

~ A ~
THEORY
~~ and ~~
PRACTICE
~~~ of ~~~
SERVICE LEVEL OBJECTIVES



Jamie Wilkinson
SRECon Asia, June 2018





★ Star this Talk 4 Stars

Published in Technology

Stats 374 Views

Share

Twitter, Facebook

Embed

Direct Link

Download PDF



share

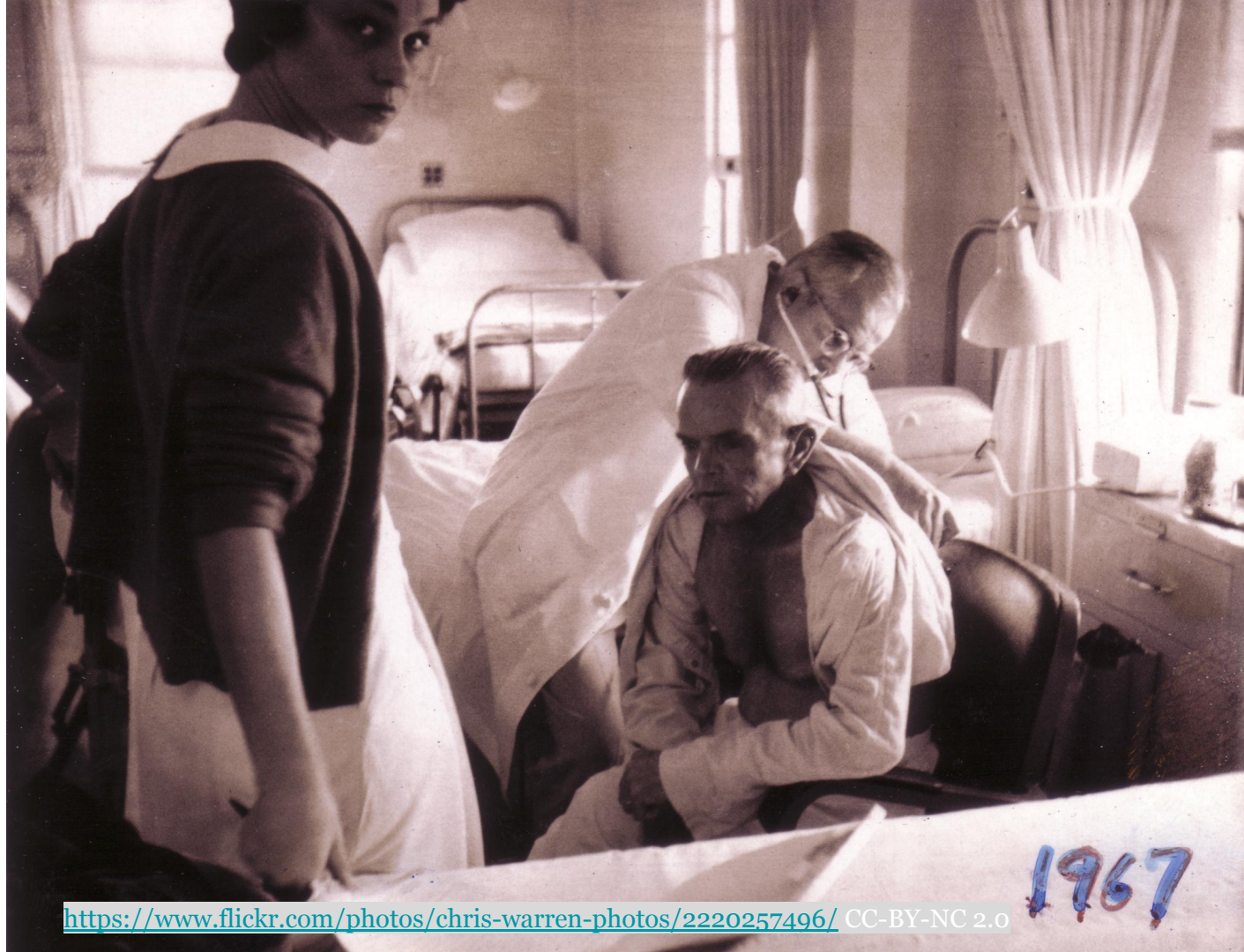
