Distributed PIV Technology: Network Storage Usage

Rodion Stepanov[1] and Andrey Sozykin[2,3]

[1] Institute of Continuous Media Mechanics UrB RAS, Perm, Russia
[2] Krasovskii Institute of Mathematics and Mechanics, Yekaterinburg, Russia
[3] Ural Federal University, Yekaterinburg, Russia
rodion@icmm.ru

Abstract. The approach to data transfer from particle image velocimetry system to a supercomputer through the network attached storage is suggested. The advantages of the approach are simple implementation and high communication speed. Connecting the particle image velocimetry system to the super-computer allows us to carry out real-time controlled experiments with feed-back and to apply computational intensive algorithms for processing.

Keywords: particle image velocimetry · supercomputer · network attached storage · high performance computing · high-speed network

1 Introduction

Particle Image Velocimetry (PIV) is a popular method of visualization of fluid and gas flows for its ability to estimate the velocity field [1]. The PIV method is widely used in hydrodynamics [3], aerodynamics [12], astrophysics [5], medical research [6], and other areas of science. The specific feature of PIV is generating large amount of imaging during the measurement process (tens or thousands of gigabytes), which then is used to compute the velocity field of the flow. Nowadays the images are processed on personal computers, which do not have enough performance to process data in the real-time. Hence, controlled experiments cannot be carried out. In addition, the relatively low performance of personal computers is not suitable for the advanced computational consuming algorithms of velocity field calculation. Widely using standard crosscorrelation algorithm meets only the minimal requirements to processing quality. Connecting the PIV system to a supercomputer removes the computational resource restriction and provides the ability to use the preprocessing procedures for noise filteration and adaptive algorithms. Effective computation distribution makes it possible to process images in the real-time and run the experiments with feedback.

The main problem to connect PIV system to a supercomputer is the lack of high-speed data transfer interfaces in modern supercomputers. Although supercomputers use high-speed network technologies (Gigabit and 10G Ethernet), the popular protocols, which are used to transfer data to supercomputer (SCP, FTP and so on), cannot utilize the full network bandwidth. In addition, the

encryption of transferred data widely used due to security consideration creates significant overhead and further decrease of the performance. The encryption is often unnecessary for connection of the experimental facility to supercomputers and, therefore, should not be used.

The common approach to data processing on the supercomputer is also problematic. According to de-facto standards, experimental data are accumulated on the local storage first, then whole data are transmitted to the supercomputer, and after that it can be processed. In this case the data processing in real-time is not possible.

The speed of data transfer to supercomputer can be increased by eliminating unnecessary intermediate elements, such as local storage of experimental facility and head node of a supercomputer. This can be done by writing experimental data directly to the supercomputer storage system. In this paper we present and test the architecture of supercomputer input/output system, which is provided with such capabilities, describe the implementation of proposed architecture in the "URAN" supercomputer, and the connection of the PIV system of Institute of Continuous Media Mechanics UrB RAS (ICMM UrB RAS) to this supercomputer.

2   PIV System

PIV method is based on the fast recording the motion of flow with small particles, which can be specially added to the medium or already be there. Velocity field is estimated by comparing two images taking in the rapid succession. A scheme of typical PIV system is shown in Fig. 1.

The base components of PIV system are pulsed laser lightening the particles, camera capturing the pairs of images at small time intervals, and personal computer, which synchronizes laser and camera and computes crosscorrelation between images. Modern cameras can generate the data stream up to 500 Mbit/s. Some PIV system uses two cameras to compute three components of velocity and produce even more data stream. The average volume of data generated during one experiment varies from 100GB to 10TB depends on the details of the experiment.

Velocity field is determined by analysing the pair of images. Nowadays, several algorithms exist for this purpose, the most popular is the crosscorrelation algorithm [7]. One of the drawbacks of the standard crosscorrelation algorithm is the requirements of using computational areas with rectangle shape and fixed boundaries. This leads to low dynamic range of velocity field values. Another disadvantage is sub-pixel interpolation procedure, which cause false peak creation near the integer-valued deviation on the probability distribution function.

