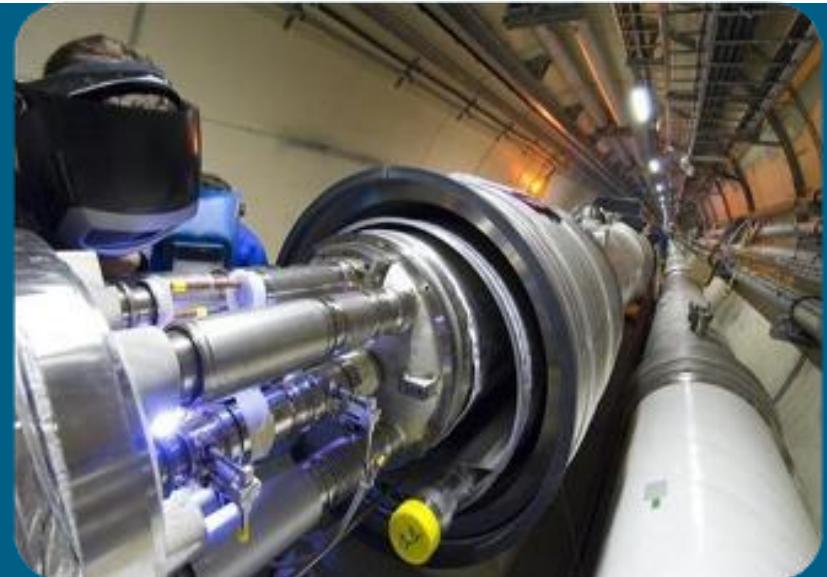




Обзор архитектур современных маршрутизаторов



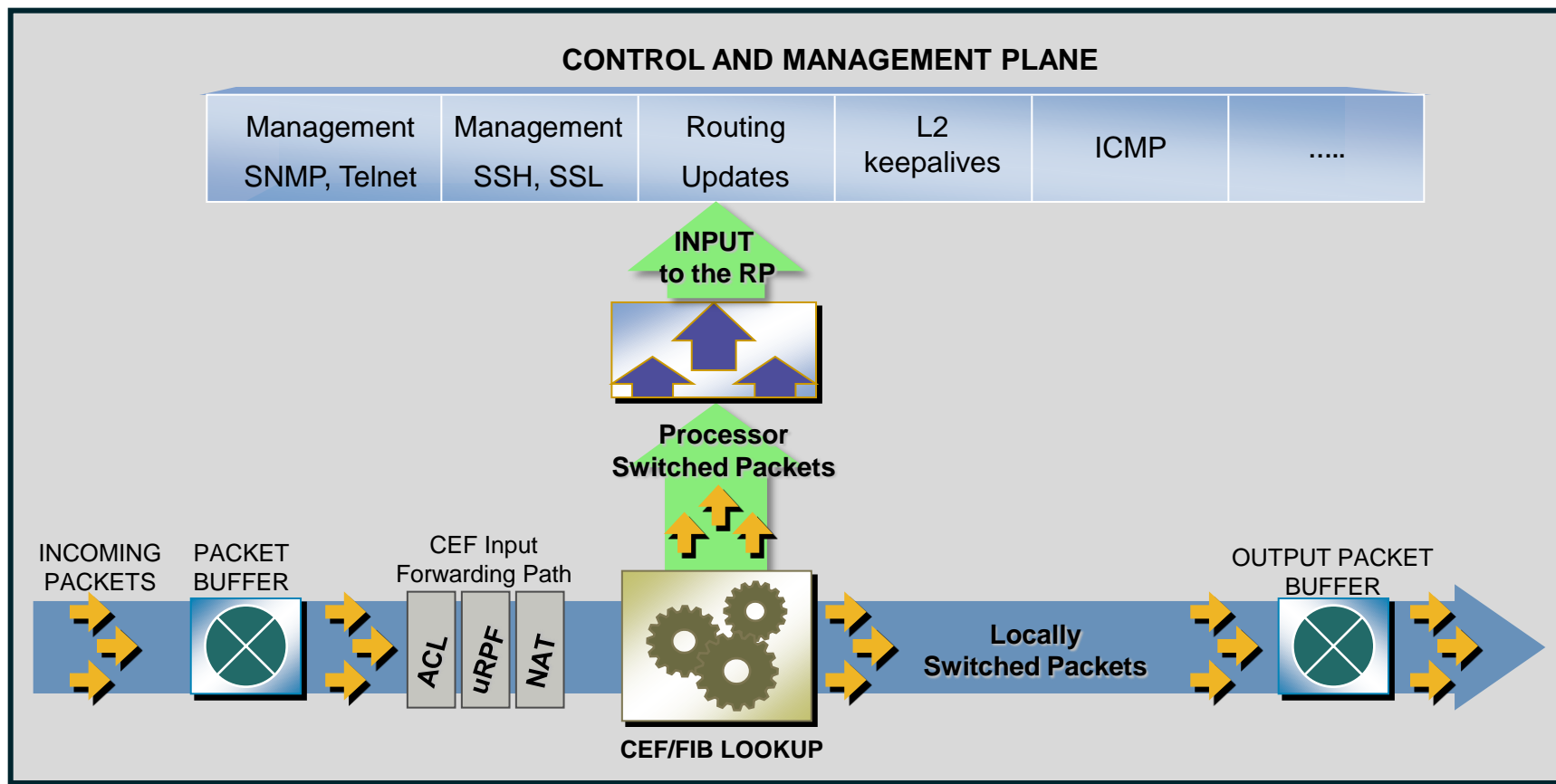
Артур Якупов
системный инженер

Программа

- Введение
- Централизованная архитектура на примере ASR1000
- Q & A

Основная задача маршрутизатора – обрабатывать трафик

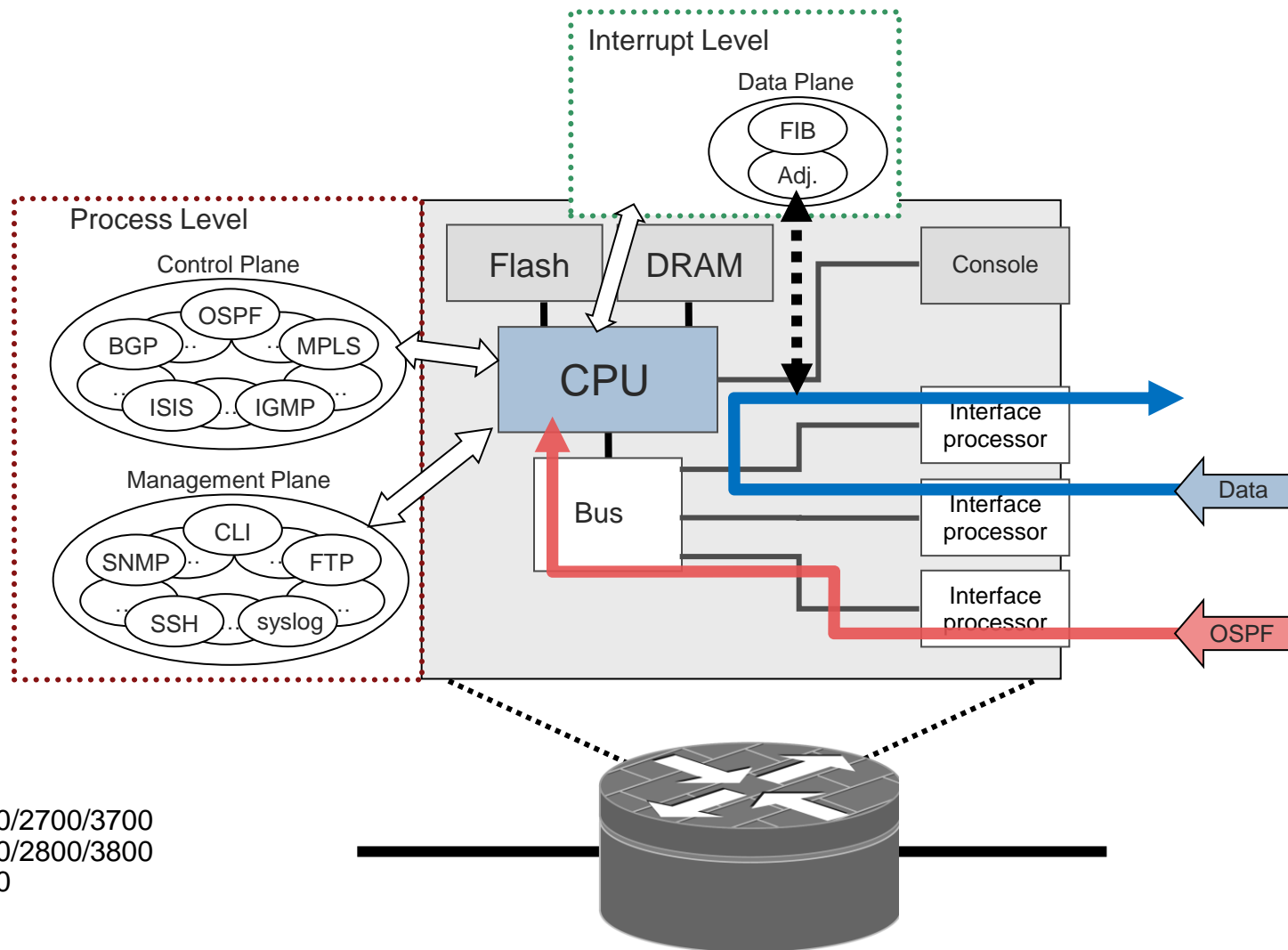
Сетевые устройства обрабатывают трафик в трех плоскостях: Data Plane, Control Plane и Management Plane...



Виды обработки трафика

- Сетевые устройства обрабатывают трафик в плоскостях Data Plane, Control Plane и Management Plane
 - Трафик Data Plane всегда проходит сквозь устройство и:
 - Обработывается аппаратно (ASIC/NP) на старших платформах
 - Коммутируется с помощью CEF на уровне прерываний на программных платформах
