



FTF | FREESCALE TECHNOLOGY FORUM
POWERING INNOVATION

Leveraging the Architectural Enhancements of Freescale's 64-bit Power Architecture® Core

FTF-NET-F0394

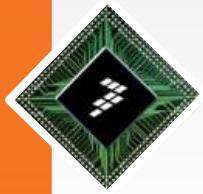
Chuck Corley & Dac Pham
Applications Engr; Core Design Mgr



June 2011

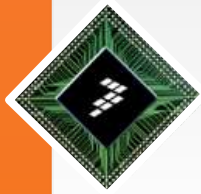
Freescale, the Freescale logo, Altivec, C-5, CodeTEST, CodeWarrior, ColdFire, C-Ware, the Energy Efficient Solutions logo, mobileGT, PowerQUICC, QorIQ, StarCore and Symphony are trademarks of Freescale Semiconductor, Inc., Reg. U.S. Pat. & Tm. Off. BeeKit, BeeStack, ColdFire+, CoreNet, Flexis, Kinetis, MXC, Platform in a Package, Processor Expert, QorIQ Converge, Qorivva, QUICC Engine, SMARTMOS, TurboLink, VortiQa and Xtrinsic are trademarks of Freescale Semiconductor, Inc. All other product or service names are the property of their respective owners. © 2011 Freescale Semiconductor, Inc.





Abstract

This session will provide an overview of the e5500 core implemented in Freescale's new 64-bit QorIQ P5020 communications processor. The extension to 64 bits in this core enhances the QorIQ communications processor family, and incorporates improvements to the floating point and integer execution units along with a larger L2 cache. In this class, you will learn how to utilize the 64-bit mode and other architectural enhancements to help impact system performance.



Agenda

- P5020 overview
 - Major blocks
 - Features
 - Benefits
- e5500 core overview:
 - 64-bit
 - Partitioning and virtualization (hypervisor support)
 - Floating-point
 - L2 cache
 - New integer instructions from Power ISA 2.06
 - Misc performance improvements from e500mc

Freescale on Facebook

Tag yourself in photos
and upload your own!



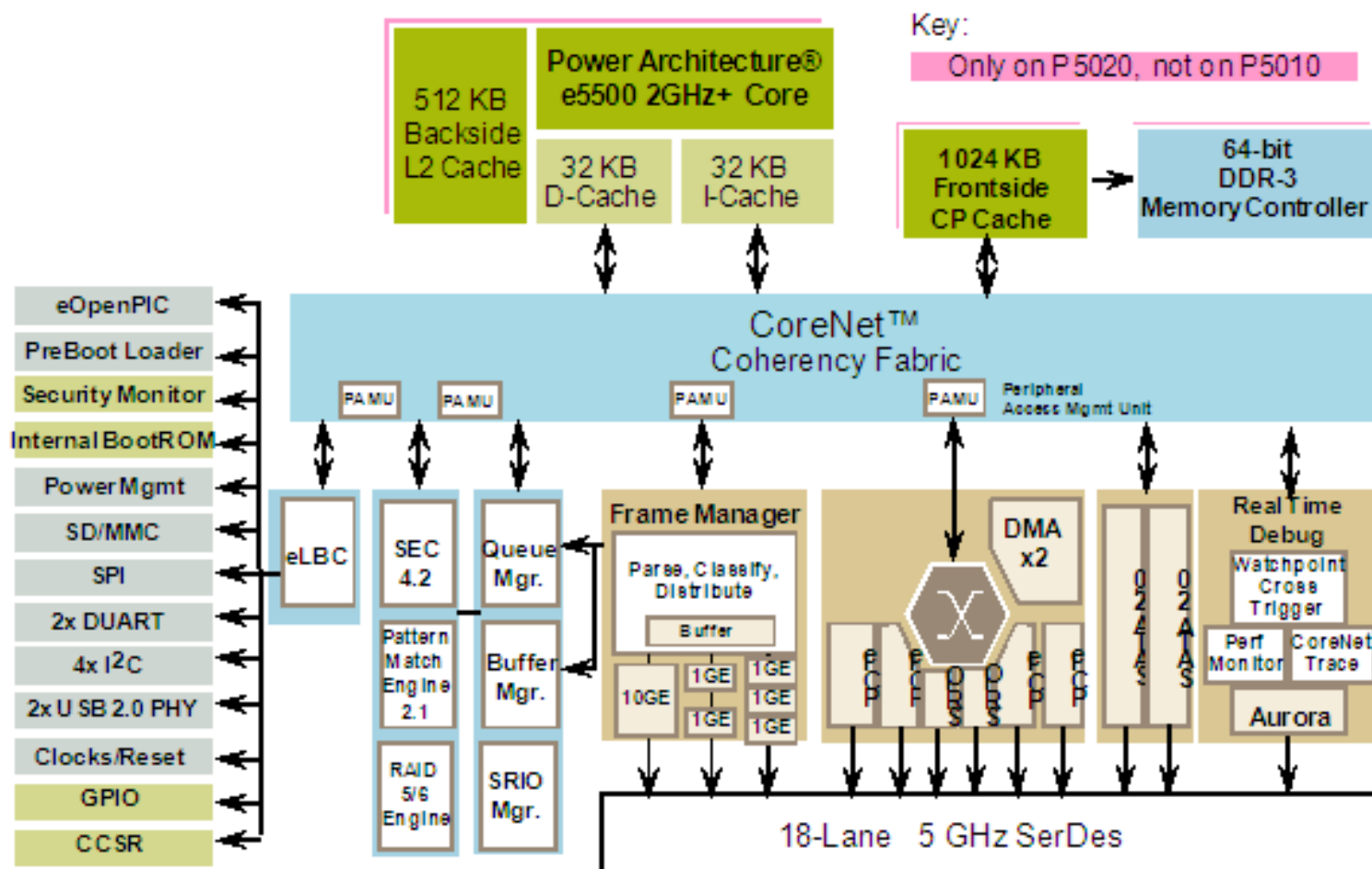
Tweeting?

