

Getting Technical with Instagram

By Howard Terry

February 19, 2013

The very popular app, Instagram, has grown to be one of the most, if not the best, photo sharing and social network that is used today by millions of people. If you don't know by now, Instagram allows people to share pictures with your followers very easy and quick. It became very popular when launched in October of 2010. As of now, Instagram has as many as 100 million current users.

Some may wonder how Instagram actually works. For the technical heads, this article will give you a quick run-down on how Instagram actually runs. Instagram thought it would be fun to give a sense of everything that powers it.

When choosing a system, Instagram goes by a few core principles; keep it simple, don't reinvent the wheel, and go with proven and solid technologies when you can. Engineers from Instagram gave us a sense of all the systems that powers Instagram at a high level. For hosting, Instagram is run by Ubuntu Linux 11.04 on Amazon EC2. Being that previous versions of Ubuntu had all kinds of unpredictable freezing episodes due to high traffic, Ubuntu Linux 11.04 has kept Instagram running just fine. The needs for Instagram are still evolving, but as of now, Instagram only uses 3 engineers.

Instagram makes it easy and quick for us to upload our pictures on the server. Every time we send our picture, it had to go through load balance machines. Instagram goes through Amazon's Elastic Load Balancer that has 3 NGINX instances behind it. Anytime you make a request, it had to go through this and once it does the application server handles our request. This application server is called Django which is on Amazon High-CPU Extra-Large machines. Gunicorn is used as their WSGI server, which is a simple and universal interface between web servers and web applications. The choice for using Gunicorn was because it was much easier to configure and less CPU-intensive.

Everything that is sent to Instagram had to be stored in data storage. Most of the data that is sent to Instagram lives in PostgreSQL and the main shard cluster involves 12 Quadruple Extra-Large memory instances. All of PostgreSQL instances are run in a master-replica setup using Streaming Replication and EBS snapshotting to take frequent backups for their systems. Databases have to be connected to the app servers, and for Instagram PgBouncer is used to connect the pools to PostgreSQL. And for the uses of the news feed, activity feed, and the sessions systems, it's being powered by Redis. And for users to share their Instagram photos to Twitter and Facebook, it is pushed through Gearman, which is a task queue system written by Daniel

For us who use Instagram, we see how easy and quick it is to share our photos with just a tap of a button. But do we every think how does everything run? It's amazing how everything that was talked about is needed to run an app as Instagram that seems like it's very simple to use. There is a lot that goes on behind the scenes in order for us to enjoy this photo sharing/social network. I hope this article gave you a better sense of how it runs because for me it's very appealing what is all needed for Instagram to work and become so popular by millions of people.