



902 Broadway, 7th Floor

New York, NY 10010

www.TheEdison.com

@EdisonGroupInc

212.367.7400

Comparison of Open Source Software vs. IBM Spectrum LSF Suite for Enterprise

**Key considerations when evaluating HPC
workload management solutions**



Printed in the United States of America

Copyright 2018 Edison Group, Inc. New York.

Edison Group offers no warranty either expressed or implied on the information contained herein and shall be held harmless for errors resulting from its use.

This report was commissioned by IBM.

The information contained in this document is based on IBM provided materials and independent research and was aggregated and validated for Edison Group, Inc. by the Edison Group Analyst team.

All products are trademarks of their respective owners.

First Publication: May 2018

Produced by: Harold Kreitzman, VP, Strategic Advisory Services

Table of Contents

| | |
|---------------------------------------------------------------------------------------------------------------|-----------|
| Executive Summary | 1 |
| IBM Spectrum LSF Suite | 3 |
| Designed with Users and Operations in Mind | 3 |
| Cost Effective | 4 |
| High Performance | 5 |
| Total Cost of Ownership | 6 |
| TCO Analysis Methodology | 6 |
| Open Source | 7 |
| Software License Management | 7 |
| Methodology by Component | 8 |
| Summarized Results | 10 |
| Cost Component Breakdown | 11 |
| Using IBM Spectrum LSF Suite for Enterprise to Address HPC Application License Related Challenges..... | 14 |
| Summarized Analysis and Results..... | 16 |
| Appendix | 18 |

Executive Summary

Today, the onslaught of data and the corresponding need to process it quickly has created significant demands on traditional IT infrastructure. To meet that need, high-performance computing (HPC) has become an integral part of the workload mix. To complicate matters, IT infrastructure is in transition as companies move towards hybrid processing models which include traditional bare metal, virtual machines, cloud and containers.

This complex infrastructure cannot function efficiently using simplistic workload scheduling tools. Each type of processing model requires a slightly different approach to achieve the highest level of efficiency, therefore creating potentially conflicting operational focus which a simple workload scheduling tool cannot satisfy.

A greater focus is needed on workload management tools to simplify operations, balance resources, manage job priorities, and performance.

Simplified operations are achieved through workflow automation, easy to use GUI, intuitive reporting dashboards, automated alerts, and customization.

Another major challenge with HPC is the high cost of software licensing. In some cases, the cost of the software can be twice as much as the cost of the HPC infrastructure. A platform that makes prudent use of software licenses through intelligent scheduling can help to significantly reduce costs.

The main premise of this paper is to illustrate how a software-defined infrastructure (SDI) platform that goes beyond traditional scheduling tools will reduce the overall cost of HPC. In a traditional sense, most customers attempt to procure as much hardware as their budget will allow. To make up for the lack of budget, many customers migrate towards open-source solutions as the general perception is that these are less costly. The reality is that in many cases, the cost of support more than offsets the software license costs.

In this document, we will introduce the concept of “infrastructure efficiency”; with a platform that not only manages complex workloads but does this in such a way that is easy to operate and monitor thus maximizing resource pools and reducing cost.

We look at an example of the total cost of ownership (TCO) of the IBM Spectrum LSF Suite for Enterprise compared to a pure open-source alternative. While there is a lot of controversy associated with the use of open-source, this will not be covered in this paper. Edison’s view is



that while unmanaged use of open-source is high-risk, open-source code integrated with proprietary offerings and managed by a legitimate technology service organization is a good business practice for both provider and end-user as it provides a way to lower cost and maintain low risk levels.

IBM Spectrum LSF Suite

IBM Spectrum LSF Suite addresses all of the challenges listed above. It is much more than a workload scheduler. It is a complete solution for HPC resource and workload management, designed for enterprise-grade productivity.

It is cost effective, delivering excellent business value by providing low TCO and high return on investment (ROI.)

