AI-агенты как второй DevOps

Как один инженер собрал команду *умных* роботов — и что из этого вышло

Евгений Дехтярёв

DevVibeOps · Collabo

ИСТОРИЯ ИЗ ПРАКТИКИ

148

production DNS-записей удалил мой AI-агент. За три минуты. В 11 зонах.

А теперь — с начала: как я вообще оказался в компании ИИ агентов

КТО Я

Немного обо мне

2012 — 2015	**Nixys** — первые шаги в Linux-администрировании
2015 — 2026	**2ГИС** — от инженера до лида инфраструктуры и DevOps · отдел в ~35 человек
2026 → now	**Collabo** — один в поле воин: 1 DevOps на 50 человек

ДОКЛАД

О чём поговорим

- Один DevOps на проекте — моя реальность
- Концепции ИИ-команды
- Как устроена команда
- Кейсы: рутина, алерты, FinOps
- Грабли: истории из жизни агентов
- Кому это полезно и куда двигаться дальше

Один DevOps на проекте

и как ходить в отпуск?

МОЯ РЕАЛЬНОСТЬ

Я один — и 50 разных потребителей

1 : 50

я · и все, кому нужна инфра

- Операционка и развитие инфраструктуры
- Продакты вайбкодят, но в прод не умеют
- Второго инженера не нанять — наше ограничение

Вопрос не «как выбить второго», а **«как одному закрыть всё»**.

ИИ УСИЛИВАЕТ ЧЕЛОВЕКА

Я + локальный ИИ — 3–4 задачи одновременно

3–4

столько контекста
влезает в голову

→

РУТИНА — СНЯТЬ С СЕБЯ

Помочь подеплоить, завести бэки, неймспейсы и VM

Дебаги, разбор инцидентов, отчёты о работе и расходах

ИИ УСИЛИВАЕТ ЧЕЛОВЕКА

Я + локальный ИИ — 3–4 задачи одновременно

3–4

столько контекста влезает в голову

→

РУТИНА — СНЯТЬ С СЕБЯ

- Помочь подеплоить, завести бэки, неймспейсы и VM
- Дебаги, разбор инцидентов, отчёты о работе и расходах

РАЗВИТИЕ — ВЛОЖИТЬ ОСВОБОДИВШЕЕСЯ ВРЕМЯ

- Новая автоматизация и инструменты
- Архитектура, которой давно не было времени заняться

ЧТО, ЕСЛИ…

А что, если те же задачи могут идти без меня?

Мы везде применяем Ai — логично попробовать собрать **свою команду ИИ-агентов**

Вдруг сработает — и снимет с меня рутину.

Концепции ИИ-команды

чтобы мы были в одном инфополе

КОНЦЕПЦИИ · 1 ИЗ 3

Два слоя: агент и оркестратор

AI-агент

Claude Code CLI

Адаптер `claude_local` — то, что реально пишет код и ходит в инструменты.

По сути — обычный Claude Code из терминала, запущенный с вашим auth-токеном.

Аналогия: **сотрудник-исполнитель**

КОНЦЕПЦИИ · 1 ИЗ 3

Два слоя: агент и оркестратор

AI-агент

Claude Code CLI

Адаптер `claude_local` — то, что реально пишет код и ходит в инструменты.

По сути — обычный Claude Code из терминала, запущенный с вашим auth-токеном.

Аналогия: **сотрудник-исполнитель**

Оркестратор · control plane

Paperclip open-source, MIT

Управляет агентами как компанией: цели, расписание, бюджеты, governance.

Вместо вас оркестрирует десятки консольных вкладок с запущенным Claude Code.

Аналогия: **«офис» с расписанием и бюджетом**

КОНЦЕПЦИИ · 2 ИЗ 3

Paperclip: компания из агентов

Company 01

Агенты с ролями, общими целями и общим бюджетом

«DevOps-команда Collabo»

КОНЦЕПЦИИ · 2 ИЗ 3

Paperclip: компания из агентов

Company 01

Агенты с ролями, общими целями и общим бюджетом

«DevOps-команда Collabo»

Goals 02

Направления, под которые агенты приоритезирует работу

Backlog Zero · Proactive Reliability · Zero-Regression Prod Deploys

КОНЦЕПЦИИ · 2 ИЗ 3

Paperclip: компания из агентов

Company 01

Агенты с ролями, общими целями и общим бюджетом

«DevOps-команда Collabo»

Goals 02

Направления, под которые агенты приоритезирует работу

Backlog Zero · Proactive Reliability · Zero-Regression Prod Deploys

Projects 03

Группируют тикеты; к ним привязаны рутины

backlog-zero · knowledge-base · development

КОНЦЕПЦИИ · 2 ИЗ 3

Paperclip: компания из агентов

Company 01

Агенты с ролями, общими целями и общим бюджетом

«DevOps-команда Collabo»

Goals 02

Направления, под которые агенты приоритезирует работу

Backlog Zero · Proactive Reliability · Zero-Regression Prod Deploys

Projects 03

Группируют тикеты; к ним привязаны рутины

backlog-zero · knowledge-base · development

Agents 04

Роль, модель, набор skills/MCP и прочие настройки

CEO · Engineer · SRE · Analyst …

Виртуальная компания — вся она описана в git и заливается через CI

КОНЦЕПЦИИ · 3 ИЗ 3

Как агенты просыпаются

ПО ТАЙМЕРУ

Heartbeat

Просыпается по таймеру, делает чеклист и засыпает.

КОНЦЕПЦИИ · 3 ИЗ 3

Как агенты просыпаются

ПО ТАЙМЕРУ

Heartbeat

Просыпается по таймеру, делает чеклист и засыпает.

ПО РАСПИСАНИЮ

Routine

Расписание (cron). Каждый запуск создаёт тикет и будит агента.

КОНЦЕПЦИИ · 3 ИЗ 3

Как агенты просыпаются

ПО ТАЙМЕРУ

Heartbeat

Просыпается по таймеру, делает чеклист и засыпает.

ПО РАСПИСАНИЮ

Routine

Расписание (cron). Каждый запуск создаёт тикет и будит агента.

ПО СОБЫТИЮ

Назначение · @-mention

Назначили тикет или @-упомянули в комментарии — агент просыпается на него.

КОНЦЕПЦИИ · 3 ИЗ 3

Как агенты просыпаются

ПО ТАЙМЕРУ

Heartbeat

Просыпается по таймеру, делает чеклист и засыпает.

ПО РАСПИСАНИЮ

Routine

Расписание (cron). Каждый запуск создаёт тикет и будит агента.

ПО СОБЫТИЮ

Назначение · @-mention

Назначили тикет или @-упомянули в комментарии — агент просыпается на него.

ПО ДЕЛЕГИРОВАНИЮ

Агент → агент

Один агент поручает работу другому. Когда она выполнена, первый просыпается и продолжает свою задачу.

Агент не работает непрерывно — он спит, просыпается по триггеру, делает дело и засыпает.

Как устроена команда

шесть агентов и три линии

КОМАНДА · ЛИНИЯ 1 ИЗ 3

PLAN — думают и планируют

Решают, что делать и сколько это стоит. Сами изменения не вносят.

CEO ОРКЕСТРАТОР

Триаж тикетов, делегирование, закрытие цикла. **Раздаёт работу — сам ничего не разрабатывает.**

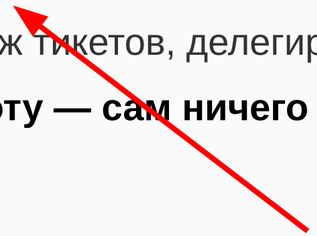
КОМАНДА · ЛИНИЯ 1 ИЗ 3

PLAN — думают и планируют

Решают, что делать и сколько это стоит. Сами изменения не вносят.

CEO ОРКЕСТРАТОР

Триаж тикетов, делегирование, закрытие цикла. **Раздаёт работу — сам ничего не разрабатывает.**



Хардкод

PLAN · CEO | BUILD | RUN

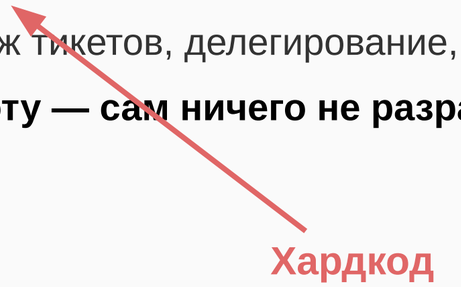
КОМАНДА · ЛИНИЯ 1 ИЗ 3

PLAN — думают и планируют

Решают, что делать и сколько это стоит. Сами изменения не вносят.

CEO ОРКЕСТРАТОР

Триаж тикетов, делегирование, закрытие цикла. **Раздаёт работу — сам ничего не разрабатывает.**



Analyst АНАЛИТИКА

Weekly brief, FinOps, квоты ресурсов. Считает, сводит цифры и предупреждает.

PLAN · CEO, Analyst | BUILD | RUN

КОМАНДА · ЛИНИЯ 2 ИЗ 3

BUILD — делают изменения

Пишут инфраструктуру и проверяют её — результат каждого шага — Merge Request.

Infra Engineer ДЕЛАЕТ

Terraform, Helm, CI/CD, BASH/Python
→ заботится о Gitflow

PLAN · CEO, Analyst | **BUILD · Engineer** | RUN

КОМАНДА · ЛИНИЯ 2 ИЗ 3

BUILD — делают изменения

Пишут инфраструктуру и проверяют её — результат каждого шага — Merge Request.

Infra Engineer ДЕЛАЕТ

Terraform, Helm, CI/CD, BASH/Python
→ заботится о Gitflow

Code Reviewer РЕВЬЮИТ

Ревью MR, critical-review checklist
→ возвращает на доработку Engineer,

пока не будет OK.

