



“Rather than experience the pain [of expanding a sharded architecture], we would rather just go where we know scalability is better.”

Dan Pollack
Senior Operations Architect
AOL

Contact Us
+1 415.501.9560
info@clustrix.com

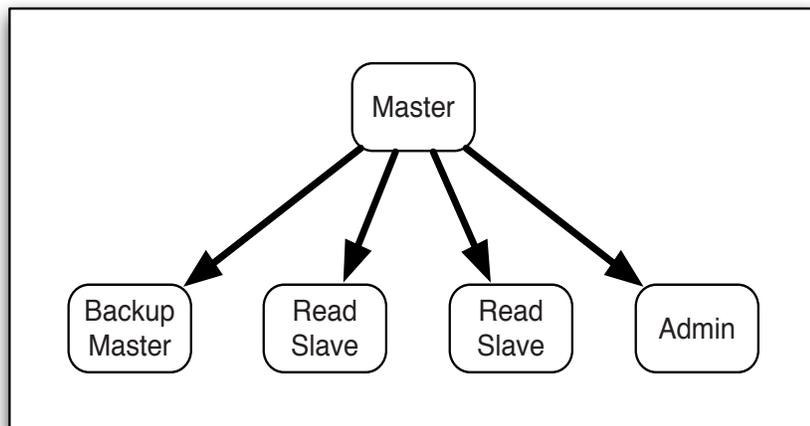
ABOUT AOL

America Online is a leading web services company offering premium and niche content sites, world-class tools and platforms.

The AOL HSS storage system enables simple web-scale storage across a wide range of internal applications. It provides a fault tolerant, high performance, high availability, and inexpensive storage infrastructure for AOL’s web properties. HSS base operations are “post file” and “get file.” HSS stores multiple copies of each file automatically for fault tolerance and performance. Files are spread across multiple data centers for site level redundancy. HSS scales to any number of objects and any aggregate performance target.

PRE-CLUSTRIX

AOL’s old HSS database infrastructure consisted of four shards spread across ten machines. Each shard consisted of a production master, a backup master, a set of read slaves, and an admin database. Implemented as separate MySQL instances the master received all updates from operations such as adding files, deleting files, changing file protection, and cleanup operations. The read slaves handled lookups for “get file” operations in production while the admin database handled housekeeping operations like scrubbing the storage system for improperly protected and orphaned files.



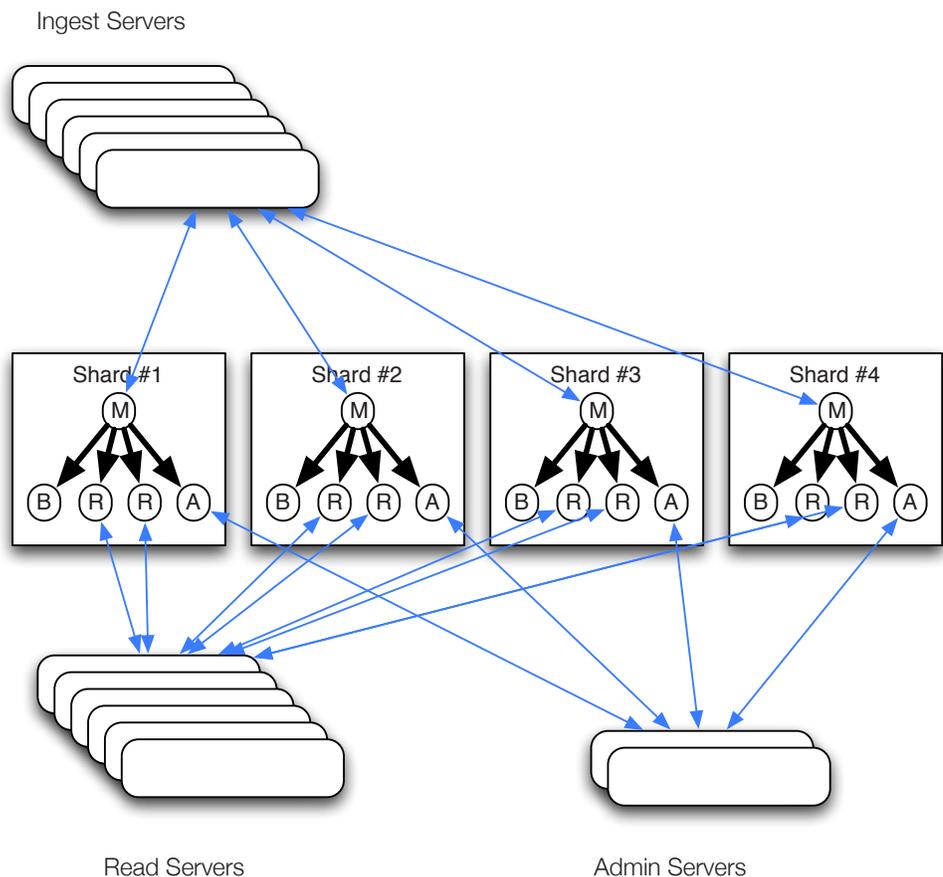


“We don’t think we can go to production with this [non-Clustrix] setup with 3B rows in the database.”

