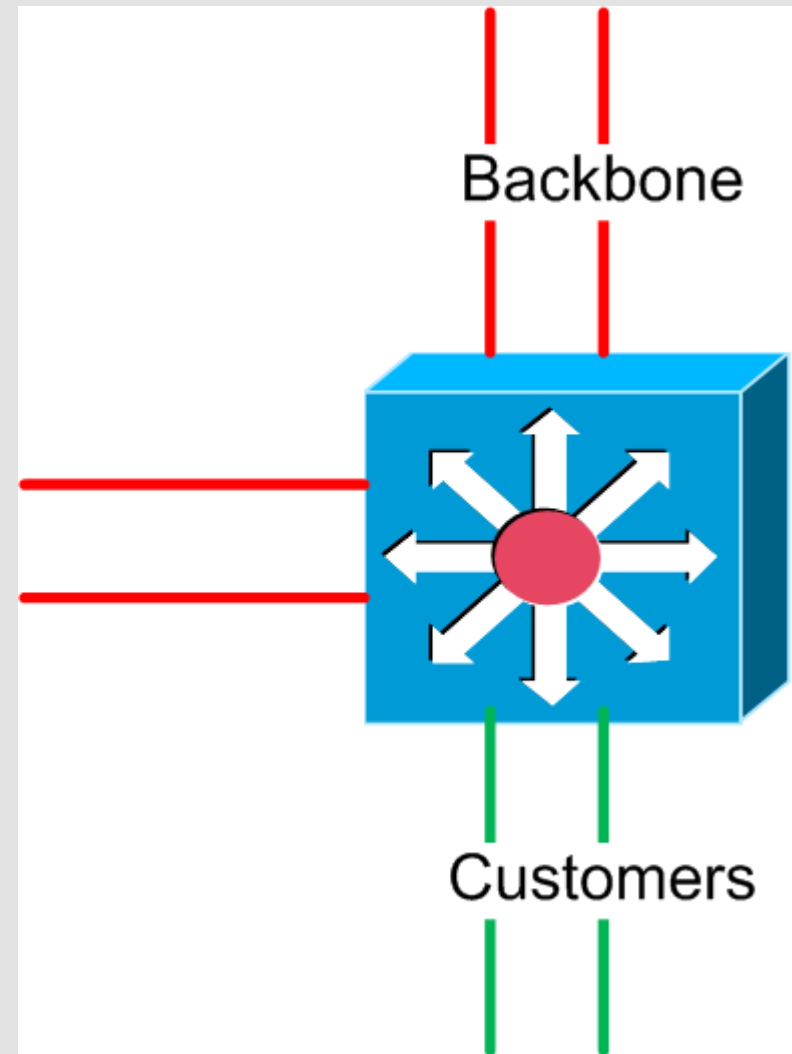
Topology until mid 2008

- DE-CIX1 and DE-CIX2 on the same campus, DE-CIX3 and DE-CIX4 in separate colocations
- combined edge and distribution in resilient star
- grown over time since 1995 from single switch, dual edge, STP resilience to fibre protection based resilience



Topology until mid 2008

- mix of edge and distribution
- hard to plan for growth in backbone bandwidth and ports for customers
- outage of distribution switch also affects customers



Agenda

- Motivation
- **Design of the infrastructure**
- Scaling the infrastructure
- Summary

Design of new infrastructure

- which topology?
 - star
 - ring
- which technology for interconnects?
 - simple dark fibre
 - DWDM
- which technology for resilience?
 - STP
 - MRP (Metro Ring Protocol)
 - Layer 0

