Chipsmall Limited consists of a professional team with an average of over 10 year of expertise in the distribution of electronic components. Based in Hongkong, we have already established firm and mutual-benefit business relationships with customers from,Europe,America and south Asia,supplying obsolete and hard-to-find components to meet their specific needs.

With the principle of “Quality Parts,Customers Priority,Honest Operation,and Considerate Service”,our business mainly focus on the distribution of electronic components. Line cards we deal with include Microchip,ALPS,ROHM,Xilinx,Pulse,ON,Everlight and Freescale. Main products comprise IC,Modules,Potentiometer,IC Socket,Relay,Connector.Our parts cover such applications as commercial,industrial, and automotives areas.

We are looking forward to setting up business relationship with you and hope to provide you with the best service and solution. Let us make a better world for our industry!

Contact us

Tel: +86-755-8981 8866 Fax: +86-755-8427 6832
Email & Skype: info@chipsmall.com Web: www.chipsmall.com
Address: A1208, Overseas Decoration Building, #122 Zhenhua RD., Futian, Shenzhen, China

Freescale Semiconductor
Product Brief

Document Number: P5020PB
Rev. 1, 02/2013

P5020 QorIQ Communications Processor Product Brief

This product brief provides an overview of the P5020 QorIQ communications processor features as well as application use cases.

The P5020 combines two Power Architecture® processor cores with high-performance datapath acceleration logic and network and peripheral bus interfaces required for control processing in applications such as routers, switches, internet access devices, firewall and other packet filtering processors, network attached storage, storage area networks, imaging and general-purpose embedded computing. Its high level of integration offers significant performance benefits and greatly helps to simplify board design.

Contents

1 P5020 Application Use Cases

1.1 Router Control Processor

The following figure shows the P5020 in a linecard control plane application, where the linecard is part of a high-end network router.

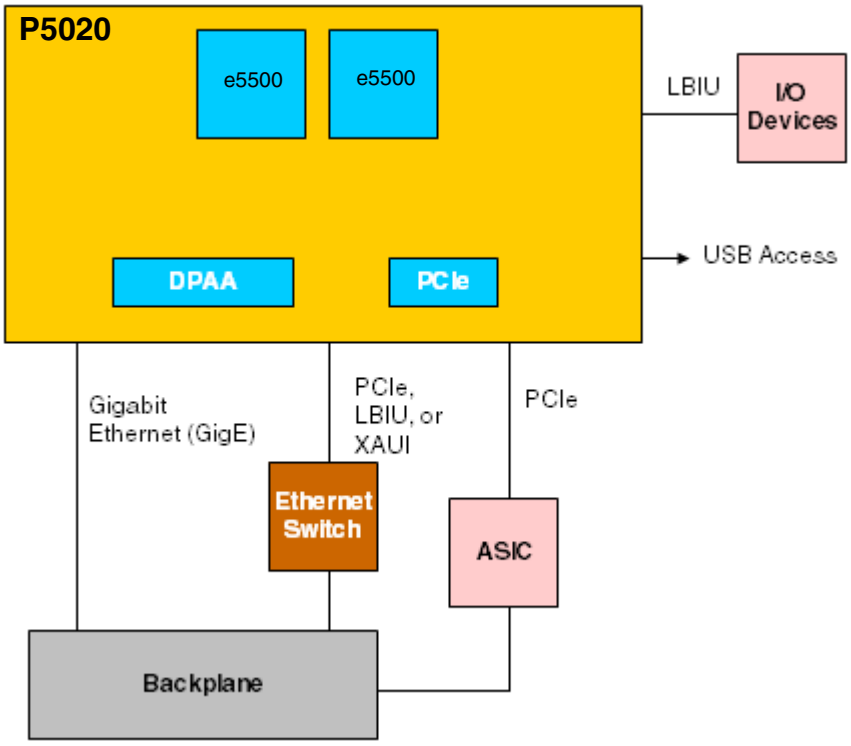


Figure 1. Control Plane Processor for a Router

1.2 DSP Farm Control Processor

The following figure shows a DSP farm enabled by the P5020 utilizing serial RapidIO.

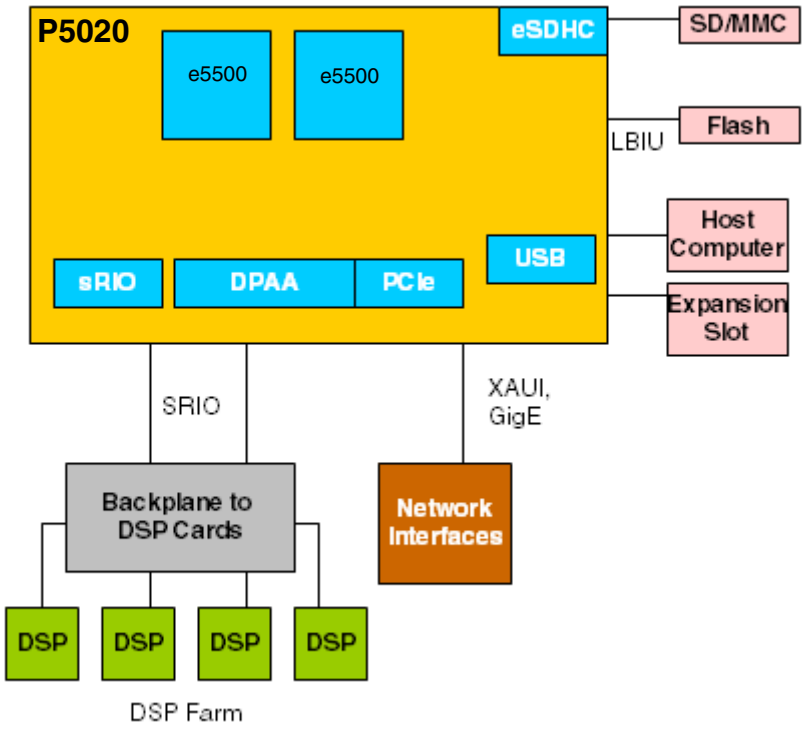


Figure 2. Control Plane Processor for a DSP Farm

1.3 SAN RAID 6 Controller

The following figure shows a RAID-enabled Disk Array Controller in an redundant active-active system for block-oriented storage systems. The P5020 Data Path Acceleration Architecture (DPAA) accelerates RAID 5/6 calculations and low-overhead data movement while optionally supporting data-at rest encryption and Data Integrity Field support.