 - Трафик же Control Plane и Management Plane всегда идет на устройство и обрабатывается CPU на уровне процессов:
 - В платформах с аппаратной обработкой трафик Control/Management Plane посылается сначала на RP/MFSC а затем к конкретному программному процессу (OSPF, BGP, SSH, и т.д.)
 - В платформах с программной обработкой трафик посылается напрямую к процессу
 - На Route Processor попадает и некоторый процент трафика data plane
 - Несмаршрутизированные пакеты поднимаются (punted) на RP, для генерации сообщений ICMP unreachable
 - Пакеты с IPv4 опциями также обрабатываются RP
 - Первые пакеты потоков multicast, сессий NAT, IOS FW, CEF Gleans, и др. тоже попадают на RP
 - А также некоторые другие...

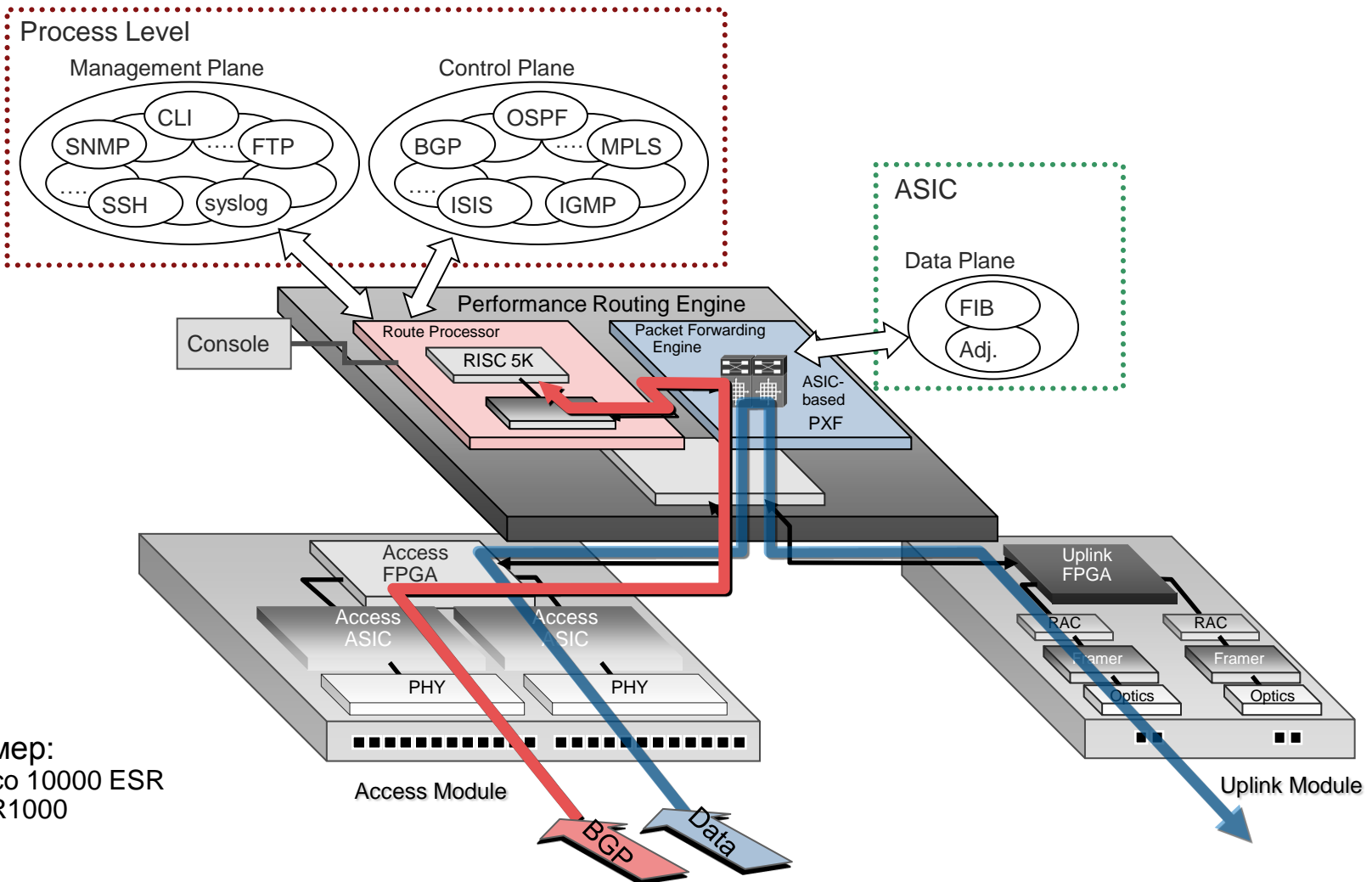
Централизованная архитектура маршрутизатора, основанная на CPU



