

Fujitsu ETERNUS CD10000

— гипермасштабируемая СХД

В конце октября 2014 г. компания Fujitsu анонсировала СХД Fujitsu ETERNUS CD10000 – систему с практически неограниченными емкостью и сроком службы, а также способную поддерживать все типы доступа (включая будущие форматы) к данным.



Александр Яковлев – менеджер по маркетингу продукции RISC/UNIX серверы и системы хранения, Fujitsu.

Введение

Новая система ETERNUS CD10000 знаменует новую эру решений с экстремально высокой емкостью (сотни петабайт), предназначенных для повседневного хранения данных, а также для управления. На момент выпуска система поддерживает емкость до 56 Пбайт (56 000 Тбайт) онлайн-данных за счет агрегации до 224 узлов хранения. В ближайшее время компания Fujitsu представит обновления, обеспечивающие значительно более высокую масштабируемость.

Fujitsu создает совершенно новую экономику для компаний, управляющих наборами онлайн-данных объемом 250 Тбайт и более (поставщики облачных и телекоммуникационных услуг, финансовые организации, средства массовой информации и компании, специализирующиеся в области бизнес-аналитики и др.), а также для любых компаний из тех сегментов, где объемы онлайн-данных растут очень быстро.

В основе новой системы хранения Fujitsu корпоративного уровня лежит ПО Inktank Ceph Enterprise с открытым исходным кодом от компании Red Hat, а также дополнительные функциональные усовершенствования (где Ceph имеет пробелы, например, VMware, SNMP), призванные обеспечить всестороннее управление всей системой из единого окна. Благодаря полному всестороннему набору сервисов обслуживания и поддержки заказчики получают обслуживание корпоративного класса. Система ETERNUS CD10000 предлагает уни-

кальное единое представление блоков, объектов и файлов в одном распределенном кластере хранения данных, что позволяет снизить сложность и затраты на управление хранением данных, а также оптимизировать пространство на физических дисках для хранения данных.

Новые требования

Особенностью современного этапа развития ИТ (в соответствии с исследованием IDC, декабрь 2012 г.) является переход от второй к так называемой третьей платформе, в которой ведущую роль играют облачные технологии, мобильность, социальные сети, большие данные, интернет вещей, а также сверхвысокие требования к масштабируемости ИТ-инфраструктуры (рис. 1).

Новые бизнес-возможности третьей платформы обещают изменить облик ИТ благодаря массовому доступу пользователей к масштабируемым веб-сервисам.

Однако использование существующих ИТ-инфраструктур, которые создавались для сред на второй платформе и работают по принципу клиент/сервер, приводит к нехватке ресурсов, необходимых для преодоления разрыва между второй и третьей платформами. Помимо этого, есть ограничения и у наследованных СХД при петабайтных требованиях масштабируемости, в частности:

- высокое время перестраивания/восстановления RAID и высокие риски потери данных при этом;
- экспоненциально высокие затраты на поддержание доступности данных;
- проблемы с выделением ресурсов (необходимость поддержания большой избыточности или их нехватка) из-за неожиданно быстрого роста объемов данных;
- экстремально высокая продолжительность миграции данных;
- проблемы при плановых простоях;
- высокая стоимость единицы хранения;
- проблемы производительности.

Архитектура Fujitsu ETERNUS CD10000

ETERNUS CD10000 в концепции Fujitsu это масштабируемая программно-определяемая СХД, которая в полной мере ориентирована на поддержку: 1) аналитики больших данных, 2) инкапсулирования данных из социальных сетей в текущие бизнес-процессы, 3) широкополосной безопасной корпоративной мобильности, 4) облачных сервисов и отвечает следующим требованиям:

– в области масштабируемости:

- практически неограниченная масштабируемость как по производительности, так и по емкости хранимых данных (сотни петабайт, теоретически – эксабайты). Низкая стоимость единицы хранимой информации;
- предотвращение “узких мест” и горячих зон”, самооптимизация;

– в области надежности:

- полное резервирование;
- самовосстановление;

