SERVER BLADES GO **MAINSTREAM**

Blade blad NOUN: 1. The flat cutting part of a sharpened weapon or tool. 2a. A sword. b. A swordsman. 3. *Archaeology* A slender, sharp-edged flake that is at least twice as long as it is wide. 4. A dashing youth. 5a. A flat thin part or section, especially one that makes contact to perform a desired action: *the blade of an oar; the blade of a hockey stick.* b. An arm of a rotating mechanism: *the blade of a propeller; the blade of food processor.* c. A long, thin, often curved piece, as of metal or rubber, used for plowing, clearing, or wiping. 6. The metal runner of an ice skate. 7. A wide flat bone or bony part. 8. The flat upper surface of the tongue just behind the tip. 9. *Botany* a. The expanded part of a leaf or petal. b. The leaf of grasses or similar plants.

—The American Heritage® Dictionary: Fourth Edition. 2000

When the authors of the American Heritage Dictionary get around to drafting the next edition of their esteemed work, they will probably find it necessary to add a new element to their definition for "blade:"

10. *Technology* A modular circuit board assembly, containing one or more microprocessors and associated high-speed memory, that is easily inserted into a rack with many similar boards. May also contain limited input/output resources. Typically intended for dense computing applications within enterprise IT infrastructures.

Server blades are moving from the "early adopter" stage to a mainstream position in the IT infrastructure. According to IDC, this transition is progressing far more rapidly than the industry's earlier shift from tower configurations to rack-optimized 1U, 2U and 3U arrangements. After just three years of product shipments, blades account for seven percent of all servers sold; IDC estimates blades will make up thirty percent of all server shipments by 2008. In this white paper, AMD examines the issues that are driving the move to blades. We provide a framework IT managers can use to determine whether blades are right for them, and the factors they should consider as they evaluate alternative bladed server systems. Finally, we close with a brief overview of the blade offerings available from top-tier and tier-two suppliers of these systems.

What Problems Are Blades Trying To Solve?

When the first bladed server systems appeared in 2001, their progenitors emphasized their ability to pack many processors into an incredibly small space. RLX, one of the pioneers in this market segment, even dubbed their product "System 324" to highlight the 324 processors they could pack into one standard 42U[1] rack. At the time, non-bladed approaches typically maxed out at. 84 processors per rack, using 42 1U dual-processor servers.

[1] One U (1U) = 1.75 inches, a common measure of height for rack-mounted components

This increased compute density appealed to IT organizations facing space constraints in their data centers, as well as those renting space in co-location centers for mission-critical operations. Unfortunately, to attain this high processor density, users had to accept serious compromises regarding the computational power of the processors they employed. These early systems typically relied on low-power Intel® Pentium III or Transmeta Crusoe processors that delivered far less performance per processor than the standard CPUs then incorporated in rack-optimized configurations.

Since those early days, each of the mainstream server suppliers has come forth with its own approach to blade computing, specifically the HP BladeSystem, the IBM BladeCenter, and the Sun Fire Blade Platform. These newer designs incorporate faster processors like the AMD Opteron™ processor and Intel Xeon that consume more power and dissipate more heat than those early low-power blades, and therefore cannot match the density of that initial RLX system.

Of course, if your datacenter can deliver 30 Kilowatts to a single rack, and remove the heat generated by that rack, Angstrom Microsystems will be glad to sell you one of their SuperBlade Systems filled with 260 AMD Opteron processors. Most datacenters don't have the wiring or HVAC facilities to support such dense configurations, and can't pack every rack with the maximum number of enclosures, or every enclosure with the maximum number of blades. In short, although blade technology has improved CPU and system density beyond what can be achieved with traditional rack-based systems, the maximization of processor density has proven to be a less compelling benefit than many users anticipated.

If density hasn't been the be-all and end-all, then what is driving the increasing adoption of blade technology? It's obvious in retrospect, even if it wasn't obvious when those first systems arrived in 2001. Blades provide a quantum increase in system manageability and reliability over earlier approaches. Some of this improvement stems from the physical structure of blade systems, which forces a level of cable management and system interconnects that was unavailable in rack-mount arrangements.

