



Международный журнал информационных технологий и энергоэффективности

Сайт журнала:

<http://www.openaccessscience.ru/index.php/ijcse/>



УДК 519.767

ХРАНИЛИЩЕ ОБЪЕКТОВ В ОБЛАКЕ: СРАВНЕНИЕ 4 ПОСТАВЩИКОВ

¹Глазунова Е.З., ²Шевырева А.Д., ³Никулина С.М.

Поволжский государственный университет телекоммуникаций и информатики (ПГУТИ), Самара, Россия (443090, Самара, Московское шоссе, 77); e-mail: ¹ glazunovaelena@ssau.ru, ² shevyreva02@mail.ru, ³ nikulinasofiaaaa@mail.ru

Работа посвящена тому, что по мере роста вашего бизнеса вам необходимо управлять изолированными, но быстро растущими пулами данных из различных источников, используемых для различных бизнес-процессов и приложений. В настоящее время многие организации имеют фрагментированные портфели хранилищ, что замедляет внедрение инноваций и усложняет программы организации. Хранение объектов поможет вашей организации избавиться от этих изолированных элементов. Он обеспечивает эффективное и крупномасштабное хранилище, в котором можно хранить данные любого типа в исходном формате.

Объектное хранилище очень подходит для облака, поскольку оно гибкое, надежное и может быть легко масштабировано до нескольких петабайт для поддержки неограниченного роста объема данных. Архитектура управляет данными и хранит их в объектной форме, в отличие от блочного хранилища, которое обрабатывает данные как логические тома, блоки и файлы хранилища, где данные хранятся в иерархических файлах.

Ключевые слова: облачные хранилища данных, информационные технологии, IT-технологии, безопасность данных, IT-провайдеры, пользователи облачных хранилищ, облачные вычисления, облачные хранилища, Amazon Web Services, Microsoft Azure, Google Cloud, IBM Cloud.

OBJECT STORAGE IN THE CLOUD: COMPARISON OF 4 PROVIDERS

¹Glazunova E.Z., ²Shevyreva A.D., ³Nikulina S.M.

Volga State University of Telecommunications and Informatics (PGUTI), Samara, Russia (443090, Samara, Moskovskoe Shosse, 77); e-mail: ¹ glazunovaelena@ssau.ru, ² shevyreva02@mail.ru, ³ nikulinasofiaaaa@mail.ru

The paper is devoted to the fact that as your business grows, you need to manage isolated but rapidly growing pools of data from various sources used for various business processes and applications. Currently, many organizations have fragmented storage portals, which slows down innovation and complicates the organization's programs. Storing objects will help your organization get rid of these isolated elements. It provides an efficient and large-scale storage in which any type of data can be stored in its original form.

Object storage is very suitable for the cloud because it is flexible, reliable and can be easily scaled to several petabytes to support unlimited data growth. The architecture manages data and stores it in object form, unlike block storage, which processes data as logical volumes, blocks and storage files, where data is stored in hierarchical files.

Keywords: cloud data warehouses, information technologies, IT technologies, data security, IT providers, cloud storage users, cloud computing, cloud storage, Amazon Web Services, Microsoft Azure, Google Cloud, IBM Cloud

Облачное хранилище позволяет сохранять данные и файлы за пределами сайта, к которым вы получаете доступ либо через общедоступный Интернет, либо через выделенное подключение к частной сети. Данные, которые вы передаете на хранение за пределы сайта, становятся ответственностью стороннего облачного провайдера. Поставщик размещает, защищает, управляет и обслуживает серверы и связанную с ними инфраструктуру и гарантирует, что у вас есть доступ к данным, когда вам это нужно.

Облачное хранилище обеспечивает экономичную и масштабируемую альтернативу хранению файлов на локальных жестких дисках или в сетях хранения данных. Жесткие диски компьютеров могут хранить только ограниченный объем данных. Когда у пользователей заканчивается хранилище, им необходимо перенести файлы на внешнее запоминающее устройство. Традиционно организации создавали и обслуживали сети хранения данных (SANS) для архивирования данных и файлов. Однако обслуживание сетей SAN обходится дорого, поскольку по мере роста объема хранимых данных компаниям приходится инвестировать в добавление серверов и инфраструктуры для удовлетворения возросшего спроса.

