

Enterprise computing

Business's digital black cloud

Jul 14th 2005 | LOS ANGELES
From The Economist print edition



New, faster computer chips are challenging the traditional structure of the huge business-software industry

FOR the past 40 years, companies around the world have grown accustomed to a doubling in computing power every 18 months to two years—fulfilling a remarkable forecast made in 1965 by Gordon Moore, one of the founders of Intel, a semiconductor powerhouse based in Silicon Valley. As their businesses have expanded, managers have been able to sleep easy in the knowledge that next year's computers would be more than able to keep pace with their needs and probably cost no more than last year's models. Alternatively, slowpokes with steady workloads have been able to replace ageing computers with flashier models costing half as much. The declining real cost of computing has been an economic boon.

Even as millions more transistors are crammed on to slivers of silicon, Moore's law continues to deliver the goods. But the tricks chipmakers such as Intel and AMD are exploiting to achieve this miracle are changing the whole approach to enterprise computing. In the process they are unleashing powerful disruptive forces. New chip architecture is allowing them to roll out ever more heavy-duty hardware at competitive prices.

The real losers in the pending upheaval could well be software suppliers. Firms such as Oracle, SAP and IBM, whose industrial-strength programs are the bedrock of business, could be badly bruised in the process. But inevitably, end-users—companies big and small that depend on enterprise software to do their various business transactions—are going to be feeling pressure as well. Over the coming year, they will have to keep their wits about them if they are to prevent their licensing costs from escalating out of control.

For many, the choice could come down starkly to accepting costlier new ways of being billed for the corporate software they depend upon for their livelihoods, or biting the bullet and switching to "open source" programs that may be free to license but have plenty of hidden costs.

The current brouhaha over software licensing has been set off by the arrival this month of large quantities of chips containing two central-processing brains (or "cores") on the same device. Chipmakers have known for some time that merely cranking up the internal speed of their computational engines was delivering diminishing returns.

Today, the co-ordinating internal clocks on some of the fastest chips beat at four billion times a second (4GHz in geek-speak). This furious internal activity gobbles up electrical power. In turn, that makes the chips scorching hot. Cooling them down so they can do their job properly has become a costly nightmare, especially when such chips are used in cheap, but poorly ventilated, "blade" computers (so-called because they are wafer-thin and plugged together in racks like packs of disposable razor blades).

The answer has been to put two or more smaller cores on a single chip. By sharing the workload, the separate cores produce less heat. But being on the same tiny piece of silicon, they still have the speed advantages that come from having all the core's ancillary components within easy reach on the same device.

Actually, dual-core processing is nothing new. IBM and Sun Microsystems have been supplying processors with two or more cores built into them for several years. But these have been thoroughbred chips for powerful Unix workstations used by scientists and engineers for cutting-edge research.

The difference today is that the dual-core approach is now being applied to the workaday processors that run the vast majority of Microsoft Windows, Linux and other popular programs designed to exploit the internal instructions used by Intel's ubiquitous Pentium processor. Such chips power not only hundreds of millions of personal computers, but also the tens of thousands of back-office servers that dish out data over computer networks to employees throughout an enterprise. It is the latter that are the mainstay of companies' IT departments everywhere.

In April, Intel and AMD announced separately that they would be delivering dual-core versions of their high-end processors later this year. AMD has been the first to ramp up production of its new device and has now started delivering its dual-core Opteron processor in commercial quantities for \$2,650 apiece. Intel is expected to start shipping dual-core versions of its Xeon and Itanium server chips in volume later this year. Meanwhile, the leading server manufacturers, including Hewlett-Packard, Sun Microsystems and IBM, have begun taking orders for their new dual-core computer systems.

With the wholesale switch to dual-core processing, some software firms feel they are about to be short-changed. If two cores on a single chip can do twice the work of a single processor, they argue, then customers paying licence fees based on the number of processors running their software (one of the most common forms of software licensing) will be getting a free ride on half the new cores being deployed. Actually, because of internal losses and design restrictions, dual-core chips tend to do the work of anything from 1.3-1.8 comparable single-core processors, depending on the applications they are running. But the free-ride argument still stands.

The core of the matter

As the dominant supplier of database software, Oracle has expressed its concern about the shift to multicore processing and is adamant: customers will be charged by the core rather than the processor. The firm actually uses two different forms of software licence. One is based on named users and is for

customers with a more or less fixed number of defined users. The other is for firms with populations of users that are hard to define, and is based on the number of processors within an enterprise that are running Oracle software. Customers can choose one or the other.

Behind the scenes, however, Oracle is actually more flexible than the image it presents. In the past it has used other means to price its software—concurrent user, named user on a single server, named user on multiple servers, processor, you name it. Some years ago, when Oracle based its pricing on the performance of the underlying hardware, licensing fees leapt in line with soaring chip speed. Faced with a customer rebellion, Oracle dropped its power-based metric in favour of a simple processor-based policy. "We've found that our customers are more satisfied when they can easily identify and predict what their licensing fees will be," says Jacqueline Woods, vice-president of global pricing and licensing strategy.

