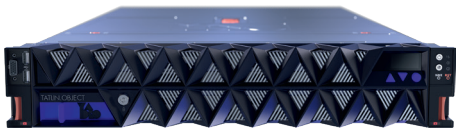
TATLIN.OBJECT



Распределенное децентрализованное объектное хранилище данных с поддержкой современных протоколов доступа, включая S3

Подходит для гарантированного хранения и работы с многопетабайтным объемом данных.

Ключевые характеристики

- Горизонтальное масштабирование до 100 узлов хранения
- Поддержка объектного доступа к данным
- Наличие SSD-кэша для ускорения работы с данными
- Поддержка протоколов S3, HTTP(S), gRPC
- Гибкие политики хранения данных
- Современный и простой HTML5 интерфейс управления
- Отказоустойчивость на уровне компонент и узлов хранения
- Возможность задать фактор репликации данных индивидуально для каждого контейнера
- Размещение объектов по узлам хранения с использованием схем избыточного кодирования
- Централизованное управление всей системой независимо от количества узлов хранения и их местоположения
- Синхронно-асинхронная репликация данных
- Поддержка мультитенантности
- Автоматизация управления через REST API
- Поддержка all-SSD узлов для повышения производительности

Архитектура
TATLIN.OBJECT — децентрализованная сеть хранения данных. Каждый узел наделен максимальной автономностью и делает все возможное, чтобы данные хранились корректно и в соответствии с заданной политикой. Такой подход позволяет масштабировать объем и производительность всей системы практически линейно, простым добавлением новых узлов хранения.

Каждый узел хранения использует все доступное локальное дисковое пространство для данных и их индексации. Избыточность и защита обеспечивается на уровне всей сети. Система TATLIN.OBJECT продолжит работать и сохранит целостность и доступ к данным после выхода из строя накопителей в соответствии с политиками хранения. Данные при этом будут эвакуированы на другие диски или узлы.

Часть узлов помимо хранения данных занимается мониторингом сети хранения и поддержанием актуального списка доступных узлов. Эта информация хранится в реплицируемой на все узлы специализированной базе данных. Таким образом, в системе нет единой точки отказа или центрального источника информации, ограничивающего производительность системы. Это позволяет размещать узлы в разных ЦОД в удаленных регионах без дополнительных операционных издержек.

Аппаратная платформа
В качестве аппаратной платформы используются вычислительные узлы на базе высокопроизводительных процессоров Intel Xeon. Аппаратные узлы способны масштабировать подсистему хранения для решения широкого круга задач. Каждый узел имеет 4 встроенных интерфейса 10/25 GbE для интеграции в современную сетевую инфраструктуру. Также предусмотрено расширение конфигурации дополнительным SSD-кэшем.

КОМПОНЕНТЫ	
Узлы хранения	От 4 до 100
Шаг расширения системы	1 узел хранения
Минимальное/максимальное количество накопителей в системе	• 24 / 1 200 × NL-SAS 16 ТБ • 24 / 1200 × SAS SSD 7,68 ТБ
Максимальная неразмеченная емкость на систему	• 19 200 ТБ (гибридные узлы) • 9 216 ТБ (all-SSD узлы)

ХАРАКТЕРИСТИКИ УЗЛА ХРАНЕНИЯ	
Форм-фактор	2U
Процессоры	2
Память	256 ГБ RAM
Минимальное / максимальное количество накопителей под хранение данных	• 6 / 12 × NL-SAS 16 ТБ • 6 / 12 × SAS SSD 7,68 ТБ
Минимальное / максимальное количество накопителей под кэш (для гибридных узлов)	0 / 4 SATA SSD 1,92 ТБ
Порты ввода-вывода	• 2 × 10/25 Гбит/с Ethernet для внутренней сети • 2 × 10/25 Гбит/с Ethernet для доступа к данным • 1 × 1 Гбит/с Ethernet для сети управления • 1 × 1 Гбит/с Ethernet для локального сервисного доступа

