

# How to design Google Docs (Episode 4)



ALEX XU

APR 23, 2022



96



3



Share



In this newsletter, we will talk about the following:

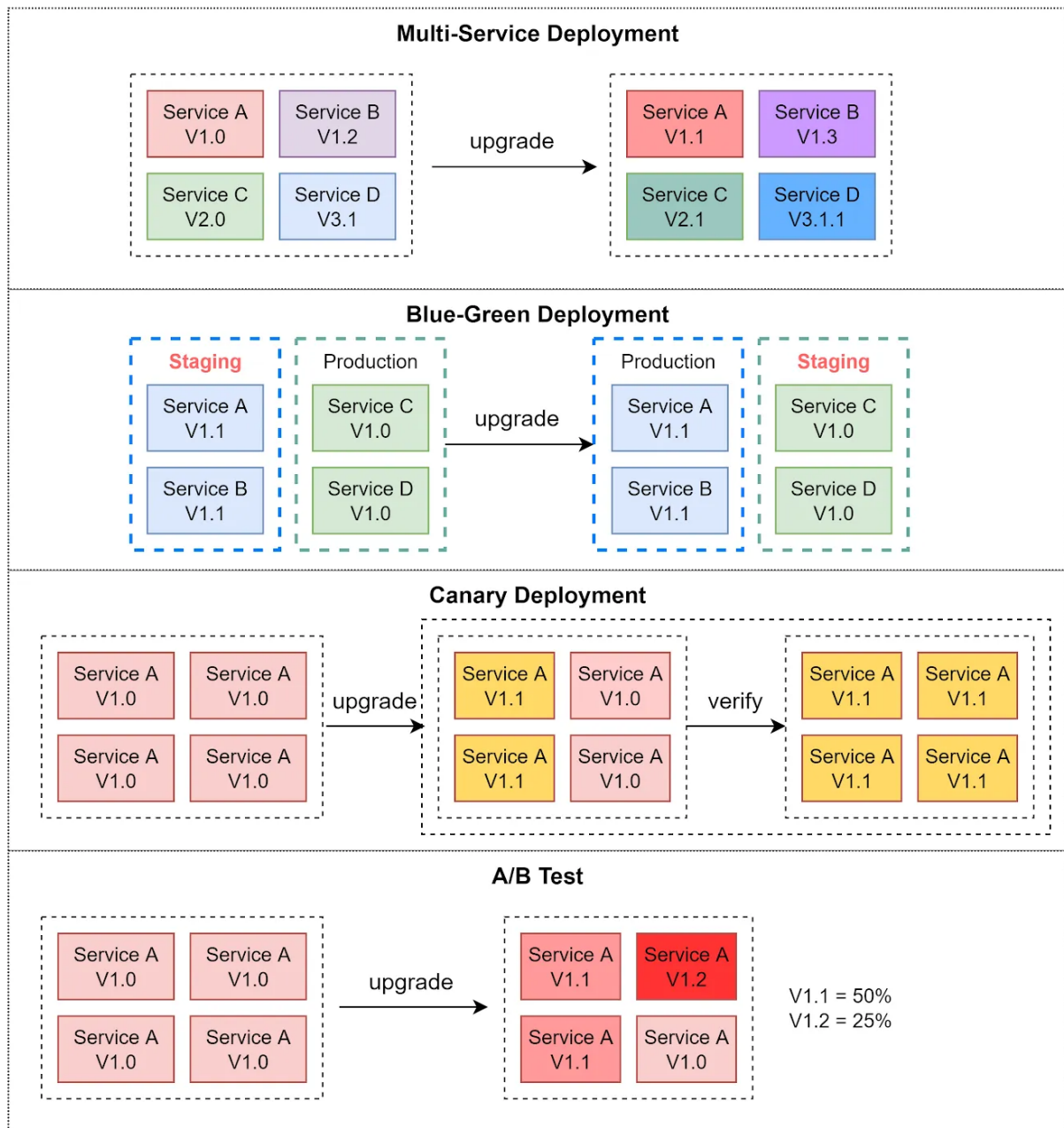
- How to deploy services
- Google doc design
- Signed Book GIVEAWAY
- Software Architecture Trends Report

## How to deploy services

Deploying or upgrading services is risky. In this post, we explore risk mitigation strategies.

