

# Инцидент-менеджмент



Сергей  
Андрюнин



**Сергей Андрюнин**

DevOps инженер

RTLabs



# План занятия

1. [Введение](#)
2. [Коммуникация](#)
3. [Реакция](#)
4. [Дежурство](#)
5. [Постмортем](#)
6. [Итоги](#)
7. [Домашнее задание](#)



# Введение



# Введение

**Инцидент** - это событие, которое вызывает нарушение или снижение качества обслуживания, и требующее экстренного реагирования.

**Инцидент считается устраненным**, когда затронутая служба или система возобновляет работу в штатном режиме.

**Инциденты могут сильно различаться по степени серьезности**, начиная от сбоя всей информационной системы до небольшого числа пользователей, имеющих периодические ошибки.

---

# Введение

**Управление инцидентами** - это процесс, обеспечивающий:

- Своевременное реагирования на незапланированные события
- Прерывание обслуживания
- Восстановления работоспособности информационных систем

---

# Введение

**Управление инцидентами можно разделить на следующие разделы:**

- Коммуникация
- Реакция
- Дежурство
- Постмортем

Для DevOps-команд данный процесс главным образом отличается эскалированием инцидентов в соответствии с профессиональными знаниями, а не занимаемыми должностями.

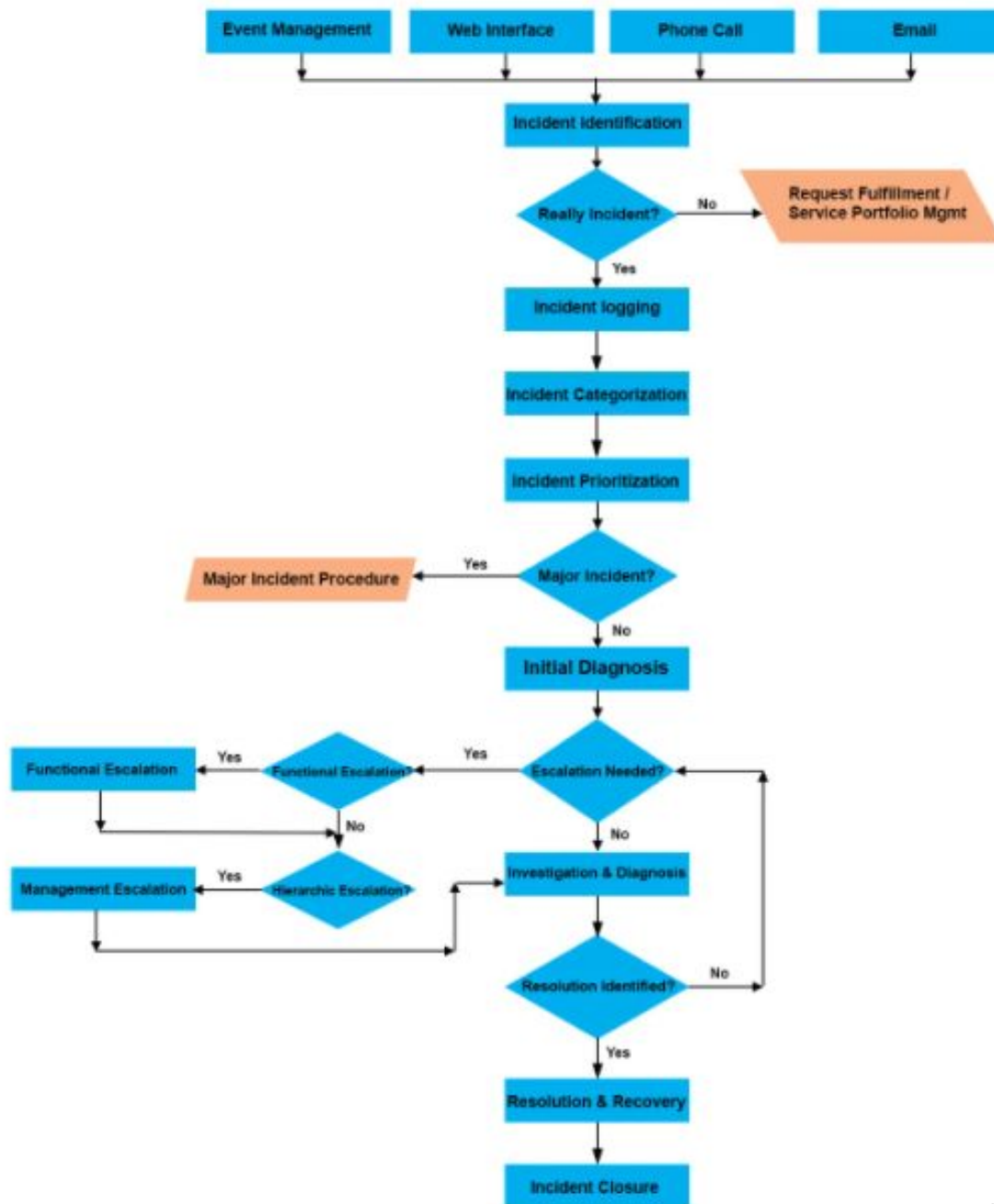
Это связано со сложностью используемых систем и невозможности внести изменения для восстановления их работоспособности.

