White Paper
Intel® Architecture

Extending the World's Most Popular Processor Architecture

New innovations that improve the performance and energy efficiency of Intel® architecture

R.M. Ramanathan
Intel Corporation

Primary Contributors
Ron Curry
Srinivas Chennupaty
Robert L. Cross
Shihjong Kuo
Mark J. Buxton
Intel Corporation

Introduction

Intel has a long history of innovation in adding new capabilities to computer architecture and enabling the industry to deliver advanced applications with greater performance and capability. From the original Intel® 8086 to the recent addition of Supplemental Streaming SIMD Extensions 3 (Supplemental SSE3) found in Intel® Core™2 Duo processors, Intel has led the charge in expanding the capabilities of the world's most popular and broadly used computer architecture—Intel® architecture. Continuing the history of innovation, this latest expansion of Intel architecture constitutes the most impactful instructions since SSE2 and represents the next major leap in Intel's fast-paced trajectory to deliver products with superior performance, capability, and energy-efficiency for years to come.

Building on the already rich Intel® 64 instruction set architecture (ISA), these new instructions will enable our microprocessors across all volume market segments to deliver superior performance and energy efficiency to a broad range of 32-bit and 64-bit applications. These new instructions include:

- **Streaming SIMD Extensions 4 (SSE4)** that will provide building blocks for delivering expanded capabilities, enhanced performance, and greater energy-efficiency for most applications.
- **Application Targeted Accelerators** that will provide a new foundation for delivering low-latency, lower power fixed-function capabilities for targeted applications.

These instructions represent another milestone in Intel's new cadence for the continuous development of next generation silicon processes and processor architecture. Applications that will benefit include those involving graphics, video encoding and processing, 3-D imaging, gaming, web servers, and application servers. High performance applications that will benefit include data mining; database; complex searching and pattern matching algorithms; audio, video, image, and data compression algorithms; parsing and state machine-based algorithms; and many more.

This paper will provide a brief background on ISA, and then give an overview of these new instructions, including SSE4 vectoring compiler and media accelerators, SSE4 efficient accelerated string and text processing, and Application Targeted Accelerators.

Leading the Instruction Set Revolution

Intel uses ISA to deliver the superior capabilities of its microarchitecture while maintaining the necessary application-level compatibility across processor generations. Good examples in maintaining instruction set compatibility include the new Intel Core 2 Duo processors. Like the previous generation Intel® Pentium® D processors, the Intel Core 2 Duo processors implement nearly identical versions of the ISA and provide application-level compatibility while having a different internal design. Nearly all applications built for Intel Pentium D processors will run on Intel Core 2 Duo processors without any modification. Even better, nearly all these applications benefit from the superior performance and energy-efficiency of these processors.

Just as Intel process technology and microarchitecture are continuously evolving at the pace of our new cadence, so are Intel instruction sets. In each new evolution:

1. Intel will optimize existing instructions to enable them to receive maximum benefit from the latest microarchitecture improvements and deliver greater performance and power efficiency to existing applications without modification.
2. Intel will also introduce new sets of instructions designed to optimize the performance and lower the power needs of a broad range of existing and new applications. To effectively get the benefit of these new instruction, existing applications will need to be recompiled with an updated compiler provided by Intel and other vendors. (See *www.intel.com/software* for more details.)

As you can see, in each case, existing software will continue to run correctly as our instruction sets evolve and new ones are added. Equally important, new applications incorporating these instructions—and existing applications recompiled to take advantage of them—will see exciting performance improvements.

Microarchitecture and Instruction Set Architecture

To better appreciate the significance of these new instructions, it helps to understand the different architectures used in developing today's modern microprocessors and their roles.

- **ISA** is the part of an overall computer's architecture related to programming, including the native data types, instructions, registers, addressing modes, memory architecture, interrupt and exception handling, and external I/O. An ISA includes a specification of opcodes (machine commands) implemented by a particular microprocessor design. Within a family of processors, ISA is often enhanced over time with new instructions to deliver superior performance and energy-efficiency while maintaining compatibility to already existing applications.
- **Microarchitecture** refers to the design, layout and implementation of ISA in silicon, including overall block design, cores, execution units and types (such as floating point, integer, branch prediction, SIMD), pipelining, cache memory design, and peripheral support. Within a family of processors, the microarchitecture is often enhanced over time to deliver improvements in performance, energy efficiency, and capabilities, while maintaining compatibility to ISA.

