

What is Data Reliability Engineering? and Why it is Crucial to any Data Organization?



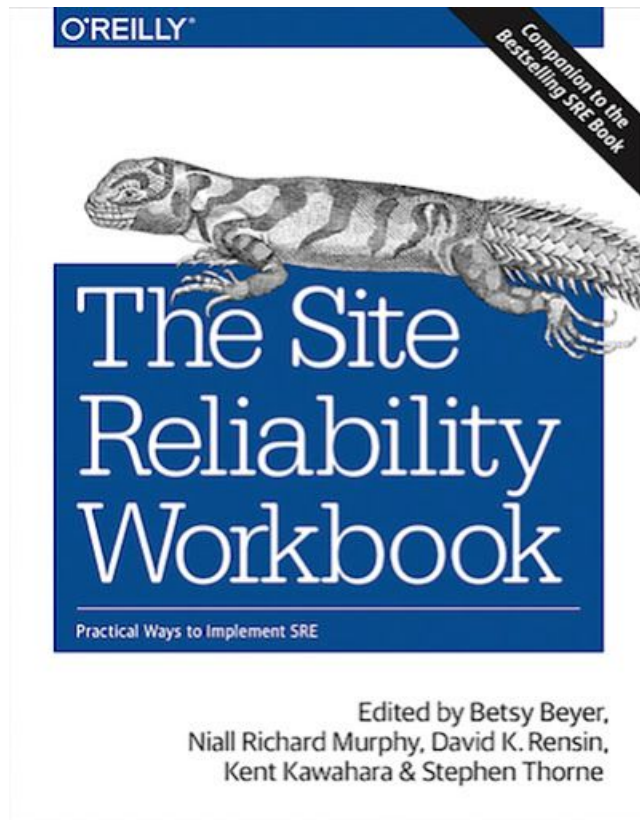
Miriah Peterson

Member Technical Staff

Tailscale

What is Data Reliability Engineering?

- Reliable Data
- Data Downtime
- Data Level Metrics
- Data Observability



Reliable Data

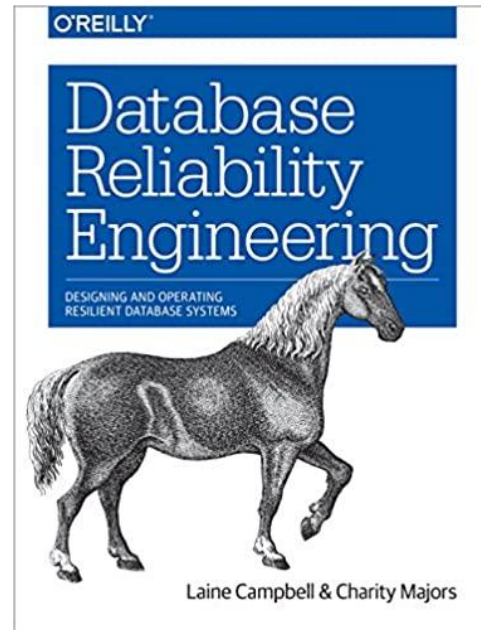
- Data Accuracy
- Data Freshness
- Data Latency
- Data Availability
- Data Duplication
- Missing Data



Reliable Data

Systems without failures, although robust, become brittle and fragile. When failures occur, it is more likely that the teams responding will be unprepared, and this could dramatically increase the impact of the incident.

- *Database Reliability Engineering, Laine Campbell & Charity Majors*



Data Downtime

Data downtime refers to periods of time when your data is partial, erroneous, missing or otherwise inaccurate.

[The Rise of Data Down Time](#), Barr Moses



Data Level Metrics

- Latency
- Traffic
- Errors
- Saturation

Golden Signals



Data Observability

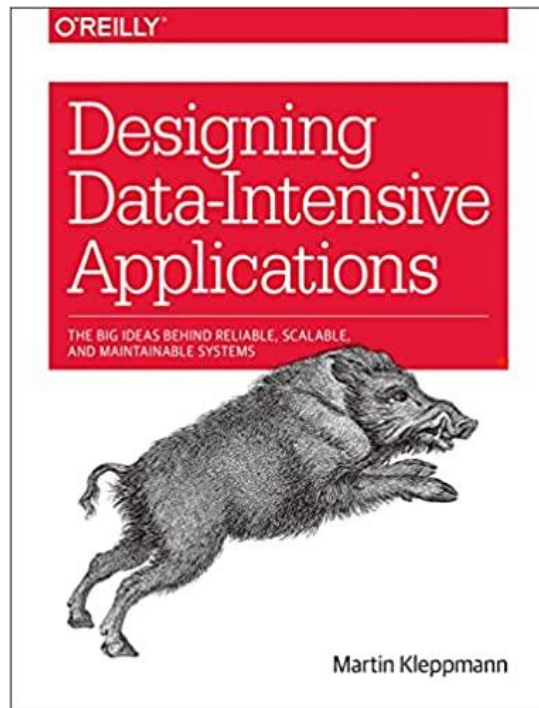
- SLAs
 - Revenue Dashboard has data with minimal freshness of 1 Day
- SLOs
 - Data Pipeline extracts data from CRM and completes transforms and analysis once a day
 - Can be manually triggered if there is an error
- SLIs
 - Errors in pipeline reported
 - Time out reported
 - Alerts of needed to be run manually

Additional Tasks

- Talk to legal, security, and compliance teams
- Audit databases
- Trace calls
- Get metrics around use and latency

Why is Reliability Engineering Crucial for Organizations

- Software Data Stack
- Enterprise Data Stack
- Modern Data stack



Software Data Stack

What is it?

Datastores for customer-facing applications

Who manages reliability?

Site Reliability Engineers

Enterprise Data Stack

What is it?

Data infrastructure/platforms

Who manages reliability?

Handled by DBAs or Database Reliability Engineers

Modern Data Stack

What is it?

Data pipelines for analytics and machine learning

ETL pipelines that perform platform critical transformations

Data Gateways that allow for many consumers of Golden Datasets

Who manages it?

Modern Data Stack

DRE skillset:

- Data ingress and egress
- Data transformations
- Data storage
- Data governance
- Data analysis

Modern Data Stack

DRE skillset:

- Automation
- Monitoring Observability
- Architecture

Lessons and Takeaways

- Data Reliability is applicable to all data centric services
- As Data democratizes across the stack reliability is crucial to organization success
- A DRE skill set is specialized and different than any other reliability engineer.

Thanks!

- Twitter: [@captainnobody1](#)
- Twitch: [@soypete01](#)
- LinkedIn: [Miriah Peterson](#)