УПРАВЛЕНИЕ И ДОСТУП	
Поддерживаемые протоколы доступа	• Поддержка S3, включая авторизацию • Полная поддержка HTTP(S), включая загрузку и работу с диапазонами байтов для проигрывания видео • Нативная поддержка gRPC
Управление системой	• Графический интерфейс управления (Web UI) • Поддержка интеграции с LDAP • Control API
Квотирование	• Дисковое пространство • Количество бакетов • Количество объектов

БЕЗОПАСНОСТЬ	
Разграничение прав доступа к данным	Настройка политик доступа IAM-пользователей и групп к ресурсам СХД
Политики хранения данных	Гибкие политики хранения на каждый контейнер с возможностью описания требований законов о персональных данных или корпоративных правил хранения информации
Парольные политики	Назначение и изменение параметров сложности паролей, срока их действия, блокировки учетных записей и др.
Защищенный доступ	Поддержка установки и управления TLS-сертификатами (для сети передачи данных, сети управления и внутренней сети)
Аудит действий администраторов	Запись действий администраторов в журнал аудита и возможность его экспортирования в SIEM-системы
Мультитенантность	Разделение СХД на изолированные наборы ресурсов с индивидуальным управлением

ОБЩИЕ ЛИМИТЫ[1]	
Количество объектов в системе	10 млрд, но при количестве >1 млрд наблюдается падение производительности
Количество объектов в контейнере	1 млрд, но при количестве >100 млн наблюдается падение производительности
Количество контейнеров в системе	Не ограничено
Количество реплик	Не превышающее количество узлов
Значения K и M в схеме Erasure Coding вида K.M	K ≤ 12; M ≤ 4 (поддерживаются не все возможные комбинации)
Количество локальных пользователей для Control Path	10 000
Количество LDAP-пользователей для Control Path после применения фильтров	1 000

ЛИМИТЫ АППАРАТНОЙ ПЛАТФОРМЫ[1]	
Количество узлов в системе	От 4 до 100 шт.
Количество узлов на площадку	От 4 до 100 шт.

[1] Значения лимитов могут быть изменены по мере развития продукта.

ЛИМИТЫ ПРОТОКОЛОВ (GRPC / S3)[1]	
Количество частей, загружаемых через MPU	2 млн (требует реализации на клиенте через SDK) / 10 тыс.
Размер одной MPU-части	Определяется реализацией на клиенте через SDK / От 5 МБ до 5 ГБ, последняя часть может быть любого размера
Максимальный размер загружаемого одиночного объекта	50 ГБ / 5 ГБ
Максимальный размер загружаемого через MPU объекта	128 ТБ / 50 ТБ
Количество версий объекта (для S3)	1 тыс.
Количество версионируемых бакетов (для S3)	Не ограничено
Размер пользовательских метаданных на 1 объект (для gRPC)	192 МБ
Размер пользовательских тэгов на 1 объект (для S3)	128 UNICODE-символов для ключа тэга и 256 UNICODE-символов для значения тэга
Количество пользовательских тэгов на 1 объект (для S3)	10 шт.

ИНТЕГРАЦИИ	
Мониторинг	Шаблоны для Zabbix 5.x, 6.x; интеграция с корпоративными системами мониторинга (Prometheus)
Визуализация	Интеграция с системой визуализации данных Grafana (версии 9.1.5 и выше)
SDK	Open-source SDK для Go
Kubernetes	Поддержка OCI для хранения контейнеров Kubernetes
Резервное копирование	Интеграция с ПО резервного копирования «Кибер Бэкап 16»
Управление данными	Интеграция с решениями по хранению электронного контента и документов «Кибер Инфраструктура 5.01», «Закрома.Хранение 1.013» и «Закрома.Архив 1.1»
Большие данные	Интеграция с ПО вычисления и обработки данных Arenadata Hadoop 3.2.4

[1] Значения лимитов могут быть изменены по мере развития продукта.

