Empowered by Innovation NEC

2011 3rd Workshop on
Data Center – Converged and Virtual Ethernet Switching
DC–CaVES 2011

Advanced FCoE: Extension of Fibre Channel over Ethernet

September 9, 2011

Satoshi Kamiya, Kiyohisa Ichino, Masato Yasuda,
Noriaki Kobayashi, Norio Yamagaki and Akira Tsuji

NEC Corporation

(kamiya@ak.jp.nec.com)

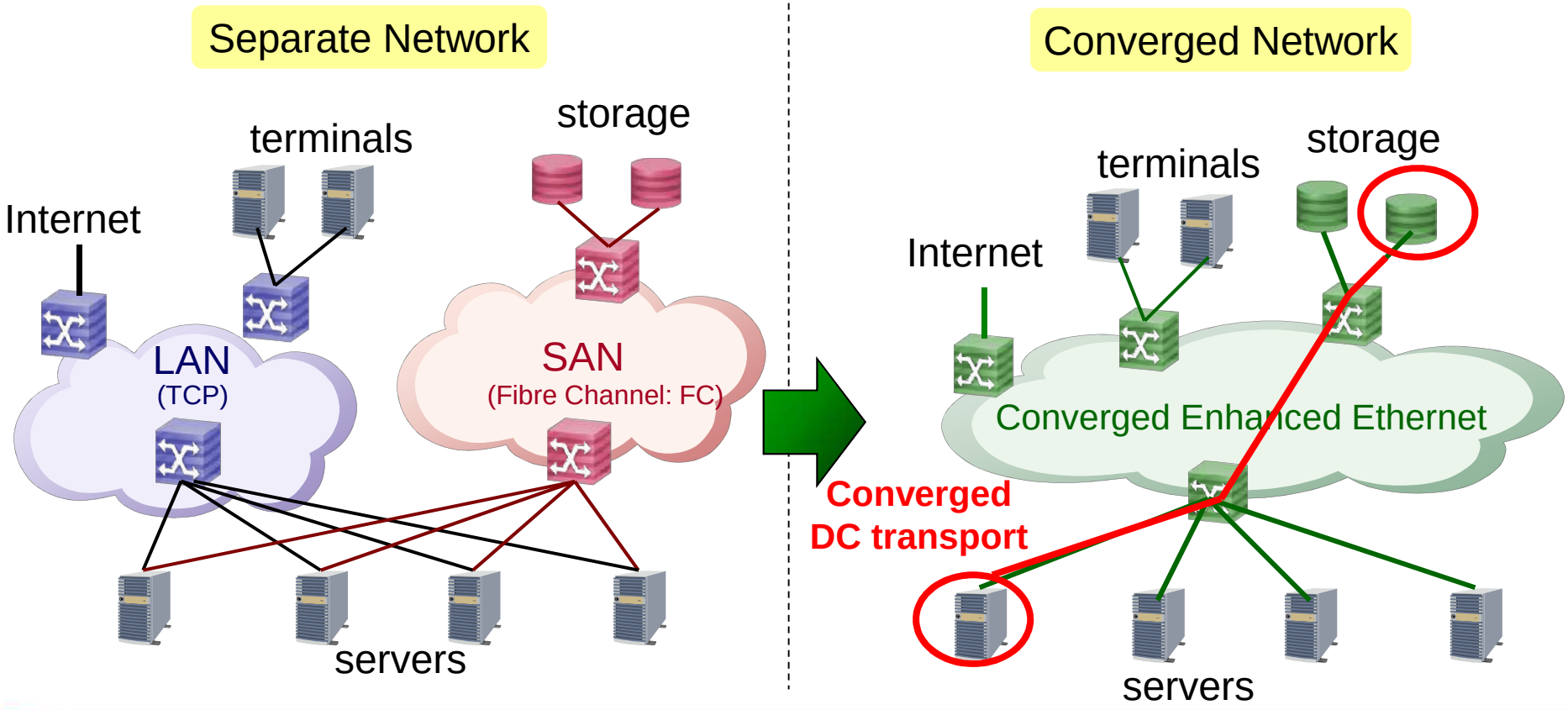
This work was partly supported by Ministry of Internal Affairs and Communications (MIC), Japan.

Outline

- Background: FCoE
- Issues of FCoE
- Proposed Architecture: “Advanced FCoE”
- Prototype Implementation and Evaluation

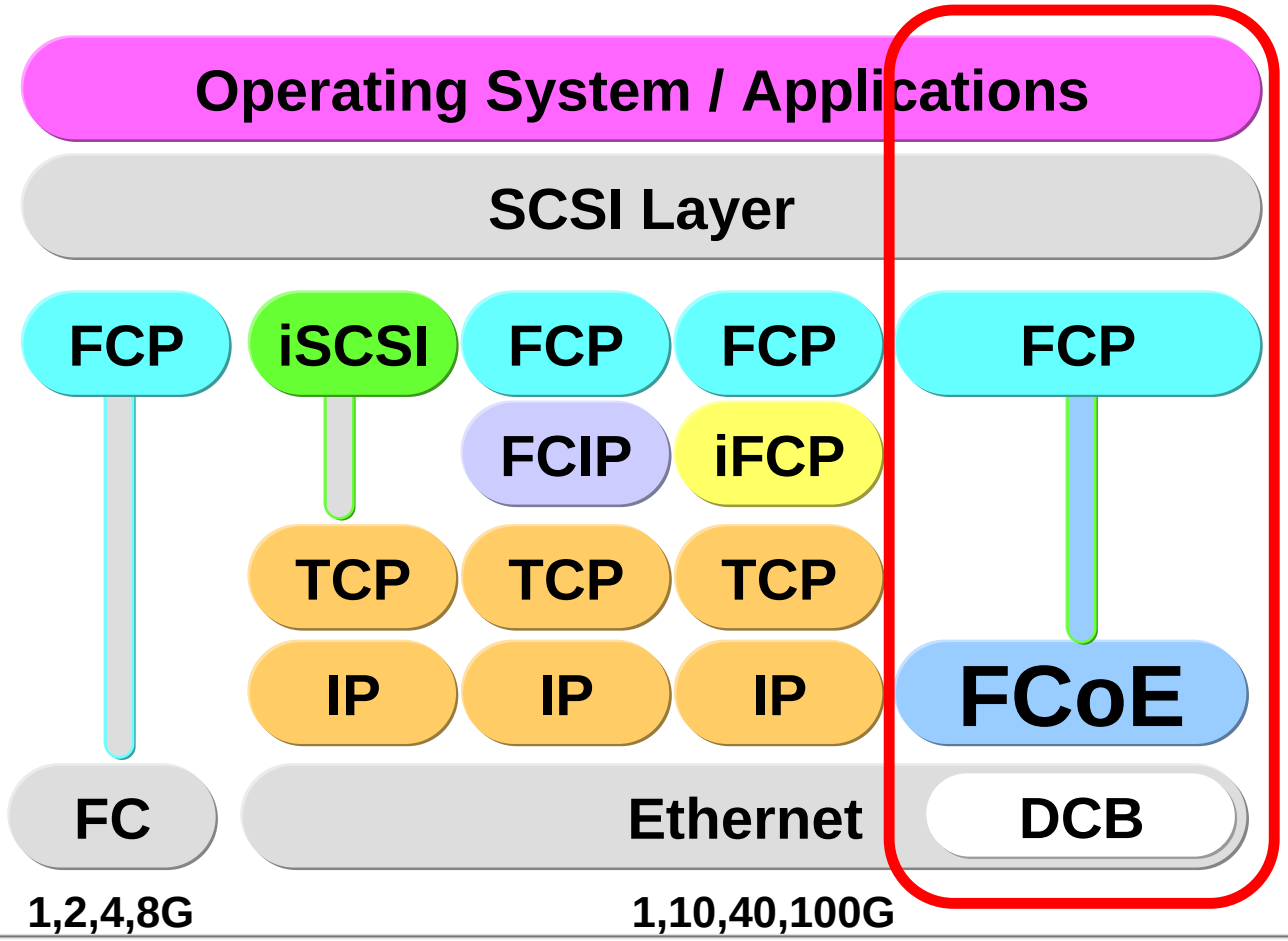
Background

- In data centers, there is a movement of I/O consolidation and Network Convergence
  - For Reduction of CAPEX and OPEX
  - Network Convergence :LAN (Ethernet) and SAN (IP-SAN, **FCoE)**