Примеры:

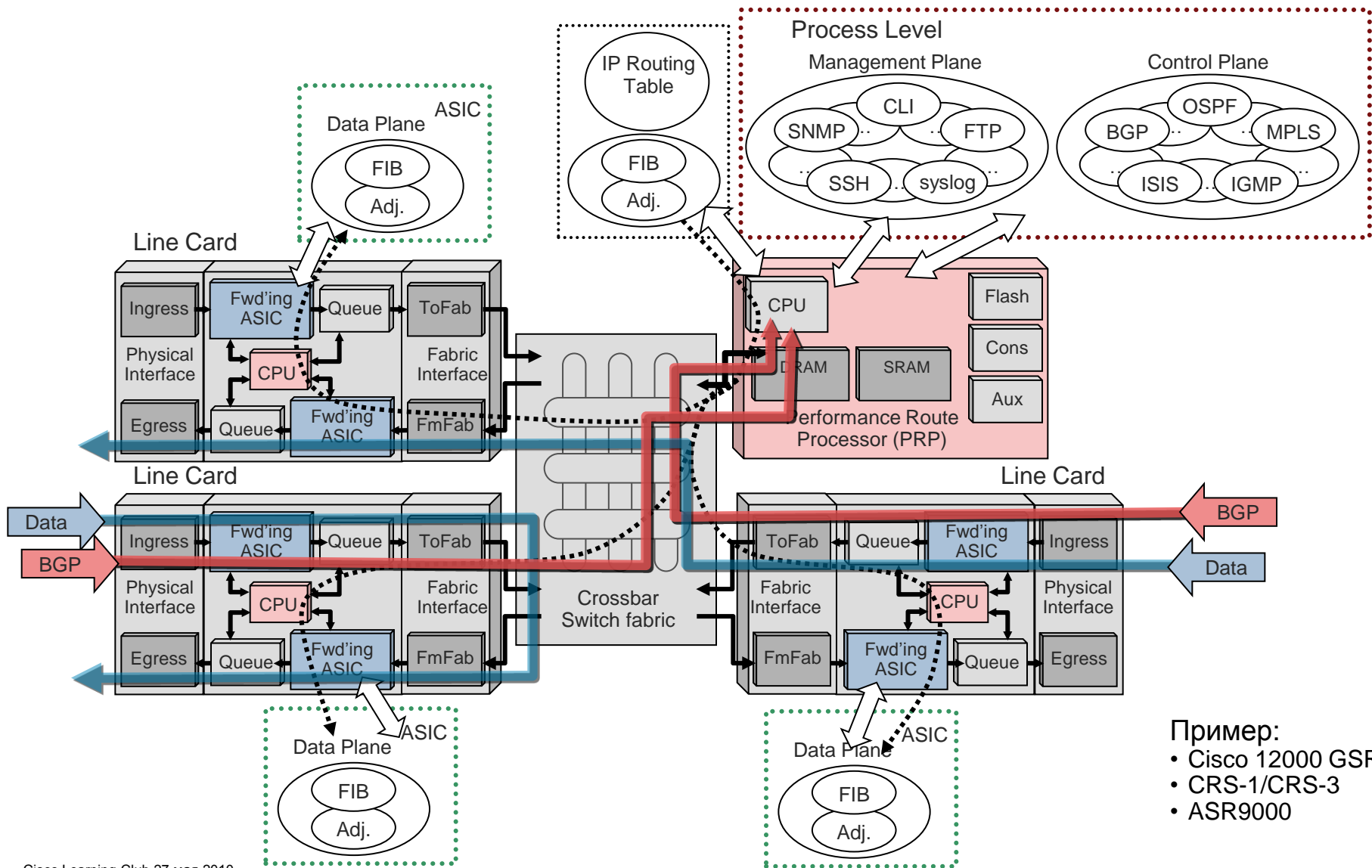
- Cisco 1700/2700/3700
- Cisco 1800/2800/3800
- Cisco 7200