# Введение

Управление инцидентами  
стандартизовано ITIL  
foundation.

ITIL инцидент-менеджмент  
охватывает не только  
непосредственно работу  
с инцидентами, но также  
расчёт KPI, финансовые  
потери и т.д.

Incident Support & Notification







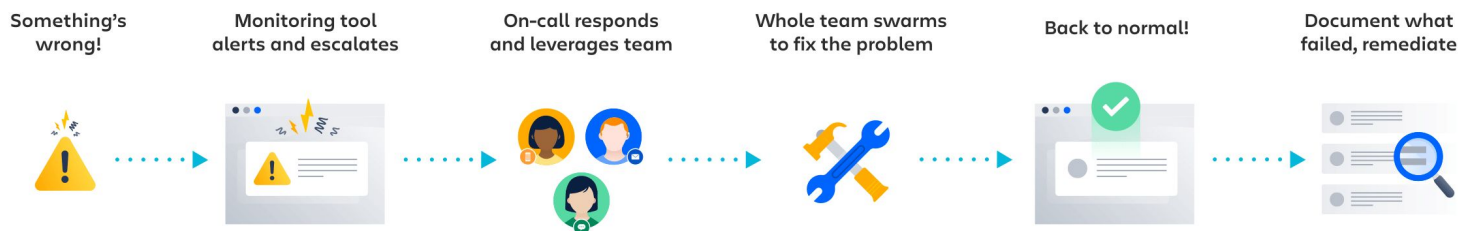
# Коммуникация

# Коммуникация

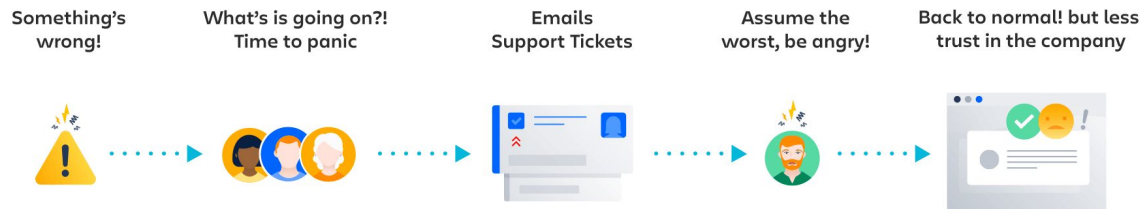
**Коммуникация (incident communication)** - это процесс оповещения пользователей сбое или снижение производительности.

Это особенно важно для веб-сервисов и программных сервисов, где предоставляется круглосуточная доступность использования.

## WHAT YOU SEE



## WHAT YOUR CUSTOMERS SEE



---

# Коммуникация

В случае отсутствия коммуникации клиенты могут усомниться в надежности вас как поставщика услуг.

Вы также потеряете будущих клиентов из-за отсутствия доверия.

Моральный дух внутри команды может пострадать и привести к снижению производительности.

Чтобы этого избежать, достаточно просто держать клиентов в курсе, сообщая о том, что происходит и что вы делаете, чтобы решить проблему. Они поймут ситуацию и будут гораздо менее негативно реагировать на всю ситуацию.

Правильно настроенная коммуникация может не только снизить негативное влияние на эмоции клиентов, но также и увеличить скорость нахождения проблемы.

---

# Коммуникация

## Определите в команде, что такое инцидент

- Разделите инциденты по уровням, согласно критичности.
- Каждый член команды должен понимать, к какому типу относится инцидент.
- Не должно быть “серых” инцидентов, которые нельзя отнести к какому-то типу (соответственно и считать за инцидент).
- Инциденты, связанные с безопасностью считаются наиболее критичными.

---

# Коммуникация

**Определите каналы коммуникации и шаблоны сообщений.** Это позволяет верно и быстро донести информацию до пользователей.

*Основные каналы коммуникаций:*

- Отдельный status-page
- Email
- Chat-tool
- Social media
- SMS

Обычно шаблон сообщения должен содержать графы для общего названия сбоя, краткого описания с указанием мест деградации сервиса и время, в течении которого обновится информация о сбое.

