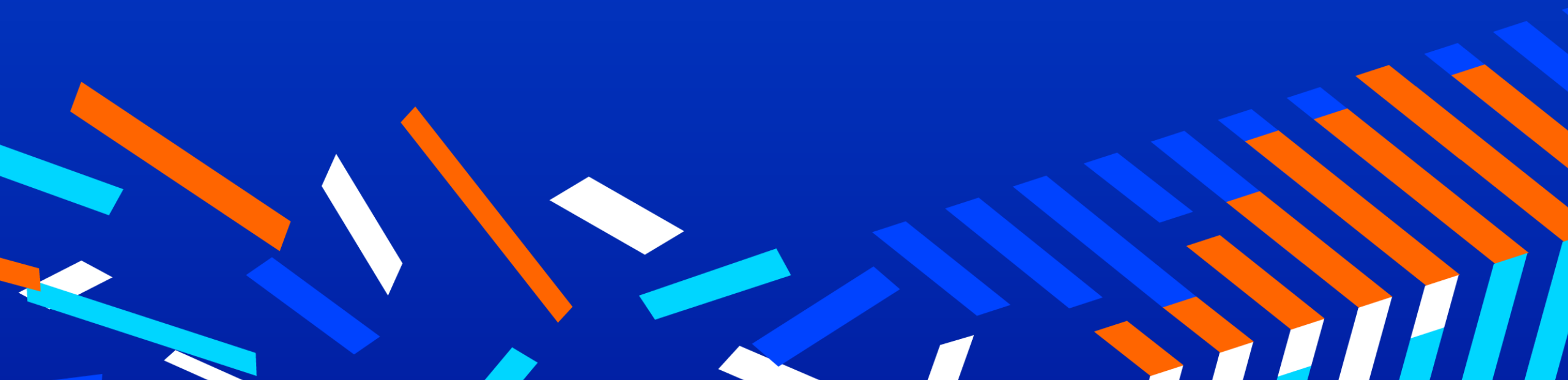


SLOs, Beyond Application Metrics

Denis Ćutić, SRE @ Infobip



How they differ?

Application metrics

Alerting signals

- RED, USE, 4 Golden Signals

Troubleshooting metrics

- HW, networking, JVM, DB, etc.

Stakeholders

- Developers
- Operations

Service Level Objectives

User journey

User satisfaction

Target

Stakeholders

- Engineering
- Operations
- Customer Support
- Product people
- Legal
- Management & C-Level

What about product metrics?

Why?



SLOs – Why?

Happier users, business and engineering teams.

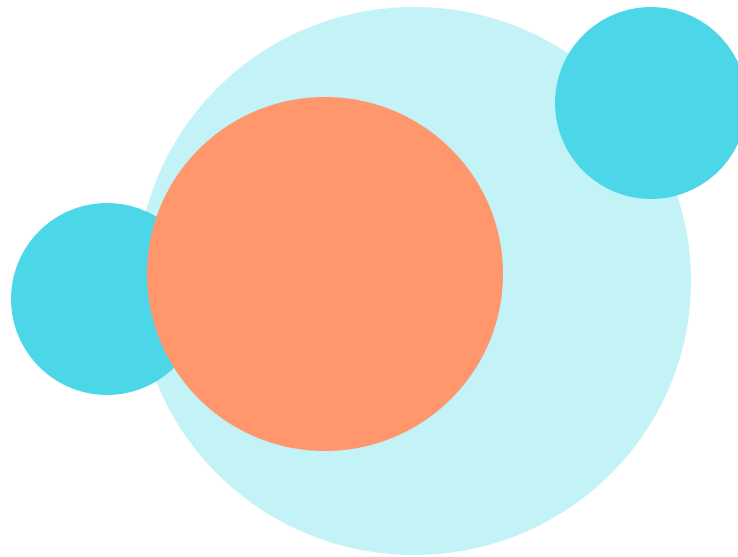
NOT: Number of nines you can append on your target.

What?



SLOs – Data

Approximation of user satisfaction.