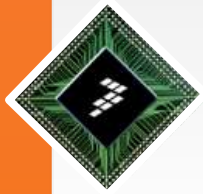
Please use hashtag
#FTF2011





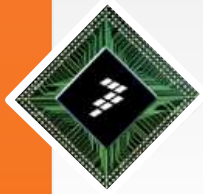
P5020 Block Diagram





P5020 Benefits

- 64-bit Power Architecture® cores
- 32KB L1 instruction cache, parity on data and tags
- 32 KB L1 data cache, parity on data and tags
- 512KB backside L2, ECC on data, parity on tags
- Full cache MESI coherency, intervention , locking, and stashing
- Hypervisor Support
- Double precision FPU
- Interprocessor signaling
- Decorated storage (fast atomic memory ops)
- Enhanced external debug
- 36-bit real address
- 20% DMIPS/MHz improvement over e500mc
- 45-nm Technology
- 2 GHz frequency



Agenda

- P5020 overview
 - Major blocks
 - Features
 - Benefits
- e5500 core overview:
 - 64-bit
 - Partitioning and virtualization (hypervisor support)
 - Floating-point
 - L2 cache
 - New integer instructions from Power ISA 2.06
 - Misc performance improvements from e500mc

Freescale on Facebook

Tag yourself in photos
and upload your own!



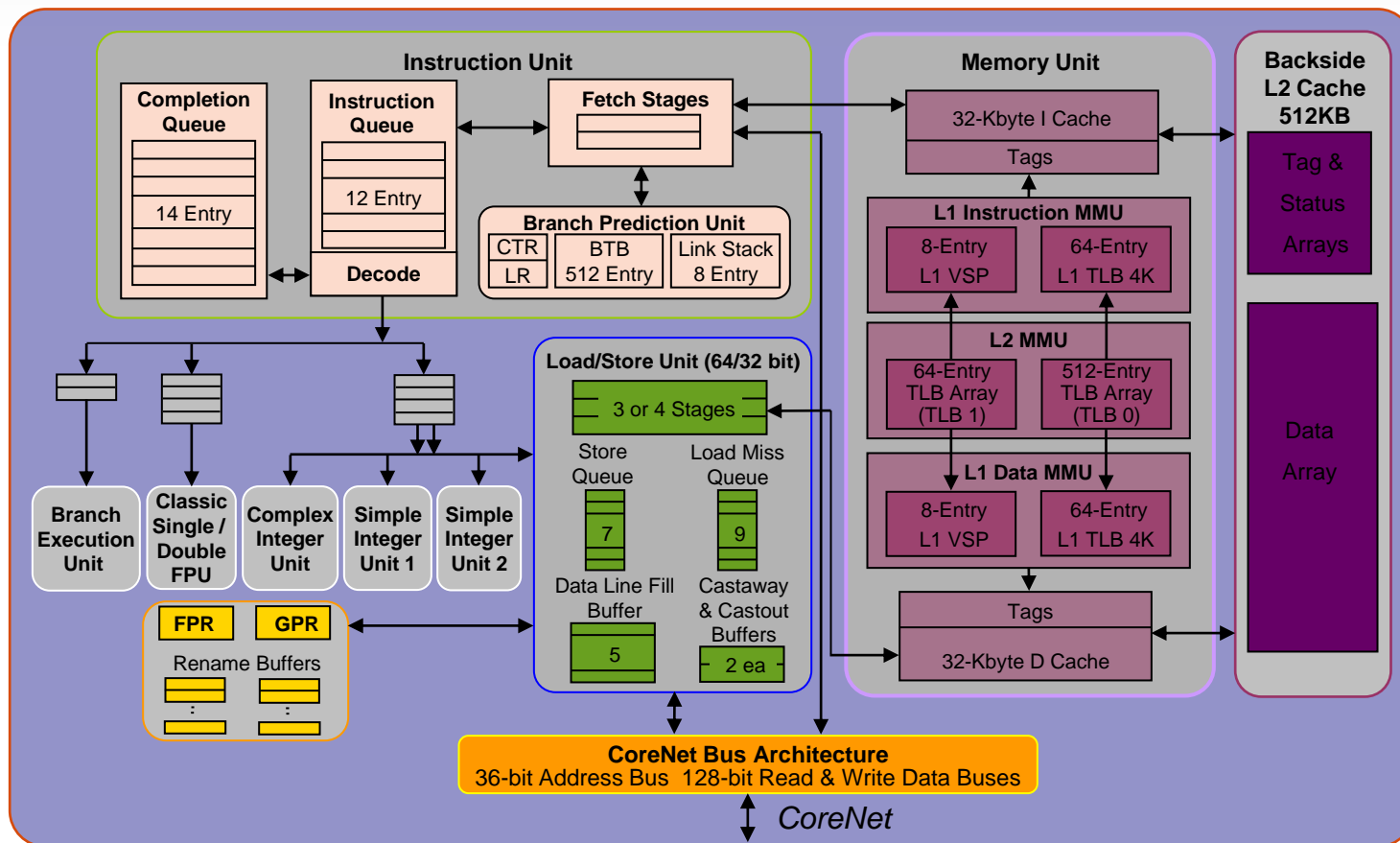
Tweeting?

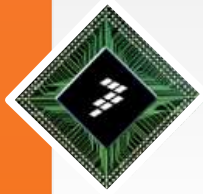
Please use hashtag
#FTF2011





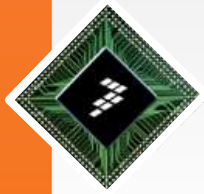
e5500 Core Block Diagram





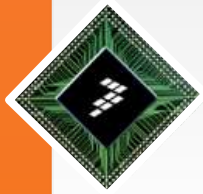
E5500 Core Features

- 64/32-bit Power Architecture
- Superscalar dual-decode, quad-issue: out-of-order execution/in-order completion
- High-Speed Classic IEEE 754 Double-Precision Floating Point
- Full Speed 512KB Backside L2 cache w/ Tag Parity and Data ECC
- 32KB I & D L1 caches with Tag and Data Parity and write shadow mode to recover from single bit errors
- 64 TLB SuperPages, 512-entry 4K Pages, 36-bit Physical Address
- Branch unit with a 512-entry, 4-way set associative, Branch Target/History Buffer and 8-entry Link Stack
- “Decorated Storage” APU provides fire-and-forget atomic updates of up to two 64-bit quantities with a single access



