are *still* scaling

How We’~~ve Scaled~~ Meet during COVID19

samschaevitz@

LeadDev Live - Staff Plus

18 September 2021

Proprietary + Confidential

Samantha Schaevitz

@samschaevitz

Senior Staff Software Engineer (SRE)

@ Google Zürich

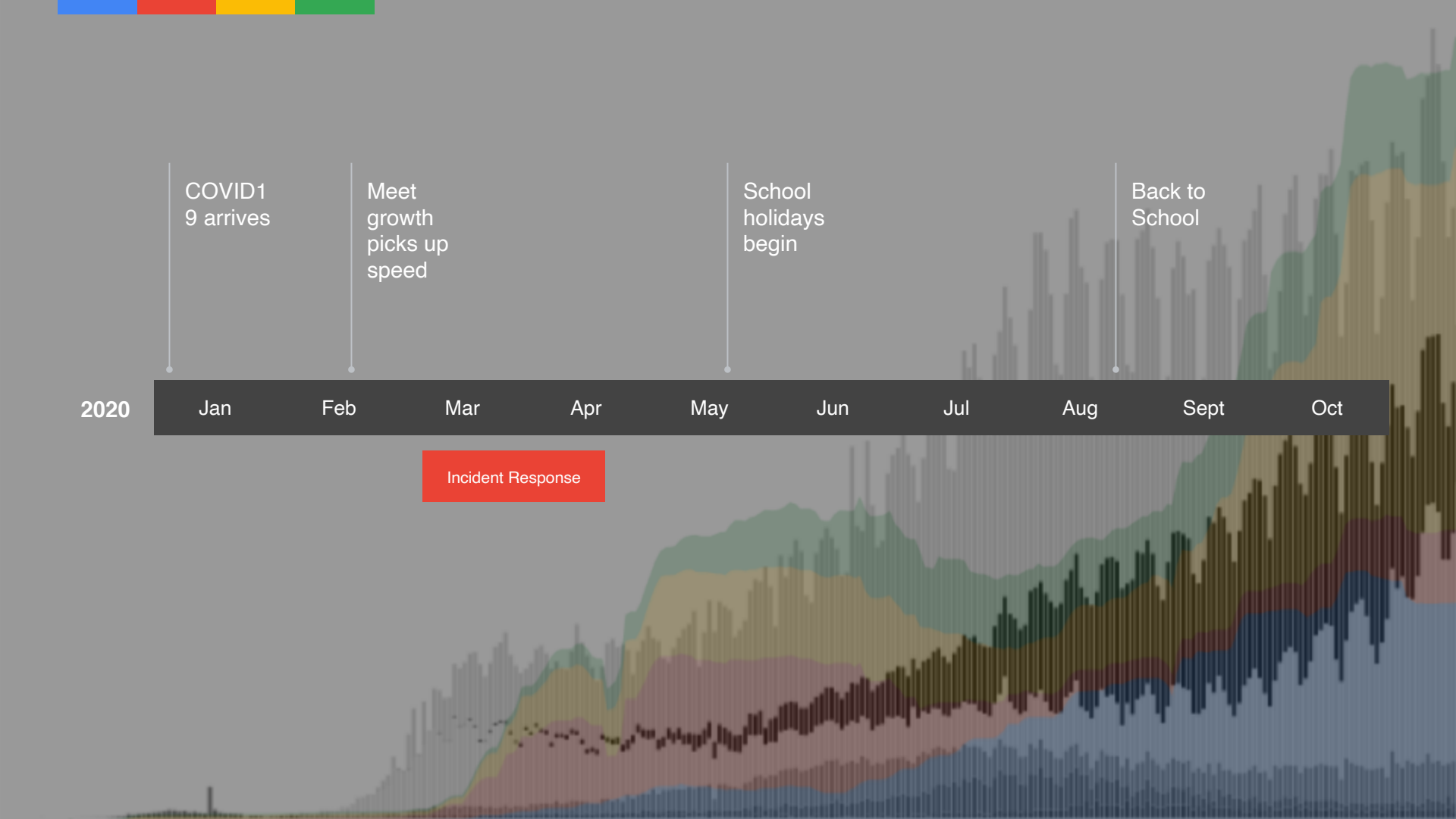
SRE üTL of Meetings

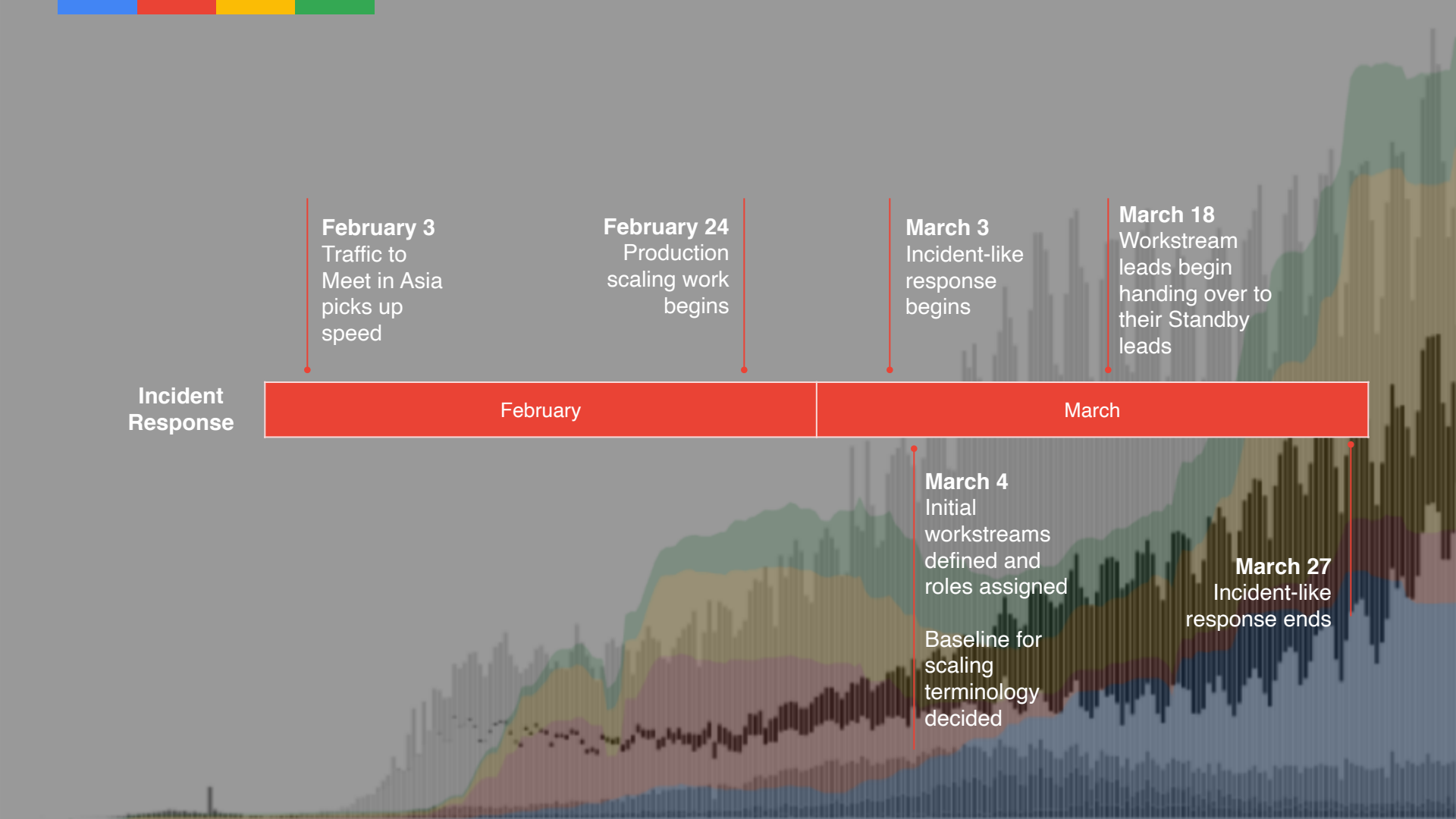
Meet, Calendar, Voice, …

Incident Commander

Meet COVID19 Scaling Incident

Google


COVID1
9 arrives
Meet
growth
picks up
speed
School
holidays
begin
Back to
School
2020
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Incident Response


February 3
Traffic to Meet in Asia picks up speed
February 24
Production scaling work begins
March 3
Incident-like response begins
March 18
Workstream leads begin handing over to their Standby leads
Incident Response
February
March
March 4
Initial workstreams defined and roles assigned
Baseline for scaling terminology decided
March 27
Incident-like response ends

Forming Storming Norming Performing Adjourning

Forming | Storming | Norming | Performing | Adjourning

Declaring an Incident*

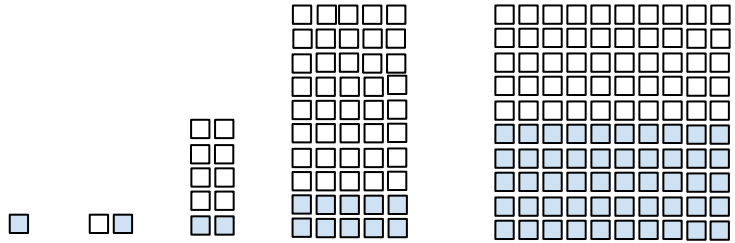
Forming | Storming | Norming | Performing | Adjourning

