

Don't be afraid, it's just an incident!

Softwerkskammer Chemnitz - January 2022 Axel Köhler - @axdotl

Who am I

- ✓ Tech Lead @Staffbase
- ✓ In software engineering since 2006
- Java background (up to JDK 11)
- ✓ Fall in love with DevOps and SRE in 2016
- ✓ Love and hate processes
- Enjoy firefighting





What is an incident?

An event that disrupts or reduces the quality of a service that requires an emergency response.

Source: <u>Atlassian Incident Management Handbook</u>



When did we recover from an incident?

We recovered from an incident when there is no more disruption or reduction in the quality of the service.



What is incident management?

The process to respond to an unplanned event or service interruption and restore the service to its operational state.

Source: <u>Atlassian Incident Management</u> Handbook



Should I be afraid of incidents?

Incidents are not a bad thing, they should be treated as unscheduled investments in the quality of our product and happen when you embrace risk to move your product forward.

How do we want to respond





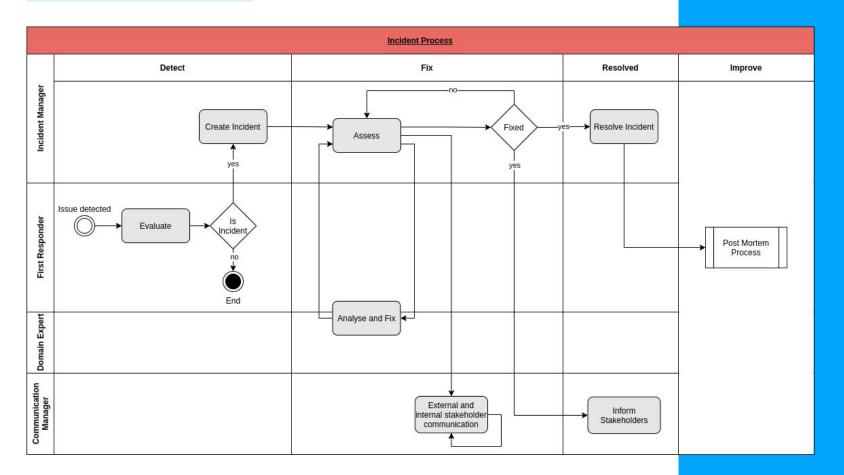
Incident Management @Staffbase

Where do we come from

- Responding reactive
- ✓ Wild communication in Slack
- Uncoordinated incident response
- Haphazard postmortems



Where are we now



Roles



- First Responder
- Incident Manager
- Domain Expert
- CommunicationsManager

First Responder

- First person who reacts to an incoming event (e.g. an alert)
- Can be everyone in the company
- Often the person on on-call duty

Domain Expert

- Analysis and develops strategies to mitigate and solve the issue
- Could be more than one
- Appointed by the Incident Manager

Incident Manager

- Driver of the incident and a leadership role
- Evaluates and coordinate
- Empowered to take all measures necessary to resolve the incident

Communications Manager

- Should be a person familiar with customer-facing communications
- Responsible for writing and sending internal and external communications
- Appointed by the Incident Manager

Staffbase

- ✓ Detect
- ✓ Fix
- Resolved
- ✓ Improve

Detect

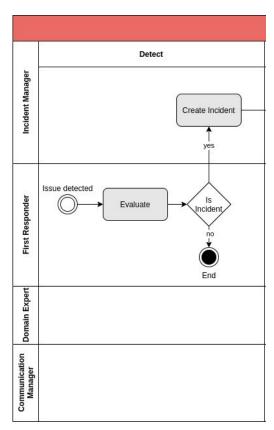
- Evaluate input event
- Result: incident is created or the process ends



- ✓ Detect
- ✓ Fix
- Resolved
- ✓ Improve

Detect

- Evaluate input event
- Result: incident is created or the process ends







- ✓ Detect
- ✓ Fix
- Resolved
- ✓ Improve

Detect

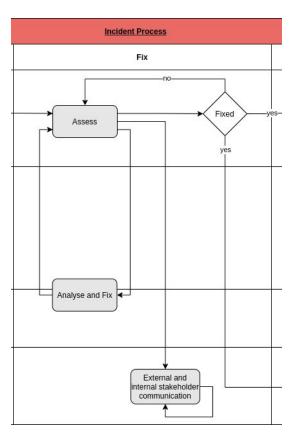
- Evaluate input event
- Result: incident is created or the process ends

Fix

- Covers all the necessary actions to solve the incident
- Iteratively until the incident is ready to resolve



