avito.tech Москва 2024

Братья водопроводчики: латаем протечки облачного хранилища

Александр Трифанов ведущий в безопасное будущее инженер

whoami

Александр Трифанов
ведущий в безопасное будущее инженер

- Замкадыш из Новосибирска
- 7+ лет занимаюсь тестированием на проникновение и разработкой решений для продуктовой безопасности
- Неравнодушен к реверс-инжинирингу, эксплуатации бинарных уязвимостей и языку Rust

hostname

Авито

В Авито **больше трех тысяч микросервисов** на Python и Go, фронтенд на JavaScript, базы данных PostgreSQL, MongoDB и Redis, автоматические тесты, обученная на данных система модерации и поисковый движок Sphinx.

У нас уже **больше 1800 инженеров**. Мы работаем в небольших кросс-функциональных командах, каждая из которых отвечает за конкретную часть сервиса. Нашим продуктом каждый месяц пользуются десятки миллионов человек, поэтому мы умеем работать с большими нагрузками.

Kudos

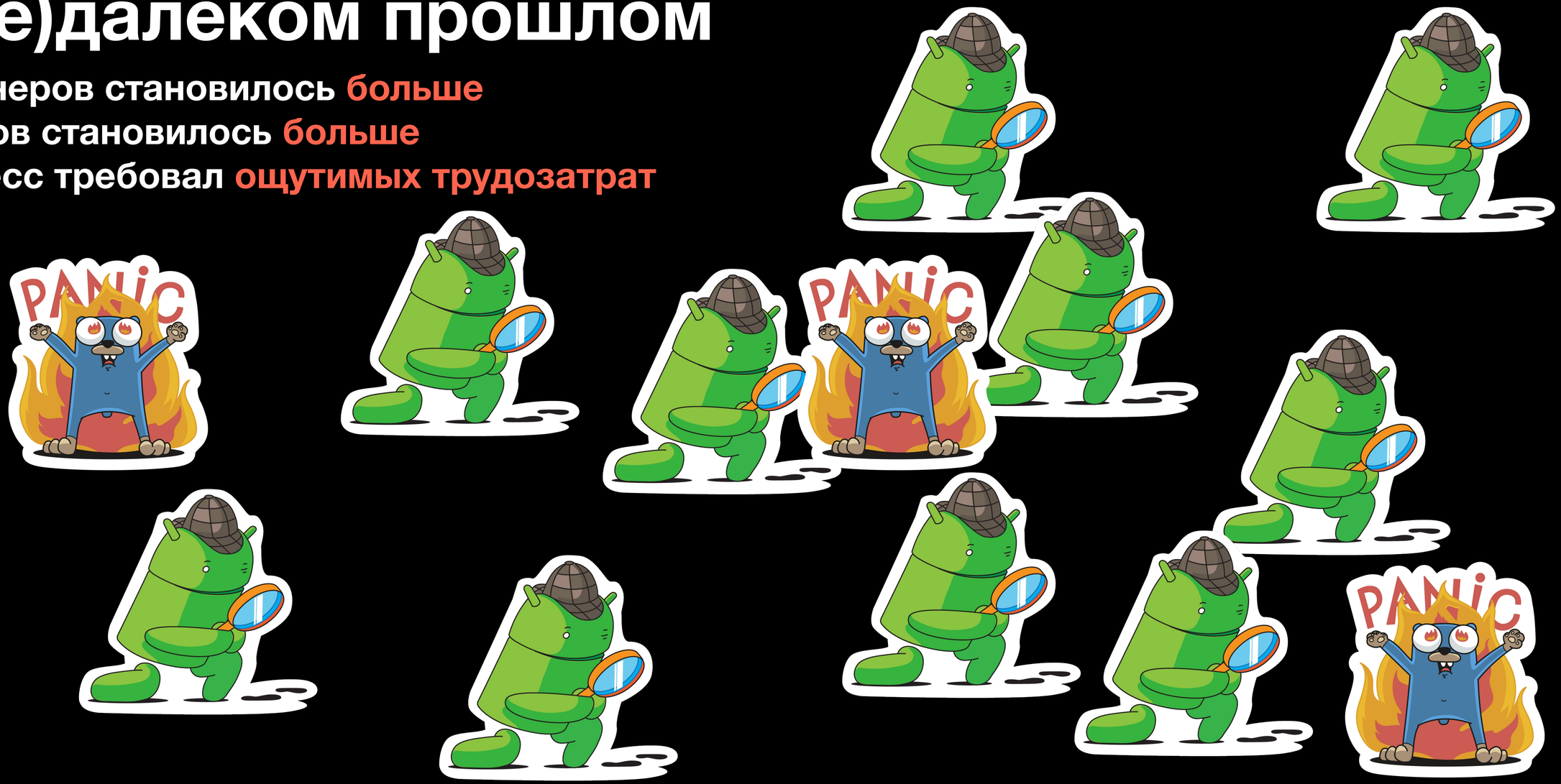
- Авторы процессов и автоматизации о которых я буду рассказывать - Артём и Николай, привет, ребята!