# Коммуникация

## Service Disruption

Title: **\*\*COMPANY\_NAME\*\*** Service Disruption

*We are currently experiencing a service disruption.*

*Our **\*\*ADD\_TEAM\*\*** team is working to identify the root cause and implement a solution. **\*\*ADD\_GENERAL\_IMPACT\*\*** users may be affected.*

*We will send an additional update in **\*\*NEXT\_UPDATE\_TIME\*\*** minutes.*

## Can't Log In

Title: Trouble logging in

*Some users may be experiencing trouble when logging in to **\*\*SERVICE\_NAME\*\***. Our **\*\*ADD\_TEAM\*\*** team is currently investigating issues related to **\*\*ADD\_SUSPECTED\_ROOT\_CAUSE\*\***.*

*We will send an additional update in **\*\*NEXT\_UPDATE\*\*** minutes.*

## Issue with the support site

Title: Support portal unavailable

*Our support portal may not be accessible for some users. We are currently investigating the issue. In the meantime if something is urgent, contact us through **\*\*SECONDARY\_SUPPORT\_EMAIL\*\*** or on twitter **\*\*ADD\_TWITTER\_HANDLE\*\***.*

*We will send an additional update in **\*\*NEXT\_UPDATE\*\*** minutes.*



# Реакция

---

# Реакция

**Реакция на инциденты** - это процесс реагирования организации на ИТ-угрозы, такие как: кибератаки, нарушение безопасности и простои информационных систем.

**Реакция на инциденты делится на следующие составляющие:**

1. Определение инцидента
2. Определение каналов коммуникации внутри команды
3. Оценка влияния и уровня критичности
4. Коммуникация с пользователями
5. Эскалация на ответственных лиц
6. Распределение ролей
7. Решение инцидента

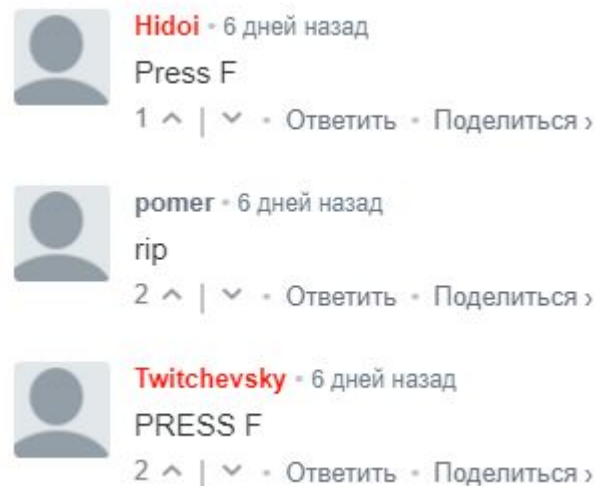


# Реакция

## 1. Определение инцидента

В идеале инцидент обнаруживается инструментами мониторинга и оповещения и превентивно передаётся команде для устранения.

Но иногда об инциденте можно узнать только от пользователей, например из Twitter или из обращений в службу поддержки.



Взято с сайта: [downdetector.com](https://downdetector.com)

# Реакция

## 2. Определение каналов коммуникации внутри команды

На данном этапе цель состоит в том, чтобы установить и сосредоточить все коммуникации команды по инцидентам в хорошо известных местах, таких как:

- корпоративный месенджер, например Slack
- видео или голосовая связь, если команда не имеет единого физического расположения

Slack и большинство других мессенджеров позволяют команде устанавливать единое инфополе для общения (например, thread).

Команда должна использовать это инфополе для информации об инциденте и передаче полезных ссылок или документов.

Оперативная связь по инциденту может происходить в голосовой или видеосвязи, в идеале - с возможностью записи для последующего разбора.

# Реакция

## 3. Оценка влияния и уровня критичности

После определения командного канала коммуникации необходимо оценить инцидент, чтобы команда могла решить, каким образом построить общение с пользователями и какие компетенции необходимо привлечь для решения инцидента.

Использование системы уровней критичности помогает быстро определить происшествие и сообщить о нем.

Даже простая фраза *«у нас может произойти super-critical»* сразу вносит понимание в серьезность вопроса, даже до получения дополнительной информации и определяет круг людей, которых нужно привлечь для эскалации инцидента.

Все уровни критичности должны быть определены *на уровне стандартов организации и задокументированы*.

\*Мы уже рассмотрели коммуникацию с пользователями, поэтому пропустим 4 шаг и перейдем сразу к 5-му.