It has broad market availability. IBM Spectrum LSF Suite is available in three editions with simplified packaging to support small environments up to large enterprise clusters.

Designed with Users and Operations in Mind

- **Centralized management and reporting portal** – a must-have for efficient operational support, and effective cost management.
- **Easy to integrate** – application workloads are typically supported by a mixture of tools and utilities that need to integrate seamlessly, *producing one consistent experience*. IBM Spectrum LSF easily integrates with other infrastructure components.
- **Operational analytics** – critical operational function. This enables IT to get-ahead of operational issues as they occur. Being pro-active eliminates end-user concerns, and complaints.
- **Workflow automation**¹ – in today's complex operational environments, automation is critical as it eliminates time-intensive manual effort, reduces errors through repeatable processes, and reduces need for highly-technical intervention.
- **Intelligent flexing**² – seamlessly burst to the cloud with automated, dynamic-hybrid cloud capabilities
- **Ease of use** - Today, IT executives are extremely concerned with user-experience, namely ease of use. To that end, IBM Spectrum LSF Suite for Enterprise is designed to simplify HPC

¹ This capability is available in: IBM Spectrum LSF Suite for HPC, IBM Spectrum LSF Suite for Enterprise ONLY. Or via the Spectrum LSF Process Manager add-on to IBM Spectrum LSF core

² This capability is available in: IBM Spectrum LSF Suite for HPC, IBM Spectrum LSF Suite for Enterprise only.

operations so a deep technical knowledge is not required. This enables operational personnel to perform workload scheduling tasks that were only available to more highly trained administrative personnel.

Cost Effective

- **Easy to install** – less time and effort needed to get the cluster configured and running either via Ansible, or via bare metal.
- **Reduces training time** – allows users to be more productive in less time.
- **Optimizes software license usage**³ – more effective management of software license use leads to significant cost savings, in many cases more than the infrastructure itself.
- **Simplified packaging** – less effort needed since it is easier to start with and easier to scale with Suites from a few nodes to large enterprise clusters.
- **Streamlined administration** - enables operations personnel to manage workloads. As stated previously, ease of use and automation leads to reduced need for highly-trained personnel, thus reducing operational costs.

³ IBM Spectrum LSF Suite for Enterprise – ONLY. Not standard in: IBM Spectrum LSF Suite for Workgroups

High Performance

- **Performance improves with each release** – Edison has been closely involved with the evaluation of IBM Spectrum LSF over the past few years⁴, “IBM Spectrum LSF significantly improves its previously documented throughput benchmark results by over 300%. In addition, scheduler cycle timing decreased by a similar 300% as compared to previous results”.
- **Architecture maximizes performance** – a lot of thought went into the design of IBM Spectrum LSF Suite for Enterprise, for example:
 - Data is moved closer to the applications that need it, thereby improving throughput performance.
 - Incorporates parallel processing taking advantage of a proven performance enhancing methodology.
 - Optimization of IO patterns to minimize impact of file system latency
 - Optimization of inter-processor communications supports growth in cluster scalability and job throughput performance.

⁴ IBM Platform LSF Meets Evolving High Performance Computing Demands, <http://bit.ly/2DNmRuO>

Total Cost of Ownership

The TCO comparison of IBM Spectrum LSF Suite for Enterprise vs. an open source competitor demonstrates that IBM Spectrum LSF Suite for Enterprise TCO is lower than the open source alternative. *Zero* software licensing cost does not necessarily mean that you will save money over other alternatives. Over the past ten years there has been a rush to move towards open-source software as the perception was that it would save money.

IBM Spectrum LSF Suite for Enterprise is built for large scale installations with complex requirements and therefore at the enterprise level has additional TCO advantages over open source.

Lastly, performance and the user experience are better with IBM Spectrum LSF Suite for Enterprise. This plus the lower TCO results in greater overall business value.

