

Технология хранения данных на основе грид-архитектуры и ее преимущества

Silverton Consulting, Inc. StorInt™ Briefing



Введение

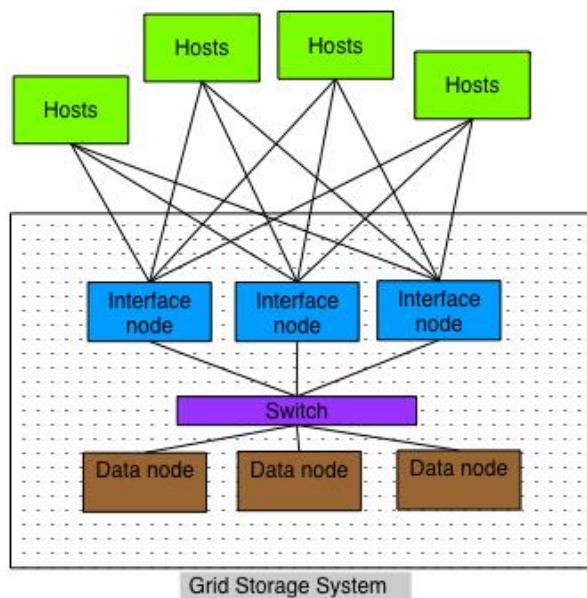
Распределенная обработка данных стала развиваться в начале этого века, когда микропроцессоры и межкластерные коммуникационные каналы стали более доступными и достаточно быстрыми — такими, что системы могли обрабатывать задачи одновременно, например, операции ввода-вывода (IO). Системы хранения на основе грид-архитектуры — это следующий этап развития распределенных вычислений. Они фокусируются на сервисах хранения данных.

Эта технология хранения данных стала весьма популярной, и поэтому возникло несколько различных решений для систем хранения на основе грид-архитектуры. Компания IBM® за последние 10 лет разработала ряд инновационных решений для грид-архитектуры хранения данных, и в прошлом году были представлены система хранения XIV® и программно определяемая система хранения IBM Spectrum Accelerate™.

Хранение данных с использованием грид-архитектуры предлагает значительные технические преимущества, такие как простота развертывания, горизонтальное масштабирование производительности подсистемы ввода-вывода, отсутствие необходимости в настройке системы хранения, и идеально подходит для многих центров обработки данных, таких как среды виртуализации, облачные сервисы и сервисы анализа данных. Мы расскажем об этих преимуществах позднее, но сначала поясним принципы работы систем хранения на основе грид-архитектуры и их отличия от стандартных систем хранения.

Масштабируемая грид-архитектура хранения данных

Системы хранения данных на основе грид-архитектуры состоят из некоторого количества составных частей, называемых модулями или узлами. Все узлы в грид-среде взаимодействуют друг с другом как кластеры, обеспечивая сервисы хранения данных. В системах хранения на основе грид-архитектуры используется один из двух вариантов развертывания:



1. **Программно определяемая система хранения данных на основе грид-архитектуры**, то есть система хранения данных разворачивается как программное решение на обычном аппаратном клиентском сервере с использованием системы хранения данных с прямым подключением как вычислительной машины. В программно определяемой грид-системе каждый экземпляр системы хранения или виртуальная машина (VM) действует как собственный контроллер или узел.

2. **Аппаратно определяемая система хранения данных на**

основе грид-архитектуры, то есть система хранения данных разворачивается на предоставленном поставщиком аппаратном решении с использованием предоставленной поставщиком системы хранения данных как вычислительной машины.

Каждый узел грид-системы соединяется с другими модулями внутри системы и объединяется в межкластерную сеть, такую как InfiniBand или высокоскоростной Ethernet, чтобы данные можно было при необходимости быстро передавать из одного модуля грид-среды в другой.

В стандартной грид-системе хранения данных у всех узлов одинаковые схемы обработки, сетевого соединения, количество накопителей данных и емкость памяти. Кроме того, у некоторых модулей грид-среды есть подключение к главному интерфейсу. Все модули грид-среды участвуют в действиях ввода-вывода главного интерфейса, и поэтому грид-система хранения данных может выполнять крупномасштабные параллельные или одновременные операции ввода-вывода.

Для упрощения одновременной обработки ввода-вывода в грид-среде главный объем данных разбивается на сегменты, которые обрабатываются псевдослучайным образом во всех узлах грид-среды. Таким образом, главный объем данных распределяется по всем узлам и накопителям данных в кластере грид-системы хранения данных. При этом все модули и накопители данных потенциально активны при обслуживании ввода-вывода хоста.

