

Keysight AI Data Center Builder AI / ML infrastructure evaluation and benchmarking

Введение

Бенчмаркинг фабрики кластера искусственного интеллекта (ИИ) и машинного обучения (МО) с реалистичными рабочими нагрузками, как правило, требует значительных вложений в вычислительные системы, такие как графические процессоры (GPU) и сетевые карты с поддержкой прямого доступа к памяти (RDMA NIC). Эти системы не только дорогостоящи, но и требуют значительных временных затрат на построение и эксплуатацию. Проблема заключается не только в финансовых расходах, но и в сложности сопровождения столь развитой инфраструктуры.

В настоящем техническом обзоре представлено, каким образом решение Keysight AI (KAI) Data Center Builder решает данные задачи за счет ускорения инновационных циклов развития сетей ИИ/МО и ввода-вывода. Это инновационное программное обеспечение использует высокоплотные генераторы трафиковой нагрузки с поддержкой портов 100–800GE для эмуляции хостов ИИ и реалистичных шаблонов трафика ИИ/МО. Вы узнаете, как KAI Data Center Builder помогает ускорить выпуск релизов, снизить затраты, упростить развертывание и обслуживание, а также повысить общую производительность.

Задачи

Для решения задач бенчмаркинга фабрики кластеров ИИ и МО с реалистичными рабочими нагрузками необходимо комплексное решение, минимизирующее как стоимость, так и сложность. Ниже перечислены пять ключевых компонентов такого решения:

Реалистичная эмуляция: Решение воспроизводит масштабные рабочие нагрузки ИИ с измеримой достоверностью, обеспечивая глубокий анализ эффективности коллективных коммуникаций.

Упрощенный бенчмаркинг: Продукт включает в себя готовые бенчмарк-приложения, что упрощает валидацию сетевой фабрики ИИ.

Гибкие испытательные движки: Пользователи могут выбирать между аппаратными генераторами нагрузки, программными конечными точками или реальными ИИ-ускорителями для сравнения результатов бенчмаркинга.

Автоматизированные методики тестирования: Данные методики помогают квалифицировать эффективность сетевой фабрики ИИ с точки зрения времени выполнения задачи (Job Completion Time), изоляции производительности, балансировки нагрузки и механизмов контроля перегрузок.

Решение: KAI Data Center Builder

Keysight AI Data Center Builder (KAI DC Builder) представляет собой надежное решение для оценки и бенчмаркинга, оптимизирующее проектирование систем ИИ/МО. Данное решение быстрее

Тел: +7 (495) 252-00-96

Email: info-site@akmetron.ru

Адрес: Москва, ул. Рабочая, д.93 стр.2, под. 2.

развертывается и эксплуатируется, обеспечивая более глубокий анализ и более высокую точность измерений на уровне отдельных потоков. На Рисунке 1 проиллюстрированы различные уровни центра обработки данных, на которых KAI DC Builder моделирует поведение рабочих нагрузок.

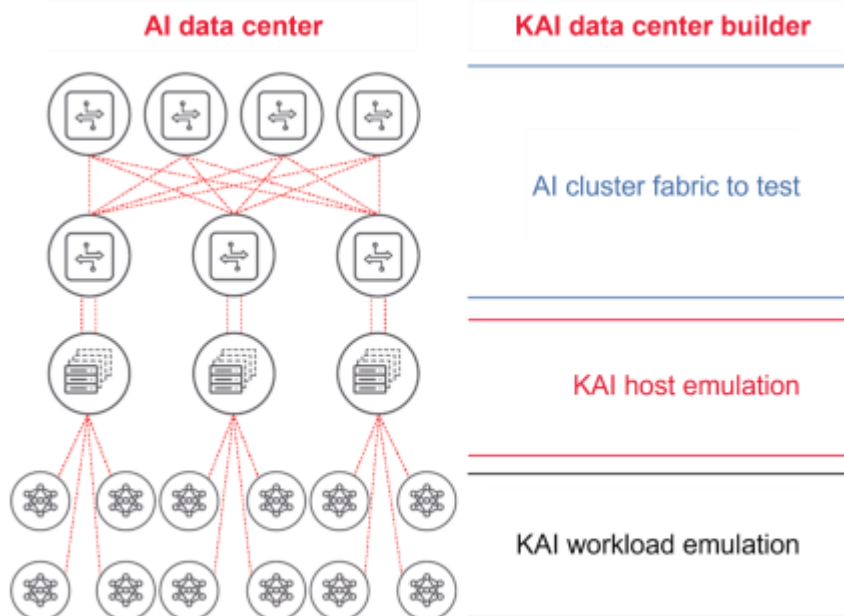


Рисунок 1. KAI DC Builder работает на нескольких уровнях для воспроизведения поведения рабочих нагрузок ИИ

KAI DC Builder позволяет размещать и запускать приложения, специализирующиеся на измерении или анализе различных аспектов инфраструктуры ИИ. Данная возможность обеспечивает детальный анализ производительности, помогает выявлять узкие места и оптимизирует общую эффективность ваших систем ИИ.