# Реакция

## 5. Эскалация на ответственных лиц

Чаще всего инцидент решают именно те, кто реагирует на него.

Также часто необходимо привлечь к инциденту другие команды - например, security команду или ops.

Необходимо заранее определять круг лиц, которых можно привлечь из смежных команд и каналы коммуникации с этими лицами.

В случае отсутствия ответа от назначенного лица смежной команды не должно быть неопределенности в замещающем его лице.

---

# Реакция

## 6. Распределение ролей

Внутри команды заранее определяется лицо, отвечающее за управлением по решению инцидента.

Данное лицо управляет процессом решения инцидента и распределяет роли среди команды эскалации.

Как только новый человек подключился к решению инцидента - управляющий назначает ему роль, в соответствии с его компетенциями и возможностями.

Каждая роль ограничена в ответственности и возможностях.

Таким образом, исключаются ситуации, когда несколько человек одновременно производят одну и ту же операцию (например рестарт службы или сервера). Также всегда определена техническая роль человека, в соответствии с его компетенцией.

---

# Реакция

## 6. Распределение ролей - ключевые роли:

- Управляющий инцидентом
  - несет общую ответственность и полномочия в отношении инцидента
  - имеет право предпринимать любые действия, необходимые для разрешения инцидента
- Технический руководитель
  - исследует причины инцидента
  - принимает решение об изменениях
  - руководит технической командой
- Менеджер по коммуникациям
  - несет ответственность за отправку клиентских оповещений об инциденте

---

# Реакция

## 7. Решение инцидента

Не существует идеального процесса для решения инцидента.

Можно лишь провести некоторую систематизацию на верхнем уровне:

- Исследуйте, что происходит. Поделитесь своими наблюдениями с командой.
- Разработайте теорию о причинах происходящего.
- Проведите эксперимент, подтверждающий или опровергающий вашу теорию.
- Повторите эти пункты до устранения инцидента.

Инцидент считается решенным, когда его влияние на бизнес исчезает.

После решения инцидента команда переходит к анализу произошедшего и написанию постмортема.



# Дежурство





# Дежурство

**Дежурство** - это практика назначения конкретных людей, которые будут доступны в определенное время для ответа в случае срочного обслуживания, даже если инцидент произошел вне его рабочего времени.

**Дежурный** - критически важная обязанность многих специалистов, которые предоставляют услуги в тех случаях, когда клиенты ожидают круглосуточной доступности.

Обычно члены команды по очереди назначаются дежурными, либо посуточно, либо только в нерабочее время.

Наряду с мониторингом и оповещением, дежурный инженер имеет право немедленно реагировать на любые перебои в доступности услуг.

# Дежурство



Взято с сайта: dtln.ru



# Постмортем

---

# Постмортем

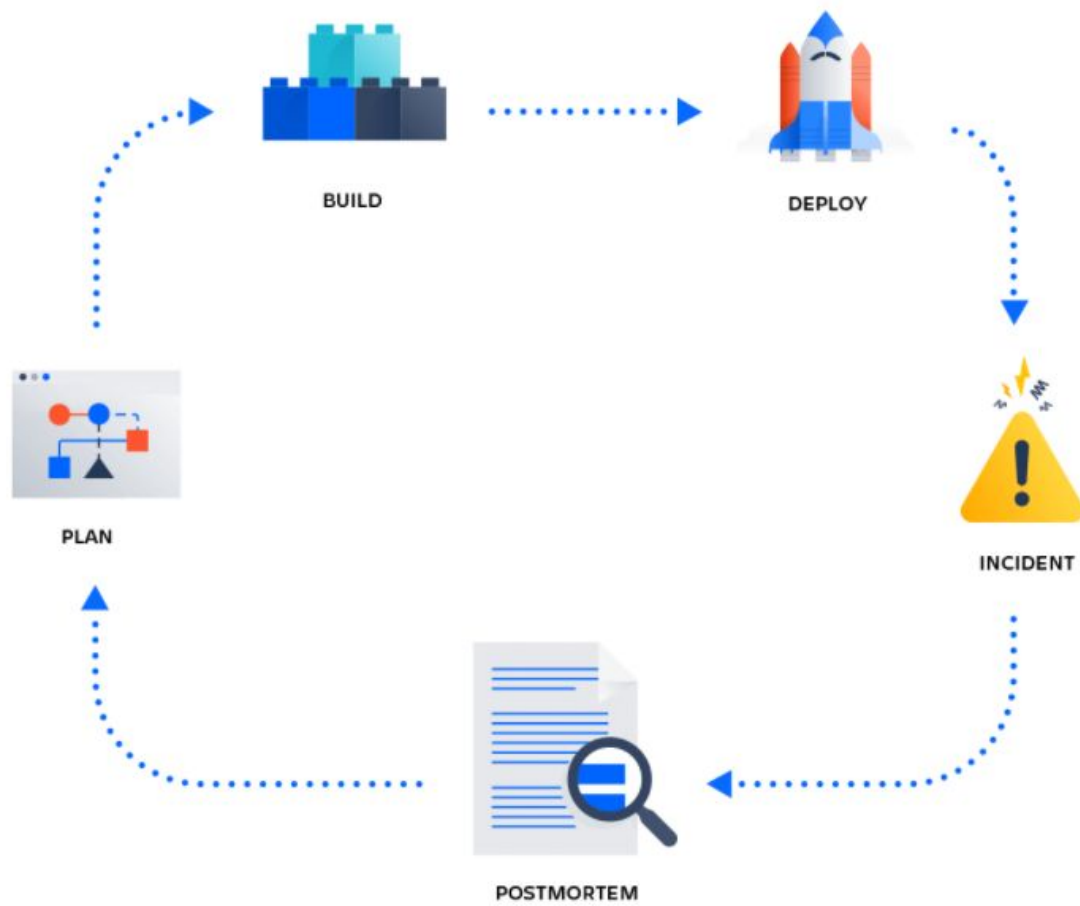
**Постмортем является “ревью” на произошедший инцидент информационной системы.**

