Host-to-Module Wiring

David Ofelt – Juniper Networks
802.3ck meeting – Spokane - 2018-09

Introduction

The ad-hoc presentation on System Vendor's Concerns (ofelt_3ck_adhoc_01_081518) referenced forwards and backwards compatibility issues with host-to-module wiring.

The presentation did not include any pictures, so the discussion was difficult to follow!

This presentation is an attempt to clarify this topic…

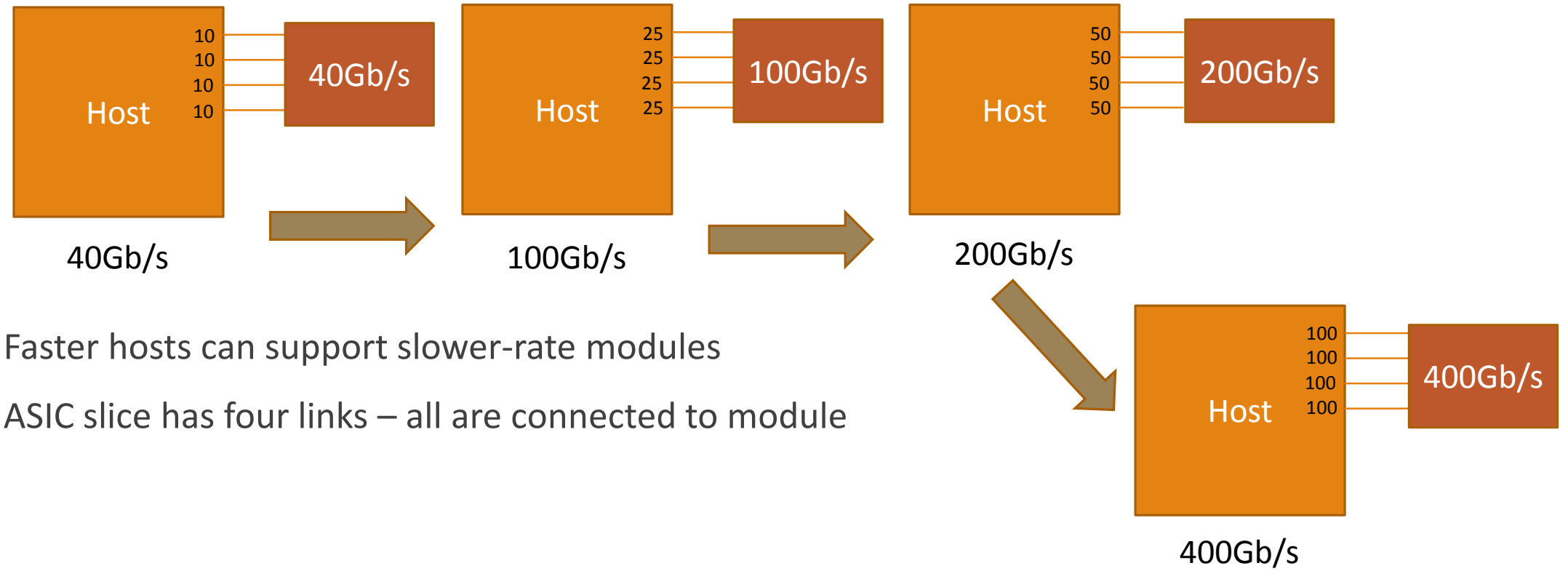
SERDES + module progression

SERDES Rate	QSFP (4 lanes)	QSFP-DD / OSFP (8 lanes)	½ QSFP-DD / OSFP (4 lanes)
10Gb/s	40Gb/s		
25Gb/s	100Gb/s	200Gb/s	
50Gb/s	200Gb/s	400Gb/s	
100Gb/s	400Gb/s	800Gb/s	400Gb/s

Empty cells are technically possible, but haven't been part of our discussions.

