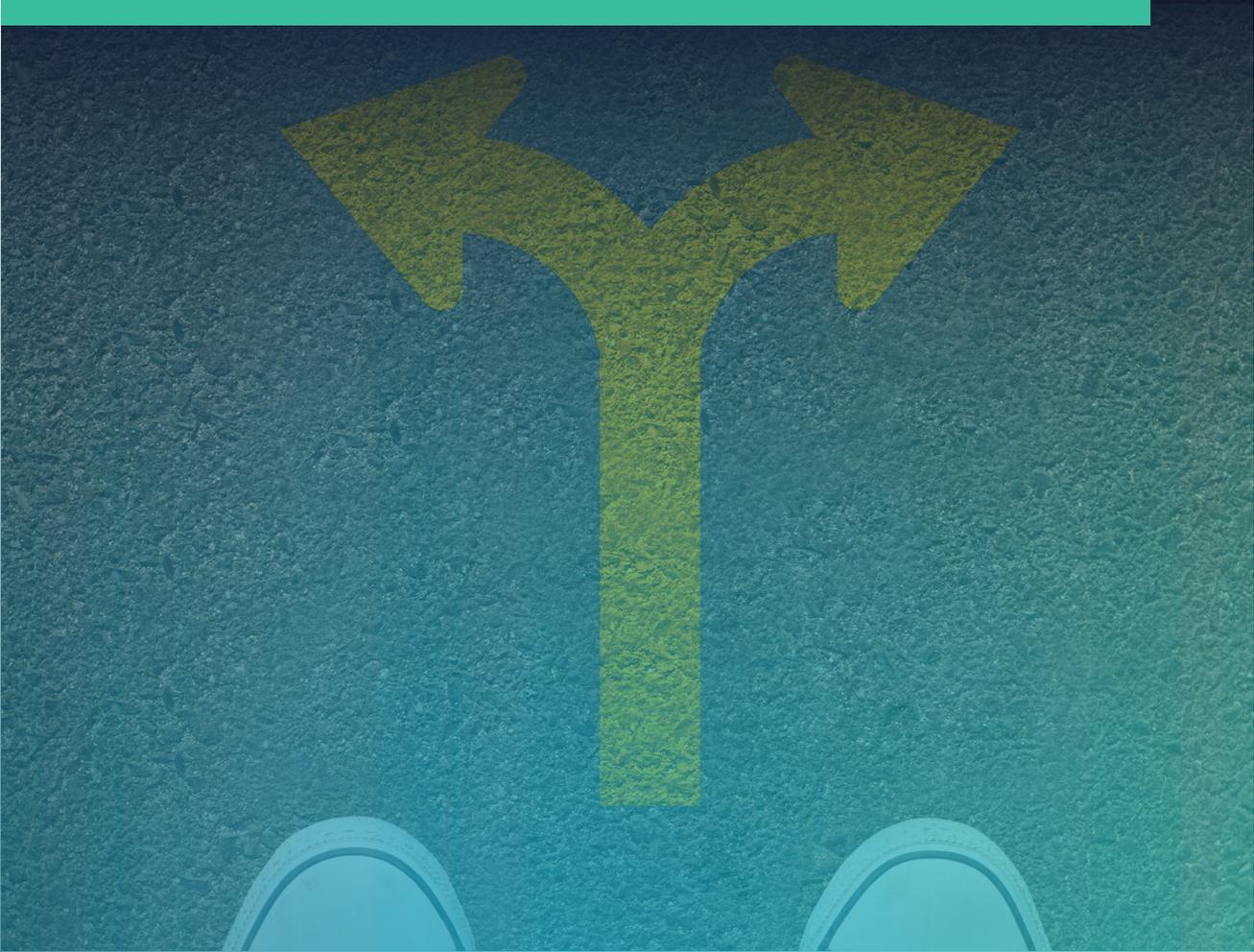


**BLAMELESS**

**SITE RELIABILITY  
ENGINEERING FOR  
BUSINESS CONTINUITY**



# Table of Contents

---

- 2**    **Principle #1:**  
Minimizing SPOFs (single points of failure)
  
- 3**    **Principle #2:**  
Preparing for staffing reductions and changes to continuity plans
  
- 4**    **Principle #3:**  
Creating margin in the system to allow for adaptive capacity
  
- 5**    **Principle #4:**  
Using SLOs (compassionately) to drive prioritization
  
