

# IBM Spectrum LSF Data Manager



*Optimize cluster throughput with cached data transfers*

## Highlights

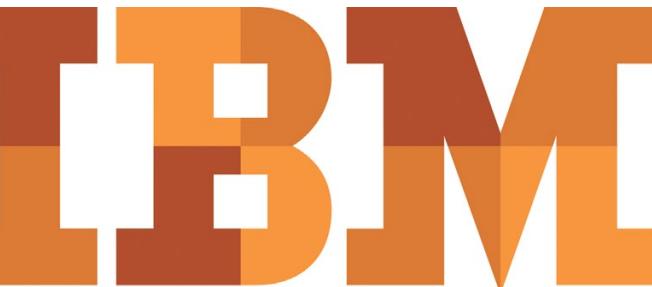
- Manage data transfers independent of cluster workloads to improve throughput and optimize use of compute resources
- Leverage a smart, managed cache to eliminate duplication of data transfers and lower storage costs
- Gain full visibility and control of data transfer jobs using IBM® Spectrum LSF scheduling policies
- Simplify administration with the ability to configure dedicated I/O nodes for inbound/outbound data transfers
- Provide user and administrator access control over cached data

As global organizations consolidate IT resources in larger, centralized data centers and deliver services through public and private clouds, application performance suffers. With business users, customers and partners all accessing the same systems and data, critical workloads can slow down as data moves from storage resources to compute resources and back again.

## Schedule data transfers to reduce costs

IBM Spectrum LSF Data Manager takes control of data transfers to help organizations improve data throughput and lower costs by minimizing wasted compute cycles and conserving disk space. IBM Spectrum LSF Data Manager automates the transfer of data used by application workloads running on IBM Spectrum LSF clusters and the cloud, bringing frequently used data closer to compute resources by storing it in a smart, managed cache that can be shared among users and workloads.

Data is staged in and out independent of workloads, freeing compute resources to work on other jobs while data is transferred back and forth among compute resources and storage. IBM Spectrum LSF Data Manager also leverages the data movement infrastructure already in place, allowing it to be easily integrated into existing cluster infrastructure.



## Smart, managed cache

IBM Spectrum LSF Data Manager leverages smart, managed cache to reuse data common to multiple workloads and accelerate time to results. Transferred files are automatically cached on the execution cluster with an optional time to live. Cached copies can be leveraged for all workloads that need access to the data, and can be shared between multiple users where appropriate. Workloads running on execution clusters can also write intermediate data to the local cache (such as restart data) for use by other jobs.

## Precise control of data movement

IBM Spectrum LSF Data Manager allows data transfers to be scheduled as jobs in IBM Spectrum LSF. With a centralized view of transfer traffic and control over both the number of concurrent transfers and the individual priority of those transfers, organizations gain precise control over data movement by workload or project. Transfers can be established that are prioritized by user, or the transfer of input data sets can be prioritized over the return of results.

## Why IBM?

IBM Spectrum Computing offers a comprehensive portfolio of software defined infrastructure solutions designed to help your organization deliver IT services in the most efficient way possible, optimizing resource utilization to speed time to results and reduce costs. These offerings help maximize the potential of your infrastructure to accelerate your analytics, high-performance computing (HPC), Hadoop, Apache Spark and cloud-native applications at any scale. The core value of the portfolio is simplifying simulations and analysis to help you uncover insights into your business or science and get higher-quality products to market faster.

## For more information

To learn more about IBM Spectrum LSF Data Manager, contact your IBM representative or IBM Business Partner, or visit:

[ibm.com/systems/spectrum-computing/products/lsf/resources.html](http://ibm.com/systems/spectrum-computing/products/lsf/resources.html)



---

© Copyright IBM Corporation 2016

IBM Systems  
Route 100  
Somers, NY 10589

Produced in the United States of America  
June 2016

IBM, the IBM logo, ibm.com, and LSF are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at [ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml)

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.



Please Recycle