Faster-rate host SERDES are assumed to be able to run at all standard slower rates

QSFP Module Progression



Faster hosts can support slower-rate modules

ASIC slice has four links – all are connected to module

Another Backwards Compatibility Option

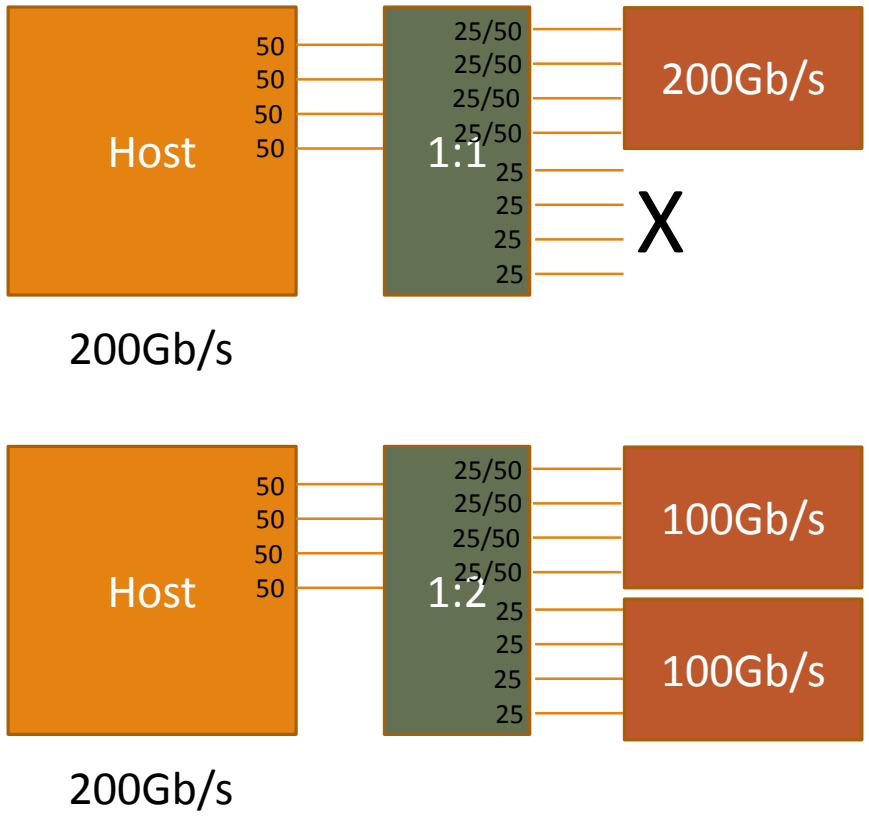
Systems at generational boundaries can hedge

- All switch bandwidth can drive previous-generation modules without stranding bandwidth
- Assumes there is enough faceplate

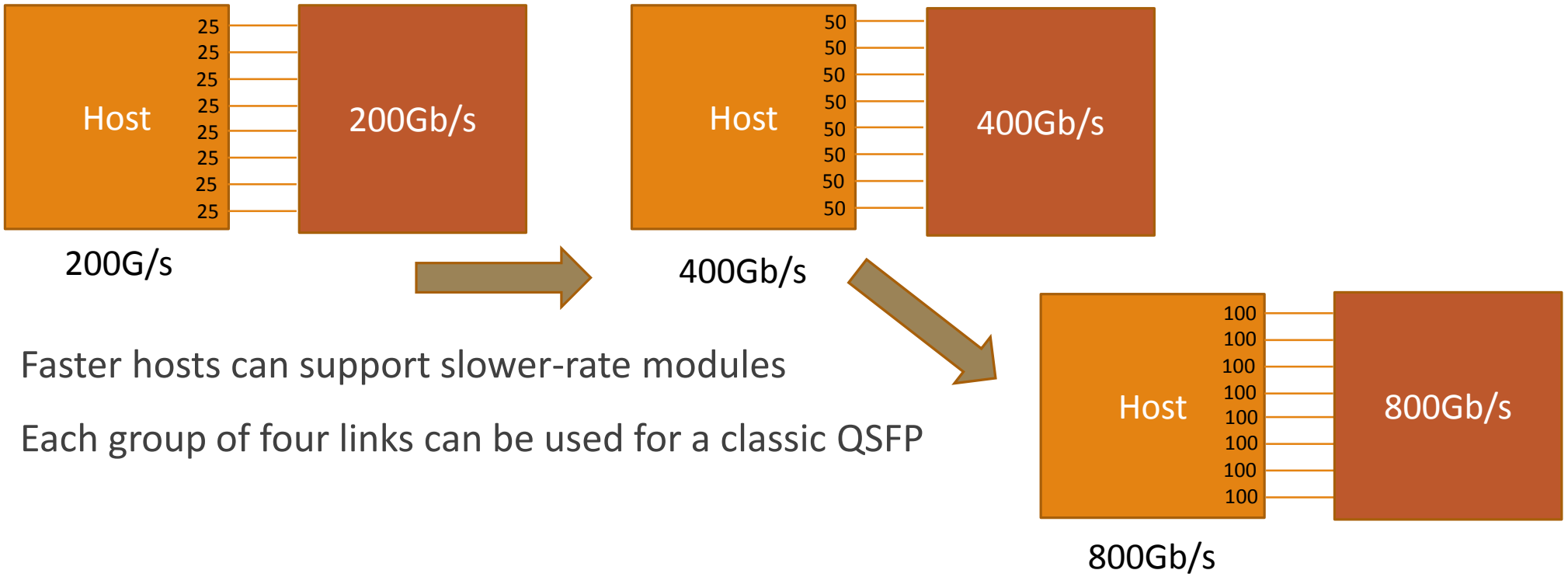
Provide two slots that can support either:

- One fast module
- Two slow modules

Board-level demux needed



QSFP-DD/OSFP Module Progression



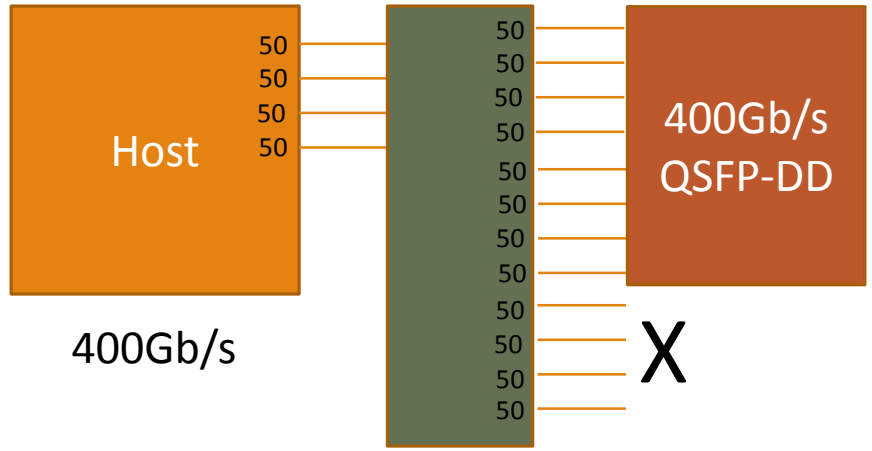
Faster hosts can support slower-rate modules

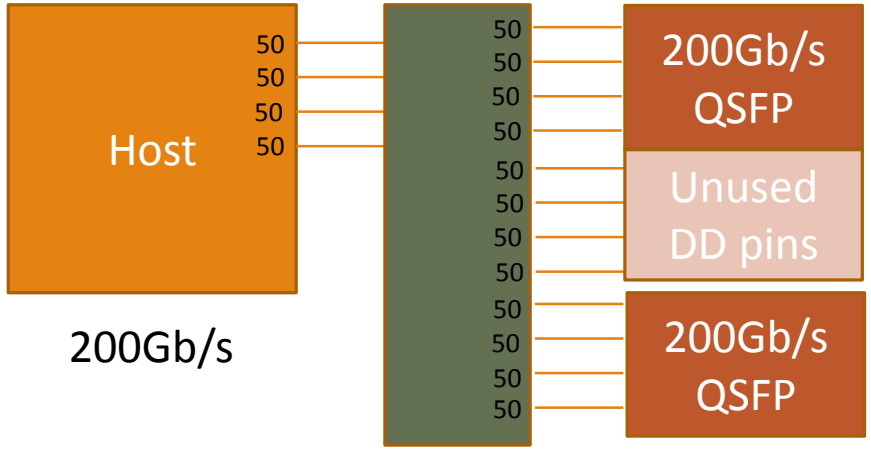
Each group of four links can be used for a classic QSFP

QSFP-DD/QSFP Backwards Compatibility

Systems can choose to support both QSFP & QSFP-DD generations without wasting host bandwidth.

