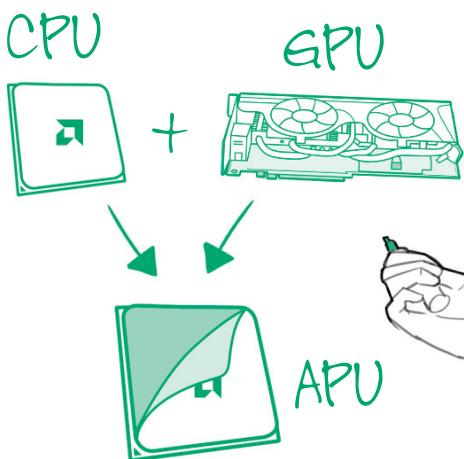


APU 101:

All about AMD Fusion Accelerated Processing Units

AND HOW THEY'RE CHANGING,
WELL, EVERYTHING.



A reference for anyone interested
in APUs and why they're important,
from the experts at AMD.





A new era in computing starts now!



An APU is an entirely new kind of processor that combines the advantages of a central processing unit (CPU) and a graphics processing unit (GPU). The result is the world's first high-performing, power-efficient and balanced APU platform ever, paving the way for applications in areas like gesture and facial recognition, search and indexing, and dynamic video enhancement that have only been imagined before. Today, only AMD has both the CPU and GPU technology leadership necessary to successfully bring this game-changing new type of processor to reality.

That's quite a set of claims, we know. And to understand why any of it is true or possible, you have to see it in the context of the history of computing and the cycle of innovation that has taken both serial and parallel processor performance to new levels over and over and over again.

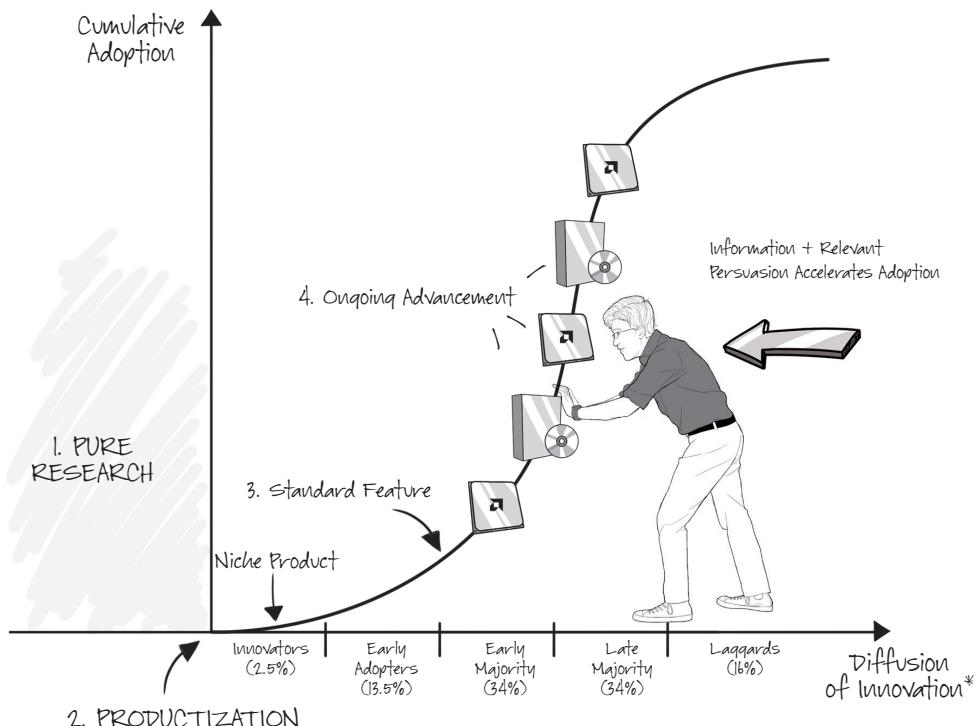
The Quest: Get Technology Out of Our Way

Creating new and better experiences that remove technological barriers is invariably what drives processor design. As semiconductor manufacturing technology enables the production of more and more transistors at a progressively lower cost, the design engineer's perpetual task is to find more ways to use those transistors to increase computing performance. The ultimate goal is to improve existing experiences – and to enable new ones – while constantly driving power consumption down.

Over the last 35 years, we've seen processor performance rise over and over again to overcome old limitations, enabling new levels of PC computing and unprecedented software innovations. Our exploration of APUs starts with a look at this history of innovation and the milestones along the way.

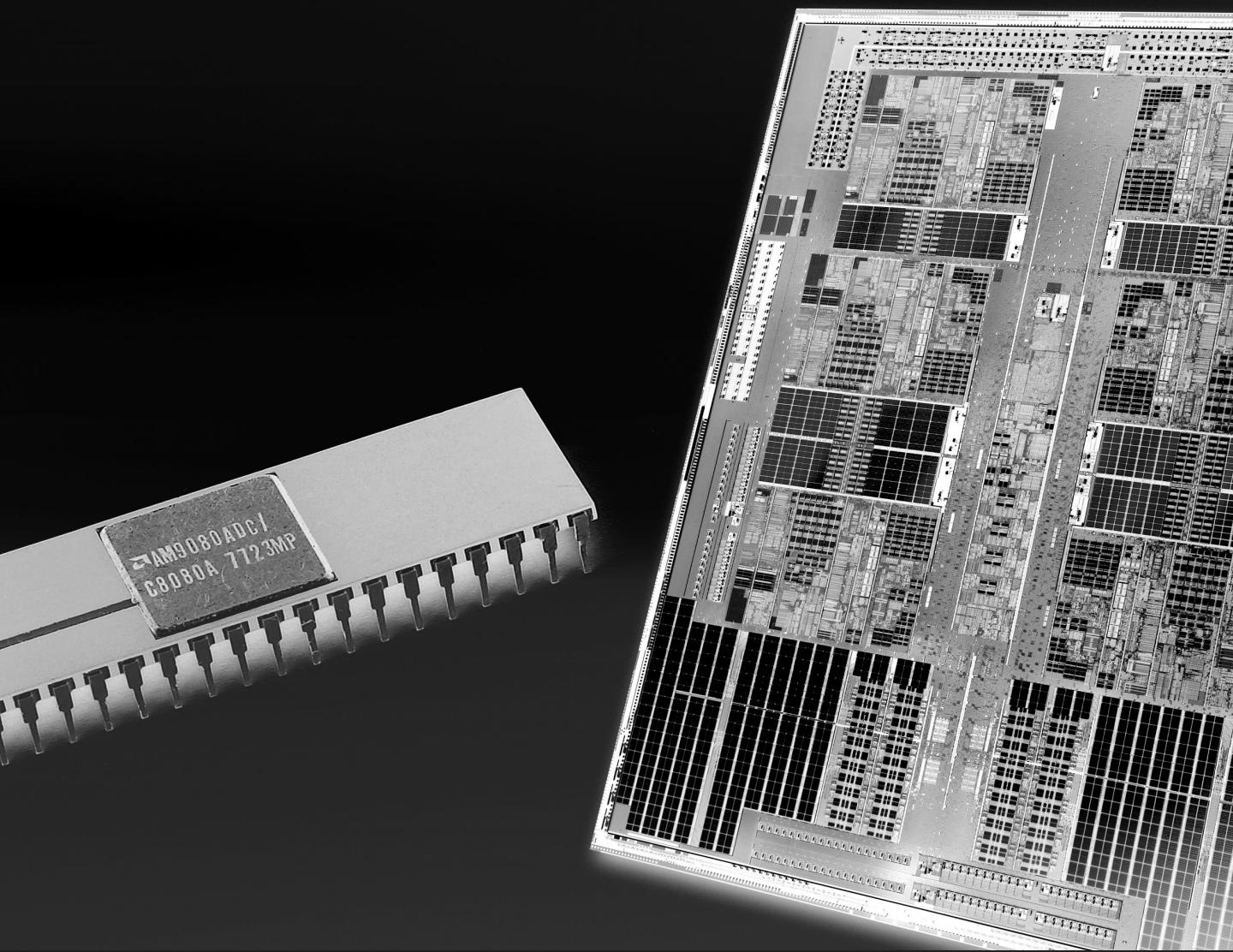
During the journey, there have been game-changing innovations. The development of the first CPU. The introduction of the GPU. The shift to multi-core computing. This is one of those game-changing times, as heterogeneous computing and the world's first and only Accelerated Processing Unit with a true, latest generation DirectX® 11 GPU take center stage in mainstream computing to push hardware and software performance to previously unimaginable levels.

FIG. 1
THE ROAD TO MAINSTREAM ADOPTION



*Everett M. Rogers, Diffusion of Innovations, 4th edition, New York Free Press, 1995

COMPUTING PLATFORM INNOVATION: THE STORY SO FAR

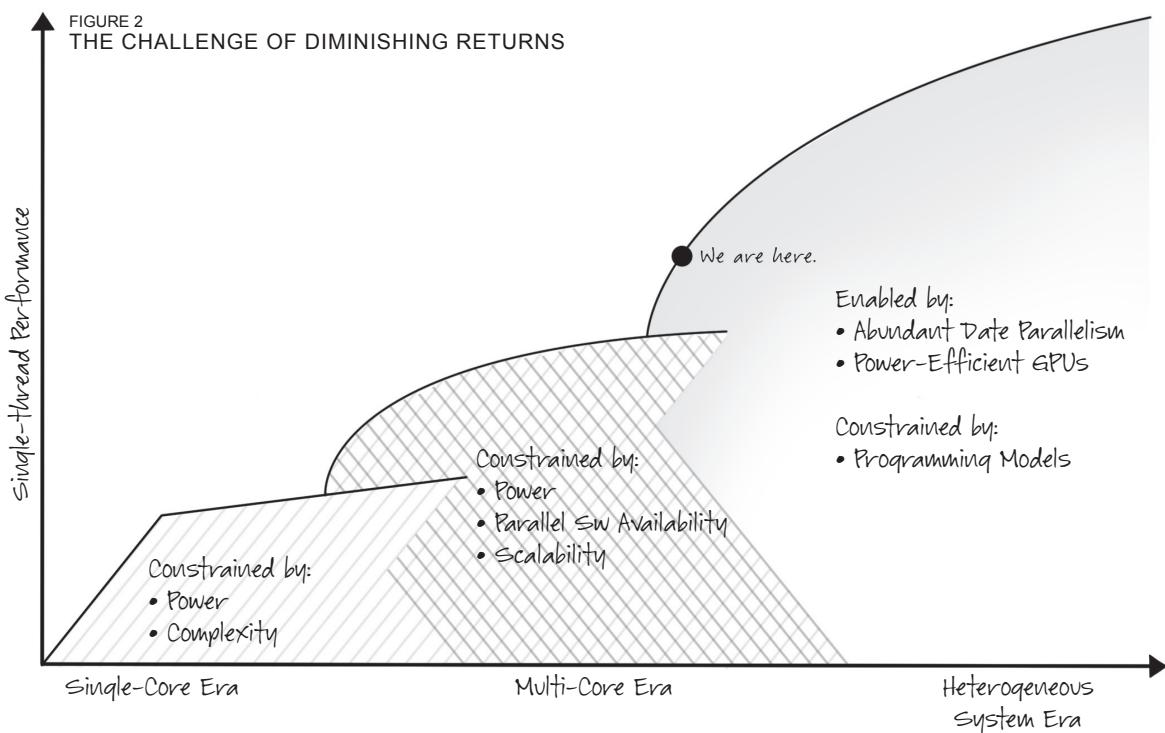


As Figure 2 illustrates, the quest for performance in processor design has taken computing through three distinct eras since the launch of the first PC – ultimately creating the environment in which APUs are now poised to take computing to new heights.

