The Three Ages of Google - Batch, Warehouse, Instant

Monday, August 29, 2011 at 9:22AM

Todd Hoff in Strategy, google

The world has changed. And some things that should not have been forgotten, were lost. I found these words from the Lord of the Rings echoing in my head as I listened to a fascinating presentation by Luiz André Barroso, Distinguished Engineer at Google, concerning Google's legendary past, golden present, and apocryphal future. His talk, Warehouse-Scale Computing: Entering the Teenage Decade, was given at the Federated Computing Research Conference. Luiz clearly knows his stuff and was early at Google, so he has a deep and penetrating perspective on the technology. There's much to learn from, think about, and build.

Lord of the Rings applies at two levels. At the change level, Middle Earth went through three ages. While listening to Luiz talk, it seems so has Google: Batch (indexes calculated every month), Warehouse (the datacenter is the computer), and Instant (make it all real-time). At the "what was forgot" level, in the Instant Age section of the talk, a common theme was the challenge of making low latency systems on top of commodity systems. These are issues very common in the real-time area and it struck me that these were the things that should not have been forgotten.

What is completely new, however, is the combining of Warehouse + Instant, and that's where the opportunities and the future is to be found-the Fourth Age.

The First Age - The Age of Batch

The time is 2003. The web is still young and HTML is still page oriented. Ajax has been invented, but is still awaiting early killer apps like Google Maps and a killer marketing strategy, a catchy brand name like Ajax.

Google is batch oriented. They crawled the web every month (every month!), built a search index, and answered queries. Google was largely read-only, which is pretty easy to scale. This is still probably the model most people have in their minds eye about how Google works.

Google was still unsophisticated in their hardware. They built racks in colo spaces, bought fans from Walmart and cable trays from Home Depot.

It's quaint to think that all of Google's hardware and software architecture could be described in seven pages: Web Search for a Planet: The Google Cluster Architecture by Luiz Barroso, Jeffrey Dean, and Urs Hoelzle. That would quickly change.

The Second Age - The Age of the Warehouse

The time is 2005. Things move fast on the Internet. The Internet has happened, it has become pervasive, higher speed, and interactive. Google is building their own datacenters and becoming more sophisticated at every level. Iconic systems like BigTable are in production.

About this time Google realized they were building something qualitatively different than had come before, something we now think of, more or less, as cloud computing. Amazon's EC2 launched in 2006. Another paper, this one is really a book, summarizes what they were trying to do: The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines by Luiz André Barroso and Urs Hölzle. Note the jump from 7 pages to book size and note that it was published in 2009, 4 years after they were implementing the vision. To learn what Google is really up to now will probably take an encyclopedia and come out in a few years, after they are on to the next thing.

The fundamental insight in this age is that the **datacenter is the computer**. You may recall that in the late 1980s Sun's John Gage hailed "the network is the computer." The differences are interesting to ponder. When the network was the computer we created client-server architectures that appeared to the outside world as a single application, but in reality they were made of individual nodes connected by a network. Wharehouse-scale Computing (WSC) moves up stack, it considers computer resources, to be as much as possible, fungible, that is they are interchangeable and location independent. Individual computers lose identity and become just a part of a service. Sun later had their own grid network, but I don't think they ever had this full on WSC vision.

Warehouse scale machines are different . They are not made up of separate computers. Applications are not designed to run single machines, but to run Internet services on a datacenter full of machines. What matters is the aggregate performance of the entire system.

The WSC club is not a big one. Luiz says you might have warehouse scale computer if you get paged in the middle of the night because you only have petabytes of data of storage left. With cloud computing

The Third Age - The Age of Instant

The time is now. There's no encyclopedia yet on how the Age of Instant works because it is still being developed. But because Google is quite open, we do get clues: Google's Colossus Makes Search Real-Time By Dumping MapReduce; Large-Scale Incremental Processing Using Distributed Transactions And Notifications; Tree Distribution Of Requests And Responses; Google Megastore - 3 Billion Writes and 20 Billion Read Transactions Daily; and so much more I didn't cover or only referenced.

Google's Instant Search Results is a crude example Luiz says of what the future will hold. This is the feature that when you type in a letter in the search box you instantly get back query results. This means for every query 5 or 6 queries are executed. You can imagine the infrastructure this must take.

The flip side of search is content indexing. The month long indexing runs are long gone. The Internet is now a giant event monster feeding Google with new content to index continuously and immediately. It is astonishing how quickly content is indexed now. That's a revolution in architecture.

Luiz thinks in the next few years the level of interactivity, insight and background information the system will have to help you, will dwarf what there is in Instant Search. If you want to know why Google is so insistent on using Real Names in Google+, this is why.

Luiz explains this change having 4 drivers:

> Applications - instantaneous , personalized, contextual
> Scale - increased attention to latency tail
> Efficiency - driving utilization up, and energy/water usage down

Hardware Trends - non-volatile storage, multi-cores, fast networks

Instant in the context of Warehouse computing is a massive engineering challenge. It's a hard thing to treat a datacenter as a computer and it's a hard thing to provide instant indexing and instant results, to provide instant in a warehouse scale computer is an entirely new level of challenge. This challenge is what the second half of his talk covers.

The problem is we aren't meeting this challenge. Our infrastructure is broken. Datacenters have the diameter of a microsecond, yet we are still using entire stacks designed for WANs. Real-time requires low and bounded latencies and our stacks can't provide low latency at scale. We need to fix this problem and towards this end Luiz sets out a research agenda, targeting problems that need to be solved:

> Rethink IO software stack. An OS that makes scheduling decisions 10s of msecs is incompatible with IO devices that response in microseconds.
> Revisit operating systems scheduling.
> Rethink threading models.
> Re-read 1990's fast messaging papers.
> Make IO design a higher priority. Not just NICs and RDMA, consider CPU design and memory systems.

"The fun starts now" Luiz says, these are still very early days, predicting this will be the:

> Decade of resource efficiency
> Decade of IO
> Decade of low latency (and low tail latency)
> Decade of Warehouse-scale disaggregation, making resources available outside of just one machine, not just a single rack, but all

machines.

This is a great talk, very informative, and very inspiring. Well worth watching. We'll talk more about specific technical points in later articles, but his sets the stage not just for Google, but for the rest of the industry as well.

Related Articles

Web Search for a Planet: The Google Cluster Architecture by Luiz Barroso, Jeffrey Dean, and Urs Hoelzle

The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines by Luiz André Barroso, Urs Hölzle

It's time for low latency by Stephen M. Rumble, Diego Ongaro, Ryan Stutsman, Mendel Rosenblum, and John K. Ousterhout

---

Article originally appeared on High Scalability (http://highscalability.com/).

See website for complete article licensing information.