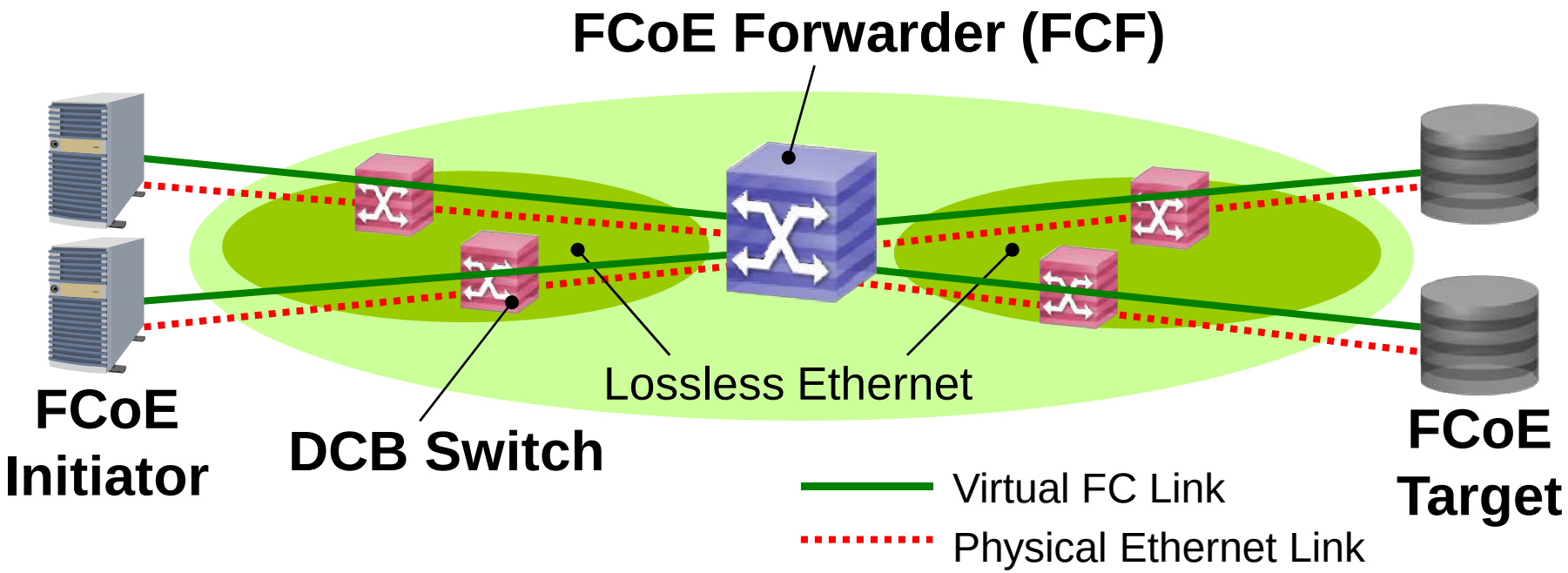
FCoE for LAN/SAN Convergence

- Simple Protocol
- Roadmap toward High-speed Ethernet (40G~100G)



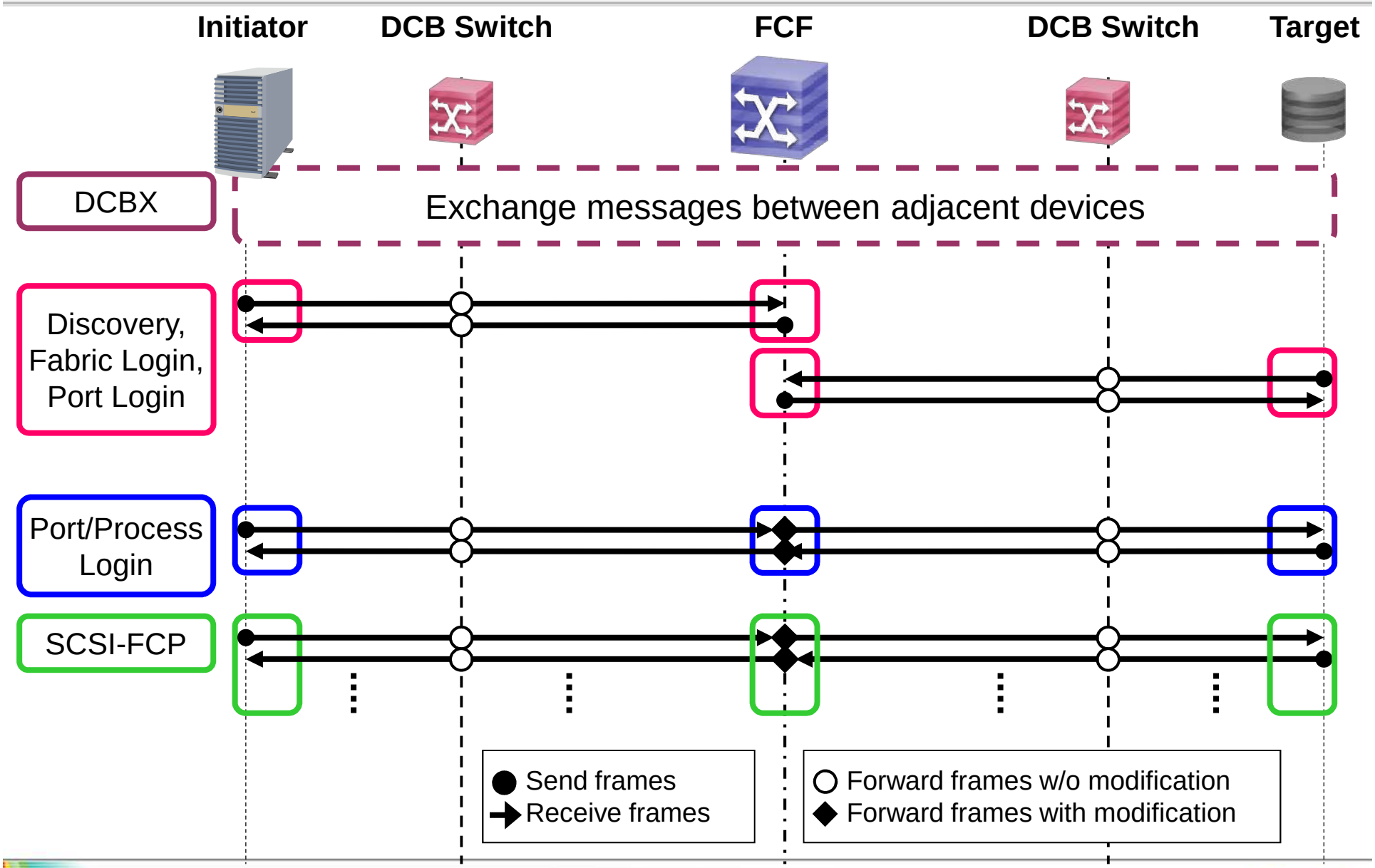
FCoE System

Consists of 4 components :
Initiator (Server), Target (Storage), FCF (FCoE Switch), DCB Switch

