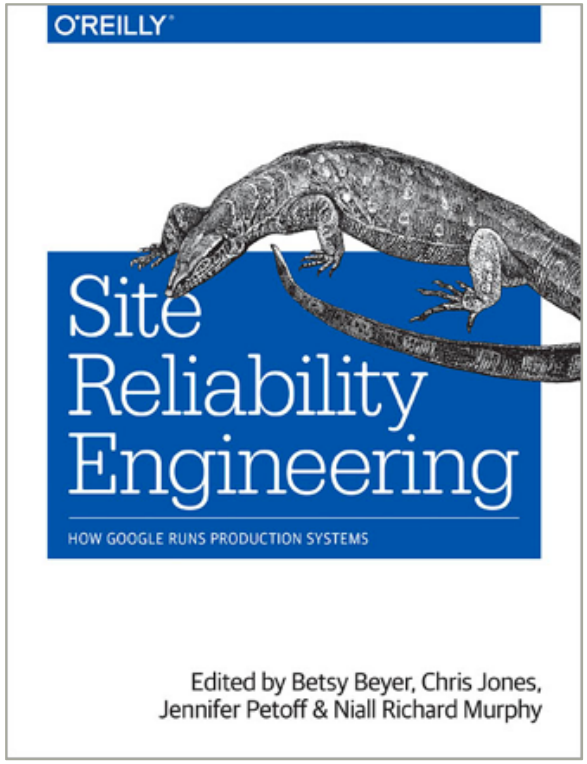
Rethinking Site Reliability Engineering for ITSM.

Jon Stevens-Hall

Principal Product Manager (BMC Software) and ITIL® 4 Contributing Author

SDI “New Ways of Working” virtual event: 7th May 2020

A brief introduction to SRE



- First published 2016
- Free online: https://landing.google.com/sre/books/

"SRE is what happens when you ask a software engineer to design an operations team"

- Some SREs at Google are software engineers hired in the standard way
- Others were *close* to being recruited as a software engineer, and brought other infra skills

- SREs help software developers run their services but also expect accountability from them.

Me

bmc | Run and Reinvent

- ITSM veteran – 23 years
- Principal Product Manager with BMC software, for our enterprise ITSM toolset, BMC Helix ITSM
- ITIL® 4 contributing author

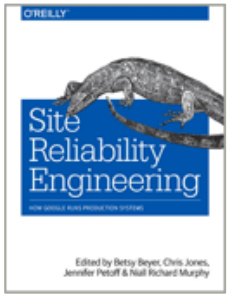
- Strong focus on DevOps and new ways of working

- Asked by SDI to talk about SRE today

Twitter: @JonStevensHall

- https://www.bmc.com/blogs/author/jonhall/
- https://medium.com/@jonhall_

Subjects covered in the SRE handbook



Risk

Eliminating Toil

Automation

Service Level Objectives

Handling Incidents

Release Engineering

Monitoring

Alerting

Communication

Distributed Systems

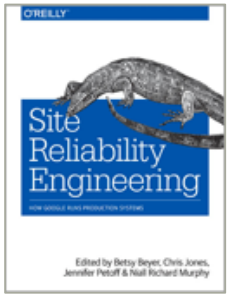
On-Call

Human Welfare

Postmortem

Scheduling

Defining Toil



- SRE teams can only spend a maximum of 50% on work considered to be “Toil”.
- Google’s SRE handbook describes “toil” as work which is:

Manual

Repetitive

Automatable

Tactical

Devoid of enduring value

Linearly scaling

Service Level Objective

SLOs are not SLAs. Not all Google services have SLAs, but some do.

SLOs define appropriate technical bounds for service performance.

