



Управление авариями в большом интернет-проекте

Андрей Губа





О себе

20+ лет в IT

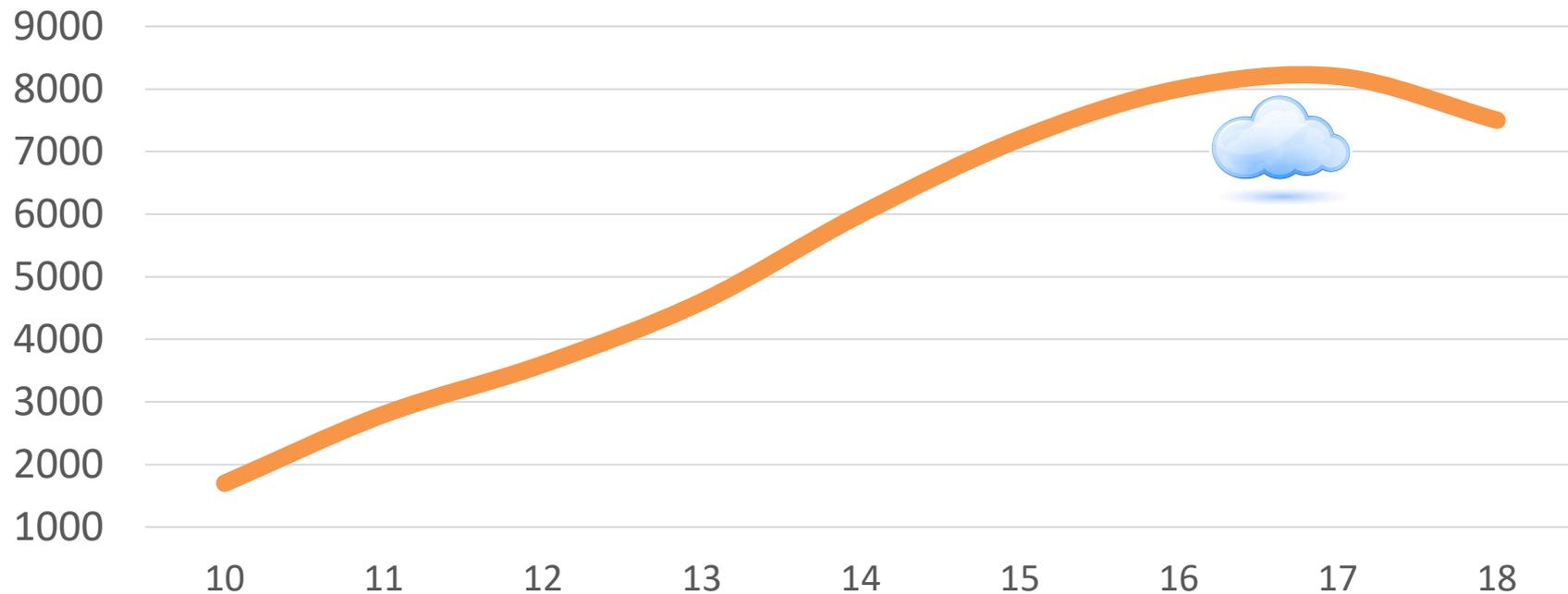
С 2008 года в компании

Технический директор



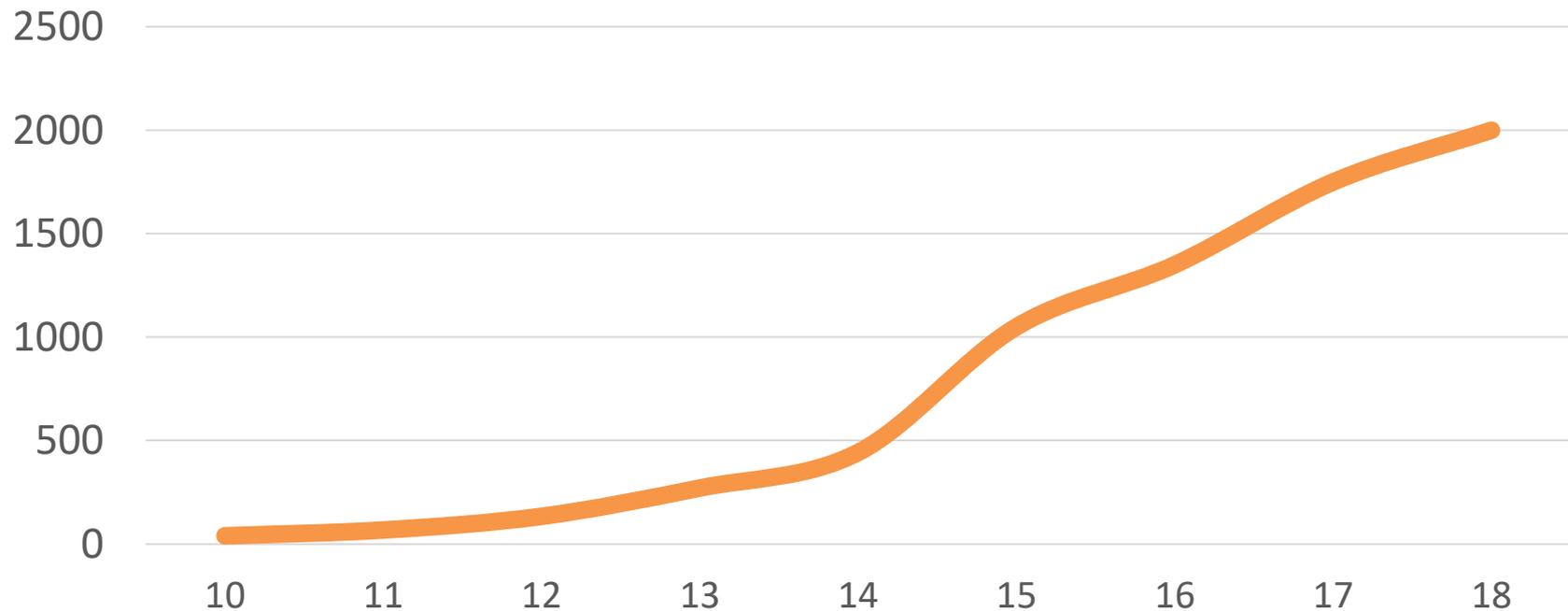
Системный администратор

Кол-во серверов



Данные: Статистика ОК, 2018

Трафик, Гбит/сек



Данные: Статистика ОК, 2018



**У кого из вас в продакшене
что-то ломается?**



**У кого из вас в продакшене
1000+ серверов?**



**У кого из вас проект
не работал 3 дня?**

4.04.2013



Альтер Эго

5 апреля 2013 в 07:34

7:34 Одноклассники все еще лежат. Такого долгого сбоя вообще, по-моему, ни у кого не было

<http://roem.ru/04-04-2013/116860/>

2006 → 2012



Сотни серверов → Тысячи серверов

Десяток сервисов → Сотни (микро)сервисов

Тысячи онлайн → Миллионы онлайн

2006 → 2012



Great Success!

2006 → 2012



Хотя увеличение технического долга как правило негативно влияет на будущее проекта, оно может быть и сознательным, компромиссным решением, продиктованным сложившимися обстоятельствами.



2006 → 2012

Частые даунтаймы:

- Выкатка кода
- Работы с данными
- Аварии



Авария в коллекторе в Москве нарушила работу операторов СВЯЗИ

18:05 19.01.2012 (обновлено: 18:18 19.01.2012)

0 227 0 0

Авария, произошедшая утром в четверг в коллекторе, идущем от Савеловской эстакады в сторону Бутырского вала, повлияла на пользователей как сотовых, так и проводных сетей связи.

МОСКВА, 19 янв - РИА Новости, Александр Бумагин. [Пожар в кабельном коллекторе на Бумажном проезде](#) под Савеловской эстакадой в Москве нарушил работу нескольких операторов связи в четверг, выяснил корреспондент РИА Новости.

Авария, произошедшая утром в четверг в коллекторе, идущем от Савеловской эстакады в сторону Бутырского вала, повлияла на пользователей как сотовых, так и проводных сетей связи.





Хотим работать при недоступности ДЦ!



Хотим работать при недоступности ДЦ!



- Распределить данные

Хотим работать при недоступности ДЦ!



- Распределить данные
- Новые системы хранения данных

Хотим работать при недоступности ДЦ!



- Распределить данные
- Новые системы хранения данных
- Мигрировать данные



Хотим работать при недоступности ДЦ!

- Распределить данные
- Новые системы хранения данных
- Мигрировать данные
- Переделать сервисы



Great Success!