Each era has posed a common challenge that gives rise to the next era: how to cope with inherent technology constraints that invariably lead to a diminishing return on the investment in performance. At that point, it becomes impossible to bring to market practical, affordable, high-performance solutions.

Single-Core Era: Launching the Personal Computing Revolution

It's easy to take office software for granted today, but years ago, the automation of activities like generating documents and creating spreadsheets revolutionized business operations everywhere. Single-core microprocessing made it possible to develop self-contained, relatively compact applications for these and other repetitive, time-consuming tasks.



Single-core processing relied on frequency scaling to achieve gains in performance. Over time, these increases in processor frequency led to more complex office software and other programs, as well as multitasking operating environments like Microsoft Windows. But as processor frequencies increased, so did their heat dissipation and power requirements. This ultimately led to untenably high energy and hardware cooling costs to maintain the pace of performance improvements year over year with a single core.

The era drew to a close with single-core microprocessing offering sufficient desktop performance for general business productivity applications and basic Internet browsing. There were even early efforts to enhance performance with the very first dedicated graphics processing units, or GPUs, which offload the task of PC display drawing from the main CPU.

But despite these gains, single-core microprocessors would always be limited by power issues. In particular, they could never support high performance in notebooks and other small-form-factor computers with far more restrictive power require-

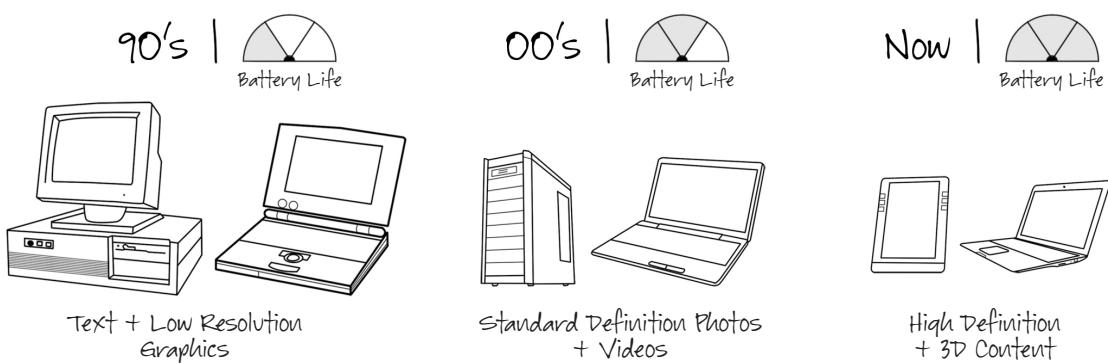
ments than desktop units. It was time for processor designers to find a new way to move forward.

Multi-Core Era: Fighting for More Performance on Two or Three (or Four) Fronts

To transcend the limitations of frequency scaling, designers shifted to creating multiple processor cores in a single design as a way of boosting performance. Multi-core processors increase performance by using additional cores to multiply the number of instructions executed – by two times, for example, with dual-core processors, or up to six times with the multiple cores in high-end desktop processors.

At the same time that multi-core processing was gaining favor, advances in semiconductor technology made it possible to achieve relatively high frequencies at power levels low enough to be practical in environments where power and heat generation would ordinarily be a barrier. These include notebooks and other small-form-factor computers, and home theater or media center PCs.

FIGURE 3
THE CHALLENGE: MAKING GREAT THINGS
COME IN SMALLER AND SMALLER AND
SMALLER PACKAGES



Heterogeneous Computing Defined

As the name implies, heterogeneous computing refers to systems that use more than one kind of processor. These are multi-core systems that gain performance not just by adding cores, but also by incorporating specialized processing capabilities to handle particular tasks. Within the context of this particular discussion, an APU is a heterogeneous system that incorporates DirectX 11 discrete-level GPU capabilities for graphics processing and other mathematically intensive computations on very large data sets, to handle visual tasks such as 3D rendering as well as certain fixed functions. Meanwhile, the APU continues to utilize a CPU to run the operating system and most traditional PC productivity applications.



By the end of 2010, nearly all new desktop computers had multi-core processors, with dual-core and even quad-core processors entering the mainstream of affordable computing. Still, multi-core processing posed some challenges of its own; the extra cores and cache memory required to fuel their instruction pipelines came at a cost of both increased processor size and, again, high power consumption.

Meanwhile, the multi-core era also saw some interesting developments in GPUs, which were growing in sophistication and complexity, spurred on by advances in semiconductor technology. GPUs have vector processing capabilities that enable them to perform parallel operations on very large sets of data – and to do it at much lower power consumption relative to the serial processing of similar data sets on CPUs. This is what allows GPUs to drive capabilities such as incredibly realistic, multiple-display stereoscopic gaming. And while their value was initially derived from the ability to improve 3D graphics performance by offloading graphics from the CPU, they became increasingly attractive for more-general purposes, such as addressing data-parallel programming tasks.

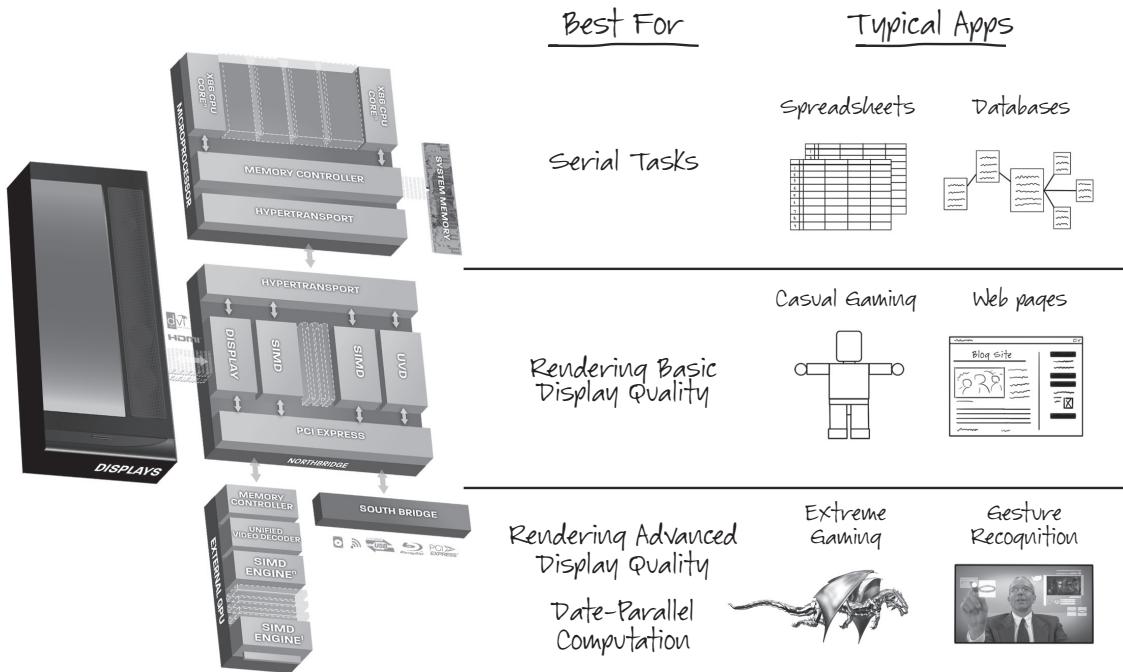
The early efforts to leverage GPUs for general-purpose computing coincided with a notable shift in consumer culture. There was a dramatic increase in the availability and quality of digital content, coupled with an increasing consumer appetite for rich visual experiences like video playback and viewing content in HD. At the same time, the emergence of mainstream operating system support for advanced multitasking began to require processing efficiency of an entirely new magnitude.

Heterogeneous Era: Setting the Stage for a Fantastic Leap Forward

The drive to improve performance and the continuing constraints on power and scalability in multi-core CPU development have led semiconductor, software and systems designers increasingly to look to the vector processing capabilities of GPUs.

Vector processors like those in advanced GPUs have up to thousands of individual compute cores, called shaders, which operate simultaneously. This makes GPUs ideally suited for computing tasks that deal with a combination of very large data sets and intensive numerical computation requirements.

FIGURE 4
THE BEGINNING OF BEAUTIFUL RELATIONSHIP



(And GPUs are rapidly advancing to do even more for less. For example, AMD's plans for its performance-class of APUs launching in 2011 will have many times the number of GPU processing cores as the low-power APUs.)

But – and this is a very important “but” – vector processing isn't the answer all the time. For example, small data arrays have overhead associated with setting up vector processing that can easily outweigh the time saved. That's why the scalar approach used by CPUs is still best for certain problems and algorithms. And that's why heterogeneous computing, which brings together the best of both CPUs and GPUs, is essential to driving faster and more powerful processor designs for new and better experiences.

The possibilities created by heterogeneous computing are truly fantastic – from flawless HD videoconferencing to heretofore unimaginable display clarity to real-time language translation and interpretation – all in lower and lower power envelopes for smaller and smaller form factors with longer and longer battery life. We'll look at all these and more – many of which are on the verge of realization – in the next section of this book.