Intel's lead in ISA extends to a broad ecosystem of operating systems, including Microsoft Windows* and Vista*, UNIX*, Linux*, and now Macintosh* operating systems. Our continuing commitment to extending our ISA for the industry includes:

- Creating architectural consistency across all operating systems through extended industry ecosystem support.
- Providing a unified approach for both 32-bit and 64-bit extensions to deliver superior innovation.
- Listening to software developers and independent software vendors (ISVs) in our development of new instructions to help developers succeed more easily with us.
- Making sure existing applications run correctly and perform better.
- Ensuring applications that use the new instructions run correctly with increased performance and energy efficiency.
- Providing ISA leadership to other architecture vendors so that the Intel ISA remains unfragmented and performs as a standard, simplifying the job of the ISV community.

A Long History in ISA

Developers know that by increasing the number of instructions processed concurrently, they can reduce the amount of time that an application will spend on code requiring many processor cycles to process data. Intel has long encouraged such coding practices to help increase overall processor throughput.

Early on, Intel began a proactive program to improve application performance on Intel processors by developing special instruction sets. Early examples include the floating point (FP) instruction set extensions defined in the 8086 chip. More recent examples include Single Instruction, Multiple Data (SIMD) and Intel® MMX™ technology. SIMD was a technique employed by Intel to achieve increased parallelism in the P5 microarchitecture through the use of special instructions that operated on multiple pieces of data simultaneously. Using Intel MMX technology instruction set, programmers had the ability to execute instructions on multiple data elements loaded into MMX technology registers that would deliver increased performance in media applications such as graphics, gaming, streaming video, and more.

In the P6 microarchitecture, Intel introduced Streaming SIMD Extensions (SSE). Designed for the Intel® Pentium® III processor, SSE extended MMX technology and allowed SIMD computations to be performed on four packed single-precision FP data elements simultaneously using 128-bit registers (named XMM0-XMM7). With the Intel NetBurst® microarchitecture (Intel® Pentium® 4 processor), Intel introduced SSE2 to extend SSE (and MMX). SSE2 provided the ability to perform more computations in parallel by extending those instructions introduced in MMX technology and SSE, and enabling support of 128-bit integer and packed double-precision FP data types. In all, SSE2 added 144 instructions that delivered performance increases across a broad range of applications.

For instance, SSE2 instructions gave software developers maximum flexibility in implementing algorithms and providing performance enhancements to software such as MPEG-2 video, MP3, 3D graphics, and more.

Intel® Architecture (IA) Instruction Sets

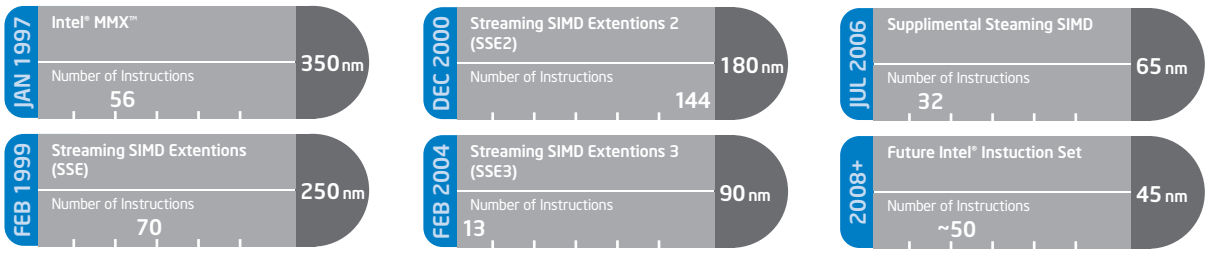
Intel has three different ISAs optimized for different market segments and applications. This enables us to provide leadership solutions from top to bottom in a variety of 64-bit and 32-bit configurations.