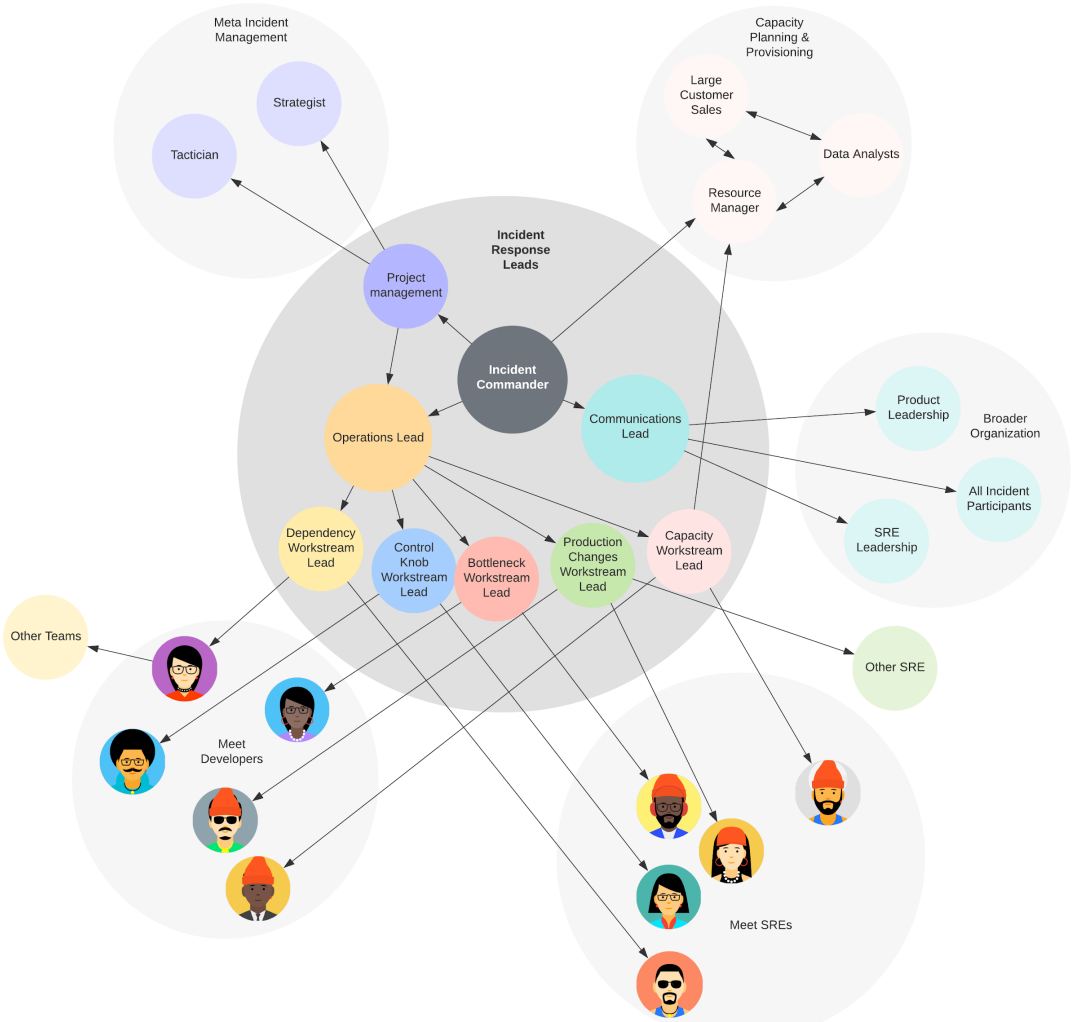
Objective #1: No outages

(but if we have one, degrade gracefully)

Objective #2: Have enough serving capacity

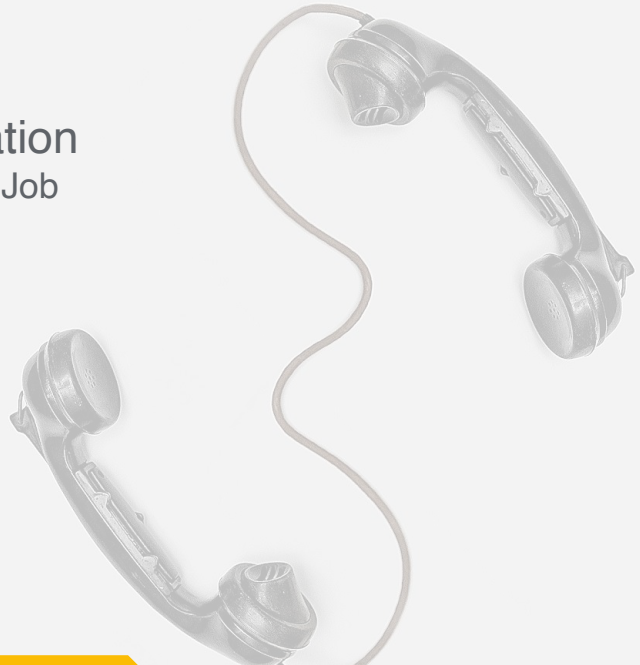
Objective #3: Forecast demand accurately

