IA-32 Architecture

Sunil Saxena
Principal Engineer
Intel Corporation

September 11, 2000

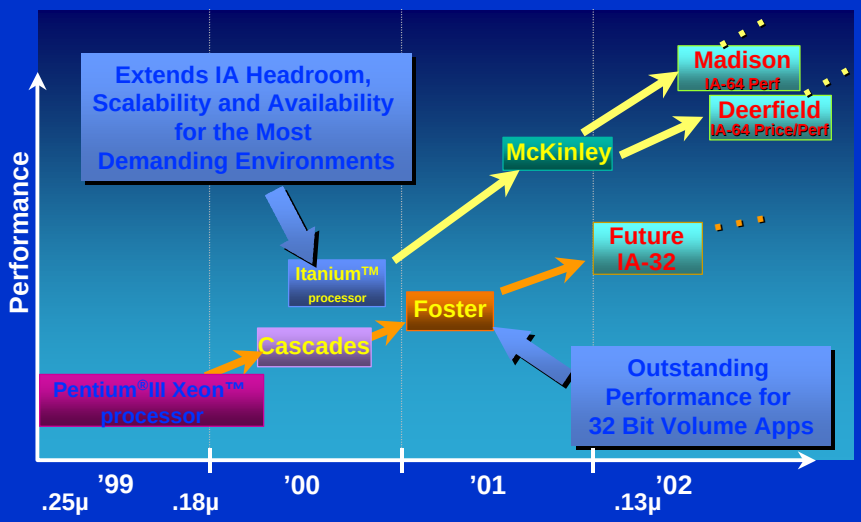
Intel Labs

Agenda

- Pentium ® III Processor New Features
- Pentium ® 4 Processor New Features
- Pentium ® 4 Processor Micro-architecture

Intel Labs

IA Processor Roadmap



Intel Labs

Pentium III Processor

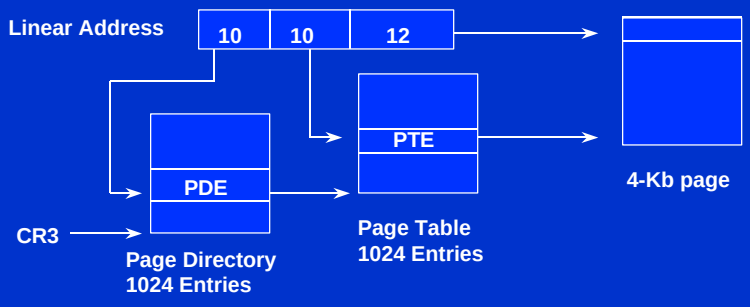
- Pentium III Processor New Features
  - 36-bit Physical Addressing
    - Physical Address Extension - PAE-36
    - Page Size Extensions - PSE-36
  - Page Attribute Table
  - Fast Floating-point save/restore
  - New Instructions
  - New Exceptions

Intel Labs

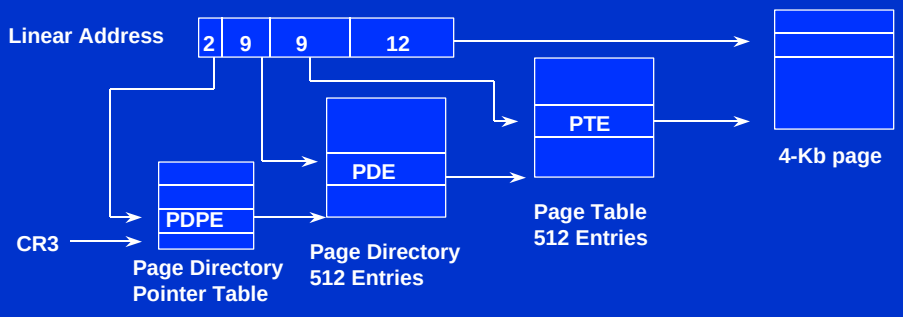
36-bit Addressing

- 36-bit Addressing
  - PSE-36
  - PAE-36
- PSE-36
  - 4GB mapped through 4K of page directories and 4MB page tables
  - Memory above 4 GB is only accessible as 4 MB pages
  - Operating system can freely use both 4KB and 4MB pages without PDE structure change
    - All 4KB pages and page tables MUST reside below 4GB boundary
    - Reduces effort needed to develop & support changes in virtual memory subsystem
- PAE-36
  - 4GB mapped through 16K of page directories and 16MB page tables
  - All Memory accessible as 4KB or 2MB pages
  - OS needs to load PDEPTRs for mapping changes on writes to CR3
  - CONFIG_HIMEM to enable more than 4 GB physical memory