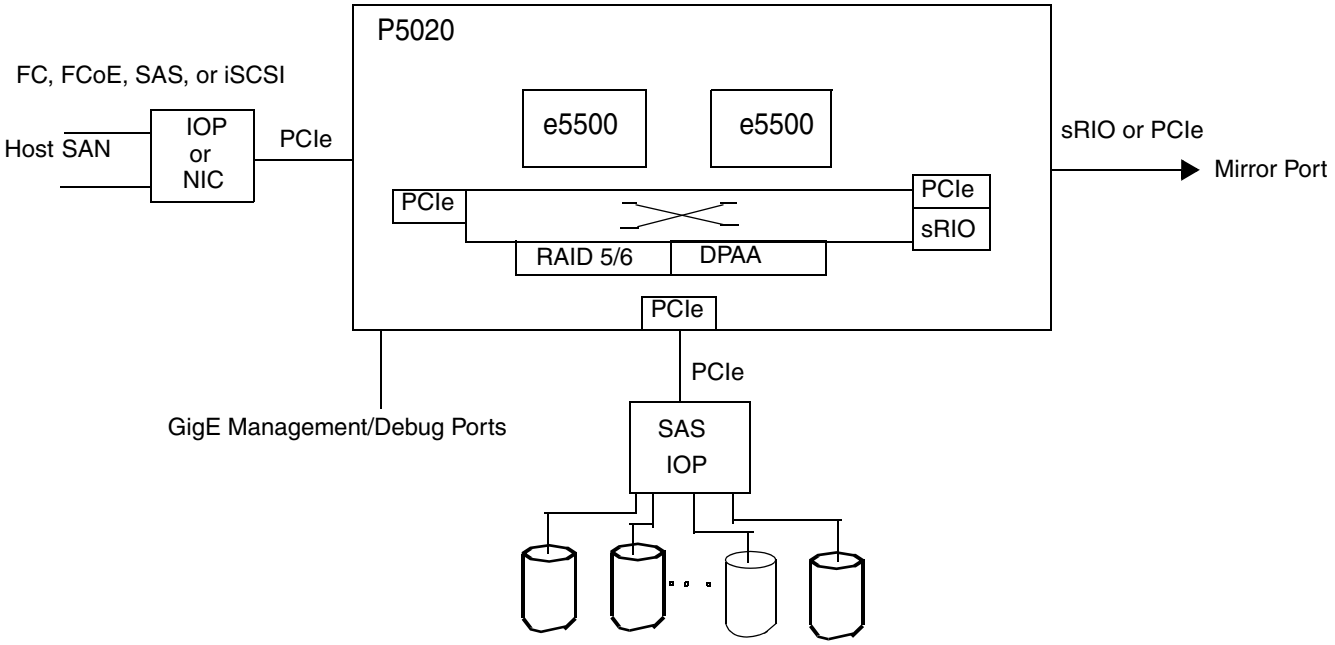


Figure 3. SAN RAID Controller

2 P5020 Dual-Core Processing Options

The device cores can run either on an OS or run OS-less using a simple scheduler.

2.1 Running on an OS

There are different multi-processing options with the device cores running on an OS:

- Symmetric multi-processing
- Cooperative asymmetric multi-processing
  - — Two copies of the same OS that are non-SMP enabled
  - — Two separate operating systems

2.2 Running OS-Less Using a Simple Scheduler

It is also possible for one or more cores to run OS-less, using a simple scheduler. This is a likely scenario when cores are performing datapath operations with bounded real-time requirements. This use case is greatly enhanced by the provisioning of a 512-Kbyte private back-side L2 cache for each e5500 core. These caches can operate as a traditional unified cache, or be set to operate as instruction only, data only, or even locked and used as memory-mapped SRAM.

CPU cores operating asymmetrically can be run at asynchronous clock rates. Each processor can source its input clock from one of the multiple PLLs inside the P5020. This allows each core to operate at the minimum frequency required to perform its assigned function, saving power. The cores are also capable of running at half and quarter ratios of their input PLL frequency and can switch between PLLs and ratios

nearly instantaneously. This allows lightly utilized CPUs to be slowed (under software control) for power savings, rather than performing more complex task migration operations.

3 Features

3.1 Block Diagram

The following figure shows the major functional units within the P5020.

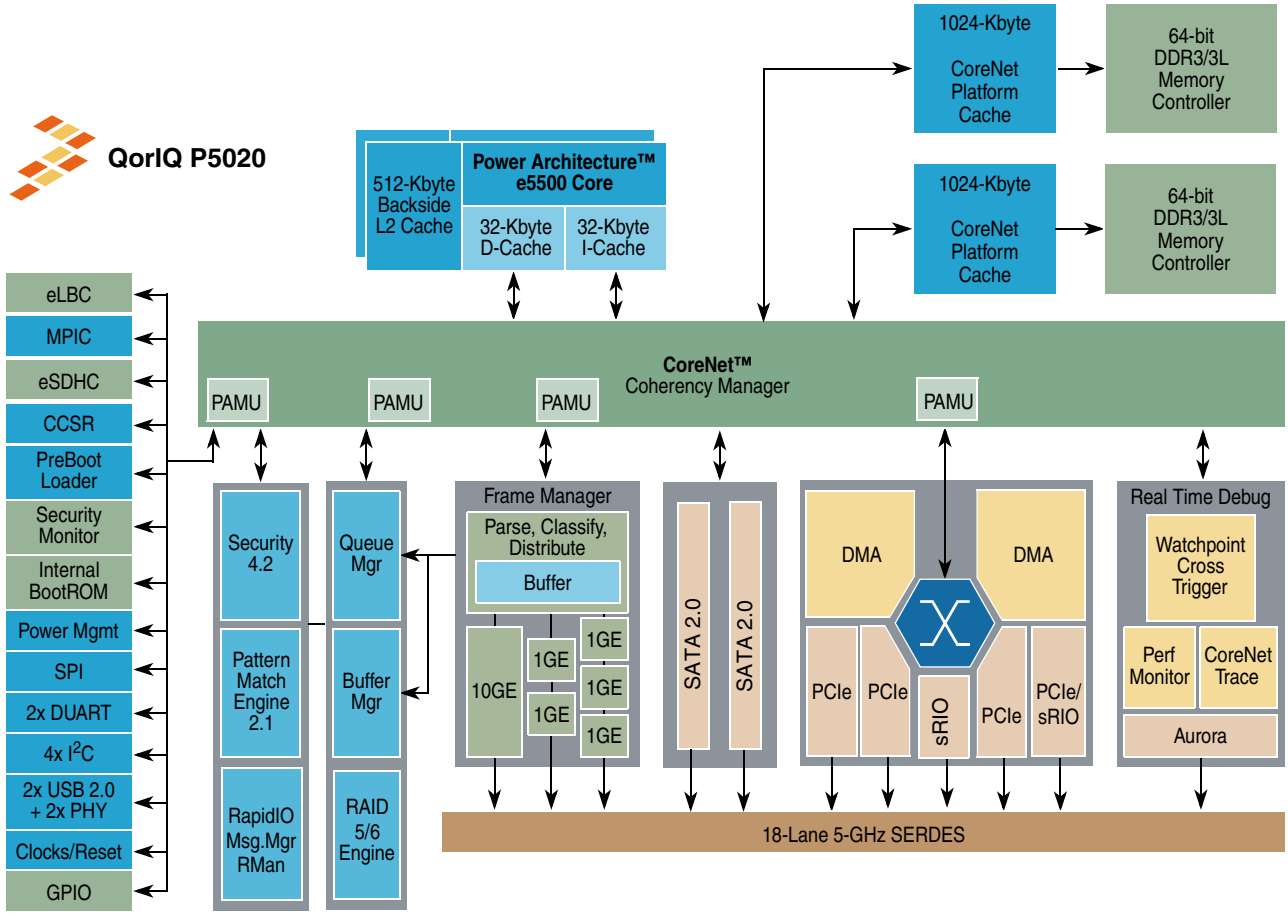


Figure 4. P5020 Preliminary Block Diagram

3.2 P5020 Features Summary

The P5020 SoC includes the following functions and features:

