СТРАТЕГИЧЕСКАЯ КОНЦЕПЦИЯ

Вычислительный ГРИД

Метакомпьютеры, федеративные базы и работа с большими вычислениями

Общий порядок доверенного вычисления
при локальном владении ресурсами и данными

Соколов Сергей Леонидович

Решение сейчас — ограниченный 90-дневный пилот

90

ДНЕЙ

3 автономных домена
3 класса нагрузки
1 независимый протокол

01	**Периметр**	задачи, baseline, данные, владельцы, STOP-критерии
02	**Подключение**	идентичность, каталоги, политики и адаптеры
03	**Сквозной запуск**	HPC, GPU/AI и федеративная аналитика
04	**Испытания**	отказ узла, отзыв прав, egress и воспроизводимость
05	**Решение**	GO, CONDITIONAL GO или STOP по измеримым KPI

Промышленное масштабирование — только отдельным решением

ПРОБЛЕМА

Изолированные мощности не образуют общий ресурс

HPC
MPI / CPU / низкая задержка

GPU
AI / ускорители / контейнеры

DATA
доменные БД / объекты / edge

Что ломает сквозной расчёт

01 разные идентичности, очереди, квоты и форматы задания

02 копирование больших массивов вместо вычисления рядом с данными

03 одна универсальная политика для MPI, HTC, AI и SQL

04 результат без версии входов, кода, среды и решения о доступе

05 раздельный учёт очереди, egress, ускорителей, энергии и стоимости

Централизация всего в одном месте переносит проблему — но не решает её

Четыре определения задают границы системы

ВЫЧИСЛИТЕЛЬНЫЙ ГРИД	федерация ресурсов разных контуров управления	**не единый кластер**
МЕТАКОМПЬЮТЕР	временная логическая машина, собранная под задачу	**не общая WAN-память**
ФЕДЕРАТИВНЫЙ СЛОЙ ДАННЫХ	каталог, политика и разрешённый запрос над автономными источниками	**не мировая БД**
БОЛЬШИЕ ВЫЧИСЛЕНИЯ	задача вне возможностей одного узла или контура	**не рекламный ярлык**

ГРИД = доверие + каталог + политика + маршрут + исполнение + учёт + доказательство

Три инварианта удерживают федерацию от централизации

01

Локальный суверенитет

Площадка сохраняет планировщик, квоты, ключи, данные и право STOP.

02

Вычисление к данным

Перемещаются код, план и разрешённая проекция; исходники остаются у владельца.

03

Проверяемый результат

Версии входов, кода, среды, политики и lineage обязательны для выпуска.

01 **default deny**

02 **policy-as-code**

03 **открытые контракты**

04 **разные backend**

05 **независимая проверка**

ГРИД связывает вопрос, вычисление и доказательство

ЭКВИЛИБРИУМ
вопрос, сценарий, риск и требования

ГРИД
план, ресурсы, допуски и исполнение

ПРОВЕРКА
корректность, безопасность и выпуск

АРХИВ
версии, lineage и разрешённый результат

СФЕРА

участники, компетенции, кейсы и знания

СПЕЦЗАЩИТА

развёртывание, эксплуатация и реагирование

ECO-PPA / ФОНДЫ

мандат, контракт, бюджет и ожидаемый результат

Ни один контур не совмещает владение данными, исполнение и окончательную самопроверку

Пять плоскостей разделяют полномочия и потоки

5 · ДОКАЗАТЕЛЬНОСТЬ	telemetry · accounting · lineage · контрольные суммы · Архив
4 · ДАННЫЕ	каталог · семантика · pushdown · staging · разрешённые проекции
3 · ВЫЧИСЛЕНИЯ	Slurm/MPI · HTCondor/workflow · Kubernetes Jobs · GPU/edge
2 · УПРАВЛЕНИЕ	каталог ресурсов · broker · orchestrator · policy engine · квоты
1 · ДОВЕРИЕ	identity · роли · атрибуты · сертификаты · отзыв · аудит

CONTROL PLANE ≠ DATA PLANE

Метакомпьютер собирается под задачу — и затем исчезает

ПАСПОРТ
цель · код · данные · SLA

КАТАЛОГ
CPU/GPU · ПО · зона · очередь

ПЛАН
время · egress · цена · риск

ДОПУСК
identity · policy · квота

ЗАПУСК
Slurm · Condor · K8s

ВО ВРЕМЯ

telemetry · checkpoints · policy events · resource accounting

ПОСЛЕ

проверка результата · манифест · отзыв прав · очистка staging

WAN не становится локальной шиной: тесно связанную MPI-задачу размещают в одном низколатентном HPC-контуре

Запрос проходит между контурами — исходники остаются

Обязательные проверки

- статистика и план запроса актуальны
- filter / projection / aggregation выполняются у источника
- объём transfer и размер результата ограничены
- семантика единиц, времени, null и типов согласована
- междоменная запись не предполагается атомарной без доказательства

ЗАПРЕЩЕНО ПО УМОЛЧАНИЮ

full scan
cross-source cross join
неограниченный egress
«единая мировая БД»

