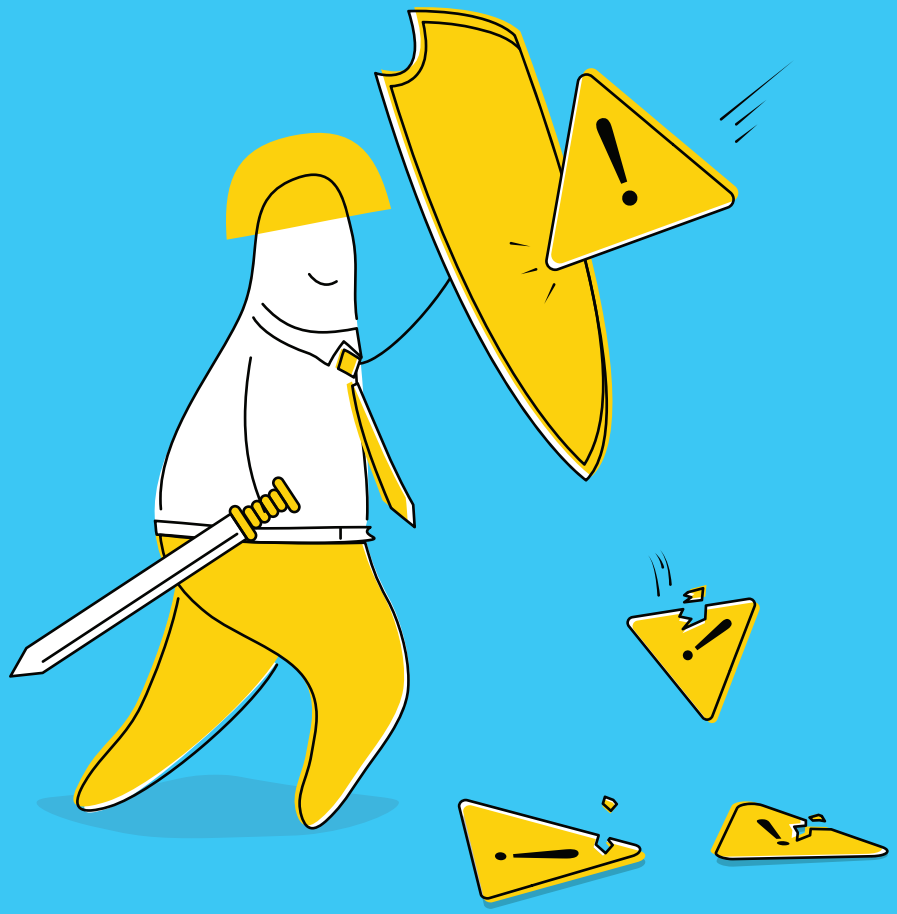
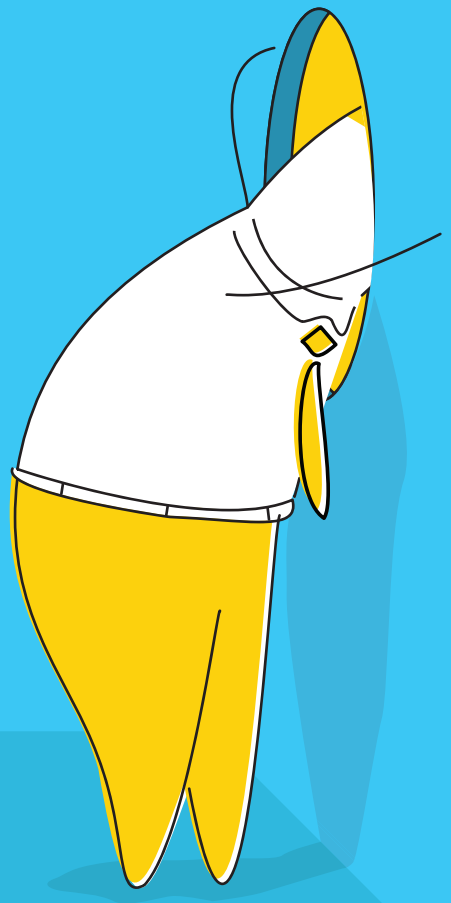
ManageEngine



# INCIDENT MANAGEMENT HANDBOOK

## HOW ZOHO HANDLES THE SPECTRUM OF IT INCIDENTS

# WHAT'S INSIDE?




## Introduction 01

Background ..... 01
Who is this guide for? ..... 03
What is an incident? ..... 03
What is incident management (IM)? ..... 03
Our incident values ..... 04
Our IM tools ..... 05

## Incident management processes 09

Desktop sprint ..... 10
Big bang ..... 21
CyberSec ..... 31

## Root cause analysis (RCA) 42

What is RCA ..... 42
Why perform RCA? ..... 42
RCA principles ..... 42
RCA process ..... 43
RCA meetings ..... 49

## Conclusion 49

“Whew! That was
a close call.
Let’s hope that never
happens again!”

# Introduction

Over the last decade, we’ve combated thousands of incidents.

As bootstrappers, we’ve experienced low-impact incidents that typically needed fewer technicians but still required a well-established incident management (IM) framework. For most incidents in the past, we relied on problem solving by individuals. However, as our IT infrastructure grew, we faced more complex and high-impact incidents, forcing us to up our IM game.

Soon, we realized there is no one-size-fits-all process to manage all the different types of incidents our organization faced. So, we took the frameworks that were most effective and added, combined, or omitted steps to handle every type of incident based on their impact and our business operations. This ensures each response is well-tailored to the challenges presented by each incident.

The result? Our incident process now extends beyond established industry frameworks. Our IM frameworks are classified based on the severity and impact the different types of incidents have on business operations.

| IM framework | | Impact | Scenarios |
|---|---|---|---|
| Desktop sprint | • Break/fix incidents that affect an individual user | • A single user is affected<br>• No critical services are involved | • Password resets<br>• Internet is slow |
| | • Low or medium-impact incidents that affect user groups or departments | • A single VIP user is affected<br>• A small group of end users are affected<br>• There is no potential for financial loss or loss of reputation | • The CEO's laptop is not working and is unable to send and receive communications<br>• A printer is not working on a particular floor |
| Big bang | • High urgency<br>• Affects service | • A certain business-critical service, application, or infrastructure component is unavailable, and the estimated time for recovery is unknown or exceedingly long<br>• Service is unavailable. Immediate restoration of service is expected | • Our high-speed network connection fails, and communication to and from outside our organization is cut off<br>• Core functionality of an application is down, affecting several customers<br>• One of our applications is down A distributed denial-of-service (DDoS) attack |
| Showstopper | • Immediate urgency<br>• Critical or red alert situations that affect business | • Affects our company's bottom line<br>• Major impact on revenue, reputation, and legal affairs | • Software bugs and vulnerabilities<br>• Malware<br>• Advanced persistent threat (APT)<br>• Ransomware<br>• Phishing and social engineering<br>• Insider threat |

When it comes to IM, there is no one-size-fits-all solution as every organization is different. What will work for your organization will depend on your business model, infrastructure, operations, the information you are protecting, your resources, and more. Recognize that some techniques only come with time and experience. This should not, however, discourage you from getting started!

## Who is this guide for?

This e-book is written for IT leaders, managers, and practitioners from a service management perspective. We will walk you through our IM processes with illustrated process flows, roles, and best practices. This guide is full of lessons we've learned through trial and error—so you don't have to.

Before we dive in, let's get the basics out of the way.

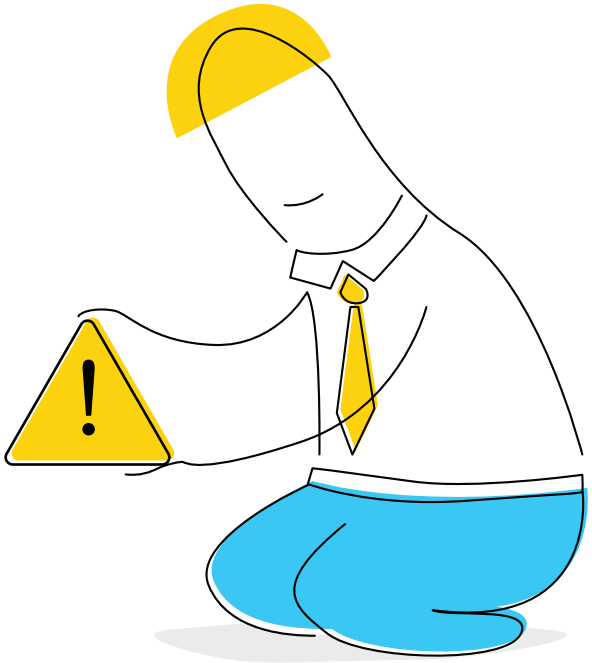
## What is an incident?

An incident is an unplanned interruption that causes, may cause, or reduces the quality of an IT service. Some classic examples are the internet running too slow, a business application going down, or a printer not working.

Truth is, we can define an incident in many ways. What matters most is that every incident should have a well-structured, timely response and resolution.

## What is incident management?

Incident management is a way to restore normal service operations as quickly as possible, minimizing any adverse impact on business operations or the user.

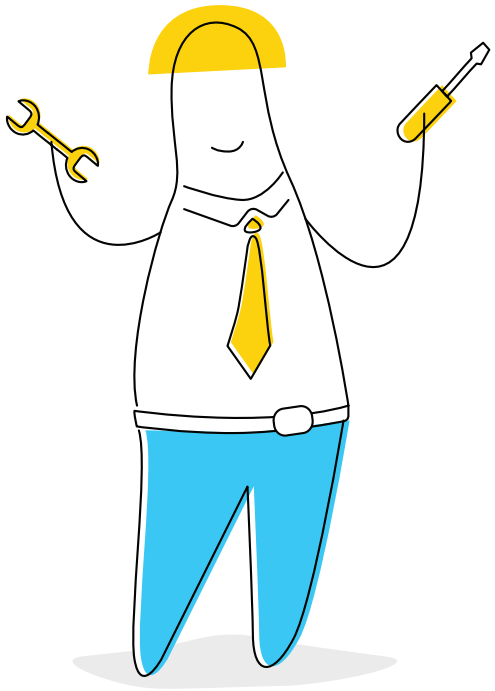


# Our incident values

| Incident Principles | Approach |
| --- | --- |
| Be proactive, not reactive | • A proactive approach: preventive maintenance is regularly performed to reduce the likelihood of failure—before users are even impacted.<br>• Monitoring tools provide visibility on the network's health and performance, and alert us about issues before they become incidents. |
| Be open, and communicate | • We communicate with our customers early on and often, to let them know that we are aware of and working on issues.<br>• A predefined group of stakeholders are automatically notified through their preferred contact methods when an incident strikes. |
| Align teams, collaborate effectively | • We have distributed teams from various time zones working together during high impact incidents, dialing in to a conference call bridge number, and utilizing communication or productivity applications to handle the incident. |
| Bounce back quickly | • Incident management could mean many things. However for us, it translates into time management.<br>• We utilize Site24x7, which lets us know as soon as something breaks. Getting ahead of issues is crucial to our IM. Sometimes, our employees turn into an alert system. They use our systems on a daily basis and will likely be the first people to notice when something does not feel right.<br>• We have an open IM system, follow protocols where needed, and work as a team to resolve the issue at the earliest. |
| Document the lessons | • We sometimes make mistakes. Who doesn't? However, we ensure that we learn from those mistakes by documenting the lessons learned. |
| Continually improve | • We deep dive into what went wrong to ensure we don’t make the same mistake twice.<br>• Sometimes, we run mock incidents to see how our IM strategy holds up, and continue to fine-tune it before the real deal. |

## Our IM tools

We utilize several tools to aid our IM processes.



## Desktop incidents

**Track & manage incidents:**
ServiceDesk Plus Cloud is customized to fit our incident management processes.

**Password management:**
Password Manager Pro is a secure vault for storing and managing shared sensitive information such as passwords, documents, and digital identities of enterprises.

