障害対応指揮の意思決定と情報共有における価値観

id:arthur-1 株式会社はてな

2024-11-13 Waroom Meetup #2 現場の声から学ぶインシデント対応

Arthurと申します

株式会社はてな Mackerel開発チーム

「オブザーバビリティの実現」チーム

テックリード

𝕏: @Arthur1__

昔のポケモンカードのデッキを落として
へコんでます

Mackerel作ってます

mackerel 製品 料金 事例 資料 サポート ブログ お知らせ

ログイン 無料トライアル お問い合わせ

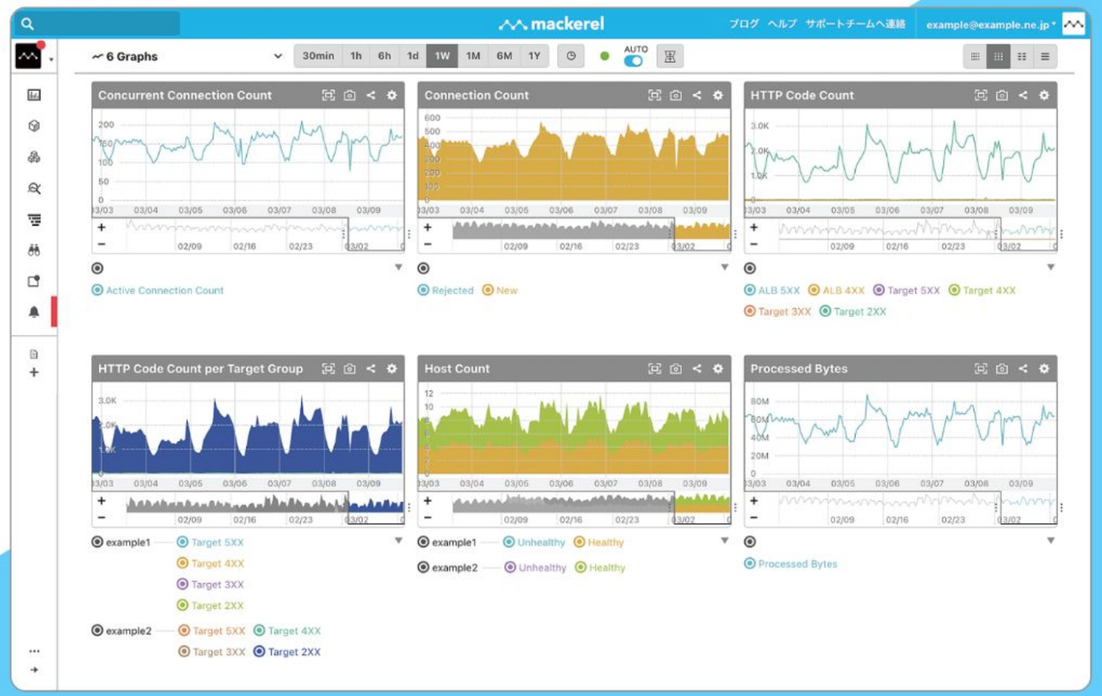
運用をイージーにする オブザーバビリティ

誰にも馴染む操作性のUIで、監視を育てるプラットフォーム。
未知の問題に立ち向かう力を開発者に。

無料トライアルをはじめる

「オブザーバビリティ」を強化して
生まれ変わったMackerelをご覧ください

詳しく見る



今日の話題

「障害対応指揮の意思決定と情報共有における価値観」

障害対応指揮官を担う機会が多い私が、障害対応において気をつけていることを色々話します

という内容ですが、ついお堅くタイトルをつけてしまいました

おしながき

- ケーススタディで学ぶ障害対応
- 障害対応中こそ気をつけたい情報共有

ケーススタディで学ぶ
障害対応

ケーススタディ

障害対応の事例※を眺めながら

指揮官としてどういう判断をするか

皆さんも一緒に考えてみましょう！

※実話かもしれないしフィクションかもしれません

Case: 1

リリース完了後にサービスの調子が悪くなったことを確認

このタイミングから、**IAMの権限不足のエラーログが出はじめている**のが分かった

ログから、どの権限が足りないかが分かっている

さあ、どうする？

Case: 2

前触れもなくサービスの調子が悪くなった

メトリックやトレース、ログを見ても原因が分からない

再デプロイしたら治りそうだと第六感が言っているが、

明確な根拠はない

さあ、どうする？

Case: 3

障害対応フォーメーションが組まれ、業務中のエンジニアがたくさん集まってきた

ところが、並行してできるオペレーションや調査の数がそこまでなく、**ただ見ているだけの人が多い**状況だ

さあ、どうする？

障害対応でやること

以下の2つを（多くの場合）並行して進める：

- 障害から復旧させ、サービスを利用可能にする
- 影響範囲の調査・ユーザーへの連絡

障害対応でやること

- 障害から復旧させ、サービスを利用可能にする
- 影響範囲の調査・ユーザーへの連絡

ユーザーができる限り早くサービスを利用できる状態にするのが**最重要課題**である

多くの場合、原因に辿り着いた上で対応が取られる

その場しのぎの暫定対応でも一旦は構わない

サービスが利用可能？

サービスが利用可能と一言で表したけれど

- 例えば重大なセキュリティの問題が発生した時、そのままサービスを提供するわけにはいかないから、結果としてサービスが利用不可能になる
- 機能は使えるけど、過去のデータが全部消えちゃってもOK？そんなこともない