- **IA-64** is for the highest end servers and computing applications. It is the ISA for the Intel® Itanium® processor family.
- **Intel® 64** is aimed at clients or servers running mainstream applications that benefit from 64-bit computing. It is the ISA for:
  - Intel® Xeon® processors
  - Intel® Core™2 Duo processors
- **IA-32** is for clients running only 32-bit mainstream applications. It is the ISA for:
  - Intel® Celeron® and Intel® Pentium® processors with pin configuration FC-PGA2
  - Ultra-low voltage processors
  - Intel® Core™ Duo processors

It is important to note that Intel® 64 is a 64-bit ISA that is a superset of and compatible with IA-32 ISA. This newer ISA allows processors to run recently written 64-bit software and access larger amounts of memory than 32-bit software.

The launch of the 90 nm process-based Pentium 4 processor saw the introduction of SSE3. SSE3 includes 13 additional SIMD instructions over SSE2 that are primarily designed to improve thread synchronization and x87-FP math capabilities. A further advancement, Supplemental SSE3, is now available in Intel Core microarchitecture. Included in Intel® Xeon® 5100 processors (server and workstation) and the Intel Core 2 Duo processors (notebook and desktop) processors, Supplemental SSE3 adds 32 new opcodes—including align and multiply-add—for yet greater performance.

Recent Intel® Processor Instruction Set Additions



Overview of SSE4 for Intel Architecture

SSE4 is Intel's largest ISA extension in terms of scope and impact since SSE2. SSE4 has several compiler vectorization primitives for even greater and more efficient media performance, as well as new and innovative string processing instructions. Beginning with the 45 nm Intel microarchitecture-based processors (codenamed Penryn) slated for production in 2007,[1] these new instructions will start to appear in most of the volume market segments, including desktop, mobile, and server.[2]

Intel has worked closely with industry partners including independent software vendors (ISVs) and operating system vendors (OSVs) to develop SSE4 as a new instruction set standard. We have translated a wide range of ISV needs into the best set of instructions for optimizing the unique capabilities, performance, and power-efficiency benefits of Intel microarchitecture for their software.

SSE4 will offer dozens of new innovative instructions in two major categories:

- SSE4 Vectorizing Compiler and Media Accelerators
- SSE4 Efficient Accelerated String and Text Processing

Building on the Foundation of Intel® Core™ Microarchitecture

The move to multi-core processing has opened the door to additional microarchitectural and instruction-level innovations that can further improve performance and energy-efficiency. A microarchitectural example is Intel® Advanced Digital Media Boost in the Intel® Core™ microarchitecture. This advance significantly improves performance when executing SSE instructions. It accelerates a broad range of applications, including video, speech and image, photo processing, encryption, financial, engineering, and scientific applications. Intel Advanced Digital Media Boost enables most 128-bit instructions to be completely executed at a throughput rate of one per clock cycle, effectively doubling, on a per clock basis, the speed of execution for these instructions as compared to previous generations. This is an example of how microarchitecture and instruction sets work hand-in-hand and complement each other to deliver the benefits to the software.

Intel's success in designing and implementing performance and power- efficient ISA extensions such as SSE3 and Supplemental SSE3 is just the start. These new extensions extend the capabilities of Intel® architecture with several new innovations that will improve the performance and lower the power of a broad range of applications.

SSE4 Vectorizing Compiler and Media Accelerators

SSE4 adds several new compiler vectorization primitives (fundamental operations from which more complex operations can be constructed) that extend the capabilities of Intel architecture by enabling performance-optimized and lower power code generation. Compilers making use of these improved compiler vectorization primitives will be able to deliver these benefits to a broad range of applications, including media and high performance computing (HPC) server applications.

The new compiler vectorization primitives include improved integer and floating-point operations, support for packed DWORD and QWORD operations, new single precision FP operations, fast register operations, performance-optimized memory operations, and more.

Applications that will benefit include those involving image processing, graphics, video processing, 2-D/3-D generation, multimedia, gaming, memory-intensive workloads, HPC workloads, and more.