Design of new infrastructure

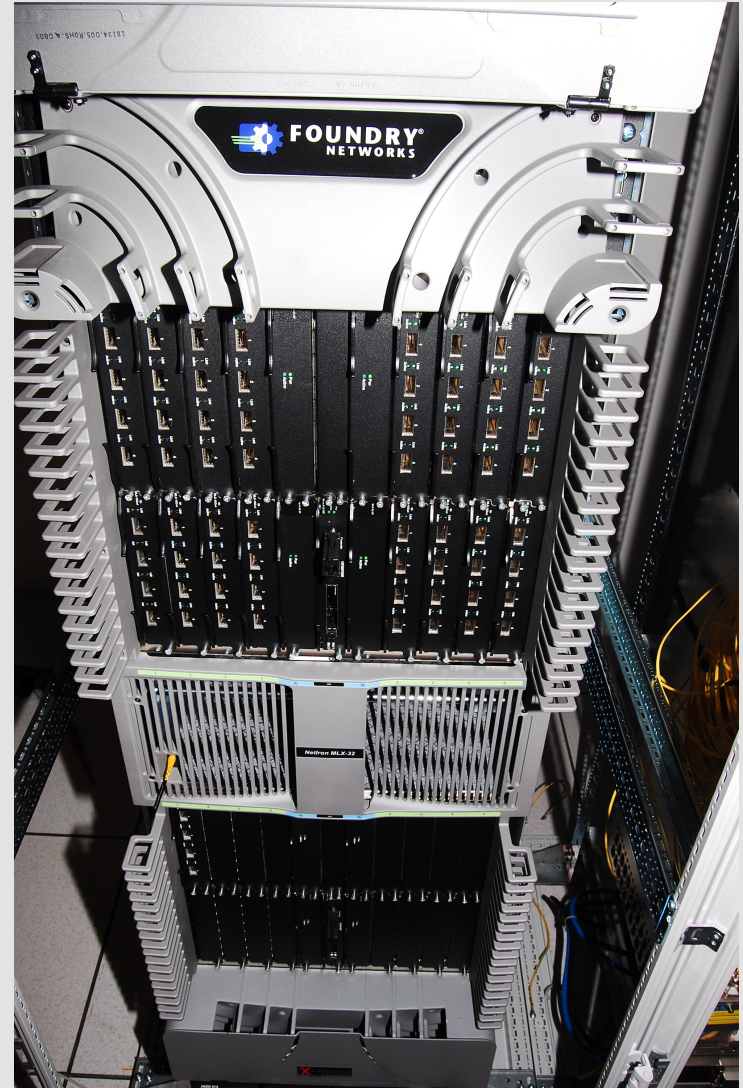
- star topology
- DWDM technology for interconnects
- Layer 0 technology for resilience
- other considerations
 - scalable, simple and robust
 - cost-efficient
 - easy migration

Building blocks

- core switches (star)
- DWDM de/muxes (interconnect)
- optical switches (resilience)

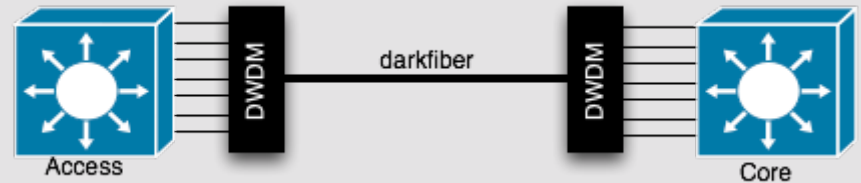
Cores

- where to locate
 - new POP?
 - existing POP?
 - which POP?
- Take existing POP's
- Hardware is 128x 10GE port Brocade MLX32



DWDM / dark fibre

- passive
- up to 32 channels
- diverse routes
- one pair per edge
- Cube Optics as a system partner