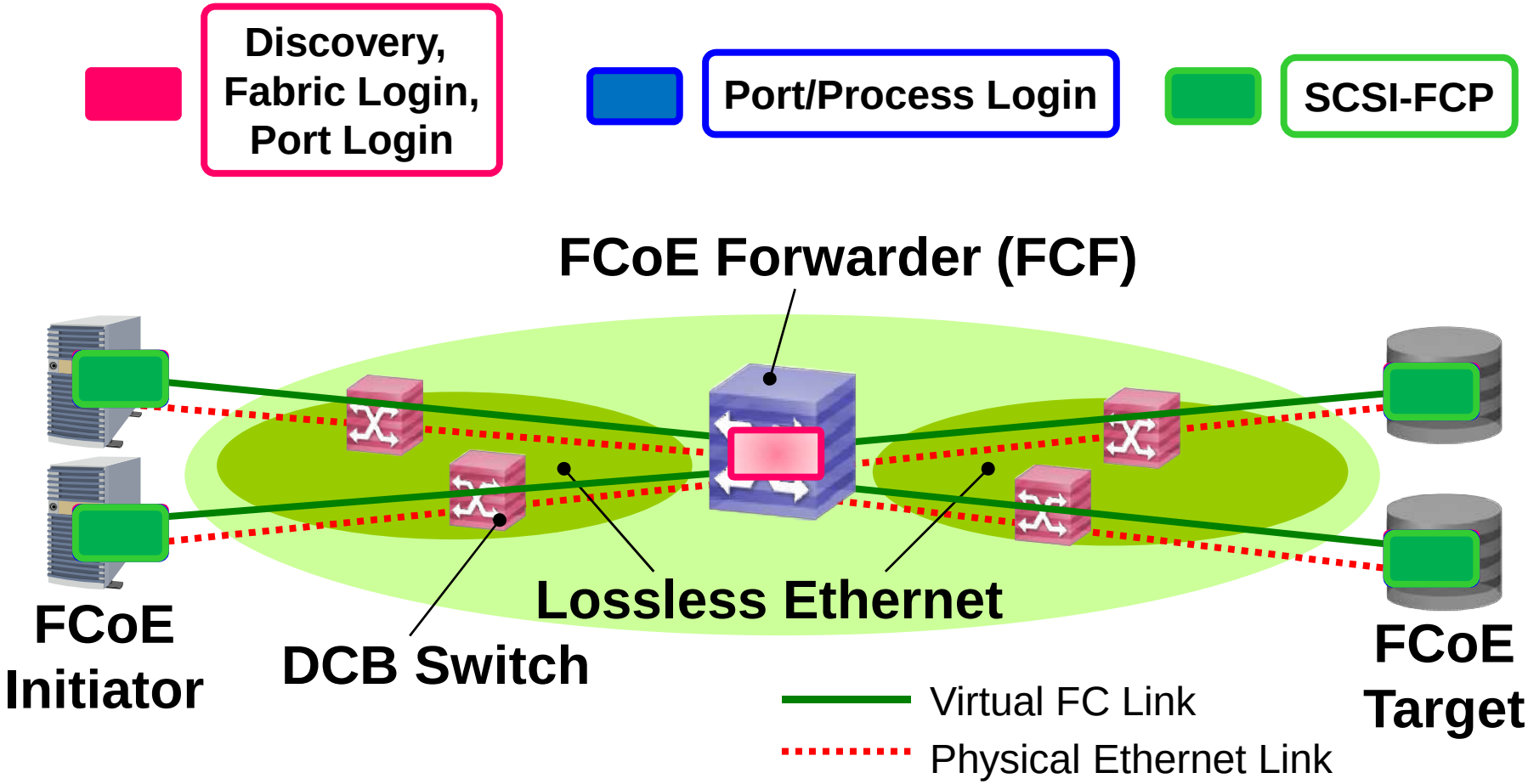


DCB : Data Center Bridging (specified in IEEE 802.1 DCB WG)

FCoE Protocol Sequence



FCoE Frame Forwarding



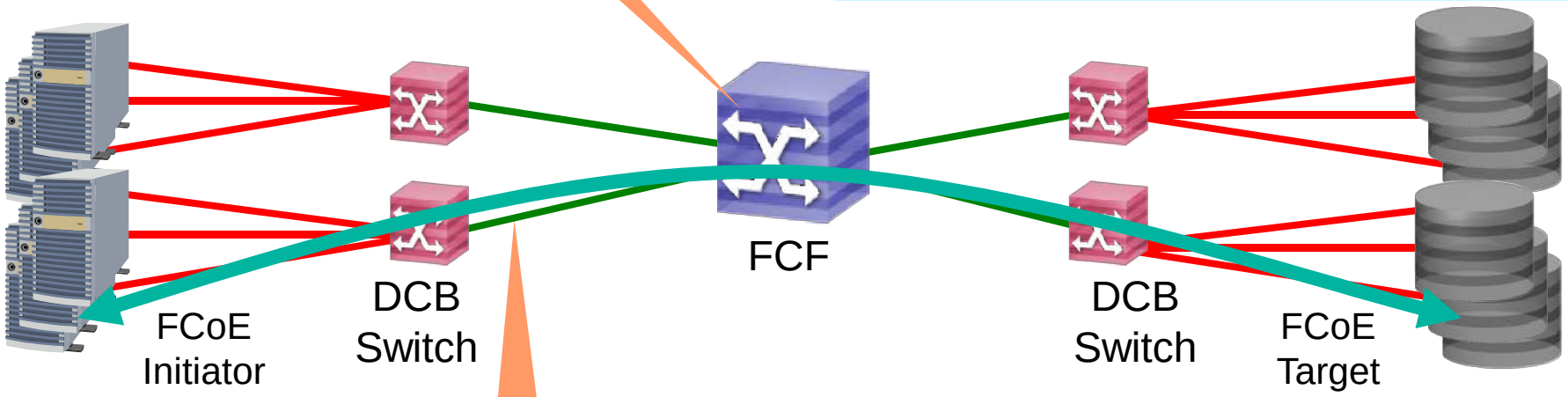
All FCoE traffic goes through FCF

Technical Issues of FCoE and Solution

Scalability Limitation

Go through all FCoE traffic in FCF.
→ FCF is bottle neck point in the system. Hard to scale

Solution: Separate U-plane and C-plane function in FCoE to realize virtually large-scale L2 Switch and "scale-free" FCoE system.



Performance Degradation by Data Loss

DCB does not support to retransmit and reorder frames. If frame loss occurs, SCSI-level timeout is long (>1sec).

Solution: Dataloss concealment using Datacenter Transport technologies (retransmission:R2D2, Packet order management)

Proposed Architecture “Advanced FCoE (AFCoE)”

- *U/C Separation*
  - Separate U-plane traffic (SCSI-FCP frames) and C-plane traffic (other frames)
- *Flat Data Transport Network by Using L2 Address*
  - Forward FCoE frames according to Ethernet MAC addresses