Sub Group	Instructions	Description	Expected Application Benefits
Packed DWORD Multiplies	PMULLD, PMULDQ	New support for four signed or unsigned 32x32 bit multiplications per instruction, as well as signed forms of 32x32->64 multiplication.	Broadly useful for improved automated compiler vectorization of data processing written in high level languages (like C and Fortran).
Floating Point Dot Product	DPPS, DPPD	Improved performance for AOS (Array of Structs) data processing through support for single and double-precision dot products.	3-D content creation, gaming, and support for languages like CG and HLSL.
Packed Blending	BLENDPS, BLENDPD, BLENDVPS, BLENDVPD, PBLENDVB, PBLENDDW	Blending conditionally copies one field in the source onto the same field in the destination. These new instructions improve the performance of blending operations for most field sizes through packing multiple operations in a single instruction.	Broadly useful for automated compiler vectorization of data processing written in high level languages (like C and Fortran), and applications such as image processing, video processing, multimedia, and gaming.
Packed Integer Min and Max	PMINSB, PMAXSB, PMINUW, PMAXUW, PMINUD, PMAXUD, PMINDS, PMAXSD	Compares packed signed/unsigned byte/word/dword integers in the destination operand and the source operand, and returns the minimum or maximum as per the instruction type for each packed operand in the destination operand.	Broadly useful for automated compiler vectorization of data processing written in high level languages (like C and Fortran), and applications such as image processing, video processing, multimedia, and gaming.
Floating Point Round	ROUNDPS, ROUNDSS, ROUNDPD, ROUNDSD	Efficiently rounds the scalar and packed single- and double- precision operands to integers, with enhanced support for Fortran, JAVA and C99 language requirements.	Image processing, graphics, video processing, 2-D/3-D applications, multimedia, and gaming.
Register Insertion/Extraction	INSERTPS, PINSRB, PINSRD, PINSRQ, EXTRACTPS, PEXTRB, PEXTRD, PEXTRW, PEXTRQ	These new instruction simplify data insertion and extraction between GPR (or memory) and XMM registers.	Broadly useful for automated compiler vectorization of data processing written in high level languages (like C and Fortran), and applications such as image processing, video processing, multimedia, and gaming.
Packed Format Conversion	PMOVSXBW, PMOVZXBW, PMOVSXBD, PMOVZXBD, PMOVSXBQ, PMOVZXBQ, PMOVSXWD, PMOVZXWD, PMOVSXWQ, PMOVZXWQ, PMOVSXDQ, PMOVZXDQ	Converts from a packed integer (from XMM register or memory) to a zero- or sign-extended integer with wider type.	Broadly useful for automated compiler vectorization of data processing written in high level languages (like C and Fortran), and applications such as image processing, video processing, multimedia, and gaming.
Packed Test and Set	PTEST	Faster branching from SIMD decisions to support conditionally vectorized code.	Useful for improved automated compiler vectorization of data processing, image and video processing, 3-D content creation, multimedia, and gaming.
Packed Compare for Equal	PCMPEQQ, PCMPGTQ	Performs SIMD compare for equality of the packed QWORDs in the destination and the source operand.	Broadly useful for automated compiler vectorization of data processing written in high level languages (like C and Fortran), and applications such as image processing, video processing, multimedia, and gaming.
Pack DWORD to Unsigned WORD	PACKUSDW	Converts packed signed DWORDs into packed unsigned WORDs using unsigned saturation to handle overflow condition. This new instruction completes the set of other instructions in this type.	Broadly useful for automated compiler vectorization of data processing written in high level languages (like C and Fortran), and applications such as image processing, video processing, multimedia, and gaming.

SSE4 Efficient Accelerated String and Text Processing

SSE4 provides new string and text processing instructions that will enhance the performance of string and text processing operations, resulting in a performance boost for a wide variety of data processing, search, and other text-based applications. These new instructions will include advanced packed string comparison instructions that can perform multiple compare and search operations in a single instruction. In general, each of these new instructions has a rich set of innovative string processing capabilities to replace operations in which several instructions were required to deliver the same functionality in the previous ISA.

Applications that will benefit include those involving databases, text search, virus scanning, string process libraries like ZLIB, Token parsing/recognizing applications like compilers, and state machine-oriented applications.