Since blades make it so easy to take system elements on or off-line physically, users needed to ensure it would be easy from a software perspective as well. This led to the notion of "stateless blades," i.e., systems that use their disks (if they have disks) only for bootstrap and paging operations, rather than for stateful storage. Blades forced suppliers and users to rethink how they could simplify the management of the hundreds of servers that might be needed to handle complex tasks, and resulted in the creation of software that monitors, allocates and provisions blade systems.

The stateless nature of blade operations fits well with virtualization, another emerging trend that's sweeping the IT community. Software virtualization technology, as presently embodied in packages like VMware, Virtual Server and Zen, allows IT managers to move applications from one physical system to another, whether to balance demand or consolidate workloads. As this technology unfolds over the next few years, IT managers will have unprecedented opportunities to match workloads and resources in ways they can only dream about today.

Many organizations are experimenting with virtualization today, but often find the software overhead these packages impose on application performance precludes deployment in production environments. AMD is working with the developers of these virtualization packages to add features to its next-generation AMD Athlon™ 64 and AMD Opteron processors that facilitate the operation and improve the performance of virtual machine monitors. The first outputs of this collaboration, which AMD has code-named "Pacifica," will appear in products the company plans to deliver in 2006.

To Blade or Not to Blade, That is the Question

Blades may not be the best solution for all computing problems. Even IDC's aggressive forecast projects that by 2008, blades will account for only 30 percent of server shipments; this means 70 percent of the servers sold that year will still fit in the traditional tower or rack-optimized arrangements. How can you tell whether blades are the right solution for your IT deployments? Here's a short list of factors that AMD thinks you should consider before you settle on your approach.

Economics

Blades present a classic trade-off between up-front investments in new equipment versus pay-as-you-go acquisition strategies. Bladed approaches typically share some system resources – equipment racks, system enclosures, power distribution, cooling and cabling, to name a few – across multiple servers. Buyers must include these shared resources in their initial system acquisition costs. Once those racks and enclosures are put in place, the incremental cost for adding additional servers shrinks, since those add-ons get a free ride on the existing infrastructure. Since suppliers have taken varied approaches to the degree of resource sharing inherent in their blade designs, the ratio of infra-structure expenses to per-blade expenses also varies. In general, enclosures need to be 30 to 50 percent filled with blades before the arrangement makes economic sense. Any IT organization invest-igating bladed solutions should do the math, based on its own capacity requirements and the arrangement of its vendors' price list.

Scale Up or Scale Out

Some applications work best in "scale up" arrange-ments that use 4-way, 8-way, even 16- or 32-way SMP configurations. Others fit well with "scale out" approaches that utilize clusters of 2-way and 4-way nodes. If your application needs a large SMP system for optimum performance, then current blade technology may be a poor fit. Most of today's blade servers support dual-processor nodes, but a few suppliers offer quad-processor SMP arrangements as well. As the year progresses, look for those 2-way and 4-way blades to incorporate dual-core processors (we'll explain why in a few minutes) that will perform like current 4- and 8-way systems. If you need more than eight powerful 64-bit processor cores grinding away to handle your application, then you will still need to consider tower and/or rack-optimized systems at this point. But stay tuned; many innovative companies are developing the technology that will enable large SMP configurations to be assembled one blade at a time.

Predictability versus Adaptability

You've been running some applications in your shop on a regular basis for years; you know how often they must run and what kind of load they put on your processing resources. Other applications are new or evolving, and must deal with unpredictable workloads. The workload monitoring and provisioning software included with many blade systems makes it easier to build adaptability into your infrastructure, and gives bladed approaches a real advantage in handling tasks with widely varying and unpredictable workloads.

Performance or Power Consumption

We noted earlier that the designers of the first blade servers to hit the market in 2001 traded away the performance of individual blades in order to minimize blade power consumption and maximize blade density. Although customers quickly recognized many of the benefits of blade computing, most were unwilling to sacrifice performance for increased density. They voted with their dollars for systems with higher performance per node and lower blade density per chassis. Performance per node matters no less in blade environments than in rack-mount or tower configurations. Higher performance per node means fewer nodes to acquire, fewer nodes to manage, and (in the case of software license fees calculated on a "per processor" basis) lower software expenses.

