MP4 Report (IDunno)

Powei Wang (poweiww2) and Krishna Phani Datta Chivukula(kc62)

Design: The IDunno system is underpinned by our MP3 SDFS system and the MP2 failure detector. We have used MP1 for debugging purposes. IDunno contains 1 coordinator, 1 standby coordinator, 1 client machine and the rest 7 VMs are used as worker machines. If a coordinator fails, the standby coordinator will be made the new coordinator. We used passive replication as there is a primary coordinator for a more responsive system. Each file is replicated at 4 nodes as 3 simultaneous failures are possible. Model type(ResNet50 or InceptionV3) and batch_size are the hyperparameters for the IDunno system.

Fair resource training: In order to enforce fair resource training, we have used 2 pretrained models ResNet50 on flowers dataset and InceptionV3 on cats_and_dogs dataset and used tensorflow and keras libraries. The system uses different test and inference data for both the models.

Fair-time inference: We have a JobScheduler class that will dynamically schedule the number of machines according to the number of jobs that are launched. Later when the server receives the inference query, it will forward this inference query to the scheduler, and the scheduler will send this query to the machine that has been launched to run this job uniformly. Therefore we can achieve fair-time inference because 1. Every machine has about the same computation resources 2. The number of machines is already scheduled by the scheduler 3. The number of queries sent to each machine are the same. The scheduler pre-calculates the time and resources for the jobs and tries to predict the future allocation based on the pre calculations and adjust the resource allocation and job assignment dynamically.

Fault tolerance: If the coordinator fails then the standby coordinator is made the coordinator. The state of the coordinator is being written periodically to the standby coordinator and checkpointed to the coordinator's disk to avoid any loss of job tracking and resource allocation. If a non coordinator VM fails then the operations will execute normally and the coordinator balances the jobs assigned to failed non coordinator VM to other live VMs in the network

train <model name>: Request the server to train the model. After the model is trained, the server will then copy the models into 3 other machines

launch <model name> <batch size>: Request the server to enter inference phase and add the job with batch size into the job list. The JobScheduler in the server will then assign appropriate machines it finds suitable to run the model.

inference <model name> <dataset name>: Request the server to use the model to inference the dataset. After the server received the request, it will add an inference command to a queue. And the consumer in the server will take one command out every time, and tell the according machine to run inference with the given data.

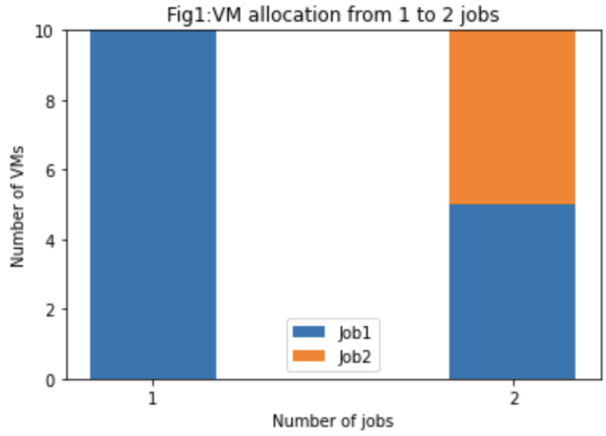
qinfo <model name>: Returns the number of processed queries of the model and the query rate within 10 seconds.

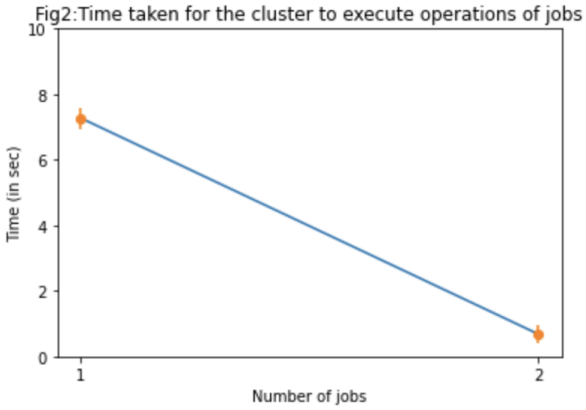
qjob <model name>: Returns the average, percentiles, standard deviation... of process time of a given job

