

NEW RIVER _____
Marketing Research

White Paper:

Scyld Beowulf™ Cluster Operating System: Best of Breed for Linux Clusters

Bill Claybrook
September 2004

Executive Summary

Penguin Computing, founded in 1998, was one of the first vendors to specialize in Linux systems. It supplied Linux adopters with servers and early on focused on enterprise customers. With the acquisition of Scyld Software in 2003, it acquired the Scyld Beowulf™ cluster operating system designed and developed by Don Becker, a pioneer in Beowulf computing. Today, Penguin Computing provides a broad range of rack-mounted servers — utility servers, enterprise servers, and cluster-optimized servers — turnkey Linux cluster solutions, and the best of breed Scyld Beowulf Linux cluster operating system (referred to as Scyld Beowulf).

The focus in this *White Paper* is on Scyld Beowulf and the benefits that it provides, as a Linux operating system-based cluster management solution, for ISVs in the high performance computing (HPC) market segment. Later on, the company will target high availability (HA) and the enterprise market for commercial applications where the cost-effective management of hundreds or thousands of loosely connected servers becomes crucial.

Penguin Computing is not a super computing company — it is not targeting the large government research laboratories with its Linux cluster technology. Instead, it is focusing on smaller cluster installations, typically between 32 and 256 processors, with the needs of Fortune 500 companies in mind. A major goal of the company in HPC is to provide high performance, low cost computing solutions to those engineering organizations that need to run simulations or synthetic case studies to fulfill their

missions. Targeted verticals include oil/gas, automotive (crash analysis, structural analysis, combustion), mechanical computer aided design/computer aided engineering (MCAD/CAE), electronic design automation (EDA), aerospace, bio-technology/life science, digital content creation (DCC), and financial services.

Scyld Beowulf is a Linux distribution, compatible from an application point of view with Red Hat Enterprise Linux (RHEL) ES 3.0. Any application that runs on RHEL ES 3.0 will run on Scyld Beowulf with no porting. Scyld Beowulf is not just for Penguin Computing servers and Linux clusters built from Penguin Computing's competitive line of AMD Opteron and Intel Xeon servers. It is equally applicable for hardware platforms from other vendors such as Dell, HP, and IBM as well as for Intel motherboards. However, if ISVs or end users prefer one stop shopping, then they can purchase an integrated cluster solution based on Penguin Computing's servers and Scyld Beowulf.

Linux clusters offer the most economic solution for compute- and data-intensive applications such as those found in HPC and for cross-over applications (applications such as portfolio management and risk analysis in financial services) as the market shifts from UNIX-based symmetric multiprocessing (SMP) to Linux clusters. Linux clusters often offer 5x to 20x price performance advantages over SMP machines. Because compute nodes in a cluster powered by Scyld Beowulf can be diskless, Linux clusters built around diskless blade servers could provide even more significant price/performance advantages.

Today, most high performance computing ISVs port their applications to Red Hat and/or SUSE Linux distributions, and then they have to interface with another vendors' Linux cluster management software to gain the benefits of Linux clusters. Scyld Software reduces the amount of work required for an ISV to port their applications to Linux cluster environments. Scyld Beowulf is both a Linux distribution and a Linux cluster management solution that provides a single system image (SSI) to ISV applications, system administrators, and users.

ISVs can use Scyld Beowulf to configure their applications for a Linux cluster environment; users and system administrators use it to submit jobs and manage clusters, respectively. No other vendor provides this triple capability in a single cluster software product. *Scyld Beowulf eliminates one of the drawbacks to the acceptance of Linux clusters as a replacement for large, expensive SMP systems — the lack of Linux cluster management solutions.*

In this New River Marketing Research (aka New River) *White Paper*, we discuss why users are moving to Linux clusters, provide Scyld Software's marketing strategy for Scyld Beowulf, and analyze the benefits that Scyld Beowulf presents to ISVs.

Market Trends — Users are Moving to Linux Clusters

The history of HPC is an interesting one highlighted by the introduction of new hardware and software architectures. The first large computers were parallel processor and vector machines, primarily from Cray. The next machines were the massively parallel

processor (MPP) and SMP computers. In 1993, the Beowulf Project spawned the development of clusters from commercial-off-the-shelf (COTS) machines. By 2000, Linux clusters were changing the price/performance curve for HPC. Today, HPC is dominated by large and expensive RISC/UNIX-based SMP servers and low cost, high performance Linux-based clusters. But Linux clusters are dominating new deliveries of HPC solutions.

