IBM System Storage

Optimizing Splunk (using SmartStore) with IBM Cloud Object Storage

Solution Brief

splunk >

IBM Cloud **Object Storage**

Integration information

Target On-Premise IBM COS Version	3.14.3.15
Target Partner Product Version	Splunk Enterprise 7.3.0

This document provides information between Splunk Enterprise's SmartStore feature that was released with Splunk version 7.2 and IBM Cloud Object Storage for both On-Premise installations and public cloud in terms of functionality.

Intended Audience

This document is intended for anyone interested in understanding the functional aspects of the integration of Splunk with IBM Cloud Object Storage (COS).

Scope

This document assumes that the readers are aware the functional aspects of Splunk Enterprise and wants to understand how Splunk's SmartStore feature works with IBM Cloud Object Storage. The intention of the document is to cover **Splunk Enterprise's SmartStore** functionality with IBM Cloud Object over the published Cloud Object Storage API.

This document only covers testing ISV specific features and functionality at the time of the qualification. Hardware, software and overall environmental differences can affect overall performance.

Increase the benefits of Splunk solutions for more insights at lower costs and an easier to manage scalable solution.
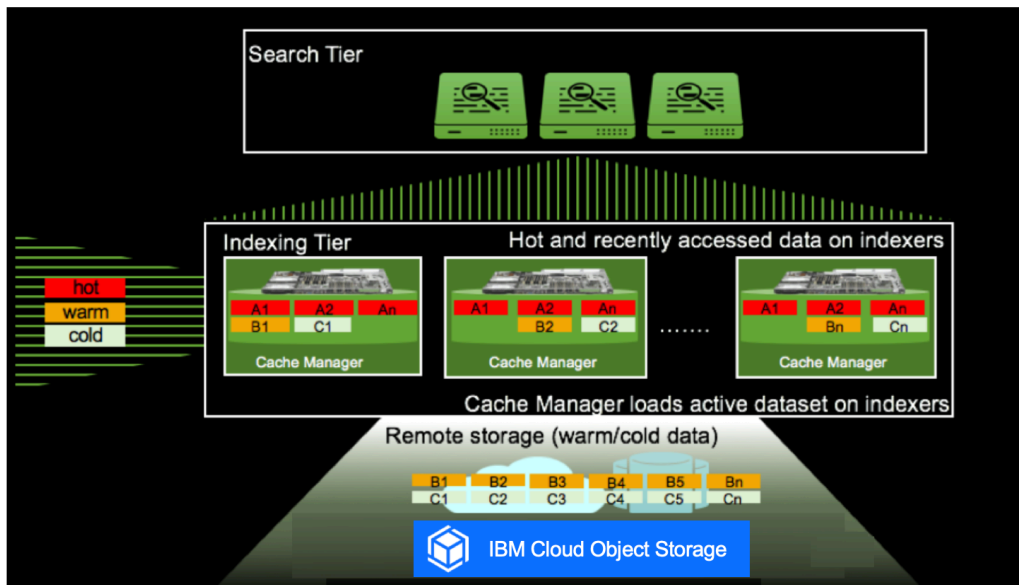
Splunk with IBM Cloud Object Storage

Generating value from data is expensive

Splunk derives insights from information and data that that can be obtained from many different sources and locations in the organization. Information Technology or IT is not the only location where valuable data is located and created as various lines of business all generate vast volumes of log data and metadata that can also be used to gain valuable insights. Splunk enables enterprises and organizations to analyze their data to make better decisions. The more data that can analyzed the better decisions and business insights can be made. The problem is that collecting and storing all that data can be complicated and expensive. Scaling data at a rapid pace is difficult and expensive for organizations. The goal of IT should be to drive down the cost of maintaining an expanding amount of data while driving up performance and creating more insights so data that is analyzed deliver better value.