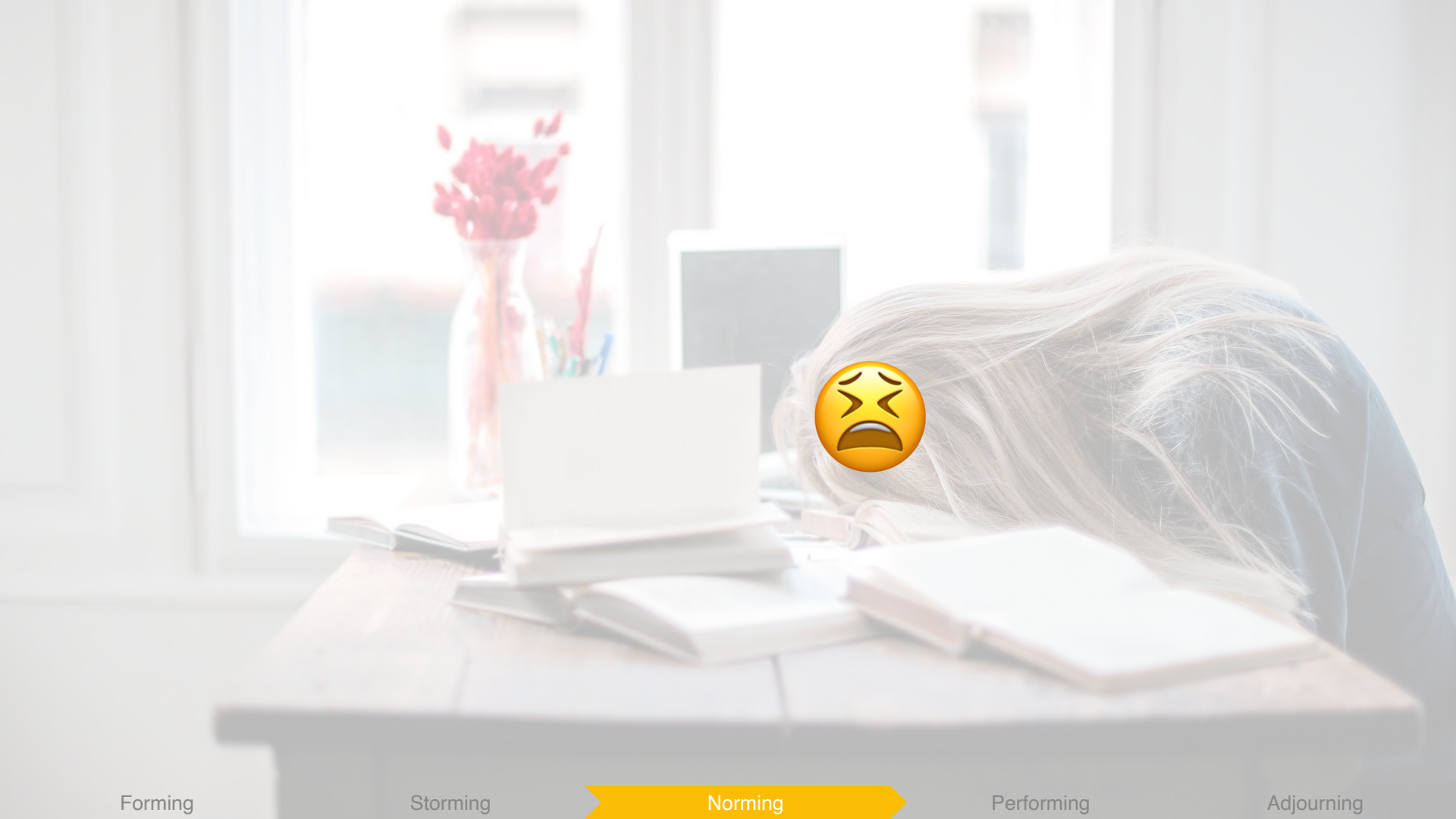

1x
2x
10x
50x
100x



Forming | Storming | **Norming** | Performing | Adjourning

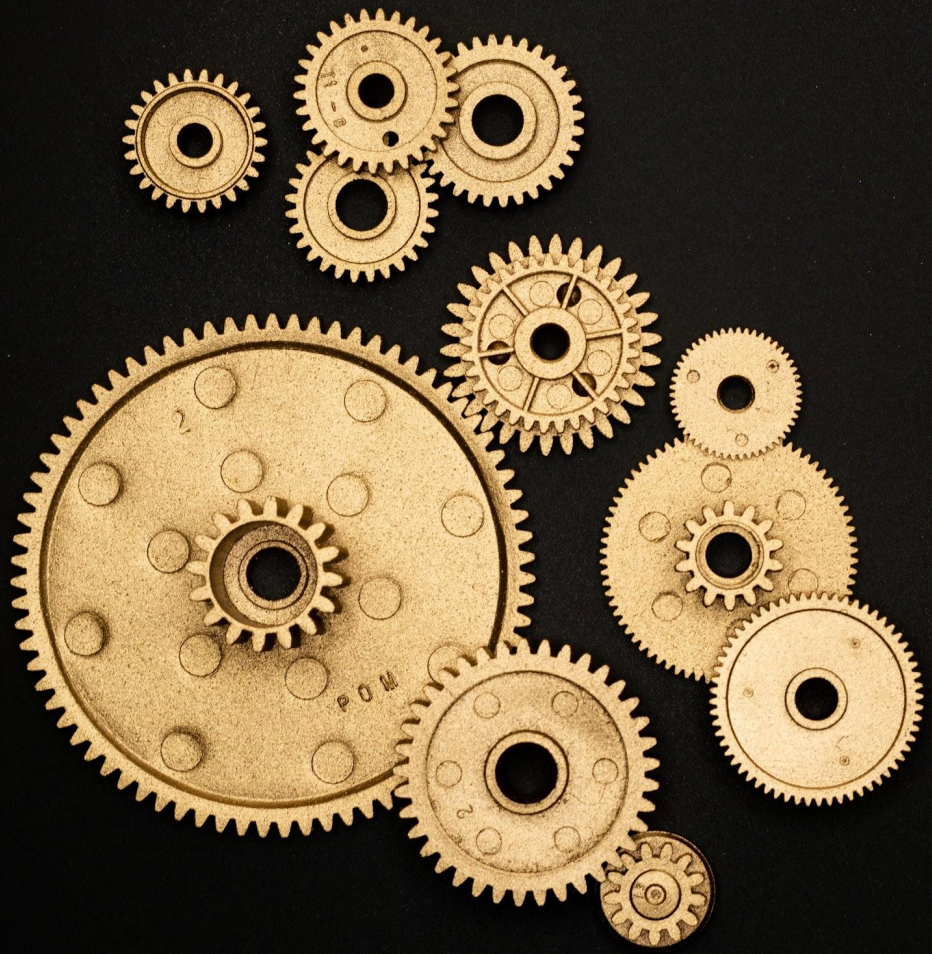
Communication
is a Full-Time Job





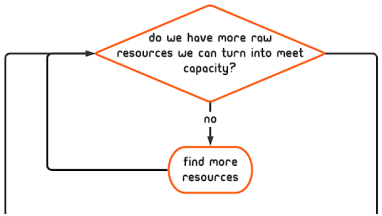
Forming Storming Norming Performing Adjourning

Forming Storming Norming **Performing** Adjourning



Workstreams

Capacity


do we have more raw resources we can turn into meet capacity?
no
find more resources

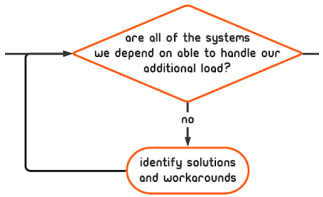
f(sessions) → compute resources

Wave 1	Estimate Resources	→	Get Resources	→	**Turn up serving Capacity**
Wave 2	Estimate Resources	→	**Get Resources**	→	Turn up serving Capacity
Wave 3	**Estimate Resources**	→	Get Resources	→	Turn up serving Capacity