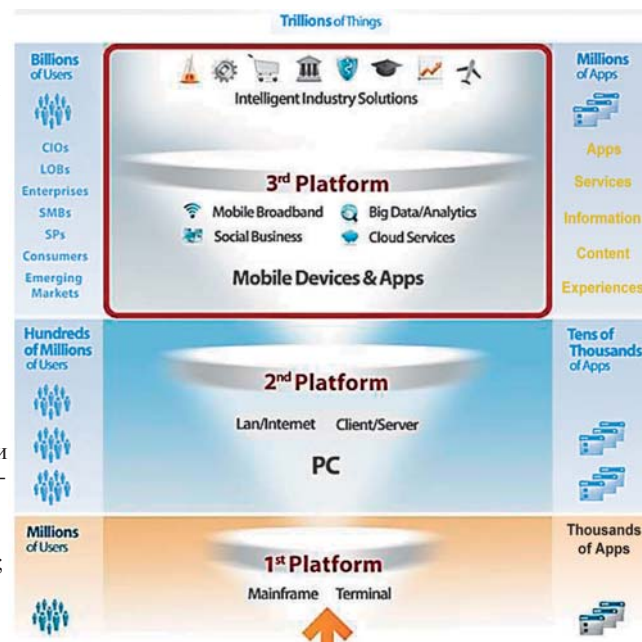


Рис. 1. Современный этап развития ИТ характеризуется переходом от второй к третьей платформе.

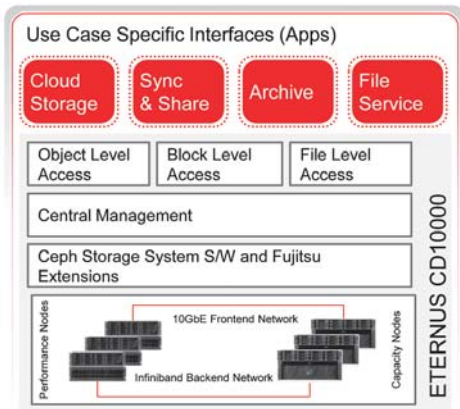


Рис. 2. Архитектура ETERNUS CD10000.

- географическая распределенность узлов системы;
 - быстрое восстановление/перестроение (rebuild);
 - нулевое время плановых и внеплановых простоев, апгрейд ПО и аппаратного обеспечения без останова;
- **в области управления:**
- централизованное эффективное управление огромным количеством хранимых данных;
 - унифицированный мультипротокольный доступ к данным (блочный, файловый, объектный и другие будущие форматы);
 - бесшовная интеграция новых единиц хранения.

Архитектура ETERNUS CD10000 представлена на рис. 2. Непосредственно ETERNUS CD10000 строится на трех типах узлов хранения (базовый и два других – “заточенные” на поддержание производительности и емкости хранения). Внешнее подключение (front end network) к узлам осуществляется через 10GbE. Сами узлы хранения дополнительно объединяются (backend network) на базе 40Gbit (IP based) Infiniband. Backend network дублируется для увеличения надежности и скорости внутреннего межсоединения между узлами хранения.

Ceph – ОС (на базе Linux) узлы хранения

Управление узлами хранения строится на Ceph File System – свободно распространяемой распределенной файловой системе. Ceph может использоваться на системах, состоящих как из нескольких машин, так и из тысяч узлов. Общий объем хранилища данных может измеряться петабайтами, встроенные механизмы продублированной репликации данных не зависят от отказа отдельных узлов и обеспечивают чрезвычайно высокую живучесть системы. При добавлении или удалении новых узлов, массив данных автоматически перебалансируется с учетом изменений¹⁾.

Ceph – open source software-defined storage платформа разработана для обеспечения объектного, блочного и файлового хранения на базе распределенных кластеризованных x86-серверах с возможностью масштабирования до экзабайтного уровня. Поддерживаются функции само-

восстановления (при сбоях/отказах и др.) и самооптимизации (по производительности).

В отличие от таких файловых систем, как GFS, OCFS2 и GPFS, в Ceph обработка данных и метаданных разделена на различные группы узлов в кластере, примерно как это сделано в Lustre, с тем различием, что обработка производится на уровне пользователя, не требуя какой-то особой поддержки от ядра. Ceph может работать поверх блочных устройств, внутри одного файла или используя существующую ФС (например, XFS)¹⁾.

Ceph использует четыре различных вида демонов:

- **мониторы кластера**, который отслеживает активные и отказавшие узлы кластера;
- **серверы метаданных**, которые хранят метаданные инодов и каталогов;
- **устройства хранения объектов (object storage devices – OSD)**, которые хранят содержимое файлов;
- **устройства передачи репрезентативного состояния (Representational state transfer – RESTful-gateways)**, которые предоставляют объектный уровень хранения как интерфейс, совместимый с Amazon S3 или OpenStack Swift APIs.

Ceph делает страйпинг индивидуальных файлов на множество узлов для достижения наибольшей производительности (подобно организации RAID0). Адаптивная балансировка нагрузки поддерживается за счет того, что часто используемые объекты реплицируются на большее число узлов.

Ceph имплементирует распределенное хранение объектов. Библиотеки программного обеспечения Ceph предоставляют клиентским приложениям прямой доступ к надежному автономному распределенному объектному хранению (reliable autonomic distributed object store – RADOS) объектно-ориентированной СХД, а также обеспечивают фундамент для ряда функций Ceph, включая: RADOS Block Device (RBD), RADOS Gateway и Ceph File System. Библиотеки RADOS обеспечивают доступ на языках C, C++, Java, Python, Ruby и PHP (рис. 3).

Блочный доступ к системе хранения объектов Ceph поддерживается через тонко разворачиваемое блочное устройство (block device) Ceph. Когда приложение записывает данные на Ceph с помощью блочного устройства, Ceph автоматически разбивает на страницы и реплицирует

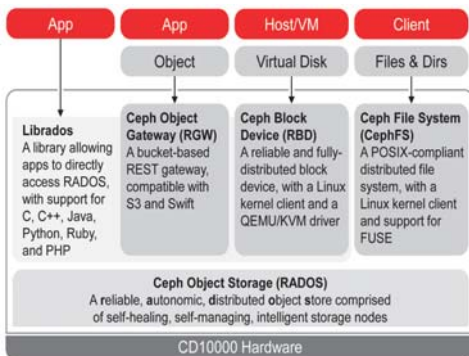


Рис. 3. Высокоуровневое представление различных способов доступа к данным в Ceph.

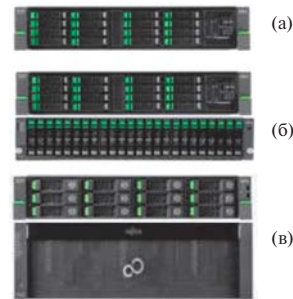


Рис. 4. Узлы хранения ETERNUS CD10000: Basic Storage Node (a), Performance Node (б) и Capacity Node (в).

данные по кластеру. Ceph’s RADOS Block Device (RBD) также интегрируется с ядром виртуальных машин (kernel virtual machines – KVM).

Block device поддерживается в виртуальных платформах, включая Apache CloudStack, OpenStack, OpenNebula, Ganeti и Proxmox Virtual Environment²⁾.

Развитие ETERNUS CD10000 разбито на 3 этапа. В первом релизе ETERNUS CD10000 обеспечивается поддержка:

- block: OpenStack/KVM;
- object: OpenStack-Swift;
- object: S3;
- file: CephFS (по специальному запросу).

Второй этап (будет представлен как Service Pack 1) предусматривает поддержку: block – VMware и file – full CephFS.

Третий этап включает поддержку: block – HyperV и file – Scale out NAS.

Узлы хранения

Узлы хранения строятся на базе x86 серверов PRIMERGY и ETERNUS JX JBODs. В первой версии ETERNUS CD10000 может масштабироваться до 224 узлов. Предусмотрено 3 типа узлов (рис. 4, табл. 1): Basic Storage Node (BSN), Performance Node (PN) и Capacity Node (CN).

Табл. 1. Типы узлов ETERNUS CD10000.