- Two e5500 cores built on Power Architecture technology, each with a private 512-Kbyte private backside cache
  - Up to 2 GHz
  - Three levels of instructions:
    - User

  - Supervisor
  - Hypervisor
- Independent boot and reset
- Secure boot capability

- Two 1-Mbyte shared CoreNet platform cache (CPC)
- Hierarchical interconnect fabric
  - CoreNet fabric supporting coherent and non-coherent transactions with prioritization and bandwidth allocation amongst CoreNet end-points
  - Queue manager fabric supporting packet-level queue management and quality of service scheduling
- Two 64-bit DDR3/3L SDRAM memory controllers with ECC and interleaving support
- Datapath acceleration architecture (DPAA) incorporating acceleration for the following functions:
  - Packet parsing, classification, and distribution
  - Queue management for scheduling, packet sequencing, and congestion management
  - Hardware buffer management for buffer allocation and de-allocation
  - Encryption/decryption (SEC 4.2)
  - RegEx pattern matching (PME 2.1)
  - RapidIO™ messaging manager (RMan)
  - RAID5/6 Engine
    - Support for XOR and Galois Field parity calculation
    - Support for data protection information (DPI)
- Ethernet interfaces
  - One 10 Gbps Ethernet (XAUI) controller
  - Five 1 Gbps or four 2.5 Gbps Ethernet controllers
- High speed peripheral interfaces
  - Four PCI Express 2.0 controllers/ports running at up to 5 GHz
  - Two serial RapidIO 2.0 controllers/ports (version 1.3 with features of 2.1) running at up to 5 GHz with Type 11 messaging and Type 9 data streaming support
- Additional peripheral interfaces
  - Dual SATA supporting 1.5 and 3.0 Gb/s operation
  - Two USB 2.0 controllers with integrated PHY
  - SD/MMC controller (eSDHC)
  - Enhanced SPI controller
  - Four I^2C controllers
  - Two Dual DUARTs
  - Enhanced local bus controller (eLBC)
- 18 SerDes lanes to 5 GHz
- Multicore Programmable Interrupt Controller (MPIC)

- Two 4-channel DMA engines

3.3 P5020 Benefits

The P5020's e5500 cores can be combined as a fully-symmetric, multi-processing, system-on-a-chip, or they can be operated with varying degrees of independence to perform asymmetric multi-processing. Full processor independence, including the ability to independently boot and reset each e5500 core, is a defining characteristic of the device. The ability of the cores to run different operating systems, or run OS-less, provides the user with significant flexibility in partitioning between control, datapath, and applications processing. It also simplifies consolidation of functions previously spread across multiple discrete processors onto a single device.

3.4 Data Path Acceleration Architecture (DPAA) Benefits

While the two Power Architecture cores offer a major leap in available processor performance in many throughput-intensive, packet-processing networking applications, raw processing power is not enough to achieve multi-Gbps data rates. To address this, the P5020 uses Freescale's Data Path Acceleration Architecture (DPAA) (see Section 3.9, "Data Path Acceleration Architecture (DPAA)"), which significantly reduces data plane instructions per packet, enabling more CPU cycles to work on value-added services rather than repetitive low-level tasks. Combined with specialized accelerators for cryptography and pattern matching, the P5020 allows the user's software to perform complex packet processing at high data rates.

3.5 Critical Performance Parameters

The following table lists key performance indicators that define a set of values used to measure P5020 operation.

Table 1. P5020 Critical Performance Parameters

Indicator	Values(s)
Top speed bin core frequency	2.0 GHz
Maximum memory data rate	1.3 GHz (DDR3/3L)[1] • 1.5-V for DDR3 • 1.35-V for DDR3L
Local bus	• 3.3 V • 2.5 V • 1.8 V
Operating junction temperature range	0–105 C
Package	1295-pin FC-PBGA (flip-chip plastic ball grid array)

Notes:
1 Conforms to JEDEC standard

3.6 e5500 Core and Cache Memory Complex

Each e5500 is a superscalar dual issue processor, supporting out-of-order execution and in-order completion, which allows the Power Architecture e5500 to perform more instructions per clock than other RISC and CISC architectures.

3.6.1 e5500 Core Features

- Up to 2.0 GHz core clock speed
- 36 bit physical addressing
- 64 TLB SuperPages
- 512-entry, 4-Kbyte pages front end
- 3 Integer Units: 2 simple, 1 complex (integer multiply and divide)
- 64-byte cache line size
- L1 caches, running at same frequency of CPU
  - 32-Kbyte Instruction, 8-way
  - 32-Kbyte Data, 8-way
  - Both with data and tag parity protection
- Supports data path acceleration architecture (DPAA) data and context “stashing” into the L1 data cache and the backside L2 cache
- User, supervisor, and hypervisor instruction level privileges
- New processor facilities
  - Hypervisor APU
  - Classic double precision floating point unit
    - Uses 32 64-bit floating-point registers (FPRs) for scalar single- and double-precision floating-point arithmetic
    - Replaces the embedded floating-point facility (SPE) implemented on the e500v1 and e500v2
    - Designed to comply with IEEE Std. 754™1985 FPU for both single- and double-precision operations
  - “Decorated Storage” APU for improved statistics support
    - Provides additional atomic operations, including a “fire-and-forget” atomic update of up to two 64-bit quantities by a single access
  - Expanded interrupt model
    - Improved programmable interrupt controller (PIC) automatically ACKs interrupts
    - Implements message send and receive functions for interprocessor communication, including receive filtering
  - External PID load and store facility
    - Provides system software with an efficient means to move data and perform cache operations between two disjoint address spaces

– Eliminates the need to copy data from a source context into a kernel context, change to destination address space, then copy the data to the destination address space or alternatively to map the user space into the kernel address space

3.6.2 512-Kbyte Private Backside Cache

- Each e5500 core features a 512-Kbyte private backside L2 cache running at the same frequency of CPU. The caches support:Write Back, pseudo LRU replacement algorithm
- Tag parity and ECC data protection
- Eight-way, with arbitrary partitioning between instruction and data. For example, 3-ways instruction, 5-ways data, and so on.
- Supports direct stashing of datapath architecture data into cache

3.6.3 CoreNet Platform Cache (CPC)

The QorIQ P5020 also contains 2x1-Mbyte of shared CoreNet platform cache, with the following features:

- Configurable as write back or write through
- Pseudo LRU replacement algorithm
- ECC protection
- 64-byte coherency granule
- Two cache line read 1024 bits per cycle at 800 MHz, 32-way cache array configurable to any of several modes on a per-way basis
  - — Unified cache, I-only, D-only
  - — I/O stash (configurable portion of each packet copied to CPC on write to main memory)
    - – Stashing of all transactions and sizes supported
    - – Explicit (CoreNet signalled) and implicit (address range based) stash allocation
  - — Addressable SRAM (32-Kbyte granularity)

3.6.4 CoreNet Fabric and Address Map

The CoreNet fabric is Freescale's next generation Interconnect Standard for multicore products, and provides the following:

- A highly concurrent, fully cache coherent, multi-ported fabric
- 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 800 MHz
- The CoreNet fabric has been designed to overcome bottlenecks associated with shared bus architectures, particularly address issue and data bandwidth limitations. The P5020's multiple, parallel address paths allow for high address bandwidth, which is a key performance indicator for large coherent multicore processors
- Eliminates address retries, triggered by CPUs being unable to snoop within the narrow snooping window of a shared bus. This results in the device having lower average memory latency

The 36-bit, physical address map consists of local space and external address space. For the local address map, 32 local access windows (LAWs) define mapping within the local 36-bit (64-Gbyte) address space. Inbound and outbound translation windows can map the device into a larger system address space such as the RapidIO or PCIe 64-bit address environment. This functionality is included in the address translation and mapping units (ATMUs).