4.04.2013





4.04 или Как правильно положить портал

- Выбрать правильное время
- Давай с твоего компа поправим!
- Баг в CFEngine (CFE-270)
- Баг в своем коде
- Быстро применить изменения на весь продакшн

4.04 Недоступность всех Linux серверов





4.04 Недоступность всех Linux серверов

- Вручную (в датацентре и удаленно используя IPMI)



4.04 Недоступность всех Linux серверов

- Вручную (в датацентре и удаленно используя IPMI)
- Полуавтоматически



4.04 Недоступность всех Linux серверов

- Вручную (в датацентре и удаленно используя IPMI)
- Полуавтоматически
- Внедрили SOL (IPMI)



4.04 Недоступность всех Linux серверов

- Вручную (в датацентре и удаленно используя IPMI)
- Полуавтоматически
- Внедрили SOL (IPMI)
- Десятки моделей серверов

4.04 Недоступность данных (tmpfs, сбойные диски)





4.04 Недоступность данных (tmpfs, сбойные диски)

Восстанавливали данные:

- С локальных дисков и накатывали логи



4.04 Недоступность данных (tmpfs, сбойные диски)

Восстанавливали данные:

- С локальных дисков и накатывали логи
- Из бэкапов



4.04 Недоступность данных (tmpfs, сбойные диски)

Восстанавливали данные:

- С локальных дисков и накатывали логи
- Из бэкапов
- Из второстепенных систем



4.04 Недоступность данных (tmpfs, сбойные диски)

Восстанавливали данные:

- С локальных дисков и накатывали логи
- Из бэкапов
- Из второстепенных систем

- Изменили частоту снэпшотов



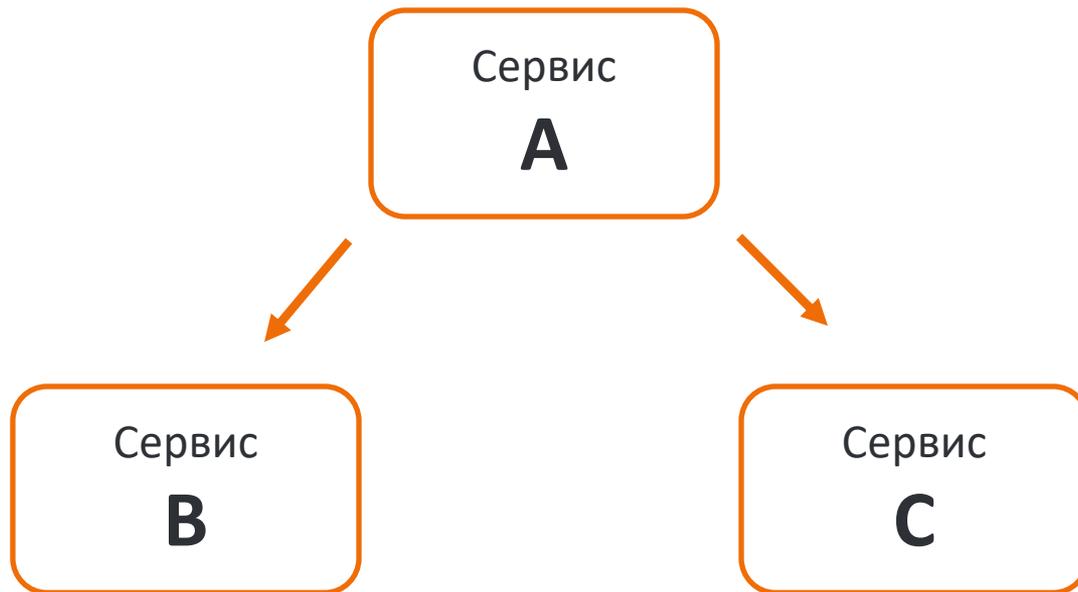
4.04 Недоступность данных (tmpfs, сбойные диски)

Восстанавливали данные:

