Get ready for the 800GE reality

Jonas Vermeulen

Technical Sales, EMEA, Webscale Business

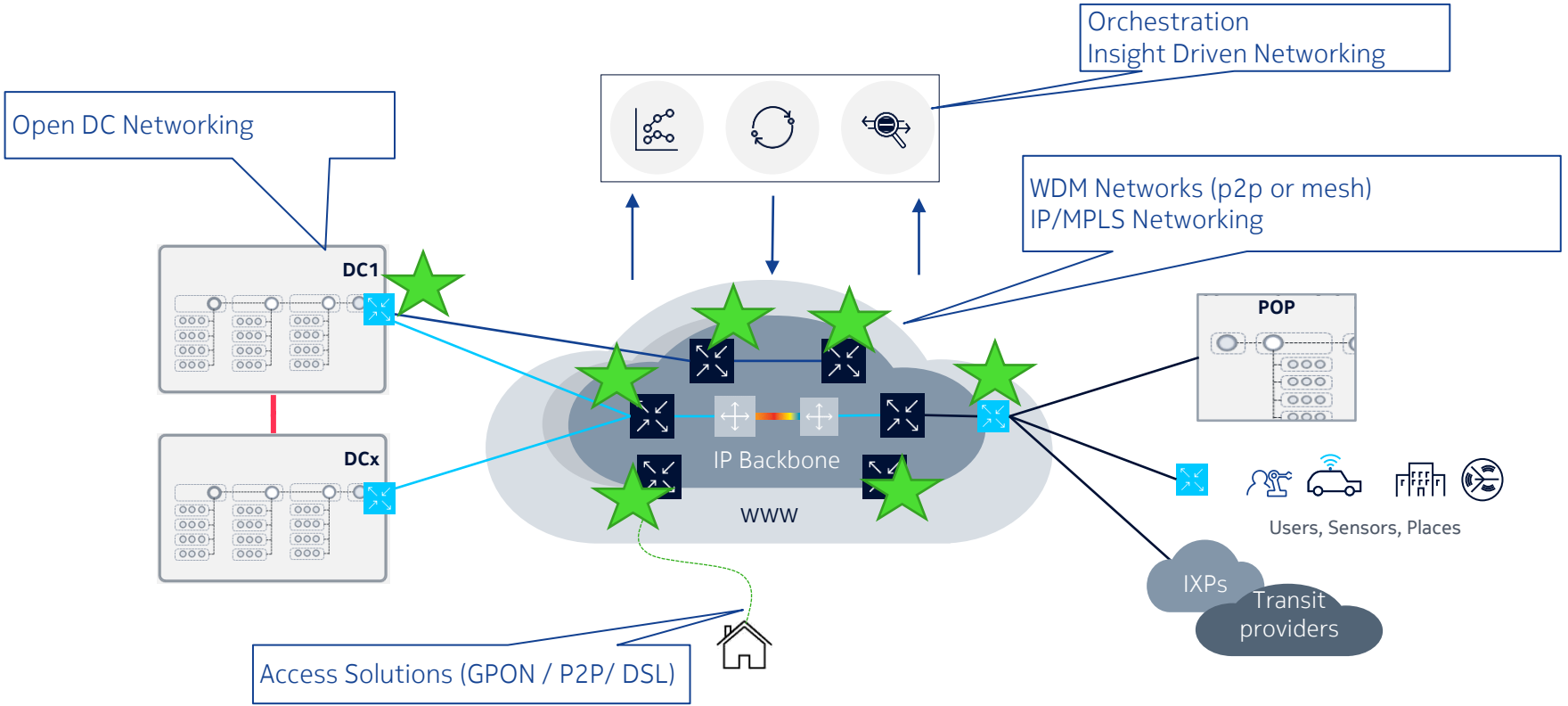
NOKIA

Nokia is taking care of your interconnectivity needs

NSP

SR-Linux

deepfield
security | performance | control



NOKIA

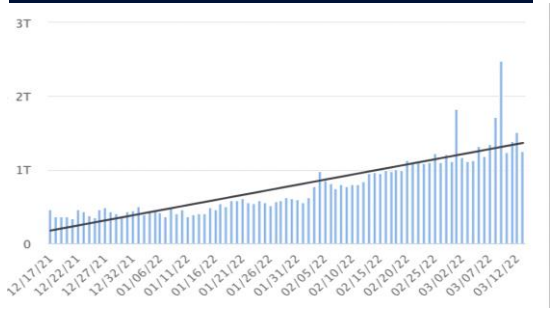
New interconnectivity needs driven by…

Accelerated **data consumption** during and after COVID

3x Acceleration in global bandwidth consumption 2022

DDoS traffic is exponential – growing faster than video or any other form of Internet content

>1T Once rare 1 Tbs attacks now daily occurrence!