3.6.5 Memory Complex

The P5020 memory complex consists of the two DDR controllers for main memory, and the memory controllers associated with the enhanced local bus controller (eLBC).

3.6.5.1 DDR Memory Controllers

The two DDR memory controllers have the following functionalities:

- Supports DDR3/3L SDRAM. The P5020 also supports chip-select interleaving within a controller. The memory interface controls main memory accesses and together the two controllers support a maximum of 64 Gbytes of main memory.
- Supports interleaving across controllers on bank, page, or cache line boundaries.
- The P5020 can be configured to retain the currently active SDRAM page for pipelined burst accesses. Page mode support of up to 64 simultaneously open pages can dramatically reduce access latencies for page hits. Depending on the memory system design and timing parameters, page mode can save up to 10 memory clock cycles for subsequent burst accesses that hit in an active page.
- Using ECC, the P5020 detects and corrects all single-bit errors and detects all double-bit errors and all errors within a nibble.
- Upon detection of a loss of power signal from external logic, the DDR controllers can put compliant DDR SDRAM DIMMs into self-refresh mode, allowing systems to implement battery-backed main memory protection.
- Supports initialization bypass feature for use by system designers to prevent re-initialization of main memory during system power-on after an abnormal shutdown.
- Supports active zeroization of system memory upon detection of a user-defined security violation.

3.6.6 PreBoot Loader (PBL) and Nonvolatile Memory Interfaces

The PreBoot Loader (PBL) is a new logic module that operates similarly to an I^2C boot sequencer but on behalf of a larger number of interfaces.

The PBL's functions include the following:

- Simplifies boot operations, replacing pin strapping resistors with configuration data loaded from nonvolatile memory.
- Uses the configuration data to initialize other system logic and to copy data from low speed memory interfaces (I^2C, eLBC, SPI, and SD/MMC) into fully initialized DDR or the 2-Mbyte CPC.
- Releases CPU 0 from reset, allowing the boot processes to begin from fast system memory.

The nonvolatile memory interfaces accessible by the PBL are as follows:

- The eLBC may be accessed by software running on the CPUs following boot; it is not dedicated to the PBL. It also can be used for both volatile (SRAM) and nonvolatile memory as well as a control and low-performance data port for external memory-mapped P5020s. See Section 3.6.7, "Enhanced Local Bus Controller."
- The serial memory controllers may be accessed by software running on the CPUs following boot; they are not dedicated to the PBL. See Section 3.6.7.1, "Serial Memory Controllers."

3.6.7 Enhanced Local Bus Controller

The enhanced local bus controller (eLBC) port connects to a variety of external memories, DSPs, and ASICs.

Key features of the eLBC include the following:

- Multiplexed 32-bit address and 32-bit data bus operating at up to 93 MHz
- Eight chip selects for eight external slaves
- Up to eight-beat burst transfers
- 8-, 16-, or 32-bit port sizes controlled by an internal memory controller
- Three protocol engines on a per-chip-select basis
- Parity support
- Default boot ROM chip select with configurable bus width (8-, 16-, or 32-bit)
- Support for parallel NAND and NOR flash

Three separate state machines share the same external pins and can be programmed separately to access different types of devices. Some examples are as follows:

- The general-purpose chip-select machine (GPCM) controls accesses to asynchronous devices using a simple handshake protocol.
- The user-programmable machine (UPM) can be programmed to interface to synchronous devices or custom ASIC interfaces.
- The NAND flash control machine (FCM) further extends interface options.
- Each chip select can be configured so that the associated chip interface is controlled by the GPCM, UPM, or FCM controller.

All controllers can be enabled simultaneously. The eLBC internally arbitrates among the controllers, allowing each to read or write a limited amount of data before allowing another controller to use the bus.

3.6.7.1 Serial Memory Controllers

In addition to the parallel NAND and NOR flash supported by means of the eLBC, the P5020 supports serial flash using SPI and SD/MMC/eMMC card. The SD/MMC/eMMC controller includes a DMA engine, allowing it to move data from serial flash to external or internal memory following straightforward initiation by software.

3.7 Universal Serial Bus (USB) 2.0

The two USB 2.0 controllers with integrated PHY provide point-to-point connectivity complying with the USB specification, Rev. 2.0. Each USB controller can be configured to operate as a stand-alone host, and USB #2 can be configured as a stand-alone device, or with both host and device functions operating simultaneously.

Key features of the USB 2.0 controller include the following:

- Compatible with USB specification, Rev. 2.0
- Supports full-speed (12 Mbps), and low-speed (1.5 Mbps) operations
- Supports the required signaling for the USB transceiver macrocell interface (UTMI).The PHY interfacing to the UTMI is an internal PHY.
- Both controllers support operation as a stand-alone USB host controller
  - Support USB root hub with one downstream-facing port
  - Enhanced host controller interface (EHCI)-compatible
- One controller supports operation as a stand-alone USB device
  - Supports one upstream-facing port
  - Supports six programmable USB endpoints

The host and device functions are both configured to support all four USB transfer types:

- Bulk
- Control
- Interrupt
- Isochronous

3.8 High-Speed Peripheral Interface Complex

All high-speed peripheral interfaces connect via 18 lanes of 5-GHz SerDes to a common crossbar switch referred to as OCeaN. Two high-speed I/O interface standards are supported: PCI Express (PCIe), and Serial RapidIO (sRIO). The P5020 integrates the following:

- Four PCIe controllers
- Two Serial RapidIO controllers
- RapidIO message manager (RMan).

3.8.1 PCI Express Controllers

Each of the four PCIe interfaces is compliant with the *PCI Express Base Specification Revision 2.0*. Key features of the PCIe interface include the following:

- Power-on reset configuration options allow root complex or endpoint functionality.
- The physical layer operates at 2.5 or 5 Gbaud data rate per lane.
- Receive and transmit ports operate independently, with an aggregate theoretical bandwidth of 32 Gbps.