HETEROGENEOUS COMPUTING:

BIG EXPERIENCES IN SMALL PACKAGES



Heterogeneous systems leverage the data parallelism and power efficiency of today's GPUs to usher in a new era of innovation in computing. They can take mainstream visual applications and elevate them to new levels, enable new pixel-intensive experiences and even introduce non-visual capabilities that were unimaginable – until now. So let's take a moment to look at the many ways GPUs continue to propel advances in mainstream computing.

Redefining everyday experiences

What's your idea of long battery life in a notebook computer or mobile device? Depending on what you're doing – watching a feature-length movie vs. just checking email, for example – you might say maybe three hours? Five? How about all-day battery life? *Real* all-day battery life.¹ The days of lugging along a second battery or scampering

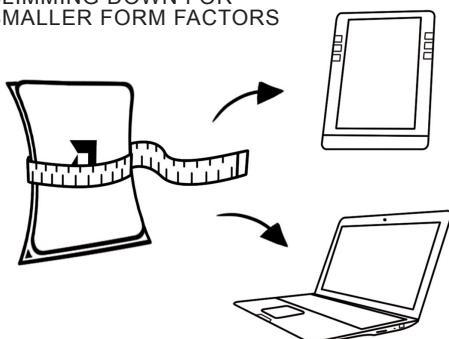
around an airport looking for AC power are ending, as heterogeneous computing sets a new standard for battery life in smaller packages, like today's ultrathin notebooks, netbooks, and tablets. The lower-power, hardware-accelerated video decoding and transcoding capabilities of today's GPUs make it possible for notebooks and mobile devices to deliver extraordinary visual performance while consuming less power.

High-definition (HD) video is another area in which today's GPUs may improve existing experiences. You've no doubt experienced both HD and standard-definition video – the former on HD equipment, and the latter on a non-HD system. But what if you didn't need content originally captured in HD to have an HD-quality experience? Standard-definition sources may be viewed in near HD-quality through the hardware-accelerated upscaling of lower-quality content – dynamically. So you'll be able to access streamed movies on more devices, or watch HD TV on a PC. Similarly, you won't need the highest-end PC powerhouse for 3D gaming.

Today's GPUs can enable ultra-high frame rates for realistic, immersive 3D gaming on mainstream (even entry-level) notebook computers. They can also make it easy and cost-effective to add stereoscopic 3D realism to 2D content, from Hollywood movies to home video.

At work, today's GPUs can make both collaborative and independent work more productive. With sup-

FIGURE 5
SLIMMING DOWN FOR
SMALLER FORM FACTORS



port for flawless HD videoconference capabilities and new forms of virtual presence, GPUs make virtual meeting as close to meeting in person as it gets. They enable visual communications that were not even remotely possible before, like bi-directional HD video chats. They can also send desktop productivity soaring by making it easier and more seamless than before to move back and forth between applications on multiple monitors – even when working with graphics-intensive content like PowerPoint presentations, product demos or simulations, and videos.

Enabling New Pixel-Intensive Experiences

The only thing more exciting than doing something better than you've ever done it is doing something entirely new. For example, what if you could log in to your computer just by looking at the screen? What if it could respond to just a gesture, not even need a touch, to know exactly what you want? And what if your HD video walkthrough of a new house could generate a 3D model of that space to see if your furniture could make that house your new home? Those are just a few of ways that heterogeneous computing could change the way people interact with computers forever. They're also just three examples of the types of software innovations that are possible now with GPU application programming interfaces (APIs).

If you're like most people, you probably find it hard to imagine something more mundane and time consuming than searching and categorizing your photos and videos based upon who is in them, where you were and what you were doing. But what if your computer could speed you through by quickly and automatically finding and tagging photos and videos for you based on the faces, places or objects that are in the images? Or by helping you sort through photo libraries to eliminate duplicates saved under different names? Or by finding the IMDb or Wikipedia entry you're looking

for based on a "look" at an actor's face? These are also pixel-intensive experiences made possible by today's GPUs. And they're enough to make current search and index capabilities seem almost unimaginably tedious and slow.

Today's GPUs can also add measurably to the enjoyment of visual content in a number of new ways. Imagine a computer that enhances your shaky hand-held video footage by automatically stabilizing HD content. Or one that delivers crystal-clear content even on room-sized screens. Or that supports holographic imagery for unprecedented realism. We foresee display resolutions so high, they deliver density and clarity that rivals what the human eye can even perceive in the analog world. This is how technology improves our lives by getting out of the way.

Creating New Non-Visual Experiences

Chances are that when you think of new applications for GPUs, you're likely to think of visual experiences. Facial and gesture recognition, photo or video indexing, and some of the others we've just described are good examples. That stands to reason, given that the "G" in GPU does stand for "graphics". But as we have discussed, GPUs can do more than just push pixels to a display. They are very capable general-purpose computing resources that can now stand side by side with the CPU and enable a broad range of compelling new experiences. For example, today's GPUs enable a level of contextually-aware computing that takes new experiences far outside the realm of the visual. We may not be at a point where "Computer: Earl Grey, hot" is going to be a reality anytime soon. But some of the very real possibilities are nearly as amazing.

FIGURE 6
MOVING BEYOND JUST GRAPHICS TO
CONTEXTUALLY AWARE COMPUTING



In ambient computing – one term for contextually-aware computing – massive amounts of data from multiple sensors enable a computer to adapt to user and situation, rather than the other way around. In the audio world, for example, consider the ever-present lag in the time it takes to transcribe or translate live audio from one language to another during video conferencing, news broadcasting and similar scenarios. People have accepted this inconvenience for years, but ambient computing enabled by advanced data parallel processing capabilities of GPUs creates the potential for audio interpretation and voice translation to other languages immediately, in real time.

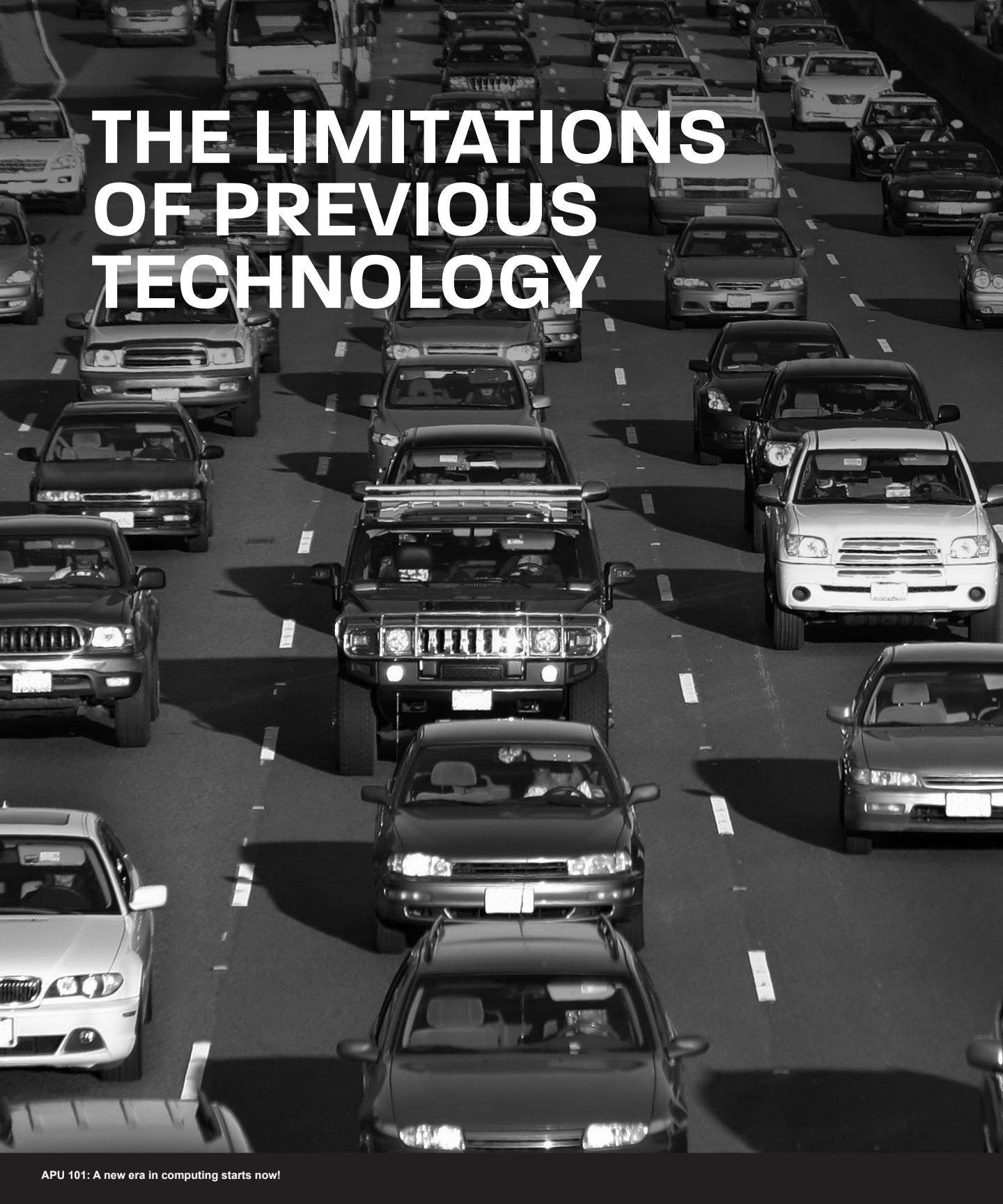
This more contextually-aware computing also opens the possibility for capabilities like speech recognition in audio recording applications – so that, for example, software will automatically determine who is speaking in a conference room and identify the different speakers in the auto-generated transcript of the conference. Another potential application is for software that will dynamically gauge the

acoustics in a room and tweak output accordingly for whatever mobile device you're using.

These kinds of audio applications translate into digital security as well – in, for example, systems that grant user access to physical locations and computer systems by recognizing and authenticating users based on voice. And while we're on the subject of security, GPU capabilities could enable targeted malicious software scans that eliminate the painfully slow runtimes associated with today's anti-virus programs.

These aren't blue-sky speculations; many of the capabilities exist today in labs all over the world. Why haven't they found their way to the mainstream yet? The answer is in the limitations of existing hardware architecture and software programming models. We'll examine these limitations in the next section of this paper, and then show how APUs – specifically APUs designed by AMD – are overcoming these limitations.