SLOs won't...

...tell you how reliable you should be.

...tell you if you are reliable enough for your business or the users.

...make sure you do not breach your SLA.

SLOs – What?

A way to **gather data** used to improve your software.

A way for the organization to **discuss reliability**.

Data to consider when making decisions.

A **process** not a project.

Data for **alerting**.

SLI
SLO
Error budget

SLIs



SLIs

Define
Implement

SLIs – Example – Send SMS message over HTTP API

POST /sms/2/text

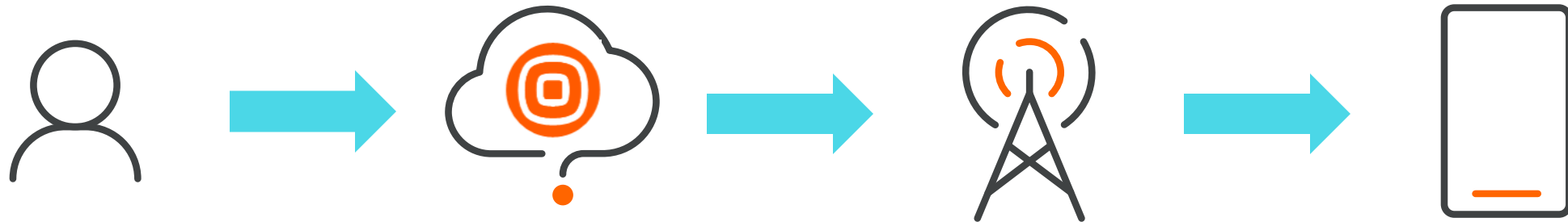
```
{  
  "to": "385912345678",  
  "text": "Welcome"  
}
```

200 OK

```
{  
  "messages": [{  
    "messageId": "12345",  
    "status": {  
      "name": "PENDING_ACCEPTED"  
    }  
  }  
]
```

SLIs - Definition

SMS messages **sent** over HTTP API are successfully **delivered**.



SLIs - Definition

SMS messages **received** over HTTP API are
successfully

processed by the platform and attempted delivery to an MNO.



SLIs - Definition

SMS messages **requests** received over HTTP API are
returning responses with a
successful response status code.

SLIs – Definition

SMS messages sent over HTTP API are successfully delivered.

SLIs – Definition - Success

Functional correctness

Timeliness

Durability

Precision

...

SLIs – Definition

How many SLIs to define?

SLIs – Implementation

How to capture the client perspective?

Where do you collect the data from?

Real user vs Synthetic monitoring

Do we filter some of the data?

SLIs – Implementation

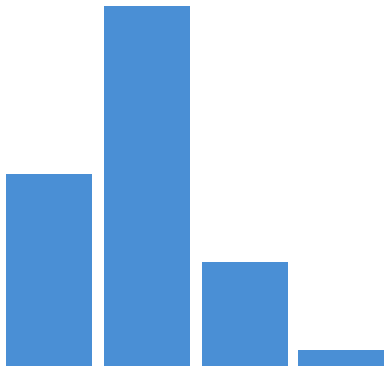
Where do we store the data?

Time-series vs Structured Event DBs

How often do you collect the data?

SLIs – Implementation

Collecting processing times
Histograms vs Percentiles



p99 = 567ms

SLIs - Definition

Percentage of SMS messages requests received over HTTP API
that are returning responses with a
response status code in the range of 200 – 499
as reported by the load balancer
and collected in the time-series DB every 30s.

SLOs



SLOs

SLI

is greater or equal to 99.9% in a 4-week rolling window.

