SRE Munich Meetup

Thu 3. May 2018

Who are we?

danrl

@danrl_com

Dan Lüdtke is the Technical Lead of SRE at eGym, former army officer, and future space traveler.

ingoa

@ingoa

Ingo Averdunk is a Distinguished Engineer in IBM and is responsible for Cloud Service Management and Site Reliability Engineering in the Cloud Adoption, Method and Solution Engineering office for IBM Cloud.

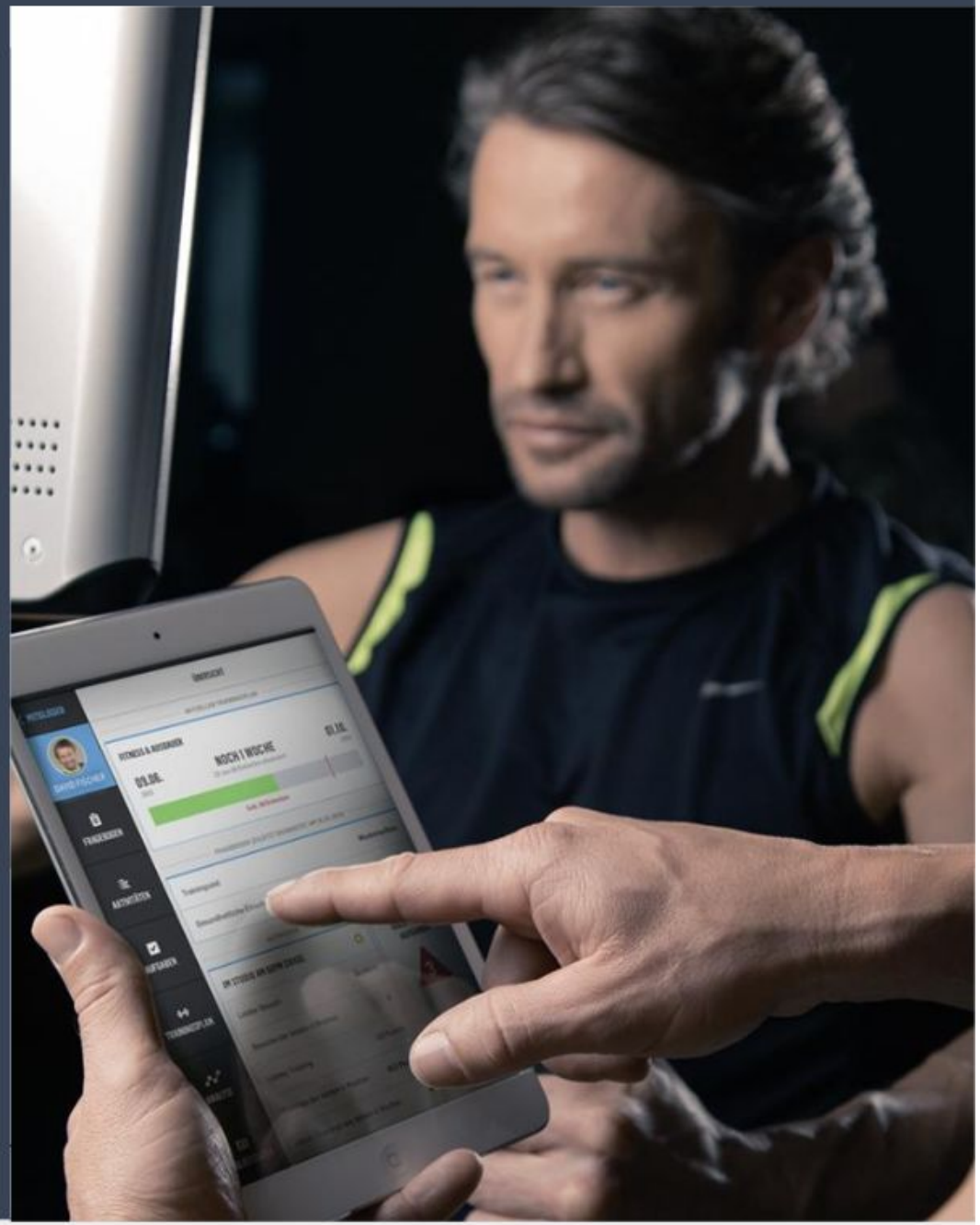
Today's Agenda

- 7:00 pm **Welcome and Kick-off** (Ingo, danrl)
  - A word from the sponsor eGym
  - An experiment: SRE MUC
- 7:30 pm **Recap SREcon 2018** (Ingo, danrl)
- 8:00 pm **Continuous performance profiling in production environments** (Dmitri Melikyan)
- 8:30 pm **Tales from On-call** / **Featured Post Mortem** (Ingo)
- 8:35 pm **Networking** + Drinks
- 9:00 pm **EOF (**Go home inspired!)

