Parallel Data Processing with the PROOF Framework

M. Schott, O. Biebel, G. Duckeck, J. Elmsheuser, J. Kennedy, G. Krobath, T. Mahmoud, B. Ruckert, D. Schaile, and C. Serfon

The PROOF (Parallel Root Facility) framework [1] was designed to perform physics analyses on parallel computer clusters based on the ROOT framework. Due to the huge amount of data expected at the ATLAS experiment, it was tested to which extent the PROOF framework could be used for physics analyses of ATLAS data. PROOF was installed and configured at ten nodes on the LRZ Tier-2. Each of the nodes has four processing cores.

The overall processing time for a given problem at a computer cluster with n cores can be described by

$$T(n) = t_o + t_d(n) + \frac{t_c}{n},$$

where t_c is the required processing time at a single node. t_o describes the required processing time of the problem, which cannot be parallelized, and $t_d(n)$ describes the time required for communication between the different nodes and the time for transferring data, which is usually bound to the maximal data transfer-rate of the underlying file system. The processing of ATLAS data can be highly parallelized, since each single recorded collision is independent from the others. Hence different nodes can process different events.

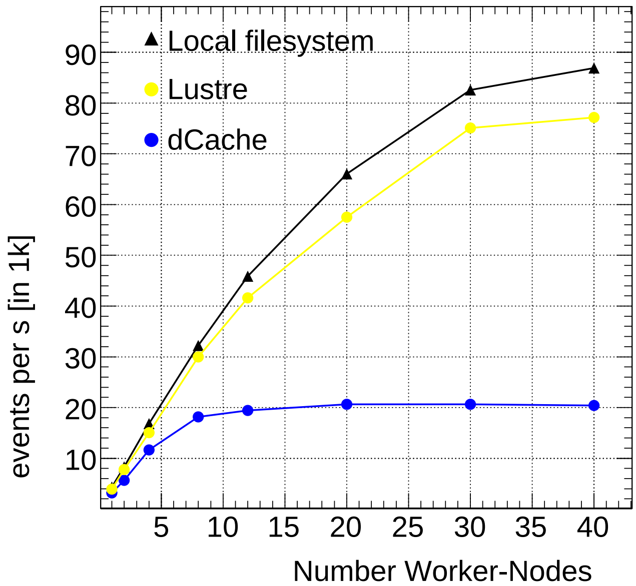

Fig. 1: Number of processed events per s vs number of CPU cores used in PROOF cluster.

Simulated ATLAS events which would have been triggered by a single muon trigger and correspond to an integrated luminosity of $100\,pb^{-1}$ have been chosen as a basis for performance tests of the PROOF system. A simple $Z \rightarrow \mu\mu$ analysis on this data was used as an example study. In a first step, the data was stored on each node separately. A scalable behavior of the number of processed events per second can be seen up to 10 processing cores, which are used for data processing in the PROOF cluster, as shown in Fig. 1. Using more cores for data processing leads to a deviation of the scalability due to the maximal data transfer-rate of the hard disks in each node.

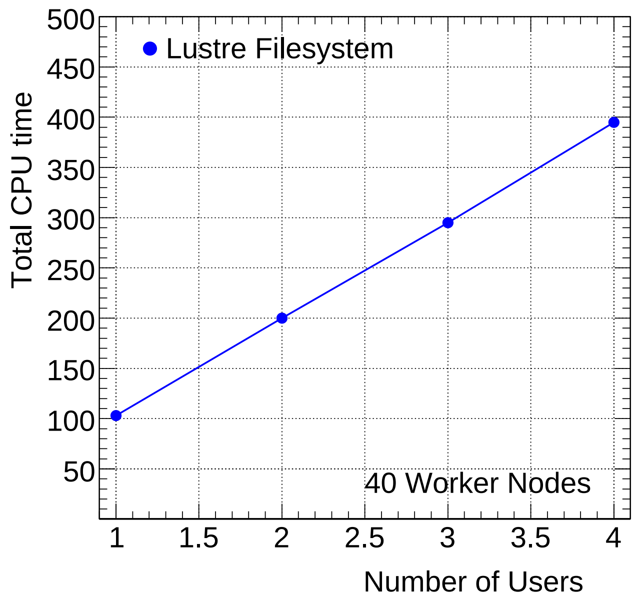

Fig. 2: Total processing time vs. number of users which perform independent analyses.

The performance of the Lustre file system and the dCache file system has also been studied. These file systems allow the PROOF framework a common access to the files, without the need of manual file replication. The number of processed events per second for both file systems is also shown in Fig. 1. While the Lustre file system showed a similar behavior as the local file system, the performance for dCache was significantly decreased. However, the dCache system at LRZ was not tuned for such an access pattern, further tests are needed for a real comparison. Several users, performing different and independent analyses, can also use the PROOF framework at the same time. A linear behavior of the total processing time of one analysis and the total number of users is expected and has been observed (Fig. 2),

The presented study showed that the PROOF cluster at LRZ provides a performant environment for physics analysis of ATLAS data during the coming years.

References

[1] F. Rademakers *et al.*, Latest Developments in the PROOF System, Proceedings of CHEP 2007, Victoria, September 2007, to be published.