Централизованная архитектура маршрутизатора, основанная на ASIC/NP



- Пример:
- Cisco 10000 ESR
 - ASR1000

Распределенная архитектура маршрутизатора, основанная на ASIC/NP

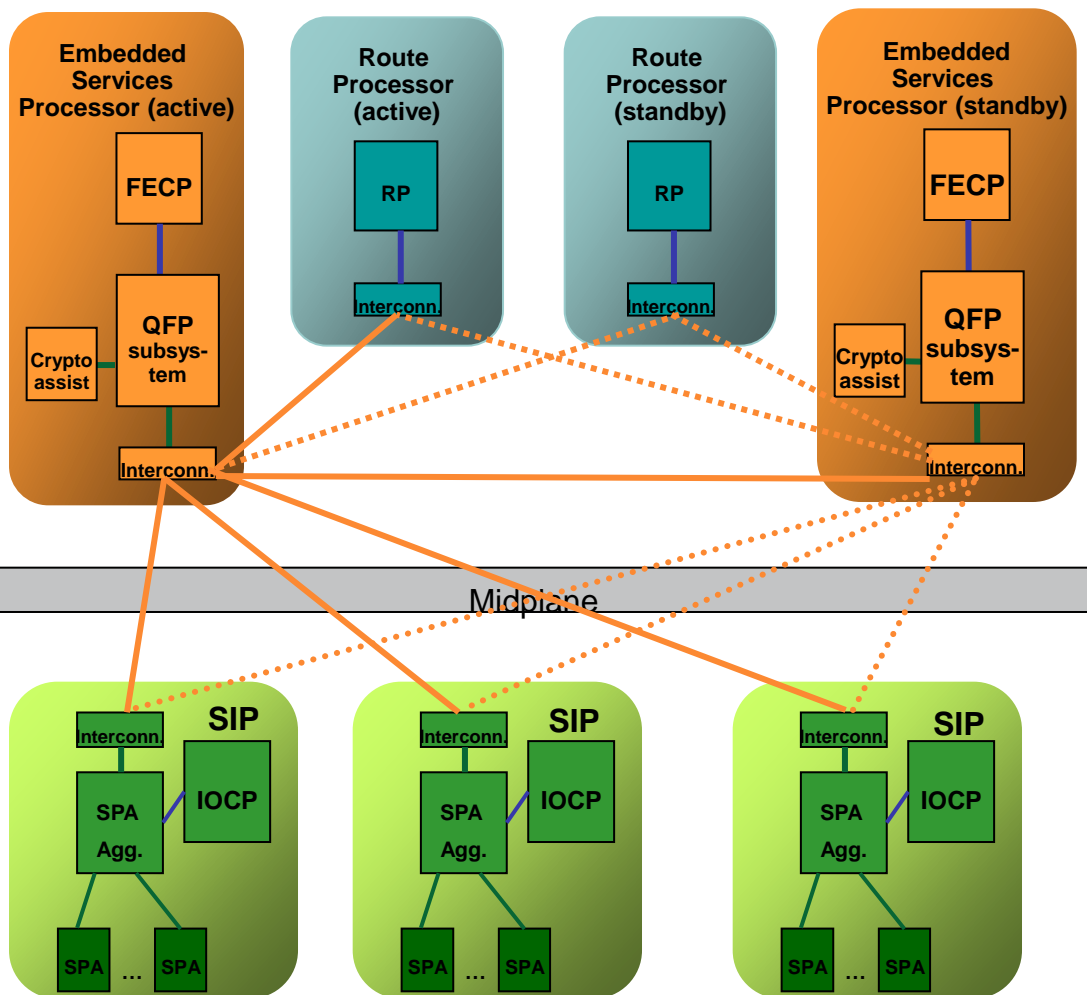


Пример:
 • Cisco 12000 GSR
 • CRS-1/CRS-3
 • ASR9000

Программа

- Введение
- **Централизованная архитектура на примере ASR1000**
- Q & A

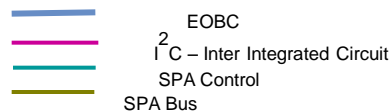
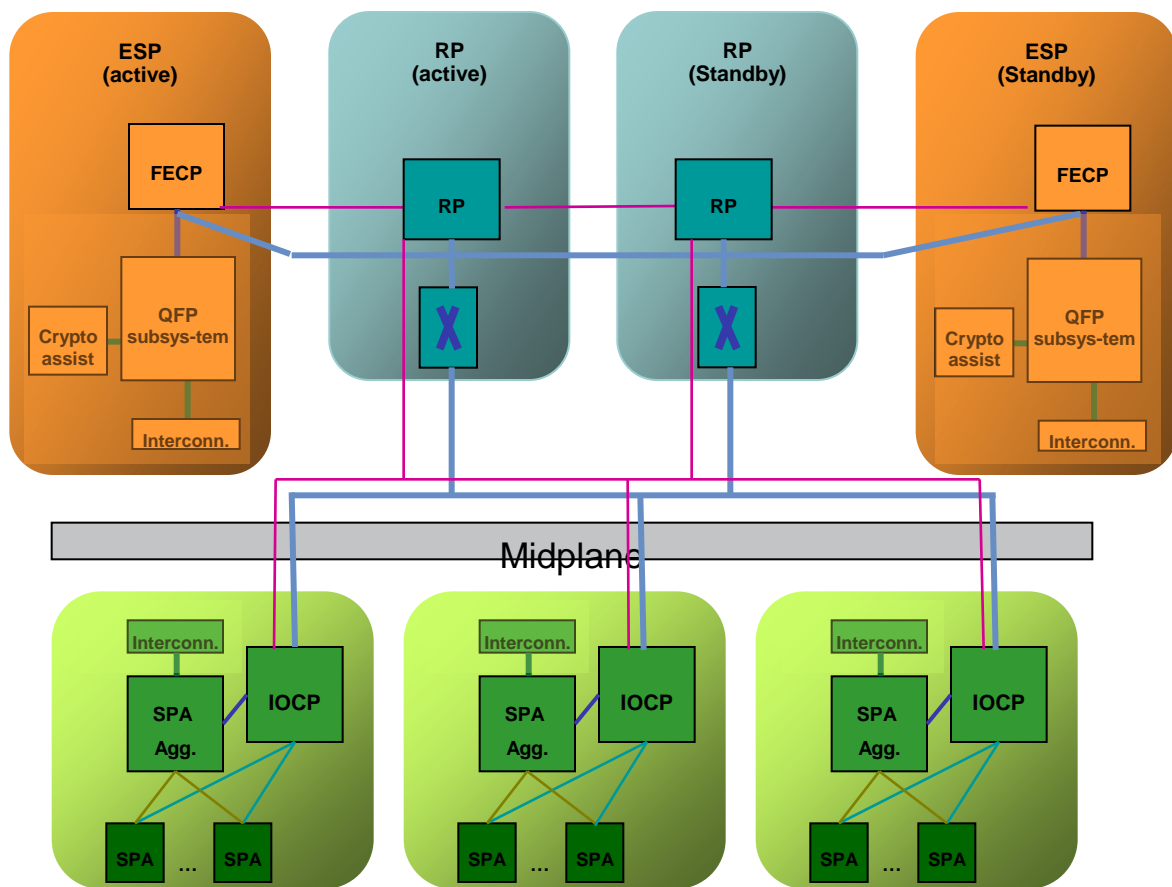
Аппаратная архитектура ASR 1000



- RP (Route Processor)
 - Control plane
 - Управление системой
- Embedded Services Processor (ESP)
 - Обрабатывает трафик и сервисы
- SPA Carrier Card (SIP)
 - Содержит SPA
- Shared Port Adapters (SPAs)
 - Внешние интерфейсы
- Централизованная архитектура
 - Весь трафик через ESP