- 6**    **Principle #5:**  
Fostering teamwork and culture in difficult times
  
- 7**    **Principle #6:**  
Adapting to remote work
  
- 8**    **Principle #7:**  
Enabling continuous improvement through work-as-done VS imagined
  
- 9**    **Principle #8:**  
Learning for the long term
  
- 10**   **Conclusion**

# Introduction

With the rise of remote work, distributed teams, black swan events, and more, organizations across the globe have experienced a massive upheaval in how business gets conducted. These unprecedented times have shown us that complex systems fail in ways that are incredibly difficult to predict, and everything is interconnected, from the vast ecosystem of SaaS vendors companies use to the disrupted schedules of daycares and schools. Now, it's more important than ever to deeply analyze your service, processes, and organization's reliability and capacity for change — especially in light of the new age of remote work.

The practice of SRE is uniquely positioned to help organizations embrace business continuity and disaster recovery. While you can't plan for every possible scenario, you can certainly improve preparedness through planning, coordination, and learning from the past. SRE takes these core elements and makes them actionable. In this guide, we've collected best practices from SRE industry leaders to help your organization through any crisis and enter the remote work era more resilient than ever.

## **Many thanks to the industry leaders who shaped the insights in this guide :**

---

**Kurt Andersen** : Senior Staff SRE of product at LinkedIn and co-author of [“What is SRE?”](#).

**David Blank-Edelman** : Cloud advocate at Microsoft, curator and editor of the book [“Seeking SRE,”](#) and cofounder of the SREcon conferences.

**Alex Hidalgo** : SRE at Squarespace and author of the forthcoming book [“Implementing Service Level Objectives”](#) (O'Reilly Media, September 2020).

**Liz Fong-Jones** : Developer Advocate at Honeycomb, labor and ethics organizer, and well-versed conference presenter.

**Dave Rensin** : SRE Director at Google, principal editor of [“The Site Reliability Workbook”](#) and contributor to [“Implementing Service Level Objectives”](#) (O'Reilly Media, September 2020).

**Craig Sebinik** : SRE at Split and co-author of [“What is SRE?”](#) and [“Salt Essentials.”](#)

**Amy Tobey**: Staff SRE at Blameless, experienced SRE and DevOps practitioner, and accomplished conference presenter.

# Principle #1:

## Minimizing SPOFs (single points of failure)

It's important to know who you can count on in a crisis, but what happens if that person gets sick or needs to tend to their children or loved ones? It's critical during times of upheaval that you can function in the event of team members needing to take PTO or sick leave. This means you'll need to determine who your SPOFs (single points of failure) are. If this person needs to take a day, or perhaps even a week off, what context will you miss?

At Google, teams practice an exercise called the "Wheel of Staycation" to ensure that SPOFs are discovered prior to a crisis situation. In [previous talks](#), Dave Rensin has also spoken on how to implement this exercise in your organization. In a nutshell, once a week, a single person from your team is selected to get a 'staycation' where they are unable to communicate with the rest of the team.

- For the person in question, this time can be used for deep project work.
- For everyone else, it's a test. Are there questions you can't get answered, or blockers in your workflow? If this is the case, you have a SPOF to eliminate.

Eliminating SPOFs is important after completing the exercise. To do this, you'll need to track the asks that your staycation team member received in one day. You could do this by setting up a staycation slack channel where all questions you would have asked your teammate are listed. Or you could simply tell your team to ask those questions via email or slack to your stay-cation team mate as usual and have that person create a list when they come back from their stay-cation. Then, the staycation team mate should be required to make sure all the context and key information the team needed is baked into process docs or confluence pages, so the SPOFs are eliminated.

This method should be practiced as a preventative initiative. According to Dave, "There are other things you can do, but the only way you can discover things like expertise SPOFs or information SPOFs is to regularly and routinely exercise them before the emergency shows up."

If you wait too long, you'll only spot SPOFs when it's a true emergency, at which time it'll be too late.