➔ Large scalability

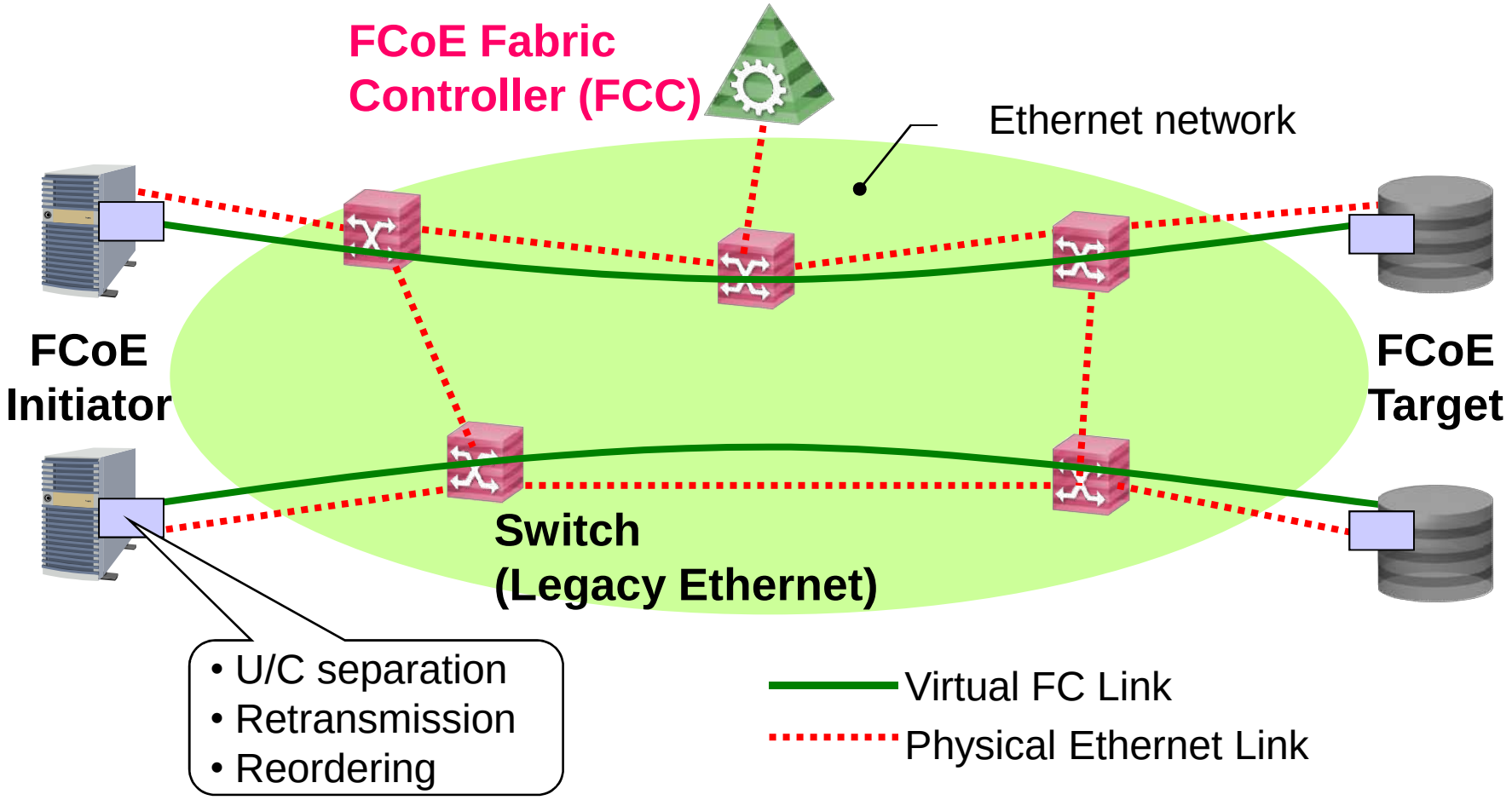
- *Reliable Ethernet Transport:* Edge based reliable Ethernet instead of lossless Ethernet provided by DCB
  - Fast retransmission function and reordering function into Ethernet layer

