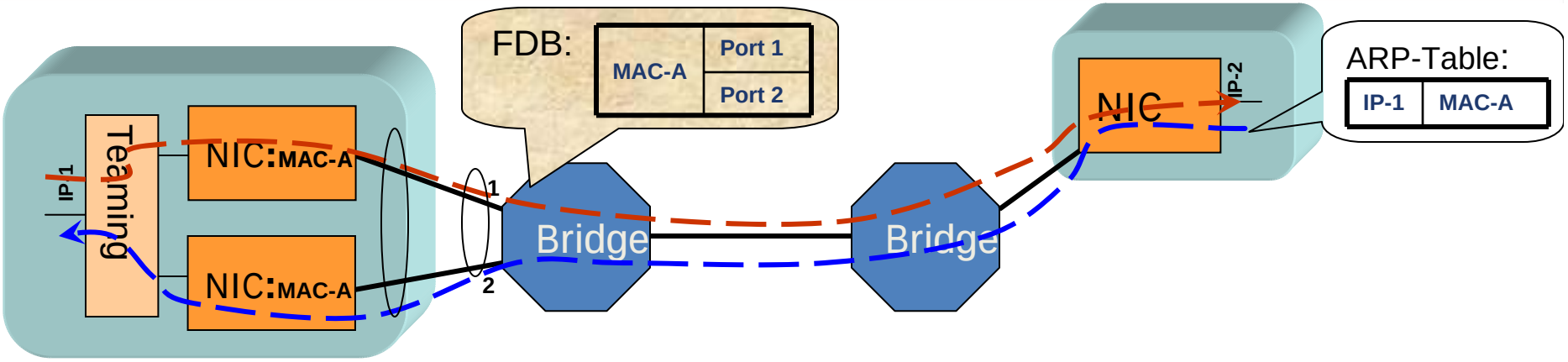
Link Aggregation Challenges

Manoj Wadekar, QLogic

Goals of presentation

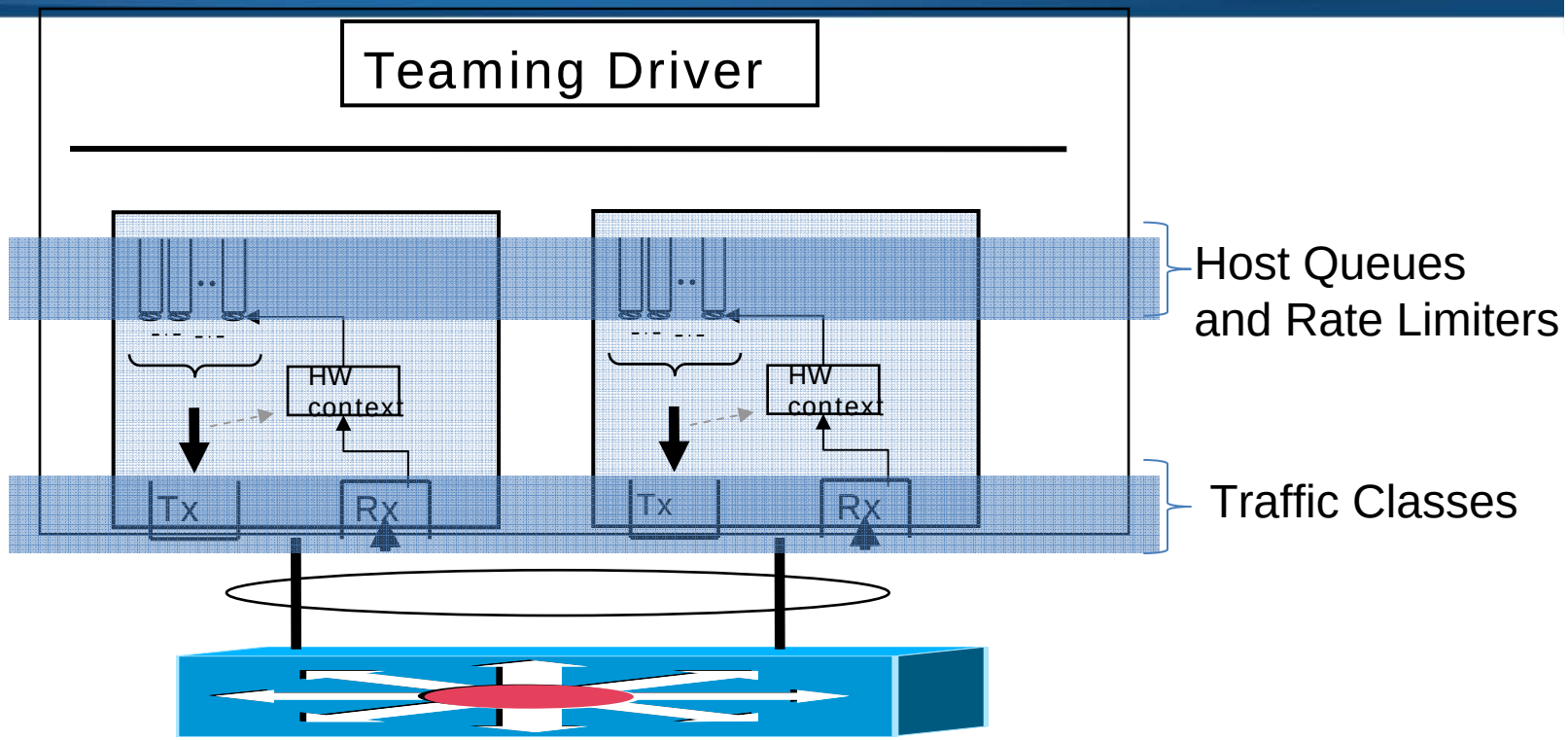
- Discuss challenges of Link Aggregation in DCB
- Trigger discussion for potential solutions

