

IBM Platform Computing Solutions for High Performance and Technical Computing Workloads

Dino Quintero

Daniel de Souza Casali

Marcelo Correia Lima

Istvan Gabor Szabo

Maciej Olejniczak

Tiago Rodrigues de Mello

Nilton Carlos dos Santos

 Cloud

Power Systems





International Technical Support Organization

**IBM Platform Computing Solutions for High
Performance and Technical Computing Workloads**

June 2015

Note: Before using this information and the product it supports, read the information in “Notices” on page vii.

First Edition (June 2015)

This edition applies to IBM Platform Symphony V7.1, IBM Platform LSF V9.1.3, IBM Spectrum Scale (formerly GPFS) V4.1, IBM Platform Application Center V9.1.3, IBM Platform HPC V4.2, IBM Platform Cluster Manager - Advanced Edition V4.2, and IBM Platform MPI V8.3 for Linux.

© Copyright International Business Machines Corporation 2015. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Notices	vii
Trademarks	viii
IBM Redbooks promotions	ix
Preface	xi
Authors.....	xi
Now you can become a published author, too!	xii
Comments welcome.....	xiii
Stay connected to IBM Redbooks	xiii
Chapter 1. Introduction to IBM Platform Computing	1
1.1 IBM Platform Computing solutions purpose	2
1.2 Cluster, grids, and clouds	3
1.3 IBM Platform Computing Services	4
1.3.1 IBM Platform High Performance Cluster.....	5
1.3.2 IBM Load Sharing Facility.....	6
1.3.3 IBM Platform Symphony	6
1.4 Benefits and industries	6
Chapter 2. Technical computing software portfolio	9
2.1 Data-centric view for technical computing.....	10
2.2 Storage management	11
2.3 Workload management.....	11
2.4 Cluster management.....	12
2.5 Virtual resource management.....	13
2.6 IBM Platform Computing Cloud Services	14
Chapter 3. Big data, analytics, and risk calculation software portfolio	15
3.1 What is big data	16
3.2 Big data analytics	18
3.2.1 Big data analytics challenge	18
3.2.2 Big data analytics solution	19
3.2.3 IBM Big Data and analytics areas with solutions	20
3.2.4 IBM Big Data analytics advantage	24
3.3 Why use an IBM Risk Analytics solutions	24
3.3.1 IBM Algorithmics software.....	25
3.3.2 IBM OpenPages software	25
3.4 Scenario for minimizing risk and building a better model	25
3.4.1 Algo Market Risk.....	26
3.4.2 IBM SPSS Statistics: Monte Carlo simulation.....	26
3.4.3 Scenario	26
Chapter 4. IBM Spectrum Scale (formerly GPFS)	29
4.1 IBM Spectrum Scale overview	30
4.2 Spectrum Scale for technical computing.....	31
4.2.1 Argonne Leadership Computing Facility.....	32
4.2.2 Jülich Supercomputing Centre	32
4.2.3 IBM Elastic Storage Server.....	32

4.3 Spectrum Scale for big data	33
4.4 Installing IBM Spectrum Scale	33
4.4.1 Introducing IBM Spectrum Scale.....	33
4.4.2 The strengths of Spectrum Scale	33
4.4.3 Preparing the environment on Linux nodes	34
4.4.4 Spectrum Scale open source portability layer.....	40
4.4.5 Configuring the cluster	42
Chapter 5. IBM Platform Load Sharing Facility product family.....	47
5.1 Overview.....	48
5.2 Platform LSF add-ons and capabilities	50
5.2.1 Using IBM Platform MapReduce Accelerator for Platform LSF	51
5.2.2 Using IBM Platform Data Manager for LSF.....	56
5.3 Using IBM Platform MultiCluster.....	59
5.4 IBM Platform Application Center.....	60
Chapter 6. IBM Platform Symphony V7.1 with Application Service Controller.....	63
6.1 Introduction to IBM Platform Symphony V7.1	64
6.2 IBM Platform Symphony: An overview	65
6.3 IBM Symphony for multitenant designs.....	69
6.3.1 Challenges and advantages	69
6.3.2 Multitenant designs	70
6.3.3 Requirements gathering	71
6.3.4 Building a multitenant big data infrastructure	71
6.3.5 Summary.....	75
6.4 Product editions	75
6.4.1 IBM Platform Symphony Developer Edition	76
6.4.2 IBM Platform Symphony Advanced Edition	78
6.5 Optional applications to extend Platform Symphony capabilities	80
6.6 Overview of IBM Platform Application Service Controller.....	80
6.6.1 Application framework integrations.....	82
6.6.2 Basic concepts	82
6.6.3 Key prerequisites	84
6.6.4 IBM Platform Application Service Controller: Application templates.....	85
6.7 IBM Platform Symphony application implementation	85
6.7.1 Planning for Platform Symphony.....	86
6.7.2 Accessing the Platform Symphony Management Console.....	89
6.7.3 Configuring a cluster for multitenancy.....	90
6.7.4 Adding an application / tenant.....	95
6.7.5 Configuring application properties.....	97
6.7.6 Associating applications with consumers	98
6.7.7 Summary.....	99
6.8 Overview of Apache Spark as part of the IBM Platform Symphony solution.....	99
6.8.1 Hadoop implementations in IBM technology.....	100
6.8.2 Advantages of Spark technology	100
6.8.3 Spark deployments	101
6.8.4 Spark infrastructure.....	102
6.8.5 Spark deployment templates	102
6.9 ASC as the attachment for cloud-native framework: Apache Cassandra.....	102
6.10 Summary.....	104
Chapter 7. IBM Platform High Performance Computing	105
7.1 Overview	106
7.2 IBM Platform HPC advantages	107

7.3 Implementation	107
7.3.1 Installing a management node	108
7.3.2 Installing a compute node	113
Chapter 8. IBM Platform Cluster Manager	117
8.1 Platform Cluster Manager - Standard Edition V4.2	118
8.1.1 Platform Cluster Manager - Standard Edition support for POWER8 nodes	118
8.1.2 LDAP integration.....	118
8.1.3 Tagging nodes	120
8.2 Platform Cluster Manager - Advanced Edition V4.2	120
8.2.1 Multitenant environment	120
Chapter 9. IBM Cloud Manager	125
9.1 IBM Software Defined Environment	126
9.2 The software-defined everything vision.....	126
9.3 OpenStack	127
9.4 Introducing IBM Cloud Manager	127
9.5 IBM Cloud Manager value points	128
Chapter 10. IBM Platform Computing Cloud Services	129
10.1 IBM Platform Computing Cloud Services: Purpose and benefits	130
10.2 Platform Computing Cloud Services architecture	131
10.3 IBM Spectrum Scale high-performance services	132
10.4 IBM Platform Symphony services	132
10.5 IBM High Performance Services for Hadoop	132
10.6 IBM Platform LSF Services	133
10.7 Hybrid Platform LSF on-premises with a cloud service scenario	134
10.7.1 Upgrading IBM Platform HPC to enable the multicluster function.....	134
10.7.2 Tasks to install IBM Platform LSF in the cloud	139
10.7.3 Configuring the multicluster feature	139
10.7.4 Configuring job forwarding	141
10.7.5 Testing your configuration.....	143
10.7.6 Hybrid cloud is ready	144
10.8 Data management on hybrid clouds	145
10.8.1 IBM Platform Data Manager for LSF.....	145
10.8.2 IBM Spectrum Scale Active File Management	145
Appendix A. IBM Platform Computing Message Passing Interface	147
IBM Platform Computing Message Passing Interface	148
IBM Platform Computing Message Passing Interface implementation	148
Appendix B. LDAP server configuration and management	151
OpenLDAP installation	152
LDAP user account management	153
Related publications	155
IBM Redbooks	155
Other publications	155
Online resources	155
Help from IBM	156

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at <http://www.ibm.com/legal/copytrade.shtml>

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX®	IBM®	POWER7®
Algo®	IBM Elastic Storage™	POWER8™
Algo Market®	IBM Flex System®	PowerLinux™
Algo Risk®	IBM Spectrum™	PureSystems®
Algorithmics®	IBM Watson™	QRadar®
BigInsights™	InfoSphere®	Redbooks®
Bluemix™	Insight™	Redbooks (logo)  ®
Cognos®	LSF®	SPSS®
DataStage®	OpenPages®	Symphony®
DB2®	Passport Advantage®	Tealeaf®
developerWorks®	POWER®	WebSphere®
Global Business Services®	Power Systems™	z Systems™
GPFS™	POWER6®	

The following terms are trademarks of other companies:

SoftLayer, and SoftLayer device are trademarks or registered trademarks of SoftLayer, Inc., an IBM Company.

Intel, Intel Xeon, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Other company, product, or service names may be trademarks or service marks of others.

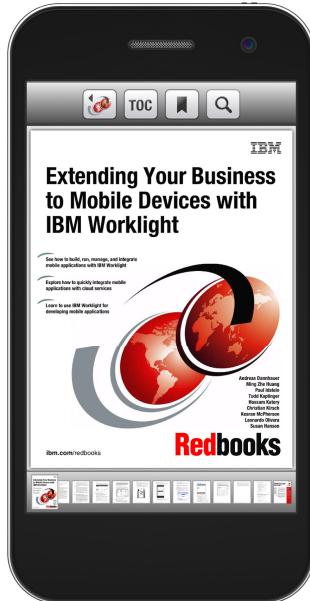
Find & read thousands of IBM Redbooks publications

- ▶ Search, bookmark, save & organize favorites
- ▶ Get up-to-the-minute Redbooks news & announcements
- ▶ Link to the latest Redbooks blogs & videos

Get the latest version of the **Redbooks Mobile App**



iOS
Download Now



Promote your business in an IBM Redbooks publication

Place a Sponsorship Promotion in an IBM® Redbooks® publication, featuring your business or solution with a link to your web site.

Qualified IBM Business Partners may place a full page promotion in the most popular Redbooks publications. Imagine the power of being seen by users who download millions of Redbooks publications each year!



ibm.com/Redbooks
About Redbooks → Business Partner Programs

THIS PAGE INTENTIONALLY LEFT BLANK

Preface

This IBM® Redbooks® publication is a refresh of *IBM Technical Computing Clouds*, SG24-8144, *Enhance Inbound and Outbound Marketing with a Trusted Single View of the Customer*, SG24-8173, and *IBM Platform Computing Integration Solutions*, SG24-8081, with a focus on High Performance and Technical Computing on IBM Power Systems™.

This book describes synergies across the IBM product portfolio by using case scenarios and showing solutions such as IBM Spectrum™ Scale (formerly GPFS™). This book also reflects and documents the IBM Platform Computing Cloud Services as part of IBM Platform Symphony® for analytics workloads and IBM Platform LSF® (with new features, such as a Hadoop connector, a MapReduce accelerator, and dynamic cluster) for job scheduling. Both products are used to help customers schedule and analyze large amounts of data for business productivity and competitive advantages.

This book is targeted at technical professionals (consultants, technical support staff, IT Architects, and IT Specialists) that are responsible for delivering cost-effective cloud services and big data solutions on IBM Power Systems to uncover insights among client data so that they can take actions to optimize business results, product development, and scientific discoveries.

Authors

This book was produced by a team of specialists from around the world working at the International Technical Support Organization, Poughkeepsie Center.

Dino Quintero is a technical Project Leader and an IT Generalist with the International Technical Support Organization (ITSO) in Poughkeepsie, NY. His areas of expertise include enterprise continuous availability planning and implementation, enterprise systems management, virtualization, and clustering solutions. He is an Open Group Master Certified IT Specialist - Server Systems. He holds a master's degree in Computing Information Systems, and a Bachelor of Science degree in Computer Science from Marist College.

Daniel de Souza Casali is an IBM Cross Systems Senior Certified and has been working at IBM for 11 years. Daniel works for the Systems and Technology Group in Latin America as a Software Defined infrastructure IT Specialist. Daniel holds an Engineering degree in Physics from the Federal University of São Carlos (UFSCar). His areas of expertise include UNIX, SAN networks, IBM Disk Subsystems, clustering cloud, and analytics solutions.

Marcelo Correia Lima is a Business Intelligence Architect at IBM. He has 17 years of experience in development and integration of Enterprise Applications. His current area of expertise is Business Analytics Optimization (BAO) Solutions. He has been planning and managing BAO Solutions lifecycle implementation, involving Multidimensional Modeling, IBM InfoSphere® Data Architect, IBM InfoSphere DataStage®, IBM Cognos® Business Intelligence, and IBM DB2®. In addition, Marcelo has added Hadoop, big data, IBM InfoSphere BigInsights™, cloud computing, and IBM Platform Computing to his background. Before working as a Business Intelligence Architect, he was involved in the design and implementation of IBM WebSphere® and Java Enterprise Edition Applications for IBM Data Preparation/Data Services.

Istvan Gabor Szabo is an Infrastructure Architect and Linux Subject Matter Expert at IBM Hungary (IBM DCCE SFV). He joined IBM in 2010 after receiving his bachelor degree in Engineering Information Technology from the University of Obudai - John von Neumann Faculty of Informatics. Most of the time, he works on projects as a Linux technical lead. His areas of expertise are configuring and troubleshooting complex environments, and building automation methodologies for server builds. In his role as an Infrastructure Architect, he works on the IBM Standard Software Installer (ISSI) environment, where he designs new environments based on customer requirements.

Maciej Olejniczak is a Cross-functional Software Support Team Leader in a collaborative environment. He works internationally with external and internal clients, IBM Business Partners, services, labs, and research teams. He is a dedicated account advocate for large customers in Poland. Maciej is an IBM Certified Expert in Actualizing IT Solutions: Software Enablement. He achieved a master level in implementing all activities that transform information technology from a vision to an actual working solution. Maciej is an Open Group Master Certified IT Specialist.

Tiago Rodrigues de Mello is a Staff Software Engineer in Brazil with more than 10 years of experience. Tiago's areas of expertise include Linux system administration, software development, and cloud computing. He is an OpenStack developer and a Continuous Integration engineer at the IBM Linux Technology Center. Tiago holds a Bachelor in Computer Science degree from the Federal University of Sao Carlos, Brazil.

Nilton Carlos dos Santos is an IT Architect and a Certified IT Specialist and has been with IBM since 2007 and has 18 years experience in the IT industry. Before joining IBM, he worked in several different areas of technology, including Linux and UNIX administration, database management, development in many different languages, and network administration. Nilton Carlos also has deep expertise in messaging, automation, monitoring, and reporting system tools. He enjoys working with open source software.

Thanks to the following people for their contributions to this project:

Richard Conway and David Bennin
International Technical Support Organization, Poughkeepsie Center

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:
ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- ▶ Use the online **Contact us** review Redbooks form found at:

ibm.com/redbooks

- ▶ Send your comments in an email to:

redbooks@us.ibm.com

- ▶ Mail your comments to:

IBM Corporation, International Technical Support Organization
Dept. HYTD Mail Station P099
2455 South Road
Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- ▶ Find us on Facebook:

<http://www.facebook.com/IBMRedbooks>

- ▶ Follow us on Twitter:

<http://twitter.com/ibmredbooks>

- ▶ Look for us on LinkedIn:

<http://www.linkedin.com/groups?home=&gid=2130806>

- ▶ Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

<https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm>

- ▶ Stay current on recent Redbooks publications with RSS Feeds:

<http://www.redbooks.ibm.com/rss.html>



Introduction to IBM Platform Computing

As the velocity of innovation increases, enterprises and organizations must have an infrastructure that can accelerate time to results for compute- and data-intensive applications. Current and future challenges for technical computing require you to leverage the proper compute power and capacity for your business. It is mandatory to reevaluate how resources must be used more efficiently instead of simply adding more arbitrarily.

This chapter introduces IBM Platform Computing and how its portfolio supports high-performance computing management in the new era of computing.

This chapter covers the following topics:

- ▶ IBM Platform Computing solutions purpose
- ▶ Cluster, grids, and clouds
- ▶ IBM Platform Computing Services
- ▶ Benefits and industries

1.1 IBM Platform Computing solutions purpose

Two user segments can be identified within the technical computing market. One segment consists of the business/application users that try to make their applications meet the business demands. The second segment is the IT organization either at the departmental level or at the corporate level that tries to provide the IT support to run these business applications more efficiently.

From the business and application user side, applications are becoming more complex. One good example is risk management type simulations that try to improve results to have more complex algorithms or to add more data.

All this complexity is driving the need for more IT resources. Clients are having trouble getting these resources because of budgetary constraints, which restricts their business opportunities. This approach is considered the *demand* side.

On the supply side, the IT organizations set up siloed data centers for different application groups to ensure service levels and availability when they are needed. Typically, an infrastructure is suboptimal because significant workload requirements drive the overall size of the infrastructure, which is overprovisioned. Unfortunately, the IT organization is constrained by budget concerns, so it cannot add more hardware.

You can either take advantage of new technologies, such as graphics processing units (GPUs), or you can try to move to a shared computing environment to simplify the operating complexities. A shared computing environment can normalize the demand across multiple groups. It effectively provides visibility into an environment that is considered a much larger IT infrastructure even though the client does not have to fund it all themselves. It provides a portfolio effect across all the demands.

Overall IT resources are fairly static, so clients want to be able to branch out to cloud service providers as needed. If clients have short-term needs, they can increase their resources, but do not always want to keep the resources on a long-term basis.

There are many demands on the business and user side. The resources are from the IT side. How do you make these two sides fit together without increasing costs?

IBM Platform Computing solutions deliver the power of sharing for technical computing and analytics in distributed computing environments.

This shared services model breaks through the concept of a siloed application environment and creates a shared grid that can be used by multiple groups. This shared services model offers many benefits, but it is a complex process to manage. At a high level, IBM provides four key capabilities across all its solutions:

- ▶ The creation of shared resource pools, both for compute-intensive and data-intensive applications, is heterogeneous in nature. These pools span physical, virtual, and cloud components. The users do not know that they are using a shared grid. They know that they can access all the resources that they need when they need them and in the correct mix.
- ▶ Shared services are delivered across multiple user groups and sites and in many cases are global. This flexibility is important to break down the silos that exist within an organization. The solution provides much of the governance to ensure that you have the correct security and prioritization and all the reporting and analytics to help you administer and manage these environments.

- ▶ Workload management is where policies are applied on the demand side to ensure that the correct workloads get the correct priorities, but then also place them in the correct resources on both the demand side and the supply side. So, the correct algorithm is provided to schedule, maximize, and optimize the overall environment to deliver service level agreements (SLAs) with all the automation and workflow. If you have workloads that depend on each other, you can coordinate these workflows to achieve a high utilization of the overall resource pool.
- ▶ Transform a static infrastructure into dynamic infrastructure. If you have undedicated hardware, such as a server or a desktop, you can bring it into the overall resource pool in a manner so that you can burst workloads both internally or externally to third-party clouds. The solution works across multiple hypervisors to take advantage of virtualization where it makes sense. You can change the nature of the resources, depending on the workload queue to optimize the overall throughput of the shared system.

1.2 Cluster, grids, and clouds

A *cluster* is typically a single application or single group. Because clusters are in multiple applications, multiple groups, and multiple locations, they became more of a grid, so in the past you needed more advanced policy-based scheduling to manage it.

In the era of cloud computing, it is now all about how using a much more dynamic infrastructure against an infrastructure that uses the concepts of on-demand self-service. Many of your grid's clients already considered their grids to be clouds. The evolution to cloud computing continues with the ability of the platform to manage the heterogeneous complexities of distributed computing. This management capability has many applications in the cloud. Figure 1-1 shows the cluster, grid, and High Performance Cluster (HPC) Cloud evolution.

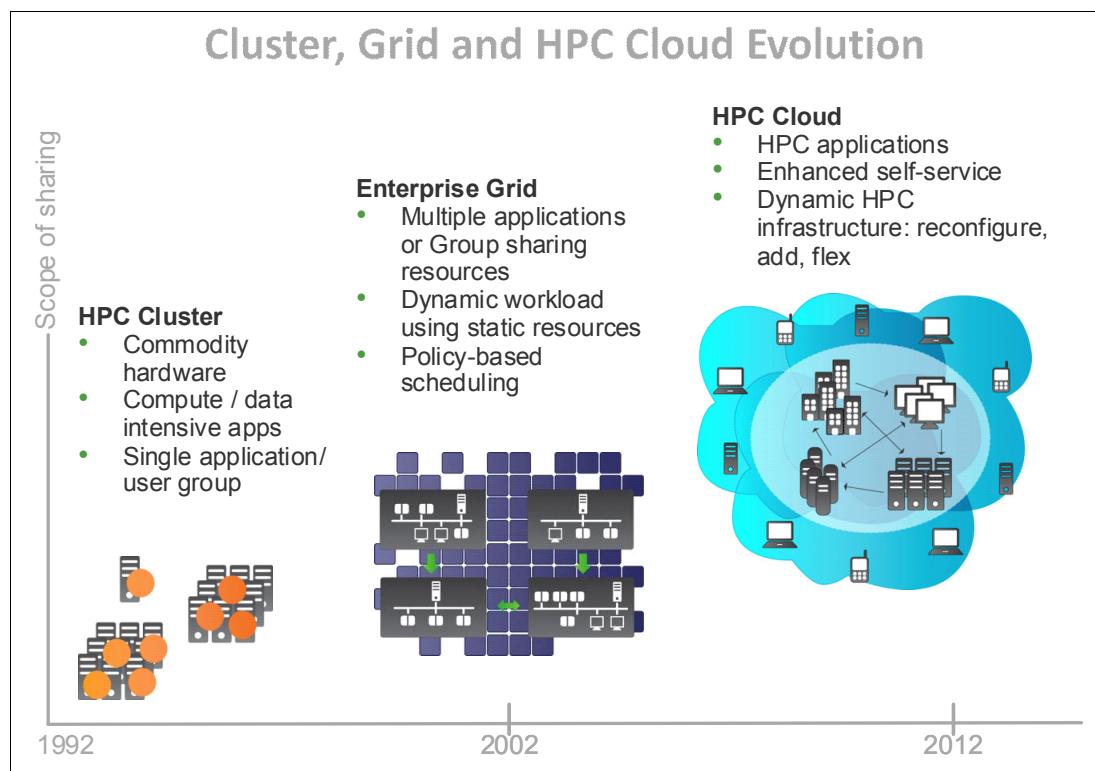


Figure 1-1 Evolution of distributed computing

Figure 1-1 on page 3 illustrates the transition from cluster to grid to clouds and how the expertise of IBM in each of these categories gives IBM Platform Computing solutions a natural position in this transition as the market moves into the next phase.

It is interesting to see the evolution of the types of workloads that moved from the world of HPC into financial services through the concepts of risk analytics, risk management, and business intelligence (BI). Data-intensive and analytical applications are increasingly adopted into the installation base of IBM Platform Computing solutions.

The application workload types become more complex as people move from clusters to grids to the much more dynamic infrastructure of cloud. There has been an evolution of cloud computing for HPC and private cloud management across the Fortune 2000 installation base.

This evolution occurs in many different industries, from the life sciences space to the computer and engineering areas and defense digital content. There is good applicability for anyone that needs more compute capacity. There is good applicability for addressing more complex data tasks when you do not want to move the data but you might want to move the compute for data affinity. How do you bring it all together and manage this complexity? You can use the IBM Platform Computing solutions capability to manage all of these areas. This capability differentiates it in the marketplace.

IBM Platform Computing solutions are viewed as the industry standard for computational-intensive design, manufacturing, and research applications.

IBM Platform Computing is the vendor of choice for mission-critical applications. Mission-critical applications are applications that can be large scale with complex applications and workloads in heterogeneous environments. IBM Platform Computing is enterprise-proven with an almost 20-year history of working with the largest companies in the most complex situations. IBM Platform Computing has a robust history of managing large-scale distributed computing environments for proven results.

1.3 IBM Platform Computing Services

Today, businesses must run more iterations, simulations, and analysis, and get business results as fast as possible. Businesses generate a vast amount of big data and must solve compute-intensive challenges. Therefore, you must maximize the potential of your computing power and the supporting infrastructure to accelerate your applications at scale, extract insights from different and complex data, and make critical decisions faster.

IBM Platform Computing is a collection of high-performance, low-latency systems management solutions and services that pools your technical computing resources, manages them efficiently across multiple groups, and gets the most out of your IT investment. IBM Platform Computing can help to optimize and manage a cluster through highly secure multi-site grids and HPC clouds. Figure 1-2 on page 5 shows the Figure 1-2 on page 5 portfolio.

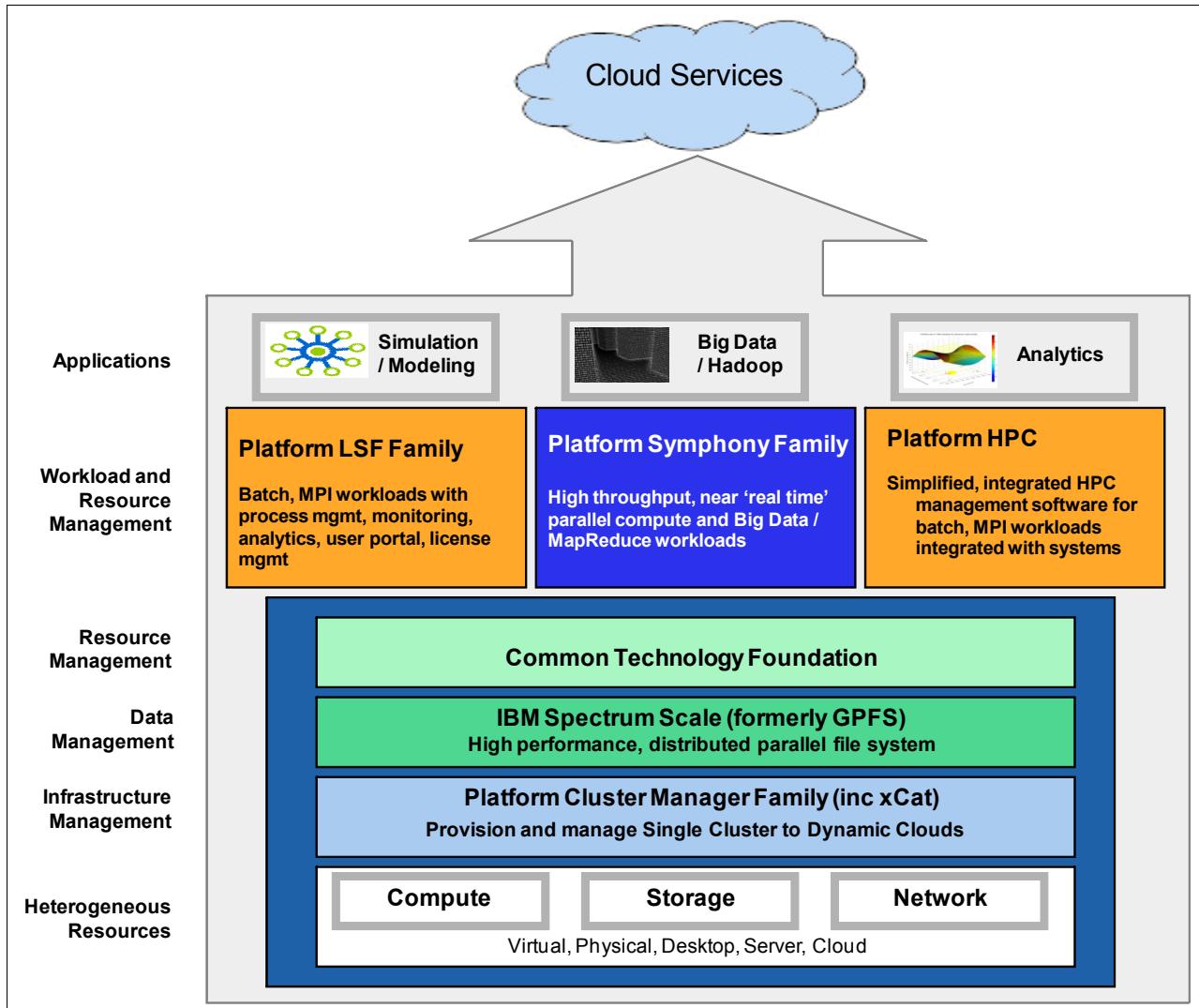


Figure 1-2 IBM Platform Computing Services

1.3.1 IBM Platform High Performance Cluster

IBM Platform High Performance Cluster (HPC) is a complete, high-performance computing management solution in a single product. Its robust cluster and workload management capabilities are accessible by using the latest design in web-based interfaces, making it powerful and simple to use. It includes a set of cluster and workload management features that help reduce the complexity of your HPC environment and improve your time-to-results. IBM Platform HPC provides a unified set of management capabilities that make it easy to harness the power and scalability of a technical cluster, resulting in shorter time to system readiness and increased user productivity, and optimal throughput. Platform HPC includes the following features:

- ▶ Cluster management
- ▶ Workload management
- ▶ Workload monitoring and reporting
- ▶ System monitoring and reporting
- ▶ IBM Platform MPI
- ▶ Integrated application scripts and templates for job submission

- ▶ Central web portal
- ▶ High availability

1.3.2 IBM Load Sharing Facility

IBM Load Sharing Facility (LSF) is a powerful workload management platform for demanding, distributed HPC environments. It provides a comprehensive set of intelligent, policy-driven scheduling features that enable you to use all of the compute infrastructure resources and ensure optimal application performance. The Platform LSF product family helps to ensure that all available resources are fully used by enabling you to take full advantage of all technical computing resources, from application software licenses to available network bandwidth. The Platform LSF family can help in the following areas:

- ▶ Reduce operational and infrastructure costs by providing optimal SLA management and greater flexibility, visibility, and control of job scheduling.
- ▶ Improve productivity and resource sharing by fully using hardware and application resources, whether they are just down the hall or halfway around the globe.

1.3.3 IBM Platform Symphony

IBM Platform Symphony delivers powerful enterprise-class management for running a wide variety of distributed applications and big data analytics on a scalable, shared grid. It accelerates dozens of parallel applications, for faster results and better utilization of all available resources.

For many enterprises, grid computing is the ideal solution to handle jobs such as analyzing big data. For grid-enabled applications, maximizing performance and scale is crucial.

However, some grid products have architectural limitations, requiring a particular operating system or specific developer tools. Because of budgetary concerns, companies want better ways to improve IT performance, reduce infrastructure costs and expenses, and meet the demand for faster answers. IBM Platform Symphony helps to control the massive compute power that is available in the current and future technical computing systems. Therefore, it is possible to achieve breakthrough results in business and research activities. Moreover, IBM Platform Symphony can address challenges in parallel application development and deployment, and in technical computing infrastructure management.

IBM Platform Symphony can deliver faster and better quality results even when there is less infrastructure available.

1.4 Benefits and industries

IBM Platform Computing manages complex calculations, either compute-intensive or data-intensive in nature, on a large network of computers by optimizing the workload across all resources. For the most complex challenges that are faced by different industries, IBM Platform Computing enables fast design modeling and analysis of large data sets and flexible high performance clusters, allowing you to achieve a wide range of benefits:

- ▶ Better IT agility of the organization
- ▶ Increased resources utilization because of the reduced number of IT silos throughout the organization
- ▶ Increased infrastructure utilization with pools of shared resources

- ▶ Reduced costs by using a heterogeneous, shared infrastructure
- ▶ Faster time-to-results
- ▶ Higher application service levels and throughput
- ▶ Simple setup and deployment to decrease the time and cost of IT administration

Regardless of the architecture, a business application sits on top of an architectural model or any type of data storage technology. Common storage problems can be addressed automatically, which avoids wasting the time of someone who needs that time to work on business functions. Because it can impact both the business and IT side, cloud computing is now a primary structural element in IT solutions. IBM Platform Computing is a powerful and integrated platform with the expertise to support key solutions of the industry. IBM Platform Computing provides solutions that enable the creation of dynamic, flexible clusters, HPC cloud environments, and big data analytics infrastructure that address compute- and data-intensive challenges that are specific to different industries:

- ▶ Aerospace and defense

Aerospace and defense companies that develop or manufacture products need speed, flexibility, agility, and control over the design cycle to meet time-to-market requirements and maximize profitability. Two key areas of the design cycle are crucial for achieving and maintaining a competitive advantage: testing and simulation. Traditionally, these processes are time-consuming because every time a new idea, component, or parts appear, they require a scale prototype and physical wind-tunnel testing. IBM Platform Computing technical and HPC applications help aerospace and defense industries with product development, critical business decisions, and breakthrough science.

- ▶ Automotive

Challenging requirements often come from the automotive industry, which has fast design and complex build environments all the time. Automotive companies need speed, agility, control, and visibility across the design infrastructure and lifecycle to meet time-to-market requirements and maximize profitability. IBM Platform HPC with grid and cloud solutions can help automotive companies transform their design chain to develop designs better, faster, and cheaper.

- ▶ Financial markets and insurance

Facing increasingly restrictive economic pressures and growing regulatory demands, financial services companies are looking for the following things:

- Better insights for trading, risk management, and customer support
- Ways to improve IT performance to meet these demands and reduce operating costs at the same time
- Greater operational agility to respond faster to market changes

IBM Platform Computing solutions can help to improve the performance of analytics to support faster, more accurate, and more reliable decision making in financial markets. Private HPC clusters, grids, and clouds allow multiple applications and lines of business to use effectively a common heterogeneous, shared infrastructure to support both compute and big data analytics.

Compute intensive analytics include the following items:

- Pricing of market and credit risk
- Compliance reporting
- Pre-trade analysis
- Back testing and new product development

- ▶ Chemical and petroleum

Chemical and petroleum organizations face huge upstream and downstream challenges. The cost of exploration is high. The cost of drilling and the consequences from drilling in the wrong location can cost hundreds of millions of dollars, in addition to the months or years that were spent to secure drilling rights and to set up an infrastructure to support the drilling. Oil and gas producers rely heavily on 3D simulations and models to help pinpoint the most promising areas for exploration. Engineers need more processing power to run these simulations. IBM Platform Computing can serve chemical and petroleum clients who are turning to high performance technical computing to accelerate time-to-results, improve infrastructure utilization, and reduce operating costs.

- ▶ Life sciences

Life sciences organizations face huge pipeline and productivity challenges. To increase discovery productivity, innovate in research and development, and compete more effectively, organizations must establish an optimized, flexible, and resilient infrastructure foundation to improve clinical development processes. With shifting regulatory burdens and the need to compress the timeline from discovery to approval, research teams need comprehensive, high-performance technical computing infrastructure solutions with the flexibility to process massive amounts of data and support increasingly sophisticated analyses. Genomic medicine promises to revolutionize medical research and clinical care. By investigating the human genome in the context of biological pathways and environmental factors, it is now possible for genomic scientists and clinicians to identify individuals at risk of disease, provide early diagnoses based on biomarkers, and recommend effective treatments. IBM Platform Computing solutions for cloud can help life sciences customers by delivering high performance resources that use the advantages of cloud.

- ▶ Education

Universities need high-performance IT environments that can process massive amounts of data and support increasingly sophisticated simulations and analyses. Researchers need computing power, agility, and scalability to analyze rapidly a wide range of structured and unstructured data and achieve deeper insights in many disciplines, ranging from astrophysics to public health. IBM Platform Computing with grid and HPC cloud solutions can help universities bring together often highly distributed IT clusters to create a shared high-performance compute and big data environment that is more agile, scalable, and cost-effective.



Technical computing software portfolio

This chapter describes the IBM Platform Computing software portfolio and how the portfolio can help you with the new high-performance computing paradigms. The entire software stack is covered, including the new data-centric view.

This chapter covers the following topics:

- ▶ Data-centric view for technical computing
- ▶ Storage management
- ▶ Workload management
- ▶ Cluster management
- ▶ Virtual resource management
- ▶ IBM Platform Computing Cloud Services

2.1 Data-centric view for technical computing

Technical computing is not just about Floating Point Operations Per Second (FLOPS) because most new applications are data-centric driven and not processor-based, as with the Linpack test. A paradigm shift is required to analyze all the data that is collected to have a more precise output generated.

Figure 2-1 shows that the Linpack benchmark is limited to special cases and is not the preferred metric to evaluate today's workloads because other tools can provide better options to perform the benchmarks.

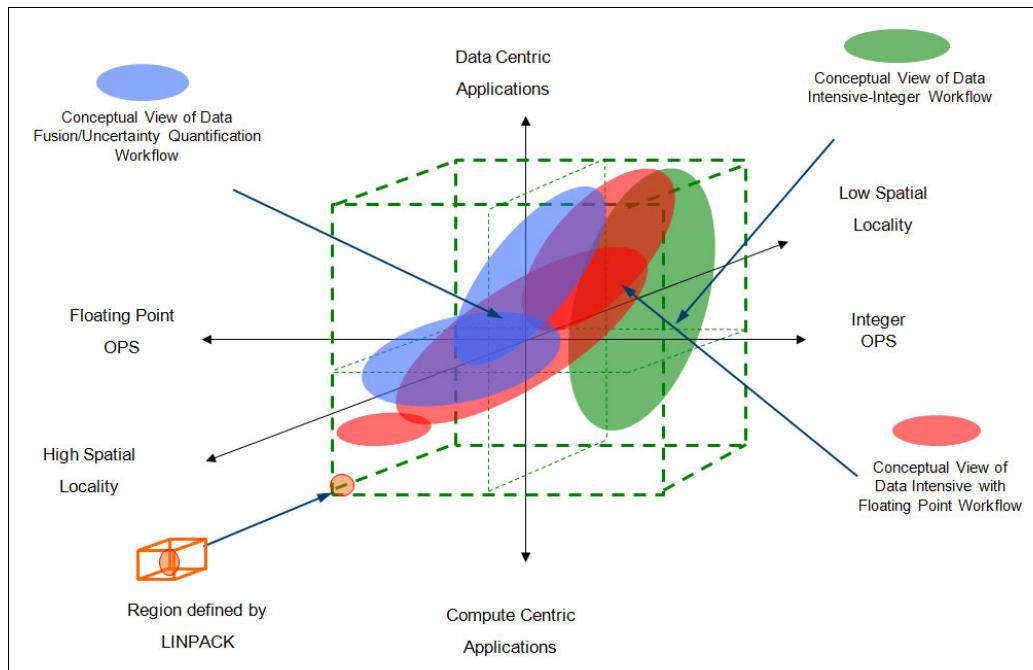


Figure 2-1 Linpack compared to different workflow concepts

Another comparison that can be performed is between Linpack and the applications on the market that are used in technical computing. This comparison is shown at Figure 2-2.

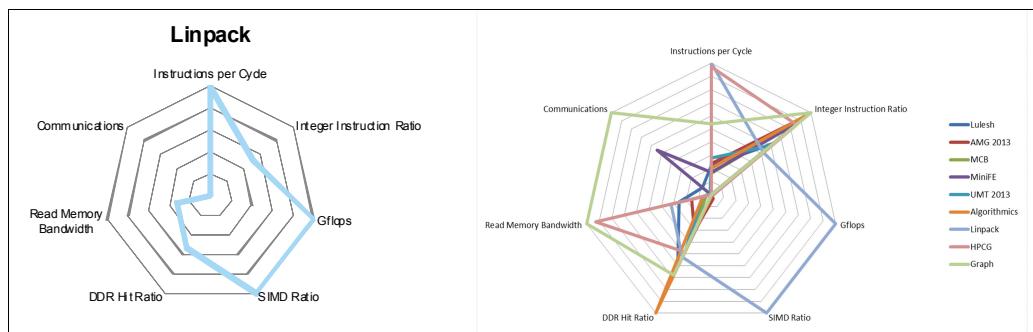


Figure 2-2 Comparison between Linpack and different workloads

Because the difference is clear, more attention must be paid to how data affects the workload, and how to make data available as fast as possible to processing points.

2.2 Storage management

Technical computing usually requires fast access to data, which provides better input/output operations (I/O) for your application and enhances the user experience. Scaling processors horizontally is common in a high performance computing (HPC) environment, so why not do the same with storage?

IBM Spectrum Scale (formerly GPFS) is a reliable technology that is used by many clients around the world and used on many Top 500 supercomputers. This solution has evolved: In addition to being a file system solution, it also provides object storage, big data storage, and storage for cloud computing environments (OpenStack Cinder and Swift for example).

Because Spectrum Scale is based on GPFS, it has a different approach to HPC environments than the approach that is used in a conventional network-attached storage (NAS) environment. The solution uses a technology that is called Network Shared Disks (NSD) in which a parallel read and write can be achieved for a single file I/O from a single NSD client. This approach is in contrast to an NAS client, as shown in Figure 2-3.

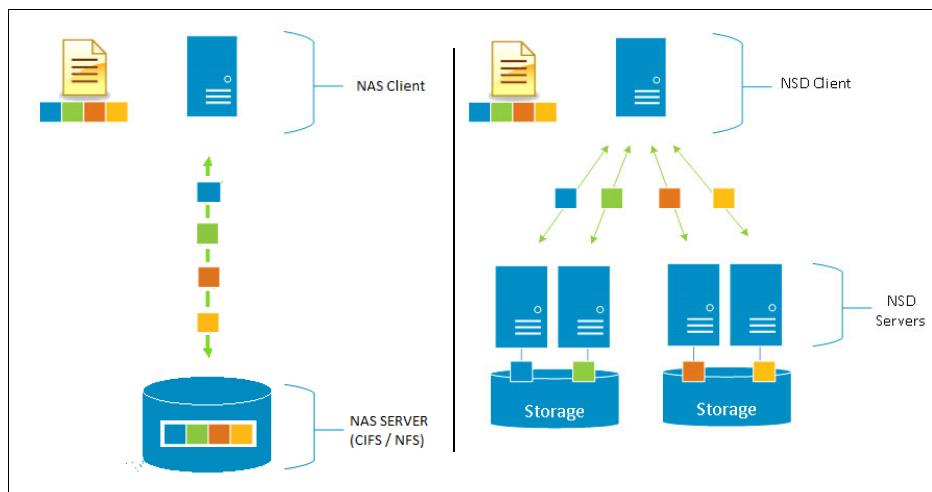


Figure 2-3 Traditional NAS and IBM Spectrum Scale comparison

Chapter 4, “IBM Spectrum Scale (formerly GPFS)” on page 29 provides more information about Spectrum Scale technology.

2.3 Workload management

Workload management is the core of any grid computing, as shown in 1.4, “Benefits and industries” on page 6. HPC needs powerful tools to manage application behavior in a cluster. IBM Platform Computing has over 20 years of history leading the management of HPC workloads.

Platform Computing started with the Load Sharing Facility (LSF), which is shown in Chapter 5, “IBM Platform Load Sharing Facility product family” on page 47, and since then the portfolio has grown to support low-latency scheduling for demanding workloads. Platform Symphony is a service-oriented grid manager that can leverage desktops and virtual servers (including GPU).

Today's workloads are not only about processing power anymore, so when large amounts of data are required to complete computations, it is preferable that your applications access the required data unhindered by the location of the data in relation to the application execution environment.

New data collection methods and technologies make most new applications data-centric. Workload management must be aware of processor slots and data location. So, the scheduling solution puts the workload near the data location or takes it to the node that is nearest to the data, as shown in Figure 2-4.

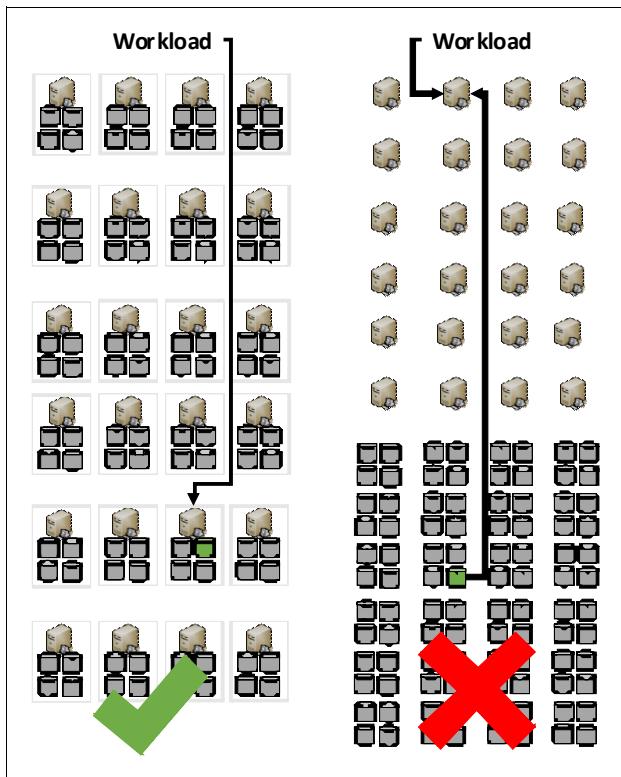


Figure 2-4 Data-centric aware workload scheduling

To deal with this new paradigm, IBM Platform Computing introduces new add-ons for the workload management portfolio: The Hadoop connector for the Platform Symphony - Advanced Edition, and the MapReduce accelerator and Data Manager for Platform LSF. For more information about big data and its relationship to HPC, see Chapter 3, “Big data, analytics, and risk calculation software portfolio” on page 15.

