Nevel

БЕЗОПАСНОСТЬ

Архитектура и границы доверия

Компоненты Platform и Guardrail, разделение контуров управления и данных, изоляция организаций и два варианта размещения.

Облако / В контуре

https://docs.nevel.sh/security/architecture/

Архитектура безопасности Nevel описывает не только соединение интерфейса с моделью. В обработке участвуют авторизация, история диалога, файлы, поиск по знаниям, подключения, Guardrail, модель, обратная обработка и аудит. Каждый компонент имеет собственный доступ к данным и может стать точкой обхода, если рассматривать только основной чат.

Для архитектурного обследования используются три понятия. **Контур данных** обрабатывает сообщения, файлы и ответы. **Контур управления** изменяет пользователей, роли, подключения, модели и политики. **Контур эксплуатации** обслуживает серверы, секреты, журналы и резервные копии. Их логическое разделение позволяет не выдавать человеку право читать документы только потому, что он управляет настройками.

Компоненты и полномочия

Компонент	Что получает	Что выполняет	Какие полномочия ему не нужны
Веб-приложение	Ввод пользователя и доступные ему данные	Отображает диалоги, проекты, файлы и результаты	Секреты провайдеров и доступ к серверным хранилищам
Сервер Platform	Идентичность, контекст пространства, выбранные объекты	Проверяет права, собирает контекст, запускает инструменты	Неограниченный доступ к чужим организациям в пользовательской операции
API Gateway	Запрос, ключ, метаданные, параметры модели	Аутентификация, лимиты, идентификатор запроса	Разрешение менять политики на основании текста запроса
DLP/NER	Текст, подлежащий отправке	Находит сущности, заменяет значения и создаёт соответствия	Управление пользователями или маршрутом вне политики
Классификатор	Содержание и обязательные метки	Определяет категорию документа и допустимость маршрута	Снятие административного запрета по просьбе модели
Model Router	Решение политики, подготовленный запрос	Выбирает разрешённый обработчик и резервный маршрут	Выбор запрещённого получателя ради доступности
Хранилище соответствий	Маски и исходные значения в границе запроса	Даёт восстановить разрешённый ответ	Публичный поиск по любому идентификатору маски
Audit	Метаданные решения и действий	Обеспечивает расследование и экспорт событий	Неограниченное накопление содержимого по умолчанию
Обработчики файлов и инструменты	Только необходимые входные данные	Извлечение, поиск, вычисления, подготовка артефактов	Неограниченная сеть, секреты других подключений

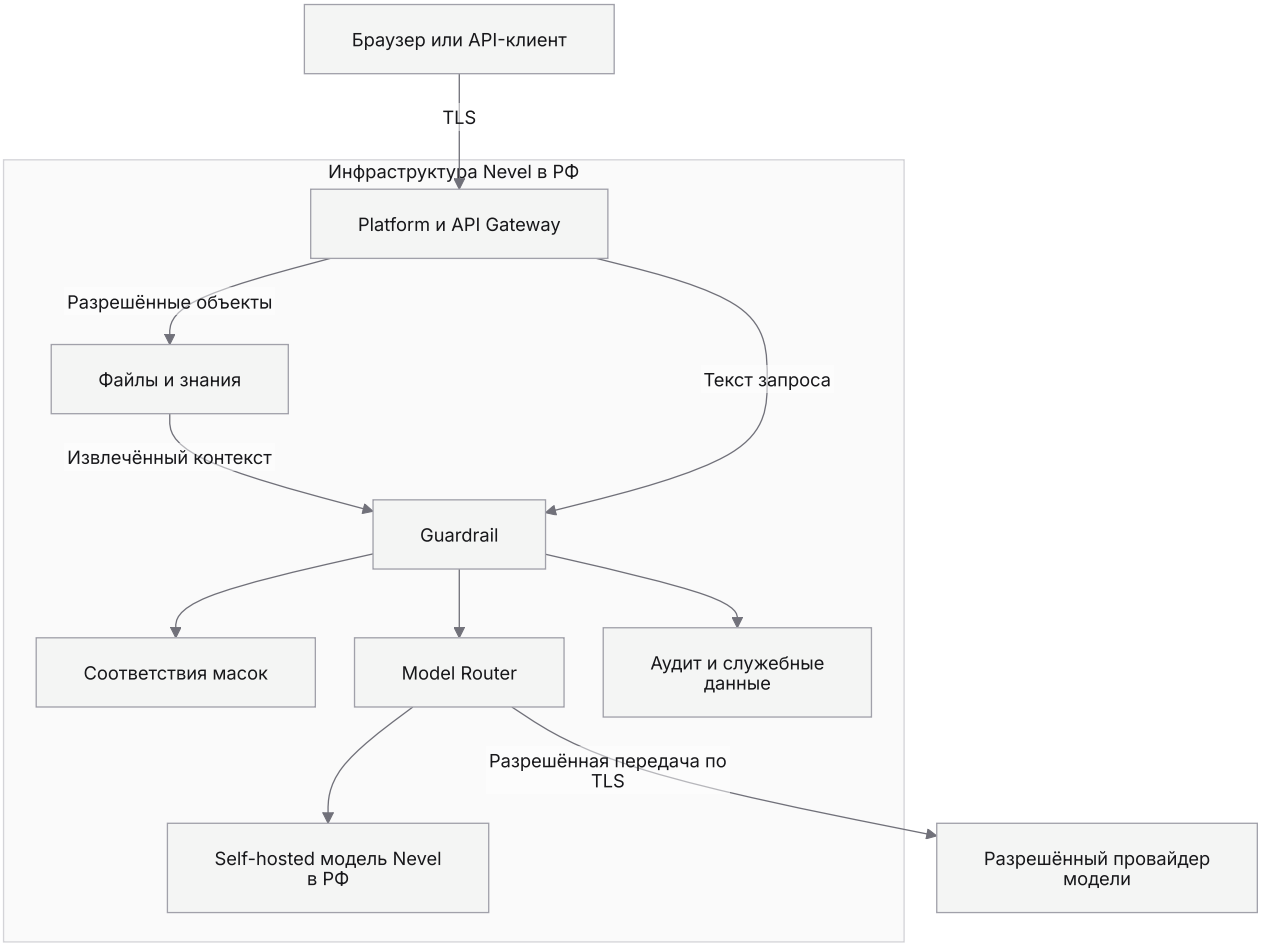
Конкретные процессы могут размещаться совместно или раздельно. Совместное размещение не отменяет разграничения сервисных полномочий и не делает внутренний сетевой вызов доверенным автоматически.

Облачная архитектура

В облачной поставке основные компоненты Platform и Guardrail, хранилища соответствий и служебные данные размещаются на российской инфраструктуре. Модельный маршрут может использовать self-hosted экземпляр Nevel на российских серверах, например DeepSeek, российского провайдера, например Яндекс или Сбер, либо разрешённую глобальную модель. Получатель и место обработки фиксируются отдельно от места Platform. Обращение к отдельному провайдеру учитывается как выход за границу инфраструктуры платформы; его география проверяется отдельно.

СХЕМА 1

Облачная архитектура



Стрелка к внешнему получателю обозначает реальную передачу подготовленного содержания. TLS защищает канал, но получатель расшифровывает запрос для обработки. Маски уменьшают объём раскрываемых идентификаторов, однако сам текст, структура таблиц, смысл документов и ответы инструментов тоже могут быть чувствительными.

Если между маршрутизатором и моделью присутствует дополнительный сервис, он включается в ту же цепочку получателей. Для каждого звена нужны назначение, география, состав получаемых данных и правила хранения. Утверждение о режиме конечной модели нельзя автоматически распространить на промежуточное звено. Результат обработки проверяется по фактическому маршруту запроса, а не по названию модели в интерфейсе.

Архитектура on-prem

При on-prem Platform, её установленные защитные компоненты и рабочие хранилища находятся в инфраструктуре заказчика. Модели размещаются независимо: внутри контура, у российского провайдера, в self-hosted инфраструктуре Nevel в РФ или у глобального провайдера. Исходящие обращения разрешаются по категории данных и конкретной цепочке получателей. Такая конфигурация остаётся on-prem, поскольку термин определяет место платформы.

Схема включает внутренний адрес модели и каждый разрешённый внешний endpoint, точки TLS, секреты, Guardrail и аудит. Если организация требует собственный AI firewall, он включается в обязательный маршрут. Российский провайдер за пределами инфраструктуры заказчика является внешним получателем по этой границе, хотя обработка может оставаться в России. Схема on-prem с несколькими маршрутами приведена в инструкции установки.

Изолированный профиль

Изолированный профиль является частным случаем on-prem. В контуре находятся приложение, шлюз, Guardrail, модели, извлечение текста, OCR, embeddings, поиск, инструменты, хранилища, секреты, аудит и резервирование. Все нужные рабочие сценарии завершаются внутри этой границы.

Внешние модели, облачный OCR, внешний поиск и телеметрия с содержимым не входят в изолированный рабочий маршрут. Если внутреннего средства недостаточно, операция завершается отказом либо выполняется другой разрешённой внутренней моделью. Недоступность компонента не меняет этот профиль на внешнюю обработку.

Обновления и поддержка являются отдельными административными операциями. Пакеты переносятся по процедуре заказчика, а доступ выдаётся конкретному исполнителю на ограниченную задачу и протоколируется. В обычном on-prem разрешённые каналы обновления также фиксируются, но постоянное отсутствие внешних соединений не является свойством всех установок.

Где существует открытый текст

Открытый текст доступен пользователю, серверу при подготовке запроса, DLP и классификатору, разрешённому обработчику файлов и модели, которая должна его обработать. В защищённом внешнем маршруте модель получает результат маскирования; в хранилище соответствий остаётся связь с исходными значениями. При восстановлении результата исходные значения снова появляются на стороне обратной обработки и в ответе уполномоченному пользователю.

Шифрование диска не скрывает данные от работающего приложения с разрешённым ключом. TLS не скрывает содержание от точки завершения TLS. Поэтому архитектурная защита включает права на сервисы, ограничение административного доступа, контроль памяти и диагностических выгрузок, минимизацию логов и разделение секретов. Подробности — шифрование и ключи.

Изоляция организаций и пространств

Граница доступа устанавливается сервером по проверенной учётной записи и членству. Переданный клиентом идентификатор организации, проекта или файла служит ссылкой на объект, но не доказательством права. Проверка выполняется при чтении, поиске, скачивании, изменении, экспорте и выполнении инструмента.

Изоляция распространяется на менее заметные объекты: кэш результатов, векторные фрагменты, соответствия масок, очередь фоновых задач и временные ссылки на файлы. Кэш без контекста полномочий может раскрыть результат, полученный ранее другим пользователем. Фоновая задача не получает бессрочное право только потому, что её когда-то создал администратор: перед выполнением учитываются актуальные права и состояние подключения.

Организационная группировка не подменяет права. Название отдела или тег пользователя помогает администрированию и аналитике, но доступ к данным определяется явным правилом. Публичность ассистента внутри пространства также не делает автоматически публичными его исходные документы или персональные подключения.

Ошибки и резервные маршруты

Недоступность основной модели допускает только резервный вариант из того же разрешённого класса обработки. Нельзя переключать документ с обязательной локальной обработкой на внешнюю модель. Тайм-аут DLP, неизвестная категория или отказ обязательного AI firewall не превращаются в разрешение отправить исходный текст.

Ошибка должна сохранять идентификатор запроса и стадию отказа, не раскрывая секреты и исходные значения в сообщении клиенту. Повторная попытка проходит те же проверки; изменение модели, тела или подключённого контекста является новым решением политики. Отмена пользователем прекращает дальнейшую работу насколько это возможно, но не означает отзыв данных, уже полученных внешним обработчиком.

Что проверяется в архитектуре

Контрольная схема содержит список компонентов, адресуемые границы, хранилища, точки TLS, сервисные идентичности, источники секретов и разрешённые исходящие соединения. Для каждой стрелки фиксируются категория данных, инициатор, получатель, протокол, возможность хранения и поведение при отказе.

Практическая проверка включает обращение к чужому объекту, работу после отзыва роли, запрос через альтернативный API-маршрут, принудительный отказ DLP и модели, а также восстановление из резервной копии. Для закрытого контура дополнительно просматриваются сетевые события при обработке файла, создании векторов, использовании инструмента и генерации артефакта. Проверка только обычного текстового ответа оставляет эти маршруты непроверенными.

Смежные статьи: движение данных, сетевая защита, файлы и RAG, модель угроз. Архитектурные механизмы не являются самостоятельным заключением о соответствии законодательству; применимость требований рассмотрена в мерах РФ.