As an alternative to the crosscorrelation algorithm for computing the velocity field, adaptive choosing of computational areas and wavelet crosscorrelation algorithms [8] can be used. However, the relatively low performance of personal computer constrains the development and application of these algorithms due to their high computational requirements. While the crosscorrelation algorithm
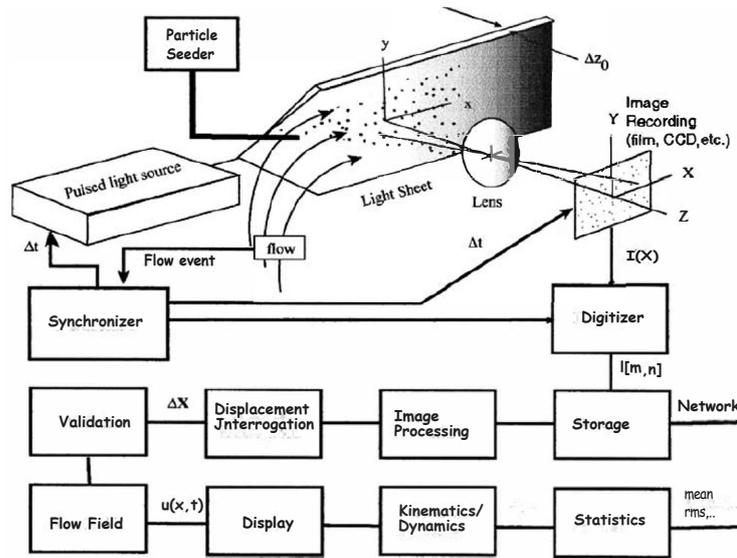
Fig. 1. Typical monoscopic PIV system [2] with a network access

requires 150 CPU cores for the real-time processing, the wavelet crosscorrelation needs 1500 CPU cores for full processing of data. But to run experiments with feedback, it is not necessary to process all images generated by the PIV system. Processing one pair of images per second is enough for control. Therefore, the wavelet crosscorrelation algorithm for the purpose of real-time controlled experiments requires only 150 CPU cores.

3 Related Work

[13] describes the real-time controlled experiments with feedback based on PIV measurements. The mineral oil has been chosen as a working fluid because it provides a relatively low flow speed up to 25 sm/s. Such low speed makes it possible to process PIV images on the personal computer with one dual-core CPU. Authors declare that the performance of their system is limited, which leads to occasional image loss and control command delays. To solve this problem and to control the flows with higher speed, the performance of image processing needs to be increased, for example, with the help a of supercomputer.

First effort to connect PIV system to a supercomputer has been made as part of "Distributed PIV" project [9]. The attempt has been made to transfer data from PIV system in ICMM UrB RAS, Perm, to the supercomputer "Chebyshev" in Moscow State University through the dedicated network channel 1Gb/s. As a result, the restrictions of standard technologies, which are used to transfer data to supercomputers, have been revealed. Particularly, maximum speed of

writing data to the supercomputer using standard network protocol is only approximately 300 Mb/s for CIFS and 500 Mb/s for FTP. These rates have been achieved by running several data transfer session simultaneously. Speed of each session has been significantly smaller. Based on the results from conducted experiments, [10] suggested the architecture and special protocol to multisession data transfer from PIV system directly to supercomputer nodes bypassing the head node. However, to use proposed protocol, special software must be developed and installed on the PIV system and the supercomputer.

Increasing data processing speed of PIV system can be achieved not only by using supercomputers, but also with the help of Field Programmable Gate Arrays (FPGA). In paper by [14], the hardware implementation of the direct crosscorrelation algorithms based on the FPGA is described. The FPGA board is installed into the case of a personal computer in the PIV system, which allows providing computational resources without creating infrastructure of data transfer to a supercomputer. Meanwhile, the FPGA programming is much more complicated comparing with the developing software for supercomputers, which impose constraints on wide FPGA using.