В далеком прошлом

- инженеров было **сильно меньше** чем 1800
- содержимое публичных файлов на гугл диске контролировалось **полуавтоматически**
- требовался **ручной труд**, но совсем **немного**

В (не)далеком прошлом

- инженеров становилось **больше**
- файлов становилось **больше**
- процесс требовал **ощутимых трудозатрат**



С чего всё началось

- у нас был процесс
- и новый стажёр со свободным временем

- так задача автоматизации процесса контроля публичных файлов на гугл диске нашла Артёма

“пароль” на диске мы тоже нашли

- это оказалась картинка с конференции в нашем офисе (я очень старался но не смог найти оригинал)
- там был плакат
- с надписью “гостевой WiFi логин:*** пароль:****”
- в тот день я узнал что гугл умеет в OCR для поиска на фотографиях

Подход быстрый

- Основная идея
  - найти публичные файлы, в которых есть ключевые слова вида “пароль” и закрыть к ним доступ

- пишем сканер, который по ключевым словам проверяет содержимое диска
- составляет репорт
- присылаем его нам

Подход быстрый

- на этом этапе мы поняли, что придется делать имперсонацию под каждую учётку и искать ключевые слова от её имени*
- репорт оказался не очень большим, его разобрали и мы пошли дальше
- репорт сделали ежемесячным

* заходит автоматика, человек приватные файлы не смотрит

Подход с полным покрытием

- Основная идея
  - найти все публичные файлы, не важно есть там ключевые слова или нет
  - присылать репорт владельцам файлов с вариантами “убрать публичный доступ” и “публичным доступ нужен, оставьте”

- Результаты:
  - поняли что стандартные инструменты от вендоров не смогли в paging и показывали заниженную статистику
  - узнали *сколько* у нас открытых файлов

Подход с полным покрытием

- Основная идея
  - найти все публичные файлы, не важно есть там ключевые слова или нет
  - присылать репорт владельцам файлов с вариантами “убрать публичный доступ” и “публичным доступ нужен, оставьте”

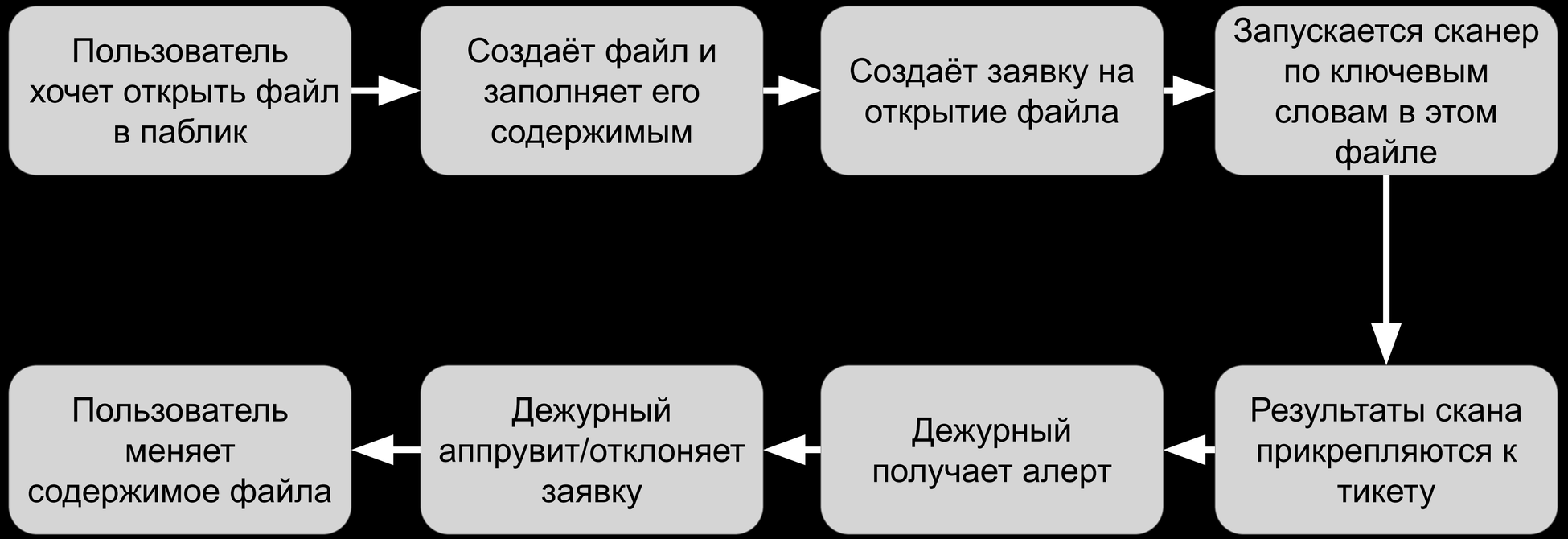
- Результаты:
  - метрика по общему числу открытых файлов немного снизилась
  - а через пару месяцев продолжила активно расти - все всегда отвечали “доступ нужен”