What's New in e5500

- Baseline: 32-bit e500mc.
- 64-bit GPRs, 64-bit computation instructions, 64-bit virtual address
 - 2X higher data throughput through registers
 - More virtual address space
 - Enabling the shift to 64-bit software.
 - Hybrid 32-bit mode to support legacy software and transition to 64-bit architecture.
 - Register settings allow users to utilize 32-bit mode or 64-bit mode, easing transition to 64-bit architecture
- New Power ISA v2.06 instructions for byte- and bit-level acceleration
 - Parity; Population count; Bit permute; Compare bytes
- Larger backside L2 cache (512K). L2 access is 11 cycles from L1 miss (2 cycles longer)
- Improved branch prediction with 8-entry Link-Stack
- High speed, fully pipelined FPU (2x faster SP, 4x faster DP)
 - New FPU supports conversion between floating point and 64-bit integer

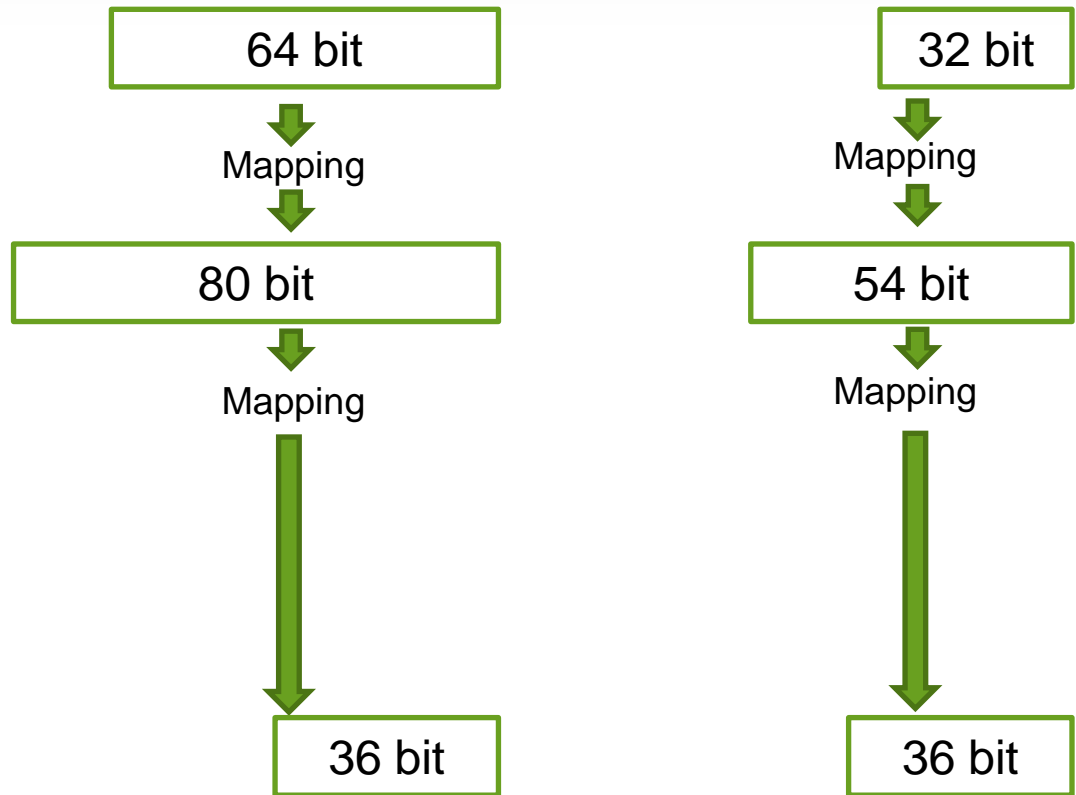


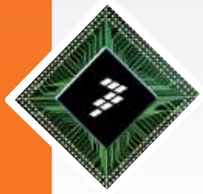
Addressing Terminology

Effective Address

Virtual Address

Real Address





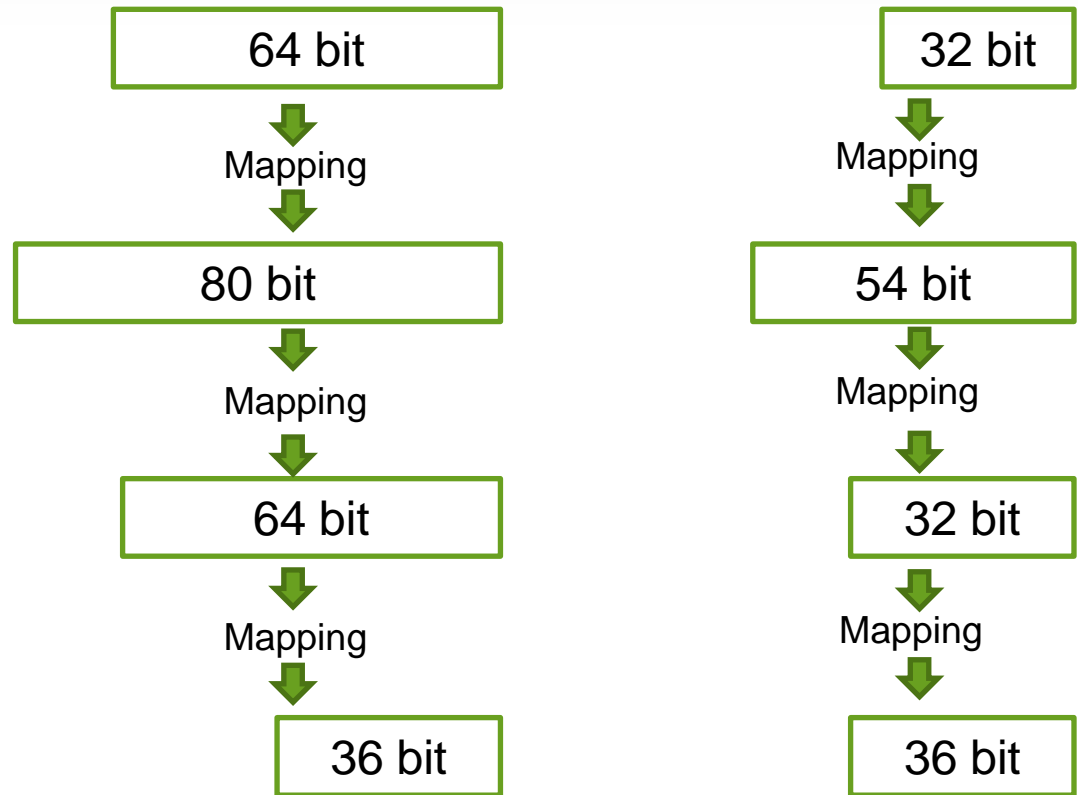
Addressing Terminology with Hypervisor

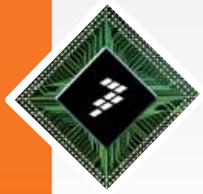
Effective Address

Virtual Address

Logical Address
(what the OS sees
when Hypervisor
is employed)

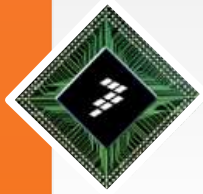
Real Address





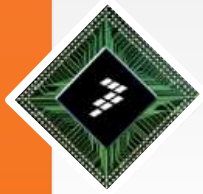
Types of Addresses (Effective Addresses)

- Effective addresses (EA)
 - What a program forms as an address used by a *load*, *store*, or *branch* instruction
 - In 64-bit mode, effective addresses are 64 bits. In 32-bit mode effective addresses are 32 bits, the upper 32 bits are treated as 0.
 - This means in 64-bit mode, a program can potentially access 2^{64} bytes in a normal flat manner. In 32-bit mode only 2^{32} bytes are available.
 - In some operating systems, user programs are further limited because the OS (linux) uses part of the effective address space. In 32-bit linux, the OS reserves 1GB of effective address space, leaving only 3GB for user programs to access.
 - Having the much larger address space in 64-bit mode allows several applications to implement large, sparse data structures which may be much more efficient to use. An example of this is a storage product/system architecture which may want to map blocks of a disk farm into a flat address space.