## A better way to optimize Splunk solutions using IBM Cloud Object Storage

Splunk recently has introduced a new efficient way to store and manage data using capability called **SmartStore**. This is a new indexer capability provides a way to use

remote object stores, such IBM Cloud Object Storage (COS) to store indexed data.   Indexed data can now leverage lower cost IBM COS and easily scale as requirement grow with minimal effect on performance.

As a deployment's data volume increases, demand for storage typically outpaces demand for compute resources. SmartStore allows you to manage your indexer storage and compute resources in a cost-effective manner by scaling those resources separately. A new remote storage tier and a **cache manager** has also been introduced to optimize performance as data moves in and out of the object storage capacity tier. These features allow data to reside either locally on indexers or on the remote IBM Cloud Object Storage tier. Data movement between the indexer and the remote storage tier is managed by the cache manager, which resides on the indexer.

With SmartStore, you can reduce the indexer storage footprint to a minimum and choose I/O optimized compute resources. Most data resides on remote storage, while the indexer maintains a local cache that contains a minimal amount of data: hot buckets, copies of warm buckets participating in active or recent searches, and bucket metadata.

You can enable SmartStore for all indexes or for a subset of indexes. This allows customers to start using the SmartStore capabilities with IBM Cloud Object Storage in a small configuration and easily scale as data requirements grow without a forklift upgrade to current resources.


Splunk with SmartStore

4
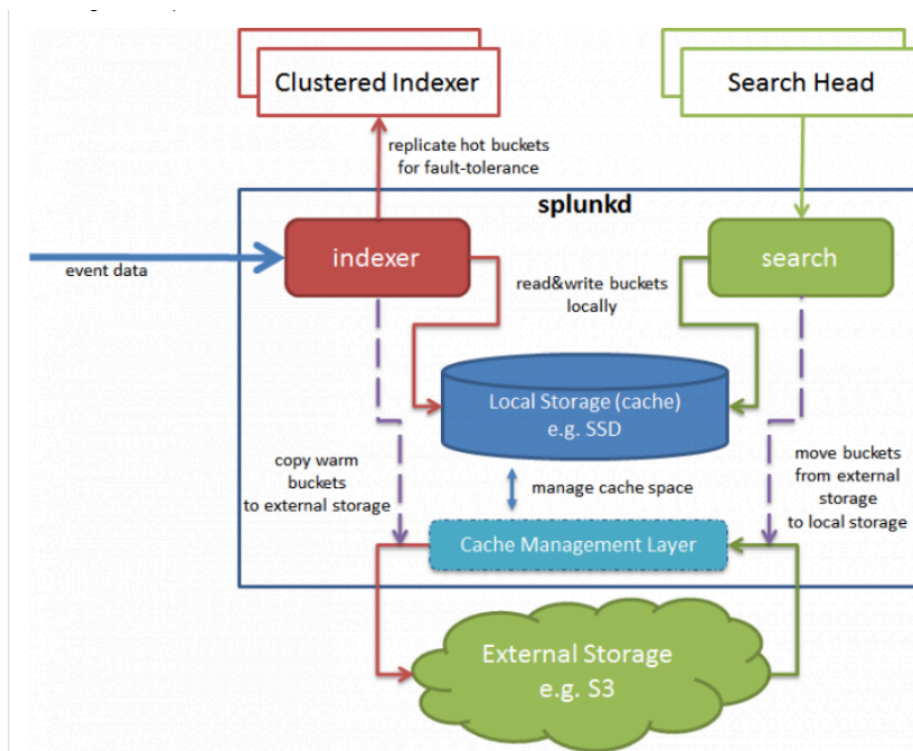
Advantages of leveraging IBM Cloud Object Storage

With this new feature, **Local Storage** can be easily extended with IBM Cloud Object Storage as a **Warm Bucket**. Since most Splunk searches will be on recently indexed data, this allows for fast searches from local storage. Splunk's Cache Management Layer is central to the SmartStore data flow, and it evicts older/less searched copies of indexed data from cache. The Cache Manager also fetches copies of indexed data from IBM COS, as necessary, to handle search requests.

