



# A Deep Dive on the **QorIQ T2080** Processor

FTF-NET-F0032

Chun Chang | Application Engineer

April 2014



External Use



# Agenda

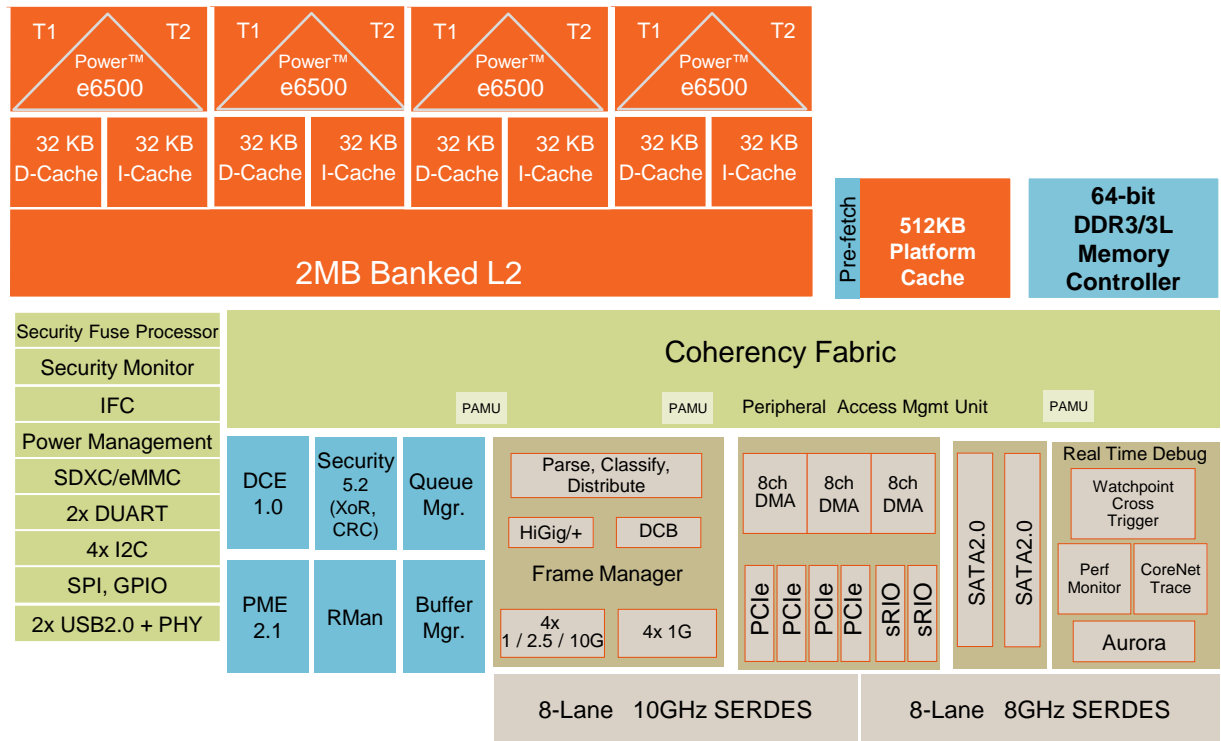
- T2080/1 Overview
- e6500 Core and Cache Hierarchy
- Power Management
- CoreNet Coherency Fabric Switch
- Data Path Acceleration Architecture (DPAA)
- SerDes Options
- Voltage ID
- eSDHC
- PCI Express
- Enablement
- Conclusion

# Agenda

- T2080/1 Overview
  - T2080 Block Diagram
  - T2081 Block Diagram
  - Duad Cores comparison



# QorIQ T2080 Block Diagram



## Datapath Acceleration

- **SEC**- crypto acceleration 10Gbps
- **DCE** - Data Compression Engine 17.5Gbps
- **PME** – Pattern Matching Engine to 10Gbps

## Processor

- 4x e6500, 64b, 1.2 - 1.8GHz
- Dual threaded, with 128b Altivec
- 2MB shared L2; 256KB per thread

## Memory Subsystem

- 512KB Platform Cache w/ECC
- 1x DDR3/3L Controllers up to 2.1GHz
- Up to 1TB addressability (40 bit physical addressing)
- HW Data Pre-fetching

## Switch Fabric

### High Speed Serial IO

- 4 PCIe Controllers: one at Gen3, three at Gen2
  - 1 with SR-IOV support
  - x8 Gen2
- 2 sRIO Controller
  - Type 9 and 11 messaging
  - Interworking to DPAA via RMan
- 2 SATA 2.0 3Gb/s
- 2 USB 2.0 with PHY

## Network IO

- Up to 25Gbps Simple PCD each direction
  - 4x1/10GE, 4x1GE or 2.5Gb/s SGMII
  - XFI, 10GBase-KR, XAUI, HiGig, HiGig+, SGMII, RGMII, 1000Base-KX

## Device

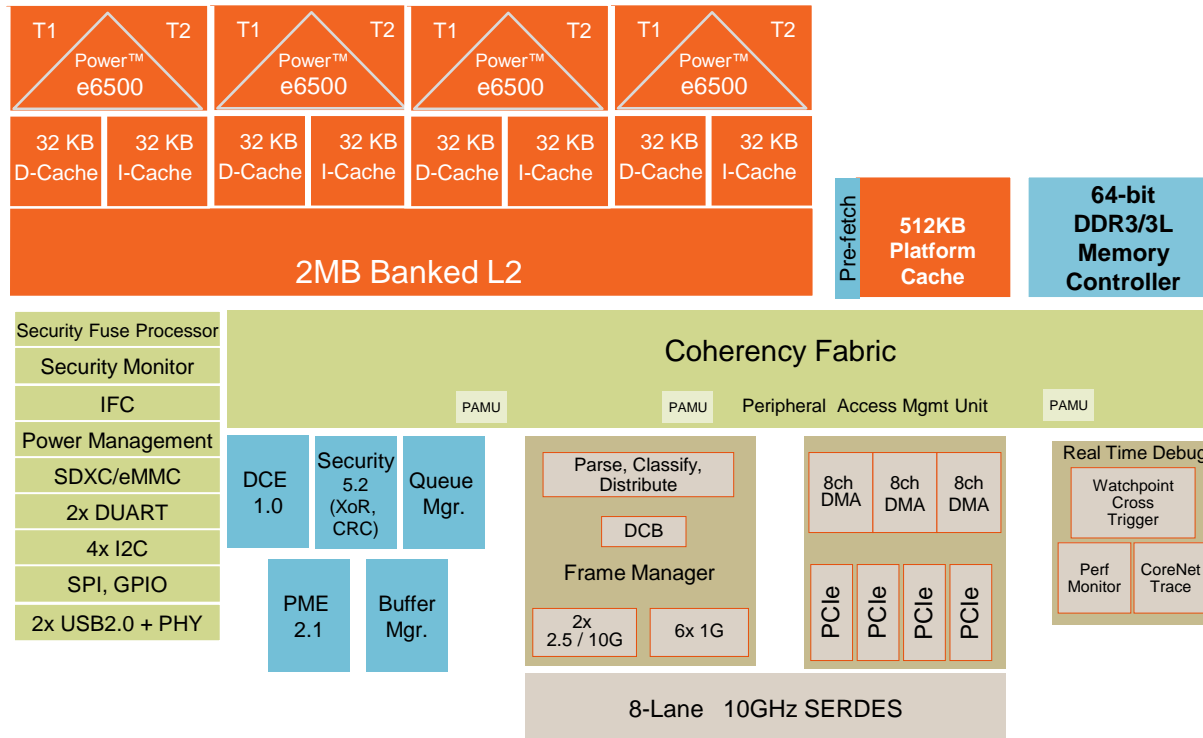
- TSMC 28HPM Process
- 25x25mm, 896 pins, 0.8mm pitch
- Power estimated at 15.2 – 25.2W (thermal) depending on frequency

**Schedule:** Q3-2013 (alpha); mid-2014 qual





# QorIQ T2081 Block Diagram



## Datapath Acceleration

- **SEC**- crypto acceleration 10Gbps
- **DCE** - Data Compression Engine 17.5Gbps
- **PME** – Pattern Matching Engine to 10Gbps

## Processor

- 4x e6500, 64b, 1.5 - 1.8GHz
- Dual threaded, with 128b Altivec
- 2MB shared L2; 256KB per thread

## Memory Subsystem

- 512KB Platform Cache w/ECC
- 1x DDR3/3L Controllers up to 2.1GHz
- Up to 1TB addressability (40 bit physical addressing)
- HW Data Pre-fetching

## Switch Fabric

## High Speed Serial IO

- 4 PCIe Controllers: one at Gen3, three at Gen2
  - 1 with SR-IOV support
  - x8 Gen2
- 2 USB 2.0 with PHY

## Network IO

- Up to 25Gbps Simple PCD each direction
- 8 MACs multiplexed over:
  - 2x 10GE, 2x 2.5Gb/s SGMII, 7x GE
  - XFI, 10GBase-KR, SGMII, RGMII, 1000Base-KX

## Device

- TSMC 28HPM Process
- 23x23mm, 780pins, 0.8mm pitch, pin compatible with T1042
- Power estimated at 18.7– 24.4W (thermal) depending on frequency

**Schedule:** samples: 2H-2014; qual Q1-15





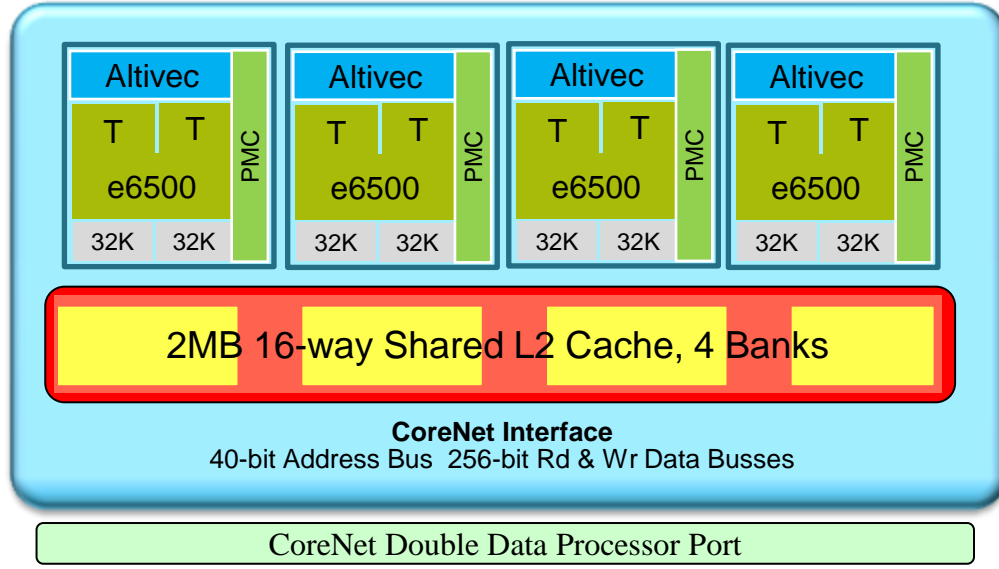
# Quad Cores Compared

|              | P2040                  | P2041                  | P3041                  | T1042                  | T2081                  | T2080                  |
|--------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| Cores        | 4x e500mc, 32b         | 4x e500mc, 32b         | 4x e500mc, 32b         | 4x e5500, 64b          | 4x e6500, 64b          | 4x e6500, 64b          |
| Threads      | 4                      | 4                      | 4                      | 4                      | 8                      | 8                      |
| Frequency    | 667MHz – 1.2GHz        | 1.2 - 1.5GHz           | 1.2 - 1.5GHz           | 1.2 - 1.4GHz           | 1.5 - 1.8GHz           | 1.2 - 1.8GHz           |
| L2           | None                   | 512kB                  | 512kB                  | 1MB                    | 2MB                    | 2MB                    |
| L3           | 1MB                    | 1MB                    | 1MB                    | 256kB                  | 512kB                  | 512kB                  |
| DDR          | 1x DDR3/3L to 1200MT/s | 1x DDR3/3L to 1333MT/s | 1x DDR3/3L to 1333MT/s | 1x DDR3L/4 to 1333MT/s | 1x DDR3/3L to 2133MT/s | 1x DDR3/3L to 2133MT/s |
| SerDes       | 10 to 5GHz             | 10 to 5GHz             | 18 to 5GHz             | 8 to 5GHz              | 8 to 10GHz             | 16 to 10GHz            |
| Enet         | 5x 1GE                 | 10GE + 5x 1GE          | 10GE + 5x 1GE          | 5x 1GE                 | 2x 1/10GE + 5x 1GE     | 4x 1/10GE + 4x 1GE     |
| PCIe Cntrls  | 3 at Gen2              | 3 at Gen2              | 3 at Gen2              | 4 at Gen2              | 3 at Gen2 + 1 at Gen3  | 3 at Gen2 + 1 at Gen3  |
| SATA2.0      | 2                      | 2                      | 2                      | 2                      | No                     | 2                      |
| USB2.0       | 2 w/ int. PHY          | 2 w/ int. PHY          | 2 w/ int. PHY          | 2 w/ int. PHY          | 2 w/ int. PHY          | 2 w/ int. PHY          |
| SRIO/Rman    | 2                      | 2                      | 2                      | No                     | No                     | 2                      |
| Aurora       | Yes                    | Yes                    | Yes                    | Yes                    | No                     | Yes                    |
| TDM/HDLC     | No                     | No                     | No                     | 2                      | No                     | No                     |
| Acceleration | SEC, PME               | SEC, PME               | SEC, PME               | SEC, PME, QE           | SEC, PME, DCE          | SEC, PME, DCE          |

# Agenda

- T2080/1 Overview
- e6500 Core and Cache Hierarchy
  - e6500 Core Diagram
  - e6500 Pipeline
  - Additional e6500 Enhancements
  - Multi-threading Implementation
  - Load-Store / L1 Data Cache
  - Shared L2 Cache
  - Platform L3 Cache

# e6500 Core Complex



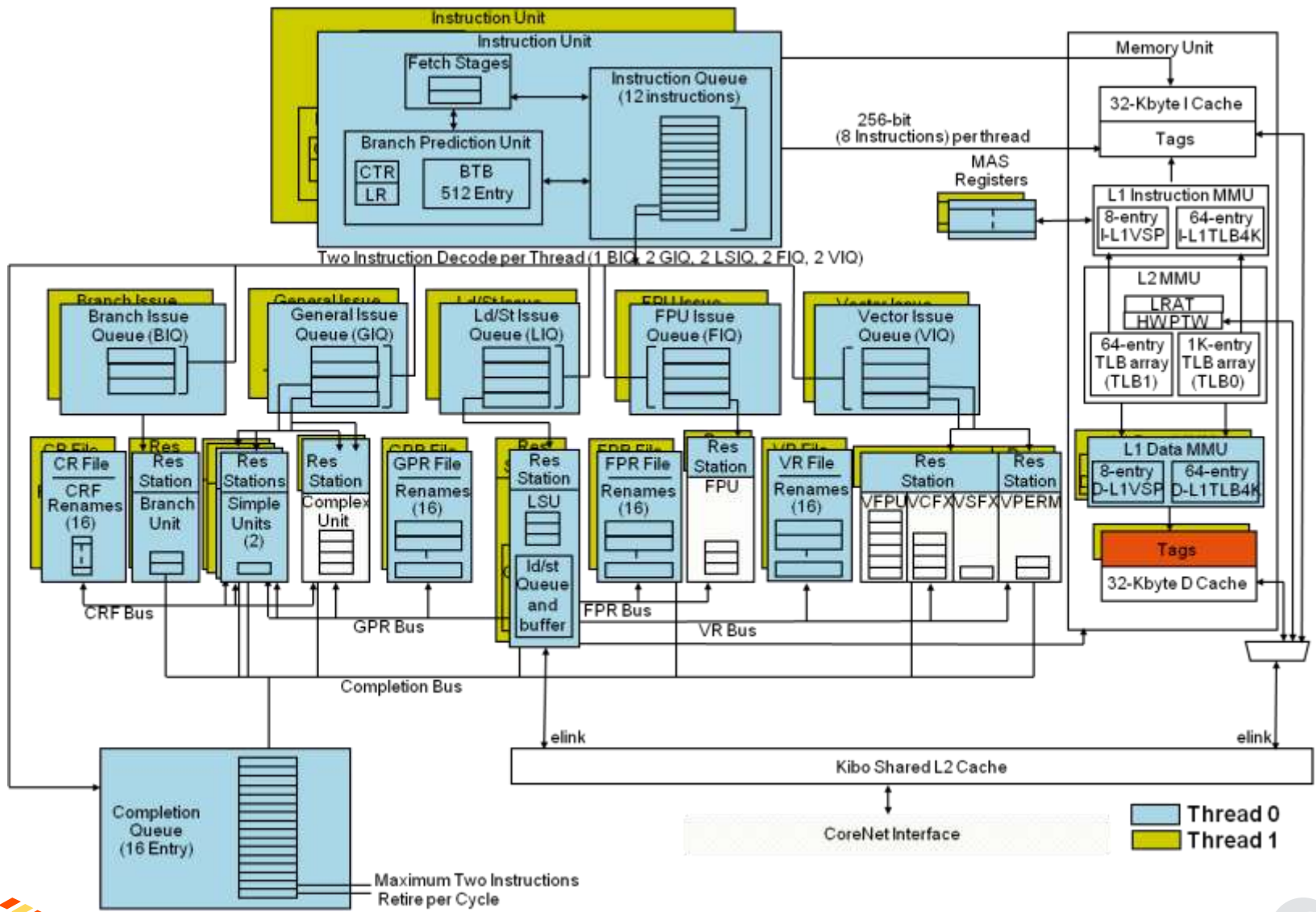
- 64-bit Power Architecture
- Up to 1.8 GHz operation
- Two threads per core
- Dual load/store units, one per thread
- 40-bit Real Address
  - 1 Terabyte physical address space
- Hardware Table Walk
- L2 in cluster of 4 cores
  - Supports Share across Cluster
  - Supports L2 memory allocation to core or thread
- Power Management
  - Drowsy: Core, Cluster, AltiVec
  - Wait-on-reservation instruction
  - Traditional modes
- AltiVec SIMD Unit (128b)
  - 8,16,32-bit signed/unsigned integer
  - 32-bit floating-point
  - 192 GFLOP (2GHz)
  - 8,16,32-bit Boolean
- Virtualization
  - Hypervisor
  - LRAT
    - Logical to Real Address translation mechanism for improved hypervisor performance

|               | <b>P3041</b><br>2.5 DMIPS<br>(1.5GHz) | <b>T2080</b><br>5.4 DMIPS<br>(1.8GHz) | <b>Improvement</b><br>from P3041 |
|---------------|---------------------------------------|---------------------------------------|----------------------------------|
| Single Thread | 3750                                  | 5940                                  | 1.6x                             |
| Core (dual T) | 3750                                  | 9720                                  | 2.6x                             |
| SoC           | 15,000                                | 38,880                                | 2.6x                             |





# e6500 Pipeline



# Additional e6500 Enhancements

- Faster FPU: 2X faster SP, 4X faster DP over e500mc
- New Power ISA v.2.06 Instr
  - instructions for byte- and bit-level acceleration: Parity, population count, bit permute, compare bytes , FPU convert to/from 64-bit integer
- Improved Branch Prediction
  - Double BTB size
  - Better branch prediction scheme (rate increases from 95% to 98%)
- Increase number of completion entries and rename registers from 14 to 16
- Re-architected the memory subsystem
  - Shared L2 cache with write-through L1 D cache and large store gather buffer per core
  - 2X L2 cache size per core, effectively more with sharing
- 40-bit real address
- PID0 field size increases from 8 to 14 bits => supports for more threads in many core systems
- Enhanced MP Performance: Accelerated Atomic Operations, Optimized Barrier Instructions, Fast intra-cluster sharing
- LRAT: Accelerate hypervisor performance (10-15% for workloads running in OS on HV)
- New power-reduction techniques
- Drowsy core with fast wake-up (<75% power of run mode)
- Option for AltiVec
- Changes for debug architecture

# Multi-threading Implementation

- Interrupts
  - Interrupts are private
  - Each thread has its own interrupt signals
- Debug
  - Almost all resources are private. Internal debug works as if they are separate cores
  - External debug has option to halt both threads when one thread debug halts
- Power Management
  - Power management control is per-thread (and the associated SoC programming model will be per-thread)
  - Actual power management will only occur when both threads reach the same power management state
  - For example, when **wait** occurs on one thread, fetching stops for that thread, but we don't go drowsy until both threads execute **wait**.

# e6500 Load-Store / L1 Data Cache

- Dual Load Store Units (LSU)
  - Each LSU is dedicated to a thread
  - Separate Data MMUs and Tags
  - Shared Data Cache
- L1 Data Cache Organization
  - 32 KB
  - 8-way set associative with PLRU replacement algorithm
- Features
  - Store Gather Buffer to optimize store bandwidth
  - Store to load forwarding to reduce stalls
  - Individual line locking with persistent locks
  - Accelerated atomic operations
  - Optimized cacheable barrier instructions

# e6500 Shared L2 Cache

- L2 Cache Organization
  - 2 MB
  - 4 banks of 512 KB each
  - 16-way set associative with configurable replacement algorithms
- L2 Cache Features
  - Individual line locking with persistent locks
  - Flexible way partitioning by thread
    - Allocation control for data read, data store, instruction read and stash
  - ECC protection for Data, Tag and Status

# Platform L3 Cache

- Platform L3 Cache Organization
  - 512 KB
  - 16-way set associative with configurable replacement algorithms
- Platform L3 Cache Features
  - Individual line locking with persistent locks
  - Flexible way partitioning by source
    - Allocation control for data read, data store, castout, decorated read, decorated store, instruction read and stash
  - Configurable SRAM partitioning
  - ECC protection for Data, Tag and Status

# Agenda

- T2080/1 Overview
- e6500 Core and Cache Hierarchy
- Power Management
  - Power Management Innovation
  - Core Power Management States
  - Cluster Power Management States
  - SOC Power Management States



# e6500 Power Management Innovation

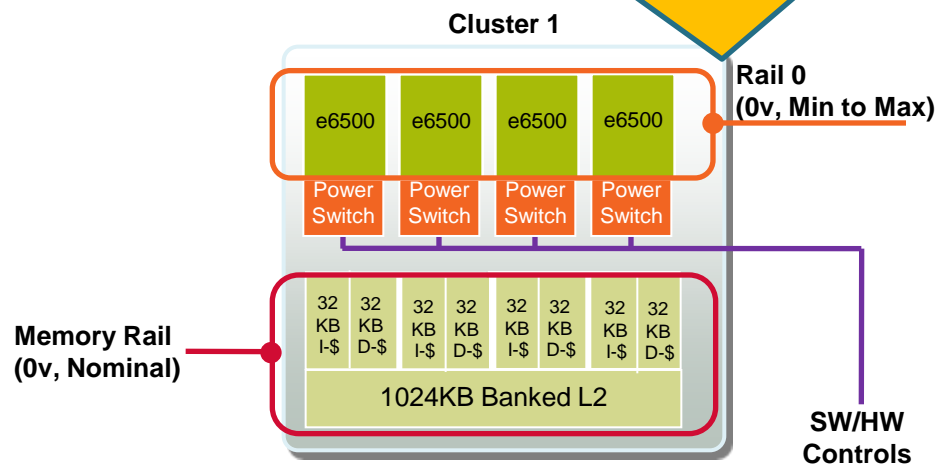
- **Wide voltage range for logic supplies to allow frequency / power tradeoff**
  - Memory arrays on a separate power supply
- **Power domain hierarchy**
  - AltiVec within core
  - Cores within cluster
  - Clusters within SoC
- **Drowsy L2 Cache**
  - Bitcell leakage reduced by ~40%
- **Drowsy Core**
  - Instantaneous wakeup response with SRPG
  - Controlled through software or waterfall power management
  - Power <75% of Run-mode
- **Deep Nap Mode**
  - State not retained
  - Power < 90% of Run-mode

**Focus**

- Reduce energy consumption under light loads
- Enable rapid return to fully loaded conditions

- Do not have to save/restore processor state to memory
- Greater than 10x improvement in wakeup response time

- Switch supports 3 modes
  - Full On
  - Drowsy Mode
  - Deep Nap Mode (Powered Off)







# e6500 Core Power Management States

| Power State | Initiated        | Description   |
|-------------|------------------|---|
| PH00        | Default          | Full-On. Global clocks running. Local clock gating based on unit usage and Dynamic Power Management (DPM)   |
| PH10        | SOC RCPM         | Previously Doze. Global clocks running and instruction fetch is stopped. Snoops still handled   |
| PH15        | SOC RCPM         | Previously Nap. Core global clocks stopped. Software must flush and invalidate caches before state entry and handle any MMU coherency issues  |
| PH20        | SOC RCPM         | New State. Core PH20 mode is core power gating with state retention   |
| PH30        | SOC RCPM         | New State. Core PH30 mode is core power gating without state retention. Interrupt is ignored. Return to PH00 requires a core reset.   |
| PW10        | Wait Instruction | Previously Wait. Global clocks running and instruction fetch is stopped   |
| PW20        | Wait Instruction | New State. Core global clocks stopped, power supply gated and state retained. Transition from PW10 to PW20 occurs completely under hardware control with no software intervention. Fast wake up based on hardware events. |



# e6500 Cluster Power Management States

| Power State | Initiated | Description  |
|-------------|-----------|--|
| PCL00       | Default   | Full-On. Global clocks running. Local clock gating based on unit usage and Dynamic Power Management (DPM)  |
| PCL10       | SOC RCPM  | Clock distribution is inhibited to cluster functional unit, Clock distribution is inhibited to cluster functional unit. The L2 cache no longer continues to participate in snooping activities. Software should always flush, and then invalidate the L2 cache prior to initiating PCL10 state to ensure that any modified data is written out to backing store. |





# T2080 SOC Power Management States

| Power State | Initiated | Description   |
|-------------|-----------|---|
| PLM10       | SOC RCPM  | LPM10 mode is a device state which at least one core is not in PH00 (full on state).  |
| PLM20       | SOC RCPM  | <p>All cores are in PH20 state.</p> <ul style="list-style-type: none"><li>– Cluster is in PCL10 state.</li><li>– Platform clock is disabled.</li></ul> <p>All clocks internal to the core are turned off as well as the clock in device logic so that only the modules which are required to wake up the device will still have a running clock. Core timebase is turned-off.</p> <p>The modules which can be used as a wake up source are internal timers, internal and external interrupts. After the core and I/O interfaces have shut down, ASLEEP pin is asserted.</p> |

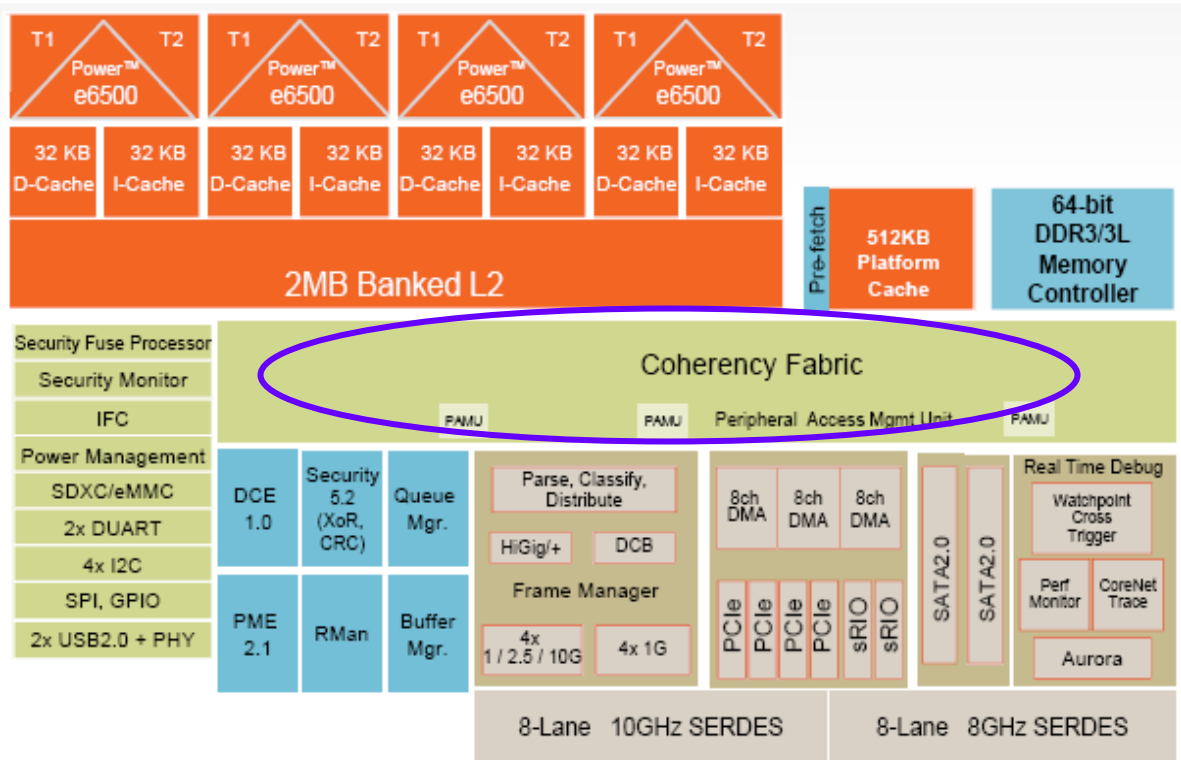


# Agenda

- T2080/1 Overview
- e6500 Core and Cache Hierarchy
- Power Management
- CoreNet Coherency Fabric Switch
  - T2080 I/O at a Glance
  - CoreNet System Bandwidth
  - Enhancements in Platform



# CoreNet Coherency Fabric Switch



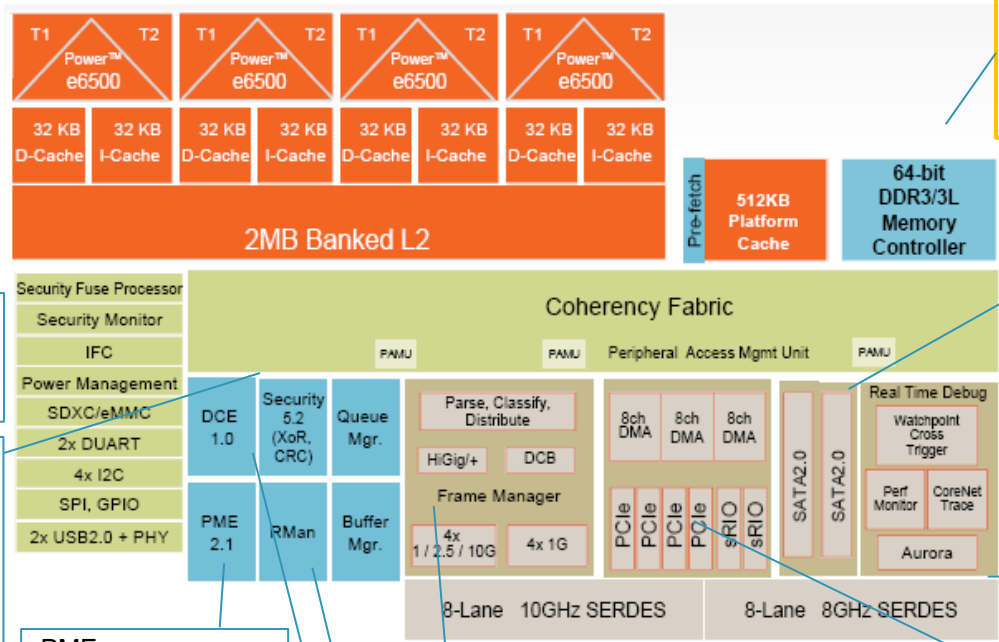
## Highly concurrent, 100% HW cache coherent, multi-ported fabric

- Overcomes limitations of bus based topologies
- **Completely eliminates retries** for busy conditions or cache coherency actions
  - Variable snoop response timing
  - Current owner always supplies data
  - Minimizes average latency in congested systems
- **Flexible point to point connectivity**
  - Point-to-point connectivity with flexible protocol architecture allows for pipelined interconnection between CPUs, platform caches, memory controllers, and I/O and accelerators at up to 700 MHz
- **Supports multiple, parallel address paths**
  - High address bandwidth: Key for large coherent multi-core processors
- **High data bandwidth**
  - Crossbar connectivity: reduced contention provides low latency
  - Variable width data path per device provides throughput and power optimization
  - Capable of sustaining multiple cache lines per cycle to the cores
- Supports future expansion to coherent multi-fabric 'clusters' on SoC's or coherent multi-chip systems





# T2080 I/O at a Glance



**DDR 3/3L**

- Up to 2.1 GHz
- Works with new Prefetch block

**SATA, USB, SD/MMC**

- Lower system cost

**Advanced Power Management**

- Power planes
- VID
- Cascade loading

**SEC 5.2**

- Enhanced RSA Operation for SSL
  - 340b RSA 25K operations
  - 680b RSA 7.8K operations
- 10Gbps DES
- 10Gbps AES
- 10G CRC

**PME 10G Regex Engine**

**DCE - Compression**

- Deflate, GZIP, Zlib
- 20Gbps

**RMAN SRIO in DPAA**

**Networking**

- 48 Gbps bandwidth
- 2x HiGig MAC
- IEEE 802.3az (EEE)
- IEEE 802.3bf (Time sync)
- Data Center Bridging (DCB)
  - Priority Flow Control
  - Enhanced Transmission Selection

**SERDES**

- Up to 10GHz
- XFI, 10Gbase-KR, XAUI, SGMII, PCIe, SRIO

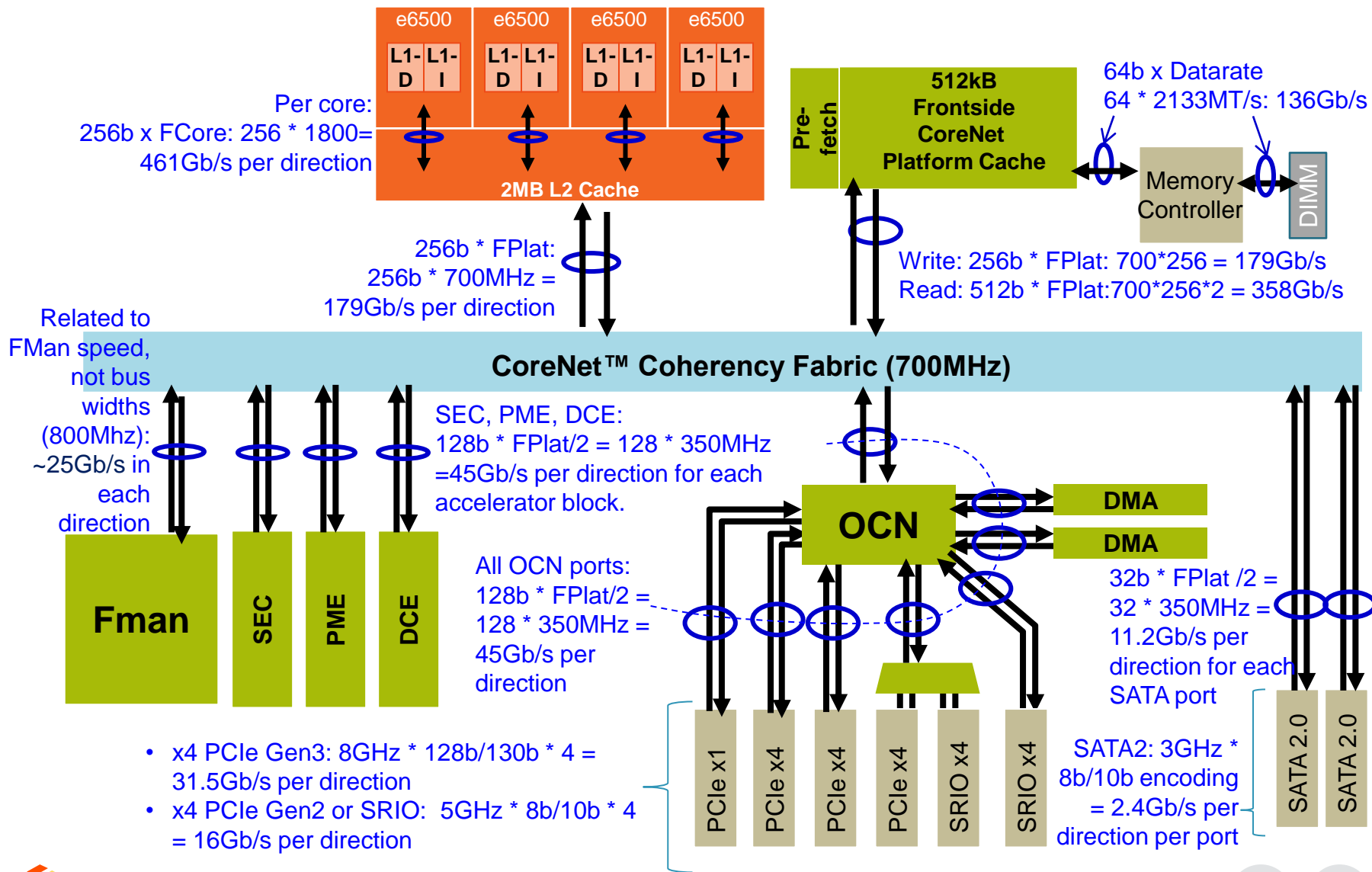
**PCIe Gen 3**

- SRIOV EP





# T2080 CoreNet System Bandwidth



## Enhancements in Platform V2 (T-Series)

- CoreNet Coherency Fabric
  - 40-bit Real Address
  - Higher address bandwidth, Larger number of active transactions
  - 2X BW increase for: Core data ports, Memory subsystem writes, many peripheral devices
  - Improved configuration architecture
  - “Safe” mode for coherency-error tolerance during multi-core software development
- Platform Cache
  - Increased Write Bandwidth
  - Increased buffering for improving throughput
  - Improved data ownership tracking for performance enhancement
- Data PreFetch
  - Tracks CPC misses
  - Prefetches from multiple memory regions with configurable sizes
  - Selective tracking based on Requesting device, Transaction type, data/instruction access
  - Conservative prefetch requests to avoid system overloading with prefetches
  - “Confidence” based algorithm with feedback mechanism
  - Performance monitor events to evaluate the performance of Prefetch in the system

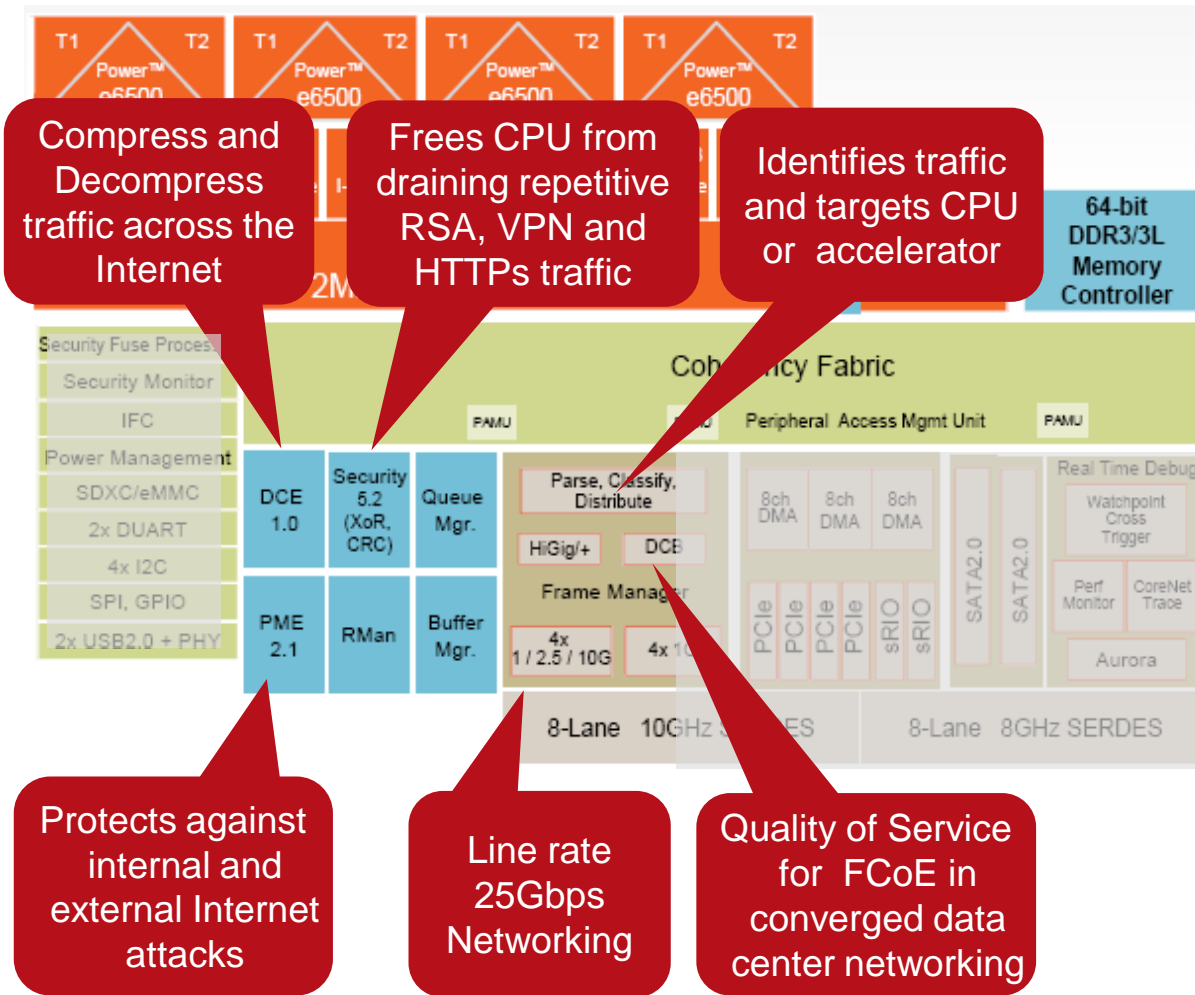


# Agenda

- T2080/1 Overview
- e6500 Core and Cache Hierarchy
- Power Management
- CoreNet Coherency Fabric Switch
- Data Path Acceleration Architecture (DPAA)
  - FMAN
  - QMAN
  - BMAN
  - RMAN
  - SEC
  - PME
  - DCE



# Enhancing Core Performance with Data Path Acceleration Architecture



| Hardware Accelerators                   |  |
|---|--|
| FMAN<br>Frame Manager                   | 25 Gbps aggregate Parse, Classify, Distribute    |
| BMAN<br>Buffer Manager                  | 64 buffer pools                                  |
| QMAN<br>Queue Manager                   | Up to 2 <sup>24</sup> queues                     |
| RMAN<br>Rapid IO Manager                | Seamless mapping sRIO to DPAA                    |
| SEC<br>Security                         | 10Gbps: IPsec, SSL<br>Public Key 25K/s 1024b RSA |
| PME<br>Pattern Matching                 | 10Gbps aggregate                                 |
| DCE<br>Data Compression                 | 20Gbps aggregate                                 |
| Saving CPU Cycles for higher value work |  |

**New** Enhanced





# DPAA Components Check List

- QorIQ P-class devices have Datapath Three-Speed Ethernet Controller (dTSEC) and 10-Gigabit Ethernet Media Access Controller (10GEC)
- QorIQ T-class devices have Ethernet Media Access Controller (EMAC)

| QorIQ Devices<br>DPAA Feature List | Revision Number |      |      |     |     |      |     |     |
|------------------------------------|-----------------|------|------|-----|-----|------|-----|-----|
|                                    | FMan            | QMan | BMan | SEC | PME | RMan | DCE | RE  |
| <b>P1023</b>                       | 4.0             | 2.0  | 2.0  | 4.2 | n/a | n/a  | n/a | n/a |
| <b>P4080/P4040 rev3</b>            | 2.0             | 1.1  | 1.0  | 4.0 | 2.1 | n/a  | n/a | n/a |
| <b>P2040, P2041<br/>P3041</b>      | 3.0             | 1.2  | 1.0  | 4.2 | 2.1 | 1.0  | n/a | n/a |
| <b>P5020, P5010</b>                | 3.0             | 1.2  | 1.0  | 4.2 | 2.1 | 1.0  | n/a | 1.0 |
| <b>T2080, T2081</b>                | 6.1             | 3.1  | 2.1  | 5.2 | 2.1 | 1.0  | 1.0 | n/a |



# DPAA New Features for T2080

A short summary of T2080 enhancements over the first generation DPAA (as implemented in the P3041) is provided below:

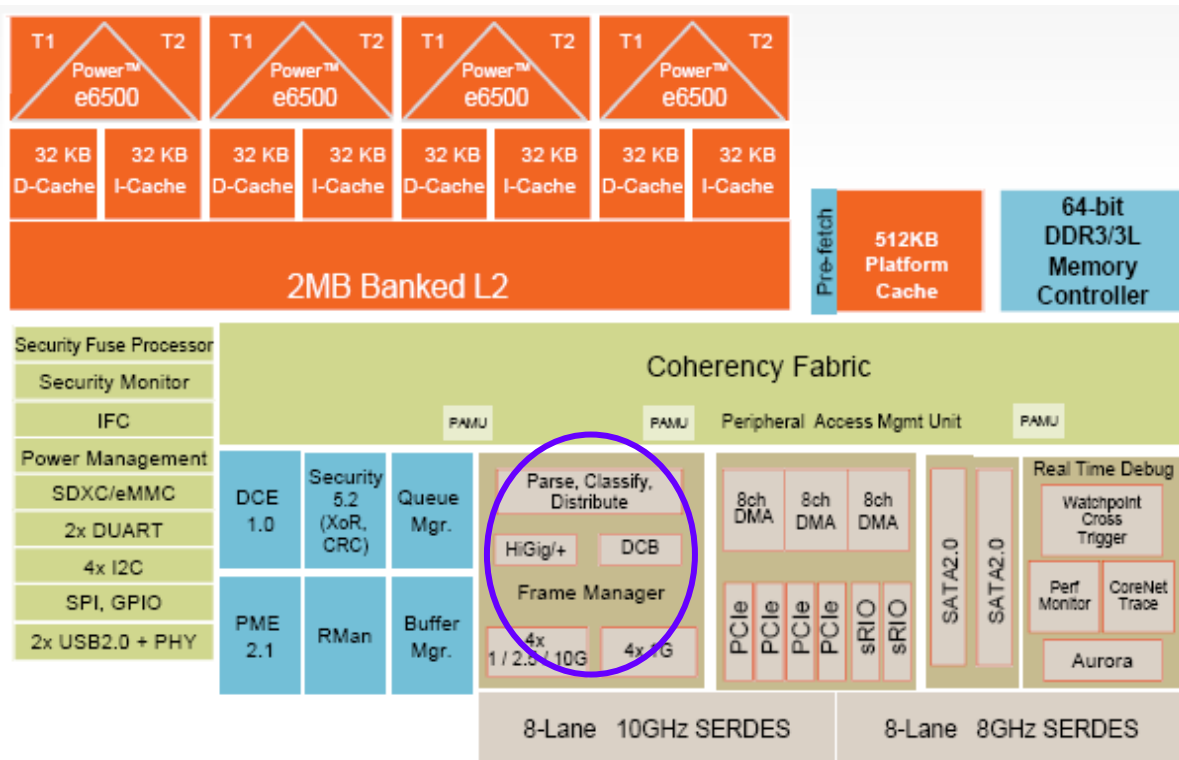
- Frame Manager
  - 2x performance increase (up to 25 Gbps per FMan)
  - Storage profiles
  - HiGig
  - Energy efficient Ethernet
- SEC 5.2
  - 2x performance increase for symmetric encryption and protocol processing
  - Up to 10 Gbps for IPsec @ Imix
  - 10x performance increase for public key algorithms
  - Support for 3GPP Confidentiality and Integrity Algorithms 128-EEA3 & 128-EIA3 (ZUK)
- DCE 1.0, new accelerator for compression/decompression
- RMan (Serial RapidIO Manager)
  - Included in P2/P3/P5/T2/T4 products
- DPAA overall capabilities
- Data Center Bridging
- Egress Traffic Shaping

## New Features for T2080 (continue...)

- T208x has a total of 50 software portals (SP), increase from 10 SP found in the P-class processors
- Supports Customer Edge Egress Traffic Management (CEETM) that provides hierarchical class based scheduling and traffic shaping:
  - Available as an alternate to FQ/WQ scheduling mode on the egress side of specific direct connect portals
  - Enhanced class-based scheduling supporting 16 class queues per channel
  - Token bucket based dual rate shaping representing Committed Rate (CR) and Excess Rate (ER)
  - Congestion avoidance mechanism equivalent to that provided by FQ congestion groups
- A total of 48 algorithmic sequencers are provided, allowing multiple enqueue/dequeue operations to execute simultaneously
- Support up to 295M enqueue/dequeue operations per second



# Frame Manager

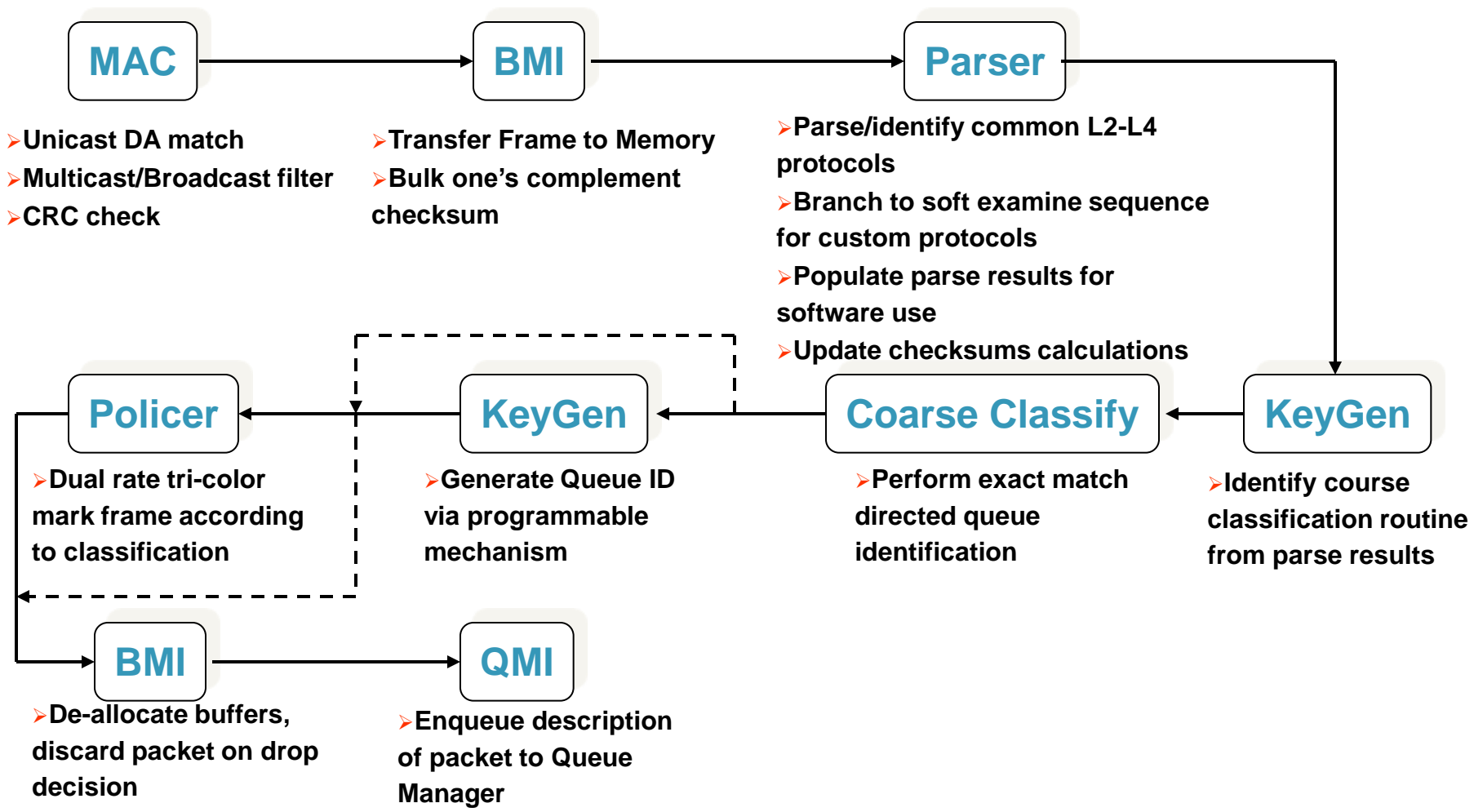


Frame Manager is responsible for moving packets into and out of the datapath

- 8 Ethernet MACs
  - 8 x 1GE
  - 4 x 2.5GE
  - 4 x 10GE
- Parse
- “Coarse” classification
- Packet distribution across queues for load spreading
- Policing

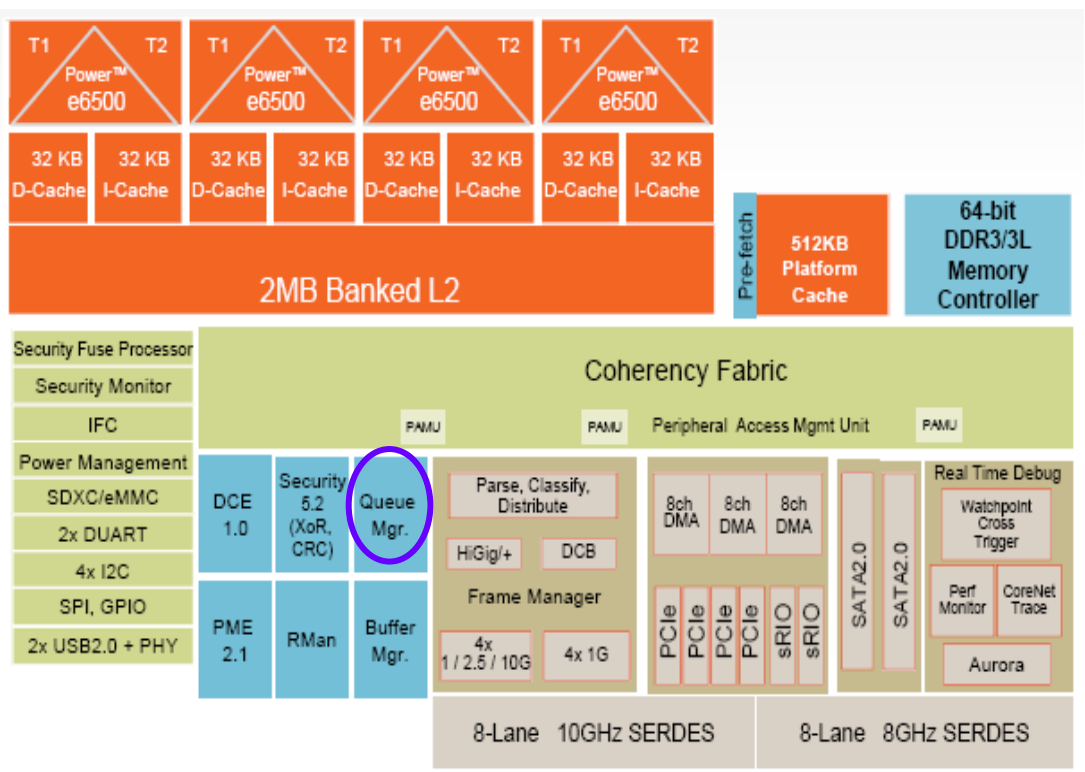


# Frame Manager Flow Chart





# Datapath Infrastructure: Queue Manager

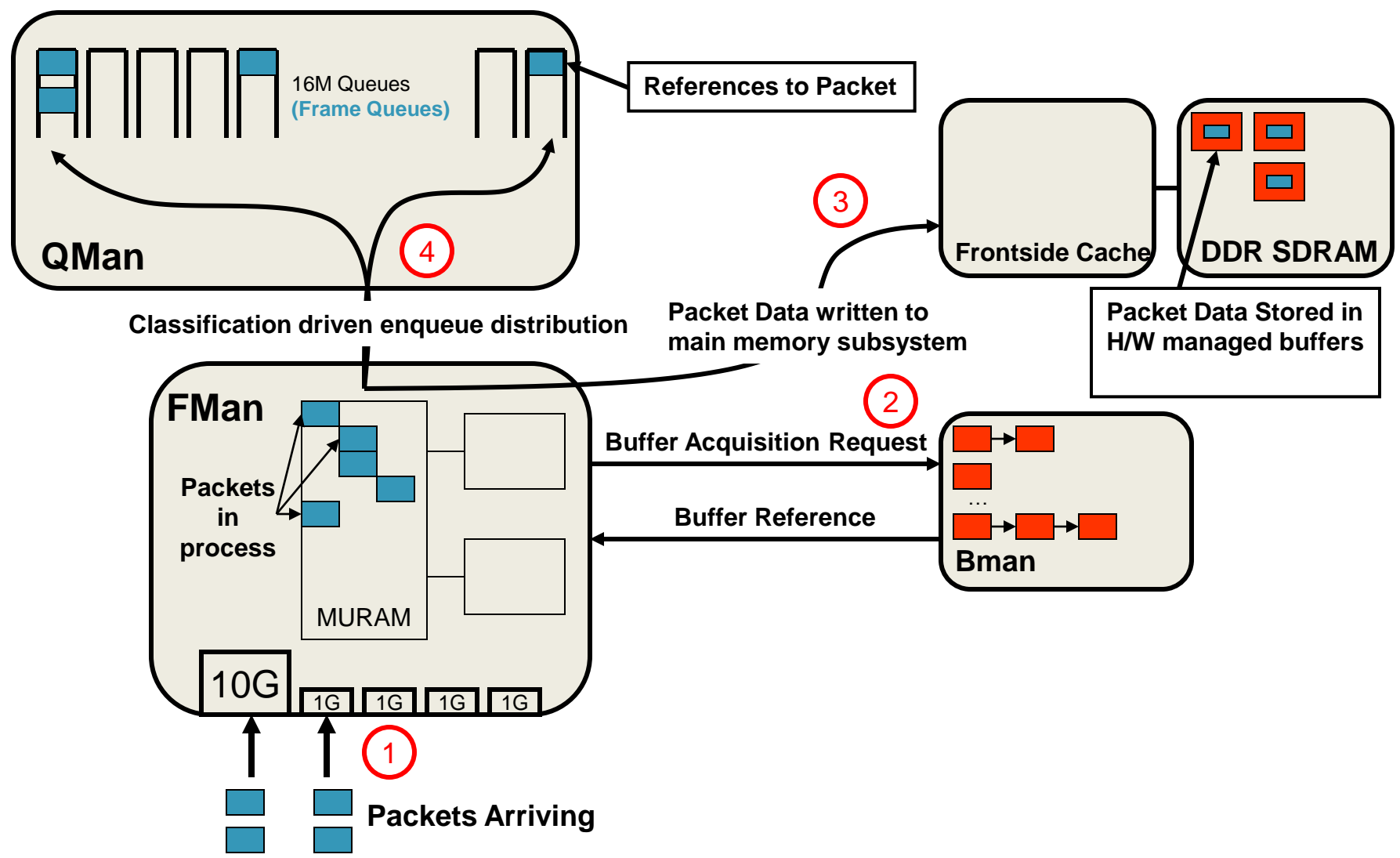


- **QMan provides a way to inter-connect DPAA components**
  - Cores (including IPC)
  - Hardware offload accelerators
  - Network interfaces – Frame Manager
- **Queue management**
  - High performance interfaces (“portals”) for enqueue/dequeue
  - Internal buffering of queue/frame data to enhance performance
- **Congestion avoidance and management**
  - RED/WRED
  - Tail drop for single queues and aggregates of queues
  - Congestion notification for “loss-less” flow control
- **Load spreading across processing engines (cores, HW accelerators)**
  - Order restoration
  - Order preservation/atomicity
- **Delivery to cache/HW accelerators of per queue context information with the data (Frames)**
  - This is an important offload for software using hardware accelerators

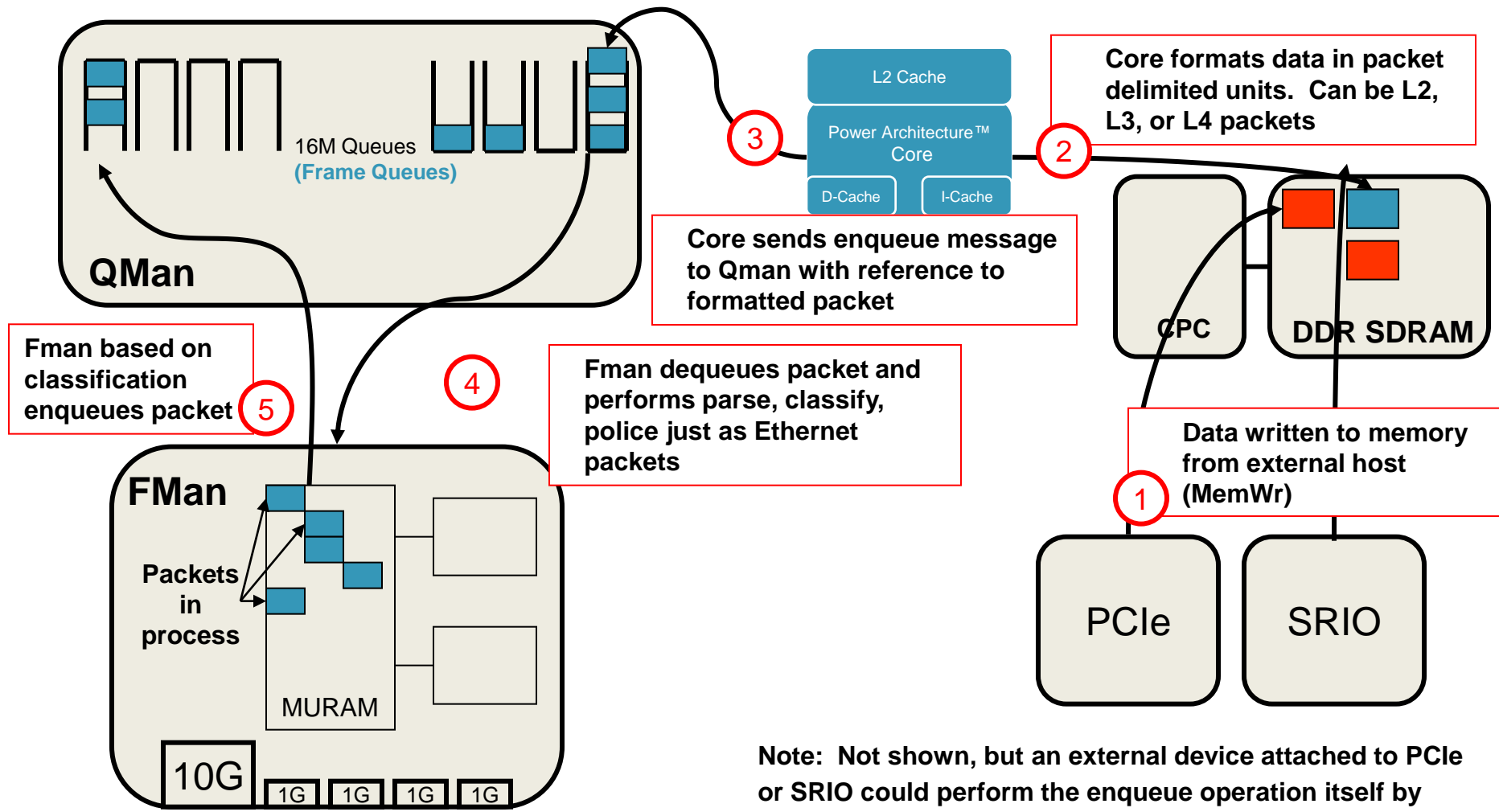




# FMan/QMan Ingress Packet Processing



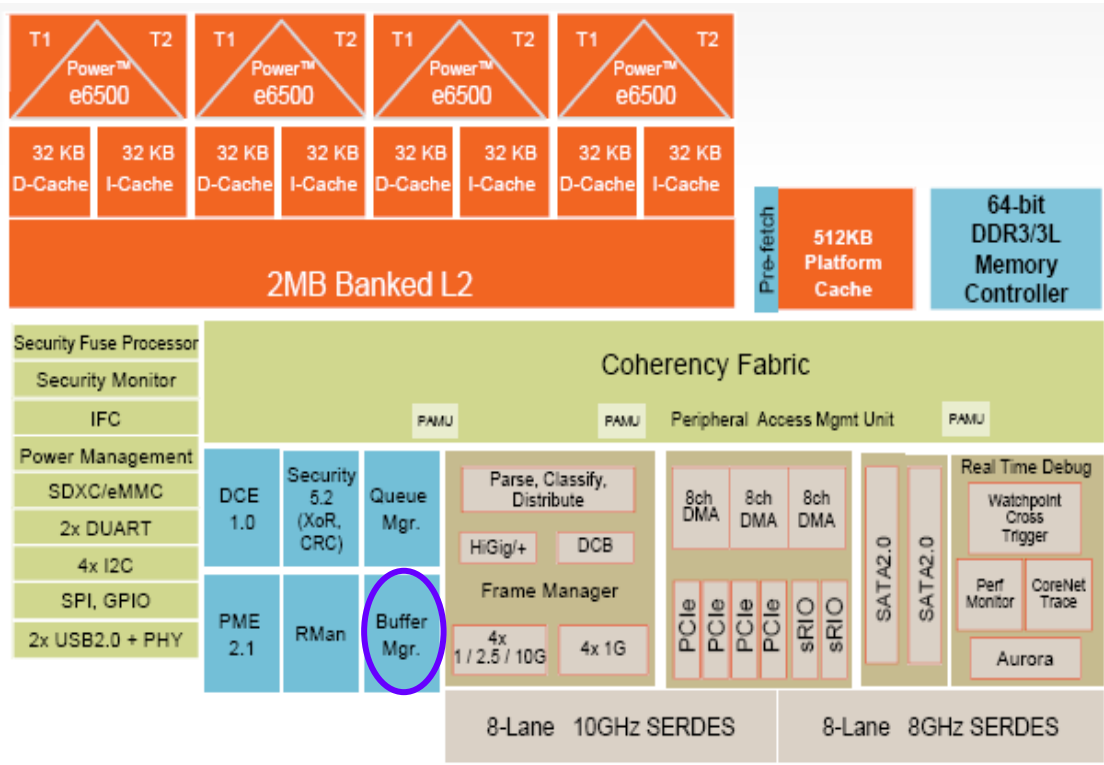
# Offline Parsing Example



Note: Not shown, but an external device attached to PCIe or SRIO could perform the enqueue operation itself by accessing a software memory mapped Qman portal



# Datapath Infrastructure: Buffer Manager

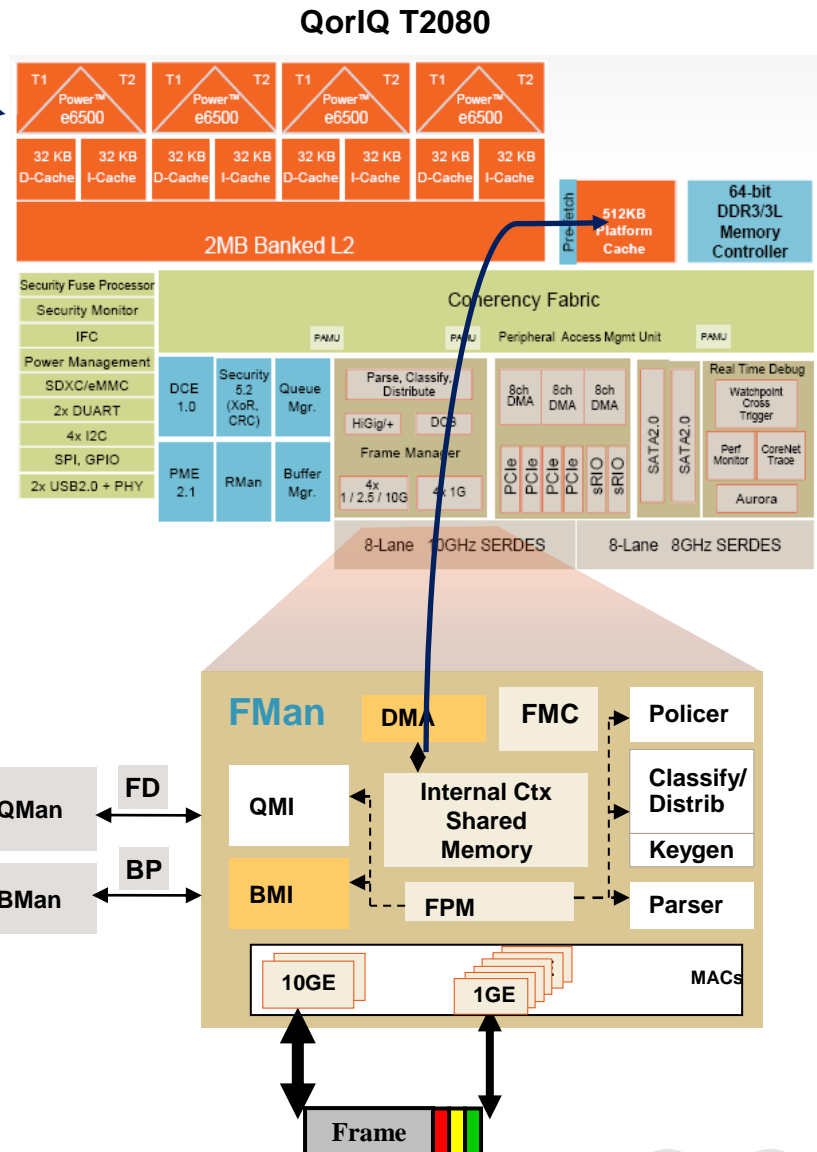
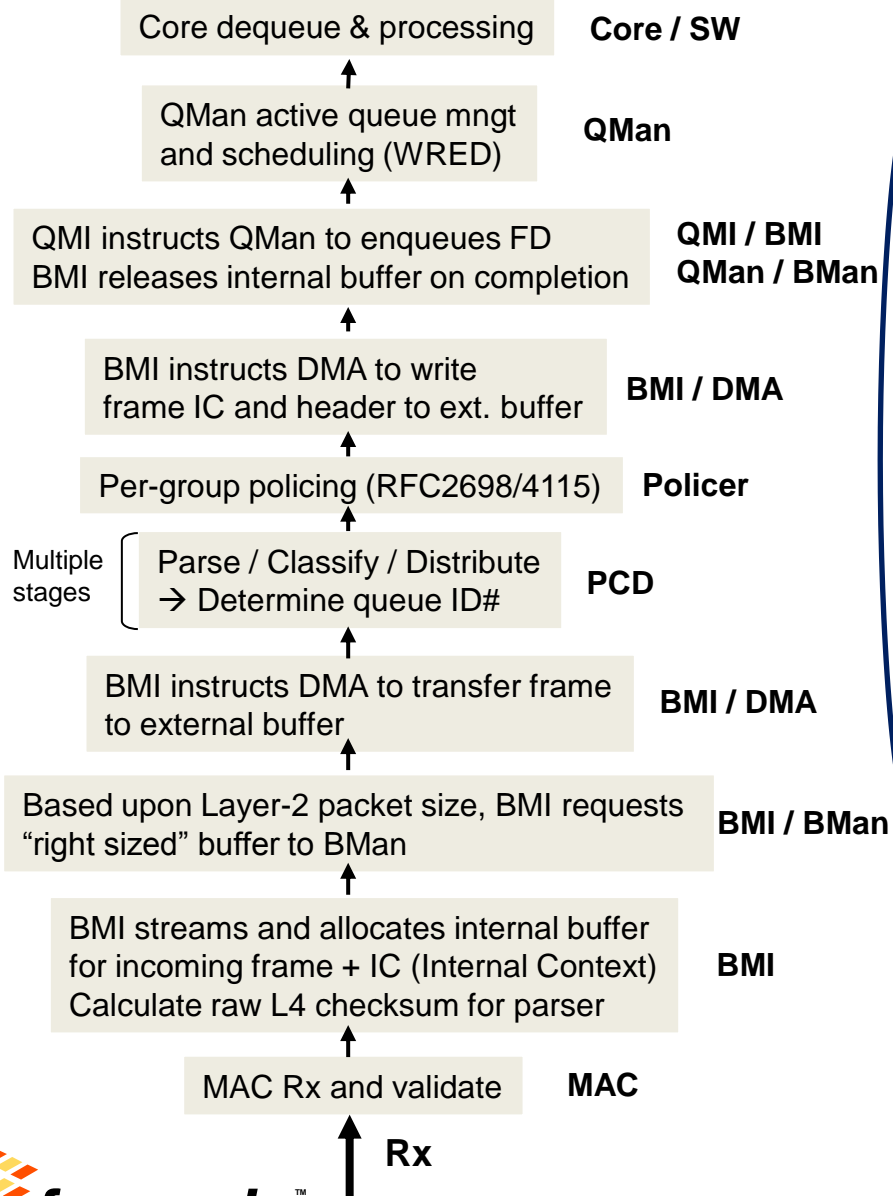


- Standardized command interface to SW and HW
  - Up to 50 software portals for software: resolves any multi-core race scenario
  - Up to 6 HW portal per HW block: simplified command for HW Accelerators
  - Up to 64 separate pools of free buffers
- BMan keeps a small per-pool stockpile of buffer pointers in internal memory
  - Stockpile of 64 buffer pointers per pool, maximum 2G buffer pointers
  - Absorbs bursts of acquire/release commands without external memory access
  - Minimized access to memory for buffer pool management.
- Pools (buffer pointers) overflow into DRAM
- LIFO buffer allocation policy
  - A released buffer is immediately used for receiving new data, using cache lines previously allocated



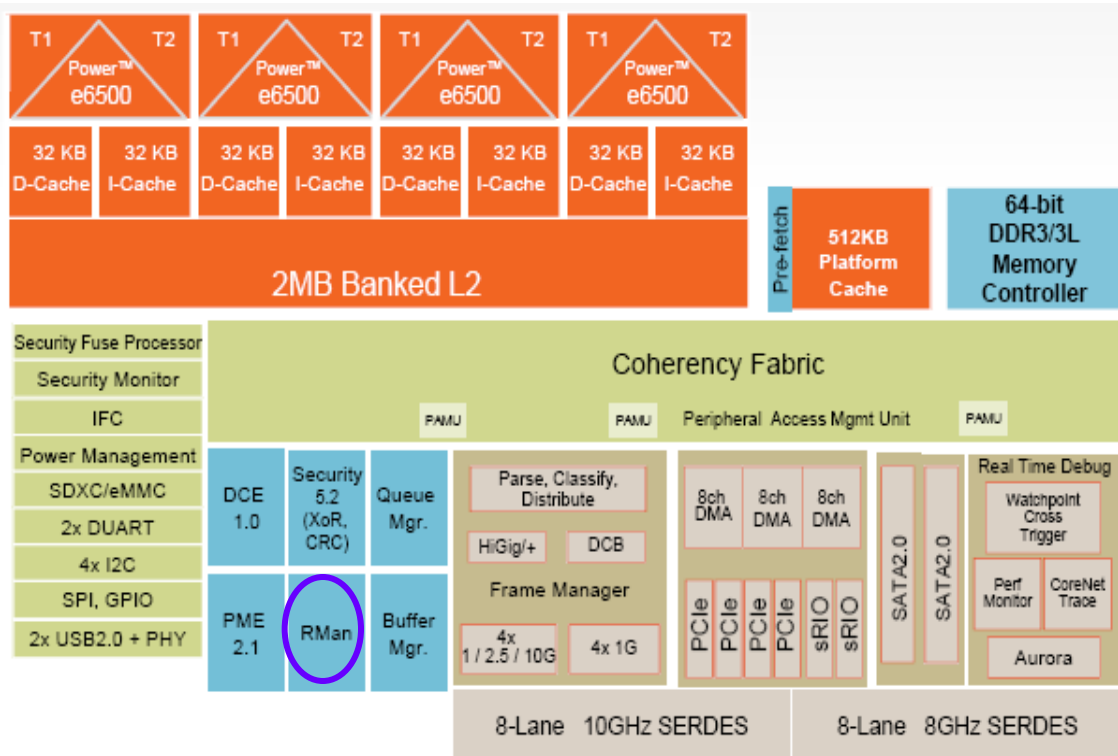


# Fman Modular Architecture Processing Pipeline





# RapidIO Message Manager

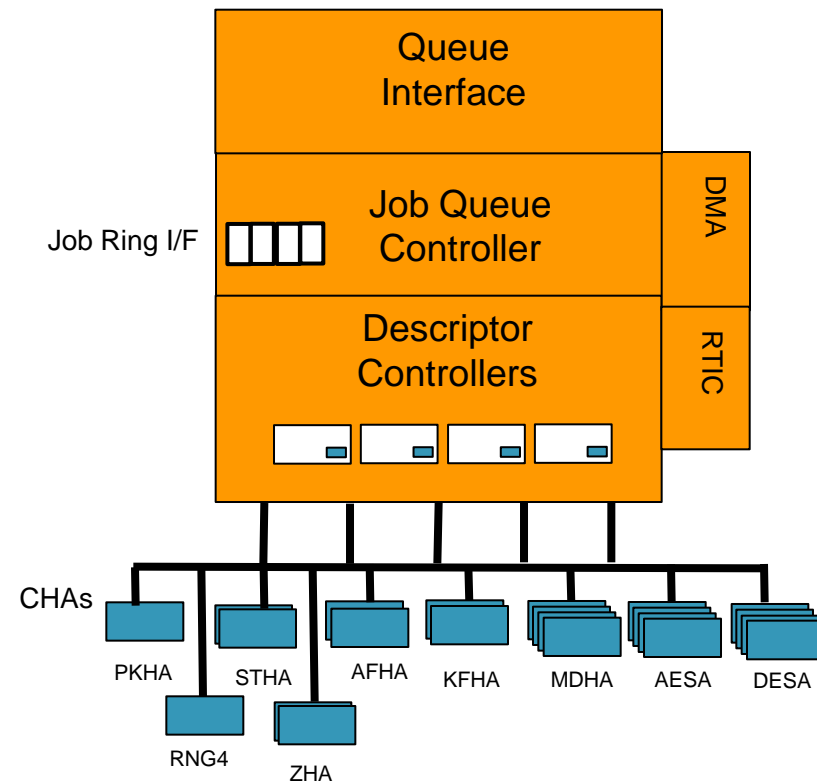


- RapidIO Rev 2.1 Compliant
- Dual controllers
- 1.25/2.5/3.125/5GBaud operation
  - 1x,2x,4x operation
- Extensive Transaction Type support
  - Type 9 Data Streaming
  - Type 10 Doorbells
  - Type 11 messaging
  - NWRITE/SWRITE
  - Port-write
- Support for hundreds of ingress/egress queues
- Robust QoS
- Direct interworking between Ethernet and RapidIO in hardware
  - No runtime CPU intervention required



# NXP Security Engine (SEC 5.2)

- (1) Public Key Hardware Accelerator (PKHA)
  - RSA and Diffie-Hellman (to 4096b)
  - Elliptic curve cryptography (1024b)
  - Supports Run Time Equalization
  
- (1) Random Number Generators (RNG4)
  - NIST Certified
  
- (2) Snow 3G Hardware Accelerators (STHA)
  - Implements Snow 3.0
  - One for Encryption (F8), one for Integrity (F9)
  
- (2) ZUC Hardware Accelerators (ZHA)
  - One for Encryption, one for Integrity
  
- (1) ARC Four Hardware Accelerator (AFHA)
  - Compatible with RC4 algorithm
  
- (2) Kasumi F8/F9 Hardware Accelerators (KFHA)
  - F8 , F9 as required for 3GPP
  - A5/3 for GSM and EDGE
  - GEA-3 for GPRS
  
- (4) Message Digest Hardware Accelerators (MDHA)
  - SHA-1, SHA-2 256,384,512-bit digests
  - MD5 128-bit digest
  - HMAC with all algorithms
  
- (4) Advanced Encryption Standard Accelerators (AESAs)
  - Key lengths of 128-, 192-, and 256-bit
  - ECB, CBC, CTR, CCM, GCM, CMAC, OFB, CFB, and XTS
  
- (4) Data Encryption Standard Accelerators (DESA)
  - DES, 3DES (2K, 3K)
  - ECB, CBC, OFB modes
  
- (4) CRC Unit
  - CRC32, CRC32C, 802.16e OFDMA CRC



Header & Trailer off-load for the following Security Protocols:  
•IPSec, SSL/TLS, 3G RLC, PDCP, SRTP, 802.11i, 802.16e, 802.1ae

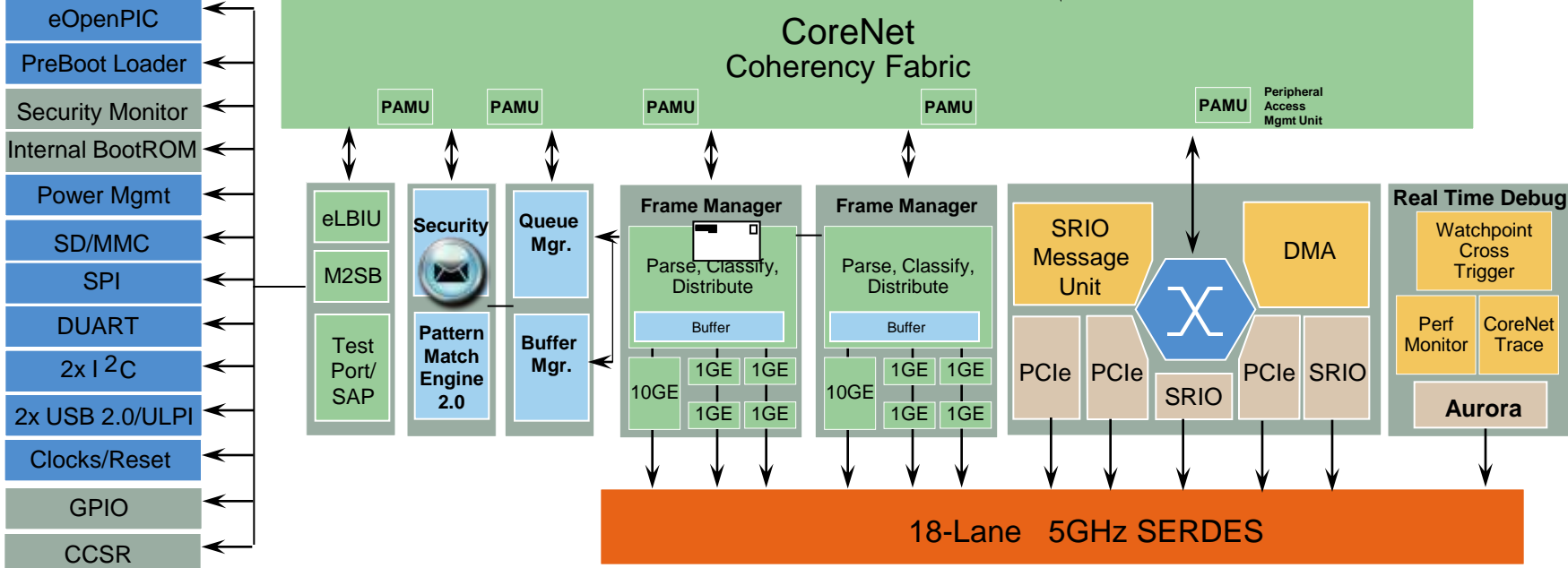
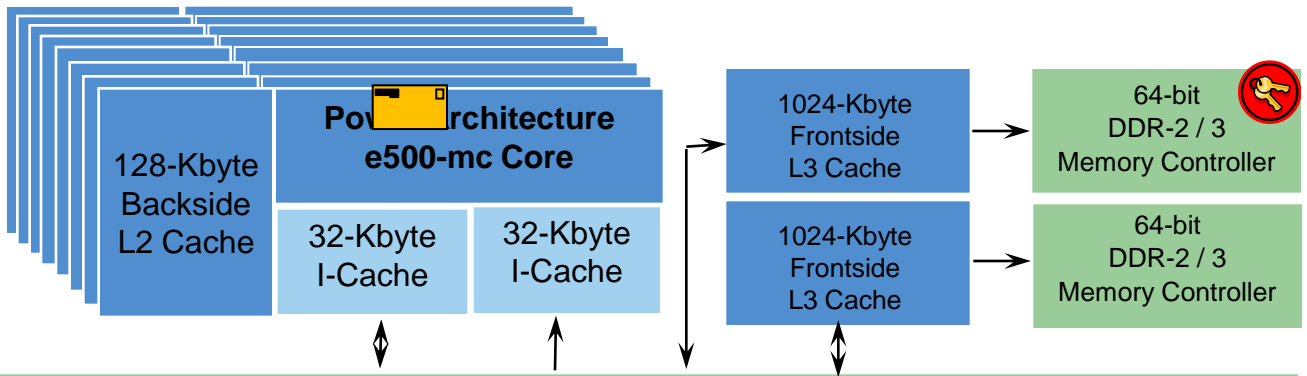




# Life of a Crypto Packet

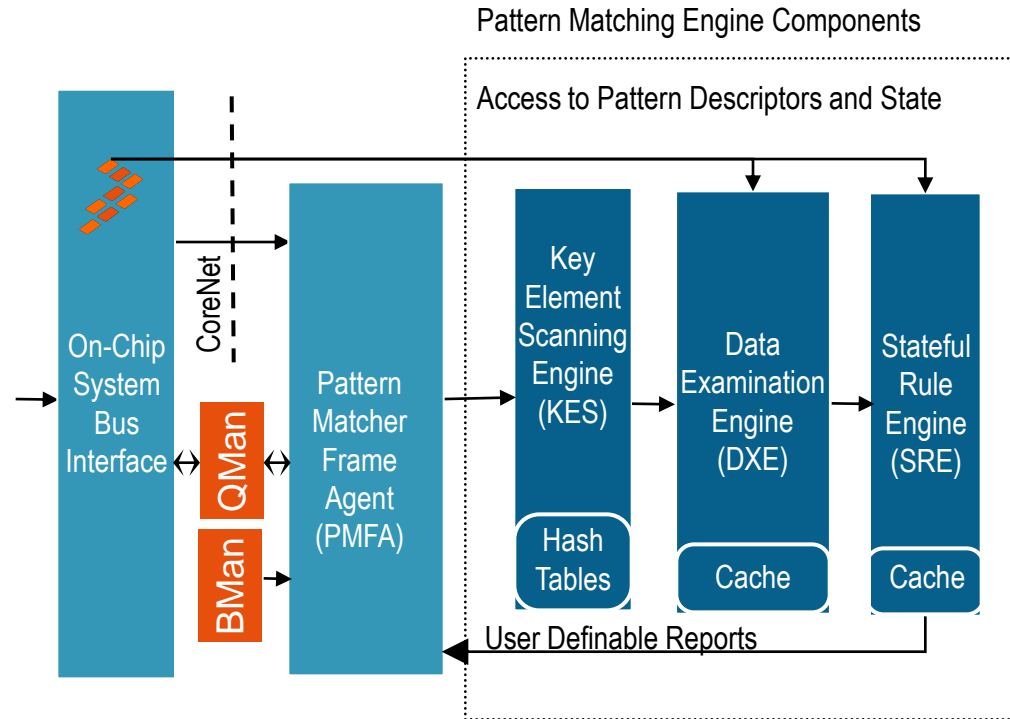


P4080E



# Pattern Matching Engine (PME) 2.1

- Regex support plus significant extensions:
  - Patterns can be split into 256 sets each of which can contain 16 subsets
  - 32K patterns of up to 128B length
  - 9.6 Gbps raw performance
- Combined hash/NFA technology
  - No “explosion” in number of patterns due to wildcards
  - Low system memory utilization
  - Fast pattern database compiles and incremental updates
- Matching across “work units”
  - Finds patterns in streamed data
- Pipeline of processing
  - PME offers pipeline of filtering, matching, and behavior base engine for complete pattern matching solution





# Life of a Packet in PME

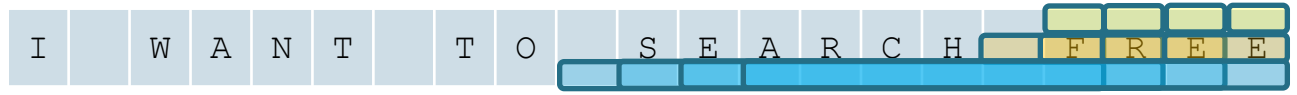


|     |                  |     |                    |
|-----|------------------|-----|--------------------|
| FD1 | 192.168.1.1:80   | TCP | 10.10.10.100:16734 |
|     | 192.168.1.1:25   | TCP | 10.10.10.100:17784 |
|     | 192.168.1.1:1863 | TCP | 10.10.10.100:16855 |

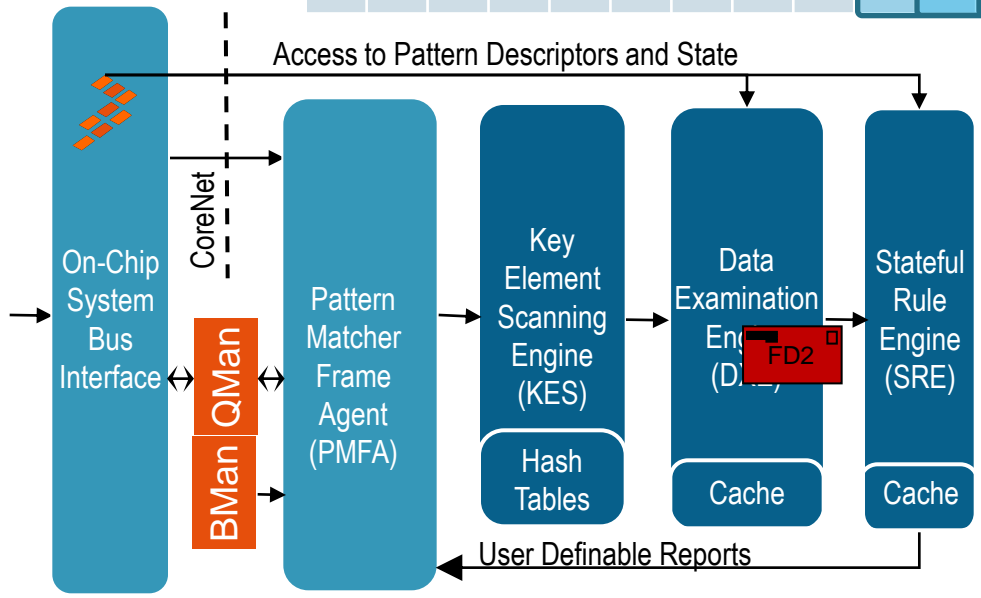


## Frame Queue: A

- flowA:FD1: 192.168.1.1:80->10.10.10.100:16734 "I want to search free"
- flowA:FD2: 192.168.1.1:80->10.10.10.100:16734 "Scale FTF 2011 event schedule"



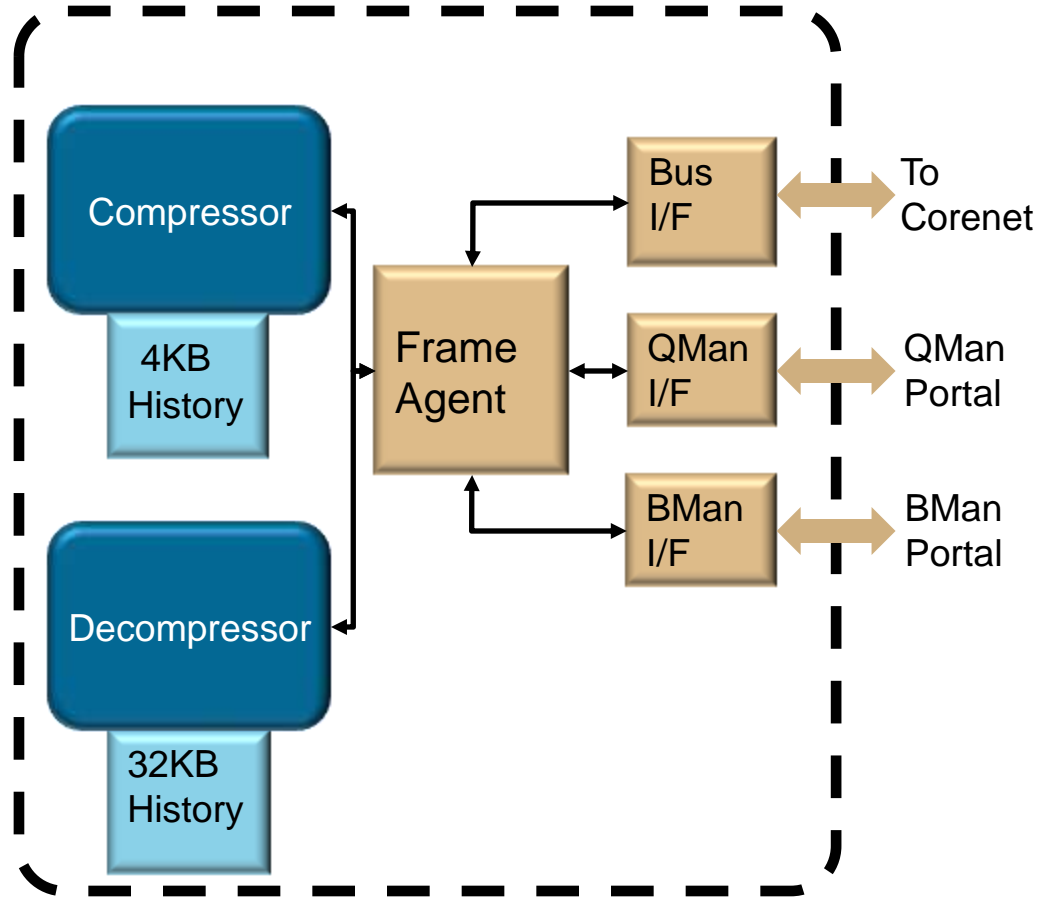
DDP  
**Patt1 /free/  
 tag=0x0001**



- Patterns
  - Patt1 /free/ tag=0x0001
  - Patt2 /freescale/ tag=0x0002
- KES
  - Compare hash value of incoming data(frames) against all patterns
- DXE
  - Retrieve the pattern with matched hash value for a final comparison
- SRE
  - Optionally post process match result before sending the report to the CPU

# Decompression and Compression Engine (DCE 1.0)

- Deflate
  - As specified as in RFC1951
- GZIP
  - As specified in RFC1952
- Zlib
  - As specified in RFC1950
  - Interoperable with the zlib 1.2.5 compression library
- Encoding
  - Supports Base 64 encoding and decoding (RFC4648).
- Operate up to 600 MHz
  - 10 Gbps Compress
  - 10 Gbps Decompress
  - 20 Gbps Aggregate



# Agenda

- T2080/1 Overview
- e6500 Core and Cache Hierarchy
- Power Management
- CoreNet Coherency Fabric Switch
- Data Path Acceleration Architecture (DPAA)
- SerDes Options
  - SerDes Lane Multiplexing
  - SerDes Supported Protocols



# T2080 SerDes Lane Multiplexing

| SRDS_PRTCL_S<br>1<br>RCW[128:135] | Lane<br>A     | Lane<br>B            | Lane<br>C            | Lane<br>D       | Lane<br>E       | Lane<br>F  | Lane<br>G  | Lane<br>H  | Parallel Port<br>availability    |
|-----------------------------------|---------------|----------------------|----------------------|-----------------|-----------------|------------|------------|------------|----------------------------------|
| 1C                                | SGMII (m9)    | SGMII (m10)          | SGMII (m1)           | SGMII (m2)      | SGMII (m3)      | SGMII (m4) | SGMII (m5) | SGMII (m6) | 2 RGMII<br>(FMAN MAC #3, #4/#10) |
| 95                                | SGMII (m9)    | SGMII (10)<br>3.125G | SGMII (m1)<br>3.125G | SGMII (m2)      | SGMII (m3)      | SGMII (m4) | SGMII (m5) | SGMII (m6) | 2 RGMII<br>(FMAN MAC #3, #4/#10) |
| A2                                | SGMII (m9)    | SGMII (m10)          | SGMII (m1)           | SGMII (m2)      | SGMII (m3)      | SGMII (m4) | SGMII (m5) | SGMII (m6) | 2 RGMII<br>(FMAN MAC #3, #4/#10) |
| 94                                | SGMII (m9)    | SGMII (10)<br>3.125G | SGMII (m1)<br>3.125G | SGMII (m2)      | SGMII (m3)      | SGMII (m4) | SGMII (m5) | SGMII (m6) | 2 RGMII<br>(FMAN MAC #3, #4/#10) |
| 51                                | XAUI (m9)     |                      |                      |                 | PCle4 (5/2.5)   | SGMII (m4) | SGMII (m5) | SGMII (m6) | 2 RGMII<br>(FMAN MAC #3, #4/#10) |
| 5F                                | HiGig (m9)    |                      |                      |                 | PCle2 (5/2.5)   | SGMII (m4) | SGMII (m5) | SGMII (m6) | 2 RGMII<br>(FMAN MAC #3, #4/#10) |
| 65                                | HiGig (m9)    |                      |                      |                 | PCle2 (5/2.5)   | SGMII (m4) | SGMII (m5) | SGMII (m6) | 2 RGMII<br>(FMAN MAC #3, #4/#10) |
| 6B                                | XFI (m9)      | XFI (m10)            | XFI (m1)             | XFI (m2)        | PCle2 (5/2.5)   | SGMII (m4) | SGMII (m5) | SGMII (m6) | 2 RGMII<br>(FMAN MAC #3, #4/#10) |
| 6C                                | XFI (m9)      | XFI (m10)            | SGMII (m1)           | SGMII (m2)      | PCle4 (5/2.5)   |            |            |            | 2 RGMII<br>(FMAN MAC #3, #4/#10) |
| 6D                                | XFI (m9)      | XFI (m10)            | SGMII (m1) 2.5G      | SGMII (m2) 2.5G | PCle4 (5/2.5)   |            |            |            | 2 RGMII<br>(FMAN MAC #3, #4/#10) |
| AB                                | PCle3 (5/2.5) |                      |                      |                 | PCle4 (8/5/2.5) |            |            |            | 2 RGMII<br>(FMAN MAC #3, #4/#10) |



# T2080 SerDes Supported Protocols

| Product | PCIe | SRIO | Aurora | SGMII | XAUI | HigGig | XFI | SATA |
|---------|------|------|--------|-------|------|--------|-----|------|
| T2080   | 4    | 2    | 1      | 8     | 1    | 1      | 2   | 2    |
| T2081   | 4    | x    | x      | 5     | x    | x      | 2   | x    |

- Numbers indicate the maximum that can be supported.

# Agenda

- T2080/1 Overview
- e6500 Core and Cache Hierarchy
- Power Management
- CoreNet Coherency Fabric Switch
- Data Path Acceleration Architecture (DPAA)
- SerDes Options
- **Voltage ID**
  - What is VID?
  - Basic Steps for System to implement VID

# What is VID?

- VID is a specific method of selecting the optimum voltage-level to guarantee performance and power targets.
- QorIQ device contains fuse block registers defining required voltage level. This eFUSE definition is accessed through the Fuse Status Register (DCFG\_FUSESR).
- Customer software will read the VID value from factory-set efuse values and configure regulator values appropriately.
- **For T2080, the core VDD value will range from 1.025V to 0.975V in 12.5mV steps**

| Power Pins | Power Islands on T2080 |
|------------|------------------------|
| VDD        | Core and Platform      |
| USB_SVDD   | USB supply             |

|                         |              |
|-------------------------|--------------|
| Start up voltage        | 1.025 ± 30mV |
| During normal operation | VID ± 30mV   |

## Basic Steps for System to implement VID

- At power up time zero, regulator must come up at default voltage as defined per product. For T2080, that is 1.025V.
- VERY EARLY in the boot code and before many high speed or other power hungry features or interfaces are turned on, the DCFG\_FUSESR register is read for the VID information. This value is translated into whatever commands to program up the new voltage value for the regulator.
- Once the regulator is sent the new values, a period of time needs to pass to allow the regulator to change values BEFORE power hungry features and higher clock rates are enabled/changed.



# Agenda

- T2080/1 Overview
- e6500 Core and Cache Hierarchy
- Power Management
- CoreNet Coherency Fabric Switch
- Data Path Acceleration Architecture (DPAA)
- SerDes Options
- Voltage ID
- eSDHC
  - New Features
  - Interface New Signals
  - Supported SD Card Modes
  - Examples

## eSDHC New Features

- Supports SDXC cards
  - Up to 2TB space
- Supports cards with UHS-I speed grade
  - Ultra high speed grade
    - SDR12, SDR25, SDR50, SDR104, DDR50
  - UHS-I cards work on 1.8V signaling
  - On board dual voltage regulators are needed to support UHS-I cards because card initialization happens at 3.3V and regular operations happen at 1.8V
  - SD controller provides a signal to control the voltage regulator. The signal is controlled via SDHC\_VS bit
- eMMC 4.5 support (HS200, DDR)

## eSDHC Interface New Signals

- **SDHC\_CMD\_DIR** - Command Line Direction Control
- **SDHC\_DAT0\_DIR** - DAT0 Line Direction Control
- **SDHC\_DAT123\_DIR** - DAT1 to DAT3 Line direction control
  - DIR signals are required to change direction of external voltage translator
  - Separate DIR signals are implemented to support card interrupt on DAT1 in single bit mode
- **SDHC\_VS** - External voltage select, to change voltage of external regulator
- **SDHC\_CLK\_SYNC\_IN** – SYNC clock input
- **SDHC\_CLK\_SYNC\_OUT** – SYNC clock output

# Supported SD Card Modes

| Mode   | 1 bit Support |          | 4 bit Support |          |
|--------|---------------|----------|---------------|----------|
|        | T1040         | SD (3.0) | T1040         | SD (3.0) |
| DS     | Yes           | Yes      | Yes           | Yes      |
| HS     | Yes           | Yes      | Yes           | Yes      |
| SDR12  | No            | No       | Yes           | Yes      |
| SDR25  | No            | No       | Yes           | Yes      |
| SDR50  | No            | No       | Yes           | Yes      |
| SDR104 | No            | No       | Yes           | Yes      |
| DDR50  | No            | No       | Yes           | Yes      |

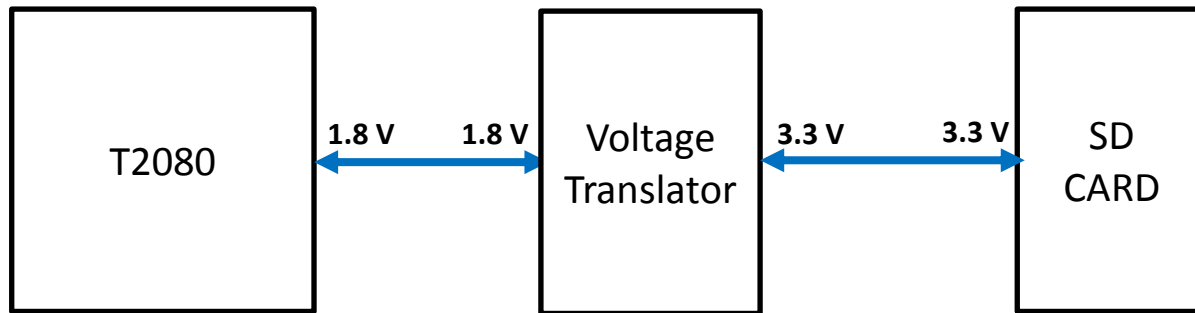


# Supported MMC/eMMC Modes

| Mode  | 1 bit Support |            | 4 bit Support |            | 8 bit support |            |
|-------|---------------|------------|---------------|------------|---------------|------------|
|       | T2080         | eMMC (4.5) | T2080         | eMMC (4.5) | T2080         | eMMC (4.5) |
| DS    | Yes           | Yes        | Yes           | Yes        | Yes           | Yes        |
| HS    | Yes           | Yes        | Yes           | Yes        | Yes           | Yes        |
| HS200 | No            | No         | Yes           | Yes        | Yes           | Yes        |
| DDR   | No            | No         | Yes           | Yes        | No            | Yes        |



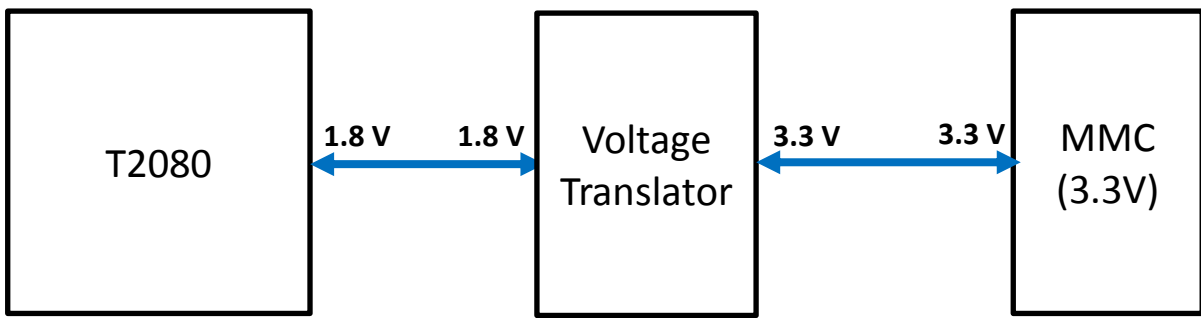
# SD Card Connections for T2080 (DS and HS Modes)



◀▶ **CMD, DAT[0], DAT[1:3], CLK, CD\_B, WP**

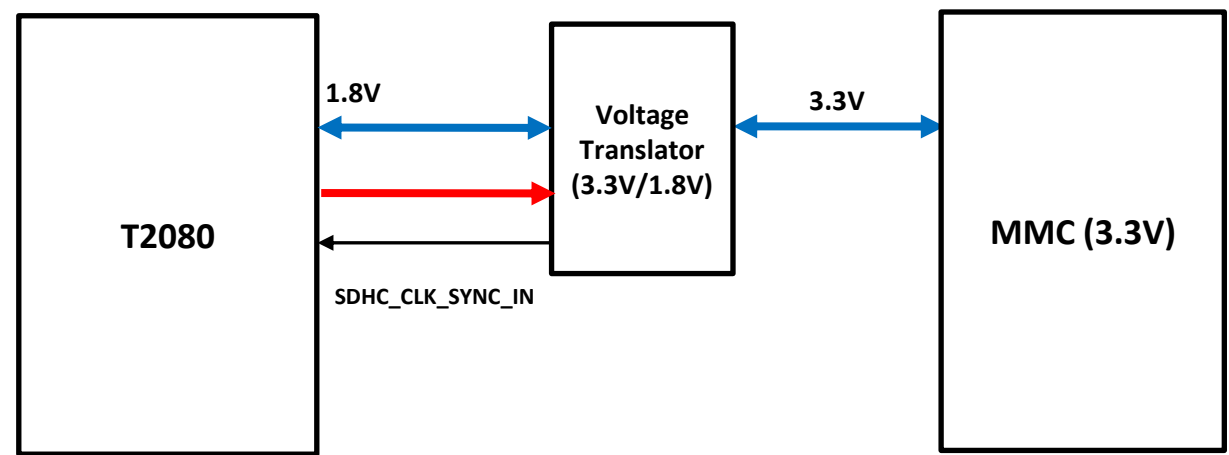
- Other signals should be left NC
- SYNC\_OUT should be pulled-down with a weak resistor or the pin should be configured for alternate functionality

# MMC Card Connections for T2080 (DS, HS, HS200 Modes)



- CMD, DAT[0], DAT[1:7], CLK, CD**
- Other signals should be left NC
  - SYNC\_OUT should be pulled-down with a weak resistor or the pin should be configured for alternate functionality
  - Voltage translator is not needed for 1.8V MMC.

# MMC (3.3V) Connections for T2080 (DDR Mode)



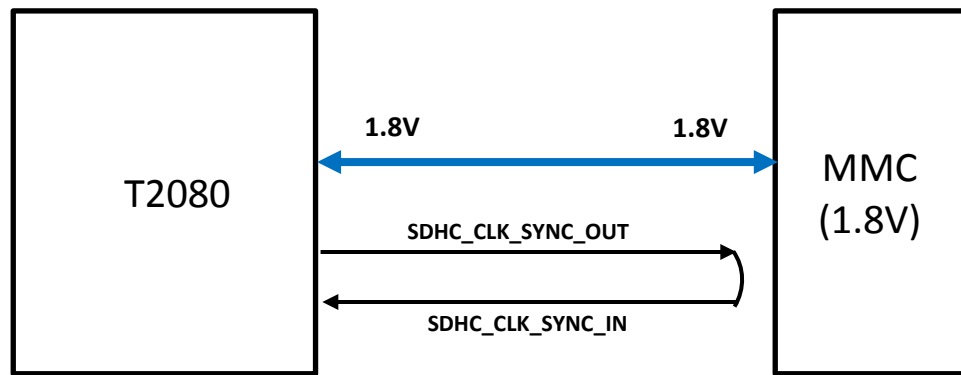
|  |                                |
|--|--------------------------------|
|  | CMD, DAT[0], DAT[1:3], CLK, CD |
|  | DIR                            |

- In DDR mode all the input signals are sampled with respect to SYNC\_IN

- Other signals should be left NC
- SYNC\_OUT should be pulled-down with a weak resistor or the pin should be configured for alternate functionality
- Voltage translator is not needed for 1.8V MMC.



# MMC (1.8V) Connections for T2080 (DDR Mode)



- ↔ **CMD, DAT[0], DAT[1:7], CLK, CD**
- Other signals should be left NC
- In DDR mode all the input signals are sampled wrt SYNC\_IN

# Agenda

- T2080/1 Overview
- e6500 Core and Cache Hierarchy
- Power Management
- CoreNet Coherency Fabric Switch
- Data Path Acceleration Architecture (DPAA)
- SerDes Options
- Voltage ID
- eSDHC
- **PCI Express**

# PCI Express

- This chip instantiates four PCI Express controllers, each with the following key features:
  - One PCI Express controller supports end-point SR-IOV
  - Two physical functions
  - 64 virtual functions per physical function
  - Eight MSI-X per either physical function or virtual function
  - Two PCI Express controllers support 2.0 (maximum lane width of x8)
  - Two PCI Express controllers support 3.0 (maximum lane width of x4)
  - Power-on reset configuration options allow root complex or endpoint functionality
  - x8, x4, x2, and x1 link widths support
  - Both 32- and 64-bit addressing and 256-byte maximum payload size
  - Inbound INTx transactions
  - Message signaled interrupt (MSI) transactions

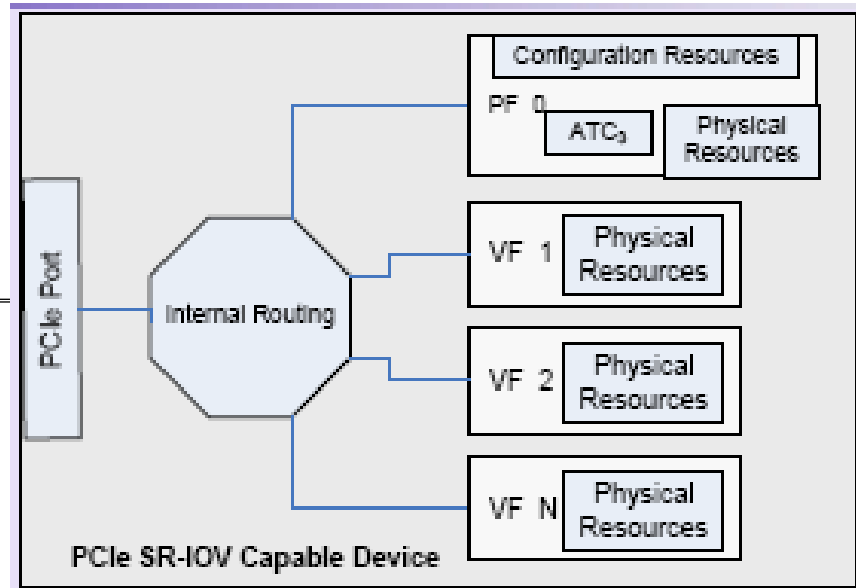
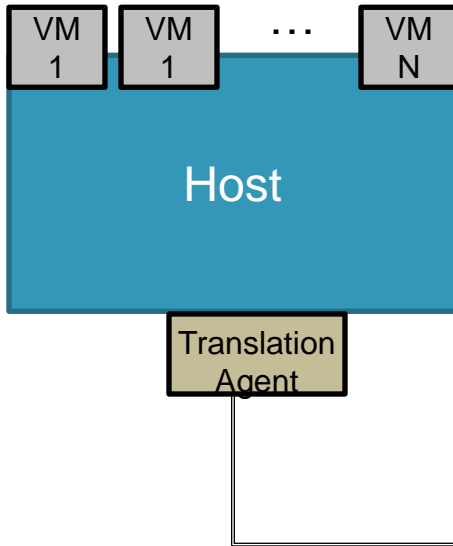
# PCIe SR-IOV End Point

Use Case: T2080 as services card, Converged Network Adapter, “Intelligent NIC”.

Single Management physical or virtual machine on host handles end-point configuration.

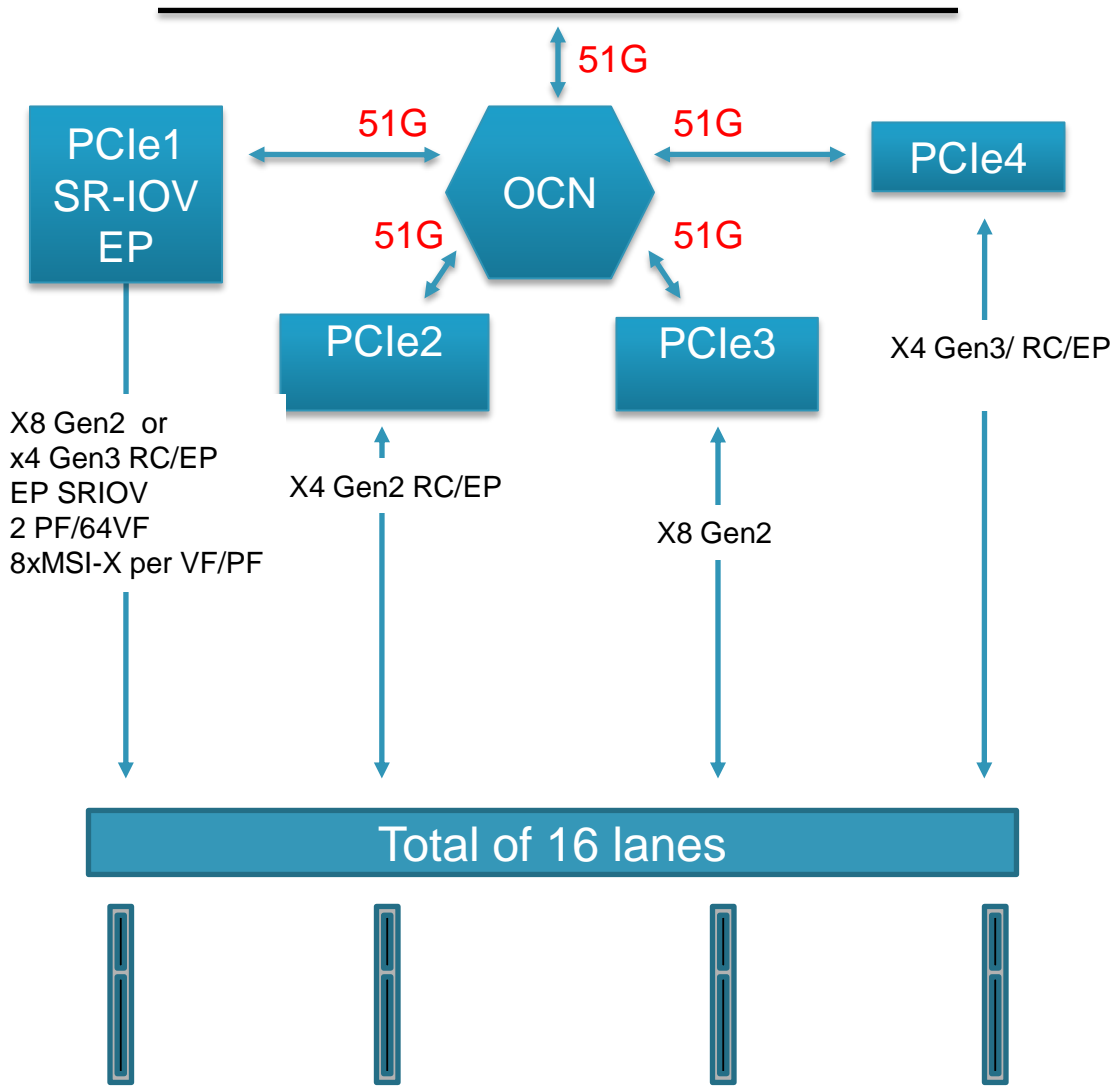
Each Virtual Machine running on Host thinks it has a private version of the services card.

Translation agent (in host or chipset) performs PAMU like address translation on behalf of the VFs.



Goal:  
Single controller (up to x4 Gen 3), 1 PF, 64 VFs

# PCIe Sub System



X8 Gen2 or  
x4 Gen3 RC/EP  
EP SRIOV  
2 PF/64VF  
8xMSI-X per VF/PF

X4 Gen2 RC/EP

X8 Gen2

X4 Gen3/ RC/EP

| 16 SERES PCIe Configuration |                    |                    |                    |
|-----------------------------|--------------------|--------------------|--------------------|
| PCIe1                       | PCIe2              | PCIe3              | PCIe4              |
| x8 <sub>gen2</sub>          | x4 <sub>gen2</sub> | x8 <sub>gen2</sub> | x4 <sub>gen2</sub> |
| x4 <sub>gen3</sub>          |                    |                    | x4 <sub>gen3</sub> |

# Agenda

- T2080/1 Overview
- e6500 Core and Cache Hierarchy
- Power Management
- CoreNet Coherency Fabric Switch
- Data Path Acceleration Architecture (DPAA)
- SerDes Options
- Voltage ID
- eSDHC
- PCI Express
- **Enablement**
  - **Software & Tools**
  - **Collaterals / Documentation**

## T2080 Software & Tools at a Glance

- Two Reference Design Boards
  - T2080 QDS
  - T2080 RDB
- Software Support
  - SDK 1.5
  - SDK support includes
    - Legacy features (refer SDK 1.4 release notes)
    - New features
    - FMAN and Linux based drivers
- QorIQ Configuration Suite
- Code Warrior based debugger, flash programmer

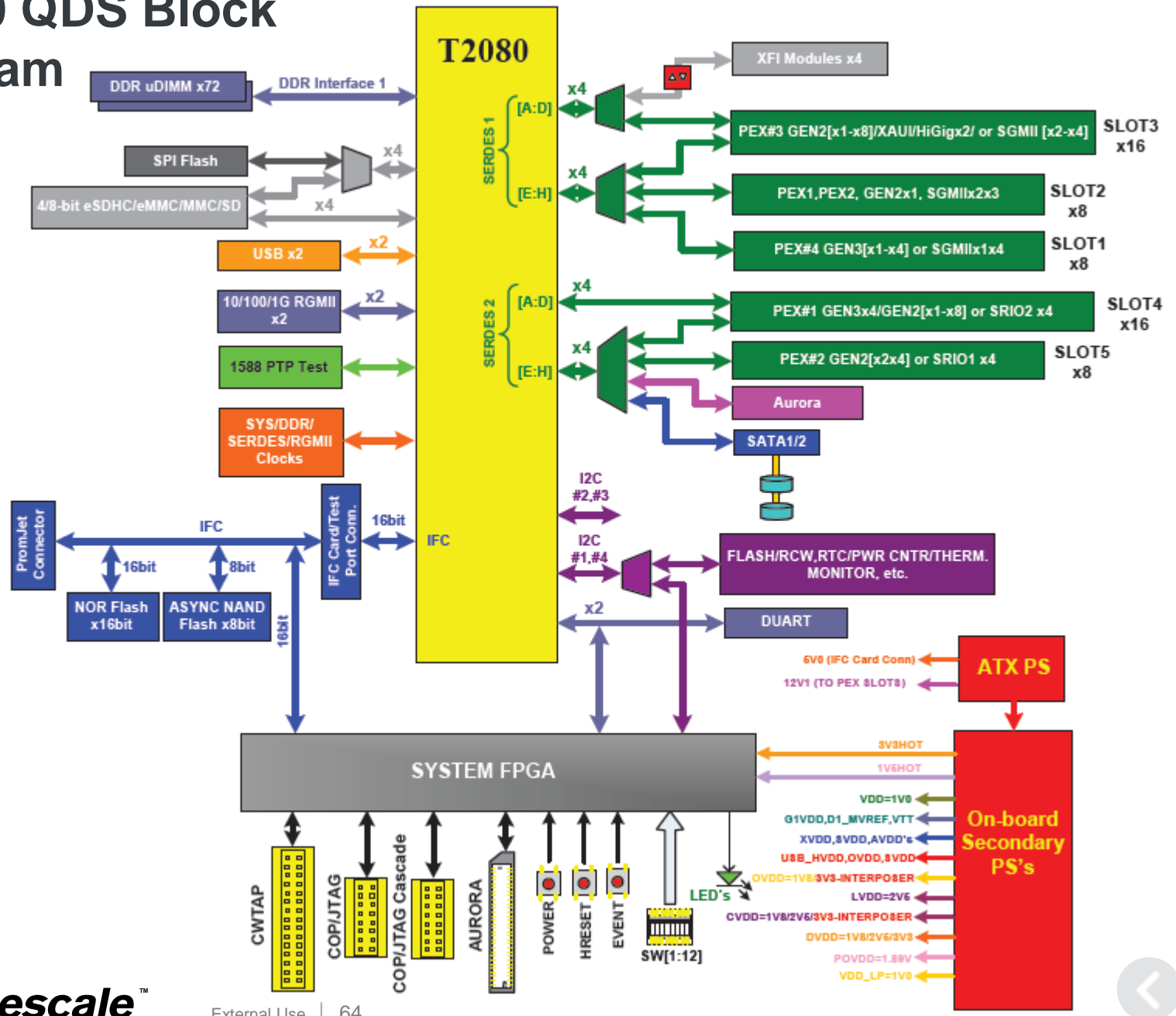


# T2080 QDS





# NXP T2080 QDS Block Diagram



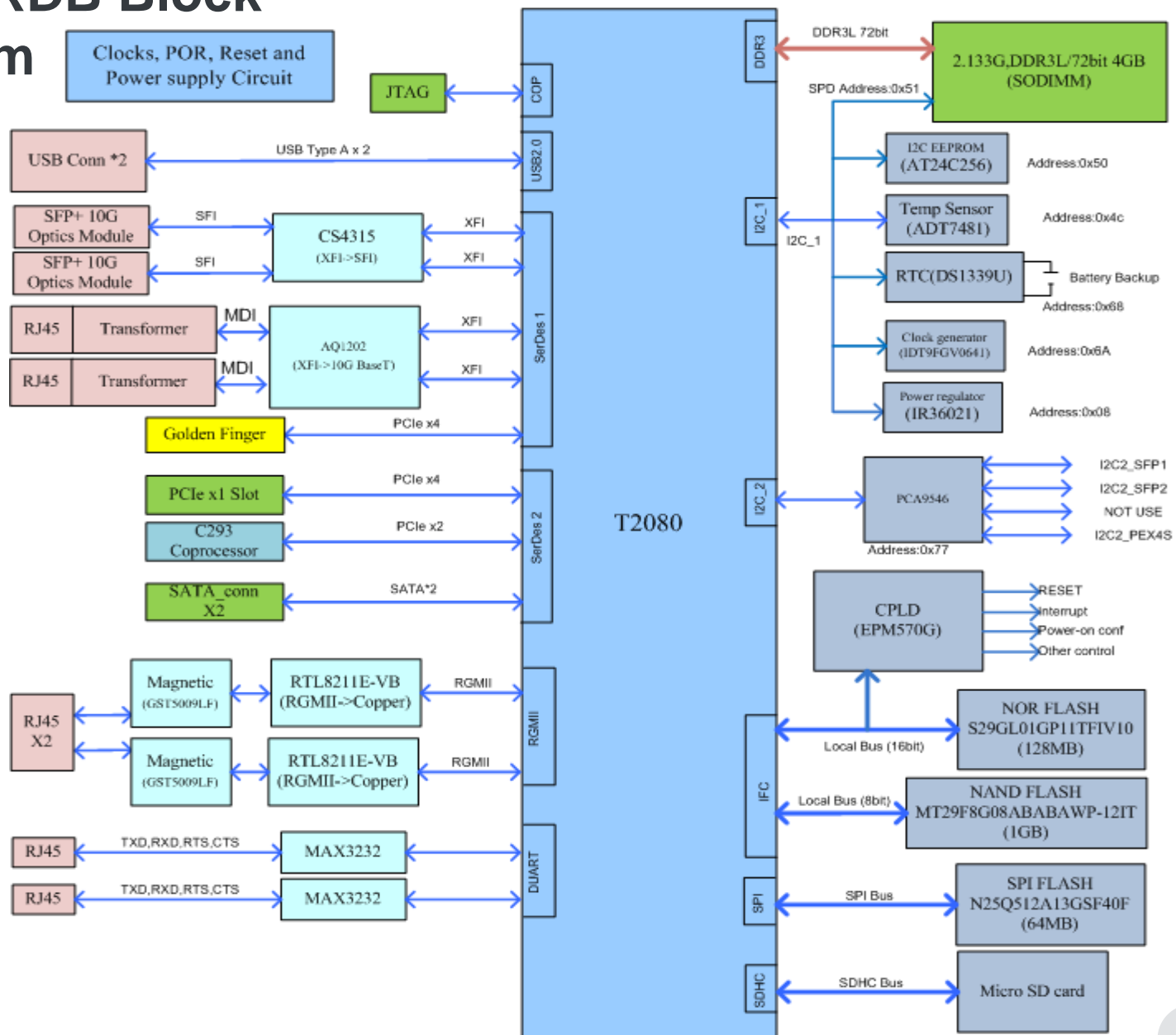


# T2080 RDB System



# NXP T2080 RDB Block Diagram

## Diagram



## Collaterals / Documentation

### On the Core:

- e6500 core Reference Manual (Rev I , 2013)

### On the SoC device:

- T2080 Fact-sheet and Product brief
- HW Spec Rev E
- Reference Manual Rev C
- Advanced Debug and Performance Monitoring Reference Manual
- Errata-sheet Rev B
- Application Notes
  - AN4804 – T2080 Design Checklist
  - AN4773 - Migration Guide from T2081 to T1040

# Agenda

- T2080/1 Overview
- e6500 Core and Cache Hierarchy
- Power Management
- CoreNet Coherency Fabric Switch
- Data Path Acceleration Architecture (DPAA)
- SerDes Options
- Voltage ID
- eSDHC
- PCI Express
- Enablement
- **Conclusion**

# QorIQ T2 Families Extend Market Leadership

- **First 64-bit embedded processor with eight virtual core and DPAA**
  - Reduces system cost, design complexity and power
- **One of the industry's most scalable, pin-compatible family of devices**
  - The T2 processor is primarily intended to succeed our successful P3041 and P2041 mid-range series of quad-core devices.
  - The T2081 is a smaller-package version of the T2080, which is pin-compatible with the quad-core T1 family.
- Ideal for **mid-range** control plane applications or mixed control and data plane applications.



# Introducing The QorIQ LS2 Family

**Breakthrough, software-defined approach to advance the world's new virtualized networks**

## **New, high-performance architecture built with ease-of-use in mind**

Groundbreaking, flexible architecture that abstracts hardware complexity and enables customers to focus their resources on innovation at the application level

## **Optimized for software-defined networking applications**

Balanced integration of CPU performance with network I/O and C-programmable datapath acceleration that is right-sized (power/performance/cost) to deliver advanced SoC technology for the SDN era

## **Extending the industry's broadest portfolio of 64-bit multicore SoCs**

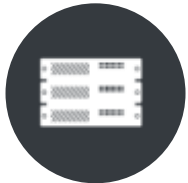
Built on the ARM® Cortex®-A57 architecture with integrated L2 switch enabling interconnect and peripherals to provide a complete system-on-chip solution

# NXP QorIQ LS2 Family

## Key Features



**SDN/NFV  
Switching**



**Data  
Center**



**Wireless  
Access**

Unprecedented performance and ease of use for smarter, more capable networks

### High performance cores with leading interconnect and memory bandwidth

- 8x ARM Cortex-A57 cores, 2.0GHz, 4MB L2 cache, w Neon SIMD
- 1MB L3 platform cache w/ECC
- 2x 64b DDR4 up to 2.4GT/s

### A high performance datapath designed with software developers in mind

- New datapath hardware and abstracted acceleration that is called via standard Linux objects
- 40 Gbps Packet processing performance with 20Gbps acceleration (crypto, Pattern Match/RegEx, Data Compression)
- Management complex provides all init/setup/teardown tasks

### Leading network I/O integration

- 8x1/10GbE + 8x1G, MACSec on up to 4x 1/10GbE
- Integrated L2 switching capability for cost savings
- 4 PCIe Gen3 controllers, 1 with SR-IOV support
- 2 x SATA 3.0, 2 x USB 3.0 with PHY





# See the LS2 Family First in the Tech Lab!



## 4 new demos built on QorIQ LS2 processors:



Performance Analysis Made Easy



Leave the Packet Processing To Us



Combining Ease of Use with Performance



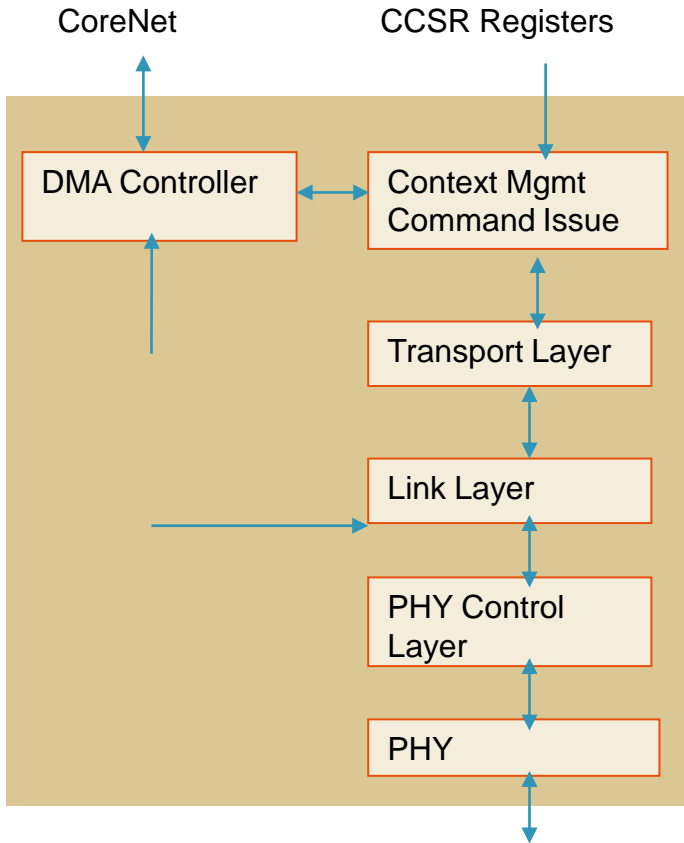
Tools for Every Step of Your Design





[www.Freescale.com](http://www.Freescale.com)

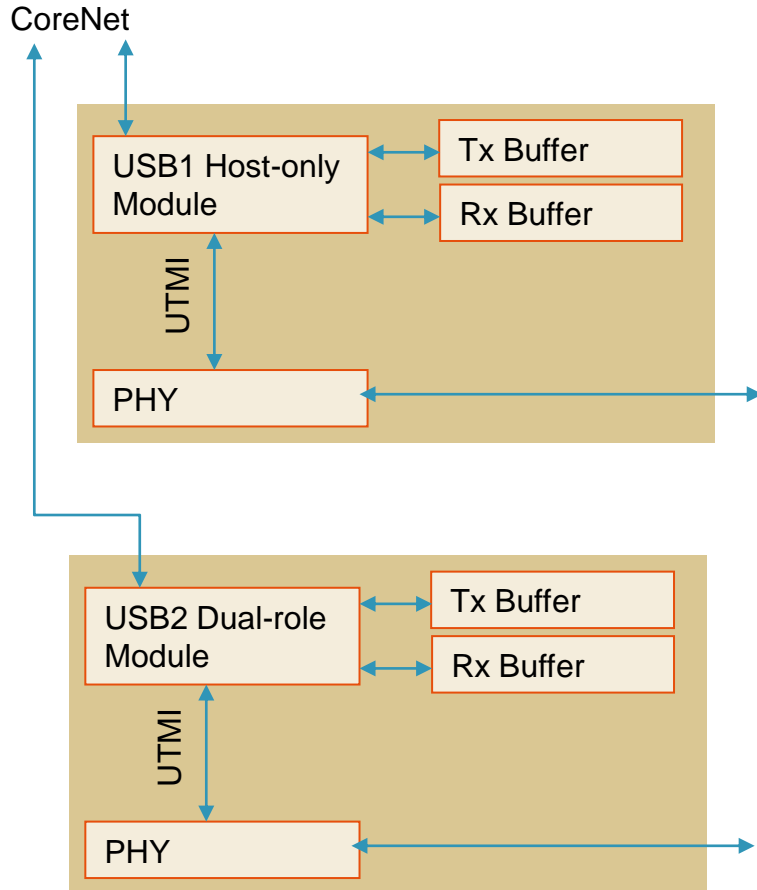
# SATA 2.0



- Compliant to Serial ATA 2.6
- Supports speeds: 1.5 Gbps (first-generation SATA), 3 Gbps (second-generation SATA)
- Supports advanced technology attachment packet interface (ATAPI) devices
- High-speed descriptor-based DMA controller
- Native command queuing (NCQ) commands
- Supports port multiplier operation
- Supports hot plug including asynchronous signal recovery

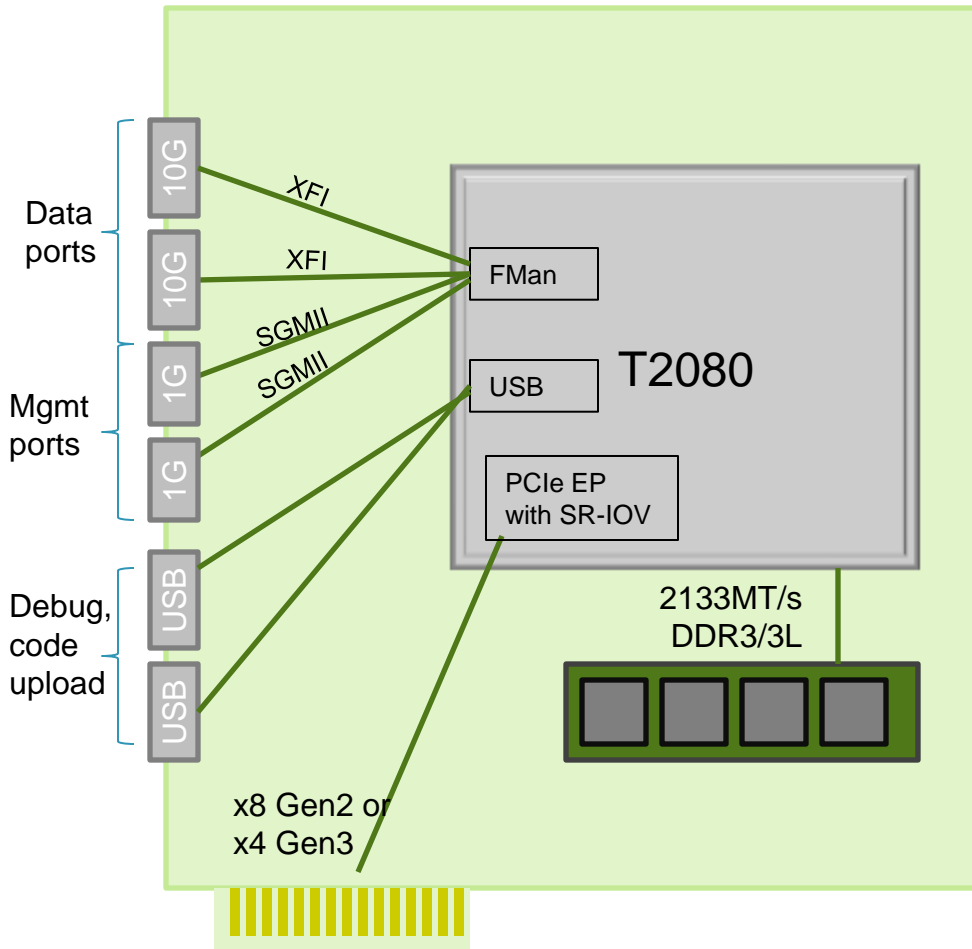
See AN111 of FTF08 for more details

# USB 2.0



- Complies with USB Specification Rev 2.0
- Operates as a standalone USB host controller
  - Enhanced host controller interface (EHCI)
- High-speed (480 Mbps), full-speed (12 Mbps), and low-speed (1.5 Mbps) operation. Low speed is only supported in host mode.
- On-chip, USB-2.0, full-speed/high-speed PHY with UTMI
- Operates as a standalone USB device
  - Supports one upstream facing port
  - Supports six bidirectional USB endpoints

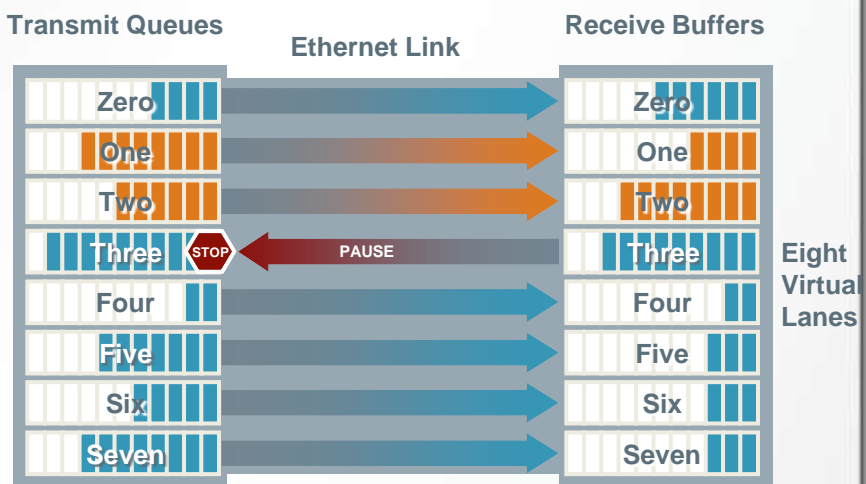
# Target Application: 20Gb/s iNIC



- Well-balanced device for 20Gb/s bi-directional application:
  - FMan moves about 25Gb/s
  - 3x DMA engines move about 20Gb/s
  - x4 Gen3 or x8 Gen2 PCIe moves 32Gb/s
- SR-IOV allows virtual machines on host to see a private iNIC
- 15.5W power fits in 30W slot-provided power budget
- Improved PCIe Endpoint capabilities support customization of Device ID, Class Code, and Vendor ID. Driver can be stored in Expansion ROM
- Offload accelerators for services cards: 10Gb/s IPSEC or Kasumi, 10Gb/s pattern matching, 17.5Gb/s data compression
- PCIe card reference board available

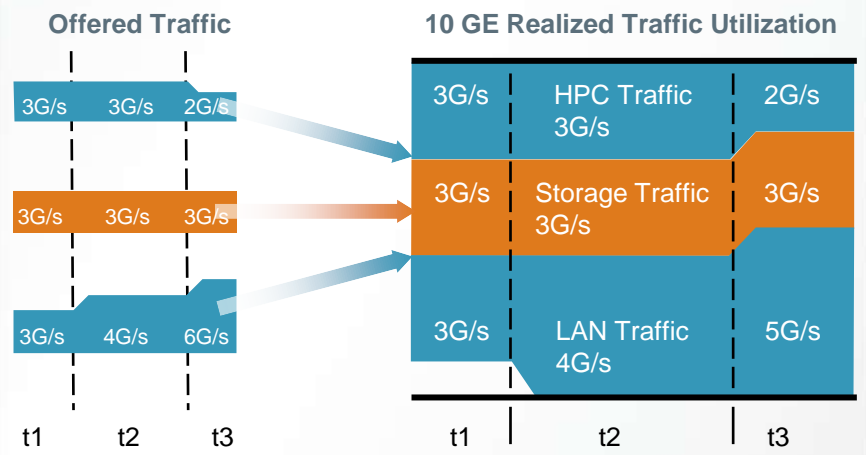
# Data Center Ethernet: PFC & Bandwidth Management

## Priority Flow Control



- **Enables lossless behavior for each class of service**
- **PAUSE sent per virtual lane when buffers limit exceeded**
- **IEEE 802.1Qbb**

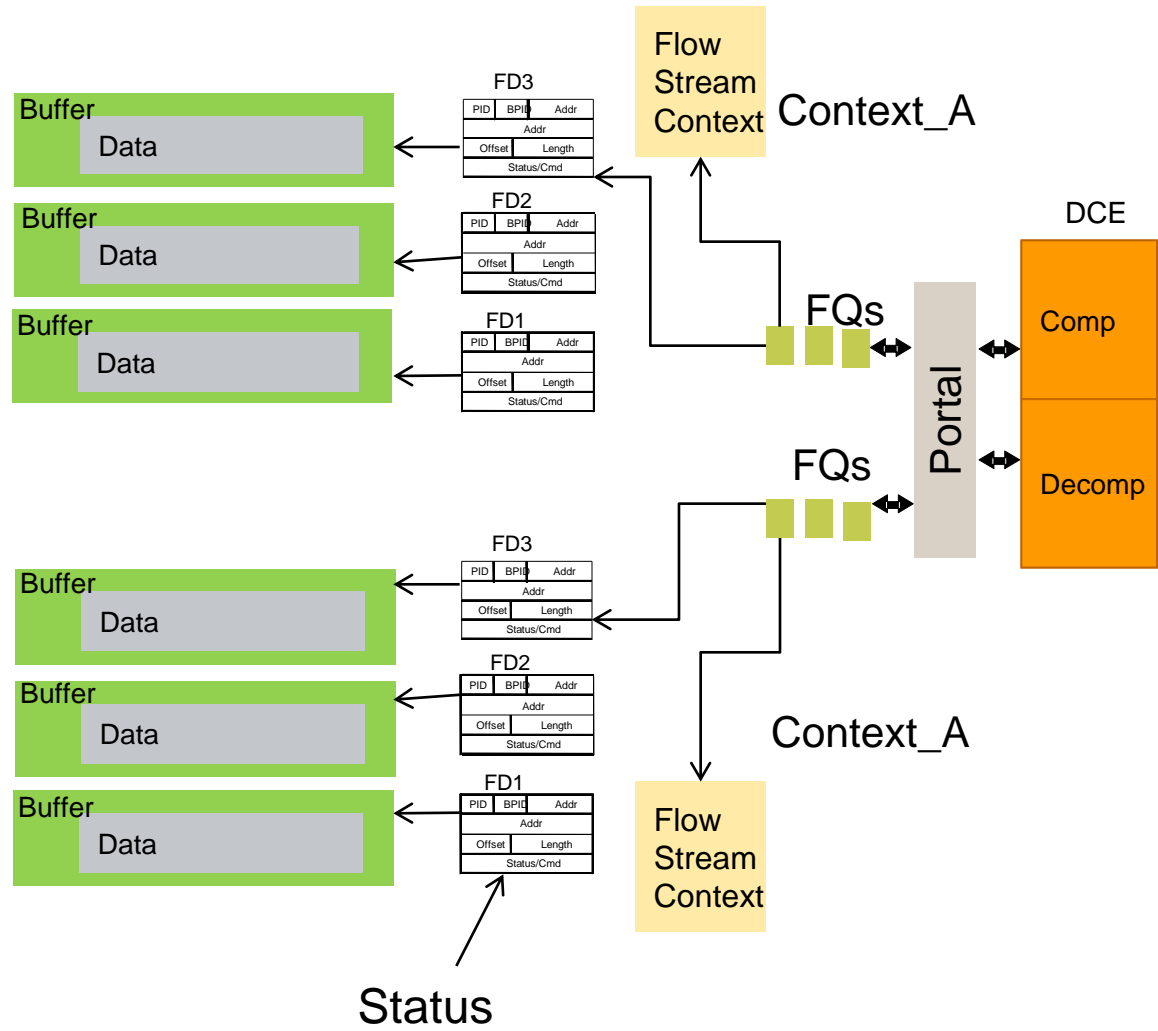
## ETS CoS-based Bandwidth Management



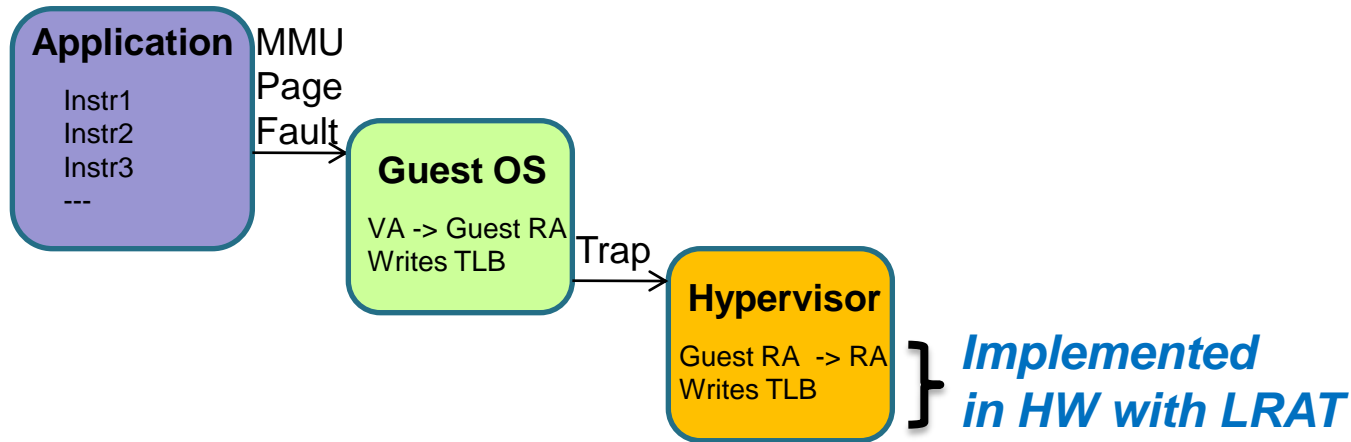
- **Enables Intelligent sharing of bandwidth between traffic classes control of bandwidth**
- **802.1Qaz**

# DCE Outputs

- DCE enqueues results to SW via Frame Queues as defined by FQ Context\_B field. When buffers obtained from BMan, buffer pool ID defined by Output FQ
- Each result is defined by a Frame Descriptor, which includes a Status field
- DCE updates flow stream context located at Context\_A as needed



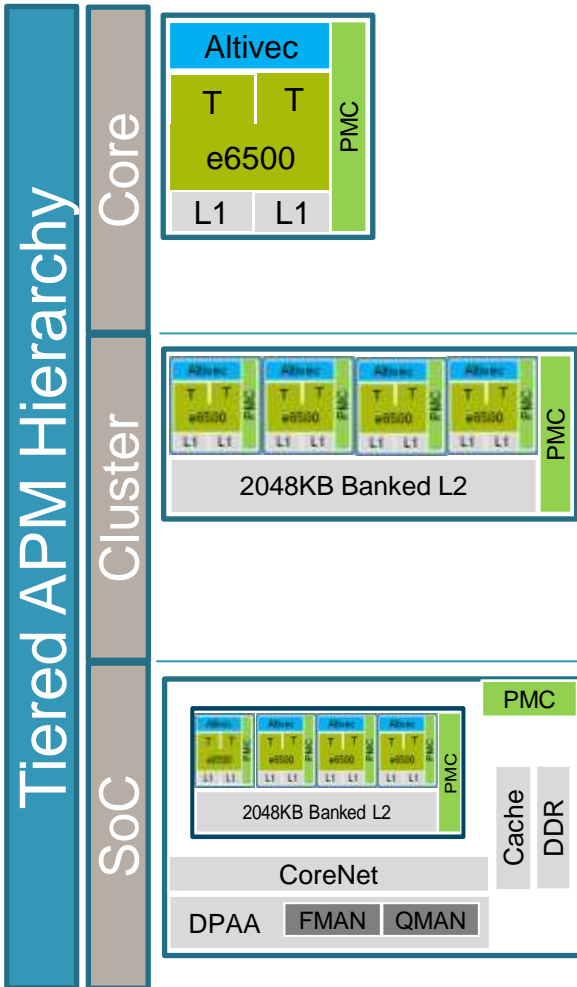
# LRAT: Hypervisor Performance Improvement



- Addition of Logical to Real Address Translation in hardware
- Benefits systems where multiple applications run on multiple OSeS running on the hypervisor
- Removes the hypervisor penalty associated with TLB faults
- Performance Improvement
  - Expect 10-15% performance increase for normal workloads
  - Greater improvement expected for benchmarks like stream or Imbench



# T2 Advanced Power and Energy Management

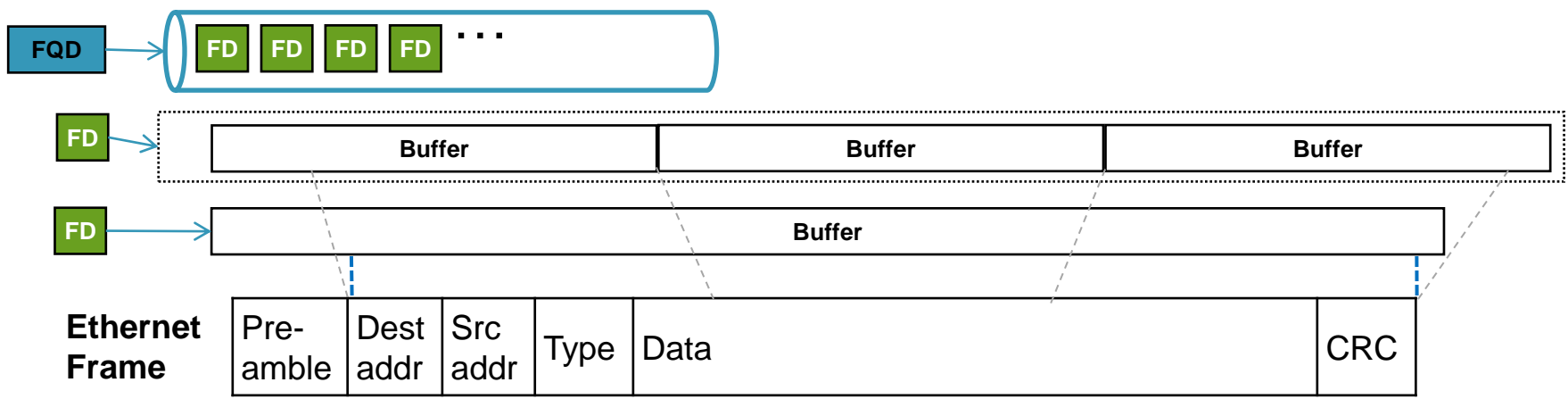


- Run, Doze, Nap
  - Wait
  - AltiVec Drowsy
    - Auto and SW controlled – maintain state
  - Core Drowsy
    - Auto and SW controlled – maintain state
  - Dynamic Clock gating
- 
- Run, Nap
    - Cores and L2
  - Dynamic Frequency Scaling (DFS) of the Cluster
  - Drowsy Cluster (cores)
  - Dynamic clock gating
- 
- SoC Sleep with state retention
  - SoC Sleep with RST
  - Cascade Power Management
  - Self Refresh
  - Dynamic clock gating
  - Energy Efficient Ethernet (EEE)

# DPAA Terminology

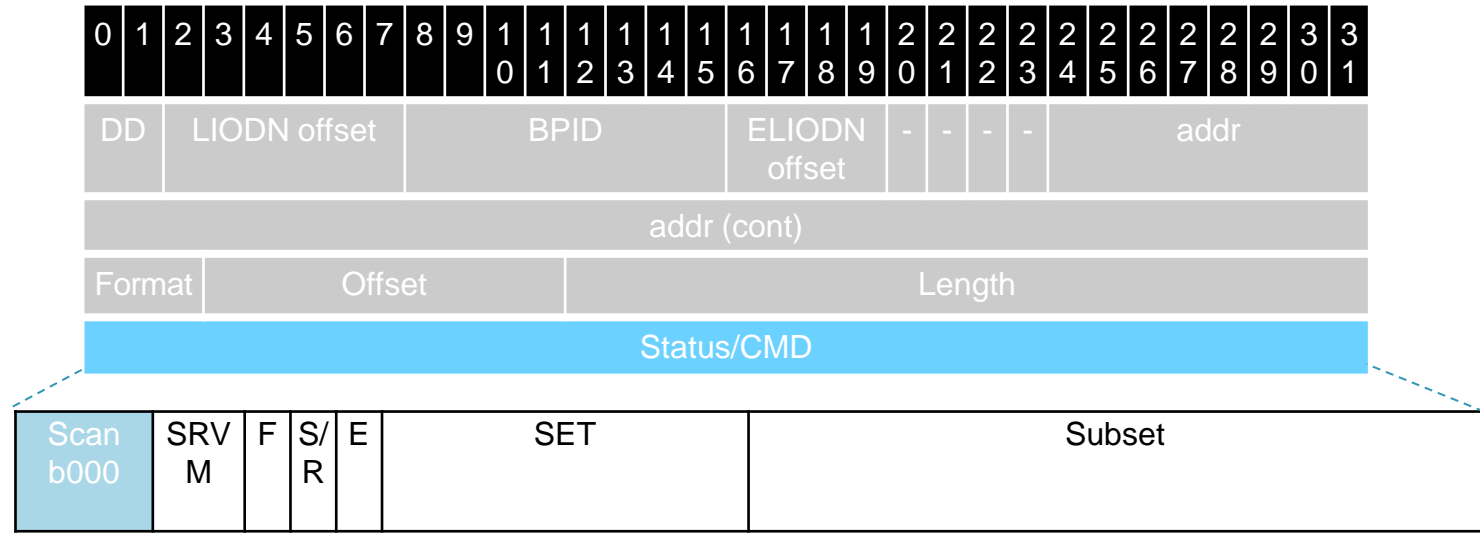
- Buffer: Unit of contiguous memory, allocated by software
- **Frame**: Buffer(s) that hold a data element (generally a packet)
  - Frames can be single buffers or multiple buffers (scatter/gather lists)
    - A “simple frame” has one delimited data element
    - A “multi-buffer frame” has two or more data elements
- **Frame Descriptor (FD)**: Proxy structure used to represent frames
- **Frame Queue**:
  - FIFO of related Frames Descriptor.(e.g. TCP session)
  - The basic queuing structure supported by QMan

**Frame Queue Descriptor (FQD)**: Structure used to manage Frame Queues



# Frame Descriptor: STATUS/CMD Treatment

- PME Frame Descriptor Commands
  - b111 NOP                      NOP Command
  - b101 FCR                      Flow Context Read Command
  - b100 FCW                      Flow Context Write Command
  - b001 PMTCC                  Table Configuration Command
  - b000 **SCAN**                Scan Command

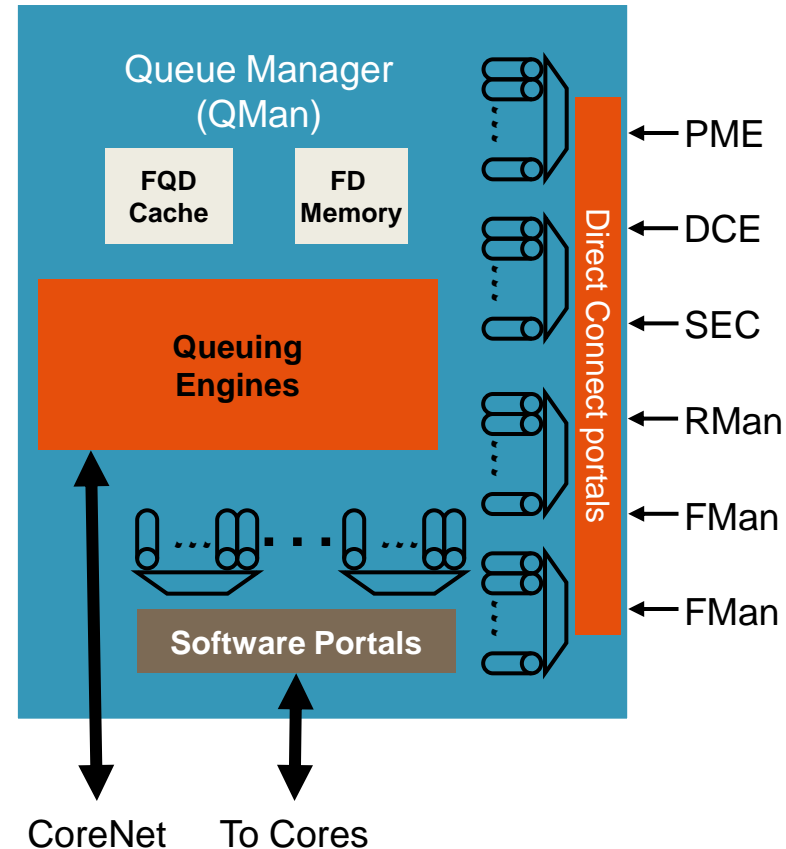


# Data Center Bridging (DCB) Overview

- QMan 1.2 (e.g. QorIQ T208x) supports Data Center Bridging (DCB)
- DCB refers to a series of inter-related IEEE specifications collectively designed to enhance Ethernet LAN traffic prioritization and congestion management
- DCB can be used in:
  - Between data center network nodes
  - LAN/network traffic
  - Storage Area Network (SAN) [e.g. Fiber Channel (loss sensitive)]
  - IPC traffic [e.g. Infiniband (low latency)]
- The DPAA is compliant with the following DCB specifications (traffic management related) :
  - IEEE Std. 802.1Qbb: Priority-based flow control (PFC)
    - To avoid frame loss, PFC Pause frames can be sent autonomously by HW
  - IEEE Std. 802.1Qaz: Enhanced transmission selection (ETS)
    - Support weighted bandwidth fairness
  - IEEE 802.1Qau: Quantized Congestion Notification (QCN)
    - End-to end congestion control mechanism

# Queue Management

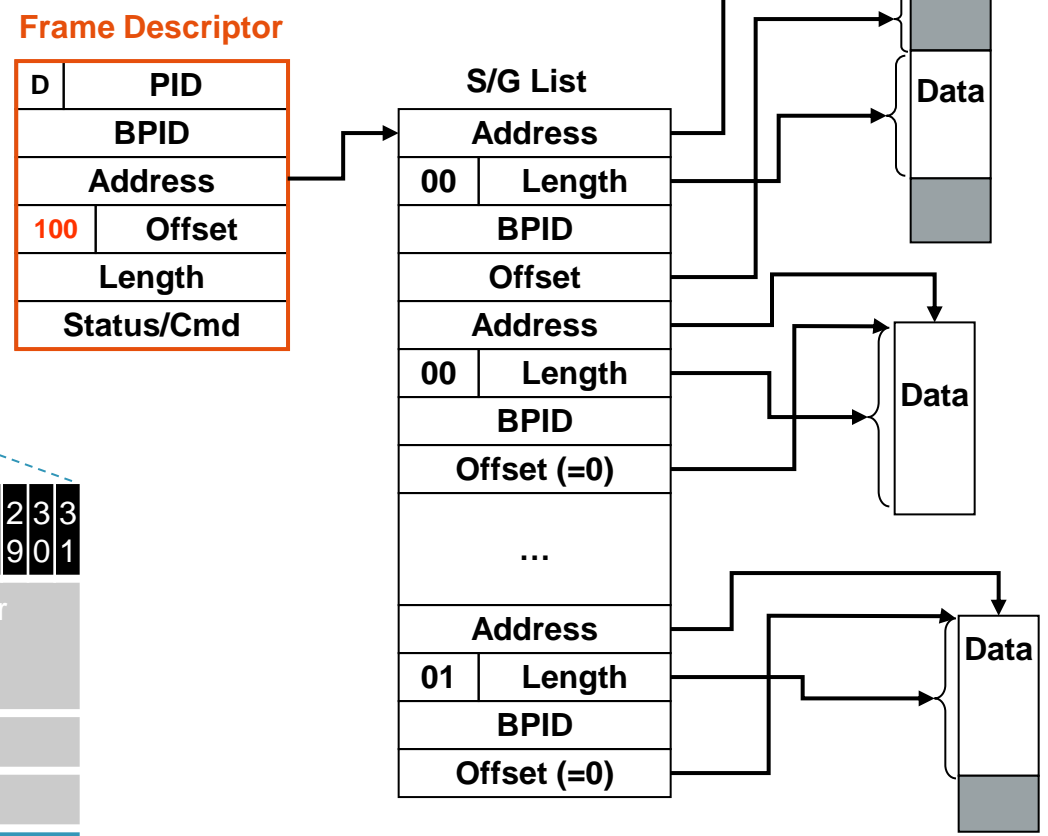
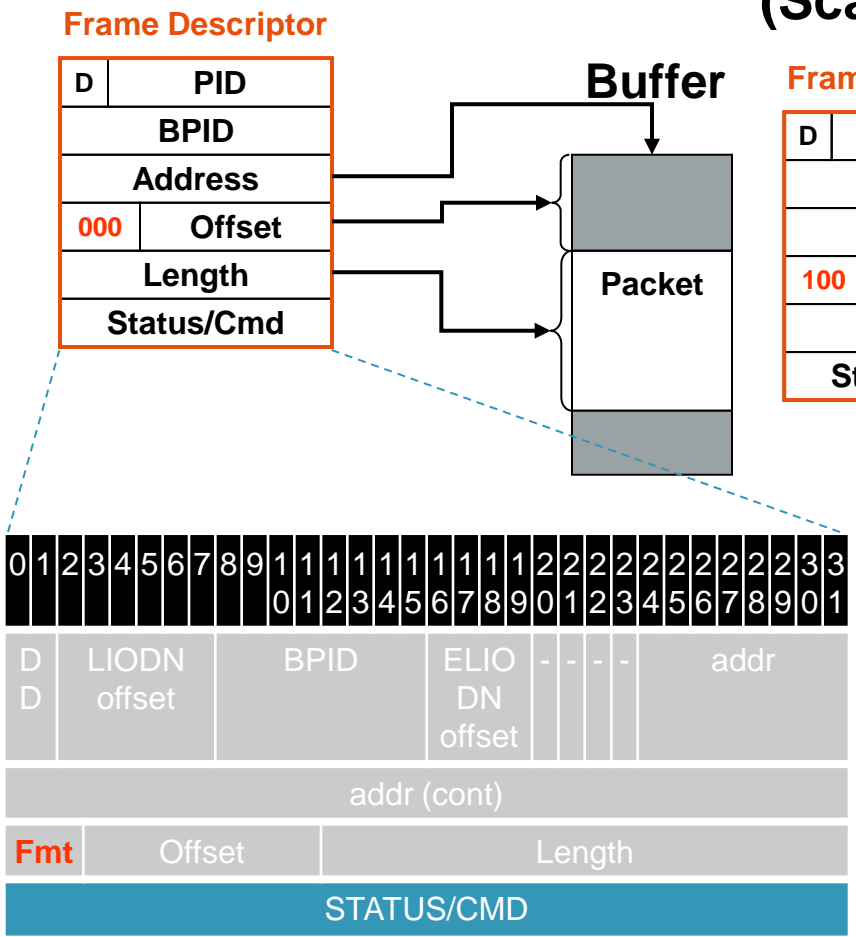
- QMan provides a way to inter-connect DPAA components
  - Cores (including IPC)
  - Hardware offload accelerators
  - Network interfaces – Frame Manager
- Queue management
  - High performance interfaces (“portals”) for enqueue/dequeue
  - Internal buffering of queue/frame data to enhance performance
- Congestion avoidance and management
  - RED/WRED
  - Tail drop for single queues and aggregates of queues
  - Congestion notification for “loss-less” flow control
- Load spreading across processing engines (cores, HW accelerators)
  - Order restoration
  - Order preservation/atomicity
- Delivery to cache/HW accelerators of per queue context information with the data (Frames)
  - This is an important offload for software using hardware accelerators



# DPAA Building Block: Frame Descriptor (FD)

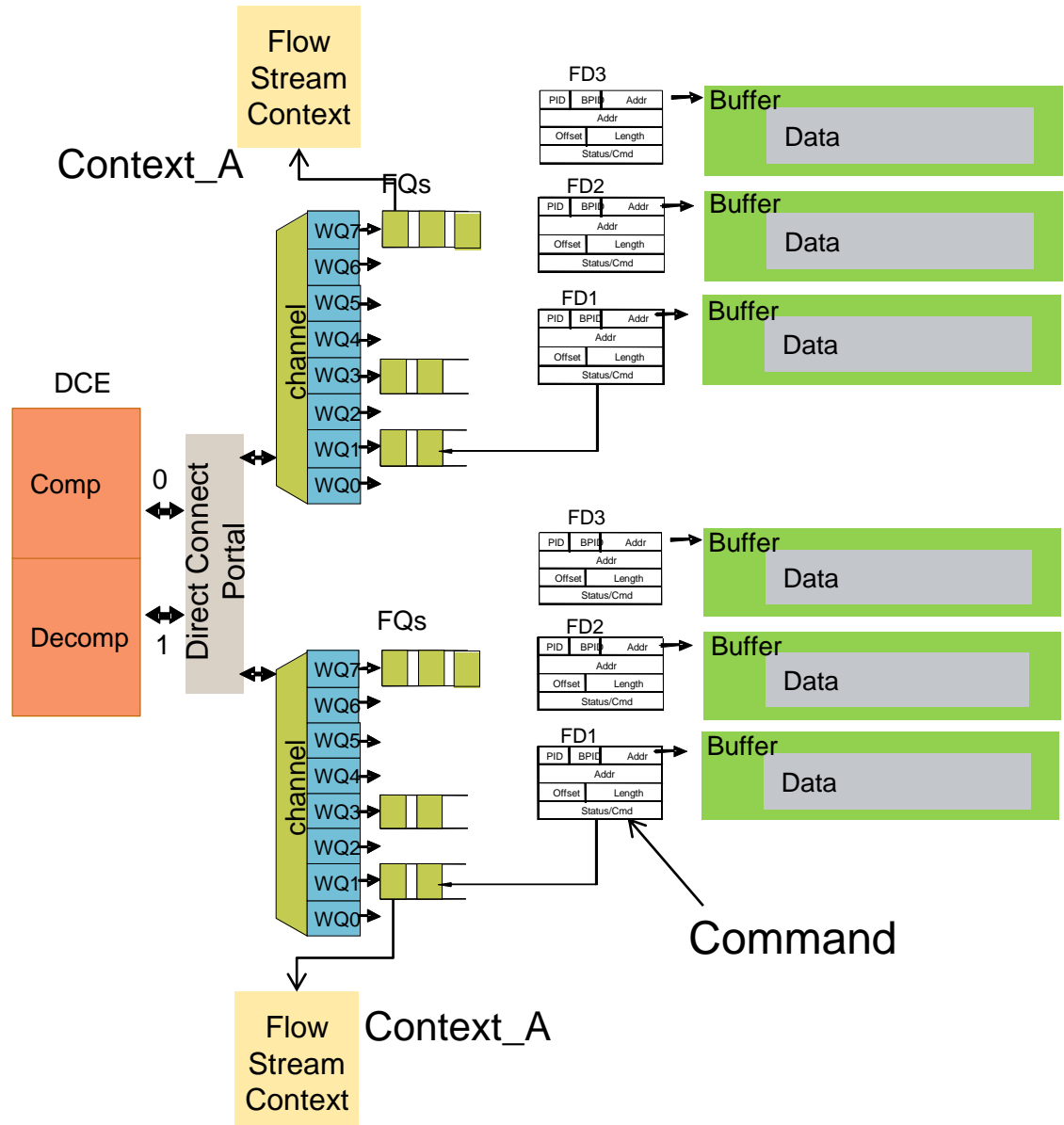
## Simple Frame

## Multi-buffer Frame (Scatter/gather)



# DCE Inputs

- SW enqueues work to DCE via Frame Queues. FQs define the flow for stateful processing
- FQ initialization creates a location for the DCE to use when storing flow stream context
- Each work item within the flow is defined by a Frame Descriptor, which includes length, pointer, offsets, and commands
- DCE has separate channels for compress and decompress





# e500mc/e6500 Caching Structure Differences

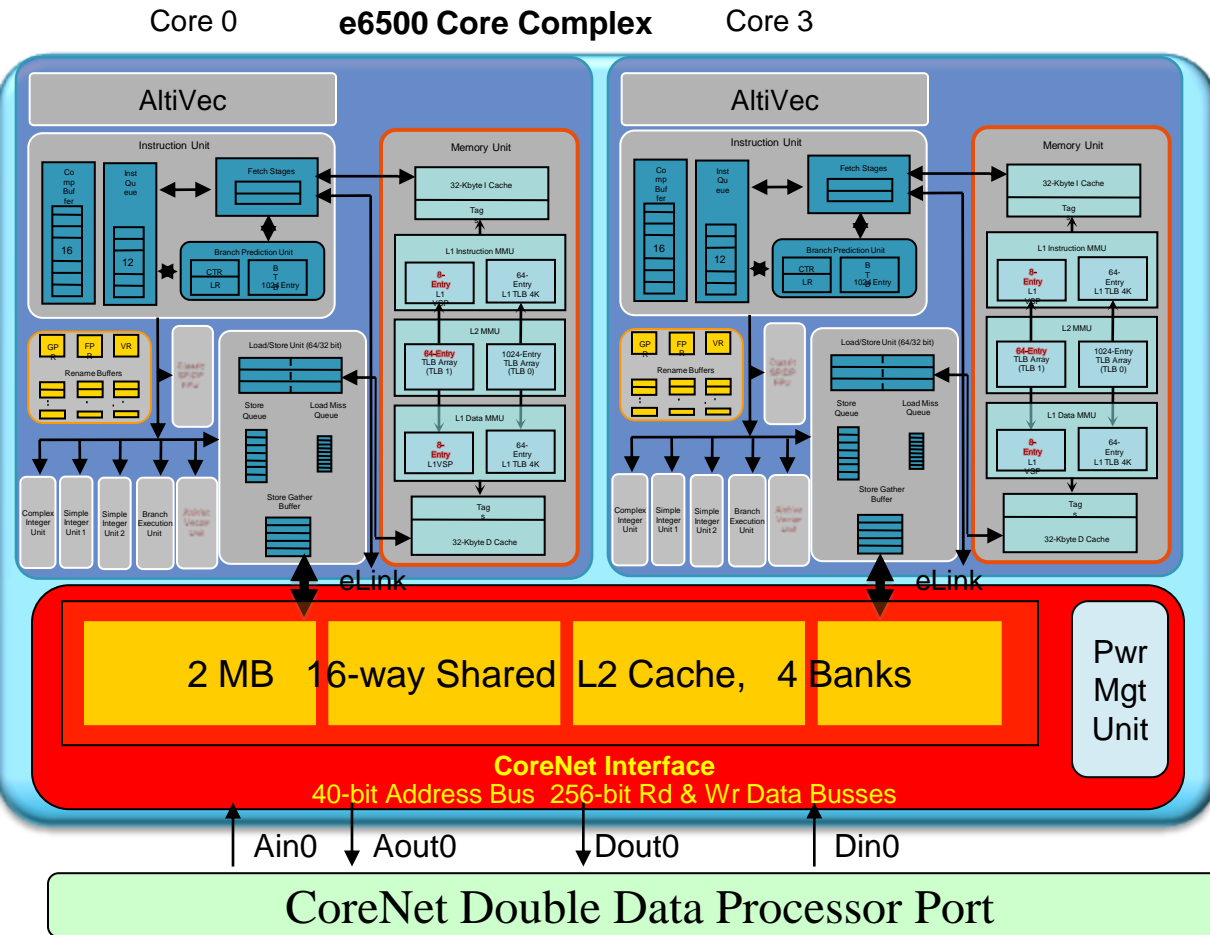
|    | e500mc                  | e6500                    | Implication  |
|----|-------------------------|--------------------------|--|
| L1 | 32kB. Can lock per core | 32kB. Can lock per core. | e6500 doesn't lock per thread.   |
| L2 | 128kB per core          | 2MB shared               | There will be a somewhat different latency profile, overall improved for e6500 |
| L3 | 1MB                     | 512kB                    |  |

- Cache changes are transparent to user application
- L1 locking is less granular in e6500





# NXP e5500 Core Complex



- 64-bit Power Architecture
- 28 nm Technology
- e5500 core features plus:
- Shared L2 in cluster of 4 cores
  - 2 MB 16-way, 4 Banks
  - Scalable from 128 KB-4 MB
  - High-performance eLink bus between coreLd/St and instruction fetch units
- Power
  - Drowsy core and caches
  - Power Mgt Unit
  - Wait-on-reservation instruction
- Enhanced MP Performance
  - Accelerated Atomic Operations
  - Optimized Barrier Instructions
  - Fast intra-cluster sharing
- AltiVec SIMD Unit
- CoreNet BIU
  - 256-bit Din and Dout data busses
- 40-bit Real Address
  - 1 Terabyte physical address space
- LRAT
  - Logical to Real Address translation mechanism for improved hypervisor performance



# P3041 vs T2080 Core Compatibility

## e6500 and e500mc Compatibility

- User code runs equally well on e6500 or e500mc
  - Interrupts per thread
  - Soft reset per thread (hard reset per core only)
  - Debug state per thread
- Changes are hidden by OS
  - L2 initialization uses a different register
  - Cache locking controlled differently
- P4080 SDK, emulated for e6500, didn't require changes
- Additional enablement for new features not present on e500mc: 64b, drowsy power manager, Altivec

## P3041 vs. T2080 DPAA Differences

- API enables minimal changes moving from P3041 to T2080
- SDK running on P3041 can be running with no changes for T2080
- Changes required to take advantage of new features:
  - Data compression engine
  - Storage Profiles
  - Data Center Bridging
  - Traffic Management
- Other
  - 8x GE sourced by single FMan on T2080 sources vs. single Fman on P3041

# Ethernet Termination

Ethernet enhancements compared to P2041/P3041:

- Storage Profile selection (up to 32 Profiles per port) based on classification. Where storage profile contains
  - LIODN offset
  - Up to four buffer pools per Storage Profile
  - Buffer Start margin/End margin configuration
  - S/G disable
  - Flow control configuration.
- IEEE802.3az (Energy Efficient Ethernet)
- IEEE802.3bf (Time sync)
- TX confirmation/error queue enhancements
  - Ability to configure separate FQID for normal confirmations vs errors
- Separate FD status for Overflow and physical error
- Egress Shaping (Definition in process)
- T2080 supports Datacenter Bridging
  - Priority Flow Control (PFC, IEEE 802.1Qbb)
  - Enhanced Transmission Selection (ETS, IEEE 802.1Qaz)
  - Data Center Bridging Exchange Notification (DCBX, currently part of IEEE 802.1Qaz, leverages 802.1AB (LLDP))