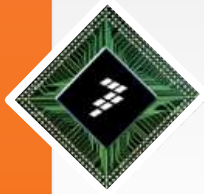
Types of Addresses (Virtual Addresses)

- Virtual addresses (VA)
 - A virtual address uses the EA and adds process context identifiers (different values for each individual process context)
 - The contents of the PID register (8 bits) [managed by the operating system]
 - The contents of either the MSR_{DS} if a data access (*load* or *store*) or MSR_{IS} if an instruction fetch [managed by the operating system]
 - The partition context (LPIDR register context) [managed by the hypervisor]
 - The guest or hypervisor context (MSR_{GS}) [managed by the hypervisor]
 - An OS provides a unique virtual address space to each process using the PID register
 - A hypervisor provides a unique chunk of virtual address spaces to each partition (OS) using the LPIDR register
 - Virtual addresses on e5500 are 80 bits (64 bit EA + 8 bit PID + MSR_{DS} or MSR_{IS} + 6 bit LPIDR + MSR_{GS} = 80 bits)



Types of Addresses (Logical/Physical Addresses)

- Logical addresses
 - Logical addresses are what a Guest operating system executing on a hypervisor thinks are real addresses
 - When a Guest writes a TLB entry that maps a virtual address to a logical address, the hypervisor intercepts the write and substitutes the actual real address
 - Logical addresses only apply to cores that have hypervisor support
 - Logical addresses on e5500 are 36 bits (same size as real addresses)
- Physical addresses (real addresses)
 - Physical addresses are the actual physical address that the core uses to tag addresses in the caches and to specify to the rest of the system when an address goes off the core (to CoreNet for example)
 - The platform on the SoC may translate the physical address to a different physical address, but this is outside the core domain
 - Physical addresses on e5500 are 36 bits



e5500 Micro-architectural Changes

1. New micro-architecture for 64 bits:

1. Improved 64-bit integer multiply and divide.

Latency		
Instruction	e500mc	e5500
32-bit mult.	4 cycles	4 cycles
64-bit mult	N/A	4 to 7 cycles
32-bit divide	4 to 35 cycles	4 to 16 cycles
64-bit divide	N/A	4 to 26 cycles

2. 64-bit add = 1 cycle

3. 64-bit shift/rotate = 2 cycles

4. Population count = 2 cycles

2. Relocated mt/mfspr ops to avoid stalls in single-cycle integer ALU.

Core Comparison (e500)

