

inloco

Post-mortems: Como ajudam a tratar incidentes e solidificar conhecimentos

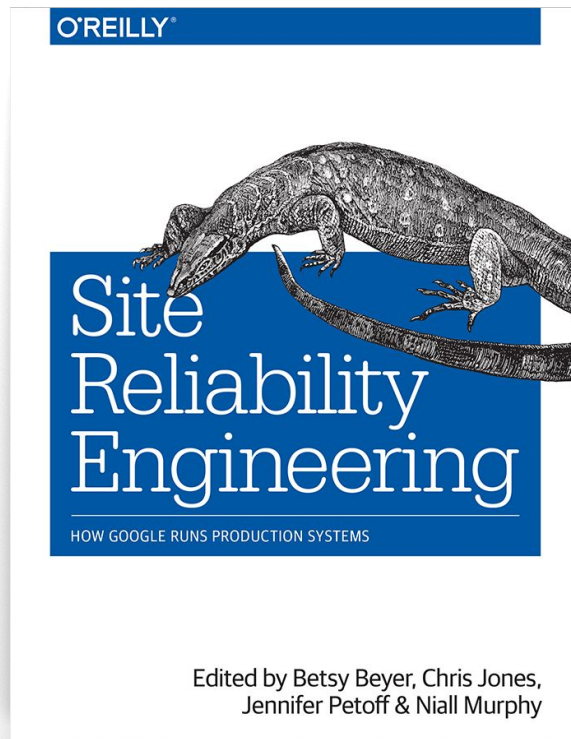
Renato Nascimento
Thiago Bastos

inloco

Perguntas no sli.do:
#DevOpsRec

De onde vem

“Fundamentally, it’s what happens when you ask a software engineer to design an operations function.”



Definindo confiabilidade

Confiabilidade é o oposto de falha.

Exemplo de SLO (*service level objective*):

Responder entre 99% a 99.95% de requests com código 2xx abaixo de 200ms.

Definindo confiabilidade

Falha é quando um sistema não atinge as condições de sucesso.

- Hardware/networking faults;
- Software faults;
- Stress/change induced failures;
- Human error.

Lidando com incidentes

Coisas quebram: é um fato.

- Precisamos de uma disciplina para lidar com incidentes da melhor forma, estressar possíveis long term solutions, e colocar esse conhecimento em uma língua comum, acessível a todo o corpo de engenheiros.

Cultura do Blameless

Remover a culpa de postmortems ajudam os engenheiros a escalar questões importantes sem receios.

É importante não estigmatizar pessoas ou times que façam postmortems frequentes.

Evitando erros humanos



[Day 4 & 5:
The Black Team](#)



[Journal Home Page](#)



[Day 5: The Maroon Team](#)

Apollo 8

Day 5: The Green Team

Corrected Transcript and Commentary Copyright © 2001-2017 by W. David Woods and Frank O'Brien. All rights reserved.

Last updated 2019-03-18

101:51:03 Borman: Houston. How do you read Apollo 8?

101:51:09 Collins: Apollo 8, Houston. Say again.

https://history.nasa.gov/afj/ap08fj/24day5_green.html

Evitando erros humanos

Create a beautiful site for your project with our [GitHub Pages generator](#).

Are you ABSOLUTELY sure?

Unexpected bad things will happen if you don't read this!

This action **CANNOT** be undone. This will delete the **jhosman/darkroom** repository, wiki, issues, and comments permanently.

Please type in the name of the repository to confirm.



Welcome to phpMyAdmin

Language

English

Log in

Username:

Password:

Como é a prática de Postmortem do Google

Postmortems são escritos apenas quando algum incidente significativo acontece

Os postmortems não procuram culpados pelos incidentes

Possuem uma cultura de incentivar postmortems

Como aplicamos a cultura de Postmortem

API que estava caindo semanalmente



Thiago Bastos 1 month ago

Só queria um fds de descanso...



Airton Sampaio 1 month ago

@bastos fofo, não se aflinja, ative uma hora extra pra ver o que está acontecendo e a gente vai te ajudar durante a semana para fazer uma solução definitiva

Como aplicamos a cultura de Postmortem

Linha do tempo dos eventos

Timeline

Start time: 2019-05-25 12:06 PM

- t0: received alert on slack because the HV API wasn't receiving requests;
- t0 + 20min: dev0 tried to make requests to the API with an api key commonly used for tests, the requests were successful;
- t0 + 35min: dev0 checked metrics and logs from all the other services to check if there were records in the AV architecture that would trigger requests to the API. There was data and it was in the normal rate;
- t0 + 36min: dev0 saw that all the records consumed by the Home Verification Service (the service that makes the request to the HV API) were being sent to the retry topic on kafka, meaning that the requests were not being successful (even after the tests being successful);
- t0 + 38min: dev0 tried to make requests to the API, from the pod of the Home Verification Service, with an api key commonly used for tests, the requests were successful;
- t0 + 58min: dev0 made a new deploy printing the result of the requests to the API and the status codes. All the requests were receiving **401 - Unauthorized**. Dev0 tried to make the tests that were successful with the same api key used in the Home Verification Service and received **401 - Unauthorized**;
- t0 + 1h 10min: dev1 gave access to dev0 for the Cassandra used in the API for authentication, two out of three nodes were dead;
- t0 + 1h 28min: dev0 restarted the Cassandra from one of the dead nodes and the node was up and normal again, the Home Verification Service started to receive successful responses from the API;
- t0 + 1h 39min: dev0 restarted the Cassandra from the last dead node and it was up and normal again, with all the requests from the Home Verification Service being successful.

End time: 2019-05-25 1:45 PM

Total time: 1h 39min

Como aplicamos a cultura de Postmortem

Pontos de ação

Action Items

- Review the max open files rule in Cassandra @ [Thiago Bastos](#);
- Search if there are issues in this version of Cassandra that causes some kind of leak in open files and double check Cassandra configuration @ [Thiago Bastos](#);
- Create an easy and fast solution to give access to EC2 machines @ [Marcelo Anabuki](#);
- Review alerts for the database @ [Tatiana Viana Padrão](#);
- Add metric of up nodes from a Cassandra cluster and add alert @ [Tatiana Viana Padrão](#);
- Add metrics for the possible status codes, returned by the HV API, in the Home Verification Service @ [Tatiana Viana Padrão](#);
- Export metrics of open files @ [Thiago Bastos](#).

Como a empresa se beneficia disso

Revisão do postmortem

Postmortem do mês

Grupo de discussão

Clube do "livro" de postmortem

Treinamento de novos SREs

An aerial photograph of a city intersection. In the center, a large white text 'Obrigado' is superimposed over the scene. The intersection features a roundabout with a colorful geometric design in the center. Surrounding the roundabout are several buildings, including a large one with a red roof and white domes, and another with a blue and white facade. A large, modern building with a glass facade is visible on the left. The street is paved with asphalt and has white lane markings. Several cars are visible on the roads. The overall scene is brightly lit, suggesting a sunny day.

Obrigado