## Major availability incidents

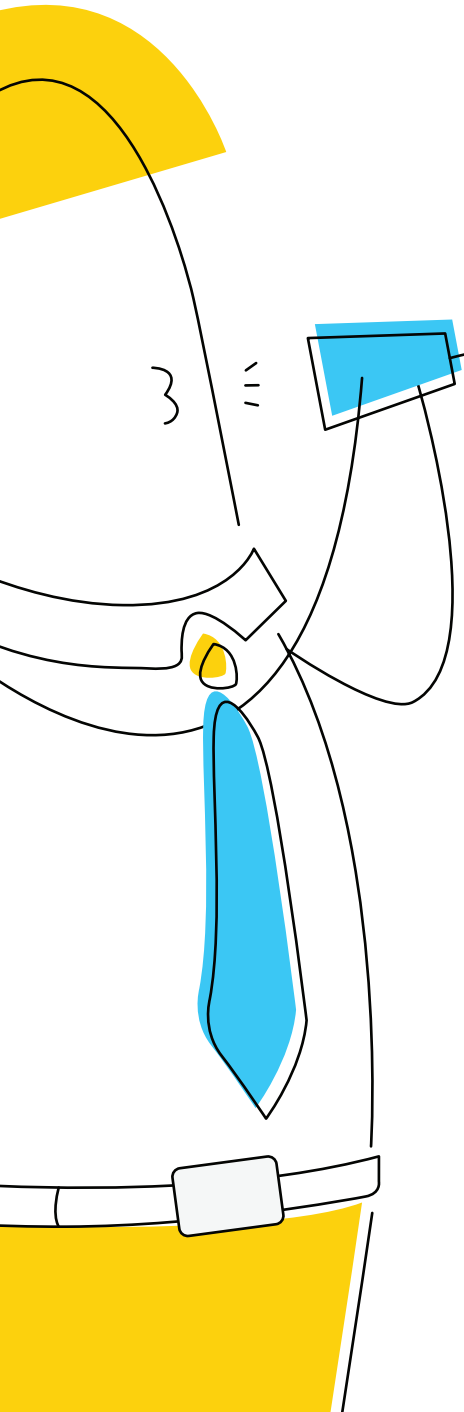
**Alerting tool:**
We use Site24x7 to monitor the availability of servers and applications.

**Password resets:**
ADSelfServicePlus is a self-service password reset tool.

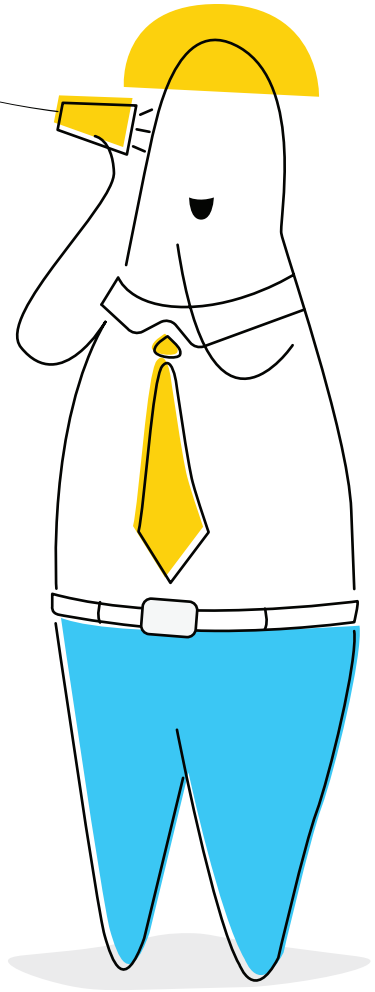
**Endpoint management:**
Desktop Central, a unified endpoint management solution, helps manage servers, laptops, desktops, smartphones, and tablets from a central location.

## Security incidents

**Bug Bounty program:**
Bug Bounty is a third-party tool for employees and individuals to report bugs, like exploits and vulnerabilities.



communication



## Note:

We also use social networking sites, messaging platforms like What's App, and phone calls as alternative ways to communicate should Cliq go down, as it's important to have alternative means of communication during a disaster.

## Documentation:

Zoho Docs is a central system for storing all incident and root cause analysis (RCA) documents.

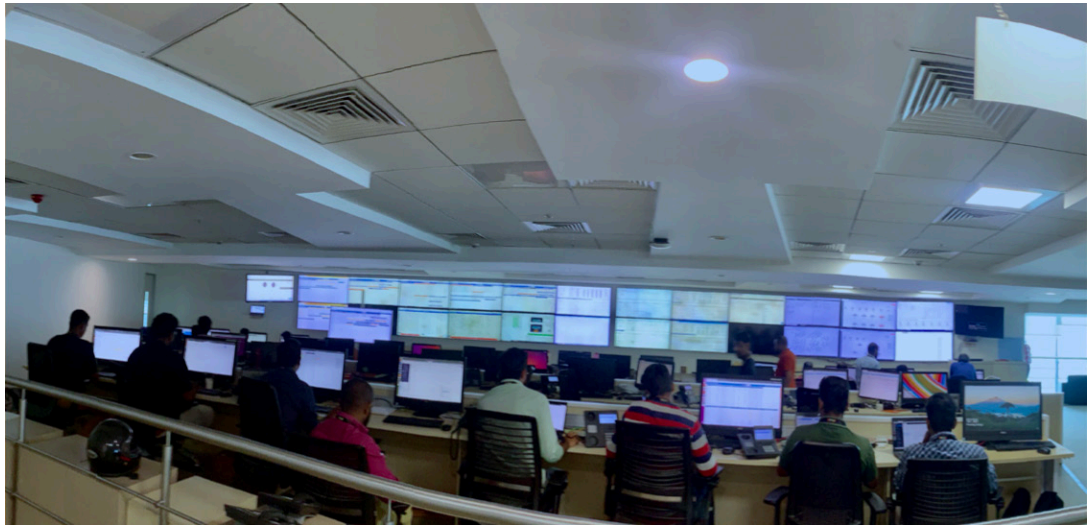
## Chat:

Zoho Cliq is a real time business messaging app that helps our employees communicate effectively with each other anytime, including during an incident.

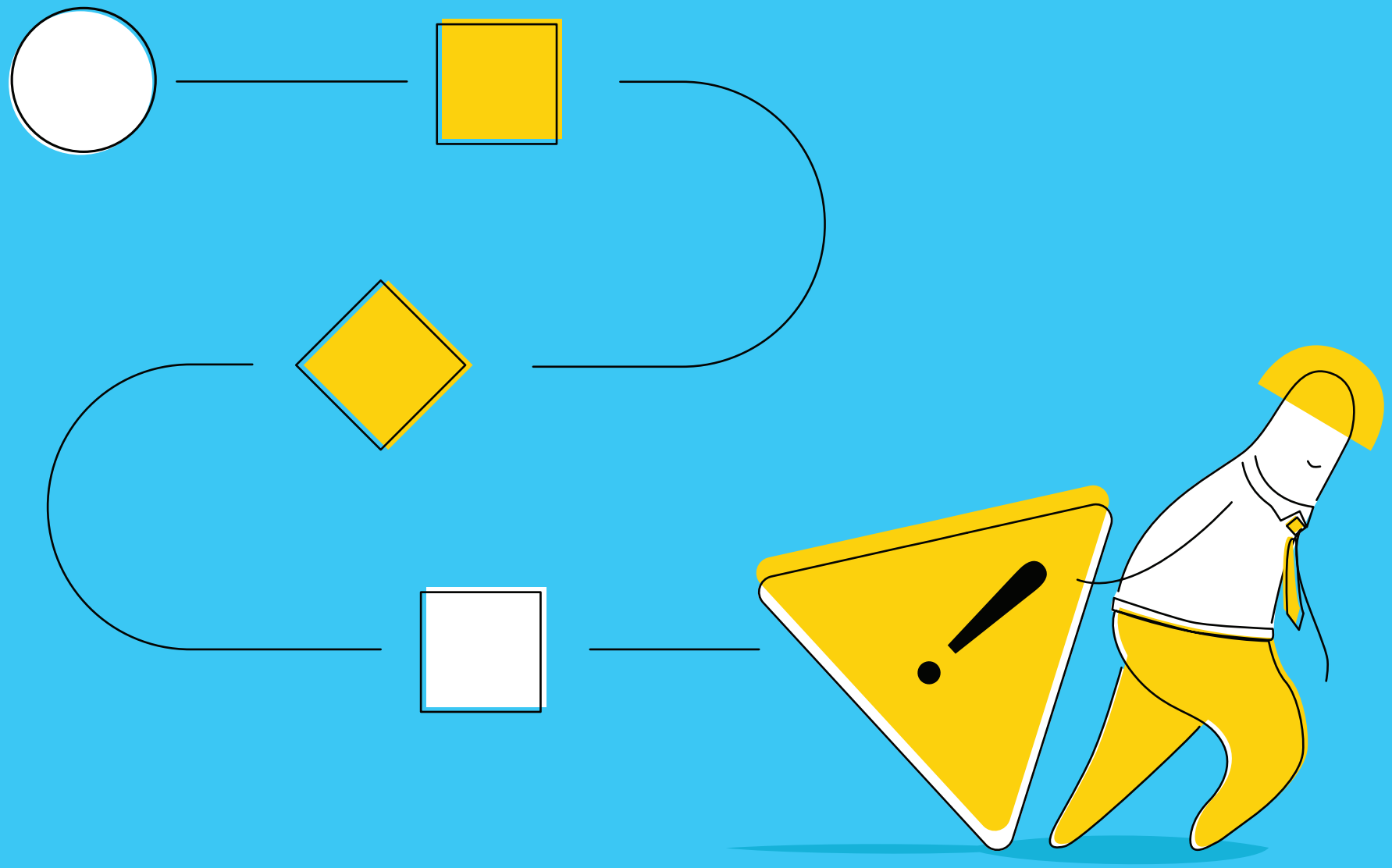
## Collaborate:

Zoho Connect is collaboration software that ensures that all teams can be on the same page when resolving incidents. Some call it the Facebook of our workplace.

# Our incident management command center (IMCC)



Our incident management command center (IMCC) is a large secure room with big, NASA-like screens of monitoring devices to provide detailed metrics and visibility, enabling our IM teams to react quickly and troubleshoot effectively during incidents. This room hosts three core teams: the network operations center (NOC) team, the Zorro team, and the central system admin team. We have dynamic access control in other work sites to perform monitoring activities.



# INCIDENT MANAGEMENT PROCESSES

**Desktop sprint**
(break/fix &
low key incidents)

**Big bang**
(major availability
incidents)

**CyberSec**
(showstoppers or
critical incidents)

# Desktop sprint

(desktop incidents)

## Teams, roles, & responsibilities

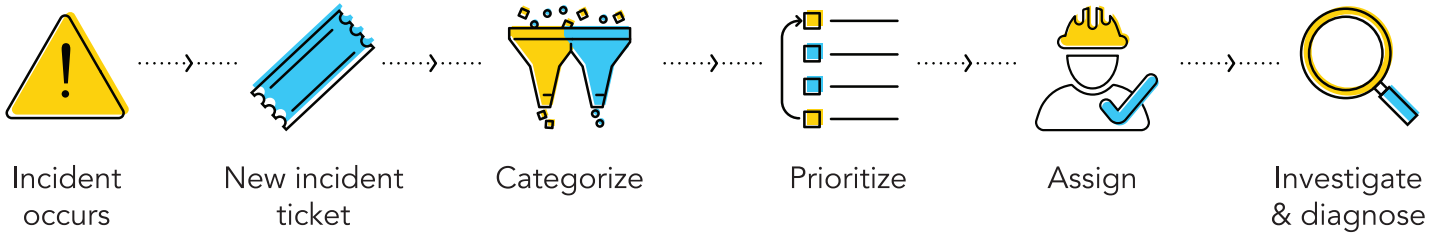
**PitStop technicians:**

Just like any IT organization, our front-line IT support team handles desktop incidents. We call our IT support center PitStop.

**Central sysadmin team:**

We have a central system administration team as part of our incident management command center overseeing all incoming incidents in our 12-story building. We place a PitStop with a technician on every floor; in the absence of a PitStop technician on a floor, the central sysadmin team handles the desktop incidents on that particular day.

Most often, incidents are routed to technicians by the incident coordinator who oversees all incoming desktop incidents using business rules in our IT service management (ITSM) tool. PitStop technicians can also self-assign tickets in the absence of the incident coordinator.

