



DE·CIX

Where networks meet

DE-CIX next generation architecture

2011-09-21

European Peering Forum 6 - Budapest

Arnold Nipper

CTO/COO

arnold.nipper@de-cix.net



Where networks meet



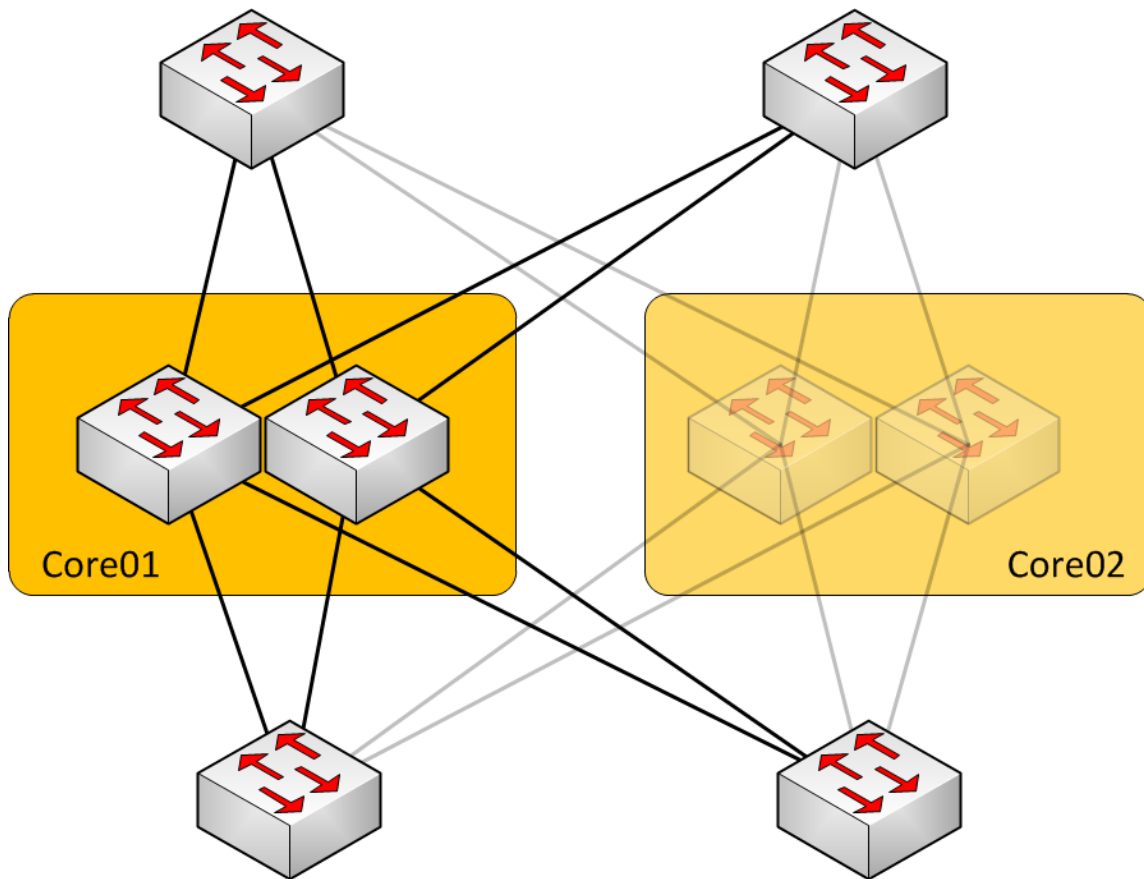
DE-CIX