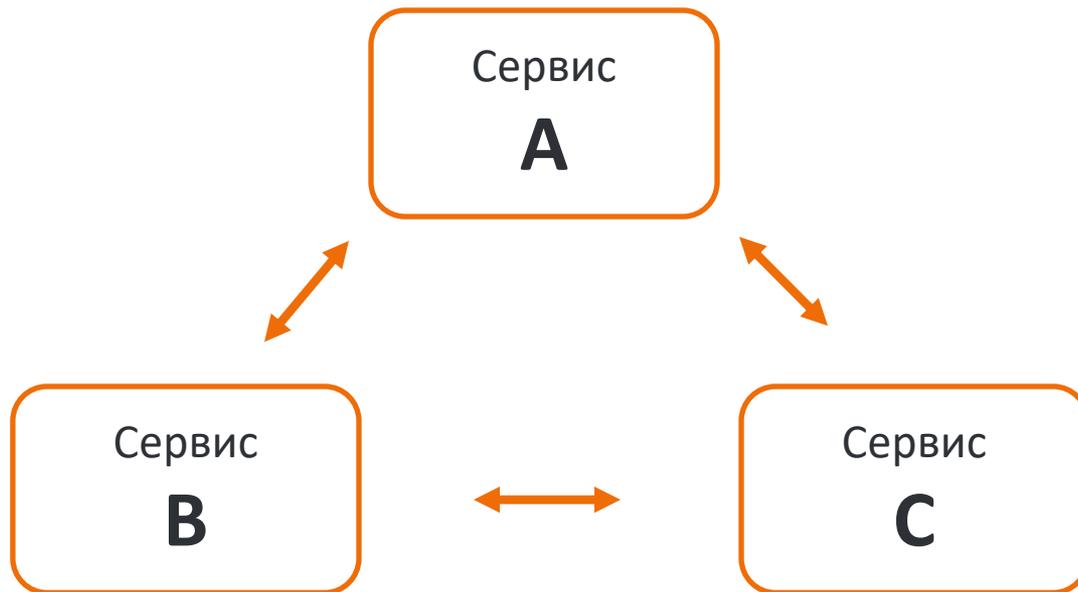
- С локальных дисков и накатывали логи
- Из бэкапов
- Из второстепенных систем

- Изменили частоту снапшотов
- Внедрили новые хранилища (Voldemort, Cassandra)

4.04 Зависимости между сервисами



4.04 Зависимости между сервисами





4.04 Зависимости между сервисами

- Починка сервисов-зависимостей



4.04 Зависимости между сервисами

- Починка сервисов-зависимостей
- Отпиливание зависимостей



4.04 Зависимости между сервисами

- Починка сервисов-зависимостей
- Отпиливание зависимостей
- Использование рубильников (выключателей в коде)



4.04 Зависимости между сервисами

- Починка сервисов-зависимостей
- Отпиливание зависимостей
- Использование рубильников (выключателей в коде)
- Внедрили тестирование зависимостей



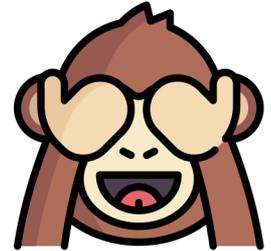
4.04 Зависимости между сервисами

- Починка сервисов-зависимостей
- Отпиливание зависимостей
- Использование рубильников (выключателей в коде)
- Внедрили тестирование зависимостей
- Доделали рубильники

4.04 Проблемы с мониторингом и статистикой



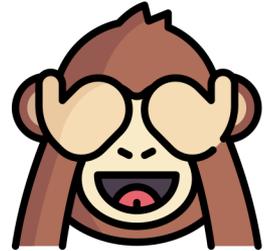
- Проверяли вручную



4.04 Проблемы с мониторингом и статистикой



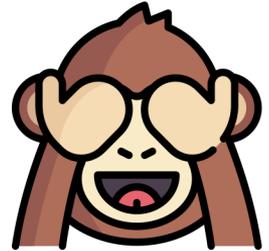
- Проверяли вручную
- Изменили логику работы и конфигурацию



4.04 Проблемы с мониторингом и статистикой



- Проверяли вручную
- Изменили логику работы и конфигурацию
- Дополнительно зарезервировали служебные системы



4.04 Проблемы с сетью (STP/portfast)





4.04 Проблемы с сетью (STP/portfast)

- Потратили много времени на поиск и устранение



4.04 Проблемы с сетью (STP/portfast)

- Потратили много времени на поиск и устранение
- Централизованно конфигурируем сеть