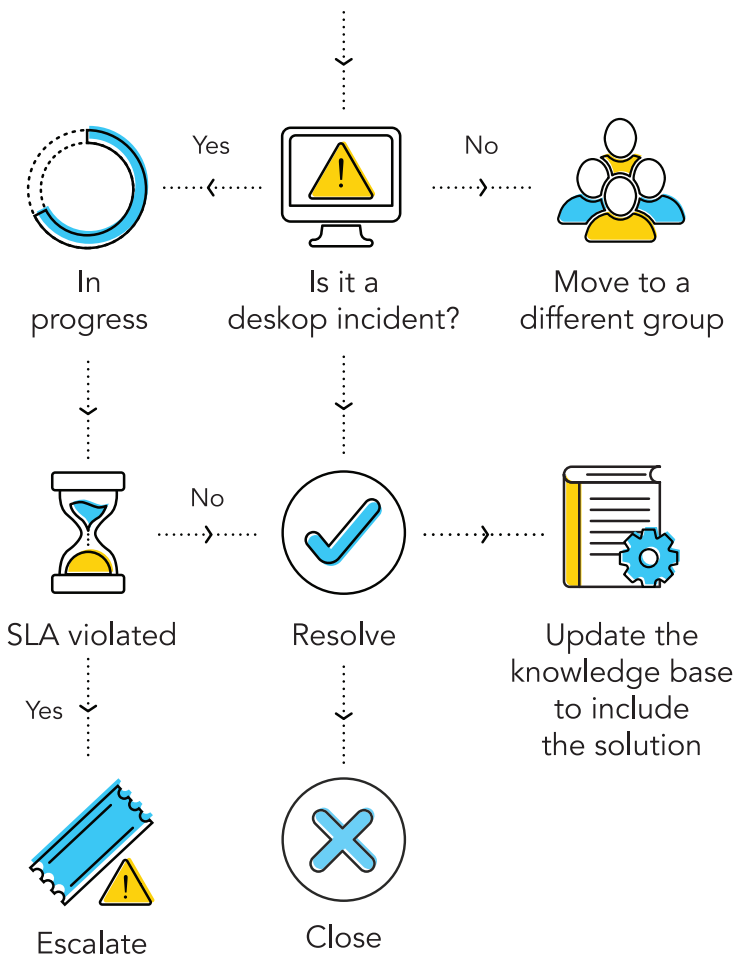



## The process

On a typical day, our PitStop technicians troubleshoot low to medium impact incidents such as password resets, printer issues, and network issues, and perform a variety of tasks including:

- Communicating service outages to all end users.
- Opening communication with end users to investigate and gather as much information on incidents as possible for quick resolution.
- Creating requests for changes or problem records.
- Adhering to the Service Level Agreements (SLAs) of incidents and escalating them as needed.
- Resolving and closing incidents.
- Providing status updates to end users throughout the incident life cycle.

To handle day-to-day incidents, we use a high-speed resolution model that you're likely familiar with. It's a simple, straightforward process that addresses hurdles and ensures seamless flow.



## New incident

An incident typically starts with our employees reporting an issue through an email, phone call, live chat, or the self-service portal in our ITSM tool. The incident is logged as an incident ticket and we fill in the following default details.

| Title | Summary of the incident |
|---|---|
| Description | Provide as much detail as possible to help the technicians diagnose the incident and provide quicker resolution. |
| Impact | Who's affected—a single user or entire business operations? |
| Urgency | How quickly must the incident be resolved? |
| Priority | What's the importance of the incident based on the impact and urgency? |
| Groups | Which resolver group handles the incident? For example, we create groups for specific issues such as hardware, software, printers, and so on. |
| Assets | What are the assets and services that are affected due to the reported incident? Is it a single asset or multiple assets? |

After logging, the incident moves to the open state, which is the first state in our incident workflow.

## Categorization

Our incident coordinator starts with assigning incidents to the right categories and subcategories for easy classification. Without categorization, the incident manager won't know how many operating system and application issues we experienced, or what actions need to be taken to reduce those incidents.

We categorize incidents for the following reasons.

- For grouping similar incidents into a common bucket to speed up the incident life cycle.
- To automatically route and assign incidents to the right teams for quick resolution
- For example, auto assign Linux-related issues to the right team.
- For problem analysis.
- To generate a well-structured report.

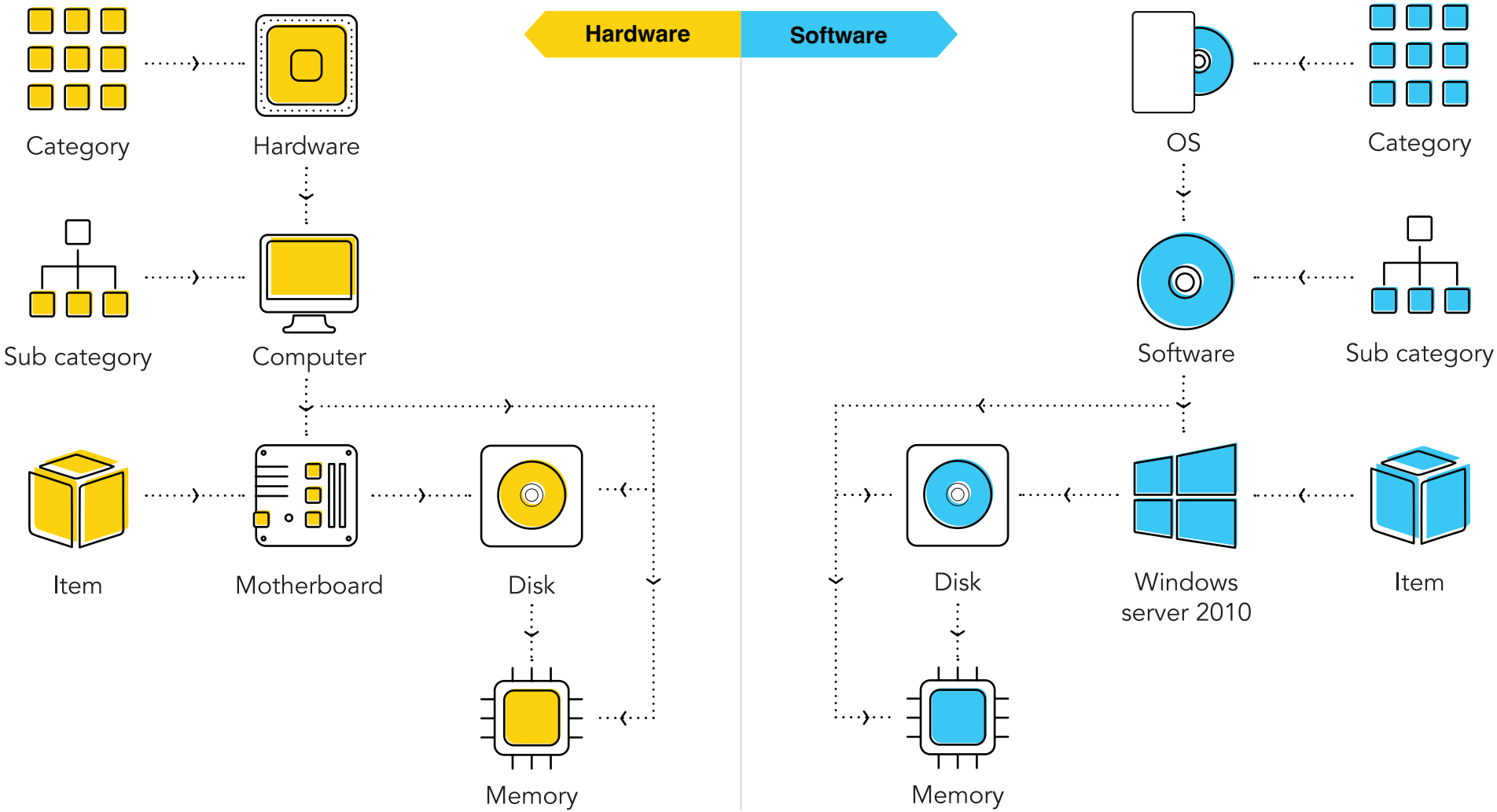
As a best practice for effective categorization, we stick to three levels of categorization. Too many levels can complicate the process, and too few could defeat the purpose. The categorization usually starts with the major category, then a sub-category, and finally the affected configuration item.

We limit the major categories to around 10-15 to keep the categories broad yet manageable. Every three to six months, our incident coordinator checks the historical records, and sorts the incidents according to the major categories to check if the incidents fall within those categories. The incident log is analyzed, and the category is determined by asking:

- How are the incidents distributed across the category tree?
- Are the major categories and sub categories well-defined?
- Are the categorization levels speeding up incident resolution?
- How many incidents are falling into the "Other" category?
- Is reporting compromised due to inefficient categorization?

Based on the answers and our business needs, the incident coordinator fine-tunes the depth of the category tree.

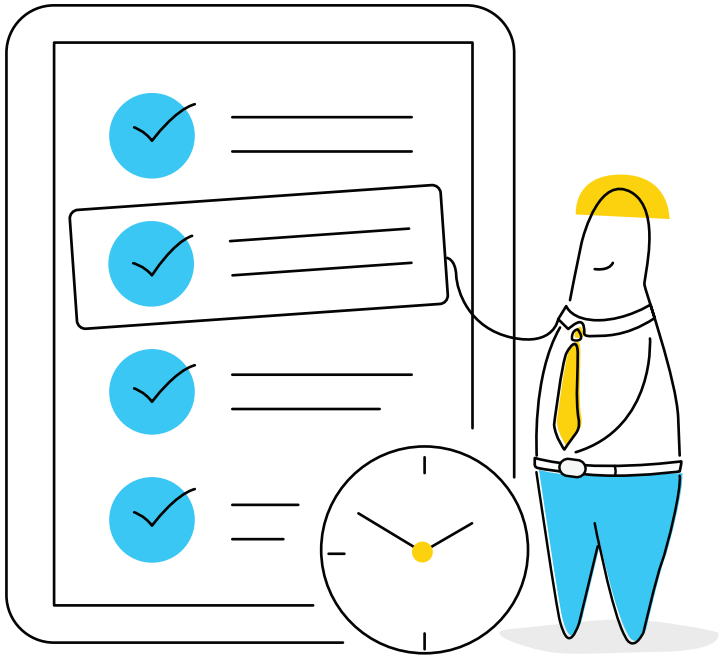
Here’s a typical example of a category tree that we use to handle hardware and software issues.



## Prioritization

Our incident coordinator starts with assigning incidents to the right categories and subcategories for easy classification. Without categorization, the incident manager won't know how many operating system and application issues we experienced, or what actions need to be taken to reduce those incidents.

While all incidents need to be resolved, some incidents have greater impact on our business and require a greater sense of urgency to resolve. We determine the priority of an incident by an incident prioritization matrix (impact x urgency) for ensuring end-user satisfaction, optimal use of resources, and minimal affect to our business operations.



To map out our priority matrix, we ask ourselves:

- How is productivity affected?
- How many users are affected—Is it a single user or a group? Are the VIP users affected?
- How many systems or services are affected?
- How critical are these systems/services to the organization?
- Are the customers affected? Is there a significant impact on revenue?
- Is there a major impact on revenue/business reputation?

The priority matrix automatically defines the priority of a particular incident based on the inputs provided (impact and urgency) by the end users when logging a ticket in our ITSM tool. In our priority matrix, the impact is listed in the y-axis, and the urgency is listed in the x-axis. We group impacts by: user, group, department, and business. For urgency, the four levels are low, medium, high, and critical.

The priority matrix provides an overview of every incident and ensures that major incidents are prioritized and addressed quickly; it also ensures low-priority incidents, like desktop incidents, are handled within an acceptable time frame.

Here are some use cases showing how we utilize our priority matrix:

| Urgency | Impact | Scenarios |
| --- | --- | --- |
| Break/fix<br>(affects individuals and small groups) | • A single user is affected<br>• No critical services are involved | • Password resets<br>• Internet is slow |
| Low key<br>(affects group/medium impact incidents) | • A single VIP user is affected<br>• A small group of end users are affected<br>• There is no potential for financial loss or loss of reputation | • The CEO's laptop is not working and is unable to send and receive communications<br>• A printer is not working on a particular floor |
| Big bang<br>(affects service) | • A certain business-critical service, application, or infrastructure component is unavailable, and the estimated time for recovery is unknown or exceedingly long<br>• Service is unavailable. Immediate restoration of service is expected | • Our high speed network connection fails and communication to and from outside our organization is cut off<br>• One of our core applications is down, affecting several customers<br>• DDoS attack |
| Critical/showstopper/ red alert situations<br>(affects business) | • Affects our company's bottom line<br>• Major impact on revenue, reputation, and legal affairs | • Software bugs and vulnerabilities<br>• Malware<br>• APT<br>• Ransomware<br>• Phishing and social engineering<br>• Insider threats |

## Assignment and routing

The incident is now assigned to a PitStop technician for further investigation and diagnosis. We accomplish this using incident rules provided by our ITSM application that define the routing order and assign incidents to selected groups. Let's say a printer on the third floor is down and an incident is logged. Our ITSM tool captures the user's location in the incident form, and because of where the incident originated, it's routed automatically to the PitStop technician on the third floor. A notification is also sent to the PitStop technician soon after the incident is routed, so the technician knows to start working on the issue.