A word from our sponsor eGym

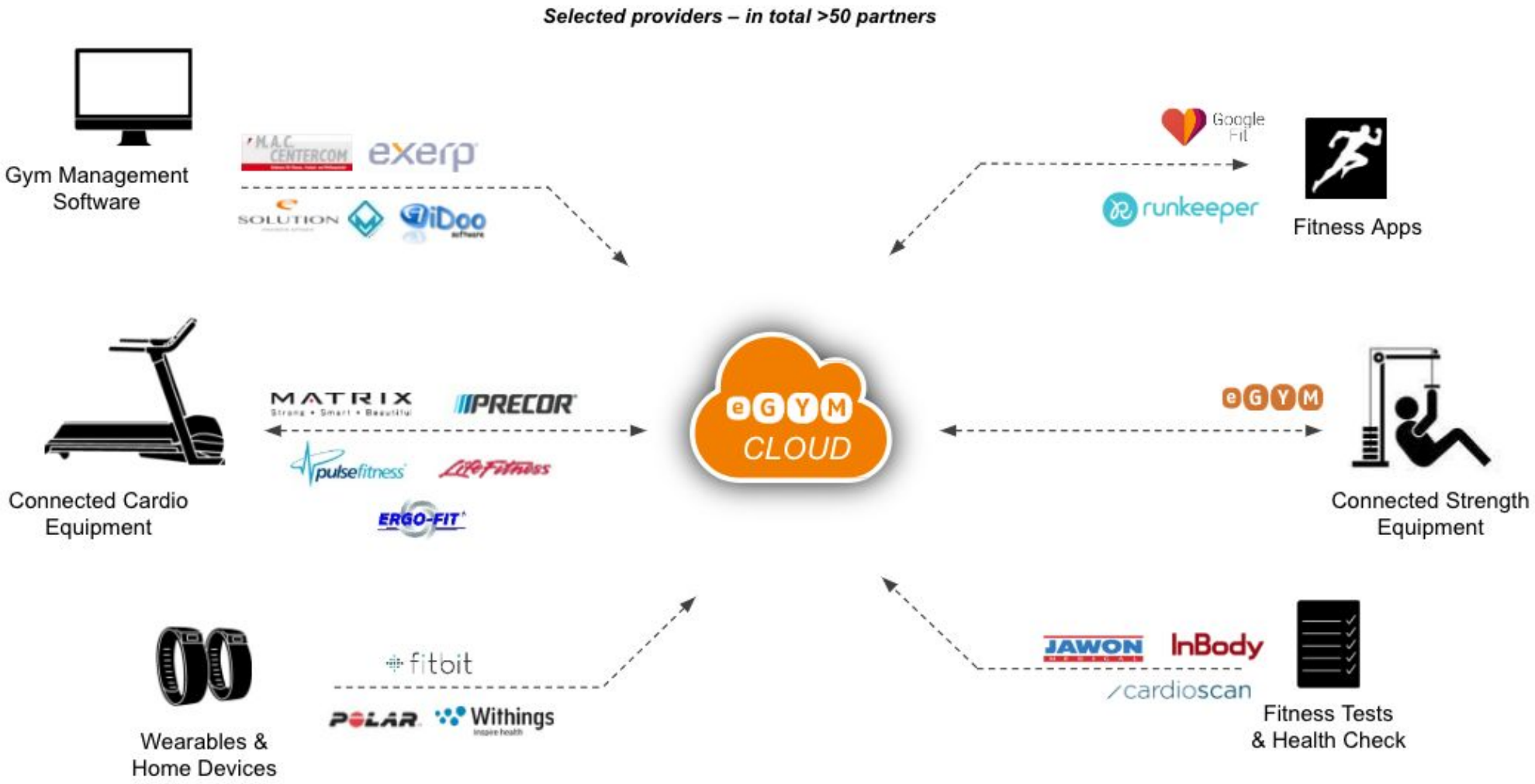
- There is a systemic problem in the fitness market…
- ...the gym only works for a subset of people
- Our mission at eGym is to **make the gym work for everyone**

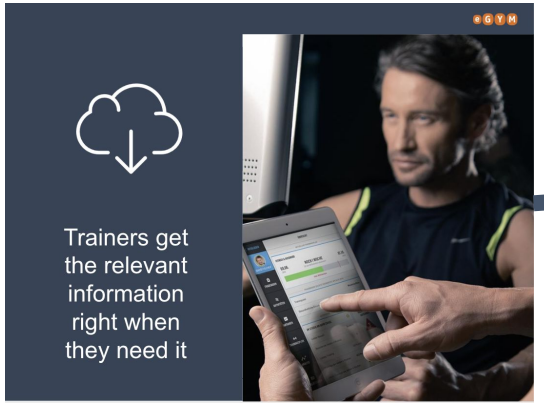
Trainers get the relevant information right when they need it

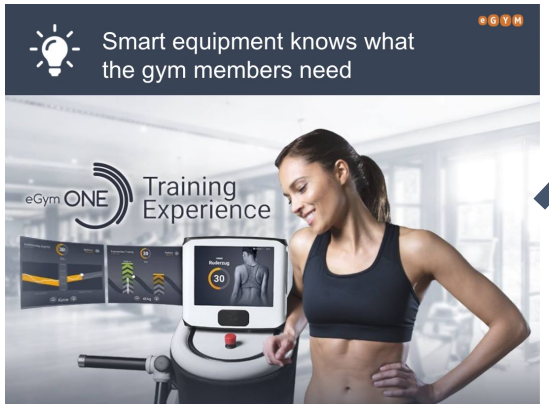


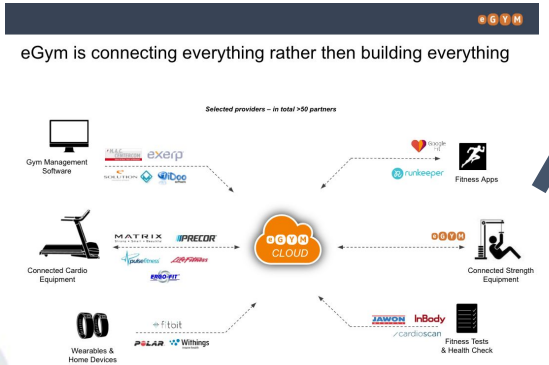
Smart equipment knows what the gym members need