4.04 Быстрое распространение конфигурации



4.04 Быстрое распространение конфигурации



- Обязательный Review



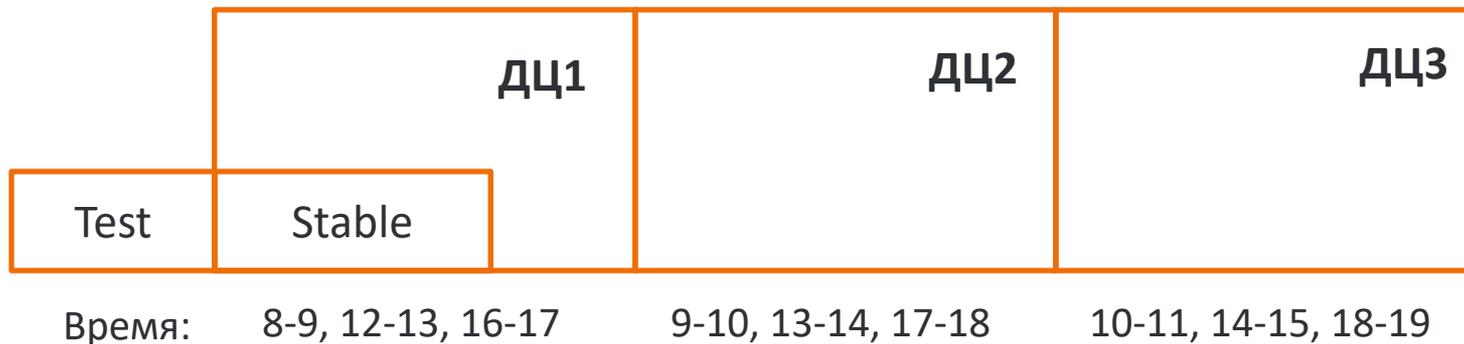
4.04 Быстрое распространение конфигурации

- Обязательный Review
- Обязательное тестирование: testing, stable, production



4.04 Быстрое распространение конфигурации

Изменили процедуру



4.04 Аварийный план





4.04 Аварийный план

- Список лиц и контактов



4.04 Аварийный план

- Список лиц и контактов
- Распределение ролей и ответственности



4.04 Аварийный план

- Список лиц и контактов
- Распределение ролей и ответственности
- Приоритезированный список сервисов



4.04 Аварийный план

- Список лиц и контактов
- Распределение ролей и ответственности
- Приоритезированный список сервисов
- Описание способов и инструментов для починки



4.04 Аварийный план

- Список лиц и контактов
- Распределение ролей и ответственности
- Приоритезированный список сервисов
- Описание способов и инструментов для починки
- Инструкция для координатора



4.04 Аварийный план

- Список лиц и контактов
- Распределение ролей и ответственности
- Приоритезированный список сервисов
- Описание способов и инструментов для починки
- Инструкция для координатора
- Инструкция на случай падения ДЦ



4.04 Аварийный план

- Список лиц и контактов
- Распределение ролей и ответственности
- Приоритезированный список сервисов
- Описание способов и инструментов для починки
- Инструкция для координатора
- Инструкция на случай падения ДЦ
- Регламент взаимодействия с партнерами

4.04 Тестирование аварий



4.04 Тестирование аварий



- Разные координаторы



4.04 Тестирование аварий

- Разные координаторы
- Разные сценарии



“Т.к. мы проводим тестирование плана действий при аварии и все события вымышленные, то показатели графиков, мониторов и содержимое некоторых логов я буду рассказывать сам, по запросу.”

Dmitry Samsonov @work, вторник 14:51

15:09

Vadim Bara

This message has been removed. 

обзваниваю Илью, АГ, Колю А.
и Лену

15:09

Вадим, **для мониторщиков VPN работает!**

15:10

Vadim Bara

сообщаю в чаты о запрете проведения всех работ

15:11

Alexander Fat

Делегирую роль координатора аварии Андрею Д

15:11

**ОСТАНАВЛИВАЕМ ТЕСТ. ПРОБЛЕМЫ В РАБОТЕ
ПОРТАЛА**

15:11



4.04 Тестирование аварий

- Разные координаторы
- Разные сценарии
- Отключение ДЦ



4.04 Тестирование аварий

- Разные координаторы
- Разные сценарии
- Отключение ДЦ
- Повторение!