Тип узлов	Basic Node	Capacity Node	Performance
Емкость (raw, TB)	12.6	252.6	34.2
SSD-кэш	800 GB PCIe SSD	800 GB PCIe SSD	2x800 GB PCIe SSD
Порты	2x10 GbE	2x10 GbE	2x10 GbE
Размер (мм) (ШxГxВ/У)	482.6x770x86.9/ 2U	482.6x980x263/ 6U	483x770x175/ 4U

BSN обеспечивает 12,6 Тбайт, используя 2.5” SAS-диски (10k rpm). PN это узел хранения плюс ETERNUS JX40, который добавляет 21,6 Тбайт на 2.5” SAS-дисках (10k rpm). Общая емкость PN – 34,2 Тбайт. CN это узел хранения плюс ETERNUS JX60 (240 Тбайт; 3.5” SATA-диски; 7,2K rpm). Общая емкость – 252,6 Тбайт.

В составе узлов хранения предусмотрено использование SSD-дисков. Пул из таких дисков может использоваться как пул быстрых дисков, так и как SSD-кэш.

Области применения

Области применения ETERNUS CD10000 обуславливаются объемами хранения 250 Тбайт и выше, а также экспоненциальным непредсказуемым ростом данных. Примеры использований:

1) см. википедия: https://ru.wikipedia.org/wiki/Ceph_File_System.

2) см. википедия: [http://en.wikipedia.org/wiki/Ceph_\(software\)](http://en.wikipedia.org/wiki/Ceph_(software)).

- провайдеры облачных сервисов;
- провайдеры телекоммуникационных услуг;
- исследовательские организации, государственные учреждения, университеты;
- компании с большим объемом исследований;
- государственные учреждения с огромными репозиториями документов;
- СМИ, ширококвещательные компании;
- задачи бизнес-аналитики, требующие быстрого доступа к большим объемам (исторических) данных.

Сравнение областей и особенностей применения ETERNUS DX и ETERNUS CD10000 приведено в табл. 2.

*Александр Яковлев,
Fujitsu.*

Табл. 2. Сравнение областей и особенностей применения ETERNUS DX и ETERNUS CD10000.

System	ETERNUS DX	ETERNUS CD10000
Storage category	Production, Business Analytics	Cloud, Content Depot
Usage area	DB, OLTP, CEM, ERP, File, Hyper-V, VMware, business analytics,	document store, cloud data services, OpenStack, big data, multimedia
Verticals	horizontal	Cloud and telecom service providers, R&D, public sector, media/broadcasting, automotive / aerospace, defense
Architecture	Scale-up, RAID	Hyper scale-out
Capacity	few TB up to 6 PB	250 TB- 50 PB
Typical Data retention time	days, weeks, months	several months, several years
Data access	file, block	object, block, (file)
External Connectivity	(SAS), Ethernet, FC, iSCSI, FCoE	Ethernet
Virtualization support	VMware, Hyper-V	KVM, (VMare planned)
DR	Short, long distance	Short distance

Fujitsu: флэш-массив для непрерывности бизнес-процессов

Март 2015 г. – Fujitsu расширила возможности компаний от высокого риска потери данных при авариях, не требуя усложнения имеющейся ИКТ-инфраструктуры, анонсировав новую систему хранения Fujitsu ETERNUS DX 200F, которая позволяет автоматически или в ручном режиме поддерживать непрерывность бизнес-процессов в случае плановых или неплановых отключений электроэнергии.

Когда речь заходит о критически важных данных, которые должны быть доступны 24 часа в сутки, устойчивая к отказам конфигурация абсолютно необходима для обеспечения непрерывной торговли в интернет-магазинах, доступа к транзакционным базам данных и бесперебойной работы виртуальных машин. Флэш-массивы могут предоставить производительность, необходимую для подобных приложений, однако без такого преимущества, как стопроцентная непрерывность бизнес-процессов. До сегодняшнего дня подобные конфигурации были обычно доступны только в сочетании с системами более высокого среднего и корпоративного уровня, которые, как правило, находятся за пределами диапазона цен для большинства компаний. Они настолько сложны, что многие ИТ-подразделения недостаточно уверенно чувствуют себя при управлении этими сложными средами. Таким образом, по мере возможности компании просто избегают их использовать – подвергая риску свой бизнес и деловую репутацию.