Linux clusters are based primarily on AMD Opteron and Intel Xeon processors (including Intel's new Xeon Extended Memory 64-bit Technology (EM64T)). New River expects this trend to continue for at least the next few years as RISC/UNIX and some vector machines are replaced by flexible, low cost Linux clusters based on industry standard servers.

Systems suppliers with the "best" industry standard system portfolio and Linux cluster strategy can expect to gain significant pieces of the overall Linux cluster market. This market is expected to increase to over \$2B in 2005 with about 70% - 80% being in the HPC market segment. The Linux cluster market in 2003 was more than one third of the overall Linux server market in terms of revenue.

The worldwide Linux cluster market is expected to grow faster than the worldwide Linux server market over the next 2 - 3 years as transition continues to take place from RISC/UNIX technology to industry standard server and Linux operating system technology. This growth will occur because buyers in HPC focus on price/performance, and Linux clusters have a 5x to 20x price/performance advantage over heritage RISC/UNIX platforms.

The price/performance advantages of industry standard-based Linux clusters is driving another market phenomenon — the rapid expansion of the HPC marketplace as companies and firms, who heretofore could not afford this type of computing, can now make an investment in Linux clusters and improve their overall productivity and product quality. The ISV community is well aware of this growth. As a result, many of the top applications providers are porting their applications from UNIX to Linux and writing new applications for Linux in anticipation of this groundswell to industry standard Linux clusters.

The transition to Linux clusters is substantiated by a recent survey of HPC scientists who indicate that 85% - 90% of HPC applications can be run on Linux clusters today, after a simple re-compilation from UNIX. In verticals such as life sciences and pharmaceuticals, 65% and 85%, respectively, use clustered platforms and more than half of these are Linux clusters. In the DCC vertical there is a similarly substantial use of clusters as most rendering algorithms lend themselves to this architecture.

One of the primary issues surrounding the use of Linux clusters has been cluster management, an area in which Scyld Software excels with its Scyld Beowulf cluster operating system. Linux cluster management solutions are designed to monitor and manage a distributed set of jobs across multiple computers.

*New River Marketing Research
16 Deer Path Road, Suite 6
Maynard, MA 01754*

*Telephone 978 793 1357
Fax 978 897 4804
www.newriverlinux.com*

Marketing Strategy

Penguin Computing and its division, Scyld Software, are focused on the HPC market segment and on crossover applications — compute- and data-intensive applications such as portfolio management and risk analysis in the financial services market segment. Later, the company will add high availability (HA) solutions to its product offerings and spread its cluster focus to include the enterprise market for commercial applications. A major goal of the company is to provide a high performance, low cost computing solution (with an emphasis on replacing UNIX-based SMP machines) via Linux clusters powered by its Scyld Beowulf operating system.

An important part of Scyld Software's strategy is to attract HPC ISVs to certify their applications on its RHEL ES 3.0 compatible Scyld Beowulf cluster operating system. Scyld Software's targeted verticals include (the dollar figures are for investments in hardware in 2003):

- MCAD/CAE (manufacturing, automobiles, aerospace, etc.) — \$1.5B
- EDA — \$.3B
- Life sciences (genome research, chemistry, pharmaceuticals, etc.) — \$1.4B
- Geosciences (oil/gas exploration, etc.) — \$.5B
- Financial services (risk analysis, portfolio management, etc.) — \$2.6B
- Entertainment (rendering, DCC, etc.) — \$.3B
- Government (fluid dynamics, large physics simulations, etc.) — \$1.5B

Today, more in-house developed codes and ISV applications run on large SMP UNIX machines than run on Linux clusters, but companies in the above verticals are buying Linux clusters at a much faster rate than UNIX systems because of the huge price/performance advantages of Linux clusters. Many ISV applications already run on Linux and/or Linux clusters. They include Fluent's computational fluid dynamics software, automotive and aerospace applications such as Nastran, Patran, and Marc from MSC, TurboBLAST from TurboworX, FAST/TOOLS from Yokogawa ISS, risk management software from Algorithmics, Maya (DCC software) from Alias!Wavefront, Renderman from Pixar, and others. Many of the applications in life sciences, entertainment, and financial services are developed by ISVs and already run on Linux or Linux clusters.