Инцидент-менеджмент в ОК



Инциденты: регистрация



Incidents / INCOKL-505797

srvd3576 - No ping

[Edit](#) [Comment](#) [Assign](#) [More](#) [Close](#) [Return to Investigation](#) [Workflow](#)

[Link](#) [Share](#) [Export](#)

Details

Type:	SW Failure	Status:	RESOLVED
Priority:	Major	Resolution:	Done
Labels:	None		
Аномалию зафиксировал:	zabbix		
Group Name:	cloud-rs-wan-dc-part1		
Hostname:	srvd3576		
Links:	CMDB , Cacti		
Rack:	H1E100		
Zabbix Name:	zabbix.dl.odkl.ru		

Triggers

	1426527	srvd3576 - No ping		27.08.2018 11:46
--	-------------------------	--------------------	--	------------------

People

Assignee:	Solomov German Assign to me
Reporter:	Zabbix Service
Responsible admin:	Narodovich Maxim
Votes:	0 Vote for this issue
Watchers:	2 Start watching this issue

Dates

Created:	Today 11:47
Updated:	Today 11:53
Resolved:	Today 11:53
Starting time:	27.08.2018 11:46
Ending time:	27.08.2018 11:52
Time zone:	Europe/Moscow



Инциденты: расследование и устранение

▼ ● **Goncharov Alexander** added a comment - 16.08.2018 14:53

[15.08.2018 20:59] Сергей Тревогода: по логам ваще не понять, в это время чисто всё, на проксиках ошибок нет, Олег только там с кассандрой запускал и было время повышено на сторадже сообщений - https://charts.odkl.ru/reports?report=4036&id_repository=35578&link=410350

но слабый довод, т к клиентских ошибок не было

[15.08.2018 21:01] Андрей Тимофеев: Были таймауты

[15.08.2018 21:01] Андрей Тимофеев: И вероятно из-за идс

[15.08.2018 22:19] Александр Гончаров: тут ошибки с 16 начались примерно

[15.08.2018 22:19] Александр Гончаров: но сейчас растут дальше

[15.08.2018 22:21] Vladislav Dolganov: Полагаю это лимиты <https://charts.odkl.ru/reports?link=410361>

[15.08.2018 22:22] Александр Гончаров: это новые лимиты выставили? раньше ж норм было

[15.08.2018 22:24] Vladislav Dolganov: там появилась новая фишка в лимитах, ее пытались включить вчера или позавчера, но оказалось, что в прошлый апдейт коммит не зашел. Вероятно не выключили обратно. Там это сделано для стикерка, насколько я знаю

[15.08.2018 22:25] Vladislav Dolganov: могу вырубить пока

[15.08.2018 22:28] Vladislav Dolganov: Давайте пока отключу, аналитики потом включают, пока не горит

[15.08.2018 22:30] Vladislav Dolganov: вырубил

▼ ● **Goncharov Alexander** added a comment - 17.08.2018 21:39

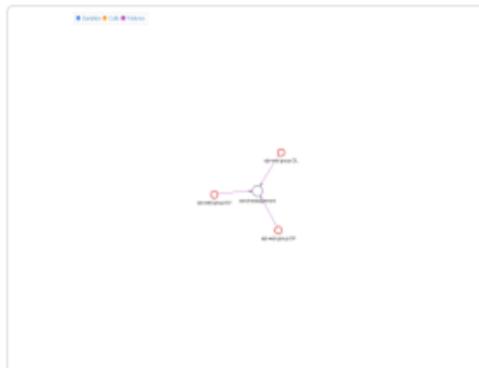
Ошибки повышены изза [XPRM-11770](#)



Инциденты: дополнение информацией



ejb-web-group-KV - sendme
16.08.2018 14:53 334 kB



Graph.png
15.08.2018 18:49 20 kB



За сообщения 16.08.2018
16.08.2018 14:53 213 kB

Issue Links



is affected by

[? XPRM-11770](#) Лимиты отправки урлов на пользов...



[OPEN](#)



Инциденты: хронология



Incidents / INCOKL-224529

Проблемы в работе портала



Edit



Comment

Assign

More ▾

Start fixing

Resolve

Workflow ▾

Хронология:

15:02 началась проблема с взбями. Взбы начали сообщать о проблеме с сессионными кэшами

15:08 обнаружена проблема с взбями. Провала на моб версии нет.

15:14 создан чат о проблеме. начался разбор. Остановлен апдейт [UPDODKL-2145](#).

15:24 отключена система сообщений.