4   Architecture

To overcome the drawbacks of existing data input interfaces of supercomputers, we suggest changing the architecture of the supercomputer input/output system. In contrast to traditional approach of transferring data to supercomputer through a head node, we propose to write data directly to the storage system in the supercomputer. The suggested architecture is presented in Fig. 2.
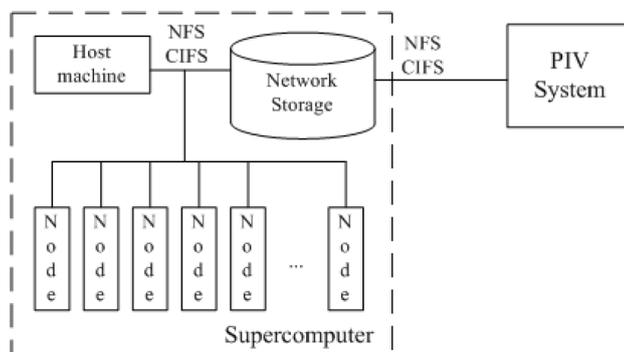

Fig. 2. Architecture of data transfer from PIV system to a supercomputer

The storage system of a supercomputer must use the Network Attached Storage (NAS) technology and contain at least two network interfaces to provide the ability to connect the PIV system. One interface is used to connect the nodes of

a supercomputer to the storage, and other one is intended to connect the PIV system.

Connection to the storage system can be established by the standard network protocol, such as NFS (Linux- and UNIX-based computers) and CIFS (Windows-based computers). One logical volume inside the storage system can be connected to nodes of the supercomputer and to the PIV system simultaneously. Storage system prevents data losses caused by concurrent access to files using the mechanisms of network sharing files protocols NFS and CIFS.

The storage is presented to the PIV system as a simple network drive or a catalog. But the images from the camera, which are written to this disk, are available not only to PIV system, but also to supercomputer nodes.

The main advantage of the proposed architecture is the transparent integration of the PIV system and the supercomputer: there are no needs to modify the experimental facility. The only one required change is to write data to the network drive instead of local drive of a personal computer in the PIV system.

5   Implementation

5.1   Trial Supercomputer Structure

The proposed architecture has been implemented in the supercomputer "URAN", which is installed in the Institute of Mathematics and Mechanics UrB RAS (IMM UrB RAS), Yekaterinburg, and used to connect the PIV system of ICMM UrB RAS to this supercomputer.

The supercomputer "URAN" has the peak performance approximately 160 TFlops and consists of the Linux-based nodes with Intel CPU and NVIDIA GPU. The storage subsystem of the supercomputer "URAN" includes the internal drives of the head node and the NAS system EMC Celerra NS-480. This NAS system has the 30 TB of usable capacities, two hardware RAID-controllers, and 8 Gigabit Ethernet network interfaces. Six of those interfaces are used to connect the supercomputer nodes to the storage subsystem, while other two are devoted to external experimental facilities. These two interfaces are connected to the Academic Network of the Ural Branch of RAS and to the Internet. Therefore, experimental facilities can transfer data by the network connection directly to the storage system of the supercomputer "URAN".

The PIV system in the ICMM UrB RAS includes two high-speed cameras, each of which generates pairs of 4 Megapixel images at frequency 15 Hz; therefore, the maximum data stream is 240 Mb/s. The PIV system is managed by a Windows-based personal computer, which is also run the ActualFlow software for processing images using the crosscorrelation algorithm. ICMM UrB RAS and IMM UrB RAS are connected by the dedicated channel of the Academic Network UrB RAS utilizing the DWDM technology. The speed of the channel physical media is 1 Gb/s, length of the channel is approximately 400 km, and round-trip time is approximately 5 ms.

The supercomputer storage system has the separate logical volume devoted to store experimental data from the PIV system. The logical volume consists

of ten 300GB Fibre Chanel disks, and has the useful capacity 1.8 TB. The logical volume is exported simultaneously by NFS and CIFS protocols. The supercomputer nodes with Linux operating systems mount this logical volume by NFS in the special directory /home3. The personal computer in the PIV system with Windows use CIFS to mount the logical volume as a network drive. When the PIV system writes data to this drive, they are became available for the supercomputer nodes in the specified catalog. Simultaneous work with the same logical volume from different operating systems by different network protocol is provided by NAS system EMC Celerra NS-480.