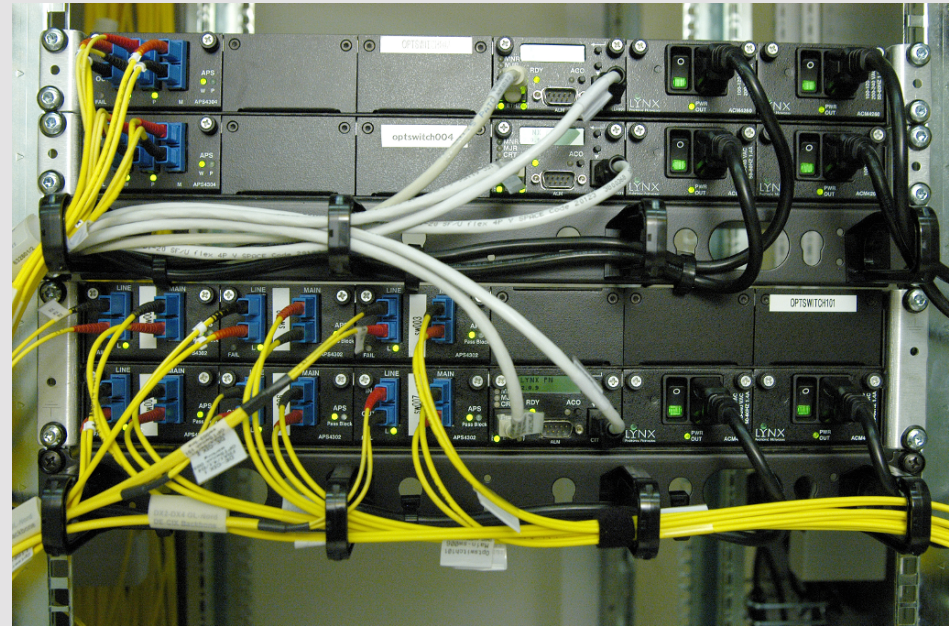


darkfiber routes

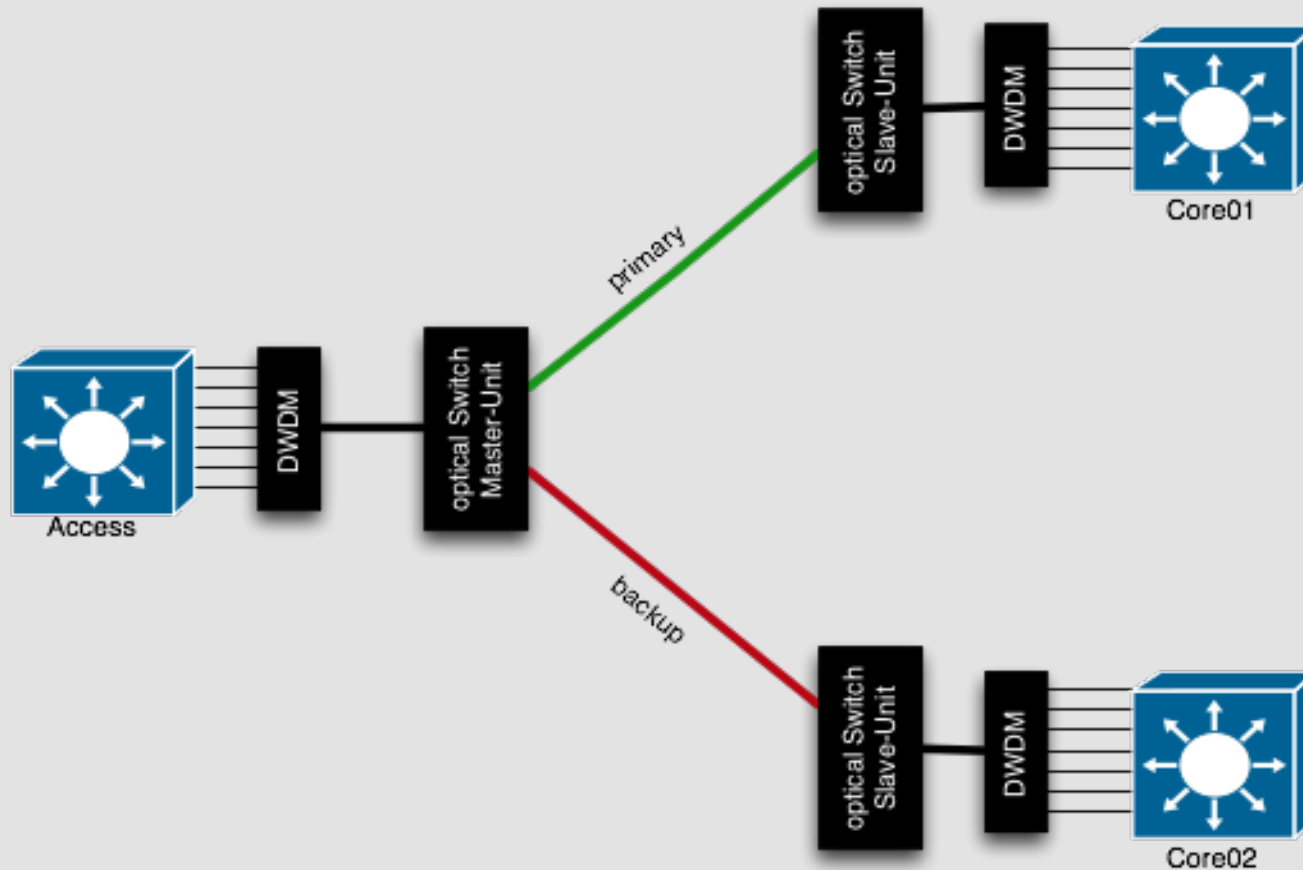


Resilience

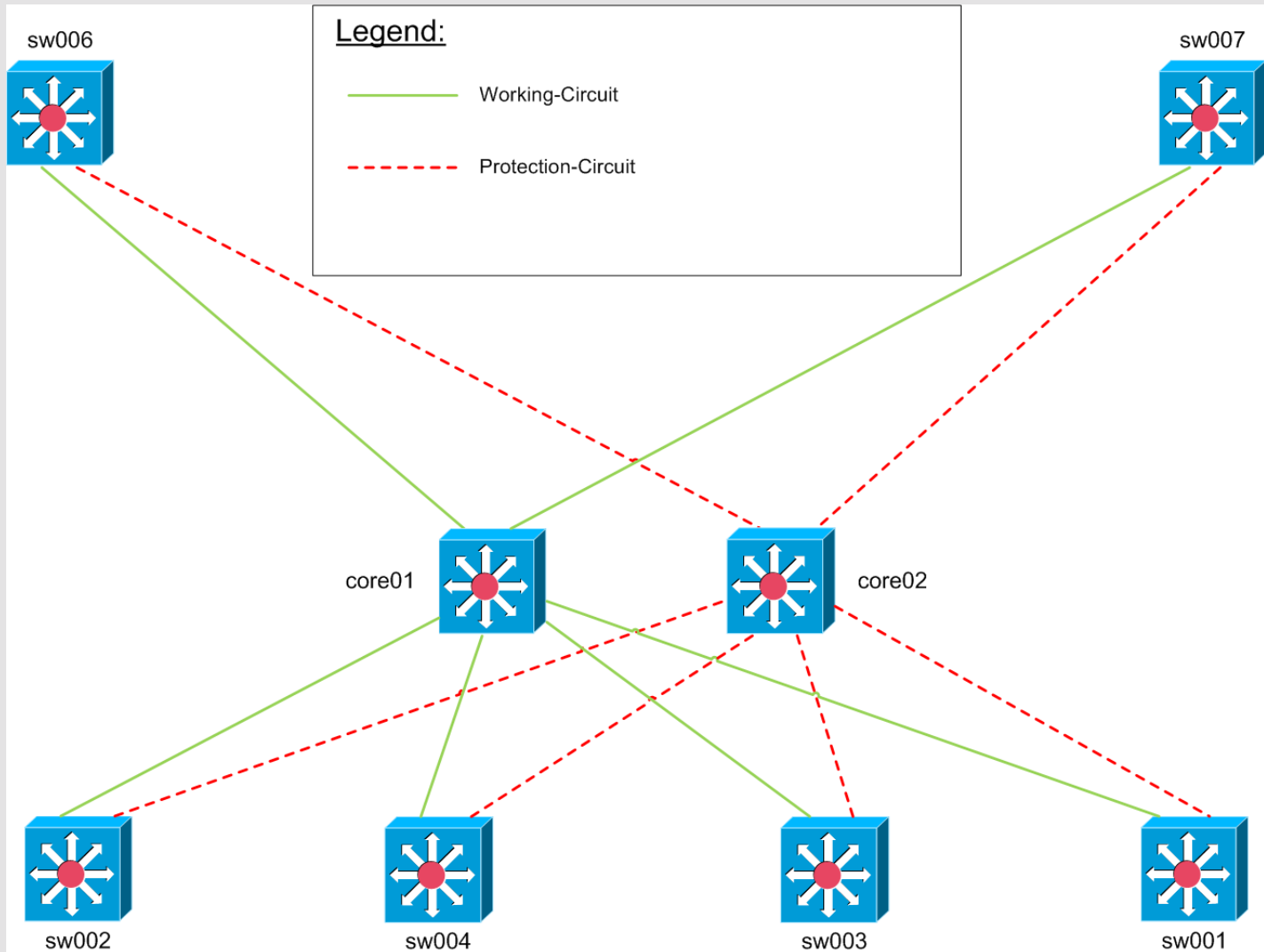
- using simple fibre switches
- power is only needed to switch fibres
- primary and backup path going via diverse fibres
- master/slave
- Lynx networks
Lightleader
- all optical switches synchronously switch over to standby core