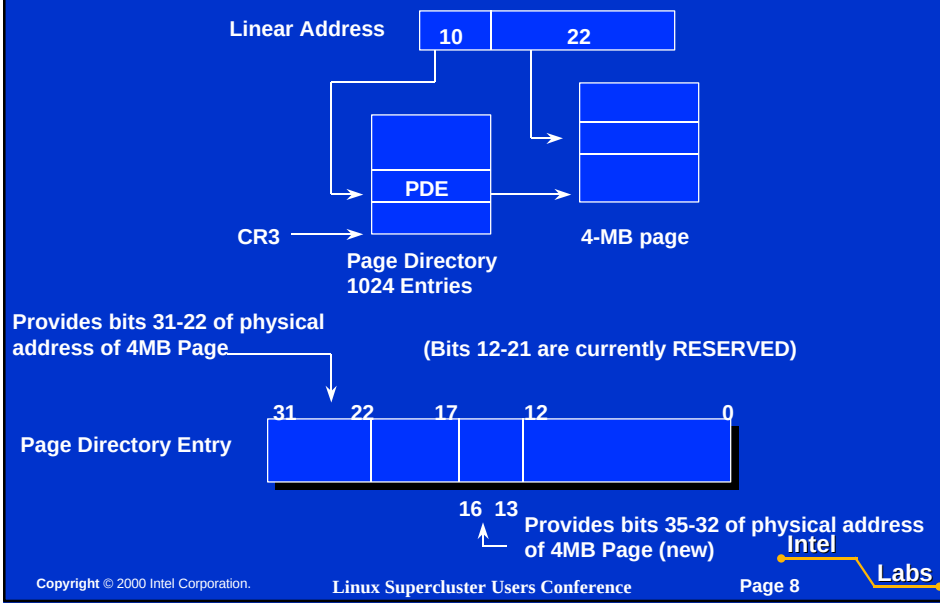
4KB Page Translation



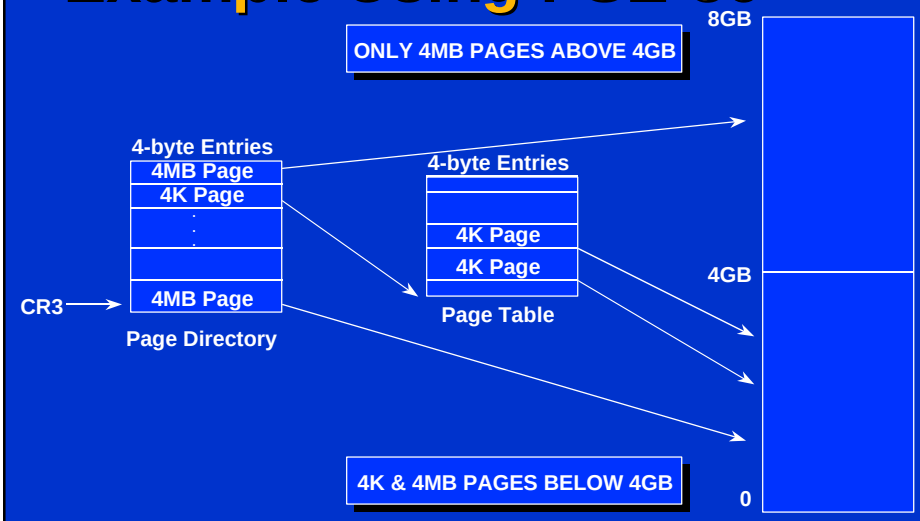
4KB PAE Translation



4MB Translation



Example Using PSE-36



Page Attribute Table (PAT)