Overview: Link Aggregation - Static



- **Traffic from all team ports carries same MAC address**
- **Adapter and Switch participate in load balancing traffic on team ports**
  - Hashing asymmetrical on two sides of the link
  - Flow mapping is not guaranteed to be symmetrical
  - E.g. IP1→IP2 traffic on teamed port 1 does not guarantee IP2→IP1 traffic will be mapped on to same port 1

Link Aggregation: Example End Station

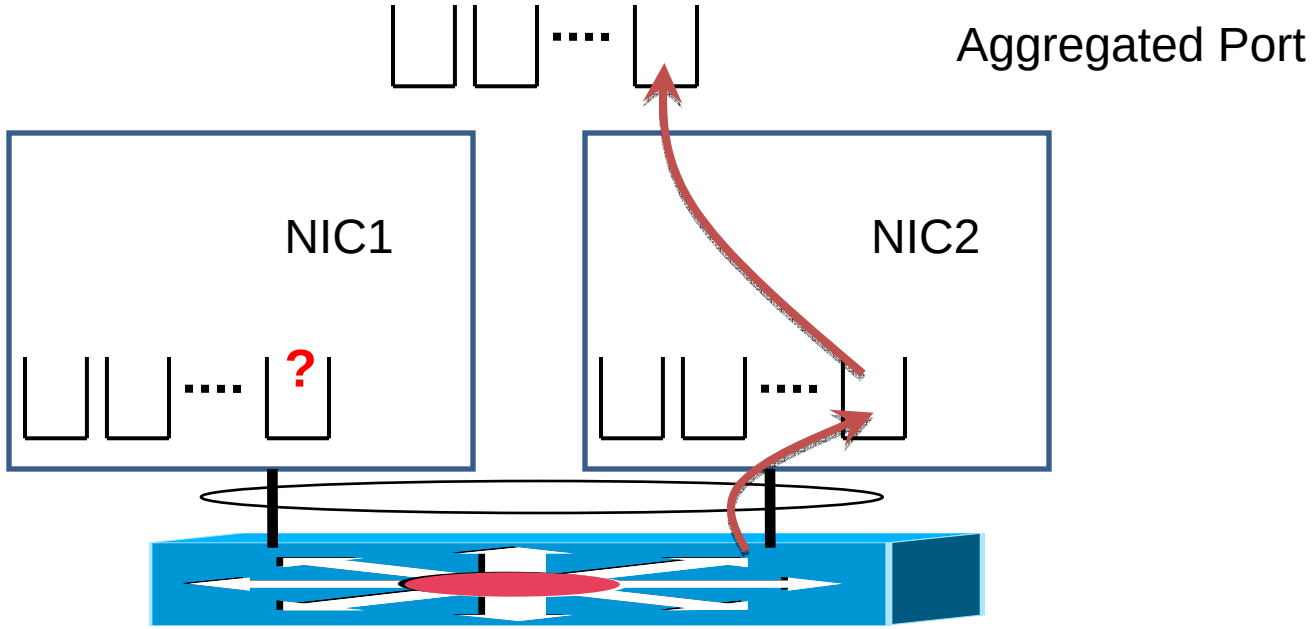


- Teaming may span multiple physical cards
