

transactions per cycle to sustain the remaining cache bandwidth. That means, on average, either the up or down leg must handle two transactions at once, which may be difficult to schedule. Thus, the ring may not be able to sustain the full bandwidth of the caches, but it will provide much greater bandwidth than Nehalem's internal bus.

The architecture easily scales downward by removing two of the CPUs and their associated cache blocks. This change reduces both the size and bandwidth of the L3 cache to match the reduction in CPU count. With only two cache blocks, the ring can sustain the full cache bandwidth. Alternatively, the architecture can be extended to eight or more CPUs. In this case, Intel could double the ring width to 512 bits to help support the greater bandwidth of the additional cache blocks.

Figure 5 shows a die photo of the chip in which the six ring stations can clearly be seen. Intel did not disclose the die size of the chip, but it is said to be about 225mm². The graphics and video unit at the bottom of the chip takes 20% of the total die area. This unit appears to consist mainly of synthesized logic, whereas the CPUs and caches use custom circuit design to maximize performance. The north-bridge logic, not counting the I/O drivers, is only 8% of the die. The four CPUs together use 33% of the die, making them just under 20mm² each. The 8MB of L3 cache (L3\$) covers 20% of the die. Most of the I/O drivers are for the two 64-bit channels of DDR3 SDRAM. Note that removing two of the CPUs and two of the cache blocks would reduce the total die area by about 30%.

Reving Up Turbo Mode

As with Intel's other multicore processors, the frequency of each CPU core can be adjusted individually, so it can be

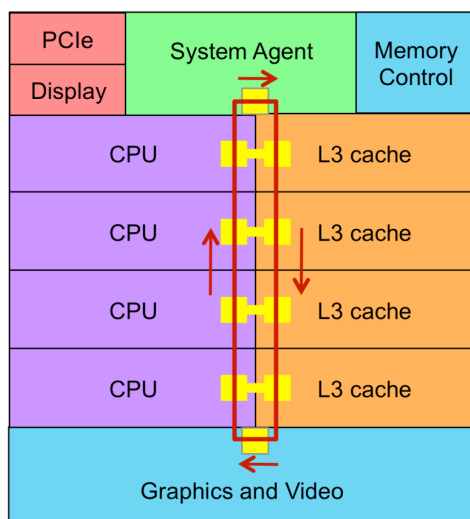


Figure 4. High-level block diagram of Sandy Bridge. The L3 cache is divided into four blocks, which are connected to the four CPU cores, the GPU, and the system agent through a ring interconnect.

turned on or off as needed. Even when its clock is turned off, however, the CPU burns leakage power. So, when the processor is truly idle, the voltage to the CPUs can be turned off, completely eliminating their power dissipation. The graphics unit is on a separate voltage plane, so it can be disabled for applications that do not require graphics. The system agent and north-bridge logic are on a third power plane, which allows the north bridge to refresh the display while the rest of the processor is asleep.

The system agent contains the power manager—a programmable microcontroller that controls the power state of all units and controls reset functions. This approach allows power-management decisions to be made more quickly than by running a driver on one of the CPUs.

If only one Nehalem CPU is active, it is allowed to exceed its rated speed, since even at a higher speed, one CPU will not exceed the multicore chip's total TDP budget. Intel calls this feature turbo boost. Sandy Bridge takes a different view of turbo boost, focusing on die temperature rather than instantaneous power dissipation.

Consider a processor that has been sitting idle for several seconds; perhaps the user is considering what digital effect to apply to a photograph. The processor and its thermal system (e.g., heat sink) will cool down. If the user

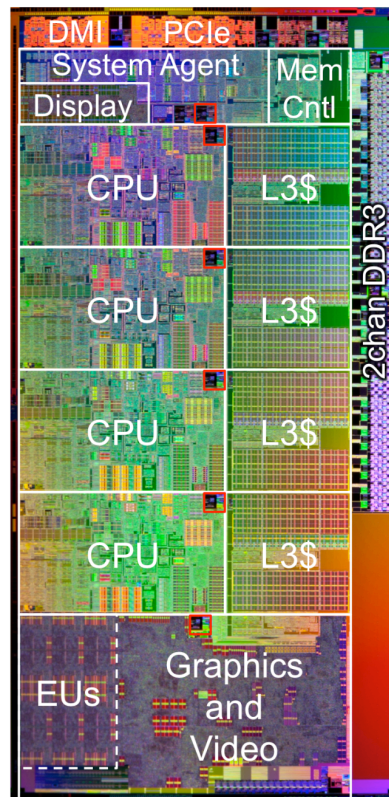


Figure 5. Sandy Bridge die photo with overlay. The red boxes show the ring stations. The 12 shaders (EUs) can be seen at the lower left. (Source: photo by Intel, overlay by The Linley Group)