1. In testing conducted by AMD performance labs the 2011 AMD C-50 Dual-Core Accelerated Processor demonstrated up to 735 minutes or 12.15 hours "all-day" battery life while idle and up to 378 minutes or 6.18 hours as an "active" metric using 3DMark '06. All testing performed using a 6-cell Li-Ion, 62.2 Whr battery AMD defines "all day" battery life as 8+ hours of idle time. The 2011 HD Internet-based netbook consisted of an AMD Dual-Core processor C-50 1.0Ghz 9W, 2GB (1x2GB) DDR3-1066 system memory, AMD Radeon™ HD 6250 Graphics with 10.1" @ 1024x600, 6-cell Li-Ion, 62.2 Whr battery. LED Backlight Windows 7 Home Premium 64-bit. BRNeB-I3



THE LIMITATIONS OF PREVIOUS TECHNOLOGY

There is enormous potential for heterogeneous computing applications to bring new and better experiences to people on a truly revolutionary scale. The problem is with how processor designers and manufacturers have attempted to execute heterogeneous computing so far. What's required from their efforts is hardware that runs fast enough and well enough to support advanced applications in high volumes. But what are often missing are the appropriate hardware architecture and software interfaces to drive successful development.

Hardware Architecture Limitations

In our initial discussion of heterogeneous computing on page 7, we talked about the specific need to bring together the best of both CPU and GPU capabilities to drive new and better experiences. However, attempts to integrate GPU and CPU capabilities in the CPU chipset (chips on the PC motherboard that surround and support the functions of the CPU) called integrated graphics processor (IGP) have proved problematic. In most cases, the graphics unit was undersized or underpowered, creating an imbalance that made it unsuitable for general-purpose computing. This was the best option available for companies with well-recognized CPU experience, but little leadership or know-how in graphics. AMD, which has leading CPU and GPU portfolios, has done the opposite: Rather than attempt to make a lower-level GPU suitable for general purposes, we have used high-end graphics to create a higher-performing

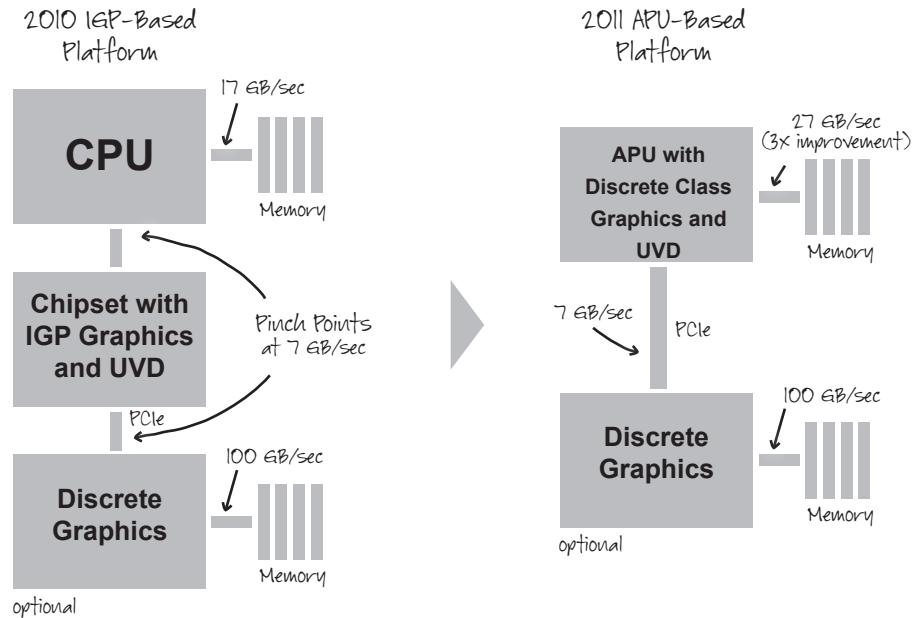
graphics processor that is easily adaptable to general-purpose computing.

AMD's use of higher-performing GPUs has helped overcome problems that IGPs typically have with inefficient, high-latency access to main memory across the PCI Express bus and through the CPU. Despite this inherent limitation, AMD's higher-performance solutions have historically delivered radically better visual experiences than competitive solutions. AMD's approach is the foundation for its AMD Fusion APU architecture, which is almost four times faster than the highest-speed PCIe 2.0 bus.

At the other end of the spectrum from companies that have relied on underpowered IGPs are those that have recognized the value of high-end graphics, but that, based on their core capabilities, have had to rely on a plug-in GPU rather than an integrated CPU-GPU approach. Plug-ins work well for display graphics for 3D gaming, but can create issues for general-purpose computing, due to the inefficiencies involved with moving data between CPU and GPU using a PCIe.

AMD, with the advantage of its long history of CPU and GPU leadership, has eliminated the need to use PCIe as the main interface between CPU and GPU in favor of a solution in which GPU and CPU capabilities are combined in one unit (Figure 7). AMD's Fusion Family of APUs includes both GPU and CPU as full-fledged, powerful, balanced processing elements.

FIGURE 7
BANDWIDTH CONSTRAINTS IN IGP CHIPSETS



General-Purpose Software Programming Limitations

Coding for GPUs has always posed a special challenge for software developers due to significant differences in hardware architectures across GPU vendors – and sometimes even across generations of GPUs from the same vendor. This can make coding directly to the hardware a one-off exercise for each specific GPU model. It is understandably impossible to sustain such an approach at the high volumes necessary for mainstream applications.

Previously, there were no standard Application Programming Interfaces (APIs) to smooth over GPU hardware differences and thus enable programmers to write just one set of code to run across multiple vendors, GPUs, and generations of GPUs. Now, however, two important development

tools have emerged to enable programming across GPUs – Open CL™ and Microsoft DirectCompute.

AMD has an OpenGL compiler that supports both its Radeon™ HD 4000, HD 5000, and HD 6000 Series GPUs and its multi-core x86 processor offerings. And software developers who use AMD platforms that include these GPUs can also use them to work on applications using new AMD APUs.

AMD also has a long history of supporting Microsoft DirectX APIs – and DirectCompute is a DirectX API introduced with the Microsoft Windows 7 operating system. Developers who are creating applications for Windows-based computers will be able to use this API with AMD APUs to take advantage of all the features of DirectX 11, the latest version of DirectX.

A second software programming limitation is that operating systems view only the CPU – and not the GPU – as a general-purpose computing device. This is understandable, because using GPUs for general purposes is a relatively new concept. But it creates problems because there are no provisions for load balancing and scheduling across a mix of scalar processor cores (CPUs) and vector processor cores (GPUs). AMD is addressing this limitation in the ongoing evolution of its APU hardware architecture and by driving increasing operating system integration between CPU and GPU.

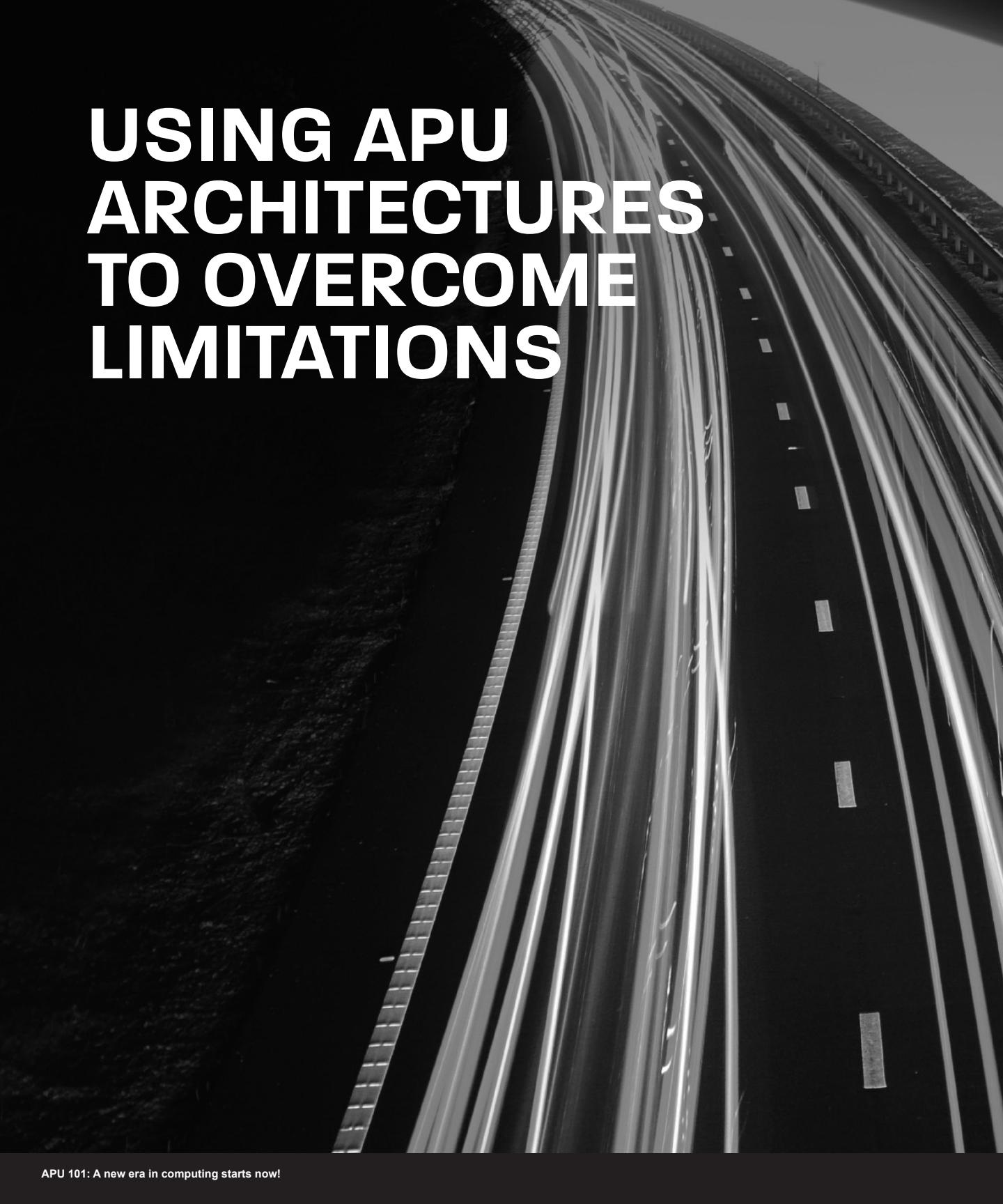
Economic Limitations

There are significant economic costs associated with the hardware and software limitations described here. The lack of standard APIs has, in part, made the development of mainstream software applications prohibitively expensive.