Control Plane

- Два различных соединения control plane не пересекающихся с dataplane соединениями
 - Ethernet out-of-band Channel (EOBC).
 - I²C – Мониторинг состояния компонент
- Управляющие соединения SPA
 - Между IOCP и SPA



Route Processor – RP1 & RP2

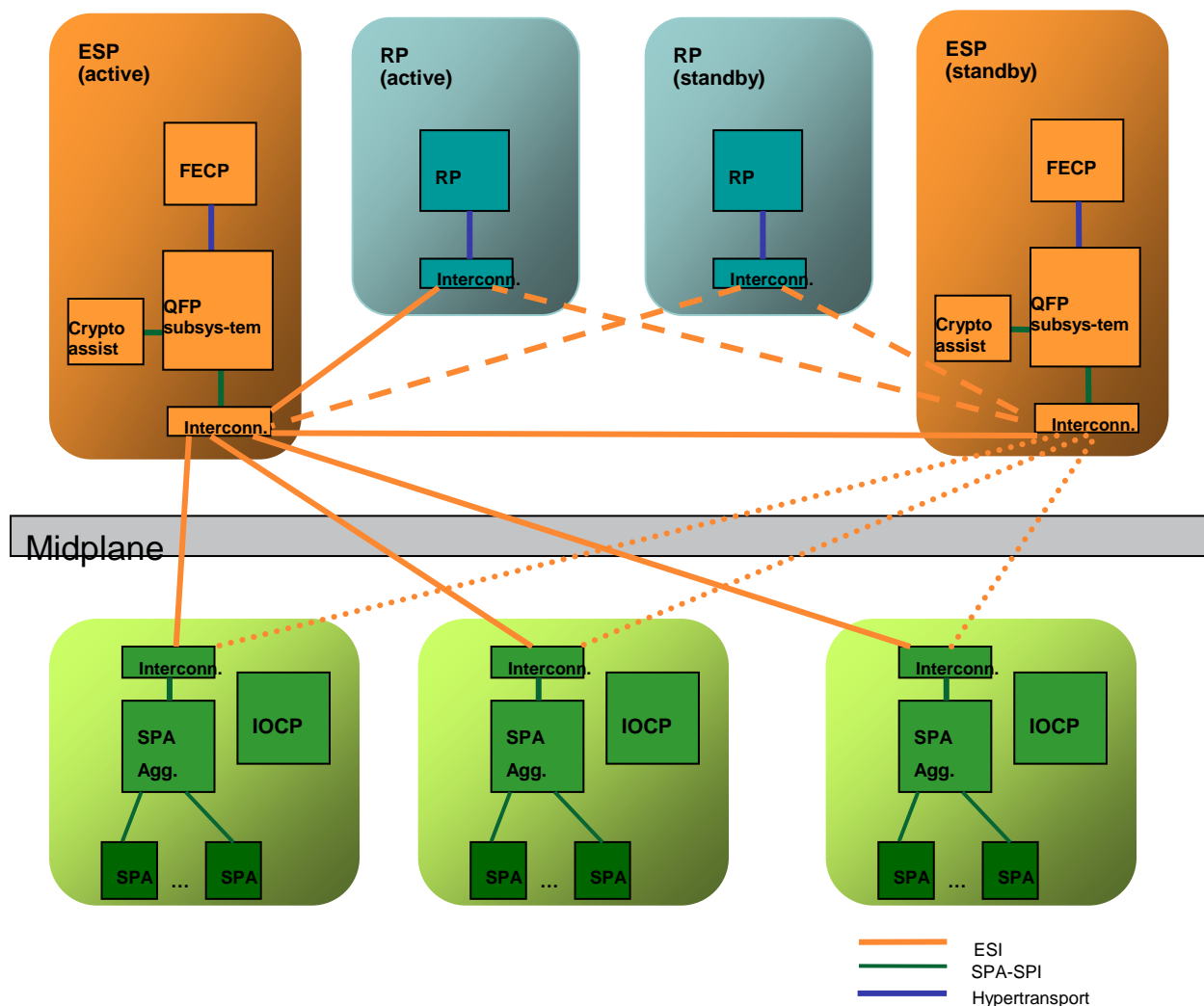
- Межузловые связи между другими картами (обмен управляющей информацией)
- Stratum-3 Clock для синхронизации (например SONET)
- Другие функции, например – управление электропитанием, обнаружение компонентов (карты) и т.д. и т.п.



| | RP1 | RP2 |
|----------------------------|------------------------|-------------------------------|
| ЦПУ | 1.5 GHz Freescale 8548 | Dual-Core Intel Xeon 2.66 GHz |
| Память | 2GB до 4GB | 8GB до 16GB |
| Ethernet порт управления | Да | Да |
| AUX порт управления | Да | Да |
| Консольный порт управления | Да | Да |
| HDD | 40Gb | 80Gb |
| Поддержка внеш. USB | Да | Да |

Dataplane

- Весь трафик обрабатывается на ESP
- Исключение: Punt path для Legacy протоколов – обрабатывается RP
- Interconnect ASIC в каждом функциональном элементе обеспечивает соединение с backplane посредством ESI
- ESI (Enhanced Serdes Interconnect) используются для пересылки трафика
- SPA соединяются с backplane через SPA-Agg ASIC и SPA-SPI линки



Линейка ESP

Июль 2010

| | ESP-5G | ESP-10G | ESP-20G | ESP-40G |
|------------------------|---------|---------|---------|---------|
| Пропускная способность | 5Gbps | 10Gbps | 20Gbps | 40Gbps |
| Базовый блок | QFP 20 | QFP 40 | QFP 40 | QFP 40 |
| Частота синхронизации | 900 Mhz | 900 Mhz | 1.2 Ghz | 1.2 Ghz |
| Шифрование | ~ 1Gbps | 3 Gbps | 8 Gbps | 10 Gbps |
| Объем памяти QFP | 256 MB | 512 MB | 1GB | 1GB |
| Объем буфера | 64 MB | 128 MB | 256 MB | 256 MB |
| Объем TCAM | 10 MB | 10 MB | 40 MB | 40MB |

«Quantum Flow Processor» - ядро ESP

▪ Packet Processing Engine (QFP-PPE)

40 PPE на одном кристалле

Каждый PPE:

- многопоточный (до 4 выделенных контекстов)
- программируется на C
- оперирует целым пакетом (header+payload)
- потребляет менее 400 мВт

▪ Buffer/Queue Subsystem (QFP-BQS)

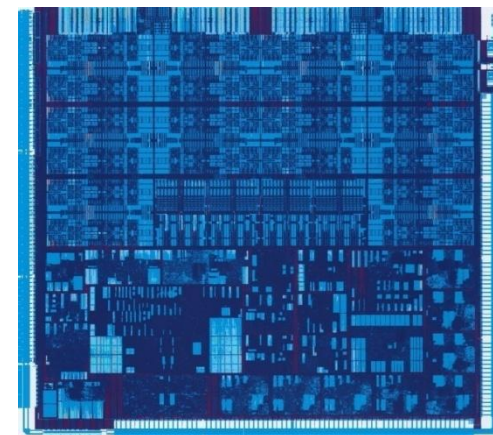
Гибкая обработка очередей разработана для более чем 5-уровневой иерархии планировщика:

Аппаратный иерархический планировщик

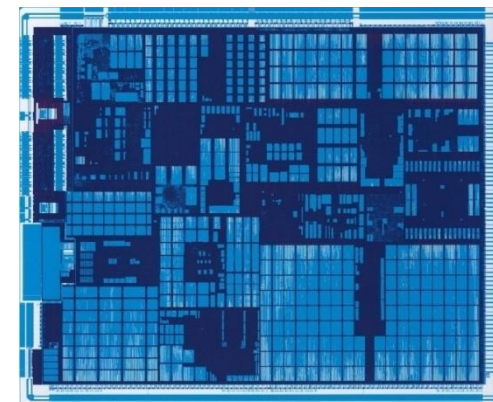
Более 100 тыс очередей с 3-мя параметрами

Полностью настраиваемые уровни на базе HQF

Приоритезация на нескольких уровнях

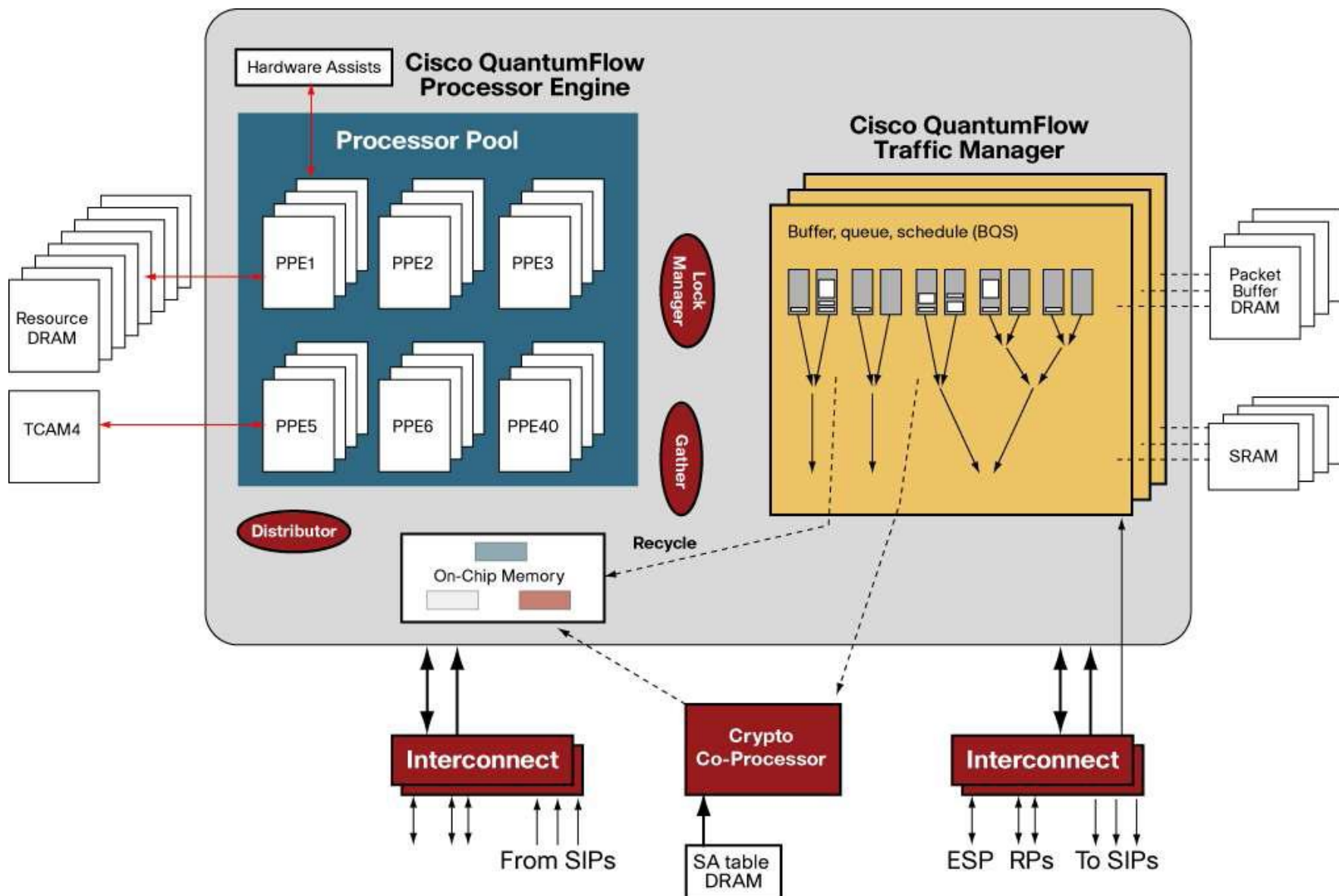


Multi-Core (40) Packet Processor

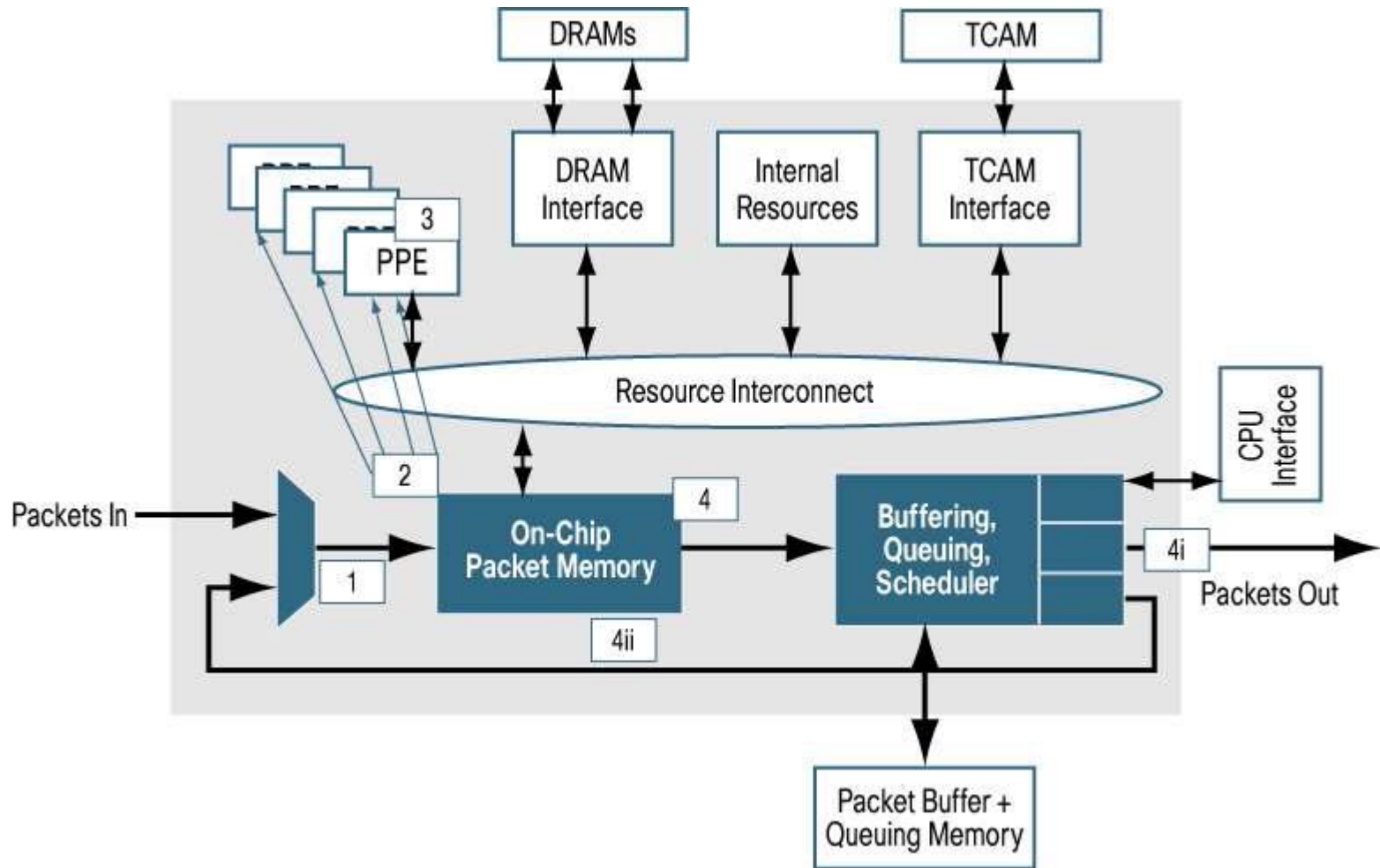


Traffic Manager (BQS)

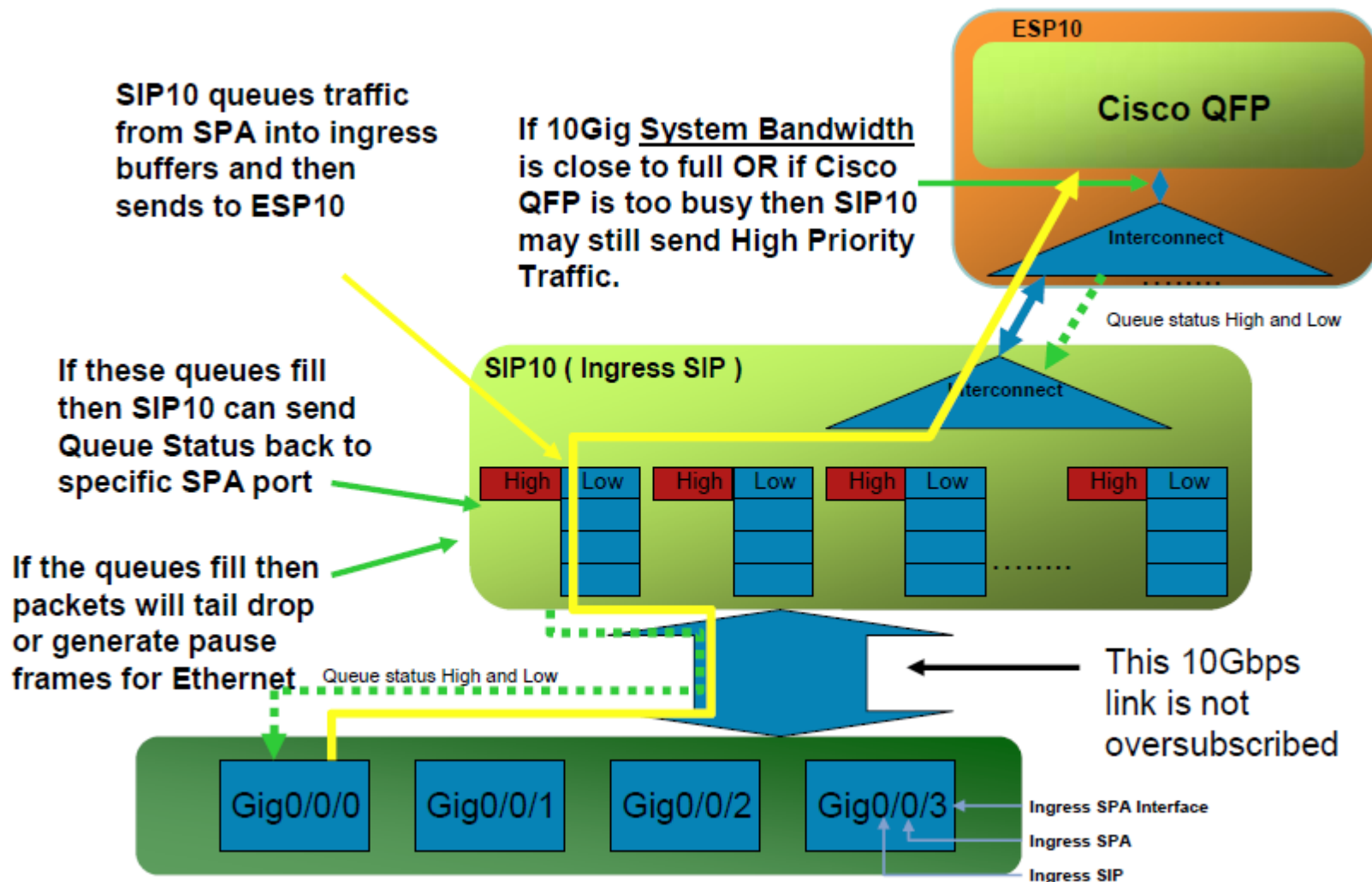
Схема организации QFP



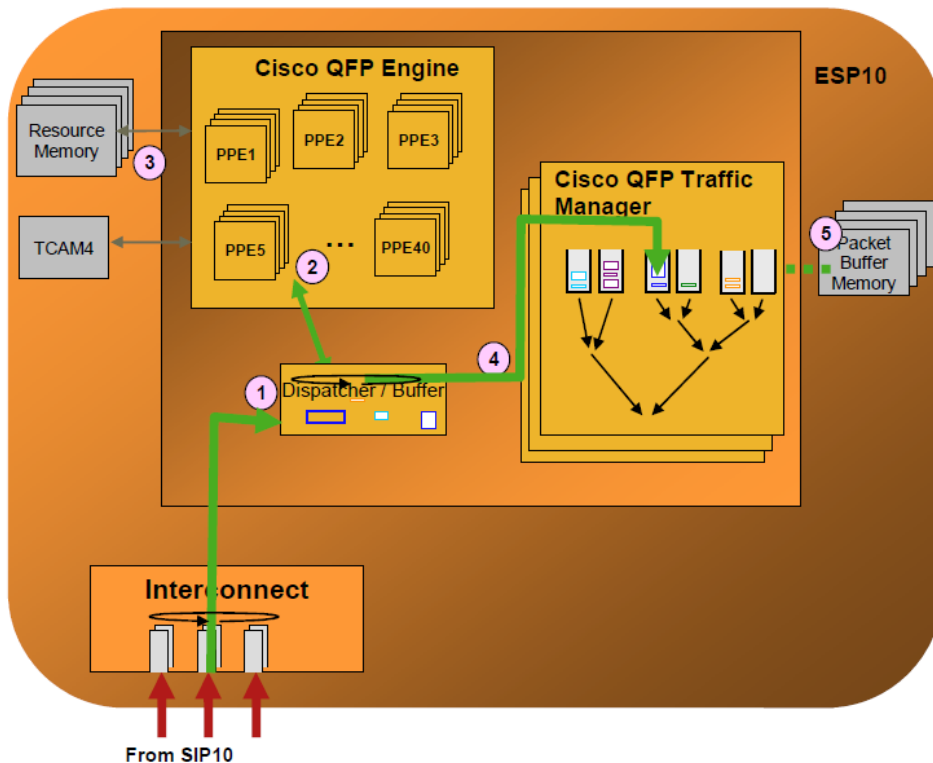
Packet flow



QoS внутри ASR1000

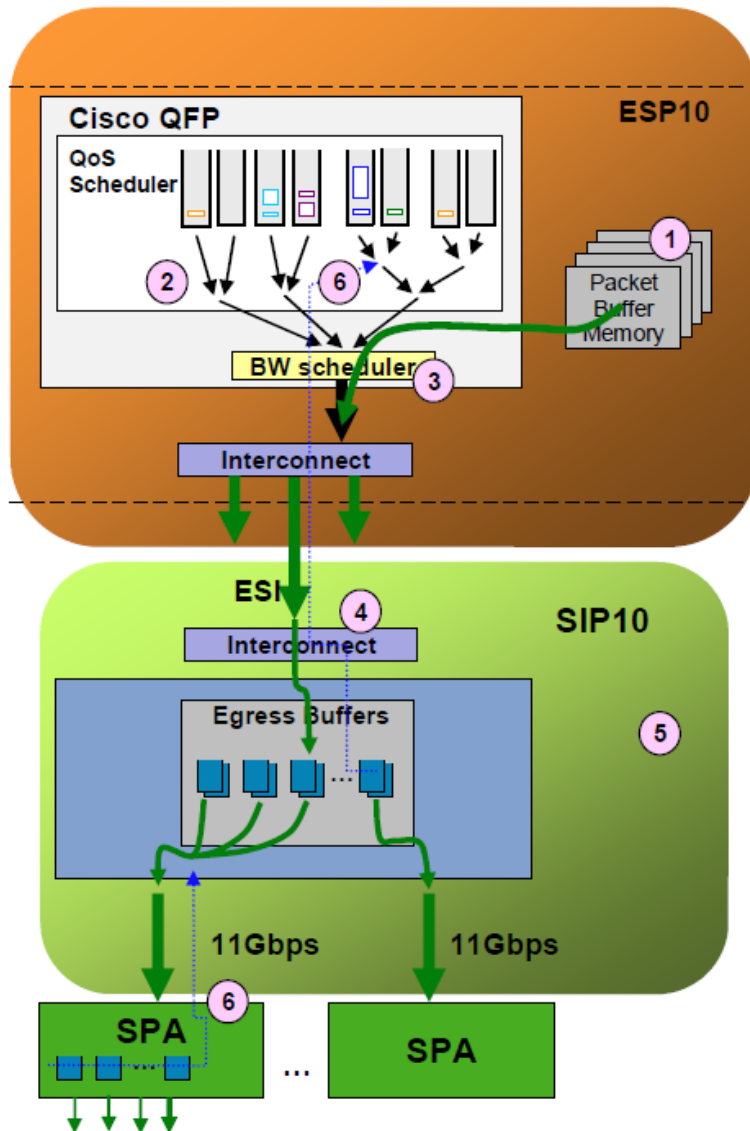


QoS внутри ASR1000 (2)



1. Входящие пакеты попадают в малый буфер в Cisco QFP
2. Назначается свободный процесс на Cisco QFP Engine и начинается обработка пакета: MAC classification, QoS classification, ACLs, forwarding lookup, WRED, cRTP и т.д.
3. Для lookup, обновления статистики, состояний и пр. имеется доступ к DRAM и TCAM.
4. Когда обработка завершена – запрос на размещение пакета в исходящей очереди
5. Содержимое пакета копируется из внутреннего буфера в большой буфер на выходе

QoS внутри ASR1000 (3)



1. Пакеты находятся в Packet Buffer Memory
2. Cisco QFP Traffic Manager Scheduler принимает решения и
 - a. Динамически создает уровни иерархии из MQC.
 - b. Может задавать minimum (bandwidth), excess (bandwidth remaining), и maximum (shape) параметры для каждой очереди.
 - c. High-priority обгоняют lower-priority пакеты.
3. Затем Bandwidth scheduler назначает Cisco QFP's output bandwidth among the SIP10s
 - a. Следит, чтобы highest priority пакеты среди разных SIP10 не были подвержены backpressure
 - b. High priority пакеты всегда имеют приоритет
4. Пакеты передаются на SIP10 simultaneously
5. Мелкие буферы на SIP10 и SPA используются для передачи пакетов сразу в несколько портов
6. Backpressure на QFP

Q AND A