AMD believes customers should never be forced to sacrifice performance to save power in servers. To this end, we have enhanced our latest AMD Opteron™ processors with the AMD PowerNow!™ technology we've offered for years in our mobile processors. AMD PowerNow! technology allows the server operating system to monitor system activity and slow the CPU down whenever the system is lightly loaded. This move reduces the CPU's power consumption by 75 percent – from approximately 93 watts when running at full speed to 32 watts when idling. That may not make much of a difference for a single processor, but that 61 watt savings per CPU adds up fast when you fill a rack with more than 100 processors. And if you want a measure of how clever our AMD Opteron processor design team is, just compare the power savings AMD gets with AMD PowerNow! technology to the power savings the "Brand X" processor supplier gets with the "Demand Based Switching" feature it recently added to its server chips. (Hint: the example they have shared publicly shows their low-power state uses about three times the power of ours.)

AMD also believes that customers can always find uses for faster processors, so we have populated our road-map with faster single-core processors, faster dual-core processors, and multi-core processors that will take performance even higher. We laid the groundwork for our dual-core AMD Opteron processor when we designed our first single-core AMD Opteron processor in the late 1990's. Systems based on our dual-core AMD Opteron processor are the only dual-core/dual processor x86 servers you will be able to buy in 2005. (For those readers who want to dig deeper, we've provided a short note that explains how these dual-core processors work their magic.)

System Uptime and Repair Considerations

Blades can reduce the maintenance and repair night-mares that often plague complex IT installations, both in data-center and in branch locations. Anyone who "has ever tried to swap a 1U or 2U server out of a 19-inch rack knows it is a job best left to professionals. The server itself is heavy and awkward to handle, and a rat's nest of cables in the back of the cabinet must be dealt with in order to remove the failing chassis. Blade enclosures replace most of these cables with fixed backplanes and hot-plug connectors that enable even non-technical personnel to remove blades from the front of the rack, just by flipping a few latches and tugging on a handle, while the rest of the system operates as usual.

One or Many System Suppliers

Some IT organizations rely on one or two system suppliers for most of their server acquisitions, while others prefer to choose from a longer list of suppliers. Those in the latter camp may find that the proprietary nature of each supplier's blade enclosures and rack arrangements places uncomfortable constraints on their ability to "mix and match" components from different suppliers in the same rack or chassis. To gain the maximum benefit from blade deployments, users should generally source all of their compute blades for a given application from a single vendor. It is even more advantageous if the blade vendor

WHY IS EVERYONE SUDDENLY TALKING ABOUT DUAL-CORE PROCESSORS?

Until recently, chip suppliers emphasized the frequency of their chips ("megahertz") when discussing speed. Suddenly they're saying megahertz doesn't matter and multiple cores are the path to higher performance. What changed? Three factors have combined to make dual-core approaches more attractive.

- First, the shift to 90nm process technology makes dual-core possible. Using 130nm technology, single-core processors measured about 200 mm2, a reasonable size for a chip to be manufactured in high-volumes. A dual-core 130nm chip would have been about 400 mm2, much too large to be manufactured economically. 90nm technology shrinks the size of a dual-core chip to under 200 mm2 and brings it into the realm of possibility.

- Second, the shift to dual-core provides a huge increase in CPU performance. Increasing frequency by 10 percent (often called a "speed bump") results in at best a 10 percent boost in performance; two speed bumps can yield a 20 percent boost. Adding a second core can boost performance by a factor of 100 percent for some workloads.

- Third, dual-core processors can deliver far more performance per watt of power consumed than single-core designs, and power has become a big constraint on system design. All other things being equal, CPU power consumption increases in direct relation to clock frequency; boosting frequency by 20 percent boosts power by a least 20 percent. In practice, the power consumption situation is even worse. Chip designers need extra power to speed up transistor performance in order to attain increased clock frequencies; a 20 percent boost in frequency might require a 40 percent boost in power. Dual-core designs apply this principle in reverse; they lower the frequency of each core by 20 percent, so that both cores combined use only a tad more power than a single-core at the higher frequency. This means the dual-core processors showing up this spring will actually boost performance by a factor of approximately 1.7 over single-core designs that fit in the same thermal envelope.