Increased focus on **power efficiency** and **sustainability**

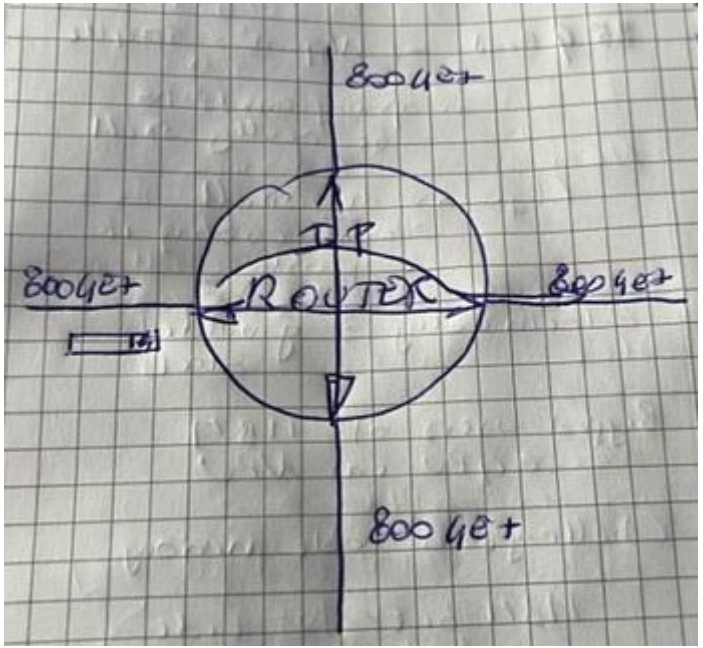
50% Emissions reduction by Nokia products & operations by 2030

How are routers (re)designed to cope with this ?

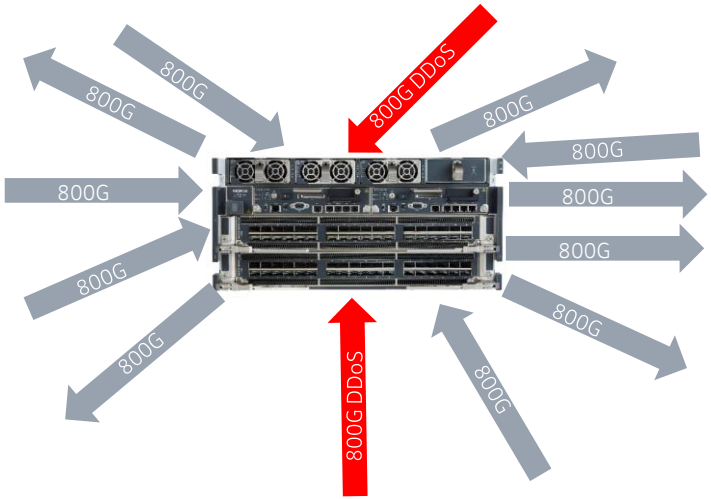
Routers are simple
.... or not so much?

A router's simple task

- Receive a packet
- Find the next-hop
- Send out the packet



Routers are simple
.... or not so much?



But ...

- Multiple Billion times per second (and growing)
- Preferably without any hick-ups (aka packet drops)

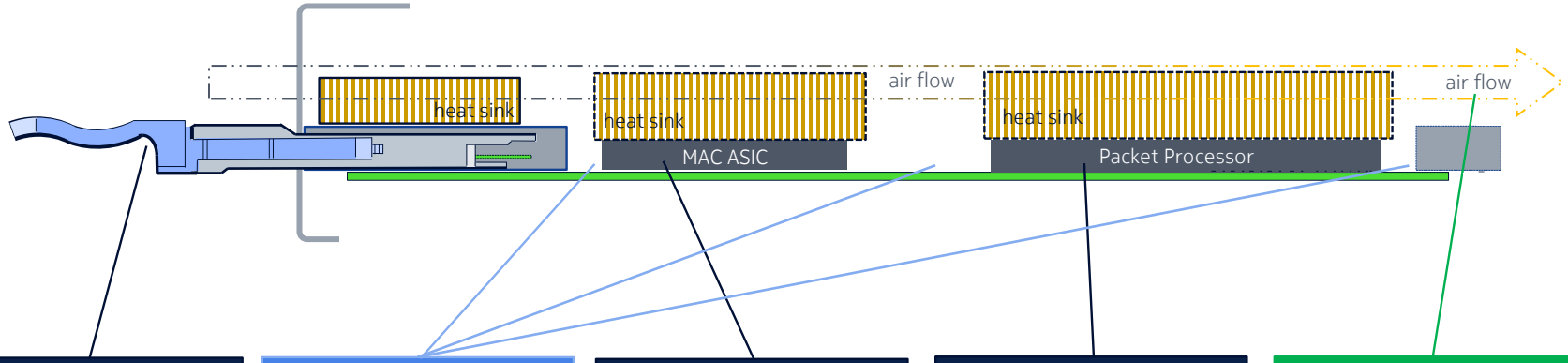
And

- Do it for more than just goodput IPv4
  (also IPv6, MPLS, VLAN, VPN, L2, PW, GRE, SDH, VXLAN, Video, MBH, Multicast, QoS, ... [ever growing list])
- Deal with “not-so-good” traffic (aka DDoS)
- With practical constraints like manageability, cost, power, ...

“From simple, uniform, best-effort IP forwarding to ... Life-over-IP”

How do we build for scale ?

