Incident Management

The Agile team's guide to harnessing change and responding to disruption
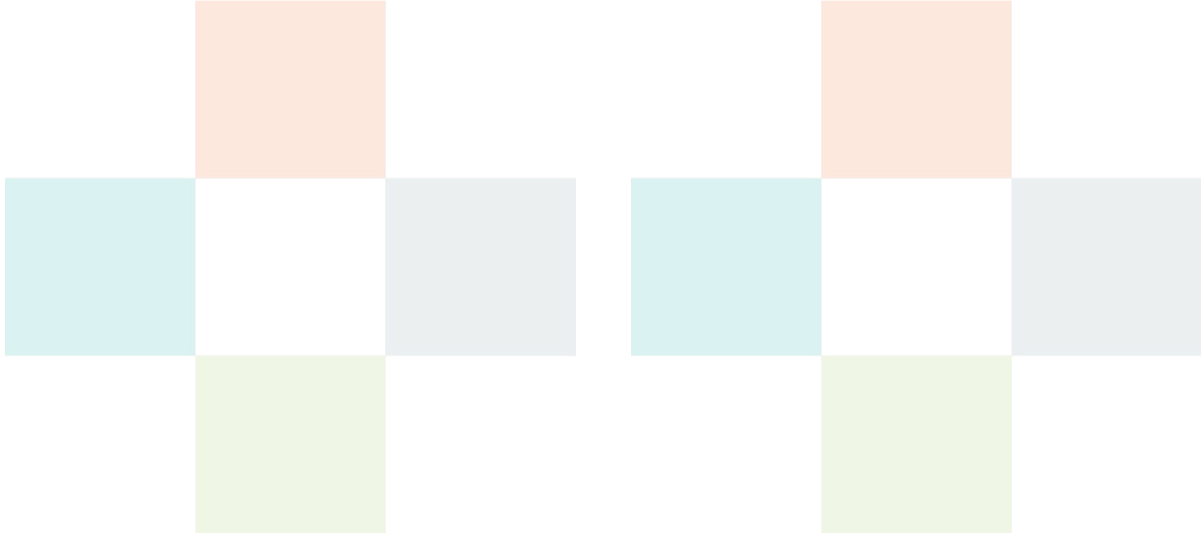
palantir.net

Hello.

Meghan Davis
Senior Project Manager
Team Lead

Overview

What we hope to cover today

Foundational concepts

+ Who is this session for?

+ What is incident management?

+ Why do we talk about incident management?

In practice

+ How do high-performing teams manage incidents?

+ Tools and best practices you can implement today

+ Continuing the learning after the incident

Foundational concepts

Who is this session for?

+ A member of a team who provides client or customer support or services.

+ Manage a team who provides client or customer support or services.

+ Agile practitioners

+ Project or Program Managers

+ Development or operations teams

What is an incident?

When we talk about an incident, as opposed to a risk, we're referring to an **event that causes disruption to or reduction in the quality of a service we provide and requires a timely response**.

Site crashes

Functionality stops working

Requested iterations

Why do we talk about incident management?

+ Incidents are costly

+ Potential to jeopardize reputation and customer attrition

+ Incidents are disruptive + confusing to the project team

Benefits of an incident management play

+ Collaboration

+ Direct communication

+ Builds trust

+ Encourages honesty and openness

In Practice

Discover → Act → Resolve → Reflect

Discover → **Act** → **Resolve** → **Reflect**

You may detect an incident in several ways:

+ Customer or client alarm

+ Customer or client filing of a Jira ticket or casually mentioning on a call

+ Monitoring of systems

+ Teammate's making an observation and flagging in Slack/email

+ Teammate's mentioning their experiencing some challenges in their work

Discover → Act → Resolve → Reflect

Tools you'll need for open, honest, frequent communication

+ Messaging tool ➜ Slack

+ Client communication tool ➜ Email / Phone

+ Some sort of shared collaboration space that provides live updates ➜ Google Doc + Atlassian's Confluence

+ Video chat as you'll need to talk to your teammates ➜ Zoom or Google Meet

The first step you take when you hear about an incident

A question, disruption, or request comes in. The person receiving this disruption should notify the full project team.

You'll want to start a video call with your team.
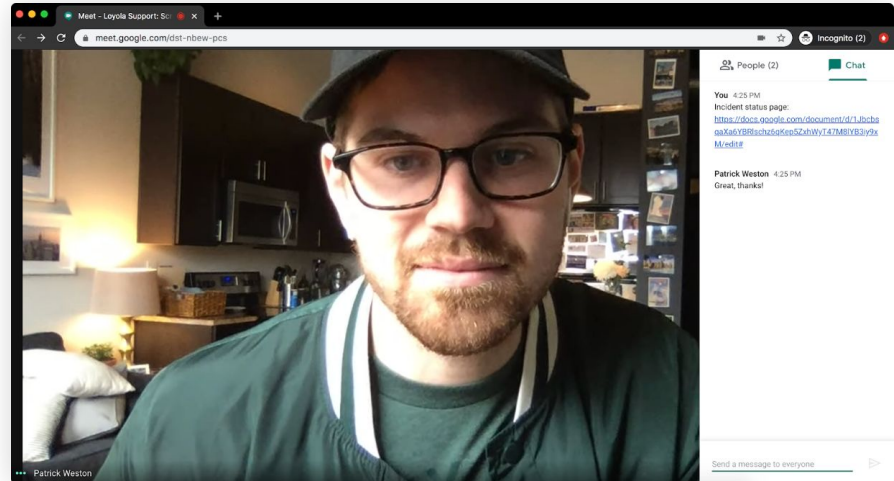
Determine the incident team

Determine the incident team, based on what you are trying to solve for.