➔ Avoid performance degradation by data loss

➔ Reduce CAPEX by using legacy Ethernet switches

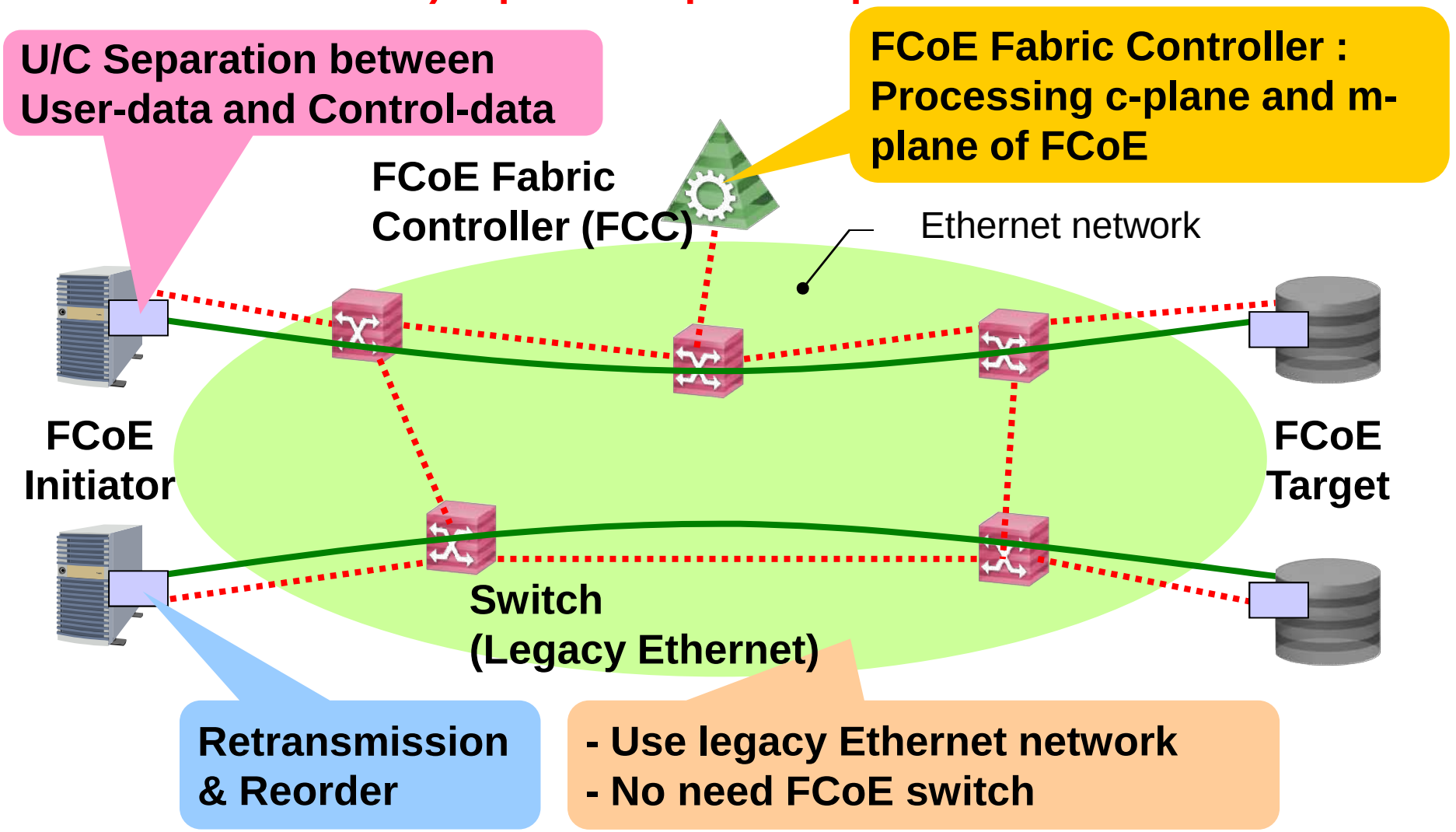
Advanced FCoE (AFCoE) Architecture

Consists of 4 components;
(1)initiator, (2) target, **(3) FCC,** and (4) switch



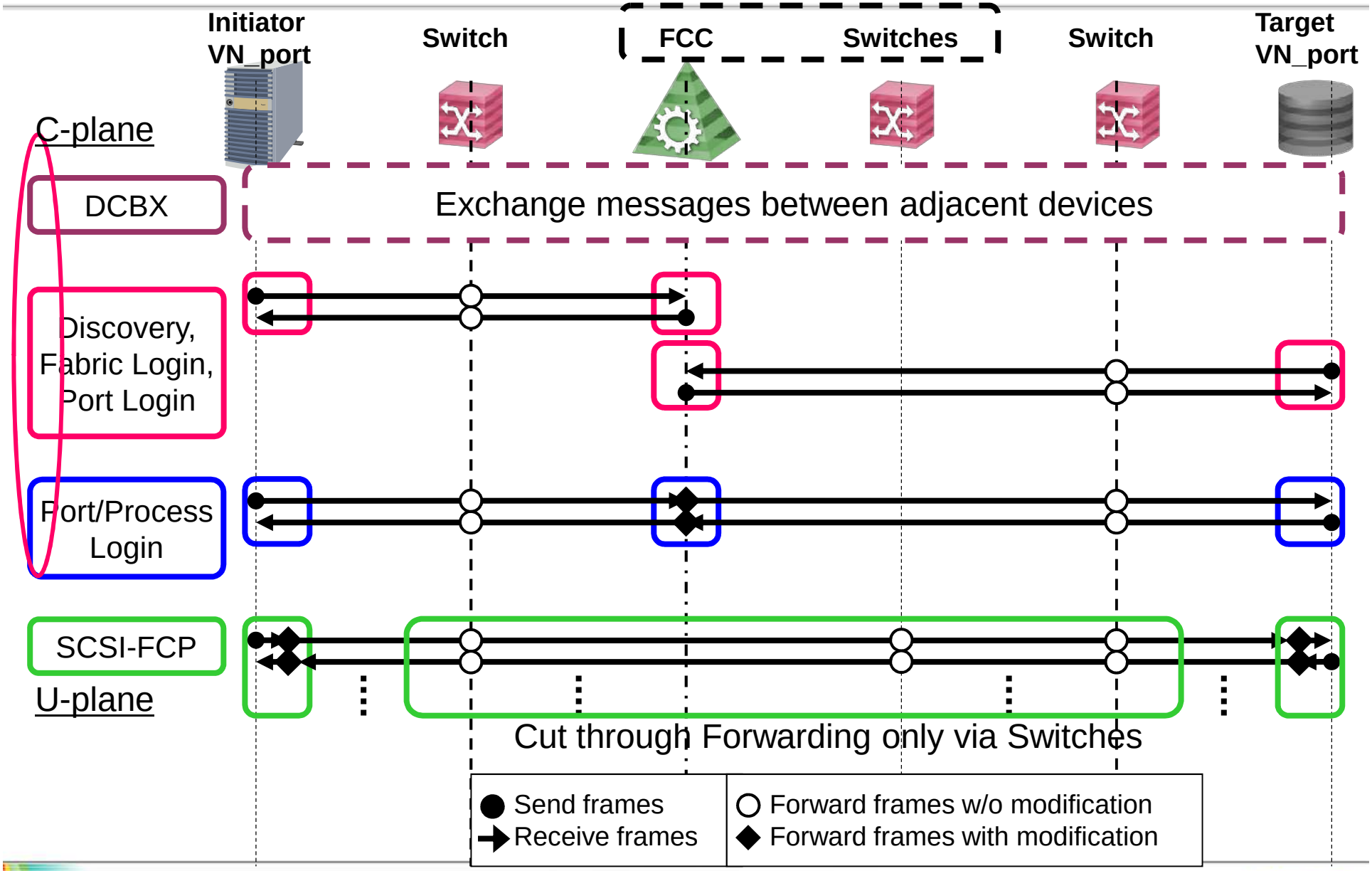
Features of AFCoE

1) U-plane / C-plane separation

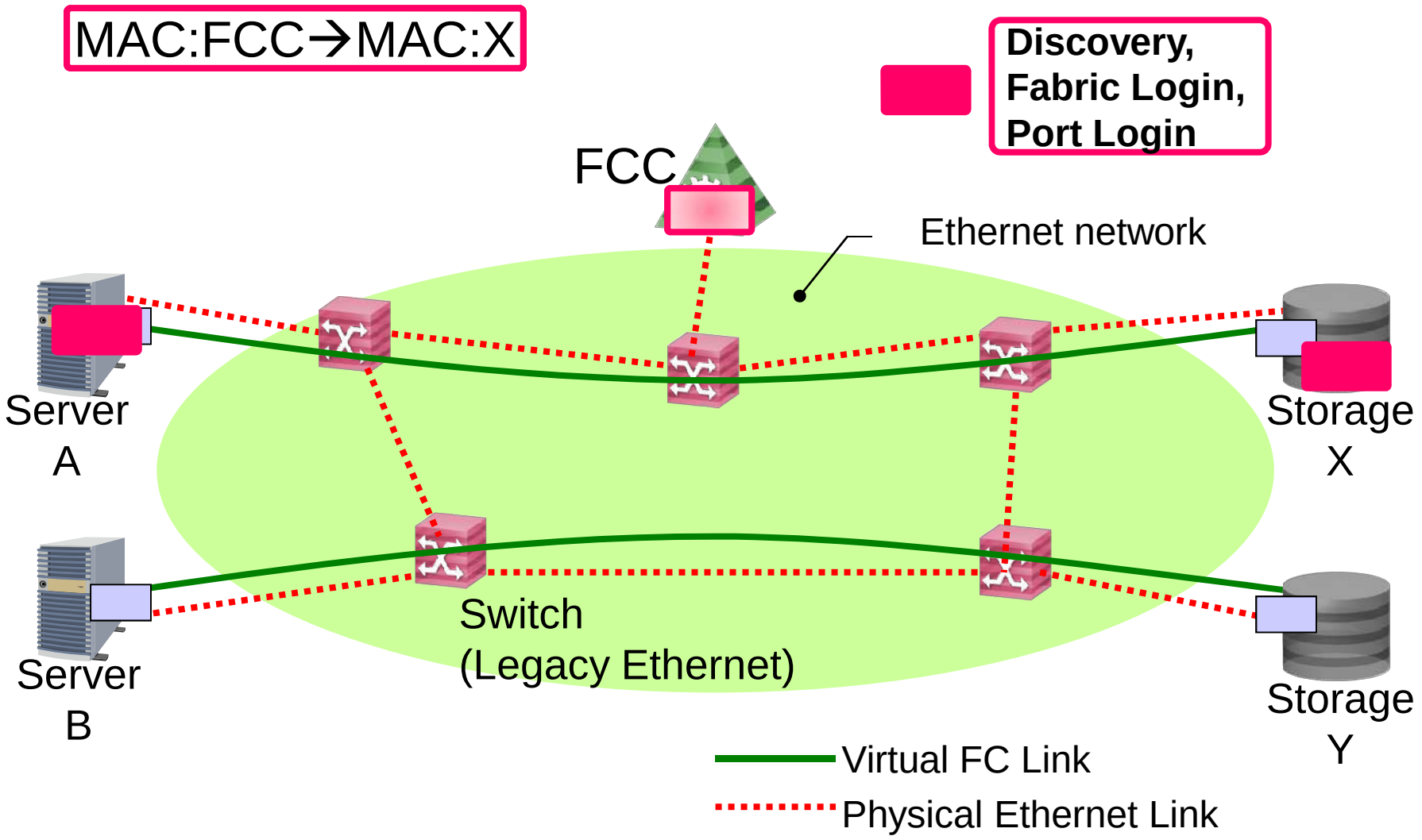


2) Edge-based Reliable Ethernet Transport

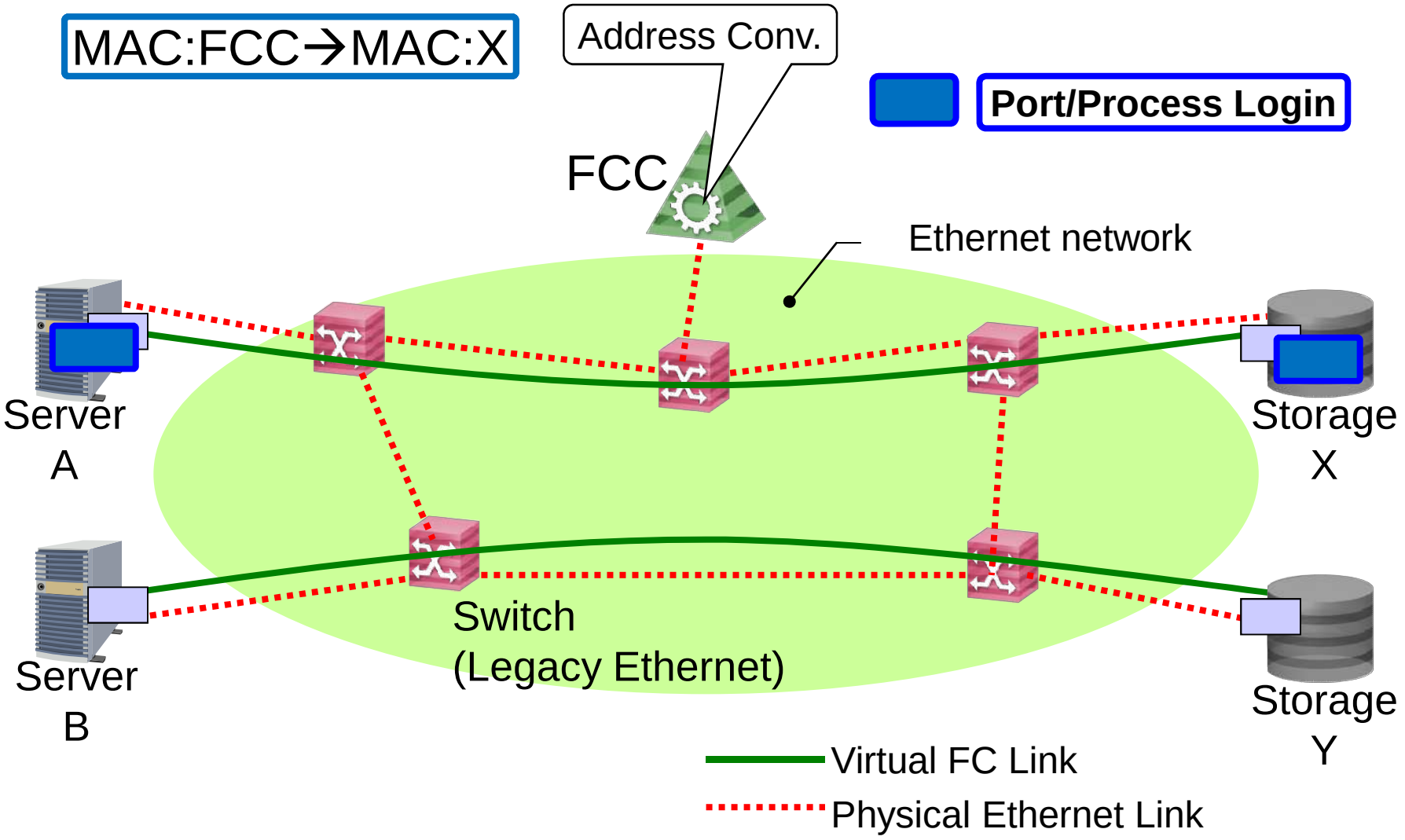
AFCoE Protocol Sequence



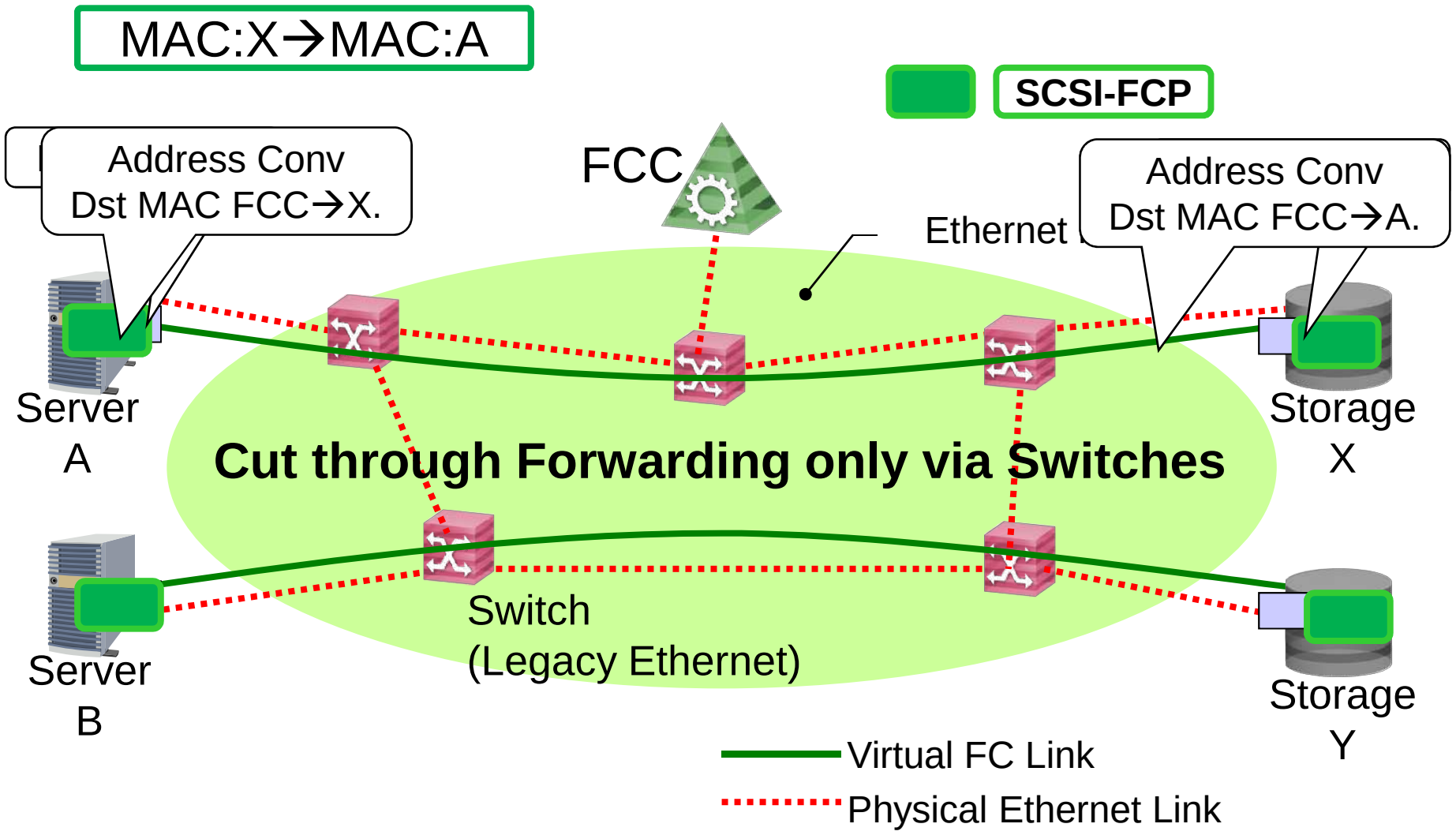
AFCoE Frame Forwarding (1)



AFCoE Frame Forwarding (2)

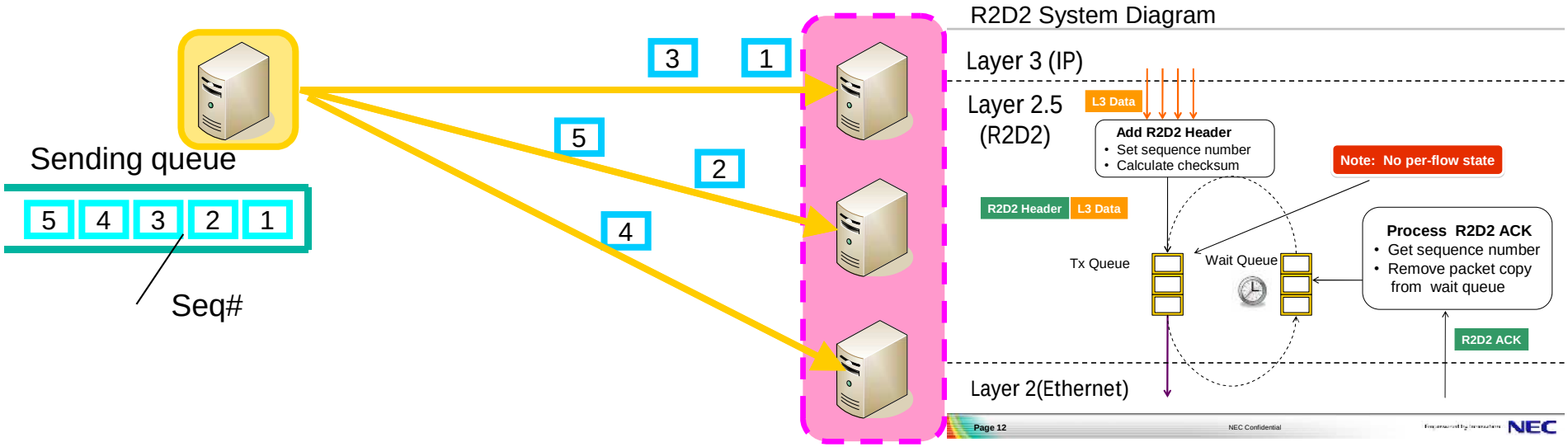


AFCoE Frame Forwarding (3)