何をMUSTとするか、現場だけで判断できないケースもある

障害対応でやること

- 障害から復旧させ、サービスを利用可能にする
- 影響範囲の調査・ユーザーへの連絡

影響範囲が分かっていなければ、復旧させようがないこともある

監視SaaSとしてはユーザーへ事象を素早く連絡できるかという点も大事にしている

[再掲] Case: 1

リリース完了後にサービスの調子が悪くなったことを確認

このタイミングから、**IAMの権限不足のエラーログが出はじめている**のが分かった

ログから、どの権限が足りないかが分かっている

さあ、どうする？

Case: 1 私はこう選択する

リリースのロールバックをしようとする

障害と権限不足のエラーの因果関係が不明なため

権限を直しても、他の原因でサービスが不安定なままかもしれない

ロールバックの利点

Binary Push※ が原因の障害では、よくある対応として
ロールバックがまず挙げられる

コンテナイメージを過去にビルド済みのものに差し替えることで、手間や時間をかけずにロールバックができる

※ cf.）https://sre.google/workbook/postmortem-analysis/

安易なロールバックに注意

安易に選択しがちだが、**リスクの評価**を必ずすること

- **一緒に巻き戻る機能やデータ**があっても大丈夫か？
- そもそもロールバックの手順は整っているか？

ロールバックを検討しているとき、実行を指示する前に**ロールバックして良いリリースかを確認**してもらっている

Pull Requestにラベルつけて「ロールバック可能」であることが一目で分かるようにできると素早く判断できそう

ロールバックを安全に行うために

デプロイと機能のリリースのタイミングを分けるフィーチャーフラグが有効

問題が起こった機能だけをロールバックできるため、
オペレーションによる影響範囲が小さくなる

Canary Releaseとテレメトリの分析・自動ロールバック
を組み合わせたProgressive Deliveryも手札に入れたい

[再掲] Case: 2

前触れもなくサービスの調子が悪くなった

メトリックやトレース、ログを見ても原因が分からない

再デプロイしたら治りそうだと第六感が言っているが、

明確な根拠はない

さあ、どうする？

Case: 2 私はこう選択する

とりあえず再度デプロイしてみる

他に打つ手がなく、再デプロイすることで新たに起こる問題もさほどないと想定されるため

リスクを評価した上で許容されるかを判断する

指揮官が何かを決めて動かなければ状況は変わらない

現実を理想に近づける場

勘や経験に頼った対応ではなく、テレメトリをドリルダウン探索して障害原因に辿り着きたいという理想はある

しかし、そんなことを嘆いていても、障害対応中はどうしようもない

障害対応が終わったら、後日ふりかえりを実施して、理想に近づくためのアクションを提案しよう

障害対応ふりかえり

Mackerel開発チームでは、障害対応後のふりかえりで、以下の観点で対策を整理している：

- **障害原因を確定**させるためのアクション
- 一時的な処置を**恒久対応**にするためのアクション
- **再発防止**するためのアクション
- **障害発生を素早く検知**するためのアクション
- **収束までの**（調査&復旧）**時間を短く**するためのアクション

[再掲] Case: 3

障害対応フォーメーションが組まれ、業務中のエンジニアがたくさん集まってきた

ところが、並行してできるオペレーションや調査の数がそこまでなく、**ただ見ているだけの人が多い**状況だ

さあ、どうする？

Case: 3 私はこう選択する

調査や復旧対応がアサインされていない人は解散して良いと伝える

障害対応フォーメーションにいる間、普段の仕事は止まってしまう

やるべき仕事が終わらなければ、やりたい仕事の優先度はどんどん下がっていき、障害が増え、負のループに

アサインの明示

指揮官は障害対応フォーメーションにいる人が何をやっているか、どういう状態なのかを定期的に把握しなければならない

何かの作業を依頼する際には、「何の目的で、何をするのか」を、特定の人を決めて明示的に指示する

宣言的に管理して、watchする状態を減らす

障害対応中こそ
気をつけたい情報共有

障害対応中のコミュニケーション

障害対応フォーメーションが組まれていて、**指揮官や作業実施者は同期的に会話**している

フォーメーションにいない人に向けて、**Slack上に現在の対応状況がポスト**されている

時には現場のエンジニアだけでは意思決定できず、**チームのディレクターや事業責任者に判断を仰ぐ**こともある

気をつけていること

情報の確からしさと詳細度を意識する

情報もたくさん飛び交い、異常事態で焦っているので、コミュニケーションのすれ違いが容易に起こり得る

情報の受信者（エンジニア？マネージャー？）が受け取りたい情報を意識する

出所が不明な情報を横に流して混乱を生まない

具体例: 障害対応フェーズprefix

Slack上での状況共有では、**障害対応のフェーズを見出しとしてつける**

詳細に踏み込みたい人だけ文章を読めば良いという構成

例：*【復旧完了】メトリックが平常時に戻ったのを確認し、14:00に復旧完了としました。現在はユーザーへの復旧連絡を進めています。*

障害対応フェーズprefixの現在

啓発しているけどチームにはまだ浸透していない

- どのフェーズでも一貫して【状況共有】という見出しがつけられているのを見る

なぜ大事かを伝えきれていないと同時に、仕組みが必要なのだろうと思う

まとめ

まとめ

障害対応は障害の迅速な復旧、具体的にはユーザーがサービスを利用できる状態にすることが最重要目標である

状況に応じてそれ以外の目標も大事にしなければならない

良い塩梅に多目的最適化できるように指揮官が決めて導く

異常事態だからこそ、整理された情報のやりとりが大事

作業者として関わるみなさんへ

障害対応指揮官だって完璧ではない一人の人間です

障害対応時に大切にしたい価値観の理解者として声を上げることで助けることができるかもしれません

そして、その価値観の重みづけは、プロダクトによって異なると思います。プロダクトを知りましょう

様々なポジション・ロールがあると思いますが、一つでも持ち帰れるものがあれば幸いです

おしまい