DE-CIX Where networks meet



DE-CIX – Current Topology



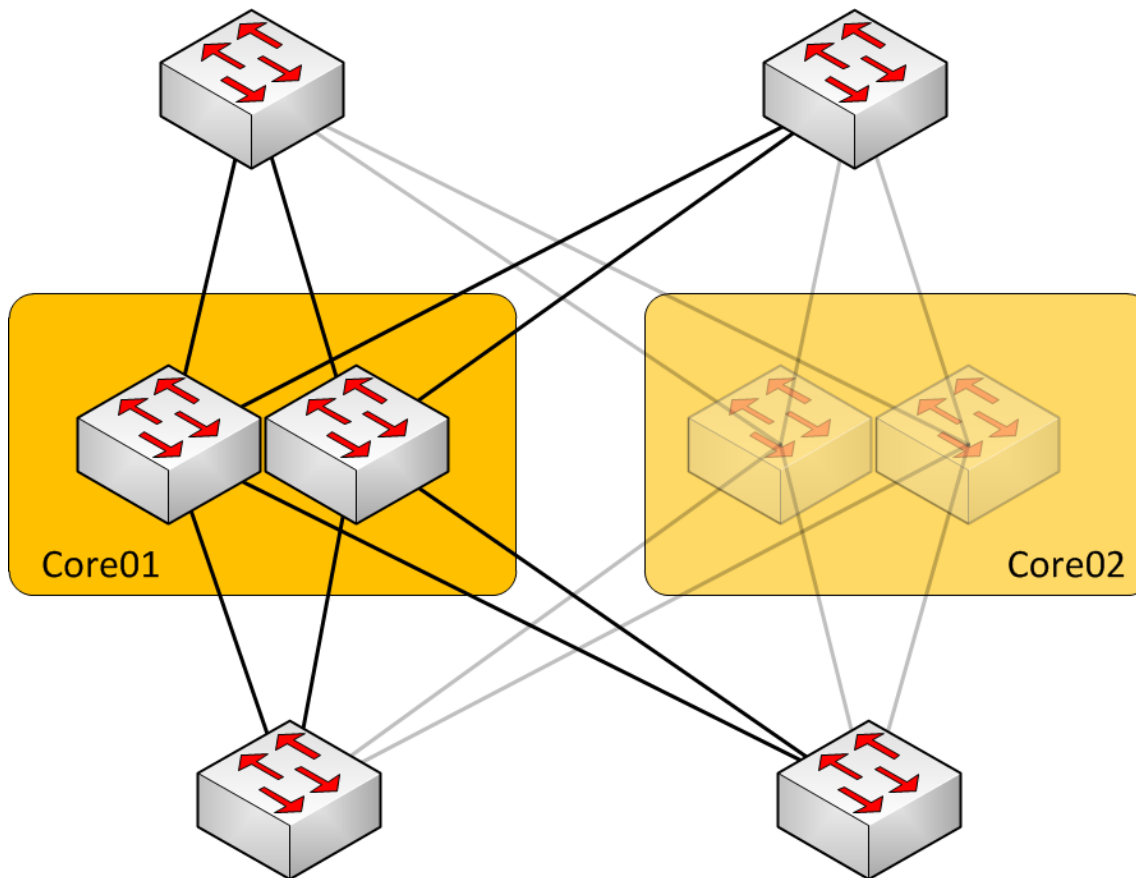
- Resilient Star
- Dual Core
- Layer 1 Redundancy
- 700+ Links have to be switched during failover
- Traffic still growing massively, now peaking at 1.5Tbps