“99.9% of RPC calls will complete in less than 100 ms"

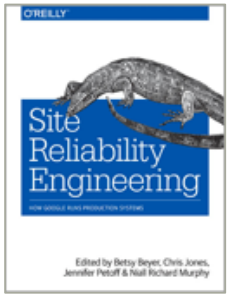
$$\text{Availability} = \frac{\text{Successful requests}}{\text{Total requests}}$$

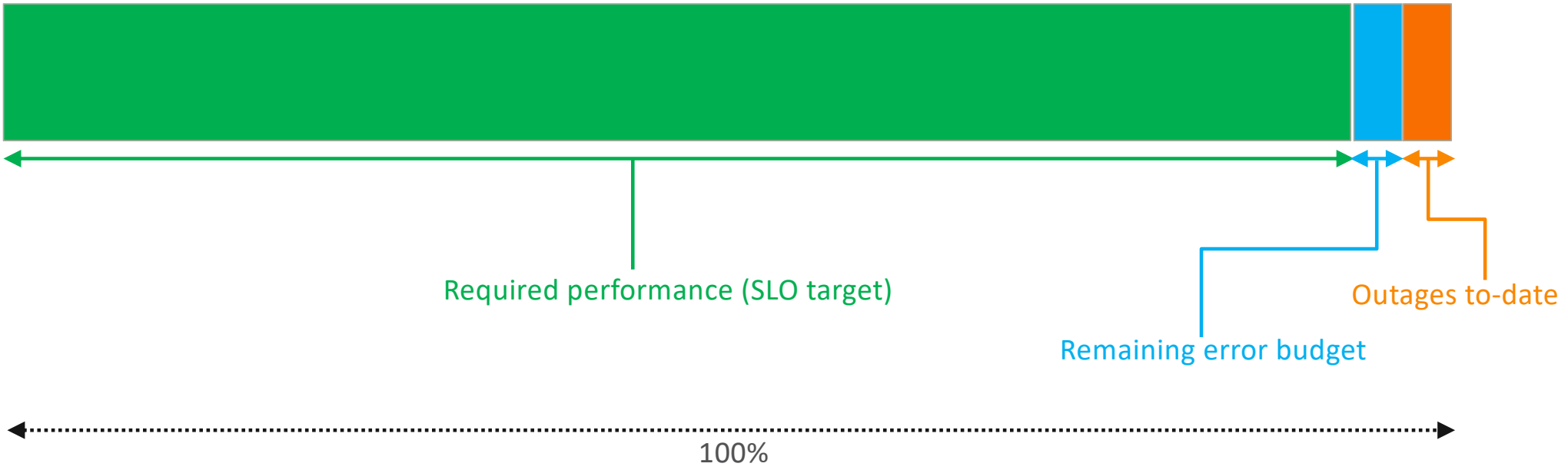
“The service will be available 99.99% of the time”

$$\text{Availability} = \frac{\text{Uptime}}{\text{Uptime} + \text{Downtime}}$$

Error Budget

Defines how unreliable the service is allowed to be, within a given time period.

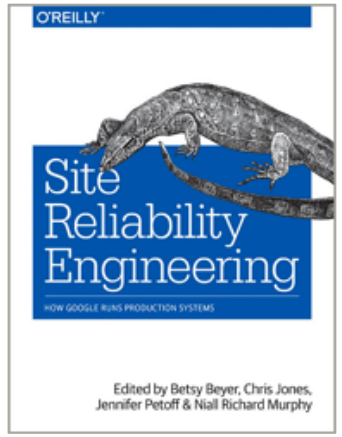




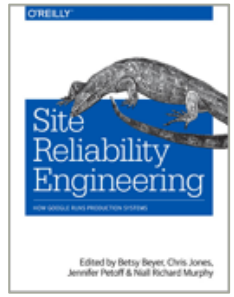
Outages are okay, within error budget tolerances