You can read more about dual-core technology at: **www.amd.com/dual-core.**

can also supply complementary elements (network switches, SAN switches, load balancers, firewalls, etc.) in a format compatible with the vendor's blade architecture.

Aggregated or Disaggregated Disk Storage

Suppliers and users have yet to reach a consensus regarding the proper amount of rotating memory to be included in blade servers. Some suppliers (notably Egenera, one of the pioneers in the field) make no provision whatsoever for storage in any of their blades; they argue that SAN or NAS storage makes more sense, even for booting and paging operations. This approach ensures that no blade in the system will ever contain persistent data unique to that blade, and enables the easy swapping or provisioning of blades for specific tasks. Other suppliers, including the market leaders HP and IBM, offer users the ability to configure blades with or without local storage resources. Their blades use 2.5 inch ATA and SCSI drives, with a shift to 2.5 inch SATA drives in the offing. The form factors for the blades offered by many smaller suppliers allow the use of one or two 3.5 inch ATA and SCSI drives. AMD believes that the disaggregated approach offers many advantages with regard to system reliability, power consumption and thermal issues, but the company also recognizes that the shift to total separation between storage and computing resources will take years to accomplish.

A Snapshot of Current Blade Alternatives from "Top Tier" Server Suppliers

The five top-tier server suppliers (HP, IBM, Sun, Fujitsu Siemens and Dell) account for more than 75 percent of all x86-based servers sold[2], so their product strategies have a great impact on the market. All five companies have strategies that are remarkably similar in concept, but dramatically different in implementation. Each offers its own unique chassis, with a backplane engineered to accept only its own blades[3]. Customers should view their adoption of one or another of these vendor-specific blade architectures as the beginning of a long-term relationship, rather than as a short-term transaction. They should be confident that their supplier's roadmap addresses their current and future computational and networking requirements, since they will find it cumbersome if they later need to add subsystems that have not been specifically adapted to fit within that supplier's blade architecture.

IBM

IBM, the market share leader, with its BladeCenter products, currently offers dual-processor blades with Xeon or PowerPC processors, but has recently demonstrated AMD Opteron processor-based blades. IBM's blades are designed to fit in a 7U chassis that accommodates up to 14 dual- or 7 quad-processor blades. At roughly 4KW per chassis, a fully populated rack can consume up to 24KW of power.

HP

HP's BladeSystem (also known as the ProLiant BL), occupies second place from a market share perspective, and offers the broadest range of blade options in the industry. Consistent with its "chip-nostic" strategy, HP's line includes 2- and 4-way Xeon blades, 2- and 4-way AMD Opteron processor-based blades, and even a 2-way Itanium blade. The company's 4-way AMD Opteron processor-based blade is particularly innovative with regard to the way it melds two 2-way blades into one 4-way blade by physically joining the HyperTransport™ technology channels of the two modules into a single network. The BladeSystem can accommodate up to six 6U/16 slot chassis in a rack, for a maximum of 192 processors per rack.

Sun

The Sun Fire Blade System uses a relatively short 3U enclosure that can accommodate up to 14 single-CPU or 7 dual-CPU blades. A fully-populated rack can contain up to 156 blades with 156 processors. Sun currently sells blades with UltrasSPARC IIi, Xeon and AMD Athlon XP processors, but has publicly indicated its intent to deliver AMD Opteron™ processor-based blades in the near future.

Fujitsu-Siemens

The Fujitsu-Siemens Primergy BX600 utilizes a 7U chassis that holds ten dual- or five quad-processor blades; either way one chassis can accommodate 20 processors and a 6-chassis rack can hold 120

2 Source: IDC Server Tracker, Fourth Quarter, 2004

3 Dell's chassis and blades bear a striking similarity to those offered by Fujitsu Siemens, but neither company makes any claims about the interoperability of their chassis with components supplied by the other. Similarly, Intel resells a version of IBM's BladeCenter, but neither IBM nor Intel provides any assurances regarding the ability of its chassis to interoperate with blades supplied by the other.

processors in all. The dual-processor BX620 includes one or two Xeon processors, and one or two 3.5-inch SCSI drives. The quad-processor BX660 supports up to four Xeon processors and four SCSI drives. All that hardware consumes massive amounts of power; each chassis is rated at 4KW, which adds up to 24KW for a fully populated rack. (That's about the same amount of power used by six electric ovens.) Fujitsu-Siemens may have had a recent peek at Intel's long-term Xeon roadmap; they just introduced a liquid-cooled rack (PrimeCenter/LC) that helps customers get the heat generated by all these blades out their datacenters.