Восемь шлюзов делают задание управляемым

J0 Запрос

J1 Класс

J2 Права

J3 План

J4 Резерв

J5 Запуск

J6 Проверка

J7 Архив

J2 кто + зачем + над чем + где + сколько + что можно выпустить

Нет обязательного атрибута, подписи, лимита или валидатора — следующий шлюз закрыт

Безопасность, наблюдаемость и воспроизводимость едины

ЗАЩИТА

identity
короткие credentials
локальные ключи
изоляция
контроль egress

НАБЛЮДАЕМОСТЬ

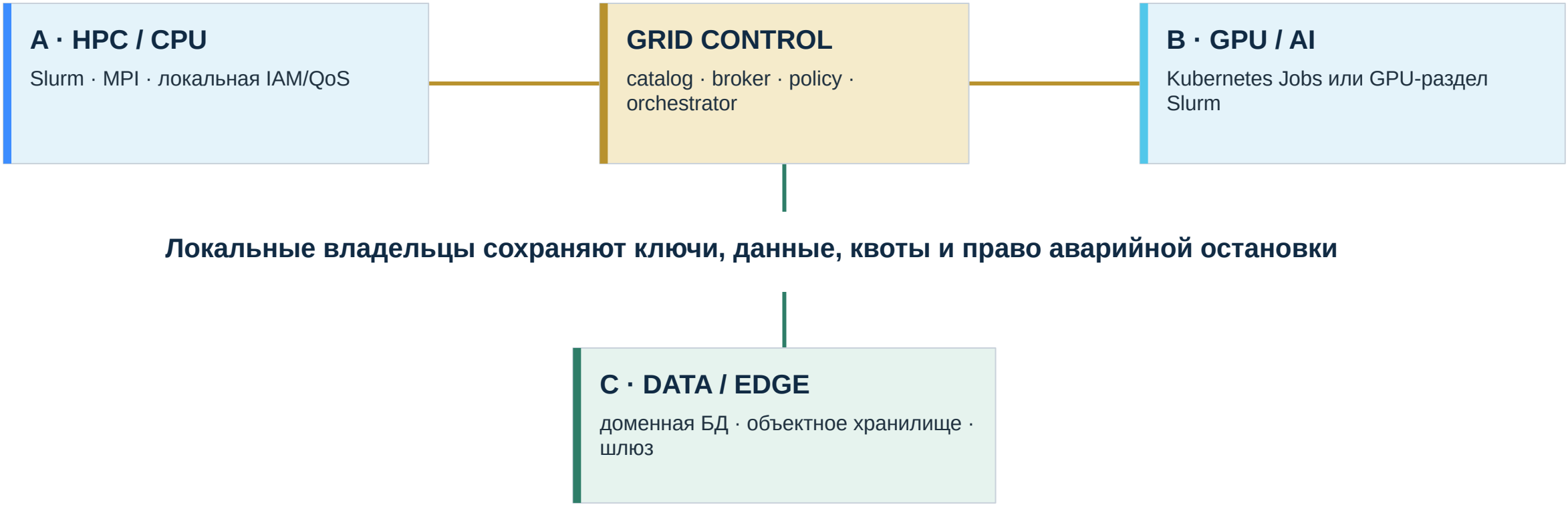
trace
metrics
logs
profiles
accounting

ВОСПРОИЗВОДИМОСТЬ

входной snapshot
версия кода
образ и параметры
policy decision
lineage

Российский контур: применимость 149-ФЗ, 152-ФЗ, 187-ФЗ и подзаконных мер квалифицируется владельцами систем до допуска данных

Три автономных домена проверяют сам принцип GRID



Три сценария проходят пять этапов за 90 дней

01

Биосферный ансамбль

HTC + агрегирование

02

GPU-анализ пространства

контейнер к данным

03

Вода · почвы · биоразнообразие

federated query + pushdown

1–15	16–30	31–55	56–75	76–90
паспорт и baseline	домены и политики	прототип и первый run	нагрузка и отказы	проверка и решение

До интеграции каждый сценарий фиксирует существующий baseline: время · ручной труд · bytes moved · стоимость · качество

GO требует одновременно пользы, контроля и воспроизводимости

GO

- 3 домена · 3 сценария
- ≥95% успешных запусков после стабилизации
- −30% time-to-first-result для 2 из 3 сценариев
- 0 неразрешённых перемещений первичных данных
- 100% операций с policy log и 100% результатов с lineage
- расхождение resource accounting ≤ ±5%

STOP

01 обход локальной политики

02 неразрешённый egress

03 утрата происхождения результата

04 критическая уязвимость без компенсации

05 невозможность независимой проверки

Пороги — проектные критерии пилота, а не отраслевой стандарт

Следующий шаг — утвердить паспорт пилота за 15 дней

ГРИД создаёт не общий склад ресурсов, а общий порядок доверенного вычисления

1. Назначить владельца пилота
2. Определить владельцев трёх доменов
3. Назначить независимого валидатора
4. Утвердить задачи, baseline и STOP-критерии

Предлагаемое решение: разрешить проектировочный пилот без промышленного допуска