eGym is connecting everything rather then building everything





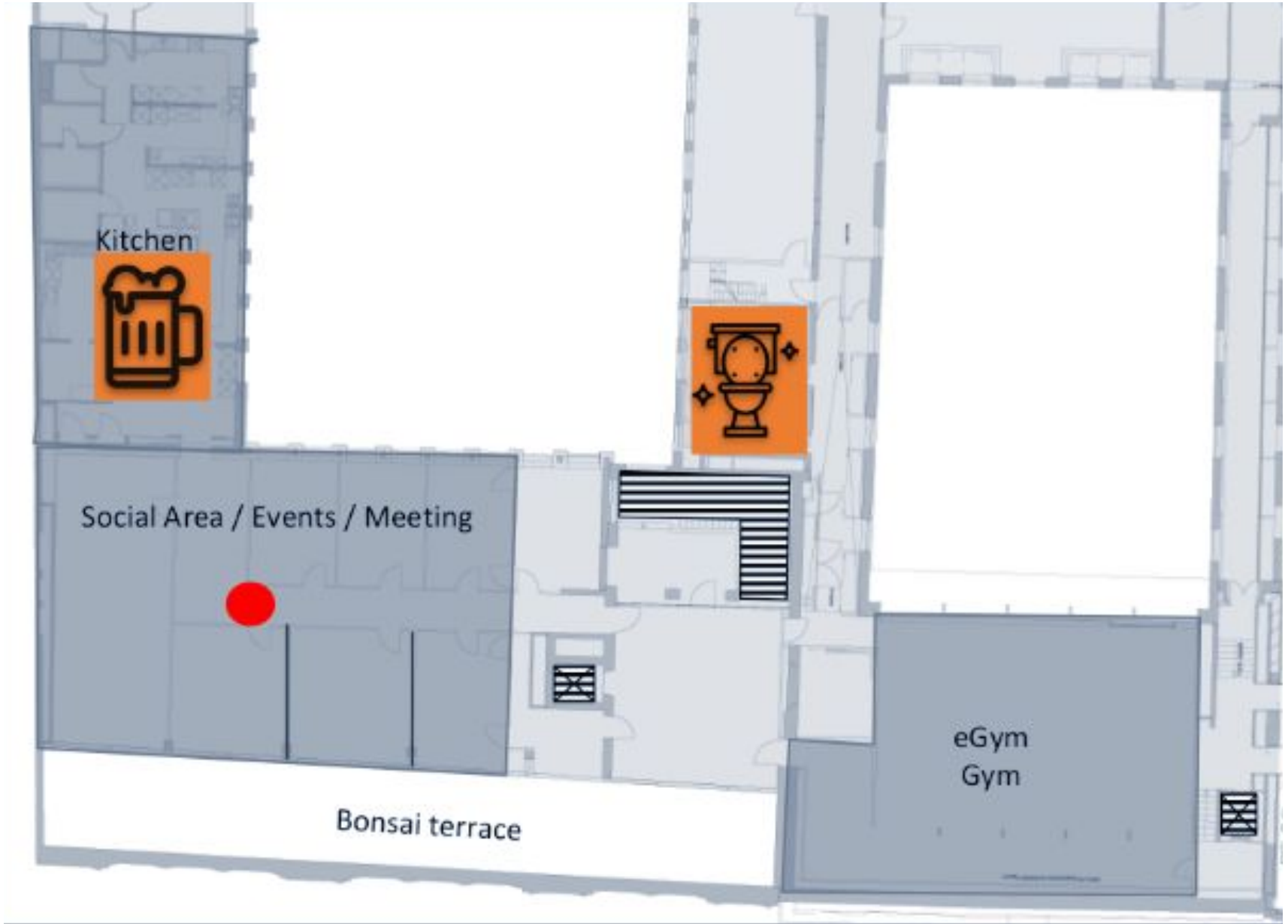




Core Team / SRE

- Run infrastructure
- Run production services
- Share knowledge and support developers
- On-call duty

We are hiring!


Kitchen
Social Area / Events / Meeting
Bonsai terrace
eGym
Gym

An Experiment: SRE MUC

- Is there a SRE community in Munich?
  - Apparently yes!
- Can we add value to Munich's SRE community by addressing their role-specific topics?
  - Without overlapping significantly with the awesome Meetups we have already, such as DevOps, Cloud Native, Microservices, etc.
  - By addressing topics like on-call, incident best practices, post mortems, non-technical SRE topics, looking into how other industries tackle 24/7 and reliability challenges

Participation: Talks

We're always looking for 20-30 minute talks (and 5-8 minute lightning talks) relating to the very broad field of Site Reliability Engineering.

Get in touch with the organizers if you'd like to present!

Participation: On-call Tales

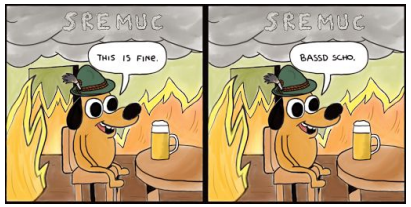
Category:
“Tales from On-call / Featured Post Mortem”

- All Industries
- All aspects of Reliability

Get in touch with the organizers if you'd like to present!

Continued Improvement – Key to SRE

Feedback on todays session



Overall rating (please circle): ☆ 1 ☆ 2 ☆ 3 ☆ 4 ☆ 5

Technical Level :

o	o	o	o	o
Way too technical	Too technical	just right	too less technical	way too less technical

Duration:

o	o	o	o	o
Way too long	Too long	just right	too short	way too short

Comments & Ideas:

_____

_____

_____

_____

Chatham House Rule

Chatham House Rule applies

Example: This indicates a slide or agenda point that is under Chatham House Rule regulation.

When a meeting, or part thereof, is held under the Chatham House Rule, participants are free to use the information received, but neither the identity nor the affiliation of the speaker(s), nor that of any other participant, may be revealed.

https://en.wikipedia.org/wiki/Chatham_House_Rule

Recap SREcon 2018

Dan
Ingo

Agenda

usenix SRECON AMERICAS

Wireless Network Information
Network: SREcon18
Passkey: usenix2018
WiFi sponsored by Scaylr

www.usenix.org/srecon18americas
#SREcon

Tuesday, March 27, 2018

7:30 am–9:00 am	Continental Breakfast, *Sponsored by Rundeck*	Grand Ballroom Foyer
9:00 am–12:30 pm	Workshop Track 1 Containers from Scratch Avishai Ish-Shalom, Aleph VC, and Nati Cohen, Here Technologies	Grand Ballroom C
9:00 am–12:30 pm	Workshop Track 2 SRE Classroom, or How to Build a Distributed System in 3 Hours Salim Virji, Laura Nolan, and Phillip Tischler, Google	Grand Ballroom GH
9:00 am–12:30 pm	Workshop Track 3 Profiling JVM Applications in Production Sasha Goldshtein, Sela Group	Grand Ballroom F
9:00 am–12:30 pm	Workshop Track 4 Incident Command for IT—What We've Learned from the Fire Department Brent Chapman, Great Circle Associates, Inc.	Grand Ballrom AB
10:30 am–11:00 am	Break with Refreshments, *Sponsored by Nutanix*	Grand Ballroom Foyer
12:30 pm–2:00 pm	Luncheon, *Sponsored by LinkedIn*	Santa Clara Ballroom
2:00 pm–5:30 pm	Workshop Track 1 Kubernetes 101 Bridget Kromhout, Microsoft	Grand Ballroom AB
2:00 pm–5:30 pm	Workshop Track 2 Chaos Engineering Bootcamp Tammy Butow, Gremlin	Grand Ballroom GH
2:00 pm–5:30 pm	Workshop Track 3 Ansible for SRE Teams James Meickle, Quantopian	Grand Ballroom F
2:00 pm–5:30 pm	Workshop Track 4 Tech Writing 101 for SREs Lisa Carey, Google	Grand Ballroom C
3:30 pm–4:00 pm	Break with Refreshments, *Sponsored by Dropbpx*	Grand Ballroom Foyer
5:30 pm–6:30 pm	Happy Hour, *Sponsored by Google*	Terra Courtyard

General Information

Attendee Badges

Exhibitors at the conference may ask to scan your badge. If you allow them to do so, they will have access to the following information you entered when registering:

- Name
- Company
- Title
- Work Email
- Work Address
- Work Phone Number

If you do not wish to share this information with our exhibitors, please do not allow them to scan your badge. A printout of the above information is available for you at the registration desk if you'd like to review it.

Attendee List

The attendee list is available for download at www.usenix.org/srecon18americas/program. Access is restricted to registered attendees. Please log in via the account you used to register for the conference.

Conference Videos

In keeping with our Open Access Policy, videos of the talks on Wednesday and Thursday will be available to everyone after the conference. *Sponsored by Indeed.*

Power Outlets

Power strips will be available in the meeting rooms on a first-come, first-served basis.

Consent to Use Photographic Images

Participation in USENIX events and related activities constitutes an agreement by the participant to USENIX's unlimited and permanent use and distribution of the participant's image in various media. Questions? Please stop by the USENIX registration desk.

SREcon18 Americas Is Mobile!

Download the USENIX app to access the schedule, speakers, sponsors, exhibitors, hotel floor plans, and more!

Download on the App Store | GET IT ON Google Play

Mobile Web version also available at www.usenix.org/sched

Wednesday, March 28, 2018

#SREcon

7:30 am–8:30 am	Continental Breakfast, *Sponsored by Squarespace*	Grand Ballroom Foyer
8:30 am–8:45 am	**Welcome and Opening Remarks** Program Co-Chairs: Kurt Andersen, LinkedIn, and Betsy Beyer, Google	Grand Ballroom ABCFGH
	Opening Plenary Session Grand Ballroom ABCFGH	
8:45 am–9:15 am	**If You Don't Know Where You're Going, It Doesn't Matter How Fast You Get There** Nicole Forsgren and Jez Humble, DevOps Research and Assessment (DORA)	
9:15 am–9:45 am	**Security and SRE: Natural Force Multipliers** Cory Scott, LinkedIn	
9:45 am–10:15 am	**What It Really Means to Be an Effective Engineer** Edmond Lau, Co Leadership	
10:15 am–10:55 am	Break with Refreshments, *Sponsored by PayPal*	Grand Ballroom Foyer

	Talks Track 1 Grand Ballroom ABC	Talks Track 2 Grand Ballroom FGH
10:55 am–11:35 am	**SparkPost: The Day the DNS Died** Jeremy Blosser, SparkPost	**Beyond Burnout: Mental Health and Neurodiversity in Engineering** James Meickle, Quantopian
11:40 am–12:20 am	**Stable and Accurate Health-Checking of Horizontally-Scaled Services** Lorenzo Saino, Fastly	**Bootstrapping an SRE Team: Effecting Culture Change and Leveraging Diverse Skill Sets** Aaron Wieczorek, U.S. Digital Service
12:20 pm–1:35 pm	Luncheon, *Sponsored by eBay*	Santa Clara Ballroom
1:35 pm–2:15 pm	**Don't Ever Change! Are Immutable Deployments Really Simpler, Faster, and Safer?** Rob Hirschfeld, RackN	**Building Successful SRE in Large Enterprises—One Year Later** Dave Rensin, Google
2:20 pm–3:00 pm	**Lessons Learned from Our Main Database Migrations at Facebook** Yoshinori Matsunobu, Facebook	**Working with Third Parties Shouldn't Suck** Jonathan Mercereau, traffiq corp.
3:05 pm–3:25 pm	**Leveraging Multiple Regions to Improve Site Reliability: Lessons Learned from Jet.com** Andrew Duch, Jet.com/Walmart Labs	**When to NOT Set SLOs: Lots of Strangers Are Running My Software!** Marie Cosgrove-Davies, Google
3:25 pm–4:05 pm	Break with Refreshments, *Sponsored by Microsoft Azure*	Grand Ballroom Foyer
4:05 pm–4:45 pm	**Lessons Learned from Five Years of Multi-Cloud at PagerDuty** Arup Chakrabarti, PagerDuty	**How SREs Found More Than $100 Million Using Failed Customer Interactions** Wes Hummel, PayPal
4:50 pm–5:10 pm	**Protect Your Data Centers with Safety Constraints** Christina Schulman and Etienne Perot, Google	**Learning at Scale Is Hard! Outage Pattern Analysis and Dirty Data** Tanner Lund, Microsoft
5:15 pm–5:35 pm	**Real World SLOs and SLIs: A Deep Dive** Matthew Flaming and Elisa Binette, New Relic	**How Not to Go Boom: Lessons for SREs from Oil Refineries** Emil Stolarsky, Shopify
5:35 pm–7:30 pm	Reception, *Sponsored by Circonus*	Terra Courtyard
7:30 pm–9:00 pm	Lightning Talks	Grand Ballroom ABCFGH

Visit the Sponsor Showcase!

Stop by the sponsor booths in Grand Ballroom DE and Lobby West.

Tuesday: 10:00 am–5:30 pm
Wednesday: 8:00 am–6:00 pm
Thursday: 10:00 am–2:00 pm

Espresso Cart available in Grand Ballroom DE during showcase hours.

USENIX Conference Policies

USENIX policies for our conferences are available online at www.usenix.org/conferences/policies:

- USENIX Event Code of Conduct
- Conference Network Policy
- Conference Submissions Policy
- Statement on Environmental Responsibility

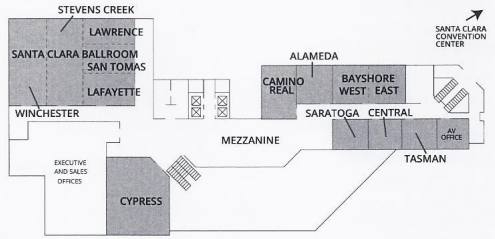
Thursday, March 29, 2018

#SREcon

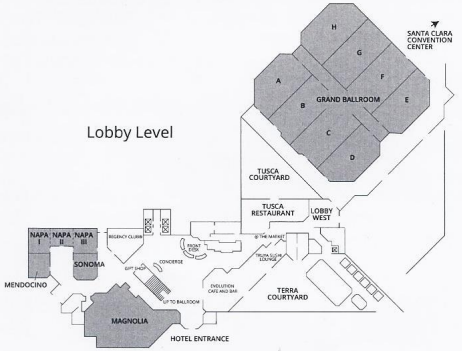
7:30 am–8:30 am	Continental Breakfast, *Sponsored by VMware and Wavefront*	Grand Ballroom Foyer

	Talks Track 1 Grand Ballroom ABC	Talks Track 2 Grand Ballroom FGH
8:30 am–9:10 am	**Containerization War Stories** Ruth Grace Wong and Rodrigo Menezes, Pinterest	**Antics, Drift, and Chaos** Lorin Hochstein, Netflix
9:15 am–9:55 am	**Resolving Outages Faster with Better Debugging Strategies** Liz Fong-Jones and Adam Mckaig, Google	**Security as a Service** Wojciech Wojtyniak, Facebook
10:00 am–10:20 am	**Monitoring DNS with Open-Source Solutions** Felipe Espinoza and Javier Bustos, NIC Labs	**Breaking in a New Job as an SRE** Amy Tobey, Tenable
10:20 am–11:00 am	Break with Refreshments, *Sponsored by Catchpoint*	Grand Ballroom Foyer
11:00 am–11:20 am	**"Capacity Prediction" instead of "Capacity Planning": How Uber Uses ML to Accurately Forecast Resource Utilization** Rick Boone, Uber	**Junior Engineers Are Features, Not Bugs** Kate Taggart, HashiCorp
11:25 am–12:05 pm	**Distributed Tracing, Lessons Learned** Gina Maini, Jet.com	**Approaching the Unacceptable Workload Boundary** Baron Schwartz, VividCortex
12:05 pm–1:20 pm	Luncheon, *Sponsored by Salesforce*	Santa Clara Ballroom
1:20 pm–2:00 pm	**Building Shopify's PaaS on Kubernetes** Karan Thukral, Shopify	**Whispers in Chaos: Searching for Weak Signals in Incidents** J. Paul Reed, Release Engineering Approaches
2:05 pm–2:45 pm	**Know Thy Enemy: How to Prioritize and Communicate Risks** Matt Brown, Google	**Architecting a Technical Post Mortem** Will Gallego, Etsy
2:50 pm–3:10 pm	**Automatic Metric Screening for Service Diagnosis** Yu Chen, Baidu	**Your System Has Recovered from an Incident, but Have Your Developers?** Jaime Woo, Shopify
3:10 pm–3:50 pm	Break with Refreshments, *Sponsored by Datadog*	Grand Ballroom Foyer
	Closing Plenary Session Grand Ballroom ABCFGH	
3:50 pm–4:30 pm	**The History of Fire Escapes** Tanya Reilly, Squarespace	
4:30 pm–5:00 pm	**Leaping from Mainframes to AWS: Technology Time Travel in the Government** Andy Brody and James Punteney, U.S. Digital Service	
5:00 pm–5:20 pm	**Operational Excellence in April Fools' Pranks: Being Funny Is Serious Work!** Thomas Limoncelli, Stack Overflow, Inc.	
5:20 pm–5:25 pm	**Closing Remarks** Program Co-Chairs: Kurt Andersen, LinkedIn, and Betsy Beyer, Google	
5:30 pm–6:30 pm	Light Happy Hour, *Sponsored by Box*	Terra Courtyard