Key technology evolutions on the router enabling 800GE



Pluggable Optics & Cage Type

- Distance
- Density
- Compatibility

SERDES

- Coding
- Speed

MAC ASIC

- Single Flow Speed
- Flexibility
- Compatibility

Packet Processor (NPU)

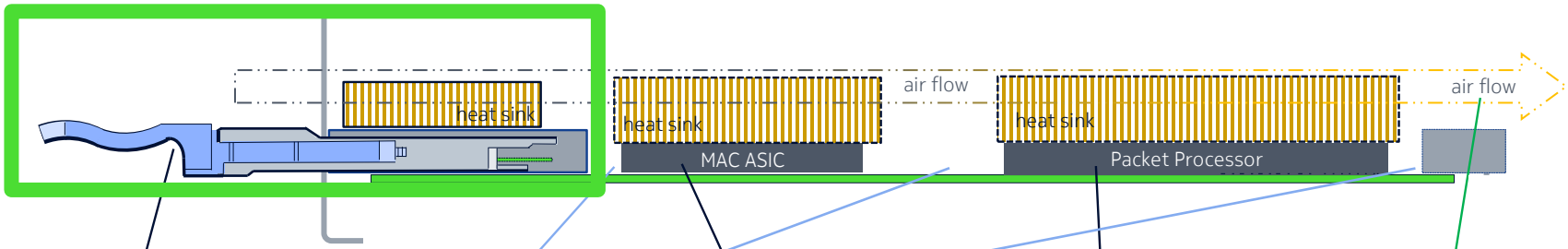
- Forwarding Speed
- Scale
- Buffering

Air Flow & Power

- Effectiveness
- Design longevity
- Power budget

How do we build for scale ?

Key technology evolutions on the router enabling 800GE



Pluggable Optics & Cage Type

- Distance
- Density
- Compatibility

SERDES

- Coding
- Speed

MAC ASIC

- Single Flow Speed
- Flexibility
- Compatibility

Packet Processor (NPU)

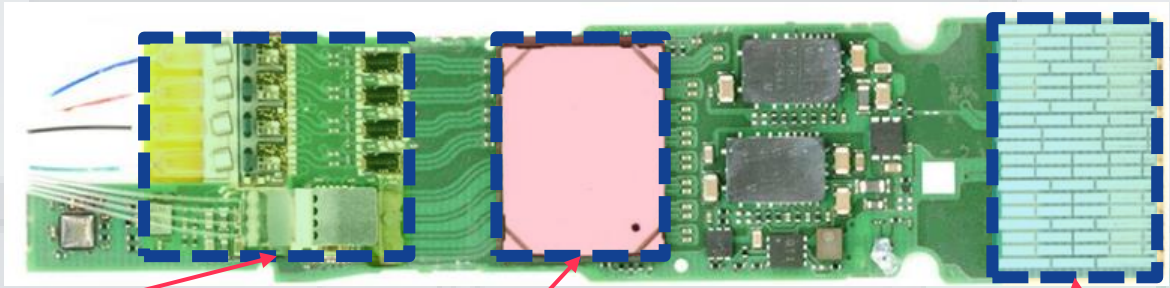
- Forwarding Speed
- Scale
- Buffering

Air Flow & Power

- Effectiveness
- Design longevity
- Power budget

Pluggables for 400G and Beyond …
Optical interface technology enablers

Formfactor
- Packaging
- Mechanics and Cooling
- **Router Interface density**



Arbitrary example photo of an optical pluggable PCB for illustration purpose

Photonics & Drivers
- Optical modulation and number of wavelengths (lambda's) are the key factor affecting cost and performance.

DSP / Multi-Link Gearbox
- Modulation/Demodulation digital signal processing
- One of KEY factors in defining power/thermal envelopes of the module

Attachment Unit Interface (AUI)
- Data transmitted over Electrical SerDes links from module to chipset

Pluggable optics and cage types

Diversity and uniformity

Cage types becoming universal

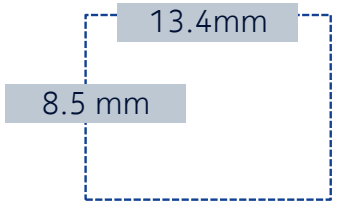
- **Narrow (100G-)**: SFP, SFP28, SFP56, SDF-DD, SFP112, SFP112-DD
- **Wide (100G+)**: QSFP28, QSFP28-DD, QSFP56, QSFP-DD 400, QSFP-DD 800

Interesting evolutions…

- **100G Single Lambda**: 800G, higher 100G/400G density, cost reduction
- **400G ZR/ZR+**: Coherent to enable “Pragmatic IPoDWDM” designs

SFP ‘Narrow’ cages

1 channel



SFP+ SFP 28 SFP-DD SFP 112

Eth Speed	10G	25G	100G	100G
Interface to ASIC (AUI)	1 x 10G XAUI	1 x 25G CAUI-1	2 x 50G 100GAUI-2	1 x **100G** 100GAUI-1
Modulation	10G NRZ	25G NRZ	100G PAM4	**100G PAM4**
Typical optical connection	LC (1 λ)	LC (1 λ)	LC (1 λ PAM4 or BiDi) or MPO LR1, FR1, DR1, SR1.2, SR2	LC (1 λ PAM4) LR1, FR1, DR1