5.2   Security

Data from the PIV system to the supercomputer "URAN" are transferred through the dedicated channel of the Academic Network of UrB RAS, which is isolated from the Internet. Inside the Academic Network of UrB RAS, network connection is further isolated by VLAN technology. Dedicated VLAN contains only the computer in the PIV system and the network interfaces of supercomputer NAS system, which are intended to connect external experimental facilities.

Each experiment running on the PIV system in the ICMM UrB RAS does not demand high security. Therefore, we make a decision do not use encryption due to its large overhead. As a result, the performance of data transfer is improved, while sufficient level of security is provided by isolation of the communication channel from the Internet.

Inside the storage system, two level of access control are implemented. The first level is the access restriction to logical volume by IP-address of experimental facility. The second level is restriction by user name and password. Mapping the usernames and file owners in the NFS and CIFS is provided by NAS system. As a result, it becomes possible that several users with different names and passwords can work with the PIV system, for example, to carry out different experiments. Files of such users are isolated from each over.

5.3   Performance Evaluation

To estimate the performance of suggested solution, we run several experiments to test the speed of data transfer to the supercomputer by different protocols. We use sequential write speed test because PIV system creates this type of load due to sequential writing of flow images recorded by camera.

The first experiment is in testing the traditional approach of transferring data to supercomputers through the head node by SCP protocol. We copy the images generated by the PIV system from the personal computer with Windows 7 to the supercomputer "URAN" by utilities pscp and WinSCP. The speed of data transfer is measured by the tools of pscp and WinSCP.

The second experiment is in writing data directly to the network storage of supercomputer from Windows 7 machine by CIFS (using both SMB1 and SMB2 protocols). To increase the performance of the data transfer through long distance channel, the Compound TCP [11] protocol has been enabled in Windows.

The performance is measured by the iozone utility (IOzone Filesystem Benchmark, http://www.iozone.org/). In the third experiment, the data are written also directly to the supercomputer storage from the computer with Ubuntu Linux 11.04 by NFS. We use the 4th version of NFS, and the size of block (both wsize and rsize) is 1MB. The performance is also measured by the iozone.

Results of the experiments are presented in Fig. 3. The number of experiments of each type is 100; the average results with confidence intervals are presented.
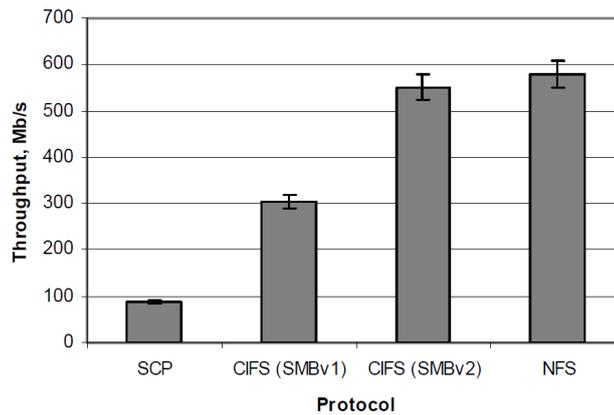

Fig. 3. Sequential write speed test

The worst results are obtained using the traditional approach to transfer data through head node by SCP protocol. Such low speed is caused by existing the intermediate element (head node of supercomputer) and encryption used by SCP that create significant unnecessary overheads.

The best results are achieved by Linux machine with NFS protocol. Its performance is 6 times more, then at SCP. Despite of such good results, this approach cannot be used in the PIV system of the ICMM UrB RAS because it includes the personal computer with Windows. But in the future Linux-based experimental facilities can be connected to the supercomputer by the NFS protocol.