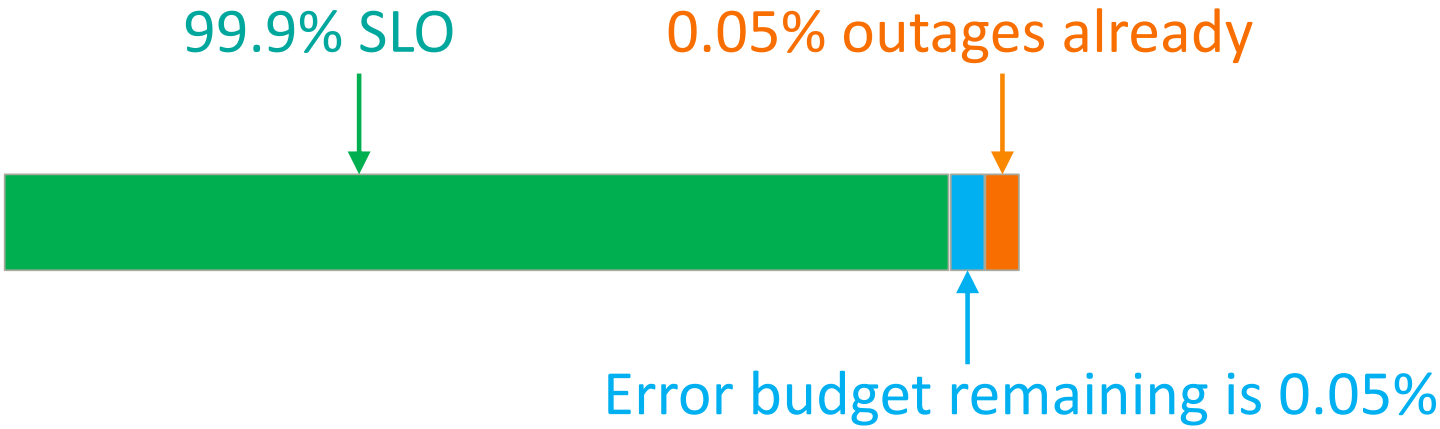
"An outage is **no longer a bad thing**.

It is an **expected** part of the process of innovation, and an occurrence that development and SRE teams **manage rather than fear**"



Error Budget





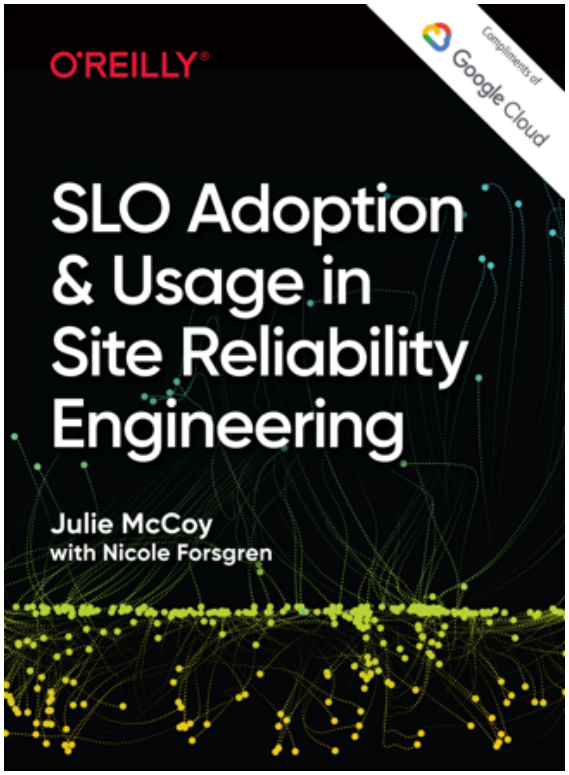
Hence an additional outage which consumes 0.02% of the quarter's availability would cost 40% of the remaining error budget.