Роль приложений заключается в сборе набора данных путем выполнения конкретного испытания (trial). Например, одно приложение может фокусироваться на измерении пропускной способности передачи данных, в то время как другое — на анализе загрузки фабрики. Каждое приложение определяет цель испытания, инициализирует его с помощью специфичной для приложения конфигурации испытания и формирует отчет о результатах на основе собранного набора данных. Взаимодействовать с приложениями KAI DC Builder можно через веб-интерфейс пользователя для визуального, управляемого пользователем рабочего процесса либо с использованием скриптов автоматизации для пакетного выполнения из автоматического конвейера непрерывной интеграции.

На Рисунке 2 показано, что KAI DC Builder располагает двумя приложениями, способными значительно повысить производительность и эффективность вашей инфраструктуры ИИ:

KAI Collective Benchmarks выполняет микро-бенчмаркинг производительности распределенных коммуникационных алгоритмов, широко используемых в масштабируемых системах ИИ и называемых коллективными коммуникациями. Типичная цель использования приложения

Тел: +7 (495) 252-00-96

Email: info-site@akmetron.ru

Адрес: Москва, ул. Рабочая, д.93 стр.2, под. 2.

Collective Benchmarks — убедиться, что сеть обеспечивает оптимальную и стабильную пропускную способность для диапазона обменов данными, ожидаемых от распределенных задач ИИ.

KAI Workload Emulation воспроизводит короткие последовательности шагов, которые реальные рабочие нагрузки ИИ, такие как предварительное или пост-обучение, проходят в течение тысяч и миллионов итераций, называемых эпохами. Эти рабочие нагрузки образуют сочетание вычислительных и коммуникационных шагов, в основном определяемое типом схем партиционирования моделей ИИ для распределения задачи по сотням или тысячам GPU. Приложение Workload Emulation предоставляет возможность экспериментировать с производительностью вариантов партиционирования моделей на различных типах сетевых топологий. Оно помогает найти оптимальное сочетание, которое минимизирует время, затрачиваемое на перемещение данных, и приоритизирует вычисления на GPU, в конечном итоге сокращая время выполнения задачи (Job Completion Time, JCT).

Сочетание этих двух приложений формирует более целостную и автоматизированную среду тестирования, снижая сложность и время, необходимые для настройки и анализа. Пользователи получают более точные аналитические выводы и могут более эффективно оптимизировать свои системы ИИ.

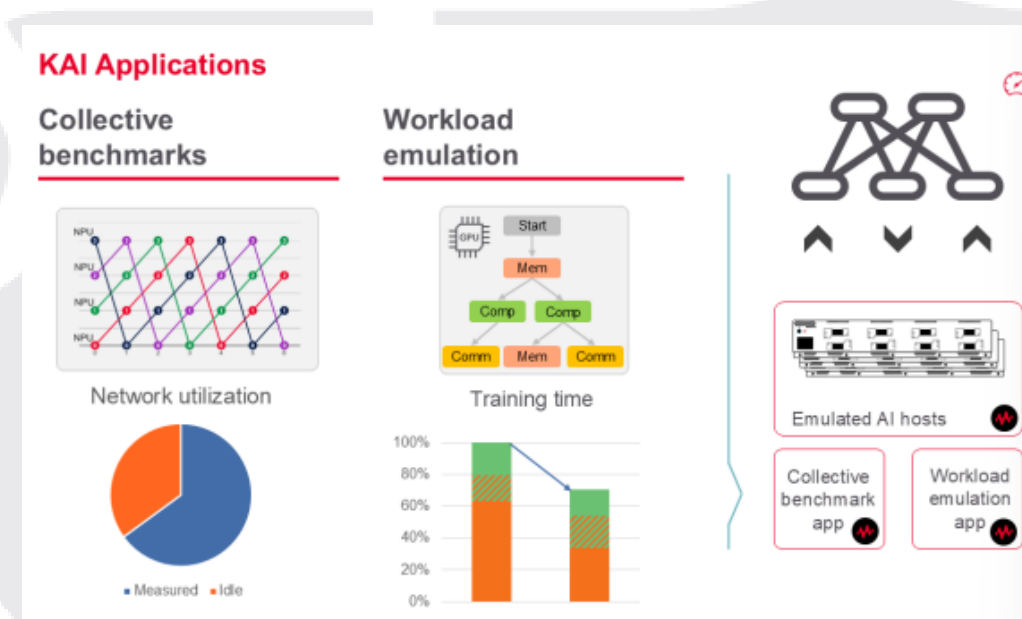


Рисунок 2. Приложения KAI Data Center Builder

Испытательные движки KAI

Приложения KAI могут выполнять испытания с использованием одного из испытательных движков KAI, к которым относятся:

Аппаратное обеспечение Keysight AresONE управляет высокоплотными трафиковыми нагрузками с

Тел: +7 (495) 252-00-96

Email: info-site@akmetron.ru

Адрес: Москва, ул. Рабочая, д.93 стр.2, под. 2.