Performance of the Windows machine with SMB version 2 is only slightly less then the performance of Linux machine. SMB2 is also provides the 6 times speedup compare to traditional SCP. The disadvantage of SMB2 is that it is available only in relatively new versions of Windows, such as Windows Vista, Windows 7, and Windows Server 2008. Early versions of Windows support only SMB1 protocol, which provides only half of the SMB2 performance (see Fig. 3). However, both SMB2 and SMB1 provide enough performance to data transfer from the PIV system of the ICMM UrB RAS to the supercomputer "URAN".

The experiments results have confirmed that writing data directly to the network storage in the supercomputer can notably speedup the data transfer.

It should be emphasized that the significant difference from the results presented in [9, 10] is that performance in our experiments has been reached in one session. Consequently, it is not necessary to run several sessions simultaneously to achieve high-speed of the data transfer.

6   Conclusions

The approach to organize the high-speed data transfer from the PIV system to the supercomputer based on direct write to the supercomputer network storage has been presented. The advantage of approach is that it can be used without modification of experimental facility. The performance testing shows that direct write to the network storage by CIFS or NFS protocols can increase the data transfer speed 6 times in comparison with the traditional data transfer through supercomputer head node by SCP or FTP protocols. In contrast to the previous work [9, 10], the speedup can be achieved in one network session without the requirements to run multiple session of the data transfer simultaneously to utilize the bandwidth of network connection.

Suggested approach has been implemented to connect the PIV system in the ICMM UrB RAS, Perm, Russia to the supercomputer "URAN" in the IMM UrB RAS, Yekaterinburg, Russia. Distance between the PIV system and the supercomputer is approximately 400 km. The connection uses dedicated GigabitEthernet channel of Academic Network of UrB RAS.

The high-speed data transfer provides the ability to process experimental data from the PIV system on the supercomputer in the real-time and control the experiment based on the results of such processing. Moreover, it is possible to use the supercomputer to implement highly accurate but computational consuming image processing algorithms. Personal computer cannot be used to run such algorithms due to low computational resources. Since the results of processing can also be written to the storage system, the user of the PIV system can visualize the experiment process using the standard existing tools. As a result, the user can monitor the experiment course and control its conditions.

Future work includes conducting the closed-loop experiments with feedback based on PIV measurements, connecting other experimental facilities to the "URAN" supercomputer, such as setup for two-phase flux control in spray of injectors for aircraft engines [4], implementing the adaptive and wavelet cross-correlation algorithms of velocity field estimation, evaluating the possibility to run this algorithms on GPU, and increasing the speed of data transfer by using 10G Ethernet network equipment.

References

1. Adrian, R.J.: Scattering particle characteristics and their effect on pulsed laser measurements of fluid flow: speckle velocimetry vs particle image velocimetry. Applied Optics 23(11), 1690–1691 (1984)
2. Adrian, R.J.: Twenty years of particle image velocimetry. Experiments in Fluids 39, 159–169 (2005)
3. Batalov, V., Sukhanovsky, A., Frick, P.: Laboratory study of differential rotation in a convective rotating layer. J. Geophys. Astrophys. Fluid Dynamics 104(4), 349–368 (2010)
4. Batalov, V., Kolesnichenko, I., Stepanov, R., Sukhanovsky, A.: The use of field measurement techniques to study two-phase flows. Vestnik Permskogo Universiteta. Mathematics. Mechanics. Informatics 5(9), 21–25 (2011)
5. Frick, P., Stepanov, R., Sokoloff, D., Beck, R.: Wavelet based faraday rotation measure synthesis. Monthly Notices of the Royal Astronomical Society Letters 401, L24–L28 (2010)
6. Hochareon, P., Manning, K., Fontaine, A., Deutsch, S., , Tarbell, J.: Development of high resolution particle image velocimetry for use in artificial heart research. In: Second Joint EMBS-BMES Conference. pp. 1591–1592. IEEE (Oct 2002)
7. Keane, R., Adrian, R.: Theory of cross-correlation analysis of piv images. Applied Scientific Research 49(3), 191–215 (JUL 1992)
8. Mizeva, I., Stepanov, R., Frick, P.: Wavelet crosscorrelations of two-dimensional signals. