

hazelcast **IMDG**

Hot Restart Store

Feature of IMDG Enterprise HD

Hazelcast Hot Restart Store at a Glance

In addition to the basic reliability options (write-through and write-behind), Hazelcast Enterprise HD IMDG has introduced a new high-performance option. This feature is known as Hot Restart Store. This persistence feature provides fast cluster restarts by storing the states of the cluster members on disk in a format specially designed for restart performance and to work in concert with SSDs.

Key Features

- Dramatically reduces restart time for huge caches by providing very fast data reload
- Ultra high-performance persistence store, optimized for SSD & mirrored in native memory
- 502,254 writes per second per SSD node
- Each node has its own independent store, so enables linear scaling across the cluster
- Restarts in 76 seconds with 100GB per node (so 76s to reload 1TB on a 10 node cluster)
- Data entirely loaded into RAM on reload – Hazelcast always operates at in-memory speed
- Persistence can be configured per data structure for JCache, Map, Web Sessions and Hibernate
- Supports partial restart which allows you to restart the cluster even with an incomplete member set
- Hot backup via Management Center or Java/REST API allows you to take a snapshot of the Hot Restart Store at a certain point in time to perform testing, QA or reproduce an issue on production data
- Supports real world operational scenarios such as: moving data to another server, replacing a server, and allowing different servers (IP addresses) to take part in the restart phase

Setup					Reading		Writing	
Storage Type	Data Size	Value Size	Level of par*lism	Partition Threads	Restart Time	Read Throughput	User Threads	Write Throughput
SSD	100 GB	1007 B	4	8	82 sec	1.4 GB/s	16	425,000 ops/sec
SSD	50 GB	495 B	4	40	49 sec	1.2 GB/s	40	400,000 ops/sec
SSD	25 GB	239 B	8	40	39 sec	0.80 GB/s	40	502,254 ops/sec

Whether the restart is a planned shutdown (including rolling upgrades across a cluster) or a sudden cluster-wide crash (e.g. power outage or total loss of a data center), Hot Restart Store allows full recovery to the previous state of configuration and cluster data. The initial release of Hot Restart Store supports the IMap and JCache interfaces, as well as Web Sessions and Hibernate. Further data structures will be delivered in subsequent releases. It is strongly recommended that Hot Restart Store be used only on SSD drives, as magnetic HDD drives are much slower (typically ~10x). This difference in capability will typically result in restart and recovery times that are not compatible with high-performance use cases.

Hot Restart Store works by each node controlling its own local snapshot. This provides linear scaling across the cluster, provided one of two best practice approaches is followed:

1. That each node is not configured with more stored data than the size of the node’s JVM heap. This is the preferred approach for on-heap storage.
2. In the case of larger stores, High-Density Memory Store should be used. This uses off-heap memory storage, and allows the creation of much larger nodes. This solves some of the limitations of garbage-collected heaps – 100GB JVM heaps are not practical to collect. This has clear best practice implications for the design of high-performance clusters – each node should have a large amount of physical RAM. In the on-heap case, we should use a large JVM heap (not larger than physical RAM) and the Hot Restart Store size configured to be the same as the heap size. In the off-heap High-Density Memory Store case, as much memory as we can should be devoted to the High-Density store.

Related Resources