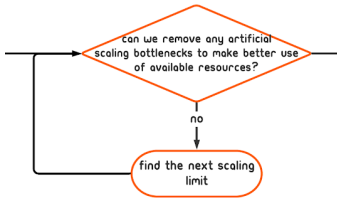
time →

Forming | Storming | Norming | Performing | Adjourning

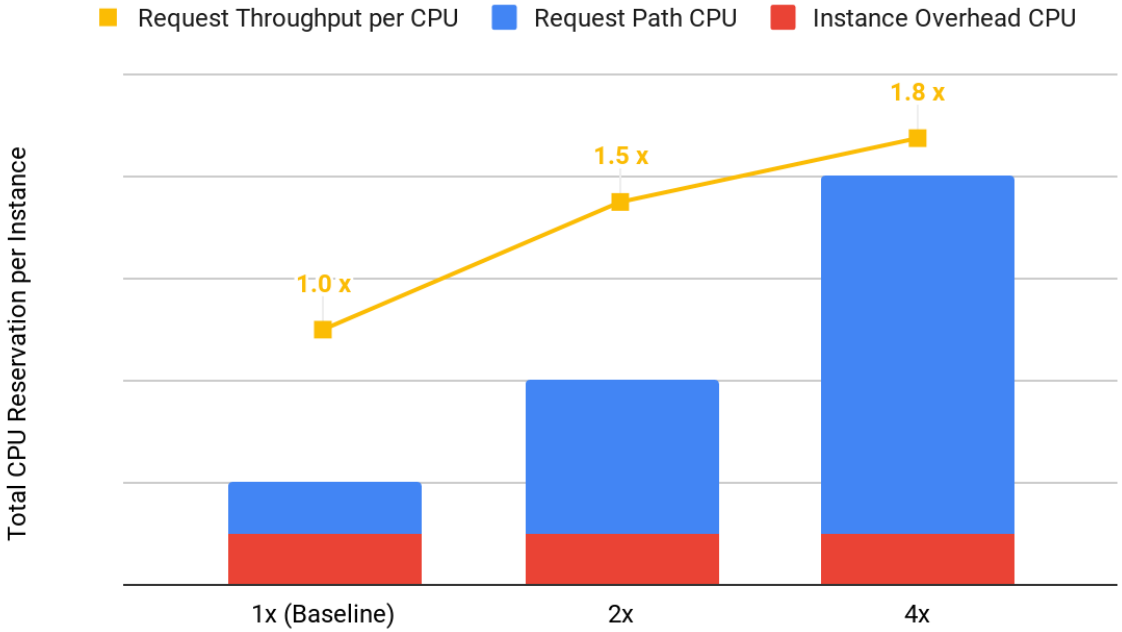
Dependencies



Bottlenecks

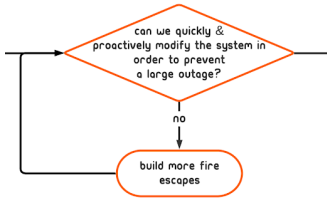


Improving Throughput by Reshaping Server Instances



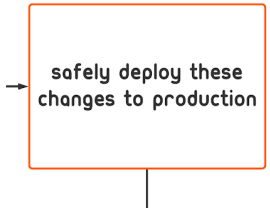
Note: The y-axes are scaled to illustrate the example in relative terms.

Control Knobs



10 HD sessions
OR

Production

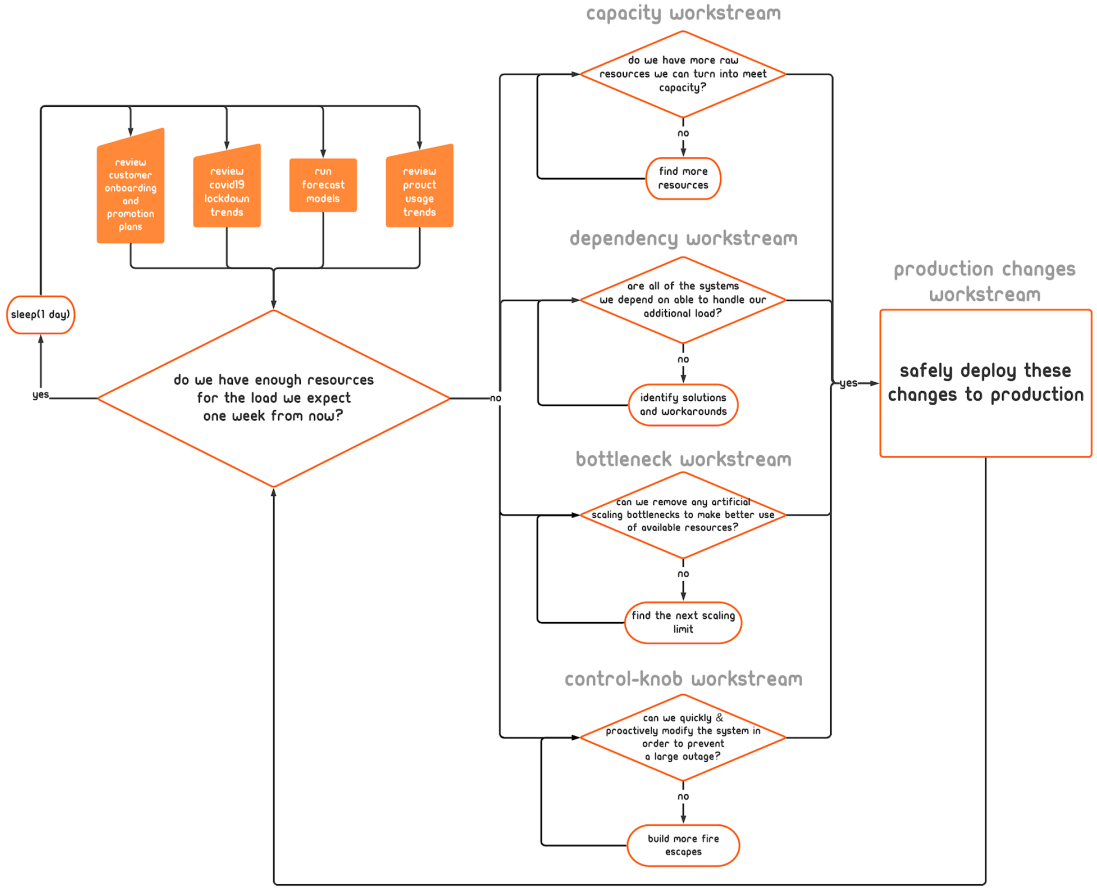


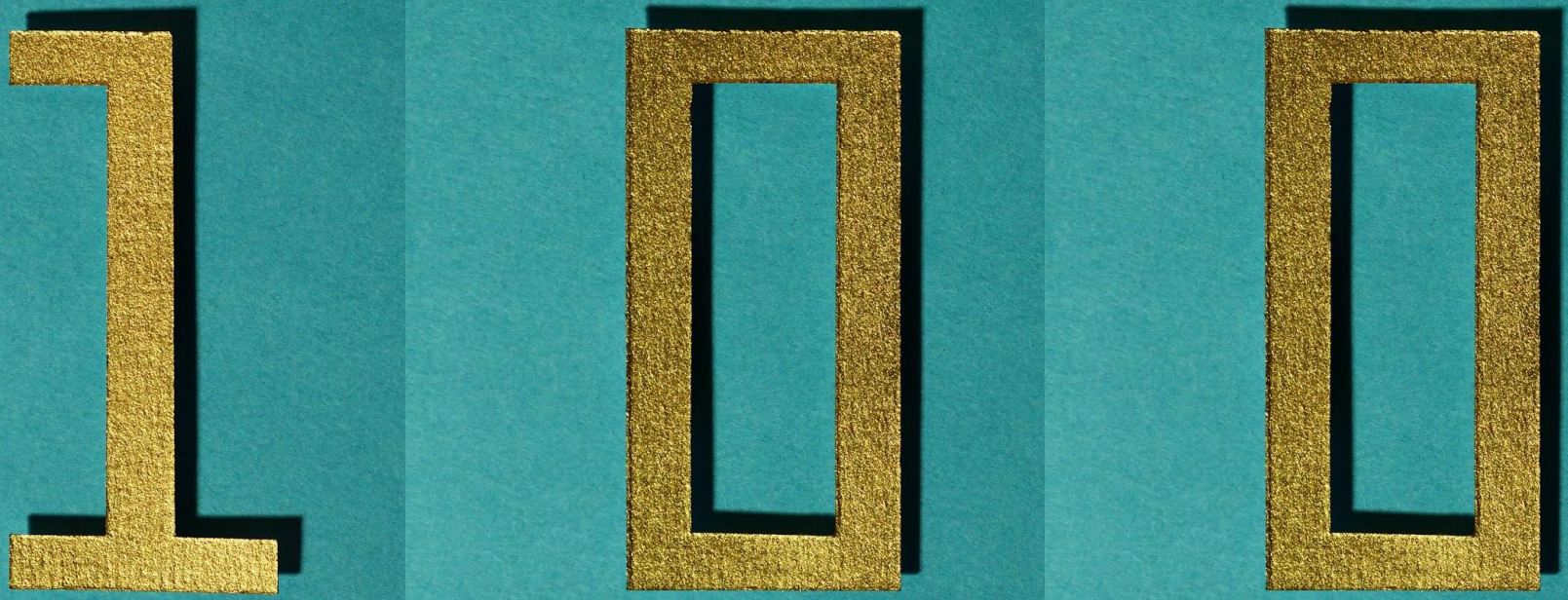
Ad-hoc

↳Documented

↳Atomic Operations = Automated

↳Flow-Control = Automated

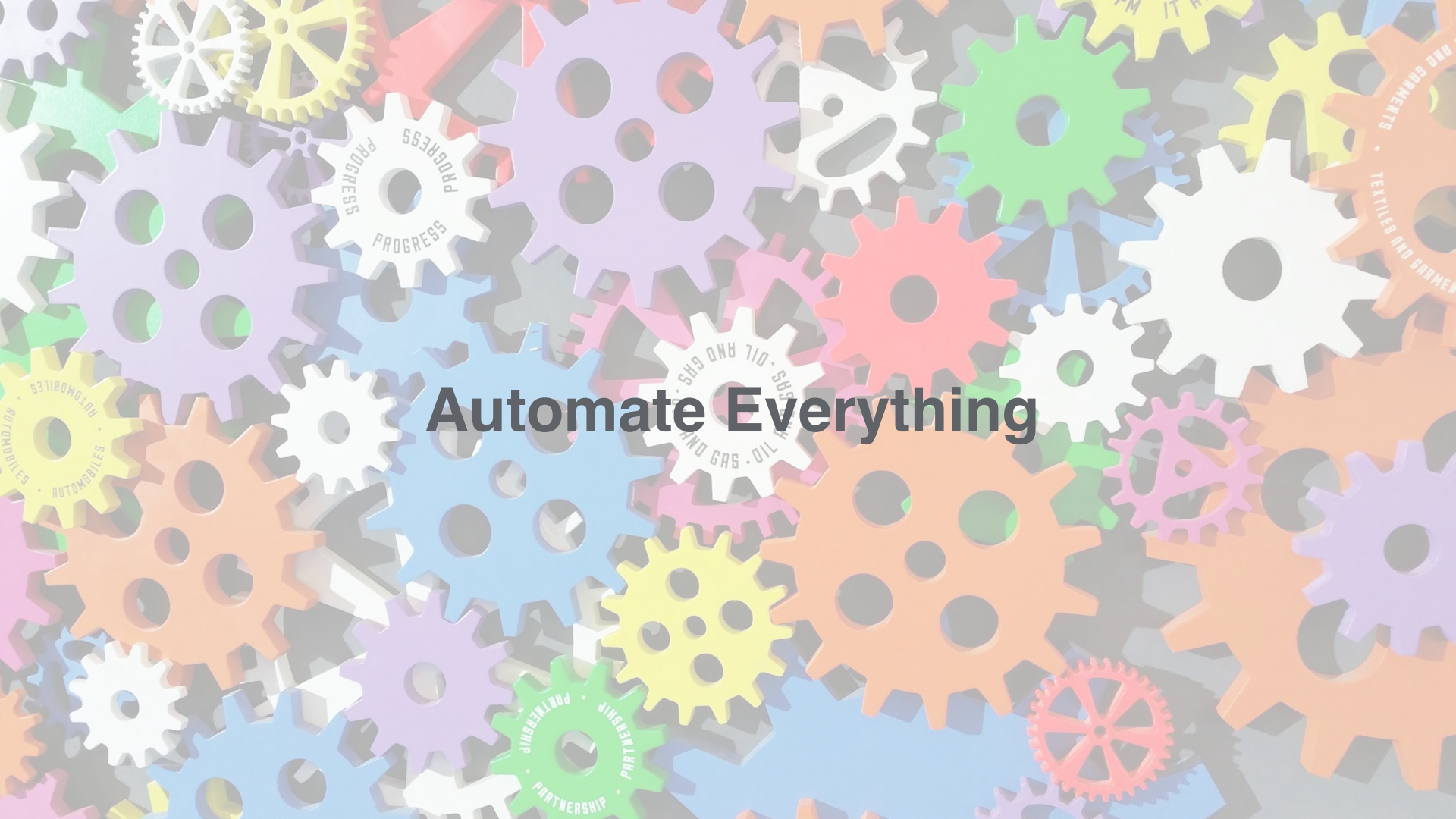

capacity workstream
do we have more raw resources we can turn into meet capacity?
no
find more resources
dependency workstream
are all of the systems we depend on able to handle our additional load?
no
identify solutions and workarounds
bottleneck workstream
can we remove any artificial scaling bottlenecks to make better use of available resources?
no
find the next scaling limit
control-knob workstream
can we quickly & proactively modify the system in order to prevent a large outage?
no
build more fire escapes
review customer onboarding and promotion plans
review covid19 lockdown trends
run forecast models
review prouct usage trends
sleep(1 day)
yes
do we have enough resources for the load we expect one week from now?
no
yes
production changes workstream
safely deploy these changes to production



million

Daily Meeting Participants

daily
→
weekly
planning cycle


Automate Everything

2 3 5

million

Daily Meeting Participants

Thanks!