That is the official line. But when customers complain to Oracle's sales representatives about the firm's aggressive pricing, they are quietly advised to talk directly to customer support. The company does not admit to any backdoor deals, saying only that it is "committed to providing customers with simple, flexible and transparent pricing."

IBM has been even more circumspect, announcing that its software licences for computers running single-core or dual-core versions of the Opteron or Xeon processors will cost the same. But IBM has yet to announce what its licensing policy will be for running its big software suites, such as DB2 and WebSphere, on computers powered by Intel's more advanced Itanium multicore processor. Even so, the announcement on Opteron and Xeon licensing was a big turnaround for the firm: it has charged customers on a per-core basis for running its software on the powerful dual-core Power5 processor it makes in house.

IBM says the boost in power that customers will get from the new dual-core chips from AMD and Intel will be only incremental. Its own dual-core Power5 is a third-generation processor for top-of-the-range servers, which deliver double the value in the highly tuned applications they are bought for. But insiders suspect that IBM's change of heart—at least where the more popular dual-core processors from AMD and Intel are concerned—was a ploy to cast Oracle in an unfavourable light, and try to steal market share.

Another software firm that is seeking to capitalise on the computer industry's current turmoil is Microsoft. When the licensing issue was first broached publicly last October, Microsoft came out uncharacteristically on the side of customers. In a move calculated to win users over, the Redmond-based company announced that it would be licensing its server software on a per-processor basis. That meant only one Microsoft licence would be needed for any dual-core Opteron or Xeon server. The same would apply to Itanium systems when they arrived, too.

Altruism played no part in Microsoft's commitment to processor-based licensing. By ensuring that customers using its Windows Server family of products (such as Microsoft SQL and Microsoft BizTalk) would not have to pay any more when they upgraded to multicore processors, Microsoft very effectively seized at least the low ground in ongoing debates about total cost of ownership (TCO) and return on investment (ROI) of computer purchases. (These abbreviations have become standard parlance among IT managers since the bursting of the dotcom bubble.)

By addressing these concerns with a cheap and easy route to multicore computing, Microsoft is positioning itself to grab more of the lucrative market for enterprise software, where, unlike the desktop market, it has faced tough competitors with beefier products and well-entrenched positions in customers'

premises. But if they do not respond in kind, Oracle, SAP, IBM, BEA, Siebel and Veritas could find themselves losing out to Microsoft's budding family of server software and its processor-based licensing. On top of that, users will not be amused by software suppliers that raise licensing fees for applications that run on top of a Windows server, especially when Microsoft has not changed the licence for the underlying platform.

Hard to be soft

Users have lots of other grouses. With some justification, they argue that software suppliers—at least the vast majority that license their products on a user or processor basis—did not raise prices during the megahertz race when computers became faster and more efficient. Likewise, they say, dual-core architecture is just another way for the hardware-makers to boost the speed of their machines, and should thus be treated the same as raising clock speed.

In short, users feel it is grossly unfair for software firms to charge more for improvements that stem entirely from buying better hardware. It is not as though IBM, Oracle and others have had to rewrite their big database or transaction programs so they can run on the new processors. In fact, the software will not even notice the difference. Some customers have likened Oracle's insistence on core-based licensing to a form of double taxation.

But users admit that change is probably inevitable. Dual-core Opterons and Xeons are only the beginning. Computers with processors that use four or more Pentium-like cores will start arriving in 2006. And industry watchers expect the trend to ever greater numbers of cores to accelerate. For instance, the New York office of Ogilvy & Mather, an advertising agency, is currently testing a computer appliance built by a small computer-maker called Azul Systems that uses 24 cores per chip.

Complicating matters even further, servers these days tend to have more than one processor inside them. The wholesale endorsement of Linux, the open-source operating system for Pentium-style processors, by the world's leading hardware and software firms (with the exception of Microsoft) has turned it into a formidable platform for enterprise servers capable of ganging dozens, even hundreds, of internal processors together into a giant "symmetric multiprocessing" unit that behaves as a single entity. Add multicores to multiprocessors, and today's software licensing policies quickly become untenable.

The software industry's licensing dilemma does not even end there. There are two other developments in computer design that could cause licensing anarchy. One is known as "partitioning and virtualisation". Broadly, this involves using a single computer to create the illusion of having multiple computers, each with its own operating system such as Unix, Linux, NetWare or Windows; and each acting as if it had exclusive access to all the real computer's resources (eg, memory, drives, network adapters, communications ports, etc), without regard for all the other operating systems installed that think likewise.