НАДЕЖНОСТЬ

Архитектура	• Масштабируемость за счет однотипных узлов хранения • Отсутствие единой точки отказа, на каждом узле доступны все необходимые сервисы • Поддержка самовосстановления системы после сбоев
Механизмы защиты данных	• На уровне каждого контейнера возможно задать уровень репликации данных • Возможность размещения объектов по узлам с использованием технологии избыточного кодирования («erasure coding»)
Реакция системы на сбои	• Возможность работы в режиме деградации с потерей всех SSD, только с объектами на HDD • При авариях система будет самовосстанавливаться и приводить фактическое хранение объектов в соответствие заданной политике по мере появления физической возможности • Формат данных, пригодный для восстановления даже после тотальной аварии на системе
Сервисные операции	• Поддержка maintenance-режима для сервисного обслуживания • Набор стандартных операций для замены компонент системы (включая HDD, SSD и т.д.) • Механизм эвакуации данных с узла хранения

ПРОИЗВОДИТЕЛЬНОСТЬ

Базовые принципы	• Эффективный движок хранения с раздельной обработкой мелких объектов и прозрачной потоковой обработкой больших объектов • Накопители каждого узла используются раздельно, формируя шарды, на которых хранятся клиентские данные • Высокая производительность каждого узла сети хранения и параллелизм обработки запросов множеством узлов • Синхронно-асинхронная репликация, позволяющая получить высокую производительность системы и катастрофоустойчивость
Кэширование данных	Использование SSD+HDD-конфигурации с кэшированием чтения и записи

ЛИЦЕНЗИРОВАНИЕ

Политика лицензирования	Лицензия с привязкой к полезной емкости системы с гранулярностью по ТБ

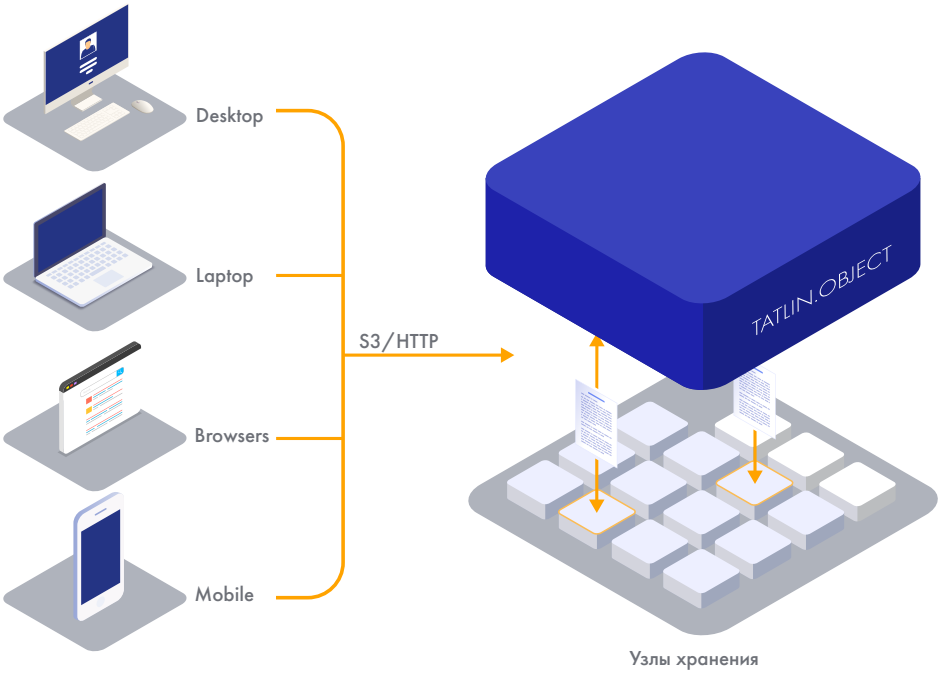
ГАРАНТИЯ И ПОДДЕРЖКА

Доступные варианты поддержки	24 × 7

ПРИМЕРЫ ПРИМЕНЕНИЯ

Работа в режиме современного S3-хранилища

Web и мобильные приложения могут напрямую загружать данные через протоколы S3 и HTTP(S) в TATLIN.OBJECT. Загруженные объекты автоматически могут быть распределены по регионам присутствия пользователей и раздаваться через кэширующие фронтенд-сервера, образуя, таким образом, CDN для проекта. При делегировании домена раздающих серверов на GeoDNS раздача будет производиться с ближайшего к потребителю фронтенд-сервера, который, в свою очередь, будет запрашивать данные с ближайшего узла TATLIN.OBJECT.

