Announcing etcd 3.4

By *Roy Schestowitz*
Created *30/08/2019 - 10:12pm*
Submitted by Roy Schestowitz on Friday 30th of August 2019 10:12:18 PM Filed under Server [1] OSS [2]

☒

etcd v3.4 includes a number of performance improvements for large scale Kubernetes workloads.

In particular, etcd experienced performance issues with a large number of concurrent read transactions even when there is no write (e.g. ?read-only range request ... took too long to execute?). Previously, the storage backend commit operation on pending writes blocks incoming read transactions, even when there was no pending write. Now, the commit does not block reads which improve long-running read transaction performance.

We further made backend read transactions fully concurrent. Previously, ongoing long-running read transactions block writes and upcoming reads. With this change, write throughput is increased by 70% and P99 write latency is reduced by 90% in the presence of long-running reads. We also ran Kubernetes 5000-node scalability test on GCE with this change and observed similar improvements. For example, in the very beginning of the test where there are a lot of long-running ?LIST pods?, the P99 latency of ?POST clusterrolebindings? is reduced by 97.4%. This non-blocking read transaction is now used for compaction, which, combined with the reduced compaction batch size, reduces the P99 server request latency during compaction.

More improvements have been made to lease storage. We enhanced lease expire/revoke performance by storing lease objects more efficiently, and made lease look-up operation non-blocking with current lease grant/revoke operation. And etcd v3.4 introduces lease checkpoint as an experimental feature to persist remaining time-to-live values through consensus. This ensures short-lived lease objects are not auto-renewed after leadership election. This also prevents lease object pile-up when the time-to-live value is relatively large (e.g. 1-hour TTL never expired in Kubernetes use case).

[3]

Server OSS

Source URL: http://www.tuxmachines.org/node/127572

Links:

[1] http://www.tuxmachines.org/taxonomy/term/147

[2] http://www.tuxmachines.org/taxonomy/term/72

[3] https://kubernetes.io/blog/2019/08/30/announcing-etcd-3-4/