Dan Pollack
Senior Operations Architect
AOL

Contact Us
+1 415.501.9560
info@clustrix.com

This architecture caused some interesting corner cases. As a robustness measure, apps that wrote files into HSS would check for accessibility before assuming success. If the app did not find the file, it would ingest again. Replication lag would delay this process. When ingesting a lot of files, the servers would get stuck in a loop of continuously uploading the same files, with the read slaves falling farther and farther behind. By eliminating read slaves, Clustrix eliminates this replication lag. Also, failures were particularly labor intensive in the pre-Clustrix architecture, and a hardware failure in one master would result in hours of juggling machines and replication configurations to fully recover. The built-in fault tolerance of Clustrix eliminates this concern.



CONSOLIDATING TO A SINGLE CLUSTRIX DATABASE

Clustrix was first applied to the admin replicas. AOL coalesced data from all the shards into a single database using the Clustrix multi-master replication. By taking replication feeds from each of the shards, AOL could consolidate data into one place for the administrative queries.

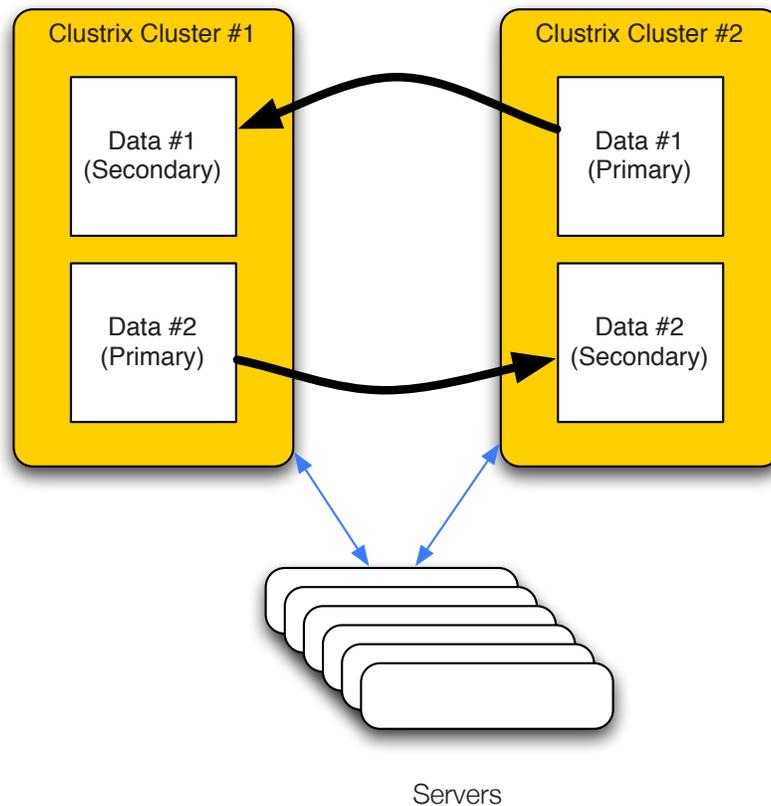


Clustrix dramatically sped up the slowest queries from over one hour on MySQL to less than one second on Clustrix. With Clustrix, AOL was able to scrub the entire HSS data set in 1.5 days, a timeframe not previously possible.

In the second stage of deployment, AOL replaced the master and read slaves with Clustrix. With the limitless scalability of the Clustrix database, AOL no longer has to maintain read slaves in the architecture. All accesses go to the same cluster and replication lag is a non-issue. For site-level redundancy, AOL deployed two clusters in two data centers. Half the data is writable in each site.



Contact Us
+1 415.501.9560
info@clustrix.com



With the Clustrix deployment, each data center has a fault tolerant, scalable database that grows with the data set and increasing load. All servers attach to the same databases and see a consistent view of the data. With the Clustrix architecture, AOL can now expand the HSS deployment throughout the company, consolidate servers, and save money and time. The AOL infrastructure can now easily scale by over an order of magnitude to incorporate new data sets, like the entire MapQuest tile set.