100G Lambda MULTI-SOURCE AGREEMENT

100G Lambda MULTI-SOURCE AGREEMENT

SFP-DD and QSFP28: 100G Single Lambda

Complementary deployment options

Single Lambda portfolio

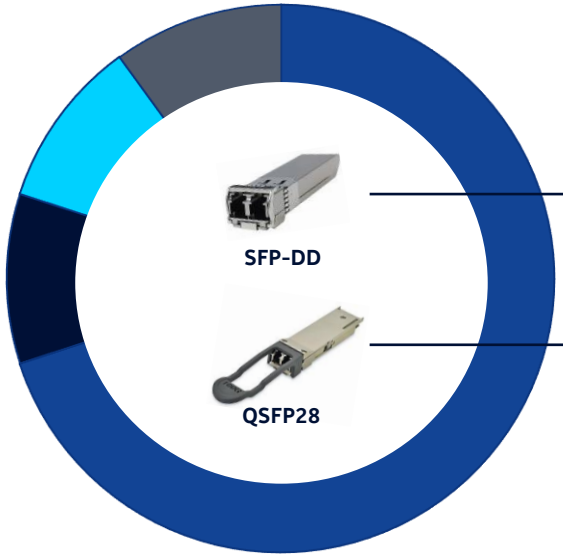
- *100G SR1.2 100M MMF*
- 100G DR 500m SMF
- 100G FR 2km SMF
- 100G LR 10km SMF

Compatible with 4x100G QSFP-DD

- 4x100G DR / FR / LR
- 8x100G DR / FR / LR

100G future proof investment

- PAM4 forward compatible

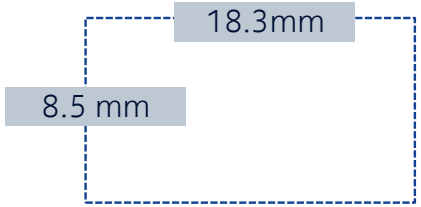


1RU Faceplate enables 48xSFPDD + 6xQSFP-DD

1RU Faceplate enables 36xQSFP28

QSFP ‘Wider’ cages

4 channel



QSFP+ QSFP28 QSFP-DD QSFP-DD800

Eth Speed	100G	25G	400G	800G
Interface to ASIC	4 x 10G XLAUI	4 x 25G CAUI-4	8 x 50G 400GAUI-8	8 x **100G** 800GAUI-8
Modulation	10G NRZ	25G NRZ	100G PAM4	**100G PAM4**
Typical optical connection	MPO LC (4 λ MUX)	MPO LC (4 λ MUX, or 1 λ)	MPO LC (4 or 8 λ MUX) 100G Lambda MULTI-SOURCE AGREEMENT	MPO LC (8 λ MUX) 100G Lambda MULTI-SOURCE AGREEMENT

QSFP-DD800G

Introducing 800G

QSFP-DD MSA

Formal announcement of QSFP-DD800 MSA May 24, 2021

- Designed to be backwards compatible with existing QSFP-DD
- Point-to-point or breakout for high density 100G applications

100G SerDes

Architected to leverage the IEEE 802.3 ck work to fit 800G

- 100G PAM4 electrical, using 50Gbaud signaling

Pluggable modules

Initial product release

- QSFP-DD800 DR8 (500 m)
- QSFP-DD800 DR8+ (2 km)
- QSFP-DD800 2x400G FR4 (2 km)

Transceiver

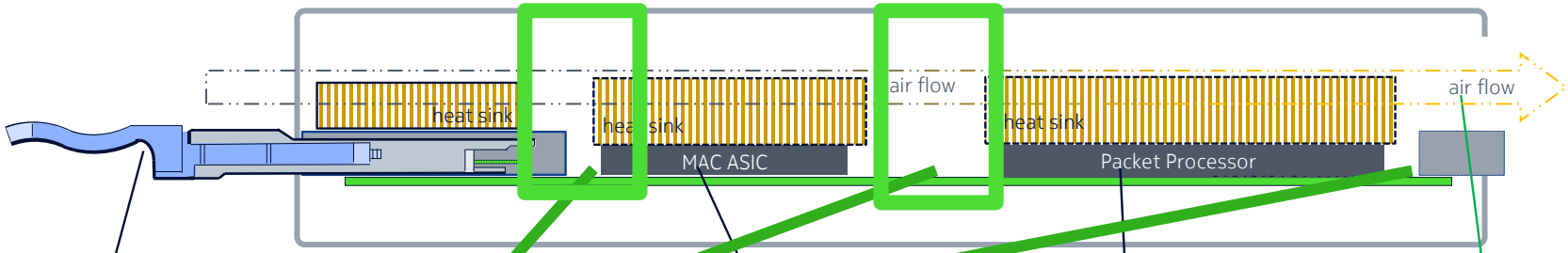
QSFP-DD form-factor

- 0/70°C case temperature
- 25% -43% power savings over 400G
- Price neutral to 400G

Clear economic and power advantages to 800G

How do we build for scale ?