Облачные сервисы хранения данных обеспечивают эластичность, что означает, что вы можете масштабировать емкость по мере увеличения объемов ваших данных или при необходимости уменьшать ее. Храня данные в облаке, ваша организация экономит, оплачивая технологию хранения данных и емкость как услугу, вместо того, чтобы инвестировать в капитальные затраты на создание и обслуживание собственных сетей хранения данных. Вы платите только за ту мощность, которую используете. Хотя ваши затраты со временем могут возрасти из-за увеличения объемов данных, вам не нужно чрезмерно расширять сети хранения в ожидании увеличения объема данных.

Как и локальные системы хранения, облачное хранилище использует серверы для хранения данных; однако данные отправляются на серверы, расположенные за пределами площадки. Большинство используемых вами серверов - это виртуальные машины, расположенные на физическом сервере. По мере роста потребностей в хранилищах поставщик создает новые виртуальные серверы для удовлетворения спроса.

Обычно вы подключаетесь к облачному хранилищу через Интернет или через специальное личное соединение через веб-портал, веб-сайт или мобильное приложение. Сервер, к которому вы подключаетесь, отправляет ваши данные на группу серверов, расположенных в одном или нескольких центрах обработки данных, в зависимости от объема функциональных возможностей облачного провайдера.

В рамках услуги поставщики часто хранят одни и те же данные на нескольких компьютерах для обеспечения избыточности. Таким образом, если сервер перейдет на техническое обслуживание или произойдет сбой, вы все равно сможете получить доступ к своим данным.

Облачное хранилище доступно в частных, общедоступных и гибридных облаках.

В данной статье рассмотрены программы по хранению объектов от некоторых ведущих мировых облачных провайдеров, а именно такие как: Amazon Web Services, Microsoft Azure, Google Cloud и IBM Cloud.

1. Хранилище объектов AWS.

AWS Storage предоставляет вам услуги, необходимые для создания решения для хранения данных, подходящего для вашего бизнеса. В этой статье предоставлена информация о сервисах хранения AWS, которые предоставляют клиентам хранилище в разных местах во время их облачных путешествий, от миграции до гибридного и встроенного в облако. В статье изложена информация об услугах, которые предоставляют сопутствующие функции, такие как передача данных и защита данных.

Основные сервисы хранения данных AWS

- Локальное хранилище по сравнению с облачным хранилищем AWS,
- Знакомство с портфолио хранилищ AWS,
- Определение правильного решения для хранения данных в облаке,
- Блочное хранилище: Amazon EBS,
- Обзор Хранилища файлов,
- Хранилище файлов: Amazon EFS,
- Хранилище файлов: Файловый сервер Amazon FSx для Windows,
- Хранилище файлов: Amazon FSx для Lustre,
- Хранилище объектов: Amazon S3.

Пограничные и гибридные решения для хранения данных

- Обзор пограничных и гибридных облачных решений,
- Решения для хранения данных Edge: семейство AWS Snow,
- Гибридные решения для хранения данных: аванпосты AWS,
- Гибридные решения для хранения данных: AWS Storage Gateway

Перенос ваших данных в облако AWS

- Внедрение служб передачи данных и миграции,
- Семейство AWS Transfer,
- AWS Синхронизация данных,
- Семейство AWS Snow,
- Служба миграции приложений AWS (CloudEndure Migration).

Защита ваших данных в облаке AWS

- Защита данных в облаке Введение,
- Резервное копирование AWS,
- Моментальные снимки собственной службы,
- Аварийное восстановление в облаке

Объединение всех сервисов хранения данных AWS воедино

- Краткое описание портфолио хранилищ AWS
- Принятие решения о хранении в AWS

AWS предоставляет широкий спектр уровней хранения для различных вариантов использования. Amazon S3 является основной платформой для размещения объектов AWS: S3 Standard предлагает холодное хранилище, а Glacier - холодное хранилище:

Стандарт Amazon S3 является общепринятым хранилищем информации и идеально подходит для многих случаев использования, включая динамические веб-сайты, облачные приложения, доставку контента, анализ данных и игры. Он обеспечивает высокую производительность, а также низкую задержку.