PLAN · CEO, Analyst | **BUILD** · Engineer, Reviewer | RUN

КОМАНДА · ЛИНИЯ 2 ИЗ 3

BUILD — делают изменения

Пишут инфраструктуру и проверяют её — результат каждого шага — Merge Request.

Infra Engineer ДЕЛАЕТ

Terraform, Helm, CI/CD, BASH/Python
→ заботится о Gitflow

Code Reviewer РЕВЬЮИТ

Ревью MR, critical-review checklist
→ возвращает на доработку Engineer, пока не будет OK.

Tester ПРИНИМАЕТ

Финальная приёмка: dry-run (plan / template), acceptance criteria.

PLAN · CEO, Analyst | **BUILD** · Engineer, Reviewer, Tester | RUN

КОМАНДА · ЛИНИЯ 3 ИЗ 3

RUN — следят за продакшеном

Наблюдают и диагностируют. Только читают прод — ничего не меняют.

SRE НАБЛЮДАЕТ

Разбор алертов, proactive scan, RCA.

ПОЧЕМУ ТОЛЬКО ЧТЕНИЕ?

Пока приложения в активной разработке, человеческие ошибки рождают алерты. ИИ принимает их **за чистую монету** и очень хочет угодить — «всё починить».

Например: **добавляет ресурсы там, где произошел обычный мисконфиг**

PLAN · CEO, Analyst

BUILD · Engineer, Reviewer, Tester

RUN · SRE

КОМАНДА

Путь задачи: от тикета до merge

Тикет
Assign = bot

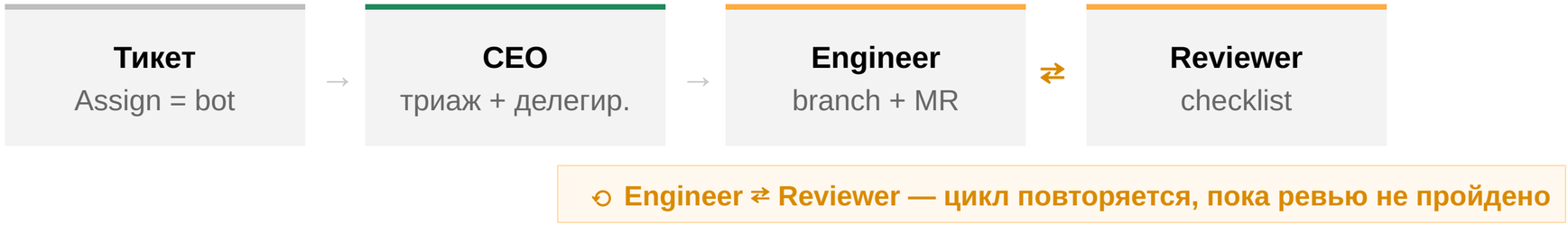
→

CEO
триаж + делегир.

- CEO берёт **только** карточки, где `Assign` = бот.

КОМАНДА

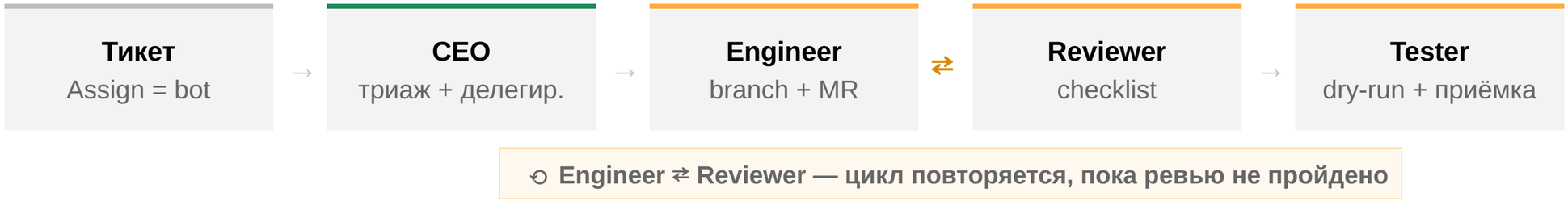
Путь задачи: от тикета до merge



- CEO берёт **только** карточки, где `Assign` = бот.
- Engineer зеркалит каждый MR комментарием обратно в тикет.
- Engineer ⇄ Reviewer — цикл повторяется, пока **Reviewer не доволен** результатом.

КОМАНДА

Путь задачи: от тикета до merge



- CEO берёт **только** карточки, где `Assign` = бот.
- Engineer зеркалит каждый MR комментарием обратно в тикет.
- Engineer ⇄ Reviewer — цикл повторяется, пока **Reviewer не доволен** результатом.
- Tester — финальная приёмка перед merge

КОМАНДА

Путь задачи: от тикета до merge



- CEO берёт **только** карточки, где `Assign` = бот.
- Engineer зеркалит каждый MR комментарием обратно в тикет.
- Engineer ⇄ Reviewer — цикл повторяется, пока **Reviewer не доволен** результатом.
- Tester — финальная приёмка перед merge
- Человек — пытается вникнуть и замержить

DevOps Kanban

All Tasks | PaperClip Tasks

Backlog 0

+ New page

Planning 1

Grafana дублирует алерты в Slack (вероятно HA-реплики без дедупа)

Dmitrii

paperclip-agents

+ New page

In progress 3

Бэкап S3-бакетов по всем ЦОД

Evgeny Dekhtyarev

paperclip-agents

5

Paperclip-agents SSH доступ до VMs

Evgeny Dekhtyarev

paperclip-agents

5

Подключить аудит-логи MKS для staging и production кластеров

Evgeny Dekhtyarev

paperclip-agents

4

+ New page

Done 14

Создать БД [redacted] и пользователя на PostgreSQL-инстансах

Mikhail

paperclip-agents

6

Создать psql базу для [redacted]

Kirill

paperclip-agents

4

Завести два бакета в s3: stg-[redacted] storage и prod-[redacted] storage (versioning OFF)

Yura

paperclip-agents

7

S3 bucket для [redacted]

Evgeny Dekhtyarev

paperclip-agents

4

psql и s3 для разделов лендинга [redacted]

Миша

paperclip-agents

paperclip-agents Jun 22

🗂️ Триаж CEO — эпик «Бэкап S3-бакетов по всем ЦОД»

Задача масштабная: 4 облака (GCP [redacted], Selectel, OCI), создание новых billable cold/archive-бакетов и кросс-облачная архитектура репликации. Это не один MR — предлагаю разбить на фазы и согласовать рамки до делегирования.

Что уже есть (переиспользуем):
• Аудиты бэкапов закрыты: GCP [redacted], Selectel — инвентаризация частично готова.
• TF-state версионирование/бэкап в работе отдельно (selectel !90).

Предлагаемая декомпозиция:
1. Фаза 1 — единая инвентаризация всех бакетов + проставление тега backup: true/false (безопасно, RO+теги, делегируется по одному облаку = один MR).
2. Фаза 2 — целевые cold-бакеты + механизм репликации по каждому облаку отдельным MR: GCP Storage Transfer, OCI CRR, Selectel rclone CronJob.
3. Фаза 3 — мониторинг в Grafana (алерт «backup:true не реплицировался >2д») — связан с задачей дашборда из бэклога.

Блокеры до старта (нужно подтверждение):
1. Бюджет: cold/archive-хранилище — новые billable ресурсы во всех 4 ЦОД. Нужен потолок затрат / одобрение Board.
2. Scope окружений: бэкапим только prod-бакеты или dev/staging тоже? Какие данные критичны в первую очередь?
3. Приоритет облаков: с какого ЦОД начинаем Фазу 2? (предлагаю GCP — native Storage Transfer, самый дешёвый путь)
4. Восстановление: тестовое восстановление — кто валидирует (DevOps только)?

После ответов разобью на per-cloud issues. Пока держу карточку на согласовании.



Evgeny Dekhtyarev Jun 22

1. Давай думать про 1000$ на все ЦОД
2. Бэкапим все бакеты где стоит флаг. Нормально что это dev/stage
3. Можешь начинать с GCP
4. Собрать бы какой-нибудь e2e тест, что работает и бэкап и восстановление с него и данные идентичны. Можно завести тестовый бакет под это в каждом из ЦОД

New

paperclip-agents Jun 22

🤖 [CEO] Забираю эпик в работу — спасибо за ответы. Зафиксированный контекст: • Бюджет ~$1000 на все ЦОД суммарно. • Скоуп: бэкапим бакеты с флагом (метка) backup; dev/stage — норма. • Порядок: начинаем с GCP. • Приёмка: e2e-тест бэкап+восстановление с проверкой идентичности данных, отдельный тестовый бакет под это в каждом ЦОД.

Paperclip DevOps Reviewer @paperclip-devops-reviewer 17 hours ago

Developer

✅ Code Review — THE-3733

Статус: Одобрено

Pipeline: success ✅

Blast radius: `[redacted]-staging / ns monitoring`, Deployment grafana, без удаления ресурсов, только конфигурация HA.

Чек-лист (post-incident 2026-04-17)

#	Пункт	Результат
0	Pipeline success	✅
1	Деструктивных операций нет	✅
2	Окружение: staging (не prod)	✅
3	Rollback plan описан	✅
4	Bot-токен, личных PAT нет	✅
5	dry-run перед apply в CI	✅
6	Blast radius указан	✅
7	Backup не требуется (нет state mutation)	✅
8	Notion source присутствует	✅

Paperclip DevOps Reviewer @paperclip-devops-reviewer 17 hours ago

Developer

✅ Code Review — THE-3733

Статус: Одобрено

Pipeline: success ✅

Blast radius: [redacted]-staging / ns monitoring, Deployment grafana, без удаления ресурсов, только конфигурация HA.

Paperclip DevOps Reviewer approved this merge request 17 hours ago

Paperclip DevOps Lead @paperclip-devops-lead 8 hours ago

Developer

@e.dekhtyarev MR одобрен ревьюером (paperclip-devops-reviewer), pipeline зелёный, blast radius разобран. Бот-аккаунты — Developer, мержить в protected `main` не могут. Прошу смержить вручную, чтобы устранить дублирование Slack-алертов (HA unified alerting). После мержа задача в Notion будет закрыта.

3	Rollback plan описан	✅
4	Bot-токен, личных PAT нет	✅
5	dry-run перед apply в CI	✅
6	Blast radius указан	✅
7	Backup не требуется (нет state mutation)	✅
8	Notion source присутствует	✅

Кейсы

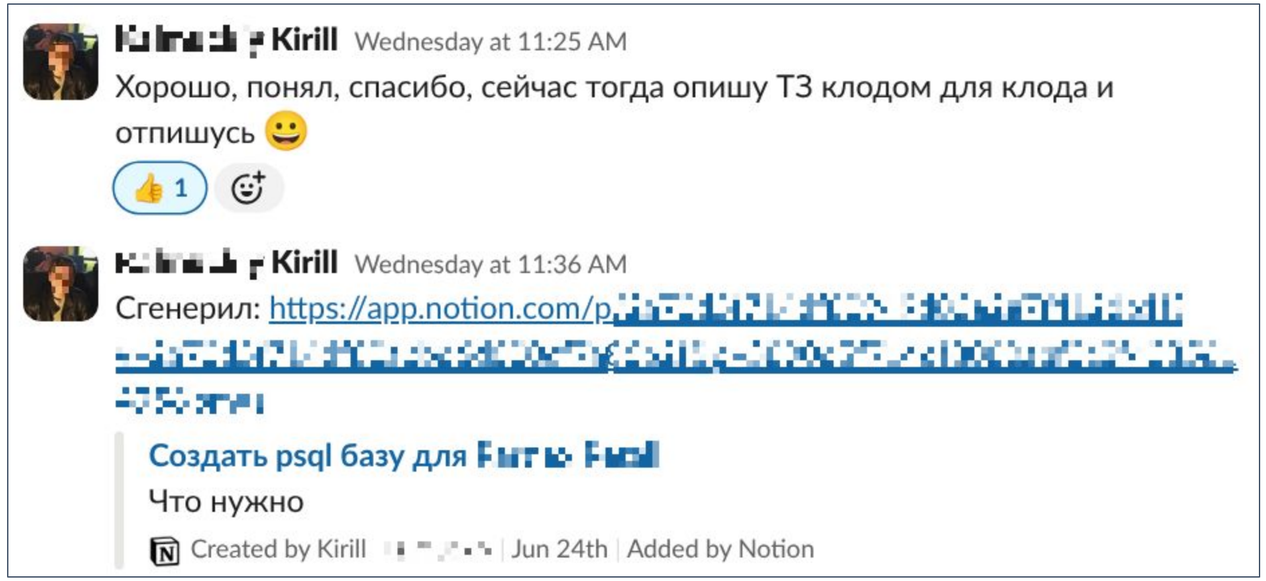
рутина · алерты · FinOps

КЕЙС 1

Рутина под ключ — только успевай моржевать

Разработчики кидают простые инфра-задачи в **Kanban (Notion)**:

бакет, БД, namespace в k8s. ИИ берёт **под ключ**.

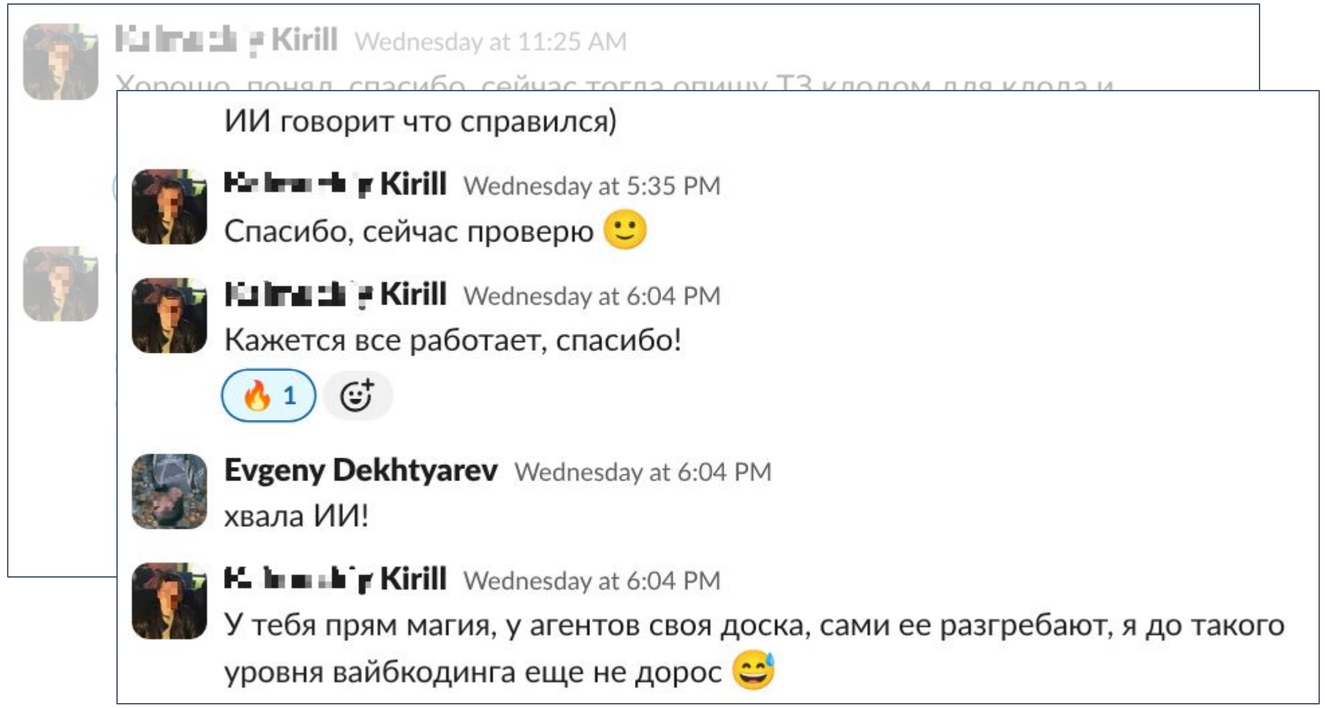


КЕЙС 1

Рутина под ключ — только успевай моржевать

Разработчики кидают простые инфра-задачи в **Kanban (Notion)**:

бакет, БД, namespace в k8s. ИИ берёт **под ключ**.

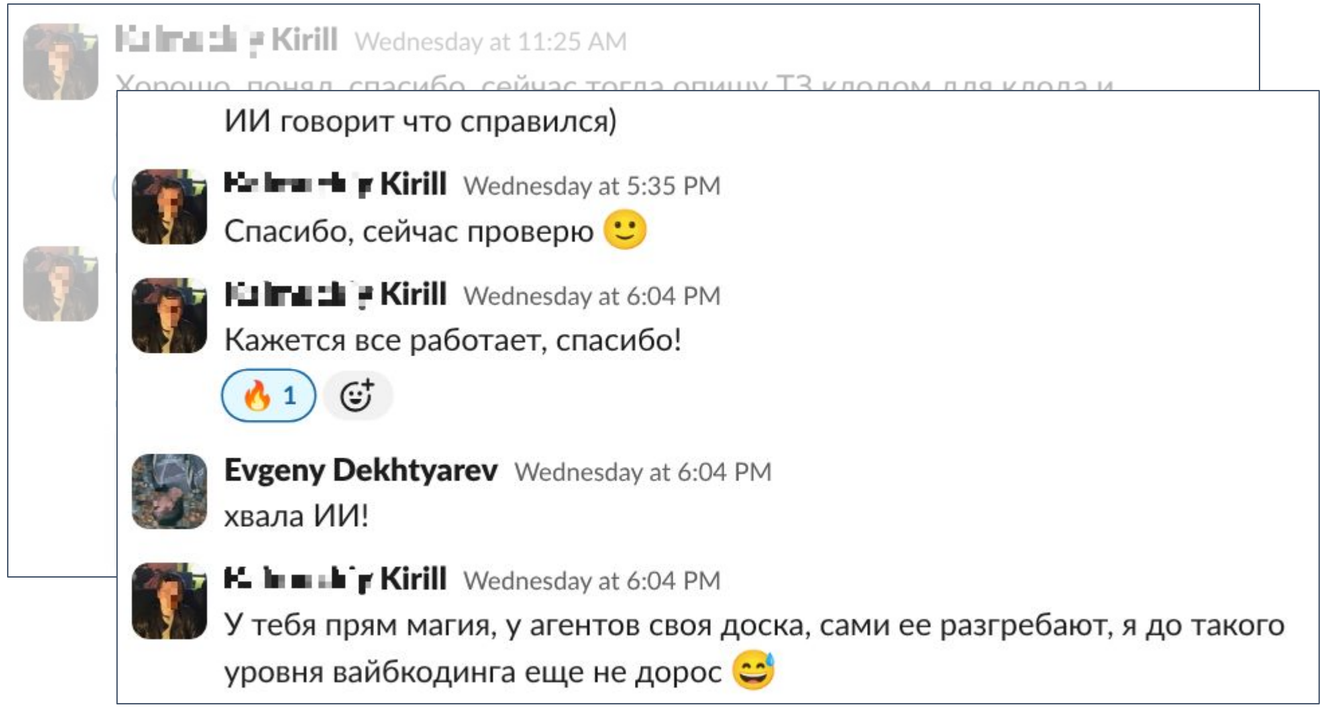


КЕЙС 1

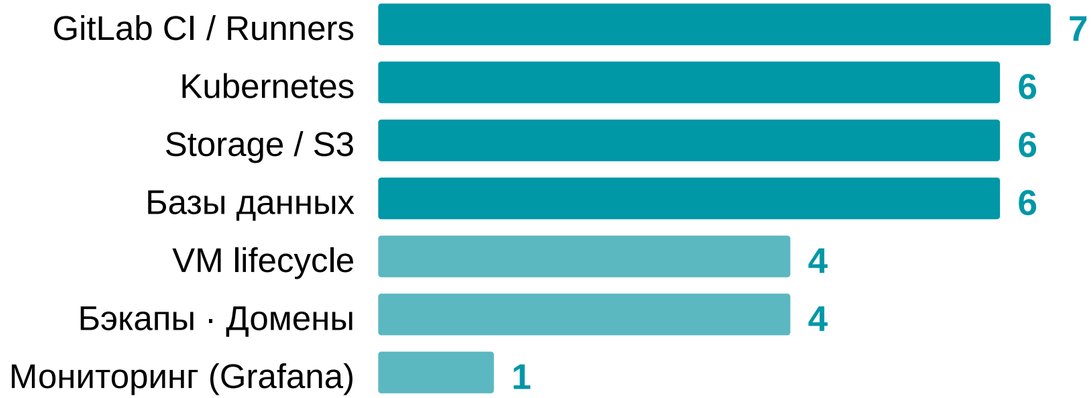
Рутина под ключ — только успевай моржевать

Разработчики кидают простые инфра-задачи в **Kanban (Notion)**:

бакет, БД, namespace в k8s. ИИ берёт **под ключ**.



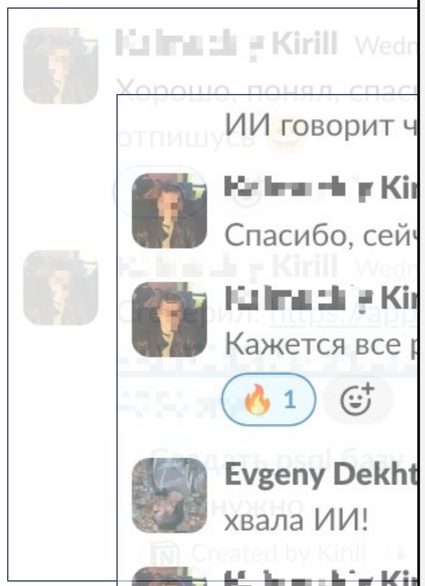
34 тикета за первый месяц



КЕЙС 1

Рутина под ключ — только успевай моржевать

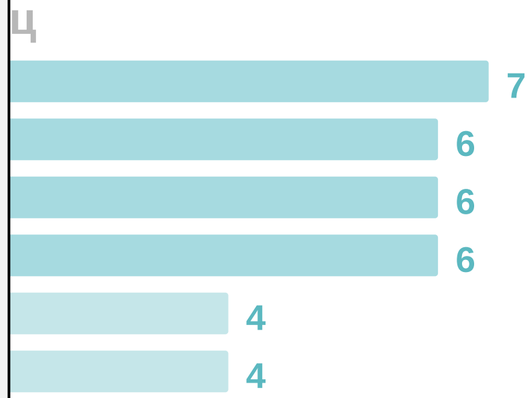
Разработчики кидают п
бакет, БД, namespace в



Полгода спустя

353

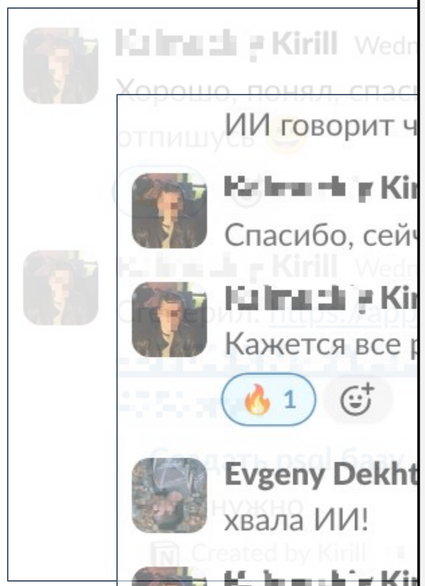
смёржено



КЕЙС 1

Рутина под ключ — только успевай моржевать

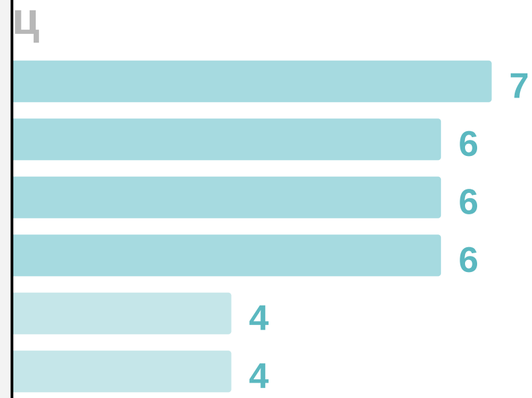
Разработчики кидают п
бакет, БД, namespace в



Полгода спустя

353
смёржено

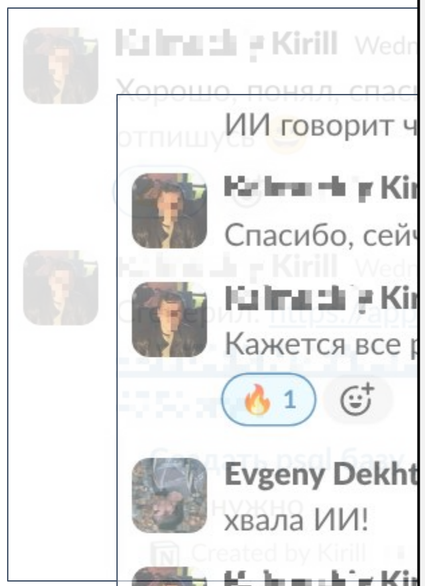
217
закрыто



КЕЙС 1

Рутина под ключ — только успевай моржевать

Разработчики кидают п
бакет, БД, namespace в



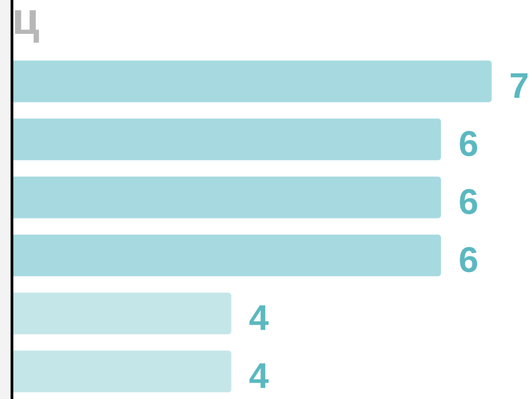
Полгода спустя

353
смёржено

217
закрыто

38% завершившихся MR закрыты без merge

Ошибки, ненужные изменения, дубликаты и замены



КЕЙС 2

Разбор алертов

1. Alertmanager публикует алерт в Slack.

КЕЙС 2

Разбор алертов

1. Alertmanager публикует алерт в Slack.
2. Routine раз в 20 минут запускает SRE-агента.

КЕЙС 2

Разбор алертов

1. Alertmanager публикует алерт в Slack.

2. Routine раз в 20 минут запускает SRE-агента.

3. Агент проверяет факты и пишет разбор в тред алерта.

У агента есть RO доступ до инфраструктуры, логов, метрик, Gitlab.

Но контекста не всегда хватает.

Обычно 1 из 3 разборов реально полезен

КЕЙС 2

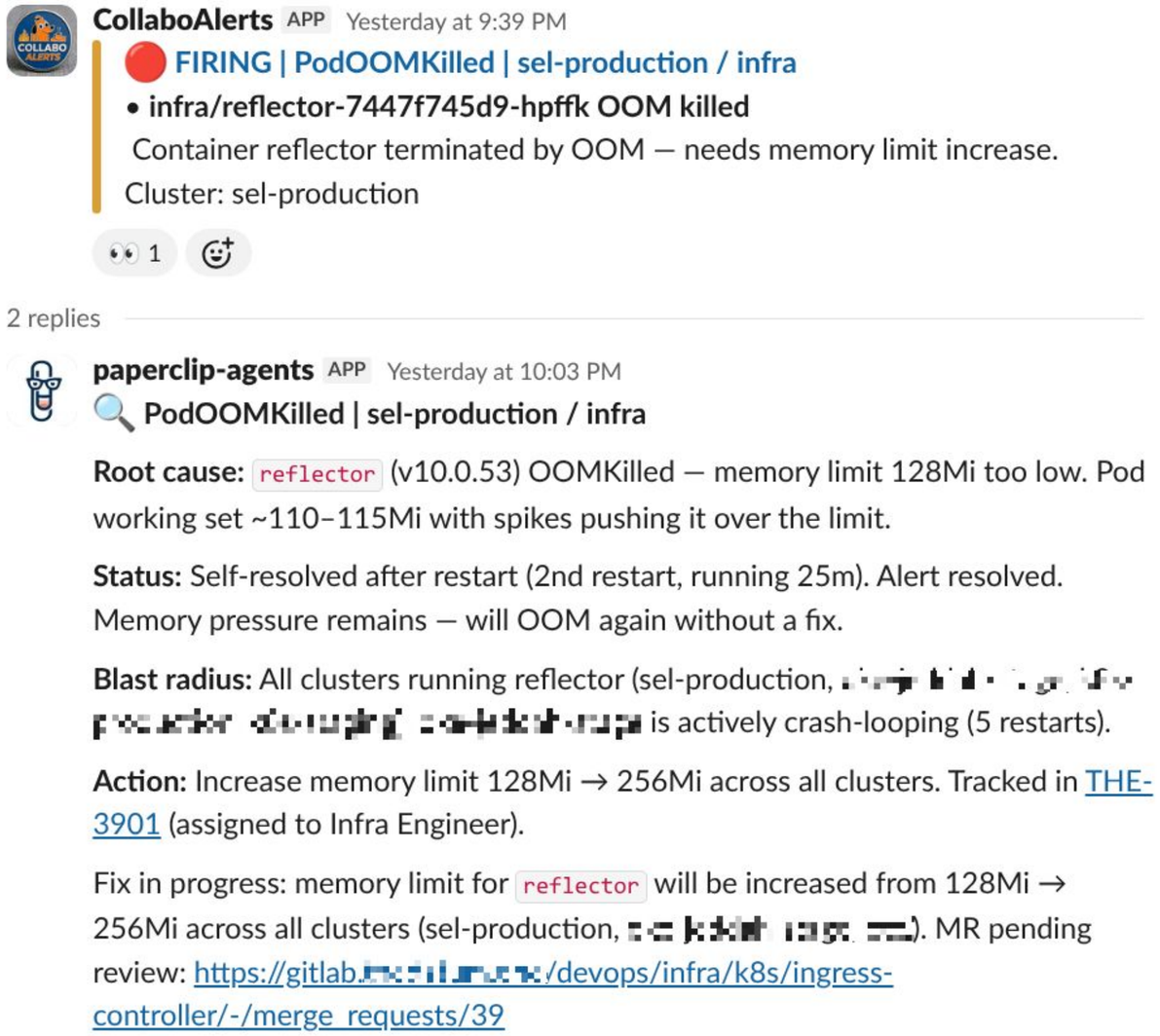
Разбор алертов

1. Alertmanager публикует алерт в Slack.
2. Routine раз в 20 минут запускает SRE-агента.
3. Агент проверяет факты и пишет разбор в тред алерта.

У агента есть RO доступ до инфраструктуры, логов, метрик, Gitlab.

Но контекста не всегда хватает.

Обычно 1 из 3 разборов реально полезен



КЕЙС 3

FinOps

Боль: руками свести биллинг разных провайдеров в таблицу

дни на один отчёт

КЕЙС 3

FinOps

Боль: руками свести биллинг разных провайдеров в таблицу

дни на один отчёт

Агент Аналитик по **Routine** собирает биллинг по всем провайдерам.

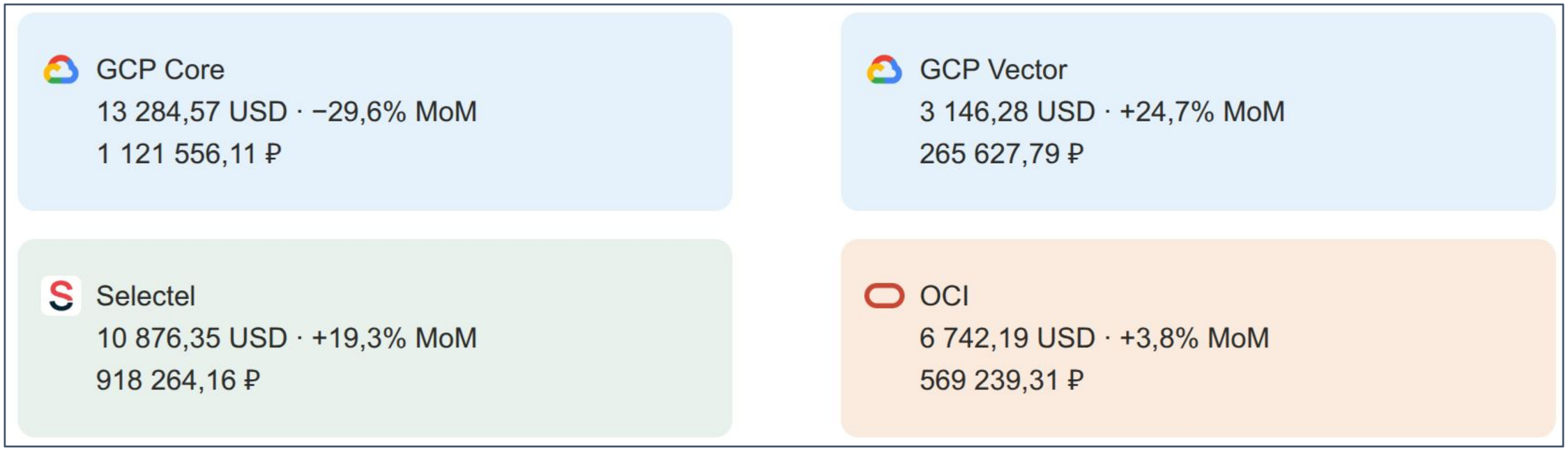
КЕЙС 3

FinOps

Боль: руками свести биллинг разных провайдеров в таблицу

дни на один отчёт

Агент Аналитик по **Routine** собирает биллинг по всем провайдерам.



КЕЙС 3

FinOp

Бухгалтерский проект	Технический проект	Всего ₽
Core	LevelUp	12 485,72 ₽
Core	Flexion / Matrix	58 742,19 ₽
Core / Orion	Orion	53 618,43 ₽
Core / Nimbus	Nimbus	218 934,56 ₽
Core / Vertex	Vertex	947 823,14 ₽
Vector	Vector	268 519,32
Vector / Infrastructure	Infrastructure: VM / DevEnv / Shared IT	578 243,91 ₽
Vector / Infrastructure	Nexus Infrastructure Platform	196 745,28 ₽
Vector / Infrastructure	Render Services	22 381,64 ₽
Vector / Client SDK	Nexus SDK	8 916,35 ₽

КЕЙС 4

Задачи из Slack

Человек

Описывает задачу в чате

КЕЙС 4

Задачи из Slack



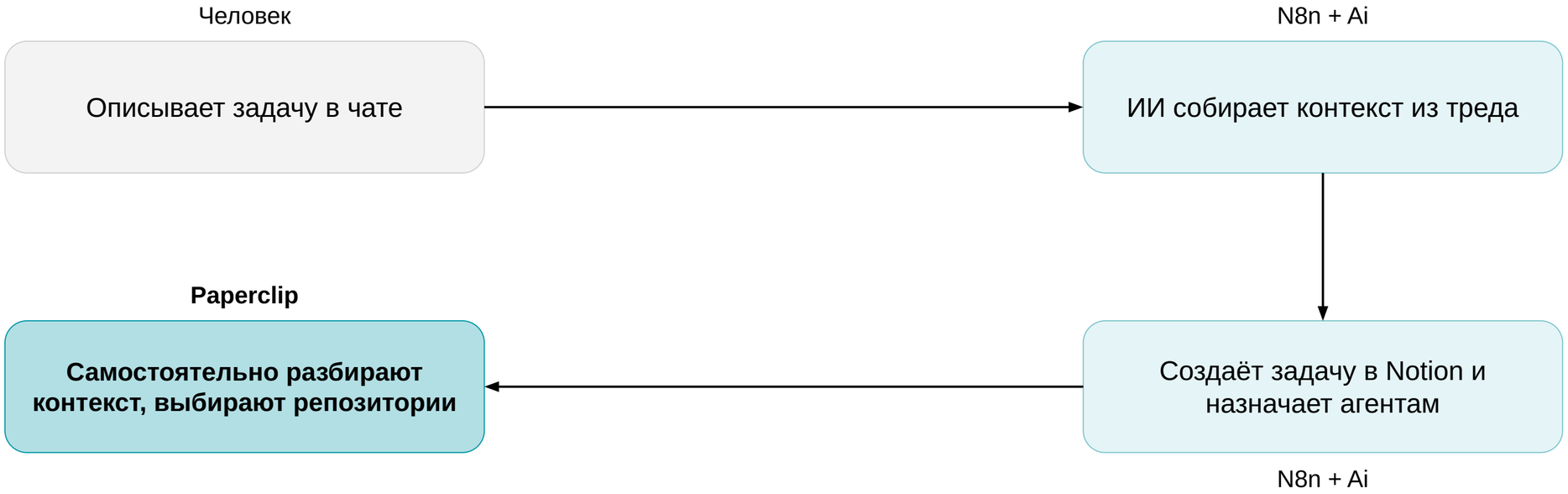
КЕЙС 4

Задачи из Slack



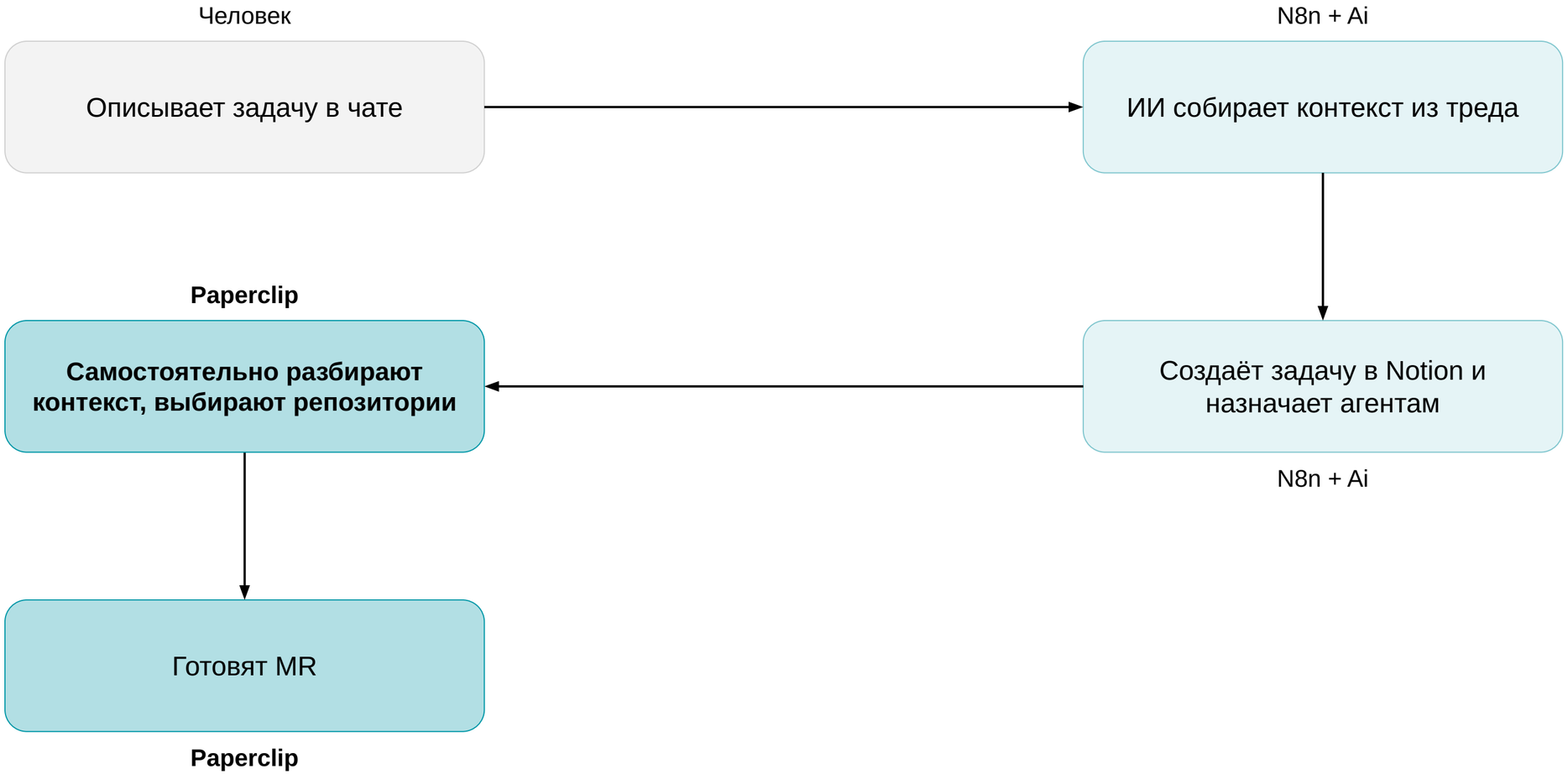
КЕЙС 4

Задачи из Slack



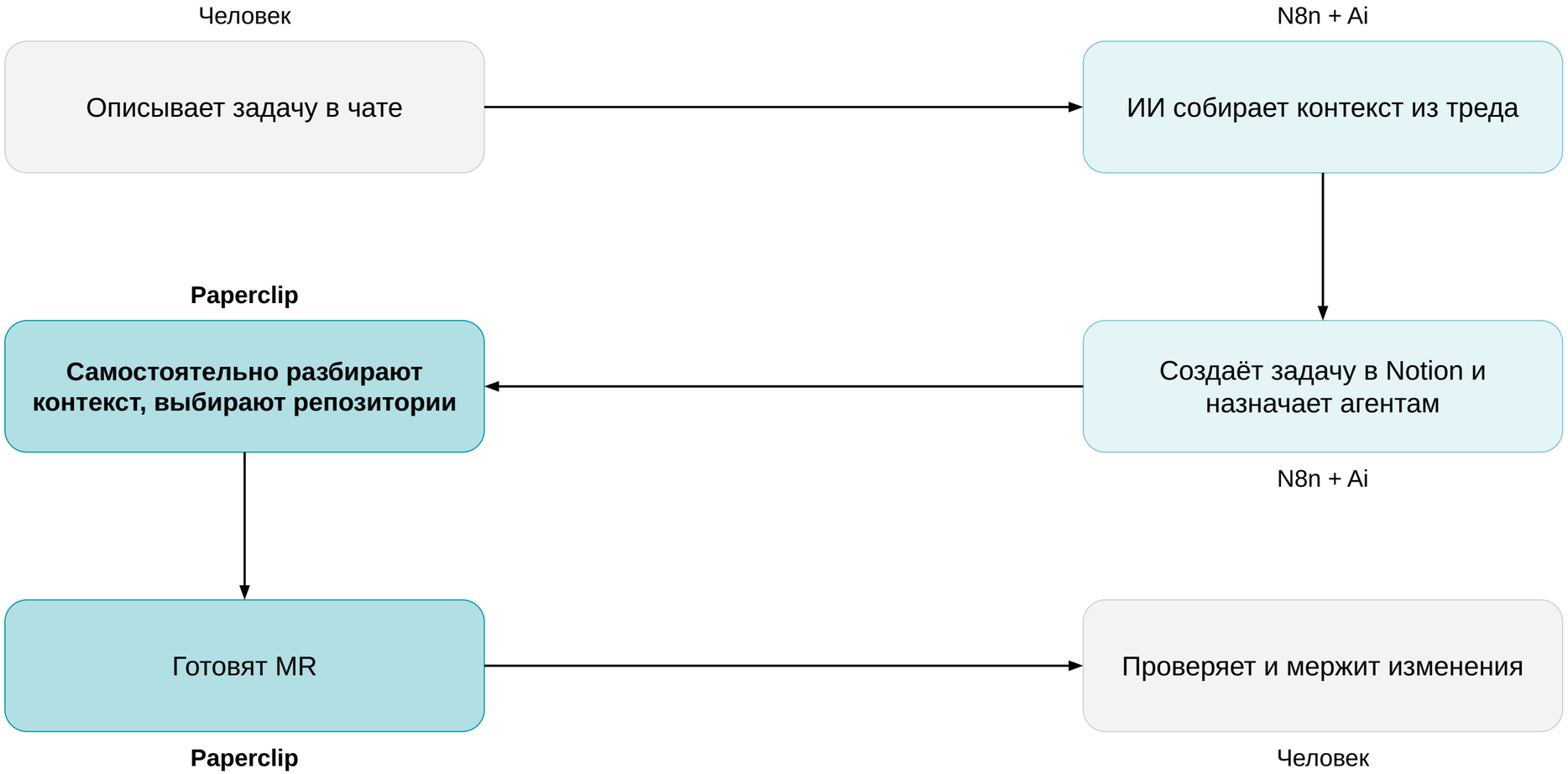
КЕЙС 4

Задачи из Slack



КЕЙС 4

Задачи из Slack



КЕЙС 4

Задачи из Slack



Evgeny Dekhtyarev Aug 12th at 10:55 AM

Мы внедрили Velero в задаче app.notion.com/p/https-velero-io... и репозитории gitlab.[redacted]/devops/infra/k8s/velero

Но ничего нигде не написали в документацию, подстраницей сюда app.notion.com/p/Kubernetes-2... .
Заведи задачу в ноушен, чтобы описать что такое velero и как его использовать в нашей инфраструктуре.
Пиши сразу конечное состояние, как будто оно есть у нас во всех кластерах и им можно уже пользоваться. По необходимости потом отдельно поправим документацию.

КЕЙС 4

Задачи из Slack



Evgeny Dekhtyarev Aug 12th at 10:55 AM

Мы внедрили Velero в задаче app.notion.com/p/https-velero-io... и репозитории gitlab.[redacted]/devops/infra/k8s/velero

Но ничего нигде не написали в документацию, подстраницей сюда app.notion.com/p/Kubernetes-2... .
Заведи задачу в ноушен, чтобы описать что такое velero и как его использовать в нашей инфраструктуре.
Пиши сразу конечное состояние, как будто оно есть у нас во всех кластерах и им можно уже пользоваться. По необходимости потом отдельно поправим документацию.

3 replies

DevOpsAi APP Aug 12th at 10:55 AM

🤖 Понял-принял. Грею нейроны, скоро вернусь с фактами.

КЕЙС 4

Задачи из Slack



Evgeny Dekhtyarev Aug 12th at 10:55 AM

Мы внедрили Velero в задаче app.notion.com/p/https-velero-io... и репозитории gitlab.[redacted]/devops/infra/k8s/velero

Но ничего нигде не написали в документацию, подстраницей сюда app.notion.com/p/Kubernetes-2... .
Заведи задачу в ноушен, чтобы описать что такое velero и как его использовать в нашей инфраструктуре.
Пиши сразу конечное состояние, как будто оно есть у нас во всех кластерах и им можно уже пользоваться. По необходимости потом отдельно поправим документацию.

3 replies



DevOpsAi APP Aug 12th at 10:55 AM

🤖 Понял-принял. Грею нейроны, скоро вернусь с фактами.

Задача создана в Notion: **Документация по Velero в разделе Kubernetes** — Status: Planning, Reporter: DevOpsAi, Assign: paperclip-agents.

Ссылка: Velero — Kubernetes

Контекст задачи (весь тред) подтянулся автоматически: описать Velero как конечное состояние — что это и как пользоваться в нашей инфраструктуре, подстраницей под разделом Kubernetes, с учётом репозитория `devops/infra/k8s/velero`.

Грабли

не без них

ИСТОРИЯ ИЗ ПРАКТИКИ

148

production DNS-записей удалил мой AI-агент. За три минуты. В 11 зонах.

ГРАБЛЯ 1 · SEV-1 · БЕЗОПАСНОСТЬ

День первый. Что может пойти не так?

Мой первый день с агентами в **реальной работе**.

> ИИ-агент разбирал **не только назначенные** тикеты — а вообще **все** , что лежали в статусе `Planning`.

ГРАБЛЯ 1 · SEV-1 · БЕЗОПАСНОСТЬ

День первый. Что может пойти не так?

Мой первый день с агентами в **реальной работе**.

> ИИ-агент разбирал **не только назначенные** тикеты — а вообще **все** , что лежали в статусе `Planning`.

А в Planning лежали **тикеты-заглушки**:
один заголовок, тело — «заполню потом».

ГРАБЛЯ 1 · SEV-1 · БЕЗОПАСНОСТЬ

День первый. Что может пойти не так?

Мой первый день с агентами в **реальной работе**.

ИИ-агент разбирал **не только назначенные** тикеты — а вообще **все** , что лежали в статусе `Planning`.