Key technology evolutions on the router enabling 400G – 800G - beyond



Pluggable Optics & Cage Type

- Distance
- Density
- Compatibility

SERDES

- Coding
- Speed

MAC ASIC

- Single Flow Speed
- Flexibility
- Compatibility

Packet Processor (NPU)

- Forwarding Speed
- Scale
- Buffering

Air Flow & Power

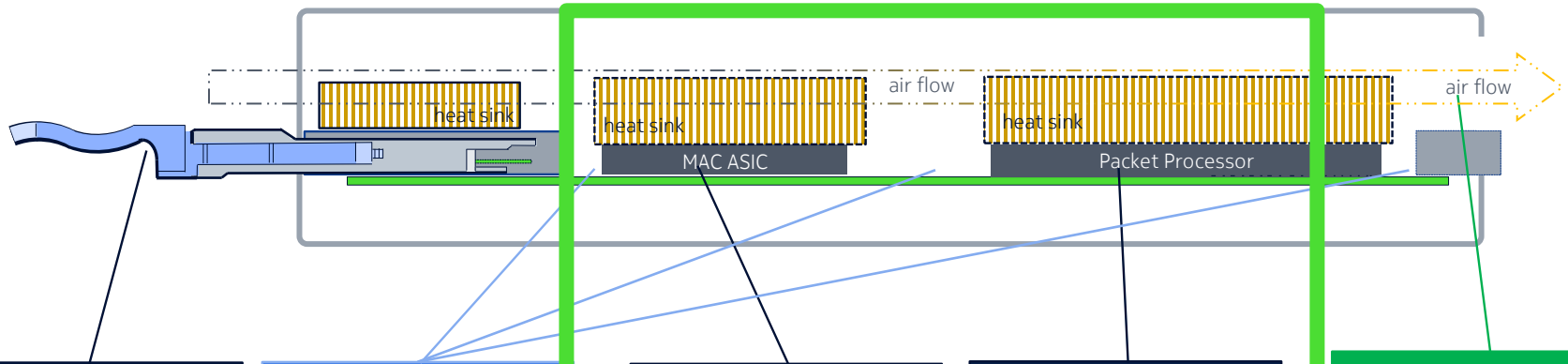
- Effectiveness
- Design longevity
- Power budget

SerDes

- Serializer/Deserializer
  - Connection between ASICs and towards cage
  - Increasing speeds of an individual lane: 10G, 28G, 56G
- Latest specification: 100G SerDes (802.3ck) for chip-to-chip or chip-to-module communication
  - Use of PAM4 modulation
  - Well-aligned with optics evolution (100G Lambda) – resulting in less active components and complexity in the transceivers
- Benefits
  - Higher I/O possible
  - Better power characteristics and cost
- Complex, but necessary evolution

How do we build for scale ?

Key technology evolutions on the router enabling 400G – 800G - beyond



Pluggable Optics & Cage Type

- Distance
- Density
- Compatibility

SERDES

- Coding
- Speed

MAC ASIC

- Single Flow Speed
- Flexibility
- Compatibility

Packet Processor (NPU)

- Forwarding Speed
- Scale
- Buffering

Air Flow & Power

- Effectiveness
- Design longevity
- Power budget

MAC ASIC and Packet Processor (NPU)

Evolving the router's data-plane to higher speeds, scale and capabilities

Enabling 800G+ interfaces requires an evolution across the main forwarding components* of the router

MAC

- 800GE and higher
- Enabling 'Universal Ports'
- Optional support for
  - MACSec
  - Flex-E
  - Intelligent Aggregation

Store

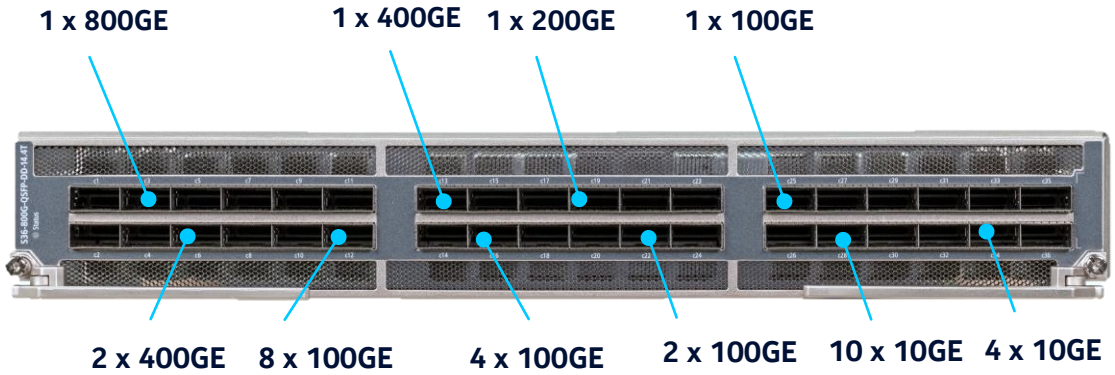
- Buffer characteristics
  - Location (ingress, egress, both)
  - Size
  - Bandwidth (full vs partial)

Forward

- Lookup/forwarding speed
- Scale
  - FIB scale
  - ACL scale
  - uRPF impact
- QoS support

