CSCI 6360: Parallel Computing Lecture Summary - 13

Anirban Das (dasa2@rpi.edu)

March 9, 2018

Summary on Parallel I/O papers n Lecture 13:

1   Incast Problem

Overall message is basically if you are running a large scale storage system it is incredibly challenging to run. Its not worth the grad students time to be in storage business unless its a enterprise class storage system.

TCP gives you most BW available , QOS, using congestion protocols and backup systems. Becomes bottleneck in data centers i.e basically with scaling. When all MPI ranks tries to fopen a single file the throughput collapses. In datacenters the equipments are really really close, hence in general we dont see such high speeds in internet.

Preconditions for TCP incast collapse: 3. What can help fix this problem.

Look at various matrix. RTOmin vs Goodput graph. Aggregate more server, less BW per server and hence links not saturated /. As long as we have a responsive rto then we could expect a high godput across all servers.

Rtt in milliseconds graph. Bottomline claims: need to use fine grained RTT things and other one is the 3 reasons for in cast.

2   I/O performance challenge at Leadership scale

Which talks about IO challenges in order to get their own storage systems to work. Mordern IO storage system in mordern HPC systems.

How did this system perform under workloadIOR benchmark