



# DCB – Datacenter Bridging

**Сергей Кацубо**

**skatsubo@solidex.by**

28 октября 2011

## DCB (Datacenter Bridging)

- Добавляет новые качества в хорошо знакомый Ethernet
- Позволяет использовать «улучшенный» Ethernet в качестве универсальной транспортной среды в центрах обработки данных

Описывается семейством стандартов IEEE

# Содержание

- **Проблемное поле**
- Новые свойства Ethernet

## Зачем

**что-то менять в Ethernet?**

## Что

**не устраивает нас на L1/L2?**

# 1. Зоопарк

- Сколько L1/L2 протоколов используется в Вашем ЦОД?
  - Только Ethernet? Поздравляю!
- Сколько **разнородных** сетей может быть в ЦОД?

Сеть	Технология	Адаптер	Пропускная способность	
LAN	Ethernet + IP	NIC	10 Мбит/с (1985)	100 Гбит/с
SAN	Fibre Channel (FC)	HBA	200 Мбит/с (1997)	5.1 Гбит/с
HPC	Infiniband	HCA	2 Гбит/с (1999)	96 Гбит/с (12X QDR)

# Разные технологии под разные задачи

- **FC:** для сети хранения данных требуются
  - Малые задержки
  - Транспорт без потерь (loseless)
- **Infiniband:** для высокопроизводительных кластеров требуются
  - Малые задержки
  - Транспорт без потерь (loseless)
  - Высокая пропускная способность

Каждая сеть – собственные технологии, оборудование, специалисты

- Можно ли отказаться от FC и использовать Ethernet в сети хранения данных?
- Можно ли отказаться от Infiniband и использовать Ethernet в среде высокопроизводительных вычислительных систем?
- LAN, SAN, HPC = Ethernet ?

**Выигрыш – простота и стоимость**

# Требования к универсальной фабрике

От Ethernet потребуются новые «умения»

1. Гарантировать **доставку без потерь**
  2. Гарантировать **минимальную задержку**
  3. Гарантировать **полосу пропускания**
- для определенных видов трафика

**А что мы имеем  
сейчас?**



## 2. Проблемы Ethernet

### Скверный характер «Best Effort»

- Задержки в сети – непредсказуемы
- Потери кадров – возможны
- Механизмы QoS, встраиваемые в коммутаторы, имеют локальный характер
  - В некоторой степени облегчают ситуацию для TCP/IP, но не решают проблему

## 2. Проблемы Ethernet (2)

### Склонность к образованию петель

- **Spanning Tree** решает проблему ценой блокирования каналов – теряем производительность сети
- Виртуализация (перемещение VM без изменения IP адреса) требует плоский L2 ЦОД – домен STP распространяется на весь ЦОД

# Итак, мотивы для изменений

Нам хотелось бы

1. Использовать Ethernet как универсальную сеть (фабрику)
2. Устранить «родные» проблемы Ethernet

# Содержание

- Проблемное поле
- **Новые свойства Ethernet**

# Совершенствование Ethernet для ЦОД

- 1. Доставка трафика без потерь (loseless)**
- 2. Передача трафика с минимальными задержками и гарантированной полосой пропускания**
- 3. Отказ от STP: параллельное использование каналов**