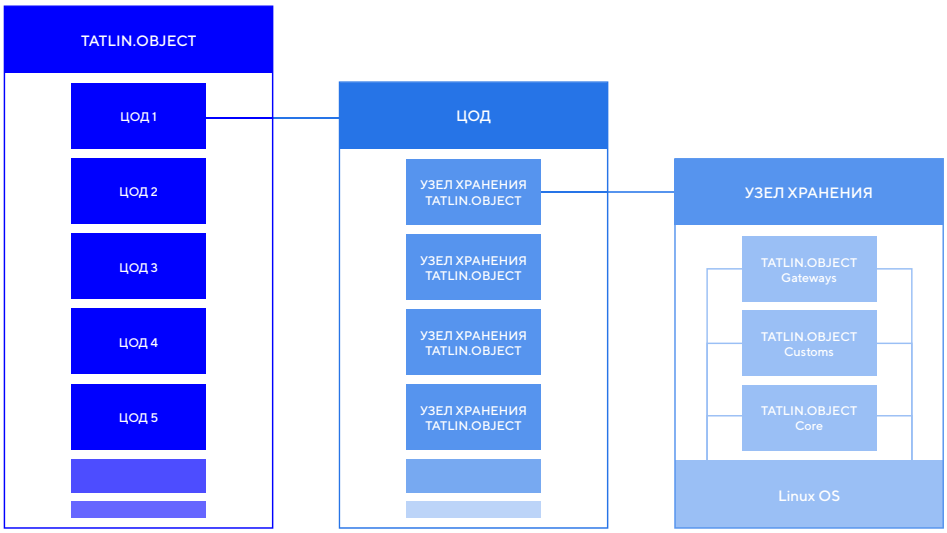


ПРИМЕРЫ ПРИМЕНЕНИЯ

Глобальная геораспределенная система хранения данных для размещения данных приложений, резервных копий и архивов

Компания или группа компаний могут использовать единую систему хранения, распределенную по нескольким площадкам. Данные пользователей и приложений располагаются на разных сайтах в соответствии с принятыми политиками хранения.

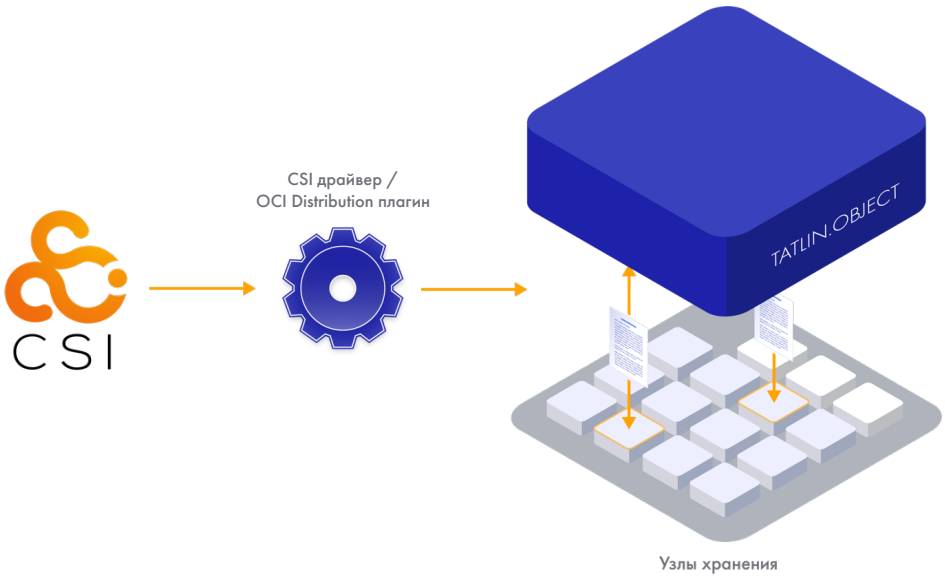
Альтернативно, одна группа узлов хранения может размещаться on-site, другие группы — в удаленных off-site-точках. Данные записываются на ближайший узел, а дальнейшая репликация происходит автоматически, в соответствии с политикой хранения, заданной для контейнера. Объект станет доступен сразу после попадания в систему, не дожидаясь репликации по всем площадкам.