Hotel Floor Plans



2nd Floor Level



Lobby Level

Key Themes

- **Containers** are hot; they become a first-class target for SRE work
- Compared to last year, this year was less emphasis on technology, and more on the methodology, process, and foremost **Experience / Lessons Learned**
- Engineering rigid continues: **Statistics & Math** become mainstream
- SRE **concepts start expanding** beyond Availability, for instance Security
- Majority of presentations still from born-on-the-cloud companies, but lots of **Enterprises** in attendance

Containers from scratch

- Workshop by Avishai Ish-Shalom and Nati Cohen
- Python, Linux, and syscalls
- Isolate a process step by step from the “host” system
  - Container
- Good explanations, helpful library
- All Open Source, free on Github
  - https://github.com/Fewbytes/rubber-docker

https://danrl.com/blog/2018/go-contain-me/

Incident Command - What We've Learned from the Fire Department

3 main roles: Incident commander , Tech lead , SME
Plus Scribe, Informed observer, Communications Lead (CL, cf Public Information Officer), Liaison

Split between TL and IC during an incident, different focus (risk to be trapped in one or the other)
- Tech lead leads SMEs to analyze and respond, focuses inward
- IC responsibility for managing the incident response, focuses outward

Practice, practice, practice
- Google “Wheels of misfortune” (scenario, dangle on master, etc)
- Gameday to test capability of org,
- Evaluation exercise to demonstrate that you can handle this
- “Name 3 people”, after 30min tell them
  "these 3 people are no longer available".
  Typically the best 3 people are named.
  See if you can do without them

Tips
- Give your emergency a name
- make first responder TL, not IC
- use a dedicated channel
- show role via display name
- share live links, not screenshots
- don’t dump long text into channel
- use chatbots to automate
- treat verbal as a sidebar
- maintain a status doc
- No freelancing (working on the problem without being part of the organized response)
- beware assumptions about roles
- use CAN reports: Conditions, Actions, Needs
- Use checklists
- Make changes cautiously
- explicitly declare end of incident

Security and SRE

SRE practice to build a performing security organization

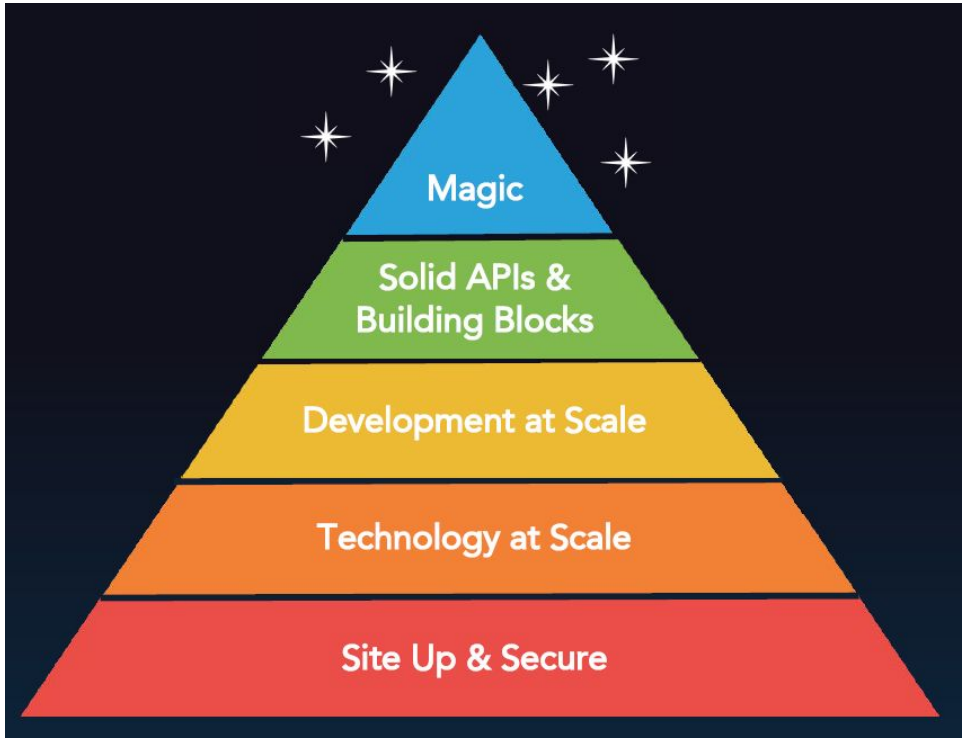
- trust but verify approach (monitoring telemetry)
- embrace the error budget, how quickly can we recover rather than just prevent. Self healing, auto remediation
- inject engineering practices (Dark Launch, Stripping of personally identifiable information, etc)

Benefits ... for security

Your data pipeline is your security lifeblood
Human in the loop is you last resort, not your first option
All security solutions must be scalable and always on

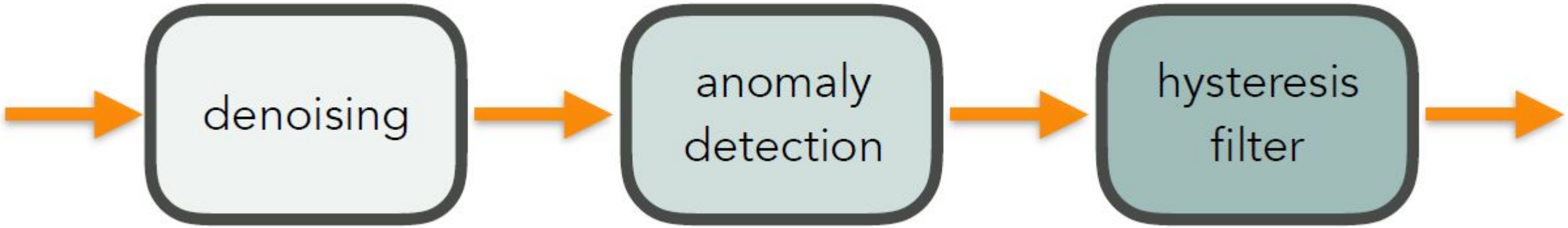
Benefits ... for SRE

Remove single points of security failure like you do for availability
Assume that an attacker can be anywhere in your system or flow
Capture and measure meaningful security telemetry