15:36 решено откатить изменения по сессионным кэшам [UPDODKL-2145](#)

15:42 включена система сообщений

15:46 произведен откат по [UPDODKL-2145](#)

16:00 начали отключать сервисы по плану аварии

• • •

17:26 возвращен ACL в lan DL

17:27 лента включена

17:34 ленту отключили из-за проблем с ней

18:05 графики возвращаются.

19:00 графики активности вернулись в норму. онлайны продолжают добираться.

19:44 тестирование портала завершено.



Инциденты: разбор и задачи

- ✓  **Arkhangelskiy Dmitry** added a comment - 27.02.2018 17:08

а разборчик был?

- ✓  **Narodovich Maxim** added a comment - 27.02.2018 19:11

разбор завтра, 28 февраля, в 15-30 МСК

- ✓  **Narodovich Maxim** added a comment - 28.02.2018 16:38 - **edited**

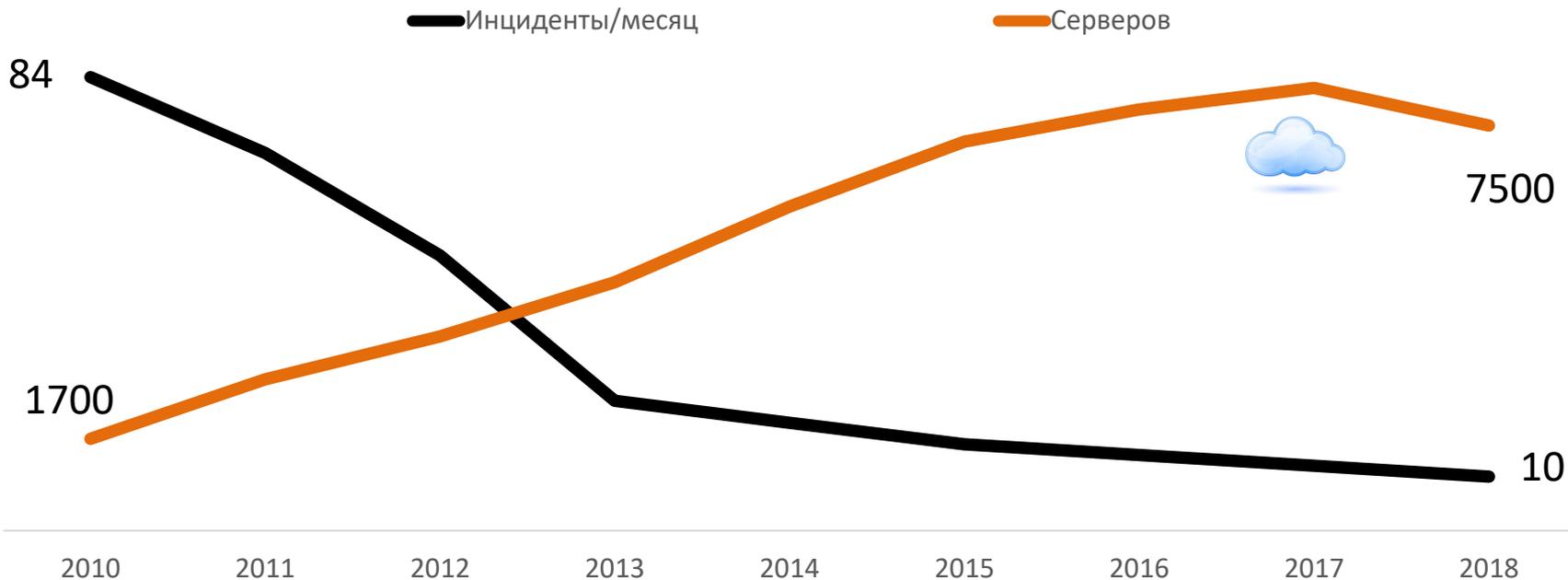
Summary:

Будет проведена работа с командой мониторинга - **ADMODKL-55345**

График был добавлен в мониторинг

Сабтаски Access Closure будут назначаться на **Yakupova Elena**, реализация за **Kurbanov Sergo**

Инциденты с эффектом для пользователей





Итого. Чтобы управлять авариями

- План
- Инструменты
- Обученные люди
- Повторение



Спасибо за внимание!

<http://habrahabr.ru/company/odnoklassniki/blog/>