- x8, x4, x2, and x1 link widths supported
- Both 32- and 64-bit addressing and 256-byte maximum payload size
- Full 64-bit decode with 36-bit wide windows
- Inbound INTx transactions
- Message Signaled Interrupt (MSI) transactions

3.8.2 Serial RapidIO

The Serial RapidIO interface is based on the *RapidIO Interconnect Specification, Revision 1.3,* with features from 2.1. RapidIO is a high-performance, point-to-point, low-pin-count, packet-switched system-level interconnect that can be used in a variety of applications as an open standard. The rich feature set includes high data bandwidth, low-latency capability, and support for high-performance I/O devices as well as message-passing and software-managed programming models. Receive and transmit ports operate independently, and with 2 x 4 Serial RapidIO controllers, the aggregate theoretical bandwidth is 32 Gbps.

Key features of the Serial RapidIO interface unit include the following:

- Support for *RapidIO Interconnect Specification, Revision 1.3* (all transaction flows and priorities)
- 1x, 2x, and 4x LP-serial link interfaces, with transmission rates of 2.5, 3.125, or 5.0 Gbaud (data rates of 2.0, 2.5, or 4.0 Gbps) per lane.
- Auto-detection of 1x, 2x, or 4x mode operation during port initialization
- 34-bit addressing and up to 256-byte data payload
- Support for SWRITE, NWRITE, NWRITE_R and Atomic transactions
- Receiver-controlled flow control
- RapidIO error injection
- Internal LP-serial and application interface-level loopback modes

3.8.2.1 RapidIO Message Manager (RMan)

The key features of the RapidIO message manager (RMan) include the following:

- Manages two inbox/outbox mailboxes (queues) for data and one doorbell message structure
- Can multi-cast a single-segment 256-byte message to up to 32 different destination DevIDs
- Has four outbound segmentation units supporting RapidIO Type 5–6 and Type 8–11

3.8.3 Serial ATA (SATA) 2.0 Controllers

The key features of each of the two SATA include the following:

- Designed to comply with Serial ATA 2.6 Specification
- Supports host SATA I per spec Rev 1.0a
  - OOB
  - Port multipliers
  - ATAPI 6+

  - Spread spectrum clocking on receive
- Support for SATA II extensions
  - Asynchronous notification
  - Hot plug including asynchronous signal recovery
  - Link power management
  - Native command queuing
  - Staggered spin-up and port multiplier support
- Support for SATA I and II data rates (1.5 and 3.0 Gbaud)
- Standard ATA master-only emulation
- Includes ATA shadow registers
- Implements SATA superset registers (SError, SControl, SStatus)
- Interrupt driven
- Power management support
- Error handling and diagnostic features
  - Far end/near end loopback
  - Failed CRC error reporting
  - Increased ALIGN insertion rates
  - Scrambling and CONT override

3.9 Data Path Acceleration Architecture (DPAA)

The DPAA provides the infrastructure to support simplified sharing of networking interfaces and accelerators by multiple CPU cores. These resources are abstracted into enqueue/dequeue operations by means of a common DPAA Queue Manager (QMan) driver. Beyond enabling multicore resource sharing, the DPAA significantly reduces software overheads associated with high-touch packet-forwarding operations. Examples of the types of packet-processing services this architecture is optimized to support are as follows:

- Traditional routing and bridging
- Firewall
- VPN termination for both IPsec and SSL VPNs
- Intrusion detection/prevention (IDS/IPS)
- Network anti-virus (AV)

The DPAA generally leaves software in control of protocol processing, while reducing CPU overheads through off-load functions, which fall into two, broad categories:

- Packet Distribution and Queue/Congestion Management
- Accelerating Content Processing

3.9.1 Packet Distribution and Queue/Congestion Management

The following table lists some packet distribution and queue/congestion management offload functions.

Table 2. Offload Functions

Function Type	Definition
Data buffer management	Supports allocation and deallocation of buffers belonging to pools originally created by software with configurable depletion thresholds. Implemented in a module called the Buffer Manager (BMan).
Queue management	Supports queuing and quality-of-service scheduling of frames to CPUs, network interfaces and DPAA logic blocks, maintains packet ordering within flows. Implemented in a module called the Queue Manager (QMan). The QMan, besides providing flow-level queuing, is also responsible for congestion management functions such as RED/WRED, congestion notifications and tail discards.
Packet distribution	Supports in-line packet parsing and general classification to enable policing and QoS-based packet distribution to the CPUs for further processing of the packets. This function is implemented in the block called the Frame Manager (FMan).
Policing	Supports in-line rate-limiting by means of two-rate, three-color marking (RFC 2698). Up to 256 policing profiles are supported. This function is also implemented in the FMan.

3.9.2 Accelerating Content Processing

Properly implemented acceleration logic can provide significant performance advantages over most optimized software with acceleration factors on the order of 10–100x. Accelerators in this category typically touch most of the bytes of a packet (not just headers). To avoid consuming CPU cycles in order to move data to the accelerators, these engines include well-pipelined DMAs. The following table lists some specific content-processing accelerators on the P5020.

Table 3. Content-Processing Accelerators

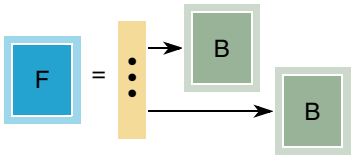
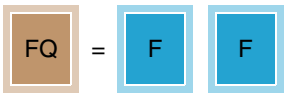
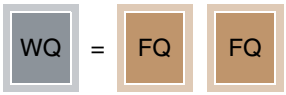
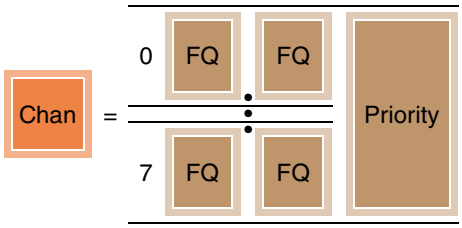
Interface	Definition
SEC 4.2	Crypto-acceleration for protocols such as IPsec, SSL, and 802.16
PME 2.1	Regex style pattern matching for unanchored searches, including cross-packet stateful patterns

Note: Prior versions of the SEC and PME are integrated into multiple members of the PowerQUICC and QorIQ family. Both of these engines have been enhanced to work within the DPAA, and also upgraded in both features and performance.

3.9.3 DPAA Terms and Definitions

The following table lists common DPAA terms and their definitions.

Table 4. DPAA Terms and Definitions

Term	Definition	Graphic Representation
Buffer	Region of contiguous memory, allocated by software, managed by the DPAA BMan	
Buffer pool	Set of buffers with common characteristics (mainly size, alignment, access control)	
Frame	Single buffer or list of buffers that hold data, for example, packet payload, header, and other control information	
Frame queue (FQ)	FIFO of frames	
Work queue (WQ)	FIFO of FQs	
Channel	Set of eight WQs with hardware provided prioritized access	
Dedicated channel	Channel statically assigned to a particular end point, from which that end point can dequeue frames. End point may be a CPU, FMan, PME, or SEC.	—
Pool channel	A channel statically assigned to a group of end points, from which any of the end points may dequeue frames.	