# Principle #2:



## Preparing for staffing reductions and changes to continuity plans

Knowing and eliminating SPOFs is key during any crisis, but it is especially important when staffing reductions come into play. These reductions can be for business or financial reasons, or due to decreased cognitive capacity as people deal with personal matters (anxiety, family needs, health challenges, etc.). Either way, you will need to be able to adapt to these reductions and make adjustments to your continuity plans. Here are three crucial steps to working through this challenge:

### 1. **Revise your on-call schedule:**

Perhaps your organization ran a weekly on-call schedule where you would take turns carrying the pager for a week. Now, with less people on board, your rotation might have gone from once a month to once every two weeks. With the additional burden of keeping services running, the challenges of WFH, and the strain that the current crisis adds to everyone, this might become overwhelming. Being on-call for a full week might be too much of a burden. It's time to talk to people about how they're feeling, take a look at the incident metrics and the time spent per engineer on incidents. If these numbers are higher than normal and team members are reporting higher levels of stress, you'll need to adjust to make the situation easier to bear. Knowing when your team members are overwhelmed and adjusting accordingly is crucial to mitigate burnout and enable long-term success.

### 2. **Know the difference between doing more with less and overworking:**

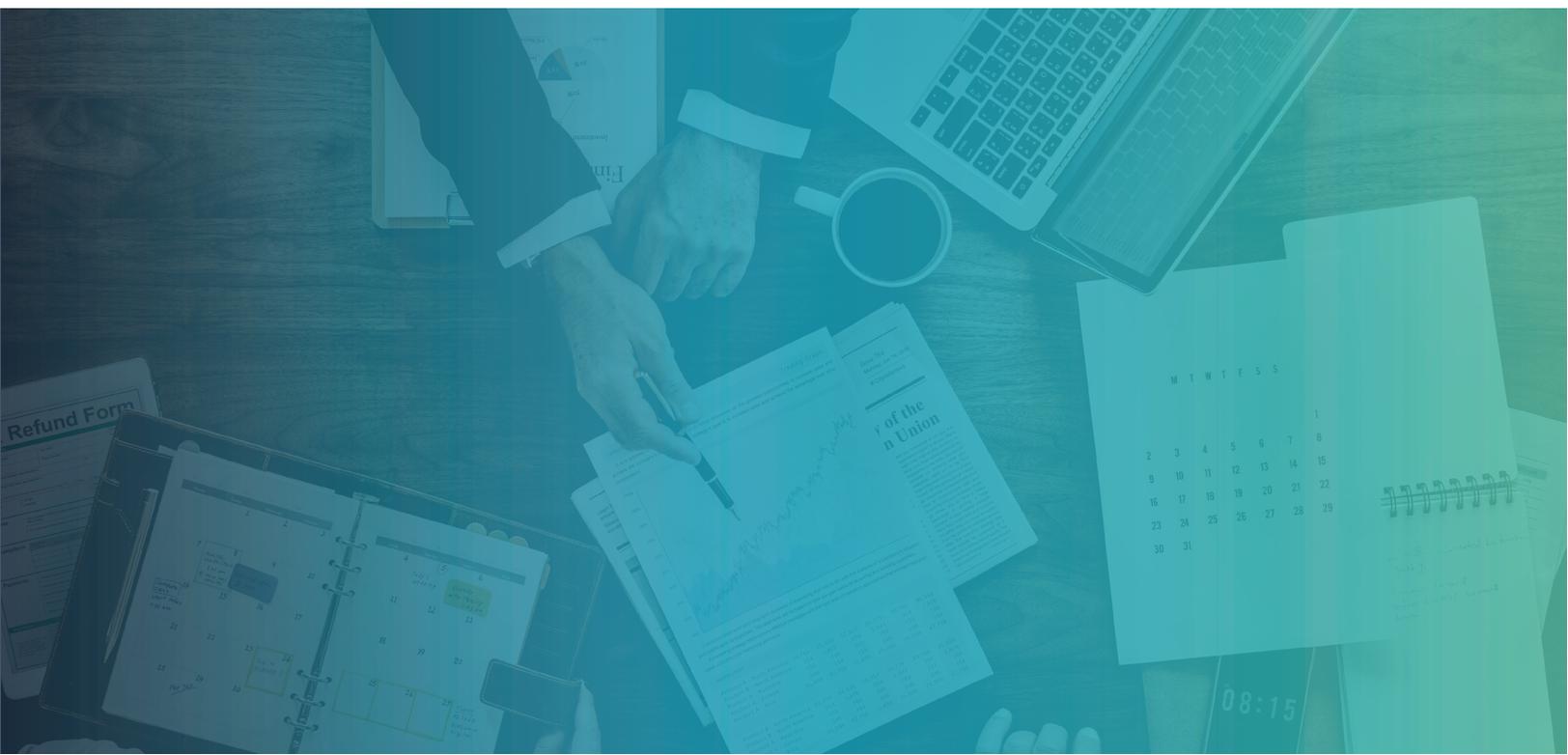
Everyone is talking about doing more with less. Many organizations are focusing on getting down to the 'essentials,' with respect to tooling, work perks, and more. However,

as capacity is reduced, it's important to acknowledge that productivity will take a hit. It's unrealistic to assume that teams can work at the same clip compared to before a black swan event. However, this does not mean sacrificing quality. Strategic prioritization is more important than ever. In times of uncertainty, focus on quality over quantity and speak with leadership about how to adjust goals and metrics for performance during this time.