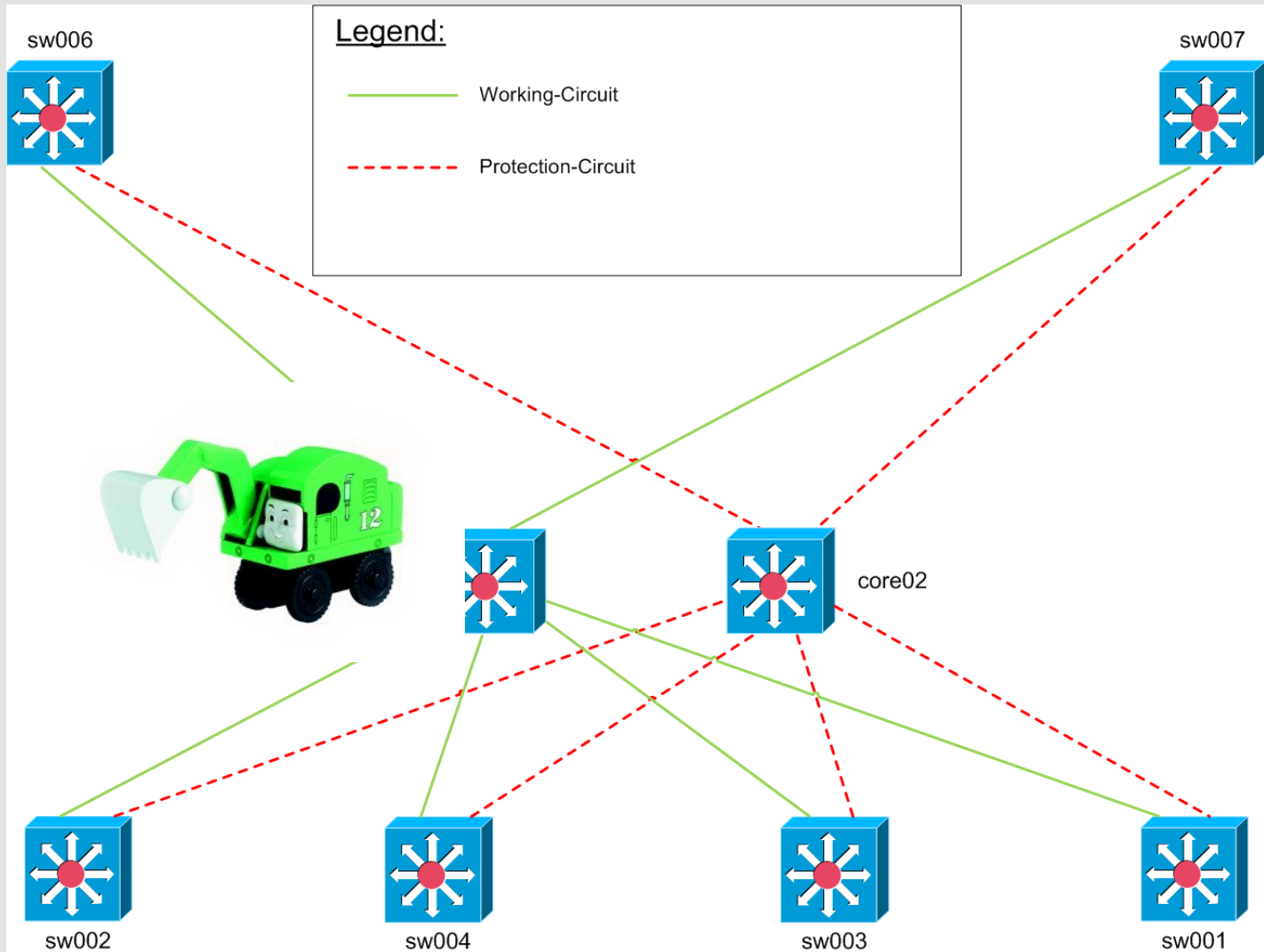
optical protection



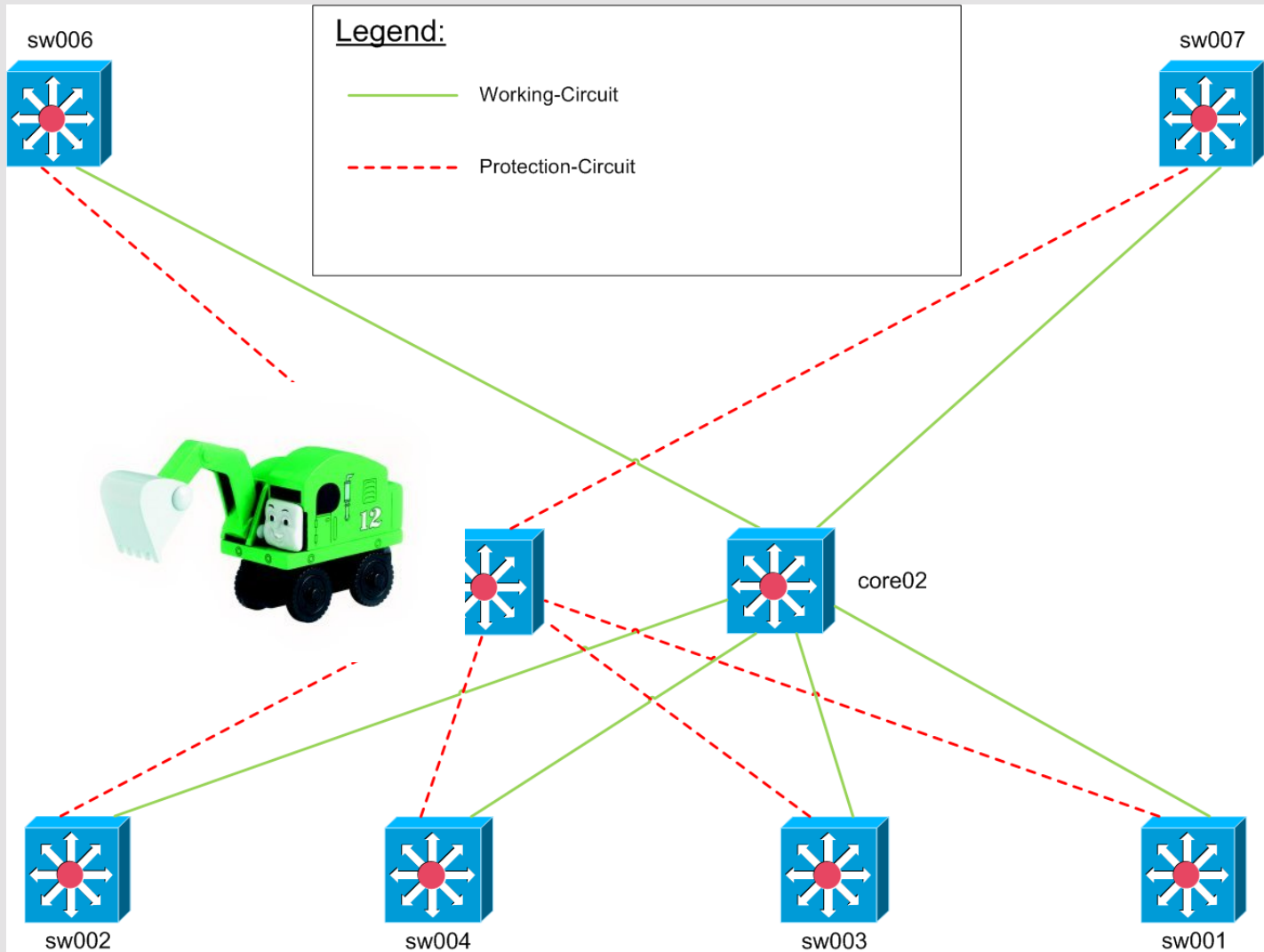
Normal operation mode 😊