Martyrs On Film: learning to hate the #oncallselfie

by Alice Goldfuss

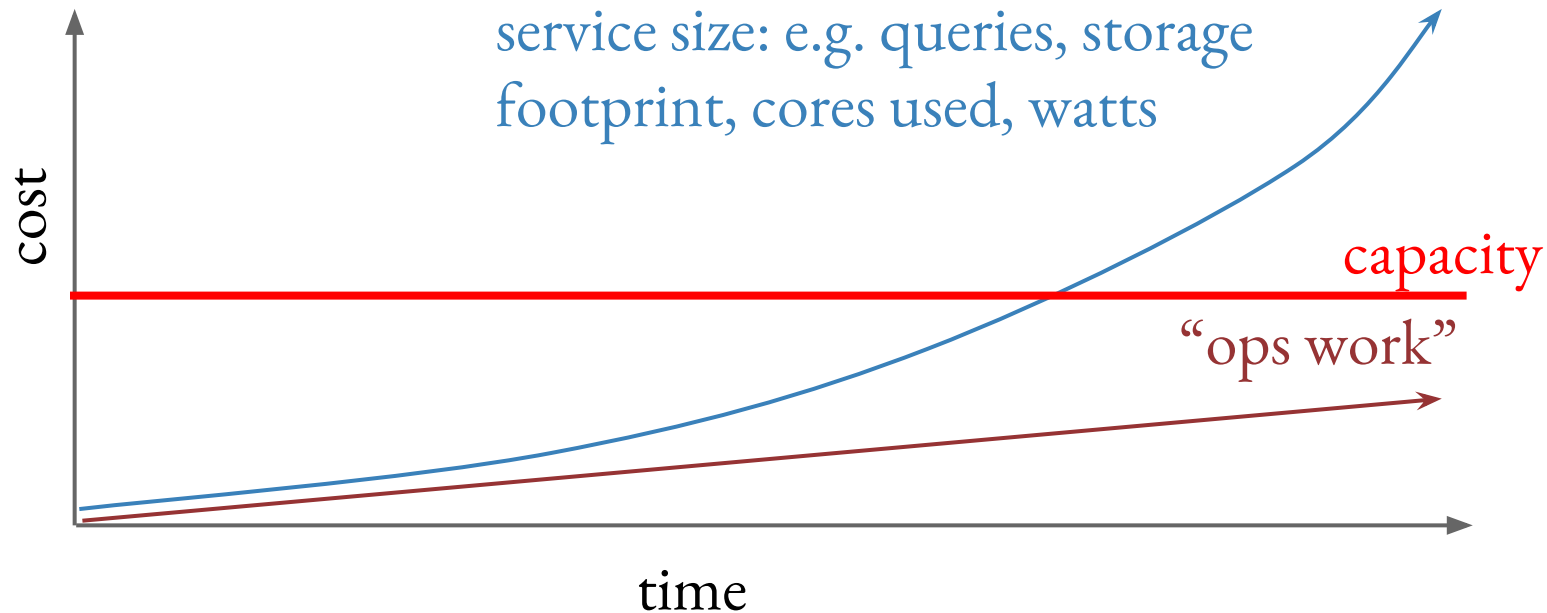
Published May 22, 2017 in Technology

Alice Goldfuss, Monitorama 2017, Used with permission

Symptom Based Alerting



Why *does* $X \forall X \in \{\text{Ops}\}$ suck?