- Physical Memory Attributes Described through the Page-Tables
  - Builds upon enhanced memory type capability provided via MTRR's in Pentium® Pro processor
  - Relaxes MTRR alignment/length requirements
- Builds upon PCD/PWT bits on IA-32 Architecture
  - These interact with effective memory type determination
- PAT Architecture
  - PAT is an 8-entry table indexed via PCD, PWT, and Resv. bits
    - Allows up to 8 memory attributes defined by the page tables
    - PAT is always enabled when Paging is used
    - Default table entries fully compatible with PCD/PWT/Resv settings
  - PAT entries R/W programmable via RDMSR/WRMSR (0x277)
    - 8 bits per entry; 3 bits for attribute with other bits reserved
    - Memory attributes as specified by Pentium® Pro processor

Page Attribute Table (PAT)

- PAT Architecture (continued)
  - PAT Memory Types interact with MTRRs
    - As architecturally specified by Pentium® Pro processor
    - Implementation specific combinations remain undefined
      - Should not be depended upon by system software
- Precautions
  - OS Uses Page Directory as a Page Table:
    - Restricted to using 4 lowest PAT entries
      - PAT bit 7 in 4K PTE is PS bit when used as a PDE
  - Memory type changes for pages require TLB invalidation
    - Follow procedure as when changing MTRRs
      - cache flush, TLB invalidation
    - PAT entries on multiple processors must be maintained in consistent manner by OS
      - All processors have same values in PAT

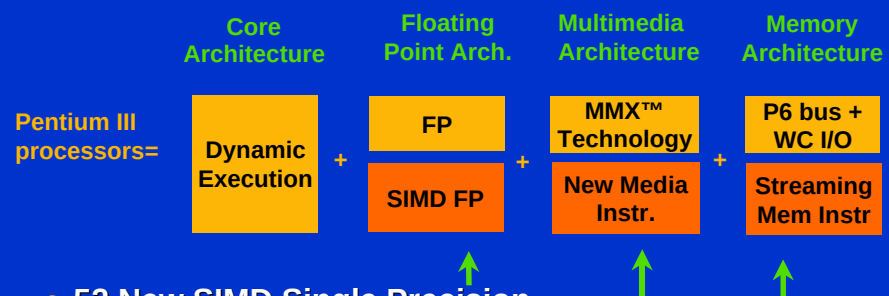
Page Attribute Table (PAT)

- Precautions (continued)
  - Page Aliasing
    - PAT maintains memory types according to linear addresses
    - Architecture allows OS to map single physical page with 2 linear addresses containing differing types
    - This may lead to undefined results and must be avoided
- PAT Uses
    - Essentially unlimited MTRRs
      - Provide support for more devices (frame buffers, RAID cards, etc…) to map memory as WC
    - Allows map system memory for specific optimizations
      - Memory shared with 3D accelerator/CPU for textures
      - Reduce eviction, read-for-ownership bus transactions and cache thrashing for common operations such as memory fill

Fast Floating Point Save/Restore

- These instructions minimize cost of saving/restoring Floating Point/MMX™ Technology State
  - Does NOT re-initialize the FPU state after saving
  - Performance improvements come from more natural format and alignment of the cpu state
- State area is larger
  - 512 bytes
  - MUST be aligned on 16 byte boundary, else GP(0) fault
  - Use of reserved fields risks incompatibility with future Intel Architecture processors
- FXSAVE does not check for unmasked exceptions (i.e. like FNSAVE)
  - FXRSTOR does not fault when loading an image that contains pending exceptions

Pentium III New Instructions



- 52 New SIMD Single Precision Floating Point Instructions
  - up to 4 FP results per cycle
  - Eight 128 bit registers
  - 4 x Single precision FP numbers
- 12 New Media Instructions
- 8 New Cacheability Instructions

Prefetching Instruction

- Prefetch gets a cacheline at a time
- Prefetch Hint (Load) Instructions
  - Instructions do not fault
  - Retires quickly to free up machine resources
  - Hints to cache at different levels
    - Store in different levels of cache hierarchy
    - Don't store in the cache hierarchy (stream)
- Potential OS tuning uses
  - e.g. TCP/IP Checksum gets ~2x speedup

Streaming Store Instruction

- Store data to memory minimizing cache pollution
- Potential OS tuning benefits
  - 128 bit registers used with streaming store to zero pages ~4x faster
  - mem copy ~2x faster using prefetch/stream together