SLOs – The nines

Any number of any number

Calculate depending on your use-case

Extract from past data

SLOs – Time Window

96.714% -> 1d per month

97.8096% -> 2d per quarter

98.08346% -> 7d per year

99.9 -> 43m 49s per month

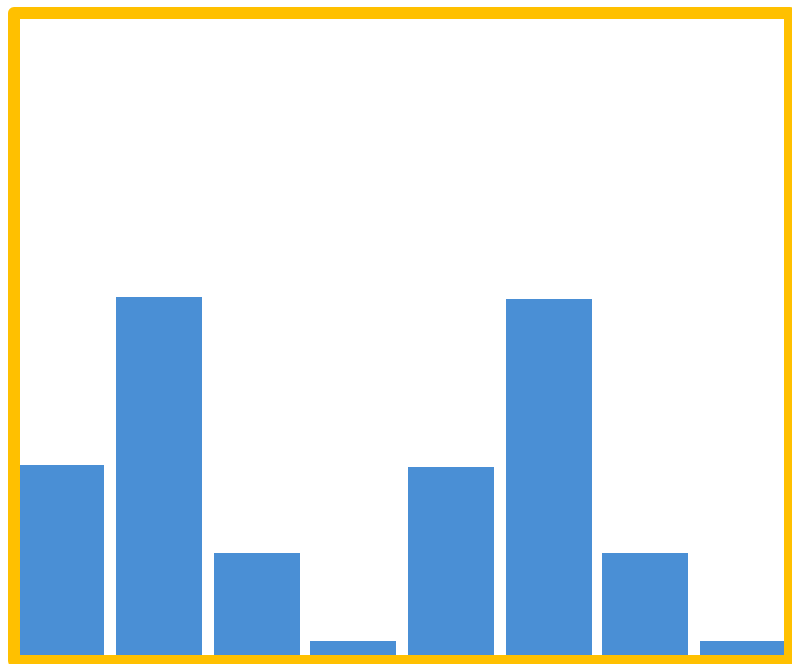
99.95 -> 21m 54s per month