ПРИМЕРЫ ПРИМЕНЕНИЯ

Хранилище образов Kubernetes

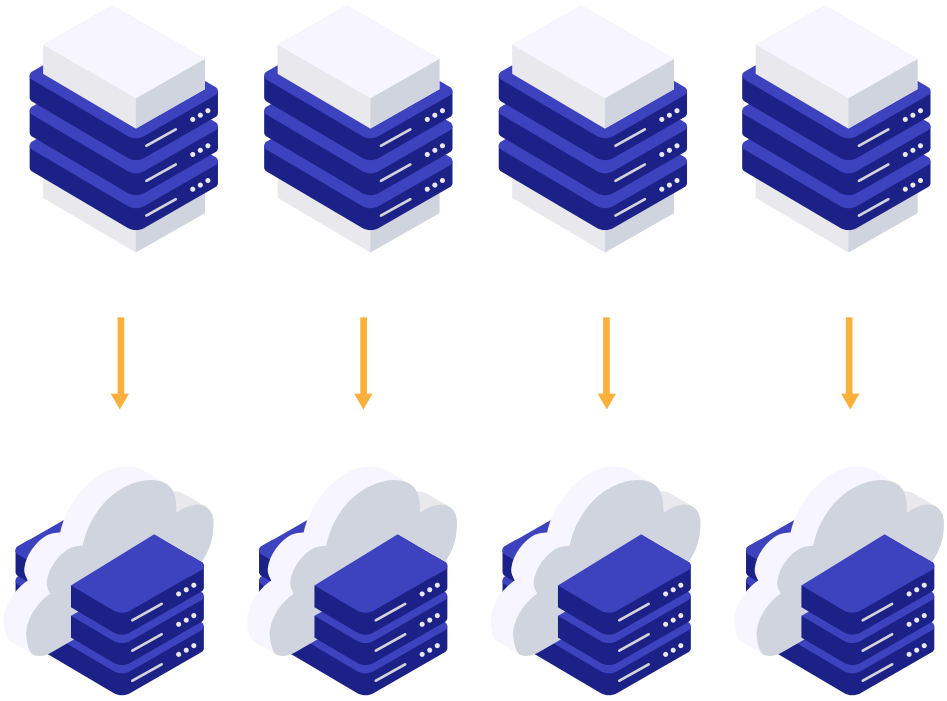
Существующие кластеры Kubernetes могут сразу переключиться на использование TATLIN.OBJECT для хранения и распространения образов контейнеров через стандартный OCI Distribution интерфейс. Для хранения образов можно применять политики, в том числе и с репликацией на другие площадки. Подключенные к общему хранилищу экземпляры OCI Distribution из разных ЦОД смогут пользоваться общей базой образов.



ПРИМЕРЫ ПРИМЕНЕНИЯ

Хранение артефактов разработки

- Универсальное хранилище для работы с исходным кодом, хранения сборок, тестовых артефактов и релизных версий приложений.
- Работа приложения с хранилищем идет по простому интерфейсу через вызовы API. Резервное хранение данных больше не является «головной болью» разработчиков приложения, т.к. СХД берет отказоустойчивое размещение объектов на себя.
- В отличие от традиционных подходов, при использовании TATLIN.OBJECT слой хранения не связан жестко со слоем обработки. Слой хранения масштабируется практически бесконечно и независимо от серверов приложения. Такую конструкцию легче проектировать и сопровождать.



ПРИМЕРЫ ПРИМЕНЕНИЯ

Предоставление услуги объектного хранения сервис-провайдерами

Сервис-провайдер может разместить узлы TATLIN.OBJECT на своих площадках и, разделив СХД на несколько наборов ресурсов (тенантов), предлагать от своего имени услугу по хранению клиентских данных в этих тенантах. Каждому потребителю услуги объектного хранения можно настроить свой адрес доступа к хранилищу и выполнять индивидуальное администрирование с установкой квот на объем хранимых данных, количество бакетов и объектов. Подсчет объема потребляемых ресурсов каждым тенантом возможен во внешней биллинг-системе посредством передачи в нее соответствующих метрик с узлов TATLIN.OBJECT.