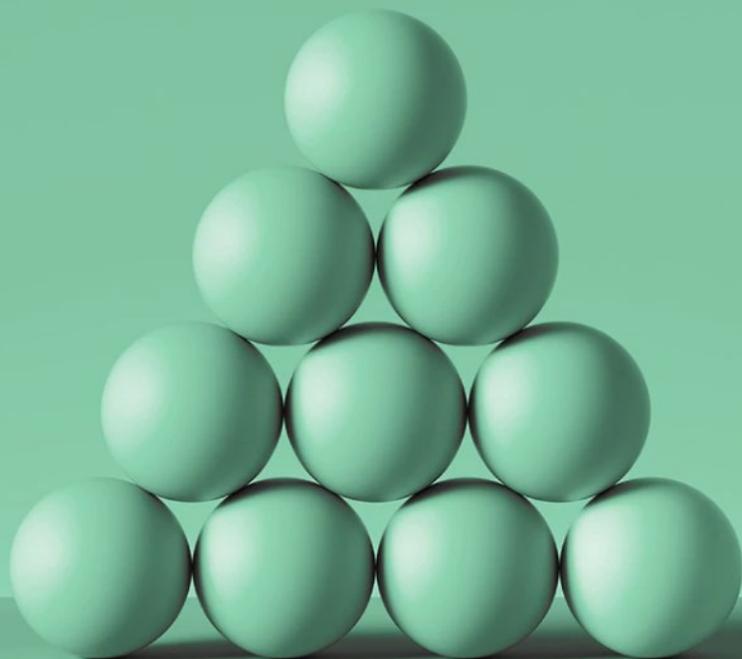


ОПИСАНИЕ РЕШЕНИЯ

ИИ NetApp EF-Series

Ускорьте аналитику данных
за счет быстрой потоковой
передачи информации
в приложения ИИ



Сложности ИИ-инфраструктуры

Искусственный интеллект (ИИ) и глубокое обучение (ГО) позволяют компаниям выявлять случаи мошенничества, чтобы улучшать отношения с клиентами, оптимизировать цепочку поставок и предоставлять инновационные продукты и услуги в условиях все более конкурсирующего рынка.

Однако, чтобы получить максимальные преимущества от глубокого обучения, сначала необходимо решить несколько ключевых задач. Варианты самостоятельной интеграции довольно сложны. Система для работы с глубоким обучением, состоящая из типовых компонентов, которые обеспечивают вычисления, хранение данных, сетевое и программное взаимодействие будет обладать сложной архитектурой и потребует дополнительного времени на развертывание.

Традиционно для ИИ использовались вычислительные системы с непосредственно подключаемой СХД. Однако, масштабирование может вызывать перебои и простои в работе существующей СХД. Сбои в работе влияют на эффективность работы инженеров и специалистов по обработке данных. Простои и низкая производительность ИИ могут вызвать цепную реакцию, которая снижает производительность разработчиков и непредсказуемо повышающую операционные расходы.

Решение

Алгоритмы с большим объемом вычислений для систем ГО подходят для архитектуры GPU. А поскольку графические процессоры на базе систем NVIDIA DGX могут проводить подобные вычисления, они теперь являются предпочтительной платформой для таких рабочих нагрузок, как высокопроизводительные вычисления (HPC), ГО, обработка видео и аналитика. Этим средам необходима инфраструктура хранения и сетевая инфраструктура, которая может обеспечивать передачу данных на графические процессоры. Им также необходим доступ к наборам данных со сверхмалой задержкой и высокой пропускной способностью.

NetApp® EF-Series AI интегрирует системы DGX A100, системы хранения NetApp EF600 NVMe и параллельную файловую систему BeeGFS с современной сетью InfiniBand. NetApp EF600 AI упрощает развертывание искусственного интеллекта, устраняя сложность структуры и необходимость в догадках. Вы можете начать с малого и без проблем масштабировать свою систему на любом этапе, от научных экспериментов и проверки концепции до производства и многого другого.

Основные преимущества

Простое развертывание

- Снижение рисков с помощью гибких проверенных решений.
- Ускоренное развертывание за счет простой архитектуры, не требующей предварительной отладки.

Обеспечение того уровня производительности и масштабируемости, который нужен вашему бизнесу

- Быстрое начало и беспроблемное расширение.
- Ускорение работы благодаря высокопроизводительным технологиям.

Оптимизация хранилища BeeGFS и службы метаданных

- Оптимизируйте различные рабочие ИИ-нагрузки в едином пространстве имен.
- Воспользуйтесь преимуществами простой и быстрой установки и удобного управления.

На рис. 1 показаны некоторые варианты семейства ИИ-решений серии EF с системами DGX A100. Совместимость блоков BeeGFS на основе EF600 была подтверждена для восьми систем DGX A100. Добавляя больше таких блоков, вы можете поддерживать множество систем DGX A100 и обрабатывать петабайты данных. У вас есть возможность изменять соотношение вычислительных ресурсов и хранилища независимо от размера озера данных, используемых моделей ГО и необходимой производительности.