Sub Group	Instructions	Description	Expected Application Benefits
Advanced String Operations	PCMPESTRI, PCMPESTRM, PCMPISTRI, PCMPISTRM	These new instructions provide a rich set of string and text processing capabilities that traditionally required many more opcodes.	Improved performance for virus scan, text search, string processing libraries like ZLIB, databases, compilers and state machine-oriented applications.

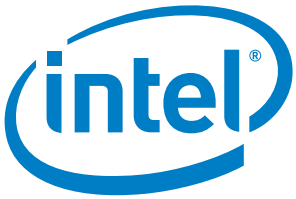
Overview of Application Targeted Accelerators

Application Targeted Accelerators extend the capabilities of Intel architecture by adding performance-optimized, low-latency, lower power fixed-function accelerators on the processor die to benefit specific applications. Such accelerators are the start of a natural evolution of adding advantageous implementations of fixed-function capabilities to the processor. Just as the evolution of silicon technology from 65 nm to 45 nm to 32 nm will enable more transistors for additional cores and cache, so too will it also enable these fixed-function on-die implementations. The benefit will be greater performance—and superior energy efficiency—in processing specific applications.

The first set of Application Targeted Accelerators will accelerate the cyclic redundancy check (CRC) of several data integrity applications. This new CRC instruction will deliver processor-based CRC for fast, efficient data integrity checks at lower cost than separate dedicated chips in upper layer data transfer protocols like Internet Small Computer System Interface (iSCSI) and Remote Direct Memory Access (RDMA) where CRCs play an important role in error detection but are also one of the biggest bottlenecks. Processor-based CRC will enable enterprise-class data assurance with high data rates in networked storage in any user environment. Without this new instruction, service providers would have to incorporate very expensive, power-consuming accelerator cards to deliver the same benefits. With the power of Intel multi-core processors based on Intel Core microarchitecture, this new CRC instruction will accelerate the performance of targeted network protocols like iSCSI and RDMA without adding additional cost. This will help enable the spread of low-cost storage area networks based on iSCSI solutions. Such networks provide an important alternative to installing much more expensive fibre channel networks and will help a wide range of businesses inexpensively solve their data storage issues.

Our second application-targeted extension provides a single instruction, POPCNT, that can be effectively used to accelerate searches involving large data sets. It works by counting the number of set bits in a data object. Applications that could benefit from this instruction include those involving genome mining, handwriting recognition, digital health workloads, and fast hamming distance/population count.

Sub Group	Instructions	Description	Expected Application Benefits
Fast CRC (Cyclic Redundancy Check)	CRC32	Finds the CRC value using a specific polynomial of a given source operand.	Fast and efficient data integrity checks in data transfer protocols for networked storage (e.g., iSCSI, RDMA).
Accelerated searching and pattern recognition of large data sets	POPCNT	Calculates the number of bits set to 1 in the given operand.	Helps to deliver higher performance in applications such as genome mining, handwriting recognition, digital health workloads, fast hamming algorithms, and others.



www.intel.com

Summary

As the largest and most impactful ISA extensions since SSE2, SSE4 and Application Targeted Accelerators are an important milestone in the Intel's fast-paced trajectory to deliver products with superior performance, energy-efficiency and expanded capabilities for years to come. Intel's leadership and ongoing work in the development of instruction set extensions for Intel architecture provide a continuing path for enhancing the performance, power efficiency and capabilities of a wide range of software. With SSE4 and Application Targeted Accelerators, we're continuing to work with ISV community to deliver instruction set extensions that truly enhance the ability of their products to provide real benefits (everything from improved performance to substantial cost savings) to their customers.

Links

www.intel.com/technology/architecture/new_instructions.htm
www.intel.com/technology

References

Intel® Core™ Microarchitecture
www.intel.com/technology/architecture/coremicro

1. Intel has not yet announced launch dates for 45 nm products.
2. Most of these instructions will be available in Penryn and some of the instructions will be available in microprocessors slated for release after Penryn.

 0906/RMR/HBD/PDF 315383-001US