The diagram below illustrates the common ones.

# How to Deploy Services?



## Multi-Service Deployment

In this model, we deploy new changes to multiple services simultaneously. This approach is easy to implement. But since all the services are upgraded at the same time, it is hard to manage and test dependencies. It's also hard to rollback safely.

## Blue-Green Deployment

With blue-green deployment, we have two identical environments: one is staging (blue) and the other is production (green). The staging environment is one version ahead of production. Once testing is done in the staging environment, user traffic is switched to the staging environment, and the staging becomes the production. This deployment strategy is simple to perform rollback, but having two identical production quality environments could be expensive.

## Canary Deployment

A canary deployment upgrades services gradually, each time to a subset of users. It is cheaper than blue-green deployment and easy to perform rollback. However, since there is no staging environment, we have to test on production. This process is more complicated because we need to monitor the canary while gradually migrating more and more users away from the old version.

## A/B Test

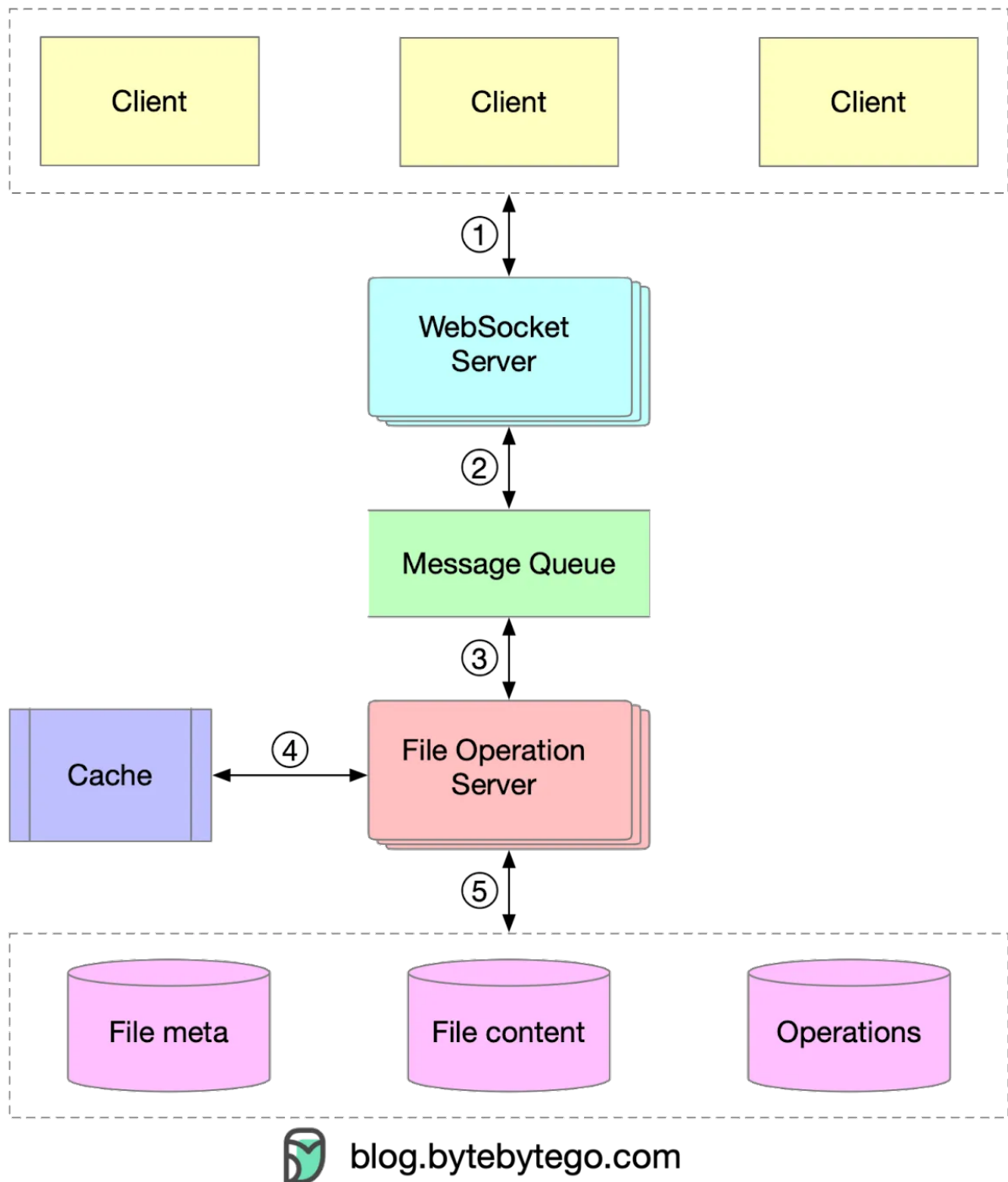
In the A/B test, different versions of services run in production simultaneously. Each version runs an “experiment” for a subset of users. A/B test is a cheap method to test new features in production. We need to control the deployment process in case some features are pushed to users by accident.

Over to you - Which deployment strategy have you used? Did you witness any deployment-related outages in production and why did they happen?

# Google doc design

One picture is worth more than a thousand words. In this post, we will take a look at how to design Google Docs

# How to Design Google Doc?



① Clients send document editing operations to the WebSocket Server.

② The real-time communication is handled by the WebSocket Server.

③ Documents operations are persisted in the Message Queue.

- 4 The File Operation Server consumes operations produced by clients and generates transformed operations using collaboration algorithms.
- 5 Three types of data are stored: file metadata, file content, and operations.

One of the biggest challenges is real-time conflict resolution. Common algorithms include:

- ◆ Operational transformation (OT)
- ◆ Differential Synchronization (DS)
- ◆ Conflict-free replicated data type (CRDT)

Google Doc uses OT according to its Wikipedia page and CRDT is an active area of research for real-time concurrent editing.

Over to you - Have you encountered any issues while using Google Docs? If so, what do you think might have caused the issue?

## Signed Book GIVEAWAY

Signed Book GIVEAWAY!

My co-author Sahn Lam and I will be giving away 10 signed copies of the System Design Interview book (Volume 2).

Shipping is on us. Yes, we do international shipping as long as USPS supports it.

[Click here](#) for more details.

## Software Architecture Trends Report

Interesting read: Software Architecture and Design InfoQ Trends Report — April 2022

Key takeaways: "Data plus architecture" is the idea that, more frequently, software architecture is adapting to consider data. This holistically includes data quality, data pipelines, and traceability to understand how data influenced decisions and AI models.

Innovative software architecture is facilitating data quality the way we've improved code quality. Catching bad data early is as important as catching bugs early.

The practice of software architecture does not belong solely to people with the job title of architect. Every engineer can actively participate in the architecture, and architects should help facilitate that process.

One positive benefit of the pandemic and the shift to remote and hybrid work is increased asynchronous communication, which can manifest as Architecture Decision Records (ADRs).

Software architects are adapting their feedback loops, which can be challenging when dealing with colleagues across many time zones or other remote work constraints. Good architects are learning from distributed working how to design better distributed systems.“ [1]

Over to you - what are some of the trends you see in software development in 2022?

[1] [Software Architecture and Design InfoQ Trends Report — April 2022](#)



96 Likes

### 3 Comments



Write a comment...



Eric Zhan May 2, 2022

What is the data flow when client needs to update its document according to the file change on cloud?

♡ LIKE (1) 💬 REPLY ↗ SHARE

...



Akhil Apr 24, 2022

This is awesome content. Thanks Alex for all your hard work. I just ordered my copy of part 2.

♡ LIKE 💬 REPLY ↗ SHARE

...

1 reply by Alex Xu

1 more comment...

---

© 2023 ByteByteGo · [Privacy](#) · [Terms](#) · [Collection notice](#)  
[Substack](#) is the home for great writing