DE-CIX Where networks meet



DE-CIX – Current Issues



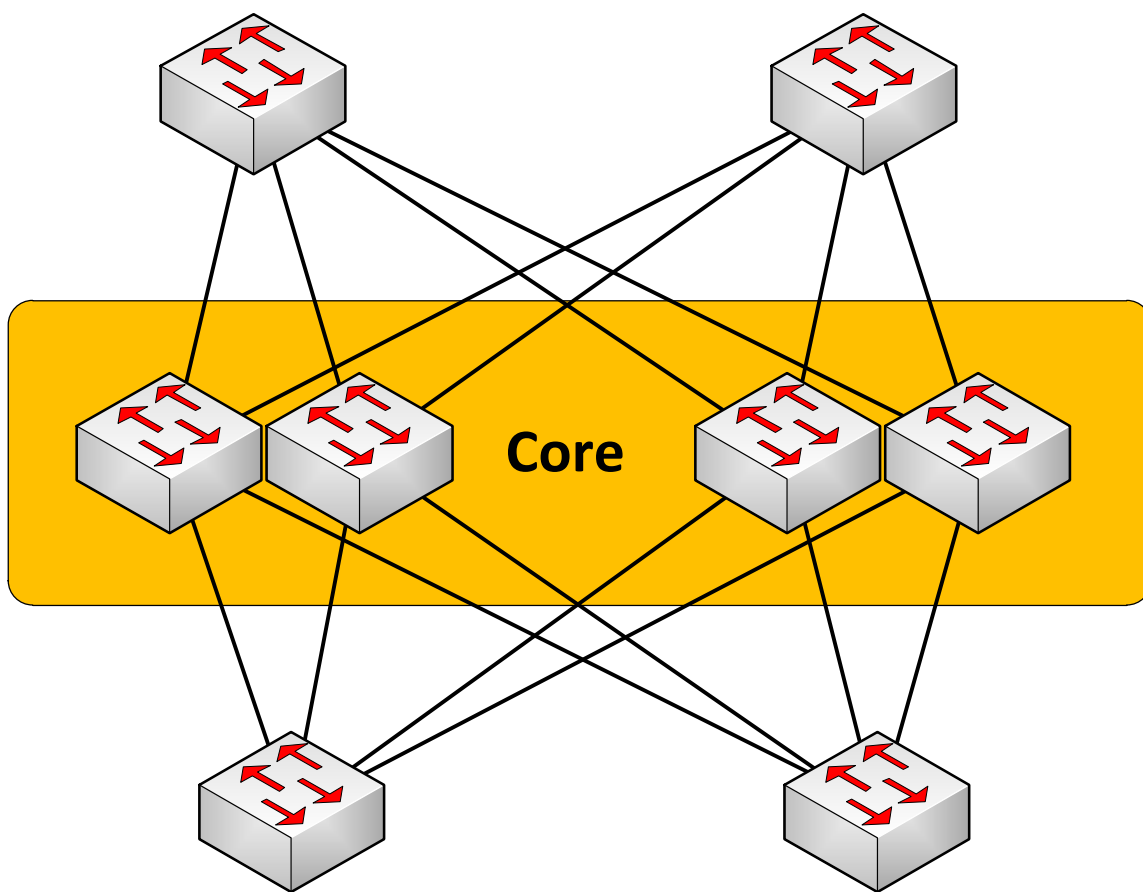
- No LACP
- MAC learning problems
- Blackholing
- No solution for loadbalancing and redundancy above Layer 1



DE-CIX Where networks meet



DE-CIX – Target Topology



- all active Core
- LACP
- core redundancy n+1
- automatic blackhole avoidance



DE-CIX Where networks meet



DE-CIX – Target Topology implementation

- MPLS / VPLS
 - additional abstraction layer
 - complex
 - difficult troubleshooting / management
- SPB (Shortest Path Bridging, 802.1aq)
 - available only later and looks not appropriate for IXPs
- TRILL
 - available and looks very promising



DE-CIX Where networks meet



TRILL – What is TRILL?

- **TRILL** – TRansparent Interconnection of Lots of Links.
A standard specified by the IETF (Internet Engineering Task Force) TRILL Working Group, RFC5556
- **RBridge** – Routing Bridge
A device which implements TRILL
- **RBridge Campus** – A network of RBridges, links, and any intervening bridges, bounded by end stations / layer 3 routers.