then begins a CPU-intensive activity, the entire processor can operate well above its rated TDP for quite some time (up to 25 seconds, Intel claims) while the thermal system heats back up. Once the chip reaches its maximum operating temperature, it must throttle back to its rated TDP level to avoid overheating.

Sandy Bridge's power manager estimates the chip's temperature and calculates the degree and duration of the turbo boost. In this design, the maximum CPU speed in turbo mode is limited only by the physical speed of the transistors and the current that can be supplied to the chip. Turbo mode should be particularly effective in mobile Sandy Bridge chips, since these chips have more headroom. Because the chip does not physically measure its own temperature, the power manager estimates the temperature on the basis of recent workloads and the thermal characteristics of the device and its cooling system. Presumably, Intel has built in a safety factor, but any unauthorized or even unintentional modification of the cooling system could wreak havoc with the power manager's calculations.

The integration of GPU and CPUs on one chip creates additional turbo opportunities. If the GPU is not in use or is operating at a low speed, the power manager can assign some of its thermal budget to the CPUs, allowing them to run at a higher clock speed. Conversely, on a graphics-intensive application, the power manager can overclock the GPU while slowing down the CPUs.

Turbo mode should provide a noticeable speedup to PC users—up to 40% in some situations. Users who often leave their system idle or in a low-power state before launching a brief burst of computation will particularly benefit. Gamers or professional users who execute long CPU-intensive jobs will not see as much gain. Benchmarking a Sandy Bridge processor could generate widely different results depending on how the test is run. Intel recommends allowing the processor to rest quietly for a

moment before running each benchmark, but this approach may not be indicative of real-world usage.

Sandy Bridge Transition in 2011

As with previous generations, Sandy Bridge will sweep through Intel's entire product line, but the transition will take some time. The initial products, which are due to appear early next year, will be in the Core i3, Core i5, and Core i7 product lines. Although Intel has not announced details, leaked roadmap slides provide the initial product configurations. According to this information, these products will offer two or four CPUs at speeds of up to 3.4GHz. The high-end 3.4GHz four-core version will top out at 95W TDP. Mobile versions with four cores at 2.5GHz will use 55W TDP, with a dual-core 2.7GHz version rated at just 35W TDP. Some models (probably the dual-CPU versions) will use 6 shaders instead of 12, reducing graphics performance.

Together, these Sandy Bridge products will quickly displace the dual-core Arrandale (mobile) and Clarkdale (desktop) processors, as Figure 6 shows. Some high-end PCs are still using the quad-core Clarkfield (mobile) and Lynnfield (desktop) processors, which are built in 45nm technology; these products can convert to the quad-core Sandy Bridge processors.

Sandy Bridge will cross into the server market in 2H11 using a platform called Romley. At IDF, Intel demonstrated Romley-EP, a dual-socket (2P) platform that succeeds Tylersburg-EP, using a version of Sandy Bridge with eight CPUs. An eight-CPU version uses more die area than the six-CPU Westmere chip, but Sandy Bridge's lower power per CPU should allow it to fit into a similar thermal budget (TDP).

For 4P and 8P systems, Intel is preparing the Romley-EX platform, which follows Westmere-EX (see [MPR 9/20/10-01](#), "Intel Unveils Atom E6xx, Westmere-EX") and the Boxboro-EX platform. The new high-end platform includes the Sandy Bridge-EX processor, which we expect to have 12 CPUs.

Intel has disclosed little about Romley but says it will use a new version of Intel's Quick Path Interconnect (QPI) to connect the processors to each other. Sandy Bridge-EP will include two QPI v1.1 links for this purpose. Intel has not specified the speed of the new version, but QPI v1.0 delivers 12.8GB/s in each direction at a clock rate of 3.2GHz. We expect the new links provide a modest increase, perhaps to 3.8GHz. The processor will still use DMI to connect to the south bridge, which in the Romley platform is code-named Patsburg.

