Load balancing is one of the crucial issues in a distributed system. Load balancing is the process of harmonizing load among the clusters of a distributed system. It aims to optimize resource utilization, maximize throughput, minimize response time, and avoid overload of any single resource. The literature survey presented in this report describes various techniques, taxonomies, goals, and components of LBA. The hybrid architecture proposed which is a combination of centralized and decentralized approach.

We have concluded from the literature survey that most of the existing DLBAs are post-active. They allow the system to go into imbalance state and then try to balance the load among the cluster by reallocation of jobs or migration of jobs from overloaded cluster to underloaded cluster. This problem motivates us to proposed pre-active DLBAs, which can balance load without job migration.

The prediction based load balancing algorithm (PDLB), pluggable to scheduler dynamic load balancing (P2S_DLBB), and the combination of PDLB and P2S_DLBB, prediction based dynamic load balancing (PP2S_DLBB) are proposed for clustered heterogeneous computational environment. The prediction algorithm predicts the status of clusters in advance. The load balancing algorithm uses the predicted status of clusters for scheduling the incoming job. The predicted information plays a vital role in the selection of the best destination cluster for incoming jobs. The P2S_DLBB algorithm finds clusters with higher numbers of free resources and schedules incoming jobs to that clusters. It steps by step balances the load between clusters.

We have simulated traditional scheduling algorithm (PS, FCFS, SJF, EDF) and proposed PDLB, P2S_DLBB, PP2S_DLBB algorithms on a real-time dataset of 3 months having 51987 jobs using real-time cluster configuration in ALEA simulator. The cluster utilization, imbalance load matric, and standard deviation are calculated as an evaluation parameter. The comparative analysis of results of all algorithms in different cluster configurations is done. The comparative results show that proposed algorithms improves the average cluster utilization and reduces the load imbalance matric as well as the standard deviation of the distributed computing environment.
The P2S_DLB algorithm is plugged to traditional FCFS, SJF, and EDF scheduling algorithms and modified dynamic scheduling algorithms FCFS_DLB, SJF_DLB, and EDF_DLB are proposed. All three scheduling algorithms without and with a combination of P2S_DLB are simulated in ALEA. The simulation results show that the P2S_DLB increases the average cluster utilization and decreases the load imbalance metric in all the three scheduling algorithms. Hence it concluded that the performance of P2S_DLB remains stable with all the tree scheduling algorithms. Out of three scheduling algorithms, the combination of EDF and P2S_DLB gives the best performance.

The PP2S_DLB algorithm performed best compared to all proposed and existing algorithms. The highest average cluster utilization is achieved using PP2S_DLB algorithm. The individual cluster utilization increases and the variation between individual cluster utilization decrease in the PP2S_DLB algorithm.

All the proposed DLBAs are simulated in ALEA simulator using homogeneous cluster configuration of Gaia cluster. The comparative results show that all the proposed DLBAs are performed better compared to PS algorithm. Therefore, we have concluded that the proposed DLBAs are giving better performance even in the homogeneous cluster environment.

At last, all the proposed DLBAs are simulated in ALEA simulator using MetaCentrum dataset. The comparative analysis of the simulation results of all the proposed DLBAs and the PS algorithm is done by different evaluation parameters. The comparative results show that the performance of all the proposed DLBAs is better compared to the PS algorithm. Therefore, we have concluded that the proposed DLBAs can balance the load for any distributed clustered environment.

As future work, the proposed algorithms can be simulated with other real-time datasets having higher numbers of jobs for checking stability in all kinds of scenario. The prediction cycle of proposed PDLB and PP2S_DLB algorithms can be increased for accurate prediction of cluster load status values. The designed algorithms can simulate in different clusters configuration. The dynamic multilevel load balancing threshold policy can be set which can change as per the dynamic nature of the incoming load. The proposed DLBAs can be implemented in real-time DS for performance evaluation.