- Transmit port for Teaming decided in Teaming driver
- ETS applied close to wire
- PFC applied close to wire
- Rate Limiters effective in HW

Link Aggregation Challenges

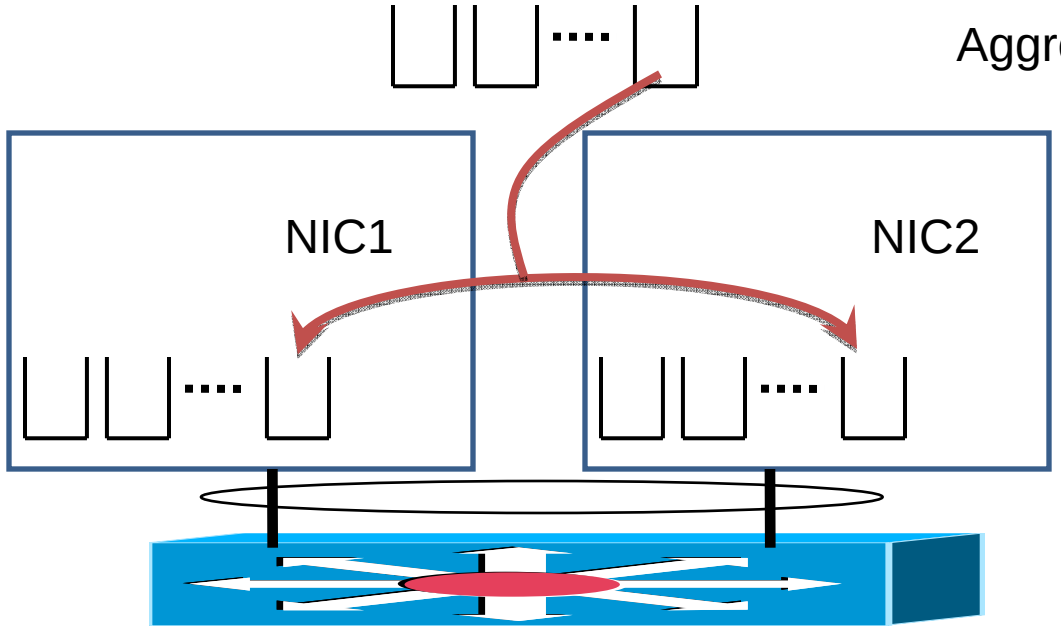
- **CN Operation (already discussed)**
  - How to deliver CNM to same port on which RL is installed?
- **PFC**
  - How to ensure single priority is Paused across aggregated link?
- **ETS**
  - How to account for aggregated link BW per TCG?
- **DCBX**
  - How does LLDP work on aggregated link?