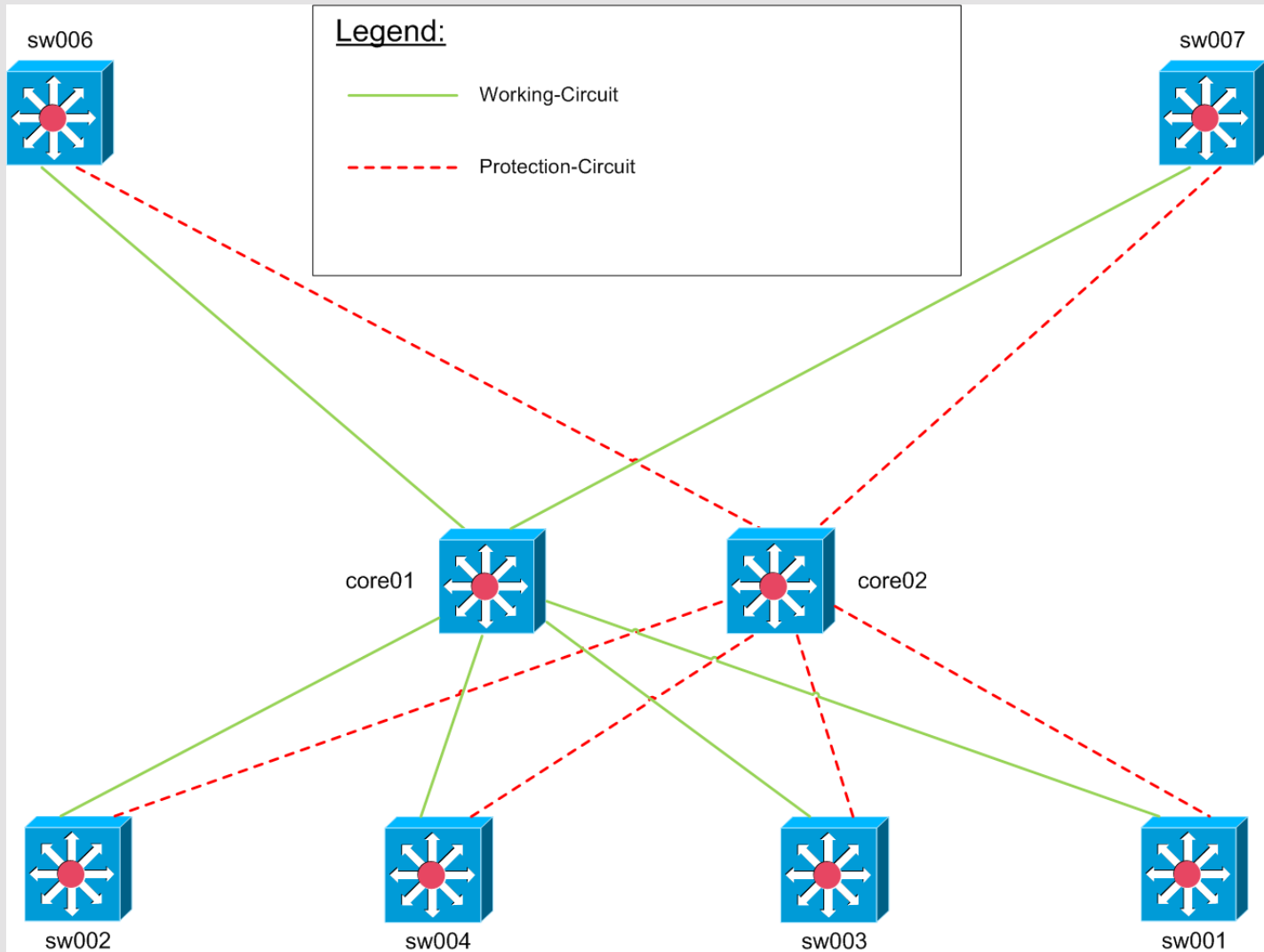
Something went wrong ☹️



Protection operation mode ☹️



Problem fixed → Normal operation mode 😊