2.4 Cluster management

A technical computing cluster environment might be difficult to manage without the correct cluster management tool. When the environment is more than 300 machines, it is difficult to know what is happening to each node in the cluster.

IBM Platform Cluster Manager is a powerful cluster management software that allows system administrators to manage a single system or even a complex cluster with multi-tenancy support by automating the layer 2 network (creating VLANs), the deployment of the operating system, and all software components. Platform Cluster Manager provides centralized monitoring with customizable alert actions on nodes, switches, and even on Spectrum Scale.

Three solutions for cluster management are available: Platform HPC (see Chapter 7, “IBM Platform High Performance Computing” on page 105) is a basic solution for small clusters that provides cluster management and includes the Express Version of Platform LSF and the Platform MPI (see Appendix A, “IBM Platform Computing Message Passing Interface” on page 147, and IBM Platform Cluster Manager, which comes in two versions: Standard Edition and Advanced Edition (see Chapter 8, “IBM Platform Cluster Manager” on page 117).

2.5 Virtual resource management

For virtual resource management, the projects that are the most active are the OpenStack projects. IBM adopted the OpenStack project and added improvements for better use in a cloud-computing environment. IBM Cloud Manager for OpenStack leverages all the properties of the OpenStack project and enhances it to cover a broader environment.

In Cloud Manager for OpenStack, KVM x86 virtual servers, Linux on IBM z Systems™, HyperV, IBM POWER® environments (through PowerVC and PowerKVM), and VMware are available. A DevOps environment for easy deployment of OpenStack with IBM DB2 database helps the customer to get the benefits sooner than trying to build the whole stack project by project. An improved resource scheduler is bundled to extend policies and capacities for online virtual server management.

The IBM enhanced environment is shown in Figure 2-5.

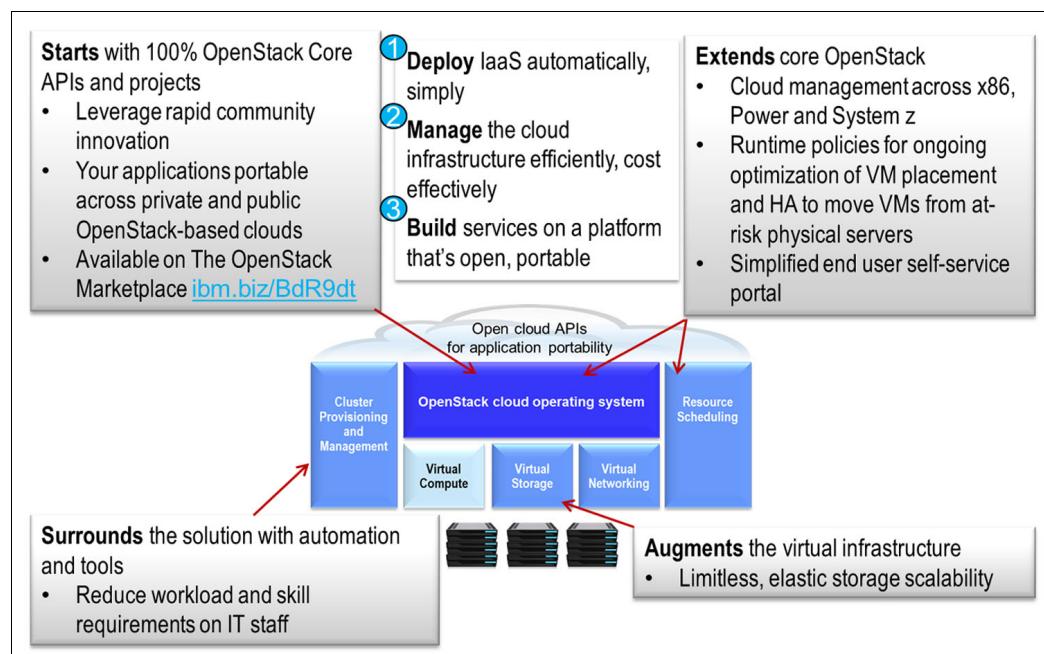


Figure 2-5 IBM Cloud Manager environment with Resource Scheduler

Chapter 9, “IBM Cloud Manager” on page 125 provides more information about this solution, and how it can help you create a resource aware cloud.

2.6 IBM Platform Computing Cloud Services

If you need elasticity for your on-premises HPC to meet peak usage or if you want to start a new cluster from scratch without having to allocate floor space or buy new infrastructure, high performance computing (HPC) in the Cloud is the correct choice for you. IBM offers this possibility with a solution that delivers a versatile, application-ready cluster in the cloud for organizations that must quickly and economically add computing capacity. The solution includes IBM Platform LSF, IBM Platform Symphony workload management software, and Spectrum Scale data management software that is delivered as a service. Both hybrid or plain cloud models might be used with the IBM SoftLayer® infrastructure in the architecture, as shown in Figure 2-6.

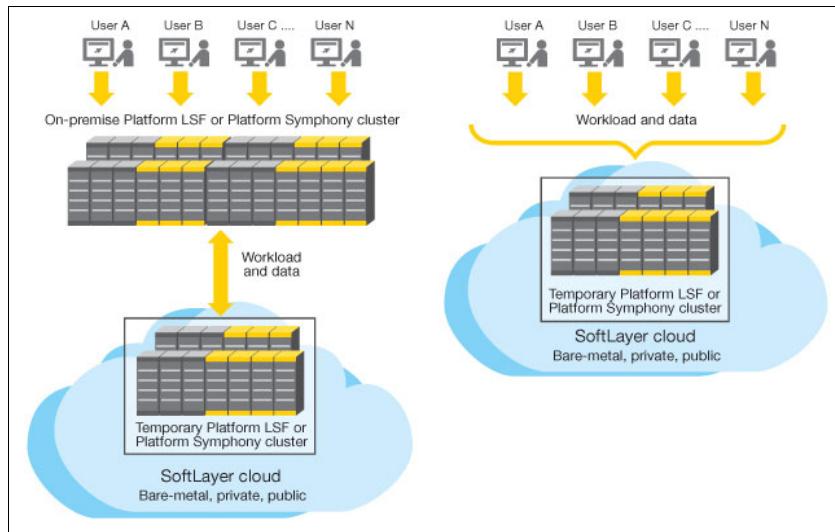


Figure 2-6 Architectures for IBM Platform Computing Cloud Services

Chapter 10, “IBM Platform Computing Cloud Services” on page 129 provides more information about this service.



Big data, analytics, and risk calculation software portfolio

This chapter introduces and describes the IBM Big Data analytics and risk calculation offerings that can help customers complement and improve their analytics solutions.

This chapter covers the following topics:

- ▶ What is big data
- ▶ Big data analytics
- ▶ Why use an IBM Risk Analytics solutions
- ▶ Scenario for minimizing risk and building a better model

3.1 What is big data

Many people are using the term *big data* to describe the latest industry trend. To help you to understand it better, this chapter provides a foundational understanding of big data, what it is, and why you should care about it.

Surveys show that 90% of the data that is generated around the world is unstructured, primarily because of the large amount of data that social networks are creating.¹ For this reason, companies have been looking for technologies to filter useful information for business. Figure 3-1 shows an example of big data sources.

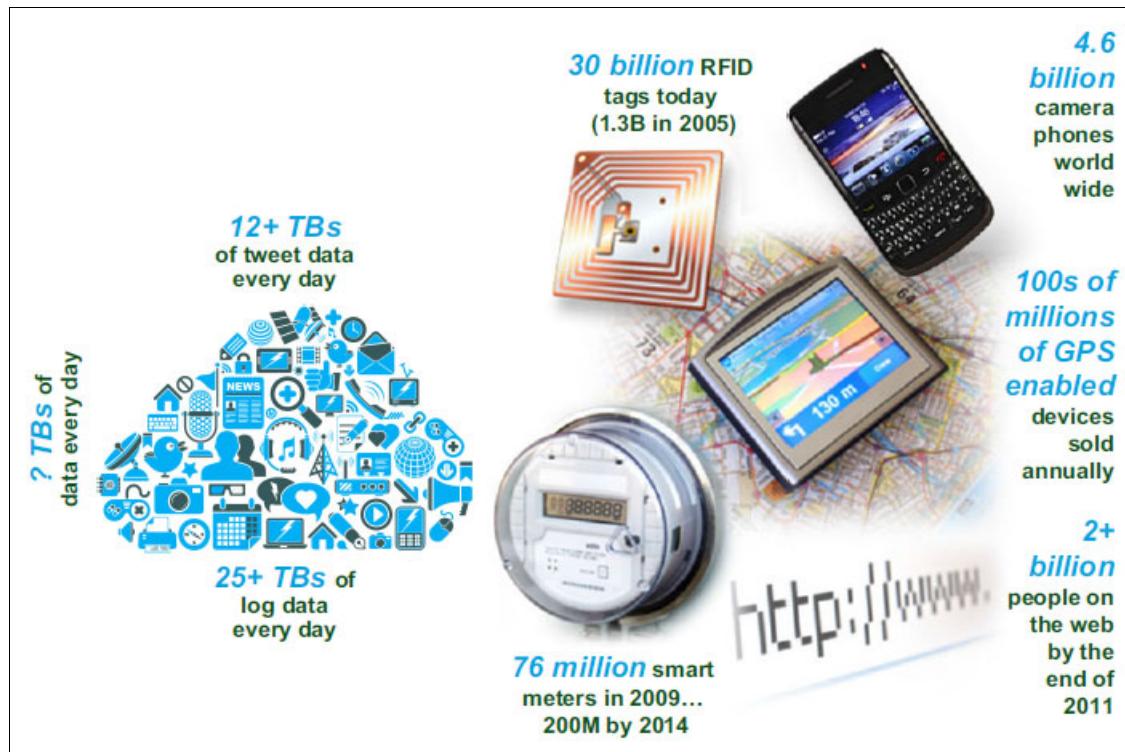


Figure 3-1 Sources of big data

A single phone call can result in hundreds of records that have all the details and tracking about it, which must be stored by the telephone service provider for the creation of plans according to the profile of each consumer. So, as the world has millions of wireless lines in service, we can imagine the vast amount of information that this sector generates. If that can be mined and analyzed in real time and accurately, it can become valuable assets for companies. Running this workload is not an easy task, and big data appears to be the solution that the market needs to solve the requirement of fast data mining that businesses require.

Telecommunications, manufacturing, retail outlets, utility companies, media, and other industries generate data every minute. This ocean of data without control is challenging certain areas of IT, and business organizations must find ways to use the data to gain a competitive advantage and better business results.

¹ Source: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4127205/>

The term *big data* was created by market analysts to describe the exponential growth of data that companies must extract and classify as useful information. The concept was disseminated to alert organizations about the need to adopt a strategy to assess unstructured data, which is beyond the control of IT. The big data challenge is not the high volume of data, but to capture the complexity of information in different media formats and use them in real time. Companies are struggling to capture content from social networks when they go live, and analyze them with other databases of call centers. This new approach is different from what was proposed by *business intelligence*, which looks at historical events to make decisions.

Big data has the advantage of analyzing events when they are happening and anticipate measures. You can, for example, monitor in real time the consumption pattern of sales of a supermarket in a given region and make correlations with climate to offer ice cream, according to the preference of buyers. If the temperature changes, you can quickly change your campaign and make other types of promotion.

In healthcare areas, it can be possible to determine when some person can have more of a chance to have a heart attack, cancer, or any other disease. Medical exams such as x-ray and magnetic resonance imaging or medical and monitoring devices might be a rich source of data to predict health problems before they occur.

The same type of analysis can be done by telecommunication operators to create plans that are tailored to mobile users by looking at consumption patterns with real-time ratings.

Today, companies make decisions based on feelings or market research, but in fact the best data comes from the companies themselves. They have exclusive information that can improve business. Big data can help them find their own wealth. Big data is a new name for an old problem. Institutes of meteorology were pioneers in adopting this approach. They looked at the satellite that watches the clouds, and used it to gather numerous historical information to see whether it might rain. So, they were able to determine trends for planting the next crop of soybeans, for example.

This type of analytic workload required a high investment for scientific supercomputers that were expensive. With the evolution of technology, this computational power has become available for commercial applications, allowing companies to adopt strategies for big data. For example, a large supermarket might notice that buyers who buy milk also buy diapers. However, in this case, the information is structured, as they are normally in a relational database.

Now, the challenge is to cross this data with unstructured data and do it quickly. Big data is considered a new service-oriented architecture (SOA), and all IT suppliers want to ride on this new wave. Many of them announced platforms to try to help companies to handle their big data more efficiently, and extract important data from social networks and other unstructured sources. Hadoop, an open source software platform, is closely associated with the movement of big data. The market is seeing devices based on Hadoop, and data warehousing is transforming into a technology that is increasingly necessary. Another solution is Spark, which is 100x times faster than Hadoop when programs run in memory and 10x faster on disk.

To read more about Spark, see this website:

<https://spark.apache.org/>

Big data is not just about the sheer volume of data that is being created. With a number of unstructured sources creating this data, a greater variety of data is now available. Each source produces this data at different rates or what is called *velocity*. In addition, you still must solve the veracity of this new information as you do with structured data. Here is where the information management industry had its epiphany: Whether your workload is largely transactional or online analytics processing (OLAP) and resource-intensive, both cases operate on structured data. Systems that are designed for the management and analysis of structured data provided valuable insight in the past, but what about all of the newer text-based data that is being created? This data is being generated everywhere you look. There is a larger volume of data, a greater variety of data, and it is being generated at a velocity that traditional methods of data management cannot efficiently harvest or analyze. To provide added insight into what is going on within your particular business arena, you must address the 4 Vs that define big data. A visual representation of the 4 Vs is shown in Figure 3-2.

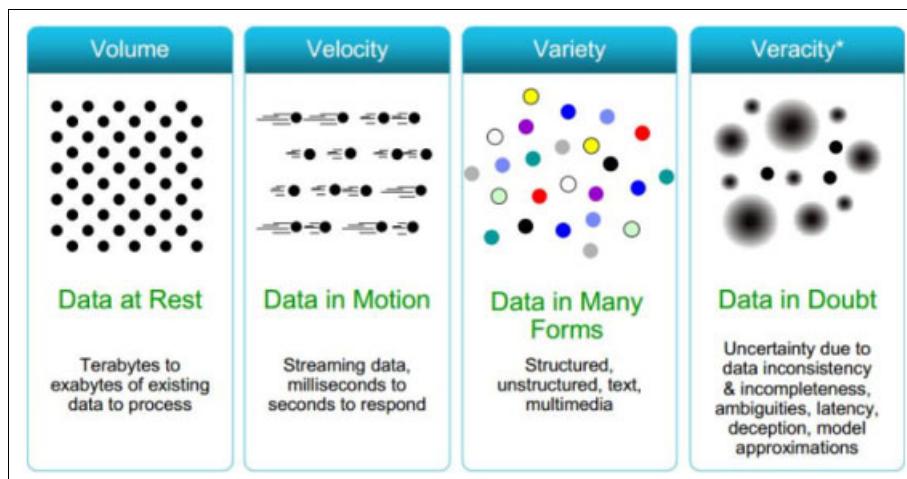


Figure 3-2 4 Vs of big data

The news about big data promises to be hot in the coming years. The market has not stopped and the data even less.

3.2 Big data analytics

With information growing at fast rates and users who demand quick and effective research of this information, your analytics workloads need a powerful base. IBM Platform Computing software improves the performance of your computing infrastructure for your most demanding analytics programs.

3.2.1 Big data analytics challenge

With more intelligent and connected devices and systems, the amount of information that you are collecting is increasing at alarming rates. In some sectors, as much as 90% of that information is unstructured and increasing at rates as high as 50% per year. To keep your business competitive, to innovate, and to get products and solutions to market quickly, you must be able to evaluate that information and extract insight from it easily and economically. For big data analytics, current alternatives do not offer the response time for statistical tasks, reducing user efficiency and delaying decision making.

3.2.2 Big data analytics solution

IBM Platform Computing software improves the performance of your most demanding applications with a low-latency solution for heterogeneous application integration on a shared multi-tenant architecture. IBM Platform Symphony V7.1 offers several editions. Figure 3-3 summarizes the features and differences between each Platform Symphony edition.

Features	Developer	Express	Standard	Advanced
Low-latency HPC SOA	Yes	Yes	Yes	Yes
Agile service and task scheduling	Yes	Yes	Yes	Yes
Dynamic resource orchestration		Yes	Yes	Yes
Standard and custom reporting			Yes	Yes
Desktop, server, and VM harvesting capability			Yes	Yes
Data affinity				Yes
MapReduce framework	Yes			Yes
MultiCluster management				Yes
Maximum hosts per cores	2 hosts	240 cores	5K hosts, 128K cores	5K hosts, 128K cores
Maximum application managers	5	300	300	

Add-ons	Standard	Advanced
Application Service Controller for Platform Symphony		Yes
Co-processor harvesting	Yes	Yes
Desktop harvesting	Yes	Yes
Server and VM harvesting	Yes	Yes
GPU harvesting	Yes	Yes

Figure 3-3 IBM Platform Symphony V7.1 editions

The software provides great resource availability and predictability. It also supports several programs and file systems, operational maturity, SLA policy control, and high resource utilization for MapReduce and applications that are not MapReduce.

With years of experience in distributed workload scheduling and management, IBM Platform Computing offers proven technology that powers objective, critical, and the most demanding workloads of many large companies. IBM Platform Symphony software offers unmatched distributed workload runtime services for distributed computing and big data statistics programs.

3.2.3 IBM Big Data and analytics areas with solutions

There are five different areas where IBM Big Data and analytics can discover fresh insights, capture the time-value of data, and act with confidence:

- ▶ Marketing
- ▶ Operations
- ▶ Finance and human resources
- ▶ New business models
- ▶ IBM IT solutions

Marketing

IBM Platform Symphony software uses marketing methods that allow a subnanosecond response and quick provisioning for a wide range of workloads. Short-running jobs have a smaller percentage of time that is spent in the provisioning and deprovisioning actions, providing a higher ratio of useful work to overhead. It also has a high job throughput rate, where the system allows more than 17,000 tasks/second to be submitted. Marketing analytics solutions from IBM can help you understand which marketing strategies and offers appeal most to your high-value customers and prospects.

Solutions

In the marketing area, here are some of IBM Platform Symphony solutions:

1. Customer analytics: Customer analytics solutions from IBM helps marketers understand and anticipate what customers want. These solutions target the preferred customers for marketing programs, predict which customers are at-risk of leaving so you can retain them, and maximize customer lifetime value through personalized up-sell and cross-sell.

Products:

- IBM Analytical Decision Management
- IBM Cognos Business Intelligence
- IBM Digital Analytics
- IBM Predictive Customer Intelligence
- IBM Social Media Analytics
- IBM SPSS® Data Collection
- IBM SPSS Modeler
- IBM SPSS Statistics
- IBM Tealeaf® CX
- IBM Tealeaf cxImpact
- IBM Tealeaf cxOverstat
- IBM Tealeaf cxReveal
- IBM Tealeaf CX Mobile

2. Marketing performance analytics: IBM Marketing Performance Analytics solutions give marketers the ability to measure ROI and eliminate the guesswork from marketing programs. Marketers rely on these solutions to access and analyze critical marketing metrics through customized reporting options, such as dashboards, KPIs, and easy to understand visualizations.

Products:

- IBM Business Intelligence
- IBM Cognos Insight™
- IBM Cognos Express
- IBM SPSS Statistics
- IBM SPSS Modeler
- IBM Social Media Analytics

3. Social Media Analytics: The IBM Social Media Analytics solution unlocks the value of customer sentiment in social media. Marketers use this solution to measure the social media impact of products, services, markets and campaigns, and use these insights to improve marketing programs and address customer satisfaction issues.

Products:

- IBM Social Media Analytics
- IBM Social Media Analytics Software as a Service (SaaS)

Operations

IBM Predictive Maintenance and Quality can reduce asset downtime by analyzing asset data in real time, and detecting failure patterns and poor quality parts before problems occur.

Features

IBM Predictive Maintenance and Quality is a preconfigured software, available either on-cloud or on-premises, which helps you monitor, maintain, and optimize assets for better availability, utilization, and performance. It analyzes various types of data, including usage, wear, and conditional characteristics from disparate sources, and detects failure patterns and poor quality parts earlier than traditional quality control methods. The goal is to reduce unscheduled asset downtime and ensure that quality metrics are achieved or exceeded. The product sends those insights, which are combined with your institutional knowledge, to provide optimized recommended decisions to people and systems. With IBM Predictive Maintenance and Quality, organizations can better optimize operations and supply chain processes, resulting in better quality products, higher profit margins, and competitive advantage.

Here are some key features of IBM Predictive Maintenance and Quality:

- ▶ Real-time capabilities: Integrate, manage, and analyze sensor and real-time information in combination with existing static data.
- ▶ Big data, predictive analytics, and business intelligence: Combine predictive modeling, decision management, workflows, and dashboards, and early warning algorithms in coordination with all types and volumes of data.
- ▶ Open architecture and data integration: Link to many systems and data sources with ready for use connectors and APIs.
- ▶ Process integration: Deliver insights and recommendations to and run work orders in existing Enterprise Asset Management (EAM) systems.

Finance and human resources

Financial analysis solutions enable analysts to create and maintain complex models of business structures, dimensions, and data sets to provide more insights into opportunities and risks.

IBM solutions for financial analysis help your finance team identify the drivers of profitability and performance. They deliver the insight that you need to make smarter decisions about revenue, profit, cash flow, and the full range of variables affecting your financial performance. Essential tasks such as variance analysis, scenario modeling, and what-if analysis are easier and faster with financial analysis solutions from IBM.

Financial analysis software from IBM helps your finance team to perform the following actions:

- ▶ Create and maintain complex, multi-dimensional models of business structures, dimensions, and data sets.
- ▶ Examine historical performance and compare it to current and forecasted performance results, then modify assumptions to test plans, budgets, and forecasts.
- ▶ Analyze profitability by product, customer, channel, region, and more to gain new insights into opportunities and risks.
- ▶ Identify the actions that are needed to better align financial and operational resources so that resources can be shifted to the most profitable areas of the business.

New business models

IBM SPSS Analytic Catalyst makes analysis and discovery of big data more accessible to business users by presenting analyses visually and by using plain language summaries.

IBM SPSS Analytic Catalyst uses the power of SPSS Analytics Server to help accelerate analytics by identifying key drivers from big data. It automates portions of data preparation, automatically interpreting results, and presenting analyses in interactive visuals with plain language summaries. The result? Statistical analysis and discovery of big data are all more accessible to business users.

SPSS Analytic Catalyst offers the following features and benefits:

- ▶ Automated key driver identification with sophisticated algorithms, automatic testing, and regression-based techniques.
- ▶ Interactive visuals and plain-language summaries of predictive analytics findings that provide insights at a glance, supporting explanations and statistical details.
- ▶ Accelerated predictive analytics in big data environments with field associations, decision trees, drill down, and functions for saving insights for later retrieval.
- ▶ Distribution in an environment that is designed for big data and massive scale.

IBM IT solutions

With IBM IT solutions, such as IBM Business Intelligence, IBM Cognos Insight, IBM Cognos Express, IBM SPSS Statistics, IBM SPSS Modeler, and IBM Social Media Analytics, you can accomplish the following tasks:

- ▶ Maximize insights, ensure trust, and improve IT economics.
- ▶ Harness and analyze all data, even real-time data streaming from the sensors and devices that make up the Internet of Things.
- ▶ Ensure the privacy and security of that data, and put in place the infrastructure to support advanced analytics.
- ▶ Take advantage of cloud-based services to accelerate innovation.

Cloud services

As you expand big data and analytics capabilities throughout your organization, you must empower all of your business users to access and analyze data for faster insight. Taking delivery of software, solutions, infrastructure, platforms, and services on the cloud can accelerate the value of big data and analytics capabilities, offering scalability with limited upfront investment. You can see the current product list here:

- ▶ Big data and analytics software-as-a-service (SaaS)
- ▶ Business process-as-a-service (BPaaS)

- Infrastructure as a service (IaaS)
- IBM Bluemix™ cloud platform

For business decisions

IBM IT solutions provide the robust big data and analytics capabilities that you can use to capture, analyze, and act on all relevant data, including the stream of data that is generated by a myriad of electronic devices. Marketers, sales managers, financial analysts, and other business users gain the insights that they need to act in real time based on timely, trusted data. Here you can see the capabilities:

- Analyze streaming data in real time as it flows through the organization.
- Make sense of unstructured data and put it into context with historical, structured data.
- Use predictive analytics and advanced algorithms to recommend actions in real time.
- Empower decision makers to act on insights in the moment, with confidence.

Here are a couple of IBM business solutions:

- IBM SPSS Modeler Gold
- IBM InfoSphere Streams

For governance and security with trusted data

To enable decision makers to act with confidence, you must ensure that the data they use is clean, timely, and accurate. There are two different areas:

- Information Integration and Governance
 - This area can help your organization understand information and analyze the data and its relationships.
 - Improve information with delivering accurate and current data.
 - Accelerate projects providing consistent information in time.
- IBM Security Intelligence with Big Data, which is shown in Figure 3-4.

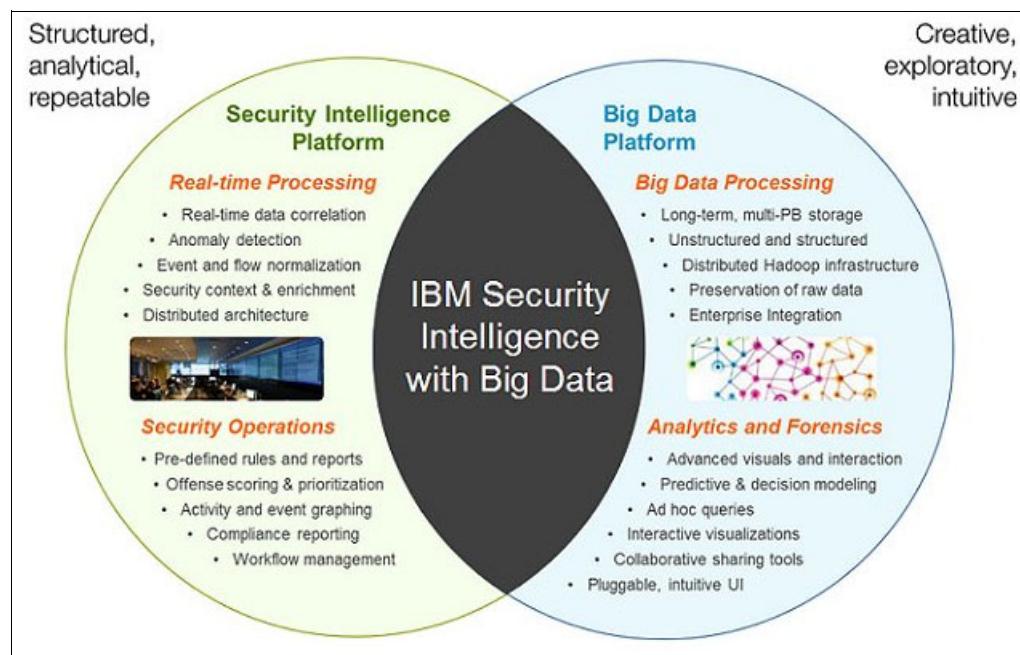


Figure 3-4 IBM Security Intelligence with Big Data

IBM Security Intelligence with Big Data combines the real-time security visibility of the IBM QRadar® Security Intelligence Platform with the custom analytics of the IBM Big Data Platform. Here are its key capabilities:

- Real-time correlation and anomaly detection of diverse security data.
- High-speed querying of security intelligence data.
- Flexible big data analytics across structured and unstructured data, including security data, email, document, and social media content, full packet capture data, business process data, and other information.
- Graphical front-end tool for visualizing and exploring big data.
- Forensics for deep visibility.

The infrastructure to maximize insights

A comprehensive big data and analytics platform like IBM Watson™ Foundations needs the support of an infrastructure that takes advantage of technologies like Hadoop to gain insights from streaming data and data at rest. Integrated, high-performance systems, whether deployed on premises or on the cloud, can reduce IT complexity and enable your organization to infuse analytics everywhere. The solutions can be seen here:

- ▶ IBM Solution for Hadoop – Power Systems Edition
- ▶ IBM BLU Acceleration Solution – Power Systems Edition
- ▶ IBM Solution for Analytics – Power Systems Edition

3.2.4 IBM Big Data analytics advantage

IBM Platform Symphony provides the following advantages for your big data analytics applications:

- ▶ Policy-driven workload scheduler for better granularity and control
- ▶ Several instances of Hadoop, other programs, or both on a single shared cluster
- ▶ Distributed runtime engine support for high resource availability
- ▶ Flexibility from open architecture for application development and choice of file system
- ▶ Higher application performance for IBM InfoSphere BigInsights workloads
- ▶ Rolling software upgrades to keep applications running

3.3 Why use an IBM Risk Analytics solutions

IBM Risk Analytics solutions can help you balance risk and opportunity, and make more informed decisions based on risk analysis.

IBM Risk Analytics solutions enable the world's most successful companies to make risk-aware decisions through smarter enterprise risk management programs and methodologies, which drives business performance and better outcomes. The combined risk management capabilities that are described in 3.3.1, "IBM Algorithmics software" on page 25 and 3.3.2, "IBM OpenPages software" on page 25 can help your company achieve profitable growth and address increasing demands for regulatory compliance in today's volatile and complex market conditions.

With IBM Risk Analytics solutions, you can improve your decision making by providing risk analysis and reduce the cost of regulatory compliance.

3.3.1 IBM Algorithmics software

IBM Algorithmics® software enables financial institutions and corporate treasuries to make risk-aware business decisions. Supported by a global team of risk experts that are based in all major financial centers, IBM Algorithmics products and solutions address market, credit, and liquidity risk, and collateral and capital management.

Here are the featured products:

- ▶ IBM Algo® Asset Liability Management
- ▶ IBM Algo Collateral Management
- ▶ IBM Algo Risk® Service on Cloud
- ▶ IBM Algo Risk

3.3.2 IBM OpenPages software

IBM OpenPages® Operational Risk Management automates the process of identifying, analyzing, and managing operational risk and enables businesses to integrate risk data into a single environment. This integrated approach helps improve visibility into risk exposure, reduce loss, and improve business performance. You can use OpenPages Operational Risk Management to embed operational risk management practices into the corporate culture, making procedures more effective and efficient. Your organization can use OpenPages GRC software to manage enterprise operational risk and compliance initiatives by using a single, integrated solution.

Here are the featured products:

- ▶ IBM OpenPages GRC on Cloud
- ▶ IBM OpenPages GRC Platform
- ▶ IBM OpenPages Operational Risk Management
- ▶ IBM OpenPages Policy and Compliance Management
- ▶ IBM OpenPages Financial Controls Management
- ▶ IBM OpenPages IT Governance
- ▶ IBM OpenPages Internal Audit Management

3.4 Scenario for minimizing risk and building a better model

This section shows how you can minimize the risk of money loss with IBM Algo Market® Risk if you want to open a new retail store. Before the scenario itself is described in 3.4.3, “Scenario” on page 26, there is a brief description about two other topics in 3.4.1, “Algo Market Risk” on page 26 and 3.4.2, “IBM SPSS Statistics: Monte Carlo simulation” on page 26.

3.4.1 Algo Market Risk

Algo Market Risk is a scenario-based solution that helps measure and manage market risk. Its Monte Carlo simulations of mark-to-market valuations allow banks and financial institutions to reduce regulatory capital requirements and increase their return on capital.

Here are some of the Algo Market Risk features:

- ▶ Advanced analytics and risk reporting delivers the highly accurate risk insights that are needed to help banks reduce their regulatory capital.
- ▶ Comprehensive instrument coverage spans 20 geographic markets and 400 financial products.
- ▶ Scenario-based portfolio optimization supports proactive, risk-informed decision making.
- ▶ Advanced computational speed integrates the front and middle office for active management of risk.
- ▶ Customizable and scalable analytics support the evolving needs of the enterprise.

3.4.2 IBM SPSS Statistics: Monte Carlo simulation

SPSS Statistics combines the power of predictive analytics with the what-if capabilities of a Monte Carlo simulation to help you accomplish the following tasks:

- ▶ Go beyond conventional what-if analysis: Explore hundreds or thousands of combinations of factors and analyze all possible outcomes for more accurate results.
- ▶ Identify the factors with the most impact: Quickly identify the factors in your model with the greatest impact on business outcomes.
- ▶ Gain competitive advantage: Knowing what is likely to happen next enables you to offer the correct products, target the correct customers, or gain other advantages over competitors who lack this insight.
- ▶ Achieve better outcomes: Because you can predict results accurately, you can adjust your business strategies and processes to help you make the correct decisions quickly and further reduce risk.

3.4.3 Scenario

In this scenario, assume that you have some existing retail data that is based on other stores and you want to use it as a starting point.

You want to know what is the likelihood that you can reach your target number in the first few months, for example, 7.5 million dollars. You use the available data to build a model that includes the following information:

- ▶ Advertising budget
- ▶ Customer confidence index
- ▶ Sales agents number
- ▶ Monthly store visits by an individual
- ▶ Previous months income

You need to analyze the data and run a simulation. Assume that you have an existing simulation plan and you use it.

You are interested in the effect of advertising, and how it can decrease the risk of loss. You set fix values, for example, \$50,000.

The simulation run thousands of times and at the end you can see the result. Based on the given parameters and the \$50,000 advertisement money, you have a 52% chance to make your numbers. With this result, you cannot convince management, so you need to tweak the outcome. You need to run the simulation again.

After you set the new budget for advertising to \$70,000 and run the simulation, you have a 68% chance, which is much more comfortable. If you increase the budget for advertising, then you can decrease the risk factor. In this way, you can maximize the chance to reach your target numbers.

For a video about this scenario, go to this website:

https://www.youtube.com/watch?v=L_8VV1yXEjc



IBM Spectrum Scale (formerly GPFS)

This chapter focuses on IBM Spectrum Scale (Spectrum Scale) for use specifically with the IBM Platform Computing stack. For more information about Spectrum Scale, see *IBM Spectrum Scale (formerly GPFS)*, SG24-8254.

This chapter covers the following topics:

- ▶ IBM Spectrum Scale overview
- ▶ Spectrum Scale for technical computing
- ▶ Spectrum Scale for big data
- ▶ Installing IBM Spectrum Scale

4.1 IBM Spectrum Scale overview

Spectrum Scale is a fully parallel software-defined storage that can manage data, objects, and files (based on GPFS technology). Spectrum Scale is largely scalable and can reach great performance. Some clients can get a throughput as large as 400 Gbps.

Spectrum Scale can reach this higher throughput by using parallelism in all layers. So, a client does not write to a single server but to as many servers as you have configured in the cluster. These servers might access different storages so that one input/output in one server does not impact the other I/O because they are independent point to point, as you can see in Figure 4-1. This is one of the many reasons that make such high transfer rates possible.

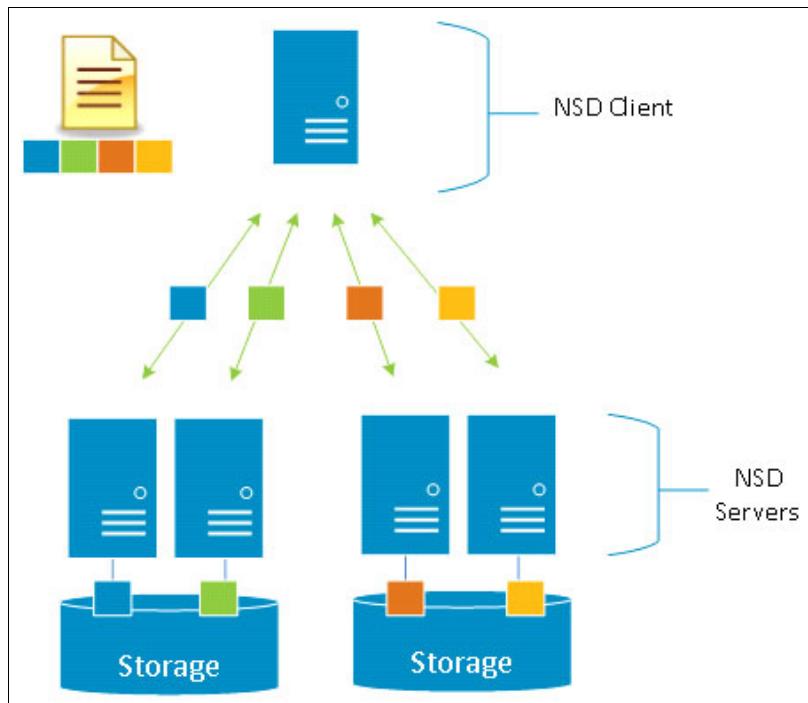


Figure 4-1 Elastic Storage parallel access

A summary of the properties and capabilities of Spectrum Scale is shown in Table 4-1.

Table 4-1 IBM Spectrum Scale properties

Elastic Storage capability	Limits
Number of files	Maximum 2^{64} files
File system size	Maximum 2^{99} bytes
Number of nodes	Maximum 16384 nodes
Servers	Add and remove online
Storage	Add and remove online
Snapshots	256 per file system and independent file set
Information Lifecycle Management	Disks and external tape pools available

Spectrum Scale is available in three different licensing options, which enable different capabilities. To decide which option is the best for your needs, see the different capabilities for each license type, as shown in Table 4-2.

Table 4-2 Licensing capabilities

Capabilities	Express Edition	Standard Edition	Advanced Edition
Facilitate data sharing with a global namespace, simplified management at scale (massively scalable file system, quotas, and snapshots), and data integrity and availability.	X	X	X
Create optimized tiered storage pools by grouping disks based on performance, locality, or cost characteristics.		X	X
Simplify data management at scale with Information Lifecycle Management (ILM) tools that include Backup and Recovery and policy-based archiving to a low-cost storage pool.		X	X
Enable worldwide data access and empower global collaboration with Active File Manager (AFM).		X	X
Provide scalable file service with simultaneous access to a common set of data from multiple servers with Clustered NFS (cNFS).		X	X
Protect data with native encryption and secure erase, NIST compliant and FIPS certified.			X

Another possibility is to buy the ready solution that was first named IBM GPFS Storage Server and now has been rebranded as IBM Elastic Storage™ Server (ESS). This solution already comes with servers, disks, and software to provide the entire stack to provide fast solutions.

4.2 Spectrum Scale for technical computing

Spectrum Scale is perfect for technical computing because it is built for fast and parallel data access.

Spectrum Scale removes data-related bottlenecks by providing parallel access to data, eliminating single filer choke points or hot spots. It removes single file locks problems by making the lock per block not at the file level, which compared to NFS is a great improvement.

Spectrum Scale also simplifies data management at scale by providing a single namespace that can be scaled simply, quickly, and infinitely (because it is not possible to buy 2^{99} bytes or more than one billion petabytes of disk capacity) by simply adding more scale-out resources, such as storage and servers.

Data lifecycle automation bridges the gaping data growth / budget chasm, bringing storage costs into line. When integrated with IBM Spectrum Protect or IBM Spectrum Archive, Spectrum Scale can uniquely manage the full data lifecycle, delivering geometrically lower-cost savings through policy-driven automation and tiered storage management. It addresses data growth by reducing your storage costs up to 90% while providing world-class reliability, scalability, and availability for technical computing data. This situation occurs because most of the data is old data that can be sent to tape, reducing dramatically the cost of maintaining data, devices, power, and cooling.

The Active File Management (AFM) feature enables Spectrum Scale to cache asynchronously data, which provides fast access to disks (low-latency writes) even over distance. Many customers are using this technology, as described in 4.2.1, “Argonne Leadership Computing Facility” on page 32.

The same technology empowers geographically distributed organizations by expanding a single global namespace literally to a global scale by placing critical data close to everyone and everything that needs it, no matter where they are in the world. Speeding data access to stakeholders around the world accelerates schedules and improves productivity.

4.2.1 Argonne Leadership Computing Facility

A good example for Spectrum Scale implementation for technical computing can be seen at the Argonne Leadership Computing Facility. They have two primary storages, one with 20 PB and the other with 7 PB both, which are being cached on a new IBM ESS delivering 400 GBps in a 13 PB file system for technical computing. For more information about this case, go to the following website:

<http://www.alcf.anl.gov/articles/alcf-storage-upgrade-aims-hands-data-management>

4.2.2 Jülich Supercomputing Centre

Another good example is the Jülich Supercomputing Centre (JSC), which uses ILM to migrate their data to tape storage by using Spectrum Protect HSM. For more information, go to the following websites:

- ▶ <http://www.fz-juelich.de/ias/jsc/EN/Expertise/Datamanagement/OnlineStorage/JUST/Configuration/Configuration.html>
- ▶ http://www.fz-juelich.de/ias/jsc/EN/Expertise/Datamanagement/OnlineStorage/JUST/Filesystems/JUST_filesystems_node.html

4.2.3 IBM Elastic Storage Server

For clients looking for a ready solution incorporating disk, server, and software, the IBM Elastic Storage Server (ESS) is a great option. This solution comes with a graphical user interface (GUI) to ease the use of the solution.

The Elastic Storage Server provides unsurpassed end-to-end data availability, reliability, and integrity with unique technologies, including IBM Spectrum Scale RAID, which uses advanced erasure coding to avoid the painful multiday rebuild times that are common with today's multi-terabyte drives, in addition to being able to withstand multiple device failures instead of the one or two failures conventional systems can withstand.

To read more about how fast and efficient Spectrum Scale RAID can work, see the documentation at the following website:

<http://ibm.co/1CmI03A>

4.3 Spectrum Scale for big data

As described in Chapter 2, “Technical computing software portfolio” on page 9, MapReduce tasks can be used by technical computing too. Another great feature of Spectrum Scale that enables the enhancement of the Hadoop open framework for MapReduce tasks is the File Placement Optimizer (FPO). This feature provides location affinity for the resource scheduling queue for better allocating processes.

FPO can be added transparently to Hadoop configurations by adding the libraries and changing configurations files to point to the new Spectrum Scale mount point.

4.4 Installing IBM Spectrum Scale

This section describes how to install Spectrum Scale.

4.4.1 Introducing IBM Spectrum Scale

Spectrum Scale is a cluster file system, which means that it provides concurrent access to a single file system or set of file systems from multiple nodes. These nodes all can be SAN-attached or a mix of SAN- and network-attached. This setup enables high performance access to this common set of data to support a scale-out solution or provide a high availability platform.

Spectrum Scale has many features beyond common data access, including data replication, policy-based storage management, and multi-site operations. You can create a Spectrum Scale cluster of IBM AIX® nodes, Linux nodes, Windows server nodes, or a mix of all three. Spectrum Scale can run on virtualized instances that provide common data access in environments, use logical partitioning, or other hypervisors. Multiple Spectrum Scale clusters can share data within a location or across wide area network (WAN) connections.

4.4.2 The strengths of Spectrum Scale

Spectrum Scale provides a global namespace, shared file system access among Spectrum Scale clusters, simultaneous file access from multiple nodes, high recoverability, and data availability through replication, the ability to make changes while a file system is mounted, and simplified administration, even in large environments.

4.4.3 Preparing the environment on Linux nodes

Before proceeding with the installation, prepare your environment by completing the following steps:

1. Add the Spectrum Scale bin directory to your shell PATH.

Ensure that the PATH environment variable for the root user on each node includes /usr/lpp/mmfs/bin.

2. Accept the electronic license agreement.

The Spectrum Scale software license agreement is shipped with the Spectrum Scale software and is viewable electronically. When you extract the Spectrum Scale software, you are asked whether you accept the license. The electronic license agreement must be accepted before software installation can continue. Read the software agreement carefully before you accept the license. See Example 4-1.

Example 4-1 Spectrum Scale files extraction and license acceptance

```
[root@pw4302-13 gpfsinstall]# ./gpfs_install-4.1.0-0_x86_64 --text-only

Extracting License Acceptance Process Tool to /usr/lpp/mmfs/4.1 ...
tail -n +456 ./gpfs_install-4.1.0-0_x86_64 | /bin/tar -C /usr/lpp/mmfs/4.1 -xvz
--exclub 2> /dev/null 1> /dev/null

Installing JRE ...
tail -n +456 ./gpfs_install-4.1.0-0_x86_64 | /bin/tar -C /usr/lpp/mmfs/4.1
--wildcards - /dev/null

Invoking License Acceptance Process Tool ...
/usr/lpp/mmfs/4.1/ibm-java-x86_64-60/jre/bin/java -cp
/usr/lpp/mmfs/4.1/LAP_HOME/LAPApp.lpp/mmfs/4.1/LA_HOME -m /usr/lpp/mmfs/4.1 -s
/usr/lpp/mmfs/4.1 -text_only
International Program License Agreement
```

Part 1 - General Terms

BY DOWNLOADING, INSTALLING, COPYING, ACCESSING, CLICKING ON AN "ACCEPT" BUTTON, OR OTHERWISE USING THE PROGRAM, LICENSEE AGREES TO THE TERMS OF THIS AGREEMENT. IF YOU ARE ACCEPTING THESE TERMS ON BEHALF OF LICENSEE, YOU REPRESENT AND WARRANT THAT YOU HAVE FULL AUTHORITY TO BIND LICENSEE TO THESE TERMS. IF YOU DO NOT AGREE TO THESE TERMS,

* DO NOT DOWNLOAD, INSTALL, COPY, ACCESS, CLICK ON AN "ACCEPT" BUTTON, OR USE THE PROGRAM; AND

* PROMPTLY RETURN THE UNUSED MEDIA, DOCUMENTATION, AND

Press Enter to continue viewing the license agreement, or enter "1" to accept the agreement, "2" to decline it, "3" to print it, "4" to read non-IBM terms, or "99" to go back to the previous screen.

1

License Agreement Terms accepted.

```

Extracting Product RPMs to /usr/lpp/mmfs/4.1 ...
tail -n +456 ./gpfs_install-4.1.0-0_x86_64 | /bin/tar -C /usr/lpp/mmfs/4.1
--wildcards - ./gpfs.base_4.1.0-0_amd64.deb ./gpfs.docs-4.1.0-0.noarch.rpm
./gpfs.docs_4.1.0-0_all.depfs.ext_4.1.0-0_amd64.deb ./gpfs.gpl-4.1.0-0.noarch.rpm
./gpfs.gpl_4.1.0-0_all.deb ./gpfs.gskit_8.0.50-16_amd64.deb
./gpfs.msg.en-us_4.1.0-0_all.deb ./gpfs.msg.en_US-4.1.0-0.noa

- gpfs.base-4.1.0-0.x86_64.rpm
- gpfs.base_4.1.0-0_amd64.deb
- gpfs.docs-4.1.0-0.noarch.rpm
- gpfs.docs_4.1.0-0_all.deb
- gpfs.ext-4.1.0-0.x86_64.rpm
- gpfs.ext_4.1.0-0_amd64.deb
- gpfs.gpl-4.1.0-0.noarch.rpm
- gpfs.gpl_4.1.0-0_all.deb
- gpfs.gskit-8.0.50-16.x86_64.rpm
- gpfs.gskit_8.0.50-16_amd64.deb
- gpfs.msg.en-us_4.1.0-0_all.deb
- gpfs.msg.en_US-4.1.0-0.noarch.rpm

Removing License Acceptance Process Tool from /usr/lpp/mmfs/4.1 ...
rm -rf /usr/lpp/mmfs/4.1/LAP_HOME /usr/lpp/mmfs/4.1/LA_HOME

Removing JRE from /usr/lpp/mmfs/4.1 ...
rm -rf /usr/lpp/mmfs/4.1/ibm-java*tgz

```

3. Check that the software package management for your operation system is correctly configured and depending upon how you installed the base operation system, you must install other packages to solve possible dependencies.
4. Clock synchronization

The clocks of all nodes in the Spectrum Scale cluster must be synchronized. If this is not done, NFS access to the data, and other Spectrum Scale file system operations, might be disrupted.

The installer creates a set of .rpm and .deb files that you use to continue the installation. The files are created in the /usr/lpp/mmfs/4.1/ directory, as shown in Example 4-2.

Example 4-2 List of packages that are extracted

```
[root@pw4302-13 4.1]# pwd
/usr/lpp/mmfs/4.1
[root@pw4302-13 4.1]# ls -l
total 41388
-rw-r--r-- 1 root root 14215820 Apr 25 2014 gpfs.base_4.1.0-0_amd64.deb
-rw-r--r-- 1 root root 14482662 Apr 25 2014 gpfs.base-4.1.0-0.x86_64.rpm
-rw-r--r-- 1 root root 271026 Apr 25 2014 gpfs.docs_4.1.0-0_all.deb
-rw-r--r-- 1 root root 292465 Apr 25 2014 gpfs.docs-4.1.0-0.noarch.rpm
-rw-r--r-- 1 root root 1541376 Apr 25 2014 gpfs.ext_4.1.0-0_amd64.deb
-rw-r--r-- 1 root root 1548454 Apr 25 2014 gpfs.ext-4.1.0-0.x86_64.rpm
-rw-r--r-- 1 root root 546506 Apr 25 2014 gpfs.gpl_4.1.0-0_all.deb
-rw-r--r-- 1 root root 573838 Apr 25 2014 gpfs.gpl-4.1.0-0.noarch.rpm
-rw-r--r-- 1 root root 4287554 Apr 25 2014 gpfs.gskit_8.0.50-16_amd64.deb
-rw-r--r-- 1 root root 4328387 Apr 25 2014 gpfs.gskit-8.0.50-16.x86_64.rpm
```

```
-rw-r--r-- 1 root root 128728 Apr 25 2014 gpfs.msg.en-us_4.1.0-0_all.deb  
-rw-r--r-- 1 root root 131514 Apr 25 2014 gpfs.msg.en_US-4.1.0-0.noarch.rpm  
drwxr-xr-x 3 root root 4096 Nov 27 16:37 license  
[root@pw4302-13 4.1]#
```

Note: The **--text-only** parameter is necessary only when you do not have a GUI.

Now, you can use the system installer depending on which system you are. Example 4-3 shows the use of **yum** to continue the installation.

Example 4-3 Perform the installation by using yum

```
[root@pw4302-13 4.1]# yum install gpfs.*  
Loaded plugins: product-id, refresh-packagekit, security, subscription-manager  
rhel6.5  
xCAT-rhels6.5-path0  
xCAT-rhels6.5-path1  
xCAT-rhels6.5-path2  
xCAT-rhels6.5-path3  
xCAT-rhels6.5-path4  
xCAT-rhels6.5-path5  
xcat-otherpkgs0  
Setting up Install Process  
Examining gpfs.base-4.1.0-0.x86_64.rpm: gpfs.base-4.1.0-0.x86_64  
Marking gpfs.base-4.1.0-0.x86_64.rpm to be installed  
Examining gpfs.docs-4.1.0-0.noarch.rpm: gpfs.docs-4.1.0-0.noarch  
Marking gpfs.docs-4.1.0-0.noarch.rpm to be installed  
Examining gpfs.ext-4.1.0-0.x86_64.rpm: gpfs.ext-4.1.0-0.x86_64  
Marking gpfs.ext-4.1.0-0.x86_64.rpm to be installed  
Examining gpfs.gpl-4.1.0-0.noarch.rpm: gpfs.gpl-4.1.0-0.noarch  
Marking gpfs.gpl-4.1.0-0.noarch.rpm to be installed  
Examining gpfs.gskit-8.0.50-16.x86_64.rpm: gpfs.gskit-8.0.50-16.x86_64  
Marking gpfs.gskit-8.0.50-16.x86_64.rpm to be installed  
Examining gpfs.msg.en_US-4.1.0-0.noarch.rpm: gpfs.msg.en_US-4.1.0-0.noarch  
Marking gpfs.msg.en_US-4.1.0-0.noarch.rpm to be installed  
Resolving Dependencies  
--> Running transaction check  
--> Package gpfs.base.x86_64 0:4.1.0-0 will be installed  
--> Package gpfs.docs.noarch 0:4.1.0-0 will be installed  
--> Package gpfs.ext.x86_64 0:4.1.0-0 will be installed  
--> Package gpfs.gpl.noarch 0:4.1.0-0 will be installed  
--> Package gpfs.gskit.x86_64 0:8.0.50-16 will be installed  
--> Package gpfs.msg.en_US.noarch 0:4.1.0-0 will be installed  
--> Finished Dependency Resolution  
  
Dependencies Resolved  
  
=====  
           Package          Arch        Version  
Repository  
=====  
Installing:  
  gpfs.base           x86_64      4.1.0-0  
/gpfs.base-4.1.
```

```

gpfs.docs          noarch      4.1.0-0
/gpfs.docs-4.1.
gpfs.ext          x86_64     4.1.0-0
/gpfs.ext-4.1.0
gpfs.gpl          noarch      4.1.0-0
/gpfs.gpl-4.1.0
gpfs.gskit        x86_64     8.0.50-16
/gpfs.gskit-8.0
gpfs.msg.en_US    noarch      4.1.0-0
/gpfs.msg.en_US

Transaction Summary
=====
Install      6 Package(s)

Total size: 61 M
Installed size: 61 M
Is this ok [y/N]: y
Downloading Packages:
Running rpm_check_debug
Running Transaction Test
Transaction Test Succeeded
Running Transaction
  Installing : gpfs.base-4.1.0-0.x86_64
  Installing : gpfs.ext-4.1.0-0.x86_64
  Installing : gpfs.gpl-4.1.0-0.noarch
  Installing : gpfs.msg.en_US-4.1.0-0.noarch
  Installing : gpfs.docs-4.1.0-0.noarch
  Installing : gpfs.gskit-8.0.50-16.x86_64
  Verifying   : gpfs.base-4.1.0-0.x86_64
  Verifying   : gpfs.gskit-8.0.50-16.x86_64
  Verifying   : gpfs.docs-4.1.0-0.noarch
  Verifying   : gpfs.ext-4.1.0-0.x86_64
  Verifying   : gpfs.gpl-4.1.0-0.noarch
  Verifying   : gpfs.msg.en_US-4.1.0-0.noarch

Installed:
  gpfs.base.x86_64 0:4.1.0-0      gpfs.docs.noarch 0:4.1.0-0
  gpfs.ext.x86_64 0:4.
  gpfs.gskit.x86_64 0:8.0.50-16  gpfs.msg.en_US.noarch 0:4.1.0-0

Complete!
[root@pw4302-13 4.1]#

```

After the Spectrum Scale base version is installed, verify whether there are any updates that are available for this product. At the time of writing, the Spectrum Scale base version is Version 4.1.0.0 and there is an update to Version 4.1.0.4, as shown in Example 4-4.

Example 4-4 Updated packages available

```
[root@pw4302-13 GPFS_update]# ls -l
total 43084
-rw-r--r-- 1 root  root    2286 Oct 31 14:25 changelog
-rw-r--r-- 1 30007 bin  14484522 Oct 28 17:27 gpfs.base_4.1.0-4_amd64_update.deb
-rw-r--r-- 1 30007 bin  14752057 Oct 28 17:23 gpfs.base-4.1.0-4.x86_64.update.rpm
-rw-r--r-- 1 30007 bin   277934 Oct 28 17:27 gpfs.docs_4.1.0-4_all.deb
```

```

-rw-r--r-- 1 30007 bin      298337 Oct 28 17:22 gpfs.docs-4.1.0-4.noarch.rpm
-rw-r--r-- 1 30007 bin      2025372 Oct 28 17:27 gpfs.ext_4.1.0-4_amd64_update.deb
-rw-r--r-- 1 30007 bin      2037264 Oct 28 17:25 gpfs.ext-4.1.0-4.x86_64.update.rpm
-rw-r--r-- 1 30007 bin      562812 Oct 28 17:27 gpfs.gpl_4.1.0-4_all.deb
-rw-r--r-- 1 30007 bin      588752 Oct 28 17:25 gpfs.gpl-4.1.0-4.noarch.rpm
-rw-r--r-- 1 30007 bin      4369504 Oct 28 17:27 gpfs.gskit_8.0.50-32_amd64.deb
-rw-r--r-- 1 30007 bin      4407428 Oct 28 17:22 gpfs.gskit-8.0.50-32.x86_64.rpm
-rw-r--r-- 1 30007 bin      136850 Oct 28 17:27 gpfs.msg.en-us_4.1.0-4_all.deb
-rw-r--r-- 1 30007 bin      139569 Oct 28 17:22 gpfs.msg.en_US-4.1.0-4.noarch.rpm
-rw-r--r-- 1 root  root     7905 Oct 20 14:14 README
[root@pw4302-13 GPFS_update]#

```

Install the updates for your system, as shown in Example 4-5.

Example 4-5 Install the updates

```

[root@pw4302-13 GPFS_update]# yum update *.rpm
Loaded plugins: product-id, refresh-packagekit, security, subscription-manager
Setting up Update Process
Examining gpfs.base-4.1.0-4.x86_64.update.rpm: gpfs.base-4.1.0-4.x86_64
Marking gpfs.base-4.1.0-4.x86_64.update.rpm as an update to
gpfs.base-4.1.0-0.x86_64
Examining gpfs.docs-4.1.0-4.noarch.rpm: gpfs.docs-4.1.0-4.noarch
Marking gpfs.docs-4.1.0-4.noarch.rpm as an update to gpfs.docs-4.1.0-0.noarch
Examining gpfs.ext-4.1.0-4.x86_64.update.rpm: gpfs.ext-4.1.0-4.x86_64
Marking gpfs.ext-4.1.0-4.x86_64.update.rpm as an update to gpfs.ext-4.1.0-0.x86_64
Examining gpfs.gpl-4.1.0-4.noarch.rpm: gpfs.gpl-4.1.0-4.noarch
Marking gpfs.gpl-4.1.0-4.noarch.rpm as an update to gpfs.gpl-4.1.0-0.noarch
Examining gpfs.gskit-8.0.50-32.x86_64.rpm: gpfs.gskit-8.0.50-32.x86_64
Marking gpfs.gskit-8.0.50-32.x86_64.rpm as an update to
gpfs.gskit-8.0.50-16.x86_64
Examining gpfs.msg.en_US-4.1.0-4.noarch.rpm: gpfs.msg.en_US-4.1.0-4.noarch
Marking gpfs.msg.en_US-4.1.0-4.noarch.rpm as an update to
gpfs.msg.en_US-4.1.0-0.noarch
Resolving Dependencies
--> Running transaction check
--> Package gpfs.base.x86_64 0:4.1.0-0 will be updated
--> Package gpfs.base.x86_64 0:4.1.0-4 will be an update
--> Package gpfs.docs.noarch 0:4.1.0-0 will be updated
--> Package gpfs.docs.noarch 0:4.1.0-4 will be an update
--> Package gpfs.ext.x86_64 0:4.1.0-0 will be updated
--> Package gpfs.ext.x86_64 0:4.1.0-4 will be an update
--> Package gpfs.gpl.noarch 0:4.1.0-0 will be updated
--> Package gpfs.gpl.noarch 0:4.1.0-4 will be an update
--> Package gpfs.gskit.x86_64 0:8.0.50-16 will be updated
--> Package gpfs.gskit.x86_64 0:8.0.50-32 will be an update
--> Package gpfs.msg.en_US.noarch 0:4.1.0-0 will be updated
--> Package gpfs.msg.en_US.noarch 0:4.1.0-4 will be an update
--> Finished Dependency Resolution

Dependencies Resolved

=====
Package           Arch        Version       Repository
=====
Updating:

```

gpfs.base	x86_64	4.1.0-4
/gpfs.base-4.1.0		
gpfs.docs	noarch	4.1.0-4
/gpfs.docs-4.1.0		
gpfs.ext	x86_64	4.1.0-4
/gpfs.ext-4.1.0-		
gpfs.gpl	noarch	4.1.0-4
/gpfs.gpl-4.1.0-		
gpfs.gskit	x86_64	8.0.50-32
/gpfs.gskit-8.0.		
gpfs.msg.en_US	noarch	4.1.0-4
/gpfs.msg.en_US-		

Transaction Summary

=====
 Upgrade 6 Package(s)

Total size: 64 M

Is this ok [y/N]: y

Downloading Packages:

Running rpm_check_debug

Running Transaction Test

Transaction Test Succeeded

Running Transaction

Updating	:	gpfs.base-4.1.0-4.x86_64
Updating	:	gpfs.gpl-4.1.0-4.noarch
Updating	:	gpfs.ext-4.1.0-4.x86_64
Updating	:	gpfs.msg.en_US-4.1.0-4.noarch
Updating	:	gpfs.docs-4.1.0-4.noarch
Updating	:	gpfs.gskit-8.0.50-32.x86_64
Cleanup	:	gpfs.gpl-4.1.0-0.noarch
Cleanup	:	gpfs.msg.en_US-4.1.0-0.noarch
Cleanup	:	gpfs.docs-4.1.0-0.noarch
Cleanup	:	gpfs.ext-4.1.0-0.x86_64
Cleanup	:	gpfs.base-4.1.0-0.x86_64
Cleanup	:	gpfs.gskit-8.0.50-16.x86_64
Verifying	:	gpfs.gpl-4.1.0-4.noarch
Verifying	:	gpfs.base-4.1.0-4.x86_64
Verifying	:	gpfs.gskit-8.0.50-32.x86_64
Verifying	:	gpfs.docs-4.1.0-4.noarch
Verifying	:	gpfs.ext-4.1.0-4.x86_64
Verifying	:	gpfs.msg.en_US-4.1.0-4.noarch
Verifying	:	gpfs.base-4.1.0-0.x86_64
Verifying	:	gpfs.gskit-8.0.50-16.x86_64
Verifying	:	gpfs.docs-4.1.0-0.noarch
Verifying	:	gpfs.ext-4.1.0-0.x86_64
Verifying	:	gpfs.msg.en_US-4.1.0-0.noarch
Verifying	:	gpfs.gpl-4.1.0-0.noarch

Updated:

 gpfs.base.x86_64 0:4.1.0-4 gpfs.docs.noarch 0:4.1.0-4

 gpfs.ext.x86_64 0:4.

 gpfs.gskit.x86_64 0:8.0.50-32 gpfs.msg.en_US.noarch 0:4.1.0-4

Complete!

```
[root@pw4302-13 GPFS_update]#
```

To check whether the Spectrum Scale package was successfully installed, run the command that is shown in Example 4-6.

Example 4-6 Check for installed packages

```
[root@pw4302-13 ~]# rpm -qa | grep gpfs
gpfs.gskit-8.0.50-32.x86_64
gpfs.base-4.1.0-4.x86_64
gpfs.msg.en_US-4.1.0-4.noarch
gpfs.ext-4.1.0-4.x86_64
gpfs.gpl-4.1.0-4.noarch
gpfs.docs-4.1.0-4.noarch
[root@pw4302-13 ~]#
```

4.4.4 Spectrum Scale open source portability layer

On Linux platforms, Spectrum Scale uses a loadable kernel module that enables the Spectrum Scale daemon to interact with the Linux kernel. Source code is provided for the portability layer so that the Spectrum Scale portability can be built and installed on various Linux kernel versions and configurations. When Spectrum Scale is installed on Linux, you must build a portability module that is based on your particular hardware platform and Linux distribution to enable communication between the Linux kernel and Spectrum Scale. For more information, see the following website:

<http://ibm.co/1CmTpjS>

Note: The Spectrum Scale kernel module should be updated any time that the Linux kernel is updated. Updating the Spectrum Scale kernel module after a Linux kernel update requires rebuilding and installing a new version of the module.

Building the portability layer

Before you start building the portability layer, check for updates to the portability layer at the IBM Support Portal: Downloads for General Parallel File System, found at the following website:

http://www.ibm.com/support/entry/portal/Downloads/Software/Cluster_software/General_Parallel_File_System

The latest kernel levels that are supported are in the Spectrum Scale FAQ in the IBM Knowledge Center at the following website:

http://www.ibm.com/support/knowledgecenter/SSFKCN/com.ibm.cluster.gpfs.doc/gpfs_faq/gpfsclustersfaq.html

One of the new features in this version of Spectrum Scale is a command that simplifies the process to build the portability layer. This command packs the necessary Spectrum Scale software plus the kernel headers on a single RPM file that can be used to distribute and install all the other compute nodes. Example 4-7 shows how to use this new tool.

Example 4-7 Portability layer building tool

```
[root@pw4302-13 GPFS_update]# mmbuildgpl --buildrpm
```

```
mmbuildgpl: Building GPL module begins at Thu Nov 27 17:13:42 EST 2014.
```

```

Verifying Kernel Header...
    kernel version = 2063299 (2.6.32-431.el6.x86_64, 2.6.32-431)
    module include dir = /lib/modules/2.6.32-431.el6.x86_64/build/include
    module build dir  = /lib/modules/2.6.32-431.el6.x86_64/build
    kernel source dir = /usr/src/linux-2.6.32-431.el6.x86_64/include
        Found valid kernel header file under
/lib/modules/2.6.32-431.el6.x86_64/build/include
Verifying Compiler...
    make is present at /usr/bin/make
    cpp is present at /usr/bin/cpp
    gcc is present at /usr/bin/gcc
    g++ is present at /usr/bin/g++
    ld is present at /usr/bin/ld
Verifying rpmbuild...
make World ...
make InstallImages ...
make rpm ...
Wrote:
/root/rpmbuild/RPMS/x86_64/gpfs.gplbin-2.6.32-431.el6.x86_64-4.1.0-4.x86_64.rpm
-----
mmbuildgpl: Building GPL module completed successfully at Thu Nov 27 17:13:59 EST
2014.
-----
[root@pw4302-13 GPFS_update]#

```

The **mmbuildgpl** tool verifies the dependencies of the kernel modules and development tools. You must install all the necessary software or the tool does not generate the RPM file. When the tool completes, the tool generates an RPM file that can be used to install the other cluster nodes. In this example, the file was created in `/root/rpmbuild/RPMS/x86_64/`.

The generated RPM file can be deployed only in machines with an identical architecture, distribution level, Linux kernel, and Spectrum Scale maintenance level.

Now that you have Spectrum Scale installed in the main server, proceed to install it in the second node by completing the following steps:

1. Install the base Spectrum Scale packages.
2. Install the updates if there any available.
3. Install the portability layer package that was created previously.

Note: Ensure that the PATH environment variable for the root user on each node includes `/usr/lpp/mmfs/bin`.

Example 4-8 shows the shortened screens of the installation.

Example 4-8 Output of the installation

```
[root@compute000 ~]# ls -l
total 20956
-rw-r--r-- 1 root root 14482662 Dec  1 14:07 gpfs.base-4.1.0-0.x86_64.rpm
-rw-r--r-- 1 root root  292465 Dec  1 14:07 gpfs.docs-4.1.0-0.noarch.rpm
-rw-r--r-- 1 root root 1548454 Dec  1 14:07 gpfs.ext-4.1.0-0.x86_64.rpm
-rw-r--r-- 1 root root  573838 Dec  1 14:07 gpfs.gpl-4.1.0-0.noarch.rpm
-rw-r--r-- 1 root root 4328387 Dec  1 14:07 gpfs.gskit-8.0.50-16.x86_64.rpm
-rw-r--r-- 1 root root 131514 Dec  1 14:07 gpfs.msg.en_US-4.1.0-0.noarch.rpm
[root@compute000 ~]# yum install *.rpm
```

```

.
.
.

Installed:
  gpfs.base.x86_64 0:4.1.0-0          gpfs.docs.noarch 0:4.1.0-0
  gpfs.ext.x86_64 0:4.1.0-0          gpfs.gpl.noarch 0:4.1.0-0
  gpfs.gskit.x86_64 0:8.0.50-16      gpfs.msg.en_US.noarch 0:4.1.0-0

Dependency Installed:
  ksh.x86_64 0:20120801-10.el6

Complete!
[root@compute000 ~]# 

[root@compute000 gpfs_files_nodes]# ls -l
total 22776
-rw-r--r-- 1 root root 14752057 Dec  1 14:04 gpfs.base-4.1.0-4.x86_64.update.rpm
-rw-r--r-- 1 root root  298337 Dec  1 14:04 gpfs.docs-4.1.0-4.noarch.rpm
-rw-r--r-- 1 root root  2037264 Dec  1 14:04 gpfs.ext-4.1.0-4.x86_64.update.rpm
-rw-r--r-- 1 root root   588752 Dec  1 14:04 gpfs.gpl-4.1.0-4.noarch.rpm
-rw-r--r-- 1 root root  4407428 Dec  1 14:04 gpfs.gskit-8.0.50-32.x86_64.rpm
-rw-r--r-- 1 root root  139569 Dec  1 14:04 gpfs.msg.en_US-4.1.0-4.noarch.rpm
[root@compute000 gpfs_files_nodes]#
[root@compute000 gpfs_files_nodes]# yum update *.rpm
.

.

.

Updated:
  gpfs.base.x86_64 0:4.1.0-4          gpfs.docs.noarch 0:4.1.0-4
  gpfs.ext.x86_64 0:4.1.0-4          gpfs.gpl.noarch 0:4.1.0-4
  gpfs.gskit.x86_64 0:8.0.50-32      gpfs.msg.en_US.noarch 0:4.1.0-4

Complete!
[root@compute000 gpfs_files_nodes]#
[root@compute000 gpfs_files_nodes]# ls -l
-rw-r--r-- 1 root root 1085401 Dec  1 13:51 gpfs.gplbin-2.6.32-431.el6.x86_64-4.1.0-4.x86_64.rpm
[root@compute000 gpfs_files_nodes]#
[root@compute000 gpfs_files_nodes]# yum install gpfs.gplbin-2.6.32-431.el6.x86_64-4.1.0-4.x86_64.rpm
.

.

.

Installed:
  gpfs.gplbin-2.6.32-431.el6.x86_64.x86_64 0:4.1.0-4

Complete!
[root@compute000 gpfs_files_nodes]#

```

4.4.5 Configuring the cluster

After you have the Spectrum Scale software installed on all nodes of the cluster, you can start configuring the disk to be available to all nodes. The Spectrum Scale cluster is created by running `mmcrcluster`, as shown in Example 4-9 on page 43.

Example 4-9 Cluster creation with basic options

```
[root@pw4302-13 4.1]# mmcrcluster  
-Npw4302-13:manager-quorum,compute000:manager-quorum -p pw4302-13 -s compute000 -C  
shareddisk -r /usr/bin/ssh -R /usr/bin/scp  
mmcrcluster: Performing preliminary node verification ...  
mmcrcluster: Processing quorum and other critical nodes ...  
mmcrcluster: Finalizing the cluster data structures ...  
mmcrcluster: Command successfully completed  
mmcrcluster: Warning: Not all nodes have proper GPFS license designations.  
    Use the mmchlicense command to designate licenses as needed.  
mmcrcluster: Propagating the cluster configuration data to all  
    affected nodes. This is an asynchronous process.  
[root@pw4302-13 4.1]#
```

Verify the cluster creation by running **mm1scluster**, as shown in Example 4-10.

Example 4-10 List of the cluster components

```
[root@pw4302-13 ~]# mm1scluster  
  
GPFS cluster information  
=====  
  GPFS cluster name:      shareddisk.pw4302-13  
  GPFS cluster id:       8865240017111139575  
  GPFS UID domain:      shareddisk.pw4302-13  
  Remote shell command:  /usr/bin/ssh  
  Remote file copy command: /usr/bin/scp  
  Repository type:      CCR  
  
  Node   Daemon node name           IP address   Admin node name   Designation  
-----  
  1     pw4302-13                  192.168.0.1  pw4302-13      quorum-manager  
  2     compute000                 192.168.0.3  compute000     quorum-manager  
  
[root@pw4302-13 ~]#
```

During the running of the **mmcrcluster** command, you can see a warning that informs you that the license was not accepted by all cluster members. Proceed to accept the license by running **mmchlicense**, as shown in Example 4-11.

Example 4-11 Accept the license on all nodes

```
[root@pw4302-13 4.1]# mmchlicense server --accept -N pw4302-13,compute000
```

The following nodes will be designated as possessing GPFS server licenses:
 compute000
 pw4302-13
mmchlicense: Command successfully completed
mmchlicense: Propagating the cluster configuration data to all
 affected nodes. This is an asynchronous process.
[root@pw4302-13 4.1]#

Now that the cluster is working, prepare the disks and the partitions that are used by the cluster. Example 4-12 shows the creation of a single partition on a disk that is installed on the main server.

Example 4-12 Create a single partition on a disk

```
[root@pw4302-13 ~]# fdisk /dev/sdb

WARNING: DOS-compatible mode is deprecated. It's strongly recommended to
switch off the mode (command 'c') and change display units to
sectors (command 'u').

Command (m for help): n
Command action
  e  extended
  p  primary partition (1-4)
p
Partition number (1-4): 1
First cylinder (1-51200, default 1): 1
Last cylinder, +cylinders or +size{K,M,G} (1-51200, default 51200):
Using default value 51200

Command (m for help): w
The partition table has been altered!

Calling ioctl() to re-read partition table.
Syncing disks.
[root@pw4302-13 ~]#
```

Example 4-13 shows an example of the stanza file that contains the settings to configure the available disks for the cluster.

Example 4-13 List the stanza file

```
[root@pw4302-13 ~]# cat stanza.txt
%pool:
pool=system

%nsd: device=/dev/sdb servers=pw4302-13,compute000 usage=dataAndMetadata
pool=system
```

This stanza file contains the basic options, and it is used with the **mmcrnsd** command, as shown in Example 4-14.

Example 4-14 Configure the disks within the nodes

```
[root@pw4302-13 ~]# mmcrnsd -F stanza.txt
mmcrnsd: Processing disk sdb
mmcrnsd: Propagating the cluster configuration data to all
affected nodes. This is an asynchronous process.
[root@pw4302-13 ~]#
```

Note: If you are creating only one pool, it should be called *system*. The scenario uses the same pool for data and metadata. In a production environment, it is highly recommended to use different pools for data and metadata to achieve the best results.

The next step is to create a file system in the disk by running **mmcrfs**, which has several options, and use the basic options, as shown in Example 4-15.

Example 4-15 Create the file system

```
[root@pw4302-13 ~]# mmcrfs bigdatafs -F stanza.txt -A yes -B 1024K -j cluster -T /mapred
```

The following disks of bigdatafs will be formatted on node compute000:

```
gpfs2nsd: size 51200 MB  
Formatting file system ...  
Disks up to size 545 GB can be added to storage pool system.  
Creating Inode File  
Creating Allocation Maps  
Creating Log Files  
Clearing Inode Allocation Map  
Clearing Block Allocation Map  
Formatting Allocation Map for storage pool system  
Completed creation of file system /dev/bigdatafs.  
mmcrfs: Propagating the cluster configuration data to all  
affected nodes. This is an asynchronous process.  
[root@pw4302-13 ~]#
```

Finally, mount the disks in the cluster, as shown Example 4-16.

Example 4-16 Mount the new file system and checking it on all nodes

```
[root@pw4302-13 ~]# mmount /mapred -a  
Mon Dec 1 15:57:10 EST 2014: mmount: Mounting file systems ...  
[root@pw4302-13 ~]#  
[root@pw4302-13 ~]# hostname  
pw4302-13  
[root@pw4302-13 ~]#  
[root@pw4302-13 ~]# df -h  
Filesystem          Size  Used Avail Use% Mounted on  
/dev/mapper/vg_pw430213-lv_root    91G   21G   66G  24% /  
tmpfs              7.7G  68K  7.7G  1% /dev/shm  
/dev/sda1           485M  39M  421M  9% /boot  
/root/Downloads/phpc-4.2.x64.iso  1.3G  1.3G    0 100% /mnt  
/root/Downloads/rhel.iso          3.6G  3.6G    0 100% /media/RHEL  
/dev/bigdatafs        50G  470M   50G  1% /mapred  
[root@pw4302-13 ~]#  
[root@compute000 gpfs_files_nodes]# hostname  
compute000  
[root@compute000 gpfs_files_nodes]#  
[root@compute000 ~]# df -h  
Filesystem          Size  Used Avail Use% Mounted on  
/dev/sda3           91G  1.6G   85G  2% /  
tmpfs              7.8G    0  7.8G  0% /dev/shm  
/dev/sda1           248M  38M  198M 17% /boot  
/dev/bigdatafs        50G  470M   50G  1% /mapred  
[root@compute000 ~]#
```



IBM Platform Load Sharing Facility product family

This chapter describes the IBM Platform Load Sharing Facility (LSF) product family.

Platform LSF is a powerful workload management platform for demanding, distributed HPC environments. It provides a comprehensive set of intelligent, policy-driven scheduling features that enable you to use all of your compute infrastructure resources and ensure optimal application performance.

Note: The Platform LSF installer package, product distribution packages, product entitlement packages, and documentation packages can be found in the IBM Passport Advantage® website:

<http://www.ibm.com/software/howtobuy/passportadvantage>

The video found at the following website provides additional help with downloading Platform LSF through IBM Passport Advantage:

http://www-01.ibm.com/support/knowledgecenter/websphere_iea/com.ibm.iea.selfassist/selfassist/1.0/download/HowtoDownloadLSF/HowtoDownloadLSF.html

This chapter covers the following topics:

- ▶ Overview
- ▶ Platform LSF add-ons and capabilities
- ▶ Using IBM Platform MultiCluster
- ▶ IBM Platform Application Center

5.1 Overview

Across enterprises of all sizes, application capabilities and data volumes continue to grow. Facing increasingly restrictive economic pressures, organizations are looking for better ways to improve IT performance, reduce infrastructure costs and expenses, and meet the demand for faster time to solution and market.

The Platform LSF product family is a powerful workload management platform for demanding, distributed, and mission-critical HPC environments. It provides a comprehensive set of intelligent, policy-driven scheduling features that enable you to take full advantage of your compute infrastructure resources and ensure optimal application performance. See Figure 5-1.

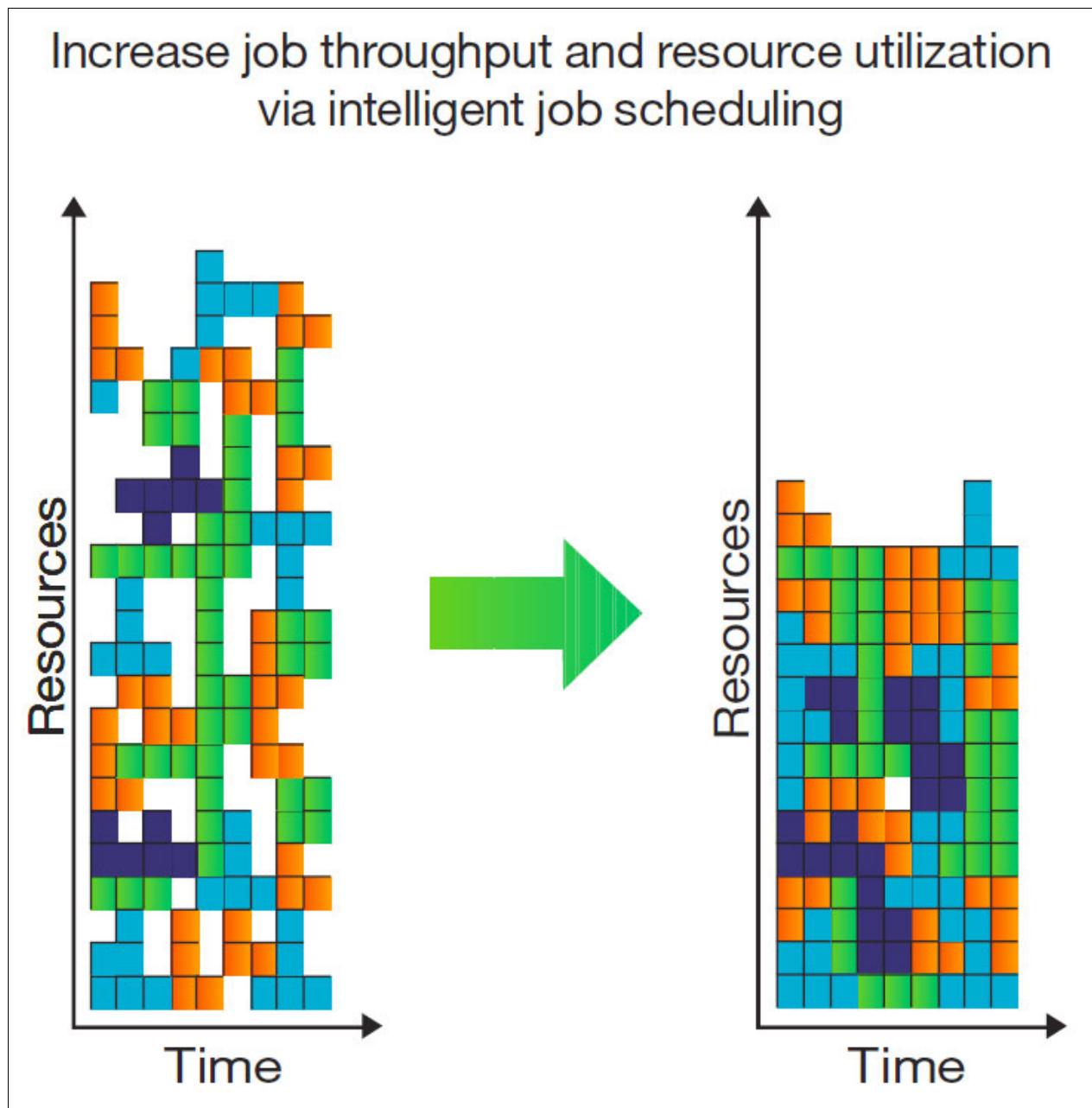


Figure 5-1 Intelligent scheduling in Platform LSF helps make optimum use of resources

A highly scalable and available architecture allows you to schedule complex workloads, and manage from workgroup to petaflop-scale resources. Optional add-ons extend Platform LSF to provide a complete set of workload management capabilities, which work together to address your high performance computing needs.

The Platform LSF product family includes the following products:

- ▶ IBM Platform LSF
- ▶ IBM Platform Application Center
- ▶ IBM Platform RTM
- ▶ IBM Platform License Scheduler
- ▶ IBM Platform Analytics
- ▶ IBM Platform Process Manager
- ▶ IBM Platform Session Scheduler
- ▶ IBM Platform Dynamic Cluster
- ▶ IBM Platform MPI

High performance computing is not easy

Across enterprises of all sizes, application capability and data volumes continue to grow, driving the need for more compute capacity and high performance management and analysis tools. Even in traditional high performance computing (HPC) environments, multiple compute silos, uneven processing, design cycle leaks, and delayed results are common. Facing increasingly restrictive economic pressures, organizations are looking for better ways to improve IT performance, reduce infrastructure costs and expenses, and meet the demand for faster time to solution and market.

The Platform LSF product family is a powerful workload management platform for demanding, distributed, and mission-critical HPC environments. It provides a comprehensive set of intelligent, policy-driven scheduling features that enable you to take full advantage of your compute infrastructure resources and ensure optimal application performance. A highly scalable and available architecture allows you to schedule complex workloads, and manage up to petaflop-scale resources.

The benefits

The Platform LSF product family helps you ensure that all available resources are fully used by enabling you to take full advantage of all technical computing resources, from application software licenses to available network bandwidth. The Platform LSF family can help in the following ways:

- ▶ Reduces operational and infrastructure costs by providing optimal SLA management and greater flexibility, visibility, and control of job scheduling.
- ▶ Improves productivity and resource sharing by fully using hardware and application resources, whether they are just down the hall or halfway around the globe.

History

Figure 5-2 illustrates the history and the future of Platform LSF.

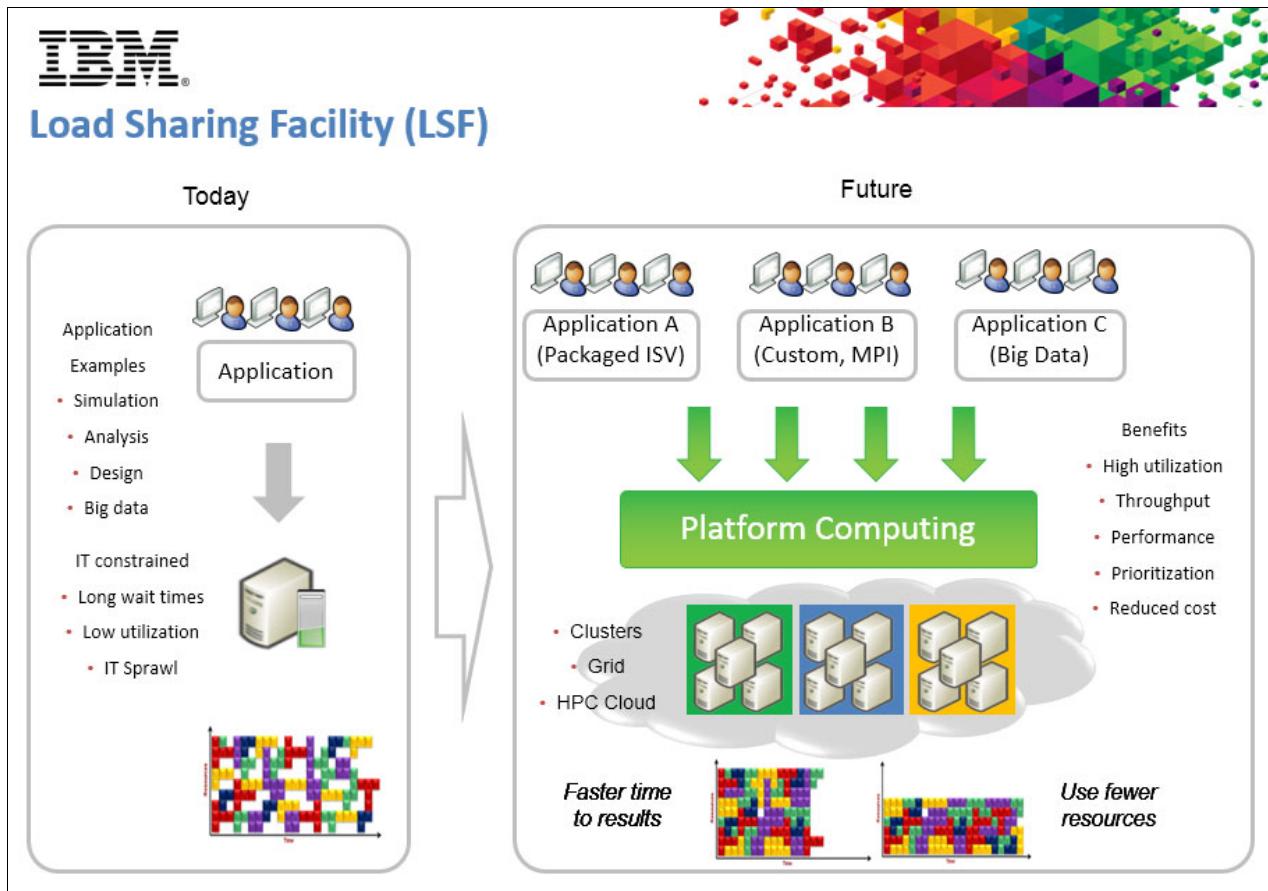


Figure 5-2 Platform LSF in historical view

5.2 Platform LSF add-ons and capabilities

Optional add-ons extend Platform LSF to provide a complete set of HPC performance computing needs, as described in Table 5-1 on page 51.

This section describes the IBM Platform Computing Load Sharing Facility functions:

- ▶ MapReduce Accelerator for LSF
- ▶ Data Manager for LSF
- ▶ Multicluster Technologies

Table 5-1 Optional add-ons extend Platform LSF

Products	Express	Standard	Advanced
IBM Platform Analytics is an advanced analysis and visualization tool for analyzing massive amounts of Platform LSF and IBM Platform Symphony workload data. It enables you to correlate job, resource, and license data from multiple clusters for data-driven decision making.		X	X
IBM Platform Application Center provides a flexible, easy to use interface for cluster users and administrators. Available as an add-on module to Platform LSF, Platform Application Center enables users to interact with intuitive, self-documenting standardized interfaces.	X	X	X
IBM Platform Data Manager provides a management framework for the scheduling and movement of data within clusters, between clusters, and to and from the cloud. An intelligent cache and out-of-band data transfers accelerates the time to solution and eliminates wasted compute cycles.		X	X
IBM Platform Dynamic Cluster turns static Platform LSF clusters into a dynamic, shared cloud infrastructure. By automatically changing the composition of clusters to meet ever-changing workload demands, service levels are improved and organizations can do more work with less infrastructure.		X	X
IBM Platform License Scheduler enables license sharing between global project teams. It ensures that license availability is prioritized by workload, user, and project, and that licenses are optimally used.	X	X	X
IBM Platform Process Manager: Complex scripts are often used to automate lengthy computing tasks. But these scripts can be risky to modify, and might depend on the expertise of a few key individuals. IBM Platform Process Manager simplifies the design and automation of complex computational processes, capturing and protecting repeatable preferred practices.		X	X
IBM Platform RTM is an operational dashboard for Platform LSF environments that provides comprehensive workload monitoring, reporting, and management. Platform RTM provides a complete, integrated monitoring facility that is designed specifically for Platform LSF environments.	X	X	X
IBM Platform Session Scheduler is designed to work with Platform LSF to provide high-throughput, low-latency scheduling in environments that run high volumes of short-duration jobs and where users require faster and more predictable job turnaround times.		X	X
Hybrid Cloud: IBM Platform Computing Cloud Services provide a ready-to-run cluster in the cloud, complete with Platform LSF workload management software, SoftLayer infrastructure, and the support of a dedicated cloud operations team. With IBM Platform Computing Cloud Services, organizations can implement a hybrid cloud environment, rapidly extending local infrastructure to physical, non-shared infrastructure in the SoftLayer cloud to accommodate quickly peaks in demand without being concerned about security or performance.			

5.2.1 Using IBM Platform MapReduce Accelerator for Platform LSF

IBM Platform MapReduce Accelerator for Platform LSF enables you to submit and work with MapReduce jobs in Platform LSF. The Hadoop MapReduce Processing framework is a distributed runtime engine for enterprise-class Hadoop MapReduce applications and shared services deployments.

IBM Platform MapReduce Accelerator for LSF (MapReduce Accelerator) is an add-on pack for Platform LSF that you use to submit and work with MapReduce jobs in Platform LSF. MapReduce jobs are submitted, scheduled, and dispatched like normal Platform LSF jobs.

The following Platform LSF commands work normally with MapReduce jobs:

- ▶ **bbot**
- ▶ **bjobs**
- ▶ **bkill**
- ▶ **bmig**
- ▶ **bmod**
- ▶ **bpost**
- ▶ **bread**
- ▶ **bqueue**
- ▶ **bresize**
- ▶ **bresume**
- ▶ **brun**
- ▶ **bstop**
- ▶ **bsub**
- ▶ **bswitch**
- ▶ **btop**

MapReduce Accelerator supports Apache Pig and Apache Hadoop Streaming jobs.

System requirements

MapReduce Accelerator is delivered in the following file:

`lsf9.1.3_pmra_linux-x64.tar.Z`

Here are the compatible Linux distributions:

- ▶ Red Hat Enterprise Linux 5 or later
- ▶ SUSE Linux Enterprise Server 10 or later

MapReduce Accelerator includes the following versions of Apache Hadoop (“Hadoop”):

- ▶ Apache Hadoop 0.20.2, 0.20.203.0, 0.20.204.0
- ▶ Apache Hadoop 0.21.0 (used by default)
- ▶ Apache Hadoop 1.0.0, 1.0.1
- ▶ Apache Hadoop 1.1.1

MapReduce Accelerator supports Apache Pig versions 0.8.1 and 0.9.2. When you are using Apache Pig, ensure that the Apache Hadoop version is set to a supported value. Supported versions of Apache Pig, and the corresponding supported versions of Apache Hadoop, are as follows:

- ▶ Apache Pig 0.8.1 (use with Apache Hadoop 0.20.x)
- ▶ Apache Pig 0.9.2 (use with Apache Hadoop 1.0.0, 1.0.1, 1.1.1)

MapReduce Accelerator is available as an add-on for all editions of Platform LSF. Purchase MapReduce Accelerator as a separate add-on, then download the distribution package from Platform LSF IBM Service Management Connect at the following website:

<http://www.ibm.com/developerworks/servicemanagement/tc/plsf/index.html>

MapReduce Accelerator supports Platform LSF versions 9.1.1, 9.1.2, and 9.1.3.

Components

MapReduce Accelerator uses elements from MapReduce and LSF. MapReduce Accelerator consists of the following components:

- ▶ **pmr**

The central management process for MapReduce Accelerator and sets up the MapReduce runtime environment for the job submission.

Use **pmr** as a **bsub** command to submit MapReduce jobs to LSF:

```
bsub [bsub_options] pmr [pmr_options]
```

- ▶ **mrsh**

The **mrsh** utility is a shell script that automatically sets up the environment, including the appropriate Java class path, to submit a MapReduce job.

Platform LSF feature interactions

MapReduce Accelerator works with the following Platform LSF features:

- ▶ MapReduce Accelerator supports the Platform MultiCluster job forwarding and resource leasing models, even with the LSF/XL feature enabled for Platform LSF Advanced Edition clusters.
- ▶ After you enable resizable jobs in LSF, MapReduce jobs can work with jobs with increasing task size, but does not support jobs that automatically decrease in task size. If you want to decrease manually the size of a job, run **bresize release**.
- ▶ If you are running MapReduce jobs on multiple hosts, specify **LSF_HPC_EXTENSIONS=CUMULATIVE_RUSAGE** in **lsf.conf** to ensure that LSF does not lose the resource usage in the first host.
- ▶ Each external message slot for a job (which can be seen by running **bread**) can contain up to 51 job messages for MapReduce job information. The default number of message slots is 128, so the default maximum MapReduce job number is 6528. If there are several MapReduce jobs in a single LSF job, increase the number of job message slots to ensure that you do not lose MapReduce job information. Define **MAX_JOB_MSG_NUM** in **lsb.params** to increase the number of job message slots to at least the following value:

(Total number of MapReduce jobs)/51 + 1

Installing MapReduce Accelerator

Run **lsvinstall** to install MapReduce Accelerator as an add-on package for Platform LSF.

Before you install it, check the following items:

- ▶ You must be running a supported version of Platform LSF.
- ▶ MapReduce Accelerator Version 9.1.3 supports Platform LSF versions 9.1.1, 9.1.2, and 9.1.3.
- ▶ Your PATH environment variable must include a path to a Java installation Version 1.4, or later.

Install MapReduce Accelerator by completing the following steps:

1. Extract the MapReduce Accelerator installer file (**lsv9.1.3_pmra_no_jre_lsvinstall.tar.Z**).
2. Edit the **install.config** file and specify the installation parameters for MapReduce Accelerator and your current LSF installation.

3. Specify the following parameters:
 - LSF_TOP: Specify the same top-level LSF installation in your existing cluster.
 - LSF_ADMIN: Specify the same value as your existing LSF cluster.
 - LSF_CLUSTER_NAME: Specify the same value as your existing LSF cluster.
 - LSF_ENTITLEMENT_FILE: Specify the path to your LSF entitlement file.
 - LSF_TARDIR: Specify the path to the location of the MapReduce Accelerator distribution file.
4. Run `lsvinstall -f install.config` to install MapReduce Accelerator.
5. Follow the prompts to install MapReduce Accelerator. You can also use the unattended installer for MapReduce Accelerator.
6. Optional: If you are running LSF V9.1.1 or V9.1.2 and intend to use `bjobs -mr` to view MapReduce job information, download the new `bjobs` binary file to replace the old `bjobs` binary file in your Platform LSF installation.

The new `bjobs` binary file for MapReduce Accelerator is available from Platform LSF IBM Service Management Connect at the following website:

<http://www.ibm.com/developerworks/servicemanagement/tc/plsf/index.html>

The MapReduce Accelerator installer creates an application that is named `pmra` in `lsb.applications` with the following configuration:

```
Begin Application
NAME = pmra
DESCRIPTION = IBM Platform LSF MapReduce
RTASK_GONE_ACTION = IGNORE_TASKCRASH
DJOB_COMMFAIL_ACTION = IGNORE_COMMFAIL
TERMINATE_CONTROL = SIGTERM
DJOB_RU_INTERVAL = 300
DJOB_HB_INTERVAL = 300
DJOB_RESIZE_GRACE_PERIOD = 30
RESIZABLE_JOBS = AUTO
POST_EXEC = mrclean.sh
End Application
```

Submitting MapReduce jobs

Run the `bsub` and `pmr` commands to submit MapReduce jobs to LSF. Before you submit a MapReduce job, set the following environment variables:

- ▶ Set the `JAVA_HOME` environment variable to specify the top-level path to the Java runtime environment (JRE).

Note: The `JAVA_HOME` file path must be accessible to all execution hosts. To ensure that the file path is accessible, either install JRE to a shared file path, or install JRE to the same local file path on each execution host. MapReduce Accelerator supports JRE V1.6 or later.

- ▶ Set the `HADOOP_VERSION` environment variable to the version of Apache Hadoop that you are using, as shown in Table 5-2 on page 55.

Table 5-2 HADOOP_VERSION values

HADOOP_VERSION value	Apache Hadoop version
21 (default value)	0.21.0
20	0.20.2
20_203	0.20.203.0
20_204	0.20.204.0
1_0_0	1.0.0 or 1.0.1
1_1_1	1.1.1

For example, if you are running Apache Hadoop 1.0.1, set HADOOP_VERSION=1_0_0 as the environment variable value.

You can set these variables as runtime environment variables on the submission host by using **setenv** or **export** before you submit the MapReduce job. You can also add the definitions into the pmr-env.sh file (in the \$LSF_ENVDIR\pmra\\$CLUSTER_NAME\9.1 directory). The environment variable definitions in the submission host override the definitions in pmr-env.sh.

About this task

Here is the general command to submit MapReduce jobs:

```
bsub [bsub_options] pmr [pmr_options] [command]
```

You can submit a single MapReduce task in the job submission by using the **mrsh** utility, which has the following syntax

```
bsub [bsub_options] pmr [pmr_options] mrsh jar jarfile.jar [classname]
[-Dproperty=value ...] [arguments]
```

- ▶ **jarfile** specifies the file name of the application that is packaged as a JAR file, which includes the MapReduce code.
- ▶ **classname** specifies the class to be started. If the class is not specified, the class that is specified by the JAR manifest is run.
- ▶ **property=value** specifies settings for a job:
 - **-Dproperty** specifies the name of a MapReduce task configuration property.
 - **value** specifies the value for the MapReduce task configuration property.

For example:

```
bsub -n 100,300 pmr mrsh jar wordcount.jar /filepath/input /filepath/output
```

You can submit a Hadoop Streaming job in the MapReduce job submission by using the **mrsh** utility, which has the following syntax

```
bsub [bsub_options] pmr [pmr_options] mrsh jarfile.jar [classname]
[-Dproperty=value ...] [arguments]
```

- ▶ **jarfile** specifies the file name of the Hadoop Streaming application that is packaged as a JAR file, which includes the MapReduce code.
- ▶ **classname** specifies the class to be started. If the class is not specified, the class that is specified by the JAR manifest is run.

- ▶ `-Dproperty=value` specifies settings for a job:
 - `property` specifies the name of a job configuration property.
 - `value` specifies the value for the job configuration property.

For example:

```
bsub -n 100,300 pmr mrsh hadoop-streaming.jar -input /filepath/input -output
/filepath/output -mapper /bin/cat -reducer /bin/wc
```

5.2.2 Using IBM Platform Data Manager for LSF

When large amounts of data are required to complete computations, it is preferable that your applications access required data unhindered by the location of the data in relation to the application execution environment. Platform Data Manager for LSF solves the problem of data locality by staging the required data as closely as possible to the site of the application.

Many applications in several domains require large amounts of data: fluid dynamics models for industrial manufacturing, seismic sensory data for oil and gas exploration, gene sequences for life sciences, and others. Locating these large data sets as close as possible to the application runtime environment is crucial to maintaining optimal utilization of compute resources.

Whether you are running these data-intensive applications in a single cluster or you want to share data and compute resources across geographically separated clusters, Platform Data Manager for LSF provides the following key features:

- ▶ Input data can be staged from an external source storage repository to a cache that is accessible to the cluster execution hosts.
- ▶ Output data is staged asynchronously (dependency-free) from the cache after job completion.
- ▶ Data transfers run separately from the job allocation, which means more jobs can request data without consuming resources waiting for large data transfers.
- ▶ Remote execution cluster selection and cluster affinity is based on data availability in a Platform MultiCluster environment. Platform Data Manager for LSF transfers the required data to the cluster to which the job was forwarded.

Platform LSF data manager

The Platform LSF data manager runs on dedicated Platform LSF server hosts. The Platform LSF data manager hosts are configured to run the Platform LSF data manager daemon (dmd). The Platform LSF data manager daemon communicates with the clusters it serves, and manages the transfer of data in the staging area.

Query the Platform LSF data manager with the **bdata** command to get information about the required data files, Platform LSF data manager configuration, cluster connections, transfer status, and other information. Data manager administrators can use **bdata** to reconfigure and shut down dmd.

Platform LSF data manager administrator

The administrator of IBM Platform Data Manager for Platform LSF must be a Platform LSF administrator for all clusters that are connected to the data manager. Platform LSF data manager administrators make sure that dmd is operating smoothly and reconfigure Platform LSF data manager as needed.

Platform LSF data manager administrators can perform the following administrative functions on the Platform LSF data manager:

- ▶ Manage the Platform LSF data manager data transfer queue in `1sb.queues`.
- ▶ Run `bdata admin reconfig` to reconfigure the Platform LSF data manager.
- ▶ Run `bdata admin shutdown` to shut down the Platform LSF data manager.
- ▶ Run `bdata tags` to list or clean intermediate files that are associated with a tag for users.

Configuring the Platform LSF data manager

Configure the Platform LSF data manager administrators with the `ADMINS` parameter in `1sf.datamanager` file.

The `1sf.datamanager` file controls the operation of Platform Data Manager for Platform LSF features. There is one Platform LSF data management configuration file for each cluster, called `1sf.datamanager.cluster_name`. The `cluster_name` suffix is the name of the cluster that is defined in the Cluster section of `1sf.shared`. The file is read by the Platform LSF data management daemon `dmd` during start and reconfiguration.

Data transfer node

A data transfer node, also referred to as an I/O node, is a Platform LSF server host in the cluster that is mounted with direct read/write access to the cluster staging area. This host can access the source of staged-in data and the destination of staged-out data.

Data transfer job

Platform LSF data manager submits transfer jobs to copy required data files for stage-in or stage-out operations. Transfer jobs run on data transfer nodes as the execution user of the job that triggered the transfer.

Transfer jobs have the following function:

- ▶ To pre-stage files that are requested in the `bsub -data` option from their source location into the staging area cache.
- ▶ Stage out files that are requested by the `bstage out` command from the staging area cache to their remote destination.

Data transfer queue

Platform LSF data manager submits transfer jobs to a transfer queue, which is configured to accept transfer jobs only.

Data transfer tool command

Transfer jobs run the transfer tool command that is specified in `FILE_TRANSFER_CMD` in `1sf.datamanager`.

Data specification file

A data specification file is a text file that is used for specifying many data requirement files for one job.

Each line in a data specification file specifies the name of the path to a source file to be transferred to the staging area before a job is submitted and scheduled. The path can point to a file or a directory.

The following example contains lines for three files. Each line specifies a host_name:file_path pair:

```
#@dataspec  
datahost:/proj/userA/input1.dat  
datahost:/proj/userA/input2.dat  
datahost:/proj/userA/input3.dat
```

Data tags

A data tag can be created for a job with a data staging requirement with the **bstage out** command. A tag allows users to transfer files from the job's current working directory to the staging area, associate those files with a chosen name, and to have the Platform LSF data manager report the existence of that tag if it is queried later.

Data queries

File-based cache query with **bdata** cache displays the job IDs of jobs that request the file under REF_JOB. The REF_JOB column is not displayed for job-based query in **bdata** cache.

How Platform Data Manager for Platform LSF works

Every Platform LSF cluster that shares a staging area also communicates to the same Platform LSF data manager instance. The clusters query the data manager for the availability of data files. If the files are not in the cache, the Platform LSF data manager stages them and notifies the cluster when a requested data for the job is ready. After files are staged, the clusters can retrieve them from the staging area by consulting the data file information that is stored in the staging area by the Platform LSF data manager.

A cluster can be implemented as single or multicluster.

Using IBM Platform Data Manager for Platform LSF

To submit and manage jobs with data requirements, use the following commands:

- ▶ **bsub**: Requests that files are staged for jobs before they are scheduled.
- ▶ **bmod**: Modifies data requirement requests for submitted jobs.
- ▶ **bstage in**: Gets requested files from the staging area during job execution.
- ▶ **bstage out**: Requests that files in the job execution environment are returned to the staging area or submission environment.
- ▶ **bdata**: Queries the status of files and data tags in the staging area cache and manages data tags that are associated with your jobs.
- ▶ **bjobs**: Queries the status of jobs with data requirements.
- ▶ **bhist**: View historical information about jobs with data requirements.

Note: For more information about these commands, go to the following website:

http://www-01.ibm.com/support/knowledgecenter/SSETD4_9.1.3/dm_using/dm_chap_using_data_manager.dita

5.3 Using IBM Platform MultiCluster

IBM Platform MultiCluster is a feature of Platform LSF.

Within an organization, sites might have separate, independently managed Platform LSF clusters. Having multiple Platform LSF clusters can solve problems that are related to the following areas:

- ▶ Ease of administration
- ▶ Different geographic locations
- ▶ Scalability

When you have more than one cluster, it is preferable to allow the clusters to cooperate to reap the following benefits of global load sharing:

- ▶ Access to a diverse collection of computing resources.
- ▶ Enterprise grid computing becomes a reality.
- ▶ Get better performance and computing capabilities.
- ▶ Use idle machines to process jobs.
- ▶ Use multiple machines to process a single parallel job.
- ▶ Increase user productivity.
- ▶ Add resources anywhere and make them available to the entire organization.
- ▶ Plan computing resources globally based on total computing demand.
- ▶ Increase computing power in an economical way.

MultiCluster enables a large organization to form multiple cooperating clusters of computers so that load sharing happens not only within clusters, but also among them. MultiCluster enables the following features:

- ▶ Load sharing across many hosts.
- ▶ Co-scheduling among clusters: The job forwarding scheduler considers remote cluster and queue availability and loads before forwarding jobs.
- ▶ Resource ownership and autonomy are enforced.
- ▶ Non-shared user accounts and file systems are supported.
- ▶ Communication limitations among the clusters are considered in job scheduling

There are two different ways to share resources between clusters by using MultiCluster. These models can be combined, for example, Cluster1 forwards jobs to Cluster2 by using the job forwarding model, and Cluster2 borrows resources from Cluster3 by using the resource leasing model.

Choosing a model

Consider your own goals and priorities when choosing the best resource-sharing model for your site:

- ▶ The job forwarding model can make resources available to jobs from multiple clusters. This flexibility allows maximum throughput when each cluster's resource usage fluctuates. The resource leasing model can allow one cluster exclusive control of a dedicated resource, which can be more efficient when there is a steady amount of work.
- ▶ The lease model is the most transparent to users and supports the same scheduling features as a single cluster.
- ▶ The job forwarding model has a single point of administration, and the lease model shares administration between provider and consumer clusters.

Job forwarding model

In this model, the cluster that is starving for resources sends jobs over to the cluster that has resources to spare. To work together, the two clusters must set up compatible send-jobs and receive-jobs queues.

With this model, scheduling of MultiCluster jobs is a process with two scheduling phases:

- ▶ The submission cluster selects a suitable remote receive-jobs queue, and forwards the job to it.
- ▶ The execution cluster selects a suitable host and dispatches the job to it.

This method automatically favors local hosts; a MultiCluster send-jobs queue always attempts to find a suitable local host before considering a receive-jobs queue in another cluster.

Resource leasing model

In this model, the cluster that is starving for resources takes resources away from the cluster that has resources to spare. To work together, the provider cluster must “export” resources to the consumer, and the consumer cluster must configure a queue to use those resources.

In this model, each cluster schedules work on a single system image, which includes both borrowed hosts and local hosts.

5.4 IBM Platform Application Center

IBM Platform Application Center V9.1.3 provides a flexible and easy to use interface for cluster users and administrators. Available as an add-on module to Platform LSF, Platform Application Center enables users to interact with intuitive, self-documenting standardized interfaces.

IBM Platform Application Center Standard Edition provides basic job submission, job and host monitoring, default application templates, role-based access control, reporting, customization, and remote visualization capabilities.

One of the interesting functions of the Platform Application Center Standard Edition product is 2D/3D Remote Visualizations. The remote console feature is disabled by default. You must make specific configurations to use any of the supported remote visualization applications.

Using NICE Desktop Cloud Visualization

You can configure Platform Application Center and Platform LSF to enable viewing of a 2D/3D Windows application from Platform Application Center by using NICE Desktop Cloud Visualization (DCV).

DCV is an advanced technology that enables technical computing users to access remotely 2D/3D interactive applications over a standard network.

The DCV protocol adapts to heterogeneous networking infrastructures such as LAN, WAN, and VPN to deal with bandwidth and latency constraints. All applications run natively on the remote machines, which can be virtualized and share physical GPU.

Users use Platform Application Center and Platform LSF to start their application and view the results remotely through DCV. Platform LSF schedules and allocates hosts that have the specific application installed.

Users do not need to know which hosts have the application installed or which hosts are available. In this way, compute resources and application licenses can be shared, increasing resource efficiencies and reducing cost.

Platform Application Center provides a default AppDCVonLinux application template. You can create custom application templates to support additional applications.

Note: For more information about setting up the Remote Visualization function, go to the following IBM Knowledge Center website:

<http://ibm.co/1KAEp6P>



IBM Platform Symphony V7.1 with Application Service Controller

This chapter describes IBM Platform Symphony (Platform Symphony) V7.1 with the Application Service Controller add-on. It also describes the advantages of new effective technology, such as Platform Symphony V7.1 working together with the Apache Spark engine for large-scaled data processing.

Platform Symphony V7.1 offers increased scaling and performance. You can use IBM Platform Application Service Controller to better manage cloud-native distributed computing environments by eliminating silos and making the most efficient use of available resources.

You can realize the following benefits:

- ▶ Faster throughput and performance
- ▶ Higher levels of resource utilization
- ▶ Reduced infrastructure and management costs
- ▶ Reduced application development and maintenance costs
- ▶ The agility to respond instantly to real-time demands
- ▶ Improved management of heterogeneous distributed applications

This chapter covers the following topics:

- ▶ Introduction to IBM Platform Symphony V7.1
- ▶ IBM Platform Symphony: An overview
- ▶ IBM Symphony for multitenant designs
- ▶ Product editions
- ▶ Optional applications to extend Platform Symphony capabilities
- ▶ Overview of IBM Platform Application Service Controller
- ▶ IBM Platform Symphony application implementation
- ▶ Overview of Apache Spark as part of the IBM Platform Symphony solution
- ▶ ASC as the attachment for cloud-native framework: Apache Cassandra

6.1 Introduction to IBM Platform Symphony V7.1

Platform Symphony provides the most powerful application framework that you can use to run distributed or parallel applications in a scaled-out grid environment. It virtualizes compute-intensive application services and processes across existing heterogeneous IT resources. You can use Platform Symphony to run pre-integrated applications that are available from various ISVs. You can take advantage of new technologies, such as running Platform Symphony and the Apache Spark engine enhancements:

- ▶ Improved scale and performance: Three times increased scalability and improved performance across core Platform Symphony and MapReduce workloads.
- ▶ Innovative data management technologies: Data bottlenecks that are removed and data movement reduced.
- ▶ Enhanced multitenancy and resource management: Runtime elasticity with new cloud-native applications addressed.
- ▶ Expanded workload management: Emerging application workload patterns managed.

The efficient, low-latency middleware and scheduling architecture of Platform Symphony delivers the performance and agility that are required to meet and exceed predictably throughput goals for the most demanding analytic workloads. Platform Symphony helps organizations realize improved application performance at a reduced total cost of ownership (TCO).

Platform Symphony can help you to obtain higher-quality business results faster, reduce infrastructure and management costs, accelerate many types of Hadoop MapReduce workloads, and combine compute- (and data-) intensive applications on a single shared platform.

It includes the following features:

- ▶ An ultrafast, low-latency grid scheduler
- ▶ Multicloud support for scalability to 128,000 service instances per cluster (typically mapped to cores)
- ▶ A unique resource-sharing model that enables multitenancy with resource lending and borrowing for maximum efficiency
- ▶ An optimized, low-latency MapReduce implementation that is compatible with IBM InfoSphere BigInsights and other big data solutions.

Platform Symphony V7.1 has been available for download since December 5, 2014. Its program number is 5725-G86.

Figure 6-1 on page 65 shows the target audience for Platform Symphony.

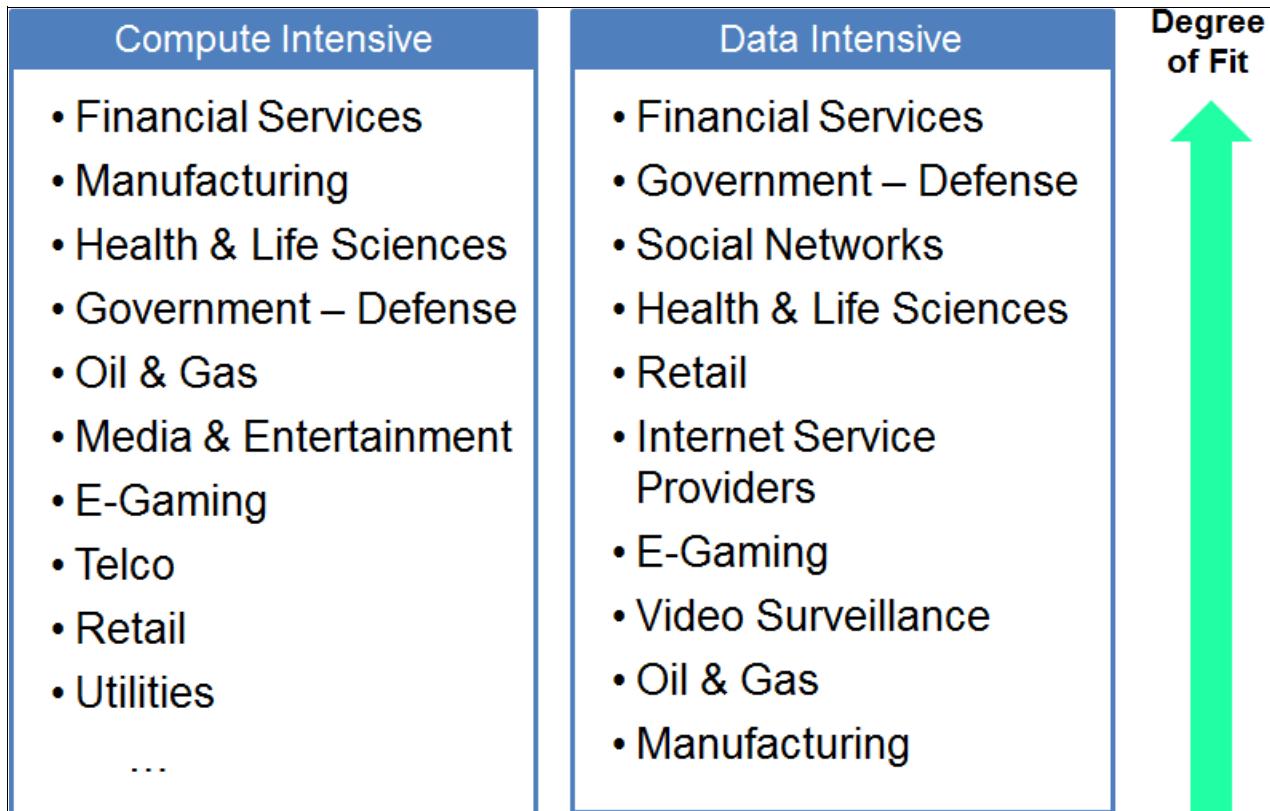


Figure 6-1 Platform Symphony - target markets

6.2 IBM Platform Symphony: An overview

Platform Symphony is enterprise-class software that distributes and virtualizes compute-intensive application services and processes across existing heterogeneous IT resources. Platform Symphony creates a shared, scalable, and fault-tolerant infrastructure, delivering faster, more reliable application performance while reducing costs.

Platform Symphony provides an application framework that you can use to run distributed or parallel applications in a scaled-out grid environment.

Note: As a quick primer to some of the terminology that is referenced in this chapter, some definitions are offered in this section. For more information, see *IBM Platform Symphony Foundations*, which is available at the following website:

<http://publibfp.dhe.ibm.com/epubs/pdf/c2750652.pdf>

Figure 6-2 illustrates the Platform Symphony application framework.

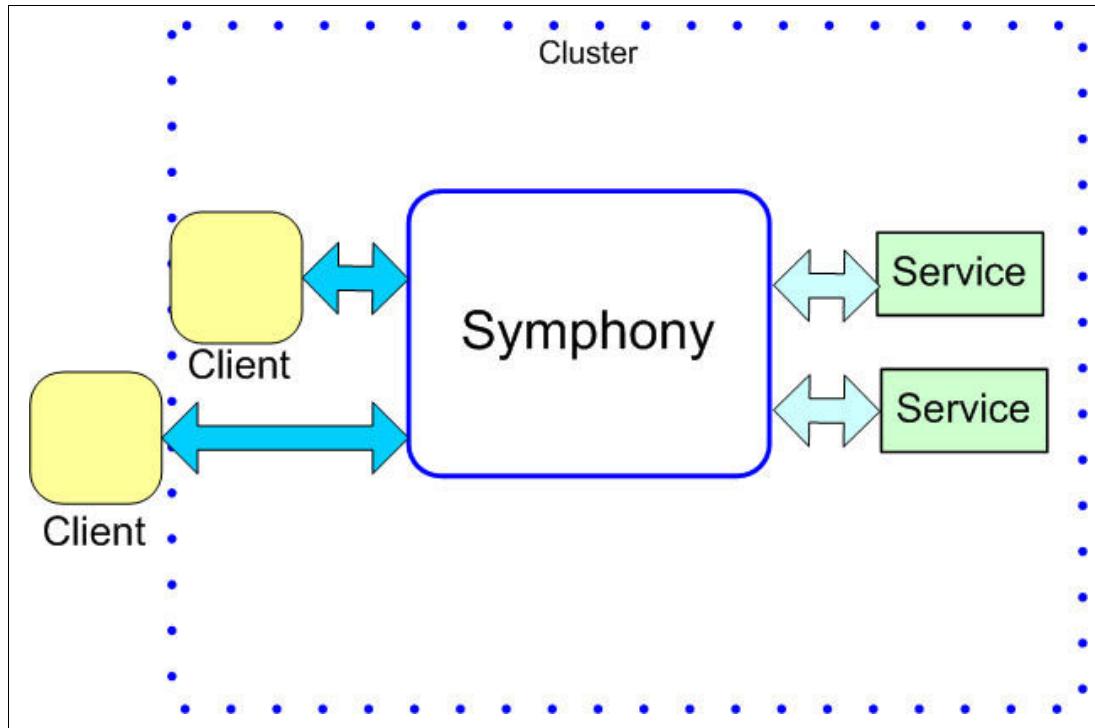


Figure 6-2 Platform Symphony application framework

Cluster

A cluster is a logical grouping of hosts that provides a distributed environment in which to run applications.

Platform Symphony

Platform Symphony manages the resources and the workload in the cluster. Using Platform Symphony, resources are virtualized: Platform Symphony dynamically and flexibly assigns resources, provisioning them and making them available for applications to use.

Platform Symphony can assign resources to an application on demand when the work is submitted, or the assignment can be predetermined and preconfigured.

Application

A Platform Symphony service-oriented application uses a client-service architecture. It consists of two programs: the client, which provides the client logic to submit work, retrieve and process results, and the service, which comprises the business logic (the computation). The service-oriented application uses parallel processing to accelerate computations. Platform Symphony receives requests to run applications from a client. Platform Symphony manages the scheduling and running of the work; the client does not need to be concerned with where the application runs.

Client

The client sends compute requests and collects results by using the Platform Symphony client APIs. The client can run on a machine that is part of the cluster, or it can run on a machine that is outside of the cluster. The client can use a service without knowledge of what programming language was used to create the service.

The client submits an input data request to Platform Symphony. Platform Symphony initiates the service that processes the client requests, receives results from the service, and passes the results back to the client.

Service

The service is a self-contained business function that accepts requests from a client, performs a computation, and returns responses to the client. The service uses computing resources, and must be deployed to the cluster. Multiple instances of a service can run concurrently in the cluster.

The service is initiated and run by Platform Symphony, upon receipt of a client request. The service runs on a machine that is part of the Platform Symphony cluster. The service runs on the cluster resources that are dynamically provisioned by Platform Symphony. Platform Symphony monitors the running of the service, and passes the results back to the client.

Platform Symphony cluster components

A Platform Symphony cluster manages both workload and resources. Platform Symphony maintains historical data, includes a web interface for administration and configuration, and also has a command-line interface (CLI) for administration.

Workload management versus resource management

A workload manager interfaces directly with the application, receiving work, processing it, and returning the results. A workload manager provides a set of APIs, or might interface with additional runtime components to enable the application components to communicate and perform work. The workload manager is aware of the nature of the applications it supports by using terminology and models consistent with a given class of workload. In a service-oriented application environment, workload is expressed in terms of messages, sessions, and services.

A resource manager provides the underlying system infrastructure to enable multiple applications to operate within a shared resource infrastructure. A resource manager manages the computing resources for all types of workload.

Enterprise Grid Orchestrator resource manager

Enterprise Grid Orchestrator (EGO) manages the supply and distribution of resources, making them available to applications. EGO provides resource provisioning, remote execution, high availability, and business continuity.

EGO provides cluster management tools and the ability to manage supply versus demand to meet service-level agreements (SLAs).

SOA middleware workload manager

SOA middleware (SOAM) manages service-oriented application workload within the cluster, creating a demand for cluster resources. When a client submits an application request, the request is received by SOAM. SOAM manages the scheduling of the workload to its assigned resources, requesting additional resources as required to meet SLAs. SOAM transfers input from the client to the service, then returns results to the client. SOAM releases excess resources to the resource manager.

Platform Management Console

The Platform Management Console (PMC) is your window to Platform Symphony, providing resource monitoring capability, application service-level monitoring and control, and configuration tools.

Historical data for reporting

Platform Symphony stores a wide variety of historical data for reporting and diagnostic purposes. Multiple reports capture and summarize the data.

How Platform Symphony supplies resources

To understand how Platform Symphony supplies resources to meet workload requests, consider the following analogy.

A bank customer does not withdraw funds directly from the bank vaults. The customer accesses an account, and requests a withdrawal from that account. The bank recognizes the customer by the account number, and determines whether the customer has sufficient funds to make a withdrawal, as shown in Figure 6-3.

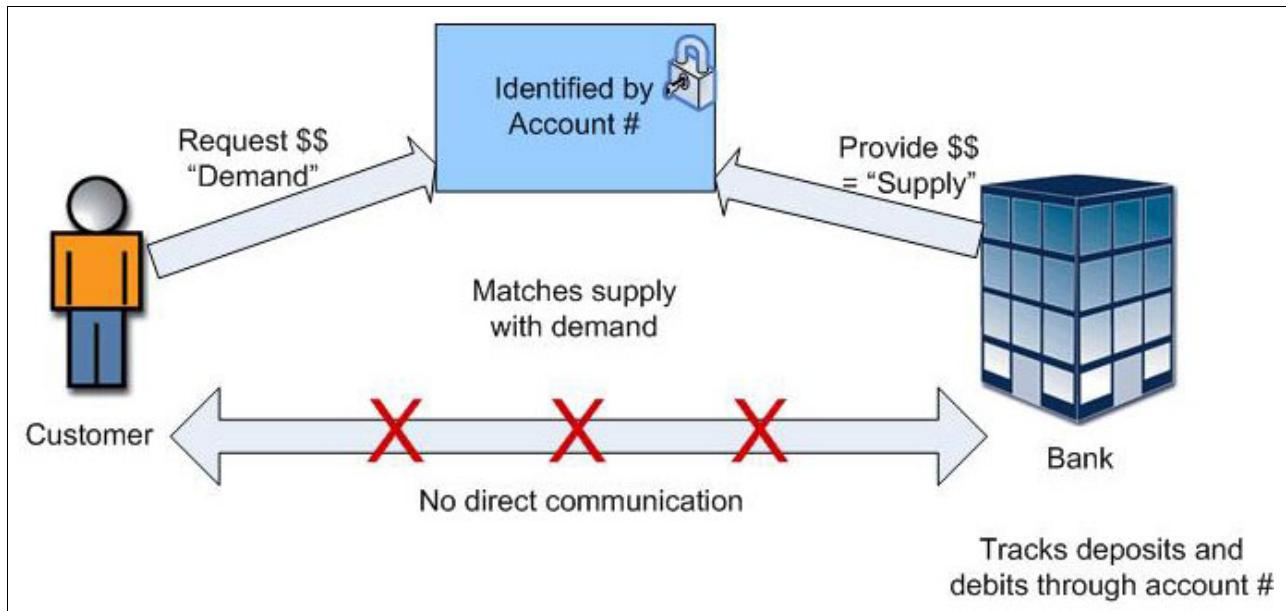


Figure 6-3 This analogy illustrates how Platform Symphony supplies resources

As shown in Figure 6-4 on page 69, when a Platform Symphony application requires resources, it does not communicate directly with EGO, and has no direct access to resources. The application is associated with a consumer, and requests resources through it. EGO recognizes the consumer, and through it, allocates resources to the application.

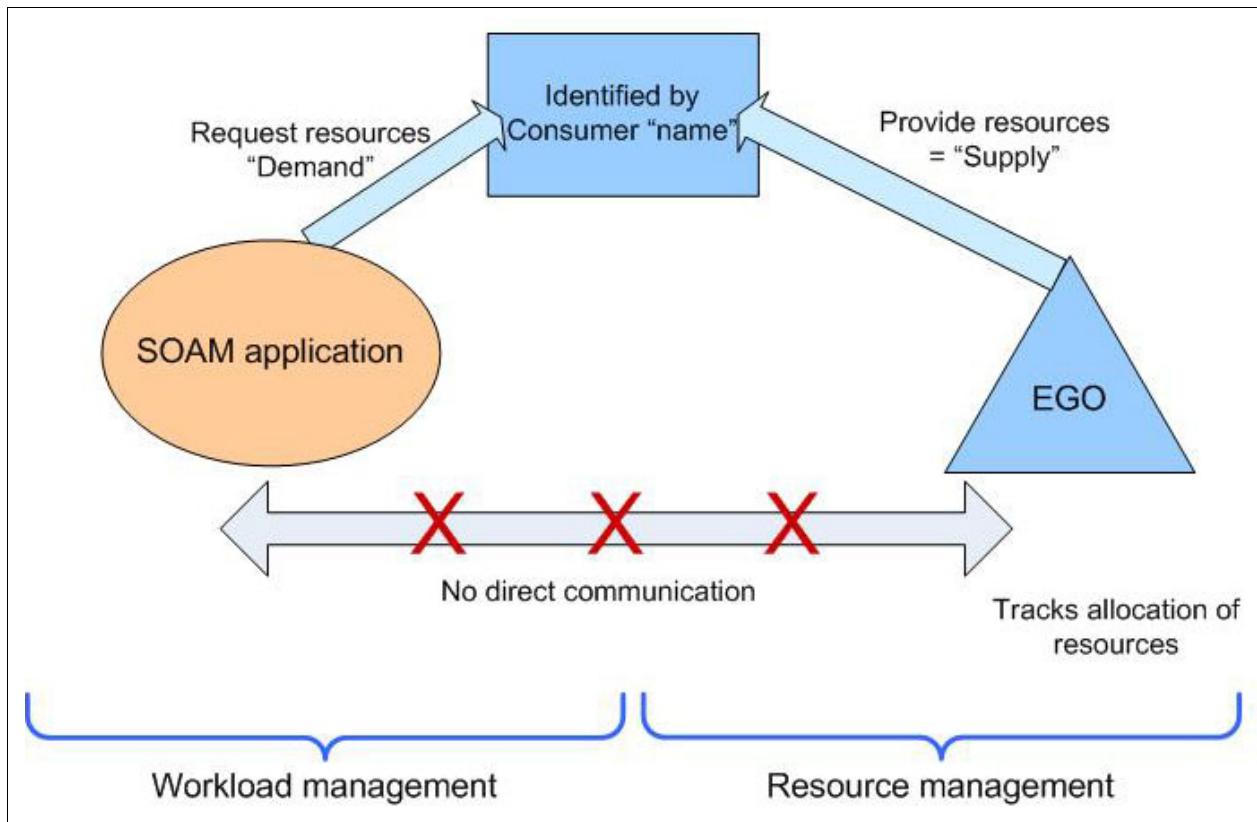


Figure 6-4 How Platform Symphony supplies resources

6.3 IBM Symphony for multitenant designs

Multitenancy is an architecture in which a single instance of a software application serves multiple customers. Each customer is called a tenant. Tenants may be given the ability to customize some parts of the application, such as color of the user interface (UI) or business rules, but they cannot customize the application's code.

Multitenancy: The narrow view

In a multitenancy environment, multiple customers share an application, running on the same operating system, on the same hardware, with the same data-storage mechanism.

Big data and analytics infrastructure silos are inefficient. Platform Symphony helps you to achieve the best results by using multitenancy.

6.3.1 Challenges and advantages

Using a shared infrastructure environment, this service reduces hardware, software, and environmental costs while maintaining a secured infrastructure through isolated LPARs and IBM's comprehensive managed services. It offers an allocation-based consumption model that further reduces costs so you pay only for what is allocated to you. The savings are obtained from using the cost of hardware and software across the entire multitenant customer base environment. In addition, the service provides dynamic capacity to meet peak workload requirements and growth as business needs change.

Here are some of the associated challenges:

- ▶ Increasing cost of analytics
- ▶ Addressing pain that is associated with the extract, transform, and load (ETL) process
- ▶ Accommodating data warehouse volume growth
- ▶ Delivering needed information in a timely manner
- ▶ Ensuring information is available when needed
- ▶ Managing the Hadoop environment

Here are some technical needs that are defined as ever-increasing expectations:

- ▶ Increased performance to support business demands
- ▶ Increased scalability to address huge and growing volumes of data
- ▶ Optimized use of existing resources for scaled performance
- ▶ Efficient data management to remove data bottlenecks
- ▶ Support for new, cloud-native application workload patterns
- ▶ Effective operational management: monitoring, alerting, diagnostic tests, and security

6.3.2 Multitenant designs

In general, multitenancy implies multiple non-related consumers or customers of a set of services. Within a single organization, this situation can be multiple business units with resources and data that must remain separate for legal or compliance reasons. Most hosting companies require multitenancy as a core attribute of their business model. This model might include a dedicated physical infrastructure for each hosted customer or logical segmentation of a shared infrastructure by using software-defined technologies.

In Platform Symphony Advanced Edition, up to 300 MapReduce runtime engines (job trackers) can coexist and use the same infrastructure.

Users can define multiple MapReduce applications and associate them with resource consumers by “cloning” the default MapReduce application. Each application has its separate and unique Job Tracker (SSM). When multiple SSMs are instantiated, they are balanced on the available management nodes.

Furthermore, inside each application, simultaneous job management is possible because of the special design that implements sophisticated scheduling of multiple sessions on the resources that are allocated for an application. This function is obtained by separating the job control function (workload manager) from the resource allocation and control (EGO). The new Apache Hadoop NextGen MapReduce (YARN), Apache Hadoop 2, has a similar feature, but this release is still in alpha stage. The stable release of Hadoop MapReduce offers only one Job Tracker per cluster.

Moreover, multitenancy is more than multiple job trackers. It is about user security, shared and controlled access to the computing resources and to the whole environment, monitoring and reporting features, and so on. These multitenancy features are addressed as they are implemented by the Platform Symphony product.

6.3.3 Requirements gathering

Requirements gathering can determine how the consumer becomes aware of and can request access to hosted services. You should be able to answer the following questions:

- ▶ Will consumers use accounts that the host creates or accounts that they use internally to access services?
- ▶ Is one consumer allowed to be aware of other consumer's identities, or is a separation required?
- ▶ Can multiple consumers share a physical infrastructure?
- ▶ Can traffic from multiple consumers share a common network?
- ▶ Can software-defined isolation meet the requirements?
- ▶ How far into the infrastructure must authentication, authorization, and accounting be maintained for each consumer?

Segmentation options that might be considered as part of a multi-tenant infrastructure are physical separation by customer (dedicated hosts, network, and storage), logical separation by customer (shared physical infrastructure with logical segmentation), data separation, network separation (VLANs), and performance separation (shared infrastructure but ensured capacity).

6.3.4 Building a multitenant big data infrastructure

Platform Symphony provides a platform for robust, multi-computer automation for all elements of a data center, including servers, operating systems, storage, and networking. It also provides centralized administration and management capabilities, such as deploying roles and features remotely to physical and virtual servers, and deploying roles and features to virtual hard disks, even when they are offline.

Platform Symphony concepts

Although you might be familiar with Hadoop and various commercial distributions, you might be less familiar with Platform Symphony. Platform Symphony is a commercial grid workload and resource management solution that shares resources among diverse applications in multitenant environments. Platform Symphony is widely deployed as a shared services infrastructure in some of the world's largest investment banks.

Session manager

Service-oriented applications in Platform Symphony are managed by a *session manager*. The session manager is responsible for dispatching tasks to service instances, and collecting and assembling results. The Platform Symphony session manager provides a function similar in concept to a Hadoop application manager, although it has considerably more capabilities. Platform Symphony implements a Job Tracker function by using the session manager. In this book, the terms Job Tracker, application manager, and session manager are used interchangeably. Although the concept of multiple concurrent application managers in Hadoop is new with YARN, Platform Symphony has always featured a multitenant design.

Resource groups

Unlike Hadoop clusters, Platform Symphony does not make assumptions about the capabilities of hosts that participate in the cluster. Although Hadoop generally assumes that member nodes are 64-bit Linux hosts running Java, Platform Symphony supports various hardware platforms and operating environments. Platform Symphony allows hosts to be grouped in flexible ways into different resource groups, and different types of applications can share these underlying resource groups in flexible ways.

Applications

The term *application* can be slightly confusing as it is applied to Platform Symphony. Platform Symphony views an application as the combination of the client-side and service-side code that comprises a distributed application. By this definition, an instance of InfoSphere BigInsights might be viewed as a single application. Examples of Platform Symphony applications are custom applications that are written in C++, a commercial ISV application such as IBM Algorithmics, Calypso or Murex, or a commercial or open source Hadoop application, such as InfoSphere BigInsights or open source Hadoop.

Platform Symphony views applications as being an instance of middleware. Various client-side tools that are associated with a particular version of Hadoop (Pig, Hive, Sqoop, and so on) can all run against a single Hadoop application definition. An important concept for those not familiar with Platform Symphony is that Platform Symphony provisions service instances that are associated with different applications dynamically. As a result, there is nothing technically stopping a Platform Symphony cluster from supporting multiple instances of Hadoop and non-Hadoop environments concurrently.

Figure 6-5 shows the result of clicking **Workload → Symphony → Applications**.

The screenshot shows the IBM Platform Symphony Advanced Edition web interface. The top navigation bar includes links for Workload, Resources, System & Services, Reports & Logs, Admin, Refresh, Date (Dec 02, 2014 16:46:38 EST), and IBM. The main menu bar has options for Workload, Resources, System & Services, Reports & Logs, and a Dashboard icon. Below the main menu, a sub-menu for 'Symphony Applications in All consumers' is displayed. This sub-menu includes buttons for Enable, Disable, Run Symping, Run Executable, and Retrieve Logs. A table lists applications with columns for Application Name, Status, Session Summary, Open, Suspended, Closed, Aborted, SSM Host, SSM PID, and Blocked Hosts. One entry, 'symping7.1', is shown with 'Enabled' status and zero values across other metrics. To the right of the table are filters and options. The bottom section of the page displays detailed information for the selected application 'symping7.1'. It shows tabs for Summary, Sessions, Services, Profile, Performance, Blocked Hosts, Resources, and Package Downloads. The 'General' tab is selected, showing details like Application Name (symping7.1), Version (7.1), Status (Enabled), Session Summary (-), Blocked Hosts (0), Consumer (/SymTesting/Symping71), and SSM-related metrics (SSM Host, SSM PID, SSM Start Time, SI Startup Failures). Other collapsed tabs include Sessions and Tasks, Resource Allocation, and Configuration.

Figure 6-5 Applications menu

Application profiles

As explained before, applications in Platform Symphony are flexible and highly configurable constructs. An application profile in Platform Symphony defines the characteristics of an application and various behaviors at run time.

Figure 6-6 shows the result of clicking **Workload** → **Symphony** → **Application Profiles**.

The screenshot shows the 'Application Profiles' page in the IBM Platform Symphony Advanced Edition. The left sidebar lists 'ITSOCluster' with various services: ClusterServices, ComputeServices, HDFS, ManagementServices, MapReduceConsumer, SampleApplications, SymExec, SymTesting, and YARN. The main panel displays the 'Applications' tab for 'ITSOCluster'. A table lists one application: 'Sympling7.1' (enabled, consumer '/SymTesting/Sympling71', Symphony Version 7.1). Global Actions and Preferences buttons are at the bottom.

Figure 6-6 Application profiles

Consumers

From the viewpoint of a resource manager, an application or tenant on the cluster is defined as something that needs particular types of resources at run time. Platform Symphony uses the term *consumer* to define these consumers of resources and provides capabilities to define hierarchical consumer trees and express business rules about how consumers share various types of resources that are collected into resource groups. The leaf nodes in consumer trees map to a Platform Symphony application.

Services

Services are the portions of applications that run on cluster nodes. In a Hadoop context, administrators likely think of services as equating to a task tracker that runs map and reduce logic. Here again, Platform Symphony takes a broader view. Platform Symphony services are generic. A service might be a task-tracker that is associated with a particular version of Hadoop or it might be something else entirely. When the Hadoop MapReduce Processing framework is used in Platform Symphony, the Hadoop service-side code that implements that Task Tracker logic is dynamically provisioned by Platform Symphony. Platform Symphony owes its name to this ability to orchestrate various services quickly and dynamically according to sophisticated sharing policies.

Sessions

A session in Platform Symphony equates to the notion of a job in Hadoop. A client application in Platform Symphony usually opens a connection in the cluster, selects an application, and opens a session. Behind the scenes, Platform Symphony provisions a Platform Symphony Session Manager to manage the lifecycle of the job. A single Platform Symphony Session Manager can support multiple sessions (Hadoop jobs) concurrently. A Hadoop job is a special case of a Platform Symphony job. The Hadoop client starts a session manager that provides JobTracker functions. Platform Symphony uses the Job Tracker and task tracker code that is provided in a Hadoop distribution, but it uses its own low-latency middleware to more efficiently orchestrate these services on a shared cluster.

Repositories

Platform Symphony dynamically orchestrates service-side code in response to application demand. The binary code that comprises an application service is stored in a Platform Symphony repository. Normally for Platform Symphony applications, Platform Symphony services are distributed to compute nodes from a repository service. For Hadoop applications, code can be distributed either through the repository service, or it can be distributed through the HDFS or Spectrum Scale FPO file system.

Tasks

Platform Symphony jobs are collections of tasks. Platform Symphony jobs are managed by a session manager that runs on a management host. The session manager makes sure that instances of the needed service are running on compute nodes / data nodes on the cluster. Services instances run under the control of a Platform Symphony Service Instance Manager (SIM). MapReduce jobs in the Platform Symphony work the same way, but in this case the Platform Symphony service is essentially the Hadoop task tracker logic. On Hadoop clusters, slots are normally designated as running either map logic or reduce logic. Again in Platform Symphony, this is fluid. Because services are orchestrated dynamically, service instances can be either map or reduce tasks. This is an advantage because it allows full utilization of the cluster as the job progresses. At the start of a job, most of the slots can be allocated to map tasks while towards the end of the job the function of slots can be shifted to perform the reduce function.

Benefits of using Platform Symphony

This section describes the benefits of implementing Platform Symphony in your environment.

Highlights

- ▶ Monitoring of the cluster and Hadoop jobs
- ▶ Configuration and management of physical resources
- ▶ Failover and recovery logic for Hadoop jobs
- ▶ Reporting framework
- ▶ Enhanced Hadoop MR processing framework

Sophisticated scheduling engine

- ▶ Priority-based scheduling
- ▶ Pre-emptive scheduling
- ▶ Fair share proportional scheduling
- ▶ Threshold-based scheduling
- ▶ Task reclaim logic
- ▶ Administrative control of running jobs

Configuration and management

- ▶ Resource group/slot-based allocation
- ▶ Consumer allocation
- ▶ Shared resources and heterogeneous application support
- ▶ GUI management console
- ▶ Real-time monitoring and management of hosts: all global assets

High availability

- ▶ Failover scenarios
- ▶ Host running job tracker fails
- ▶ Host running map task fails
- ▶ Host running reduce task fails
- ▶ Job recovery
- ▶ Services failover

Enhanced MapReduce implementation

- ▶ Low latency with immediate map allocation
- ▶ Fast workload allocation
- ▶ Small impact to starting jobs
- ▶ Platform provides the tools for meeting necessary SLO and business continuity

6.3.5 Summary

Platform Symphony supports advanced multitenancy. With advanced multitenancy, customers can share a broader set of application types and scheduling patterns on a common resource foundation. Key advantages are better performance, better resource utilization, multitenancy/shared services, and agile workload scheduling.

Figure 6-7 shows how Platform Symphony supports many advanced IT products today.

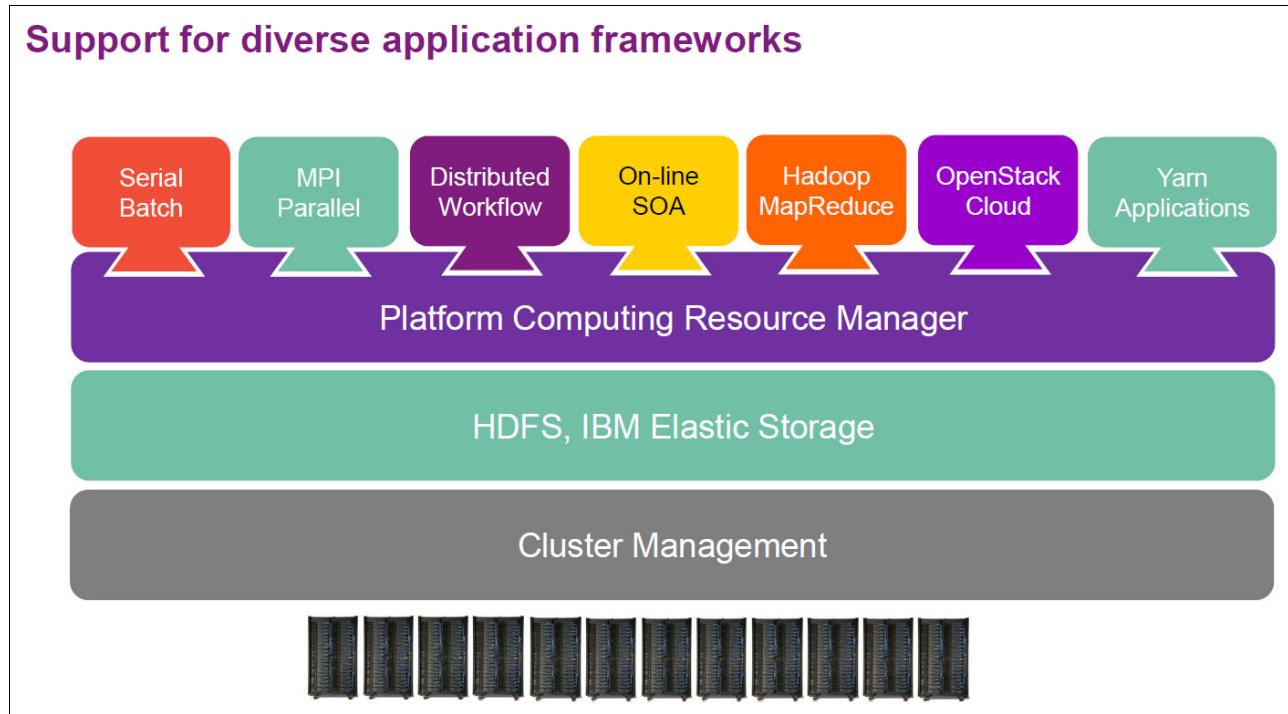


Figure 6-7 Support for diverse application frameworks

Important: Multitenant capabilities are enabled by licensing Platform Symphony Advanced Edition.

6.4 Product editions

Platform Symphony is available in four different editions that are tailored to different business requirements:

- ▶ Developer: Build and test applications without needing a full-scale grid
- ▶ Express: The ideal solution for departmental clusters
- ▶ Standard: Enterprise-class performance and scalability
- ▶ Advanced: Ideal for distributed compute- and data-intensive applications requiring Hadoop MapReduce, or benefiting from the advanced capabilities of the Application Service Controller for the Platform Symphony add-on

Table 6-1 summarizes the features that are associated with each Platform Symphony edition.

Table 6-1 IBM Platform Symphony features

Features	IBM Platform Symphony Edition			
	Developer	Express	Standard	Advanced
Low-latency HPC SOA	X	X	X	X
Agile service and task scheduling	X	X	X	X
Dynamic resource orchestration		X	X	X
Standard and custom reporting			X	X
Desktop, server, and virtual server harvesting capability			X	X
Data affinity				X
Hadoop MapReduce Processing framework	X			X

Product add-ons are optional and serve to enhance the functions of the Standard and Advanced editions. Table 6-2 shows the IBM Platform Symphony add-ons that are associated with each Platform Symphony edition.

Table 6-2 IBM Platform Symphony add-ons

Add-ons	IBM Platform Symphony Edition			
	Developer	Express	Standard	Advanced
Platform Application Service Controller				X
Desktop harvesting			X	X
Server and virtual server harvesting			X	X
Graphics processing units (GPU)			X	X
IBM Spectrum Scale			X	X
IBM Spectrum Scale-Shared Nothing Cluster (SNC)				X

6.4.1 IBM Platform Symphony Developer Edition

Platform Symphony Developer Edition (DE) provides an environment for application developers to grid-enable, test, and run their service-oriented applications.

Platform Symphony DE provides a complete test environment, simulating the grid environment that is provided by Platform Symphony. Developers can test their client and services in their own cluster of machines before deploying to the grid.

Platform Symphony DE provides the following features:

- ▶ Easy-to-use APIs and rich design patterns to seamlessly grid-enable all types of service-oriented applications with minimal changes.
- ▶ A Hadoop MapReduce Processing framework to run MapReduce applications with minimal changes. The Hadoop MapReduce Processing framework in Platform Symphony DE provides the following features:
 - Different modes for debugging MapReduce applications:
 - The stand-alone mode, in which the entire MapReduce workflow runs in a single Java process on the local host.
 - The pseudo-distributed mode, in which each MapReduce daemon runs in separate Java processes on the local host.
 - Support for distributed file systems, such as the open source Apache Hadoop Distributed File System (HDFS), Cloudera's Distribution Including Apache Hadoop (CDH), Appistry Cloud IQ, and IBM Spectrum Scale.
 - A command-line utility that is called `mrsh` that automatically sets up the environment when you submit MapReduce jobs.
 - A MapReduce service class definition that you can customize to implement custom lifecycle event handlers.
 - A Java class wrapper that defines buffers as data containers, enabling you to copy and view large files as a sequence of bytes.
- ▶ A web interface for monitoring and controlling your test environment for Platform Symphony and MapReduce workloads.
- ▶ An IBM Knowledge Center for easy access to documentation.

To run a Platform Symphony workload on the grid, the application developer creates a service package and adds the service executable file into the package: no additional code changes are required.

The Platform Symphony DE does not include the EGO resource management component. It does include an EGO stub to simulate basic EGO resource distribution.

Platform Management Console

The Platform Management Console (PMC) is your web interface to Platform Symphony and IBM Platform Application Service Controller. For Platform Symphony, the PMC provides a single point of access to the key system components for cluster and workload monitoring and control, configuration, and troubleshooting. For IBM Platform Application Service Controller, the PMC also provides a single point of access to manage and monitor your application instances.

Cluster and workload health dashboards

The Dashboard window appears when you log in to the PMC. This window provides a quick overview of the health of your cluster. It shows a summary of the workload in the cluster, a summary of hosts utilization and status, and links to key pages in the PMC.

Note: The *Dashboard* displays only when the console is used to access the grid. It does not appear in Platform Symphony DE.

The *Common Tasks* menu is available at the upper right of the Dashboard window and provides links to key pages in the PMC, such as Platform Symphony Workload, Resources, Cluster Settings, System Logs, and Reports. Based on your entitlement, extra pages, such as MapReduce workload and Application Service Controller workload, also appear.

6.4.2 IBM Platform Symphony Advanced Edition

Across a range of industries, organizations are collecting tremendous volumes of data, generated by a wide variety of sources, often at extreme velocities. Analyzing this big data can produce key insights for improving the customer experience, enhancing marketing effectiveness, increasing operational efficiencies, reducing financial risks, and more. IBM Platform Symphony Advanced Edition software can help address the challenges of achieving outstanding performance for analyzing big data while controlling costs.

Advanced Edition includes a best-of breed runtime engine for MapReduce applications that is fully compatible with popular MapReduce distributions, including Hadoop and Spark MapReduce. It delivers enterprise-class distributed computing capabilities for the MapReduce programming model. It meets enterprise IT requirements by delivering high resource utilization, availability, scalability, manageability, and compatibility. This all leads to the ability to deliver a higher quality of service that is aligned to customer service level requirements at a lower total cost.

As a platform for distributed MapReduce workloads, Platform Symphony Advanced Edition provides an open application architecture for both applications and file systems. It provides client- and server-side APIs for both MapReduce and non MapReduce applications supporting multiple programming languages. Also, its open architecture supports connections to multiple data types and storage file systems, including full compatibility with the open source Hadoop Distributed File System (HDFS). A high-level view of the architecture of Platform Symphony Advanced Edition and the Hadoop MapReduce Processing Framework is shown in Figure 6-8 on page 79.

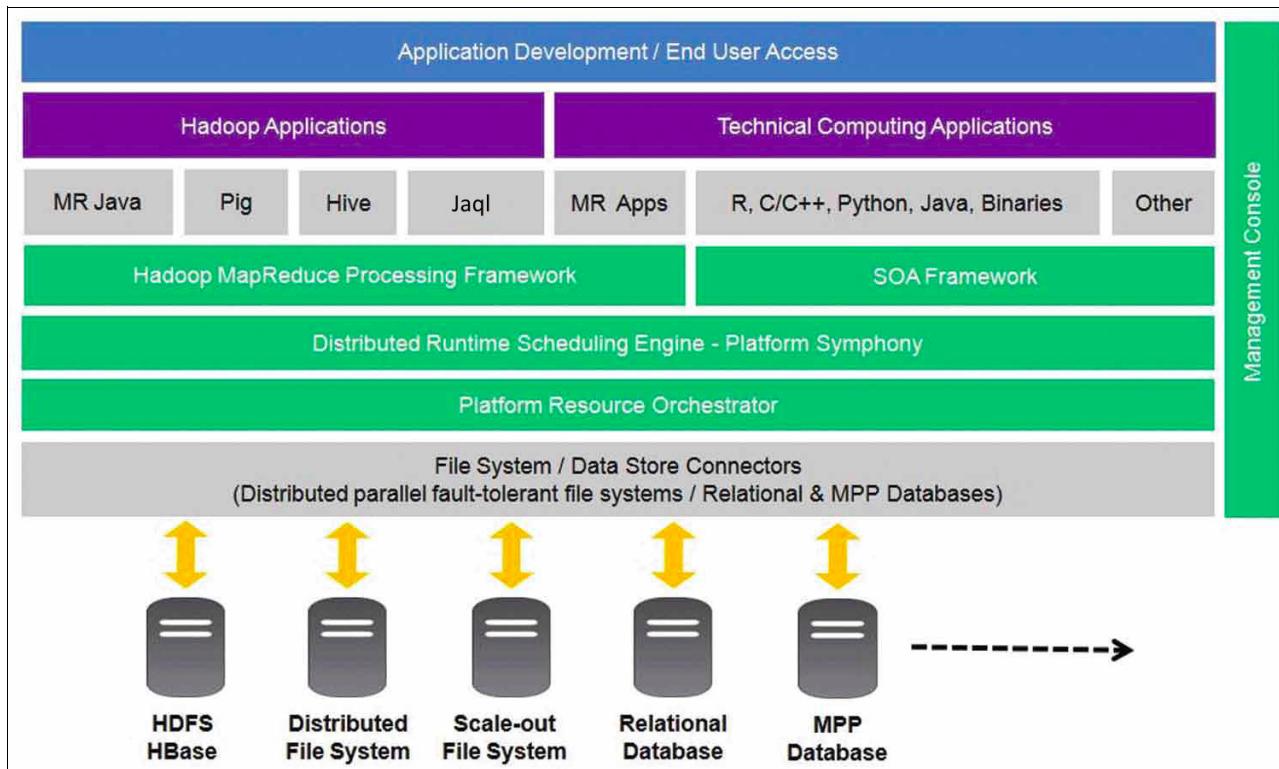


Figure 6-8 Platform Symphony Advanced Edition MapReduce framework

Platform Symphony Advanced Edition provides a variety of client- and server-side APIs to facilitate easy application integration and execution. These APIs include MapReduce APIs that are fully compatible with open source Hadoop, and various capabilities that support commercial application integrations. These services allow developers to use the open source Hadoop logic and projects and easily port the resulting applications into the Platform Symphony MapReduce architecture. It also provides developers with a much richer set of tools to avoid performance bottlenecks and optimize performance by taking advantage of advanced Platform Symphony features, such as multi-core optimization, direct data transfer, and data affinity.

6.5 Optional applications to extend Platform Symphony capabilities

Several add-on tools and complementary products can be used with both Platform Symphony Standard and Advanced Editions. They are all designed to help you do more while spending less.

- ▶ IBM Platform Symphony Desktop Harvesting: This add-on harnesses the resources from available idle desktops and adds them to the pool of potential candidates to help complete tasks. Platform Symphony services do not interfere with other applications running on the desktops, and harvested resources are managed directly through the integrated management interface.
- ▶ IBM Platform Symphony Server/VM Harvesting: To take full advantage of more of your enterprise's resources, you can use this add-on to tap idle or underutilized servers and virtual machines (VMs). Instead of requiring new infrastructure investments, Platform Symphony locates and aggregates these server resources as part of the grid whenever additional capacity is needed to handle larger workloads, or when the speed of results is critical.
- ▶ IBM Platform Symphony GPU Harvesting: To unleash the power of general-purpose graphic processing units (GPUs), this tool enables applications to share expensive GPU resources more effectively and to scale beyond the confines of a single GPU. Sharing GPUs more efficiently among multiple applications, and detecting and addressing GPU-specific issues at run time helps improve service levels and reduce capital spending.
- ▶ IBM Platform Analytics: IBM Platform Analytics is an advanced analysis and visualization tool for analyzing the massive amounts of workload and infrastructure usage data that is collected from Platform Symphony clusters. You can easily correlate job, resource, and license data from multiple Platform Symphony clusters for data-driven decision making.
- ▶ IBM Platform Application Service Controller: The Application Service Controller, available only in the Advanced Edition, extends the Platform Symphony grid to provide a shared-service backbone for a broad portfolio of distributed software frameworks. By enabling a wide variety of applications to share resources and coexist on the same infrastructure, the Application Service Controller helps organizations reduce cost, simplify management, increase efficiency and improve performance.

The next section describes the IBM Platform Application Service Controller extension.

6.6 Overview of IBM Platform Application Service Controller

IBM Platform Symphony V7.1 and IBM Platform Application Service Controller help you exceed performance goals with a fast, efficient grid and analytic computing environment. Version 7.1 offers increased scaling and performance. IBM Platform Application Service Controller enables you to better manage cloud-native distributed computing environments.

IBM Platform Application Service Controller Advanced Edition is a generalized service controller for complex, long-running application services.

IBM Platform Application Service Controller extends the Platform Symphony grid to enable a shared-service backbone for a broad portfolio of distributed software frameworks. Designed specifically to address the requirements of a new generation of distributed application workloads that stem from the wide adoption of *born-on-the-cloud* technology, it increases resource utilization, minimizes application silos, and offers increased resiliency and high availability.

IBM Platform Application Service Controller is available for Platform Symphony Advanced Edition. Application Service Controller offers the following benefits:

- ▶ Increased utilization of existing hardware resources:
 - Reduce server idle time across a broader set of distributed applications, including a new generation of cloud-native workloads
 - Share resources across applications, users, and lines of business
 - Defer the need for incremental capital investment
- ▶ Increased application performance:
 - Obtain bare metal performance with dynamic runtime elasticity: Manage demand at run time rather than at build time
 - Gain application isolation without virtual machines
 - Reduce application wait time
- ▶ Increased resiliency and high availability
- ▶ Improved management efficiencies: Reduced administration impact for visualization, monitoring, alerting, reporting, application deployment, and lifecycle management.

IBM Platform Application Service Controller Version 7.1 is supported on the following operating system platforms:

- ▶ Windows
- ▶ Linux on Power
- ▶ Linux on System x

Application Service Controller lifecycle

With IBM Platform Application Service Controller, you can create application instances by using an application template. Figure 6-9 illustrates the basic tasks that are typically associated with using IBM Platform Application Service Controller.

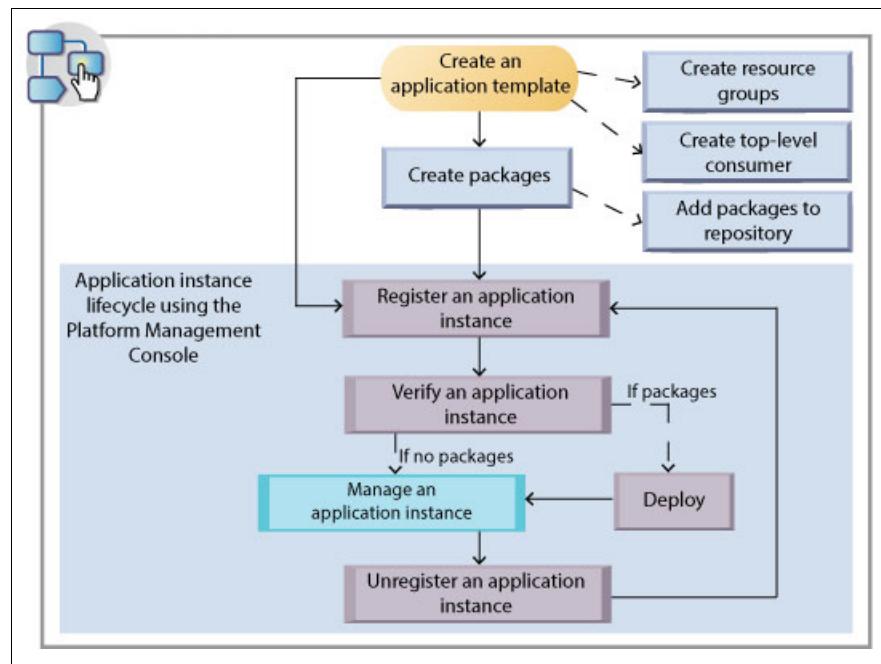


Figure 6-9 Basic tasks that are associated with IBM Platform Application Service Controller

Note: IBM Platform Application Service Controller features can be configured through a RESTful API.

First, you must create an application template. Next, you create the packages that are based on the application template. When you register an application instance, you can add the created packages to the Platform Symphony repository and you can specify consumers that you want to use.

Alternatively, if you want to define your own consumers and resource groups that are available to the application instance, you can create the resource groups, consumers, and add the packages to the repository ahead of time so that they can be used by multiple application instances.

After you register the application instance, you must verify that it was registered correctly. If the application instance has packages, you must deploy the application instance first, and then manage it.

If the application instance does not include packages, you start managing it, as there is no need for deployment. If you must update your application template, you must unregister it.

6.6.1 Application framework integrations

IBM Platform Application Service Controller can integrate with any distributed application framework to manage and run them on a scalable, shared grid.

The following applications are some of the application frameworks that are integrated with IBM Platform Application Service Controller:

- ▶ Apache Hadoop
- ▶ Apache Spark
- ▶ Apache YARN
- ▶ Cassandra
- ▶ Cloudera
- ▶ Hadoop
- ▶ Hortonworks
- ▶ MongoDB

Note: For more information and the latest set of application frameworks integrations, go to the following websites:

- ▶ <http://ibm.github.io/>
- ▶ <https://hub.jazz.net/learn/>

6.6.2 Basic concepts

To understand IBM Platform Application Service Controller, you must understand the concepts that are described in this section.

Application instance

An application instance is a collection of services and service groups that is associated with a top-level consumer. You can monitor and manage an application instance and drill down to manage the related services and service instances. You create (register) an application instance from an application template.

Application Service Controller service

Application Service Controller services can be part of an application instance or independent. If you create a service, select the type as ASC, to enable IBM Platform Application Service Controller features. An Application Service Controller service can be either stateful or stateless.

Application template

An application template is defined in YAML Ain't Markup Language (YAML) and contains all of the parameters, resources, and outputs that are required to register application instances.

Consumer

A consumer is a unit within the representation of an organizational structure. The structure creates the association between the workload demand and the resource supply.

EGO Service Controller

The EGO service controller (**egosc**) is the first service that runs on top of the EGO kernel. It functions as a bootstrap mechanism for starting the other services in the cluster. It also monitors and recovers the other services. It is analogous to **init** on UNIX systems or Service Control Manager on Windows systems. After the kernel starts, it reads a configuration file to retrieve the list of services to be started. There is one **egosc** per cluster, and it runs on the master host.

Process information manager

Process information manager (PIM) collects resource usage of the process that runs on the local host.

Platform Management Console

The PMC is your web interface to IBM Platform Application Service Controller. The PMC provides a single point of access to manage and monitor your application instances.

Resources

Resources are physical and logical entities that are used by application instances to run. Processor slots are the most important resource.

Resource group

A resource group is a logical group of hosts. A resource group can be specified by resource requirements in terms of operating system, memory, swap space, CPU factor, and so on. It can be explicitly listed by host names.

Service

A service is a self-contained business function that accepts one or more requests and returns one or more responses through a well-defined, standard interface. The service performs work for a client program. It is a component that can perform a task, and is identified by a name. Platform Symphony runs services on hosts in the cluster.

The service is the part of your application instance that does the actual calculation. The service encapsulates business logic.

Service instance

When a service is running on a host, service instances are created. When the service is stopped, there are no longer service instances.

Stateful Application Service Controller service

An Application Service Controller service that typically stores data locally to a disk on the host on which it runs. IBM Platform Application Service Controller aims to keep the service running on that host, and enables optional decommission of the service to correctly handle the data for that service when it is removed.

Stateless Application Service Controller service

An Application Service Controller service that does not store data locally on the host on which it runs. A stateless service can be safely restarted on a different host if necessary. By default, all Application Service Controller services are stateless.

6.6.3 Key prerequisites

To deploy IBM Platform Application Service Controller, a Platform Symphony Advanced Edition license is required. In addition, you must have the following prerequisites:

- ▶ A physical grid computing environment that consists of any of the following servers:
 - IBM Power Systems
 - IBM PureSystems®
 - Similar servers from third-party companies
- ▶ Cluster nodes that are preinstalled with supported operating environments
- ▶ Cluster nodes that are connected through a fast Internet Protocol network infrastructure
- ▶ Management hosts on the cluster that ideally share a common network file system (enable recovery of grid sessions in case of failure)

Hardware requirements

Platform Symphony V7.1 is supported on Lenovo System x iDataPlex and other rack-based servers, and non IBM x64 servers. Also supported are IBM Power Systems servers running PowerLinux™ operating environments. PowerLinux support is for Big Endian only.

IBM Power System servers running AIX can integrate with Platform Symphony, but from a client perspective only.

Other platforms include the following ones:

- ▶ Microsoft Windows 64-bit
- ▶ Linux x86-64
- ▶ Linux on IBM POWER
- ▶ Solaris x86-64
- ▶ IBM AIX 64-bit: C++ software development kit (SDK) and Java client
- ▶ SPARC Solaris 10-64: C++ and Java SDK
- ▶ Co-Processor Harvesting: Client, SDK, and compute nodes

Software requirements

Here is a high-level summary of operating environments that are supported by Platform Symphony:

- ▶ Microsoft Windows Server 2008 SE, 2008 EE, 2008 R2 SE, and 2008 R2 EE (64-bit)
- ▶ Windows HPC Server 2008 and 2008 R2 (64-bit)
- ▶ Windows Server 2012 Standard and Datacenter, and 2012 R2 Standard and Datacenter (64-bit)
- ▶ Windows 7 and 8 (64-bit)
- ▶ Red Hat Enterprise Linux (RHEL) AS 5, 6, 6.4, and 6.5 (x86-64)

- ▶ RHEL AS 7 (x86-64)
- ▶ CentOS 6 (x86-64)
- ▶ SUSE Linux Enterprise Server (SLES) 10 and 11 (x86-64)
- ▶ SLES 11 SP2 (x86-64)
- ▶ RHEL on IBM POWER6®, POWER7®, and POWER8™
- ▶ Oracle Solaris SPARC 10-64 and x86-64 11 (64-bit) (with limitations)
- ▶ IBM AIX V7.1 (64-bit) (with limitations)

IBM Knowledge Center: You can find additional details in the *Supported System Configurations* document at the following IBM Knowledge Center website:

<http://www.ibm.com/support/knowledgecenter/>

6.6.4 IBM Platform Application Service Controller: Application templates

IBM Platform Application Service Controller provides application template samples to help you create your own application templates.

The following Application Service Controller application template samples are available to customize:

- ▶ `asc_sample_minimal.yaml`
- ▶ `asc_sample.yaml`

There are also application template samples that are available to customize that are specific to the following application frameworks:

- ▶ Ambari
- ▶ Cassandra
- ▶ Hadoop 2.4.1
- ▶ Hadoop 2.4.1 with Docker support
- ▶ MongoDB
- ▶ Spark
- ▶ Spark and HDFS with Docker support
- ▶ ZooKeeper

All of the application template samples are available in the following directory:

`$EGO_CONFDIR/../../asc/conf/samples.`

Note: To use the application template samples, IBM Platform Application Service Controller must be started as the root user. Changes in the scripts are required if another user is used to start IBM Platform Application Service Controller.

6.7 IBM Platform Symphony application implementation

Platform Symphony service-oriented applications consist of a client application and a service. When the application runs, a session is created that contains a group of tasks. The application profile provides information about the application.

This section provides details about how to deploy Platform Symphony.

6.7.1 Planning for Platform Symphony

This section describes the necessary planning steps for Platform Symphony.

For this book, the environment is configured as a mixed-cluster environment with IBM Platform Symphony V7.1 on four virtual machines and a test Hadoop MapReduce Processing framework with WordCount application, which is suitable for production use or small-scale application testing.

Components of the solution

Here are the solution components:

- ▶ Two virtual guests that are hosted by X3850X5 7145-AC1 running VMware ESXi
- ▶ Two virtual guests that are hosted by IBM PowerLinux 7R2 8246-L2C
- ▶ Red Hat Enterprise Linux Server release 6.5 (Santiago) x86_64-bit
- ▶ Apache Hadoop release 1.1.1 x86_64-bit
- ▶ IBM Platform Symphony V7.1 x86_64-bit
- ▶ IBM Spectrum Scale V4.1.0 with Elastic Storage (based upon IBM General Parallel File System or GPFS technology) V4.1.0 with fix pack GPFS_STD-4.1.0.4 x86_64-bit
- ▶ WordCount v1.0
- ▶ Oracle Java jre-6u45-linux-x64 on Intel based platform and ibm-java-jre-6.0-16.2-linux-ppc64 on IBM PowerLinux

Installation prerequisites

You must set some variables and fulfill the following prerequisites before you can start installing the Platform Symphony V7.1:

- ▶ Choose the root operating system account for installation. This choice provides the flexibility to use different execution accounts for different grid applications.
- ▶ Set the grid administrator operating system account to `egoadmin` by running the following command. This account was created in the Lightweight Directory Access Protocol (LDAP) before starting the installation process.
`useradd egoadmin`
- ▶ Grant root privileges to the cluster administrator, and set up the cluster and a host. You should see the following message:
A new cluster <ITSOCluster> has been created. The host <pw4302-12> is the master host.
- ▶ You must increase cluster scalability if the number of processors in the cluster, plus the number of client connections to the cluster, exceeds 1000. As root, run the following command:
`/opt/ibm/platformsymphony/profile.platform`

- ▶ Add this line to the `/etc/security/limits.conf` file:
`* hard nofile 6400`
- ▶ Before you start the EGO configuration, you must connect your IP with your host name in `/etc/hosts` as root to avoid host name errors. Shut down the `iptables` service to avoid connection failures.
- ▶ Set the following variables:
`export CLUSTERADMIN=egoadmin`

```
export CLUSTERNAME=ITSOCluster  
export JAVA_HOME=/usr/java/latest  
export SIMPLIFIEDWEM=N
```

- ▶ To run **egoconfig** and complete the cluster configuration, you must log in as egoadmin. The cluster uses configuration files under the directory indicated by (EGO_CONFDIR/.../...). The value of the environment variable EGO_CONFDIR changes if the cluster keeps configuration files on a shared file system. When your documentation refers to this environment variable, substitute the correct directory.
- ▶ Configure a mixed cluster environment for IBM Platform Symphony: This is done on PowerKVM and Intel Linux platforms. For this scenario, configure the first host as the master host and the second host as the master candidate for failover. Do not set the cluster to do failover so that the cluster uses configuration files under the installation directory:
EGO_CONFDIR=EGO_TOP/kernel/conf
- ▶ Enable automatic start, grant root privileges to the cluster administrator, and start EGO.

The following settings were used for the Platform Symphony installation in this book:

Workload Execution Mode (WEM): Advanced
Cluster Administrator: egoadmin
Cluster Name: ITSOCluster
Installation Directory: /opt/ibm/platformsymphony
Connection Base Port: 7869

After installation, you can run **egoconfig setbaseport** on every host in the cluster to change the ports that are used by the cluster.

If you want to add more compute hosts, you must follow the Installation on PowerKVM or the Installation on Intel Linux.

Ports

The default base port that is used by Platform Symphony is 7869. Use the default value unless you have systems that run other services through that port. Platform Symphony requires seven consecutive ports that start from the base port, for example, 7869 - 7875. Ensure that all ports in that range are available before installation.

Important: On all hosts in the cluster, you must have the same set of ports available.

If you must set a different base port, use the **BASEPORT** environment variable when you define the cluster properties for installation. For example, to use 17869 as the base port, define **BASEPORT=17869** in the **install.config** file.

Platform Symphony also requires more ports for services and daemons. Table 6-3 describes the required ports for each service.

Table 6-3 Additional port requirements

Service	Required ports
Web server	8080, 8005, and 8009
Service director	53
Web service	9090

Loader controller	4046
Derby database	1527

Workload execution mode

At installation, it is necessary to decide whether a single user (non-root) is the primary user of the grid. If so, use the Simple Workload Execution Mode (WEM) approach where the Platform Symphony applications run under one user account.

Otherwise, to provide better flexibility to allow different applications and users to run applications from the grid, use the Advanced WEM approach. Platform Symphony applications run under the workload execution account of the consumer, which is a configurable account. Different consumers can have different workload execution accounts.

Do not let the Advanced name discourage you from using this installation because the default values from Platform Symphony can run most workloads.

Cluster name

You must customize the installation if you want to specify your own unique cluster name. Do not use a valid host name as the cluster name.

Important: The cluster name is permanent; you cannot change it after you complete the installation.

To specify the cluster name and not use cluster1, set the environment variable CLUSTERNAME=<Name>.

Multi-head installations

Platform Symphony requires a configuration parameter named OVERWRITE_EGO_CONFIGURATION. If this parameter is set to Yes (the default is No), the Platform Symphony default configuration overwrites the EGO configuration. For example, it overwrites EGO ConsumerTrees.xml, adds sd.xml in the EGO service conf directory, and overwrites the EGO Derby DB data files.

If you plan a multi-head cluster (a cluster that runs both Platform Symphony and IBM Platform Load Sharing Facility (LSF)), it is acceptable for IBM Platform LSF and Platform Symphony workloads to share EGO resources in the cluster. In this case, you must avoid overwriting the EGO configuration.

The environment that is planned in this section is single-headed, so ensure that the variable OVERWRITE_EGO_CONFIGURATION is set to Yes.

Software packages

Ensure that you have all the required software packages and entitlement files available, as shown in Table 6-4.

Table 6-4 Software packages and entitlement file list

Type	File name
Platform Symphony package	symSetup7.1.0_1nx26-1ib23-x64.bin
EGO package	ego-1nx26-1ib23-x64-3.1.0.rpm
SOAM package	soam-1nx26-1ib23-x64-7.1.0.rpm

6.7.2 Accessing the Platform Symphony Management Console

The Platform Symphony console is on the same host if you follow the installation recommendations in this chapter. Port 8080 is the default. You can log in to the Platform Symphony management console at the following address:

`http://<master-host>:8080/platform`

The default administrator login for Platform Symphony is "Admin / Admin". Figure 6-10 shows the Platform Symphony V7.1 login window.



Figure 6-10 IBM Platform Symphony login window

In production clusters, there normally are multiple Platform Symphony management hosts. Setting up multiple hosts is covered in 6.7.3, “Configuring a cluster for multitenancy” on page 90. For more information, see the *Platform Symphony V7.1 Installation Guide*, found at: http://www-01.ibm.com/support/knowledgecenter/SSGSMK_7.1.0/sym_kc/sym_kc_installing.dita?lang=en

If you are having trouble connecting to the Platform Symphony web console, run the following command:

```
egosh service view WEBGUI
```

This command shows details about the web service.

The WEBGUI services should be started automatically by EGO, but if it becomes necessary to start or stop the service, you can run the following command:

```
egosh service start WEBGUI  
egosh service stop WEBGUI
```

To log on to the EGO service, run the following command

```
egosh logon
```

Enter Admin / Admin as the user name and the password when you are prompted:

The WEBGUI service is implemented by using Apache TomCat. If there are problems with the WEBGUI, you can inspect the logs at \${EGO_TOP}/gui/logs/catalina.out for information about what might be wrong with the service.

If you cannot connect to the Platform Symphony console, this might be because of your firewall configuration. You can disable your firewall temporarily to see whether this is the cause by running the following command:

```
service iptables stop
```

After a user logs in to the Platform Symphony console on port 8080, the user sees the main Platform Symphony dashboard. This view is mostly used to monitor the high-level status of the various applications and tenants on a Platform Symphony cluster. Figure 6-11 illustrates the main dashboard view.

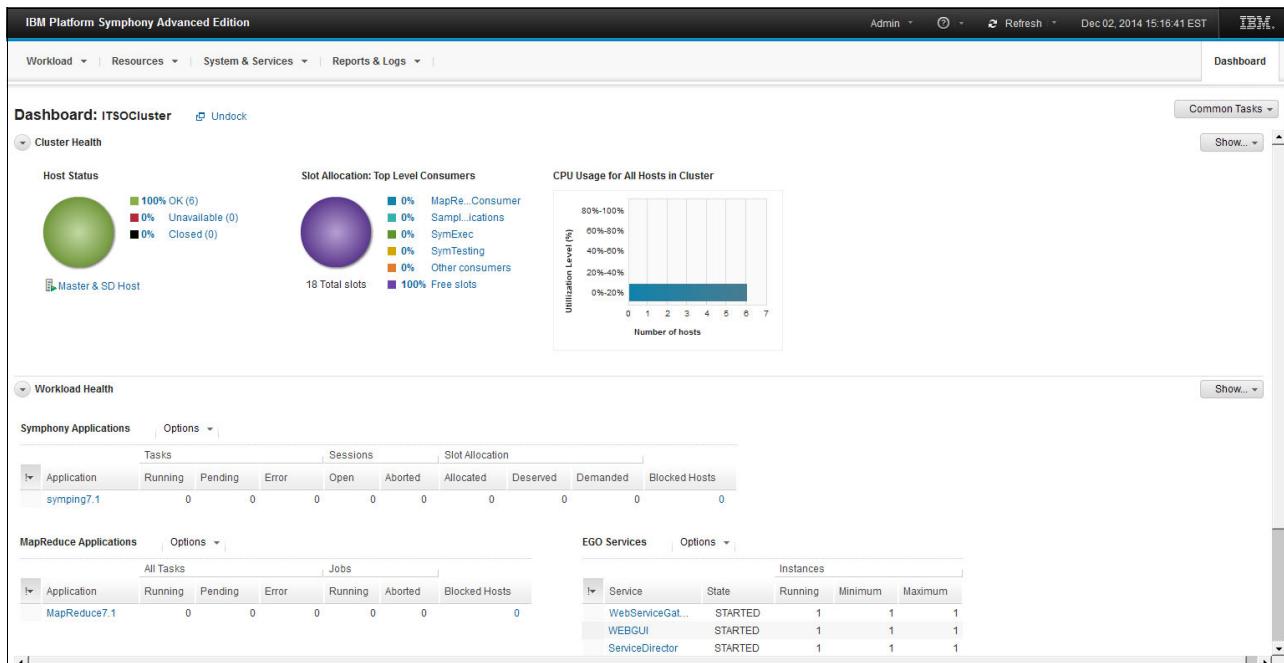


Figure 6-11 Platform Symphony dashboard view after login

6.7.3 Configuring a cluster for multitenancy

Platform Symphony has two different workload execution modes:

- ▶ Simple Workload Execution Mode
- ▶ Advanced Workload Execution Mode

This is normally an installation option with Platform Symphony. Enterprise Edition installation automatically installs Platform Symphony in Advanced Workload Execution Mode (WEM). In Advanced WEM, core Platform Symphony services run as root, and application administrators can control the user ID under which clustered applications run. Platform Symphony is frequently deployed in secure environments, where these capabilities are important.

Configuring OS groups for the multitenant environment

For users that use Platform Symphony (both named users and the user IDs that applications run through impersonation), these IDs must be part of the OS group that owns the Platform Symphony installation.

Users and security

To allow users to use resources when running their applications in a managed way, Platform Symphony implements a hierarchical model of consumers. This tree of consumers allows association of users and roles on one hand with applications and grid resources on the other. Policies for the distribution of resources among multiple applications that are run by different users can be configured this way to share the resources in the grid. MapReduce applications and other non MapReduce applications, such as standard SOA compute-intensive applications inside Platform Symphony, can use the same infrastructure. In addition, a multi-head installation of both Platform LSF and Platform Symphony is supported. This installation allows batch jobs from LSF, and compute-intensive and data-intensive applications from Platform Symphony to share the hardware grid infrastructure. A security model is enforced for the authentication and authorization of various users to the entitled applications and to isolate them when they try to access the environment.

You can create user accounts inside the Platform Symphony environment, as shown in Figure 6-12, and then assign them to either predefined or user created roles. User accounts include optional contact information, a name, and a password.

The screenshot shows the IBM Platform Symphony Advanced Edition interface. On the left, the 'Dashboard: ITSOCluster' is visible with sections for Cluster Health (green), Host Status (purple), and Workload Health. The 'System & Services' menu is open, showing 'Symphony Entitlement Cluster' with 'Users' selected, which then has 'Accounts' selected. A sub-menu 'User Accounts' is displayed with a 'Create New User Account' button highlighted. To the right, a modal window titled 'User Property - Mozilla Firefox IBM Edition' is open, showing fields for 'User Name:' (empty), 'Specify optional attributes' (Email, Telephone, Description), 'Specify a password (*)' (Password, Confirm Password), and a 'Required fields' section. An arrow points from the 'Create New User Account' button on the main dashboard to the 'Create New User Account' button in the modal window.

Figure 6-12 Creating user accounts

Platform Symphony has four predefined user roles that can be assigned to a user account:

- ▶ Cluster administrator
- ▶ Cluster administrator (read only)

A user with this role can perform any administrative or workload-related task, and has access to all areas of the Platform Management Console and to all actions within it.

This user role allows read-only access to any cluster information, but cannot perform any add, delete, or change action.

- ▶ Consumer administrator

Users with this role are assigned to a top-level consumer in the consumer hierarchy, and can administer all subconsumers in that branch of the tree.

- ▶ Consumer user

Consumer users are assigned to individual consumers on the tree, and have access and control only over their own workload units.

To submit a workload for an enabled application, a user must have the appropriate roles and permissions. When a user account is added to more roles, the permissions are merged. To configure such a setup, you need an administrator role with the correct permissions.

Sharing resources

An application can be used only after it is registered and enabled. You can register an application only at a leaf consumer (a consumer that has no subconsumers). Only one application can be enabled per consumer. Before you can register an application, you must create at least one consumer, and deploy the service package of the application to the intended consumer. You can deploy the service package to a non-leaf consumer so that all applications registered to child leaf consumers can share the service package. A service package is created that puts all developed and compiled service files and any dependent files that are associated with the service in a package.

Resource distribution plan

In this step, you relate the resources themselves to the consumer tree and introduce the resource distribution plan that details how the cluster resources are allocated among consumers. The resource orchestrator distributes the resources at each scheduling cycle according to this resource distribution plan. The resource plan accounts for the differences between consumers and their needs, resource properties, and various policies about consumer ranking or prioritization when allocating resources.

You must initially assign bulk resources to consumers in the form of resource groups to simplify their management. Later, you can change this assignment. Resource groups are logical groups of hosts. A host in a resource group is characterized by a number of slots. The number of slots is a variable parameter. When you choose a value for it, the value must express the degree of specific workload that the host can serve. A typical slot assignment is, for example, the allocation of one slot per processor core.

After it is created, a resource group can be added to each top-level consumer to make it available for all the other subconsumers underneath. Figure 6-13 on page 93 shows an example of a consumer tree with all its top-level consumers and their assigned resource groups and users. Platform Symphony provides a default top-level consumer, MapReduceConsumer, and a leaf-consumer.

Consumer Name	Resource Group	Roles	Users	User Groups
PClusterServices	InternalResourceGroup	Consumer Admin	Admin	Actions
PComputeServices	MapReduceInternalResourceGroup	Consumer Admin	Admin	Actions
PHDFS	DataNodeRG, NameNodeRG, SecondaryNodeRG	Consumer User, Consumer Admin	Guest, Admin	Actions
PManagementServices	ManagementHosts	Consumer Admin	Admin	Actions
PMapReduceConsumer	ComputeHosts, ManagementHosts	Consumer User, Consumer Admin	Guest, Admin	Actions
PSampleApplications	ComputeHosts, ManagementHosts	Consumer Admin	Admin	Actions
PSymExec	ComputeHosts, ManagementHosts	Consumer Admin	Admin	Actions
PSymTesting	ComputeHosts, ManagementHosts	Consumer Admin	Admin	Actions
PYARN	InternalResourceGroup, ManagementHosts	Consumer User, Consumer Admin	Guest, Admin	Actions

Figure 6-13 Platform Symphony consumer tree

The concepts that are used inside a resource distribution plan are ownership, borrowing and lending, sharing, reclaiming of borrowed resources, and rank:

- ▶ Ownership: The ensured allocation of a minimum number of resources to a consumer.
- ▶ Borrowing and lending: The temporary allocation of owned resources from a lending consumer to a consumer with an unsatisfied demand.
- ▶ Sharing: The temporary allocation of unowned resources from a “share pool” to a consumer with an unsatisfied demand.
- ▶ Reclaiming: Defines the criteria under which the lender reclaims its owned resources from borrowers. The policy can specify a grace period before starting the resource reclamation, or the policy can specify to stop any running workload and reclaim the resources immediately.
- ▶ Rank: The order in which policies are applied to consumers. Rank determines the order in which the distribution of resources is processed. The highest ranking consumer receives its resources first, borrows resources first, and returns borrowed resources last.

Figure 6-14 shows the resource plan.

Consumer		Model type: Ownership		Model type: Share	
		Owned Slots	Consumer Rank	Lend Limit	Borrow Limit
ITSOCluster	SymTesting	18	0	<input checked="" type="checkbox"/>	1
	Sympling71	0	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>
	Total	0	-	-	-
SampleApplications	Balance	0	-	-	-
	SOASamples	0	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>
	EclipseSamples	0	50	<input type="checkbox"/>	<input checked="" type="checkbox"/>
SymExec	Total	0	-	-	-
	Balance	0	-	-	-
	SymExec71	0	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>
MapReduceConsumer	Total	0	-	-	-
	Balance	0	-	-	-
	MapReduce71	0	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Total	Balance	0	-	-	-
	Balance	18	-	-	-

Figure 6-14 Resource plan

The first allocation priority is to satisfy each consumer's reserved ownership. Remaining resources are then allocated to consumers that still have demand. Unused owned resources from consumers willing to lend them are then allocated to demanding consumers that are entitled to borrow them. The resource orchestrator then allocates the unowned resources from the share pool to consumers with unsatisfied demand and entitled to this type of resources. The resources from the "family" pool (any unowned resources within a particular branch in the consumer tree) are allocated first. After the family pool is exhausted, the system distributes resources from other branches in the consumer tree. The free resources in the shared pools are distributed to competing consumers according to their configured share ratio. A consumer that still has unsatisfied demand and has lent out resources reclaims them at this stage.

Owned resources are reclaimed first, followed by the entitled resources from the shared pool that is used by consumers with a smaller share-ratio. This is the default behavior. The default behavior can be changed so that owned resources are recalled first before trying to borrow from other consumers.

The resource orchestrator updates the resource information at a frequency cycle that is determined by EGO_RESOURCE_UPDATE_INTERVAL in ego.conf. Its default value is 60 seconds. At each cycle, the resource orchestrator detects any newly added resource or unavailable resource in the cluster, and any changes in workload indexes for the running jobs.

Each resource group must have its own plan. Also, you can define different resource plans for distinct time intervals of the day, allowing you to better adapt them to workload patterns. At the time interval boundary, the plan change might determine important resource reclamation.

Enabling Platform Symphony repository services

By default, when Platform Symphony is installed, the repository service in Platform Symphony is disabled. The function of the repository service is to store the application services and distribute the code that implements services dynamically to service instances on the cluster.

The MapReduce framework in Platform Symphony by default distributes the application service code (specifically the application logic that implements the task tracker function and JAR files that implement map and reduce logic) by copying them to HDFS with a high block replication factor so that the files are accessible on all nodes.

If you are planning to add and remove application profiles in Platform Symphony or consumers, you must start the Platform Symphony repository service. Otherwise, you encounter errors because some of these services assume that the repository service in Platform Symphony is running. This task can be done through the web interface by clicking **System & Services → EGO Services → Services**. This action shows a list of system services that EGO is managing. Figure 6-15 illustrates the system services view.

Name	State	# of Inst	Min Inst	Max Inst	RG	Host Name	Actions
DataNode	DEFINED	0	1	5000	-	-	Actions
derbydb	STARTED	1	1	1	ManagementHosts	pw4302-i2	Actions
GPFSSmonitor	DEFINED	0	1	1	ManagementHosts	-	Actions
MRSS	STARTED	6	1	5000	MapReduceInternalResourceGroup	pw4302-i2, ...	Actions
NameNode	DEFINED	0	1	1	-	-	Actions
plic	STARTED	1	1	1	ManagementHosts	pw4302-i2	Actions
purger	STARTED	1	1	1	ManagementHosts	pw4302-i2	Actions
RS	STARTED	1	1	1	ManagementHosts	pw4302-i2	Actions
RSA	STARTED	6	1	10000	InternalResourceGroup	pw4302-i2, ...	Actions
SD	STARTED	1	1	1	ManagementHosts	pw4302-i2	Actions
SecondaryNode	DEFINED	0	1	1	-	-	Actions
ServiceDirector	STARTED	1	1	1	ManagementHosts	pw4302-i2	Actions

Figure 6-15 System services

6.7.4 Adding an application / tenant

Fundamental to the design of open source Hadoop is the idea that there is only a single instance of a Hadoop cluster. Platform Symphony supports multiple applications that share a cluster. It is also flexible enough to support multiple instances of an application environment.

You might want to add the following tenants:

- ▶ A native Platform Symphony application that is written to the Platform Symphony APIs
- ▶ A batch-oriented workload (when Platform LSF is installed as an add-on to Platform Symphony)
- ▶ A distinct Hadoop MapReduce Processing framework environment

- ▶ Third-party applications
- ▶ A separate Hadoop MapReduce application instance that shares resources between applications but shares Hadoop binary files and a file system instance.

Click **Workload → MapReduce → Application profiles**. The Add Application window opens, as shown in Figure 6-16.

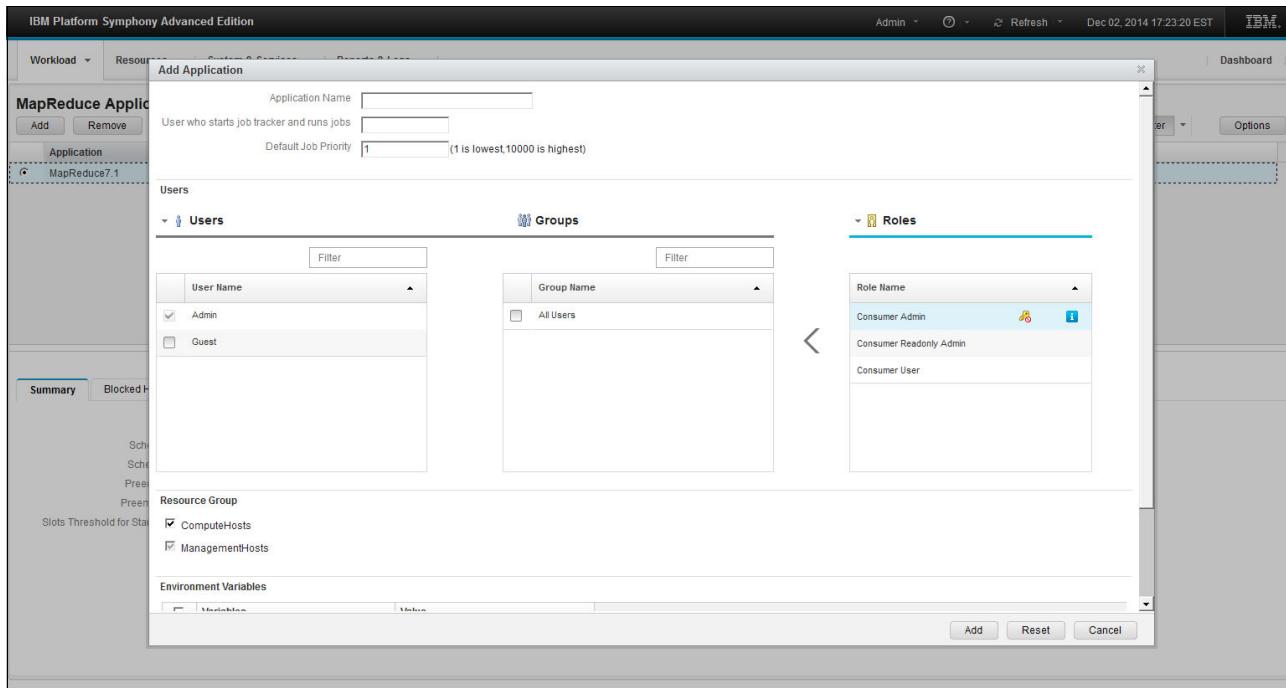


Figure 6-16 Add Application window

There is already an application profile that is defined for MapReduce. It is installed automatically with Platform Symphony. To add an application profile to support a new tenant, click **Add**.

The following parameters must be completed:

- ▶ Application name
- ▶ The user ID that starts the Job Tracker and runs jobs. This is the impersonation feature. You must define under what operating system ID that the application will run.
- ▶ Platform Symphony has 10,000 priority levels. By default, you can go to submit your application jobs as having a low priority and increase it if necessary.
- ▶ Configure user accounts that have access to this application. You should provide all users in a specialized group access to the application along with the named operating system and Platform Symphony users.

Based on this information, Platform Symphony adds an application with a set of reasonable defaults for a Hadoop MapReduce job.

The next step is to edit the configuration of the tenant as necessary to suit the unique needs of the application. Click **Workload → MapReduce → Application Profiles**, where you can define as many separate applications as you want.

6.7.5 Configuring application properties

When new applications profiles are created for each new application, a default template is used to represent reasonable settings for a MapReduce workload. The next step is to configure application profiles to meet the unique requirements of each application workload.

Application profiles are covered in detail in *Managing the Platform Symphony Cluster And Application*, which can be found at the following IBM Knowledge Center:

http://www-01.ibm.com/support/knowledgecenter/SSGSMK_7.1.0/sym_kc/sym_kc_managing_cluster_application.dita?lang=en

To configure application properties for Sqoop, modify the application profile by clicking **Workload** → **MapReduce** → **Application Profiles** from the top menu of the MapReduce applications window. Select the application profile definition for the application that was created earlier and select **Modify**.

A new window opens that allows detailed settings for the application to be changed. The web interface affects the application service profile definitions that are stored in the \$EGO_TOP/data/soam/profiles directory on the Platform Symphony master host. Enabled profiles are in a subdirectory called “enabled” and disabled profiles are in a directory called “disabled”.

The first tab in the interface, which is called Application Profile, is where you adjust application profile settings. The second tab, which is called Users, is where you can modify the users and groups that have access to the application profile.

Some important tips about application profiles:

- ▶ Application profile names must be unique.
- ▶ An application profile can be associated with only a single consumer.
- ▶ In the consumer tree, MapReduce applications are by default placed under the MapReduceConsumer tree.

The application profile can be viewed in an Advanced Configuration, a Basic Configuration, or in a Dynamic Configuration Update mode.

In the General settings area, there are settings such as where metadata that is associated with jobs and job history are stored, the default service definition to be used (MapReduce for MapReduce applications), and resource requirements.

The Platform Symphony application profile definition provides precise control over how MapReduce workloads run, and this is useful to advanced users.

A nice feature of Platform Symphony is that because the execution logic is provisioned dynamically, slots are interchangeable between mappers and reducers. Settings allow this situation to be configured along with preferences for default ratios between mappers and reducers and precise configuration on a per resource group basis.

In Platform Symphony, multiple service definitions can exist for each application, and the service definition section provides granular control over this capability. This is a useful for applications that are written to Platform Symphony native APIs and might be useful for Hadoop developers. Platform has already implemented a service that is called *RunMapReduce* that is started by service-instance managers to handle MapReduce workloads. The process of starting this service is automatic for the MapReduce service.

Heterogeneous applications support

Platform Symphony supports heterogeneous applications. It does not matter whether application clients or services are written in C/C++, Java, scripting languages, or even C# in Microsoft .NET environments. The versatility to handle all types of workloads is what makes Platform Symphony powerful as a multitenant environment.

Another unique capability that Platform Symphony brings to Hadoop is the notion of *recoverable sessions*. This concept does not exist in open source Hadoop, where the JobTracker is implemented in a simplistic way. If the JobTracker fails at run time in standard Hadoop, the job must be restarted.

The Platform Symphony SOAM middleware has long supported the notion of journaling transactions so that Hadoop MapReduce jobs become inherently recoverable. If the software service running the JobTracker logic fails (and restarts on the same host or a different host), the Platform Symphony job can recover from where it left off. This is a major advantage for customers that have long-running Hadoop jobs that must complete within specific batch windows.

This and other points of configurability are important for specific workloads. As another example, if you have execution logic where the reducer is multithreaded, you can control the ratio of reducer services to slots so that a reducer has multiple slots of which it can take advantage.

6.7.6 Associating applications with consumers

In the Platform Symphony architecture, resources are not allocated to the applications directory. They are allocated to consumer definitions that in turn map to applications.

This is an important distinction because although the application space is *flat* (if you have multiple applications and flavors of applications of different types), the structure of consumers is hierarchical. Most organizational structures are hierarchical:

- ▶ A bank might have several lines of business, each with various departments or application groups.
- ▶ A service provider might have multiple tenant customers, and might provide different application services for each tenant.
- ▶ A government agency might have different divisions, each running different applications with a particular need to segment data access.

Platform Symphony allows consumer trees to be set up in flexible ways to accommodate the needs of almost any organization. A key concept to understand is that the leaf-nodes of consumer trees are linked to the application definitions.

To view consumer definitions, from the MapReduce window in Platform Symphony, click **Resources** → **Resource Planning** → **Consumers**. This is the interface that is used to manage the consumer tree.

Setting up the consumer tree is reasonably straightforward. The left side pane is used to control where you are on the tree and the right side of the interface allows you to perform operations relative to that segment on the tree. Note the hierarchical notion of consumers in Platform Symphony.

Advanced users might find it easier to edit manually the consumer tree. Platform Symphony stores consumer tree definitions in the following file:

\$EGO_TOP/kernel/conf in the file ConsumerTrees.xml

If you manually edit this file, you must restart EGO services to bring the web-based view into synchronization with the actual contents of the XML files where these settings persist.

After editing the `ConsumerTrees.xml` file, while logged in as the cluster administrator, stop and restart EGO services to make sure that changes are reflected in the Platform Symphony console.

6.7.7 Summary

This section described a customer use case involving a multitenant implementation of IBM Platform Symphony that permits the following situations:

- ▶ Concurrent execution of different Hadoop applications (including different versions of code) on the same physical cluster.
- ▶ Dynamic sharing of resources between tenants in a fashion that maximizes performance and resource utilization while respecting individual SLAs.
- ▶ Support for applications other than Hadoop MapReduce to maximize flexibility and allow capital investments to be repurposed for multiple requirements.
- ▶ Security isolation between tenants, removing a major barrier to sharing in many commercial organizations.

These advances are significant. While Hadoop is advancing, competing open source and commercial distributions are many years away from offering true multitenancy and practical solutions for supporting multiple workloads on a shared infrastructure.

The economic arguments in favor of resource sharing are compelling. Analytic applications are increasingly composed of multiple software components that rely on distributed services. Rather than deploying separate *silos* of application infrastructure, Platform Symphony provides the option to consolidate these different application instances on a common foundation, thus increasing infrastructure utilization, boosting service levels, and helping reduce costs.

6.8 Overview of Apache Spark as part of the IBM Platform Symphony solution

The focus of this section is to describe a Platform Symphony feature that is called Spark as Adaptive MapReduce that users might choose to deploy at the time of installation.

The earlier releases of IBM Platform Symphony Advanced Edition include an Apache Hadoop-compatible MapReduce implementation that is optimized for low latency, reliability, and resource sharing, which has been demonstrated, in an audited benchmark, to deliver on average four times the performance of open source Hadoop.

IBM Platform Symphony MapReduce

IBM Platform Symphony MapReduce is an enterprise-class distributed runtime engine that integrates with open source and commercial, for example, IBM InfoSphere BigInsights and Cloudera CDH3, Hadoop-based applications. The IBM Platform Symphony MapReduce Framework addresses several challenges that typical Hadoop clusters experience. With it, you can incorporate robust HA features, enhanced performance during job initiation, sophisticated scheduling, and real-time resource monitoring. Typically, stand-alone Hadoop clusters, which are often deployed as resource silos, cannot function in a shared services model. They cannot host different workload types, users, and applications.

IBM Platform Symphony V7.1 revolutionizes big data analysis through the Apache Spark platform. Apache Spark is a general-purpose cluster computing system, a processing engine for Hadoop data that is built around speed, ease of use, and sophisticated analytics. It provides high-level APIs in Java, Scala, and Python, and an optimized engine that supports general execution graphs. It also supports a rich set of higher-level tools, including Spark SQL for SQL and structured data processing, MLlib for machine learning, GraphX for graph processing, and Spark Streaming.

In addition to simple “map” and “reduce” operations, Spark supports SQL queries, streaming data, and complex analytics, such as machine learning and graph algorithms ready for use. Better yet, users can combine all these capabilities seamlessly in a single workflow.

6.8.1 Hadoop implementations in IBM technology

Hadoop is the *de facto* standard for large-scale data processing across nearly every industry and enterprise, with numerous vendors providing Hadoop “distributions” that are coupled with enterprise-grade support services.

In 2009, the IBM Information Management division created a Hadoop implementation that is called InfoSphere BigInsights that includes Apache Hadoop and various other open source components, and IBM-developed tools that are aimed at simplifying management, application development, and data integration. Although InfoSphere BigInsights customers continue to use the Hadoop MapReduce API and higher-level tools such as Pig, Hbase, and Hive, they have the option of using proprietary components in addition to or in place of the open source Hadoop components.

Adaptive MapReduce reimplements the standard Hadoop JobTracker, TaskTracker, and Shuffle services on a low-latency grid middleware implementation that is provided by IBM Platform Computing. Adaptive MapReduce provides even better production-oriented benefits than Hadoop’s grid management and scheduling components. One of those benefits is superior performance.

Hadoop scales out computation and storage across cheap commodity servers and allows other applications to run on top of both of these servers (Spark is one of these applications). Spark runs on top of existing Hadoop clusters to provide enhanced and additional functions.

Although Hadoop is effective for storing vast amounts of data cheaply, the computations it enables with MapReduce are highly limited. Hadoop MapReduce can run only simple computations and uses a high-latency batch model. Spark provides a more general and powerful alternative to Hadoop MapReduce, offering rich functions such as stream processing, machine learning, and graph computations.

Spark is 100% compatible with Hadoop Distributed File System (HDFS), HBase, and any Hadoop storage system, so your existing data is immediately usable in Spark.

6.8.2 Advantages of Spark technology

Spark is intended to enhance, not replace, the Hadoop stack. From day one, Spark was designed to read and write data from and to HDFS and other storage systems. Hadoop users can enrich their processing capabilities by combining Spark with Hadoop MapReduce, HBase, and other big data frameworks.

Comparison of IBM InfoSphere BigInsights Enterprise Edition with Adaptive MapReduce and Apache Hadoop

In an audited benchmark that was conducted on October 2013 by the Securities Technology Analysis Center (STAC), InfoSphere BigInsights for Hadoop was found to deliver an approximate 4x performance gain on average over open source Hadoop.

In jobs that are derived from production Hadoop traces, InfoSphere BigInsights accelerated Hadoop by an average of approximately 4x. The speed advantage of InfoSphere BigInsights was closely related to the shuffle size. Much of the InfoSphere BigInsights advantage appears to be because of better scheduling latency.

In a pure corner-case test of scheduling speed, this InfoSphere BigInsights configuration outperformed the Hadoop configuration by approximately 11x in warm runs. Default settings for the Hadoop core and for InfoSphere BigInsights were used. Nevertheless, it is possible that different settings for Hadoop or InfoSphere BigInsights might achieve different results.

Note: The full report (document IML14386USEN) can be downloaded from the IBM website at:

<http://ibm.co/1bDFq1R>

6.8.3 Spark deployments

This book has constantly focused on making it as easy as possible for every Hadoop user to take advantage of Spark's capabilities. There are three ways to deploy Spark in a Hadoop cluster: Stand-alone, YARN, and Spark In MapReduce (SIMR). Figure 6-17 illustrates possible Spark deployments in a Hadoop cluster.

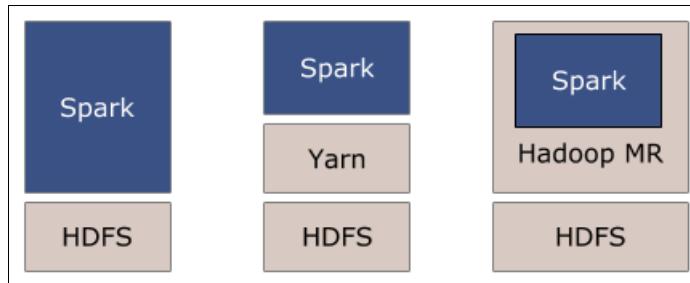


Figure 6-17 Spark in a Hadoop cluster - stand-alone, YARN, and SIMR

Stand-alone deployment

With the stand-alone deployment, you can statically allocate resources on all or a subset of machines in a Hadoop cluster and run Spark side-by-side with Hadoop MR. The user can then run arbitrary Spark jobs on the HDFS data. Its simplicity makes this the deployment of choice for many Hadoop 1.x users.

Hadoop YARN deployment

Hadoop users who have already deployed or are planning to deploy Hadoop YARN can run Spark on YARN without any preinstallation or administrative access required. Users can easily integrate Spark into their Hadoop stack and take advantage of the full power of Spark, and other components running on top of Spark.

Spark In MapReduce

For the Hadoop users that are not running YARN yet, another option, in addition to the stand-alone deployment, is to use SIMR to start Spark jobs inside MapReduce. With SIMR, users can start experimenting with Spark. This tremendously lowers the barrier of deployment, and lets virtually everyone play with Spark.

6.8.4 Spark infrastructure

Spark Core is the underlying general execution engine for the Spark platform in which all other functions are built. It provides in-memory computing capabilities to deliver speed, a generalized execution model to support a wide variety of applications, and Java, Scala, and Python APIs for ease of development.

Spark provides simple and easy-to-understand programming APIs that can be used to build applications at a rapid pace in Java, Python, or Scala. Data scientists and developers alike can benefit from Spark by building rapid prototypes and workflows that reuse code across batch, interactive, and streaming content. For example, users can load tables in Spark programs by using Shark, call machine learning library routines in graph processing, or use the same code for batch and stream processing.

6.8.5 Spark deployment templates

IBM Platform Application Service Controller can integrate with many distributed application frameworks to manage and run them on a scalable, shared grid.

You can also have your application instances running in Docker containers with tools that are provided by IBM Platform Application Service Controller. There is a website that is called IBM developerWorks® Services where you can find the most recent examples and templates that can be used to improve your daily work:

<http://www.ibm.com/developerworks/>

For the latest set of application frameworks integrations, go to the IBM Application Service Controller DevOps Services website at:

<https://hub.jazz.net/user/ibmasc>

Note: An IBM Application Service Controller (ASC) application template to deploy quickly Spark with one HDFS cluster can be found at the following website:

<https://hub.jazz.net/project/ibmasc/asc-spark-hdfs-docker/overview>

6.9 ASC as the attachment for cloud-native framework: Apache Cassandra

IBM Platform Application Service Controller allows you to deploy, run, and manage complex long-running application instances in the Platform Symphony cluster; these can be application servers, InfoSphere BigInsights instances, MongoDB, Cassandra, HBase, and so on. You can monitor and manage application instances and drill down to manage the related services and service instances.

DataGrid solutions are still used to optimize data distribution in HPC environments, with NoSQL solutions being used mostly as inbound/outbound data stores for the core compute engines. It is likely that they will progressively start to override the DataGrid market, mainly for cost reasons. It is becoming also increasingly complex to maintain two different technologies that are dedicated to data management.

NoSQL solutions are starting to be adopted across the board in investment banking either by using MongoDB, Cassandra, or HDFS/Hadoop as part of their compute stack.

The Application Service Controller uses proven technologies that are widely deployed at scale in some of world's largest production clusters to enable increased asset utilization and improved application performance.

The Application Service Controller is designed to be flexible and accommodate distributed cloud-native frameworks, such as Hadoop, Apache Cassandra, and MongoDB.

Cassandra is a massively scalable open source NoSQL database. Cassandra is perfect for managing large amounts of data across multiple data centers and the cloud. It delivers continuous availability, linear scalability, and operational simplicity across many commodity servers with no single point of failure, along with a powerful data model that is designed for maximum flexibility and fast response times.

Cassandra has a *masterless* architecture, meaning all nodes are the same. It provides automatic data distribution across all nodes that participate in database cluster. There is nothing programmatic that a developer or administrator must do or code to distribute data across a cluster because data is transparently partitioned across all nodes in a cluster.

Cassandra also provides customizable replication. This means that if any node in a cluster goes down, one or more copies of that node's data is still available on other machines in the cluster. Replication can be configured to work across one data center, many data centers, and multiple cloud availability zones.

Cassandra supplies linear scalability, meaning that capacity may be easily added by adding nodes online.

Note: Companies running their applications on Cassandra have realized benefits that have directly improved their business. Check out how businesses have successfully deployed Apache Cassandra in their environments at the following websites:

<http://planetcassandra.org/apache-cassandra-use-cases/>

Cassandra architecture

Cassandra is designed to handle big data workloads across multiple nodes with no single point of failure. Its architecture is based on the understanding that system and hardware failures occur. Cassandra addresses the problem of failures by employing a peer-to-peer distributed system across homogeneous nodes where data is distributed among all nodes in the cluster. Each node exchanges information across the cluster every second. A sequentially written commit log on each node captures write activity to ensure data durability. Data is then indexed and written to an in-memory structure, called a *memtable*, which resembles a write-back cache.

Cassandra is a row-oriented database. The Cassandra architecture allows any authorized user to connect to any node in any data center and access data by using the CQL language. For ease of use, CQL uses a similar syntax to SQL. From the CQL perspective, the database consists of tables. Typically, a cluster has one keyspace per application. Developers can access CQL through `cqlsh` and through drivers for application languages.

Client read or write requests can be sent to any node in the cluster. When a client connects to a node with a request, that node serves as the coordinator for that particular client operation. The coordinator acts as a proxy between the client application and the nodes that own the data being requested. The coordinator determines which nodes in the ring should get the request based on how the cluster is configured.

Cassandra and multitenancy

Most users of Cassandra stand up a cluster for each application or related set of applications because it is much simpler to tune and troubleshoot. There has been work done to support more multitenant capabilities, such as scheduling and authorization. However, the traditional path is definitely single-tenant.

6.10 Summary

For the most recent information about Platform Symphony, consult the following IBM Knowledge Centers. They are updated daily.

- ▶ Platform Symphony V7.1 documentation:
http://www-01.ibm.com/support/knowledgecenter/SSGSMK_7.1.0/sym_kc/sym_kc_welcome_71.html?lang=en
- ▶ Release notes for IBM Platform Symphony V7.1:
http://www-01.ibm.com/support/knowledgecenter/SSGSMK_7.1.0/sym_release_notes/newfeatures.dita?lang=en

If you are a developer: Learn about what is new, what has changed, and the limitations, known issues, and documentation updates for Platform Symphony and Platform Symphony Developer Edition by going to the following website:

http://www-01.ibm.com/support/knowledgecenter/SSGSMK_7.1.0/sym_kc/sym_kc_release_notes.dita?lang=en



IBM Platform High Performance Computing

This chapter introduces and describes the IBM Platform High Performance Computing (HPC) product offering. Technical computing users without IT support for their applications often need to become experts about how to administer workloads on their clusters. As a result, these domain experts are spending time and effort managing infrastructure rather than focusing on producing results. They either struggle with building, managing, and supporting a cluster infrastructure themselves, or compromise performance by running their applications on a workstation, which adversely impacts speed to solution and competitiveness.

This chapter covers the following topics:

- ▶ Overview
- ▶ IBM Platform HPC advantages
- ▶ Implementation

7.1 Overview

Platform HPC provides a set of technical and high performance computing management capabilities in a single product. The rich set of ready to use features empowers IT managers and users by reducing the complexity of deploying, managing, and using their computing environment and improving their time to results while reducing costs. Figure 7-1 shows the relationship of the components of Platform HPC.

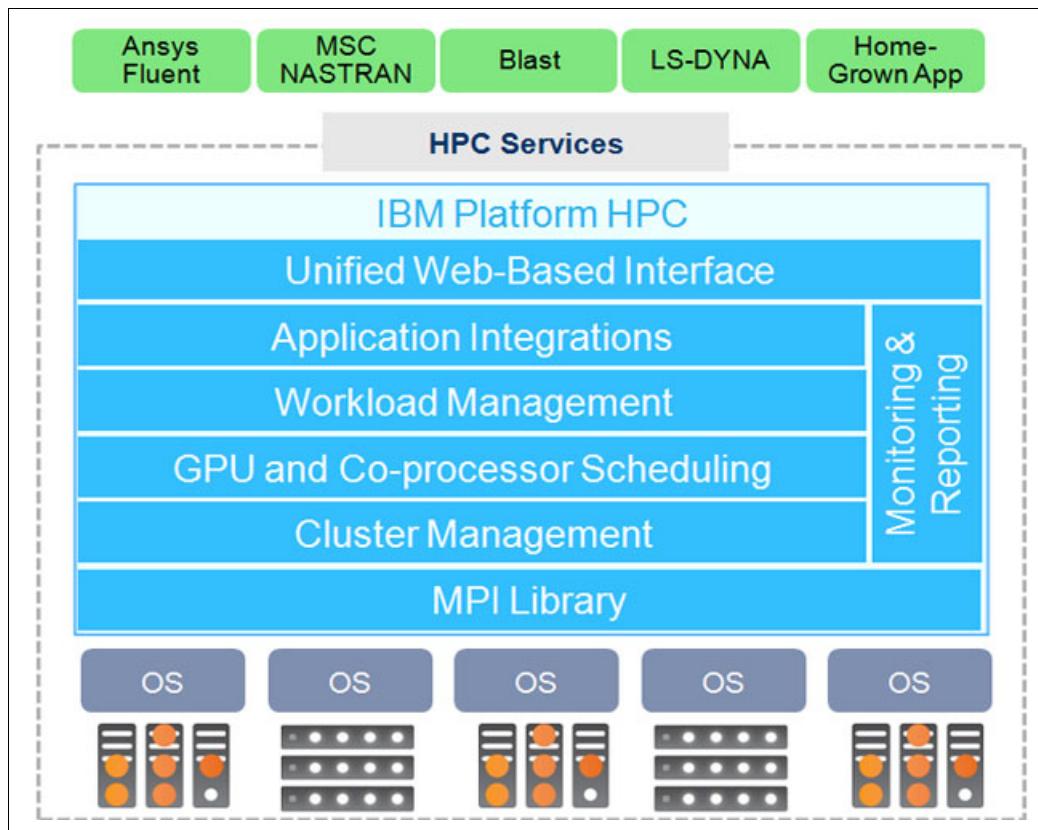


Figure 7-1 IBM Platform HPC components

Platform HPC allows technical computing users in industries such as manufacturing, oil and gas, life sciences, and higher education to deploy, manage, and use their HPC cluster through an easy to use web-based interface. This interface minimizes the time that is required for setting up and managing the cluster for users and allows them to focus on running their applications rather than managing infrastructure. Platform HPC provides full management capabilities from cluster provisioning, monitoring, and management to workload scheduling and reporting. All of the functions that are required to operate and use a cluster are installed at once and are tightly integrated. The product is designed to deliver faster time to system readiness, ease-of-use, and improved application throughput.

7.2 IBM Platform HPC advantages

Platform HPC provides robust cluster and workload management capabilities, although some HPC cluster solutions combine multiple tools and integrate them in a package that is not integrated, certified, or tested together. Platform HPC provides a unified set of management capabilities that help you harness the power and scalability of your HPC cluster, resulting in optimal resource utilization and application throughput. It simplifies the application integration process so that users can focus on running their applications instead of managing the cluster. With Platform HPC, users can take advantage of the following features:

- ▶ A complete solution that achieves faster time to cluster readiness and faster time to results.
- ▶ Maintain the cluster, apply patches and upgrades, and monitor and report cluster health.
- ▶ Straight-forward cluster deployment and provisioning process.
- ▶ An easy to use web-based interface for simplified cluster management, application integration, and workload submissions.
- ▶ A sophisticated workload scheduler to improve application throughput with advanced scheduling policies.
- ▶ Ensure fair use by multiple users, and thus avoiding application conflicts.
- ▶ Submit, manage, and monitor jobs.
- ▶ Isolate problems and troubleshoot.

Platform HPC also includes the following features:

- ▶ Cluster and management, including integrated xCAT
- ▶ Ready to use management of IBM hardware, including IBM NeXtScale, IBM System x iDataPlex, IBM Flex System® x86 nodes, and IBM Intelligent Cluster
- ▶ Intel Xeon Phi co-processor and NVIDIA GPU scheduling and monitoring
- ▶ IBM Platform MPI libraries
- ▶ Integrated application scripts and templates
- ▶ Unified web portal

7.3 Implementation

Platform HPC uses a single unified installer for all of the standard elements of the product. Instead of having to install the Platform Cluster Manager, Workload Manager (LSF), and MPI library, this unified installer approach speeds up implementation and provides a set of standard templates from which a cluster can be built quickly.

This section demonstrates the basic steps about how to install Platform HPC on two machines, one master server, and the first node. Using the minimum required hardware, the procedure covers a simple scenario that can be used either for learning purposes or to help improve a complex scenario with hundreds of nodes. The step-by-step procedures assume that the user is familiar with basic Linux administration and also have some skills in network management.

7.3.1 Installing a management node

This section describes how to install the management node.

Hardware requirements

Here are the minimum hardware requirements for the management node:

- ▶ 100 GB free disk space
- ▶ 4 GB of physical memory (RAM)
- ▶ At least one static Ethernet configured interface (this example uses two Ethernet interfaces)

Software requirements

One of the following operating systems is required:

- ▶ Red Hat Enterprise Linux (RHEL) 6.5 x86 (64-bit)
- ▶ SUSE Linux Enterprise Server (SLES) 11.3 x86 (64-bit)

Before you install the management node, you must configure the operating system. Check that the following conditions are met:

1. Check that /opt has a least 4 GB.
2. Check that /var and /install have at least 40 GB each.
3. Use a fully qualified domain name (FQDN) for the management node.
4. The package openais-devel must be removed manually if it is already installed.
5. Make sure that shadow passwords authentication is enabled.
6. Ensure that IPv6 is enabled for remote power and console management.
7. Ensure that the operating system time is set to the current real time.

Note: Do not run a `yum update` (RHEL) or `zypper update` (SLES) before installing Platform HPC. You can update the management node's operating system after installation.

Specific Red Hat Enterprise Linux prerequisites

Check that the following prerequisites are satisfied:

1. The 70-persistent-net.rules file is created under /etc/udev/rules.d/.
2. Stop the NetworkManager service.
3. Disable SELinux.
4. Ensure that the traditional naming scheme ethN is used.
5. Install the package net-snmp-perl.

Specific SUSE Linux Enterprise Server prerequisites

Check that the following prerequisites are satisfied:

1. Disable AppArmor.
2. Install the createrepo and perl-DBD-Pg packages.

Network considerations

A production environment requires that the network devices are configured properly to avoid installation issues (although this is not necessary for this basic scenario).

Performing the installation

Platform HPC can be installed by using a quick installation or custom installation method. The quick installation method quickly sets up basic options with default options. The custom installation method provides added installation options and enables the administrator to specify additional system configurations. This exercise covers the custom installation method.

If you have downloaded the installation media, you must burn a DVD media first or mount the ISO into a directory. To begin the installation, change to the directory where you have the Platform HPC software and run the installer, as shown in Example 7-1.

Example 7-1 First installation screen

```
[root@pw4302-16 mnt]# ./phpc-installer
    Preparing to install 'phpc-installer'... [ OK ]
    Enter the path to the product entitlement file
[/mnt/entitlement/phpc.entitlement]:
    Parsing the product entitlement file... [ OK ]

=====
Welcome to the IBM Platform HPC 4.2 Installation
=====

The complete IBM Platform HPC 4.2 installation includes the following:
1. License Agreement
2. Management node pre-checking
3. Specify installation settings
4. Installation

Press ENTER to continue the installation or CTRL-C to quit the installation.
```

When the installation begins, the installer automatically checks the hardware and software configurations, and also prompts you for the product entitlement file. If no error messages are displayed, the installation continues through the next steps. Press Enter to proceed to the license agreement. You must accept it to continue the process, as shown in Example 7-2.

Example 7-2 License agreement

```
=====
Step 1 of 4: License Agreement
=====

International Program License Agreement

Part 1 - General Terms

BY DOWNLOADING, INSTALLING, COPYING, ACCESSING, CLICKING ON
AN "ACCEPT" BUTTON, OR OTHERWISE USING THE PROGRAM,
LICENSEE AGREES TO THE TERMS OF THIS AGREEMENT. IF YOU ARE
ACCEPTING THESE TERMS ON BEHALF OF LICENSEE, YOU REPRESENT
AND WARRANT THAT YOU HAVE FULL AUTHORITY TO BIND LICENSEE
TO THESE TERMS. IF YOU DO NOT AGREE TO THESE TERMS,
* DO NOT DOWNLOAD, INSTALL, COPY, ACCESS, CLICK ON AN
"ACCEPT" BUTTON, OR USE THE PROGRAM; AND
* PROMPTLY RETURN THE UNUSED MEDIA, DOCUMENTATION, AND
```

Press Enter to continue viewing the license agreement, or enter "1" to accept the agreement, "2" to decline it, "3" to print it, "4" to read non-IBM terms, or "99" to go back to the previous screen.

After the license agreement is accepted, the installer verifies whether the prerequisites are met. If necessary, you can cancel the installation and fix any issues that are found and restart the installation again. The next step is to choose the installation method, type the number 2, and press Enter for a custom installation. Example 7-3 shows the installer checking the prerequisites and prompting for the installation method.

Example 7-3 Prerequisite verification

```
=====
Step 2 of 4: Management node pre-checking
=====

Checking hardware architecture... [ OK ]
Checking OS compatibility... [ OK ]
Checking free memory... [ OK ]
Checking if SELinux is disabled... [ OK ]
Checking if Auto Update is disabled... [ OK ]
Checking if NetworkManager is disabled... [ OK ]
Checking if PostgreSQL is disabled... [ OK ]
Checking for DNS service... [ OK ]
Checking for DHCP service... [ OK ]
Checking for available ports... [ OK ]
Checking management node name... [ OK ]
Checking static NIC... [ OK ]
Probing DNS settings... [ OK ]
Probing language and locale settings... [ OK ]
Checking home directory (/home) ... [ OK ]
Checking mount point for depot (/install) directory... [ OK ]
Checking required free disk space for opt directory... [ OK ]
```

```
=====
Step 3 of 4: Specify installation settings
=====
```

```
Select the installation method from the following options:
1) Quick Installation
2) Custom Installation
Enter your selection [1]: 2
```

Note: You can cancel the installation at any time by pressing **Ctrl+c**. The installer confirms whether you really want to stop the installation process. If so, then the installation stops and the installer reverts any changes that it might have been done so that you can start a fresh installation later.

You must choose a mount point for the depot where you keep the images and kits (you need at least 40 GB). Example 7-4 on page 111 shows an example. Enter your selection and press Enter.

Example 7-4 Mount point selection

Select a mount point for the depot (/install) directory from the following options:

- 1) Mount point: '/' Free space: '78 GB'
Enter your selection [1]: 1
-

The installer also must know from where it should look for the operating system files. You can point to an ISO file or the DVD media. Example 7-5 shows it is using the DVD media.

Example 7-5 Select the source media for the operating system

The OS version must be the same as the OS version on the management node.
From the following options, select where to install the OS from:

- 1) CD/DVD drive
- 2) ISO image or mount point

Enter your selection [1]: 1

The next steps are used to configure the network settings. First, configure the installer with the interface it should use for the provisioning network. The management node uses this network to communicate with the nodes. Then, decide which network range it should use. Example 7-6 shows the options.

Example 7-6 Network ranges

Select a network interface for the provisioning network from the following options:

- 1) Interface: eth0, IP: 172.16.20.165, Netmask: 255.255.252.0
- 2) Interface: eth1, IP: 192.168.0.165, Netmask: 255.255.255.0

Enter your selection [1]: 2

Enter IP address range used for provisioning compute nodes
[192.168.0.3-192.168.0.200]:

This sample scenario does not choose the node discovery option. You must set which interface the management node uses to connect to a public network, and choose whether you want to enable the specific rules for the management node firewall to the public interface. As you probably already have a firewall in your network, you can enter N as the answer. Example 7-7 illustrates these options.

Example 7-7 First steps of the network configuration

Do you want to provision compute nodes with node discovery? (Y/N) [Y]: n

The management node is connected to the public network by:

- 1) Interface: eth0, IP: 172.16.20.165, Netmask: 255.255.252.0
- 2) It is not connected to the public network

Enter your selection [1]: 1

Enable Platform HPC specific rules for the management node firewall to the public interface? (Y/N) [Y]: n

In the next steps, continue configuring the network. Determine whether you want to enable a BMC network. Because this scenario is adding one node, enter N. Then, determine the DNS settings. You can use a different name server or simple use the default options. Example 7-8 shows the default options for this example.

Example 7-8 Finish the network configuration

Enable a BMC network that uses the default provisioning template (Y/N) [N]:

Enter a domain name for the provisioning network [private.dns.zone]:

Set a domain name for the public network (Y/N) [Y]: n

Enter the IP addresses of extra name servers that are separated by commas [192.168.0.1]:

To complete the installation, enter a valid time server to synchronize the management node. If you use an external time server, you can verify the time if you have the correct firewall rules in place. Then, determine whether you want to export the home directory on the management node so that it can be visible to all other nodes. Finally, choose whether you want to change the root password for compute nodes and the Platform HPC database. Example 7-9 shows the last options in the installation process.

Example 7-9 Final questions in the installation process

Enter NTP server [pool.ntp.org]:

Synchronizing management node with the time server... [OK]

Do you want to export the home directory on the management node and use it for all compute nodes? (Y/N) [Y]: n

Do you want to change the root password for compute nodes and the default password for the Platform HPC database? (Y/N) [Y]: n

The next screen shows a summary of all your choices. You can change any of your previous choices. If you are satisfied with the options, press 1 to begin the installation. Example 7-10 shows the summary of the options that are selected.

Example 7-10 Summary of the configuration settings

=====

Platform HPC Installation Summary

=====

You have selected the following installation settings:

Provision network domain:	private.dns.zone
Provision network interface:	eth1, 192.168.0.0/255.255.255.0
Public network interface:	eth0, 172.16.20.0/255.255.252.0
Depot (/install) directory mount point:	/
OS media:	CD/DVD drive
Network Interface:	eth1
eth1 IP address range for compute nodes:	192.168.0.3-192.168.0.200
eth1 IP address range for node discovery:	N/A
Enable firewall:	No
NTP server:	pool.ntp.org

```

Name servers:           192.168.0.1
Database administrator password: ****
Compute node root password: ****
Export home directory:   No
=====

```

Note: To copy the OS from the OS DVD, you must insert the first OS DVD into the DVD drive before beginning the installation.

To modify any of the above settings, press "99" to go back to "Step 3: Specify installation settings", or press "1" to begin the installation.

Note: The default user name and password is *phpcadmin*. Change it as soon as possible.

7.3.2 Installing a compute node

The first time you log in to the web interface, you see a dashboard with an overview of the resource health of all configure nodes. Currently, you have only the management node. Before you start adding the first compute node, you might want to look at the menu in the left pane to see all the menu options. At the upper right of the page, you see a help menu that explains the web portal interface and its components.

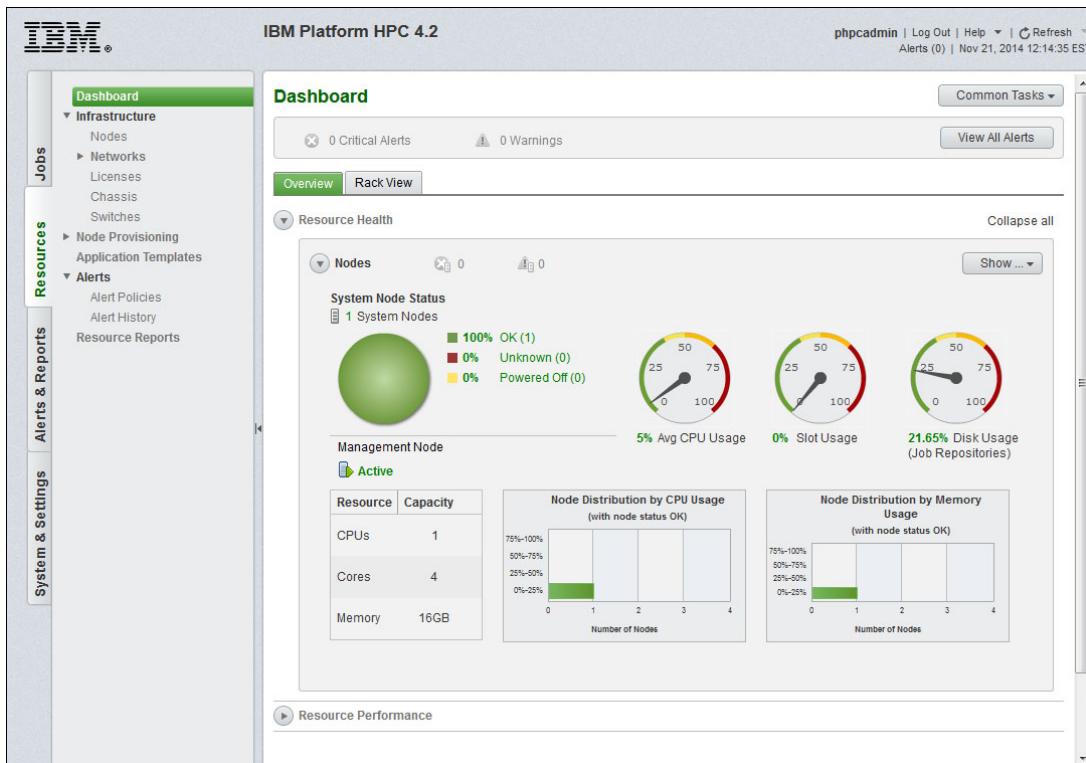


Figure 7-2 Web interface after the installation of the management node

The next step is to specify how the nodes are discovered. In this example, the machine has a MAC address of 00:50:56:82:15:30. Use this information to create a txt file with this MAC address. Example 7-11 shows the content of the sample txt file. Although other parameters can be set in this file, this is the minimum required information.

Example 7-11 Sample file with the first compute node

```
_hostname_:
mac=00:50:56:82:15:30
```

To add the first compute node, click **Resources** → **Infrastructure** → **Nodes**, and click **Add**. A dialog box opens. In the Node Group, select **compute**, and in the Select provisioning template, select **rhels6.5-x86_64_stateful_compute** and click **Next**. Figure 7-3 shows the example of these selections.

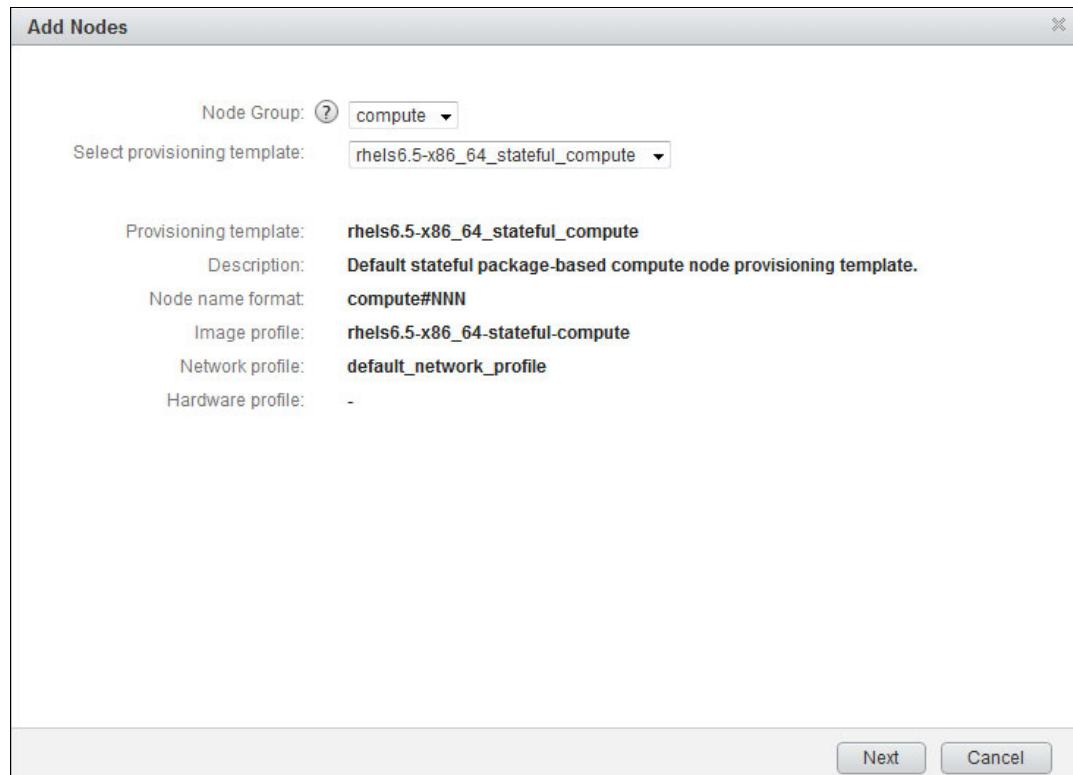


Figure 7-3 First step to add a compute node

Select **Import node information file** and click **Browse** to select the file that you created. You can also add tags to this node to help identify it if necessary, and then click **Add**, as shown in Figure 7-4 on page 115.

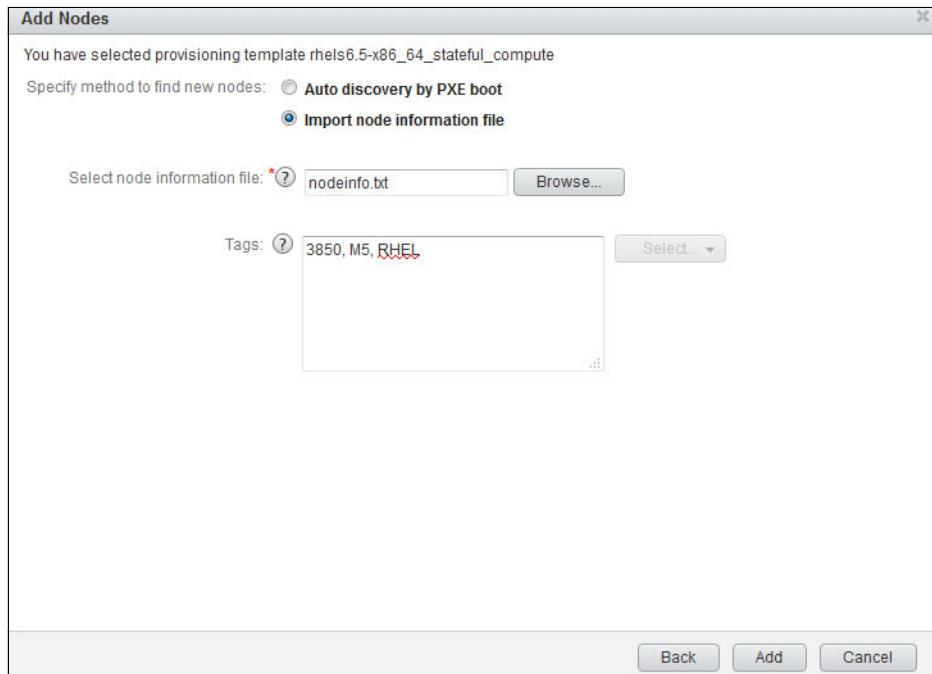


Figure 7-4 Second step to add a compute node

Note: Check that the new compute node is set to start over the network interface.

After the file has been imported, the new compute node is available in the Node List with the provision status *defined*. Proceed with the restart of this server and the installation starts automatically. Figure 7-5 shows the management server and two compute nodes, one defined and another already installed.

Nodes: List												
	Name	Power...	Workload...	Total...	Job ...	Provision Status	Synchronization...	Node Type	CPU...	Free ...	Node Groups	
<input type="checkbox"/>	pw4302-i3	-	OK	4	0	provisioned	synced	Manage...	5%	12924	-	
<input type="checkbox"/>	compute001	-	-	0	0	defined	-	Compute...	-	-	compute	
<input type="checkbox"/>	compute000	-	OK	4	0	provisioned	synced	Compute...	0%	15720	compute	

Figure 7-5 Node list



IBM Platform Cluster Manager

IBM Platform Cluster Manager is a powerful and easy-to-use software for managing complex clusters and high performance computing (HPC) data centers. Platform Cluster Manager provides useful features that allow administrators to manage hardware and software, such as the following items:

- ▶ Operating system deployment automation
- ▶ HPC clusters deployment
- ▶ System maintenance

This chapter covers some of the features of the latest version of Platform Cluster Manager - Standard Edition and Advanced Edition. To illustrate these features, this chapter goes step-by-step through the features in configuration and sample scenarios.

The chapter covers the following topics:

- ▶ Platform Cluster Manager - Standard Edition V4.2
- ▶ Platform Cluster Manager - Advanced Edition V4.2

8.1 Platform Cluster Manager - Standard Edition V4.2

Platform Cluster Manager - Standard Edition uses a centralized user interface that allows system administrators to manage a complex cluster as a single system. Platform Cluster Manager empowers users to add customized features for a specific environment and several other features:

- ▶ A kit framework for easy InfiniBand driver deployment
- ▶ Monitoring capability for visualizing the performance and conditions of the cluster
- ▶ Allows monitoring of non-server components, such as chassis, network switches, IBM Spectrum Scale, GPU, and co-processors
- ▶ Adds management node automatic failover capability to ensure the continuity of cluster operations

Platform Cluster Manager V4.2 has new features:

- ▶ LDAP support.
- ▶ IBM POWER8 nodes support.
- ▶ Node tags.
- ▶ The `1parid` parameter is added to the node information file.
- ▶ Nodes that are configured for switch based provisioning can be replaced without specifying a MAC address (automatically retrieved from a switch).
- ▶ *Node power status* is a new node attribute that indicates the power status of a node.

This section describes how to set up user authentication against an LDAP server and how node tags works.

Note: For more information about how to install Platform Cluster Manager - Standard Edition, see Chapter 4, “IBM Big Data implementation on an IBM High Performance Computing cluster”, in *IBM Platform Computing Solutions Reference Architectures and Best Practices*, SG24-8169, and the *Platform Cluster Manager Standard Edition V4.2 Release Notes*.

8.1.1 Platform Cluster Manager - Standard Edition support for POWER8 nodes

Platform Cluster Manager - Standard Edition V4.2 now supports POWER8 nodes.

When using IBM Power Systems, Platform Cluster Manager - Standard Edition shows the CPU socket number for each compute node that is listed in the web portal, as shown in Figure 8-1 on page 120. The number is different from the CPU socket number that is produced by the `lscpu` command.

8.1.2 LDAP integration

As described in the *Platform Cluster Manager Standard Edition V4.2 Release Notes*, the LDAP integration is added to the new version and now provides the system administrators ready to use support for user account management through LDAP.

To enable this feature, run the following commands:

- ▶ # source /opt/pcm/bin/pcmenv.sh
- ▶ # pcmadmin system ldap --enable

The first command loads some required environment variables, and the second command runs a script that guides you on how to enable the LDAP authentication in Platform Cluster Manager - Standard Edition. Example 8-1 shows the output of these commands.

Example 8-1 Enabling LDAP authentication on Platform Cluster Manager - Standard Edition

```
[root@pcmsepkm ~]# pcmadmin system ldap --enable
To enable LDAP authentication, you must stop both WEBGUI and PCMD services before
continuing.
Continue? (Y/N) [N]: Y
Service WEBGUI is already stopped.
Service PCMD is already stopped.
Type the URL of the LDAP server (for example, ldap://LDAP_server:389):
ldap://localhost:389
Type the base domain where users and groups will be retrieved (for example,
dc=example,dc=com): dc=platform,dc=itso,dc=ibm,dc=com
Type the distinguished name of the LDAP user mapped to IBM Platform Cluster
Manager (for example, uid=pcmuser,ou=user,dc=platform,dc=itso,dc=ibm,dc=com):
cn=Manager,dc=platform,dc=itso,dc=ibm,dc=com
Type the password for the mapped user:
Enable base domain LDAP users login this node through SSH? (Y/N) [N]
Verifying LDAP configuration...
Installing LDAP client required packages...
Configuring pcmd...
Enable LDAP client setup for compute nodes...
IBM Platform Cluster Manager has been successfully configured to retrieve user
information from LDAP.
Logs can be found in /opt/pcm/pcmd/log/pcmd.log
Start up PCMD service by running 'pcmadmin service start --service PCMD'
Start up WEBGUI service by running 'pcmadmin service start --service WEBGUI'
[root@pcmsepkm ~]# pcmadmin service start --service PCMD
Service PCMD is already started.
[root@pcmsepkm ~]# pcmadmin service start --service WEBGUI
Service WEBGUI is already started.
```

Note: The LDAP server must be previously configured. Use the IP address or the full host for the LDAP server instead of localhost. For more information about how to install a basic LDAP server, see Appendix B, “LDAP server configuration and management” on page 151.

Now, Platform Cluster Manager - Standard Edition is configured to use LDAP authentication for new nodes. To enable LDAP authentication for the existent nodes, run the following command:

```
# updatenode compute
```

All nodes are updated and configured. This installation is taking into account that the HOME directory is exported (NFS) during the Platform Cluster Manager - Standard Edition installation. The user accounts must reflect the POSIX information that is in the management node, such as home directory, UID number, and GID number.

8.1.3 Tagging nodes

Platform Cluster Manager - Standard Edition provides a handy feature called tags. Tags are descriptors that are used to identify nodes. Nodes can be tagged with one or more tags. After a tag is created, it can be reused in a system to enable tracking of similar nodes. Figure 8-1 shows how to add tags to a node. Tags are single words.

The screenshot shows the 'Nodes: List' screen in the IBM Platform Cluster Manager Standard Edition 4.2 interface. On the left, there's a navigation sidebar with sections like Dashboard, Infrastructure (Nodes, Networks, Licenses, Chassis, Switches), Node Provisioning (Provisioning Templates, OS Distributions, Kit Library), Alerts, and Resource Reports. The 'Nodes' section is currently selected. The main area displays a table of nodes with columns for Name and Power Status. The node 'powernode2' is selected and has a context menu open. The menu includes options like Reinstall, Replace, Remove, Relocate, Synchronize, Update Firmware, Location LED, and Tags. The 'Tags' option is highlighted with a green background. A sub-menu for Tags is also open, containing 'Add' and 'Remove All' options. Below the table, a specific node details view is shown for 'Node Name: powernode2' with tabs for Summary, Performance, Alerts & Events, and Properties.

Figure 8-1 Add a tag to a node in Platform Cluster Manager - Standard Edition

8.2 Platform Cluster Manager - Advanced Edition V4.2

IBM Platform Cluster Manager - Advanced Edition has all of the Platform Cluster Manager - Standard Edition features plus the following additional capabilities:

- ▶ Multiple clusters
- ▶ Multitenancy

This section describes how a multitenant environment can be created and how it can be used to isolate servers that belong to a specific customer account.

8.2.1 Multitenant environment

A multitenant environment allows system administrators to create different accounts with different levels of access and resource limitations per account.

As a requirement for setting up a multitenant environment, Platform Cluster Manager - Advanced Edition requires a configured and activated LDAP server before the account management is set up.

To enable LDAP authentication, the system administrator must follow the steps that are described in 8.1.2, “LDAP integration” on page 118.

Users can authenticate on any provisioned server under a cluster against the LDAP server. After a server is provisioned under a cluster, users who have access to these servers are authenticated by using the LDAP server.

As an example of how a multitenant environment can be useful, this section shows how it can be implemented for a hypothetical scenario:

- ▶ Company A offers to its clients an HPC environment for general purposes. Clients are free to take advantage of Platform Cluster Manager - Advanced Edition to provision clusters and run their applications without any intervention by Company A during the process.
- ▶ Company A can manage multiple administrator accounts and clients are isolated from each other.

To implement this scenario, complete the following steps:

1. Create a user account in Platform Cluster Manager - Advanced Edition by clicking **System & Settings** → **Accounts** → **New**, as shown in Figure 8-2.

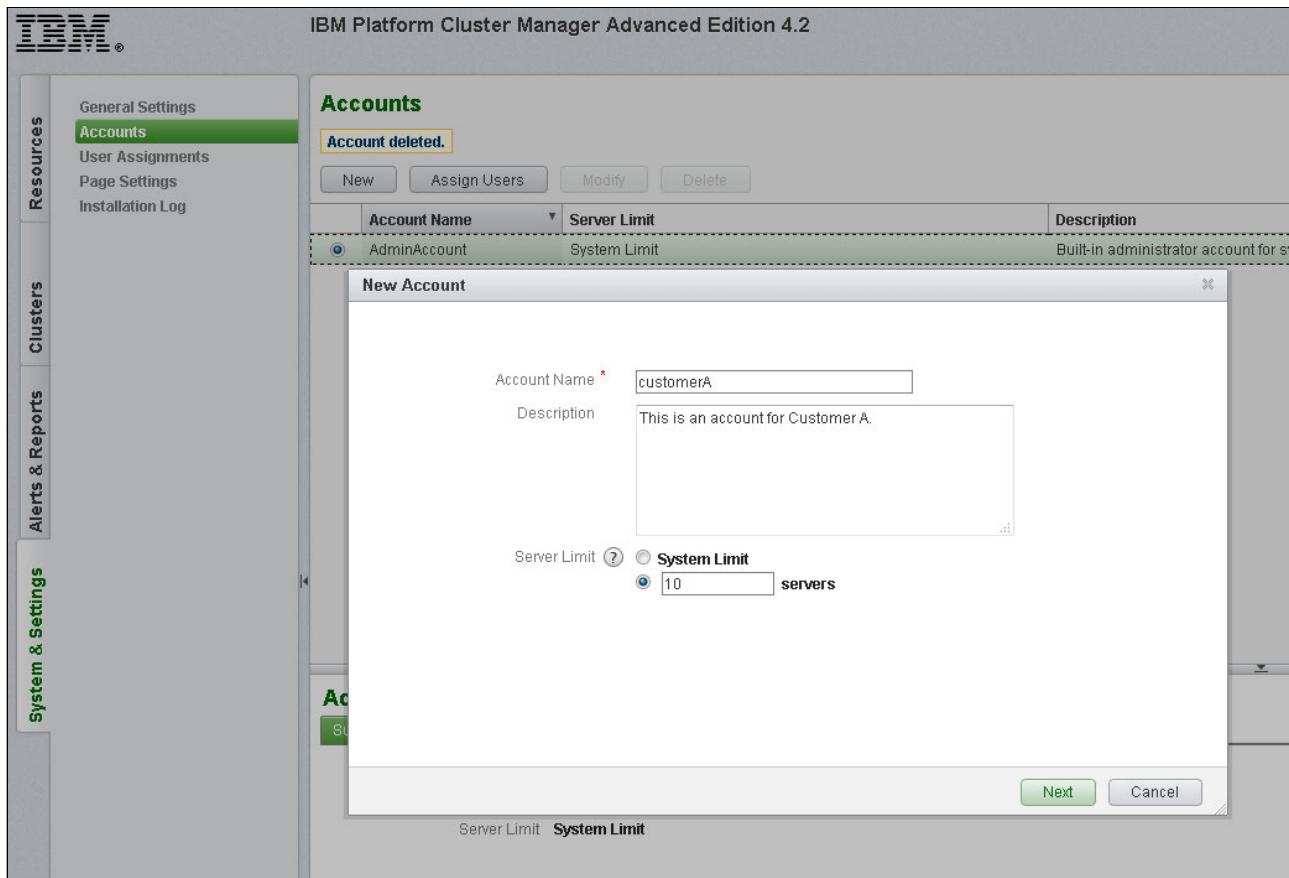


Figure 8-2 User account creation for a multitenant environment

Figure 8-2 shows that you can specify the number of servers that are allowed per account or simply use the system limit.

- Specify which group or groups to which the user account belongs. This step is required, and the groups from the LDAP server are shown in Figure 8-3.

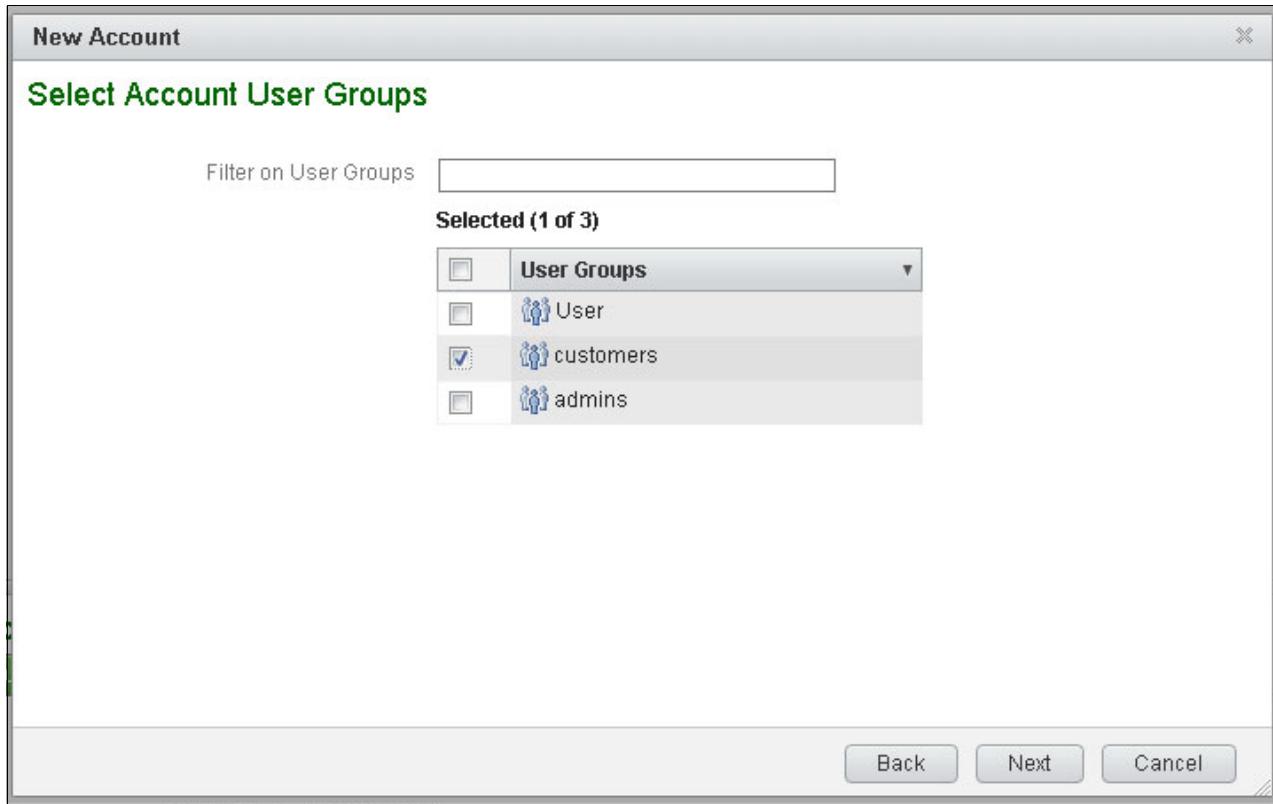


Figure 8-3 Select groups for the user account

- Select which account administrator must manage the account that is being created. This step is *not required*. If no account administrator is selected, only the Platform Cluster Manager - Advanced Edition administrator can manage the account.

The confirmation window (Figure 8-4 on page 123) shows all the information that was specified.

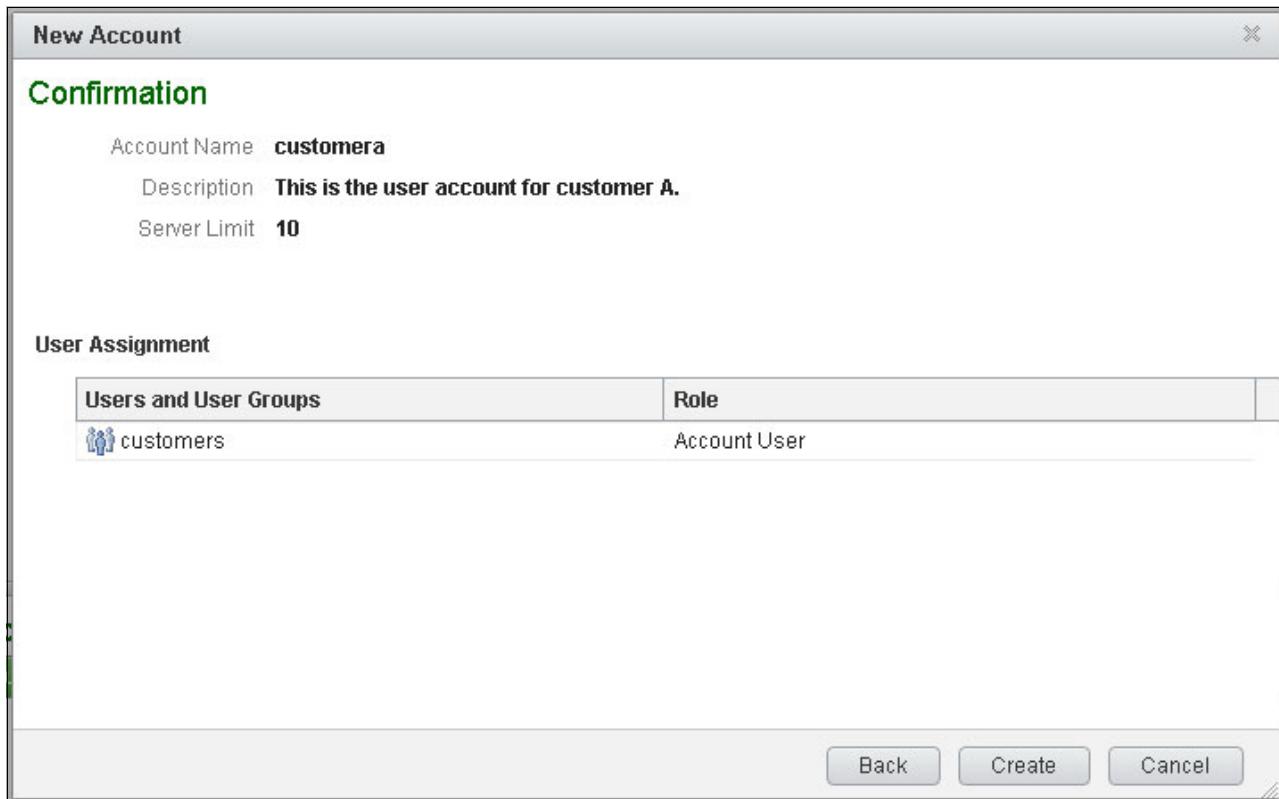


Figure 8-4 Confirm all information before creating the user account

You have created the user account that has a name that matches the UID of the LDAP account. The customera user can now log in to the Platform Cluster Manager - Advanced Edition portal by using its LDAP account password.

Note: Platform Cluster Manager - Advanced Edition requires only user accounts creation for those users who manage clusters and nodes. Any valid LDAP user account can be used to authenticate to the servers within a cluster.



IBM Cloud Manager

As described in Chapter 1, “Introduction to IBM Platform Computing” on page 1, cloud computing is a critical IT component that becomes more important as the velocity of innovation increases, and enterprises and organizations must have an infrastructure that can accelerate time to results for compute- and data-intensive applications.

This chapter introduces IBM Cloud Manager, its foundational concepts, architecture, and how it can help support complex cloud environments in a single dashboard.

IBM Cloud Manager provides the required tools, mechanisms, and features to create, manage, and operate different cloud environments.

For more information, see *IBM Software Defined Environment (SDE)*, SG24-8238.

This chapter covers the following topics:

- ▶ IBM Software Defined Environment
- ▶ The software-defined everything vision
- ▶ OpenStack
- ▶ Introducing IBM Cloud Manager
- ▶ IBM Cloud Manager value points

9.1 IBM Software Defined Environment

Investments in enterprise virtualization, centralized administration, and hardware with enhanced management and optimization functions have laid the groundwork for a new era in business responsiveness. The technical capabilities now exist, and the time is right to take the next step by using those technologies to enable a fully programmable IT infrastructure that can sense and respond to workload demands automatically.

IBM calls this idea a *software-defined environment* (SDE). It is a new approach for holistic, simplified IT management in which software provisions and configures entire infrastructures based on real-time workload needs. SDE is a term that was coined by IBM for its *software-defined everything* vision. The IBM Software Defined Environment group is the latest evolution of what first began as the application, integration, and middleware group inside the IBM Software group.

An SDE optimizes the entire computing infrastructure, compute, storage, and network resources so that it can adapt to the type of work that is required. In today's environment, resources are assigned manually to workloads, which happens automatically in a SDE. In SDE, workloads are dynamically assigned to IT resources based on application characteristics, best-available resources, and service-level policies to deliver continuous, dynamic optimization and reconfiguration to address infrastructure issues. Underlying all of this infrastructure are policy-based compliance checks and updates in a centrally managed environment.

By dynamically assigning workloads to IT resources based on various factors, including the characteristics of specific applications, the best-available resources, and service-level policies, a SDE can deliver continuous, dynamic optimization and reconfiguration to address infrastructure issues.

9.2 The software-defined everything vision

Software-defined everything is a phrase that denotes the grouping of various software-defined computing technologies into one overarching framework and architecture. The umbrella of software-defined everything technologies includes, among other terms, software-defined networking (SDN), software-defined computing, software-defined data centers (SDDC), software-defined storage (SDS), and software-defined storage networks.

With software-defined everything, the computing infrastructure is virtualized and delivered as a service. In a software-defined everything environment, management and control of the networking, storage, and data center infrastructure is automated by intelligent software rather than by the hardware components of the infrastructure.

So, integration, automation, and optimization are enablers to cloud delivery and analytics. SDE can accelerate business success by making a happy marriage between workloads and resources so that you have a responsive, adaptive environment.

With IBM Software Defined Environment, infrastructure is fully programmable so that it can rapidly deploy workloads on optimal resources and instantly respond to changing business demands:

Software	Abstracted and virtualized IT infrastructure resources that are managed by software.
Defined	Applications that automatically define infrastructure requirements and configuration.

Environments	An IT infrastructure that extends multiple environments to go beyond the data center.
---------------------	---

For more information about IBM Software Defined Environment, see *IBM Software Defined Environment (SDE)*, SG24-8238.

9.3 OpenStack

OpenStack is a global collaboration of developers and cloud-computing technologists working to produce an infrastructure open source cloud computing platform for public and private clouds.

For more information about OpenStack, see the following website:

<https://www.openstack.org/>

9.4 Introducing IBM Cloud Manager

IBM Cloud Manager is an easy to deploy and use cloud management software offering that is based on OpenStack with IBM enhancements and support.

Managing today's complex cluster environment is a time-consuming and costly effort for many technical and high performance computing (HPC) data centers. Adding to the challenge is the management of multiple clusters as data centers grow in size. Isolated clusters can create major inefficiencies in a technical computing environment and hinder the ability for organizations that require substantial compute- and data-processing capabilities to compete. The solution helps alleviate this complexity with tools for the self-service creation and management of flexible clusters.

Platform Cluster Manager includes many tools that you need to get clusters up and running quickly. For clients with diverse application and user requirements, Platform Cluster Manager - Advanced Edition automates assembly of multiple high-performance technical computing environments on a shared compute infrastructure that is used by multiple teams. The software creates an agile environment for running both HPC and analytics workloads. By doing so, it allocates the correct resources to the correct workloads, and consolidates disparate cluster infrastructures and multiple workload schedulers, resulting in increased resource utilization, the ability to meet or exceed service level agreements (SLAs), and reduce infrastructure and management costs.

For clients with a single HPC cluster deployment, IBM Platform Cluster Manager - Standard Edition delivers the capability to provision, run, manage, and monitor quickly technical computing clusters with ease and scalability. The latest release of IBM Platform Cluster Manager - Standard Edition offers new flexible monitoring capabilities for servers, chassis, network switches, IBM Elastic Storage, GPU and co-processors, and customized devices. It also adds management node automatic failover capability to ensure cluster operation continuity.

For more information about IBM Cloud Manager, see *IBM Software Defined Environment (SDE)*, SG24-8238.

9.5 IBM Cloud Manager value points

The following are the IBM Cloud Manager value points:

- ▶ Enables rapid IT response to the ever-changing demands of business through self-service provisioning of infrastructure services, as users can redeploy virtual servers with an easy to use interface.
- ▶ Yields improved virtualization operational efficiency and greater overall business effectiveness. Administrators capture and manage standard VM images with support for common business processes.
- ▶ Provides the capability to track and correlate the cost of infrastructure to department usage through basic usage metering, so Organizations & Managed Service Providers (MSPs/CSPs) can align service to expense.
- ▶ Supports production-grade cloud operations and interoperability at scale through enhanced foundation and full OpenStack API compatibility.
- ▶ Open computing cloud alternative to proprietary vendors, with world-class support from IBM.
- ▶ Hybrid capability to IBM SoftLayer through an IBM Global Business Services® offering.



IBM Platform Computing Cloud Services

This chapter provides a description of IBM Platform Computing Cloud Services and describes a scenario about how IBM Platform LSF multicloud and IBM Spectrum Scale Active File Manager help manage the usage of the cloud services in an efficient way.

This chapter provides factual and comparative economics of deploying a solution on-premises versus in the cloud, and includes information about the benefits of on-premises versus in the cloud solution deployments.

This chapter covers the following topics:

- ▶ IBM Platform Computing Cloud Services: Purpose and benefits
- ▶ Platform Computing Cloud Services architecture
- ▶ IBM Spectrum Scale high-performance services
- ▶ IBM Platform Symphony services
- ▶ IBM High Performance Services for Hadoop
- ▶ IBM Platform LSF Services
- ▶ Hybrid Platform LSF on-premises with a cloud service scenario
- ▶ Data management on hybrid clouds

10.1 IBM Platform Computing Cloud Services: Purpose and benefits

Engineering, scientific, financial, or research workloads are not the only demanding workloads for technical and high performance computing (HPC) infrastructures. Big data challenges are solved by using the same method, distributing the workload across multiple machines within a technical computing cluster.

Meeting all these demands can be especially challenging for organizations that have seasonal or unpredictable demand spikes, need access to additional compute or storage resources to support a growing business, or are starting to use these technologies. The time that it takes to respond to a critical market analysis, a product release, or a research study can be impacted by resource availability, which affects competitiveness and profitability.

Organizations can quickly and efficiently overcome these challenges by combining market-leading workload management from IBM Platform Computing with the efficiency and cost benefits of cloud computing.

Platform Computing Cloud Services running on the SoftLayer cloud delivers a versatile, high-performing cloud-based environment to fuel your organization's growth if you are engaged in the following activities:

- ▶ Seeking to meet variable workload demands
- ▶ In need of clustered resources, but do not have the budget or in-house skills to deploy and use a technical computing infrastructure
- ▶ Running out of data center space and must continue to grow compute and storage capacity
- ▶ Considering to provide applications on a pay-per-use basis, but do not have the infrastructure or time to create a service

If any of these activities are important to you, you can count on the benefits that are delivered by the Platform Computing Cloud Services offering to meet your needs:

- ▶ Ready-to-use IBM Platform LSF and IBM Platform Symphony clusters in the cloud reduce time to results and accelerated time to market.
- ▶ High-performance file system with IBM Spectrum Scale that is delivered as a service improves data management and provides seamless transfer between on-premises and cloud infrastructures.
- ▶ Non-shared physical machines, InfiniBand interconnect, the latest processor technology, and your choice of SoftLayer data center leads to optimal application performance and security.
- ▶ Integrated workload management with both on-premises and on-cloud infrastructures simplifies management and the user experience, and full support from IBM technical computing experts reduces administrative impact.

10.2 Platform Computing Cloud Services architecture

Platform Computing Cloud Services is built on top of SoftLayer. SoftLayer deploys the infrastructure in its data centers in the form of *Points of Delivery (PODs)*, which are groups of thousands of machines, petabytes of storage, and all the networking, firewalls, power distribution, internet connectivity that is needed to support this infrastructure. Theoretically, a customer might use an entire POD, which is more than 60,000 processor cores on a single cluster. A customer can request other PODs to meet demands. Usually, a cluster starts much smaller than at the POD scale, and flexes up or down as a client's needs dictate. All configurations that are delivered by Platform Computing Cloud Services deliver exclusive, non-shared server usage for the client. The Platform Computing Cloud Services solution offers a true cloud-based consumption model: Pay by the hour or by the month for all elements of the service.

Platform Computing Cloud Services is a purpose-built Software as a Service (SaaS) where clients can use ready-to-use clusters that are available either for usage by the hour or by the month. The service is owned and operated by the IBM Platform Computing team, which ensures that the HPC experts are available to provide management and support of your chosen environment. The service uses IBM Platform Computing HPC management and scheduling tools (Platform LSF and Platform Symphony), which provide optimum performance and user experience.

The SaaS architecture counts on the two Platform Computing schedulers, Platform LSF and Platform Symphony, for either traditional HPC clusters or service-oriented architectures (SOA). The offering is provided with or without IBM Spectrum Scale (formerly GPFS), which can ease and reduce the data transfer that is needed to and from the cloud by using the Spectrum Scale Active File Management (AFM) facility for caching only the needed files to run the workload on the remote site in the case of a hybrid cloud. The Platform Computing Cloud Services high-level architecture is shown in Figure 10-1.

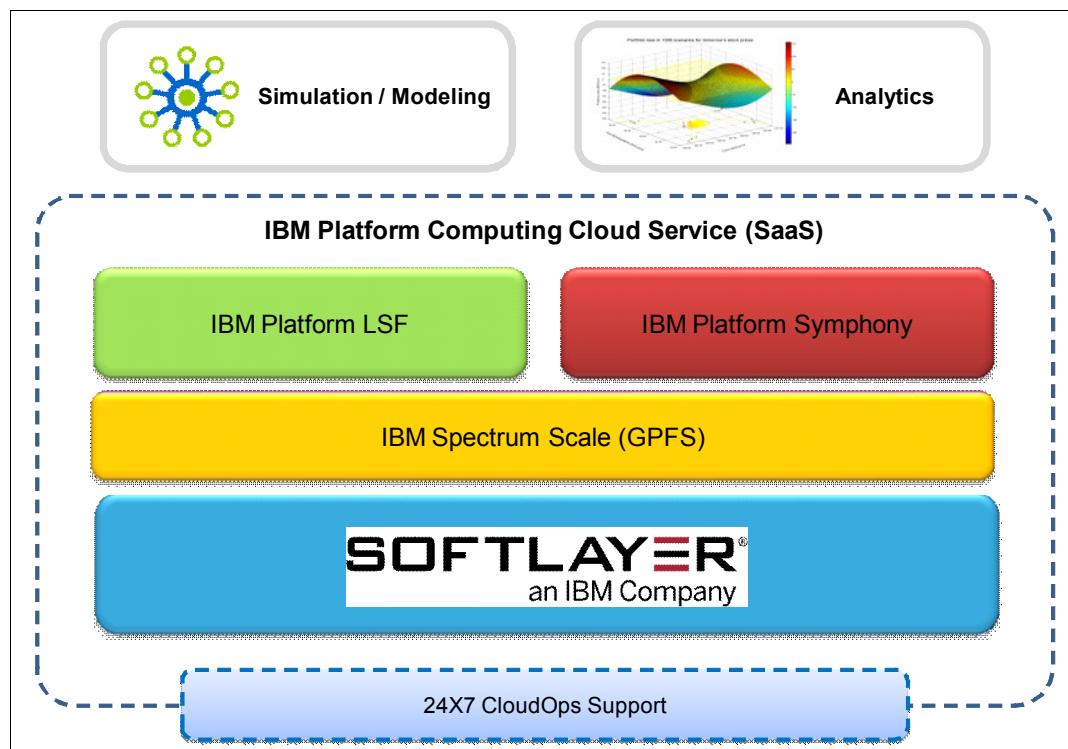


Figure 10-1 IBM Platform Computing Cloud Services - high-level architecture

10.3 IBM Spectrum Scale high-performance services

For clients that consider adding storage capacity or who require more performance and scalability than a network file system (NFS) can provide, IBM Spectrum Scale is now available as a service on the SoftLayer cloud as part of Platform Computing Cloud Services.

Optimized for technical computing and analytics workloads, Spectrum Scale in the cloud enables seamless transfer of files between local and cloud-based resources by using the Spectrum Scale AFM feature.

With the addition of Spectrum Scale in the cloud, Platform Computing Cloud Services enable speedy deployment of fully supported, ready-to-run technical computing or analytic environments in the cloud.

Organizations that use Platform Computing Cloud Services can easily meet additional resource demands without the cost of purchasing or managing an in-house infrastructure, which minimizes the administrative burden and quickly addresses evolving business needs.

10.4 IBM Platform Symphony services

Although a benefit of IBM Platform Symphony is its ability to support diverse applications in a multitenant environment while ensuring service levels, performance tests show that IBM Platform Symphony also helps to provide better performance and efficiency, and superior management and monitoring.

If you do not have a specific application to run on Platform Symphony, for example, but you need a service environment for your Hadoop workload, see 10.5, “IBM High Performance Services for Hadoop” on page 132.

For more information about how IBM Platform Symphony can help improve your Hadoop workload, see the following website:

<http://www.ibm.com/systems/platformcomputing/products/symphony/highperfhadoop.html>

10.5 IBM High Performance Services for Hadoop

IBM High Performance Services for Hadoop is suitable for organizations that are looking for a fully supported, ready-to-run Hadoop environment for production use, or as a development and testing environment. This service enables customers to deploy quickly and easily Hadoop workloads on ready-to-run clusters on the SoftLayer cloud, complete with a bare metal SoftLayer infrastructure, a private network, and your choice of data center to help achieve optimal performance and security.

An experienced and dedicated cloud operations team configures, deploys, and supports the cloud-based infrastructure and the software, which helps minimize the administrative burden on your organization and the need to develop the skills to design and manage a Hadoop environment.

IBM High Performance Services for Hadoop delivers a Hadoop-ready cluster as a service on SoftLayer and helps deliver the following benefits:

- ▶ Rapid access to Hadoop clusters in the cloud for both product use and development testing
- ▶ Optimal performance with bare metal resources
- ▶ Security through physical isolation and choice of data center location
- ▶ Reduced capital expenditure
- ▶ Minimal user and administrator impact
- ▶ Easy adoption of public cloud technology and resources

IBM High Performance Services for Hadoop can deliver the following benefits:

- ▶ More capability and lower costs: Easily meet demand without the upfront costs of purchasing an in-house infrastructure or the ongoing cost of infrastructure management.
- ▶ Match resources to demand while helping reduce capital expenditures: Help minimize administrative costs by using a skilled cloud operations team with deep Hadoop expertise.
- ▶ Security:
 - Help achieve security through physical isolation with a dedicated virtual local area network (VLAN).
 - Upload data securely through a virtual private network (VPN) or Multi-Protocol Label Switching (MPLS) to gateway servers.
 - Use your SoftLayer data center of choice for regulatory compliance.
- ▶ Faster time to results:
 - Accelerate Hadoop MapReduce workloads with dedicated bare metal servers.
 - Optimize I/O performance with 10-Gb Ethernet networking.

10.6 IBM Platform LSF Services

IBM Platform LSF is a powerful workload management platform for demanding, distributed HPC environments. It provides a comprehensive set of intelligent, policy-driven scheduling features that enable you to use all of your compute infrastructure resources and ensure optimal application performance.

Platform LSF helps to ensure that all available resources are fully used by enabling you to take full advantage of all technical computing resources in the cloud. Platform LSF helps to ensure that the computing power in the cloud is fully used, and it helps to manage application software licenses usage, which is expensive for demanding workloads.

The IBM Platform Computing LSF Cloud Service provides the following features:

- ▶ A single source for end-to-end cluster support with access to technical computing experts to eliminate the skills barrier for using clustered resources.
- ▶ Dedicated bare-metal servers and InfiniBand interconnect for applications that require the full capacity of a non-virtualized, and parallel computing environment.
- ▶ Control of data center locality, enabling organizations to choose the location where workloads run to protect their information and meet data regulations.
- ▶ Non-shared physical machines and dedicated network for workloads requiring maximum security.

10.7 Hybrid Platform LSF on-premises with a cloud service scenario

A transparent user experience that manages workloads between an on-premises cluster and in the cloud can be achieved with IBM Platform LSF Multicloud and IBM Spectrum Scale AFM.

Note: If you already have Platform LSF Standard Edition, skip 10.7.1, “Upgrading IBM Platform HPC to enable the multicloud function” on page 134. Otherwise, see *IBM Platform Computing Integration Solutions*, SG24-8081 and *IBM Platform Computing Solutions Reference Architectures and Best Practices*, SG24-8169 for information about how to implement IBM Platform HPC.

10.7.1 Upgrading IBM Platform HPC to enable the multicloud function

To start the upgrade, you need the name of your Platform LSF installation directory (LSF_TOP), the Platform LSF administrators (LSF_ADMINS), and the cluster name (LSF_CLUSTER_NAME) available. If you do not have this information, run the commands that are shown in Example 10-1 to gather the information.

Example 10-1 Gather information for Platform LSF Standard Edition upgrade

```
[root@homecluster etc]# grep LSF_TOP $PCMD_TOP_LOCAL/etc/lsf.install.config
LSF_TOP="/shared/ibm/platform_lsf"
[root@homecluster etc]# grep LSF_ADMINS $PCMD_TOP_LOCAL/etc/lsf.install.config
LSF_ADMINS="phpcadmin root"
[root@homecluster etc]#
[root@homecluster etc]# grep CLUSTER $PCMD_TOP_LOCAL/etc/lsf.install.config
LSF_CLUSTER_NAME="phpc_cluster"
[root@homecluster etc]#
```

The information that is gathered by the commands in Example 10-1 is necessary to upgrade IBM Platform HPC and install IBM Platform LSF Standard Edition into the cluster.

To start, add the parameters to the `install.config` file from your Platform LSF installation directory, as shown in Example 10-2. You must add the path for the Platform LSF Standard Edition entitlement file of the installation configuration file, for example:

```
LSF_ENTITLEMENT_FILE="/tmp/phpc/platform_lsf_std_entitlement.dat"
```

Example 10-2 Configuration file to install Platform LSF (install.config)

```
#####
#          LSF 9.1.3 INSTALL.CONFIG FILE
#####
#
# Name:      install.config
#
# Purpose:   LSF installation options
#
# $Id$
#
# File Format:
#     o Options (without # sign) can only appear once in the file.
```

```

#      o Blank lines and lines starting with a number sign (#) are ignored.
#
# Option Format:
#      o Each disabled example looks like this:
#          # -----
#          # LSF_OPTION_NAME="EXAMPLE_VALUE"
#          # -----
#
#      o An enabled option looks like this:
#          # -----
#          LSF_OPTION_NAME="ACTUAL_VALUE"
#          # -----
#
# Instructions:
#   1. Edit install.config to specify the options for
#      your cluster. Uncomment the options you want and
#      replace the EXAMPLE values with your own settings.
#      Note that the sample values shown in this template
#      are EXAMPLES only. They are not always the default
#      installation values.
#
#   2. Run ./lsfinstall -f install.config
#
#
#*****
#      PART 1: REQUIRED PARAMETERS
#      (During an upgrade, specify the existing value.)
#*****
#
# -----
LSF_TOP="/shared/ibm/platform_lsf"
# -----
# Full path to the top-level installation directory {REQUIRED}
#
# The path to LSF_TOP must be shared and accessible to all hosts
# in the cluster. It cannot be the root directory (/).
# The file system containing LSF_TOP must have enough disk space for
# all host types (approximately 300 MB per host type).
#
# -----
LSFADMINS="phpcadmin root"
# -----
# List of LSF administrators {REQUIRED}
#
# The first user account name in the list is the primary LSF
# administrator. It cannot be the root user account.
# Typically, this account is named lsfadmin.
# It owns the LSF configuration files and log files for job events.
# It also has permission to reconfigure LSF and to control batch
# jobs submitted by other users. It typically does not have
# authority to start LSF daemons. Usually, only root has
# permission to start LSF daemons.
# All the LSF administrator accounts must exist on all hosts in the
# cluster before you install LSF.
# Secondary LSF administrators are optional.
#

```

```

# -----
LSF_CLUSTER_NAME="phpc_cluster"
# -----
# Name of the LSF cluster {REQUIRED}
#
# It must be 39 characters or less, and cannot contain any
# white spaces. Do not use the name of any host, user, or user group
# as the name of your cluster.
#
#
#*****
#          PART 2: PRIMARY PARAMETERS FOR NEW INSTALL
# (These parameters are ignored if they are already defined in the cluster.)
#*****
#
# -----
# LSF_MASTER_LIST="hostm hosta hostc"
# -----
# List of LSF server hosts to be master or master candidate in the
# cluster {REQUIRED when you install for the first time or during
# upgrade if the parameter does not exist.}
#
# You must specify at least one valid server host to start the
# cluster. The first host listed is the LSF master host.
#
#
# -----
LSF_ENTITLEMENT_FILE="/tmp/phpc/platform_lsf_std_entitlement.dat"
# -----
# You must specify a full path to the LSF entitlement file.
#
#
...

```

To perform the update after you enter all the environment variables in the configuration file (see Example 10-2 on page 134), run **lsfinstall**, as shown in Example 10-3.

Example 10-3 Run the lsfinstall command

```
[root@homecluster lsf9.1.3_lsfinstall]# ./lsfinstall -f install.config

Logging installation sequence in /tmp/phpc/lsf9.1.3_lsfinstall/Install.log

International Program License Agreement

.

.

.

Press Enter to continue viewing the license agreement, or
enter "1" to accept the agreement, "2" to decline it, "3"
to print it, "4" to read non-IBM terms, or "99" to go back
to the previous screen.
```

Read and accept the license agreement to proceed with the installation. After you finish reading and agreeing to the terms, press the number 1 key, and Platform LSF checks for the prerequisites. If the prerequisites are met, the installer prompts for the distribution .tar file to be used, as shown in Example 10-4.

Example 10-4 Platform LSF preinstallation check and distribution selection

LSF pre-installation check ...

```
Checking the LSF TOP directory /shared/ibm/platform_lsf ...
... Done checking the LSF TOP directory /shared/ibm/platform_lsf ...
```

You are installing IBM Platform LSF - 9.1.3 Standard Edition.

Checking LSF Administrators ...

```
LSF administrator(s):      "phpcadmin root"
Primary LSF administrator: "phpcadmin"
```

Checking the configuration template ...

CONFIGURATION_TEMPLATE not defined. Using DEFAULT template.

```
  Done checking configuration template ...
```

```
  Done checking ENABLE_STREAM ...
```

Checking the patch history directory ...

```
... Done checking the patch history directory /shared/ibm/platform_lsf/patch ...
```

Checking the patch backup directory ...

```
... Done checking the patch backup directory /shared/ibm/platform_lsf/patch/backup
```

...

Searching LSF 9.1.3 distribution tar files in /tmp/phpc Please wait ...

1) linux2.6-glibc2.3-x86_64

Press 1 or Enter to install this host type: 1

The installation proceeds without further prompts until a message is displayed that is similar to the one that is shown in Example 10-5.

Example 10-5 Installation completed successfully

You have chosen the following tar file(s):

```
lsf9.1.3_linux2.6-glibc2.3-x86_64
```

Checking selected tar file(s) ...

```
... Done checking selected tar file(s).
```

Pre-installation check report saved as text file:
/tmp/phpc/lsf9.1.3_lsfinstall/prechk.rpt.

... Done LSF pre-installation check.

.

.

.

```
Creating lsf_quick_admin.html ...
... Done creating lsf_quick_admin.html
```

lfsinstall is done.

To complete your LSF installation and get your cluster "phpc_cluster" up and running, follow the steps in "/tmp/phpc/lzf9.1.3_lfsinstall/lsf_getting_started.html".

After setting up your LSF server hosts and verifying your cluster "phpc_cluster" is running correctly, see "/shared/ibm/platform_lsf/9.1/lsf_quick_admin.html" to learn more about your new LSF cluster.

After installation, remember to bring your cluster up to date by applying the latest updates and bug fixes.

Note: For the latest release information about IBM Platform LSF Version 9.1.3, see the "IBM Platform LSF" topic in the IBM Knowledge Center at the following website:

http://www.ibm.com/support/knowledgecenter/SSETD4_9.1.3/lsf_welcome.html

For the latest information about IBM Platform HPC, see the "IBM Platform HPC" topic in the IBM Knowledge Center at the following website:

http://www-01.ibm.com/support/knowledgecenter/SSENRW_4.2.0/release_notes/release_notes.dita

Now, restart Platform HPC services to enable the new Platform LSF entitlement, as shown in Example 10-6.

Example 10-6 Restart Platform HPC services

```
[root@homecluster platform_lsf]# service phpc stop
Stopping Web Portal services                                [  OK  ]
Stopping PERF services                                     [  OK  ]
Stopping Rule Engine service                             [  OK  ]
Stopping PCMD service                                    [  OK  ]
Stopping Message broker                                 [  OK  ]
Stopping the LSF subsystem                               [  OK  ]

Stopping Platform HPC Services:                           [  OK  ]
[root@homecluster platform_lsf]# service phpc start
Checking for xcatt service started                         [  OK  ]

Starting the LSF subsystem                               [  OK  ]
- Waiting for EGO service started ..                  [  OK  ]

Cluster name : phpc_cluster EGO master host name : homecluster EGO master version
: 1.2.10
- Waiting for PCM master node online .....           [  OK  ]

Starting PERF services                                  [  OK  ]
Starting Message broker                               [  OK  ]
Starting PCMD service                                 [  OK  ]
```

```
Starting Rule Engine service [ OK ]
Starting Web Portal services [ OK ]
Starting Platform HPC Services: [ OK ]
[root@homecluster platform_lsf]#
```

Your cluster is ready to be configured as a multicluster server.

10.7.2 Tasks to install IBM Platform LSF in the cloud

The IBM Cloud Services team installs and configures the Platform LSF cluster for you in the cloud. You set up only your connection to the new cluster and configure your new multicluster feature.

Note: The connection to the cloud network can be done with a VPN or with MPLS. To connect to the cloud, you must add, to your Domain Name System (DNS) or hosts file, only the name of the master and master candidates of the new cloud cluster, and exchange the SSH keys between the hosts.

10.7.3 Configuring the multicluster feature

After you have exchanged the SSH keys, enable the multicluster feature. Copy both of your cluster definitions files (on-premises and in the cloud) on both \$LSF_TOP/conf/ directories, as shown in Example 10-7.

Example 10-7 Copy the cluster definition files between the master nodes

```
[root@homecluster conf]# scp
softlayer:/usr/share/lsf/conf/lsf.cluster.HPC_Services
/shared/ibm/platform_lsf/conf/
lsf.cluster.HPC_Services
100% 1801      1.8KB/s  00:00
[root@homecluster conf]# scp
/shared/ibm/platform_lsf/conf/lsf.cluster.phpc_cluster
softlayer:/usr/share/lsf/conf/
lsf.cluster.phpc_cluster 100% 2897 2.8KB/s   00:00
```

Now edit the \$LSF_TOP/conf/lsf.shared file and check that all the clusters are defined in the cluster stanza file, as shown in Example 10-8.

Example 10-8 LSF shared configuration file containing both clusters

```
# $Revision$Date$
#
# -----
# T H I S   F I L E: Is shared by all clusters in the LSF system.
#
# This file contains all definitions referenced by individual
# lsf.cluster.<clustername> files. The definitions in this file can be
# a superset, i.e., not all definitions in this file need to be used in
# other files.
#
# See lsf.cluster(5) and "LSF User's and Administrator's Guide".
# -----
```

```
Begin Cluster
ClusterName          # Keyword
phpc_cluster
HPC_Services
End Cluster
```

Note: Make the lsf.shared file the same on both clusters.

Now, as shown in Example 10-9, add a module to the lsb.modules file in the local cluster to see resources in the remote cluster. In this case, the file is in the homecluster server at the following path:

```
/install/shared/ibm/platform_lsf/conf/lbatch/phpc_cluster/configdir/lsb.modules
```

Example 10-9 Add schmod_mc to lsb.modules

```
# $Revision$Date$

# Define plug-ins for Scheduler and Resource Broker.
# SCH_PLUGIN column specifies the share module name for Scheduler, while
# RB_PLUGIN specifies the share module name for Resource Broker
# A Scheduler plug-in can have one, multiple, or none RB plug-ins
# corresponding to it.
# SCH_DISABLE_PHASES specifies which phases of that scheduler plug-in
# should be disabled, i.e., deactivated. A scheduler plug-in has four phases:
# pre processing, match/limit, order/alloc, post processing. Scheduler
# will not start disabled phases over jobs
# Note all share modules should be put under LSF_LIBDIR

Begin PluginModule
SCH_PLUGIN           RB_PLUGIN           SCH_DISABLE_PHASES
schmod_default       ()                  ()
schmod_fcfs          ()                  ()
schmod_fairshare     ()                  ()
schmod_limit          ()                  ()
schmod_mc            ()                  ()
schmod_parallel       ()                  ()
schmod_reserve        ()                  ()
schmod_preemption     ()                  ()
schmod_advrsv         ()                  ()
schmod_ps              ()                  ()
#schmod_dc             ()                  ()
End PluginModule
```

Restart the Platform LSF services on both the on-premises cluster and the cloud cluster, as shown in Example 10-10.

Example 10-10 Restart Platform LSF services

```
[root@homecluster conf]# lsdadmin lrestart all
```

```
Checking configuration files ...
No errors found.
```

```
Do you really want to restart LIMs on all hosts? [y/n] y
Restart LIM on <homecluster> ..... done
```

```
[root@homecluster conf]# badmin mbdrestart
Checking configuration files ...
No errors found.

MBD restart initiated

[root@softlayer ~]# lsadmin limrestart all
Checking configuration files ...
No errors found.

Do you really want to restart LIMs on all hosts? [y/n] y
Restart LIM on <softlayer> ..... done
[root@softlayer ~]# badmin mbdrestart

Checking configuration files ...

No errors found.

MBD restart initiated
[root@softlayer ~]#
```

To check whether the multicluster feature is correctly configured and these clusters are enabled to access each other, run **lscusters** and **bclusters** at the prompt to get OK status responses from both clusters, as shown in Example 10-11.

Example 10-11 Check the configuration

```
[root@homecluster ~]# lscusters
CLUSTER_NAME STATUS MASTER_HOST ADMIN HOSTS SERVERS
phpc_cluster ok homecluster phpcadmin 1 1
HPC_Services ok softlayer lsfadmin 1 1
[root@homecluster ~]# bclusters
[Job Forwarding Information ]
No local queue sending/receiving jobs from remote clusters

[Resource Lease Information ]
No resources have been exported or borrowed
[root@homecluster ~]#
```

10.7.4 Configuring job forwarding

This scenario shows how to change the high priority queue to send jobs to the IBM Platform Computing Cloud Services cluster. To do this task, change the `high_priority` stanza in the `lsb.queues` file at the local cluster (`phpc_cluster`). In this scenario, the master node `homecluster` path is the following one:

/install/shared/ibm/platform_lsf/conf/lbatch/phpc_cluster/configdir/lsb.queues

This scenario does not preempt running jobs because the idea is to show how to send jobs to the cloud instead of interrupting a running job, so comment the PREEMPTION line and add SNDJOBS_TO to point to the remote cluster. Then, change the description to state the usage of the queue. All changes are shown in bold in Example 10-12.

Example 10-12 lsb.queues on the local cluster

```
Begin Queue
QUEUE_NAME      = high_priority
PRIORITY        = 43
NICE            = 10
SNDJOBS_TO     = receive@HPC_Services
#PREEMPTION     = PREEMPTIVE
#RUN_WINDOW
#CPULIMIT       = 8:0/SunIPC      # 8 hours of host model SunIPC
#FILELIMIT      = 20000
#DATALIMIT      = 20000          # jobs data segment limit
#CORELIMIT      = 20000
#PROCLIMIT      = 5              # job processor limit
#USERS          = user1 user2 user3
#HOSTS          = all
#ADMINISTRATORS   = user1 user3
#EXCLUSIVE      = N
#PRE_EXEC       = /usr/local/lsf/misc/testq_pre >> /tmp/pre.out
#POST_EXEC      = /usr/local/lsf/misc/testq_post |grep -v "Hey"
#REQUEUE_EXIT_VALUES = 55 255 78
DESCRIPTION    = Jobs submitted for this queue are scheduled as urgent\jobs. Jobs in this queue can jobs in this queue can be forwarded to the Cloud Services Cluster.
End Queue
```

In similar fashion, configure the receiving side to handle the jobs coming from the high priority queue. The lsb.queues files in the remote cluster, for this scenario, can be found in the SoftLayer host of the HPC_Services Cluster at the following path:

```
/usr/share/lsf/conf/lsbatch/HPC_Services/configdir/lsb.queues
```

In lsb.queues, add a stanza at the end of the file, as shown in Example 10-13.

Example 10-13 lsb.queues on the remote cluster

```
Begin Queue
QUEUE_NAME=receive
RCVJOBS_FROM=high_priority@phpc_cluster
PRIORITY=70
NICE=20
End Queue
```

Now, reconfigure the queues on both sides, as shown in Example 10-14.

Example 10-14 Reconfigure the queues

```
[root@softlayer ~]# badmin mbdrestart
```

```
Checking configuration files ...
```

```
No errors found.
```

```
MBD restart initiated  
[root@softlayer ~]#
```

```
[root@homecluster ~]# badmin mbdrestart
```

```
Checking configuration files ...
```

```
No errors found.
```

```
MBD restart initiated  
[root@homecluster ~]#
```

Example 10-15 shows how to check the job forwarding status configuration for the local and remote queues.

Example 10-15 Check job forwarding

```
[root@homecluster ~]# bclusters  
[Job Forwarding Information ]  
LOCAL_QUEUE      JOB_FLOW      REMOTE      CLUSTER      STATUS  
high_priority    send         receive     HPC_Service ok  
  
[Resource Lease Information ]  
No resources have been exported or borrowed  
[root@homecluster ~]#
```

```
[root@softlayer ~]# bclusters  
[Job Forwarding Information ]  
LOCAL_QUEUE      JOB_FLOW      REMOTE      CLUSTER      STATUS  
receive          recv         -           phpc_clust ok  
  
[Resource Lease Information ]  
No resources have been exported or borrowed  
[root@softlayer ~]#
```

10.7.5 Testing your configuration

Now, test the new configuration by using the command-line interface (CLI) or the graphical user interface (GUI) from Platform HPC. If you have Platform LSF with Platform Application Center, you can use this interface as well.

Note: To submit the job to the cloud, the user must have authority to run jobs on the receiving queue.

This scenario uses the CLI to submit the jobs. Example 10-16 shows how to run the **bsub** command to a dummy sleep job.

Example 10-16 Submit jobs to the respective queues

```
[root@homecluster ~]# bsub -q high_priority sleep 50
Job <857> is submitted to queue <high_priority>.
[root@homecluster ~]#
[root@homecluster ~]# bsub -q medium_priority sleep 50
Job <858> is submitted to queue <medium_priority>.
[root@homecluster ~]#
[root@homecluster ~]# bsub -q medium_priority sleep 50
Job <859> is submitted to queue <medium_priority>.
[root@homecluster ~]#
[root@homecluster ~]# bsub -q high_priority sleep 50
Job <860> is submitted to queue <high_priority>.
[root@homecluster ~]#
```

In this case, four jobs are submitted in a row, but the only queue that can forward jobs is the `high_priority` one. There are only two slots in the on-premises environment, so submit three jobs in the `medium_priority` queue and the last job in the `high_priority` queue. Only the job with the high priority runs in the cloud.

Example 10-17 shows the running jobs and the pending jobs in their respective queues.

Example 10-17 Jobs running in the cluster

```
[root@homecluster ~]# bjobs
JOBID  USER    STAT  QUEUE      FROM_HOST   EXEC_HOST   JOB_NAME   SUBMIT_TIME
857    root    RUN   high_prios homecluster homecluster *813545588 Mar 31 10:52
858    root    RUN   medium_prio homecluster homecluster *813554430 Mar 31 10:52
860    root    RUN   high_prios homecluster softlayer@H *813563747 Mar 31 10:53
859    root    PEND  medium_prio homecluster homecluster *813557373 Mar 31 10:52
```

As you can see, the first preference is to use available slots in the home cluster, and after there is no resource that is available, then only the high priority job goes to the cloud, even after being submitted after the last medium priority job.

Note: This is an example on how Platform LSF sends jobs from only a configured queue. Platform LSF is a powerful tool that helps you do advanced scheduling, and provides the best policies to suit your business needs.

10.7.6 Hybrid cloud is ready

The previous sections described the steps to configure a hybrid cloud in a few steps, although with the help of IBM Platform Computing Cloud Services, customers do not need to worry about configuring and managing a cloud infrastructure.

After following the five simple steps that were described in previous sections, you can have extra capacity added to receive jobs from your existing environment. Nevertheless, if you need assistance to configure a hybrid cloud environment, contact the IBM Platform Computing Services team for help and assistance.

10.8 Data management on hybrid clouds

Two easy ways to manage data across hybrid clouds are by implementing IBM Platform Data Manager for LSF and IBM Spectrum Scale AFM. Both technologies optimize data transfer needs to reduce costs and time to results because only the required data is moved at the correct time.

10.8.1 IBM Platform Data Manager for LSF

Platform Data Manager for LSF automates the transfer of data that is used by application workloads running on Platform LSF clusters and in the cloud. Frequently used data that is transferred between multiple data centers and the cloud can be stored in a smart, managed cache closer to compute resources. This smart data management helps to improve data throughput and minimizes wasted compute cycles, which helps you lower storage costs in the cloud.

With Platform Data Manager, the following actions occur:

- ▶ Data is staged in and out independently of workloads, freeing compute resources while data is transferred behind the scenes.
- ▶ A smart, managed cache reuses transferred data and avoids duplication of data transfers, sharing cached copies with all workloads that need access to the data, and among multiple users where appropriate.
- ▶ Data transfers are scheduled as jobs in Platform LSF and are subject to Platform LSF scheduling policies that are established by administrators, including priority.

For more information about IBM Platform Data Manager for LSF, see the following website:

<http://www.ibm.com/systems/platformcomputing/products/lsf/datamanager.html>

10.8.2 IBM Spectrum Scale Active File Management

AFM is a scalable, high-performance, file system caching layer that is integrated with Spectrum Scale. You can use AFM to create associations from a local cluster to a remote cluster or storage, and to define the location and flow of file data to automate the management of the data to implement a single namespace view across sites around the world.

AFM masks wide area network (WAN) latencies and outages by using Spectrum Scale to cache massive data sets, allowing data access and modifications even when a remote storage cluster is unavailable. In addition, AFM performs updates to the remote cluster asynchronously, which allows applications to continue operating while not being constrained by limited outgoing network bandwidth.

The AFM implementation uses the inherent scalability of Spectrum Scale to provide a multinode, consistent cache of data that is in a home cluster. By integrating it with the file system, AFM provides a Portable Operating System Interface (POSIX)-compliant interface, making the cache transparent to applications. AFM is easy to deploy, as it relies on open standards for high-performance file serving and does not require any proprietary hardware or software to be installed at the home cluster.

For a step-by-step configuration information, see the following website:

<http://ibm.co/1bPKBfY>



A

IBM Platform Computing Message Passing Interface

This appendix introduces the IBM Platform Computing Message Passing Interface (MPI), and describes how it is implemented.

This appendix covers the following topics:

- ▶ IBM Platform Computing Message Passing Interface
- ▶ IBM Platform Computing Message Passing Interface implementation

IBM Platform Computing Message Passing Interface

IBM Platform MPI is a high-performance and production-quality implementation of the Message Passing Interface standard. It supports the broadest range of industry-standard platforms, interconnects, and operating systems to help ensure that parallel applications can run on any platform. It fully complies with the MPI-2.2 standard and provides enhancements, such as low latency and high-bandwidth point-to-point and collective communication routines, over other implementations. IBM Platform MPI V8.3 for Linux is supported on Intel/AMD x86 32-bit, AMD Opteron, and EM64T servers that run CentOS 5, Red Hat Enterprise Linux AS 4, 5, and 6, and SUSE Linux Enterprise Server 9, 10, and 11 operating systems.

For more information about IBM Platform MPI, see the *IBM Platform MPI User's Guide*, SC27-4758.

IBM Platform Computing Message Passing Interface implementation

To install IBM Platform MPI, you must download the installation package. The installation package contains a single script that, when you run it, decompresses itself and installs the MPI files in the designated location. There is no installation manual that is available, but the installation is as simple as running the script in the installation package.

Help: For more information about how to use the installation script, run the following command:

```
sh platform_mpi-08.3.0.0-0320r.x64.sh -help
```

When you install IBM Platform MPI, even if you give an installation directory as input to the script, all files are installed under the /opt/ibm/platform_mpi directory. Example A-1 shows the installation log of a successful installation. Example A-1 provides the shared directory /gpfs/fs1 as the installation root. After the installation, the files are available in the /gpfs/fs1/opt/ibm/platform_mpi directory.

Example A-1 IBM Platform MPI - installation log

```
[root@i05n45 PlatformMPI]# sh platform_mpi-08.3.0.0-0320r.x64.sh  
-installdir=/gpfs/fs1 -norpm
```

```
Verifying archive integrity... All good.  
Uncompressing platform_mpi-08.3.0.0-0316r.x64.sh.....  
Logging to /tmp/ibm_platform_mpi_install.JS36  
International Program License Agreement
```

Part 1 - General Terms

```
BY DOWNLOADING, INSTALLING, COPYING, ACCESSING, CLICKING ON  
AN "ACCEPT" BUTTON, OR OTHERWISE USING THE PROGRAM,  
LICENSEE AGREES TO THE TERMS OF THIS AGREEMENT. IF YOU ARE  
ACCEPTING THESE TERMS ON BEHALF OF LICENSEE, YOU REPRESENT  
AND WARRANT THAT YOU HAVE FULL AUTHORITY TO BIND LICENSEE  
TO THESE TERMS. IF YOU DO NOT AGREE TO THESE TERMS,
```

```
* DO NOT DOWNLOAD, INSTALL, COPY, ACCESS, CLICK ON AN
```

"ACCEPT" BUTTON, OR USE THE PROGRAM; AND

* PROMPTLY RETURN THE UNUSED MEDIA, DOCUMENTATION, AND
PROOF OF ENTITLEMENT TO THE PARTY FROM WHOM IT WAS OBTAINED

Press Enter to continue viewing the license agreement, or
enter "1" to accept the agreement, "2" to decline it, "3"
to print it, or "99" to go back to the previous screen.

1

Installing IBM Platform MPI to /gpfs/fs1/
Installation completed.

When you install IBM Platform MPI on the shared directory of a cluster, avoid using the local rpmdb of the server where you are installing MPI. You can use the **-norpm** option to extract all of the files to the installation directory and disable interaction with the local rpmdb.

If you are not installing IBM Platform MPI on a shared directory, you must install it in all hosts of the cluster that run applications that use MPI. The installation must be done in the same directory in all hosts.

Before you can start using IBM Platform MPI, you must configure your environment. By default, MPI uses Secure Shell (**ssh**) to connect to other hosts, so if you want to use a different command, you must set the environment variable **MPI_REMSH**. Example A-2 shows how to set up your environment and run **hello_world.c** (an example program that ships with IBM Platform MPI) to run on the cluster with four-way parallelism. The application runs the hosts i05n47 and i05n48 of the cluster.

Example A-2 IBM Platform MPI - Running a parallel application

```
[root@i05n49 PlatformMPI]# export MPI_REMSH="ssh -x"
[root@i05n49 PlatformMPI]# export MPI_ROOT=/gpfs/fs1/opt/ibm/platform_mpi
[root@i05n49 PlatformMPI]# /gpfs/fs1/opt/ibm/platform_mpi/bin/mpicc -o
/gpfs/fs1/helloworld /gpfs/fs1/opt/ibm/platform_mpi/help/hello_world.c
[root@i05n49 PlatformMPI]# cat appfile
-h i05n47 -np 2 /gpfs/fs1/helloworld
-h i05n48 -np 2 /gpfs/fs1/helloworld
[root@i05n49 PlatformMPI]# /gpfs/fs1/opt/ibm/platform_mpi/bin/mpirun -f appfile
Hello world! I'm 1 of 4 on i05n47
Hello world! I'm 0 of 4 on i05n47
Hello world! I'm 2 of 4 on i05n48
Hello world! I'm 3 of 4 on i05n4
```



B

LDAP server configuration and management

This appendix shows how to configure a simple LDAP server and how to manage user accounts to be used in the IBM Platform Cluster Manager or any other IBM Platform product offer that supports LDAP authentication.

For this tutorial, the following assumptions are taken into account:

- ▶ Red Hat Linux 6.5 is installed and configured to install packages from DVD or from the repositories.
- ▶ User accounts are configured with their corresponding HOME directories in the LDAP server.

This appendix covers the following topics:

- ▶ OpenLDAP installation
- ▶ LDAP user account management

OpenLDAP installation

For this tutorial, install OpenLDAP as the LDAP server for Red Hat Linux 6.5. Different applications can be used, but the instructions for the installation and configuration might differ; adjust accordingly.

Install the following RPMs by running the following command:

```
yum install -y openldap openldap-servers openldap-clients
```

After the installation completes, open and edit the /etc/openldap/slapd.d/cn=config/olcDatabase={0}config.ldif file and change the lines as shown in Example B-1.

Example B-1 Edit /etc/openldap/slapd.d/cn=config/olcDatabase={0}config.ldif

```
olcRootDN: cn=Manager,dc=platform,dc=itso,dc=ibm,dc=com
```

Change the DN to reflect your scenario.

Next, edit the /etc/openldap/slapd.d/cn=config/olcDatabase={2}bdb.ldif file and change the lines as shown in Example B-2.

Example B-2 Edit /etc/openldap/slapd.d/cn=config/olcDatabase={2}bdb.ldif

```
olcSuffix: dc=platform,dc=itso,dc=ibm,dc=com
olcRootDN: cn=Manager,dc=platform,dc=itso,dc=ibm,dc=com
olcRootPW: {SSHA}vM0c7VqI1vNW1AyzQDQVd7DW4xxa5YF6
olcAccess: {0}to attrs=userPassword,shadowLastChange by self write by
dn.base="cn=Manager,dc=platform,dc=itso,dc=ibm,dc=com" write by anonymous auth by
anonymous search by * none
olcAccess: {1}to * by dn.base="cn=Manager,dc=platform,dc=itso,dc=ibm,dc=com" write
by self write by * read
```

Hint: You can run the following command to generate the SSHA hash:

```
slappasswd -h {SSHA} -s <plain text password>
```

Because you want to test your LDAP server without specifying the bind DN all the time, you can edit the /etc/openldap/ldap.conf file, which tells the LDAP client which base DN to use and which URL, as shown in Example B-3.

Example B-3 Edit /etc/openldap/ldap.conf

```
URI ldap://localhost:389
BASE dc=platform,dc=itso,dc=ibm,dc=com
TLS_CACERTDIR /etc/openldap/cacerts
```

Check that you edit the information in Example B-3 according to your setup. After changing these files, restart the LDAP server by running the following command:

```
# service slapd restart
```

Reminder: Check your firewall or SELinux configuration. By default, OpenLDAP expects you to use both TCP ports 389 and 636.

OpenLDAP should be up and running. Now you can test it by running the following command:

```
# ldapsearch -x
```

You do not need to specify the host or the base DN because the `ldap.conf` file was edited to contain this information.

LDAP user account management

OpenLDAP uses the LDAP Data Interchange Format (LDIF) specification that is used to describe the directory information or modification of a particular directory. The LDIF format is used to import or export the data to or from an LDAP server. For this tutorial, create an LDIF file to import some users into your LDAP directory.

Note: For more information about the LDIF specification, see RFC2849 at the following website:

<http://tools.ietf.org/html/rfc2849>

Example B-4 shows how an LDIF file can be written. It contains all the required attributes to create a user account by following the Portable Operating System Interface (POSIX).

The LDIF file requires that some object classes be loaded before you specify the attributes. The objects have the specification for each attribute or entry. When creating POSIX accounts, specify `objectClass: posixAccount` for each entry.

Example B-4 LDIF for creating user accounts and an Organization Unit (OU) called “team”

```
dn: ou=team,dc=platform,dc=itso,dc=ibm,dc=com
objectClass: top
objectClass: organizationalUnit
ou: team

dn: cn=Tiago Mello,ou=team,dc=platform,dc=itso,dc=ibm,dc=com
objectClass: top
objectClass: posixAccount
objectClass: organizationalPerson
objectClass: inetOrgPerson
objectClass: person
cn: Tiago Mello
gidNumber: 100
homeDirectory: /home/tmello
sn: Mello
uid: tmello
uidNumber: 1001
givenName: Tiago
mail: tmello@myemail.com
userPassword:: e2NyeXB0fSQxJFN1MFdNOExYJG9CQkFQNWyvZHg4QnowcDRuZ2F1eTA=
```

Change the `uidNumber` and `gidNumber` to match the existent user account in the LDAP system. The user HOME directory can be automatically created on the first login, but that topic is not described in this document.

To import the accounts into the directory, run the following command:

```
# ldapadd -Y EXTERNAL -H ldapi:/// -f accounts.ldif
```

Now, you can list all the content of your directory by running the following command:

```
#ldapsearch -x
```

You can also edit an existent entry and modify some attributes, as shown in Example B-5.

Example B-5 LDIF for modifying an existent LDAP entry - changeaccount.ldif

```
dn: cn=Tiago Mello,ou=team,dc=platform,dc=itso,dc=ibm,dc=com
changetype: modify
replace: userPassword
userpassword: {SSHA}vM0c7VqI1vNW1A
```

The following command makes the change that is described in the LDIF file:

```
ldapmodify -D "cn=Manager, dc=platform,dc=itso,dc=ibm,dc=com" -w rootdnpasswd <
changeaccount.ldif
```

Related publications

The publications that are listed in this section are considered suitable for a more detailed description of the topics that are covered in this book.

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Some publications referenced in this list might be available in softcopy only.

- ▶ *IBM Platform Computing Solutions*, SG24-8073
- ▶ *IBM Platform Computing Solutions Reference Architectures and Best Practices*, SG24-8169
- ▶ *Implementing an Advanced Application Using Processes, Rules, Events, and Reports*, SG24-8065
- ▶ *Implementing IBM InfoSphere BigInsights on IBM System x*, SG24-8077

You can search for, view, download, or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

Other publications

This publication is also relevant as a further information source:

- ▶ *IBM Platform MPI User's Guide*, SC27-4758

Online resources

These websites are also relevant as further information sources:

- ▶ Algorithmics Software
<http://www-01.ibm.com/software/analytics/algorithmics/>
- ▶ Big Data and the Speed of Business
<http://www-01.ibm.com/software/data/bigdata/industry.html>
- ▶ IBM Platform Computing
<http://www-03.ibm.com/systems/platformcomputing/products/index.html>

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Redbooks

IBM Platform Computing Solutions for High Performance and Technical Computing Workloads

(0.2"spine)
0.17" <-> 0.473"
90<->249 pages



SG24-8264-00

ISBN 0738440752

Printed in U.S.A.

Get connected