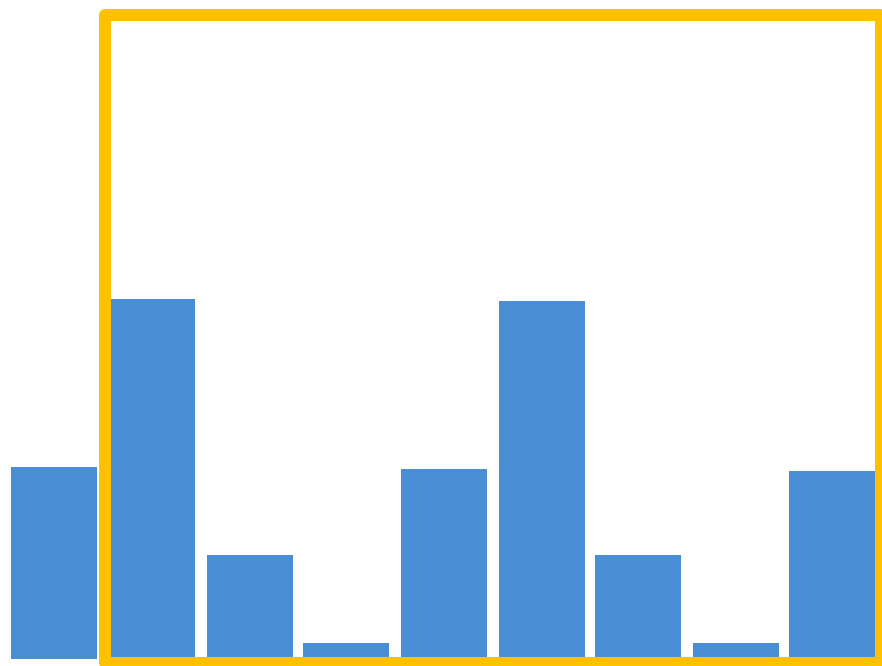
99.99 -> 4m 22s per month

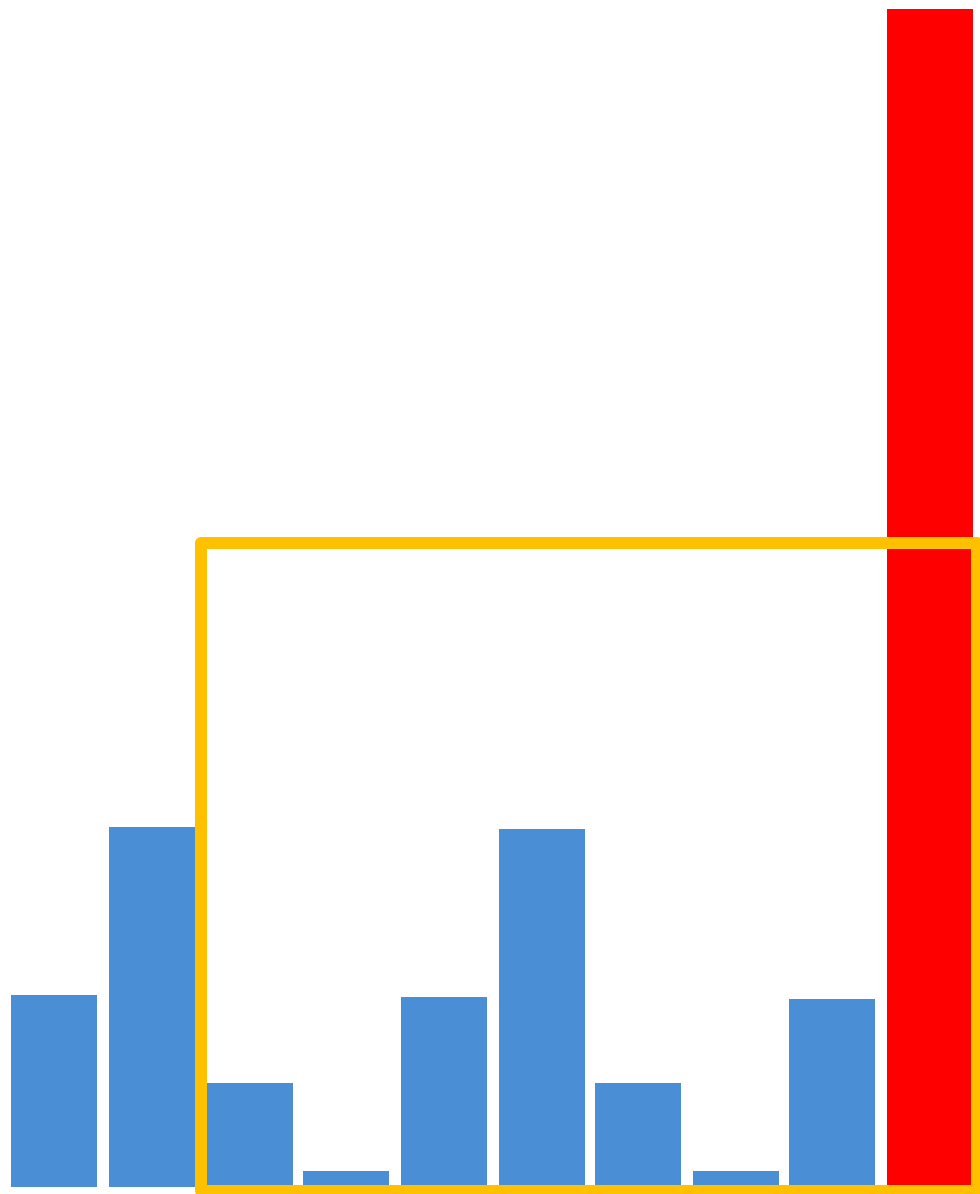
SLOs – Event-based