The Benefits of Scyld Beowulf for ISVs

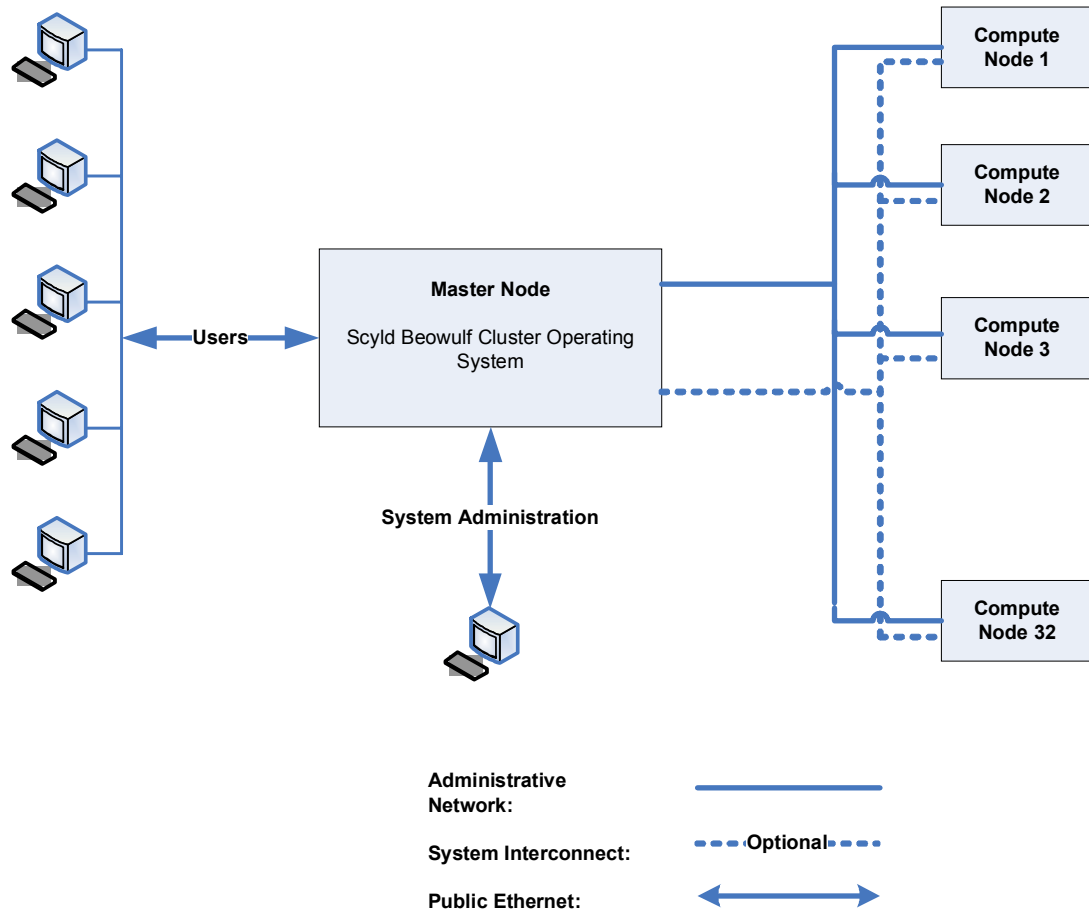
Figure 1 gives a pictorial view of where Scyld Beowulf fits into a Linux cluster architecture. A major benefit for ISVs who use Scyld Beowulf is that they can certify their applications on Scyld Beowulf and then use it to configure and schedule their applications.

*New River Marketing Research
16 Deer Path Road, Suite 6
Maynard, MA 01754*

*Telephone 978 793 1357
Fax 978 897 4804
www.newriverlinux.com*

Scyld Beowulf effectively sits in the middle — between users and systems administrators and the cluster hardware. ISVs use Scyld Beowulf to configure their applications for a Linux cluster environment and users and system administrators use it to submit jobs and manage the cluster, respectively.

Figure 1. A Beowulf Cluster Architecture with Scyld Beowulf at its Center



Source: New River Marketing Research, September 2004

Users who deploy a Scyld Beowulf-enabled application need cluster hardware, Scyld Beowulf, and an application. Users who do not deploy a Scyld Beowulf-enabled application need hardware, Linux, a third party cluster management solution, and an application. When ISVs cluster-enable their applications using Scyld Beowulf, they reduce the costs of the end user for deploying and managing their applications. The cost of Scyld Beowulf is typically between 10% and 20% of the cost of the cluster itself.

