

Compute farms

MANY LINUX BOXES MAKE FAST WORK

ALAIN WIEDMER

Clustering is a technology that enables you to perform computing tasks faster than they would run on a single machine. Linux is the ideal operating system for this application.



A compute farm is a lot of computers strung together to run power-hungry applications in record time. When you do this, you need to use a reliable and secure operating system. There's a tendency to go for the reassurance of using well-marketed and supported software, even though it ties you into a particular vendor's expensive hardware. But what about Linux? Just because its commercial backing isn't yet as great as that

of its rivals and the technology hasn't been cemented into our minds by clever and persistent marketing – the so-called "Microsoft effect" – it doesn't mean that Linux isn't mature enough to do the job. But is Linux the answer for every compute farm, or is it still a niche choice for those who have advanced technical expertise and those suffering processing budget constraints?

Linux-based enterprise server sales are riding high, suggesting that the technology is increasingly being deployed by users keen to take advantage of Linux as a more cost-effective, open source alternative to standard proprietary operating systems. Linux-based enterprise server sales worldwide are growing at a compound annual growth rate of 57% according to IDC, having increased from \$320 million to be worth \$4.1 billion today, and an estimated \$7.4 billion by 2002. In comparison, total server sales worldwide encompassing every operating system has a

compound annual growth rate of 17% for the same period. Moreover, the latest Eetimes EDA 2000 research study shows that there will be more than a three-fold increase in the use of Linux over the next two years, from 11% to 38%.

Phenomenal growth

Much of this phenomenal growth in the use of Linux will come from its use in compute farms. Clustering Linux-based systems gets around the problem that Linux is not as scalable as other Unix-based operating systems. Remove the issue of scalability and Linux is easily comparable in quality to any other mature operating system and far outstrips the rest in terms of cost-effectiveness.

Running on Intel architecture, the performance of which is comparable to that of 32-bit Unix and RISC-based systems, Linux-based compute farms are easily the most cost-effective way of getting the power needed to run big or complex applications. Linux can run on any make of PC or server right up to the biggest RAID arrays of SCSI devices and Gigabit Ethernet interfaces. That means you can use your existing kit or buy the cheapest on the market without being constrained by your choice of manufacturer.

Being free to choose hardware that delivers a low cost of ownership means that the decision to throw it out after 18 months is not such a financial headache. In fact, clustering or farming processors is a cost-effective method of extending the life of hardware, because new hardware can be added and obsolete hardware retired "on the fly." More and more industries are realising the benefits

of using server farms to undertake CPU intensive computing tasks. As a consequence, Linux is steadily becoming an operating system of choice mostly because of the value for money it delivers. The most popular server in the world, the open-source Apache web server, runs on Linux, as do media servers such as the RealServer, encoders, secure enterprise information management systems and many e-mail, news and other types of server.

But for Linux, being open source software is as much of a minus as it is a plus. Whilst it means that if you've got a problem you can quickly fix it yourself, it also means that if you get a problem you *have* to fix it yourself. This puts off some users from betting their business on Linux. They worry about issues of support and whether or not staff are sufficiently technically skilled to tinker with the software when tailoring needs or problems arise.

Mature and secure

Nevertheless, Linux is a mature and secure network operating system with all the robustness, familiarity and power of Unix. Also, hardware manufacturers such as IBM and SGI are increasingly offering support contracts that cover different flavours of Linux so you can have the reassurance of an all-inclusive hardware and software services contract.

Linux certainly delivers on the functionality requirements of a compute farm. The new 2.4 kernel adds excellent multiprocessor support and very fine-grained security capabilities down to Internet Protocol level. "Because Linux is built from the ground up as a secure network operating system you need never visit your server unless it is to change its hardware," says John Hayward-Warburton, an independent consultant and multimedia producer whose clients include Hyperion Records. "I have servers that have never been visited since before the beginning of this year and have been able to keep them fully up-to-date with operating software, kernels, etc, from my desk on a hilltop in Herefordshire." He adds, "I will not hand control of my clients' valuable data and hardware over to any closed source company."

While Linux is reliable, customisable, and comes free from license fees, it doesn't usually come pre-installed on systems. But management is easy using your own custom-built tools scripted in shell languages, Perl or Python, or those developed by commercial enterprises. "That does mean that you need more technical knowledge than you would of your average operating system, especially if you intend to tinker with the software," says Bill McMillan, a technical specialist in product management with Platform Computing. "If things fall apart then it's your problem!"

Platform is the developer of LSF, a distributed resource management package for computing clusters. The company backed Linux in response to increasing customer demand. Indeed, LSF is used by



chip design company Transmeta, which employs the father of Linux, Linus Torvalds, as a software engineer. Transmeta operates a pool of 600 high-performance processors for the design of its Crusoe range of chips, which calls for complex simulations and builds. Without LSF Suite, the simulations would take around a week to run instead of hours, and between eight and ten times the amount of computing power would be needed, according to Transmeta's IT manager, Ray Borg. He adds, "If I turn LSF off today, my week-long job could take three or four weeks."

Most of Transmeta's machines run Linux but some run Sun Microsystems' Solaris operating system. LSF ensures that jobs run transparently and reliably on the most appropriate machine within the cluster. At Transmeta, LSF ties Linux and Solaris together in a seamless fashion so that the pool becomes one giant processor. "Linux takes advantage of resources a lot more smoothly and efficiently than other software brains," says Borg.

Predictably, Linux has gained an early foothold in sectors where staff are more techno-savvy, such as in the rendered film-resolution market and in electronics design (EDA). EDA users have pushed for Linux because they want more flexibility out of their operating system so they can tinker with configurations, which the likes of Windows NT don't allow. Other sectors are really only behind in their adoption of Linux because they're waiting for necessary software to be ported to it. However, with Oracle and IBM backing Linux, appropriate enterprise solutions won't be long in coming.

"In our mind, there's now a level of agreement in the industry about what is required from a base operating system, and Linux meets that requirement," says John Fleming, SGI's UK marketing manager. "For users, if they want to be totally in control of their own destiny and build their own software and hardware compute farm, clearly Linux is the only choice. For the rest, Linux is still the best option, but using the services of a vendor like ourselves which offers a Linux conceptual framework, commodity components and assembly."

He concludes, "Linux for compute farms is here and it's here to stay. The question that's more open is: will Linux invade the desktop space? There the answer is a little less obvious, but it's coming." ■

[links]
A cluster of SGI computers
running Linux

Info

Platform Computing

<http://www.platform.com/>

SGI

<http://www.sgi.com/>

THE AUTHOR
ALAIN WIEDMER IS VP
SALES EMEA FOR
PLATFORM COMPUTING