Dell

Dell's recently-introduced PowerEdge 1855 Blade System features a 7U chassis that holds up to ten dual-processor blades with Xeon processors and SCSI hard drives. The PowerEdge chassis and blades bear a striking similarity to those of the Fujitsu-Siemens BX600, although neither company has ever acknowledged an OEM arrangement. Like the BX600, the PowerEdge chassis is rated at 4KW, although Dell has yet to introduce liquid-cooled racks. Dell often claims it uses only Intel processors in its products because its customers don't ask for AMD processor-based systems. If you are a Dell customer in need of AMD Opteron processor-based blades for your PowerEdge 1855, you might want to drop a note to Kevin_Rollins@Dell.com.

A View of Blade Alternatives from "Tier-Two" Server Suppliers

More than half a dozen suppliers with brands few would recognize offer their own versions of server blade systems. Given their more limited resources and lower sales volumes, few of these vendors can afford to develop the proprietary backplanes and sophisticated enclosures that characterize the top-tier suppliers' blade systems. Instead, most of these companies have pursued a strategy akin to what one of them, Rackable Systems, has labeled "Open Blades." They utilize industry-standard ATX or extended ATX motherboards designed for use in white-box rack-mount servers.

These suppliers add value in the form of the cabinets they design to distribute power, manage cables, direct airflow, and simplify blade insertion. They also provide the assembly and test services needed to deliver the final configuration at the customer's site. Some suppliers centralize the fans that pull heat out of the enclosures; others centralize the power supplies that feed individual blades.

The buyers of these systems can choose from a smorgasbord of components that support virtually any microprocessor they prefer; indeed they can even mix and match blades with different motherboards and processors in the same rack. The experience of ordering an "open blade" systems is not unlike what one goes through when dining in a restaurant. The restaurant's designer provides ways for patrons to enter and leave the facility, and ensures that smoke from the kitchen exits via a flue and not through the dining room. The waiter may make some recommendations, but ultimately diners get to choose appetizers, entrees, dessert, and beverages, with no requirement that everyone at the table gets the same meal. The chef acquires the appropriate components and delivers the completed meal in a relevant time frame. Now that we've discussed what these suppliers have in common, let's review what each does in a unique manner.

Angstrom Microsystems

Angstrom Microsystems' SuperBlade currently holds the record for the most processors packed into a single rack, although they use a much larger rack than most. Angstrom can pack up to ten chassis enclosures into one rack, arranging five chassis on the front and five more accessed from the rear in a back-to-back arrangement. Each chassis holds up to 13 dual-processor blades, giving Angstrom a grand total of 260 AMD Opteron processors. Angstrom relies on very large cabinet-mounted fans to draw air through each of its blades, thus eliminating the need for array of small fans often found inside blades or rack-mount servers.

Rackable Systems

Rackable Systems' Scale Out Series offers the second-most dense configurations, with the ability to incorporate up to 92 dual-processor blades in one of its extra-large racks. Rackable orients its blades horizontally, rather than vertically, and can load four blades into each horizontal slot in its rack – placing the blades side by side and back to back. Rackable is one of the few suppliers to centralize the rack's power supplies, and can optionally distribute DC power to each blade. This approach allows the company to focus on more efficient power conversion, which the company claims can reduce overall power consumption by up to 30 percent.

Verari Systems

Verari Systems' BladeRack comes in a close third behind Rackable in the blade density sweepstakes. Verari can position four enclosures, each with eleven dual-processor blades, in the front half of its cabinet, and another four in the back half, for a total of 176 blades in all. Verari (which recently changed its name from the less memorable "RackSaver') emphasizes its unique bottom-to-top airflow design that pulls cool air in from a sub-floor plenum and expels warm air into a ceiling plenum, thus eliminating the need for individual, failure-prone fans in the blades themselves.