We expect the server versions of Sandy Bridge to have some minor changes in addition to the greater number of CPUs. The on-die graphics unit will probably be omitted. Whereas Westmere-EP has three DRAM channels for six CPUs, Sandy Bridge-EP could use four channels to provide a better match for its eight CPUs. Although the initial

	2010	2011
Laptop	Arrandale	Mobile Sandy Bridge
Desktop / 1P Server	Clarkdale	Sandy Bridge
2P Server	Westmere-EP	Sandy Bridge-EP
	Tylersburg-EP	Romley-EP
4P Server	Westmere-EX	Sandy Bridge-EX
	Boxboro-EX	Romley-EX

Figure 6. Intel processor roadmap. In early 2011, PC users will convert from Intel's Arrandale and Clarkdale processors to Sandy Bridge. In 2H11, servers will convert to the Romley platform, which supports new versions of Sandy Bridge.

Sandy Bridge processors will integrate PCI Express 2.0, the server versions provide a good opportunity for Intel to debut the faster PCI Express 3.0. The new specification, which is due to be completed in late 2010, doubles PCI Express throughput to 8Gbps per lane, or 16GB/s for a 16-lane slot.

Intel has already taped out its first 22nm processor. Following the tick-tock plan, Ivy Bridge is a 22nm shrink of Sandy Bridge that is scheduled to enter production in late 2011. As a “tick,” Ivy Bridge is not expected to add many new features to the platform, but the 22nm process will significantly reduce power per CPU and enable Intel to put more CPUs on a chip. Following Ivy Bridge is Haswell, the next tock. Little is known about Haswell, but it is said to be a completely new CPU microarchitecture.

Intel Steamrolls Bulldozer

AMD “coincidentally” disclosed its next-generation Bulldozer architecture just one month before Intel disclosed Sandy Bridge. Thus, it is natural to compare the two designs. Although AMD has beefed up its microarchitecture, Bulldozer matches up better against Nehalem/Westmere than it does against Sandy Bridge, as Table 2 shows. Intel has again upped the ante by expanding the reorder capabilities of its CPU. Sandy Bridge’s micro-op cache is a unique feature that offers performance and power-savings opportunities. We also suspect Intel’s branch prediction is superior, although neither vendor has disclosed enough information to be sure. Bulldozer’s 16KB data cache may also hinder it on certain single-thread applications.

On dual-threaded applications, Bulldozer’s unique two-core module offers clear benefits over Intel’s Hyper-Threading design. Sandy Bridge must share its execution units and reorder buffer between the two threads, whereas Bulldozer offers each thread a separate set of resources. But AMD counts each Bulldozer module as two cores, so all we can say is that two Bulldozer cores are better than one Sandy Bridge.

Although AMD has been talking about integrating graphics for years, Intel has actually been doing it, first with the two-chip module of Arrandale and Clarkdale and then with the single-die Sandy Bridge. AMD’s first Fusion chip, code-named Ontario (see [MPR 8/30/10-02](#), “Bobcat Snarls at Atom”), is scheduled to debut in 4Q10, only slightly before Sandy Bridge. We do not expect the first Bulldozer processors to integrate graphics. Sandy Bridge should negate any gains AMD might have achieved from its Fusion project.

Sandy Bridge products will appear in early 2011, whereas we expect the first Bulldozer products to ship in mid-2011. Although this schedule gives Sandy Bridge a head start of several months in the desktop and mobile markets,

Bulldozer is headed straight for the server market in the Valencia and Interlagos processors. Thus, Bulldozer may actually appear in servers before Sandy Bridge does (in the Romley platform).

Whereas Sandy Bridge is the clear winner on a core-to-core basis, a Bulldozer core will use much less power, allowing AMD to operate Valencia at a higher clock speed within the same thermal envelope. Sandy Bridge-EP should handily outperform Valencia on SPECint per watt, although the Bulldozer chip may narrow the gap from current products.

Rising to the Next Level

Sandy Bridge brings several architectural innovations to Intel’s platforms. PC users will see the biggest improvements in multimedia performance, particularly video encoding but also graphics and image processing. On these types of applications, which are becoming increasingly popular, users will see an improvement of 2× or better compared with Intel’s previous graphics solutions.