### 3. **Become comfortable with being uncomfortable:**

During this crisis, all of our previous continuity plans have proven to be insufficient. As unknown unknowns are near impossible to account for, no amount of planning will be able to completely prevent impact, and even the most thorough continuity plans will require adaptability. In short, you'll need to become comfortable with being uncomfortable. As Liz Fong-Jones has shared, "You can't enumerate every single possible thing that's going to go wrong. The playbook strategy is not necessarily going to work super well because you cannot anticipate what the next black swan is going to be. So we have to focus on making our organizations of people more resilient."

While all of these things are easier said than done, they are a good starting point for pivoting in this new reality. But how do you learn and enable adaptability? Let's take a look.



# Principle #3:

## Creating margin in the system to allow for adaptive capacity

Flexibility is crucial right now, but it's difficult to create during a crisis if it didn't exist prior to it. Organizations with less toil built into their processes are set up to succeed better than those with toil-intensive processes. This means it's more important than ever to build some "margin" into our processes in order to remain flexible.

Brain space is at a premium during a crisis. With stress levels mounting, cognitive capacity is diminished. While teams may be too busy putting out fires to focus on automation, it's actually more important than ever to decrease the cognitive load teams are facing. Additionally, automation can help build a buffer between the loss of productivity teams face during this crisis and the need to perform at an increased capacity. This can also increase the likelihood of the 50/50 engineering and toil split, giving you more room for innovation despite the constraints on resources.

Your team will also function better with decreased strain and toil. Richard Cook from [Adaptive Capacity Labs](#) notes that during this crisis, "Social spaces will become more tightly coupled. The effects of events and strains at work will transfer to home and vice versa. The influence of work on home (and home on work!) is usually moderated via social conventions. As stress saps energy it becomes more difficult to maintain boundaries."

When toil becomes overwhelming, teams will lose energy and productivity. Automation helps build margin for your teams to recharge, take time with their families, and deal with this difficult time in a healthy way.

# Principle #4:



## Using SLOs (compassionately) to drive prioritization

Margin can be built into your processes in other ways, too. One useful method is through error budgets and SLOs. SLOs are powerful tools to help align teams on how to prioritize engineering work against new features vs. reliability needs. This shared agreement is even more important now than ever. Richard Cook from [Adaptive Capacity Labs](#) predicts that during this crisis, “Tribalism will increase. Past success in producing a “no blame” and “learning” environment will come under severe pressure as the strain accumulates. Groups that previously worked in harmony may be at odds. Willingness to share productivity across groups will be sapped by the loss of resources and decreased performance.”

As teams experience unprecedented strain and are hit simultaneously with increases in unplanned work as well as reduced capacity, a game of tug of war could erupt. This means that even policies and metrics of success must change during this time. As such, SLOs and error budgets should be established with the team’s context in mind. As Alex said, “The best way to use the concept of an error budget isn’t that you have to actually have measurements, but rather that the concepts behind it give you a different way of thinking about things. And to have good discussions with people with that data and to help you make decisions based upon that.”

He also stressed the importance of revisiting a target whenever necessary: whether that’s due to an incident, change in code base, or a massive black swan event. Relaxing your error budget and compassionately setting flexible SLOs can help facilitate your team’s adaptive capacity, while improving shared prioritization of the work that matters most.

# Principle #5:



## Fostering teamwork and culture in difficult times

With decreased resources, increased stress and cognitive load, and social distancing policies, many teams are under extreme pressure. Without over-communication and special attention paid to organizational culture, teams can become fractured, anxious, or disillusioned. As Liz points out, the most important thing about fostering teamwork during uncertain times is having a people-first mentality: “If it turns out that if your service is going down 1%, it’s probably actually acceptable if it means that people stay home, or that people take their kids to the hospital if they’re getting high fevers.” At the heart of a successful team is shared compassion and trust to back up teammates who might be struggling.