Например, если хост записывает данные в том, то данные попадают в грид-среду через узел с интерфейсом хоста и собираются в кэше этого узла. В момент переноса данных из оперативной памяти этот узел разбивает данные на сегменты и выполняет расчет, определяя, в каком узле будут

содержаться эти данные. Затем сегмент передается в соответствующий модуль и там записывается в один из накопителей этого узла. В процессе чтения происходят такие же процессы. Таким образом, грид-система хранения данных размещает данные при записи и собирает данные при чтении, обслуживая процесс ввода-вывода хоста. Более того, поскольку данные хоста распределены по всем накопителям данных, то для указания новых томов хранения единственное, что должен сделать администратор, — это убедиться, что для выполнения запроса ресурсов достаточно.

В грид-системе хранения данных процедуры защиты данных включают зеркалирование данных, когда данные хоста автоматически записываются на два или три различных накопителя, расположенных в различных модулях. В результате отказ одного накопителя или узла не становится помехой при доступе к данным хоста.

Помимо дисковых накопителей данных в грид-системах могут использоваться флеш-накопители как расширение кэша данных. Такое использование флеш-накопителей увеличивает объем данных хоста, который можно быстро считать из кэша.

Поскольку все узлы унифицированы, заказчик легко может расширить емкость грид-системы и увеличить ее производительность, добавив больше модулей. Если грид-система обнаруживает новый узел, она автоматически перераспределяет сегменты данных хоста по новой, расширенной сетке, распределяя процесс ввода-вывода хоста по всем модулям.

Отличие грид-систем хранения данных от традиционных систем

Обычные широко доступные системы хранения данных построены на **архитектуре с двумя контроллерами** с присоединенным хранилищем данных. В такой архитектуре у каждого из контроллеров есть интерфейс хоста и кэш данных.

В архитектуре с двумя контроллерами и без широкого чередования данные тома хоста размещаются только на определенных накопителях, присвоенных одному контроллеру. Для конфигурации томов хоста заказчик должен определить, какой контроллер будет использоваться для этих данных. Более того, в архитектуре с двумя контроллерами защита данных обычно строится на различных уровнях RAID, что означает, что данные тома хоста присваиваются не только определенному контроллеру, но и определенному набору накопителей данных в группе RAID, закрепленному за этим контроллером.

Как правило, такие присвоения существуют, пока оператору не понадобится переместить данные, поскольку такая конфигурация данных тома хоста может приводить к появлению «горячих дисков» — накопителей данных, производительность которых страдает вследствие слишком активных операций ввода-вывода. Если в архитектуре с двумя контроллерами возникают проблемы с операциями ввода-вывода, администратор хранилища должен настроить систему, перемещая данные хоста в другие RAID-группы или **в другой контроллер**.

В целом в некоторых конфигурациях один из контроллеров может страдать проблемами «горячего контроллера». В этом случае оператору может понадобиться переместить данные из целой RAID-группы с одного контроллера на другой или **в другую систему хранения** для устранения проблемы.

В отличие от поддержки любых операций в грид-среде хранения данных, архитектура с двумя контроллерами поддерживает или операции «активный-активный» или «активный-пассивный». В системах с конфигурацией «активный-активный» оба контроллера обслуживают операции ввода-вывода хоста все время, а в системах с конфигурацией «активный-пассивный» только один контроллер обслуживает ввод-вывод, а второй в это время простаивает. В случае неисправности контроллера в конфигурации «активный-активный» другой контроллер перехватывает доступ к накопителям данных неисправного контроллера. Таким образом, второй контроллер занимается всеми процессами ввода-вывода хоста, но использование только одного контроллера снижает производительность ввода-вывода вдвое.

Производительность системы хранения с двумя контроллерами невозможно горизонтально масштабировать. К этим системам можно добавлять накопители данных, кэш-память и интерфейсы хоста, но новые вычислительные возможности добавить нельзя. Более того, данные должны оставаться в одной группе RAID на одном или другом контроллере, и поэтому производительность ввода-вывода для одного тома хоста не повышается при добавлении емкости. Дополнительные кэш-память и интерфейсы хоста могут немного увеличить производительность, но добавление емкости не увеличивает производительность ввода-вывода без перемещения данных тома хоста.