эмуляцией RoCEv2 для точного воспроизведения коммуникационных шаблонов ИИ/МО в масштабе на тестируемой фабрике. RoCEv2 — транспортный протокол для прямого доступа к памяти (RDMA) поверх Ethernet — используется узлами кластера ИИ для коллективных коммуникаций.

Программное обеспечение конечных точек Keysight работает на серверах общего назначения, оснащенных сетевыми картами RDMA NIC, но без GPU, с тем чтобы включить сетевые карты и их конфигурацию в состав тестируемой системы.

Преимущества программного решения:

- Совместная настройка NIC+Fabric
- Стоимость
- Новые транспортные протоколы
- Production & Cloud

Преимущества аппаратного решения:

- Изолированная валидация фабрики
- Пропускная способность 800G
- Глубокий анализ сетевых показателей

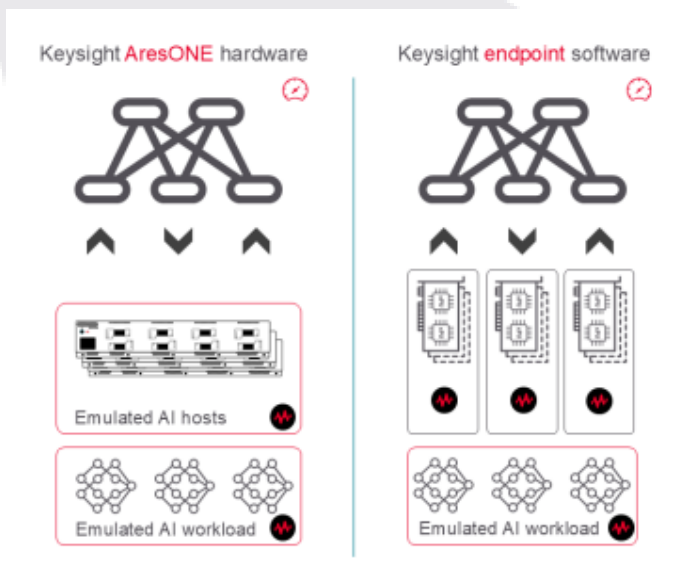


Рисунок 3. Варианты испытательных движков KAI — аппаратные устройства AresONE и программные конечные точки



KAI Collective Benchmarks

Приложение Keysight Collective Benchmark может выполнять микро-бенчмаркинг типовых алгоритмов коммуникаций ИИ на предоставленной пользователем сетевой фабрике ИИ, как

Тел: +7 (495) 252-00-96

Email: info-site@akmetron.ru

Адрес: Москва, ул. Рабочая, д.93 стр.2, под. 2.

показано на Рисунке 4. Вы можете измерять производительность фабрики и вносить улучшения в проект в лабораторной или промежуточной среде без подключения вычислительных узлов ИИ с GPU-ускорителями к тестируемой фабрике за счет использования эмуляции трафика RoCEv2 с аппаратным обеспечением AresONE.