## Open communications

After an incident is assigned, a PitStop technician opens communications with the affected end user. The technicians ask and answer questions, and provide end users with regular updates before, during, and after the incident. It's important for PitStop technicians to communicate well with end users at every step.

We primarily use three methods of communication:

- An email thread starts soon after the technician initiates a conversation with the end user within the ITSM tool, ensuring that all communication are in one place. Regular notifications and updates are sent to the affected end users until the incident is resolved and closed.
- We use announcements in our ITSM tool to publish help desk-related information across the organization, or to particular end user groups with regard to server issues, service updates, license renewal, and so on. It's important that PitStop technicians and end users in our company remain cognizant of incident details.
- For quicker resolution and more details about the incident, the PitStop technician calls the end users on their desktop or mobile phones.

## Escalation

The incident now moves to in-progress status and shows the life cycle stage of the ticket. The PitStop technician updates the status to keep the end user informed and to stick to the applicable SLAs. If the PitStop technician is unable to resolve the ticket, it's escalated to the incident coordinator who reassigns the ticket to a technician with a more advanced skill set.

For desktop incidents with low priority, the SLA is usually set to three to five days and end users should receive a response within four hours, and for a medium priority incident, it's set to one day and end users should receive a response within two hours.

## Closure

When no escalation is required, the PitStop technician can close the ticket; this is the final step in the incident life cycle. This involves logging the resolution into the ITSM tool for future reference before closing it. Once closed, incidents are still accessible to the PitStop technicians and the incident coordinator so that if the end user calls back, the technician can view the history and reopen the incident if necessary.

## Best practices for desktop incidents

- Have multiple channels for ticket creation to enable end users to raise tickets easily via email, chat, portal, and phone call.
- Encourage end users to find answers even before they approach a technician for help with self-service.
- Have technicians utilize mobile apps to manage your help desk and respond to employee requests—even when they're away from their desk.
- Automate user management by integrating with the company's Active Directory.
- Sort your end users into groups that are based on their department and managed by the service desk.
- Proactively manage frequently recurring incidents like password resets by using self password reset tools that allow the system admins to provide employees with access to a web-based self-service portal so they can securely reset their passwords.
- Automate activities that improve the efficiency and productivity of the team, including routine desktop incidents like categorization, prioritization, and assignment.
- Have a knowledge base in place to enable technicians to search for existing solutions, so they can efficiently resolve issues.
- Don't keep your end users waiting endlessly. Adhere to your SLAs.
- Keep end users notified at every stage of the incident life cycle.
- Automate notification activities to save time.

# Big bang

(major availability incidents)

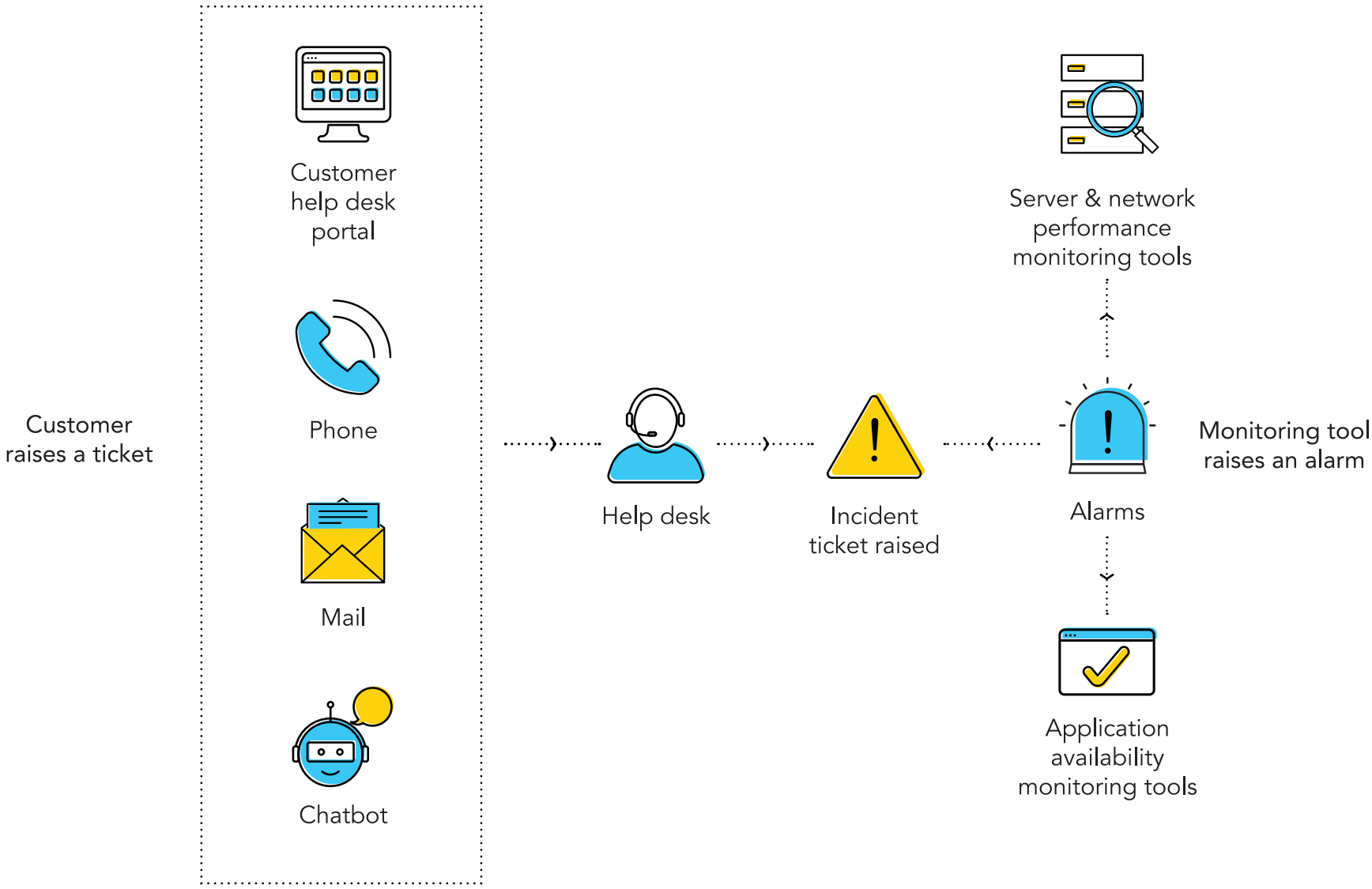
Any incident that affects many users, deprives the business of one or more crucial services, and demands a fast and efficient response is considered a major incident. In the world of cloud technology, achieving 99.99 percent availability has become the standard. At Zoho, our commitment to our customers is to ensure 99.99 percent availability.

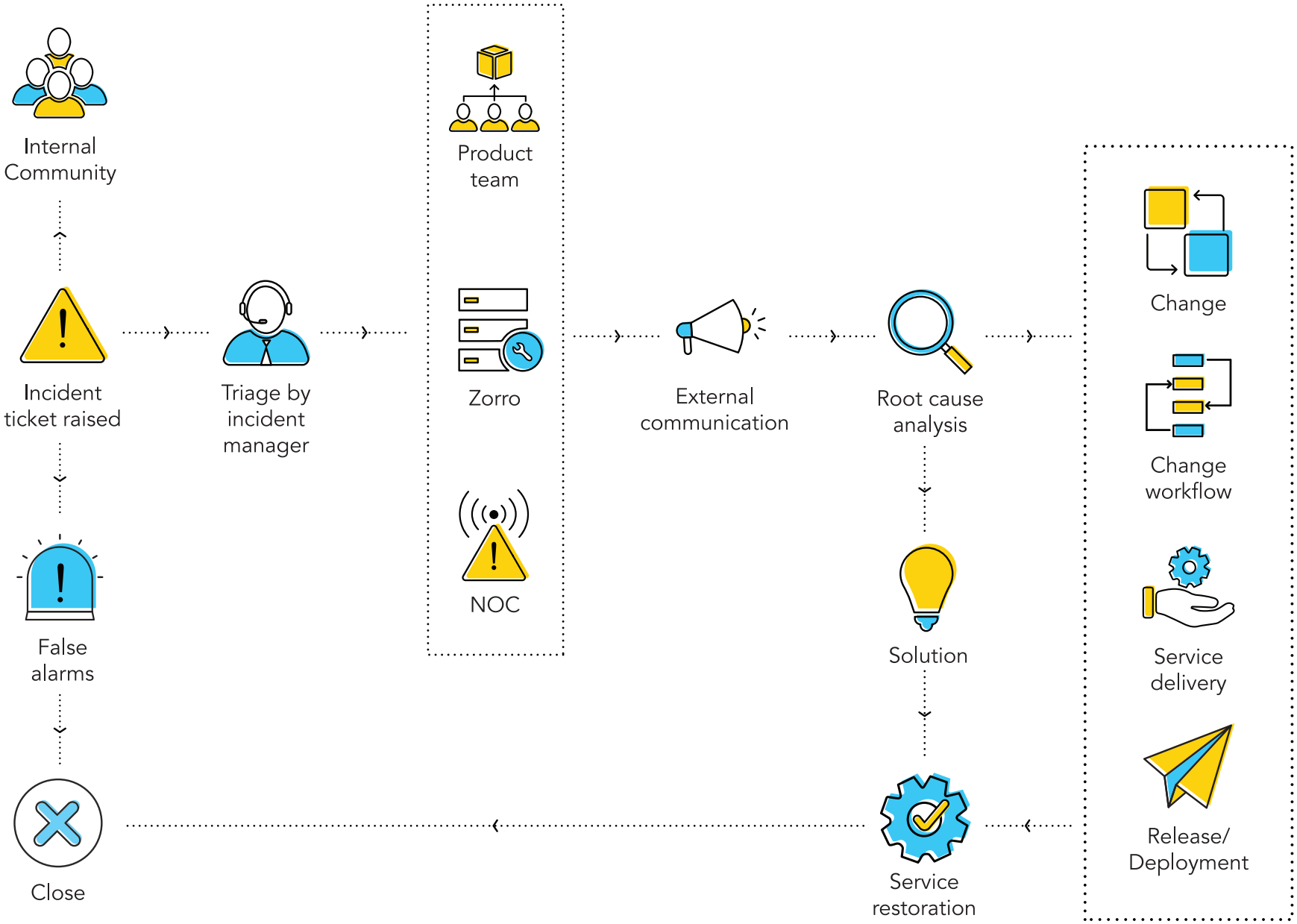
Customers can check the availability of our services at our Zoho Status Page. When a major availability incident hits, we follow the big bang IM process; this includes facilitating collaboration, aligning stakeholders, informing customers, and ultimately working on the incident continuously until resolution.

This section deals with three different availability issues:

- Network issues
- Physical server issues
- Application issues

The figure below shows the process we follow during an availability incident.



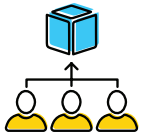

Internal Community
Incident ticket raised
False alarms
Close
Triage by incident manager
Product team
Zorro
NOC
External communication
Root cause analysis
Solution
Service restoration
Change
Change workflow
Service delivery
Release/ Deployment

# Teams, roles, & responsibilities

**Incident response team (IRT)**

**Incident manager:**

Serving as the captain of the ship who oversees the incident, the incident manager works with the NOC, Zorro, and product teams, and their respective incident coordinators to resolve issues and maintain SLAs.



**Engineering & development (product teams):**

For an application-related incident, the individual product team is the primary point of contact for the incident manager. The product team's engineers are typically the group that resolves issues during availability incidents.

**Software as a service (SAS) team:**

Handles the inventory of data center assets.

**Network operations center (NOC):**

Handles network availability incidents.

**Incident coordinator:**

A designated incident coordinator is assigned to every product team and is responsible for assessing and coordinating an availability incident.

**Servers & maintenance (Zorro):**

When an incident is identified as a server-related incident, the Zorro team assists. The Zorro team handles provisioning and maintenance of the servers in the data centers.

**Service delivery (SD) team:**

Handles the pushing of updates to all Zoho applications.

**External communications manager:**

The incident manager acts as our external communications manager, providing customers with frequent updates on outages.

## Detect

**Site24x7** is an availability monitoring tool we use to monitor our application availability across various locations. This application integrates seamlessly with our ITSM tool, recognizes the unavailability of applications, and sends proactive alerts to create incidents in our ITSM tool. In case of false alarm, the incident is closed.

## New incident

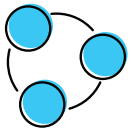
We've configured our major IM process in our help desk tool using a request life cycle (RLC). Whenever an incident is created, notifications are sent to the incident manager and the incident coordinators of the Zorro, NOC, and concerned product teams. These notifications include the incident ticket number, description, and the incident priority. Once the incident is logged as a ticket in our ITSM tool, it's in the open state—the initial state in our RLC.

## Communicate with stakeholders

The incident manager, after taking the necessary information from the alert and the incident coordinator, opens communications with internal stakeholders.

**ITSM tool:**

From the incident ticket, an email is sent to the incident management, NOC, Zorro, and product teams to begin initial investigation.



**Zoho Connect:**

Connect is a team collaboration software, like an internal Facebook-like application, that connects all stakeholders and enables open discussions during an incident. We have a group called Incidents which includes over 1,200 members including responders, key stakeholders, and decision makers to ensure greater transparency and coordination.

**Zoho Cliq:**

A business collaboration chat software that enables the incident manager, incident coordinators, product teams, and other stakeholders to give quick updates, share files, and search for a contact or conversation from the past. Group chat enable us to contact and add more responders and resolvers as needed to work through incidents faster.

**Document the outage:**

A conference call or a discussion thread is not enough to help everyone see what's going on and what lies ahead. The stakeholders and customers need meaningful progress reports, reassurances that the incident can be fixed, and no surprises. The incident manager keeps an incident state document on Zoho Writer to provide a clear place to see how, why, and when the incident occurred, the actions taken or underway, shared data, and an understanding of the clear path forward to resolution.

This document can be edited, commented on, and shared across the organization. The incident manager shares this document on the Connect thread corresponding to the incident, and also uses it for root cause analysis (RCA). This is also a great way for the incident manager to record key observations and decisions that happen in unrecorded conversations on other mediums, such as chat and discussion threads.

## Assess

We handle many kinds of major availability incidents and we bring in multiple teams to accomplish the fix. The response given to these incidents depends on many factors such as coordination, communication, and management. For successful incident response, all these factors must work together. To optimize the response, we need a common language to communicate, and the order by which teams have to be involved and tasks executed has to be well-defined.

Once an availability incident comes from Site24x7, triage between teams begins. Our incident manager acts as the triage officer, bringing together the NOC, Zorro, and product teams. A channel that includes NOC, Zorro, and the product team is created on Cliq to identify if the issue is related to the network (NOC), server (Zorro), or a product, so that the ticket can be delegated to the right teams and resolved.

The incident manager starts with assessing the incident by asking a few questions in order to communicate the right information to the stakeholders and customers.

- When did the outage happen?
- Is the outage visible to customers?
- How many customers are affected?
- How many support tickets are in?
- Which team (NOC, Zorro, or product) handles the fix?
- Is the team equipped with the right resources on that particular day?
- Does the resolver team agree to its communication schedules and protocols before getting to work?

Once the incident ownership has been identified, and it receives a priority that has been deemed a major incident—meaning the incident is both urgent and has an impact on the organization—the incident manager sends the initial external communication.

## Communicate externally

The incident manager is now reasonably clear on the incident and the team's involvement, and has to get the word out to the customers as quickly as possible. The incident manager gets help updating the blog on the unavailability from the communications team.

During an outage, we make a blog announcement that includes details such as the date and time of occurrence, the nature of the incident, and the remedial actions with frequent updates. Whenever customers try to access the service during an outage, they are redirected to the blog announcement so they can stay updated on the happenings.

In addition to the blog, an announcement post is also made in the community during an outage where we provide frequent updates and answer customer questions. Customers can also check for the availability of the service on our status page.

## Delegate

The incident manager works with the incident coordinator in the NOC, Zorro, and product teams to manage all incident operations, application of resources, and responsibilities of everyone involved. Once the NOC, Zorro, and the respective product teams get back through the Cliq channel and the team ownership has been identified, a set of tasks is automatically triggered through the request life cycle to the team that owns the incident as shown below.

## Send follow up

The incident manager regularly pings the resolver team to receive quick updates about the progress of the incident, which they will forward to the customer. Short, concise details—that include the start of the downtime, a short description of the known cause of the downtime, estimated time for restoration, and the scheduled time for the next status update—are frequently updated in the forum and the blog to keep customers informed.

## Resolve and close

After the incident no longer affects customers, it's considered resolved, and either a technician will manually close the ticket, or after sufficient time has passed, the ticket state will change to closed automatically. The incident manager sends out the final internal and external communications, and initiates the RCA using the incident state document as the base.

Here's our checklist for resolving (and closing) tickets:

- ✓ Is the incident resolved to the satisfaction of the ticket owners?
- ✓ Are the resolvers taking care of cleanup tasks?
- ✓ Have all the related tasks been closed and relevant users notified?
- ✓ Has the incident manager notified all parties?
- ✓ Most importantly, have the customers been notified of the resolution?
- ✓ Have all stakeholders agreed to major incident closure?
- ✓ Has RCA been recorded and initiated?
- ✓ Has the RCA meeting agenda been sent to the resolver groups?
- ✓ Has the service desk been notified of closure?

We check these off to complete the major incident process as cleanly as possible and to ensure that we didn't miss anything.

# Best practices for major incidents

**Clearly define a major incident:**

We call it a major incident when it affects many users, deprives the business of one or more crucial services, and demands a response beyond the routine incident management process. Sometimes, a high priority desktop incident can be perceived as a major incident. A VIP's laptop broken during a user conference is a high priority incident, but it's certainly not a major incident. To avoid any confusion, you must define a major incident clearly based on factors such as urgency, impact, and severity.

**Have a communication plan in place:**

The communication plan should include the details of the event (how, when, action plan, estimated time of fix, and update interval time), the parties involved, and how often to communicate.

**Set up SLAs:**

Set up separate response and resolution SLAs with clear escalation paths. If the team is short-staffed for the day, don't hesitate to pull in the required resources from other teams to work on the resolution and ensure the SLA is not impacted.

**Have an exclusive IM process:**

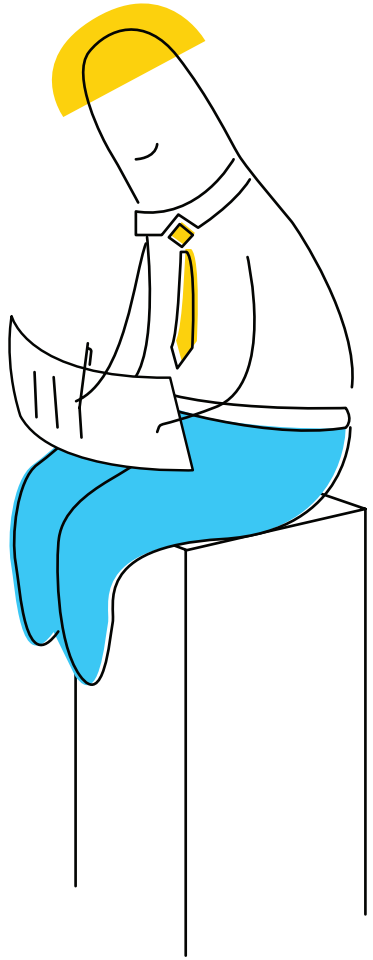
Separate workflows or processes for major incident management will help you efficiently deal with various types of incidents, such as service unavailability or performance issues and hardware or software failure, so you can ensure seamless resolution.

**Bring in the right resources and teams:**

Ensure that the right team and resources are working on incidents with clearly defined roles and responsibilities

**Document the process for continual service improvement:**

As a best practice, our incident manager captures details such as the number of personnel involved in the process, their roles and responsibilities, the communication channels, tools used for the fix, approval and escalation workflows, and the overall action plan used for the response and resolution in the incident state document. The stakeholders, including top management, evaluate this document to ensure continual service improvement.



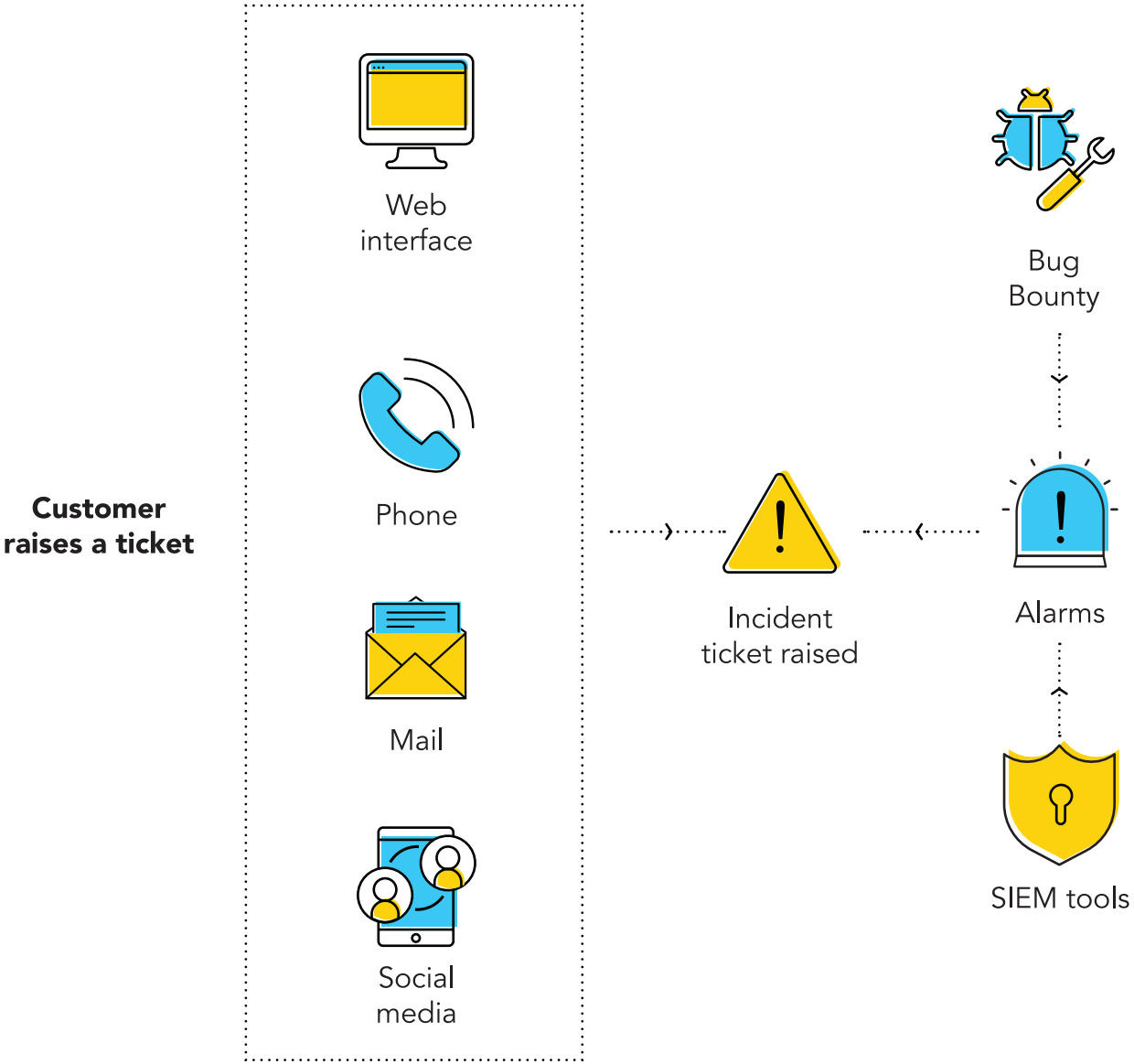
# CyberSec

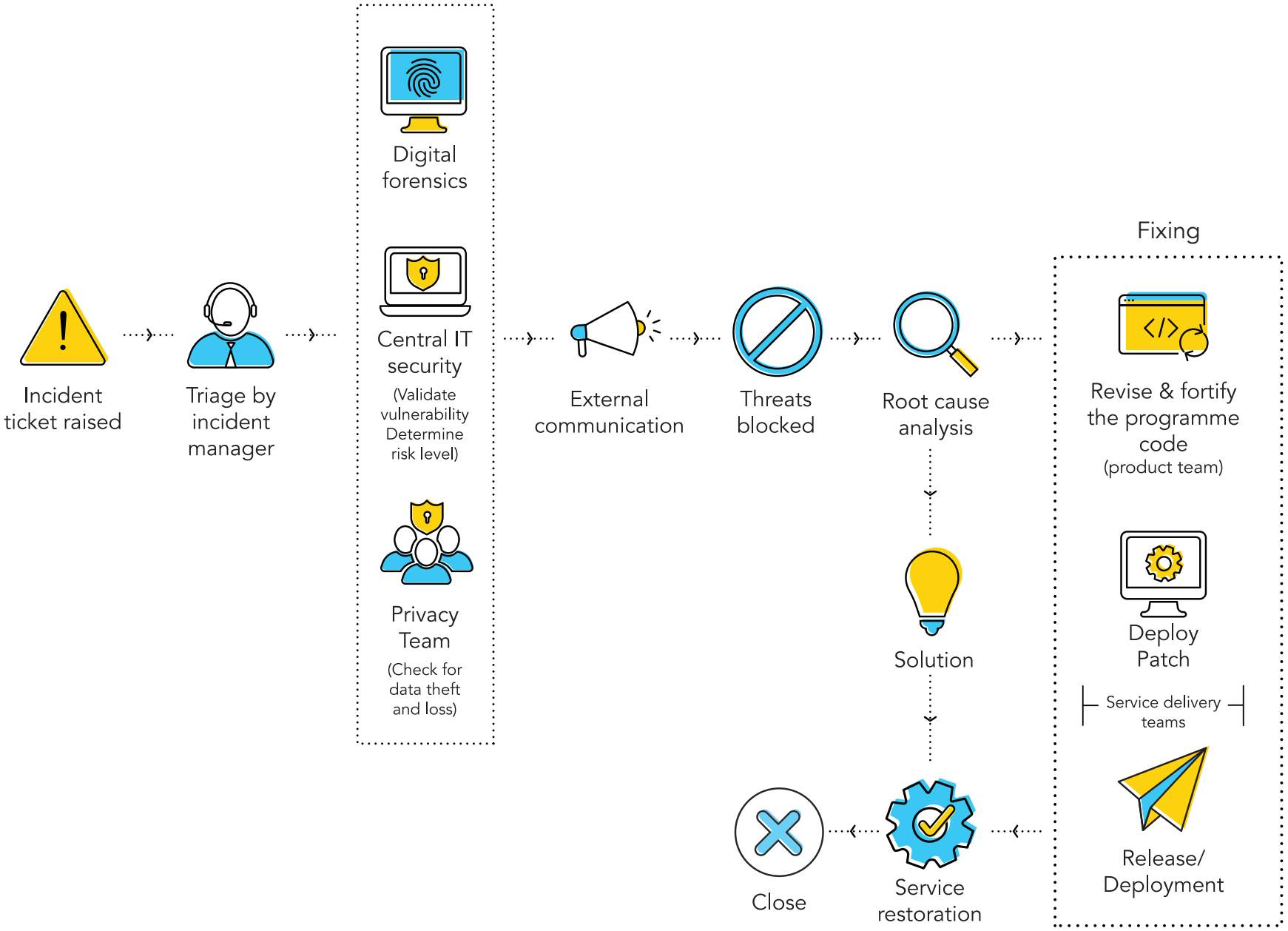
(Showstoppers/security incidents)

Security is the bedrock in our organization. Our IT security team protects the confidentiality, integrity, and availability of our systems and data, securing our organization against malware, APTs, ransomware, phishing, social engineering, insider attacks, and other security threats. We have a group of white hat hackers called the red team that is continually attempting to circumvent our security checks. The security coordinators are involved in the product development life cycle, ensuring that security is built into every product we develop.

Anticipating that one day we could be the target of a cyberattack, we want to be prepared. The CyberSec process is a mature vulnerability detection, containment, coordination, and recovery process that makes our company resilient against cyber threats. It can help us recover from security breaches by minimizing the exposure time and impact of threats on data, applications, and our IT infrastructure.

The below figure is our CyberSec process flow.




Incident ticket raised
Triage by incident manager
Digital forensics
Central IT security
(Validate vulnerability Determine risk level)
Privacy Team
(Check for data theft and loss)
External communication
Threats blocked
Root cause analysis
Solution
Service restoration
Close
Fixing
Revise & fortify the programme code
(product team)
Deploy Patch
Service delivery teams
Release/ Deployment

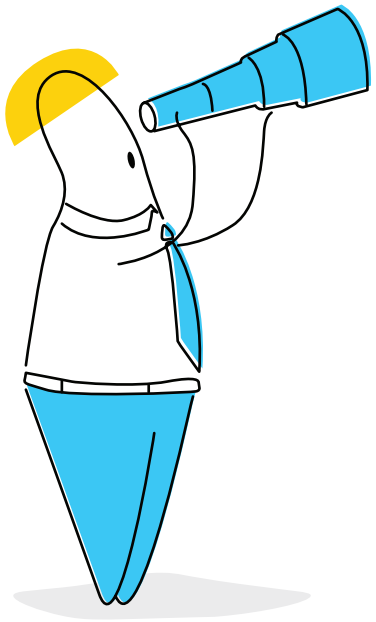
# Teams, roles, & responsibilities

| Team | Roles | Responsibilities |
|---|---|---|
| Incident management | Incident response and incident coordinators | Manages the cybersecurity incident from detection to resolution |
| Central IT security team | IT security | Continuously monitors and analyzes the security proceduresof our products |
| Security incident response team (SIRT) | The SIRT includes the IM team (incident manager, incident coordinators) and the central IT team | The SIRT assesses the impact, ensures SLA, coordinates with privacy, legal, and product teams and top management during crisis situations |
| Top management | Business decision makers | Take inputs from the incident manager, assess the business impact, and make key business decisions; for example, deciding if the internet connection of a compromised system should be shutdown. If yes, when is the appropriate time? Also decides when to contact the authorities |
| Legal team | Legal department/legal advisors | Assesses the contractual and judicial impact of an incident. Gives assurance that the incident response activities meet legal and regulatory standards along with the organization's policy boundaries. Guides the company on the steps of filing a complaint |
| Privacy team | Data protection officer (DPO) | Handles data privacy, administers privacy regulations, and answers to regulators in case of a data loss |
| Product team (every product team has a security coordinator) | Product engineers | Fixes vulnerabilities and releases product updates |
| Red team | White hat hackers | Attempts to break in to our applications, bypass security protocols, and expose potential cybersecurity vulnerabilities to ensure the security of our products |

# The process

## Detect

Our employees have the greatest potential to help the organization detect and identify cybersecurity incidents. They play a significant role in detecting security threats. Every member of our organization is aware of the ways to alert our security team when they notice something abnormal on their computer or mobile device.



We also have a Bug Bounty program to encourage and reward employees who detect and report a security vulnerability to Zoho. Our employees can report a security incident through:

- The self-service portal
- A toll free number
- Bug Bounty app
- Social media
- Email

## Communicate internally:

*Refer Big Bang - Page 21*

## Assess

When a security alarm sounds, it's crucial to first assess what happened, pull together the details, perform a business impact analysis (BIA), and take the right measures. The SIRT, which includes the incident manager and the central IT security team, begin triage to collect and analyze information before acting.

The incident manager convenes a meeting that includes the central IT security team and the privacy team along with the data protection officer (DPO) and opens a Cliq channel for follow-up discussions and updates. The central IT team collects all available information and conducts a forensic investigation to examine the magnitude and depth of the attack; the privacy team identifies any data privacy breach.

The SIRT asks the questions below to assess the incident and its impact.

| Incident manager | Central IT team | Privacy team |
|---|---|---|
| • Who identified and reported the incident?<br>• When was the incident identified and reported?<br>• Where was the incident discovered or located?<br>• What impact does the incident have on business operations?<br>• Has there been data loss and is privacy involved?<br>• What is the extent of the incident with the network and applications? | • Is it a security attack?<br>• Was the attack successful?<br>• What is the source IP score?<br>• What is the destination IP score?<br>• What is the threat feed score?<br>• What is the vulnerability score?<br>• What assets were compromised?<br>• What are the associated vulnerabilities?<br>• How should the organization respond to this attack?<br>• What could happen if the incident is not contained?<br>• Is it necessary to preserve evidence?<br>• What sources of evidence should the organization acquire?<br>• Where will the evidence be stored?<br>• How long should the evidence be retained? | • Has the user account involved been compromised?<br>• Was the data equipment in the device encrypted?<br>• Did the compromised account have access to sensitive information?<br>• What activities did the attacker conduct?<br>• What is the attack density?<br>• What is the number of individuals potentially affected?<br>• Was this event associated with any other event or an artifact?<br>• What is the level of risk to individuals and the company?<br>• What are the types of controls in place to mitigate the risks? |

## Contain

A security incident is analogous to a forest fire, and it has to be contained as quickly as possible. The SIRT quarantines the infected or compromised networks and devices affected by viruses or other malware, and installs security patches to resolve malware issues or network vulnerabilities.

When the incident is identified as the result of a software vulnerability, the SIRT disables the feature used in the exploit, writes a custom firewall rule blocking specific requests targeting the vulnerability, or even temporarily uninstalls the software as preventive actions. The attacking IP address is also blocked to prevent any further attempts.

Meanwhile, the incident manager gathers the necessary details, and opens up external communications.

## Communicate externally

Communicating externally is a key step in cybersecurity incident response. The incident manager works with top management to control the flow of communication to ensure the right information is provided at the opportune time.

For example, an internal hacking attempt will likely not warrant communication with the media or the authorities. On the contrary, if the incident involves exposure or theft of sensitive customer records, then it might be mandatory to report to the media and the authorities in light of the GDPR and other consumer privacy regulations.

| Who? | What? |
|---|---|
| Customers | Details of the incident include: the date and time of occurrence, a description of the issue stating if any customer data was lost or stolen, steps being taken to mitigate the risks, and estimated time for recovery. |
| Media | Sometimes, media attention cannot be avoided and our company spokesperson issues a statement about the incident and its impact to show our commitment and capability to manage the incident. |
| Police | In cases of criminal intent, the SIRT, legal team, and top management work together to report the incident to law enforcement authorities. |

## Delegate

Once the investigation is concluded and necessary steps taken to contain the attack, the incident moves to the delegation state. The SIRT delegates the resolution responsibility to product engineers to revise and fortify the program code, ensuring resolution of the vulnerability.

## Resolve

Eradication and recovery is accomplished as a singular step. The eradication phase includes a more permanent fix for infected systems. If the threat gained entry from one system and proliferated into other systems, then the SIRT aims to remove the threat and erase all traces of the attack from our devices and network via antivirus software, hardware replacement, or network reconstruction. Our objective is to return systems to "business as usual." Here's our SIRT's eradication checklist:

- ✓ Have all infected systems been hardened with new patches?
- ✓ Do the systems and applications have to be reconfigured?
- ✓ Have all possible entry points of the attack been patched?
- ✓ Have all processes to eradicate the threat(s) been covered?
- ✓ Are any additional defense measures needed to eradicate the threat(s)?
- ✓ Have all malicious activities been eradicated from affected systems?

In case of a software vulnerability, the incident is delegated to the engineering team of that particular product. The product engineers fix the vulnerabilities and release the software update.

## Review

After a resolution is implemented by the product team, it's typically verified and reviewed by the head of engineering, the SIRT, and the privacy team. On approval by the SIRT and privacy teams, the incident moves to closure. This step ensures that nothing was missed, and the incident has been fixed and prevented from recurring. Tempting as it may be to skip this step, ensuring there is a review of the implemented fix is strongly recommended.

## Close

All cybersecurity incidents, like any other incident, need to be properly closed. The incident manager closes the incident. As a post incident measure, the incident manager sends out communication to all affected parties, media, and the authorities confirming that the threat has been contained.

### Here's our checklist for closing tickets:

- ✓ Is the incident resolved to the satisfaction of resolver groups?
- ✓ Are the resolvers taking care of clean up tasks?
- ✓ Have all the related tasks been closed and notified?
- ✓ Has the incident manager notified all parties?
- ✓ Most importantly, have the customers been notified of the resolution?
- ✓ Have all stakeholders agreed to security closure?
- ✓ Has RCA been recorded and initiated?
- ✓ Has RCA been initiated by the incident manager?
- ✓ Has the service desk been notified of closure?

The final phase of the security incident response life cycle involves RCA and feedback loops. The incident manager initiates RCA, as it's crucial to learn from each incident for continual service improvement.

## Best practices for security incidents

**Have a well-defined incident response process:**

Have an actionable process to identify, address, and manage the aftermath of a security breach or cyberattack in a way that limits damage and reduces recovery time and costs. Ensure the incident response plan aligns with the company policies.

**Clearly define the teams, roles, and responsibilities:**

Identify the teams that are largely responsible for each phase or step (e.g., containment, eradication, and recovery) in the incident response process. Identify the key people from the respective departments and teams, who will serve as their backup, and how to reach them day or night. As a best practice, we create a RACI chart that helps us identify the people who are responsible, accountable, consulted, or informed (RACI) for defined activities before and after an incident.

**Take stock of your data:**

It is important to assess your organization's data to know what needs the most protection during a data breach.

**Have a communication plan in place:**

Having defined lines of communication to engage stakeholders and manage communication between the security incident response team and other groups is crucial to successful incident recovery. A communication plan ensures that everyone follows protocols during an emergency in contacting stakeholders, partners, service providers, media, authorities, and customers.

**Gather evidence against the attacker:**

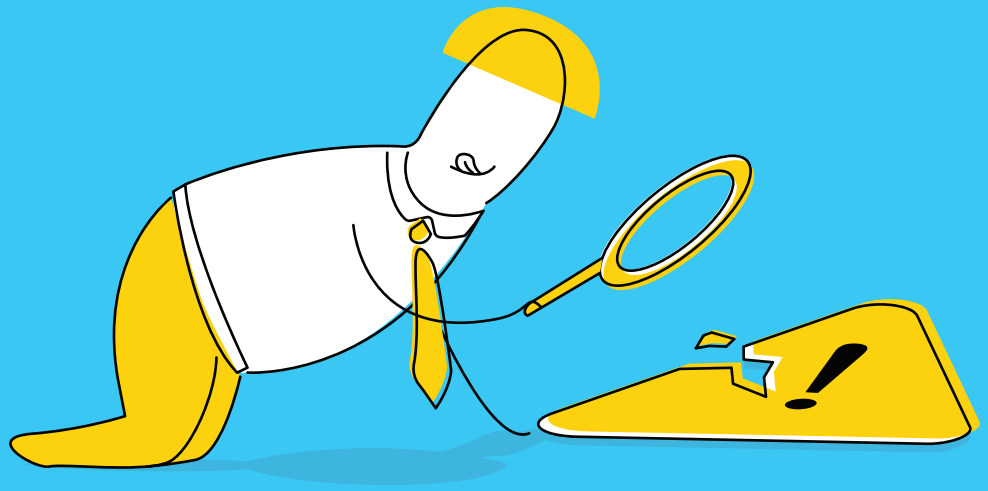
Taking corrective actions during a crisis, such as taking systems off the network or cleaning up systems, can result in alerting an attacker or destroying vital evidence. The organization has a lot to lose in the court of law if evidence is destroyed or insufficient. Collect the evidence in a forensically sound manner so that it can be credibly presented in the court of law.

**Stay calm during a crisis:**

Dealing with security incidents can be quite stressful. It's necessary to stay calm when an attack occurs, and follow the incident response plan.

**Conduct root cause analysis:**

A post-incident analysis involving the incident response team and other resolver groups can help provide insight into the source of the issue and prevent recurrence.



# **ROOT CAUSE ANALYSIS** (RCA)

## What is RCA?

Root cause analysis (RCA) is a systematic approach that drills deep to identify the root cause of an incident by repeatedly asking "why" questions until no additional diagnostic responses can be provided. It typically involves an analysis or a discussion soon after an incident has occurred. An additional resource, the incident state document, serves as a written record of what happened before and during the incident, and gives answers to the questions necessary to conduct a root cause analysis.

The incident state document, also known as an incident report, is the best place to start with root cause analysis. However, it is critical to dig deeper than just what the form states. At Zoho, we create a problem record from within the incident ticket to perform a fully fledged RCA through our ITSM tool.

## Why perform RCA?

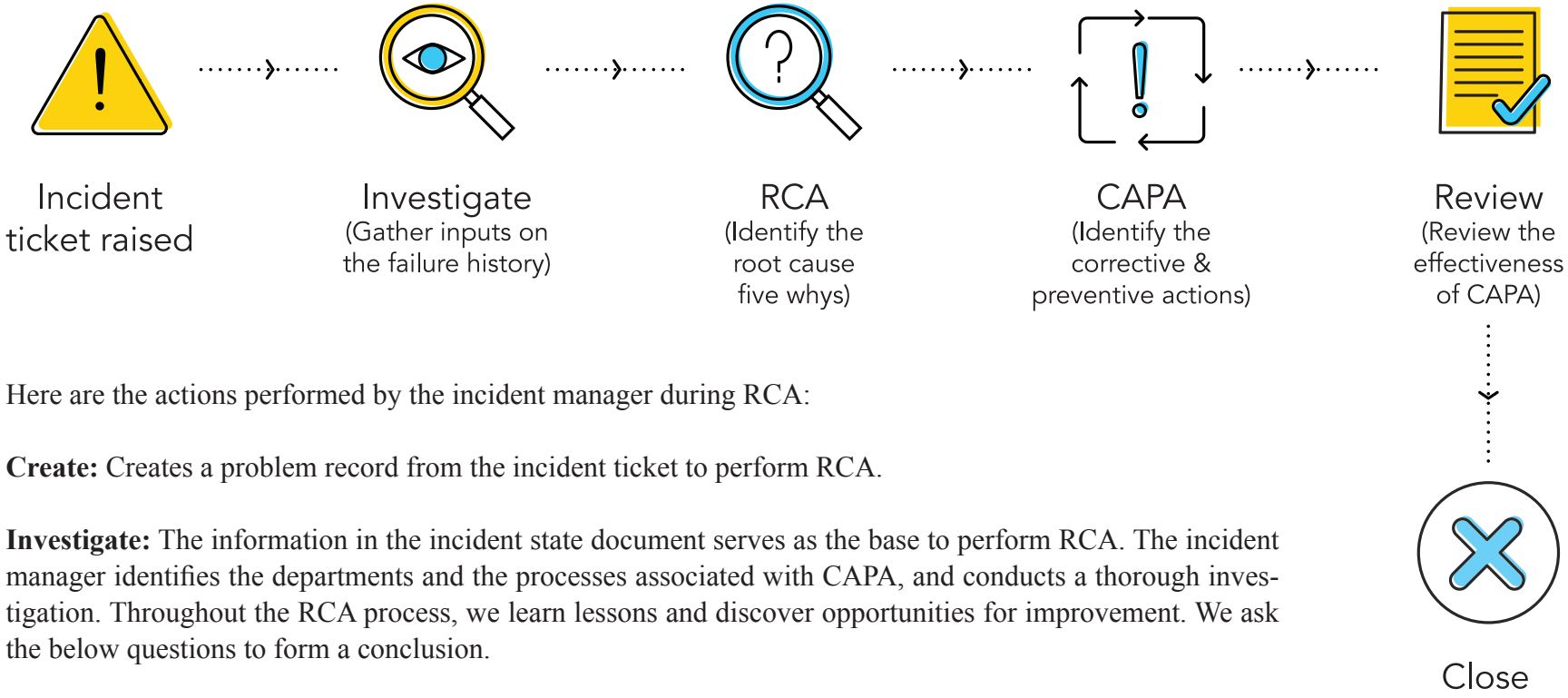
At Zoho, we never let a good crisis go to waste. We see an unfortunate event as an opportunity to learn from our mistakes, identify where the processes or systems failed, and be better prepared to handle similar incidents in the future.

## RCA principles

- RCA is conducted to determine the factors that resulted in the incident and to take corrective actions rather than to simply treat the symptoms.
- A successful RCA is performed systematically with conclusions backed by real evidence.
- Most often, there is more than just one root cause for an incident.
- "If you are not making mistakes, then you are not doing anything." At Zoho, we believe in learning from our mistakes. Having a "blameless" RCA process allows our employees and teams to give the exact details of their approach such as the actions they took and the assumptions they made when handling the incident.

## RCA process

Corrective Action Preventive Action (CAPA) is our structured approach to investigate, identify the root cause, take corrective action, and prevent the recurrence of the root cause(s).




Here are the actions performed by the incident manager during RCA:

**Create:** Creates a problem record from the incident ticket to perform RCA.

**Investigate:** The information in the incident state document serves as the base to perform RCA. The incident manager identifies the departments and the processes associated with CAPA, and conducts a thorough investigation. Throughout the RCA process, we learn lessons and discover opportunities for improvement. We ask the below questions to form a conclusion.

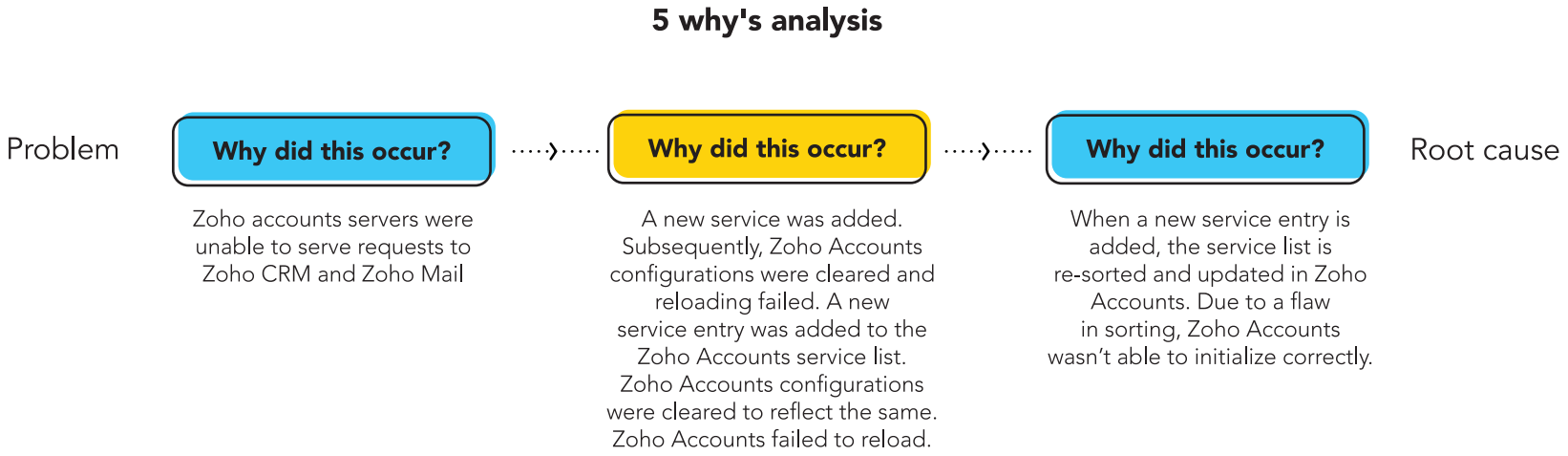
| Stage | Questions | Use case |
| --- | --- | --- |
| Incident summary | • When was the incident first noticed? (date and time of the incident)<br>• Where did the incident occur? (Location of the incident, network, server, product, and others)<br>• What type of incident? (failure/issue reported)<br>• What's the real issue and what's happening? (observations from involved teams)<br>• Affected parties (stakeholders, customers, or both) | **Root cause statement:** Zoho Accounts servers were up but were unable to serve any requests resulting in Zoho CRM and Zoho Mail facing accessibility issues.<br><br>**Incident summary:** This availability incident was triggered on 22-Jan-2019 15:31 IST and ended on 22-Jan-2019 at 15:52 IST. The incident was detected by Site24x7 and affected Zoho CRM and Zoho Mail services.<br><br>The event was mitigated by taking the following actions:<br><br>**Temporary fix (immediate):** The problem-causing service entry was removed immediately within 15 minutes after the outage. Zoho Accounts servers were up and services were accessible within 20 minutes.<br><br>**Permanent fix (next day build update):** When a new service is added, we clear the cache, which clears the JVM cache and repopulates it. As an alternate step, we are now repopulating the newly added service in the JVM cache so that even if repopulation fails in the future, the older service list will be used. |
| Impact | • How long did the impact last and how was it mitigated?<br>• What did the customers see?<br>• How many were involved or impacted? (for example: customers of a suite or product)<br>• How many support tickets were raised? | The downtime lasted for 21 minutes,  and Zoho CRM and Zoho Mail customers were not able to access the services.<br><br>20 support tickets were raised following the incident through phone call, email, and chat. |