Scyld Beowulf Extends Linux to Clusters

Generally, ISVs do not port to an operating system platform unless there is sufficient customer demand to warrant a port. This is the reason why few ISVs ported to Linux until the last two years. Today, there is high user demand in some market segments for Linux clusters because of the significant price/performance advantages that they hold over other platform solutions, most notably UNIX-based SMP platforms. Porting applications in these market segments (oil/gas, automotive, financial services, etc.) to Linux is one of the steps. *Linux, however, does not have the capability to configure, manage, and schedule applications in a Linux cluster because it is not designed for a cluster environment.* This same type of phenomena also occurs in the embedded Linux area in which vendors extend Linux to handle real-time processing requirements.

Why can't Linux configure and schedule HPC applications? HPC applications are generally classified as being *serial* or *parallel*. Serial applications are traditionally single machine applications; that is, they are applications that run on one processor. Parallel applications run asynchronously as a single application that has been partitioned into two or more tasks that run on different processors. Scyld Beowulf provisions and manages parallel applications from a single point to greatly simplify administration and end-user access.

Many applications in the HPC space run as serial applications, and they are often considered to be the least suitable for Linux clusters. But this is not always the case. Many HPC designers perform what are called parametric simulations, i.e., the same simulation application is run with slightly different sets of input data. For example, the same airfoil is simulated at different air speeds; different airfoils are simulated at the same air speed, etc. This design methodology works perfectly with the architecture of a Beowulf cluster. Each of the simulations can be dispatched by Scyld Beowulf to run on a different compute node. Because they have no need to communicate with one another, a fast interconnect is in general not required, nor is it necessary to utilize an expensive SMP system. The advantage of Scyld Beowulf in this context is that the management of a serial application running on multiple nodes (with different sets of data) occurs under a single point of administration and usage, greatly simplifying both the system administrator and engineer's efforts.

Linux, without parallel library extensions, does not have the capability to effectively configure and schedule parallel applications. Scyld Software understands that a single scheduling algorithm¹ is generally not appropriate for all applications, and they also understand that ISVs know how best to map the tasks (processes) of their applications to nodes in a cluster. For example, an application task may only run well with one process per node regardless of the number of processors per node (for today's architectures, many compute nodes are configured with dual processor servers). Most schedulers would not schedule processes in this manner. They would schedule two processes per node if the node were a two-processor server.

¹ Scheduling consists of job initiation, assigning jobs to processors, monitoring jobs, job control, exit status reporting, and environment setup/tear down.

To resolve this scheduling issue, Scyld Software has implemented scheduling in Scyld Beowulf via a library, making scheduling extensible. This allows an ISV to substitute its scheduler(s) for its applications, or they can use the default scheduler in Scyld Beowulf. Because an ISV application is aware that it is going to be run on a cluster, there are more opportunities to enhance scheduling than on site-wide schedulers such as those found in Platform Computing's LSF and open source offerings such as Sun Grid Engine (SGE). However, Scyld Beowulf easily permits distributed scheduling algorithms such as those found in LSF and SGE to be added. This means that any ISV application that has been integrated with LSF or SGE (and runs on Linux) can be run on a Scyld Beowulf cluster, and the cluster can be managed via Scyld Beowulf.

Note that LSF, PBSPro, and GridServer do not have the capability to configure and manage clusters. They are essentially distributed schedulers. That is, when LSF is deployed, it requires a third party management solution to manage the underlying cluster.

Scyld Beowulf is a Linux Distribution

Release 29 of the Scyld Beowulf cluster operating system runs on the master node (see Figure 1) in a Scyld Beowulf cluster. It is a complete Linux distribution, compatible with RHEL ES 3.0 libraries and utilities, and inherits most of the same source code tree as RHEL ES 3.0. The installation process is identical to the standard Red Hat Linux installation with the exception of one additional screen. It is effectively a distribution with extensions to the kernel to provide a SSI of a Linux cluster to applications, ISVs, users, and system administrators. Any Linux application that runs on RHEL ES 3.0 will run on Scyld Beowulf with no alterations.

Compute nodes (see Figure 1) in a Scyld Beowulf cluster run a minimal Linux kernel that is full featured for the benefit of applications — a full-featured library and kernel. As far as an application is concerned it is running on a full Linux kernel. Using the minimal kernel, a compute node can be dynamically configured and provisioned in a few seconds by Scyld Beowulf. Scyld Beowulf controls the entire cluster. It configures compute nodes as they are powered up, runs the filesystem, and monitors the health of compute nodes. Scyld Beowulf also stores all application binaries on the master node, and compute nodes get binaries from the master node when the job starts.

The lightweight kernels can provide performance improvement during application execution. Performance improvement has been measured between 5% and 10% in some application environments. Applications run faster on Scyld Beowulf cluster compute nodes than on cluster nodes with a full Linux install because compute nodes run no services, just applications; no scheduling interference; no memory management contention (less memory fragmentation), etc.

