



Perforce Disaster Recovery at Google

Plans and Experiences

Rick Wright
Perforce Administrator
Google, Inc.

About Google



Google's mission is to organize the world's information and make it universally accessible and useful.

- Headquartered in Mountain View, CA
- More than 20,000 employees worldwide
- More than 60 offices in more than 20 countries

Outline

Google

- Perform at Google
- Disaster Planning Overview
- Data Protection
- Procedures
- How to Apply These Principles

Perform at Google

Google

Main Server Hardware

- 4 quad-core processors
- 256GB RAM
- Metadata stored on RAM-based storage system (900GB total available)
- Versioned files stored on network attached storage system

Perforce at Google

Google

Perforce Installation

- 11 total Perforce instances
- 6 active replicas of main server
- 2 standbys for each Perforce instance
- More than 60 proxies worldwide
- Main server has 600GB of metadata
- Main server recently passed changelist 10,000,000

Disaster Planning Overview

Google

What is a "disaster?"

- Can be single machine, single building or whole site
- Hardware failure
- DB corruption
- Major network disruptions
- Localized power outage/fire/etc.
- Full-scale natural disaster that incapacitates a region

Disaster Planning Overview

Google

Infrastructure and Plans

- 3 sets of identical hardware
 - Master server
 - Local standby
 - Remote standby
- Versioned files stored on NAS and mirrored to standby systems
- No single point of failure
- The recovery plan is regularly tested and updated

Data Protection

Google

Protecting the metadata

- Weekly checkpoints
- Back-up of checkpoints and journals to mirrored NAS
- Replication to standby servers
- Tape backup as last resort

Data Protection

Google

Replication to standby servers

- g4jrep (like p4jrep)
- Replication to local standby lags main server by 10-60 seconds
- Replication to remote standby lags main server by 10-20 minutes
- Replication also creates local copy of journal on standby machines

Data Protection

Google

Protecting the versioned files

- Files stored on NAS
- Snapshots of filesystem
- Changes mirrored synchronously to second local NAS
- Changes mirrored asynchronously to remote NAS (10-20 minute delay)
- Tape backup as last resort

Procedures

Google

Quarterly failover

- Fail over from current server to local standby
- Failover document updated every time
 - In Perforce, but synced elsewhere
 - Cut & Paste commands
- Rotated among Perforce admins
- 30 minutes of downtime for main server

Procedures

Google

Annual diasaster recovery test

- Fail over to remote standby
- Perforce fully supported by admins not located at company HQ in Mountain View, CA
- May need to do failovers with only minutes or a couple hours of warning

Procedures

Google

Communication

- Regular communication with users for all major and most minor outages
- "Perforce Weather Report" section of internal website for quick updates
- Team uses internal IRC chat
- Documentation about who else to contact

Procedures

Google

Checks and Monitoring

- Weekly checkpoint recovery
- Monitor connectivity to server
- Monitor replication lag
- Monitor p4d processes
- Backup check
- Monitor other things like disk space
- Most monitors will automatically page the on-call Perforce admin

Procedures

Google

Real disasters

- Differences between real and simulated disasters
- Real disasters can have data loss of up to 20 minutes
- We have had real disasters, but fortunately without data loss

How to apply these principles

Google

- Have a plan
- Determine your risk tolerance
 - How long can you be down?
 - How long for a checkpoint restore?
 - How much data can you lose?
- *Do it once*
 - Better yet, do it regularly
- Decide that planning for a disaster is important

Conclusion

Google

- All sites are different, but all should have a plan
- Continually improve your plan
- Plan for disasters of all sizes

Q & A

Google

Q & A
Thank You
rickw@google.com