Appro

Appro's recently announced XtremeBlade accommodates 'only' 72 dual-processor blades (144 processors in all), but its ability to interconnect the HyperTransport™ technology links on up to four blades in order to create 2- , 4- or 8-way SMP nodes sets it apart from other blade suppliers. Appro has also moved aggressively to incorporate the latest technology (Infiniband, SATA drives and PCI Express™) in its blade configurations. Unlike most "tier-two" offerings, XtremeBlade relies on unique Appro-designed motherboards, rather than off-the-shelf ATX designs. This approach differentiates Appro's technology, but constrains the menu of choices Appro can offer its customers. For example, the ability to combine several two-way blades into a single n-way SMP arrangement works only with Appro's AMD Opteron processor-based XtremeBlades, and not with the Xeon versions of those blades.

Western Scientific

Although competitors have pursued wider and deeper racks in order to run up the number of processors they can accommodate in a single rack, Western Scientific emphasizes its use of standard 19-inch racks in its TeraBlade offering. This approach limits TeraBlade to 'only' 55 dual-processor blades (110 processors in all) per rack, but it makes it easier for Western Scientific's customers to integrate devices from other vendors into its TeraBlade configurations.

Egenera

Egenera's BladeFrame system marches to the beat of a different drummer; its approach differs conceptually from that of its "tier-two" peers as well as top-tier competitors. Most tier-two suppliers focus on maximizing density and minimizing per-blade cost for price-sensitive customers in the high-performance computing market. Egenera, one of the pioneers of blade computing, argues that most datacenters cannot accommodate the power and thermal requirements of these dense configurations, and offers systems with processor densities comparable to those attainable with 1U rack-optimized servers. Egenera's notion of a "Processor Area Network" (PAN) is what truly separates BladeFrame from other alternatives. Individual blades contain no storage or network interfaces. Instead, each blade in the PAN contains a single physical I/O device – a high speed link that ties the blade into a centralized switch fabric, which in turn is linked to external SAN, NAS, and Ethernet switches. Egenera provides virtual I/O drivers that hide the details of its switched environment from the applications and operating systems that run on individual blades. BladeFrame supports 24 horizontally oriented blades per rack, and includes both dual- and quad-processor blade configurations. BladeFrames with 24 quad-processor blades can accommodate up to 96 processors; the four-way arrangement of these nodes makes them better suited for a variety of "scale-up" tasks.

Conclusion

We hope this document has peeled away the marketing hype surrounding blade computing and given you a framework to decide whether blades can help your IT organization meet its objectives. Blades are no panacea, but they can simplify your environment, especially if you have a large number of tasks amenable to scale out approaches. Although we at AMD would love for you to pack every square foot of your datacenter with hundreds of our AMD Opteron™ processors, we urge you to be wary of brochures that promise such densities, unless your facilities contain far more power feeds and HVAC capabilities than usual. Instead, you should view blades as a means to help you deploy and manage your servers and thus create a more reliable, responsive and adaptive environment.

Although most of our system partners offer blade systems based on both AMD Opteron processor-based and Intel Xeon technology, we hope we've helped you understand why we believe AMD Opteron processor-based blades will often be a better solution for your computing requirements. In the unlikely event our message until now has been too subtle, our sales department asked us to be as clear as we can on these points. So here's the unvarnished sales pitch:

- AMD Opteron processor-based blades offer higher performance than brand "X" so you will need fewer of them. This lowers your acquisition and operational expense, and may save you money on software licenses as well.

- AMD Opteron processor-based blades use less power than brand "X" so they run cooler. You pay less for the electricity to run them, and less for the electricity to cool them.

- AMD Opteron processor-based blades are the only ones you can buy today that offer dual processor/dual-core capability.

- AMD's Direct Connect Architecture is more than a buzzword. Direct Connect Architecture boosts performance in two-way configurations, and provides the raw material that allows clever system designers create modular two-way SMP blades that can be snapped together like building blocks to form four-way blades.