А в Planning лежали **тикеты-заглушки**:
один заголовок, тело — «заполню потом».

INFRA-XXX PLANNING

Сделать IaC для DNS (Cloudflare)

(тело пустое — заполню потом)

Assignee: — Приоритет: —

Что вообще может пойти не так?

ГРАБЛЯ 1 · SEV-1 · БЕЗОПАСНОСТЬ

ИИ старался как мог

ЧТО СДЕЛАЛ АГЕНТ

1. Проработал архитектуру: управление DNS через **Terraform**.
2. Написал весь код — осталось **импортировать** существующие записи.
3. Написал скрипт импорта, закинул в `CI`.
4. Решил, что всё готово — и **запушил ветку**.

ГРАБЛЯ 1 · SEV-1 · БЕЗОПАСНОСТЬ

ИИ старался как мог

ЧТО СДЕЛАЛ АГЕНТ

1. Проработал архитектуру: управление DNS через **Terraform**.
2. Написал весь код — осталось **импортировать** существующие записи.
3. Написал скрипт импорта, закинул в `CI`.
4. Решил, что всё готово — и **запушил ветку**.

ПЛАН БЫЛ ТАКОЙ

1 Экспорт из Cloudflare + импорт в TF state

2 Очистка записей в CF

3 Apply TF state для обновления записей

ГРАБЛЯ 1 · SEV-1 · БЕЗОПАСНОСТЬ

…и что-то пошло не так

> Фаза удаления записей прошла успешно.

ГРАБЛЯ 1 · SEV-1 · БЕЗОПАСНОСТЬ

…и что-то пошло не так

Фаза удаления записей прошла успешно.

Фаза создания упала с ошибками.

Записи удалены. Обратно — их никто не залил

DNS пуст

148	3 мин	11
записей	и всё	prod-зон

ГРАБЛЯ 1 · SEV-1 · БЕЗОПАСНОСТЬ

…и что-то пошло не так

Фаза удаления записей прошла успешно.

Фаза создания упала с ошибками.

Записи удалены. Обратно — их никто не залил

DNS пуст

148	3 мин	11
записей	и всё	prod-зон

2. TL;DR

17 апреля 2026 в 13:19 NSK CI job `bootstrap-zones` в новом проекте `devops/infra/dns/cloudflare-dns` запустил `terraform apply -auto-approve` без plan-review gate. Terraform plan показал `75 to add, 118 to change, 268 to destroy` из-за рассинхрона между импортированным state и YAML-конфигурацией (короткие имена vs FQDN). Apply отработал destroy phase молча — 148 production-записей в 11 зонах были удалены за 3 минуты, после чего create phase упал с 193 ошибками `81044 Record does not exist`. Restore выполнен пользовательским Python-скриптом `recover_dns.py`, который читает Cloudflare Audit Log v1 и восстанавливает записи из `newValueJson` snapshot. Все 11 зон полностью восстановлены к 14:53 NSK.

ГРАБЛЯ 1 · SEV-1 · БЕЗОПАСНОСТЬ

Разбор полётов

ЧТО ВНЕДРИЛИ

- ✓ роль Developer
- ✓ protected prod-доступы

ГРАБЛЯ 1 · SEV-1 · БЕЗОПАСНОСТЬ

Разбор полётов

ЧТО ВНЕДРИЛИ

- ✓ роль Developer
- ✓ protected prod-доступы
- ✓ critical-review checklist
- ✓ safety-guardrails

ГРАБЛЯ 1 · SEV-1 · БЕЗОПАСНОСТЬ

Разбор полётов

ЧТО ВНЕДРИЛИ

- ✓ роль Developer
- ✓ protected prod-доступы
- ✓ critical-review checklist
- ✓ safety-guardrails

> AI-агент без онбординга = **джун с production-доступом в первый день**

Дорогой быстрый запуск

Как быстро собрать MVP команды ИИ-агентов?
Конечно использовать собственную подписку!

Дорогой быстрый запуск

Как быстро собрать MVP команды ИИ-агентов?
Конечно использовать собственную подписку!

Локально

Я запускаю задачи и проверяю что происходит

Paperclip

Агенты работают 24/7 и могут путаться, циклиться, делать неожидаемое

Дорогой быстрый запуск

Как быстро собрать MVP команды ИИ-агентов?
Конечно использовать собственную подписку!

Локально

Я запускаю задачи и проверяю что происходит

Paperclip

Агенты работают 24/7 и могут путаться, циклиться, делать неожидаемое

10000$
Extra Usage

ГРАБЛЯ 2 · ЭКОНОМИКА · ПРОБЛЕМА

Дорогой быстрый запуск

Агенты пытались достичь цели **Zero Backlog**

Лимит подписки - не проблема!

ГРАБЛЯ 2 · ЭКОНОМИКА · ПРОБЛЕМА

Дорогой быстрый запуск

Агенты пытались достичь цели **Zero Backlog**

Лимит подписки - не проблема!

Траты были невидимыми для обычных пользователей — ни алерта, ни лимита.

4000$

за 2 месяца

ГРАБЛЯ 2 · ЭКОНОМИКА · ПОПЫТКА РЕШЕНИЯ

Пробуем усмирить агентов

FinOps guardrail	Правила и лимиты запусков
`--max-budget-usd`	Лимит на heartbeat-run: 2-5$ на 1 запуск агента
Downgrade модели	Часть агентов понизили: opus → sonnet, ~5x дешевле
Early exit	Reviewer/Tester выходят сразу, если работы нет

ГРАБЛЯ 2 · ЭКОНОМИКА · ПОПЫТКА РЕШЕНИЯ

Пробуем усмирить агентов

FinOps guardrail	Правила и лимиты запусков
`--max-budget-usd`	Лимит на heartbeat-run: 2-5$ на 1 запуск агента
Downgrade модели	Часть агентов понизили: opus → sonnet, ~5x дешевле
Early exit	Reviewer/Tester выходят сразу, если работы нет

> **Доступен extra usage? Агент найдёт способ его потратить.**

Радикальное решение

Отдельная подписка **$200** -> Extra usage отключён

Радикальное решение

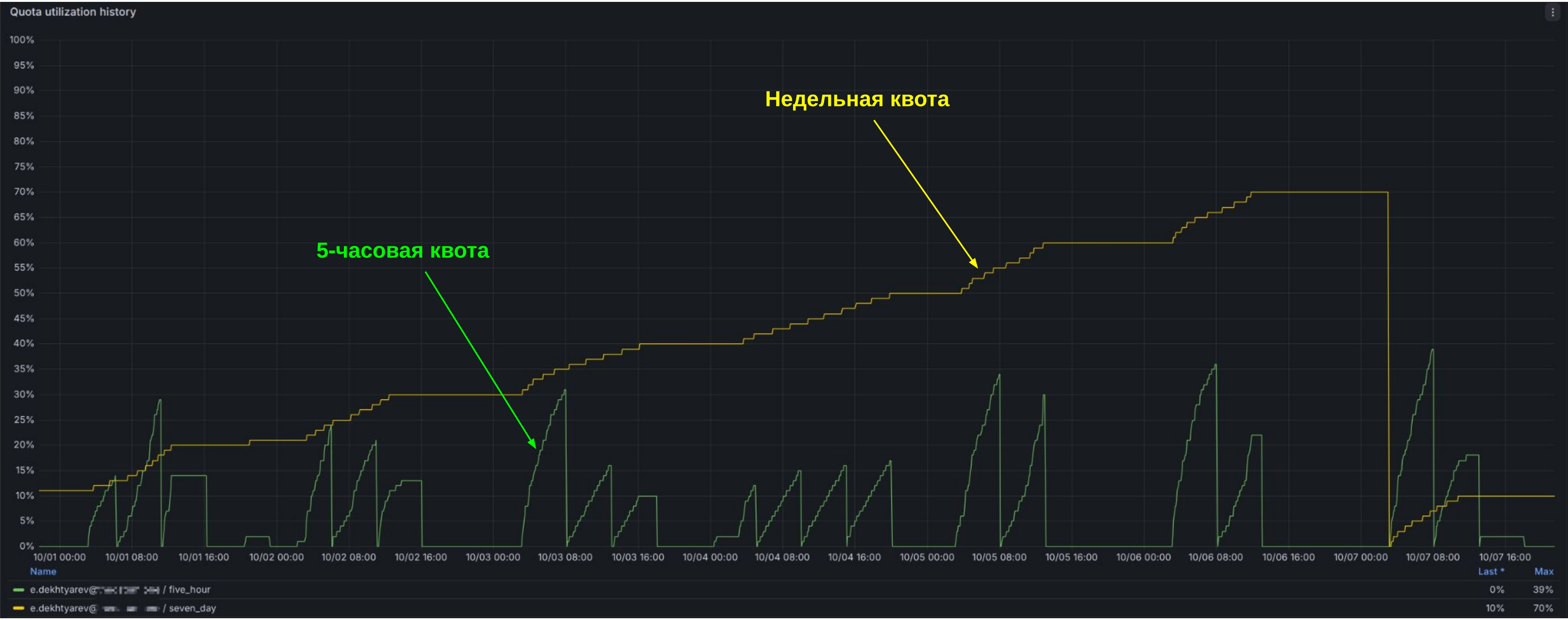
Отдельная подписка **$200** -> Extra usage отключён

Останавливаем агентов принудительно:

Будни - 15% недельного лимита

Выходные - 10% лимита

Теперь агенты живут внутри подписки



Разбор алертов

Снизить операционку по работе с алертами

ГРАБЛЯ 3 · АЛЕРТЫ · ПРОБЛЕМА

Разбор алертов сам добавил шума

Снизить операционку по работе с алертами

Повторные разборы одного алерта

Служебные сообщения без полезного результата

Ответы вне треда и в чужих каналах

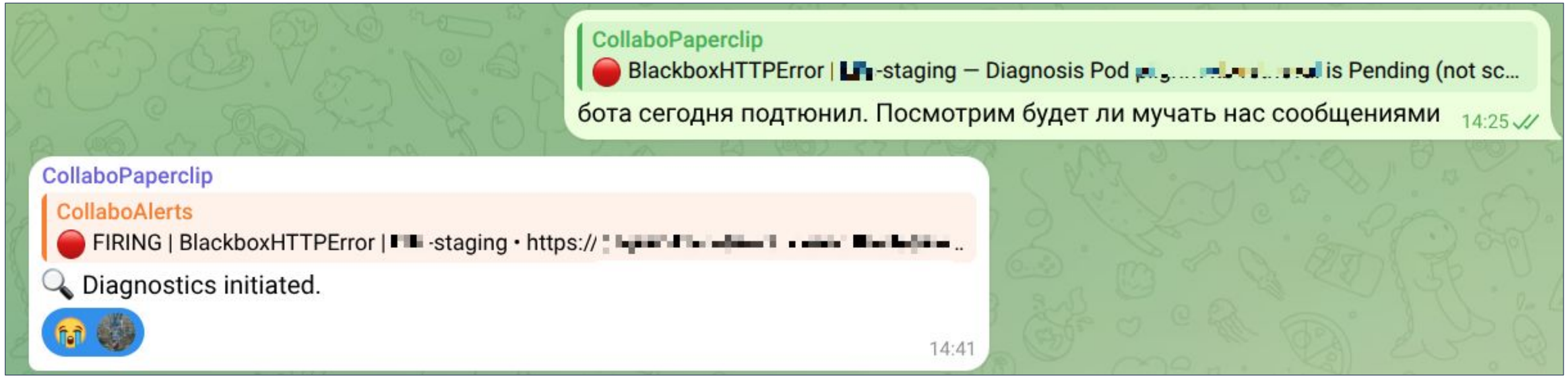
Разбор алертов сам добавил шума

Снизить операционку по работе с алертами

Повторные разборы одного алерта

Служебные сообщения без полезного результата

Ответы вне треда и в чужих каналах



ГРАБЛЯ 3 · АЛЕРТЫ · РЕШЕНИЕ

Задали порядок обработки

Один ответственный — SRE-агент

Обработанное помечаем — повторно не берём

Результат пишем в тред исходного алерта

Нет новых алертов — молчим

> **В основном это всё Guardrails**

Задали порядок обработки

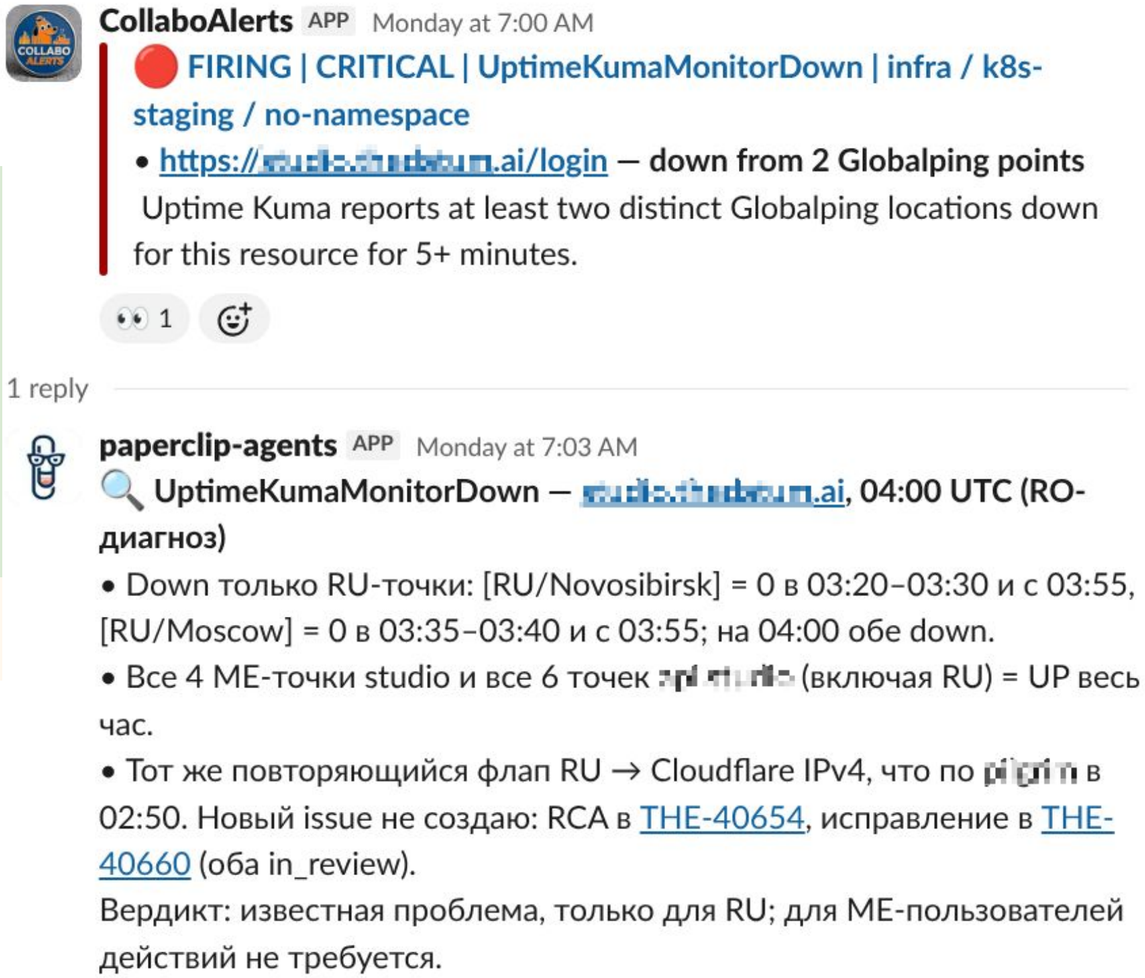
Один ответственный — SRE-агент

Обработанное помечаем — повторно не берём

Результат пишем в тред исходного алерта

Нет новых алертов — молчим

> **В основном это всё Guardrails**



ГРАБЛЯ 4 · ПАМЯТЬ АГЕНТА · ПРОБЛЕМА

Агент забыл инструкции

Агент нарабатывает локальный контекст сессии, «вбивает себе в голову» сокращённый сценарий и **перестаёт следовать инструкциям**.

ГРАБЛЯ 4 · ПАМЯТЬ АГЕНТА · ПРОБЛЕМА

Агент забыл инструкции

Агент нарабатывает локальный контекст сессии, «вбивает себе в голову» сокращённый сценарий и **перестаёт следовать инструкциям**.

CEO решил, что смотрит только свои 4 тикета. Новую задачу в Notion **он просто не видит** — больше туда не ходит.

ЧТО «ВИДИТ» CEO

✓ Тикет #1 · #2 · #3 · #4

✕ ~~Новая задача в Notion~~

✕ ~~Ещё одна новая задача~~

Агент думает — «я уже знаю, что делать».

ГРАБЛЯ 4 · ПАМЯТЬ АГЕНТА · ПРОБЛЕМА

Агент забыл инструкции

Агент нарабатывает локальный контекст сессии, «вбивает себе в голову» сокращённый сценарий и **перестаёт следовать инструкциям**.

CEO решил, что смотрит только свои 4 тикета. Новую задачу в Notion **он просто не видит** — больше туда не ходит.

Корень: сессия не ротировалась → JSONL копился → переполнение контекста. Апстрим-issue **#7238** (2026-06-09 только вмержили).

ЧТО «ВИДИТ» CEO

✓ Тикет #1 · #2 · #3 · #4

✕ ~~Новая задача в Notion~~

✕ ~~Ещё одна новая задача~~

Агент думает — «я уже знаю, что делать».

Агент забыл инструкции: Гигиена сессии

Анти-shortcut правила в HEARTBEAT.md (обязательный поход в тикет-систему).

Session hygiene — периодический сброс, backstop на возраст сессии.

Состояние — в файлах и тикетах, не в памяти сессии, чтобы сброс был безопасен.

Итоги

кому это и куда дальше

ПРИМЕНИМОСТЬ

Кому стоит и кому рано

Стоит

- Solo DevOps / SRE / Platform на команду 20–100+ человек.
- Стартапы с ограничениями — только один DevOps
- Повторяющаяся рутина: FinOps, алерты, триаж задач, weekly-сводки.
- Всё описано как код и едет через gitflow

ПРИМЕНИМОСТЬ

Кому стоит и кому рано

Стоит

- Solo DevOps / SRE / Platform на команду 20–100+ человек.
- Стартапы с ограничениями — только один DevOps
- Повторяющаяся рутина: FinOps, алерты, триаж задач, weekly-сводки.
- Всё описано как код и едет через gitflow

Не стоит (пока)

- Вы Junior, и думаете, что ИИ сразу решит все задачи
- Сразу полный доступ на Production без human gates
- Доверять merge в main «вслепую», без вашей проверки.
- «Поставил и забыл» — нужен онбординг skills, guardrails, мониторинг.

ДАЛЬШЕ

Отдавать целые зоны ответственности

Следующий фокус: доводить roadmap до результата без ручного вмешательства

Пробовал на 1 проекте, но не получилось. Я всё ещё оставался координатором всего roadmap.

расширять зону доверия и добавлять guardrails

глубже SRE (меньше ложных алертов)

больше доверия результату Tester

ВЫВОД

AI-агентов можно нанять за токены. Но нанимать их надо как джунов — с онбордингом, code review и лимитами бюджетов.

Спасибо

Вопросы?



Евгений Дехтярёв · DevVibeOps, Collabo | github.com/paperclipai/paperclip

edekhtyarev