Подход системный

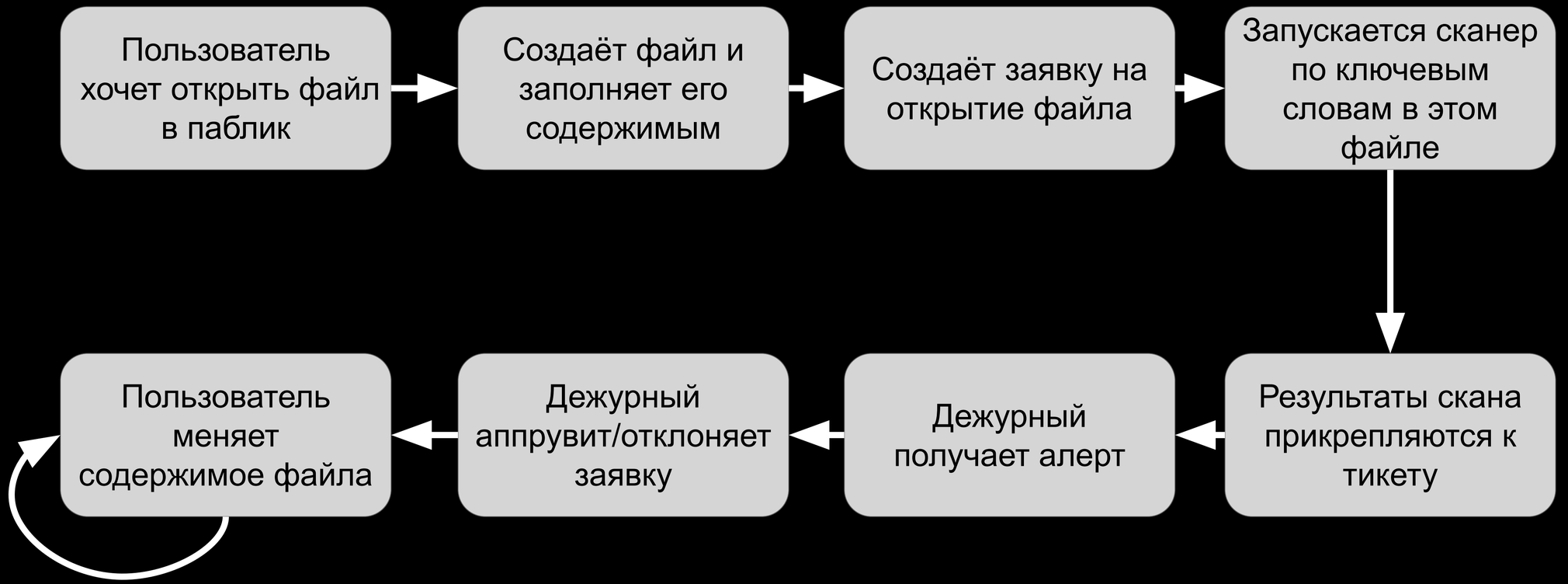
- Основная идея
  - забрать возможность открывать доступ к файлам*
  - * и придумать процесс аппрува для файлов, которым такой доступ нужен