* Different implementations/combinations possible

The ease of having Universal Ports



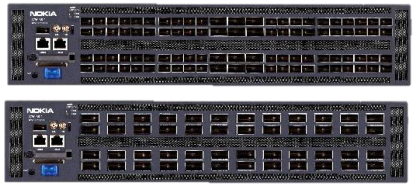
Your choice to use each port at whatever speed and breakout you like

Use of universal ports in Appliance-Based routers

7750 SR-1 (Modular)

600G – 1.5T, up to 3T iA
2 MDA slots, flexibility with SFP-DD, QSFP28, QSFP-DD400 and CFP2

1 x 100GE (S)	1 x 400 GE
2 x 50GE	1 x 200 GE
1 x 50GE	4 x 100 GE
1 x 40GE	2 x 100GE
4 x 25GE	
1 x 25GE (S)	+ CFP2
10 x 10GE	
4 x 10GE	
1 x 10GE (S)	



7750 SR-1 (Fixed)

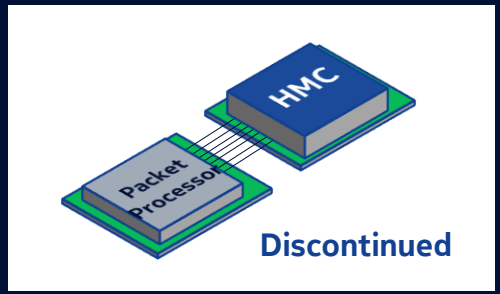
2.8T or 6.0T, up to 19.2T iA
SFP-DD or QSFP-DD800 options

1 x 100GE (S)	1 x 400 GE	1 x 800 GE
2 x 50GE	1 x 200 GE	2 x 400 GE
1 x 50GE	4 x 100 GE	8 x 100GE
1 x 40GE	2 x 100GE	
4 x 25GE		
1 x 25GE (S)		
10 x 10GE		
4 x 10GE		
1 x 10GE (S)		

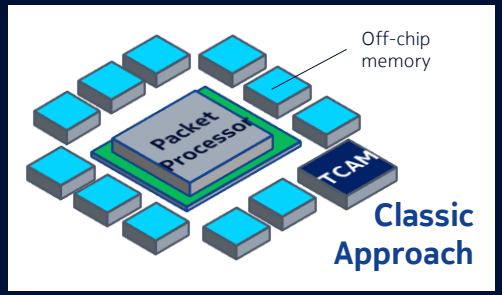
Your choice to use each port at whatever speed and breakout you like

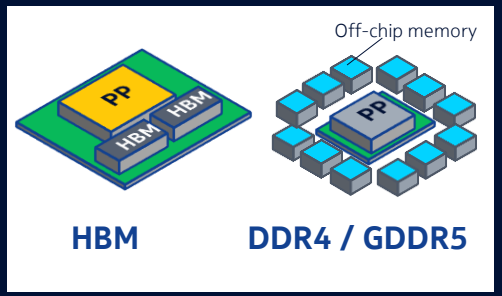
Next-generation packet processing

Hybrid Memory Cube (HMC)

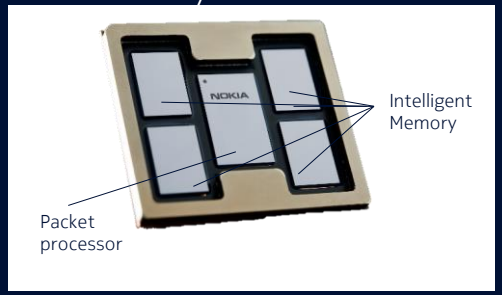


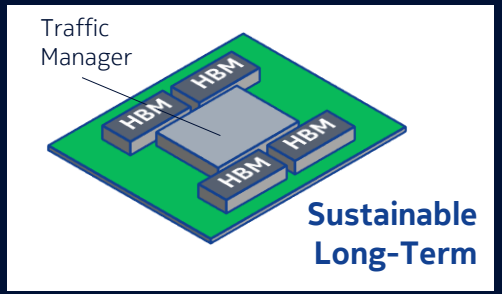
DDR / RLDRAM / QDR / TCAM / HBM Partial Buffered