DE-CIX Where networks meet



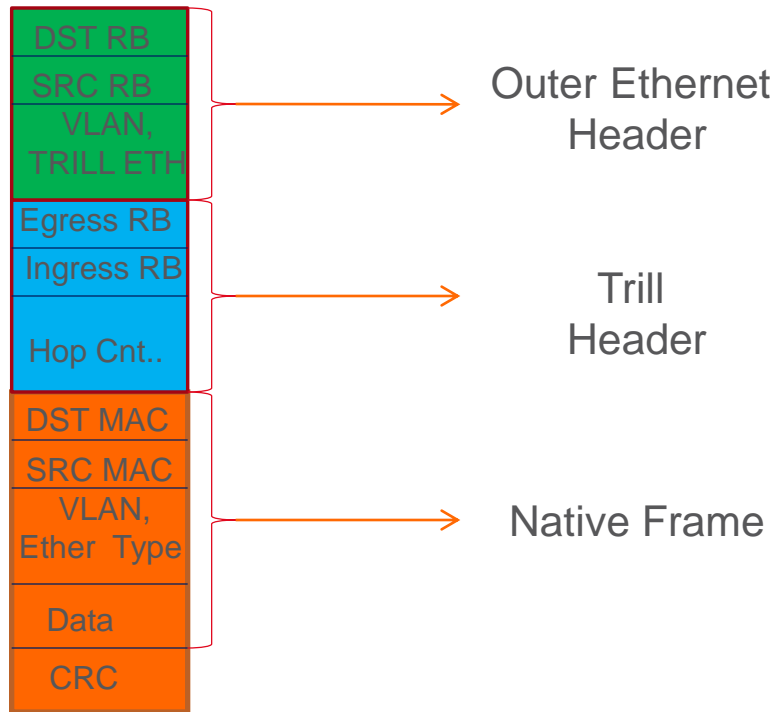
TRILL – What is TRILL?

Basically a simple idea:

- Encapsulate native frames in a transport header providing a hop count.
- Switch the encapsulated frames using FIB buildt by IS-IS.
- Decapsulate native frames before delivery.
- Backwards compatible