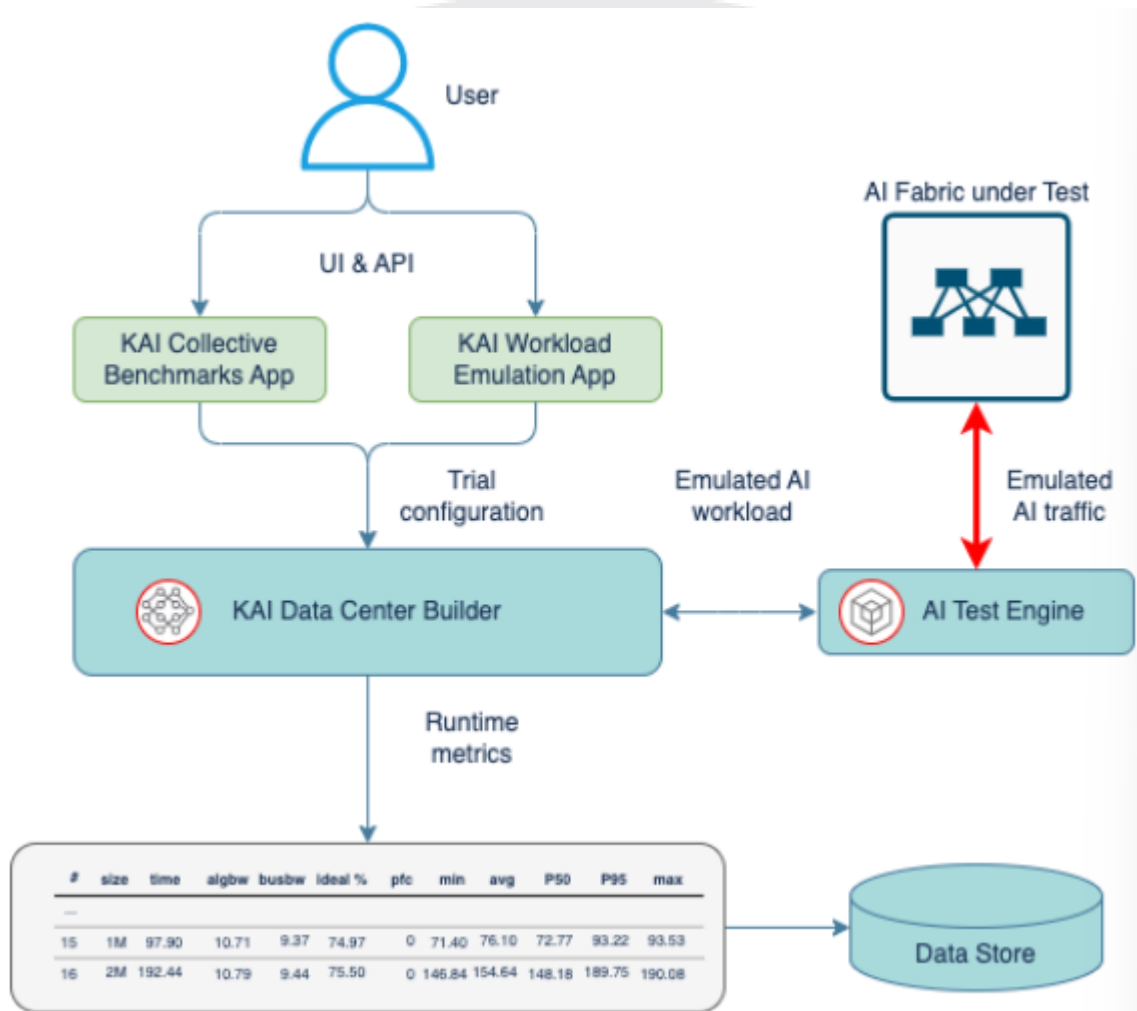


Рисунок 4. Рабочий процесс приложения KAI Collective Benchmarks с аппаратным обеспечением AresONE

Параметры конфигурации

Вы можете параметризовать конфигурацию испытания для KAI Collective Benchmarks, чтобы отразить размер и свойства коллектива ИИ, на котором выполняется бенчмаркинг, а также профиль рабочей нагрузки ИИ, который будут эмулировать генераторы трафиковой нагрузки.

Параметр	Значения	Описание
Конфигурация логического кластера	<ul style="list-style-type: none"> ·Количество GPU ·GPU на вычислительный узел ·Интерконнект NPU 	Описывает конфигурацию системы ИИ для эмуляции рабочей нагрузки
Коллективная операция: тип	<ul style="list-style-type: none"> ·AlltoAll ·AllReduce ·AllGather ·ReduceScatter ·Gather ·Broadcast 	Тип коллективной операции
Коллективная операция: алгоритм	<ul style="list-style-type: none"> ·Parallel ·Unidirectional Ring ·Bidirectional Ring ·Halving-Double 	Алгоритмы коллективных операций: применимость алгоритмов зависит от типа операции
Предустановка рабочей нагрузки: размер данных	<ul style="list-style-type: none"> ·Начальное значение ·Значение шага итерации ·Конечное значение 	Диапазон размеров данных для выполнения бенчмаркинга
Предустановка рабочей нагрузки: транспортный профиль	<ul style="list-style-type: none"> ·Транспортный протокол ·Количество Q-пар на пару рангов ·Размер сообщения RDMA 	Характеристики транспортного профиля; используется для точной настройки на уровне Q-Pair
Предустановка инфраструктуры: профиль NIC	<ul style="list-style-type: none"> ·Скорость интерфейса NIC ·MTU NIC ·Настройки IP-интерфейса NIC ·Значение класса трафика / DSCP ·Параметры PFC ·Параметры DCQCN 	Конфигурация контроля перегрузок и классификации трафика для эмулируемых NIC

Измеряемые метрики

Приложение KAI Collective Benchmarks формирует отчет об измерениях в формате, общепринятом среди специалистов по ИИ. На основе конкретных параметров коллектива ИИ и рабочей нагрузки оно собирает набор метрических данных и представляет их в табличном формате. Каждая строка представляет измерения для определенного размера данных, которыми коллектив должен обмениваться.

Как показано на Рисунке 5, путем итерации по настраиваемому диапазону размеров данных приложение KAI Collective Benchmarks выполняет бенчмаркинг производительности фабрики ИИ, используя комбинацию ключевых показателей эффективности (KPI), специфичных для коллективных коммуникаций, таких как время завершения, алгоритм и пропускная способность

Тел: +7 (495) 252-00-96

Email: info-site@akmetron.ru

Адрес: Москва, ул. Рабочая, д.93 стр.2, под. 2.

шины, с дополнительным анализом распределения статистических метрик по отдельным парам очередей RoCEv2 (Q Pairs).

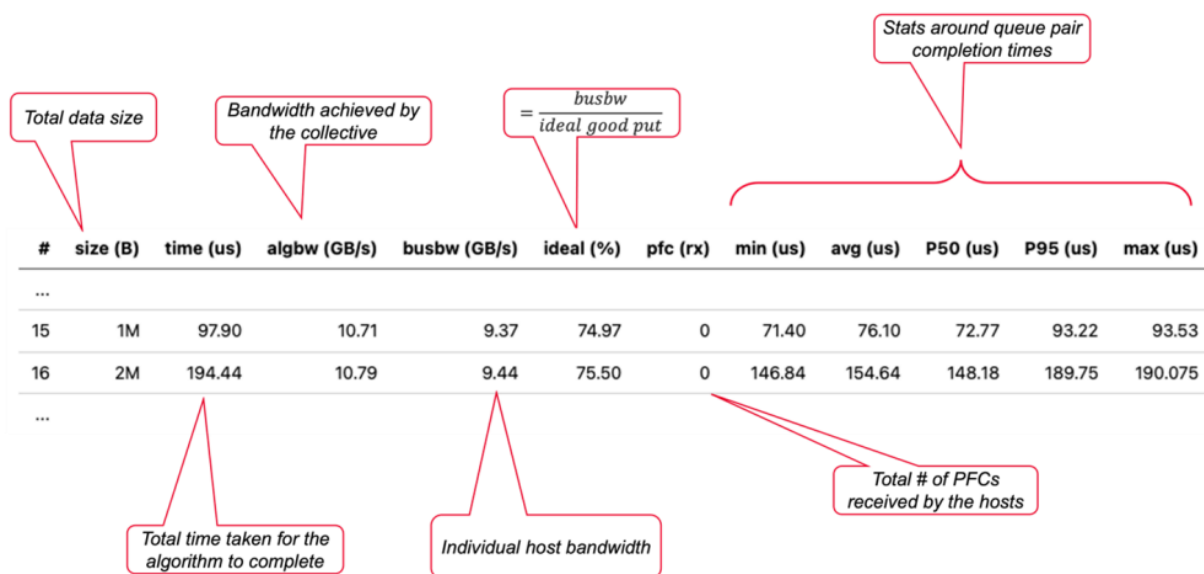


Рисунок 5. Пример измеряемых метрик

Trial Summary											
	Data Size	Collective	Completion Time (ms)	Algbw (GB/s)	Busbw (GB/s)	Ideal (%)	Min FCT (ms)	P50 FCT (ms)	P95 FCT (ms)	Max FCT (ms)	
0	16.00 MB	ALL_REDUCE-1	6.58	2.55	4.46	9.10	0.05	0.05	0.05	0.05	
1	32.00 MB	ALL_REDUCE-2	7.20	4.66	8.16	16.65	0.09	0.09	0.09	0.09	
2	64.00 MB	ALL_REDUCE-3	11.44	5.87	10.27	20.95	0.17	0.18	0.29	0.33	
3	128.00 MB	ALL_REDUCE-4	10.99	12.21	21.37	43.60	0.35	0.35	0.39	0.40	
4	256.00 MB	ALL_REDUCE-5	16.51	16.26	28.45	58.04	0.69	0.70	0.95	1.12	
5	512.00 MB	ALL_REDUCE-6	26.16	20.52	35.91	73.27	1.39	1.40	1.62	1.85	
6	1,024.00 MB	ALL_REDUCE-7	46.03	23.33	40.83	83.29	2.77	2.79	3.03	3.16	
7	2,048.00 MB	ALL_REDUCE-8	84.70	25.36	44.37	90.52	5.54	5.57	5.81	6.00	
8	4,096.00 MB	ALL_REDUCE-9	163.36	26.29	46.01	93.86	11.08	11.16	11.42	11.55	
9	8,192.00 MB	ALL_REDUCE-10	318.55	26.97	47.19	96.27	22.15	22.29	22.58	22.85	

Рисунок 6. Пример сводного отчета об испытании KAI Collective Benchmarks

Эмуляция рабочих нагрузок KAI

Операторы ИИ используют различные стратегии параллелизма, также называемые партиционированием моделей, для ускорения обучения моделей ИИ. Практические демонстрации показывают, что согласование партиционирования с топологией и конфигурацией кластера ИИ обеспечивает исключительную производительность обучения. На этапе проектирования кластера ИИ критические вопросы наилучшим образом решаются путем экспериментирования. Например, многие вопросы касаются эффективности перемещения данных между графическими процессорами (GPU), в том числе:

- Проектирование вертикального масштабирования (scale-up) интерконнектов GPU внутри хоста ИИ

Тел: +7 (495) 252-00-96

Email: info-site@akmetron.ru

Адрес: Москва, ул. Рабочая, д.93 стр.2, под. 2.

или стойки.

- Проектирование горизонтального масштабирования (scale-out) сети, включая пропускную способность на GPU и топологию.
- Настройка балансировки нагрузки сети и контроля перегрузок.
- Настройка параметров фреймворка обучения для оптимизации производительности и эффективности.

Приложение KAI Workload Emulation воспроизводит сетевые коммуникационные шаблоны реальных задач обучения ИИ. Его целями являются ускорение экспериментирования, снижение порога входа, необходимого для достижения профессионального уровня, и предоставление более глубокого анализа причин наблюдаемого снижения производительности. Получение подобных аналитических данных было бы затруднительно при использовании реальной задачи обучения ИИ для экспериментов.

Запуск KAI Workload Emulation в KAI DC Builder позволяет:

- Экспериментировать с параметрами партиционирования модели и планированием рангов на инфраструктуре ИИ.
- Понимать вклад коммуникаций внутри и между партициями в общее время выполнения задачи (JCT).
- Выявлять типы коллективных операций, демонстрирующих низкую производительность, и детализировать анализ для обнаружения узких мест.
- Использовать данные об утилизации сети, хвостовой задержке, перегрузке и их влиянии на JCT.

Библиотека рабочих нагрузок

KAI DC Builder включает библиотеку рабочих нагрузок ИИ — от классификации изображений до больших языковых моделей (LLM), таких как Llama, как показано на Рисунке 8. Он предлагает набор популярных схем партиционирования моделей, включая параллелизм данных (DP), полностью сегментированный параллелизм данных (FSDP) и трехмерный (3D) параллелизм. Вы можете импортировать трассировки выполнения ваших рабочих нагрузок в библиотеку с использованием формата MLCommons Chakra.

На Рисунке 7 проиллюстрированы рабочие нагрузки в библиотеке, подготовленной Keysight, которая содержит подробные метаданные, документирующие среду и ключевые гиперпараметры, использованные для снятия их трассировок выполнения. Библиотека предоставляет возможности для исследования поведения рабочих нагрузок и понимания типов и временных характеристик коллективных операций, а также вычислений.

Тел: +7 (495) 252-00-96

Email: info-site@akmetron.ru

Адрес: Москва, ул. Рабочая, д.93 стр.2, под. 2.

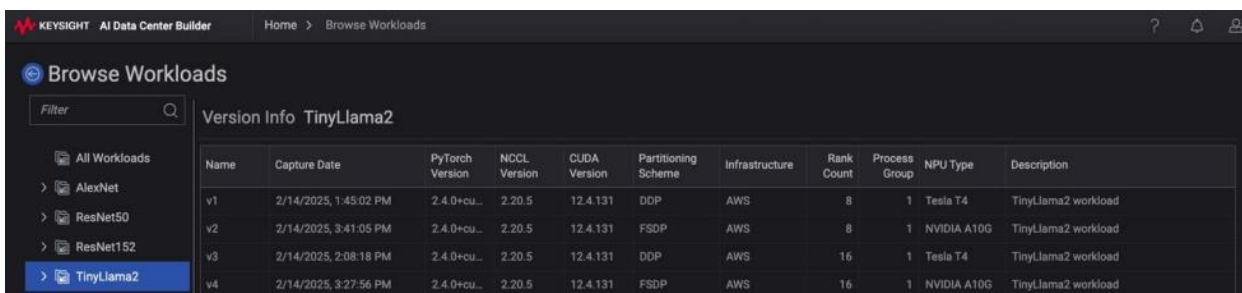


Рисунок 7. Библиотека рабочих нагрузок KAI Data Center Builder

Каждая рабочая нагрузка может содержать трассировки, собранные с каждого отдельного ранга (NPU), которые фиксируют шаги, пройденные задачей на данном вычислительном узле, как показано на Рисунке 8. За исключением простейших случаев, таких как DDP, поведение различных рангов может различаться в зависимости от их положения в партиции.

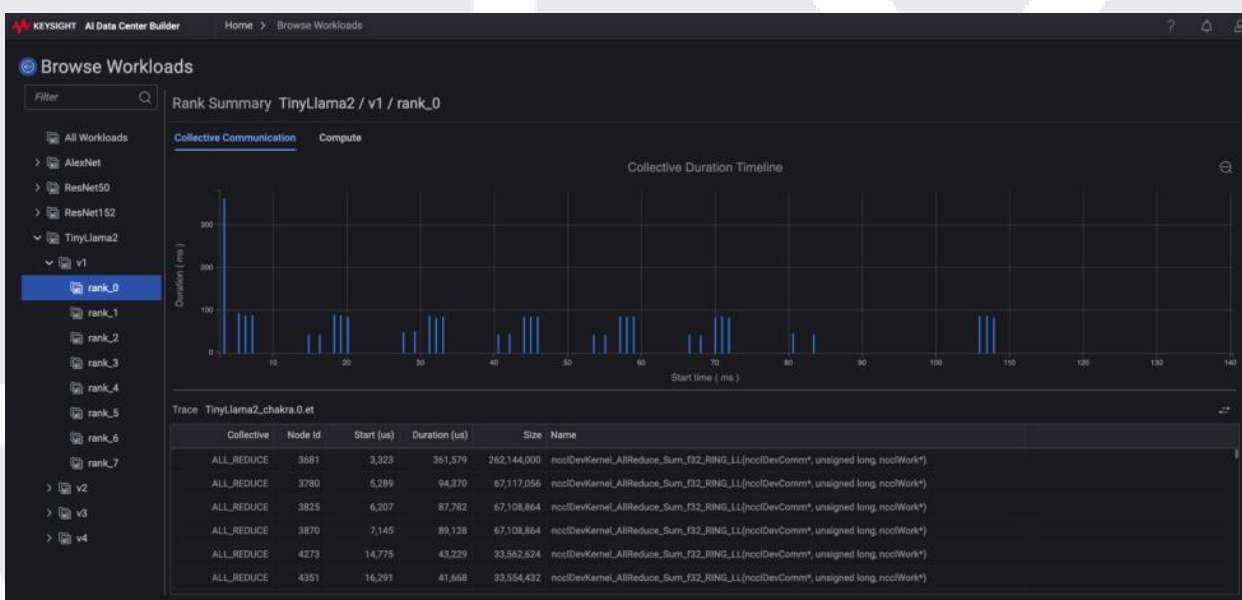


Рисунок 8. Детализация рабочей нагрузки: временная шкала длительности коллективных операций

Параметры конфигурации — для Workload Emulation

После выбора трассировок рабочих нагрузок из библиотеки пользователь назначает их набору эмулируемых NPU (рангов) для воспроизведения. Как и в приложении KAI Collective Benchmarking, в конфигурации имеется раздел, определяющий конфигурацию логического кластера путем указания количества хостов ИИ, количества NPU на хост и наличия интерконнектов NPU. Ваш испытательный стенд должен располагать достаточными ресурсами эмуляции (аппаратные тестовые порты или RDMA NIC) для поддержки количества элементов, определенного в разделе логической инфраструктуры, как показано на Рисунках 9, 10 и 11.

Тел: +7 (495) 252-00-96

Email: info-site@akmetron.ru

Адрес: Москва, ул. Рабочая, д.93 стр.2, под. 2.

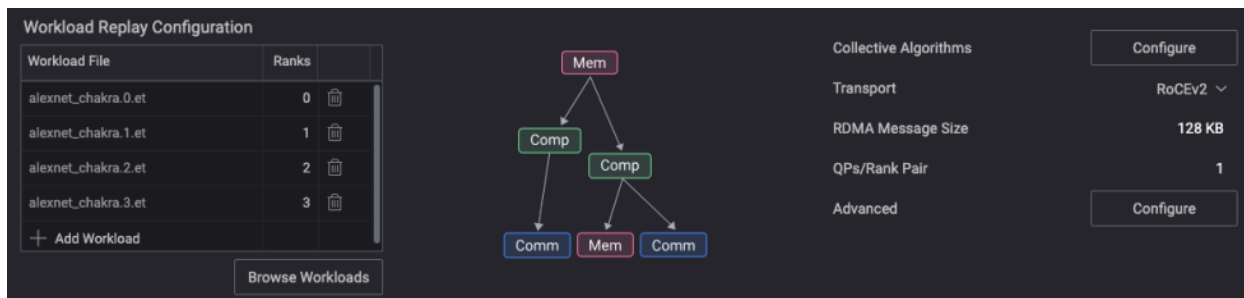


Рисунок 9. Планирование рабочих нагрузок

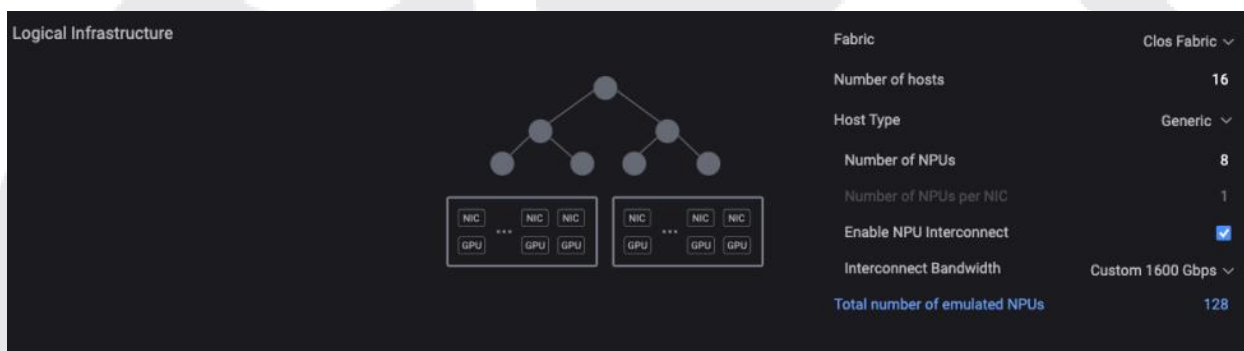


Рисунок 10. Моделирование инфраструктуры ИИ с восемью стойками, 16 хостами ИИ и восемью GPU

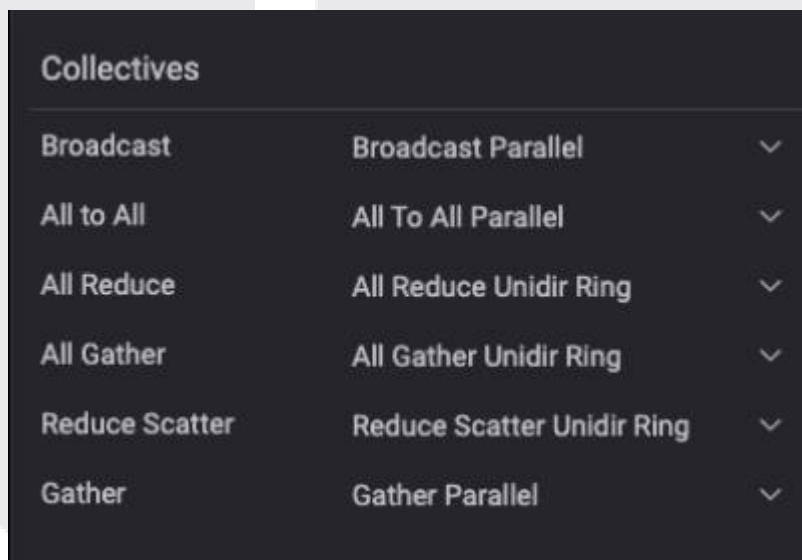


Рисунок 11. Параметризация коллективной операции

Измеряемые метрики — для Workload Emulation

После выполнения испытания Workload Emulation приложение выдает отчет об общем времени

Тел: +7 (495) 252-00-96
 Email: info-site@akmetron.ru
 Адрес: Москва, ул. Рабочая, д.93 стр.2, под. 2.

завершения для каждой трассировки рабочей нагрузки, суммируя время для всех рангов, на которых она была запланирована. Дополнительно приложение предоставляет данные о производительности каждой коллективной операции для всех рабочих нагрузок в испытании, включая такие детали, как характеристики операции, время завершения, относительное время начала и информация о партиции модели, к которой относится операция, как показано в Таблице 1.

№	Workload ID	Collective ID	Размер данных (ГБ)	ССТ (с)	Ранги	Начало (с)	Группа РР	Группа ТР	Группа ДР
1	LLM_3DP_1	AllGather_1	0,125	0,32	8	3,56	-	1	-
2	LLM_3DP_1	AllGather_2	0,125	0,29	8	4,12	-	2	-
3	LLM_3DP_1	AllReduce_1	2,0	1,34	16	5,43	-	-	1

Таблица 1. Сводка измерений эмуляции рабочих нагрузок

Резюме

Keysight AI Data Center Builder обеспечивает сочетание высокоточной эмуляции рабочих нагрузок ИИ, готовых бенчмарк-приложений и инструментов анализа наборов данных для значительного улучшения бенчмарк-производительности сетевой фабрики кластера ИИ/МО.

Для ускорения проектирования сетей ИИ/МО Keysight AI Data Center Builder:

- Эмулирует масштабные рабочие нагрузки ИИ с измеримой достоверностью — обеспечивает глубокий анализ производительности коллективных коммуникаций.
- Упрощает процесс бенчмаркинга — валидирует сетевую фабрику ИИ с помощью готовых бенчмарк-приложений, созданных в партнерстве с крупнейшими операторами ИИ и поставщиками инфраструктуры ИИ.
- Выполняет заданные поведенческие модели ИИ/МО — обеспечивает обмен между пользователями и заказчиками для воспроизводимости экспериментов.
- Предлагает выбор испытательных движков — выбор между эмуляцией конечных точек RoCEv2 на высокоплотных генераторах трафиковой нагрузки AresONE и программными конечными точками на реальных ИИ-ускорителях для сравнения результатов бенчмаркинга.

KAI DC Builder обеспечивает масштабную валидацию и экспериментирование с проектированием фабрики с использованием тестового оборудования AresONE-S 400GE и AresONE-M 800GE реалистично и экономически эффективно. Данное решение дополняет использование GPU для тестирования рабочих нагрузок ИИ/МО, позволяя операторам ИИ сократить расходы, которые они были бы вынуждены полностью направить на системы бенчмаркинга на базе GPU, в пользу более

Тел: +7 (495) 252-00-96

Email: info-site@akmetron.ru

Адрес: Москва, ул. Рабочая, д.93 стр.2, под. 2.

масштабируемого, надежного и интегрированного решения Keysight AI Data Center Builder.

Узнать больше

Ознакомьтесь с техническим описанием Keysight AI Fabric Test Solution для получения подробных сведений о поддерживаемых параметрах эмуляции трафика RoCEv2, контроля перегрузок и измерений.

Узнайте больше о: KAI Data Center Builder и решениях Keysight AI.
Остались вопросы? Обратитесь к нашим продуктовым экспертам.

Keysight помогает инноваторам расширять границы инженерного искусства, быстро решая задачи проектирования, эмуляции и тестирования для создания наилучших потребительских решений. Начните свой инновационный путь на www.keysight.com.

Данная информация может быть изменена без предварительного уведомления. © Keysight Technologies, 2023 – 2025. Опубликовано в США, 2 апреля 2025 г., 3123-1809.EN

Тел: +7 (495) 252-00-96

Email: info-site@akmetron.ru

Адрес: Москва, ул. Рабочая, д.93 стр.2, под. 2.