To embrace a people-first culture, collaboration is key. According to Liz, it’s important “to not silo knowledge, which means that you have to have more than one person with a working knowledge of how something works. How do we do things, why do we do things? That’s one kind of redundancy mechanism that we can employ to guard against the possibility of someone going to the hospital and not being available.”

Not only does this create a safeguard against SPOFs, it also encourages communication.

Beyond logistically replicating a team structure while working from home, it’s also important to create remote-first spaces for social engagement. Without coffee breaks, water cooler chats, and group lunches to organically bond with one another, it can be difficult to reinforce bonding and positive organizational culture. Distributed teams can try the following ideas to keep up the team camaraderie:

**1. Virtual water coolers or coffee breaks:**

Setting up 15-30 minute windows where everyone can hop onto Zoom to catch up or engage in casual 1x1s can make a big difference in how your teams gel. [Donut](#) is a great way to keep the remote connections going.

**2. Team lunches:**

This can be your mid-day boost to keep energy high. Scheduling a weekly team lunch can maintain a sense of normalcy with those struggling with social distancing.

**3. Happy hours:**

Wind down the week the right way by setting up a happy hour with your team. Consider enforcing a “no work talk allowed” policy to keep things focused on relationship building.

**4. Team games:**

Teams that play together stay together. Taking the time for friendly competition with trivia, Dungeons & Dragons, or [Jackbox](#) games can lift spirits and get some laughs.



# Principle #6:



## Adapting to remote work

Prior to COVID-19, moving an organization to remote work was a calculated, planned decision rather than a mandate. Teams would have significant advance notice to get equipment, processes, and documentation in place prior to the switch.

However, in a context where organizations need to shift to remote work overnight, how employees adapt to remote work often is sink-or-swim.

According to Liz, teams should ensure their “tooling supports remote workflows for people to not need to be shoulder surfing each other in order to collaborate.” She adds, “The sooner companies can adapt to that reality with solutions that enable people to collaborate without being physically in the same place, the better prepared they’ll be for any scenario, including hiring remote employees, or this crisis dragging on longer. And yes, including people not wanting to go back to the office, too.”

Remote work might become the new way of life for many organizations, possibly even permanently. This will require a heightened level of collaboration. DevOps and SRE best practices can certainly help. While having similar roots, SRE and DevOps are two different things, and remote work functions better with both practices in place. As David Blank-Edelman says, “SRE and DevOps, though complementary modern operations practices, focus on different things. The way I have come to understand it is if the keyword for SRE is reliability, one possible keyword for DevOps would be delivery.” You need both to keep an organization functioning.

In addition to adopting DevOps and SRE best practices, there are 5 additional things organizations can do to help [remote work](#) work for them.

- 1. Get ahead of complications with dry runs.** Before meetings, make sure that you're equipped by testing your camera, microphone, and internet speed for sending and receiving video. Plan a space to take calls in without distractions that has good lighting and a neutral background. The more things you can figure out ahead of time, the more you can focus on the work that matters.
- 2. Have "runbooks" with plenty of plan Bs.** Failure is inevitable. Expect that things will go wrong, and plan as a team for what to try next. If one video service crashes, what's the backup? If there's a miscommunication and someone doesn't join the call, what's the protocol? Once the plans have been decided, ensure people are aligned on it with a team-wide email.
- 3. Communicate effectively.** Set up roles responsible for communicating information to stakeholders, from coworkers to management to customers. Conversations that happen naturally in the office should be codified and scheduled to stop things from falling through the cracks. Be sure to solidify decisions made in emails to create an organized and permanent record.
- 4. Plan to learn.** By recording meetings and Slack conversations, and monitoring KPIs during remote work periods, areas of friction can be identified and smoothed out. Your team can even create a remote work retrospective to analyze exactly where procedures can be improved. Work through a timeline of incidents related to remote work, and discuss where systems can be changed to stop recurrence.
- 5. Stay blameless.** You will run into challenges when transitioning to remote work. Remember: your coworkers have the best intentions, and failures result only from systemic issues that you can (and should!) work together to change.



