CS425 MP2 Group 41 Report

James Hsu chsu41, Zitong Chen zitongc2

Design:

We implemented the distributed group membership service using the heartbeating style of failure detection. Our transmissions are all UDP based to reduce the bookkeeping overhead. Our VM's are organized in a virtual ring structure, where they are ordered by their IP prefixes. Each machine will maintain a heartbeat list to monitor three of its immediate successors. They will also keep their individual member lists based on received information from other nodes in the ring. The member list updates will be included in the heartbeat messages, so each node will still be able to know whether a not-directly-monitored machine is alive or not. This approach can scale to N regardless of the magnitude of N, because each node will only have to monitor three successors nodes.

We have written two programs, introducer.go and machine.go, for the MP. Intoducer.go will only be run by the 10th machine (fa19-cs425-g41-10.cs.illinois.edu), which will serve as the contact point when other VM's want to join the group. Machine.go will be run by all VM's to use the group service.

MP1's distributed grep program allowed us to efficiently verify that a churn event has been reflected correctly in all the machines at once.

Joining:

A VM can join the group by running mahine.go, and then typing "j\n" at the prompt. It will send a join request to the introducer, which will notify other nodes in the group. Note that the introducer has to be up before this can happen.

Leaving:

A VM can leave the group by typing "q\n" at the prompt after it has joined the group. It will notify its heartbeat receivers about leaving the group, and they will remove the leaving node from their membership lists and pass down the information.

Failure:

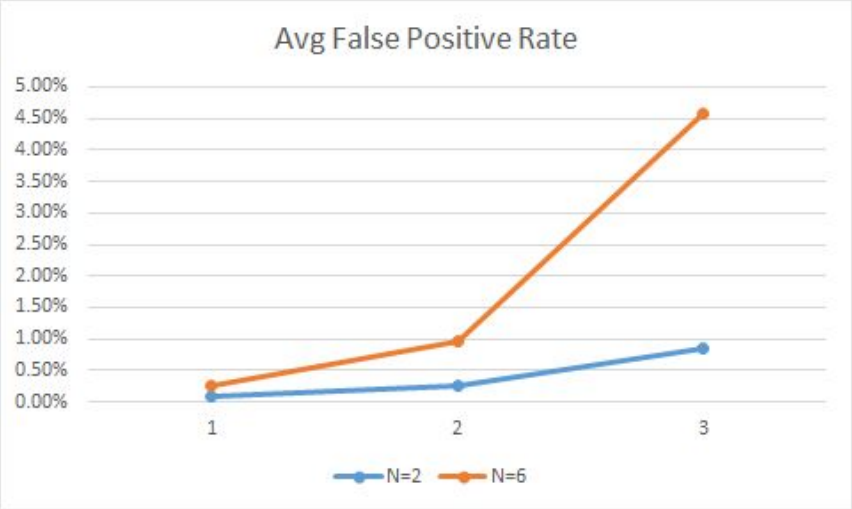
A VM will be marked as failed if any of its heartbeat receivers do not receive a new heartbeat within 4 seconds from the last heartbeat. Upon failure, a node will be treated the same way as when it leaves, but the log messages will be different. We currently set our heartbeat interval at 0.5 second.

Displaying membership list & logging:

A VM can display its up-to-date membership list at runtime by typing "l\n" after it has joined the group. All the churn messages will be logged in the file churn_log.txt. Whenever a machine rejoins the group after a fail/leave, its previous log file will be overwritten.

Evaluation (captured by iftop):
Background bandwidth usage (6 nodes): 360 Bps
Bandwidth usage (node join, leave, fail): 460 Bps, 308.5 Bps, 333.8 Bps
False positive rate stats:

Drop %	3%			10%			30%		
	avg	std	c.i.	avg	std	c.i.	avg	std	c.i.
N=2	0.10%	0.06%	0.05%	0.25%	0.14%	0.14%	0.86%	0.34%	0.30%
N=6	0.17%	0.08%	0.09%	0.73%	0.34%	0.40%	3.72%	2.84%	2.53%



We expected a positive correlation between the number of nodes and the false positive rate because if there are more machines, each node will have more heartbeat receivers. As long as any one of a given node's receivers doesn't receive a heartbeat within a given timeout, it will mark the monitored node as failed. With message drop rates increased, we see an even more aggressive increase in the false positive rate.