

Best Practices in Cluster Management

Rick Friedman

SCALI Inc.
e-mail: rfriedman@scali.com

(Rec. 7 April 2006)

Abstract: Albert Einstein defined success as ten percent inspiration and ninety percent perspiration. Although he had no inkling for the emergence of Linux Clusters for High Performance Computing (HPC) his words are ring true for designing, building and managing compute clusters. This paper focuses on best practices in implementing Cluster Management software, a paradigm for designing, building and managing compute clusters. The learnings are based on the successful experiences of environments using Scali's leading Cluster Management solution, Scali Manage.

Key words: clusters, cluster management, Linux, scaling, installation, configuration, monitoring, HPC

1. LINUX COMPUTING IS HERE A Disruptive Computing Paradigm

Businesses, research labs and universities requiring HPC have increasingly adopted Linux cluster architectures to meet their ever increasing demand for computing power, application speed and "always-on" availability. Clusters have proven to be an excellent solution for compute-intensive technical and business computing tasks requiring significant data crunching.

Due to its many benefits, Linux clusters have enjoyed widespread adoption in performance driven sectors like Financial Services, Government, Energy, Life Sciences and Advanced Manufacturing. Linux clusters are used for a wide variety of demanding applications. Examples include derivative trading, drug design, risk analysis, weapons simulation, petroleum exploration and design automation.

As penetration has moved beyond the early adoption phase, the decision to utilize Linux-based clusters has become an easy choice. It's no longer a question of whether to deploy them, but rather how to implement cluster computing to deliver high ROI, improved system performance and minimal business risk. Fundamentally, Linux computing success is about how clusters are implemented and managed and not about the choice of hardware or software tools.

2. SCALING OUT LINUX CLUSTERS Compelling Benefits but Hidden Costs and Challenges

For small, 2 to 8 node clusters, with minimal or no scheduling requirements, cluster management is relatively

straightforward. However, clusters tend to have a strong growth impetus, given their attractive economies, as well as, the usual insatiable application demand for computing power.

At a certain point, often around 16-64 nodes, cluster management begins to be problematic and expensive due to growing workload scheduling requirements, increasing cluster size and the increasing (and inherent) heterogeneity of software and hardware components. At this point, every node that is added increases management attention, magnifies complexity and business risk and multiplies both direct (overhead, labor) and indirect (lateness to market, hassle) costs.

Adding to the complexity is a myriad of clustering tools and applications, interconnects, MPIs, compilers & libraries and management tools.

The pain around cluster management is shared across industries, processing needs and user roles and represents a significant user concern.

According to a 2004 IDC user study Cluster Management rates as one of the key challenges to implementing clusters.

Large cluster sizes makes it even tougher to deploy applications quickly, fully control and monitor all resources or respond to rapid changes to user or business requirements. All of these issues conspire to reduce the organization's return on IT assets and overall business performance.

Consequently, as new clusters scale or existing Linux infrastructures grow, there is an ongoing and critical need for software management solutions that can quickly and economically scale, manage and troubleshoot clusters.