Nokia Intelligent Memory Fully Buffered

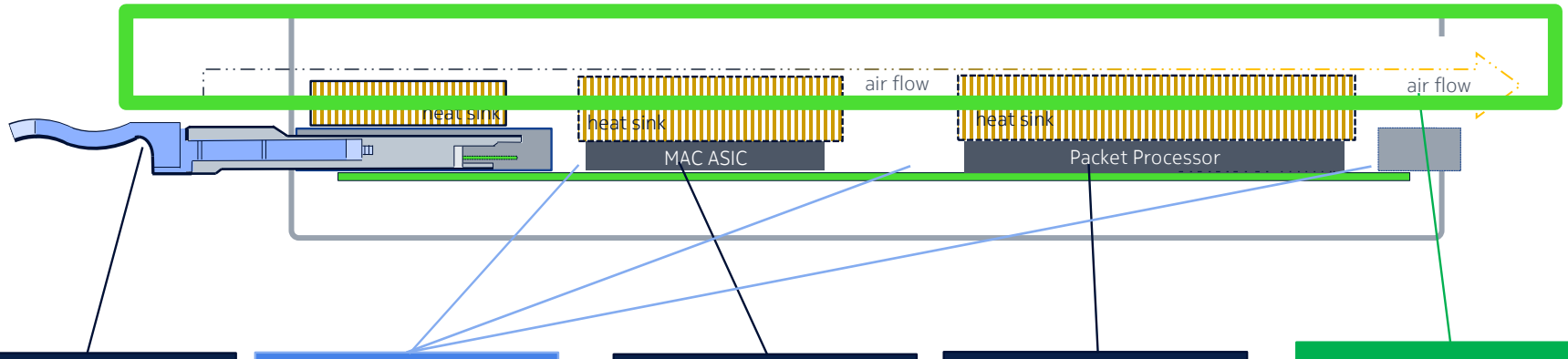




FP4/FP5: Multi-dimensional, deterministic scale

How do we build for scale ?

Key technology evolutions on the router enabling 400G – 800G - beyond



Pluggable Optics & Cage Type

- Distance
- Density
- Compatibility

SERDES

- Coding
- Speed

MAC ASIC

- Single Flow Speed
- Flexibility
- Compatibility

Packet Processor (NPU)

- Forwarding Speed
- Scale
- Buffering

Air Flow & Power

- Effectiveness
- Design longevity
- Power budget

Stacking more linecards into a chassis => Chassis System architecture
Design Considerations

