

2024-05-23 Mackerel Drink Up 出張版@福岡



Mackerel開発チームの障害対応演習 ——新卒エンジニアが障害対応指揮官 を務めるに至るまでのステップ

株式会社はてな Mackerel 開発チーム アプリケーションエンジニア
Arthur (id:arthur-1)

ja.mackerel.io

自己紹介

id:arthur-1

株式会社はてな Mackerel 開発チーム
アプリケーションエンジニア

 @Arthur1__

> 2022 年新卒入社

顧客要望やdogfoodingから特定したMVPを検証可能な形で実装し、リリースの効果を可視化し続けることで仮説検証型開発を支えています



Mackerel開発チームの状況

- 自分も含めチームに開発メンバーがたくさん増えキャッチアップ不足
- 新規開発やインフラ更改でシステムの様子が大きく変化

限られた人だけが障害対応で指示し手を動かしていた

障害対応の学びの難しさ

障害対応の目標は障害（ユーザーがサービスを正常に利用できない状態）をいち早く復旧すること

障害対応はいつ起こるか、そしてその場に居合わせる人が誰かわからない

→本番の障害対応をスキルアップの場としにくい

自分も最初はこうだった

新卒入社後すぐの自分

- システム構成がわからないからグラフを見ても何の数値かわからない、飛び交う言葉もわからない
- 問題を認識しても、どうすれば良いか決断できない

→今では障害対応の指揮官ができるまでに成長した

Mackerel開発チームのやり方

障害対応のふり返し

- 経験を一般的な学びに昇華する手段
- 個人の過ちと捉えず、仕組みで解決する機会

障害対応演習 ⇐ 今回のトピックはこれ

- 災害と同じように定期的な訓練が必要

Mackerelの障害対応演習の設計

最近Mackerel開発チームSREのid:heleenによって
アウトプットされたものがあるので、ぜひご覧ください

- Mackerel で行った障害対応演習を紹介します
<https://developer.hatenastaff.com/entry/2024/04/30/162658>
- 障害対応をちょっとずつよくしていくための演習の作りかた
<https://speakerdeck.com/heleen/troubleshooting-exercises>

障害対応演習のメニュー

これらのメニューを、**チームやシステムの状況を元に組み合わせ**て実施する

- オペレーションの手順確認
- 障害対応フローの確認
- 検証用の環境での実践

オペレーションの手順確認

障害対応で起こり得る手動オペレーションの手順を学び、

実際に手を動かして練習する：

- アプリケーションのロールバック
- データベースのスナップショットからの復元
- アプリケーションをメンテナンスモードにする

インフラ構成が変化する（例えばクラウドシフト）際に特に有効

オペレーション手順確認の効果

- オペレーションができる作業者が増える
- 手順が記されたドキュメントが用意される
- マニュアル通りにやればできるという安心感を得る

私はSREと一緒に手順確認の演習を準備して実施（解説）
する経験をした

→手順の意味や具体的な方法を深く知る機会に

オペレーション手順確認の準備

私はSREと一緒に手順確認の演習を準備して実施（解説）
する経験をした

→手順の意味や具体的な方法を深く知る機会に

時間の制約なく、自分がわからないところに向き合える

障害対応フローの確認

座学

- 障害対応フォーメーションを組んでから解散するまでの一連の流れをおさらいする
- 各ステップで各ロールが**どんな心がけをすべきか、それはなぜか**を知る

ロールプレイ

- 発生した障害や原因というシナリオが決まっていて、座学で学んだ通りに障害対応フローを疑似体験する

ルールプレイのコツ

シナリオの細部まで決まっているとは限らない

細かいところは実際の障害対応の様子を思い出しながらでっち
上げる と 停滞しない

メタ発言が減ると没入感、障害対応さながらの緊張感も高まる

例) 「実装内容に不備がある」というシナリオのケース:

数分調査するふりをした後、「リリース後に500系のレスポンスが増
加した様子が確認されたので、ロールバックしてみませんか」と発言

障害対応フロー確認の効果

実際の障害対応中で、スキルが高まったと実感する機会がたくさん：

- 「誰かやってください」ではなく「**～～さん**にお願いします」と**明確な指示**を出せる
- 「**原因究明より復旧優先**しませんか」という声かけが自然とできる
- 集まった情報から冷静に状況を分析し判断できる

■ 検証用の環境での実践

検証環境が実際に壊された状況から、アラートなどの手がかりを元に影響範囲を特定しシステムを復旧する

これまで紹介した2つの演習で得た知識を**実践する反復によって真にチームに身に付く**

■ 検証用の環境での実践例

これらによってサービスがおかしくなった状態から原因を突き止め何とかする：

- コンテナイメージをレジストリから削除
- Security Groupの変更
- DBクラスタの削除

謎解き・脱出ゲームが好きなので普通に楽しい

■ 検証環境での実践の効果

本番で行う障害対応演習と何ら変わらず、ふり返りによって監視運用や障害対応をより良くできる

本当に起きたら困る、**ビジネス的に大きなインパクトがあるケースでのコミュニケーションの練習もできた**

まとめ

- チームの状況を踏まえ、目的意識を持って障害対応演習のメンバーを組みましょう
- 自信を持って手動オペレーションができるように練習機会を用意しましょう
- チームメンバーにうまく采配し、素早く復旧に持っていけるような指揮官を増やしましょう
- 新メンバーの知識を増やすには、障害対応演習の準備を一緒にやるのも効果的

宣伝

明日2024-05-24（金）

「Road to SRE NEXT @福岡」にはてなスタッフ

id:cohalz, id:cd_sioremonが登壇します！

<https://sre-lounge.connpass.com/event/314694/>

福岡の皆さん、現地でお会いしましょう

**ご清聴いただき
ありがとうございました！**