LinkedIn's Engineering Hierarchy of Needs

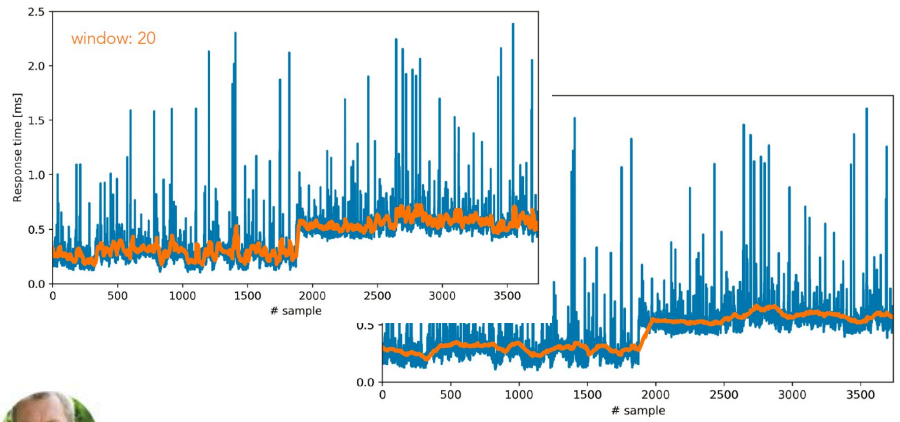
Stable & Accurate Health-Checking of Horizontally-Scaled Services

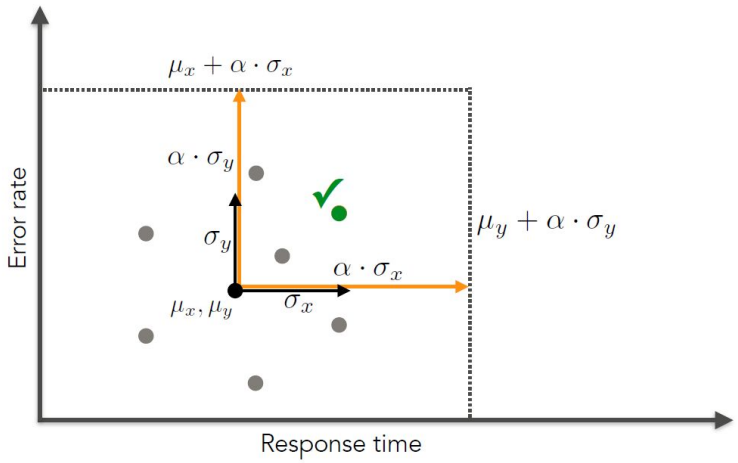


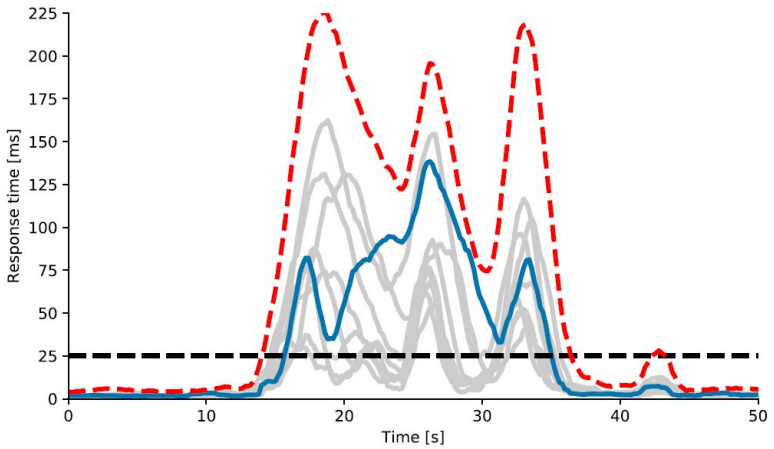
- Moving Average (MA)
- Weighted MA
- Low-pass filtering
- Rolling quantile
- Karhunen-Loève transform
- Subspace projection

- Simple thresholding
- Hypothesis testing
- Conditional entropy
- Distributional thresholding
- Mahalanobis distance
- Kullback-Leibler divergence
- Pattern matching / Clustering

- Sharp hysteresis
- Continuous hysteresis
- Finite State Machine
- Fuzzy logic program







Five Years of Multi-Cloud at PagerDuty

Multi Cloud = having the same product or service spread across multiple cloud provider

Lessons learned

- portability \o/
- teams build Reliability in, because they know they have to run it on different providers
- right sizing is hard (infrastructure across providers can't be matched exactly 1:1)
- deep technical expertise required (LB, databases, applications, HA systems)
- complexity overhead
  - = abstract away providers via Chef (different APIs, different instance sizing)
  - = even less control over the network
- cannot use hosted services (i.e. RDS, document store)

Building a successful SRE in large enterprises - One year later

Recap from 2017 goo.gl/T83gcf
- Reliability is the most important feature
- Our users decide our reliability, not our monitoring / logs
- if you run a platform, then reliability is a partnership
- all popular systems eventually become platforms

Therefore we have to "do SRE " with your customers, too

Lessons Learned

- Enterprise love SRE
- willingness is the thing (single most relevant item)
- Start with the error budget
- Do one application first
- SRE is great for regulated industries
- you don't have to eat it all at once
- Not everyone makes it the whole way - and that's ok

Leaping from Mainframe to AWS: Technology Time Travel in the Government

- Highly relatable (for me)
- U.S. Digital Service
  - Internal “Consultants” helping government agencies to improve digital services
  - Change Agent
- Requesting a VM
  - AWS: *click*
  - GOV: six months! forms, paper, patience
- Launching login.gov for the Trusted Traveler Program (TTP) of CBP
  - 9months
  - Github, OSS, CI-CD pipelines
  - Major bug at launch day -> site taken offline
  - Bug fixed, back online → Celebrated Success! ¯_(ツ)_/¯