| Stage | Questions | Use case |
| --- | --- | --- |
| Response | • Who responded and when?<br>• What was the response time? | The incident was detected by customers, and the incident coordinators of Zoho CRM and Zoho Mail teams responded to the incident informing the incident manager and involving the heads of product teams and other stakeholders.<br><br>The incident was responded to within 15 minutes of occurrence, and a temporary fix was provided. |
| Recovery | • How was the service restored?<br>• What surprises did the resolver groups have to handle?<br>• What circumstances were not anticipated?<br>• Were there any useful workarounds or solutions that cropped up during the crisis? | **Temporary fix (immediate):** The problem-causing service entry was removed within 15 minutes after the outage. Zoho Accounts servers were up and services were accessible within 20 minutes.<br><br>**Permanent fix (next day build update):** Whenever a new service is added in Zoho Accounts, a list in which all services are stored will be cleared. When populating the now empty list, we try to sort the services obtained from the database. Since this list is no longer needed, we have completely removed it from our codebase. |
| Timeline | • A detailed incident timeline, in chronological order, timestamped with the time zone. | 22-Jan-2019 15:27 IST: A new service entry was added to the accounts.<br><br>22-Jan-2019 15:30 IST: Zoho Accounts configurations were cleared to reflect the new entry. Subsequently, Zoho Accounts servers failed to reload.<br><br>22-Jan-2019 15:31 IST: Zoho Accounts servers were down and the services were inaccessible.<br><br>22-Jan-2019: 15.51 IST: A temporary fix was executed in 20 minutes.<br><br>22-Jan-2019 15:52 IST: Zoho Accounts became stable and services were accessible again. |

| Stage | Questions | Use case |
|---|---|---|
| Lessons learned | • What could be done to prevent this class of incident from recurring?<br>• If we had to do it all over again, what would we do differently? | The sorting of services was part of the old UI. As the listing was not required anymore, we removed it from the codebase.<br><br>We also identified and removed similar functions where the same sorting algorithm is employed so that this downtime does not occur in the future. |

## RCA

The incident manager determines the root cause of the incidents using the "5 why's" technique that involves repeatedly asking the question "why?" until the root cause is identified. The purpose is not to place blame, but to uncover why an incident occurred in the first place.




Note: Sometimes it may take just three "why?" questions to reach the root cause; often, it requires more. It takes a while to master the art of questioning, but when the right questions are asked, the root cause can be identified quickly. In this case, the root cause was identified with just three questions.

# CAPA

Simply put, corrective actions are based on an adverse event that happened in the past. Preventive actions are based on thwarting an adverse event in the future. Corrective Action Preventive Actions, typically referred to as CAPA, are integral parts of our continual improvement process.

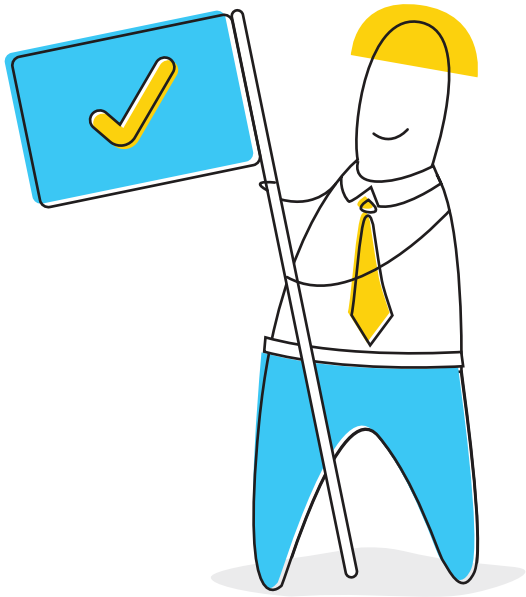
The success of RCA requires careful management of the action plan. So the next stage of the RCA process is to establish a proposed action plan defining the list of corrective actions and preventive actions. The action plan should define the time frame in which the actions will be completed and who handles each task.

Here's our checklist to ensure a systemic action plan:

- ✓ Are there corrective actions listed that are not supported by the analysis?
- ✓ Are the corrective actions clear and appropriate for the cause?
- ✓ Are the corrective actions listed in order of priority?
- ✓ If a third party is involved, will the action items be delivered within the intended time frame?
- ✓ Are the corrective actions likely to cause unintended consequences?
- ✓ Are the corrective actions under the control of management?
- ✓ Are the corrective actions likely to prevent recurrence?
- ✓ Has the department/action owner agreed to do the corrective action?
- ✓ Does each corrective action have a clear owner and due date?

The preventive actions process is to build in safeguards and process changes to prevent non-conformance. As a proactive measure, we:

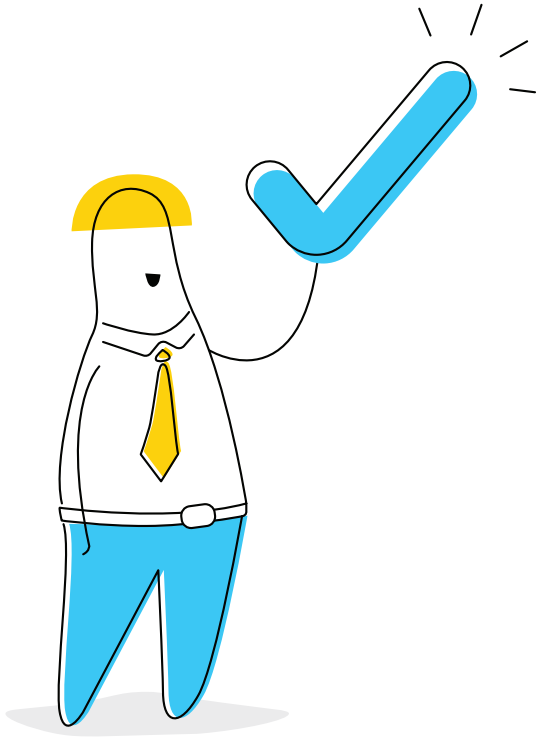
- Analyze processes and services for negative trends that could escalate an incident.
- Perform risk analysis to uncover latent hazards.
- Conduct training programs to enhance our employees' skills and to be better prepared during an incident.
- Introduce disaster recovery, security, and contingency plans for unpredictable crisis situations.
- Set up preventive maintenance to ensure our services are always safe, available, and performing optimally.
- Perform audits to assist in streamlining processes and to deliver quality service.



## Review

Finally, the RCA goes to management for approval to make the changes and prevent repeat problems. The incident manager establishes thorough follow-ups with resolver groups to ensure the corrective steps are effective and recurrence has been prevented.

The below checklist can be used by all IT teams to evaluate the overall quality of an incident response plan.



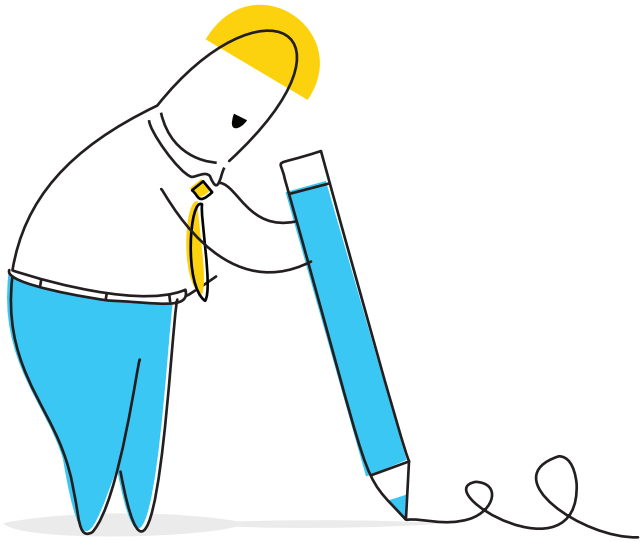
- ✓ Did the incident response plan help resolve the incident, or did the organization rely on “off-plan” activities?
- ✓ Is there a clear summary document for quickly understanding the incident?
- ✓ Is the entire incident analysis fact-based?
- ✓ Was the IT architecture robust enough to limit the impact between internal systems?
- ✓ How well did the associated teams, for example HR, legal, product, and so on, engage in assessment and communication?
- ✓ Was the data protection policy and practices adequate to identify and prioritize critical data?
- ✓ How effective was the communication plan?
- ✓ Have we asked “why” enough times to determine the root cause?
- ✓ Is there a clear link between facts, causes, and corrective actions?
- ✓ Did the analysis identify if the incident occurred previously?
- ✓ Were the resolvers identified earlier to handle this type of incident, or pulled in later based on their knowledge?
- ✓ Were the risks to the organization evaluated and managed?
- ✓ Has the RCA gone through the approval mechanism?

## RCA meetings

We conduct RCA meetings to get to the bottom of the issue, take necessary corrective actions to fix the issue permanently, and take preventive actions. The most important guideline for our RCA meetings is to learn and continually improve, not to assign blame or to vent.

Here are some tips to ensure an effective RCA meeting:

- Schedule a date and time that's convenient for all meeting participants, keeping in mind the team members who work in shifts and other distributed teams.
- Develop and stick to a meeting agenda that doesn't exceed two hours.
- Reserve a conference/meeting room with sufficient seating for all resolver groups, stakeholders, and top management.
- Schedule and invite participants (we use Zoho Calendar) one to two days prior to the RCA meeting, emphasizing the importance of the meeting and including the meeting agenda.
- Keep a written record of how long the meeting ran.



## Conclusion

Incident management processes are meant to shield organizations against adverse events. This is especially true for organizations like Zoho Corporation that rely heavily on the internet and computer networks, and deal with a vast amount of personal data.

An effective incident response policy focuses on four key aspects: risk management, regular audits, preventive measures, and most importantly, employee training. At Zoho, we have the right people, processes, and tools in place to stay ahead of future cyberattacks.

Now that you've seen how Zoho handles incidents, we hope your organization can design and pursue a similar strategy keeping in mind your business operations, task force, and company culture.

**Zorro, Zoho's infrastructure operations team, uses ManageEngine's IT management solutions to manage 18 data centers across four continents to serve 100 million users.**

**Over 280,000 organizations across 190 countries, including 9 out of every 10 Fortune 100 companies, use ManageEngine to manage their IT operations.**

## IT service management

- IT asset management with CMDB
- Knowledge base with user self-service
- Built-in and custom workflows
- Orchestration for all IT management functions
- Reporting and analytics
- Service management for all departments

## Unified endpoint management

- Desktop management
- Mobile device management
- Patch management
- OS and software deployment
- Remote desktop support
- Web browser security
- Monitoring and management of peripheral devices

## Identity & access management

- Identity governance and administration
- Privileged identity and access management
- AD and Azure AD management and auditing
- SSO for on-premises and cloud apps with MFA
- Password self-service and sync
- Office 365 and Exchange management and auditing
- AD and Exchange backup and recovery

## IT security management

- Unified SIEM for cloud and on-premises
- AI-driven user and entity behavior analytics
- Firewall log analytics
- SSH key and SSL certificate management
- Endpoint device security
- Data leakage prevention and risk assessment
- Regulatory and privacy compliance

## IT operations management

- Network, server, and application performance monitoring
- Bandwidth monitoring with traffic analysis
- Network change and configuration management
- Application discovery and dependency mapping
- Cloud cost and infrastructure monitoring
- End-user experience monitoring
- AIOps

## Advanced IT analytics

- Self-service IT analytics
- Data visualization and business intelligence for IT
- Hundreds of built-in reports and dashboards
- Instant, flexible report creation
- Out-of-the-box support for multiple data sources

## About ManageEngine

ManageEngine is a division of Zoho Corporation that offers comprehensive on-premises and cloud-native IT and security operations management solutions for global organizations and managed service providers. Established and emerging enterprises—including nine of every 10 Fortune 100 organizations—rely on ManageEngine's real-time IT management tools to ensure the optimal performance of their IT infrastructure, including networks, servers, applications, endpoints and more. ManageEngine has 18 data centers, 20 offices and 200+ channel partners worldwide to help organizations tightly align their business to IT.

## About the author

Meghna Reddy brings over a decade of expertise in IT, focusing on IT Service Management (ITSM) and content strategy. Known for her unique approach to turning complex technical information into engaging stories, Meghna excels in creating long-form content that makes intricate IT concepts accessible to diverse audiences. Beyond her professional pursuits, Meghna is dedicated to animal rescue and advocates for the voiceless.

ManageEngine

www.manageengine.com