The integration of Splunk's SmartStore with IBM Cloud Object Storage allows clients to minimize the amount of data on local storage, while maintaining the fast indexing and search capabilities characteristic of Splunk Enterprise deployments. Splunk claims that except in a few uncommon scenarios, indexers return search results for SmartStore enabled indices with **speeds similar to those for non SmartStore indices**.

Here's another way to show how Splunk's SmartStore feature decouples compute from storage, and shows IBM Cloud Object Storage storing Splunk's Warm buckets on cost effective IBM Cloud Object Storage:
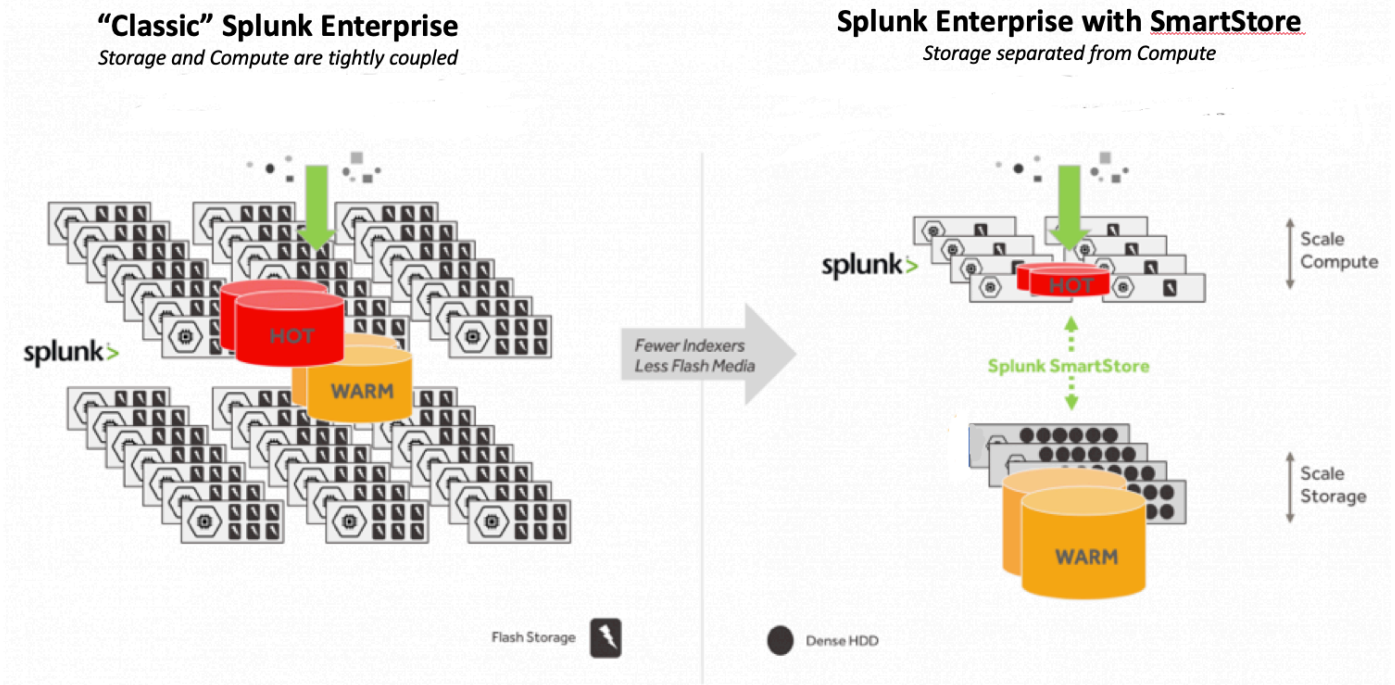
Splunk before and after SmartStore

IBM offers a market leading solution for the data center for Splunk data as well as providing a cloud infrastructure in the data center using object storage. IBM Cloud Object Storage provides data that is accessible from any network accessible location with simplicity and efficiency to lower cost for growing and modern data requirements. With proven scalability to EB capacity but efficiency to start in small 72TB configurations, our solution provides a strong foundation to keep your data safe and available for more than a lifetime.