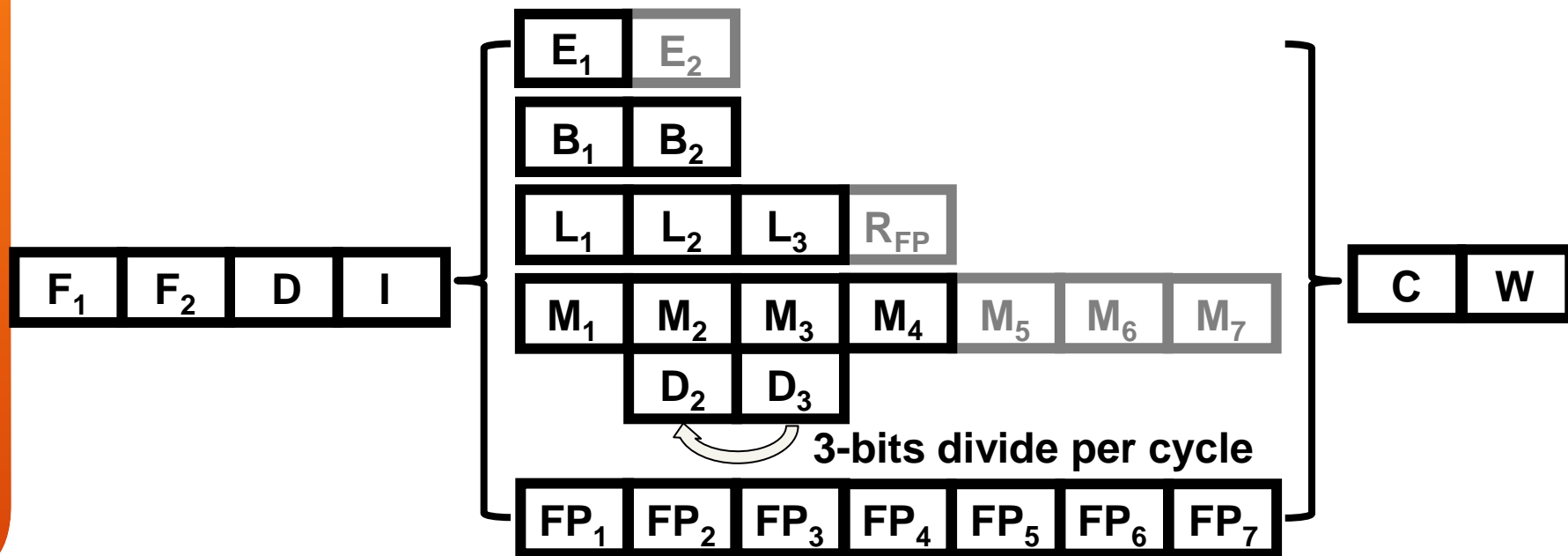
	e500v1 and v2	e500mc	e5500 ⁴
Max Frequency	1.5GHz	1.5 GHz	2GHz
Dhrystone	2.4	2.5	3.0
Pipeline depth / Width	7 / 2	7 / 2	7 / 2
Integer Units	3	3	3
GFLOPs	SP FP = 2 OP/cycle ⁵ DP FP = 1 OP/cycle ⁵	SP FP = 1 OP/cycle ² DP FP = 0.5 OP/cycle ²	SP FP = 2 OP/cycle ² DP FP = 2 OP/cycle ²
Floating-Point	Embedded	Classic	Classic
Vector support	SPE	<none>	<none>
Cache line size	32 bytes	64 bytes	64 bytes
L1 I and D caches	32K 8-way PLRU	32K 8-way PLRU	32K 8-way PLRU
Backside private cache	<none>	128KB 8-way backside L2 per core, PLRU replacement	512KB 8-way backside L2 per core, PLRU replacement
Frontside shared cache	256-1024KB 8-way L2	2MB 32-way CPC ¹	2MB 32-way CPC ¹
Branch direction prediction	512-entry, two-bit	512-entry, two-bit	512-entry, two-bit

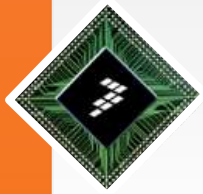
1. P4080 Implementation
2. Pre-silicon calculation
3. Includes private backside cache

4. 64-bit core with 36 bit physical addressing
5. V2 core with SPE only

e5500 Microarchitecture Overview

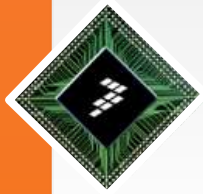
- 7 stage 64-bit integer execution pipeline (8 stages for 64-bit shifts & rotates)
- Quad-fetch, dual-decode, quad-issue, out-of-order execution, dual-complete
- 5 cycle branch miss-prediction recovery
- 3 cycle integer load-use latency
- 4 cycle 32-bit multiplies (up to 7 cycles for 64-bits)
- Un-pipelined divides with 3 result bits per cycle and early-out
- 7 cycle fully-pipelined single & double-precision floating-point





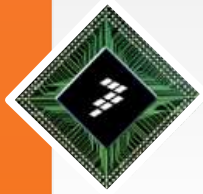
Partitioning and Virtualization

- e5500 supports the same partitioning and virtualization as e500mc
 - Support for 64 logical partitions
 - Run different OSes in each partition under control of a hypervisor
 - Partitions can be a mix of 32-bit or 64-bit; in 64-bit partitions the OS can further support a mix of 32-bit and 64-bit user programs



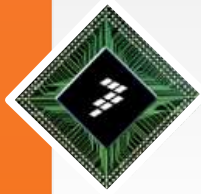
Floating Point

- e5500 supports the same floating point model as e500mc, e600, and e300
 - Significant performance improvement over e500mc at same frequency
 - 2X faster on single precision
 - 4X faster on double precision
 - Additional instructions that perform convert to/from integer double word to floating point



L2 Cache

- L2 cache, 512KB
 - Was 128KB on e500mc
 - Larger area more suitable to control plane applications and larger working sets of data
 - To accommodate the size increase, 2 extra cycles of latency are required for L2 cache accesses over e500mc



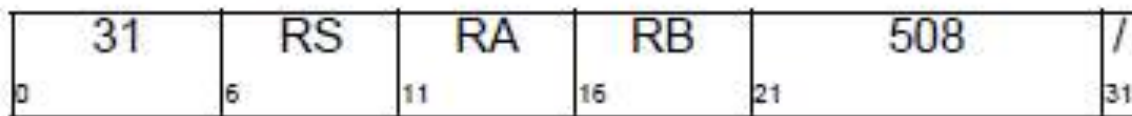
New Integer Instructions

- Compare bytes
 - ***cmpb*** – can compare up to 8 bytes in a GPR
- Population count
 - ***popcntb, popcntw, popcntd*** – computes the number of bits set in each byte, word, or doubleword of a GPR
- Bit permute
 - ***bpermd*** – permutes bits selected by a control GPR from bytes in a source GPR, producing an 8-bit result
- Parity
 - ***prtyw, prtyd*** – exclusive-or the least significant bits of each byte in the word (or both words) in a GPR, producing a 1 bit result for each word



New Instruction: cmpb – Compare Bytes

cmpb RA,RS,RB



- Each byte of the contents of register RS is compared to each corresponding byte of the contents in register RB. If they are equal, the corresponding byte in RA is set to 0xFF. Otherwise the corresponding byte in RA is set to 0x00.