The CPU microarchitecture improvements are fairly subtle, given the evolutionary design approach. We expect performance per megahertz to improve very modestly, perhaps by 10% on most applications. Many of the design changes, particularly the L0 micro-op cache, aim to reduce power more than to increase performance. Intel has not quantified these gains, but we estimate they could trim CPU power by 10% at the same clock speed. Because most of the company’s products are power limited, Intel can turn power efficiency into higher clock speeds at the same TDP levels, providing an effective performance boost of 20% for Sandy Bridge compared with Westmere. In certain situations, the improved turbo boost will magnify this gain by as much as 40%, which can translate directly into longer battery life for notebook computers.

	AMD Bulldozer	Intel Westmere	Intel Sandy Bridge
Building Block	Dual CPU	CPU core	CPU core
Integrated Graphics?	Not disclosed	In package	On chip
Instruction Extensions	SSE1–SSE4.2, AES, AVX	SSE1–SSE4.2, AES	SSE1–SSE4.2, AES, AVX
Number of Threads	2 threads	2 threads	2 threads
Instr Decode Rate	4 per cycle	4 per cycle	4 per cycle
Micro-Op Issue Rate	4 per cycle*	4 per cycle	4 per cycle
Instr Reordering	128 micro-ops*	128 micro-ops	168 micro-ops
FP Operations (SP)	16 per cycle	8 per cycle	16 per cycle
L0 (micro-op) Cache	None	None	1.5K micro-ops
L1 Cache (I/D)	64KB/16KB†	32KB/32KB	32KB/32KB
L2 Cache	Undisclosed	256KB	256KB
IC Process	32nm	32nm	32nm
First Production	Mid-2011‡	Early 2010	Early 2011

Table 2. Comparison of AMD Bulldozer and Intel Sandy Bridge CPUs. Greater instruction reordering and an L0 cache should improve Sandy Bridge’s performance over Bulldozer’s. *Dedicated amount for each thread; †L1 data cache dedicated for each thread. (Source: vendors, except ‡The Linley Group estimate)

For More Information

Intel has not yet announced any products using the Sandy Bridge architecture. We expect the first Sandy Bridge products to ship in early 2011. For more information on Sandy Bridge, download Intel's IDF presentations from www.intel.com/go/idfsessions (search for Sandy Bridge). For complete information on the AVX extensions, access <http://software.intel.com/en-us/avx>.

The integration of graphics has several effects. It increases performance while greatly reducing power: we expect a performance gain of about 2x at one-fifth the power compared with the Arrandale/Clarkdale graphics unit. Integration is likely to reduce manufacturing cost compared with the current two-chip module (or with an external GPU chip). Graphics integration also makes PC makers less likely to choose a third-party graphics solution. Although the performance of Intel's graphics has generally been modest at best, the improvement in Sandy Bridge makes it acceptable for most PC users. Gamers and others who care about maximizing graphics performance will still opt for an external graphics card.

Intel's rollout schedule brings Sandy Bridge to the PC market in early 2011, but it will not reach the server market until late that year. In targeting the PC market first, Intel satisfies the bulk of its customer base and should extend its lead over AMD in PCs. But AMD's focus on the server market may allow Bulldozer to beat Sandy Bridge to that market. Furthermore, AMD's strategy of making Bulldozer a smaller, lighter core may prove more power efficient than Intel's bulkier Sandy Bridge. AMD has been successful with server customers by undercutting Intel on price; Bulldozer may provide it with an additional advantage in that market.

Intel's track record of bringing new processor architectures to market is impressive and unrivaled. No other processor vendor can completely remake its product line on an annual basis, yet Intel has done just that in each of the past four years. With Sandy Bridge, the company is poised to go five for five. The new processor shows that Intel is thinking outside of the CPU box, delivering new graphics and multimedia capabilities to support emerging usage models. Taking advantage of the growing transistor budget provided by its industry-leading fab technology, the company is integrating new features along with additional system functions, improving efficiency. Sandy Bridge will carry Intel into the next generation of PCs. ♦

To subscribe to *Microprocessor Report*, access www.MPRonline.com or phone us at 408-945-3943.