Простое развертывание

Технология искусственного интеллекта эволюционирует довольно быстро, поэтому разработка эффективной ИИ-инфраструктуры также является сложной задачей. Систем EF-Series AI позволяет быстрее начать работу благодаря проверенной на практике эталонной архитектуре. Оптимизация конфигурации и развертывания.

Обеспечение того уровня производительности и масштабируемости, который нужен вашему бизнесу

Обучающие процессы для глубокого обучения требуют высокой вычислительной мощности. Ускоренная обработка обучающих образов поможет сократить общие расходы на вычисления, при этом ускоряя внедрение технологий искусственного интеллекта и повышая производительность.



Рисунок 1. Строительные блоки BeeGFS на основе серии EF с системами NVIDIA DGX A100

Система DGX A100, созданная с использованием новой архитектуры NVIDIA Ampere, повышает производительность обучения почти в шесть раз по сравнению с предыдущим поколением. Вы получаете эквивалент вычислительной инфраструктуры ЦОД для аналитики, обучения и вывода в эксплуатацию, объединенной в единую систему. По сравнению с системами на основе центральных процессоров, DGX A100 занимает в 25 раз меньше места, потребляет в 20 раз меньше энергии и в 10 раз дешевле.

Инвестирование в передовые вычислительные технологии требует ультрасовременной СХД, которая может обрабатывать тысячи обучающих образов в секунду. Этого можно добиться с помощью высокопроизводительных сервисов обработки данных, способных обеспечивать самые высокие рабочие нагрузки, связанные с глубоким обучением.

Флеш-массив NetApp EF600 обеспечивает постоянный доступ к данным в режиме, близком к реальному времени, при одновременной поддержке любого количества рабочих нагрузок. Для быстрой и непрерывной подачи данных в ИИ-приложения системы хранения EF600 обеспечивают до 2 миллионов операций ввода-вывода в секунду, время отклика менее 100 микросекунд и пропускную способность последовательного считывания 42 Гб/с в едином корпусе. Благодаря надежности систем хранения EF600 на уровне 99,9999% данные для операций искусственного интеллекта доступны в любое время и в любом месте.

Оптимизация хранилища BeeGFS и службы метаданных

BeeGFS — это параллельная файловая система, которая обеспечивает гибкость, необходимую для выполнения разнообразных и постоянно растущих рабочих нагрузок ИИ. Системы хранения NetApp серии EF дополняют хранилища BeeGFS и службы метаданных, облегчая RAID и другие задачи, например, мониторинг дисков и отслеживание износа.

NetApp и NVIDIA: продвигаем инновации вместе

Система DGX A100 — это платформа ГО нового поколения, для работы с которой требуются современные системы хранения и управления данными. Поскольку она сочетает в себе DGX A100 со стандартными блоками BeeGFS на основе систем NetApp EF600, такую архитектуру можно реализовать практически в любом масштабе. Например, можно связать одну систему DGX A100 с одним строительным блоком BeeGFS. Или у вас может быть до 140 систем DGX A100 с масштабируемым количеством строительных блоков BeeGFS, которые представляют собой единое пространство имен. Благодаря превосходной облачной интеграции и программно-определяемым возможностям портфеля продуктов, решения NetApp позволяют успешно реализовывать ГО-проекты на периферии, в собственном ЦОД и в облаке.

Компоненты решения

- Системы NVIDIA DGX A100
- Флеш-массивы NetApp EF600
- NVIDIA Mellanox Quantum QM8700
- Программный стек NVIDIA DGX
- Параллельная файловая система ThinkParQ BeeGFS

О компании NVIDIA

Изобретение компанией NVIDIA графического процессора в 1999 г. повлекло рост рынка игровых ПК, заставило производителей переосмыслить современную компьютерную графику и произвело революцию в области параллельных вычислений. Недавно глубокое обучение на основе GPU поспособствовало развитию современного искусственного интеллекта — следующей эпохи вычислений — где графический процессор выполняет роль мозга компьютеров, роботов и автомобилей с автономным управлением, наделяя их возможностью воспринимать и понимать окружающий мир. Более подробную информацию можно найти на сайте www.nvidia.com.

О компании NetApp

В отличие от многих NetApp не является экспертом широкого профиля, но это истинный специалист в своей области. Наша специализация — помочь вам извлечь максимум преимуществ из ваших данных. Мы предлагаем облачные сервисы по хранению данных корпоративного класса, которые не подведут вас в нужный момент, а также инструменты для удобной и гибкой работы с вашими ЦОД в «облачном» формате. Нами передовые решения подойдут для любой пользовательской среды и совместимы с крупнейшими публичными облаками.

Компания NetApp всегда специализировалась на разработке программного обеспечения для работы с облаками и обработки данных. Поэтому именно наши продукты помогут вам создать собственную фабрику данных или упростить управление вашей облачной инфраструктурой и наладить взаимодействие между ее элементами. А если вы поставщик цифровых решений, будь то данные, сервисы или приложения, то мы обеспечим их точную и безопасную поставку клиентам в любых условиях. www.netapp.com/ru