PFC Challenges



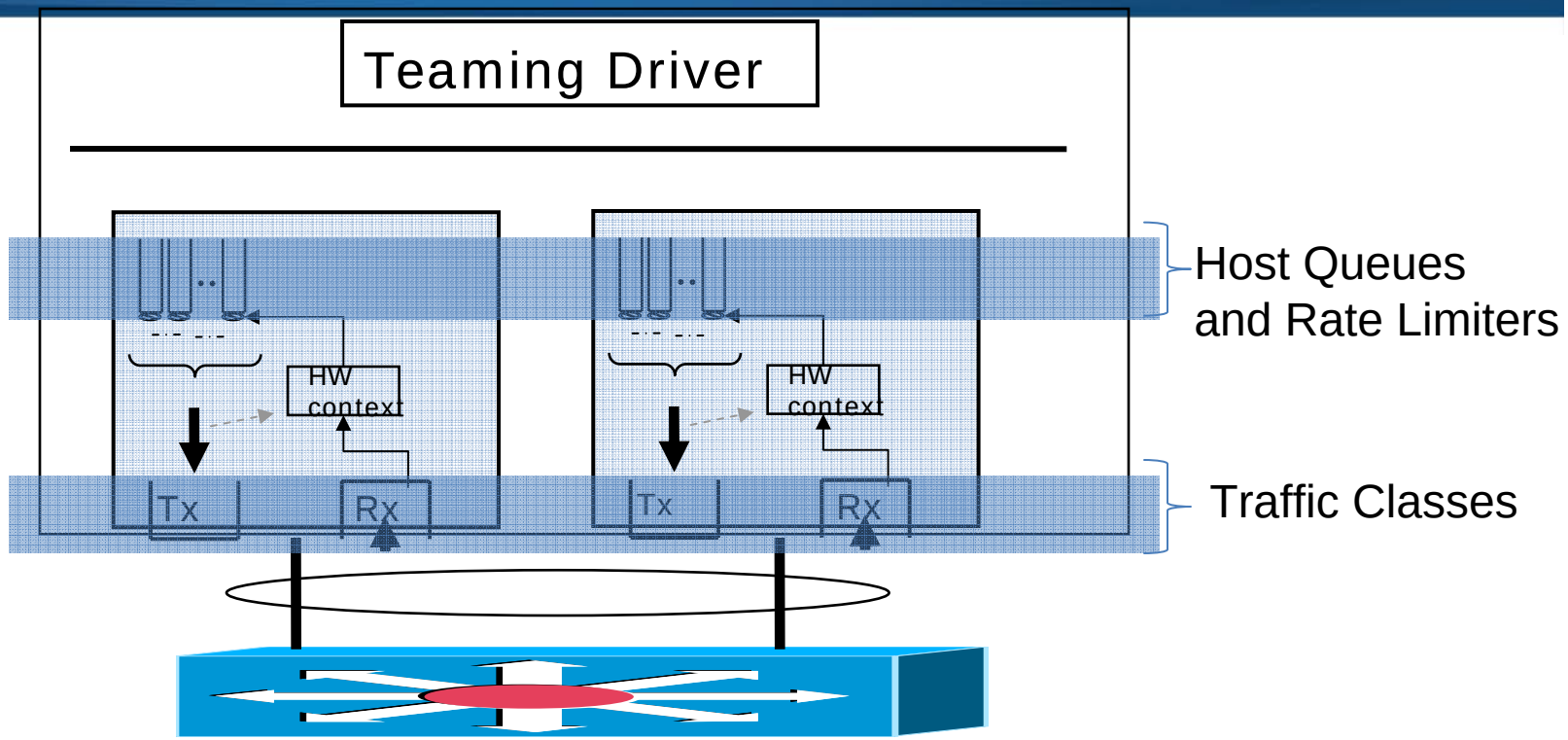
- **PFC operates on physical port**
- **Aggregated port options:**
  - A. Pause given priority on all ports if PFC received on any aggregation-member port
    - Can not satisfy response-time requirements
  - B. Require that PFC should be generated independently on all aggregation member ports