1) a)Here the resources means, the number of VMs assigned to execute the jobs

The ratio of resources used when the cluster(1→2) $= \frac{Num\ VMs\ for\ 2\ jobs}{Num\ VMs\ for\ 1\ job} = \frac{5}{10} = 0.5$

The above ratio changes as the coordinator tries to balance the resources across the jobs fairly. Though initially the ratio is 0.5 finally the ratio is maintained such that the query rates are within 15% of each other at the end of inference of all the jobs





Initially, the coordinator assigns the jobs initially equally to the jobs it has in the system. After assigning the jobs the coordinator then calculates the inference times for each job. The coordinator gradually reduces the resources to the job which is taking less inference time and increases the resources to the job taking more inference time to ensure fair time inference. The coordinator assigns the resources in the above mentioned dynamic fashion to ensure that the query rates are within 15% of each other at the end of inference of all the jobs.

b) Time taken for the cluster to start executing the queries of job2

Time taken for the cluster(1→2) to start executing the queries of the second job = 7.27846 sec

Time taken for the cluster to execute queries of jobs	Mean (sec)	Standard deviation (sec)
first job (0→1)	7.27846	0.32034
second job (1→2)	0.67846	0.27373

As you can see from Fig2(line plot) the time taken for the cluster to execute the queries from (0→1) is much higher than the (1→2) case

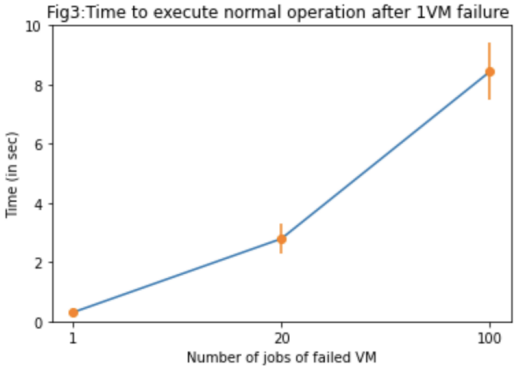
This is expected as in the case of (0→1) IDunno takes time for initiating the coordinator and other nodes in the cluster whereas the in case of (1→2) the cluster takes much less time as the initiation time is not considered here

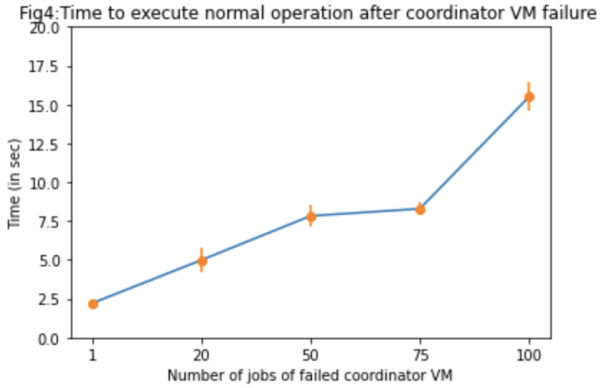
2) **Time taken for the cluster to resume "normal" operations after failure of 1 (non coordinator) VM = 0.30073 sec**

Time taken for the cluster to resume	Number of jobs	Mean(sec)	Standard deviation (sec)

normal operation after failure of 1VM			
Non coordinator VM failure	1	0.30073	0.06243
	20	2.79205	0.50213
	100	8.43868	0.95760

From Fig3 below the time taken for the cluster to resume the “normal” operation after failure of 1 non coordinator VM is increasing as the number of jobs is increasing. This is desired, as the number of jobs increase, the coordinator will take more time to balance the failed VM jobs to other live VMs





3) **Time taken for the cluster to resume “normal” operations after coordinator VM fails = 2.20940 sec**

Time taken for the cluster to resume normal operation after failure of coordinator VM	Number of jobs	Mean(sec)	Standard deviation (sec)
Coordinator VM failure	1	2.20940	0.29932
	20	4.98705	0.80392
	50	7.82952	0.69220
	75	8.29103	0.42921
	100	15.49202	0.92919

From Fig4 above the time taken for the cluster to resume to execute “normal” operation after failure of 1 coordinator VM is increasing as the number of jobs is increasing. This is desired, as the number of jobs increase, the standby coordinator will take more time to reconcile the state of the failed coordinator.