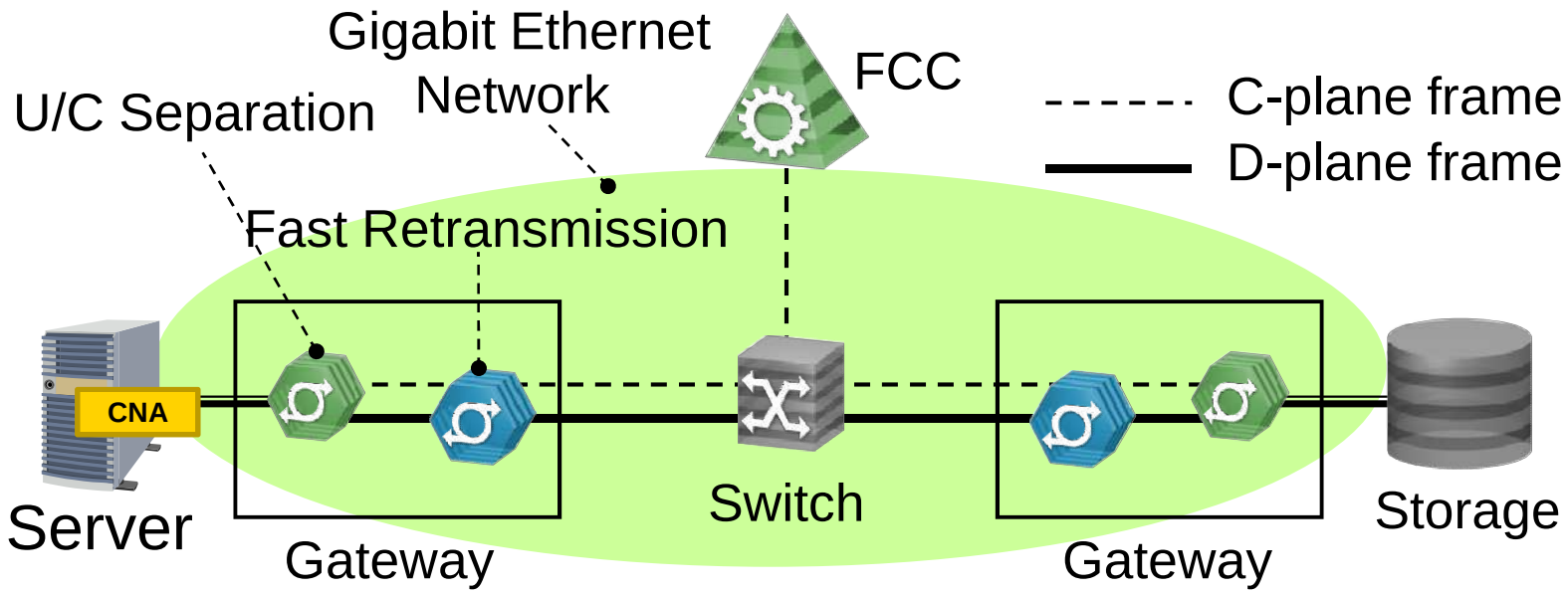
Edge-based Reliable Ethernet Transport

- Fast Retransmission and Reordering Function instead of DCB
- R2D2 : Rapid Reliable Data Delivery - Rapid Retransmission technology
  - Conceal packet loss
- Reordering
  - Packet reordering function
- Easy to Implementation
  - Enable low cost NIC. No change switch.



[Ref] B. Atikoglu, M. Alizadeh, J. S. Yue, B. Prabhakar and M. Rosenblum, R2D2: Rapid and Reliable Data Delivery in Data Centers, April 2010, “http://forum.stanford.edu/events/posterslides/R2D2RapidandReliableDataDeliveryinDataCenters.pdf.”

Prototype Implementation

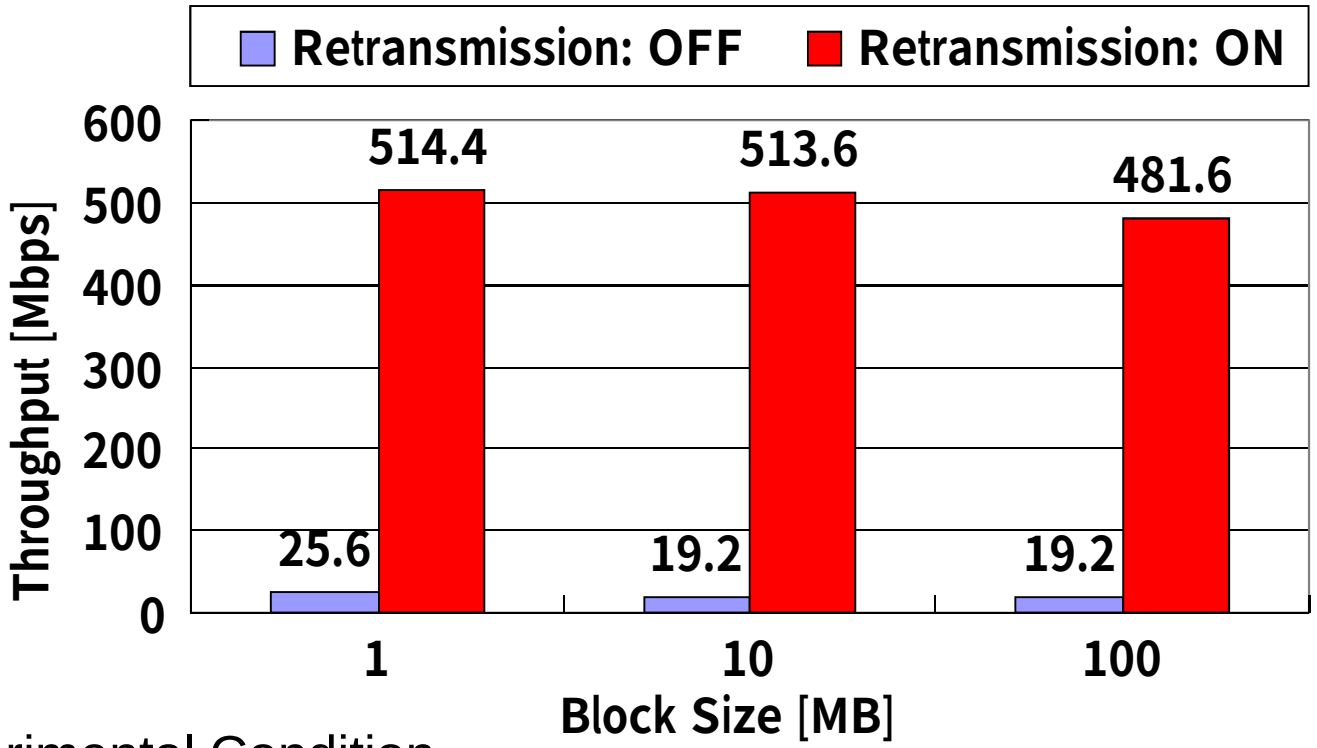


Equipment	Vendor, Model	OS	
Server	IBM x3650M2 (CNA: Qlogic QLE8142-SR)	RHEL5.3 (32bit)	Xeon Quadcore 2.53GHz 4GB
Storage	NetApp FAS3140	ONTAP 7.3.2P5	Mobile Celelon 2.2GHZ 4GB (NVRAM 512MB)
Switch	NEC QX-S5828T	-	
Gateway	NEC Mate MY33A/E7	CentOS5.5 (64bit)	Core 2Duo 3.33GHz 4GB
FCC	NEC Mate MY24A/B4	Fedora 13 (32bit)	Core 2 Duo 2.4GHz 2GB

We confirmed the whole sequence of AFCoE.