Compute nodes in a cluster can be booted (that is, the lightweight kernels can be installed) and run without any disks attached². As a result, Scyld Beowulf is the perfect

² Compute nodes can also be booted and run with disks and other local media attached.

cluster operating system to power diskless blade servers — only the blade server acting as the master node and running Scyld Beowulf would require access to a local disk or shared disk(s) in a chassis. A cluster built using a rack of diskless blade servers and powered by Scyld Beowulf could provide very appealing price/performance characteristics.

Today, Scyld Beowulf clusters support the notion of a single master node per cluster. However, by mid-2005, they will support multiple master nodes per cluster with more and more sophisticated fail-over mechanisms supported in the roadmap. Initially, one master will be responsible for configuring a compute node, but subsequent to that other master nodes can schedule applications to run on it. In addition, future releases of Scyld Beowulf will support a virtualized environment. This will be a critical feature and a major competitive advantage when large farms of loosely connected servers, as opposed to relatively closed-coupled clusters, will be managed by Scyld Beowulf.

Summary of Scyld Beowulf Benefits

In summary, Scyld Beowulf and Scyld Software deliver the following benefits for ISVs:

- Scyld Beowulf is hardware independent; it runs on AMD and Intel hardware available from most server vendors
- Scyld Beowulf-enabled ISV applications greatly simplify the deployment and management of applications
- Scyld Beowulf-enabled ISV applications save users money (enabling lower total cost of ownership) because the deployment system, Scyld Beowulf, and the management solution are one and the same, reducing training and system administration costs — ISVs who Scyld Beowulf-enable their applications reap the benefits of customer satisfaction
- Scyld Beowulf is a Linux distribution compatible with RHEL ES 3.0
- Scyld Beowulf is designed for cluster environments
- Scyld Beowulf provides a SSI for applications, users, and systems administrators
- Scyld Beowulf prevents software versioning skew because updates for all elements of a cluster are done via Scyld Beowulf on the master node
- Scyld Beowulf provides scalable performance
- Scyld Software has leading Beowulf cluster expertise with Don Becker, one of the pioneers in Beowulf cluster technology

- The Executive management team at Penguin Computing/Scyld Software includes executives with many years of experience in Fortune 50 companies and startups in HPC, high availability, Linux, and cluster technology

Conclusions

Penguin Computing/Scyld Software's Linux cluster strategy for reducing cost/complexity is singly focused — deliver lower cost Linux clusters solutions than the competition with the best of breed Linux cluster solution in the industry. The key ingredient in this strategy is Scyld Software's Scyld Beowulf cluster operating system.

Scyld Beowulf has no competition in the HPC market when it comes to configuring, deploying, and managing Linux clusters. It is RHEL ES 3.0. Any application that runs on RHEL ES 3.0 runs on Scyld Beowulf. When Scyld Beowulf powers a Linux cluster, the compute nodes in the cluster run a minimal Linux operating system to simplify provisioning and provide an infrastructure for scalable application execution performance. The minimal Linux operating system also runs any Linux application that runs on RHEL ES 3.0.

Some Linux cluster management solutions cannot scale because of inherent bottlenecks in the underlying implementation. This is not the case with Scyld Beowulf. Scyld Beowulf permits compute nodes to be dynamically added to a cluster with little increase in subsequent cluster management overhead. With Scyld Beowulf's single point of administration, system administrators can add new nodes with little intervention. Because, Scyld Beowulf is easy to use, and because it is both a Linux distribution and a cluster management solution, customers benefit from a lower total cost of ownership (TCO) — lower system administration, training, and software-licensing costs.

Scyld Beowulf provides a simple, cost-effective way for ISVs to cluster-enable their applications. Standard Linux distributions such as RHEL and SLES (SUSE Linux Enterprise Server) are not designed to cluster-enable applications without extensions. Scyld Beowulf provides these extensions while delivering a Linux distribution that runs any applications that run on RHEL.

HPC and enterprise ISVs are cognizant of the large price/performance advantages afforded by Linux cluster technology. Heretofore, the lack of Linux cluster management solutions has been a drawback to the deployment of Linux clusters. Scyld Beowulf solves that problem and others elegantly and effectively. Today, Scyld Software is working with ISVs in HPC market segments, such as oil/gas, aerospace, bio-technology and life sciences, MCAD/CAE, financial services, EDA, and DCC, to cluster-enable their applications using Scyld Beowulf.