In addition, discrete graphics cards that deliver the high levels of performance necessary to support new and better user experiences can add substantially to the price of a PC. For this reason, they have achieved attach rates of roughly only 25% globally. And while those rates may be higher in Europe and China, the overall volumes of high-performance GPUs have not always been sufficient to justify further investment by independent software vendors (ISVs).

Finally, the economic gains made possible by the multi-core CPU era are tapering off as power and scalability continue to limit performance gains. GPUs and APUs represent a more attractive area for parallel programming investment today. Their scalability enables significant improvements in performance and thus provides the potential for greater returns.

USING APU ARCHITECTURES TO OVERCOME LIMITATIONS



APUs are emerging as the next logical step in the cycle of innovation, as the cycle moves squarely into the era of heterogeneous computing. Combining the advantages of CPUs and GPUs in one unit, APUs enable the kinds of unprecedented experiences described on pages 11-13 of this book. They offer up many of the capabilities needed to realize these experiences; at the same time, APUs overcome the architectural and economic limitations of previous technologies.

In the following pages, we will explore the convergence of trends that has made this exactly the right time for APUs. First, though, let's take a look at the basics of these processing units, how they work, and how they are both similar to and entirely different from CPUs and GPUs.

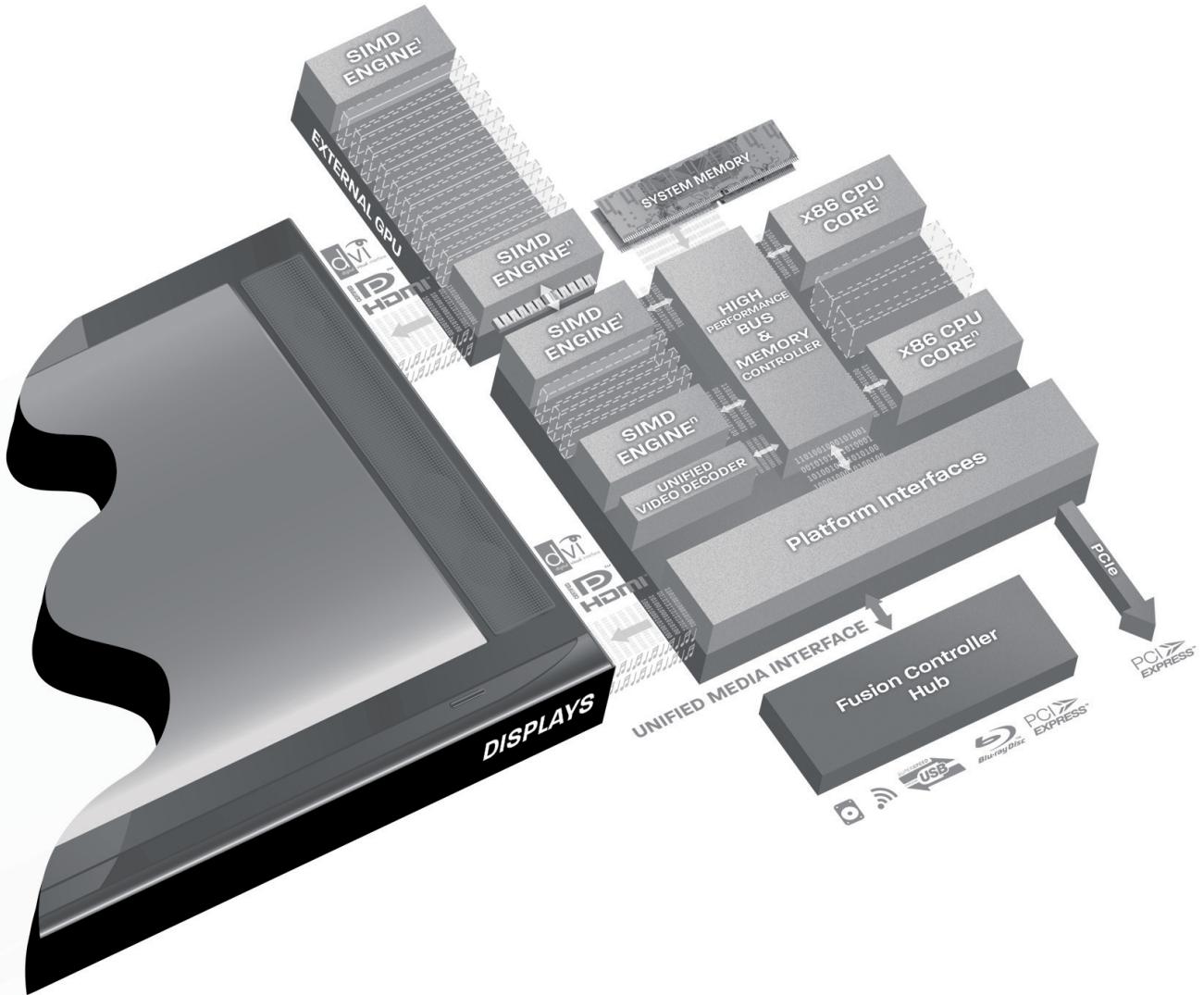
Just the Facts: What is an APU?

At the most basic level, APUs combine general-purpose x86 CPU cores with programmable vector processing engines – thus bringing together sophisticated CPU scalar processing with the large-scale, parallel-based vector processing that is traditionally associated with GPUs. AMD designed its APUs to combine CPU and GPU strengths, thus providing software developers with unprecedented flexibility to utilize whichever approach is best suited to the task when they develop new applications.

In AMD APUs, a programmable x86 CPU and the vector processing architecture of a GPU are joined together on a single piece of silicon by a high-performance bus. Both have local access to high-speed memory. AMD APUs also include a variety of other system elements such as memory controllers, I/O controllers, specialized video decoders, display outputs, and bus interfaces.

The appeal of AMD APUs is that they include both scalar and vector hardware as full-fledged processing elements. Other processor designers have attempted to lash together a CPU and a very basic graphics unit in a single package, but the result has been an imbalance in which the GPU lacks the performance and power to fully take on today's graphics processing challenges. And no other processor company has attempted such a marriage of CPU and GPU using truly programmable GPUs, let alone GPUs that can be programmed using high-level industry-standard tools like OpenCL, DirectCompute and DirectX 11. AMD, with its extensive patents and engineering expertise in x86 processor technology and its industry-leading GPU technology, is best situated to deliver on the full potential of combined CPU and GPU strengths.

FIGURE 8
THE APU: BALANCED COMPUTING FOR A NEW ERA



How AMD APU Architectures Address Key CPU and GPU Limitations

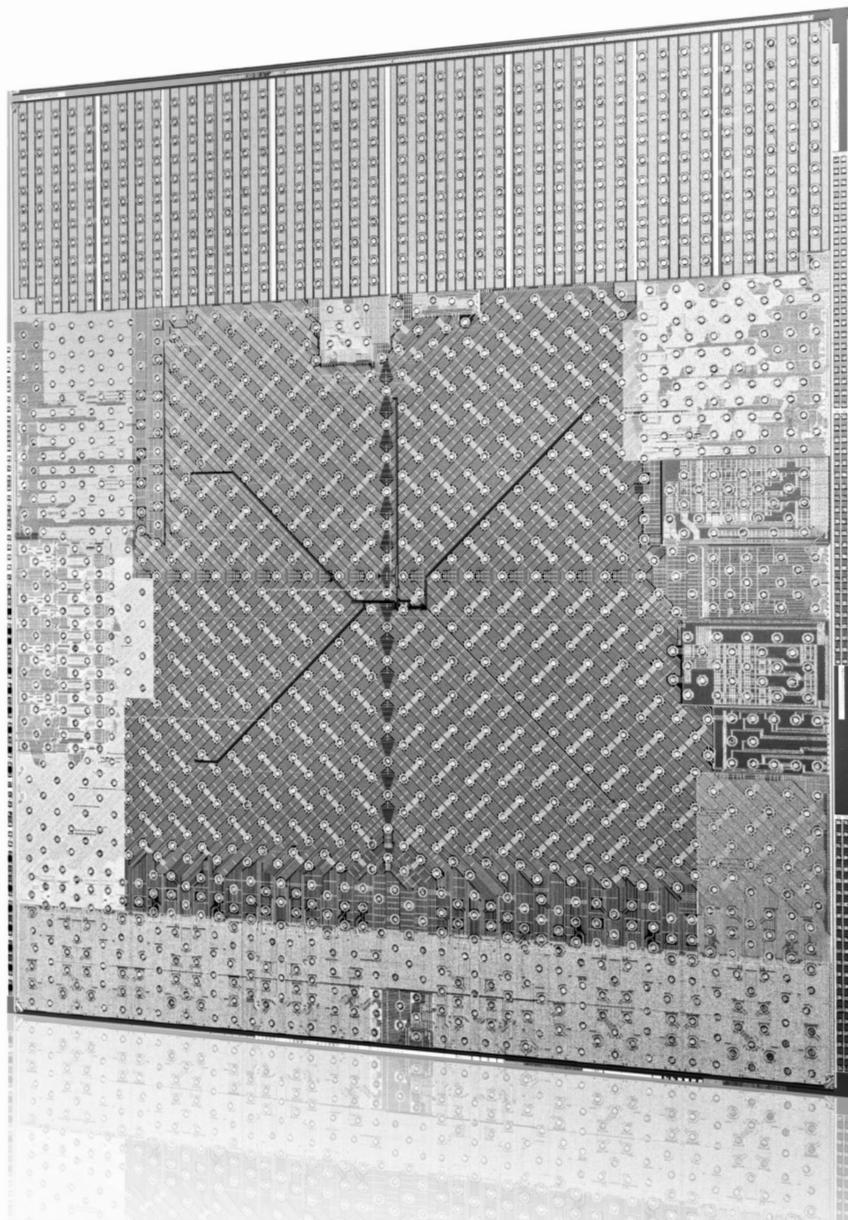
In the first APUs that AMD has introduced, all the major system elements – x86 cores, vector processing engines, and Unified Video Decoder (UVD) for HD decoding – attach directly to the bus (and, therefore, the memory). This eliminates one of the fundamental issues with IGPs historically. IGPs have typically been imbalanced two-chip solutions pairing a CPU with a less capable GPU, due to transistor budget constraints. With two chips, designers had to use a chip-to-chip crossing between memory controller and CPU or GPU.