# Principle #7:



## Enabling continuous improvement through work-as-done VS imagined

Once you've figured out the logistics of working remotely, the actual work comes into focus. But your day-to-day may look significantly different. Prior to moving to remote work, you might have had your toil ratio at 50% or less and spent the majority of your time innovating. Perhaps you were shipping at an exceptional velocity. But now, you may be spending a lot of time on toil-heavy work, reducing capacity to keep up development velocity. Does your manager know this, as well as your team? What are the reasons behind this?

Kurt Andersen explains it this way: "Work-as-done versus work-as-imagined is a concept that comes out of postmortems or incident retrospectives, and a learning-from-incidents mindset where one undertakes an exploratory journey to understand, as best you can, what actually happened as opposed to what you think happened."

An example he raises is the vast amounts of time spent reviewing architecture documents that are essentially pre-implementation designs. He says, "One of the big problems is that those don't ever get updated to reflect the as-built system."

When teams must not only manually dig through documentation, but find that it is not up-to-date, toil increases dramatically.

As Kurt Andersen said, "If you, as a senior director/VP of SRE, have this theoretical concept of SRE as 50% engineering and no more than 50% toil, that's a great construct to have in mind. But if that is your imagination and it doesn't align with the way that work is actually being done by the teams, you are setting your teams up for burnout and frustration and yourself for frustration when you can't align with what is actually hurting the teams and costing them time, effort, and brain space."

Managers or team leads need to stay alert to changes in operation and toil levels, and respond compassionately. As Kurt said, “If you want to implement something to make your SRE teams’ lives better, it’s really important to understand what those lives actually consist of before you try to make them better, not just what you hypothetically want them to be... if that your imagination doesn’t align with the way that work is actually being done by the teams, you are setting your teams up for burnout and frustration and yourself for frustration.”

Communication here is key. Managers and leads need to be ready to listen, even more so when teams are remote. Strong communication is crucial to surface insights that make your teams and organization more resilient.



# Principle #8:



## Learning for the long term

The organizations that are able to learn, pivot, and make improvements are the ones that succeed despite extreme circumstances. Organizations that struggle to embrace learning are less likely to recover from failure. Dave Rensin gave his thoughts on how this will be reflected given the current circumstances: “The ones who don’t learn will self-select for extinction.”

This means making adjustments to culture and processes. Dave also notes, “The most useful thing we can do from here is to ask what principles, what practices, what cultural norms do we want to drive into our companies, so that the next generations don’t have to remember the specifics of this incident in order to get the value of what we learned from it.” But how do we bake continuous learning into our processes?

Craig Sebenik has an example. He spoke with us about how he has learned to vet and onboard new SaaS vendors through a constantly evolving process, iterating on best practices from larger organizations. . According to Craig, “One of the biggest things you can do is try to take lessons from the big companies, not because they’re necessarily doing things right, but because they’ve seen all kinds of weird things. But don’t take them wholesale. Take what they say, figure out what applies to you. Take those pieces and continue to evolve them. Don’t remain static.”

Iteration is key, especially after unexpected events like incidents. Treating crises as learning opportunities can help organizations adapt and prepare for the future of work.





# Conclusion

Richard Cook from [Adaptive Capacity Labs](#) predicts that “The next 6 to 12 months will be a period of intensive learning. The organizational buffeting that is in store for us will reveal much about how well and how poorly our systems can function. It is essential for survival to learn quickly what does and does not work, what new vulnerabilities are cropping up, and how to reshape the tech to meet these demands.”

In times of uncertainty, weaknesses in the system are brought to the fore. The first instinct is to regard them as failures, but they should also be acknowledged as opportunities. What remains constant is the need for growth, learning, compassion, and a focus on the well-being of humans. If we can anchor on these values, we will set up our teams and organizations for success, regardless of how unexpected circumstances take hold in the future.

For more reading on embracing business continuity, check out these resources:

- [Thought Leadership Panel: What is a “real” SRE?](#)
- [How SREs can Embrace Resilience During Crises](#)
- [SRE for Business Continuity in the Face of Uncertainty](#)

## About Blameless

Blameless helps teams drive resiliency across the entire software lifecycle by operationalizing Site Reliability Engineering best practices. As teams share unified context, they are empowered to treat incidents as unplanned investments and opportunities. Leading organizations such as Procore, Under Armour, and Box trust Blameless to collectively embrace a culture of continuous improvement, do more with less, and protect their customers. Headquartered in San Mateo, Calif., Blameless is backed by Lightspeed Venture Partners and Accel. To learn more about Blameless, visit [our website](#) or [schedule a demo](#) today.