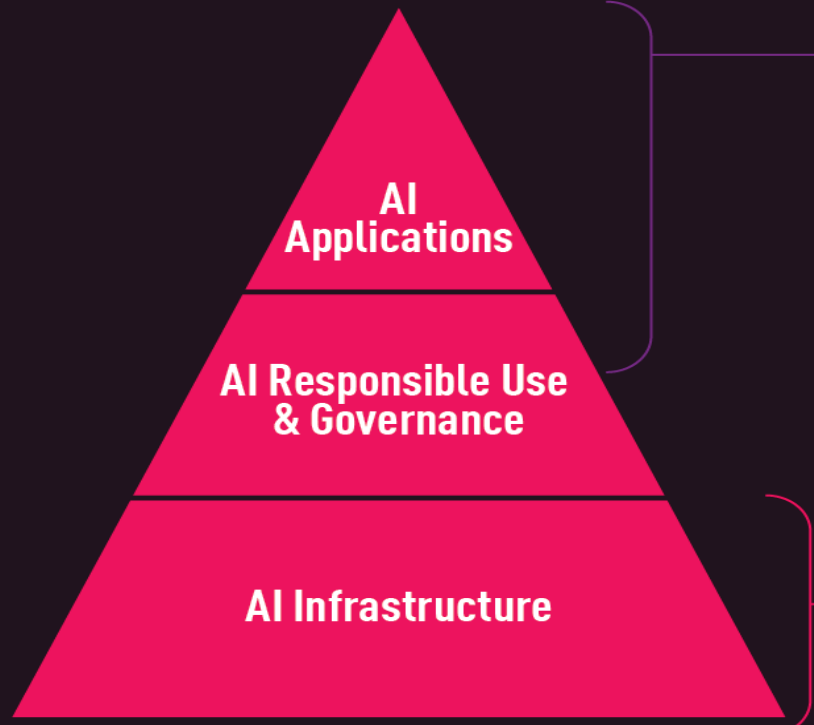
CHECK POINT™

Защитите каждое взаимодействие с ИИ от сотрудников до агентов



Комплексная защита AI-фабрик и дата-центров
С использованием многослойной AI-безопасности



1. AI Native Application Security
AI Security + WAF | Prompt Defense + API security

2. Perimeter & Network Security
AI Firewall + Hyperscale | Zero Trust + Prompt Inspection

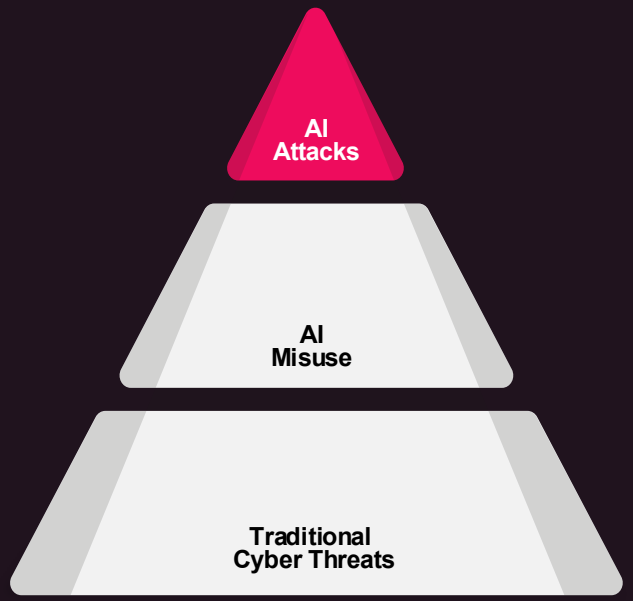
3. Distributed Host Level Security
Check Point AI Factory Firewall + NVIDIA BlueField-3 DPU

4. Hardware Accelerated AI Security Integration
NVIDIA DOCA Argus + Check Point ThreatCloud

5. Workload & Container Security
Microsegmentation | Kubernetes Runtime Security

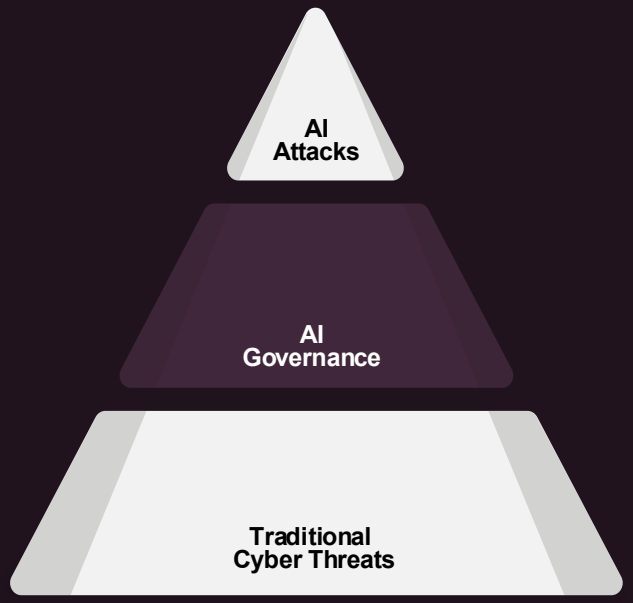
OWASP GenAI SECURITY PROJECT
MITRE ATT&CK
MITRE ATLAS
NIST SSDF
ISO
ISO 42001:2023
NIST AI RMF 1.0
EU AI Act

Типы рисков и примеры — атаки на ИИ



Тип риска	Пример	Влияние на бизнес
Prompt Injection **LLM01 - ASI01**	Внедрённый контент заставляет GenAI/агента экспортировать записи или одобрять действия	Нарушение безопасности, мошенничество, регуляторные риски
Sensitive Info Disclosure **LLM06 - ISO27091**	RAG возвращает конфиденциальные контракты или персональные данные клиентов (PII)	Штрафы, утрата интеллектуальной собственности, расходы на уведомления
Insecure Output Handling **LLM02**	SQL/JS, сгенерированные LLM, приводят к удалённому выполнению кода (RCE), XSS или манипуляции данными	Компрометация системы, простой
Training Data / Model Poisoning **LLM03 - ATLAS**	Фраза-триггер вызывает обход политики или целенаправленное распространение дезинформации	Инцидент безопасности, ущерб репутации
Supply Chain & Plugins **LLM05/07 - ASI04**	Троянский плагин/модель крадёт ключи или извлекает данные.	Компрометация всего парка устройств, расходы на реагирование на инциденты (IR)
Model DoS / Cost Abuse **LLM04**	Противник исчерпывает квоту → нарушаются SLA и растут счета.	Потеря доходов, отток клиентов
Model Theft / Extraction **LLM10/ATLAS**	Конкурент клонирует модель с помощью API-скрейпинга	Кража интеллектуальной собственности, снижение прибыли

Типы рисков и примеры — злоупотребление ИИ



Тип риска	Пример	Влияние на бизнес
Hallucination (Validity & Reliability)	Вымышленные данные о политике или финансах используются при принятии решений	Юридическая ответственность, финансовые потери
Toxicity / Bias (Fairness)	Процесс найма/отбора непреднамеренно отдаёт предпочтение одной группе	Иски о дискриминации, ущерб бренду
Relevance / Context Failure	Бот поддержки отвечает информацией о нерелевантных продуктах	Отток клиентов, снижение производительности
Accuracy / Overreliance	Аналитик/клиницист принимает небезопасную рекомендацию	Инцидент безопасности, меры со стороны регуляторов



NIST AI RMF 1.0

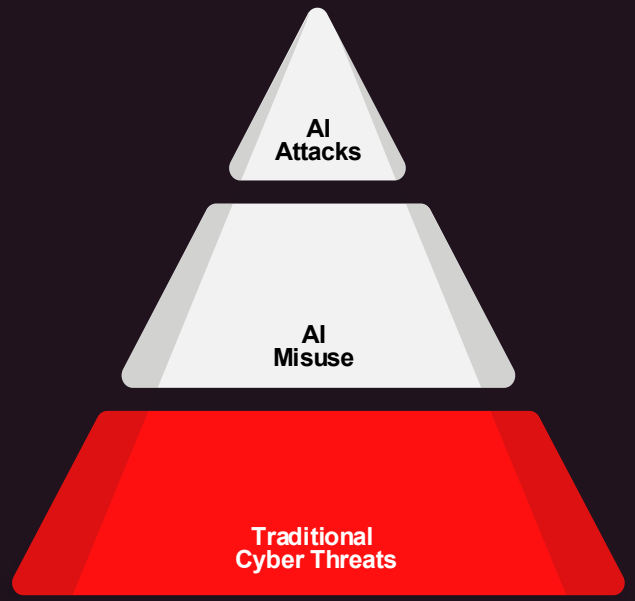


EU AI Act



ISO 42001:2023

Типы рисков и примеры — инфраструктура



Тип риска	Пример	Влияние на бизнес
Ransomware	Обучение/вывод недоступны; требуется восстановление и реконструкция	Простой системы, влияние на доход
Exposed Mgmt Plane	Захват кластера; выполнение вредоносных заданий	Потеря данных, сбои по соблюдению требований
CI/CD Pipeline Compromise	Скомпрометированные (с «закладкой») изображения отправлены в продуктивную среду	Инцидент в цепочке поставок, ущерб клиентам
Lateral Movement	Доступ к наборам данных и моделям по всему кластеру	Утрата интеллектуальной собственности, простой системы
Kubernetes Control/Runtime	Вредоносные поды на GPU-узлах	Нарушение работы сервиса, компрометация
GPU/NIC Firmware or Driver Attack	Доступ к памяти между VM/GPU.	Утечка данных, влияние на целостность
Identity & Privilege Abuse	Несанкционированный доступ к критически важным активам	Мошенничество, риск нарушения соответствия требованиям
Runtime Injection (Container/DLL)	Экспроприация ключей/учётных данных; сохранение доступа (persistence)	Нарушение безопасности, перебои в работе бизнеса

Полноценная платформа для безопасности ИИ

LAKERA

Lakera Red

Противодействие и тестирование GenAI на основе оценки рисков

- Выявлять противоречивые крайние случаи
- Обнаруживать слабые защитные меры и ограничения
- Предоставлять результаты для устранения проблем и аудитов

GenAI Protect

Выявлять, оценивать и управлять рисками при использовании ИИ сотрудниками

- Выявлять скрытое использование ИИ
- Оценивать риски с разбором активности AI
- Применять политики для предотвращения потери данных
- Отслеживать внедрение AI с течением времени

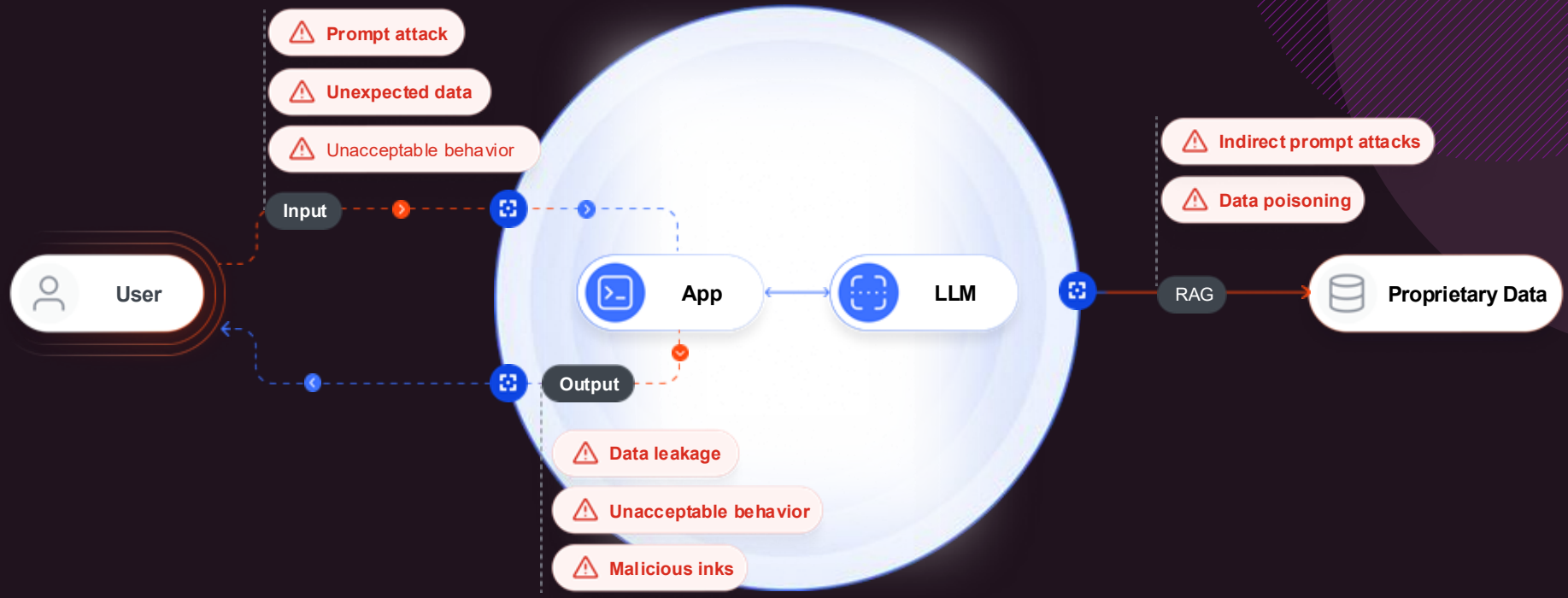
LAKERA

Lakera Guard

Мониторинг и защита AI-приложений и агентов в режиме реального времени.

- Обнаруживать и предотвращать обходы защиты (jailbreak)
- Выявлять вредоносных участников и заражённый контент
- Применять защитные ограничения и политики

Выявление и снижение рисков с постоянно эволюционирующей защитой



Check Point является **предпочтительным партнёром** на каждом этапе вашего пути **внедрения AI-приложений**

Обнаружение

Выявляйте **скрытые риски AI и его использование** в вашей организации

Создание

Выявляйте уязвимости и получайте **практические рекомендации по безопасности до развертывания**

Эксплуатация

Защита в реальном времени с передовой точностью и минимальным числом ложных срабатываний

Масштабирование

Должно быть масштабируемым для обработки **значительного роста** в соответствии с прогнозируемым увеличением

LAKERA

Lakera поддерживает усилия по соблюдению требований и нормативов с помощью:

OWASP | EU Artificial Intelligence Act | NIST

AI Factory Firewall

CHECK POINT

Чем отличаются AI-фабрики?

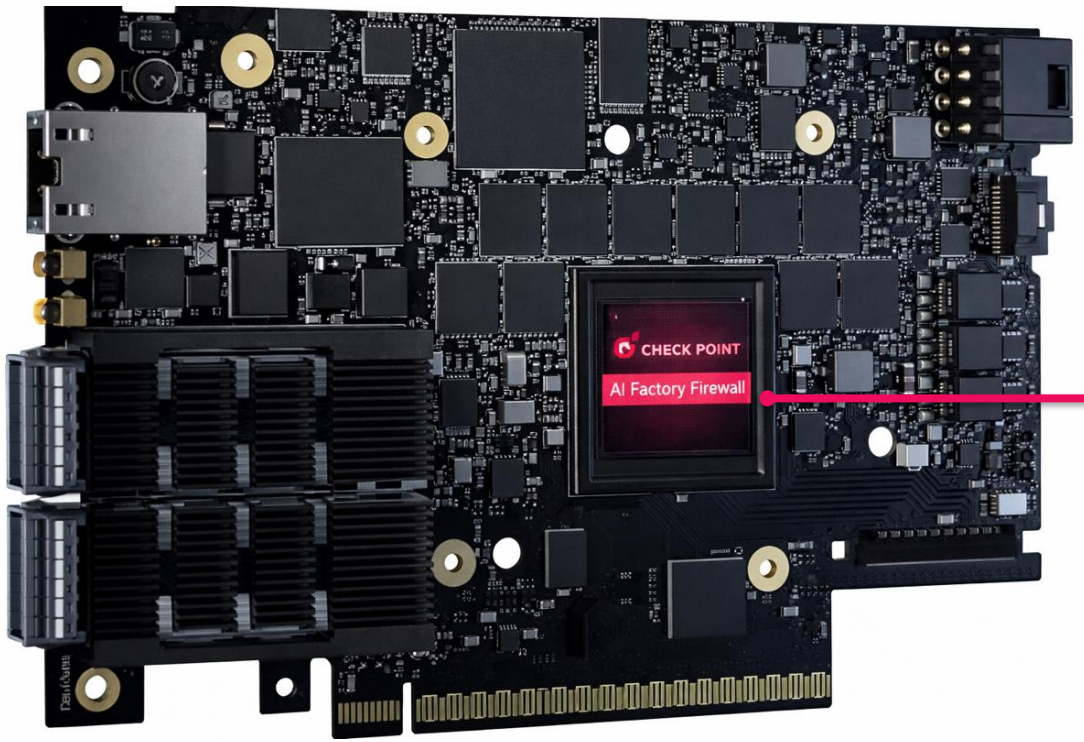
Почему безопасность AI-инфраструктуры особенно сложна

1. **Длительно выполняющиеся нагрузки**
   AI-тренировки и inference-задачи могут длиться неделями или месяцами
2. **Постоянная загрузка на уровне 100%**
   Работает как болид Формулы-1 — на пределе возможностей и без запаса ресурсов
3. **Открытость среды по замыслу**
   AI-среды целенаправленно открыты для разработчиков и исследователей



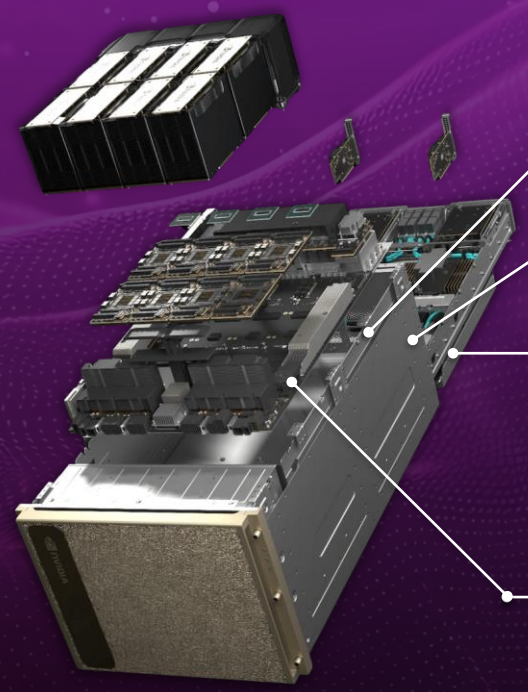
CHECK POINT

Безопасность, встроенная в GPU-кластеры NVIDIA®



CHECK POINT™

AI Factory
Firewall Container



BlueField-3/4 DPU
Check Point AI Cloud Protect

BlueField-3/4 SuperNIC

CPUs

GPUs

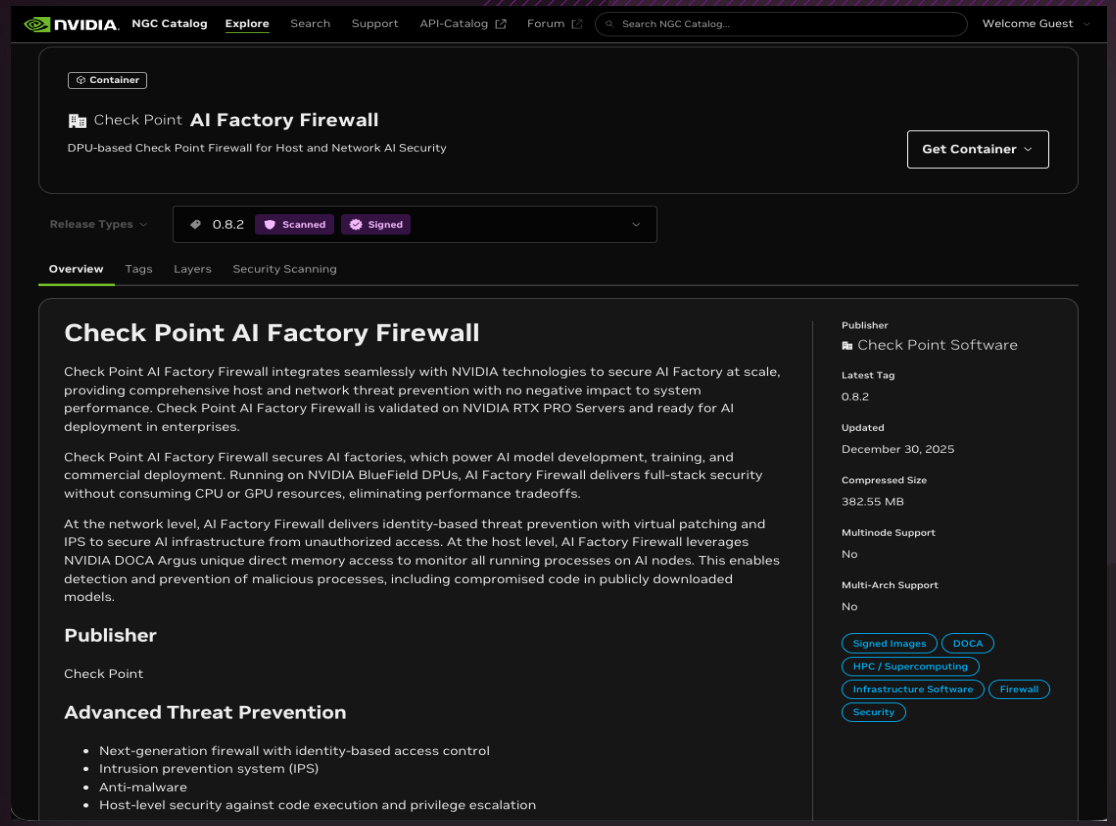
LARGE LANGUAGE MODEL (LLM)



DGX/HGX

AI Factory Firewall
Контейнерный Firewall, работающий на DPU

- Network Threat Prevention
- Host & OS Security (Via ARGUS)
- AI training and Inference Security
- Zero Performance Impact

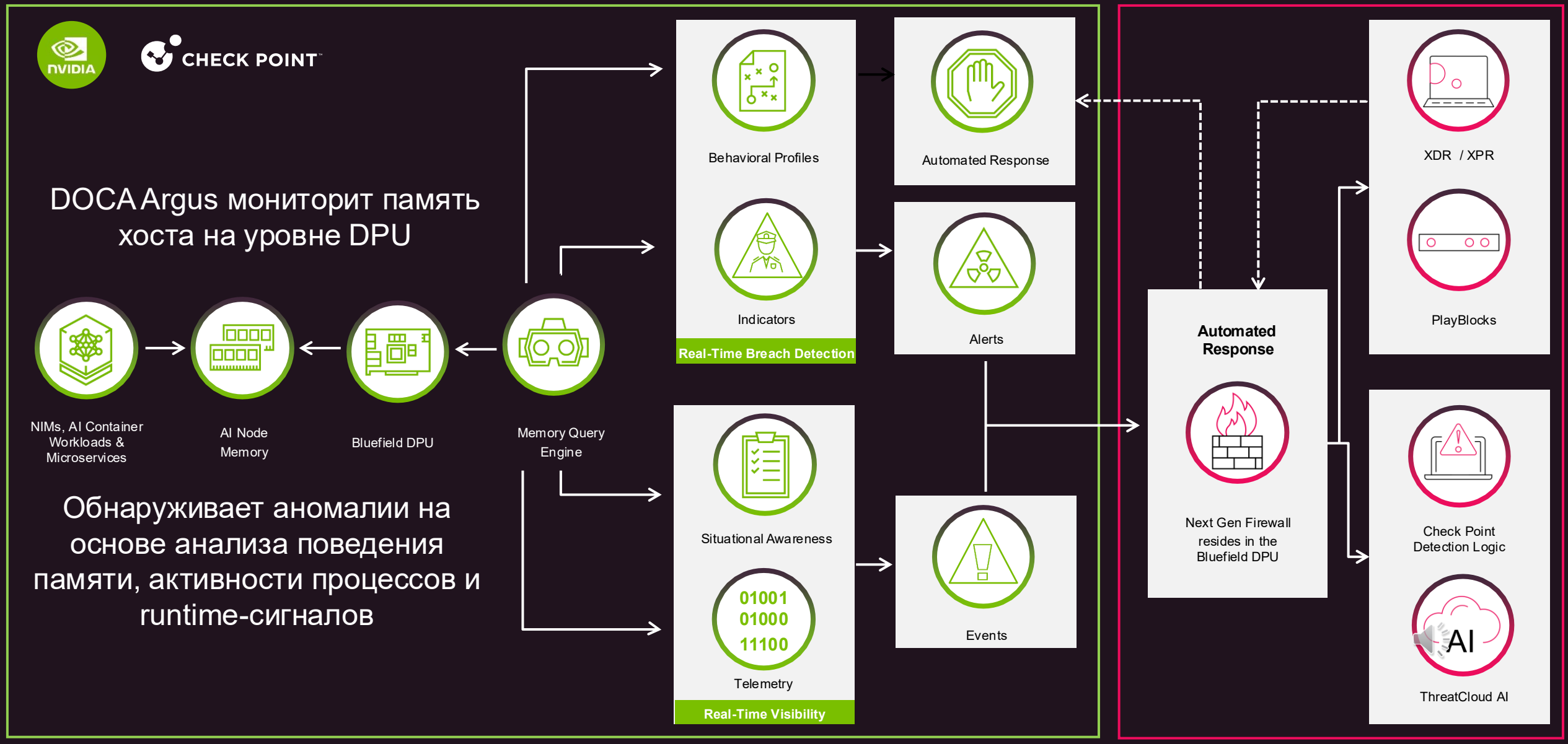


Интегрировано с NVIDIA DSX Air
Облачная среда моделирования

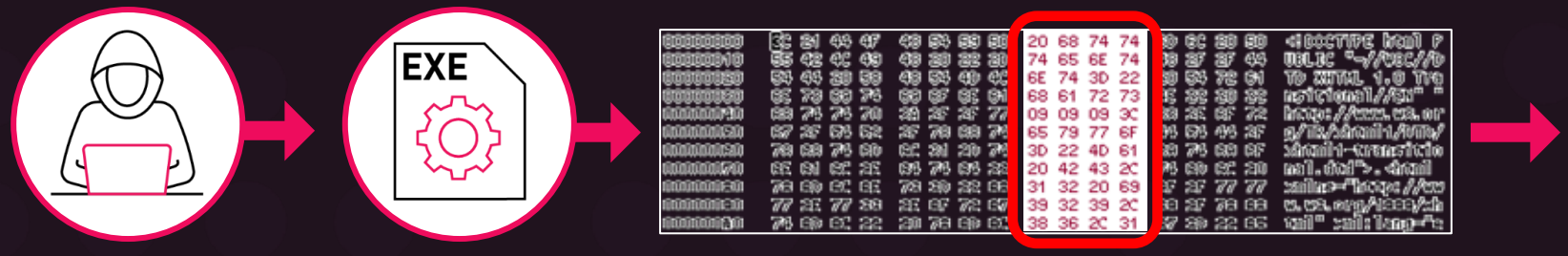
CHECK POINT

Комплексная защита на уровне хоста

DOCA Argus Workload Threat Detection



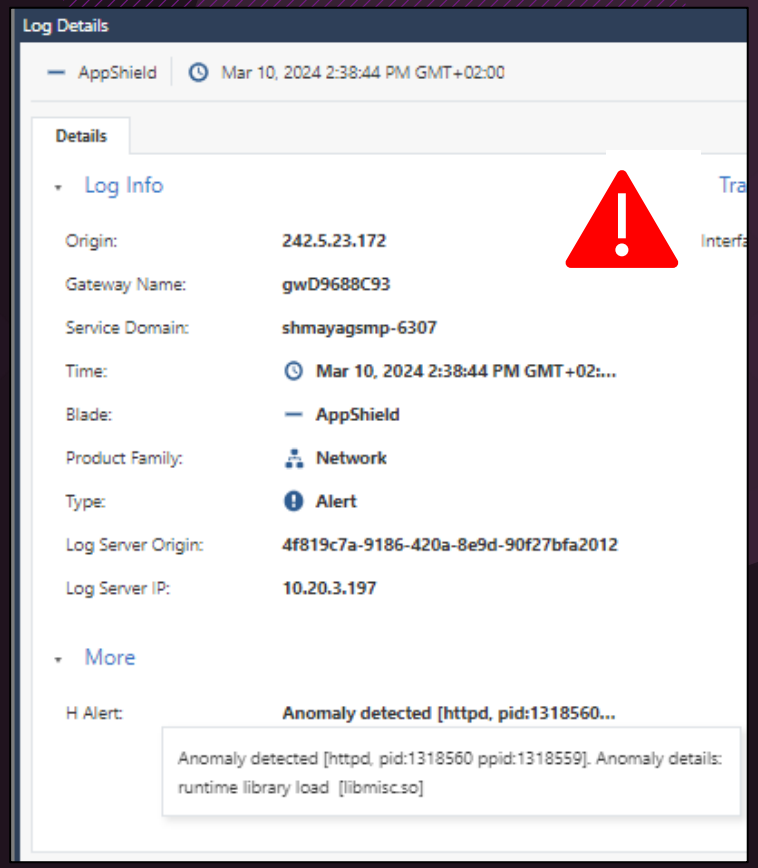
Комплексная защита на уровне хоста



Атакующий

httpd

Memory Overrun

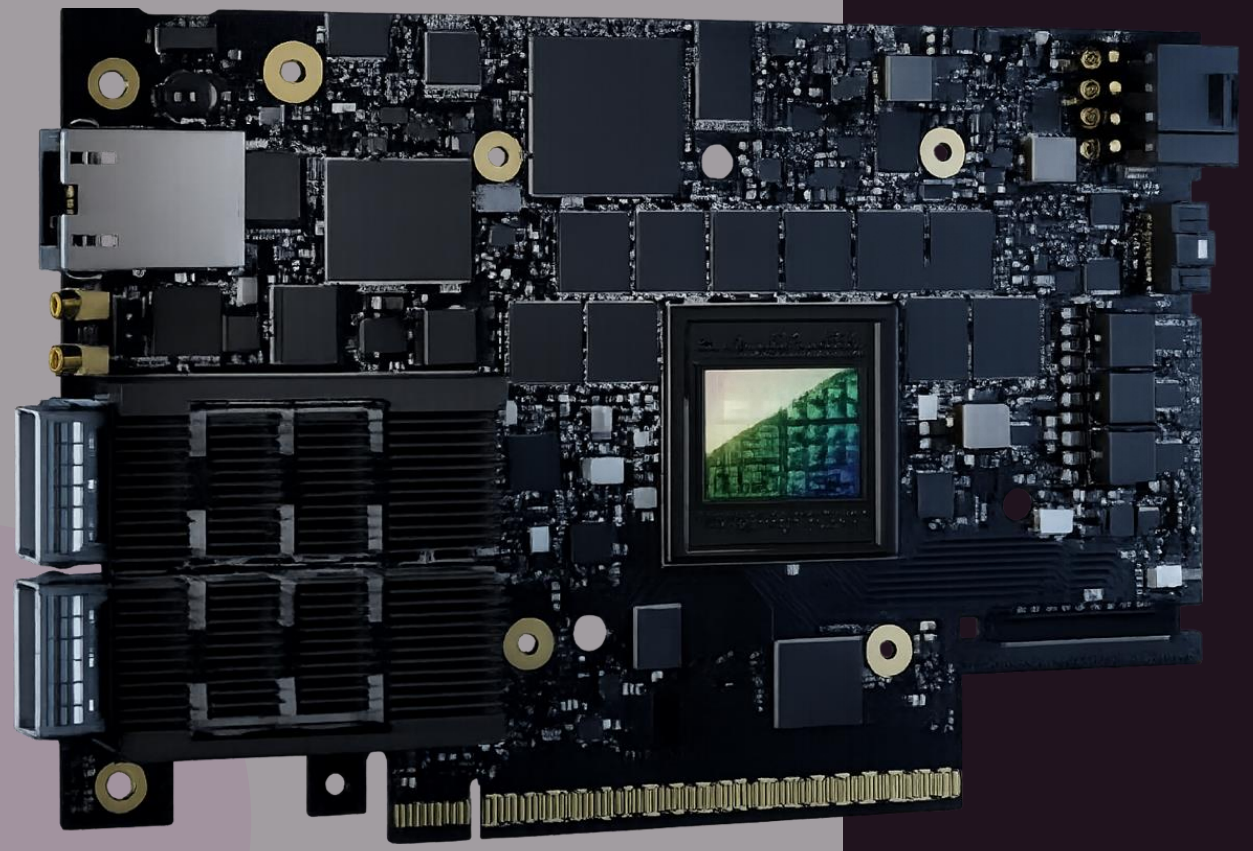


Уникальная защита процессов и приложений на уровне хоста, предотвращающая выполнение кода и повышение привилегий

CHECK POINT

Zero Impact on AI Performance

Check Point Firewall Работает на Nvidia BlueField-3/4 DPU

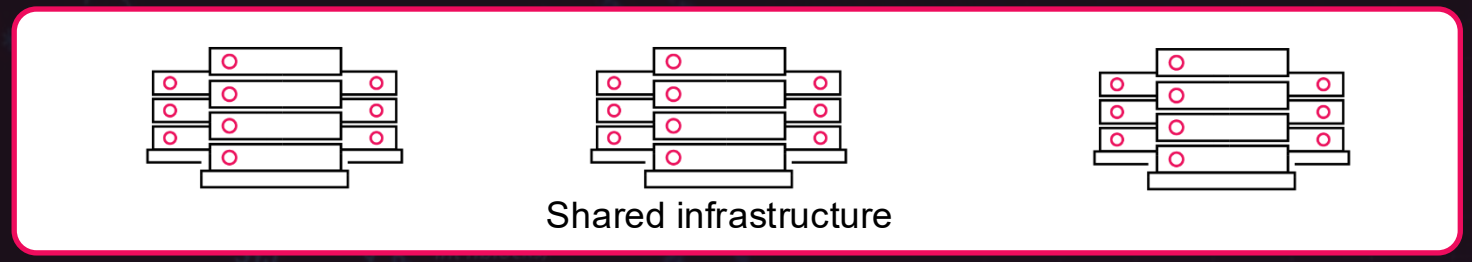


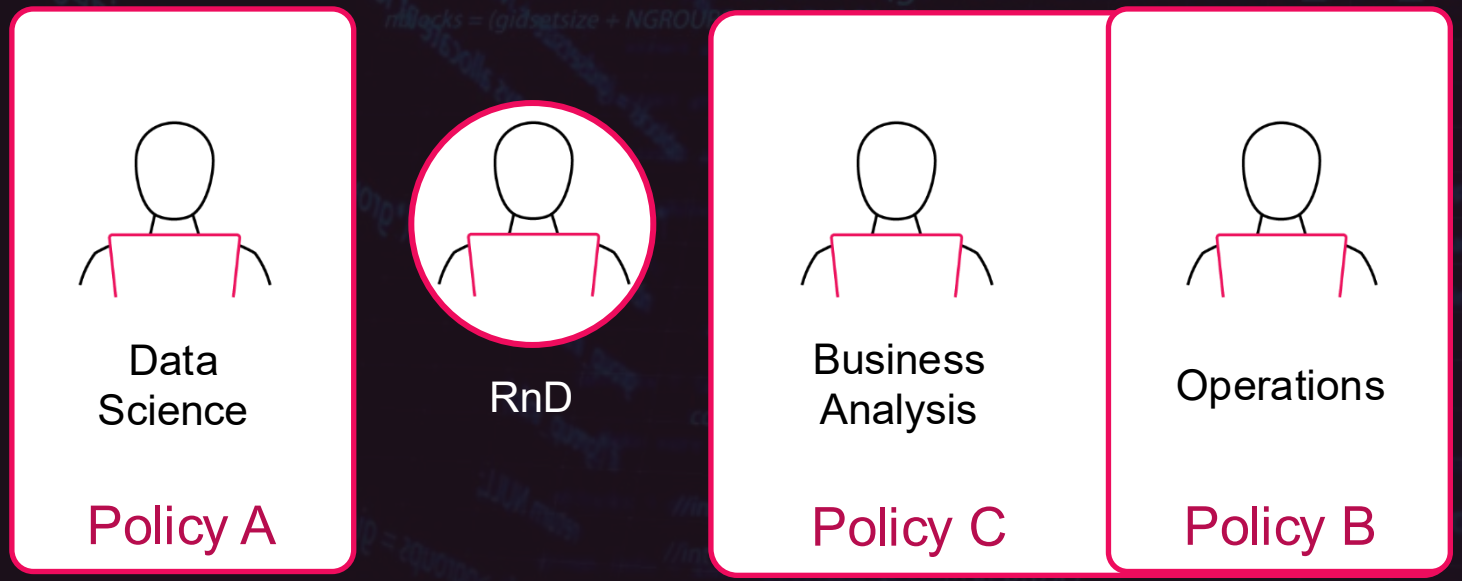
Независима от хоста и невидима для атакующих

Функции безопасности выполняются на DPU, поэтому CPU и GPU полностью выделены под задачи ИИ

CHECK POINT

Масштабируемое управление





Мультитенантное управление безопасностью
с неограниченным масштабированием и готовыми политиками

Расширенный API для автоматизации и оркестрации

Спасибо!



CHECK POINT™