These transfers tended to shorten battery life due to their effect on memory latency and to their power consumption. A common path to memory helps avoid this problem.

To further enhance system performance, the APU architecture allows for a multi-GPU configuration in which a discrete GPU can be added to the design, adding even more graphics processing power (Figure 8). The system can then scale as needed depending on application demands, harnessing all available resources and enabling unprecedented performance.

Opportunities for Software Developers

Just as AMD has woven well-balanced x86 cores and GPU cores into a single hardware component, software developers can now begin to weave high-performance vector algorithms into new applications without being constrained by the traditional computational limitations of scalar processors. Software from ISVs who take full advantage of the enhanced capabilities offered by APUs will be able to execute well beyond the capability of packages that lack support for combined CPU and GPU features.



Trends from two directions are converging today to drive the development of APUs and of software that takes full advantage of APU capabilities. Within the industry there have been important technical advances, including improved API availability and easier integration of GPUs and specialized processors into single, cohesive designs. From the market, end users' increased interest in visual content sources, demand for smaller form factors with long battery life, and desire for more intuitive user interfaces have all made clear the need for processors with far greater levels of graphics support than ever before.

APIs: Lowering the Investment Threshold for Software Innovation

APIs standardize the parameters that software developers use to render images, which basically just means that developers don't need to write code over and over again to accommodate differences in GPU hardware.

The recent introduction of industry-standard OpenCL and DirectCompute APIs has kicked open the door for developers to use GPU hardware to accelerate applications. With these APIs, they can create highly dynamic data-parallel software applications much more easily and cost-effectively – because they don't have to constantly reinvent the wheel to meet different GPU requirements.

OpenCL is a cross-platform standard for parallel computing coordinated by the Khronos Group, while DirectCompute is a new set of Microsoft DirectX APIs that facilitates GPU computing applications. DirectCompute is ideal for developers who already typically focus on Windows and use DirectX APIs for graphics, while OpenCL is an ideal platform for developers who already use OpenGL APIs to handle graphics.

- DirectCompute features data structures that are compatible with DirectX APIs, helping simplify the process of adding GPU acceleration for physics or artificial intelligence tasks to DirectX 11 applications.
- OpenCL supports both data-parallel and task-parallel execution models and uses data structures that are compatible with OpenGL APIs, thus simplifying the task of adding GPU compute acceleration to OpenGL applications.

Manufacturing Advances: Enabling Cost-Effective Processor Integration

Processor designers who want and need to integrate GPUs and other specialized processing units into their designs were until recently constrained by the fact that semiconductor manufacturing processes were simply not geared toward this type of integration.

Now, however, semiconductor manufacturing advances are making it much more cost-effective to integrate high-performance GPUs – as well as Universal Video Decoders (UVDs) and other fixed-function specialized processors – into processors.

This development allows the insertion of massive numbers of GPU cores into processors, making it possible to design and produce processors targeting entry-level and mainstream price points.

On the Internet and Beyond: Increasingly Visual Digital Content Sources

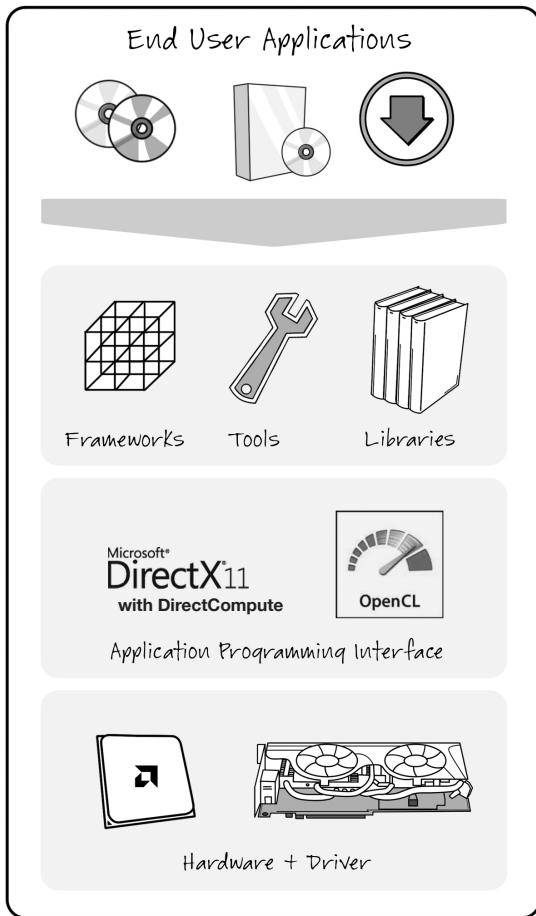
Digital content is growing increasingly visual, whether on the Internet or from other sources. According to industry reports, YouTube now gets more than 2 billion video views per day.¹ People are adding 50 million digital media files to their personal content libraries every day.² And they're uploading more than 1,000 pictures to Facebook every second.³

Increasingly, much of the visual content people receive is in HD; in fact, nine billion video files in consumers' personal libraries today are HD. Not surprisingly, people are expected to buy a half-billion HDMI-enabled consumer devices in 2013.²

All of this presents an amazing opportunity for APUs to improve users' experiences. AMD is working with ISVs now to make Internet browsing better by enabling a 2D-accelerated option to speed the viewing experience and page loads. We're also investigating how to use UVDs in APUs to offload and refine video content and then display the far-higher-quality result on a large screen. Those are just two examples of how the trend toward more and more visual experiences is driving the development of APUs and software that takes advantage of them.

It's also worth noting that it's not just the display quality aspect of visual experiences where APUs can have an impact. All that growth in visual content is also making functions like search and index, as well as capabilities such as photo and video cleanup, more important to users than ever before. And these, too, are areas where APUs can dramatically speed up and simplify processes.

FIGURE 9
AMD LEADING COMPLETE ECOSYSTEM ENABLEMENT



Form Factors: Getting Smaller All the Time

The continuing demand for smaller PC form factors is driving a need for greater integration and miniaturization, which ties directly into the highly space- and power-efficient nature of APUs. When space is at a premium, hardware designers need to be able to reduce the number of components necessary to deliver a great user experience.

At the same time, though, today's designers need to be able to meet higher performance demands than ever. Users will always want the same increasingly high-end, highly visual applications available on desktops in their notebooks and netbooks – especially since, for many users, their only computer is a notebook or other portable device. With their integration of CPU, GPU, and specialized units on one processor, APUs can dramatically reduce the amount of physical space needed inside small-form-factor systems.

APUs can explode the possibilities for these smaller systems. You may see AMD APU-powered ultrathin notebooks, netbooks, tablets, and slates – in both expected and unique designs. For example, building on AMD APU support for multiple displays, AMD could enable a notebook with multiple pull-out displays for a panoramic work surface. The processing capability to support them would be there, leaving only the design challenge of how they would physically be incorporated into the system. Or think even smaller, and consider the possibility of a slate design that does more than ever – with built-in projection capabilities, wireless HDMI, and the latest Wifi connectivity.

User Interfaces (UIs): The Growing Fascination with the Possibilities

People have always been fascinated by the idea of interfaces that require little more than a look, word, or gesture for a computer to recognize a user and what he or she is trying to do. Once the stuff of science fiction and fantasy, sophisticated user interfaces no longer belong solely to the realm of the future.

Recently, media outlets ranging from Computerworld (Nov. 27, 2010) to Variety (Oct. 27, 2010) have reported on everything from desktops being taken over by new gesture-driven interfaces to TV remotes being made obsolete by voice and gesture technologies.

APUs are poised to capitalize on this trend by enabling contextually aware computing, in which massive amounts of data from multiple sensors enable a computer to adapt to the user and situation in ways that were previously impossible.

Toward a More Balanced Computing Architecture

Today, the trend is toward declining performance returns on investments in applying additional transistors to CPU cores and cache. Certainly, in relation to the performance gained and user experience achieved, processor designers would do poorly to continue to invest exclusively in multi-core solutions. This presents the ideal opportunity to seek greater performance returns by investing in larger and more capable GPUs and fixed-function blocks. APUs strike a balance between scalar and vector processing that balances the investment across both for processor and software designers alike.

1. YouTube online fact sheet, retrieved December 22, 2010 http://www.youtube.com/t/fact_sheet
2. CEA Market Research Report – Amassing Digital Fortunes: A Digital Storage Study, May 2008
3. Facebook online fact sheet <http://www.facebook.com/press/info.php?statistics> retrieved 7 December 2009



A black and white photograph showing a close-up of a road roller's drum as it moves over a dark, textured surface, likely asphalt. The drum is metallic and shows signs of wear. A single bolt and nut are visible on the side of the drum. The background is blurred, suggesting motion.

AMD PAVING THE ROAD TO BROAD MARKET ADOPTION

AMD is eager to see software and hardware developers begin adopting AMD Fusion APUs on a large scale. The sooner they start using APUs to develop new applications and form factors, the sooner today's users will be able to begin enjoying the next generation of computing experiences. To that end, AMD is taking bold steps to make it as economically attractive as possible for developers to adopt APUs.

APUs Everywhere

AMD has designed the first generation of APUs to support high-volume market segments where discrete-level programmable GPUs were the exception, not the norm. This sets the stage for a much wider spectrum and volume of platforms relevant to coding for GPUs as general-purpose computing devices. This approach should drive a rapidly expanding installed base of systems with new high-performance graphics – a base that builds on AMD's existing installed base of discrete graphics cards as the worldwide market share leader in GPUs.

The solid foundation established by this approach should provide attractive revenue opportunities for ISVs to innovate in data-parallel applications and for hardware developers to explore innovative form factors.

In hardware design and development, the presence of powerful new graphics capabilities means that APUs can deliver the high levels of media performance needed to support media-rich capabilities even in the smallest form factors.

Funding Innovation through the AMD Fusion Fund

AMD recently announced a new funding program to accelerate development of software applications, hardware systems, and other solutions based on AMD APUs. Through the fund, AMD invests in companies eager to create new solutions that take advantage of APU features and get them to market quickly. For example:

- Software applications that are newly enabled or greatly enhanced by the synergy between CPU and GPU in APUs
- Software tools that enable developers to easily develop, scale, optimize, or test new applications
- Technology or intellectual properties that can serve as foundational building blocks for applications
- Form factors or devices that expand the scope for deployment of new applications

The AMD Fusion Fund is AMD's way of encouraging companies that embrace the innovative potential of APUs. Without this kind of incentive, it could be difficult for many companies with ideas about new computing experiences that could be enabled by APUs to move forward and begin to realize those ideas.