IBM Cloud Object Storage is a very parallel system meaning its an any to any to any architecture. There is no single point of failure or bottle and our system is balanced throughout our architecture. Any application can access any accessor (through an IP address) and the access spreads the data to any Slicestor.
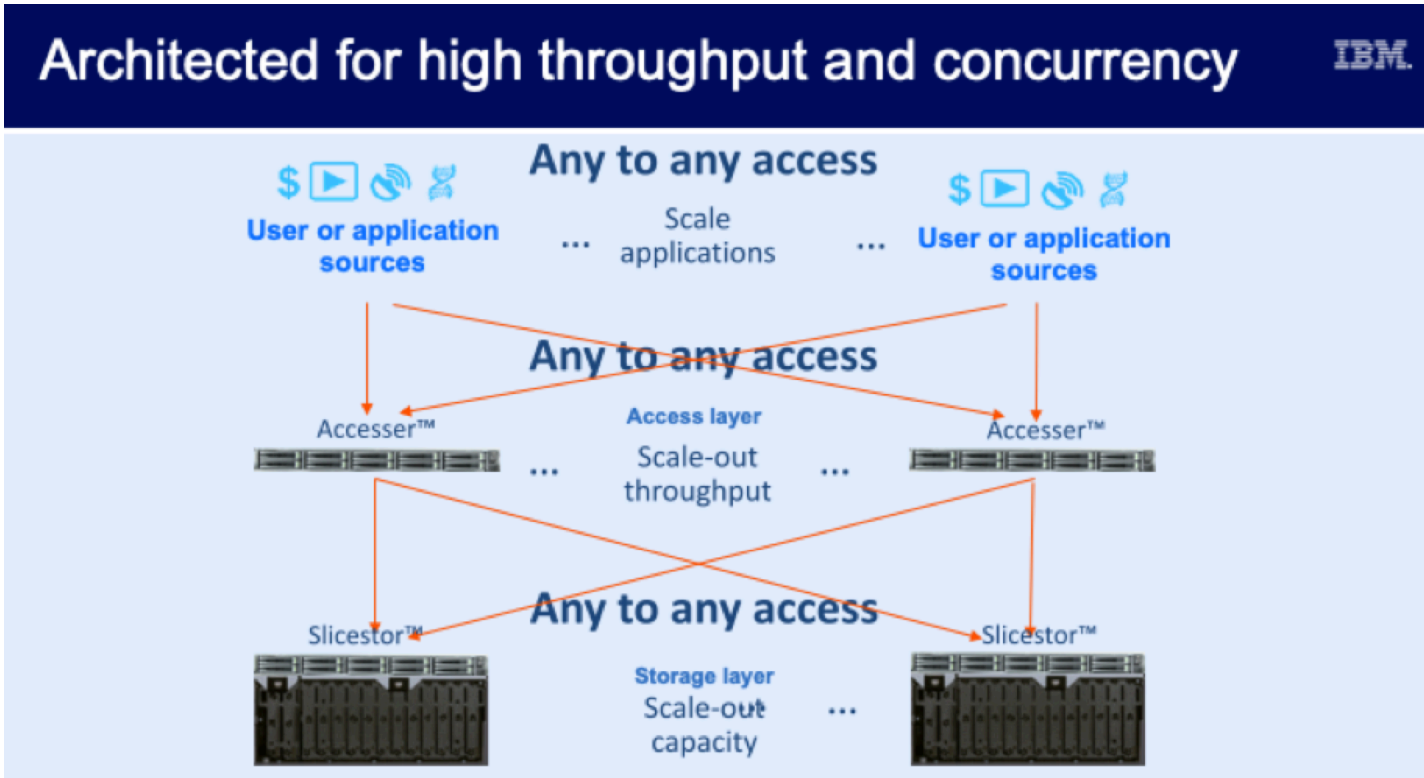
Architected for throughput and concurrency