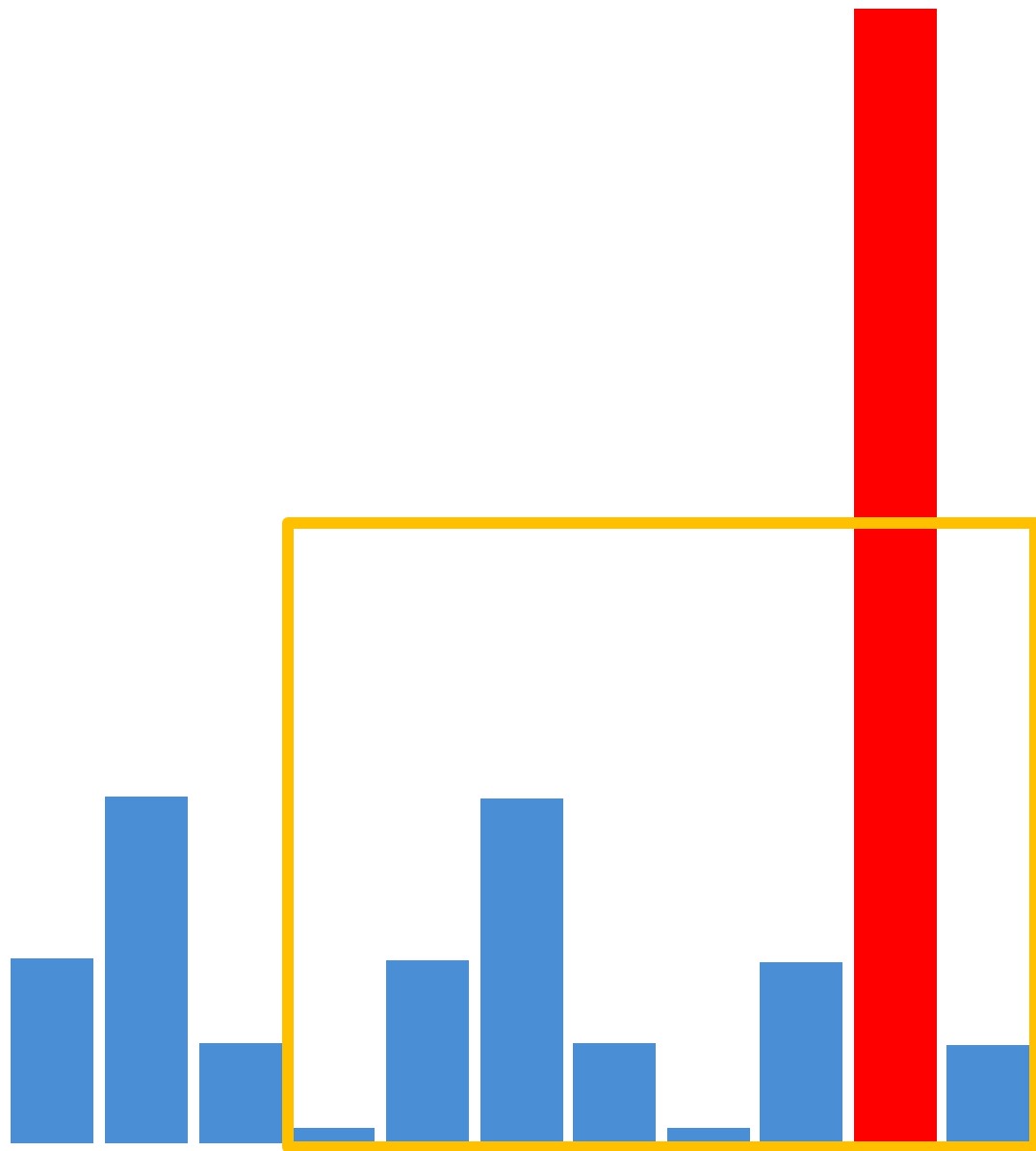
SLI

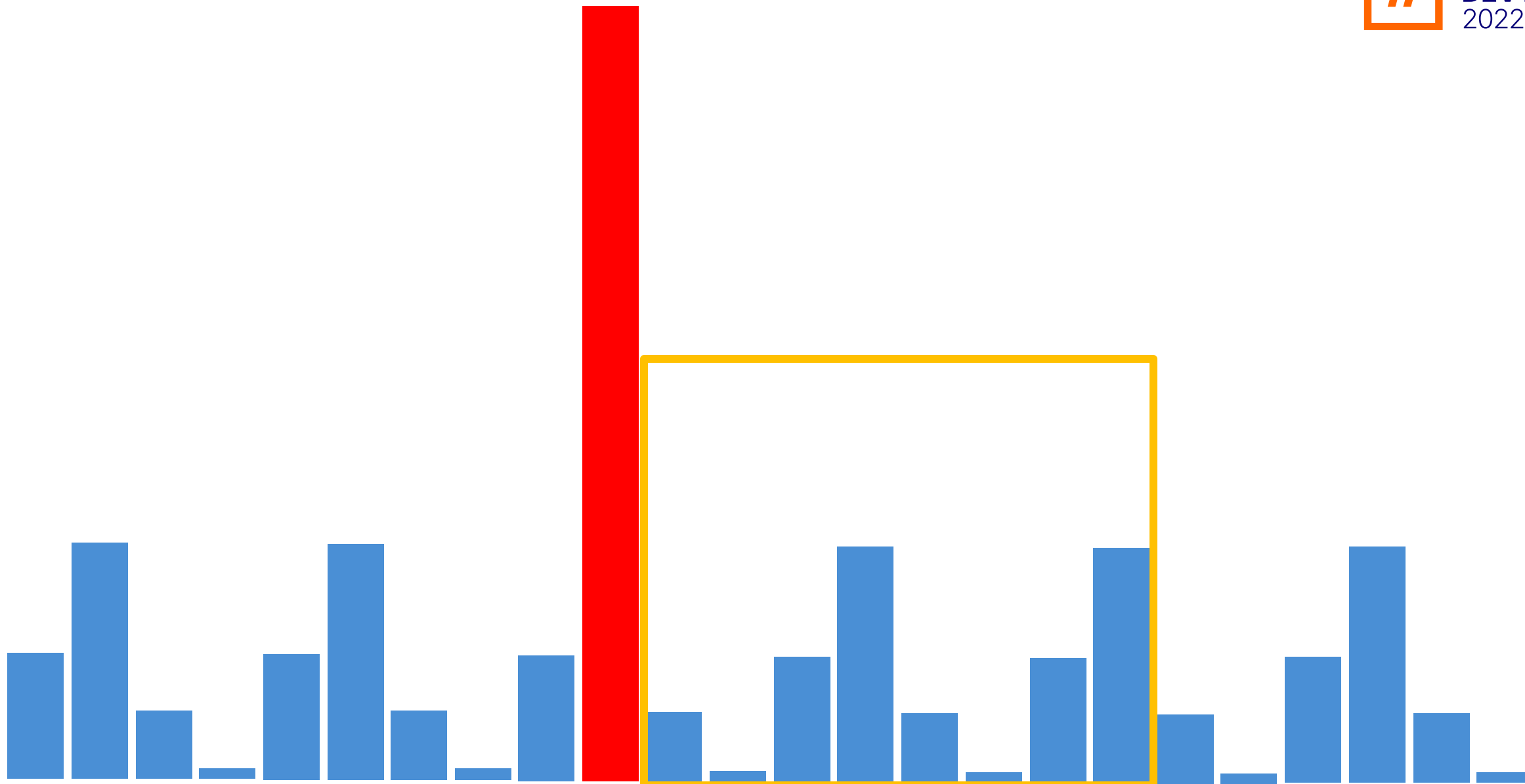
is greater or equal to 99.9% for batches of 1M requests.