1. Consider which area(s) of the project is impacted.

2. Who is the decision maker? Who is the Incident Manager?

3. Who will be consulted and informed?

4. Who will manage client communications and keep the client informed?

5. Communicate these roles to the full team.

Those not part of the incident team are not brought into decision making or discussion meetings. Team members progress with meaningful work that will advance the project.

Status page

The Incident Manager will next set up what is called an incident status page, even if the disruption is not fully understood at the time.

The incident status page should be captured in a shared space, such as Google Drive or Confluence.

Palantir.net | Team Healthcare
Incident Status Page
March 1, 2020

Project team: Team Healthcare

Core deciders: Rebecca, Stephen, Elvio

Notes + client communications: Meghan

Date: March 1, 2020

A brief summary of the incident: On March 2, 2020, Rebecca attempted to do a deployment to production for the find a doctor project. When she proceeded with the deployment, it appeared the database server went down. The client's site was down.

Resolution: The Team Healthcare site is back up and running successfully at 9:52 am EST (8:52 am CT) on March 1, 2020. Team Healthcare rolled back the deployment initiated this morning and are in communication with the client about the next steps.

Resolved: The Team Healthcare site is successfully back up and running. We rolled back the deployment from this morning.

March 1, 2020 @ 9:52 am EST (8:52 am CT)

Good morning again,

Following up on our email from a few minutes ago: the site is now back up and running on the code and database version prior to today's deploy.

The site is back up and running.

We will continue to investigate what went wrong with this deploy and follow up when we have more information.

Status page

Goal of the doc:

+ Get the message/updates out there, while still primarily focusing on solving the issue

+ Share first, improve later

+ Clear, plain language & screenshots

+ Updated with brief, honest, frequent updates

Palantir.net | Team Healthcare
Incident Status Page
March 1, 2020

Project team: Team Healthcare

Core deciders: Rebecca, Stephen, Elvio

Notes + client communications: Meghan

Date: March 1, 2020

A brief summary of the incident: On March 2, 2020, Rebecca attempted to do a deployment to production for the find a doctor project. When she proceeded with the deployment, it appeared the database server went down. The client's site was down.

Resolution: The Team Healthcare site is back up and running successfully at 9:52 am EST (8:52 am CT) on March 1, 2020. Team Healthcare rolled back the deployment initiated this morning and are in communication with the client about the next steps.

Resolved: The Team Healthcare site is successfully back up and running. We rolled back the deployment from this morning.

March 1, 2020 @ 9:52 am EST (8:52 am CT)

Good morning again,

Following up on our email from a few minutes ago: the site is now back up and running on the code and database version prior to today's deploy.

The site is back up and running.

We will continue to investigate what went wrong with this deploy and follow up when we have more information.

Act: Status page

March 1, 2020 @ 9:25 am EST (8:25 am CT)

Update - Restarting the service did not work. The site remains down. While attempting to restart the service we noticed XYZ which prompted us to ask Jeremy if it was indeed safe to restart the mysql service on the database server (ie. that it would not affect other servers or services).

Jeremy replied: *Yes, you may proceed.*

March 1, 2020 @ 9:21 am EST (8:21 am CT)

Investigating - Palantir believes the issue here is that the database service has stopped and needs to be restarted (and not a lack of access to database backups). Palantir is proceeding with restarting the service.

March 1, 2020 @ 9:20 am EST (8:20 am CT)

Troubleshooting - Rebecca, Stephen, and Elvio begin troubleshooting together.

March 1, 2020 @ 9:00 am EST (8:00 am CT)

Investigating - Rebecca attempted to do a deployment to production for the find a doctor project. When she proceeded with the deployment, it appeared the DB server went down. The Team Healthcare site is down.

March 2, 2020 @ 8:24 am EST (7:24 cm CT)

Engage in the advice process

+ Decision maker should seek out various perspectives

+ Gather advice from everyone who will be meaningfully impacted or those who hold subject matter expertise

+ *Decision maker forms a written proposal.*

+ Make a call on how to move forward

Act: Status page

Palantir.net | Team Healthcare

Incident Status Page
March 1, 2020

Project team: Team Healthcare

Core deciders: Rebecca, Stephen, Elvio

Notes + client communications: Meghan

Date: March 1, 2020

A brief summary of the incident: On March 2, 2020, Rebecca attempted to do a deployment to production for the find a doctor project. When she proceeded with the deployment, it appeared the database server went down. The client's site was down.

Resolution: The Team Healthcare site is back up and running successfully at 9:52 am EST (8:52 am CT) on March 1, 2020. Team Healthcare rolled back the deployment initiated this morning and are in communication with the client about the next steps.