Our solution takes innovative technology and brings business benefits to your organization now and provides the foundation for a modern data infrastructure for years into the future. IBM Cloud Object Storage is turning storage challenges into business value. With IBM's object storage platform, you can select the best configuration and approach to address the unique application, data and workload requirements for your business. The platform is backed by over 800 patents and leverageable by current and new modern applications such as containers, micro services and cloud native applications.

	Slicestor™ 12	Slicestor™ 53	Slicestor™ 84	Slicestor™ 106
Total rack space	3U	5U	5U	5U
Disk configurations	12	14, 28, 41, 53	28,56, 84	64, 78, 92, 106
Drive sizes	4, 8, 12, 14, 16 TB	4, 8, 12, 14, 16 TB	4, 8, 12 TB	4, 8, 12, 14, 16 TB
Efficiency Options (IDAs. X/X/X)	1000's of combinations to support optimized levels of availability, performance and capacity			
Node capacity raw (min/max)	48TB/192TB	56TB/848TB	112TB/1.01PB	256TB/1.70PB

Flexibility to optimize configurations for cost, performance, capacity, availability and efficiency

Investment Protection	Lower Costs	High Performance	New Efficiency
No migration required; Mix configurations	Low cost options	Scale-out throughput	Consolidated flexibility

IBM Cloud Object Storage Capacity Storage

# The benefits of Splunk SmartStore with IBM COS

- Reduced storage cost. Your deployment can take advantage of the economy of IBM Cloud Object Storage, instead of relying on costly local storage.
- Access to high availability (up to 99.999999) and data resiliency (up to 99.9999999999999) of IBM Cloud Object Storage.
- The ability to scale compute and storage resources separately, thus ensuring that you use resources efficiently.
- Simple and flexible configuration with per-index settings.
- A bootstrapping capability that allows a new cluster or standalone indexer to inherit data from an old cluster or standalone indexer.

SmartStore offers additional advantages specific to deployments of indexer clusters:

- Fast recovery from peer failure and fast data rebalancing, requiring only metadata fixups for warm data.
- Lower overall storage requirements, as the system maintains only a single permanent copy of each warm bucket.
- Full recovery of warm buckets even when the number of peer nodes that goes down is greater than or equal to the replication factor.
- Simplified upgrades.

IBM Cloud Object Storage optimizes Splunk storage with flexibility

3 Site Configuration (geo-dispersed)
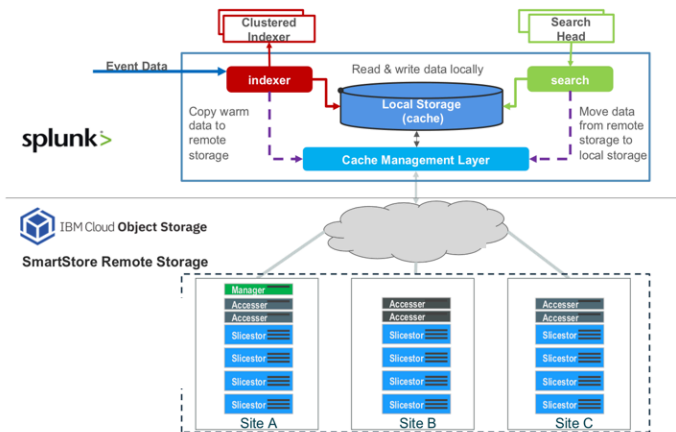
2 Site Configuration (mirrored)
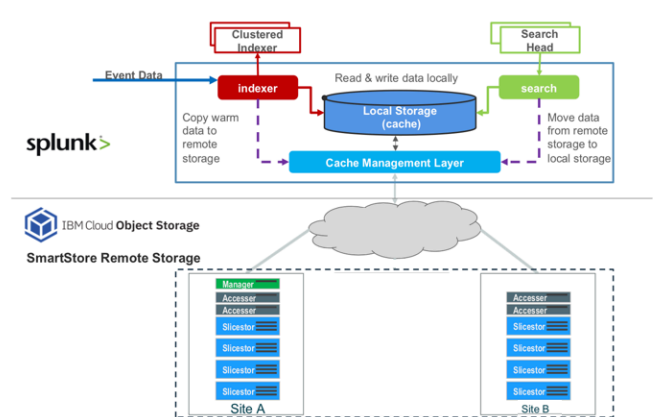
Flexible configurations

