Infrastructure - Bug #8642

Story # 8639 (New): Replication performance is too slow to service demand

Replication tasks apparently not deleted from backing store in a timely fashion

2018-07-04 13:24 - Dave Vieglais

Status:	New	**Start date:**	2018-07-04
Priority:	Normal	**Due date:**	
Assignee:	Dave Vieglais	**% Done:**	0%
Category:	d1_replication	**Estimated time:**	0.00 hour
Target version:			
Milestone:	None	**Story Points:**	
Product Version:	*		

Description

The replication task queue ~~implemented in Hazelcast uses postgres as a backing store~~ to preserve state between restarts.

Actually, it looks like the queue is entirely in Postgres. ReplicationManager holds an instance of ReplicationTaskQueue, which in turn uses the ReplicationDao interface, for which the implementation is ReplicationDaoMetacatImpl.

Observing event log messages and correlating with code where those events are emitted, it appears that tasks are not being deleted from the backing store.

```
ReplicationManager

removeReplicationTasksForPid(); log message = "removing replication tasks for pid: ..."
  ReplicationTaskRepository taskRepository.delete(tasks)
```

Sequence of log messages for a single PID, starting with removeReplicationTasksForPid():

```
[ INFO] 2018-07-04 13:19:21,996 [pool-6-thread-1]  (ReplicationManager:removeReplicationTasksForPi
d:779) removing replication tasks for pid: ess-dive-112ed52c7689908-20180328T192607490

[ INFO] 2018-07-04 13:19:22,095 [pool-6-thread-1]  (ReplicationManager:createAndQueueTasks:390) Ad
ded 0 MNReplicationTasks to the queue for ess-dive-112ed52c7689908-20180328T192607490

[ WARN] 2018-07-04 13:19:22,096 [pool-6-thread-1]  (ReplicationManager:requeueReplicationTask:794)
 In Replication Manager, task that should exist 'in process' does not exist.  Creating new task fo
r pid: ess-dive-112ed52c7689908-20180328T192607490
```

ReplicationTaskRepository is an interface that extends org.springframework.data.repository.PagingAndSortingRepository. An instance is created by repositoryFactory.getReplicationTaskRepository() using the ReplicationRepositoryFactory passed to the ReplicationManager constructor.

The implementation of ReplicationRepositoryFactory is ReplicationPostgresRepositoryFactory

History

#1 - 2018-07-04 14:41 - Dave Vieglais

- Description updated

#2 - 2018-07-05 12:11 - Dave Vieglais

- Description updated

The sequence of log messages seems to imply that delete() does work, however logic elsewhere in ReplicationManager seems to recreate the deleted task.