TCO Analysis Methodology

Total cost of ownership models review the costs associated with the acquisition, installation, licensing, maintenance, and bricks-and-mortar infrastructure needs of a HPC workload management solution. It is Edison's experience that this level of detail is required to obtain a realistic perspective of the total cost impact of a selection decision.

Edison strongly suggests that the actual realized costs are tracked throughout the life of the project. It is our experience that most organizations make decisions based on initial information and then do not follow-through to see if their decisions were the correct ones.

Within this study average discounts are applied. The cost components reviewed include:

- Cluster hardware and maintenance
- Workload management software
- Installation fees
- Power costs
- Unanticipated costs
- IT administrative personnel
- End-user productivity opportunity cost

Total cost of ownership is analyzed using a four-year time horizon. Capital costs, such as hardware, are depreciated, straight-line, over a four-year period. Operating costs include all other spending to sustain operations, including all maintenance, power, space, and staffing costs and are included in the TCO analysis. Some costs, such as installation, are one-time only.

Before reviewing the details of the TCO analysis, it is important to put certain elements into perspective. For the purposes of illustrative TCO comparison, Edison compared IBM Spectrum LSF Suite for Enterprise against an open source alternative. Open source was chosen vs. a proprietary alternative since there is a persistent market perception that open source is always a less expensive alternative to any solution that contains proprietary software.

Open Source

Open source scheduling software is free; hardware, supporting utilities, tools and personnel resource requirements are not. In addition, the lack of a transparently integrated suite of supporting tools, utilities and function/feature enhancements place a greater reliance on highly technical resources to bridge whatever gaps exist or occur based on what is required.

The biggest challenge with open source software is that the technician needs to become intimately familiar with the open-source code itself, the interfaces, release dates (if any), open issues, fixes, associated tools and utilities to provide a complete function/feature set, integration with those tools and utilities, and to be responsible for maintaining and integrating the code. No plug-and-play here. It's not simply configure-and-go. Technicians have to dedicate a significant amount of time keeping on top of any open source environment.

All of this time and effort translates into hidden or opportunity costs that need to be included in the analysis, i.e., increased time and costs spent on system administration, troubleshooting, training, software certification, software maintenance, quality assurance and testing.

The other significant, frequently overlooked aspect of open source software is that each component is developed as a stand-alone component with a non-standard, frequently changing APIs. These APIs are then used to integrate with other components. The reality is that high performance and stability are not guaranteed as the components are not necessarily developed in conjunction with each other.

Software License Management

Software licensing costs are directly affected by inefficiently performing application workload schedules. Edison has found that the licensing of these software applications can be 2X more

than the HPC infrastructure itself. IBM Spectrum LSF Suite for Enterprise excels at software license management, therefore making it possible to perform the same amount of work as an open source alternative with a smaller number of software licenses. IBM Spectrum LSF Suite for Enterprise provides the tools to do analysis of license usage allowing customers to make more educated decisions on software license purchases.

As an example, ISV application licenses for CAE and Virtual Product development are quite expensive some costing 10's of thousands of dollars per server. This where IBM Spectrum LSF Suite for Enterprise can help not only manage hardware efficiency but application software utilization.

Methodology by Component

The foundations for these analyses were based on a typically-sized enterprise HPC cluster environment. For the purposes of this analysis, a typically-sized enterprise HPC cluster of 500 nodes was used for the open source cluster and 400 nodes (a 20% reduction) was used for the IBM Spectrum LSF cluster. This ratio is based on comparative performance testing⁵. In an earlier Edison study, IBM Spectrum LSF was shown to be significantly more efficient at scheduling so less hardware is required to get the same amount of work done.

All information used to drive this analysis was collected from several customer interviews.