New Exceptions

- Interrupt vector 19 used to invoke unmasked exception handlers
  - Provide larger (512 bytes of state) context record to handler
  - Handlers need to account for SIMD nature of Pentium III SSE numeric exceptions
    - One instruction can generate multiple exceptions
- Exceptions are precise (reported when detected)
- Pentium III SSE Instructions architecturally separate from x87-FP
  - Pentium III SSE Instructions do not report x87-FP/MMX™ Technology exceptions
  - New handlers must include IEEE filter to decode and emulate exception raising SIMD instructions

Pentium® 4 Processor

- Pentium 4 Processor New Features
  - SSE2 Instructions
  - Enhanced Prefetch Instructions
  - System Bus and Cache Enhancements
- OS Recommendation
- New Instruction support

Pentium 4 Architecture Overview

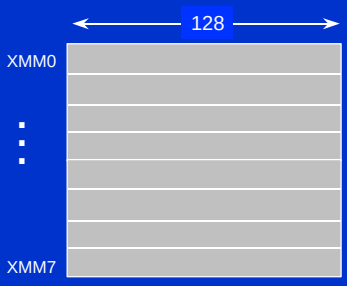
- Willamette is the next generation IA-32 processor microarchitecture
- New micro-architecture
  - ~1.4x average performance of Pentium® III processor family on same process
  - Enables faster processor speeds (1 GHz+)
  - Trace Cache for Instruction Decode
- Willamette New Instructions
- New platform (chipsets, AGP4X)

Pentium 4 New Instructions

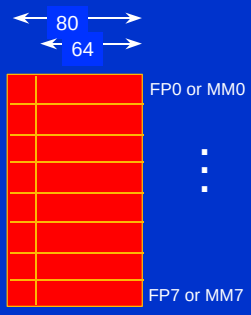
- New 128 bit arithmetic instructions
  - Extend MMXä technology instructions from 64 bit to 128 bit data type
    - Operates on XMM registers instead of MMX/x87-FP registers
  - New 128-bit integer and SIMD-Integer instructions
    - Memory operands MUST be 128-bit aligned! Will cause Exception during executions if not aligned.
  - Packed 32 * 32 bit Multiply
  - Packed 64 bit Add/Subtract
  - Shift, Shuffle, Unpack, Move, Conversion
- New SIMD Double Precision FP instructions
  - Full complement of FP arithmetic operations
  - Packed/Scalar DP ⇔ SP conversions
- New cache / memory management instructions
  - Cache line flush instruction
  - Fences (LFence / MFence)
  - New streaming store instructions

Streaming SIMD Extensions 2

Floating Point Registers
(Scalar/packed SIMD-SP-FP,
SIMD-DP-FP, 128-bit Integer)

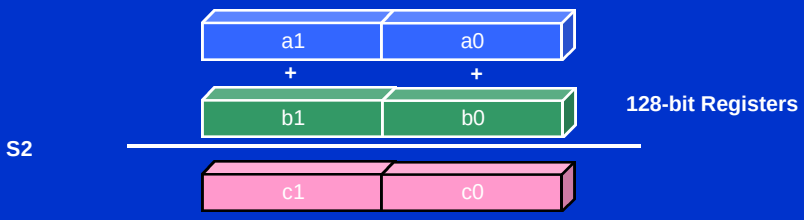


Integer / x87 Registers
(64-bit Integer, x87 data)



Example SIMD Add (ADDPD)

- Effectively performs two double precision ops in one cycle
- a1+b1=c1 in parallel with a0+b0=c0
- Useful for matrix operations



Prefetches

- The Intel® Pentium® 4 processor has automatic prefetches which
  - Work on large buffers
  - Have Sequential access
- Even fewer prefetches necessary

Use sequential access to buffers and get prefetches "for free"

But…

- prefetch instructions may still be the best solution in some cases
  - PrefetchNTA reduces cache evictions of useful data (1.1-1.15x gain)
  - Benefits unusual (ie, non-contiguous) data access patterns
  - Can maximize read bandwidth to system memory
  - Increase fetch-ahead distance since memory-latency/computation delta increases