- Результаты:
  - метрика по общему числу открытых файлов почти перестала расти

Процесс аппрува: файлы

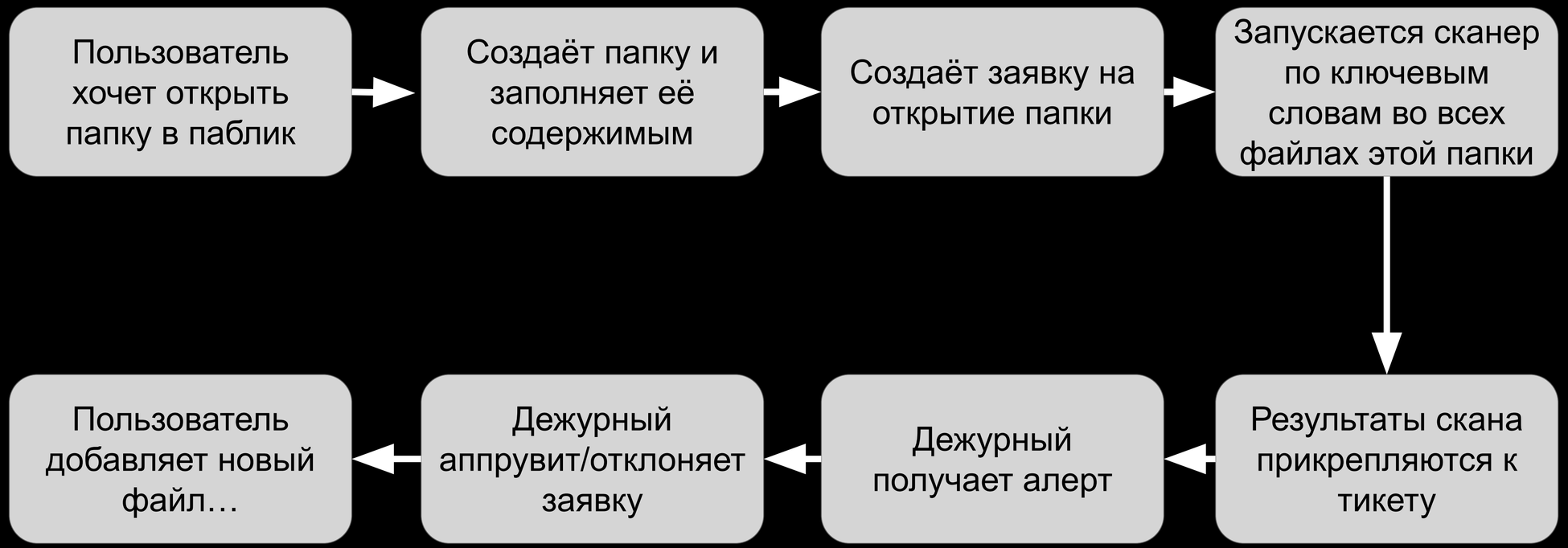


Процесс аппрува: файлы

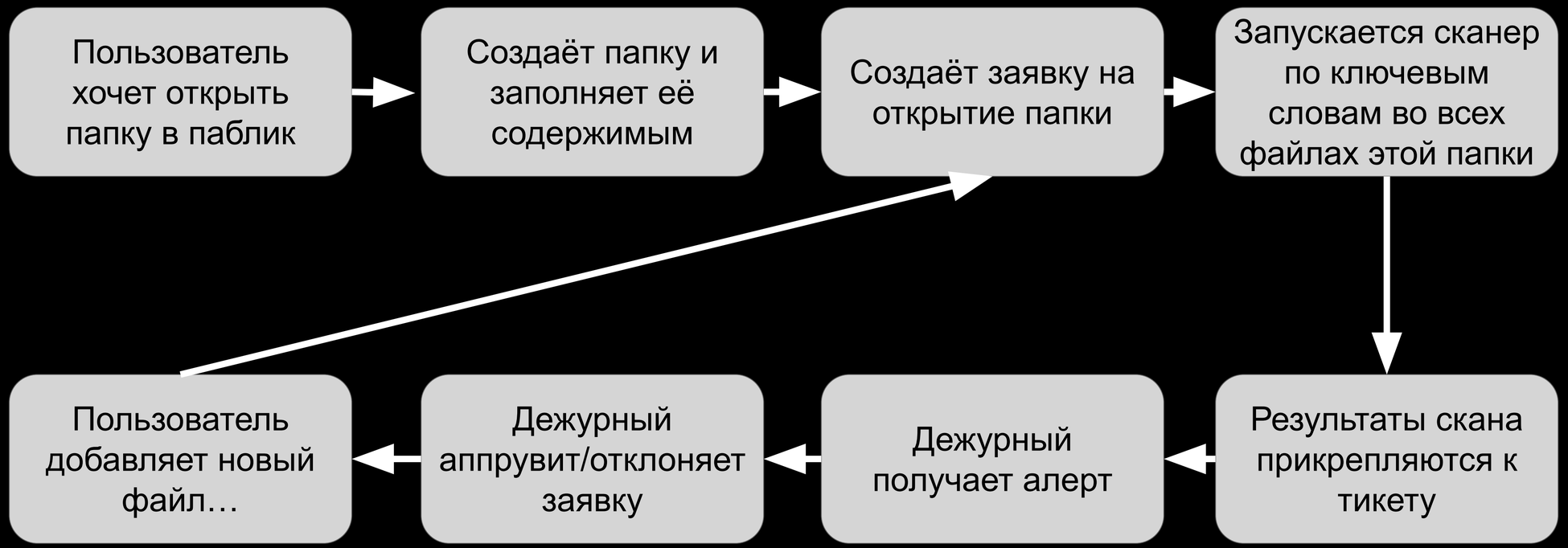


Покрывается другим процессом

Процесс аппрува: папки



Процесс аппрува: папки



Подход к техдолгу

- Основная идея
  - взять все оставшиеся открытые файлы
  - по каждому файлу слать алерт владельцу
  - если действий от владельца не было - закрывать доступ

- Результаты:
  - за несколько месяцев открытых файлов стало в ~100 раз меньше

Edge-case: Спам

- Когда изучали статистику - видели что у каждого пользователя обычно не более 5 публичных файлов
- после запуска алертов на каждый файл в поддержку пришло 3 человека, т.к. их “заваливало” письмами (счёт шёл на десятки)
- решили что проблема массовая - остановили всё, пошли разбираться
- выяснили что у нас есть всего 4 человека с тысячами публичных файлов, а у остальных действительно около 5

- Результаты:
  - рейтлимиты на количество писем одному пользователю в день:)

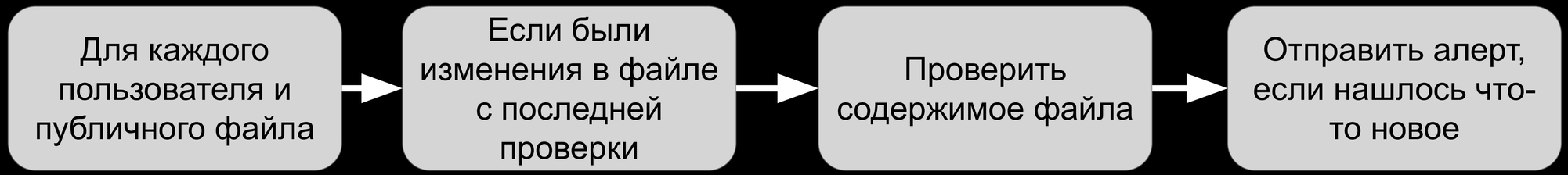
Edge-case: два домена

- Когда разбирались почему открытых файлов стало *сильно* больше - узнали что гугл через интерфейс не умеет расшаривать файлы на два домена
- А нам надо и все жмут “всем по ссылке”
- Узнали что через API гугл прекрасно шарит файлы на два домена
- Завернули этот кейс в автоматику

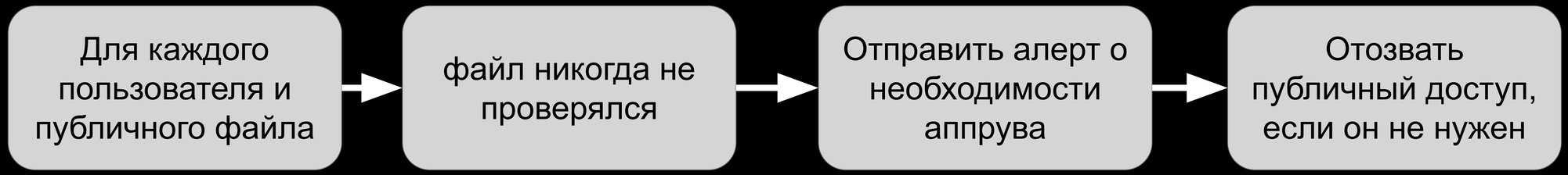
Итоговая схема

- 2 режима: техдолг и новые файлы
- режим техдолга покрывает кейс изменений в файлах

- для всех файлов



- для старых файлов



Результаты

- получили универсальный процесс контроля публичных документов
- закрыли доступ к огромному количеству архивных документов
- сумели найти трейд офф между контролем и удобством

Безопасность должна быть удобной

Александр Трифанов

ведущий в безопасное будущее инженер

Процесс вокруг инструмента иногда важнее инструмента

Никто не любит лишнюю работу

Не протекайте:)