Appendix C - Example Incident State Document

- 1. Table of Contents
- 2. Foreword
- 3. Preface
- 4. Part I Introduction
- 5. 1. Introduction
- 6. 2. The Production Environment at Google, from the Viewpoint of an SRE
- 7. Part II Principles
- 8. 3. Embracing Risk
- 9. 4. Service Level Objectives
- 10. <u>5. Eliminating Toil</u>
- 11. <u>6. Monitoring Distributed Systems</u>
- 12. 7. The Evolution of Automation at Google
- 13. <u>8. Release Engineering</u>
- 14. 9. Simplicity
- 15. Part III Practices
- 16. 10. Practical Alerting
- 17. 11. Being On-Call
- 18. 12. Effective Troubleshooting
- 19. 13. Emergency Response
- 20. 14. Managing Incidents
- 21. 15. Postmortem Culture: Learning from Failure
- 22. 16. Tracking Outages
- 23. 17. Testing for Reliability
- 24. 18. Software Engineering in SRE
- 25. 19. Load Balancing at the Frontend
- 26. 20. Load Balancing in the Datacenter
- 27. 21. Handling Overload
- 28. 22. Addressing Cascading Failures
- 29. 23. Managing Critical State: Distributed Consensus for Reliability
- 30. 24. Distributed Periodic Scheduling with Cron
- 31. 25. Data Processing Pipelines
- 32. 26. Data Integrity: What You Read Is What You Wrote
- 33. 27. Reliable Product Launches at Scale
- 34. Part IV Management
- 35. 28. Accelerating SREs to On-Call and Beyond
- 36. 29. Dealing with Interrupts
- 37. 30. Embedding an SRE to Recover from Operational Overload
- 38. 31. Communication and Collaboration in SRE
- 39. 32. The Evolving SRE Engagement Model
- 40. Part V Conclusions
- 41. 33. Lessons Learned from Other Industries
- 42. 34. Conclusion
- 43. Appendix A. Availability Table
- 44. Appendix B. A Collection of Best Practices for Production Services
- 45. Appendix C. Example Incident State Document
- 46. Appendix D. Example Postmortem
- 47. Appendix E. Launch Coordination Checklist
- 48. Appendix F. Example Production Meeting Minutes
- 49. Bibliography

Example Incident State Document

Shakespeare Sonnet++ Overload: 2015-10-21

Incident management info: http://incident-management-cheat-sheet

(Communications lead to keep summary updated.)

Summary: Shakespeare search service in cascading failure due to newly discovered sonnet not in search index.

Status: active, incident #465

Command Post(s): #shakespeare on IRC

Command Hierarchy (all responders)

• Current Incident Commander: jennifer

• Operations lead: docbrown

• Planning lead: jennifer

• Communications lead: jennifer

• Next Incident Commander: to be determined

(Update at least every four hours and at handoff of Comms Lead role.) **Detailed Status** (last updated at 2015-10-21 15:28 UTC by jennifer)

Exit Criteria:

- New sonnet added to Shakespeare search corpus **TODO**
- Within availability (99.99%) and latency (99%ile < 100 ms) SLOs for 30+ minutes **TODO**

TODO list and bugs filed:

- Run MapReduce job to reindex Shakespeare corpus DONE
- Borrow emergency resources to bring up extra capacity DONE
- Enable flux capacitor to balance load between clusters (Bug 5554823) **TODO**

Incident timeline (most recent first: times are in UTC)

- 2015-10-21 15:28 UTC jennifer
 - Increasing serving capacity globally by 2x
- 2015-10-21 15:21 UTC jennifer
 - Directing all traffic to USA-2 sacrificial cluster and draining traffic from other clusters so they can recover from cascading failure while spinning up more tasks
 - MapReduce index job complete, awaiting Bigtable replication to all clusters
- 2015-10-21 15:10 UTC martym
 - Adding new sonnet to Shakespeare corpus and starting index MapReduce
- 2015-10-21 15:04 UTC martym
 - Obtains text of newly discovered sonnet from shakespeare-discuss@ mailing list

- 2015-10-21 15:01 UTC docbrown
 - Incident declared due to cascading failure
- 2015-10-21 14:55 UTC docbrown
 - Pager storm, ManyHttp500s in all clusters

previous

Appendix B - A Collection of Best Practices for Production Services

<u>next</u>

Appendix D - Example Postmortem

Copyright © 2017 Google, Inc. Published by O'Reilly Media, Inc. Licensed under CC BY-NC-ND 4.0