System Bus & Cache Enhancements

- The Pentium 4 system bus is an evolutionary extension of the P6 bus
- 3.2 GByte/sec data transfer rate
  - 100MHz quad pumped data bus - similar to AGP-4X
  - Source synchronous 64 bit data bus
- Caches
  - Trace cache for decoded instructions
  - 128 byte cache lines with 64 byte sectors
  - 256K on-die, 2nd level write-back, unified data and instruction cache
- APIC
  - Messages now sent over front side bus
  - Physical destination mode expanded to 8-bits
  - ISR, IRR, TMR implementation increased to 256 bits
  - Remote read is no longer supported

OS Recommendations

- All spin-loops should include the PAUSE instruction
  - Backwards compatible with prior IA-32 processors
  - Significant performance benefit in future IA-32 processors
  - Already done in 2.4-test* kernels
- Cache line size is 128 bytes with 64 byte sectors
  - Impact to hot locks
    - Hot locks should be on separate 64 byte sectors
  - Impact to data structure alignment
    - 128 byte line allocation in cache
- Use Non-execution based Timing Loops!
  - Already done in 2.4-test* kernels

Pentium 4 New Instruction Support

- FXSAVE/FXRSTOR support for Pentium 4 state
  - Already done if enabled for Pentium® III processor (Internet Streaming SIMD Extensions)
  - No New State!
  - Already done in 2.4-test* kernels
- New Exception Handlers
  - Double Precision SIMD capable
  - IEEE Compliant
- Prefetch and Streaming Store Optimizations
  - Integer state streaming store instruction `MOVNTi`
    - For zeroing, memcpy, etc.
    - Does not use FP state so DNA is avoided

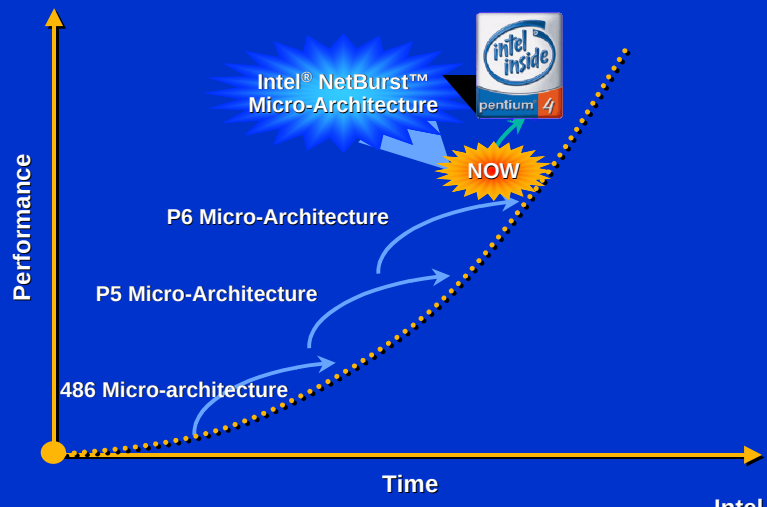
Pentium® 4 Processor Micro-architecture

Next Generation IA-32 Micro-architecture

Agenda

- IA-32 Processor Roadmap
- Design Goals
- Frequency
- Instructions Per Cycle
- Summary

Intel® Pentium® 4 Processor



Intel® Pentium® 4 Processor Design Goals

- Deliver world class performance across both existing and emerging applications
- Deliver performance headroom and scalability for the future

Micro-architecture that will Drive Performance Leadership for the Next Several Years

CPU Architecture 101

Delivered Performance =
* Instructions Per Cycle
Frequency

Frequency

- What limits frequency?
  - Process technology
  - Microarchitecture
- On a given process technology
  - Fewer gates per pipeline stage will deliver higher frequency