Стандарт Amazon S3 -ограниченный доступ (Amazon S3 Standard is) - это альтернатива более ограниченному хранилищу данных для доступа, такого как аварийное восстановление и долгосрочное резервное копирование.

Amazon Glacier: эта высоконадежная система хранения, оптимизированная для часто недоступных или "холодных" данных, таких как данные с истекшим сроком службы, хранящиеся для обеспечения совместимости и резервного копирования. Данные хранятся для долгосрочного хранения в нетронутом виде и зашифрованы. [2]

2. Хранилище объектов Azure

Microsoft предлагает большие двоичные файлы хранилища Azure для хранения объектов в облаке. Хранилище больших двоичных объектов подходит для хранения неструктурированных данных любого формата, таких как двоичные файлы или текст. Сюда входят видео, изображения, документы, аудио и т. Д. Хранилище Azure обеспечивает высокую степень целостности, гибкости и изменчивости данных.

Большое двоичное хранилище используется для отправки документов или изображений непосредственно в браузер, хранения файлов для распределенного доступа, потоковой передачи аудио и видео, хранения файлов, восстановления после сбоя, хранения данных для восстановления и резервного копирования, а также архивирования, чтобы их можно было анализировать с помощью локальной службы или службы, размещенной в Azure.

Azure имеет несколько уровней хранения, в том числе:

- Уровень *горячего доступа*— для информации, которая активно используется или предполагается к активному использованию и подготовлена для обработки и последующего переноса на уровень холодного хранения.
- Уровень *прохладного доступа*— для данных, которые должны оставаться на уровне прохладного доступа более 30 дней. Сюда входят наборы данных для аварийного восстановления и кратковременного резервного копирования, устаревший мультимедийный контент, который должен быть немедленно доступен при использовании, и большие наборы данных.
- Уровень доступа к архиву— для данных, которые будут храниться на уровне архива более 180 дней и могут выдерживать многочасовые задержки при извлечении.

Хранилище больших двоичных объектов поддерживает Azure Data Lake Storage Gen2, корпоративное решение Microsoft для анализа больших данных в облаке. Azure Data Lake Storage Gen2 предлагает иерархическую файловую систему, а также преимущества хранилища больших двоичных объектов, в том числе:

- Недорогое многоуровневое хранилище
- Высокая доступность
- Сильная консистенция
- Возможности аварийного восстановления

Хранилище больших двоичных объектов предлагает три типа ресурсов:

- Учетная запись хранилища
- Контейнер в учетной записи хранения
- Капля в контейнере
- На следующей диаграмме показана взаимосвязь между этими ресурсами.

Диаграмма (Рисунок 1), показывающая взаимосвязь между учетной записью хранилища, контейнерами и большими двоичными объектами.



Рисунок 1 – Взаимосвязь между объектами данных

Учетная запись хранения предоставляет уникальное пространство имен в Azure для ваших данных. Каждый объект, который вы храните в хранилище Azure, имеет адрес, который включает в себя ваше уникальное имя учетной записи. Комбинация имени учетной записи и конечной точки хранилища больших двоичных объектов формирует базовый адрес для объектов в вашей учетной записи хранения.

Примечание: уровень архивированного хранилища недоступен на уровне учетной записи хранилища, а только на уровне больших двоичных файлов. Azure также предлагает матрас премиум-класса, предназначенный для работы, требующей быстрого и последовательного реагирования. [1]

3. Облачное хранилище Google

Вы можете использовать облачное хранилище Google для хранения данных в облаке Google. Облачное хранилище обычно используется для хранения неструктурированных данных. Вы можете добавлять объекты любого вида и размера, объемом до 5 ТБ.