3.9.4 Major DPAA Components

The Data Path Acceleration Architecture (DPAA) includes the following major components:

- Section 3.9.4.1, “Frame Manager (FMan)
- Section 3.9.4.2, “Queue Manager (QMan)
- Section 3.9.4.3, “Buffer Manager (BMan)
- Section 3.9.4.6, “RapidIO Message Manager (RMan)

- Section 3.9.4.4, “Security Engine (SEC 4.2)
- Section 3.9.4.5, “Pattern Matching Engine (PME 2.1)

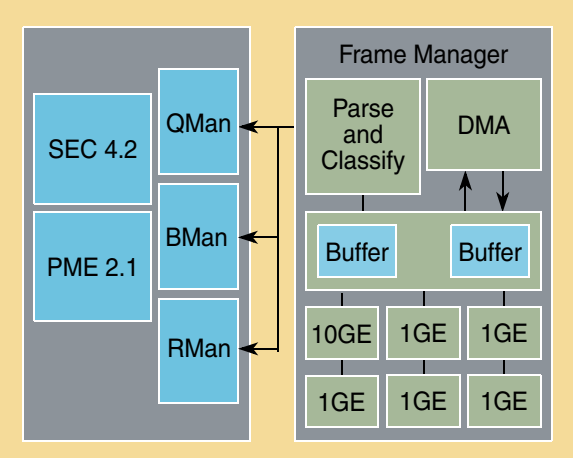


Figure 5. QorIQ Data Path Acceleration Architecture (DPAA)

3.9.4.1 Frame Manager (FMan)

The Frame Manager (FMan) combines the Ethernet network interfaces with packet distribution logic to provide intelligent distribution and queuing decisions for incoming traffic. This integration allows the FMan to perform configurable parsing and classification of the incoming frame with the purpose of selecting the appropriate input frame queue for expedited processing by a CPU or pool of CPUs.

3.9.4.1.1 FMan Network Interfaces

The FMan integrates five data path, tri-speed Ethernet controllers (dTSECs) and one 10-Gbit Ethernet controller.

Note that the more basic parsing and filing capability found in prior PowerQUICC eTSECs is removed from the MACs themselves, and aggregated in the more flexible and robust parsing and classification logic described in Section 3.9.4.1.2, “FMan Parse Function.”

The Ethernet controllers support the following:

- Programmable CRC generation and checking
- RMON statistics
- Jumbo frames of up to 9.6 Kbytes

They are designed to comply with IEEE Std 802.3®, IEEE 802.3u, IEEE 802.3x, IEEE 802.3z, IEEE 802.3ac, IEEE 802.3ab, and additionally the 1Gbps MACs support IEEE-1588 v2 (clock synchronization over Ethernet).

The dTSECS are capable of full- and half-duplex Ethernet support (1000 Mbps supports only full duplex); the 10-Gbit MAC is a single-speed full duplex. It supports IEEE 802.3 full-duplex flow control (automatic PAUSE frame generation or software-programmed PAUSE frame generation and recognition).

When all SERDES are otherwise allocated, it is possible to enable two of dTSECs by means of RGMII or RMII physical interfaces.

3.9.4.1.2 FMan Parse Function

The primary function of the packet parse logic is to identify the incoming frame for the purpose of determining the desired treatment to apply. This parse function can parse many standard protocols, including options and tunnels, and supports a generic configurable capability to allow proprietary or future protocols to be parsed.

There are several types of parser headers, shown in the following table.

Table 5. Parser Header Types

Header Type	Definition
Self-describing	Announced by proprietary values of Ethertype, protocol identifier, next header, and other standard fields. They are self-describing in that the frame contains information that describes the presence of the proprietary header.
Non-self-describing	Does not contain any information that indicates the presence of the header. For example, a frame that always contains a proprietary header before the Ethernet header would be non-self-describing. Both self-describing and non-self-describing headers are supported by means of parsing rules in the FMan.
Proprietary	Can be defined as being self-describing or non-self-describing

The underlying notion is that different frames may require different treatment, and only through detailed parsing of the frame can proper treatment be determined.

Parse results can (optionally) be passed to software.

3.9.4.1.3 FMan Distribution and Policing

After parsing is complete, there are two options for treatment (see Table 6).

Table 6. Post-Parsing Treatment Options

Treatment	Function	Benefits
Hash	• Hashes selected fields in the frame as part of a spreading mechanism • The result is a specific frame queue identifier. • To support added control, this FQID can be indexed by values found in the frame, such as TOS or p-bits, or any other desired field(s).	Useful when spreading traffic while obeying QoS constraints is required
Classification look-up	• Looks up certain fields in the frame to determine subsequent action to take, including policing • The FMan contains internal memory that holds small tables for this purpose. • The user configures the sets of lookups to perform, and the parse results dictate which one of those sets to use. • Lookups can be chained together such that a successful look-up can provide key information for a subsequent look-up. After all the look-ups are complete, the final classification result provides either a hash key to use for spreading, or a FQ ID directly.	• Useful when hash distribution is insufficient and a more detailed examination of the frame is required • Can determine whether policing is required and the policing context to use

Key benefits of the FMan policing function are as follows:

- Because the FMan has up to 256 policing profiles, any frame queue or group of frame queues can be policed to either drop or mark packets if the flow exceeds a preconfigured rate.
- Policing and classification can be used in conjunction for mitigating Distributed Denial of Service Attack (DDOS).
- The policing is based on two-rate-three-color marking algorithm (RFC2698). The sustained and peak rates as well as the burst sizes are user-configurable. Hence, the policing function can rate-limit traffic to conform to the rate the flow is mapped to at flow set-up time. By prioritizing and policing traffic prior to software processing, CPU cycles can be focused on the important and urgent traffic ahead of other traffic.

3.9.4.2 Queue Manager (QMan)

The Queue Manager (QMan) is the main component in the DPAA that allows for simplified sharing of network interfaces and hardware accelerators by multiple CPU cores. It also provides a simple and consistent message and data passing mechanism for dividing processing tasks amongst multiple CPU cores. The QMan features are as follows:

- Common interface between software and all hardware
  - Controls the prioritized queuing of data between multiple processor cores, network interfaces, and hardware accelerators
  - Supports both dedicated and pool channels, allowing both push and pull models of multicore load spreading
- Atomic access to common queues without software locking overhead
- Mechanisms to guarantee order preservation with atomicity and order restoration following parallel processing on multiple CPUs
- Two-level queuing hierarchy with one or more Channels per Endpoint, eight work queues per Channel, and numerous frame queues per work queue
- Priority and work conserving fair scheduling between the work queues and the frame queues
- Lossless flow control for ingress network interfaces
- Congestion avoidance (RED/WRED) and congestion management with tail discard and up to 256 congestion groups with each group composed of a user-configured number of frame queues.

3.9.4.3 Buffer Manager (BMan)

The buffer manager (BMan) manages pools of buffers on behalf of software for both hardware (accelerators and network interfaces) and software use. The BMan features are as follows:

- Common interface for software and hardware
- Guarantees atomic access to shared buffer pools
- Supports 32 buffer pools. Software and hardware buffer consumers can request both different size buffers and buffers in different memory partitions.
- Supports depletion thresholds with congestion notifications
- On-chip per pool buffer stockpile to minimize access to memory for buffer pool management
- LIFO (last in first out) buffer allocation policy that optimizes cache usage and allocation

3.9.4.4 Security Engine (SEC 4.2)

The SEC 4.2 is QorIQ's fourth generation crypto-acceleration engine. In addition to off-loading cryptographic algorithms, the SEC 4.2 offers header and trailer processing for several established security protocols. The SEC 4.2 includes several Descriptor Controllers (DECOs), which are updated versions of the previous SEC crypto-channels. DECOs are responsible for header and trailer processing, and managing context and data flow into the CHAs assigned to it for the length of an operation.

