CS425 DS MP2 -Distributed Group Membership Report

Contributors: Shubham Singhal & Nirupam K N (ss77, nirupam2)@illinois.edu

1. Design & Implementation: In this Distributed Group Membership Protocol, we consider that all machines are in a ring topology and are linked to some of its neighbours in that ring. It is mentioned that at most three machines can fail simultaneously, so we have each node communicating to 4 of its neighbours i.e. 2 Predecessors and 2 Successors in pseudo ring topology.

Node Joins: Status of each node is set to "INIT" in the membership list before any node joins the system. New node sends a "JOIN" message to an Introducer. Introducer adds this new node to its membership list and broadcasts to all the members in its membership list by setting current status as "RUNNING".
Node Leave: Node leaves will send a "Leave" message to all the neighbours in its membership list by setting current status as "LEFT".
Node Failure: Failure is detected by monitoring the heartbeat time interval with threshold limit. Whenever the heartbeats are not received within that threshold interval, that node is marked as "FAILED" and sent to its neighbours. Eventually this message is propagated to all the nodes in the network.

When the program is executed, three threads are spawned namely commandListener, sendHeartbeats & monitorHeartbeats. commandListener thread - Processes each message (Initialize/ Join/ Leave/ Failure/ Heartbeat). sendHeartbeats thread - Periodically send Heartbeats to its neighbours. monitorHeartbeats thread - Monitors Heartbeats to check the aliveness of its neighbours.

Scalability: This protocol can be scaled up to big values of N. Any node is only sending heartbeats to 4 other nodes irrespective of size N. To add further, upon joins and failures, the entire membership list is not propagated across the nodes, but just the Hostname of the failed node. By adding more number of nodes - Bandwidth and Failure detection time of individual node is still constant even though N increases.
General Message Format: Each message contains a message type (Initialize /Join /Leave /Failure /Heartbeat), a sender hostname and timestamp.
Membership list Message Format: Each entry contains Hostname, Timestamp and MessageType.
JSON message format:
Message{ID:Id; Hostname: Hostname; IpAddress: Ipaddress; Timestamp: Time ; Status: Status}

2. Use of MP1 in MP2:
We used logging mechanism to log debug/info/error messages into a file and used MP1 to grep useful information from the log files to debug several issues and to calculate the false positive rates.

3. Measurements:

Background Bandwidth Usages for 6 Machines (assuming no membership changes): Each node sends heartbeats to 4 nodes and it will also receive 4 heartbeats from its 4 neighbors. Therefore, it is 6*(4+4)*(33+9+3) = 2160 Bytes per second. Along with this, UDP header length will also contribute to background bandwidth usage.

Average Bandwidth Usages:

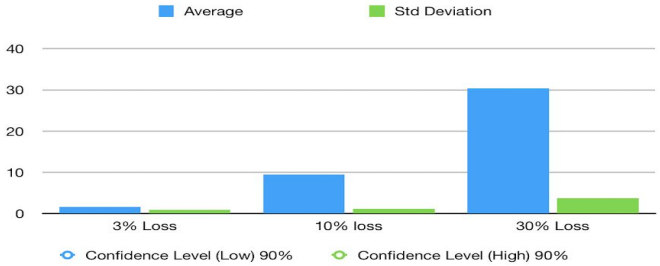
Join: When a node joins, bandwidth usage increases due to join messages. 1 message to Introducer and 4 message to other nodes. Therefore, it is (33+4+3)*5= 200 Bytes per second. So, additional +200 Bytes along with background usage.

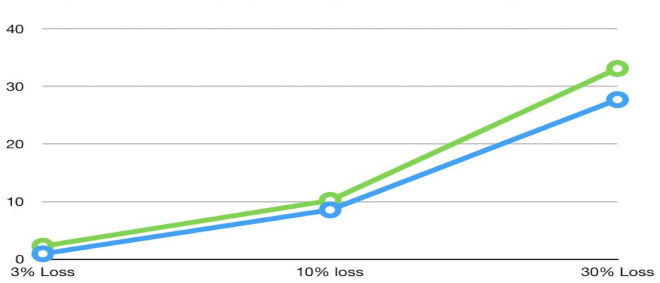
Leave: When a node leaves, bandwidth usage increases due to leave messages. Therefore, it is (33+5+3)*4 = 164 Bytes per second. So, additional +164 Bytes along with background usage.

Fail: When a node fails, bandwidth usage increases due to fail messages. Therefore, it is (33+6+3)*4 = 168 Bytes per second. So, additional +168 Bytes along with background usage.

N=6 (Out of 100 Packets)

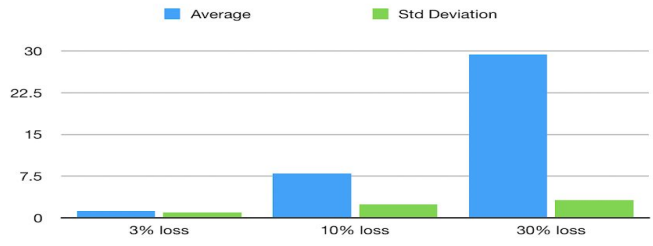
	3% Loss	10% loss	30% Loss
Reading1	2	11	29
Reading2	1	9	36
Reading3	1	10	26
Reading4	3	8	31
Reading5	1	9	29
Average	1.6	9.4	30.4
Std Deviation	0.8944	1.1402	3.7018
Confidence Level (Low) 90%	0.942	8.56	27.7
Confidence Level (High) 90%	2.26	10.2	33.1
Confidence Interval (90%)	1.6 ± 0.658	9.4 ± 0.839	30.4 ± 2.72





N=2 (Out of 100 Packets)

	3% loss	10% loss	30% loss
Reading1	1	10	32
Reading2	0	4	33
Reading3	1	8	28
Reading4	1	8	25
Reading5	3	10	29
Average	1.2	8	29.4
Std Deviation	0.9797	2.44948	3.2093
Confidence Level (Low) 90%	0.479	6.2	27
Confidence Level (High) 90%	1.92	9.8	31.8
Confidence Interval	1.2 ± 0.721	8 ± 1.8	29.4 ± 2.36



False Positive Rate: We simulated this by dropping the heartbeat messages and by measuring the false positive rate by counting the number of times the machine has been marked as failure in the logs falsely by using distributed log query. As the message loss rate increases, the false positive rate also increases. At 3% Loss and 10% Loss, the false positive rates are somewhat similar. When N=6, the standard deviation value is slightly larger than the standard deviation value when N=2. When a node fails, to disseminate the failure quickly - message is sent to all its neighbours leading to logging of more false positive failure messages.