Google (GCS). В нем есть четыре категории для резервного копирования и архивирования, а также для хранения высококачественных элементов. Все пять классов обеспечивают высокую выносимость и низкую задержку:

- *Горячее (высокопроизводительное) хранилище*— GCS обеспечивает региональное и межрегиональное хранилище для информации с высокочастотным доступом.
- *Мультирегиональное хранилище*- позволяет хранить информацию, к которой часто обращаются по всему миру, включая потоковое видео, контент веб-сайта или мобильные и игровые приложения.
- *Региональное хранилище*— обеспечивает частый доступ к информации в соответствующем регионе экземпляра Google Compute Engine или Google Cloud DataProc, например data analytics.
- *Хранилище Nearline (cool)*— для данных, доступ к которым требуется не реже одного раза в месяц, а несколько раз в год. Подходит для резервного копирования и длинного мультимедийного контента.
- *Холодное хранилище (cool)*— для данных, доступ к которым требуется реже одного раза в год. Подходит для архивирования данных и аварийного восстановления.

4. Облачное хранилище объектов IBM Cloud

Облачная система хранения объектов развернута в нескольких конфигурациях. Каждый узел состоит из облачных серверов.

Программное обеспечение для хранения объектов, работающее на сервере промышленного стандарта.

Программное обеспечение для хранения облачных объектов совместимо с широким спектром серверов из многих источников, включая физическое или виртуальное устройство. Кроме того, IBM проводит сертификацию конкретных серверов, которые клиенты хотят использовать в своей среде, чтобы обеспечить быструю первоначальную установку, долгосрочную надежность и предсказуемую производительность.

IBM Cloud предоставляет масштабируемое и гибкое облачное хранилище с возможностью архивирования неструктурированных данных на основе политик. Эта служба облачного хранилища предназначена для архивирования данных, например, для долгосрочного хранения данных, которое редко доступно, включая мобильные и веб-приложения, а также для резервного копирования и аналитики.

IBM имеет четыре уровня хранилища данных, встроенных в возможности высокоскоростной передачи данных Aspera. Это упрощает передачу данных и хранилища в облако, а также выполнение запросов на месте.

Уровни классов хранения объектов IBM Cloud Object Storage:

- *Стандартное хранилище*— для активных рабочих нагрузок, требующих высокой производительности и низкой задержки, а также для данных, требующих частого и многократного доступа в течение месяца. Сценарии использования - это, например, активные хранилища контента, аналитика, потоковая передача мобильного и веб-контента, совместная работа и DevOps.
- *Хранилище Vault* — для менее активных рабочих нагрузок, которым требуется доступ в режиме реального времени по требованию, но нечасто, не чаще одного раза в месяц. Варианты использования включают хранение цифровых активов и резервное копирование.
- *Холодное хранилище*— для холодных рабочих нагрузок, где данные нуждаются в доступе по требованию в режиме реального времени, когда это необходимо, но в основном архивируются. Например, данные, к которым обращаются несколько раз в год. Общие варианты использования включают долгосрочное резервное копирование, сохранение больших наборов данных, таких как устаревший медиаконтент и научные данные.
- *Гибкое хранилище*— этот уровень класса используется для динамических рабочих нагрузок (комбинируя холодные и горячие рабочие нагрузки) и данных на основе шаблонов доступа. Типичные варианты использования включают когнитивные рабочие нагрузки, облачную аналитику и пользовательские приложения. [3]

Плюсы и минусы облачного хранилища объектов

Ниже приведены некоторые из ключевых преимуществ и недостатков хранения объектов в облаке.