Capacity Prediction instead of Capacity Planning

Predicting
- empirical
- repeatable
- scalable
- grounded in data
- expectation of success

Example: choosing the best model, evaluated multiple options:
- rides on trip
- drivers on trip
- drivers online
- completed trips (has highest correlation to CPU consumption)

2 questions
1. Knowledge about how a service or platform behaves under all conditions and demands
2. Knowledge about behavior on future conditions and demands

Steps to perform model:
1. consider what drives your service resource consumption
2. Gather data and build aligned datasets
   if not available right now, begin to ingest and store it
3. Build a predictive model via machine learning methods
   Scikit learn (http://scikit-learn.org/), R Libraries, TensorFlow
5. Store the weights, accuracy scores and metadata
6. Apply the inputs

The History Of Fire Escapes

- History lesson on deadly fire tragedies in and around NYC
  - How contingency plans failed
  - How it influenced politics and regulations
  - How it did not really work out well most of the time
- Entertaining!
  - People invited crazy things to escape fires → Bad tooling :)
  - Automated responses such as sprinklers
  - Failure domains such as interior fire partitions
- What can we learn from history here?
  - Prevent the spark (safety measures)
  - Automatically fix it (like the sprinklers)
  - Contain it (failure domains)
  - If disaster strikes: Have fire escapes ready (rollbacks, tooling, etc.)

Know thy enemy, How to prioritize and communicate risk

what are the risks - **prioritize and communicate**
SLO / **Error Budget** our primary tool for prioritizing our work
Prioritizing Risk: Intuition vs **System** (open to review, feedback, break into details; expose any biases)

3x3 matrix Likelihood (frequent, common, rare) vs. Impact (catastrophic, damaging, minimal)
useful for communication, less useful for prioritization (items tend to be in the middle)

Expected Cost = Probability (Likelihood) * Cost (Impact)
Likelihood
- quantified as MTBF
- Ideally from historical data
- Pragmatically we estimate (ETBF)

Impact
- quantified as MTTR (typically minutes)
- How much of your error budget will the risk consume?
- ETTD (estimated time to detection)
- ETTR (estimated time to resolution)
- % of Users

Risk Name	ETTD (mins)	ETTR (mins)	% Users	ETBF	Bad mins/year
Operator accidentally deletes database	5	480	100	1460	**121**
Bug in new release breaks uncommon request type	1440	30	2	90	**119**
Physical failure of hosting; implement back-up/DR plan	5	720	100	1095	**242**
Overload causes 15% slow requests at peak each day	0	60	15	1	**3287**
No lame-ducking/health-checks; restarts drop requests	0	1	100	7	**52**

What it means to be an effective engineer

Effective engineers:

- **build simple things first**
- Invest in iteration speed
- **prioritize aggressively**
- validate ideas early and often
- work hard and get things done
- build infrastructure for their relationships
- explicitly design their alliances
- explicitly share their assumptions
- **build trust** by making implicit things explicit

Effective engineers work hard and get things done & focus on high-leverage activities & build infrastructure for their relationships

Your System Has Recovered from an Incident, but Have Your Developers?

We make sure that systems are recovered ?
Are we doing the same level of care to the people (ops and dev) ?

Doctors: peer support and counseling can help

Stand-up comedians
Understand how to mentally get back to a better place
- hobbies, people you are about, talk to someone

Olympians face incredibly high-stress situations
What happens when you failed on a global stage?
Self compassion - regulate their stress and emotions

State rumination
- do you find it hard to stop thinking about problem after
- do you have positive or negative thoughts when you reflect
- Does thinking about the problem tend to make the problem worse

Some other interesting sessions

SPARKPOST

The Day the DNS Died

Jeremy Blosser, Principal Operations Engineer
jblosser@sparkpost.com
@SparkPost

https://tinyurl.com/spdnstalk

@SparkPost

Ansible for SRE Teams

•••

Presented by James Meickle
SREcon 2018
March 27, 2018

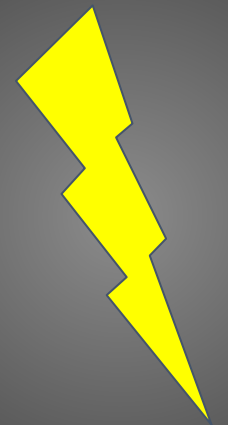
Antics, drift and chaos

Lorin Hochstein
Chaos Team, Netflix
@lhochstein

CHAOS ENGINEERING BOOTCAMP

TAMMY BUTOW, GREMLIN
SRECON AMERICAS 2018

Lightning



Talks

References and Links

All presentations/video/voice available at
https://www.usenix.org/conference/srecon18americas/program

Some summary blogs:
https://michael-kehoe.io/post/srecon-americas-2018-day-1/
https://michael-kehoe.io/post/srecon-americas-2018-day-2/
https://michael-kehoe.io/post/srecon-us-day-3-what-im-seeing/
https://bridgetkromhout.com/speaking/2018/srecon/
https://noidea.dog/blog/srecon-americas-2018-day-1
https://noidea.dog/blog/srecon-americas-2018-day-2
https://noidea.dog/blog/srecon-americas-2018-day-3
http://willgallego.com/2018/04/02/no-seriously-root-cause-is-a-fallacy/

Questions?

Continuous performance profiling in production environments

Dmitri Melikyan

Dmitri is a software engineer and the founder of StackImpact, where he is working on performance profiling and monitoring tools.

Chatham House Rule applies

Tales from On-call

(a.k.a. Featured Post Mortem)

Questions?

Chatham House Rule applies

Networking

everyone!