TRILL – Frame Encapsulation

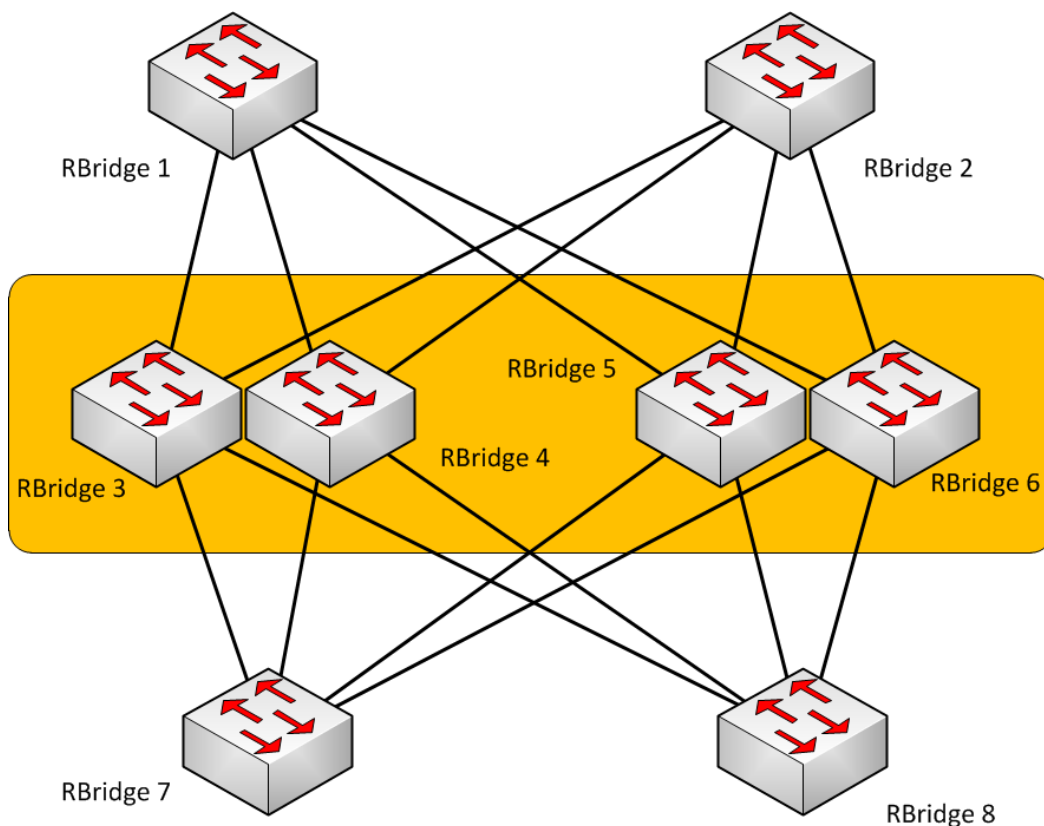




DE-CIX Where networks meet



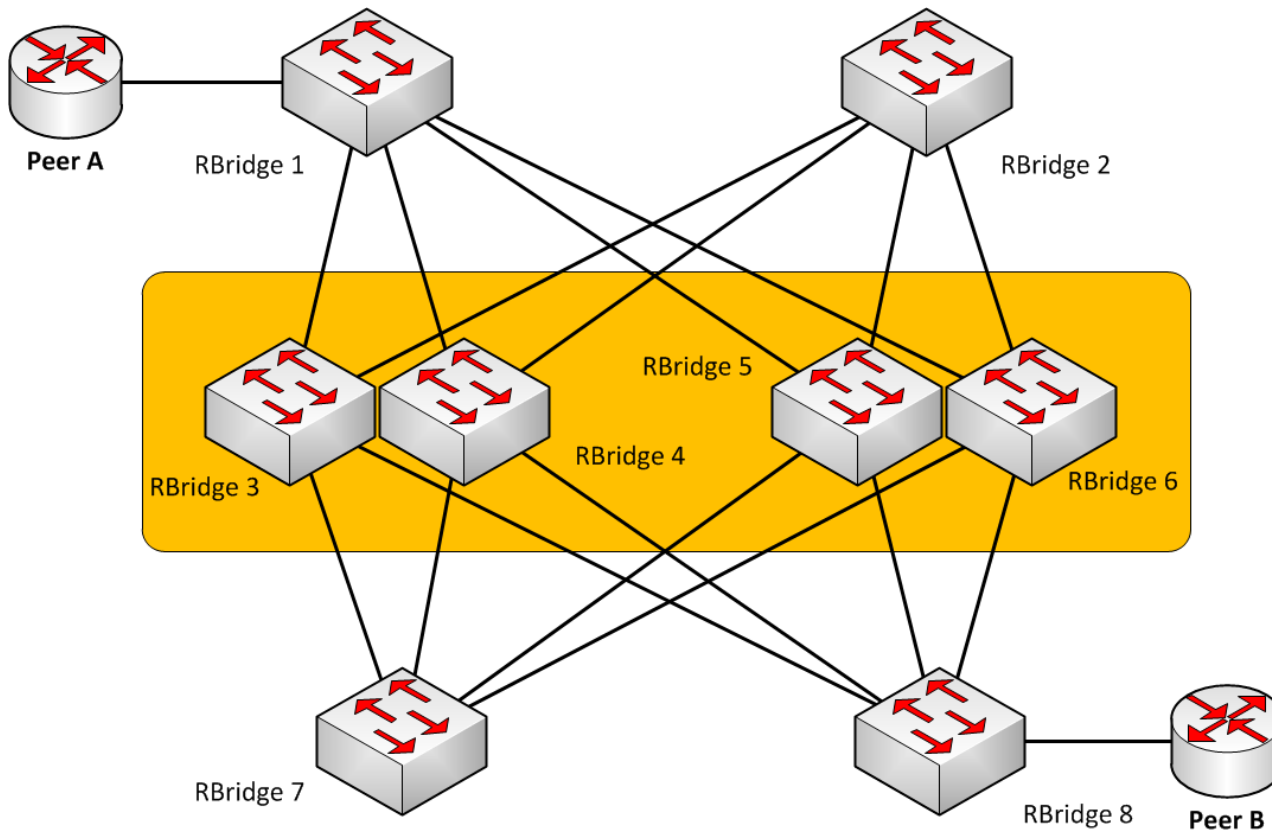
TRILL @ DE-CIX



- All devices TRILL enabled
- No unused redundant links
- ECMP load balancing
- Big Edge switches
- Small core switches