Error Budget shapes ways of working

Balancing innovation and reliability

- Large budget: developers can take more risks
- Budget nearly drained: developers tend to slow down
- Budget used up: releases temporarily halted

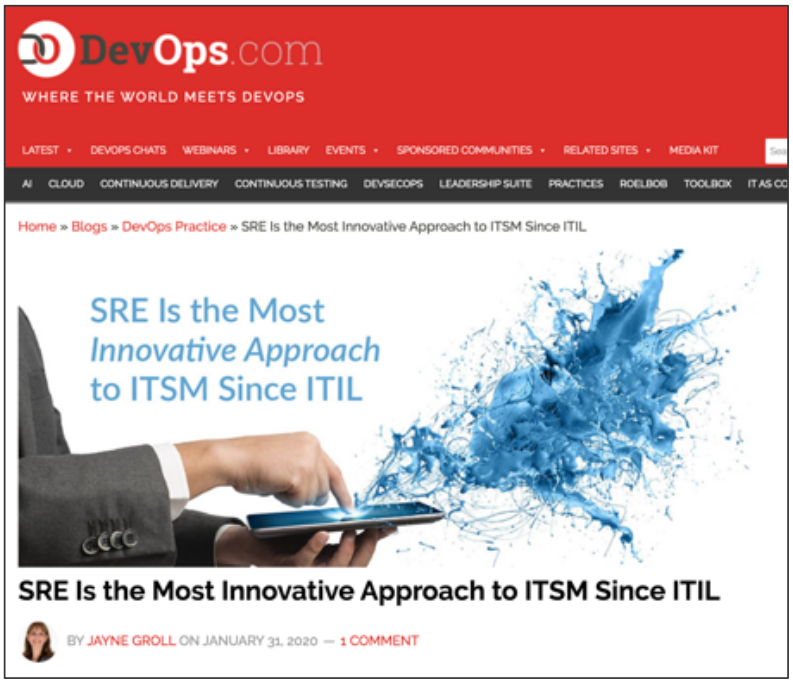
How does SRE align to ITSM as we know it?



"Site Reliability Engineering (SRE) is an ***emerging IT service management (ITSM) framework"***

(First line of the report, published 1st April 2020)

Will SRE *replace* ITSM as we know it?

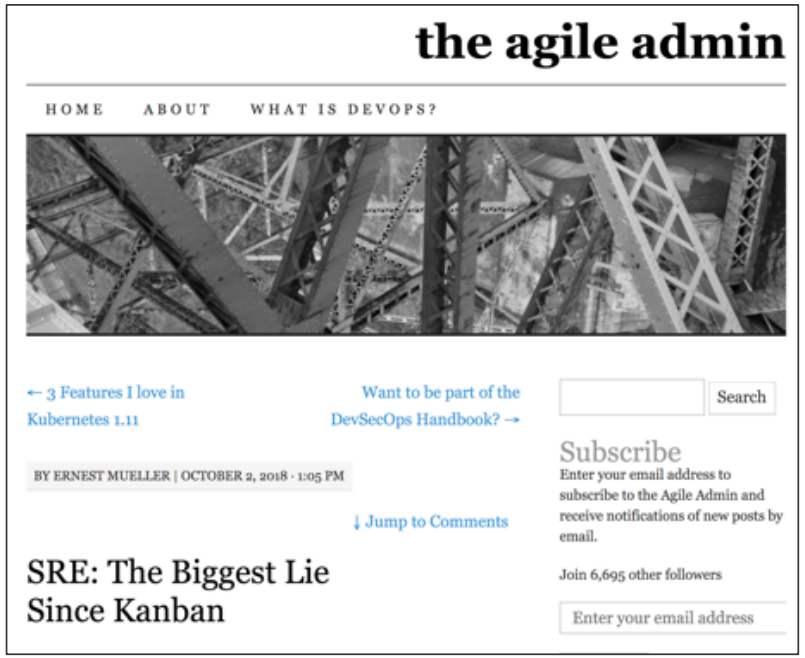


"As a service management alternative, SRE **updates traditional ITSM activities** with innovative and self-organizing concepts such as…

- Management to service level objectives
- Error budgets
- Toil reduction
- Release engineering
- Monitoring/observability
- Embracing risk"

Jayne Groll, DevOps.com, 2020

A counterpoint: Most of us are not Google.