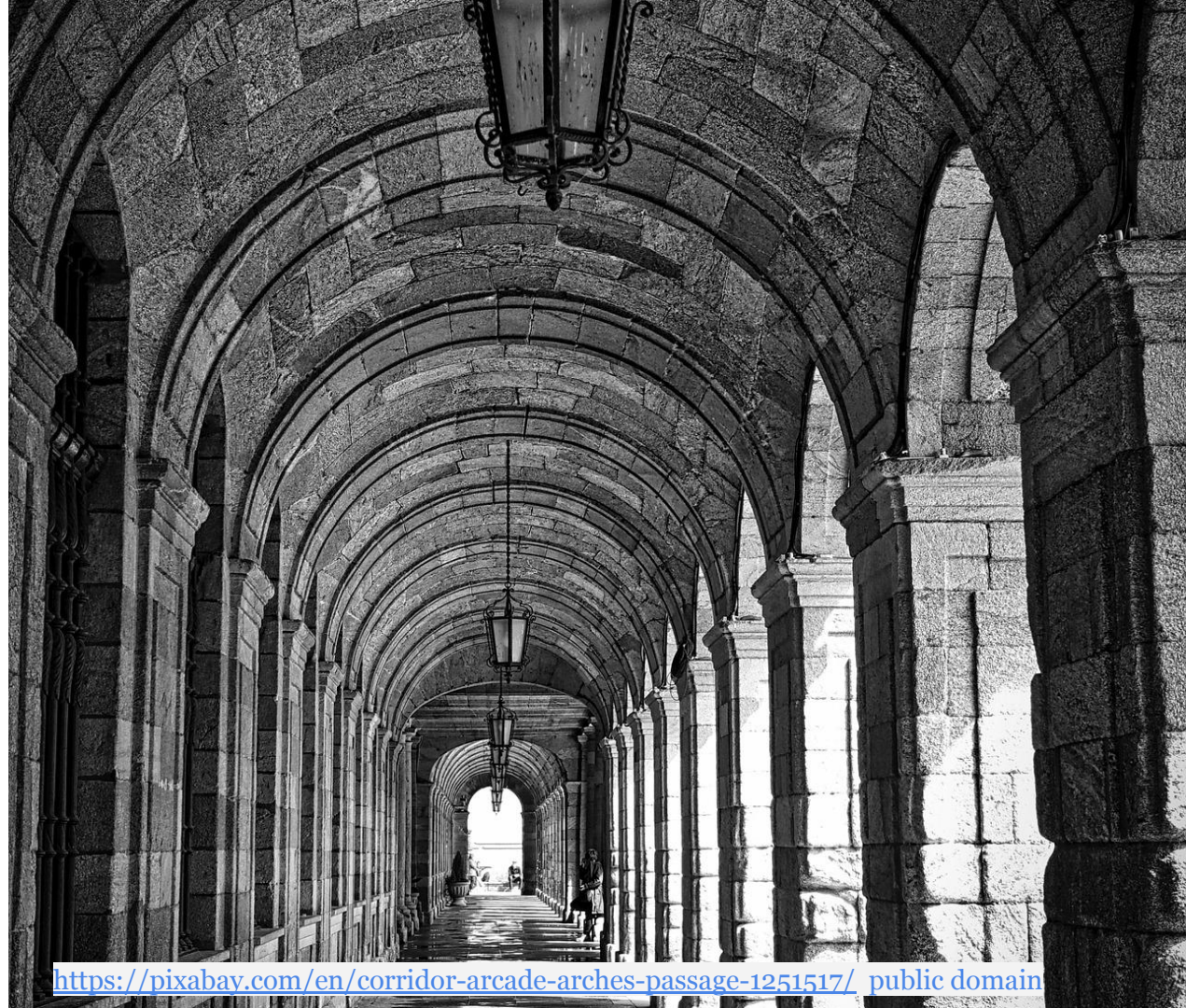
the cost of maintenance must scale sublinearly with the growth of the service



IMAGE CREDITS from Taiyo no Yusha Faibado (The Brave Fighter of Sun Fighbird) Episode 3 - All Members In! Space Police! Screenshot from: Hirano, Yasushi (Writer), & Yatabe, Katsuyoshi. (1991) All Members In! Space Police [Series Episode] In S. Imai, Y. Honna, T. Takayuki, Taiyo no Yusha Faibado [Brave Fighter of Sun Fighbird]. Tokyo, Japan. Takara and Sunrise

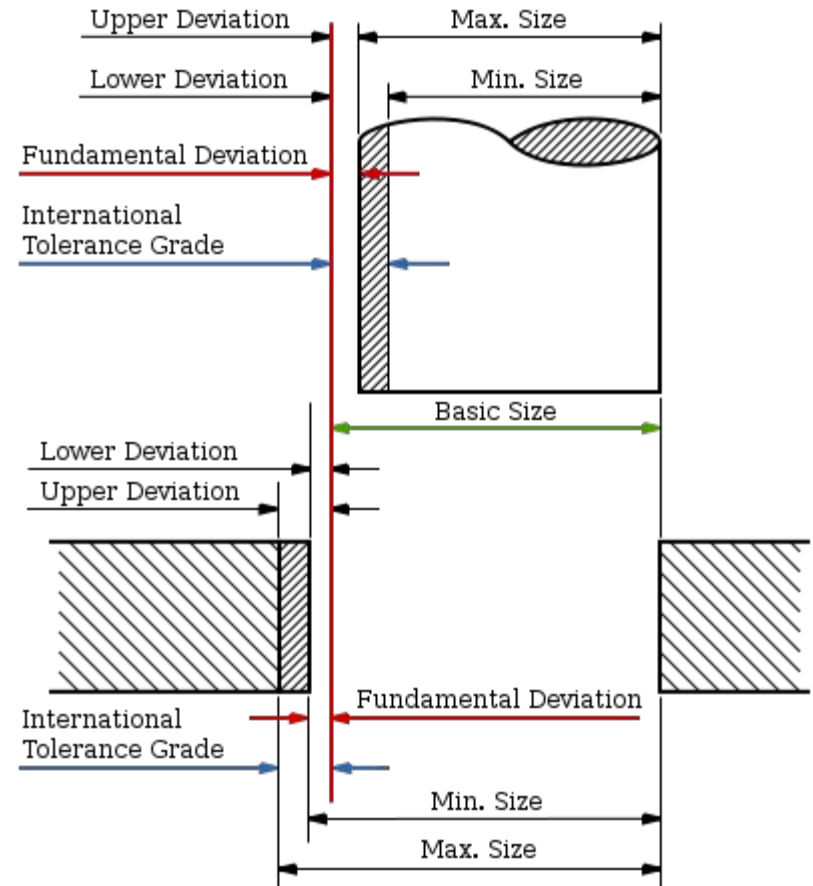
ATtribution CREDITS
Phillipines Department of Science and Technology
<https://www.facebook.com/DOSTph/photos/a.1778676488821528.1073741834.1124649117557605/1786984204657423/?type=3&theater>

**What makes this a
symptom?**





Engineering Tolerance





99%

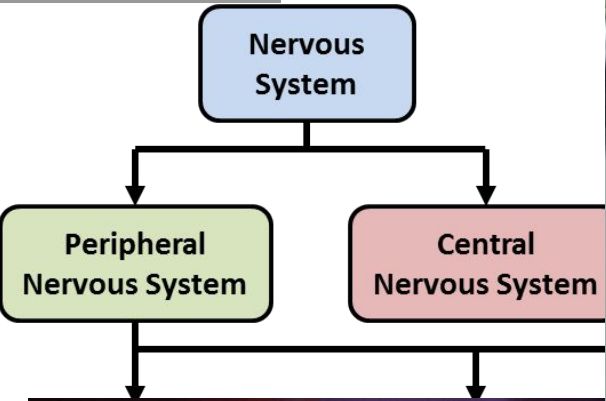
99.5%

99.9%

100%

Availability “Tolerance”

By Jenna Fair [CC BY-SA 3.0
(<https://creativecommons.org/licenses/by-sa/>
Wikimedia Commons



By G.E. Ulrich, USGS. Cropping by Hike395 (ta
(USGS) [Public domain], via Wikimedia Comm



Queensland Newspapers Pty Ltd [Public
media Commons

SLAs, SLOs, SLIs

- **SLI** → **Indicator**: a measurement
 - distribution of response time over 10 minutes
 - response error ratios over 10 minutes
- **SLO** → **Objective**: a goal
 - 99.9th percentile response latency below 5ms
 - lower than 1% rate of errors
- **SLA** → **Agreement**: economic incentives
 - or we get paged

“As a mechanical engineer in an R&D lab I frequently ask myself, what is a reasonable tolerance to set on this part?”

Does your service have an SLO?

The answer may surprise you!

Posted on May 31, 2018, at 1:53 p.m.



Jamie Wilkinson

Site Reliability Engineer

A **symptom** is anything that can be measured by the **SLO**.

A **symptom-based alert** is an alert when the **SLO is in danger** of being missed.

For availability SLAs we often talk about *system uptime*:

$$\textit{availability} = \frac{\textit{uptime}}{(\textit{uptime} + \textit{downtime})}$$

How do you measure uptime of a distributed system?