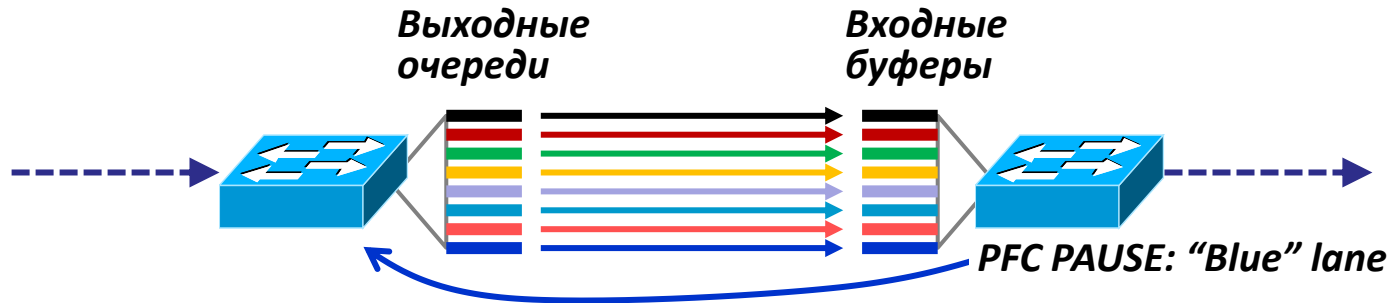
# Datacenter Bridging

Свойство	Назначение
<b>Priority-based Flow Control (PFC)</b> IEEE 802.1Qbb	Позволяет сети приостанавливать передачу определенных классов трафика. Обеспечивает доставку без потерь (loseless).
<b>Enhanced Transmission Selection (ETS)</b> IEEE 802.1Qaz	Управляет передачей трафика с соблюдением приоритета и гарантированной полосы пропускания для каждого класса.
<b>Data Center Bridging Exchange Protocol (DCBX)</b> IEEE 802.1Qaz	Сигнальный протокол для согласования настроек (PFC, ETS) между соседними устройствами.
<b>Congestion Notification</b> IEEE 802.1Qau	Идентифицирует потоки, приводящие к перегрузкам очереди, и отправляет уведомления к источнику трафика.

# Замена для STP

Свойство	Назначение
<b>Layer 2 Multipath</b> IEEE 802.1aq SPB IETF RFC 5556 TRILL	Позволяет использовать несколько параллельных путей в сети одновременно.

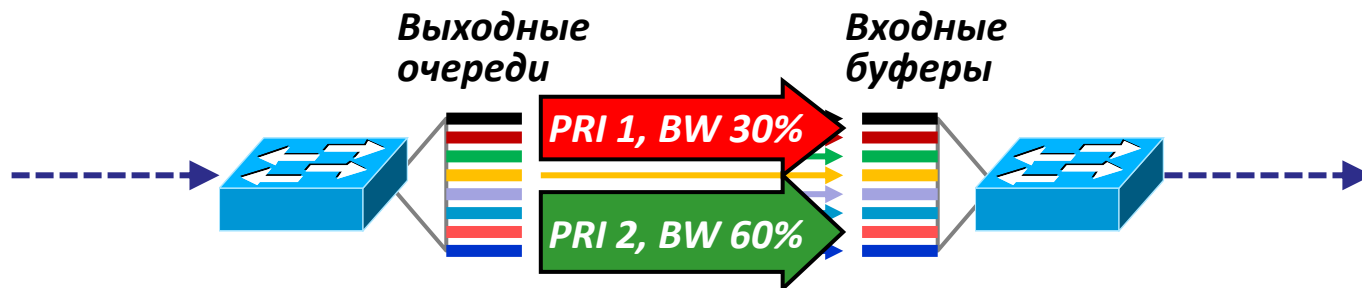
# Priority-based Flow Control (PFC)



- Разделение канала на 8 виртуальных линий (Virtual Lane) согласно значению dot1p
- Отправка PAUSE для данной линии при исчерпании соответствующего ей буфера
- Можем, например, организовать
  - 1 линию без потерь (SAN)
  - 7 линий – в традиционном режиме сброса трафика при переполнении буфера

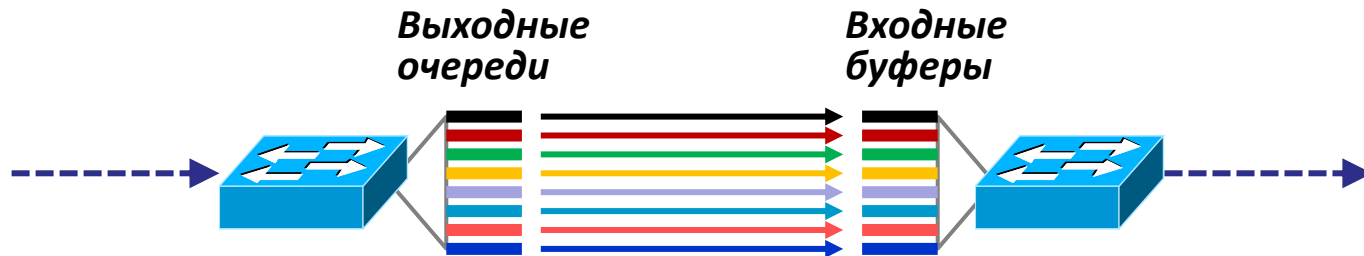


# Enhanced Transmission Selection (ETS)



- Передача трафика на основе приоритета (Priority) и минимальной полосы (Bandwidth)
- Разделение пропускной способности канала между группами Virtual Lane

# Data Center Bridging Exchange Protocol (DCBX)



- **Согласование конфигурации DCB между соседними коммутаторами**
  - Параметры управления потоком (PFC) и передачей трафика (ETS)
  - Логическое состояние Virtual Lane
- **Кадр DCBX базируется на LLDP (802.1AB)**

# SOLIDEX

**Вопросы ?**

**Сергей Кацубо**  
[skatsubo@solidex.by](mailto:skatsubo@solidex.by)



**SOLIDEX<sup>®</sup> PI**

**Internetworking Solutions**