Преимущества грид-среды хранения данных

Технические преимущества грид-среды хранения данных по сравнению с традиционными системами с двумя контроллерами:

- **Нет горячих точек** — грид-система автоматически распределяет данные хоста по модулям и накопителям грид-среды. Данные тома хоста не содержатся в одном узле или накопителе, поэтому не возникает проблем с «горячими» накопителями или контроллерами.

- **Автоматическая настройка производительности** — администраторам грид-систем хранения данных больше не нужно перемещать данные для регулировки производительности, поскольку данные постоянно распределяются по всем модулям и накопителям данных.
- **Равномерная производительность ввода-вывода** — грид-системы хранения данных обеспечивают унифицированную предсказуемую производительность ввода-вывода при разнообразных прикладных нагрузках вследствие отсутствия горячих точек и благодаря автоматической настройке таких систем.
- **Беспроблемная конфигурация томов** — конфигурация тома хоста в грид-системе хранения данных стала намного проще, теперь заказчик не должен решать, какой контроллер, группу RAID и защиту данных использовать для данных тома хоста.
- **Более быстрое восстановление** — грид-система хранения данных может вернуть полную работоспособность быстрее, поскольку все накопители и узлы принимают участие в считывании зеркалированных данных с неисправных накопителей и узлов и в записи этих данных в новые места.
- **Линейная масштабируемость производительности** — производительность ввода-вывода в грид-системе хранения данных может изменяться линейно при добавлении большего количество модулей, поскольку вычислительные возможности и емкость добавляются совместно, в отличие от систем с двумя контроллерами, где добавляется только емкость, а нагрузка на вычислительные системы увеличивается. При расширении вычислительных возможностей, кэш-памяти и накопителей грид-система получает дополнительные ресурсы, необходимые для улучшения производительности любых операций ввода-вывода хоста.

Указанные выше технические преимущества становятся реальной выгодой для заказчика: беспроблемная система хранения данных, которой легко пользоваться, просто конфигурировать и легко управлять. При размещении и сборе данных в процессе ввода-вывода в грид-системе хранения данных автоматически обрабатывается расположение данных, что исключает проблемы с размещением данных и производительностью (т.е. система автоматически определяет, какой физический накопитель ассоциируется с какими данными хоста, какой контроллер или RAID-группа перегружены).

Кроме того, в отличие от систем с двумя контроллерами заказчики, работающие с грид-системами хранения данных, легко могут

масштабировать производительность одновременно с емкостью. Как уже было сказано ранее, добавляемый в грид-среду узел включает процессоры, кэш-память и накопители данных, а производительность ввода-вывода увеличивается линейно, поскольку система автоматически перераспределяет данные хоста по новым узлам.

Система хранения данных на основе грид-архитектуры в IBM



IBM предлагает три решения для грид-систем хранения данных: IBM Spectrum Accelerate™, IBM XIV™ и IBM FlashSystem A9000/IBM FlashSystem A9000R™. Все грид-системы хранения данных от компании IBM имеют схожую функциональность, но их производительность различна. Все функции грид-решений IBM изначально были реализованы в XIV, а теперь внедряются в Spectrum Accelerate.

- **IBM Spectrum Accelerate** – программно определяемая грид-система хранения данных с расширенными функциями хранения, например, зеркалирование с использованием двух или трех узлов, дифференциальные моментальные копии, асинхронная или синхронная репликация. Spectrum Accelerate отличается удобным графическим пользовательским интерфейсом мирового уровня с API-интерфейсом RESTful для автоматизации центров обработки данных и расширенной поддержкой VMware и сервисов хранения Microsoft.

Spectrum Accelerate также существует как гибридное облачное решение с сервисами IBM SoftLayer.

- Система хранения IBM XIV** – аппаратно определяемая грид-система хранения данных, обеспечивающая хранение либо только на дисках, либо на дисках и флеш-накопителях с межкластерной сетью InfiniBand и источником бесперебойного питания. XIV поддерживает сжатие Real-time Compression™ и накопители с поддержкой автоматического шифрования для большей эффективности и безопасности хранения данных.
- Системы хранения IBM FlashSystem A9000 и IBM FlashSystem A9000R** – аппаратно определяемые грид-системы хранения данных с высокой скоростью операций ввода-вывода, с узлами грид-среды, включающими обеспечивающие малые задержки флеш-накопители с Spectrum Accelerate, а также высокопроизводительные узлы флеш-накопителей с технологией IBM FlashCore™ мирового класса. IBM FlashSystem A9000 имеет один узел флеш-накопителей с тремя контроллерами грид-среды, а IBM FlashSystem A9000R поддерживает несколько флеш-модулей и контроллеров грид-среды.