Постмортем позволяет систематизировать знания о инциденте и содержит ответы на вопросы:

- почему инцидент произошел
- влияние инцидента
- какие действия были предприняты для смягчения инцидента и разрешения
- что нужно сделать, чтобы предотвратить повторение инцидента

Постмортем может быть интегрирован как опциональный шаг в ваш CI/CD пайплайн и быть связан с определенной сборкой или коммитом.

# Постмортем



---

# Постмортем

**Постмортем содержит следующие разделы:**

- Краткое описание инцидента (краткая выжимка о инциденте)
- Предшествующие события (что произошло перед инцидентом)
- Причина инцидента (из-за чего возник инцидент)
- Воздействие (на что повлиял инцидент)
- Обнаружение (когда и как инцидент был обнаружен)
- Реакция (кто ответил на инцидент, кто был привлечен, какие каналы коммуникации были задействованы)
- Восстановление (описание действий по устранению инцидента и поведение системы)
- Таймлайн (последовательное описание ключевых событий инцидента с указанием времени)
- Последующие действия (что нужно предпринять, чтобы инцидент не повторялся)

# Постмортем

Краткое описание инцидента	Около 2х часов ночи увеличилась утилизация ЦПУ на сервере. Вследствие этого часть сервисов стала недоступна.
Предшествующие события	Была установка патча 1.1 на сервис pretty
Причина инцидента	В патче 1.1 был допущен баг, в котором происходила многократная повторяющаяся запись на диск файлов.
Воздействие	Также на сервере с сервисом pretty был сервис корзины интернет магазина. Заказ товаров был недоступен для 100% пользователей в течении 15 минут.
Обнаружение	Инцидент был замечен дежурным инженером. Затем были привлечены ответственные разработчики.
Реакция	Ответственные разработчики устранили инцидент за 15 минут.
Восстановление	Был установлен минорный патч 1.1.1, устраняющий данную проблему и система перешла к штатной работе. Нагрузка сервера упала сразу после установки патча, корзина стала доступна для заказов
Таймлайн	01:55 прилетел алёрт о утилизации ЦПУ 01:57 дежурный инженер проинформировал о сбое ответственные лица 02:00 была проведена инспекция кода последнего патча и найдена проблема 02:05 был сделан коммит, устраняющий проблему 02:09 прошла автоматическая сборка и выкатка сервиса pretty 02:10 система заработала в штатном режиме
Последующие действия	Была заведена задача в issue tracker (ZADCH-123) по добавлению дополнительных модульных тестов на узкое место кода.



# Итоги



---

# Итоги

В данной лекции мы узнали:

- Что такое инцидент-менеджмент
- Как должна происходить коммуникация с пользователями при инциденте
- Как происходит реакция на инцидент со стороны команды
- Что такое дежурство
- Что такое постмортем и каким образом он составляется

---

# Домашнее задание

Давайте посмотрим ваше [домашнее задание](#).

- Вопросы по домашней работе задавайте **в чате** мессенджера .
- Задачи можно сдавать **по частям**.
- Зачёт по домашней работе проставляется после того, как **приняты все задачи**.

**Задавайте вопросы и  
пишите отзыв о лекции!**

**Сергей Андрюнин**