Resolved: The Team Healthcare site is successfully back up and running. We rolled back the deployment from this morning.

March 1, 2020 @ 9:52 am EST (8:52 am CT)

Good morning again,

Following up on our email from a few minutes ago: the site is now back up and running on the code and database version prior to today's deploy.

The site is back up and running.

We will continue to investigate what went wrong with this deploy and follow up when we have more information.

The Practice

Discover → Act → Resolve → Reflect

Review the status page

Palantir.net | Team Healthcare

Incident Status Page

March 1, 2020

Project team: Team Healthcare

Core deciders: Rebecca, Stephen, Elvio

Notes + client communications: Meghan

Date: March 1, 2020

A brief summary of the incident: On March 2, 2020, Rebecca attempted to do a deployment to production for the find a doctor project. When she proceeded with the deployment, it appeared the database server went down. The client's site was down.

Resolution: The Team Healthcare site is back up and running successfully at 9:52 am EST (8:52 am CT) on March 1, 2020. Team Healthcare rolled back the deployment initiated this morning and are in communication with the client about the next steps.

Resolved: The Team Healthcare site is successfully back up and running. We rolled back the deployment from this morning.

March 1, 2020 @ 9:52 am EST (8:52 am CT)

Good morning again,

Following up on our email from a few minutes ago: the site is now back up and running on the code and database version prior to today's deploy.

The site is back up and running.

We will continue to investigate what went wrong with this deploy and follow up when we have more information.

Closing communications, wrap up

+ Incident report (written in plain language)

  ○ Issue summary

  ○ Timeline

  ○ Root cause analysis

  ○ Recommendations

+ Share with clients, your project team, key stakeholders, whole company

+ Hop on a brief call, if the client would like to discuss or needs clarification

Resolve: Closing communications, wrap up

Palantir.net | Team Healthcare

Internal - 3/1/2020 Incident Report

Issue summary

Timeline

Root cause

Palantir's recommendation
    Next steps
    Until the transition is complete
    Appendix: Resources for reference

Issue summary

On March 1, 2020, Rebecca attempted to do a deployment to production that contained the work for the Find a Doctor project. Palantir was not able to complete the necessary database operations during the deployment and further, the database service stopped during the deployment, causing the site(s) to go down.

The incident lasted from approximately 7:24 am CT to 8:52 am CT, when the incident was resolved and the site(s) was restored. The impact of the site(s) being down, as we understand it, was minimal as it was outside of Team Healthcare's business hours. There were no reports from key stakeholders at Team Healthcare of the site being down.

The root causes of the incident were XYZ and, upon further research and investigation, XYZ.

Timeline

7:24 am CT - Rebecca reaches the "database operations" portion of the deployment during which she receives an error message.

Resolve: Closing communications, wrap up

[DATE] - [INCIDENT NAME]

Created by Meghan Davis
Just a moment ago • 📊 Analytics

Here is the documentation that was generated during the [DATE] incident.

– Incident Status Page (The status page outlines the issue encountered in clear, plain language. It was updated by [TEAMMATE NAME] with frequent, honest, brief updates, and reviewed by the full incident team afterwards to add in detail. It served as the central place where we were communicating about the incident, avoiding having conversations in Slack or email).

– Incident post-mortem (Once the issue was resolved, the incident team took some time away from the computer. We regrouped in the afternoon of the same day to reflect on the incident, together. We identified things that worked well, things that did not go smoothly, and recommendations for the client).

– Incident de-brief with client [TEAMMATE NAME] + [TEAMMATE NAME] met with [CLIENT NAME + ROLE], on [DATE] to discuss the incident. This doc captures the notes and recommendations from that conversation.

– Internal incident report - Shared with [COMPANY NAME] stakeholders on [DATE] BY [TEAMMATE NAME].

– Client communication following incident - This [DATE]. email recaps the incident, communicates the decisions made in our client debrief back to [CLIENT NAME], outlines risk, and captures action items/next steps.

Share

THE PRACTICE

Discover → **Act** → **Resolve** → **Reflect**

Why reflect?

+ Encourages support and collaboration

+ Fosters team communication, growth, and openness

+ Increases the chance that a teammate will come forward if they recognize an incident in the future

The blameless post mortem

+ Begin with the assumption that all team members acted from a place of good intent
+ Keep communication fact-based, honest, objective
+ Opportunity to step back, breath, and reflect on the incident.
    ○ Reflect on what happened
    ○ Reflect on why it happened
    ○ Examine how the team responded
    ○ Identify what can be done to prevent repeat incidents
    ○ Identify ways to improve future responses
+ Document in a shared space (Confluence)
+ Share out with your project team, whole company

Thank you.

Meghan Davis
Senior Project Manager
Team Lead

Questions?

PLEASE PROVIDE YOUR FEEDBACK!

mid.camp/6331

The top rated sessions will be captioned, courtesy of Clarity Partners

CONTRIBUTION DAY
Saturday 10am to 4pm

You don't have to know code to give back!

New Contributor training 10am to Noon
with **AmyJune Hineline** of Kanopi Studios