TRILL – Unicast Paket Flow



Peer B
Peer A
Ether Type
Data
CRC

MAC Table

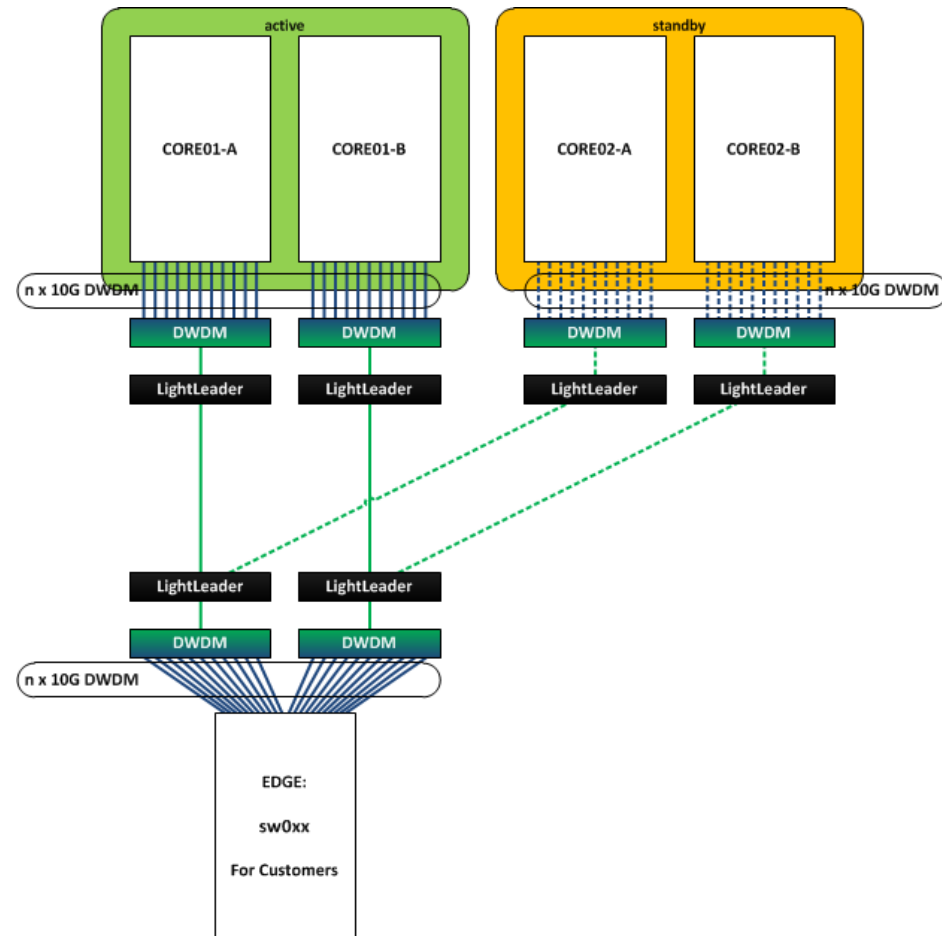
MAC DA	Egress
Peer B	RBr 8
....

Dest	Next Hop	Link
RBr1	-	-
RBr2	RBr5	c
...
RBr8	RBr4	b



DE-CIX Production Network 2011

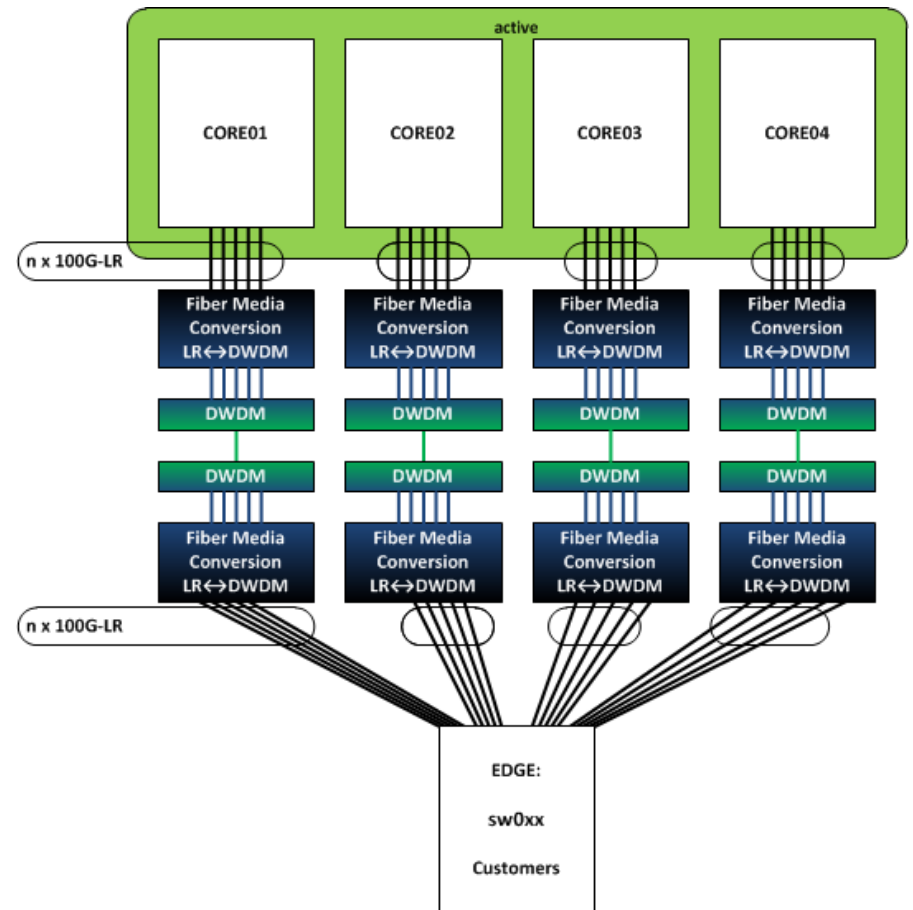
- Passive DWDM equipment with up to 32 channels
- DWDM-XFP optics used at Core and Edge Switches
- failover is done by fast, mechanical optical switches (Lynx Lightleader)
 - in case s.th. breaks
 - scheduled every two months
 - post processing is more and more time consuming (adjusting light levels)
- provisioning of additional links can not be fully tested on the standby links





Optical Network Design

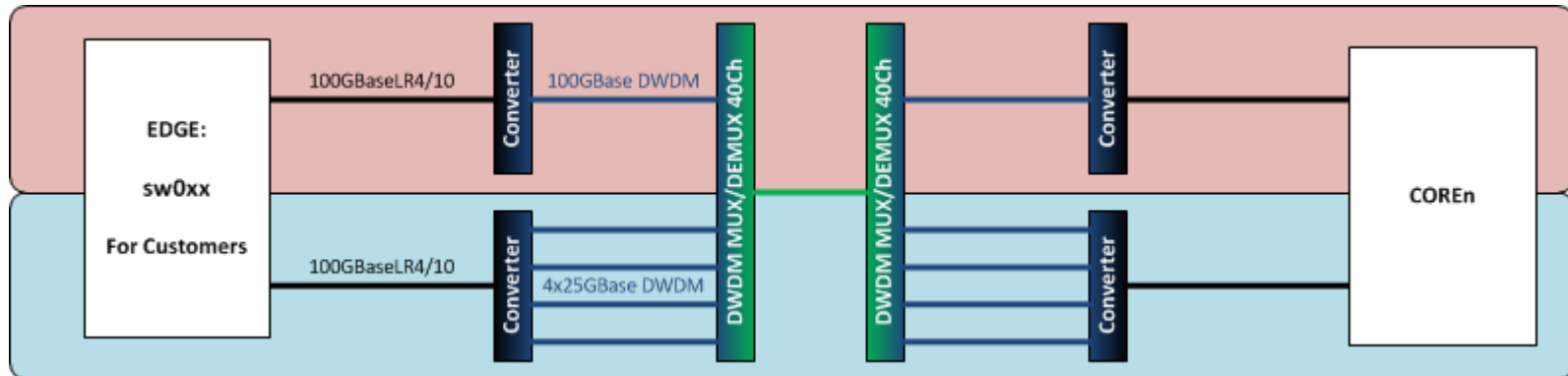
- Using 100G connections instead of 10G connections
- active DWDM equipment to be more flexible on client side
- new DWDM equipment with up to 80 x 100G connections per Dark Fiber
- all links permanently active





solution approaches

- active DWDM with coherent optics (100G per wave)
 - expensive because of complex technology used
 - available by the middle of 2012