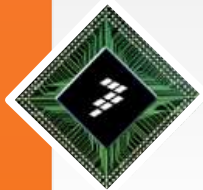
Example:

cmpb r3,r4,r5

r4 = 0xdeadbeef_00112233

r5 = 0xdeadbeef_00000000

r3 = 0xffffffff_ff000000



New Instruction: *popcntb*, *popcntw*, *popcntd* – Population count

popcntb RA, RS

31	RS	RA	///	122	/
0	6	11	16	21	31

- A count of the number of one bits in each byte/word/double word of register RS is placed into the corresponding byte/word/double word of register RA. This number ranges from 0 to 8/32/64, inclusive.

Example:

popcntb r3,r4

r4 = 0x00112233_44556677

r3 = 0x00020204_02040406



New Instruction: *bpermd* – Bit Permute

`bpermd RA,RS,RB`
 [Category: Embedded.Phased-in, Server]

31	RS	RA	RB	252	/
0	6	11	16	21	31

- Eight permuted bits are produced. For each permuted bit *i* where *i* ranges from 0 to 7 and for each byte *i* of RS, do the following.
 If byte *i* of RS is less than 64, permuted bit *i* is set to the bit of RB specified by byte *i* of RS; otherwise permuted bit *i* is set to 0.
- The permuted bits are placed in the least-significant byte of RA, and the remaining bits are filled with 0s.

Example:

`bpermd r3,r5,r4`

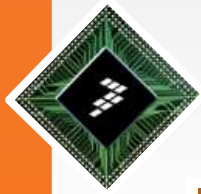
`r4 = 0x00112233_44556677`
`r5 = 0x20212324_30313233`
`r3 = 0x00000000_00000046`



Programming Note – Bit Permute

- The fact that the permuted bit is 0 if the corresponding index value exceeds 63 permits the permuted bits to be selected from a 128-bit quantity, using a single index register. For example, assume that the 128-bit quantity Q, from which the permuted bits are to be selected, is in registers r2 (high-order 64 bits of Q) and r3 (low-order 64 bits of Q), that the index values are in register r1, with each byte of r1 containing a value in the range 0:127, and that each byte of register r4 contains the value 64.
- The following code sequence selects eight permuted bits from Q and places them into the low-order byte of r6.

	<code># will select 8:11 and 92:95</code>	<code>r1 = 0x08090A0B_5C5D5E5F</code>
	<code># from bits 0:127</code>	<code>r2 = 0xFFEEDDCC_BBAA9988</code>
	<code># where r2 holds 0:63</code>	<code>r3 = 0x77665544_33221100</code>
	<code># and r3 holds 64:127</code>	<code>r4 = 0x40404040_40404040</code>
<code>bpermd r6,r1,r2</code>	<code># select from high order half of Q</code>	<code>r6 = 0x00000000_000000E0</code>
<code>xor r0,r1,r4</code>	<code># adjust index values</code>	<code>r0 = 0x48494A4B_1C1D1E1F</code>
<code>bpermd r5,r0,r3</code>	<code># select from low order half of Q</code>	<code>r5 = 0x00000000_00000004</code>
<code>or r6,r6,r5</code>	<code># merge the two</code>	<code>r6 = 0x00000000_000000E4</code>



New Instruction: *prtyw*, *prtyd* – Parity

prtyd RA,RS
[Category: 64-bit]

31	RS	RA	///	186	/
0	6	11	16	21	31

- The least significant bit in each byte of the contents of register RS is examined. If there is an odd number of one bits the value 1 is placed into register RA; otherwise the value 0 is placed into register RA.
- *prtyw* performs the same parity operation on each 32 bit word in RS, placing the result in each word of RA.

Example:

prtyd r3,r4

r4 = 0x00415101_00617100
r3 = 0x00000000_00000001



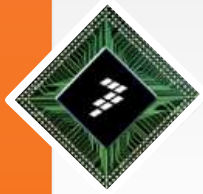
Programming Note – Parity

The *Parity instructions* are designed to be used in conjunction with the *Population Count instruction* to compute the parity of words or a doubleword. The parity of the upper and lower words – $2 * 32$ bits - in RS can be computed as follows.

	RS = 0x10111313_00000000
popcntb RA, RS	RA = 0x01020303_00000000
prtyw RA, RA	RA = 0x00000001_00000000

The parity of 64 bits in RS can be computed as follows.

	RS = 0x10111313_00000000
popcntb RA, RS	RA = 0x01020303_00000000
prtyd RA, RA	RA = 0x00000000_00000001



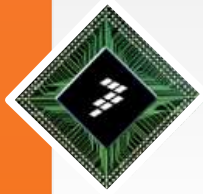
Frequency and Power Improvements

- Higher clock frequency attained via:
 - a. Logic optimization of critical paths
 - b. Selective usage of faster (but more leaky) gates on critical paths
 - c. Gate placement optimization (manually and semi-automated)
- Power Optimization
 - a. More clock-gating to reduce dynamic power
 - b. Extensive use of low-leakage gates on non-critical paths to reduce leakage power



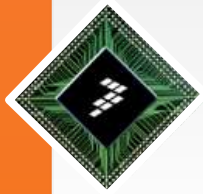
Misc Performance Improvements

- Performance improvements increase DMIPs/MHz by 20% over e500mc
- Frequency is 2GHz, a 33% improvement over e500mc
- Enhanced branch prediction
 - Prediction of upper 32-bits of fetch address
 - Code which jumps between more than 8 4GB segments of effective address space can incur some penalty (OSes may want to limit how they allocate executable addresses)
 - Link stack for better prediction of function call and return
- Some 64-bit operations naturally improve performance
 - 64-bit load and store can move data faster
 - 64-bit arithmetic operations, especially multiply
- Improved performance on multiply and divide operations



More Information

- Power ISA 2.06, available at power.org, is a good reference for understanding more details about new e5500 instructions. The architecture therein is described as a 64-bit architecture.
- The P5020 product website is:
http://www.freescale.com/webapp/sps/site/prod_summary.jsp?code=P5020



Summary and/or Q&A

- P5020 is the first 64b device from Freescale
 - Addresses control plane or processing requirements that exceed the 4GB addressing limit of a 32b processor.
- 45 nm technology and e5500 design enhancements enable higher core frequency while staying under 30 Watts.
- e5500 core enhances double-precision floating point performance
- The high clock frequency, larger L2 cache, improved per-cycle performance, 64b architecture, and good power efficiency in a pin-compatible package make the P5020 a very competitive device.