Mechanical design of huge significance

Midplane vs. Orthogonal Direct Cross Connect

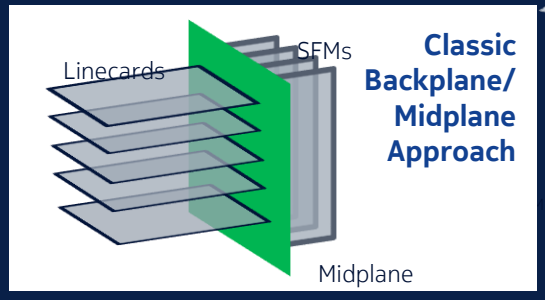
Line card pitch & orientation

Cooling design

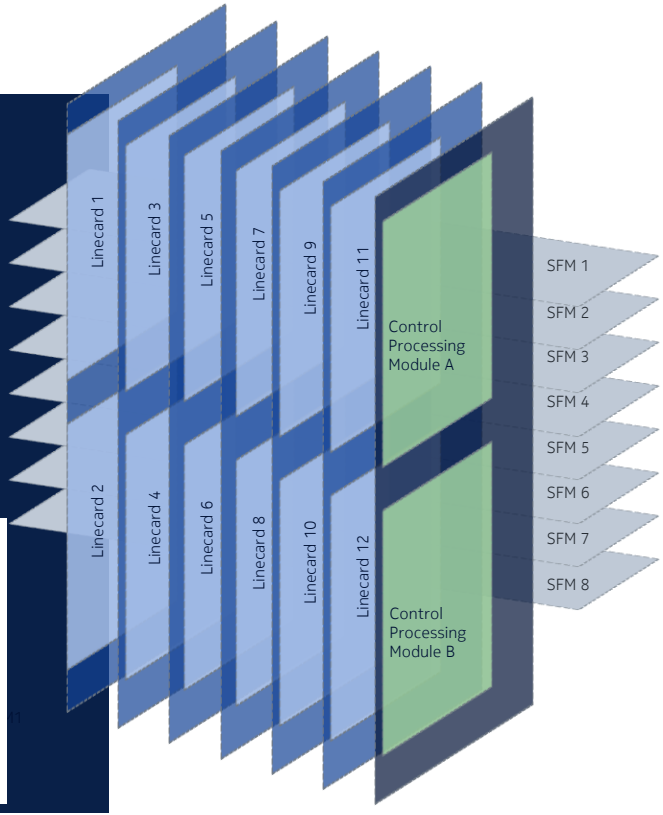
Power design

Impacts

- Density
- Power consumption
- Optics support

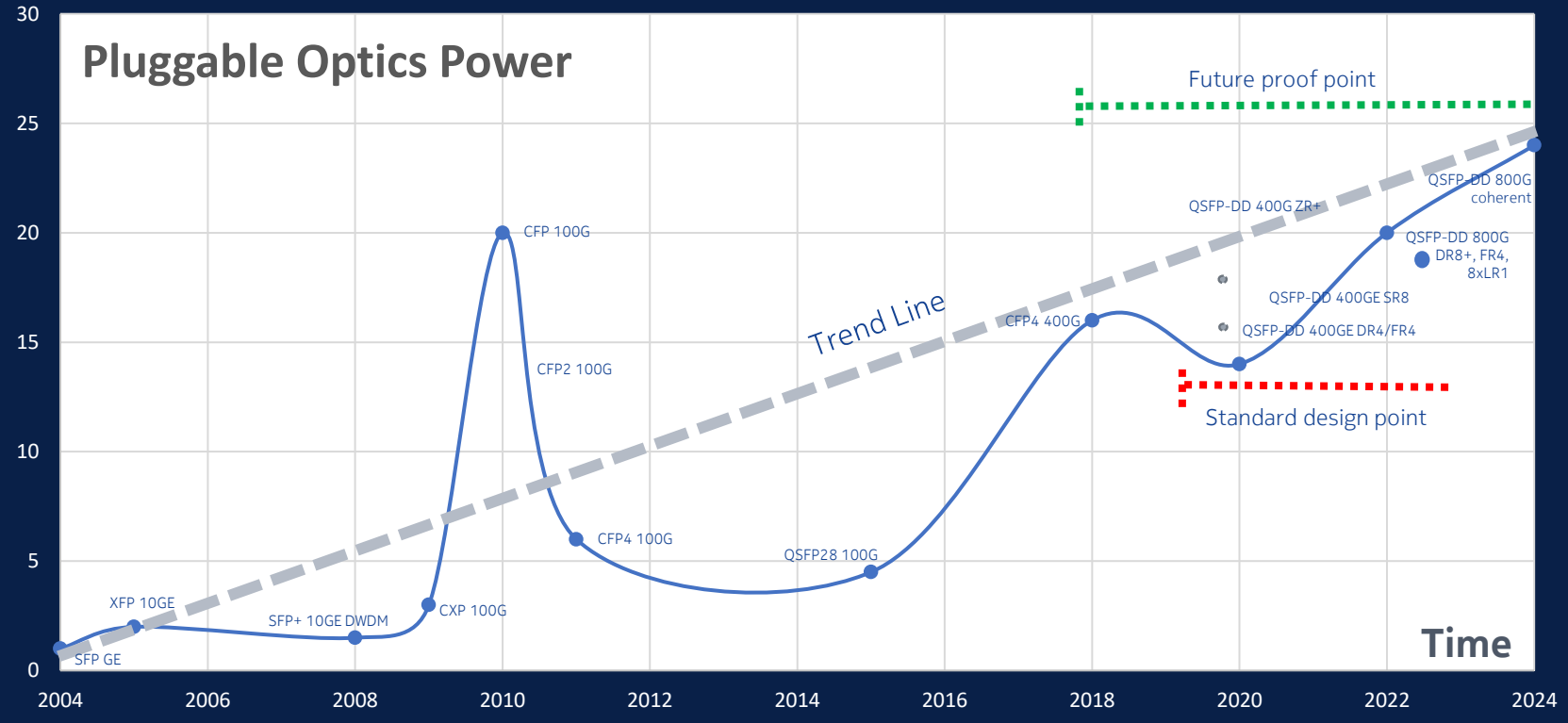


Classic Backplane/ Midplane Approach



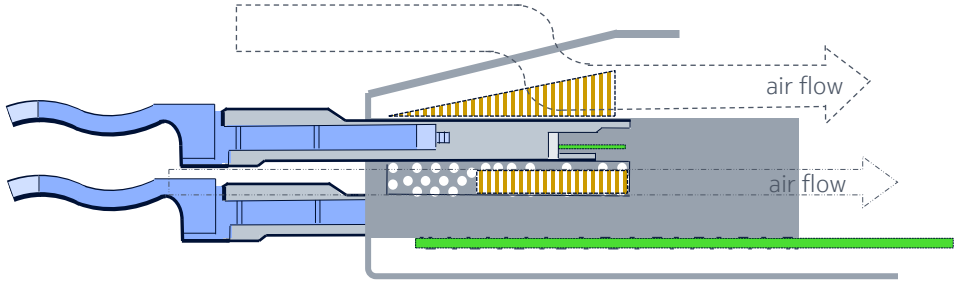
800G Optics Evolution

Cooling today's and tomorrow's optics

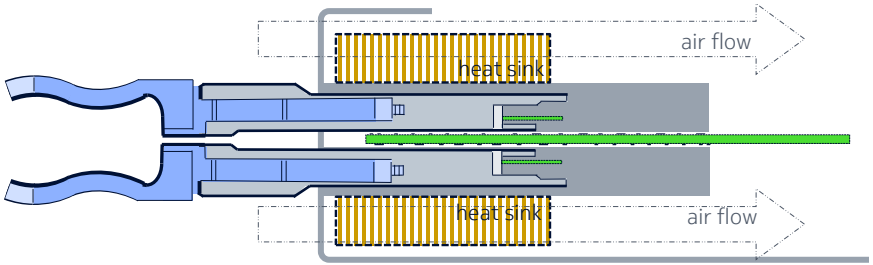


Optics cooling design

Single sided PCB



Dual sided PCB



Stacked SFP Cages

- Classic DC design
- Large heat sink only on top cage
- Bottom cage always hotter - imbalanced optical performance
- DD Design point ~13W optics in all cages at 40C
- Limits applicability to future optics
- Fans might have to run faster

Belly-to-Belly SFP Cages

- Future proof design
- Large dedicated heat sink per cage
- Even cooling to all cages
- Cooling to 28W+ in all cages at 40C

Enabling 800G and beyond on IP routers

Design choices along the datapath

Platform

Mechanical design
Power
Cooling

Dataplane & chipset interconnect

Forwarding
MAC
SERDES

Pluggable Optics

SFPDD-100, QSFP28, QSFP56-DD, QSFPDD-800

NOKIA