ETS Challenges



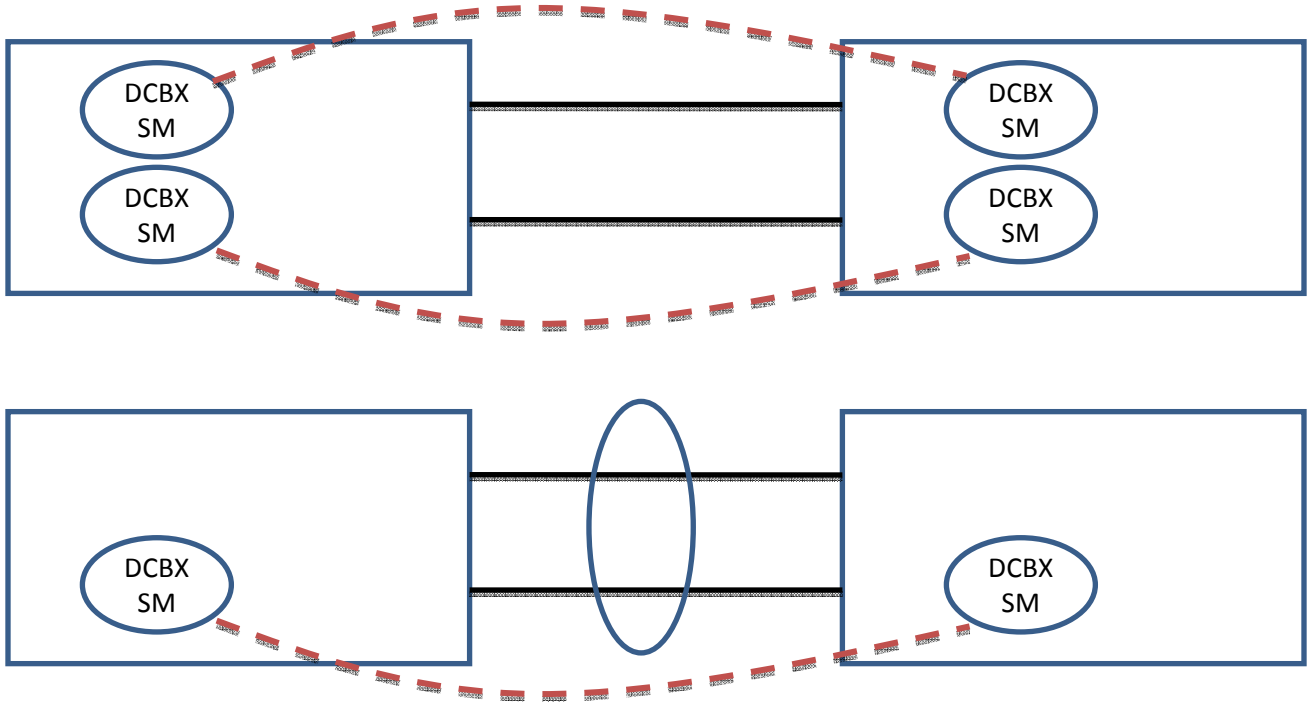
- **ETS operates on traffic classes on each physical port**
- **How to ensure bandwidth across aggregation-member ports follows ETS assignment?**

CN Challenges



- **Multiple ports are aggregated and HW stores context in HW**
  - E.g. Offload information (FCoE, TCP, iSCSI, iWARP) etc.
- **Multiple ports are aggregated and CN reaches wrong port**
  - CN needs to be handled expeditiously
- **Can cause Large latency to CN handling, performance impact to offload functionality**

DCBX Challenges



- **Single DCBX instance required across aggregated-member ports?**
  - LLDP runs on each aggregated-member port
  - PFC, CN, ETS negotiations across physical ports
  - DCBX needs to be done in or above teaming driver
  - BTW : which port is used for protocol?