- **Hardware** – is broken down into number of nodes per cluster and hardware maintenance, estimated around 6% of hardware cost per year. Hardware is refreshed every four years, so hardware refresh is not a factor in these analyses.
- **Workload Management Software** – depending on alternative, namely open source or IBM Spectrum LSF.
- **Installation Fees** – cost associated with installation, training, configuration, etc.
- **Power** – is based on power requirements for hardware and power required for facilities, namely HVAC, lighting, etc.
- **Unanticipated Costs** – missing components, consultant time, etc.
- **IT Administrative Personnel** – cluster set-up, user support, application maintenance, support incidents, etc.

⁵ IBM Platform LSF Meets Evolving High Performance Computing Demands, November 2015, Edison Group, http://www.theedison.com/pdf/2015_Samples_IBM_Platform_LSF_TB.pdf

- **Business Opportunity Cost** – costs associated with losses due to preventable job failures, training, operational processing bottlenecks, etc.

Summarized Results

The following table summarizes the results by component. A detailed review follows.

| Cost Components | IBM Spectrum LSF Suite for Enterprise | Open Source Cluster | Difference \$ |
|----------------------------------------|---------------------------------------|---------------------|----------------------|
| Hardware | \$5,360,000 | \$6,700,000 | \$(1,340,000) |
| Workload Management Software | 900,000 | - | 900,000 |
| Initial Installation, Configuration | 20,127 | 16,884 | 3,243 |
| Power | 1,660,279 | 2,075,349 | (415,070) |
| Unanticipated Costs | 20,000 | 80,400 | (60,400) |
| IT Administrative Personnel | 119,404 | 286,353 | (166,949) |
| End-user Productivity Opportunity Cost | 6,137 | 2,545,464 | (2,539,327) |
| Totals | \$8,085,947 | \$11,704,450 | \$(3,618,503) |

Table 1 – Summarized TCO Analysis

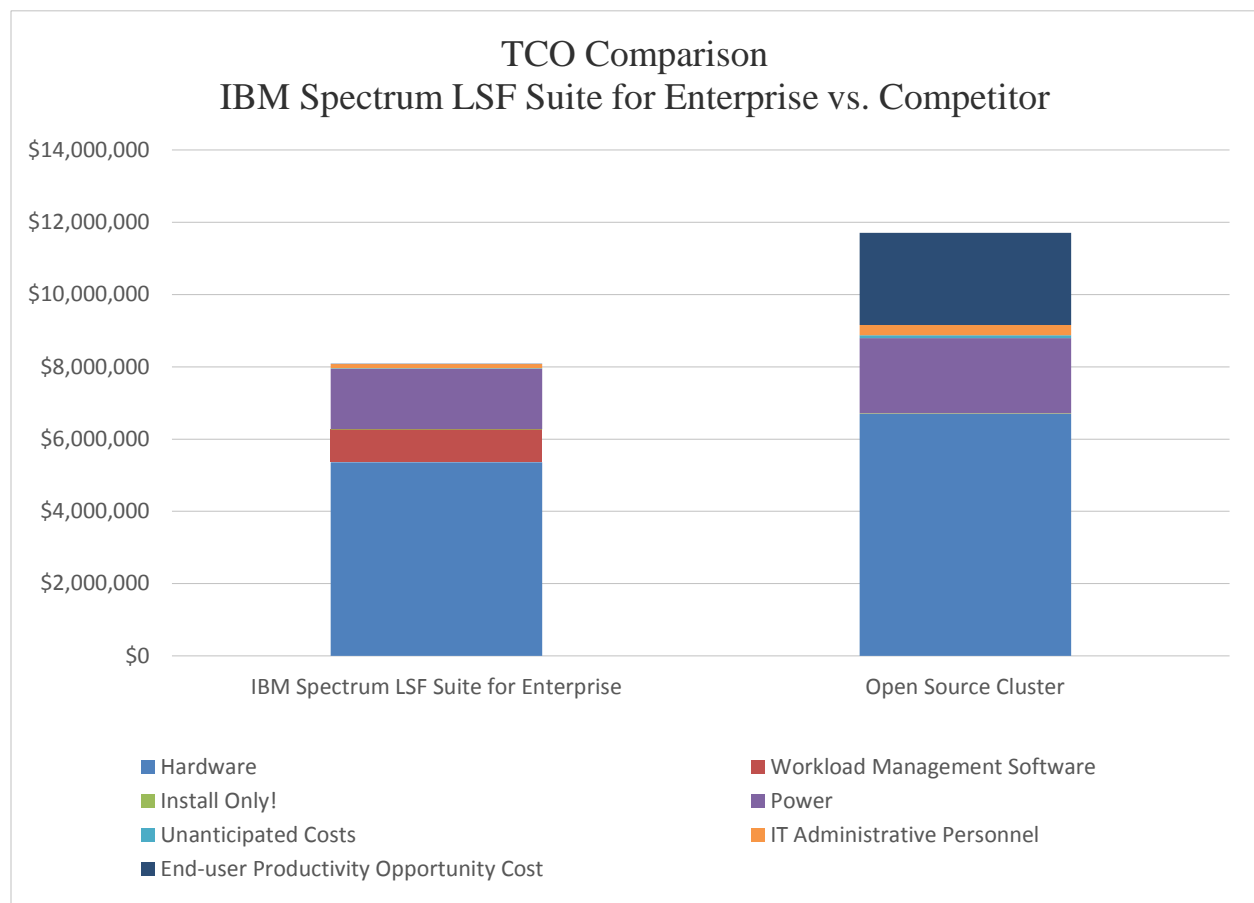


Chart 1 – Summarized TCO Results

As presented above, the net savings of IBM Spectrum LSF Suite for Enterprise over open source alternative is \$3.6 million or 31% over a four-year period. The most significant contributors are

reduced hardware costs, end-user productivity and power costs. The major savings using open source is due to a \$0 cost for software licenses.

Beyond the actual TCO over a four-year period is three specifics that need to be highlighted:

- **Smaller infrastructure footprint** – less hardware is needed. This translates to lower hardware and power costs.
- **Better performance** – as a result of a fully integrated SDI and a performance-aware architecture IBM Spectrum LSF is more efficient which translates to lower infrastructure costs.
- **Less reliance on highly-trained personnel** – IBM Spectrum LSF Suite for Enterprise is fully integrated, has easy to use GUI's and does not rely on highly trained technicians to install, configure and operate.

Cost Component Breakdown

Hardware

As per the summary table:

| Cost Components | IBM Spectrum LSF Suite for Enterprise | | Open Source Cluster | |
|-----------------|---------------------------------------|---------------|---------------------|---------------|
| | Cost | / No of Nodes | Cost | / No of Nodes |
| Hardware | \$ 5,360,000 | 400 | \$ 6,700,000 | 500 |

The difference between the hardware costs is based on the number of nodes required to support the HPC infrastructure. While the baseline open source configuration requires 500 nodes, the more efficient IBM Spectrum LSF Suite for Enterprise processing only requires 400 nodes. This results in a 20% savings.

Workload Management Software

Open Source software is free while there are costs associated with the acquisition of IBM Spectrum LSF Suite for Enterprise. The licensing cost of IBM Spectrum LSF Suite for Enterprise is based on the number of cores represented by the hardware.

| Workload Management Software | Four Year Discounted Price |
|---------------------------------------|----------------------------|
| IBM Spectrum LSF Suite for Enterprise | \$900,000 |
| Totals | \$900,000 |

Initial Installation, Configuration

This represents any service and/or personnel costs required to initially install, configure and train administrative personnel. Ongoing maintenance, customer support and incident support is covered under IT Administrative costs. IBM Spectrum LSF Suite for Enterprise includes installation services provided by IBM Lab Services as well as effort to install and configure clusters. Open Source only includes effort to install and configure clusters.