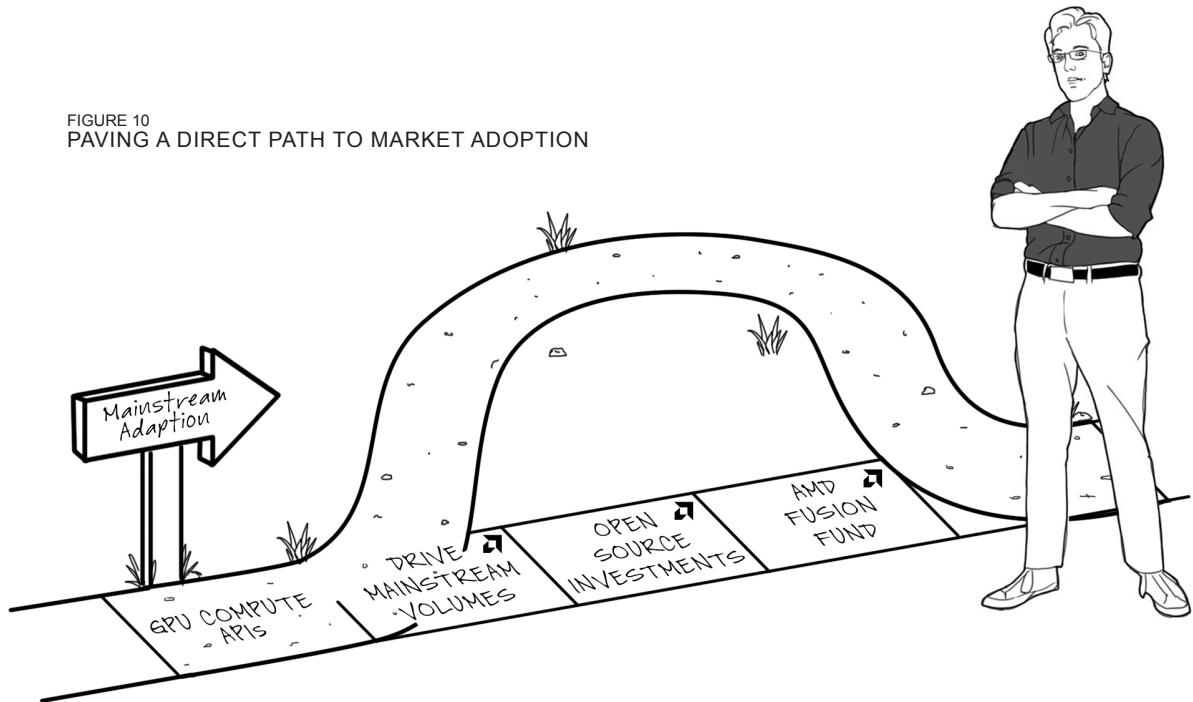
There's no limit to who can apply for assistance from the AMD Fusion Fund; start-ups, entrepreneurs, software tools companies, software design firms, original design manufacturers (ODMs), and component companies are all invited to apply for consideration. By funding their efforts, AMD looks forward to getting more APU-based solutions to market more quickly.

Investing in the heterogeneous computing software ecosystem

AMD is investing in the development of software libraries for open source communities, enabling even more opportunities to develop applications with APUs. The overarching goal of this investment in the heterogeneous-computing software ecosystem is to enable more widespread, productive software development for new, APU-enabled end-user applications.

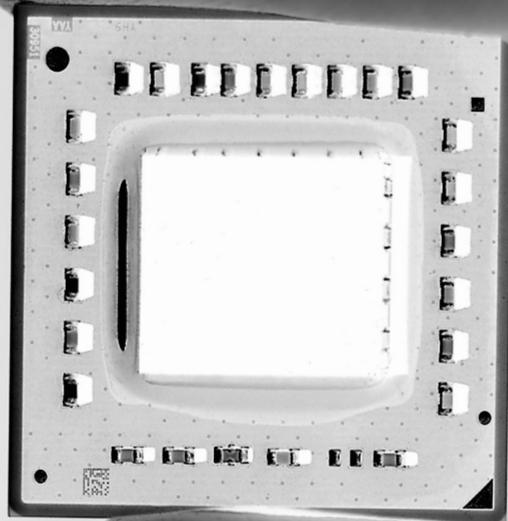
Based on open industry standards, this ecosystem facilitates more focused and parallel development through well-defined abstractions, and encourages aligned contributions from established as well as new companies, as well as from academic institutions.

FIGURE 10
PAVING A DIRECT PATH TO MARKET ADOPTION



AMD FUSION FAMILY OF APUs:

THE FIRST GENERATION



In keeping with AMD's strategy of using volume to seed the market for APUs (as described in the preceding section of this book), the first generation of processors in the AMD Fusion Family of APUs is designed to facilitate development of new applications and is built at a cost structure designed to encourage rapid and widespread adoption. AMD began mass production of the AMD C-Series and E-Series Fusion APUs in the fourth quarter of 2010 (shipping to customers in that same quarter), and will launch the AMD Fusion A-Series performance-class APUs during 2011.

A New Era in HD Visual Performance

AMD Fusion C-Series and E-Series APUs are both part of AMD's platform for unprecedented HD visual performance in desktops and notebooks. Optimized for both cost and power, this platform is focused on providing precisely the right balance of CPU and GPU performance to enable sharp, clear HD online experiences, along with cool, quiet computing, in a sleek, affordable desktop or notebook.

Desktops built around E-Series APUs enable users to enjoy video, photos and websites in amazing HD quality, with the clarity of 1 billion+ colors and full 1080p resolution for photos and videos – as well as faster, easier photo and video searching, viewing, editing, and transfer. Smaller form factors built on C-Series and E-Series APUs bring these same benefits to mainstream and ultrathin notebooks and HD netbooks.

AMD Fusion C-Series and E-Series APUs for Desktops and Notebooks

AMD's platform for HD visual performance in desktops and notebooks consists of an AMD C-Series or AMD E-Series APU that is connected by a high-performance bus to a DDR3 memory channel and multiple analog and digital display interfaces.

- AMD C-Series APUs power both desktops and notebooks. They consist of one or two low-power x86 CPU cores running at up to 1.2GHz clock speed and supporting out-of-order execution, along with a third-generation UVD and an Accelerated Parallel Processing (APP)-enabled array of SIMD engines. AMD C-Series APUs allow for 80 parallel processing units, or shaders, running at 280MHz and supporting the full DirectX 11 API and OpenGL for 3D graphics, as well as DirectCompute and OpenCL for general-purpose programming of the vector processing engine. These APUs have an amazingly low power consumption profile of just 9W maximum Thermal Design Power or TDP.

In the notebook arena specifically, AMD C-Series APUs for HD netbooks reflect a complete redefinition of personal computing in a small form factor, delivering 10 times the performance of previous-generation system graphics. They enable full 1080p HD video quality, Adobe Flash acceleration and DirectX 11 graphics, raising the bar for an advanced visual experience in both netbooks and

high-end tablets. Netbooks built on AMD C-Series APUs will be branded as AMD HD Internet systems.

- AMD E-Series APUs are designed for notebooks, desktops, All-in-Ones, and small-form-factor systems, bringing all-day battery¹ life to high-performance systems. They utilize two advanced x86 CPU cores, running at up to 1.6GHz clock speed and supporting out-of-order execution. AMD E-Series APUs also include 80 parallel processing units running at 500MHz and supporting the full DirectX 11 API and OpenGL for graphics, as well as DirectCompute and OpenCL for general-purpose programming of the vector processing engine.

AMD E-Series APUs represent a significant step up in notebook and All-in-One experiences and capabilities relative to the C-Series-based AMD HD Internet systems. In mainstream opening price point (OPP) notebooks, these APUs deliver twice the graphics performance of 2010 Intel CULV platforms within the same power envelope. In ultrathin notebooks, they bring an amazing online HD experience to a thinner design. Ultrathin notebooks powered by AMD E-Series APUs also deliver twice the performance of a 2010 Pentium T4400-based system.² Whether in mainstream or ultrathin notebooks, E-Series APUs offer full 1080p HD video quality, Adobe Flash acceleration and DirectX 11 graphics – add all-day battery life.¹

Key Features

New Low-Power X86 Cores: Small, Efficient and Strong

AMD C-Series and E-Series APUs both use innovative new x86 CPU cores (code-named “Bobcat”). These are very low-power, high-performance CPUs that can deliver today’s mainstream performance in a very small area – less than half the silicon area of

what was previously possible – while requiring very little power. In fact, by die area, the AMD E- and C-Series APUs are more GPU than CPU!

Synthesizable and easy to reuse, these cores are fully ISA-compliant, with SSE3 multimedia instructions and AMD-V virtualization support. One particularly interesting feature is the Neural Net Logic Branch Predictor, which is designed to improve the ability of Series C and Series E APUs to deliver performance for today’s workloads.

DirectX 11: A New Generation of Visualization Plus Powerful Programming Capabilities

For software developers who primarily use Microsoft DirectX APIs, AMD Fusion C-Series and E-Series APUs have DirectX 11 APIs, which succeed DirectX 10 APIs as the most advanced form of the DirectX APIs available. Moving up to DirectX 11 is critical because DirectX 11 offers a number of new features that collectively improve the graphic fidelity of games and 3D applications while at the same time enhancing the ability of the CPU and GPU to work together efficiently. For end users, the move to DirectX 11 brings the potential to realize significant increase in performance as application developers harness the power of new DirectX 11 features such as tessellation, multi-threading and parallel-processing technology.

Tessellation is a technique that special effects producers have used for years to add detail to 3D images; with DirectX 11, it’s now available to PC developers for real-time rendering. It’s truly critical for more realistic environments and characters than ever before. Multi-threading support in DirectX 11 enables rendering tasks assigned to the CPU to be efficiently spread across multiple cores, enabling a better balance of performance between the CPU and the GPU. Compute shaders are programs that

FIGURE 11
THE FIRST-GENERATION AMD FUSION APUS

AMD Fusion E-Series	Model #	CPU Cores + Speed	AMD Radeon GPU Cores + Speed
	E-350	 1.6 GHz	 80 500 MHz
AMD Fusion C-Series	Model #	CPU Cores + Speed	AMD Radeon GPU Cores + Speed
	C-50	 1.0 GHz	 80 280 MHz

are executed on the graphics processor to not only enhance image quality but also accelerate non-graphics applications such as video transcoding and artificial intelligence.