Session materials will be posted @

www.freescale.com/FTF

Look for announcements in the FTF Group
on LinkedIn or follow Freescale on Twitter



Introducing the New QorIQ AMP Series

Leading Performance, Power and Scalability

Advanced Multiprocessing (AMP) Series enables **new levels of performance** through intelligent integration balanced with a focus on **power efficiency**

AMP Up Performance – Get 4x performance improvement with efficient, high-performance cores and application accelerators

Save Power – Reduce system power using advanced 28nm process technology and cascading power management

Ultimate Scalability – Single to 24 core virtual machines, the QorIQ portfolio offers a broad range of supported applications and software compatibility

QorIQ Processing Platforms

T5xxx (Future)

P5010, P5020

- Up to 2.5 GHz
- Up to 6 cores
- Up to 6 MB L2 Cache



**Service Provider
Routers**



**Network
Admission Control**



**Storage
Networks**

T4240 **T4xxx, T3xxx** (Future)

P4080, P4040, P3041

**T4240
Sampling
Q1'12**

- Up to 2.0 GHz
- Up to 24 virtual cores
- 50 Gbps IP forwarding
- High amount of off-load



**Metro Carrier
Edge Router**



**Aerospace
& Defense**



**Access
Gateway**

T1xxx, T2xxx (Future)

P101x(8), P102x(5),
P2x(4)

- Up to 1.6 GHz
- Up to 8 virtual cores
- Less than 10W
- Price competitive



**Integrated
Services Router**

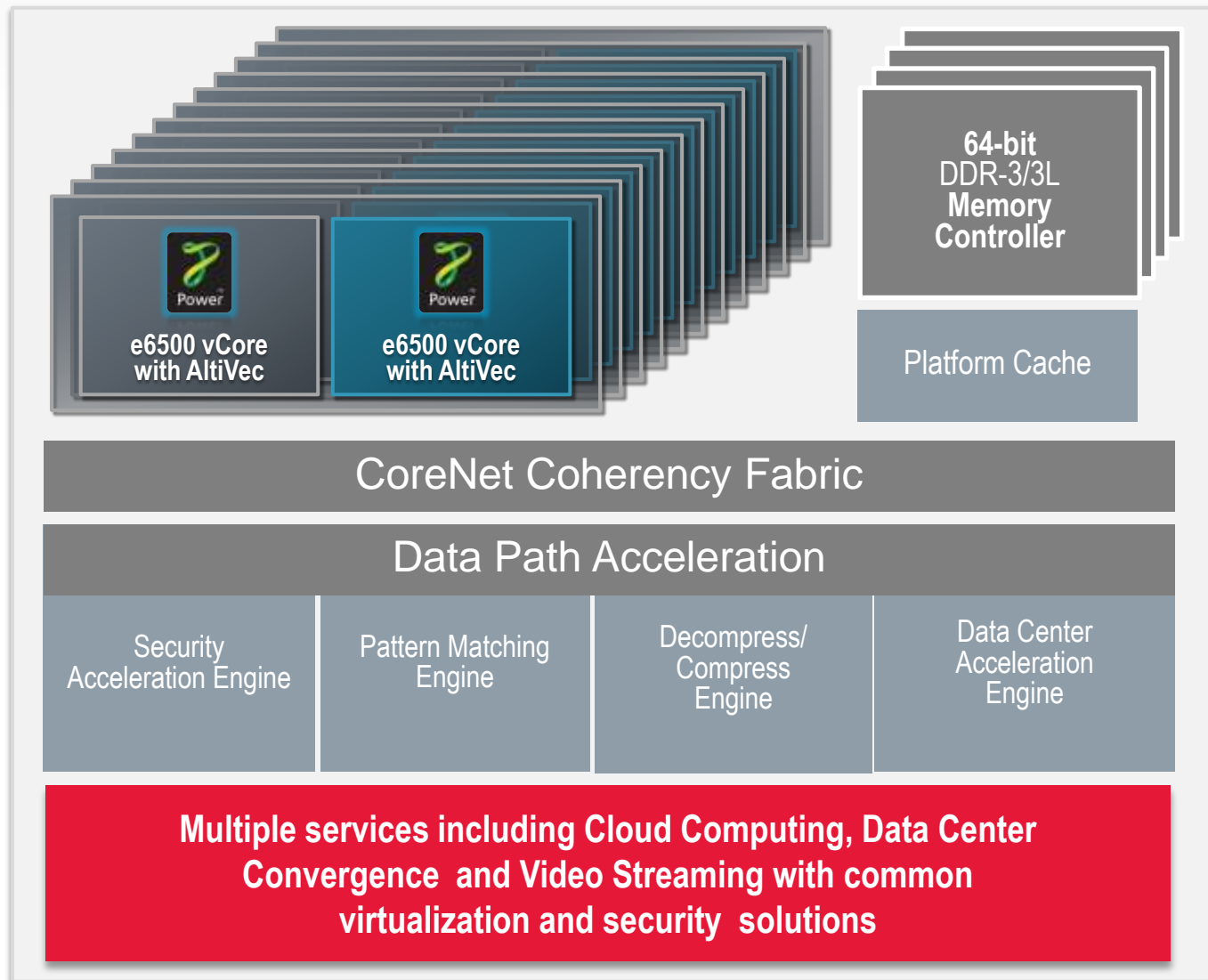


**Network
Attached Storage**



**Media
Gateway**

QorIQ Advanced Multiprocessing (AMP) Series Block Diagram







FTF | FREESCALE TECHNOLOGY FORUM
POWERING INNOVATION

Additional Material on 64 bit Addressing



Freescale, the Freescale logo, Altivec, C-5, CodeTEST, CodeWarrior, ColdFire, C-Ware, the Energy Efficient Solutions logo, mobileGT, PowerQUICC, QorIQ, StarCore and Symphony are trademarks of Freescale Semiconductor, Inc., Reg. U.S. Pat. & Tm. Off. BeeKit, BeeStack, ColdFire+, CoreNet, Flexis, Kinetis, MXC, Platform in a Package, Processor Expert, QorIQ Converge, Qorivva, QUICC Engine, SMARTMOS, TurboLink, VortiQa and Xtrinsic are trademarks of Freescale Semiconductor, Inc. All other product or service names are the property of their respective owners. © 2011 Freescale Semiconductor, Inc.