An intelligent cache manager ensures that, for most search use cases, SmartStore provides similar performance to local storage configurations.

Using SmartStore with IBM Cloud Object Storage allows your Splunk instances to store/access petabytes of storage without having to add additional Splunk indexers.

Splunk with SmartStore changes the deployment model by rendering indexers stateless for warm and cold data, making the indexer highly elastic and simplifying the management of a Splunk deployment.   Not only is it easier to scale Splunk indexers as necessary, Splunk with SmartStore also allows easy migration of your indexers to higher performance hardware, without having to worry about the typical storage migration.

Splunk with SmartStore also provides more options than the normal Splunk cluster backup and restore processes by allowing an entire set of indexers in the cluster to be shutdown/replaced and later revived by bootstrapping the data from IBM Cloud Object Storage.

A customer success

Multi-billion dollar enterprise leverages IBM COS for Splunk

IBM Cloud Object Storage allowed customer to bring more business insights with less cost

Splunk SmartStore
solution leverages IBM COS S3 API for easy access

Growing to 10PB+
after success with PB capacity, this customer has already scaled capacity over 3x and is bringing more insights to business with more data

Performance
was not effected when moving to Splunk SmartStore with IBM COS

Solution components
IBM Cloud Object Storage
Splunk SmartStore

Business challenge: As data analysis and AI was fast becoming a growing business requirement for this large enterprise the IT organization was being stretched with storing more data but without corresponding increase in budgets.

Business impact: After migrating Splunk data to the new Splunk SmartStore and IBM COS the customer was able to increase data and storage capacity at a much lower cost than previous primary storage capacity and maintain current performance levels.

IBM customer leverages IBM COS for Splunk solution

CUSTOMER Quotes:

> **"If older data is needed, Splunk will go to COS and pull back any data and the reads are superfast – did not notice any performance hit at all"**
>
> **"Less than half the cost for twice the amount of Splunk Index servers"**

For more information reference http://ibm.com/marketplace/cloud-object-storage-system

Notices

This information was developed for products and services offered in the US. This material might be available from IBM® in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

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.

For license inquiries regarding double-byte character set (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing to:

Intellectual Property Licensing
Legal and Intellectual Property Law IBM
Japan, Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

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 jurisdictions 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 provide in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Director of Licensing IBM
Corporation
North Castle Drive, MD-NC119 Armonk, NY
10504-1785
US

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided   by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

The performance data discussed herein is presented as derived under specific operating conditions. Actual results may vary.

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.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

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. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and ibm.com® are trademarks or registered 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 www.ibm.com/legal/copytrade.shtml.

Accesser®, Cleversafe®, ClevOS™, Dispersed Storage®, dsNet®, IBM Cloud Object Storage Accesser®, IBM Cloud Object Storage Dedicated™, IBM Cloud Object Storage Insight™, IBM Cloud Object Storage Manager™, IBM Cloud Object Storage Slicestor®, IBM Cloud Object Storage Standard™, IBM Cloud Object Storage System™, IBM Cloud Object Storage Vault™, SecureSlice™, and Slicestor® are trademarks or registered trademarks of Cleversafe, an IBM Company and/or International Business Machines Corp.

Other product and service names might be trademarks of IBM or other company