As per the summary table:

| Cost Components | IBM Spectrum LSF Suite for Enterprise | Open Source Cluster | Difference \$ |
|---------------------------------------|---------------------------------------|---------------------|---------------|
| Installation (Using IBM Lab Services) | \$ 20,127 | \$16,884 | \$3,243 |

Power

As per the summary table:

| Cost Components | IBM Spectrum LSF Suite for Enterprise Cost / No of Nodes | Open Source Cluster Cost / No of Nodes |
|-----------------|----------------------------------------------------------|----------------------------------------|
| Power | \$1,660,279 / 400 | \$2,075,349 / 500 |

Power is needed for hardware and supporting facility requirements, HVAC, UPS, and lighting. The more hardware (nodes) required, the more power needed. As presented above, the reduced hardware required to support an Open Source equivalent (400 vs 500 nodes) saves costs associated with powering and supporting hardware infrastructure including the facility itself. A 20% hardware savings results in a 20% power savings and associated costs.

Unanticipated Costs

As per the summary table:

| Cost Components | IBM Spectrum LSF Suite for Enterprise | Open Source Cluster | Difference \$ |
|---------------------|---------------------------------------|---------------------|---------------|
| Unanticipated Costs | \$20,000 | \$80,400 | (\$60,400) |

Unanticipated costs are estimated unbudgeted costs associated with open source solution, i.e., missing components, consultant time, etc.

IT Administrative Personnel

IT administrative costs are a direct result of how much time is needed to support users, maintain applications and support incidents. Specific assumptions and associated costs are as follows:

| Cost Components | IBM Spectrum LSF Suite for Enterprise | | Open Source Cluster | |
|-------------------------|---------------------------------------|--------------|---------------------|--------------|
| | Cost | Hours | Cost | Hours |
| User support | 81,043 | 960 | 202,608 | 2,400 |
| Application Maintenance | 27,014 | 320 | 67,536 | 800 |
| Incident Support | 11,346 | 136 | 16,208 | 192 |
| Totals | \$ 119,403 | 1,416 | \$286,352 | 3,292 |

Note, a lot more effort needs to be expended when implementing and supporting an open source cluster. This translates to more time and money being spent.

End-user Productivity Opportunity Cost

End user productivity opportunity cost is based on three items:

- Time lost due to preventable job failures
- Avoidable training time learning interfaces
- Loss of productivity waiting for administrators

| Cost Components | IBM Spectrum LSF Suite for Enterprise | | Open Source Cluster | |
|-------------------------------------------------------|---------------------------------------|-----------|---------------------|---------------|
| | Cost | Hours | Cost | Hours |
| Preventable Job Failures (approximately \$111/hr.) | \$ 2,224 | 20 | \$ 1,334,640 | 12,000 |
| User support (approximately \$111/hr.) | 2,224 | 20 | 400,392 | 3,600 |
| Application Maintenance (approximately \$84/hr.) | 1,688 | 20 | 810,432 | 9,600 |
| Totals | \$ 6,136 | 60 | \$2,545,464 | 25,200 |

Based on customer input, it takes a lot more effort to support open source. The net impact over a four-year period is estimated to be 18,000 hours for open source vs. 60 hours for IBM Spectrum LSF Suite for Enterprise.

Using IBM Spectrum LSF Suite for Enterprise to Address HPC Application License Related Challenges.

For customers in industrial markets such as automotive, aerospace and the semiconductor industry, the most expensive spend in an HPC enterprise deployment is not hardware or a workload management system. In fact, the cost of the IBM Spectrum LSF Suite for Enterprise is typically only 1-3% of the cost of the entire environment which includes, computer hardware, storage, electricity and administrative headcount.

But the cost to the business of not efficiently managing licenses for additional applications, such as Computer Aided Engineering (CAE) applications, by selecting the wrong workload scheduler is always overlooked in hardware procurement decisions. This can be an omission that can cost millions in additional application license software spending.

IBM Spectrum LSF Suite for Enterprise includes powerful software license optimization capabilities. This improves HPC efficiency by 10% to 20% and can reduce the overall overspending for application licenses by increasing utilization and better tracking of the use of these very expensive assets. Some ways IBM Spectrum LSF Enterprise Suite can prevent overspending are:

- The ability to compare forecasted license capacity to utilization to refine future forecasting for projects
- An intelligent way to arbitrate licenses in a cross-cluster manner, so that jobs are not starved or licenses are not wasted
- Not having to understand when the application in the job will check out the actual license and not suffer the loss of productivity due to “duration”
- A way to properly manage users to maintain a predictable license environment and prevent abuse
- A way to simplify submission while maintaining visibility into license usage/complex license requirements
- A simple methodology to prevent double counting (wasted licenses) of licenses for jobs that checks out and ins licenses without having to define duration and predict job’s runtime
- A way to prevent erroneous job dispatch, a more robust way to handle the short comings of application provided license management systems

- Ability to allow engineering managers to look who is taking up licenses.

| Cost Components | IBM Spectrum LSF Suite for Enterprise | Open Source Cluster | Difference \$ |
|----------------------------------------|----------------------------------------------|----------------------------|----------------------|
| CAE Application Software Licenses | \$19,200,000 | \$24,000,000 | (\$4,800,000) |
| Hardware | 5,400,000 | 6,700,000 | (\$1,300,000) |
| Workload Management Software | 984,000 | 0 | \$984,000 |
| Initial Installation, Configuration | 20,127 | 16,884 | \$3,243 |
| Power | 1,660,279 | 2,075,349 | (\$415,070) |
| Unanticipated Costs | 20,000 | 80,400 | (\$60,400) |
| IT Administrative Personnel | 119,404 | 286,353 | (\$166,949) |
| End-user Productivity Opportunity Cost | 6,137 | 2,545,464 | (\$2,539,327) |
| Totals | \$27,409,947 | \$35,704,450 | (\$8,294,503) |

Summarized Analysis and Results

The increased reliance on real-time business data has made high-performance computing an integral part of the workload mix. Traditional workload scheduling software cannot efficiently handle this level of workload complexity.

While there are some alternatives that support HPC-based workload scheduling, a review of functions, features, ease-of-use, and cost narrows the field of choices.

In this paper Edison compared the total cost of ownership for IBM Spectrum LSF Suite for Enterprise against an open source alternative. Contributing to the costs were a number of components including:

- Cluster hardware and maintenance
- HPC workload and resource management software
- Installation fees
- Power costs
- Unanticipated costs
- IT administrative personnel
- End-user productivity opportunity cost.

Inherently, the differences between the two alternatives directly affected specific cost components, which in turn, affected the end result, TCO. Specifically:

- Hardware cost is less for IBM Spectrum LSF Suite for Enterprise since it requires less hardware (400 nodes) to perform the same function as open source (500 nodes), i.e., it is more efficient with better performance.
- Software cost is free for open source, not free for IBM Spectrum LSF Suite for Enterprise.
- Installation, training, configuration costs are not applicable to IBM Spectrum LSF Suite for Enterprise as they are included in the software cost. They are applicable for open source.
- Power costs are greater for open source as it requires more hardware to perform the same tasks.
- Unanticipated costs are costs associated with software integration, missing components, consultant time associated with open source.

- IT administrative personnel was significantly greater for the open source alternative than IBM Spectrum LSF Suite for Enterprise. This was due to the additional effort required by highly-trained personnel to support open source.
- End-user productivity opportunity cost is what effort it would take to support either environment. Much more open source support was needed on a ratio of 420:1.

The bottom line is that it has been determined that using IBM Spectrum LSF Suite for Enterprise instead of an open source alternative results in a 31% savings over a four-year period.

The above TCO analysis considered a large-scale HPC environment containing 400-500 nodes. Organizations making use of HPC may start with small environments, but they typically grow in size over time. So the above considerations are also important for people starting out in HPC with an eye to future growth.

Edison recommends that IT executives seriously consider IBM Spectrum LSF Suite for Enterprise over open source alternatives because of potential cost savings due to smaller infrastructure footprint, better performance and less reliance on highly-trained personnel.

Appendix

IBM Spectrum LSF Suite for Enterprise provides a complete solution for systems and workload management designed to maximize productivity of administrator and users through a fully integrated UI, while helping to control costs through advanced workload scheduling. IBM Spectrum LSF Suite is available in 3 editions with increasing levels of capabilities for small organizations through to large enterprise clients and provides the following benefits:

- Enhanced user and administrator productivity through simplification
- Integrated, powerful reporting dashboard
- Faster time to cluster readiness - streamlined installation process

IBM Spectrum LSF Suites: Simplified packaging to accelerate time to results

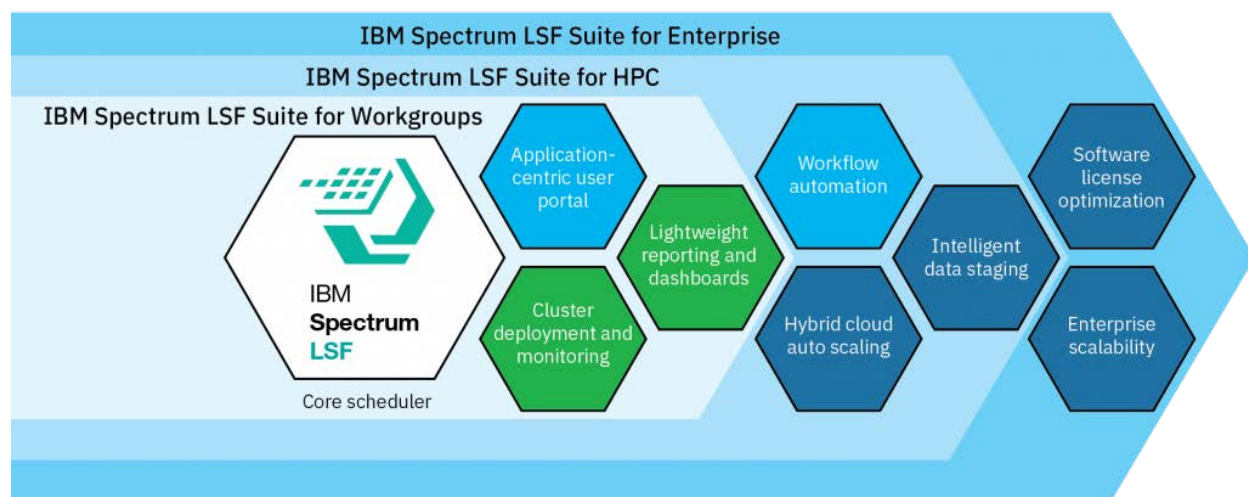


Figure 1: Increasing capabilities of the IBM Spectrum LSF Suite for Enterprise by edition

In addition to the LSF Suite for Enterprise editions, IBM Spectrum LSF family products are available a la carte:

IBM Spectrum LSF Analytics

Advanced analysis of technical computing workloads

IBM Spectrum LSF

Complete workload management for demanding, mission-critical computing environments

IBM Spectrum LSF Application Center

Application-centric workload submission and management

IBM Spectrum LSF Data Manager

Optimize cluster throughput with cached data transfers

IBM Spectrum LSF Explorer

A powerful, lightweight reporting solution for IBM Spectrum LSF clusters

IBM Spectrum LSF License Scheduler

Enables license sharing between global project teams ensuring license availability is prioritized by workload/user/project.

IBM Spectrum LSF Process Manager

Run and manage business-critical workflows on a shared, distributed compute environment

IBM Spectrum LSF RTM

Advanced reporting, tracking and monitoring for IBM Spectrum LSF environments

IBM Spectrum LSF Session Scheduler

High-throughput computing enabling more complex simulations to be run in less time