Performance Evaluation: Reliable Ethernet in AFCoE

Reliable Ethernet improves FCoE throughput under lossy situation

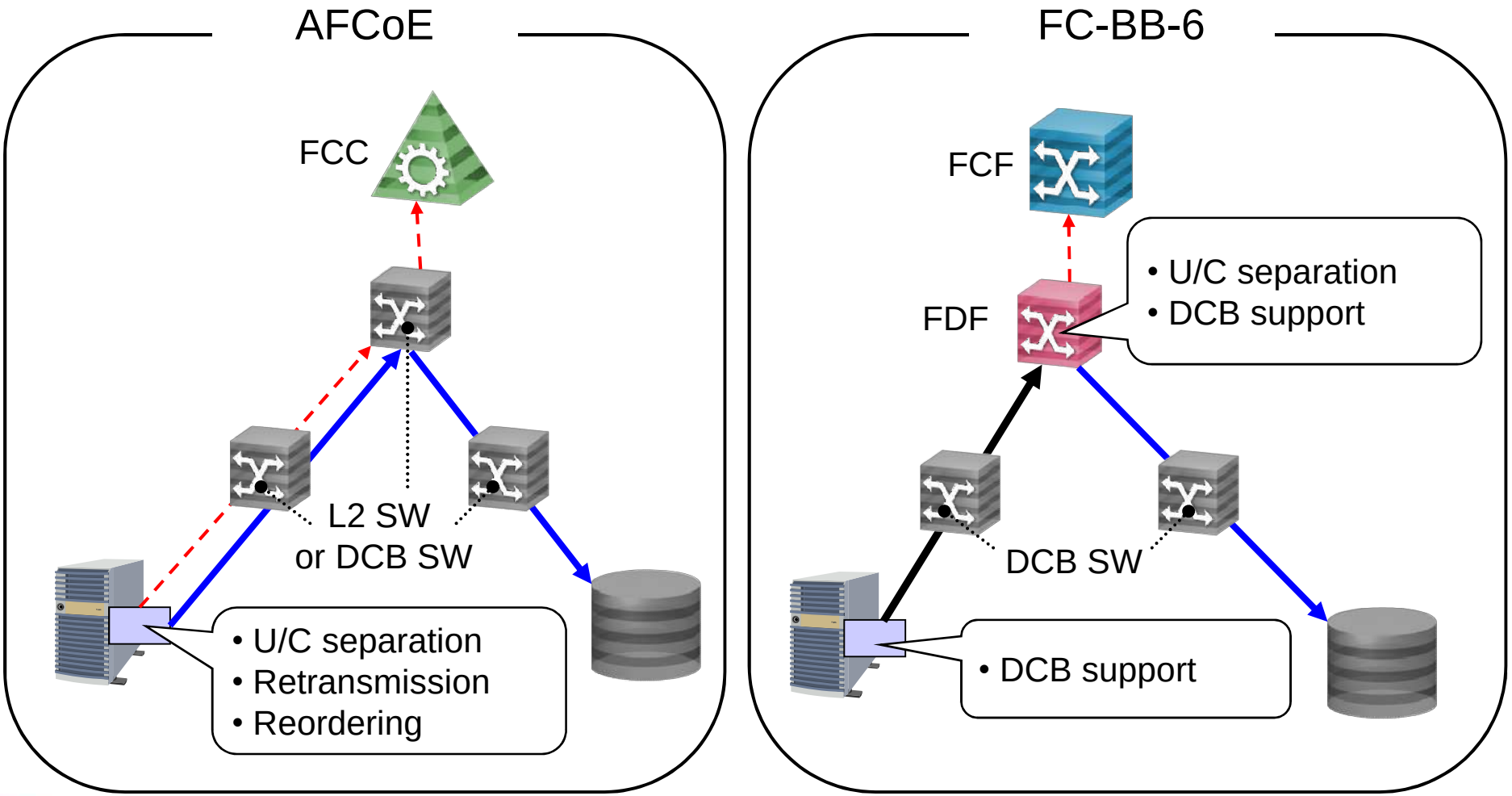


Experimental Condition

- D-plane line speed: 1000Mbps
- Packet drop rate: 1% (random drop)
- FC link timeout: 10 seconds
- Retransmission timeout: 100 micro seconds

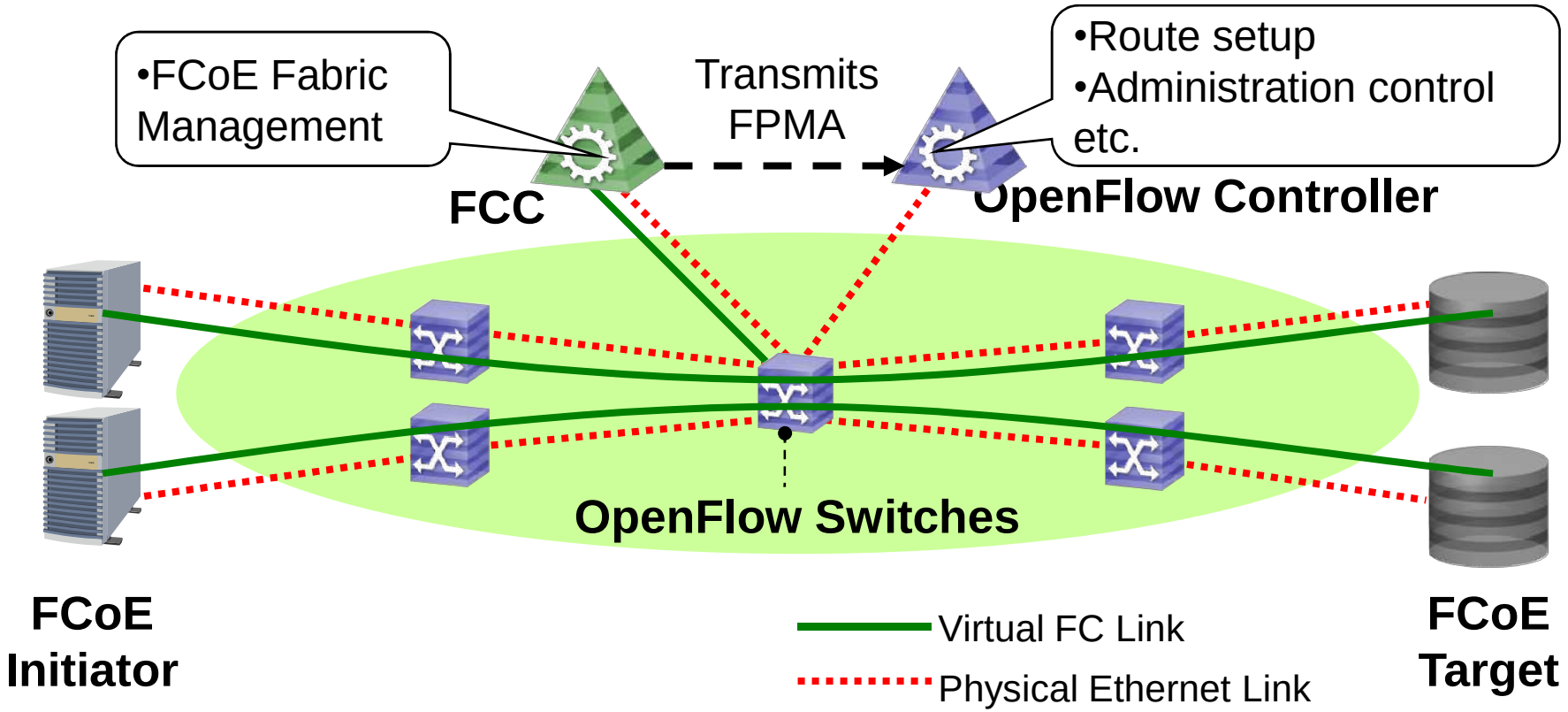
Comparison between AFCoE and FC-BB-6

AFCoE does not need FDFs (FC-aware switches).
AFCoE makes network simple and flat compared to FC-BB-6.



Next Enhancement: OpenFlow-based AFCoE System

- LAN/SAN unified management with OpenFlow
- Makes FCoE (SAN) network more efficient :
  - Multi-path setup for redundancy and bandwidth
  - Rapid reroute in network failure



Conclusion

- Advanced FCoE (AFCoE) has been proposed
  - Enhanced FCoE system
  - Addresses FCoE's issues: Scalability and performance degradation
- Confirmed correct operations of basic AFCoE system

- Future Work
  - Evaluations in more complex network
  - OpenFlow-based AFCoE system

Empowered by Innovation