SREference

# On-call cheatsheet

## When you are paged

1. **Acknowledge** so it does not escalate, then open the alert's runbook.
2. **Check user impact** on the SLO dashboard: who is affected, how badly, since when.
3. **Declare an incident** if you need a second team, customers can see it, or it is unsolved after an hour. Open one channel, set a provisional severity, name the incident commander.
4. **Post the first update** soon after declaring, and say when the next one will come.
5. **Look at recent changes**: deploys, config, feature flags, traffic.
6. **Mitigate first** with a known-safe rollback, failover or feature switch-off, or by shedding load. Check for data migrations before rolling back, and the replica's health before failing over. Find the cause later; keep the evidence.

**Incident roles and communication →** sreference.com/incident-roles/

## Severity at a glance

| LEVEL | IMPACT | RESPONSE |
|---|---|---|
| **SEV1** | Core function down or badly degraded for most users; data loss or breach | Page 24/7, dedicated IC, update every 30 min, status page |
| **SEV2** | Major degradation for many users, or a core feature down for some | Page 24/7, IC assigned, update every 60 min, status page if customers notice |
| **SEV3** | Limited impact or an easy workaround | Owning team, business hours |
| **SEV4** | No user impact and no imminent risk | Ticket |

A starting point, not a standard. Classify by impact, not cause; when in doubt, go higher. A risk without impact, such as losing the last spare replica, still needs action now.

**Incident severity levels →** sreference.com/incident-severity-levels/

## Escalation and handing over command

- **No acknowledgement?** The page goes to the secondary, then a manager or the wider team, after a fixed timeout such as 5 minutes.
- **Need another team?** Declare an incident and page their on-call. Do not wait to be sure.
- **Handing over command:** brief the next IC, say "You are now the incident commander, OK?", wait for a yes, and announce it in the channel.
- **Long incident?** Rotate responders and the IC every few hours; tired people make mistakes.

**On-call basics →** sreference.com/on-call/
**Incident roles →** sreference.com/incident-roles/

## Internal update

```
[SEV?] [Service]: [what users see]. Update [n],
[HH:MM] UTC

Impact: [who is affected, how, since when]. [What
still works.]
Current state: [what we know, what we are doing].
Next steps: [next action and owner].
Help needed: [none / team or person].
Next update: [HH:MM] UTC or sooner if anything
changes.
IC: [name]   Comms: [name]   Channel: #inc-[id]
```

Send one on schedule even with no news. Say what users see, not component names.

**More incident templates →** sreference.com/incident-roles/#templates

## Burn-rate alerts

```
burn rate = error rate ÷ (1 - SLO)
```

| ACTION | WINDOWS | BURN | ERROR RATE | EMPTY IN |
|---|---|---|---|---|
| **Page** | 1h / 5m | 14.4× | 1.44% | 2d 2h |
| **Page** | 6h / 30m | 6× | 0.6% | 5d |
| **Ticket** | 3d / 6h | 1× | 0.1% | 30d |

Error rate shown for a 99.9% SLO over 30 days; "empty in" is how long the budget lasts at that burn. Both windows must be over the threshold for the alert to fire. A 14.4× or 6× page is at least SEV2.

**Burn-rate alerting →** sreference.com/burn-rate-alerting/

## Downtime allowed

| SLO | PER 30 DAYS | PER DAY |
|---|---|---|
| **99%** | 7h 12m | 14m 24s |
| **99.5%** | 3h 36m | 7m 12s |
| **99.9%** | 43m 12s | 1m 26s |
| **99.95%** | 21m 36s | 43.2s |
| **99.99%** | 4m 19s | 8.64s |

Full outage at steady traffic. A partial outage spends the budget more slowly.

**Availability table →** sreference.com/availability-table/

Sources: Google, Site Reliability Engineering, 1st ed., ch. 14 "Managing Incidents" and ch. 11 "Being On-Call"; The Site Reliability Workbook, ch. 5 "Alerting on SLOs" and ch. 9 "Incident Response". Severity levels and update cadences are starting points to adapt.