

# EFFICIENTLY STORE, ORGANIZE AND ACCESS UNSTRUCTURED DATA

Unstructured data is on a soaring growth curve and companies are struggling to keep up. Even with the cost per gigabyte of storage coming down, the speed of data growth is sending the costs of traditional storage through the datacenter roof. But the cost of storing all this data isn't the only issue. High availability, fast performance, and ease of scale can be a challenge, especially across multiple datacenters.

Public cloud storage provides an answer to these challenges. But the ingress and egress costs and the issues of availability are eating into the benefits. Maybe you need your own private cloud storage or just more reliable object storage behind your application.

Riak® S2 (formerly Riak CS) provides simple, available, distributed object storage for public, private, or hybrid clouds. Riak S2 is more cost effective than traditional storage with petabyte-scale and is compatible with Amazon S3 and OpenStack Swift for easy integration into existing production workloads. Riak S2 is software-defined object storage that helps you efficiently store, organize, and access your unstructured data, and it is architected for high availability, fault-tolerance, ease of operations, and massive scale.

“ Cloud-based object storage is ideal for storing our customer's growing business-critical data, and we have relied on the excellent performance, cost efficiency, and high reliability of Riak S2. ”

– Tsutomu Taguchi, Senior Managing Director,  
business group of I-NET Corp

## RIAK S2 BENEFITS

### FLEXIBLE OBJECT STORAGE

Easily store and manage small files or large files up to terabytes in size. Multi-part upload handles even your largest files.

### MASSIVE SCALABILITY

Automatic distribution of data and the ease of adding nodes mean near-linear performance increase as your object store grows.

### RESILIENCY

A masterless, multi-node architecture is fault tolerant to ensure no data loss in the event of network or hardware failures.

### GLOBAL AVAILABILITY

A distributed architecture with advanced local and multi-datacenter replication means your data is always available.

### GET TO THE CLOUD

For business continuity or to meet your peak demand, Riak S2 provides object storage in private, public, and hybrid clouds..

### USAGE AND ACCESS TRACKING

Multi-tenancy with per-user storage usage and access tracking makes it easy to support billing or chargebacks.

### DEVELOPER EMPOWERMENT

Extensive documentation and powerful APIs are easy to use. S3/Swift API plus Riak KV make Riak S2 a powerful object store for your application.

### OPERATIONAL SIMPLICITY

Easy-to-run and easy-to-scale up and down. Operations are powerful yet simple.

## CHALLENGES OF UNSTRUCTURED DATA

Traditional storage architectures have been around for decades. They were built to scale in the datacenter and require bigger, faster servers to keep up. Riak S2 has a distributed architecture designed to run and scale on commodity hardware. Riak S2 is uniquely built to handle a variety of unstructured and Big Data needs. It is optimized for large document stores, storing rich media files, providing backup and archive storage, plus much more. Whether you need object storage for your applications or are building a service offering, Riak S2 provides an easy-to-scale, easy-to-operate solution.

### DOCUMENT STORE

Document stores are used to store a variety of content from emails and attachments to financial documents to application log files. These objects are stored in their original formats: XML, PDF, proprietary/binary word processor formats, or even plain text. Riak S2 is designed for these specialized document stores. It is architected for peak performance and to scale easily. You can store content that can be consumed by a variety of web sites or locations; or you can store and analyze thousands of log files. Riak S2 powers the document store for your business applications.

### RICH MEDIA STORE

Videos, images, and other rich media are creating petabytes of unstructured data. Movies, on-demand broadcasts, games, marketplace digital images, MRIs/CAT scans, and music are just a few examples of the rich media content that is flooding storage. Riak S2 is object storage built to scale and meet the demand of these very large, fast growing rich media stores. It is designed to scale horizontally with commodity hardware and is specifically designed with large object support and multipart upload. Also, if you provide storage as a service offering, Riak S2 provides multi-tenancy and per-user reporting. Whether you need to keep this data in house or share it across the globe, Riak S2 provides the object storage required to meet your business needs.

### BACK UP AND ARCHIVE

RTO and RPO are as ancient as PDAs when it comes to your critical business data. Your critical data has to be fully recoverable all the time, immediately. If you are under compliance rules, your archive needs to be immutable and easily accessible by auditors. By providing highly available, fault-tolerant storage for your real-time backups and archives, Riak S2 meets your business needs. Easily store VMs, relational database backups, and media file copies all in Riak S2. Riak S2 also functions as a backup solution for any all-flash array provider offering Amazon S3 compatible backup targets.

“ Object storage is radically simpler than traditional SAN and even NAS, it scales much better from a capacity standpoint and it's especially well-suited for cost-effectively storing the reams of unstructured data — think files, videos, music and images — that are being created in this 'big data' era. ”

— Simon Robinson, 451 Research

## RIAK S2 FOR YOUR OBJECT STORAGE

Riak S2 is easy-to-use object storage software built on top of Riak, Basho's distributed database. Riak S2 is designed to provide simple, available, distributed cloud storage at any scale, and can be used to build cloud architectures — be they public, private or hybrid — or as storage infrastructure for heavy-duty applications and services. Riak S2's API is Amazon S3 and SWIFT compatible and supports per-tenant reporting for use cases involving billing and metering.

### WHY USE RIAK S2 ENTERPRISE?

#### AVAILABILITY

Riak S2 replicates data, defaulting to three replicas per object, to ensure data is always available for your application.

#### FAULT TOLERANT STORAGE

Riak S2 is fault tolerant with no master node. Data remains available even in the event of node failure or network partition.

#### LARGE OBJECT STORAGE

Riak S2 can store any conceivable data type, such as images, text, video, documents, VMs, database backups, or software binaries. Riak S2 can store objects into the terabyte size range.

#### SCALABILITY

Riak S2 scales easily. Just add nodes to add capacity. Riak S2 automatically redistributes data in the system, preventing storage hot spots and decreasing the operational burden of adding new nodes.

#### CLOUD SERVICES

Riak S2 is designed as a highly available, fault-tolerant, low-latency foundation for your public, private or hybrid cloud services.

#### S3 COMPATIBLE

Riak S2 has a built-in Amazon S3 interface with S3 ACL support. You can import and extract data currently residing in Amazon S3 and use your existing S3 tools and frameworks to manage your data.

#### MULTI-DATACENTER REPLICATION

Riak S2 provides multi-datacenter replication for active backups, availability zones, data locality and more. All with low-latency storage wherever your users are and maintains availability even in the event of site failure.

#### MULTI-TENANCY

Riak S2 can be deployed for multi-tenancy and with the Riak S2 reporting API, you can access per-tenant usage data and statistics. Use this for accounting, subscription, chargebacks, plugins with billing systems, efficient multi-department utilization, and much more.

#### OPERATIONAL SIMPLICITY

Interfaces for user creation, disablement and credential management; plus querying stats on access and storage.

#### Masterless



Multiple nodes operate to form a single cluster, making Riak S2 masterless

#### Scalable



Scale out, up, and down predictably and linearly by adding nodes to the cluster

#### Global



Replicate data across multiple datacenters, globally, leveraging Riak S2 Enterprise

## RIAK S2 CUSTOMERS

Businesses building innovative ways to collect and use unstructured data that need massive scalability, high available, fault tolerance, and a powerful yet operationally simple object storage platform will benefit from Riak S2.

Some of the most data intensive applications and businesses are using Riak S2:



View more of our users at [basho.com/riak-users/](http://basho.com/riak-users/)

“ We offer our customers affordable and scalable cloud storage solutions built on Riak S2. The enhanced Amazon S3 compatibility and ability to scale well into the multi-petabyte level helps us better support the rapid growth we are seeing in our storage business. ”

– Makoto Oya, Vice Director of IDC Frontier, Yahoo



## ABOUT BASHO TECHNOLOGIES

Basho is a distributed systems company dedicated to developing disruptive technology that simplify enterprises' most critical data management challenges. Basho has attracted one of the most talented groups of engineers and technical experts ever assembled devoted exclusively to solving some of the most complex issues presented by scaling distributed systems.

Basho's distributed database, Riak® KV, the industry leading distributed NoSQL database, and Basho's cloud storage software, Riak® S2, are used by fast growing Web businesses and by one third of the Fortune 50 to power their critical Web, mobile and social applications. The Basho Data Platform helps enterprises reduce the complexity of supporting Big Data applications by integrating Riak KV and Riak S2 with Apache Spark, Redis, and Apache Solr. Basho is the organizer of RICON — a distributed systems conference. Riak is the registered trademark of Basho Technologies, inc.



BASHO TECHNOLOGIES, INC  
10900 NE 8TH STREET  
SEATTLE, WA 98004

617.714.1700 // [WWW.BASHO.COM](http://WWW.BASHO.COM)