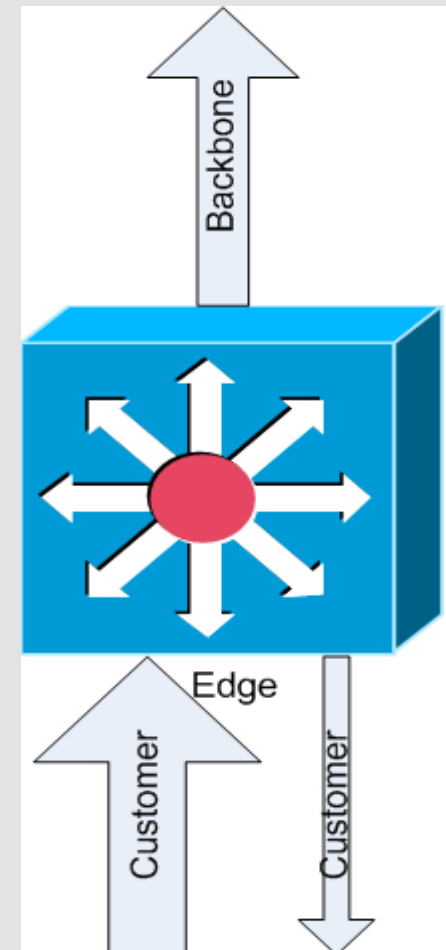
Agenda

- Motivation
- Design of the infrastructure
- **Scaling the infrastructure**
- Summary