Support for the Latest DDR3 Memory

AMD Fusion C-Series and E-Series APUs support the latest DDR3 memory, which delivers three times the bandwidth between GPU and memory and 59% more bandwidth between CPU and memory as compared to 2010 AMD platforms. DDR3 memory also reduces power consumption by 30% compared to DDR2 memory.

Coming Soon: AMD A-Series High-performance APUs

AMD's high-performance APU is code-named Llano and will be the AMD A-Series. Slated for volume production in the first half of 2011, the AMD A-Series APUs mark a significant technology shift for AMD

as the company transitions its entire mainstream client computing processor offerings to APUs. AMD A-Series APUs offer higher performance, more CPU cores, and a significantly larger GPU than AMD C-Series and E-Series APUs. In addition, AMD A-Series APUs include AMD Turbo CORE technology, which varies the performance of the CPU and GPU to respond to the demands of a particular application, and AMD Power Gating, which improves power efficiency by turning on and off CPU and GPU resources based upon application loads.

AMD A-Series APUs are slated to be the basis for AMD's new mainstream desktop and notebook platforms for 2011, which are designed to enable visually brilliant experiences and high performance for end users. Software applications developed for these platforms can use AMD A-Series APUs to improve video quality regardless of the source – delivering sharper HD than ever, and bringing HD clarity to online video. Other applications can make it easier than ever to shoot, edit and store video, and to quickly search and tag images.

The Future of AMD Fusion APUs

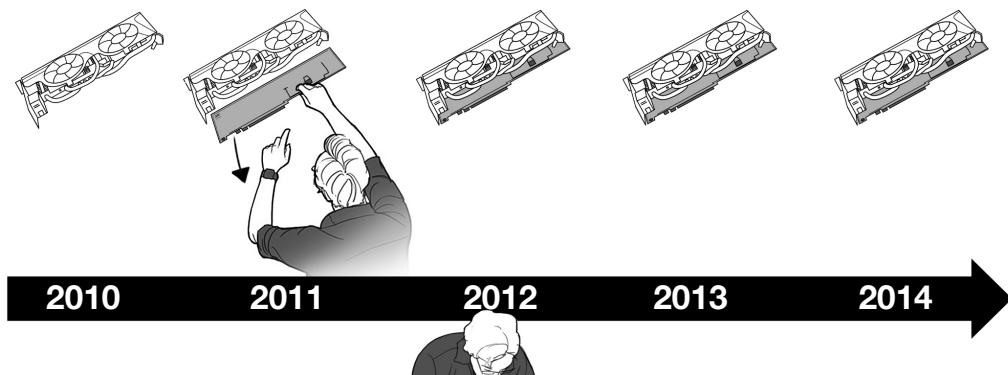
With AMD C-Series, E-Series, and A-Series APUs, AMD's HD Internet and VISION lineup can provide the best possible visual performance across a full range of solutions including:

- Multi-purpose tablets
- Netbooks
- Ultrathin notebooks
- Mainstream and performance notebooks
- Home theater PCs
- All-in-one desktops
- Small-form-factor desktops
- Mainstream desktops

Velocity: Driving the Pace of Innovation

"Velocity" is AMD's term for the way visual computing technologies are now driving the pace of innovation at AMD. AMD has been innovating GPU architectures every year, introducing significant new technology developments about twice as frequently as the company's competitors. The new discrete GPUs that result from this will provide new technology every year to incorporate into APUs (Figure 12). This means that AMD will be able to introduce new APUs with a general design cadence similar to that of the company's GPUs, with many of the core features enabled by new GPU technology.

FIGURE 12
VELOCITY AND BALANCED COMPUTING DELIVER EVERY YEAR

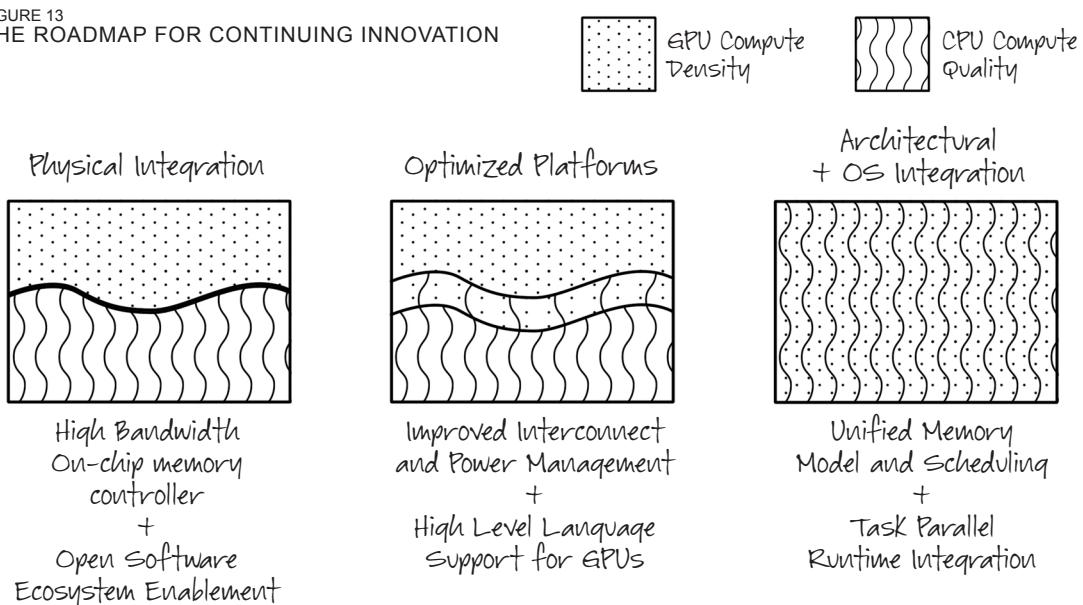


New leadership in discrete graphics every year

Feeds the following year APU with a new more advanced GPU

Visual technologies set the pace for AMD mainstream innovation

FIGURE 13
THE ROADMAP FOR CONTINUING INNOVATION



Discrete Graphics: Fueling Innovation from the Top

APUs will no doubt alter the graphics landscape, enhancing user experiences at all price points and enabling new experiences and innovations that we are now only beginning to imagine. However, these innovations will be fueled by the advances in GPU technology that are occurring in the discrete graphics space, where new and existing vertical markets and niche consumer markets will first benefit from AMD's continuously advancing GPU architectures.

Discrete graphics will remain AMD's first point for the introduction of new features, architectural innovations, improvements in memory subsystems and other advances. Discrete graphics will also provide synergy with APUs, increasing overall performance. As more applications take advantage of GPUs for both graphics rendering and application acceleration, more powerful GPUs will become must-haves for applications in every area from gaming to general productivity.

The Ongoing Evolution of AMD Fusion Architecture

Beginning with its ground-breaking integration of CPU and GPU, AMD Fusion APU architecture will evolve through the era of heterogeneous computing – from the physical integration of processors, to platform optimization, to the full integration of architecture and operating system (Figure 13). Hardware will evolve along the way, from on-chip memory to fully unified memory and scheduling, while software travels the path from enabling the open software ecosystem to achieving full task-parallel runtime integration. These ongoing innovations will enrich the heterogeneous-computing era with processing and performance improvements for years to come.

And it all starts now!

1. In testing conducted by AMD performance labs the 2011 AMD C-50 Dual-Core Accelerated Processor demonstrated up to 735 minutes or 12.15 hours "all-day" battery life while idle and up to 378 minutes or 6.18 hours as an "active" metric using 3DMark '06. All testing performed using a 6-cell Li-Ion, 62.2 Whr battery AMD defines "all day" battery life as 8+ hours of idle time. The 2011 HD Internet-based netbook consisted of an AMD Dual-Core processor C-50 1.0Ghz 9W, 2GB (1x2GB) DDR3-1066 system memory, AMD Radeon™ HD 6250 Graphics with 10.1" @ 1024x600, 6-cell Li-Ion, 62.2 Whr battery. LED Backlight Windows 7 Home Premium 64-bit. BRNeB-I3

2. Graphics Performance is 3DMark Vantage E performance. AMD E-Series 18W APU, 2x2GB DDR3-1066, 250GB HD, Windows 7 Ultimate (8.65 beta driver) compared to Intel Pentium T4400 Processor with G45 graphics, 2x2GB DDR3-1066, 160GB HD, Windows 7 Ultimate, 8.15.10.1995

This book contains forward-looking statements concerning, among other things, the timing of the planned introduction of AMD Fusion platforms and the features and performance of new product and technology releases which are made pursuant to the safe harbor provisions of the Private Securities Litigation Reform Act of 1995. Forward-looking statements are commonly identified by words such as "would," "may," "expects," "believes," "plans," "intends," "projects," and other terms with similar meaning. Investors are cautioned that the forward-looking statements in this book are based on current beliefs, assumptions and expectations, speak only as of the date of this book and involve risks and uncertainties that could cause actual results to differ materially from current expectations. The material factors that could cause actual results to differ materially from current expectations include, without limitation, the following: AMD will be unable to develop, launch and ramp new products and technologies in the volumes and mix required by the market; AMD's third party wafer foundries will be unable to manufacture its products on a timely basis with acceptable quality, at acceptable manufacturing yields and using competitive technologies; AMD will be unable to maintain the level of investment in research and development that is required to remain competitive; demand for computers and consumer electronics products and, in turn, demand for AMD's products will be lower than currently expected; customers will stop buying AMD's products or materially reduce their demand for its products; there will be unexpected variations in market growth and demand for AMD's products and technologies in light of the product mix that it may have available at any particular time or a decline in demand; or the effect of political or economic instability internationally on sales or production. Investors are urged to review in detail the risks and uncertainties in the company's Securities and Exchange Commission filings, including but not limited to the Quarterly Report on Form 10-Q for the quarter ended September 25, 2010.

AMD (NYSE: AMD) is a semiconductor design innovator leading the next era of vivid digital experiences with its ground-breaking AMD Fusion Accelerated Processing Units (APUs). AMD's graphics and computing technologies power a variety of devices including PCs, game consoles and the powerful computers that drive the Internet and businesses.

For more information, visit <http://www.amd.com>.





©2011 Advanced Micro Devices, Inc. All rights reserved. AMD, the AMD Arrow logo, Radeon, and combinations thereof are trademarks of Advanced Micro Devices, Inc. Other names are for informational purposes only and may be trademarks of their respective owners. 49390-A