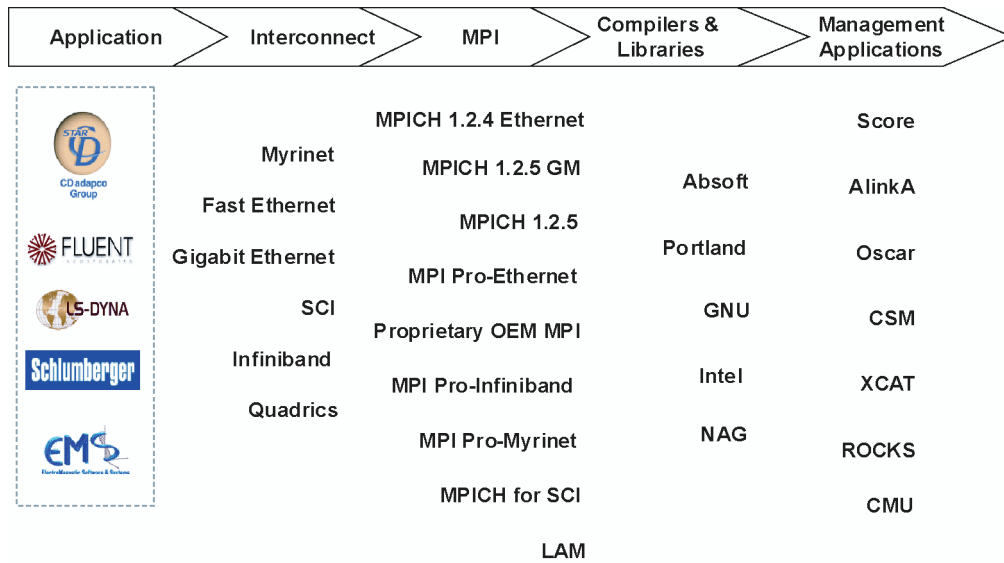


Fig. 1. Complexity of a cluster’s software environment

Priority Ranking	Share of Problems
Complexity of implementing parallel algorithms	21%
System management capability	17%
Application availability/solution. Lack of maturity of the solution	16%
Interconnect complexity/cable, cards, switches	11%
Interconnect Bandwidth/ Latency	9%
I/O performance	9%
Facility issues, space, power, cooling, noise	4%
3rd Party Software costs	1%
Other	11%
Total	100%

Fig. 2. IDC study on Cluster Implementation Issues

According to the 2004 IDC study on Cluster Management: *Clusters are enabling more and more work to be undertaken and therefore are a major driver in the growth of the HPC marketplace...there still are clearly large hurdles that clusters will need to overcome in order to expand the addressability of the technology...[including] the management of a large number of nodes and greater levels of parallelization of cluster applications as well as the conversion of applications to cluster operating environments.*

After deciding Linux clusters are the way to go, knowing how to plan and implement a cluster is equally crucial.

This effort requires a significant amount of upfront analysis and a detailed, yet realistic, implementation methodology.

3. IMPLEMENTING LINUX CLUSTERS Document, Plan, Test, Launch

The process of implementing clusters is no different than the usual project planning methodology including a review of goals, resources, staffing and features. There are, however, some guidelines that can shorten time to implementation and results. For example, Scali has identified a number of best practices for deploying, scaling and managing clusters.

In general, investing more time up front in planning and testing will bear rewards later on in terms of reduced re-work, cost and deployment time.

Baseline Information Needs

Before commencing an implementation, two important lists should be created. The first is a hardware inventory which includes specific information on minimum hardware requirements, performance tests and specifications on how to set up and configure nodes. The second is an application inventory that would include information on install applications, fix run errors and run & test performance.

Hardware Proof

To minimize risk and the time associated with “going live”, it is useful to set up a mini proof of concept to ensure that you do not run into unanticipated problems during

actual deployment. The proof of concept does not take very much time but will enable you to verify cluster software installation as well as review the performance and reliability of your network and hardware. You may want to customize the installation requirements to support hardware or other software unique to your cluster, and performance-testing tools to optimize performance. A management solution can handle replication of those requirements across additional nodes.

Application Proof

All your applications need to be tested on your mini proof of concept to ensure high performance and robustness for your stakeholders. As you probably have many different kinds of off-the-shelf, proprietary and custom applications in the environment, you could expect this process to take some time with many fits and starts. It is critical to configure, run and test the applications to highlight missing functionality or corrupted operation instead of relying only on 3rd party documentation. This process should also be repeated for all software, to ensure a complete environment.

Following configuring and testing, the next important phase is implementing the cluster. Successful cluster implementation should incorporate the following six steps:

1. **Communicate with cluster users'** critical information on policies, system information, support, and application performance.
2. **Brief the operator** on the design philosophy, system architecture, user interface as well as management tools, scripting and nomenclature.
3. **Hardware migration** prepares your system for running your application software. This may include power testing of any new hardware (i.e. master node, compute nodes, power systems), PXE boots or changes to nodal BIOS settings.
4. Once the new nodes are active, your system is ready for **application migration and testing**.
5. Upon completion of application proofing, **commissioning can begin**. At this stage, your users and operators should be made aware the cluster can now run production jobs and be aware of any missing or delayed features.

Completion of these steps helps ensure a robust cluster implementation.

5. CHOOSING CLUSTER MANAGEMENT SOFTWARE

Proven Commercial Software Offers Superior Value

Given the variety of possible tools, what should an IT Manager be looking for in a cluster management solution?

The ultimate vision of cluster management is to manage the cluster as if it was a single server, making it easy for applications to take advantage of the aggregate power of the individual nodes. However, this is not merely a simple decision of choosing a piece of cluster management software and a methodology for deploying it. Without the proper planning and due diligence, the wrong choice of cluster management software or implementation process can easily result in higher cost, longer implementations and slower time to results, negating much of the promise of cluster computing.

Effective cluster management software is predicated on a synergistic combination of manageability, enterprise-class installation and configuration flexibility.

A key consideration in choosing a management solution is the importance of growing or updating the cluster over time. For many installations, clusters are built with an expectation that they will be able to grow and support the latest technology in the future. This can mean a mix of hardware, processor, operating system, or interconnect types. Understanding the importance of this within your environment is a key factor. For most organizations with a greater sensitivity to application and business risk, and a high level of mission critical work, the less risky and less costly approach would be to deploy market-proven and supported commercial software solutions.

Powerful software bundles conveniently integrate the key elements of the hardware and system software stack together with application management tools into one convenient and tested package. Much of the value is around enhanced convenience, expert commercial support, comprehensive documentation, predictive development, 3rd party support and integrated functionality. This package would include the fundamental cluster components: Management & Monitoring Tools, Workload Management and Installation & Configuration software.

Successful commercial solutions base their value proposition around single point management control, rapid node installation, cross-platform configuration flexibility and advanced monitoring and troubleshooting.

Single Point Management Control

Clusters are complex, cumbersome organisms that defy simple deployment and control. Without centralized control, these infrastructures have a difficult time ensuring maximum up time; effective scheduling; optimal resource allocation for maximum performance and rapid cluster scale out to meet growing computing demands. With the requirements of having all the elements working tightly together, the need for a single "view" into the environment is critical. By being able to have a complete overview of the environment, it is possible to understand where the bottlenecks, issues, and opportunities are.

Rapid Installation

The ability to scale out a cluster quickly has an important impact on time to results and overall competitiveness. As IT assets and budgets become more dependent on business performance, the ability to rapidly deploy and manage clusters is seen as fundamental to market success. For example, many market-driven financial services firms facing tight market-entry windows for new products have wasted precious weeks of revenue and hundreds of thousands of integration dollars trying to get clusters built and managed without the right tools. As well, global car makers undertaking 7-24 design are often faced with sudden and unexpected increases in computing power demand. Too often, crucial design time is squandered while incremental cluster nodes and applications are added to the existing infrastructure.

Effective cluster management software preempts the aforementioned business risk, enabling dynamic enterprises to rapidly scale out clusters. This capability should enable multiple installation options, is node-neutral (x86, EM64T, Itanium, Opteron) and facilitates simple, automated installs of 3rd party applications and libraries.

Configuration Flexibility

To minimize Total Cost of Ownership, deployment time and improve reliability, the solution needs to effectively cope with the complex, heterogeneous nature of every organization's IT infrastructure. Growing clusters often contain multiple processors, operating systems, kernels, interconnects (Ethernet, Myrinet, Infiniband, SCI), as well as, a variety of hardware form factors (Blades and Servers) and industry-standard Linux distributions. In many complex environments this heterogeneity is a major barrier to application deployment (slower time to value),

cluster scale outs (inability to cope with unexpected surges in demand), and timely troubleshooting (high cost of failure). A cluster management solution must be able to efficiently support a wide variety of different hardware and software vendors, both today and in the future.

Detailed Monitoring

Another key component of implementing cluster management is being able to effectively monitor resources for system optimization, planning and troubleshooting once the cluster is up and going. Being able to quickly and simply monitor system performance enables the user to quickly preempt and troubleshoot resource limitations before their impact is magnified. Additionally, being able to track the on going performance of an environment enables improvement in the performance and reliability of the applications. A strong monitoring environment should allow for both proactive monitoring (being able to open a window and review the system) and reactive monitoring, where the system generates notifications of any issues that come up.

6. SUMMARY

Successful cluster implementations are usually the result of careful planning and anticipation of the complexity of deployment and on-going management issues. Following a strong planning process that understands the hardware requirements, the user expectations, and the potential future uses for a cluster can impact success considerably. Leveraging software tools, particularly for clusters of 16 nodes or greater significantly increases the likelihood of consistency, reliability and understanding of the total cluster environment.



RICK FRIEDMAN is the Vice President of Product Marketing at Scali, a leader in Linux Cluster Management solutions. While he currently carries the title of Marketing, he is in fact really an engineer with a BSEE from University of Penn, and experience in semiconductor design, networking and servers. Additionally, he spent a number of years in the high performance applications areas of electronic design automation and pharmaceutical clinical trials. His combination of both hardware and software, as a vendor and as a user, provides a unique perspective.