- Board-level mux parts again are required
- More face-plate area required



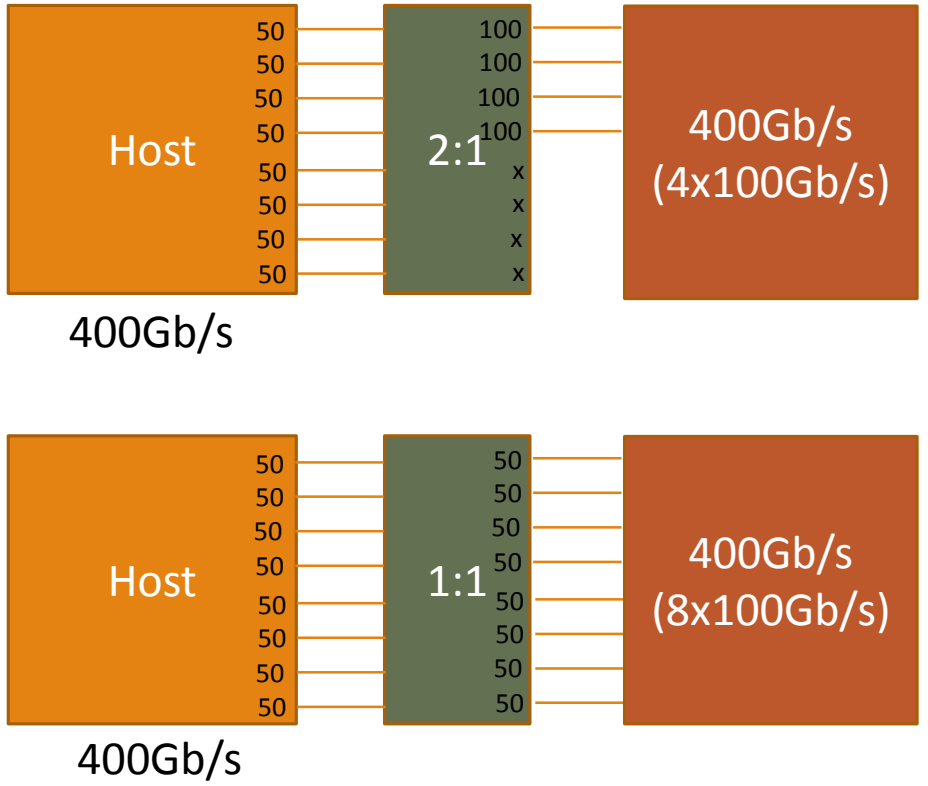


4x100Gb/s QSFP-DD/OSFP

If 8-lane module only has 4x100Gb/s SERDES:

- Older host ASIC needs board-level 2:1 mux

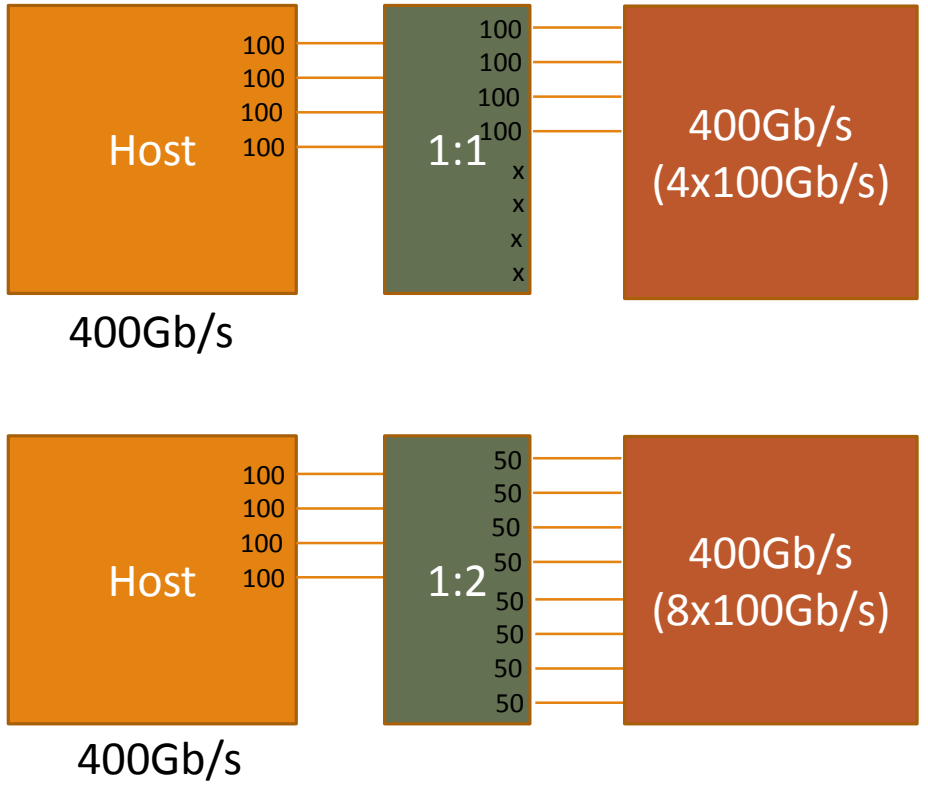
Older host without mux can't use module



4x100Gb/s Host

If host only has 4x100Gb/s SERDES

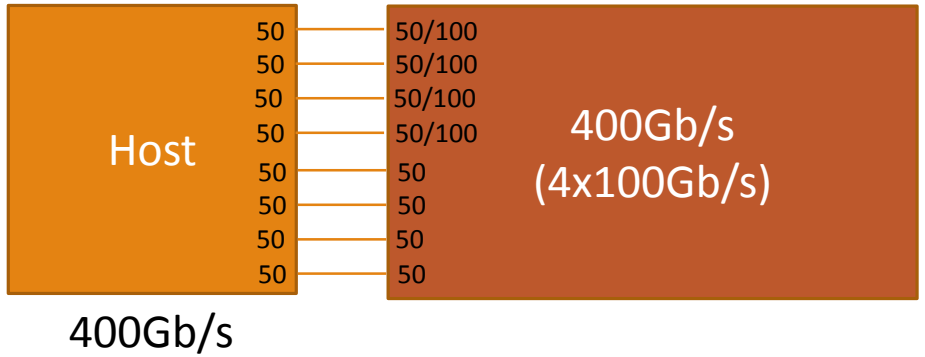
- Older 8x50Gb/s modules need a board-level de-mux

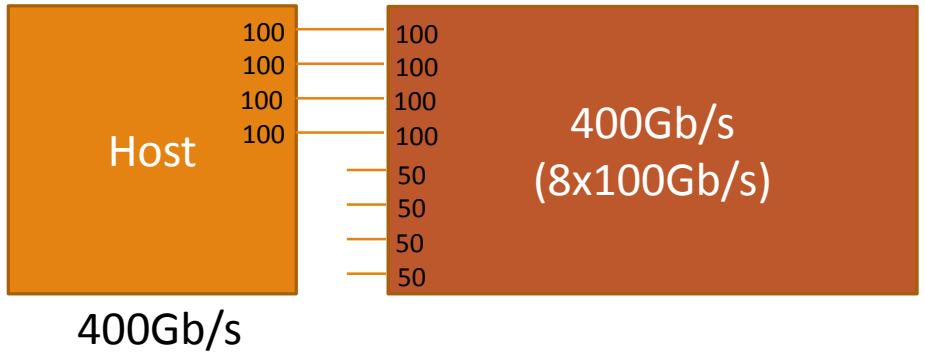


Forward-looking module

Once 100Gb/s-based modules are available- a dual mode 400Gb/s module is possible.

- 8x50Gb/s for older hosts
- 4x100Gb/s newer





Discussion

Many options exist for interconnection host ASICs and the various generations of modules

Not all options may be acceptable for all users:

- Single-application, custom, homogeneous boxes with few suppliers may be fine with limiting options
- System vendors trying to deploy to a heterogenous, broad, user space with a large diversity of suppliers may find some of the more limiting options difficult to use.

End up with complicated economic feasibility and broad-market potential discussions

Thanks!