Frequency is driven by Microarchitecture

Intel Labs

Netburst™ Micro-architecture Pipeline vs P6

Basic P6 Pipeline

Intro at 733MHz .18µ

1	2	3	4	5	6	7		10
Fetch	Fetch	Decode	Decode	Decode	Rename	ROB Rd	Rdy/Sc	Exec

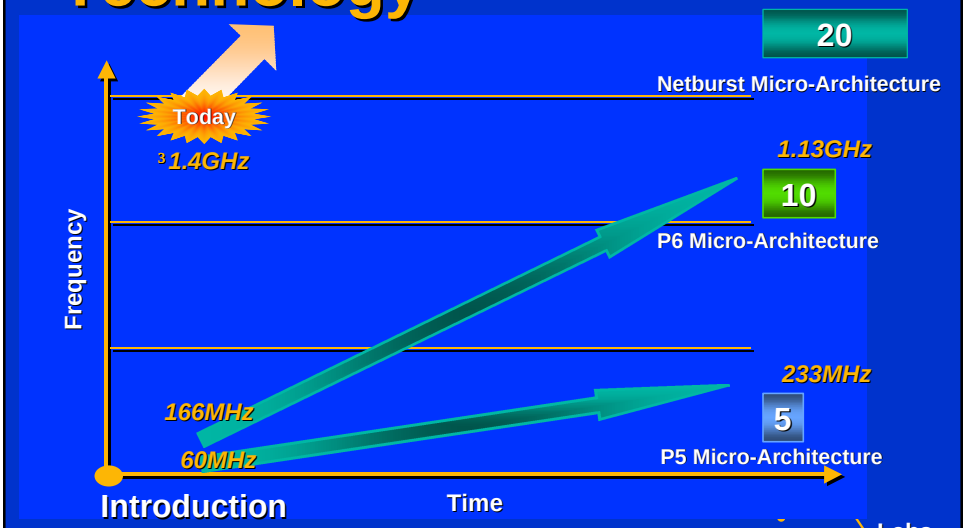
Basic Pentium® 4 Processor Pi

Intro at 1.4GHz .18µ

1	2	3	4	5	6	7	8	9	10	11	12	13	14	3
TC Nxt IP		TC Fetch		Drive	Alloc	Rename		Que	Sch	Sch	Sch	Disp	Disp	

Hyper pipelined Technology enables industry leading performance and clock rate

Hyper Pipelined Technology



CPU Architecture 101

Delivered Performance =
Freq
Instructions Per Cycle

Improving Instructions Per Cycle

- Improve efficiency
  - Branch prediction
  - Do more things in a clock
- Reduce time it takes to do something
  - Reducing latency

Improving Instructions Per Cycle

- Improve efficiency
  - Branch prediction
  - Do more things in a clock
- Reduce time it takes to do something
  - Reducing latency

Branch Prediction

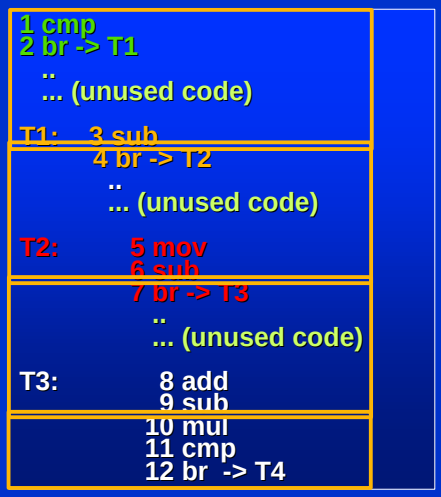
- Accurate branch prediction is key to enabling longer pipelines
- Dramatic improvement over P6 branch predictor:
  - 8x the size (4K)
  - Eliminated 1/3 of the mispredictions
- Proven to be better than *all* other publicly disclosed predictors
  - (g-share, hybrid, etc)

Execution Trace Cache

- Advanced L1 instruction cache
  - Caches “decoded” IA-32 instructions (uops)
- Removes decoder pipeline latency
- Capacity is ~12K uOps
- Integrates branches into single line
  - Follows predicted path of program execution

Execution Trace Cache feeds fast engine

Execution Trace Cache



Trace Cache Delivery

1 cmp	2 br T1	3 T1: sub
4 br T2	5 mov	6 sub
7 br T3	8 T3:add	9 sub
10 mul	11 cmp	12 br T4

Advanced Dynamic Execution

- Extends basic features found in P6 core
- Very deep speculative execution
  - 126 instructions in flight (3x P6)
  - 48 loads (3x P6) and 24 stores (2x P6)
- Provides larger window of visibility
  - Better use of execution resources

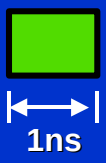
Deep Speculation Improves Parallelism

Improving Instructions Per Cycle

- Improve efficiency
  - Branch prediction
  - Do more things in a clock
- Reduce time it takes to do something
  - Reducing latency

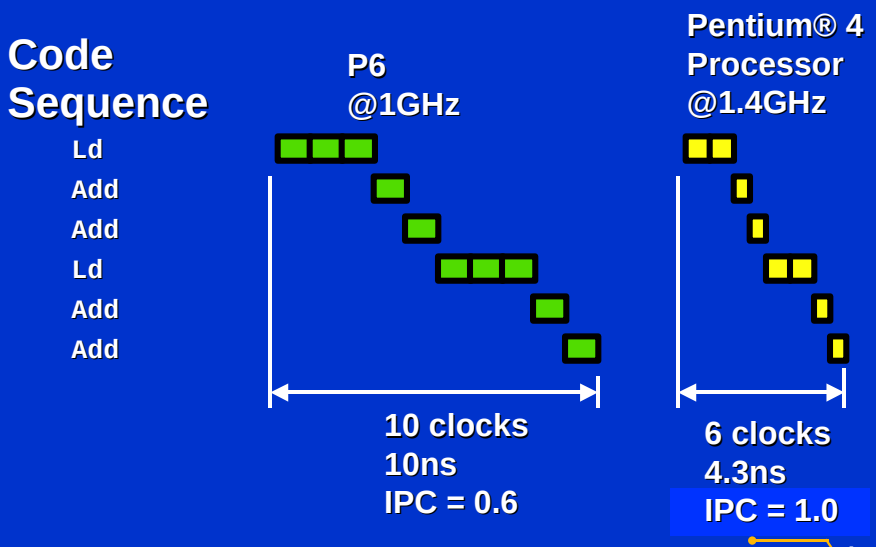
Rapid Execution Engine

- Dramatically lower ALU latency
- P6:
  - 1 clock @ 1GHz



- P4P:
  - ½ clock @ >1.4GHz

Example with Higher IPC and Faster Clock!



Recap

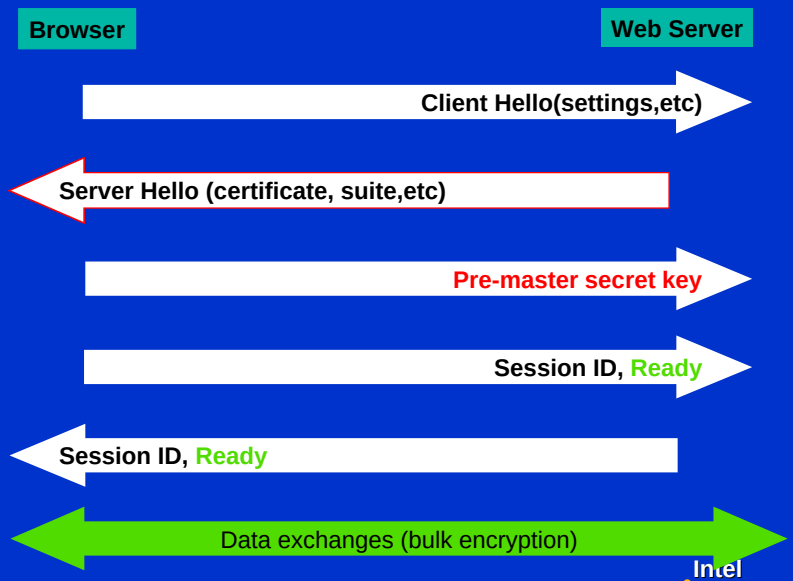
	Pentium®III Processor	Pentium® 4 Processor	Relative Improvement
Frequency	1 GHz	> 1.4 Ghz	> 1.4
Adder Speed	1 ns	< .36 ns	> 2.8
L1 Cache Speed	3 ns	< 1.42 ns	> 2.1
L1 Cache Size	16 KB	8 KB	0.5
L1 Cache Bandwidth	16 GB/sec	> 44.8 GB/sec	> 2.8
L2 Cache Bandwidth	16 GB/sec	> 44.8 GB/sec	> 2.8
Uop Fetch Bandwidth	3 billion/sec	> 4.2 billion/sec	> 1.4
Adder Bandwidth	2 billion/sec	> 5.6 billion/sec	> 2.8
Branch targets	512	4092	8
Instructions In flight	40	126	3.15
Loads in flight	16	48	3
Stores in flight	12	24	2

Example - Security and e-Commerce

- Secure transactions enable e-Commerce
- SSL is the standard for secure Web transactions
  - Protocol for secure communication
  - Built upon a core set of algorithms
    - Public-key encryption
      - RSA, DSA, Diffie-Hellman, etc.
    - Message digest
      - SHA-1, MD5, etc.
    - Digital signature
    - Bulk encryption
      - RC4, DES, 3DES, AES

Security Impacts Performance

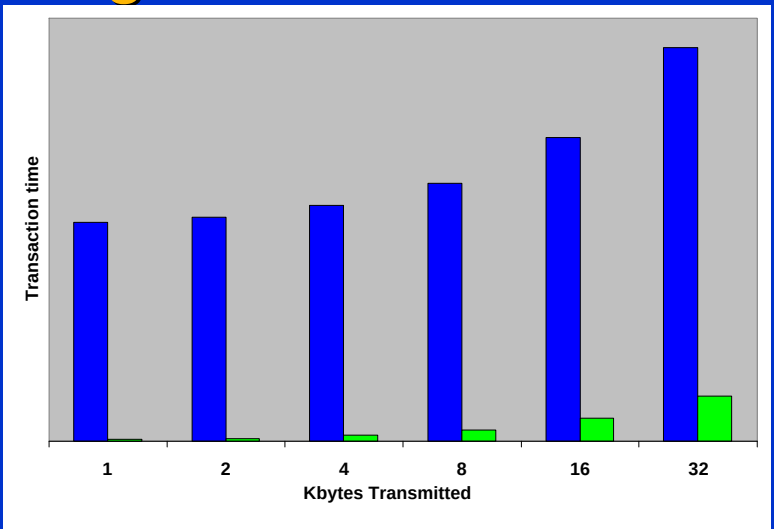
SSL - The basics



Security Impacts Performance

The high cost of SSL

Source - Intel Lab research



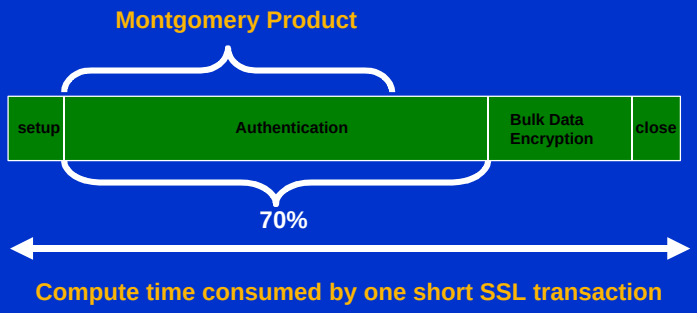
Secure transactions are orders of magnitude slower than non-secure

Intel Labs

Identify Key Algorithms

Computation in SSL

- Goal: Increase the number of secure transactions
  - Identify ***server*** performance issues in SSL
  - One server may deal with hundreds of clients



Compute time consumed by one short SSL transaction

Source - Intel Lab research

Intel Labs

Architectural Features

- New instructions in SSE2
  - PMULUDQ (32x32=>64)
  - PADDQ (64+64=>64)
  - PSHUFD (Re-arrange DWORDs)
- All pipelined
- SIMD

Increase size and reduce number of individual multiplications

Timings

Algorithm	Bits	Lang	Clocks	Ratio
Naïve 32-bit	1x16	C	51000	19.07
Optimized ASM using MUL	1x32	asm	28750	10.75
Using Pentium 4 New Instruct	2x32	asm	2675	1.00

Almost 20x performance gain versus naïve implementation

Summary

- Expect 400+ 1024-bit RSA Decrypts/second
- Breakthrough performance on public key algorithms for Intel® Pentium® 4 processor
  - The right architecture
  - The right instruction set

Pentium® 4 processor delivers more secure transactions to more users

Intel Labs