- active DWDM with 4x25GBase transponder optics
 - use 4 wave length per 100G on DWDM
 - less expensive due to less complex used technology
 - available by the middle of 2012

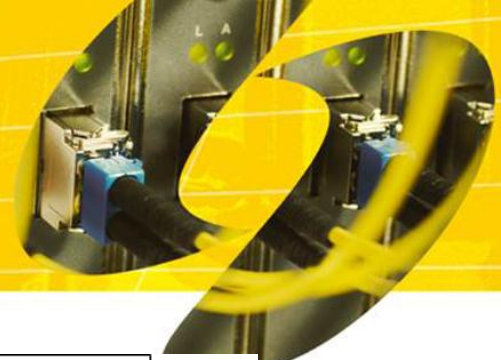


DE-CIX Where networks meet



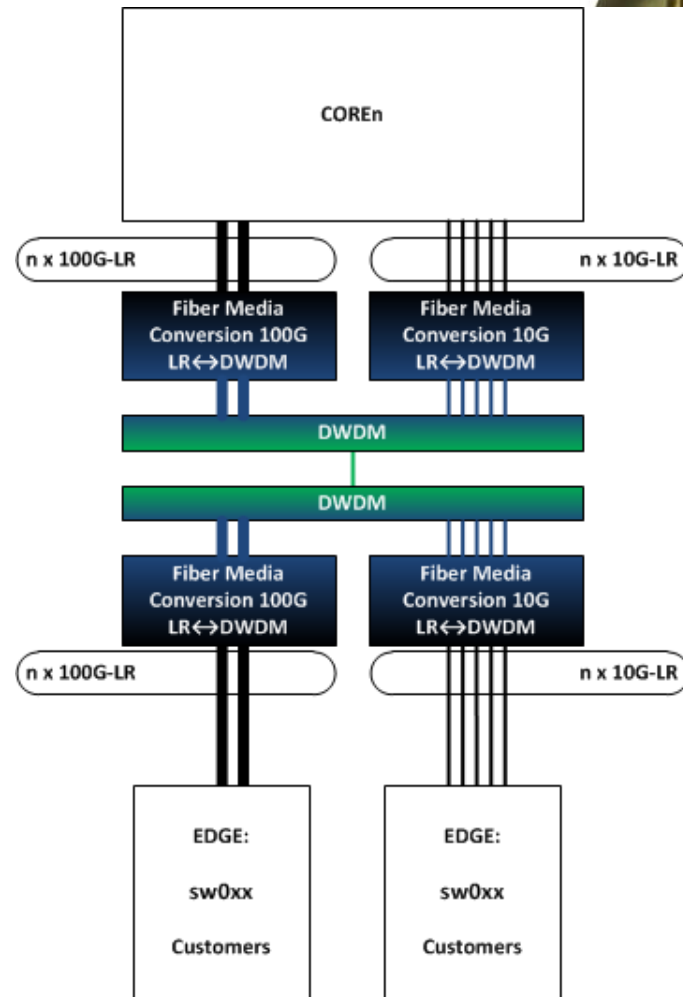
solution approaches (cont'd)

- 100GBase DWDM-Transceiver directly used within the switch
 - not available in the medium term
 - high dependency on form factor
- on Client side 100GBase-LR or 10x10MSA 2km



Possible Set-Up at DE-CIX

- usage of active DWDM
 - 10G and 100G in parallel for smooth transition





DE-CIX Where networks meet



Thank you

Join DE-CIX now!

DE-CIX Competence Center
Lindleystrasse 12
60314 Frankfurt/Germany

Phone +49 69 1730 902 - 0
info@de-cix.net



DE-CIX Competence Center @
Kontorhaus Building
Frankfurt Osthafen (Docklands)