The DECOs can perform header and trailer processing, as well as single pass encryption/integrity checking for the following security protocols:

- IPsec
- SSL/TLS
- SRTP
- IEEE Std 802.1AE™MACSec
- IEEE 802.16e WiMax MAC layer
- 3GPP RLC encryption/decryption

In prior versions of the SEC, the individual algorithm accelerators were referred to as Execution Units (EUs). In the SEC 4.2, these are referred to as Crypto Hardware Accelerators (CHAs) to distinguish them from prior implementations. Specific CHAs available to the DECOs are listed below.

- Advanced encryption standard unit (AESA)
- ARC four execution unit (AFHA)
- Cyclic redundancy check accelerator (CRCA)
- Data encryption standard execution unit (DESA)
- Kasumi execution unit (KFHA)
- SNOW 3 G hardware accelerator (STHA)
- Message digest execution unit (MDHA)
- Public key execution unit (PKHA)
- Random number generator (RNGB)

Depending on the security protocol and specific algorithms, the SEC 4.2's aggregate symmetric encryption/integrity performance is 5 Gbps, while asymmetric encryption (RSA public key) performance is ~5,000 1024b RSA operations per second.

The SEC 4.2 is also part of the QorIQ Trust Architecture, which gives the P5020 the ability to perform secure boot, runtime code integrity protection, and session key protection. The Trust Architecture is described in Section 3.10, "Avoiding Resource Contentions Using the QorIQ Trust Architecture."

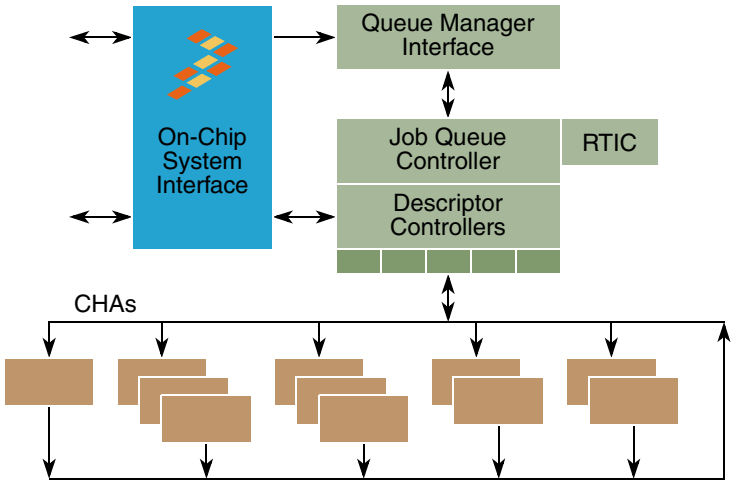


Figure 6. SEC 4.2 Block Diagram

3.9.4.5 Pattern Matching Engine (PME 2.1)

The PME is a self-contained hardware module capable of autonomously scanning data from streams for patterns that match a specification in a database dedicated to it. The PME 2.1 is an updated version of the PME used in previous members of the PowerQUICC family. Specific updates include the following:

- QMan interface supporting the DPAA Queue Interface Driver
- 2x increase in the number of patterns supported (16 Kbytes to 32 Kbytes)
- Increase in number of stateful rules supported (8 Kbytes to 16 Kbytes)
- Raw scanning performance is ~ 5 Gbps.

Patterns that can be recognized, or "matched," by the PME are of two general forms:

- Byte patterns are simple matches such as "abcd123" existing in both the data being scanned and in the pattern specification database.
- Event patterns are a sequence of multiple byte patterns. In the PME, event patterns are defined by stateful rules.

3.9.4.5.1 PME Regular Expressions (Regex)

The PME specifies patterns of bytes as regular expressions (regex). The P5020 (by means of an online or offline process) converts Regex patterns into the PME's pattern specification database. Generally, there is a one-to-one mapping between a regex and a PME byte pattern. The PME's use of regex pattern matching offers built-in case-insensitivity and wildcard support with no pattern explosion, while the PME's NFA-style architecture offers fast pattern database compilation and fast incremental updates. Up to 32,000 regex patterns are supported, each up to 128 bytes long. The 32,000 regex patterns can be combined by means of stateful rules to detect a far larger set of event patterns. Comparative compilations against DFA style regex engines have shown that 300,000 DFA pattern equivalents can be achieved with ~8000 PME regexes with stateful rules.

3.9.4.5.2 PME Match Detection

Within the PME, match detection proceeds in stages. The key element scanner performs initial byte pattern matching, with handoff to the data examination engine for elimination of false positives through more complex comparisons. As the name implies, the stateful rule engine receives confirmed basic matches from the earlier stages, and monitors a stream for addition for subsequent matches that define an event pattern.

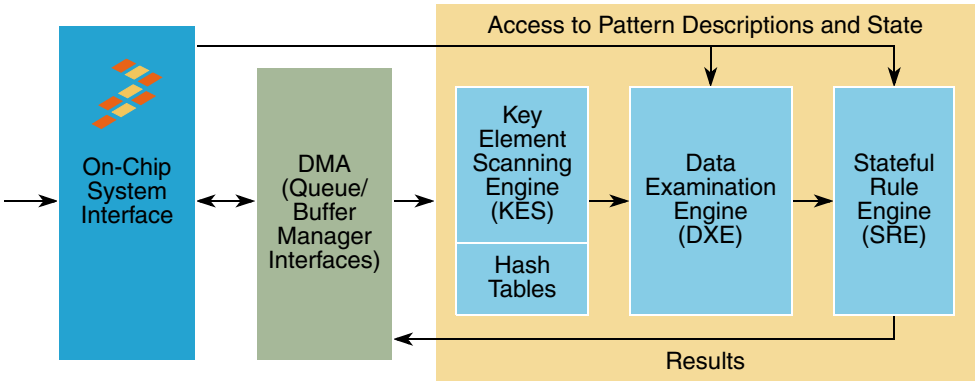


Figure 7. PME 2.1 Block Diagram

3.9.4.6 RapidIO Message Manager (RMan)

The RapidIO message manager (RMan) produces and consumes Type 8 Port-write, Type 9 Data Streaming, Type 10 Doorbells and Type 11 Messaging traffic and is capable of producing Type 5 NWRITE and Type 6 SWRITE transactions.

For inbound traffic, the RMan supports up to 17 open reassembly contexts as a arbitrary mix of Type 9, and Type 11 traffic.

As ingress packets arrives at the RMan, they are compared against up to 64 classification rules to determine the target queue. These rules support Type 8, 9, 10 and 11 transaction types. They may be wildcarded and are configured as masks over selected header fields. The following fields are maskable as part of each classification rule:

Transaction types:

- RapidIO port
- Source ID
- Destination ID
- Flow level

Type 9 messaging-specific fields:

- Class-of-service (CoS)
- StreamID

Type 11 messaging-specific fields:

- Mailbox

- Extended mailbox
- Letter

Should the packet remain unclassified, the traffic is retried with an error in the case of Type 10 and 11 traffic and dropped in the case of Type 9 traffic. Dropped traffic is logged and upon a threshold can assert an error interrupt.

