SysTutorials

Colossus: Successor to the Google File System (GFS)

By Eric Ma | In Storage systems, Systems | Updated on Oct 7, 2018

Colossus is the successor to the Google File System (GFS) as mentioned in the paper on Spanner at OSDI 2012. Colossus is also used by spanner to store its tablets. The information about Colossus is slim compared with GFS which is published in the paper at SOSP 2003. There is still some information about Colossus on the Web. Here, I list some of them.

Storage Architecture and Challenges

On Faculty Summit, July 29, 2010, by Andrew Fikes, Principal Engineer.

The slides. Some interesting points:

- Storage Software: Colossus
  - Next-generation cluster-level file system
  - Automatically sharded metadata layer
  - Data typically written using Reed-Solomon (1.5x)
  - Client-driven replication, encoding and replication
  - Metadata space has enabled availability analyses
- Why Reed-Solomon?
  - Cost. Especially w/ cross cluster replication.
  - Field data and simulations show improved MTTF
  - More flexible cost vs. availability choices

GFS: Evolution on Fast-forward

An interview with Google's Sean Quinlan by the Association for Computer Machinery (ACM).

View the interview.

Some important info:

- "We also ended up doing what we call a "multi-cell" approach, which basically made it possible to put multiple GFS masters on top of a pool of chunkservers."
- "We also have something we called Name Spaces, which are just a very static way of partitioning a namespace that people can use to hide all of this from the actual application." … "a namespace file describes"
- "The distributed master certainly allows you to grow file counts, in line with the number of machines you're willing to throw at it." … "Our distributed master system that will provide

> for 1-MB files is essentially a whole new design. That way, we can aim for something on the order of 100 million files per master. You can also have hundreds of masters."
> - BigTable "as one of the major adaptations made along the way to help keep GFS viable in the face of rapid and widespread change."

Google File System II: Dawn of the Multiplying Master Nodes Comments on GFS2 (colossus)

by Cade Metz in San Francisco.

The article and some excerpt.

Related posts:

- Reference: Special HTML Characters
- A Simple CPU and Memory Performance Test of Xen Dom0 and DomU
- Dynamics CRM Consultants Explain How To Filter Sub Branch Lookup Field Value in CRM
- New Linux Kernel 5.0: Features and Improvements
- Windows 7 64-bit fails to install on VirtualBox / Linux with status code 0xc0000225
- Vim Tutorial for Beginners: vimtutor

Eric Ma

Eric is a systems guy. Eric is interested in building high-performance and scalable distributed systems and related technologies. The views or opinions expressed here are solely Eric's own and do not necessarily represent those of any third parties.

All 875 posts by Eric Ma

6 comments:

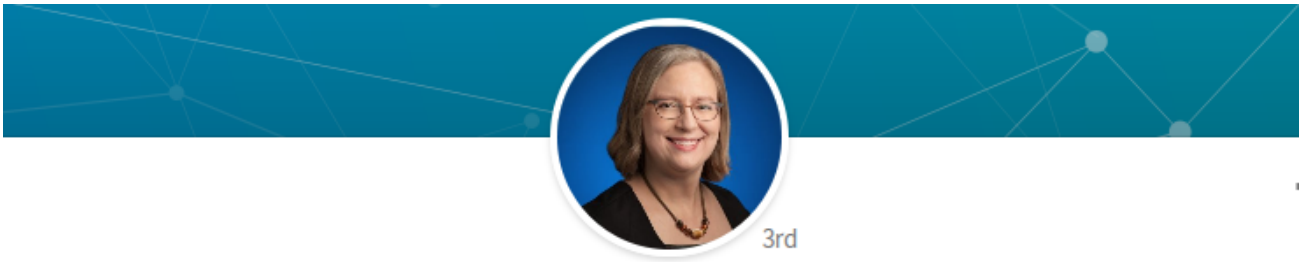
Pingback: Large-scale Data Storage and Processing System in Datacenters | Cloud Computing

Pingback: Google and evolution of big-data | Useful Stuff

Pingback: Large-scale Data Storage and Processing System in Datacenters - SysTutorials

Clara says:
Feb 28, 2017 at 8:10 pm

This page is linked by a director from Google https://www.linkedin.com/in/mbinde/ for reference to Colossus !



3rd

Melissa Binde

Director, Engineering

Google • Swarthmore College

Mountain View, California • 500+

Send InMail Connect

Watch our SRE Q&A (1 June 2016): https://www.youtube.com/watch?v=x-J0IRFF5b8

Experience

Director, Storage, Analytics and Cloud ML SRE

Google

Jan 2015 – Present • 2 yrs 2 mos • Mountain View, CA

I run the SRE team for our Storage, Analytics, and Machine Learning products:

Google Cloud Storage Products: https://cloud.google.com/products/storage/
Google Cloud Analytics Products: https://cloud.google.com/products/big-data/
Google Cloud Machine Learning Products: https://cloud.google.com/products/machine-learning/

We've also published many papers on our internal products, including:

Colossus: http://www.highlyscalablesystems.com/3202/colossus-successor-to-google-file-system-gfs/
Spanner: http://research.google.com/archive/spanner.html
Bigtable: http://research.google.com/archive/bigtable.html
Megastore: http://research.google.com/pubs/pub36971.html
F1: http://research.google.com/pubs/pub41344.html
Availability in Globally Distributed Storage Systems: http://research.google.com/pubs/pub36737.html

Pingback: Separation of storage and compute in BigQuery – Cloud Data Architect

Eric Z Ma says:
Jul 25, 2018 at 6:43 pm

Glad this article is referenced by “Google Cloud Platform” article “Opinionated Managed Storage Engine” at https://medium.com/google-cloud/the-12-components-of-google-bigquery-c2b49829a7c7 – “Colossus is Google’s successor to GFS”.