Slice and dice

Virtualisation is actually an old computer trick, stemming from IBM's mainframe computers of the 1970s. It resurfaced in the late 1990s thanks to some nifty software developed by a company called VMware that lets users slice up a Pentium-style computer as if it were many different machines running different operating systems and software. That started off as being a handy way for IT managers to test new software configurations before installing them company-

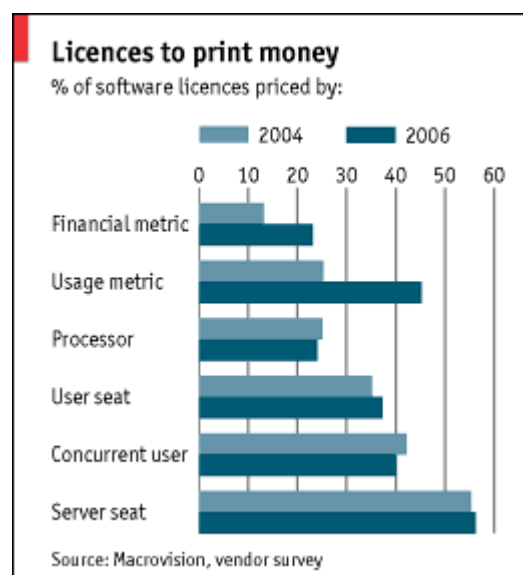
wide. By 2002, however, customers had begun to realise that a lot of the servers they had hurriedly acquired during the great Y2K panic at the turn of the millennium “were running with no more than 5% to 10% utilisation,” says Raghu Raghuram, senior director for strategy and market development at VMware.

Today, VMware's ESX and GSX servers and Microsoft's Virtual Server are being used to get more out of users' hardware investments, by allowing existing machines to run best-of-breed software, no matter what operating system it was designed for. While virtualisation may be great for hardware ROI, it wreaks havoc on software licensing policies.

Then there is the industry-wide trend to “rapid provisioning”. This is a way of providing capacity-on-demand, explains Amy Konary, director of software pricing and licensing at IDC, a computer consultancy. The idea is to make whole computers, parts of hard-drives (ie, partitions) or even individual cores instantly available, along with an operating system and entire stack of software applications, for a particular task that has cropped up and needs urgent attention.

Hewlett-Packard and IBM have developed new types of licences that allow some software to be used in such an intermittent manner. But the rest of the industry is still agonising over how to license rapidly provisioned machines. One thing is for sure, says Ms Konary: users are unlikely to accept any requirement that makes them pay for a full software licence for such momentary use.

Clearly, the days of licensing software on a simple one licence per installation or on a straight per-user or per-processor basis are ending. A recent survey* conducted by a handful of software firms and trade associations in Silicon Valley found that while software suppliers were pushing for licensing models based on annual subscriptions, companies overwhelmingly preferred single, one-time payment methods, be they on a per-user or per-processor basis (see chart). This was surprising because users have been decrying the high up-front cost of software and have been urging suppliers to provide much better value for money.



But given all the imponderables, it has become extremely hard, if not impossible, to quantify what the value of any given piece of software is. What is known is that negotiating licences is not a trivial exercise. John Fowler, executive vice-president of Sun's network systems group, finds that companies spend typically between eight and 12 weeks planning and discussing software licences with their suppliers. In its bid to answer the value conundrum, Mr Fowler's firm has adopted the simplest of financial metrics. It charges firms a straight \$140 times the number of employees on the customer's payroll for using its proprietary software. Why \$140? Because it seems to correlate with the price that the company and its customers think is good value for having no hassles. The simple subscription gives customers the unrestricted right to run Sun's software on as many computers, by as many people, and as often, as they like.

Others think that a better way of delivering value is to adopt some kind of utility model for software pricing—allowing users to pay only for what they use. Macrovision likens computer upgrades to a city upgrading its water supply system (hardware) to allow residents to use more water (software). The question, then, is should residents who do not want to use more water pay more for at least having the opportunity to do so? Opinion is divided.

The maintenance model

But if Margaret Lewis, AMD's senior software strategist, is right, software pricing could be in for some radical rethinking. In her view, the open-source model of software licensing looks like being surprisingly attractive. This involves software firms licensing their products for nothing, while earning their keep from maintenance and support. Software firms such as Red Hat, Novell, Mandriva and a host of other Linux distributors already make their living that way.

Indeed, more and more suppliers of proprietary software are beginning to think along similar lines. Sun has recently made its flagship product, the rock-solid Solaris operating system, available with a form of open-source licence that permits it to be downloaded free of charge. Meanwhile, IBM has re-invented itself as a successful service company thanks to the way it has embraced open-source Linux and J2EE (Java 2 Enterprise Edition).

The clincher is that if software firms continue to think they can cash in on every new increase in computer performance, they will only encourage more and more customers to defect. And today, unlike a decade ago, open-source software has become just too good to be ignored. MySQL or PostgreSQL, two powerful open-source databases running on Linux, have become attractive alternatives to commercial products such as Oracle 9 or DB2 running on some proprietary flavour of Unix from Hewlett-Packard, Sun or IBM. The same goes for open-source servers such as Apache, JBoss and Samba.

Today, customers have a plethora of alternatives that should give enterprise-software firms pause for thought. Finally, of course, there is Microsoft. No one in the enterprise-software business should underestimate its determination to own their market as well as the desktop business it already dominates.

* "Key trends in software pricing and licensing", available from SoftSummit (www.softsummit.com).