802.1 Architecture and Implementation challenges

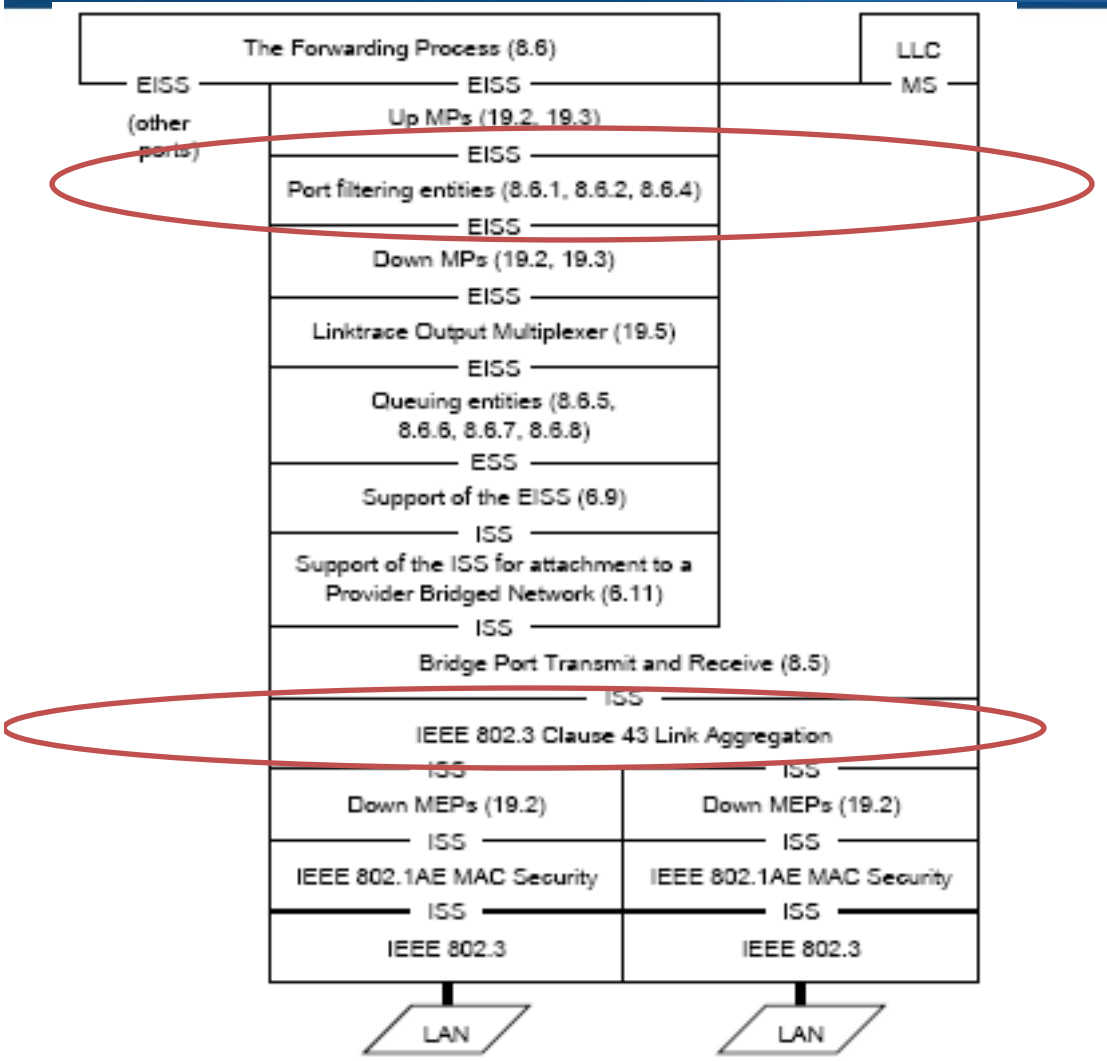


Figure 22-8—Maintenance Point placement relative to other standards

Per port Queues

Link aggregation is below Per port queues

This can't be achieved for aggregation across physical card


QLOGIC®
The Ultimate in Performance