Error Budget



Error Budget

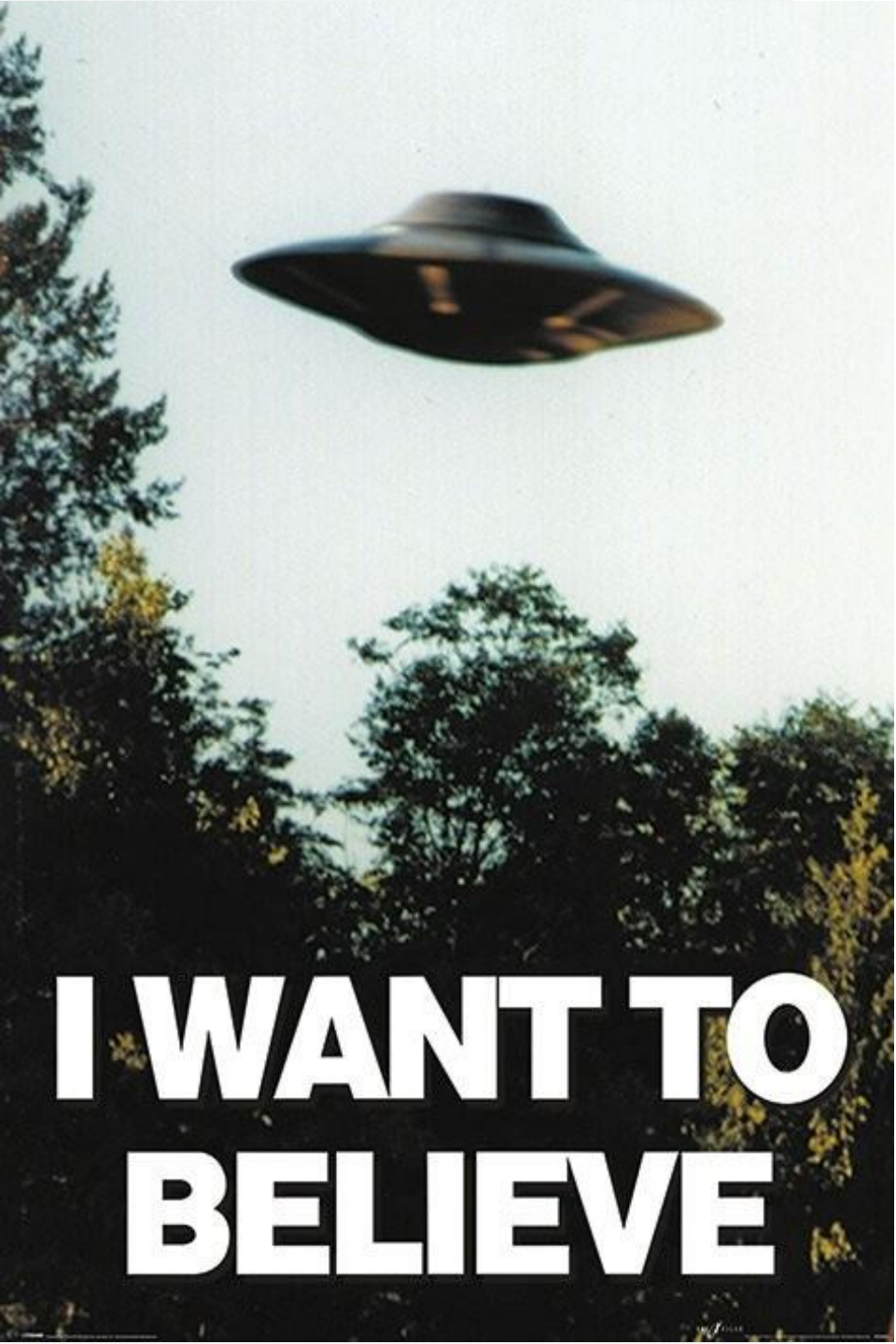
Error Budget = 1 – SLO

Budget burning rate

Over multiple windows

The Truth





- Buy-in
 - ▶ Leadership/Management ✓
 - ▶ Product ✓
 - ▶ Developers ✗
- Nines
 - ▶ Market
 - ▶ Marketing
- Hard
 - ▶ Perfection
 - ▶ Compromise
- Not many resources about how to do it (properly)
- SLIs bring the most value

O'REILLY®

Implementing Service Level Objectives

A Practical Guide to SLIs, SLOs & Error Budgets



Alex Hidalgo

O'REILLY®

SRE with Java Microservices

Patterns for Reliable Microservices
in the Enterprise



Jonathan Schneider
Foreword by Josh Long

A **strategy** for achieving your reliability targets.

Hope is not a strategy

Q&A



Thank you

denis.cutic@infobip.com