- ✓ Detect
- ✓ Fix
- ✓ Resolved
- ✓ Improve



Fix

- Covers all the necessary actions to solve the incident
- Iteratively until the incident is ready to resolve



- ✓ Detect
- ✓ Fix
- Resolved
- ✓ Improve

Detect

- Evaluate input event
- Result: incident is created or the process ends

Fix

- Covers all the necessary actions to solve the incident
- Iteratively until the incident is ready to resolve

Resolved

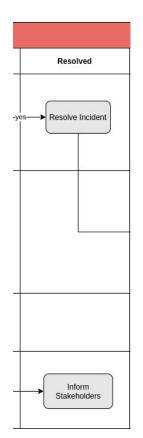
- Takes care of resolving the incident itself and informing all stakeholders
- At the end of this stage, we are back to regular work



- ✓ Detect
- ✓ Fix
- ✓ Resolved
- ✓ Improve

Resolved

- Takes care of resolving the incident itself and informing all stakeholders
- At the end of this stage, we are back to regular work





- ✓ Detect
- ✓ Fix
- Resolved
- ✓ Improve

Detect

- Evaluate input event
- Result: incident is created or the process ends

Fix

- Covers all the necessary actions to solve the incident
- Iteratively until the incident is ready to resolve

Resolved

- Takes care of resolving the incident itself and informing all stakeholders
- At the end of this stage, we are back to regular work

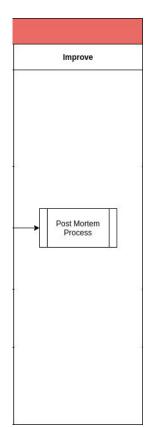
Improve

- Ensures, that we're learning from our failures
- Triggers the subsequent Postmortem
 Process





- ✓ Detect
- ✓ Fix
- Resolved
- ✓ Improve



Improve

- Ensures, that we're learning from our failures
- Triggers the subsequent Postmortem
 Process

Take a break





The technical side

Incident Management Stack

- ✓ Prometheus
- Opsgenie
- Statuspage
- Slack
- ✓ Google Meet
- Alarm tool (internal)



Time for a demo





Postmortem @Staffbase





- Preparation
- Meeting
- ✓ Review & Publish
- ✓ Implement

Preparation

- Incident Manager schedules a meeting with everyone involved during the incident
- Send preparation form (5-Why)
- Incident Manager prepares meeting based on gathered information

Review & Publish

- Use template so that all postmortem documents look roughly the same
- Git based review process
- Once PR is merged, the postmortem is published and available for everyone at Staffbase
- After the postmortem is published, the incident can be closed.

Meeting

- Happens asap (max 2 weeks)
- ✓ Timeboxed max 1.5h
- Collect and prioritize action items
- Jira items (labels pm-action-item and inc-xxx)

Implement

- Incident Manager set together with
 Product Owners due dates
- Important and urgent items has to be implemented within 6 weeks



- Information via Statuspage
 https://status.staffbase.com/incidents/
 7y84xnhjrv86
- See more complex <u>example</u>



Analytics is not available on the US infrastructure

Incident Report for Staffbase

Resolved

This incident has been resolved. We are sorry for the trouble

this has caused.

Please contact us at support@staffbase.com should you

experience any further issues. We're happy to help!

Posted 3 months ago. Oct 04, 2021 - 09:52 CEST

Identified

Customers hosted on the US infrastructure can't open the

Analytics Interface and see errors when loading the Reach value or open the Post statistics on the Experience studio.

We already identified the problem and are working on a fix.

We apologize for the interruption. In case you have any questions about this topic, please let us know at

support@staffbase.com.

Posted 3 months ago. Sep 30, 2021 - 18:27 CEST

This incident affected: Common Services (Plugins, Push Notifications...) (Common Features).



Example - Internal

- All postmortems are available for all employees on a dedicated webpage
- Chronological and searchable
- ✓ See <u>example</u>

INC-169: Analytics data missing on prod-us1

Basic Information

Data	2024 00 20
Date:	2021-09-30
Incident Manager:	Theobald Gerard
Involved People:	Theobald Gerard, Cedar Geena, Diane Rudi, Rex Briscoe, Mervin Rhetta, Stephan Ada
Slack Channel:	#inc-169
Opsgenie Incident:	#169 - Analytics Data is missing in prod/main-us1
Blast Radius:	12 Zendesk Tickets



We're done... well, almost



Takeaways for successful incident management

- Establish a lightweight process
- Automate and standardize critical paths in your process
- ✓ Define clear roles and responsibilities
- ✓ Setup clear paths of communication
- Ensure a blameless culture
- Learn and improve from your incidents



Q&A