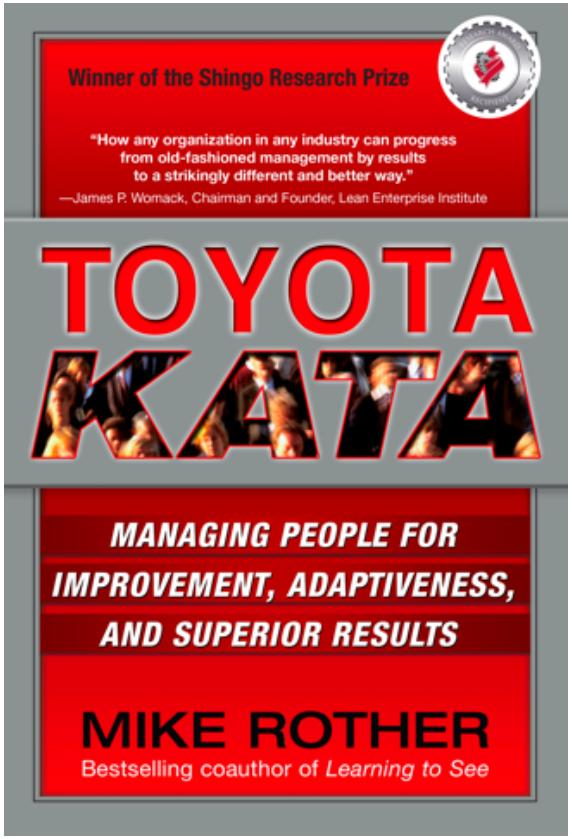


"If you think spinning up an ops team and calling it SRE is 'an implementation of DevOps' *you've swallowed the worst poison pill the DevOps talk circuit can deal to you."*

- Unless you are Google scale, a separate production operations team is inefficient
- DevOps showed that "throw over the wall from Dev to Ops" was wrong
- Throwing over the wall from Dev to SRE doesn't improve that

The Agile Admin – Blog (2018).

One size doesn’t fit all



“Toyota opens its factory doors to us again and again, but I imagine Toyota’s leaders may also be shaking their heads and thinking, ‘Sure, come have a look. But why are you so interested in the solutions we develop for our specific problems?

Why do you never study how we go about developing those solutions?’”

Mike Rother – Toyota Kata (2009)

What's in a name? Rethinking SRE for ITSM

- **"Site Reliability Engineering"…**
  - In many (most?) cases, IT services aren't "sites"
  - A lot of ITSM work isn't software engineering
  - Reliability is a broad term

SRE is already an established profession

Site Reliability Engineer (SRE)

F FULLSTACK RECRUITMENT LIMITED
London, UK

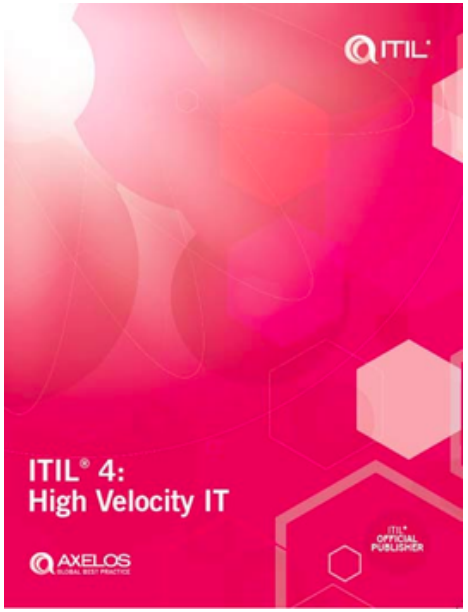
Essential requirements:
- Expert Linux and Networking experience (Physical and Cloud)
- Infrastructure as Code (YAML, Terraform)
- Cloud Infrastructure Orchestration (AWS/EKS, Docker, Kubernetes)
- Monitoring (Prometheus, Splunk)
- Cloud Database Management (Aurora DB, Cassandra)
- Queue Management (SQS, Kafka)
- Agile product development experience
- Good communication skills
- Collaboration in a diverse and distributed team