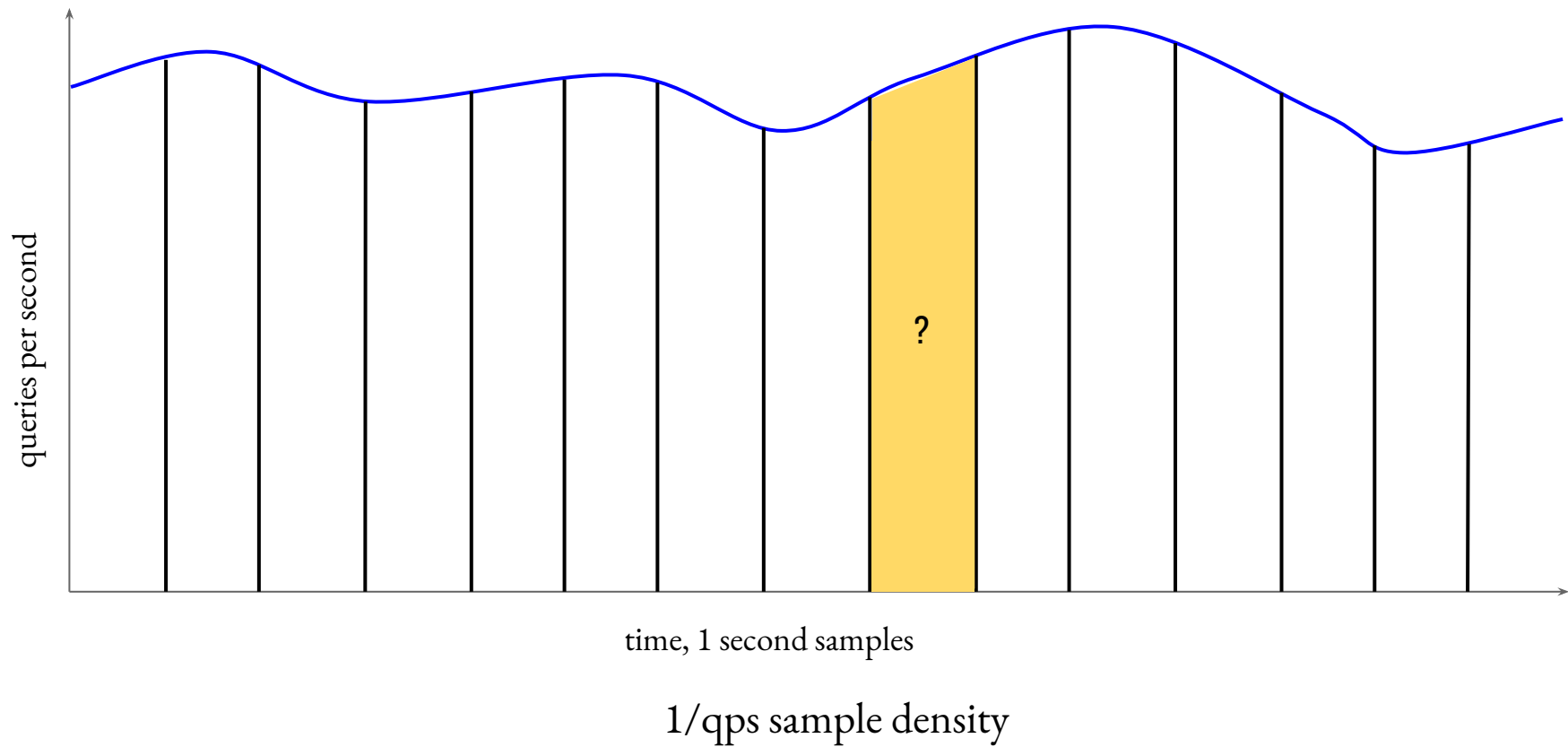
By David Hall [CC BY-SA 4.0 (<https://creativecommons.org/licenses/by-sa/4.0/>)], from Wikimedia Commons

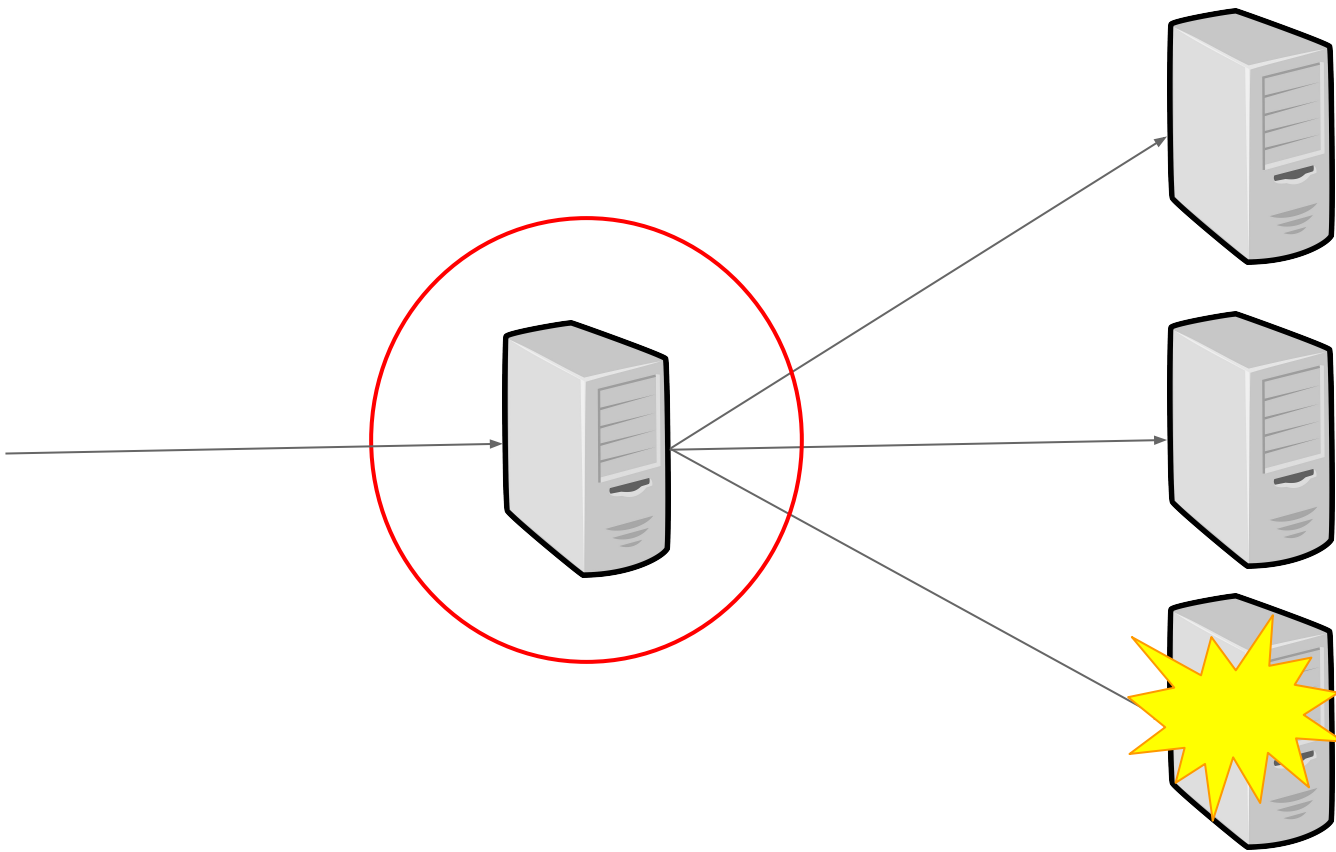
Another way to calculate this is with a *request success rate*:

$$\textit{availability} = \frac{\textit{successful requests}}{\textit{total requests}}$$

Defining SLOs in terms of request success rate makes it easier to measure an error budget

```
var responses = prometheus.NewCounterVec(
    prometheus.CounterOpts{
        Name: "responses",
        Help: "total errors served"},
    []string{"code", "user"})
...
responses.WithLabelValues(
    http.StatusText(400),
    GetUser(req)).Add(1)
```



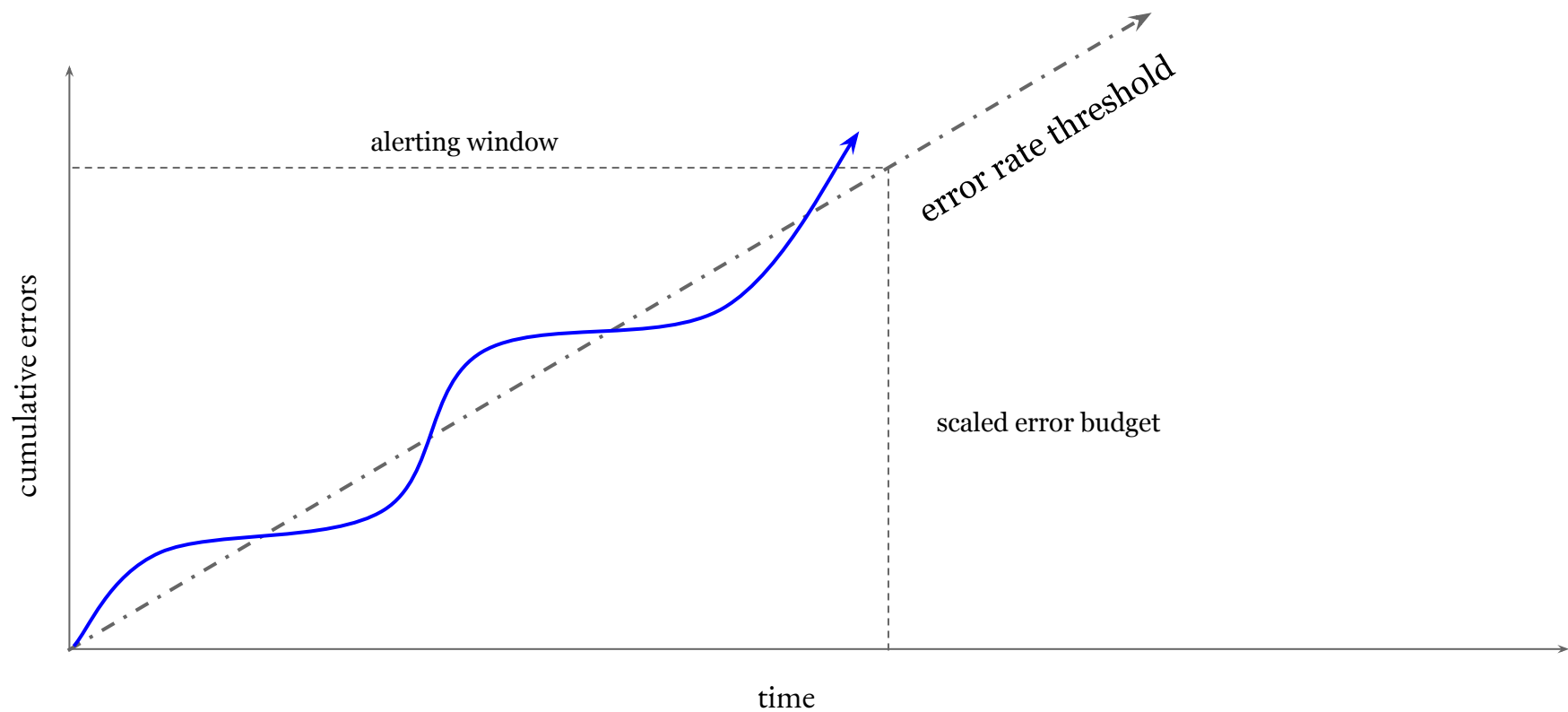


record: error_ratio_by_user

```
expr: sum by (job, user)(  
    rate(responses{code!~"200"}[10s]))  
    / on (job, user)  
    sum by (job, user)(rate(responses[10s]))
```

alert: ErrorRatioTooHigh

```
expr: error_ratio_by_user > 0.01
```



SLO burn

Burn rate maths

Average QPS rate: 1000

SLO: 99% over 1 week

= 604,800,000 total queries

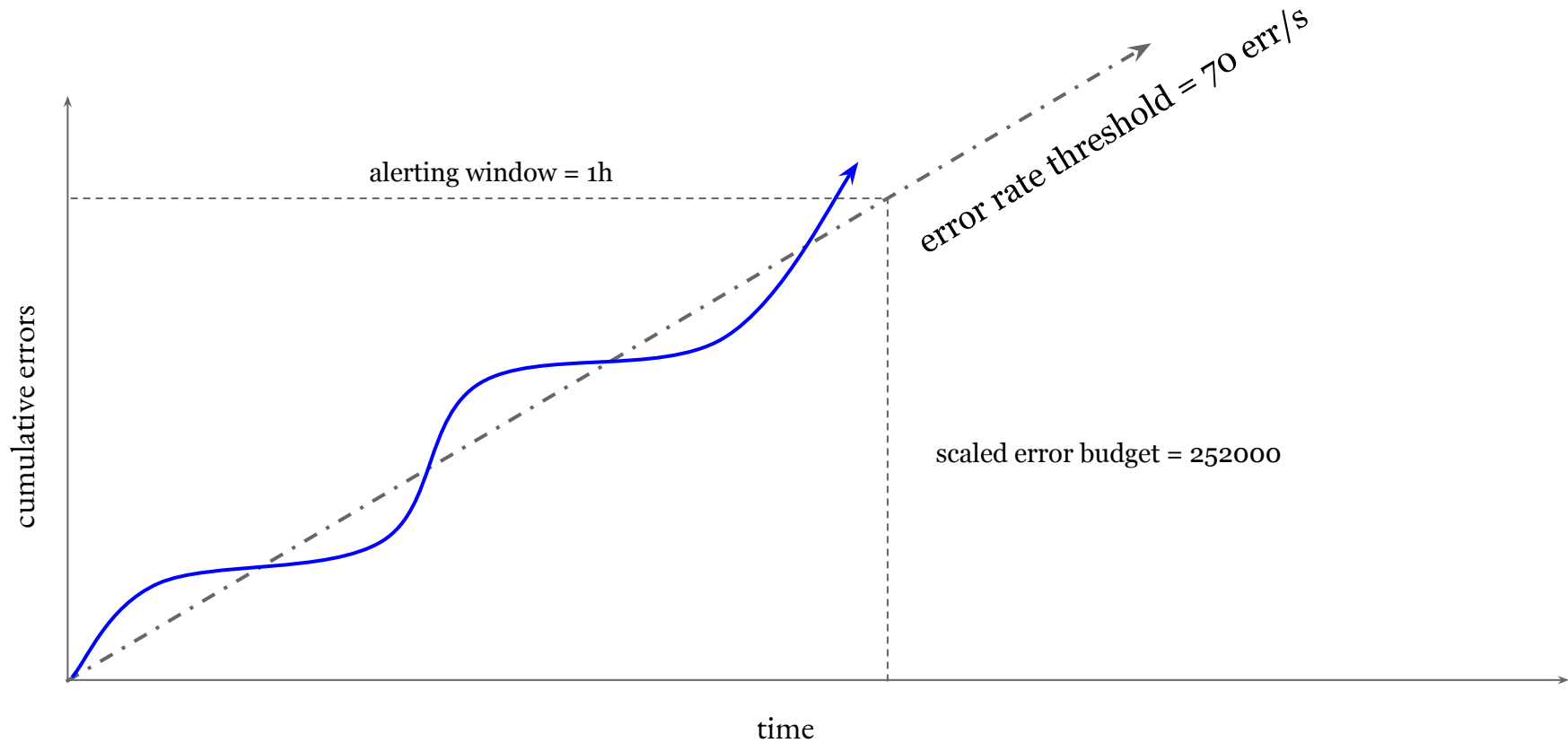
= 6,048,000 permissible errors

Take 1 hour moving average of errors

Page if error budget is going to be exhausted in less than
24 hours

= 6,048,000 errors consumed per day

= 70 err/s = 252,000 errors in 1 hour



Page if 15m rate over 70.

SLO Fast Burn

expr: delta(errors[1h]) > (expected_events *
error_budget / burn_period)

=

expr: delta(errors[1h]) > ((1000 qps * 7d) *
0.01 / 24h)

=

expr: delta(errors[1h]) > 70



**“one of the most powerful
context-sensitive incredibly adaptive
anomaly-detecting and responding agents
in the world”**

-- John Allspaw, Monitorama 2013

1. Symptom-based alerts are good for your health
2. SLO is defined by you, customers, and system
3. SLO implies error budget, informs engineering tolerance
4. Page only on SLO risk, because that's what matters