К ключевым преимуществам объектного хранилища относятся:

- Данные хорошо распределены, что обеспечивает их более высокую устойчивость к аппаратным сбоям или катастрофам. Таким образом, он доступен, даже если различные узлы выходят из строя.
- Объекты хранятся в едином адресном пространстве, что сводит к минимуму проблемы сложности и масштабируемости.
- Защита данных встроена в эту архитектуру в виде технологии кодирования стирания или репликации.
- Объектное хранилище наиболее подходит для облачного хранения и статических данных. Распространенные варианты использования объектного хранилища включают архивирование и резервное копирование в облаке — эта технология лучше всего работает с данными, которые чаще считываются, чем записываются.
- Хранилище объектов развилось до такой степени, что оно масштабируется на уровне эксабайт и представляет триллионы объектов. Использование виртуальных машин или стандартного оборудования позволяет легко добавлять узлы, а дисковое пространство используется более эффективно.
- Системы хранения объектов с помощью идентификаторов объектов (OID) или идентификаторов могут получить доступ к любой части данных, не зная, на каком физическом устройстве хранения, каталоге или файловой системе они находятся. Абстракция позволяет устройствам хранения объектов работать с аппаратными средствами хранения, сконфигурированными в архитектуре распределенных узлов. [4]

К основным недостаткам объектного хранилища относятся:

- Системы хранения объектов недостаточно устойчивы для систем реального времени, включая транзакционные базы данных. Нежелательным вариантом использования для хранения объектов является среда или приложение с высокой скоростью транзакций.
- Хранение объектов не гарантирует, что запросы на чтение будут выдавать самую последнюю версию данных.
- Эта технология не всегда подходит для приложений с высокими требованиями к производительности.
- Облачное хранилище часто оказывается более дорогим, потому что вам нужно платить за хранилище на постоянной основе. При использовании оборудования на месте вы платите один раз, и хранилище принадлежит вам. [5]

Перенос хранилища объектов в локальное хранилище с помощью Cloudian.

Cloudian HyperStore - это устройство хранения объектов большой емкости, полностью совместимое с Amazon S3. Это позволяет легко настроить решение для хранения объектов в локальном центре обработки данных, используя преимущества облачного хранилища объектов при гораздо меньших затратах.

Hyperstone может хранить до 1,5 петабайт на устройстве 4U, что позволяет хранить до 18 петабайт на одной полке центра обработки данных. Hyperstore поставляется с полностью резервируемым источником питания и охлаждением, а также функциями повышения производительности, такими как твердотельный накопитель емкостью 1,92 ТБ для метаданных и порты Ethernet емкостью 10 Гб для быстрой передачи данных.

Вывод: в заключении можно сказать, что вычислительная мощность может увеличить емкость хранилища данных.

Требования ввода-вывода не должны проходить через центральный контроллер, что позволяет действительно глобальной системе хранить большое количество объектных данных, физически хранящихся повсюду и принимаемых через Интернет или огромную сеть. [6]

Список литературы

1. Облачное хранилище данных. URL: <http://b23.ru/y2o3> (дата обращения: 05.02.2023).
2. Облачные вычисления. URL: <http://b23.ru/y2ok> (дата обращения: 09.02.2023).
3. SkyDrive. URL: <http://b23.ru/y2op> (дата обращения: 09.02.2023).
4. Гордеев И. Обзор бесплатных облачных хранилищ данных. URL: <http://b23.ru/y2oy> (дата обращения: 08.02.2023).
5. Солодовникова А. «Облачный бум», или какое облачное хранилище выбрать. URL: <http://b23.ru/y2oh> (дата обращения: 09.02.2023)
6. Грищачук С. Облачные хранилища данных. URL: <http://b23.ru/y2oa> (дата обращения: 07.02.2023)

References

1. Cloud data storage. URL: <http://b23.ru/y2o3> (date of access: 05.02.2023).
 2. Cloud computing. URL: <http://b23.ru/y2ok> (date of access: 09.02.2023).
 3. SkyDrive. URL: <http://b23.ru/y2op> (date of access: 09.02.2023).
 4. Gordeev I. Review of free cloud data storage. URL: <http://b23.ru/y2oy> (date of access: 08.02.2023).
 5. Solodovnikova A. «Cloud boom», or which cloud storage to choose. URL: <http://b23.ru/y2oh> (date of access: 09.02.2023)
 6. Gritsachuk S. Cloud data warehouses. URL: <http://b23.ru/y2oa> (date of access: 07.02.2023)
-