Desirable skills/experience:
- Other languages: Java, Kotlin, JavaScript, Python, etc
- Redis, SQL & NoSQL databases, Elastic Search
- Trading or Sporting experience

Job advert for ”disruptive Fintech”, retrieved May 2020

ITIL® 4 maps SRE principles to 14 practices

(but at a very high level)



- Availability management
- Capacity and performance management
- Change enablement
- Incident management
- Infrastructure and platform management
- Monitoring and event management
- Problem management
- Service design
- Software development and management
- Deployment management
- Organizational change management
- Release management
- Service configuration management
- Service validation and testing

Example of an “SRE-like” change for IT support

Before:

- **Current service desk performance: 98% call pickup, on a target of 90%**
- Agents fully utilised on calls

Applying SRE thinking:

- Schedule a proportion of each agent’s time to be *away* from calls
- Use that time for proactive reduction of toil *(e.g. building self service offerings, developing knowledge content)*
- The previous 8% over-performance in call pickup is spent as **“error budget”**

SRE applied to Change Enablement

Before:

- Level of change approval determined by perceived risk of change type
- Pressure from DevOps to remove approval barriers (particularly CAB)

Applying SRE thinking:

- Dynamically assess approval level based on confidence level in team
  - Implementation success rate of previous changes
  - Issues caused by previous changes
  - Adherence to guardrails in the DevOps toolchain
- **Minimal intervention, applied more smartly**

Summary of issues

- Site Reliability Engineering is a bad name for many aspects of ITSM.
- SRE is already defined (and recruited) as a specialist operations developer role.
- Some key, novel themes have great potential but need redefinition for much of ITSM.

Proposal

- “Rebrand” the SRE concept for ITSM: New name, new thinking.
- Identify new ways to apply key novel principles such as reliability and toil reduction.
- Invest in ITSM initiatives such as KCS which share good principles with SRE.

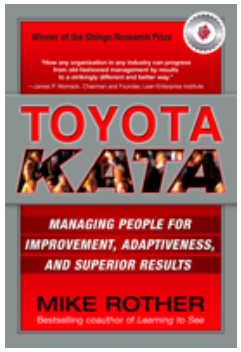
Referenced materials and further reading

https://theagileadmin.com/2018/10/02/sre-the-biggest-lie-since-kanban/

@ernestmueller

https://devops.com/sre-is-the-most-innovative-approach-to-itsm-since-itil/

@JayneGroll



http://www-personal.umich.edu/~mrother/Homepage.html

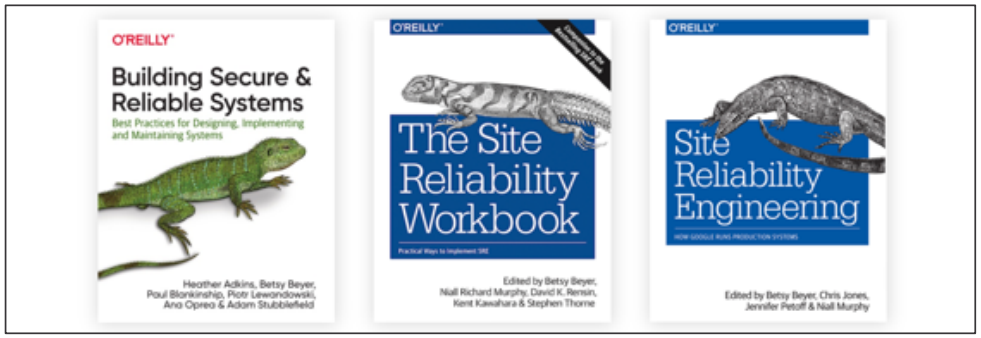
@RealMikeRother

https://www.bmc.com/blogs/sre-vs-itops/

@DevDroid9

https://medium.com/@JonHall_/what-can-site-reliability-engineers-teach-service-desks-about-toil-e1df38ca284e

@jonstevenshall



https://landing.google.com/sre/books/

@srebook


bmc
Run and Reinvent