Classification allows Type 9, 10 and 11 traffic to be distributed across 64 possible Frame queues. A single dedicated inbound Type 8 Port-write Frame queue is provided.

For all outbound traffic types (Type 8, 9, 10 and 11), the Datapath Acceleration Architecture allows a very large number of outbound Frame queues effectively limited by system, software and performance constraints.

3.9.4.7 RAID5/6 Engine

The P5020 includes a RAID5/6 Engine for storage applications, which significantly extends the capability and performance of earlier PowerQUICC RAID (XOR) functionality. The RAID5/6 Engine supports a variety of storage-related functions such as Move, Generate XOR, RAID 6 Parity, Fill and Check. The following table summarizes the functions supported by the engine.

Table 7. RAID5/6 Engine Supported Functions

Function	No. of Sources	No. of Destinations	Command Options	
			Scatter/Gather	DIF
No Op	—	—	—	—
Single Source Move	1	1	Y	N
Multicast Move	1	2	Y	N
Add DIF	1	1 or 2	Y	Y
Remove DIF	1	1 or 2	Y	Y
Update DIF	1	1 or 2	Y	Y
Generate Q Parity	2–16	1	Y	Y
Generate Q and Q Parity	2–16	2	Y	Y
Fill Pattern	—	1	Y	Y
Check Pattern	1	—	Y	Y
Fill LFSR	—	1	Y	N
Check LFSR	1	—	Y	N
Compare	2	—	Y	Y
Gather DIF	1	1	Y	Y

The RAID5/6 Engine supports commands with between 1 and 16 sources for relevant functions. A simple DMA move operation is supported along with a two-destination multicast move that duplicates the source data. Both of these simple operations are the foundation for commands that support Data Protection Information (DIF) insertion, updating and checking. A single RAID5/6 parity generate function is

supported which calculates Galois field (GF) based parity calculation for (where MULT = 1 performs simple XOR) up to 16 sources. A variant supports calculation of two GF multiplies for use in calculating XOR and RAID 6 Parity simultaneously without reading the input data twice. This command calculates two GF multiplications across the sources and writes them to two destinations. The GF primitive polynomial is programmable and thus supports common polynomials such as 0x11D and 0x14D.

In addition to classic storage acceleration, the RAID5/6 Engine provides some additional helpful functions including the ability to fill or check a region based on a 128-bit value, incrementing value or using a LSFR algorithm. A compare function is provided that compares two regions of memory and reports the result to a result queue.

The RAID5/6 Engine supports ANSI T10 Data Protection Information and is capable of checking, adding, removing and updating the Data Integrity Fields (DIF). All Reference and Application Tags seen during an operation may be set to an initial value or that value can be incremented as blocks are processed by the engine. Reference Tag, Application Tag can be configurable disabled/enabled from DIF function on per command basis. It also supports IP checksum-based guard generation and checking (RFC 793), in addition to the T10 CRC based guard.

3.10 Avoiding Resource Contentions Using the QorIQ Trust Architecture

Consolidation of discrete CPUs into a single, multicore SoC and potential repartitioning of legacy software on those cores introduces many opportunities for unintended resource contentions to arise, but the QorIQ Trust Architecture can reduce the risk of these issues.

3.10.1 QorIQ Trust Architecture Benefits

A system may exhibit erratic behavior if the multiple CPUs do not effectively partition and share system resources. While it can be challenging to prevent unintended resource contention, stopping malicious software is much more difficult. Device consolidation combined with a trend toward embedded systems becoming more open (or more likely to run third-party or open-source software on at least one of the cores) creates opportunities for malicious code to enter a system.

The P5020 offers a new level of hardware partitioning support, allowing system developers to ensure software running on any CPU only accesses the resources (memory, peripherals, etc.) that it is explicitly authorized to access. This may not seem like a challenge in an SMP environment, because the OS performs resource allocation for the applications running on it. However, it is a very difficult problem to overcome in AMP environments where there may be multiple instances of the same OS, or even different OSes running on the various CPU cores. Even OS protections in an SMP system may be insufficient in the presence of malicious software.

3.10.2 e5500 Core MMU and Embedded Hypervisor

The P5020's first line of defense against unintended interactions amongst the multiple CPUs/OSes is each core's MMU, which are configured to determine which addresses in the global address map the CPU is able to read or write. If a particular resource (such as a portion of memory, peripheral device, and so on) is dedicated to a single CPU, that CPU's MMU is configured to allow access to those addresses (on

4-Kbyte granularity); other CPU MMUs are not configured for access to the other CPU's private memory range. When two CPUs need to share resources, their MMUs are both configured so that they have access to the shared address range.

This level of hardware support for partitioning is common today, however, it is not sufficient for many core systems running diverse software. When the functions of multiple discrete CPUs are consolidated onto a single, multicore SoC, achieving strong partitioning should not require the developer to map functions onto cores that are the exclusive owners of specific platform resources. The alternative, a fully open system with no private resources, is also unacceptable. For this reason, the core MMU also includes embedded Hypervisor extensions.

Each core MMU supports three levels of instructions:

- User
- Supervisor (OS)
- Hypervisor: An embedded Hypervisor micro-kernel (provided by Freescale as source code) runs unobtrusively beneath the various OSes running on the CPUs, consuming CPU cycles only when an access attempt is made to an embedded Hypervisor-managed shared resource.The embedded Hypervisor determines whether the access should be allowed, and if so, proxies the access on behalf of the original requestor. If malicious or poorly tested software on any core attempts to overwrite important P5020 configuration registers (including CPU MMUs), the embedded Hypervisor blocks the write. Other examples of embedded Hypervisor managed resources are high- and low-speed peripheral interfaces (PCIe, UART) if those resources are not dedicated to a single CPU/partition.

3.10.3 Peripheral Access Management Unit (PAMU)

The P5020 includes a distributed function collectively referred to as the peripheral access management unit (PAMU), which provides address translation and access control for all bus masters in the system (PME, SEC, FMan, and so on). The PAMU access control can be one of the following:

- Absolute—The FMan, PME, SEC, and other bus masters can never access memory range XYZ.
- Conditional—Based on the Partition ID of the CPU that programmed the bus master

Being MMU-based, the embedded Hypervisor is only able to stop unauthorized software access attempts. Internal components with bus mastering capability also need to be prevented from reading and writing to specific memory regions. These devices do not spontaneously generate access attempts, but, if programmed to do so by buggy or malicious software, any of them could overwrite sensitive configuration registers and crash the system.

3.10.4 Secure Boot and Sensitive Data Protection

The core MMUs and PAMU allow the device to enforce a consistent set of memory access permissions on a per-partition basis. When combined with embedded Hypervisor for safe sharing of resources, the P5020 becomes highly resilient when poorly tested or malicious code is run. For system developers building high reliability/high security platforms, rigorous testing of code of known origin is the norm.