Edge

- more efficient to use big edges

| #switches | %local traffic |
|-----------|----------------|
| 1 | 100 |
| 2 | 50 |
| 3 | 33 |
| 4 | 25 |
| 5 | 20 |
| 6 | 16 |
| 7 | 14 |
| 8 | 13 |
| 9 | 11 |
| 10 | 10 |



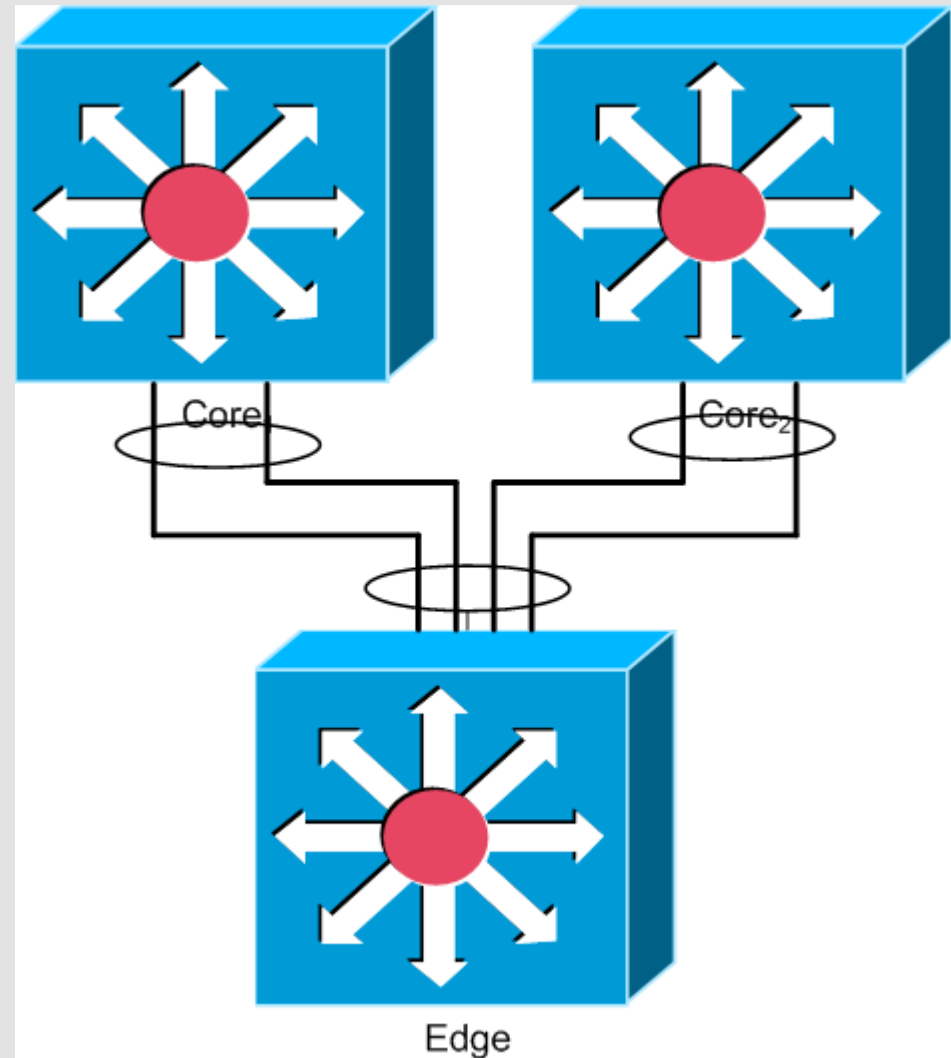
Edge

- simply add another switch
- Virtual chassis
 - Easy to grow customers across edges



Core

- use bigger switches
- „stack“ core switches when you can't get bigger ones
- will upgrade core to 2 Brocade MLX32 per core soon
- scales up to 32 core members (~41TBps)



Summary

- Building resilient, scalable highbandwidth IXP is possible
- Scaling is possible up to >40 Tbps which will safely take us to when 100Gbps ports are widely available

Thanks!

DE-CIX Management GmbH

Lindleystr. 12

60314 Frankfurt a.M.

info@de-cix.net

+49 69 1730 902 0



DE-CIX office @ Frankfurt docklands