Implementation and Evaluation of Fault Tolerance Mechanisms for MPI Applications

Fault-tolerance enables applications and users to cope with occurring failures. This is desirable in High Performance Computing where job aborts usually result in high costs. Checkpoint/Restart provides a common strategy to increase fault-tolerance of applications. This thesis analyzes a tool called Distributed MultiThreaded CheckPointing (DMTCP). This tool is of interest because of its ability to checkpoint large distributed applications with a negligible runtime overhead. The results reveal that the checkpoint duration is almost independent of the amount of processes being checkpointed. Furthermore, applications launched under the control of DMTCP only experience little runtime overhead.