Система хранения ETERNUS DX 200F, буква «F» в названии которой означает «флэш», оперативно разрешает все проблемы производительности критически важных для бизнеса приложений, не требуя сложной настройки. Новейшее дополнение к семейству бизнес-ориентированных СХД ETERNUS DX использует

популярную систему управления, что обеспечивает высокую совместимость и упрощает администрирование. Флэш-массив предварительно сконфигурирован и содержит от 5 до 24 твердотельных накопителей общей емкостью до 38,4 Тбайт. Твердотельные накопители позволяют повысить скорость доступа к данным в 100 раз и более и значительно превосходят жесткие диски, особенно там, где имеют значение задержки доступа к данным, способные представлять реальную проблему в самых ресурсоемких бизнес-приложениях.

В отличие от очень дорогих решений с флэш-массивами на базе проприетарных системных архитектур со специально встроенными флэш-компонентами, ETERNUS DX 200F использует высокопроизводительную архитектуру систем хранения серии ETERNUS DX RAID и ценовые преимущества стандартных твердотельных накопителей. Как уже доказано отраслевыми тестами производительности систем хранения данных (*официальный результат теста SPC -2, проведенного в феврале 2015 года, доступен по адресу: http://www.storageperformance.org/results/benchmark_results_spc_2_active/#b00071*), такое сочетание обеспечивает лучшее на рынке ускорение приложений по существенно более низкой цене».

В то время как многие флэш-системы гарантируют лишь ограниченный уровень высокой готовности и восстановления данных после аварий, СХД ETERNUS DX 200F поддерживает синхронизацию данных между двумя системами. Используя функциональность кластера ETERNUS Storage Cluster, можно сконфигурировать прозрачную систему обработки отказов, которая позволяет пользователям комбинировать производительность флэш-решения с полной устойчивостью к авариям, что в настоящее время является уникальным для современного рынка.

Переход на резервную систему или вторую площадку происходит автоматически во время отказа – абсолютно незаметно для пользователей и без необходимости вмешательства системных администраторов. Доступ ко всем приложениям происходит в режиме реального времени, а все

системы в вычислительной среде с высоким уровнем готовности могут продуктивно работать при выполнении штатных операций. Данное решение также поддерживает ручную обработку отказов, которая помогает в случае запланированных отключений электроэнергии, проведения тестов на аварийное восстановление и выполнения обновлений продолжить работу без прерываний.

ЦРТ: новая биометрическая платформа VoiceKey 2.0

Март 2015 г. – VoiceKey 2.0 разработан дочерней компанией ООО "Центр речевых технологий" – "ЦРТ-инновации" и используется для аутентификации клиентов при дистанционном обслуживании. Подтверждение личности с помощью голоса и лица позволяет отказаться от PIN-кодов и паролей.

VoiceKey 2.0 – это безопасный доступ к мобильному приложению или личному кабинету на сайте, защита персональной и корпоративной информации на ПК или в мобильном приложении, удобная верификация в контакт-центре, в том числе в IVR. VoiceKey снижает риск мошенничества и сокращает время обслуживания клиента за счет применения мультимедийной аутентификации там, где раньше она была невозможна – в телефонном канале.

Новую биометрическую платформу от ЦРТ отличают поддержка голосовой и лицевой биометрии с возможностью подключения других модальностей (палец, ладонь, радужка и др.), масштабируемость, языконезависимость, гибкая интеграция с мобильными и web-приложениями, большинством популярных АТС, IVR и CRM, возможность настройки прав доступа.

Аутентификация производится по парольной фразе (статической или динамической) и по свободной речи, в онлайн- и оффлайн-режимах. Причем точность верификации по парольной фразе увеличена в 2 раза. В VoiceKey 2.0 встроен сервер съема звука, добавлены поддержка виртуализации и новая подсистема безопасности.