



## Couchbase at LinkedIn

LinkedIn, the world's largest professional social media network, was quickly outgrowing its storage systems. The kind of traffic the global site was experiencing was growing exponentially, and their existing database architecture simply could not handle it. Scaling became a huge, important component of their backend system. The site reliability engineering (SRE) team soon realized they would need to find a backend solution that could scale to meet its massively growing requirements, while also providing a zero downtime and high, consistent performance.

They chose Couchbase.

### LinkedIn by the Numbers

- 300+ million members
- 3+ million company pages and 2+ million groups with over 70+ millions follows
- 60 clusters of Couchbase running at LinkedIn
- Up to 100 servers in a cluster, both single and multi-tenant clusters
- 10 separate services with one or more clusters in 3 data centers
- 1K queries per second to the servers with latency in the 99<sup>th</sup> percentile

## LinkedIn's Caching Challenge

With billions of page hits per day and global, round the clock traffic, LinkedIn is the largest professional social media network in the world. Part of running a service 24 hours a day, 7 days of week is that it's crucial to manage and monitor your site at all times.

LinkedIn's legacy storage systems proved difficult to keep up with such demands. A strong caching technology was crucial for LinkedIn to provide the performance its users required, and replicating entire databases into memory was not scalable. While the team built infrastructures around storage systems to deal with the custom caches, Memcached and EHCACHE, it became increasingly difficult to manage. Reliable data replication was impossible, and caching invalidations were prevalent.

Increasing the number of servers to support its growth made caches on individual servers untenable. Even using memcached was limiting with the tens of millions of rows and seed data required. Caches would become cold when the server rebooted and errors would come up on the site as the database got overloaded. It was difficult to move data across data centers.

LinkedIn required a database able to support the hundreds of millions of global users that could scale, provide the high performance of a cache, manage/monitor, and always be on.

## Enter Couchbase

After evaluating MongoDB, Redis and others to replace Memcached, the LinkedIn team selected Couchbase.

Memcached, a widely used technology at LinkedIn, was built-in with Couchbase (the memcached protocol). Couchbase had replication and cluster expansion abilities, both crucial for LinkedIn's database management.

Couchbase also offered other key capabilities:

- Memory latency for operations
- Asynchronous writes to disk
- Partitioning
- Warm caches

It was easy and quick for the SRE to adopt Couchbase within its architecture. All that was required was an implementation to their Java interface. Management and monitoring around clusters was crucial for LinkedIn, and Couchbase's rich set of tooling allowed them to do easily do just that.

Couchbase acted like the team expected it to act, something the team had struggled with in the past with other database technologies. This was extremely important to LinkedIn as they made the switch.

## Couchbase at LinkedIn today

As they continue to ingratiate Couchbase within its architecture, the LinkedIn SRE team has built processes and tools around Couchbase to make it easy to transition from Memcached to Couchbase.

With its elastic scalability, consistent high performance, and always-on availability, Couchbase Server was the easy choice for LinkedIn.

You can see the full LinkedIn video presentation at <https://www.youtube.com/watch?v=UyGlgzc7i4w>

## About LinkedIn

Our mission is simple: connect the world's professionals to make them more productive and successful. When you join LinkedIn, you get access to people, jobs, news, updates, and insights that help you be great at what you do.