Рекомендуемые сферы применения для систем хранения данных на основе грид-архитектуры от IBM

Среды виртуализации

IBM Spectrum Accelerate предлагает поддержку новых сервисов хранения данных VMware, таких как VMware Virtual Volumes (VVOLs), управление хранением с vSphere APIs for Storage Awareness (VASA) и vStorage APIs for Array Integration (VAAI). Благодаря такой широкой поддержке сервисов VMware любая грид-система хранения данных от IBM прекрасно подходит для сред VMware VM. Что не менее важно, VMware смешивает потоки ввода-вывода из нескольких VM в запросы ввода-вывода одного сервера. Отсутствие «горячих» точек вследствие широкого распределения данных по грид-среде, автоматическая настройка, постоянная производительность ввода-вывода и линейная масштабируемость производительности помогут обслуживать эту требовательную смешанную среду ввода-вывода. Благодаря расширенной поддержке автоматизации репликации Microsoft Hyper-V с System Center Virtual Machine Manager (SCVMM), Microsoft Azure Site Recovery, мониторингу системы с помощью Microsoft System Center Operations Manager (SCOM) и схожих с Hyper-V характеристикам

производительности ввода-вывода, грид-системы хранения данных от IBM отлично подходят для заказчиков Microsoft Hyper-V.

Облачные среды и среды поставщиков сервисов

В отличие от архитектур с двумя контроллерами линейная масштабируемость производительности вместе с емкостью, автоматическая подстройка и поддержка нескольких арендаторов в грид-системах хранения данных IBM позволяют облачным средам, поставщикам управляемых услуг и другим организациям, предоставляющим «все как сервис» (XaaS), быстро и легко размещать данные хоста в грид-системах IBM. Более того, простота управления хранением и поддержка API-интерфейса RESTful делают грид-систему хранения данных IBM гораздо более простой в настройке, автоматизации и постоянной эксплуатации. Эти преимущества, а также равномерная производительность ввода-вывода особенно важны для облачных сред и поставщиков услуг, у которых приложения и данные постоянно изменяются.

Аналитические среды

Широко распространенные функции ввода-вывода в грид-системах хранения данных отличаются более высокой производительностью по сравнению с архитектурами с двумя контроллерами. Они предпочтительны для заказчиков, у которых системы анализа и приложения должны работать с одними и теми же данными хоста. Отсутствие «горячих» точек означает, что операции ввод-вывода аналитического приложения будут меньше пересекаться с прикладными операциями ввода-вывода.

Производительность ввода-вывода грид-системы хранения данных достигает пиковых значений, когда данные приложений и анализа размещаются в грид-системах IBM FlashSystem A9000/IBM FlashSystem A9000R на основе флеш-памяти. Работая с системой хранения Spectrum Accelerate, заказчики могут размещать свои аналитические приложения на тех же серверах, где расположены модули грид-системы.

Заключение

Благодаря своим характеристикам, системы хранения данных на основе грид-архитектуры отлично подходят для корпоративной архитектуры, облачных сред и поставщиков услуг. С технической точки зрения, классическая архитектура с двумя контроллерами не может соперничать с неизменной простотой конфигурации, масштабируемой производительностью и автономной настройкой грид-системы хранения данных. Эти особенности выражаются в таких практических преимуществах для заказчиков, как линейная масштабируемость производительности и бесперебойное хранение данных.

Не менее важно то, что предлагая программно определяемые системы Spectrum Accelerate, системы XIV (как на основе только дисков, так и гибридные с дисками и флеш-накопителями) и системы хранения IBM FlashSystem A9000/IBM FlashSystem A9000R на основе флеш-памяти,

корпорация IBM предлагает самый широкий спектр систем хранения данных на основе грид-архитектуры на рынке на сегодняшний день.

Тот факт, что IBM добавила технологию мирового уровня FlashSystem в грид-системы хранения данных, еще раз подтверждает множество преимуществ такой архитектуры. Ну и, наконец, заказчики корпоративных центров обработки данных, облачных сред и поставщики услуг смогут выбрать подходящее для себя решение хранения данных на основе грид-архитектуры от IBM.

Silverton Consulting, Inc. является консалтинговой фирмой, расположенной в США, и предлагает продукты и сервисы для хранения данных.



Отказ от обязательств: Данный документ был разработан по заказу International Business Machines Corporation (IBM). Хотя в документе используются материалы из различных открытых источников, включая IBM, в нем не обязательно отражаются позиции таких источников касательно обсуждаемых проблем.