Translation (e5500)

MSR_{GS}

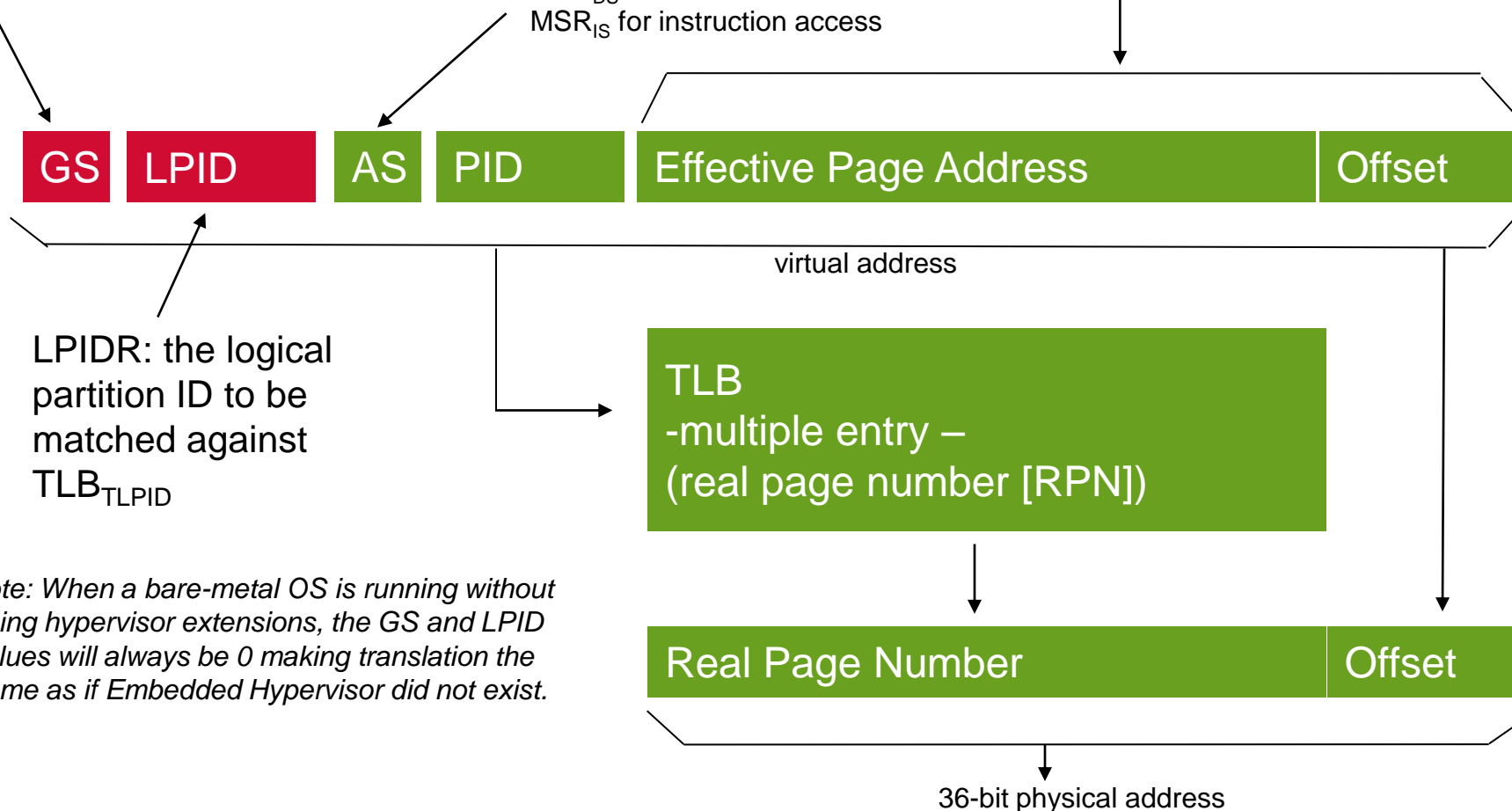
0 = Hypervisor access

1 = Guest access

MSR_{DS} for data accesses

MSR_{IS} for instruction access

64-bit effective address

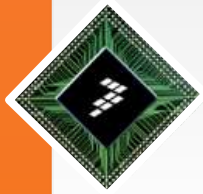


Note: When a bare-metal OS is running without Using hypervisor extensions, the GS and LPID Values will always be 0 making translation the Same as if Embedded Hypervisor did not exist.



External PID Load / Store

- Load/store access to other address space and contexts
 - Allows loads and stores to be performed in a context different from the current executing context
 - The external load context is independent from the external store context
- Supervisor instructions to perform external loads, stores and cache management operations
 - *lbepx, stbepx, lhexp, sthexp, lwepx, stwepx, ldexp, stdep* ...
 - The External PID Load Context (EPLC) register defines the load context
 - The External PID Store Context (EPSC) register defines the store context
 - The external context includes the PID value, the AS (MSR_{DS}) value and the privilege (MSR_{PR})
 - For the hypervisor, the LPID value and the GS value are also available in the external context
 - Indexed forms only
- All instruction fetches are translated in the current context only



External PID Load / Store Uses

- Useful for kernel *copyin()* and *copyout()* functions
 - True address space separation with high performance
- Kernel address space can be completely independent of user process address spaces
 - Allows user process address spaces to be 2^{64} (64-bit mode) or 2^{32} (32-bit mode)
- Easily validate user addresses passed to the kernel
- Provide a fast method of copying data from one context to another
 - Setup EPLC to be the copy-from context
 - Setup EPSC to be the copy-to context
- For hypervisor, useful for performing same type of operations between partitions

External PID Address Translation (e5500)

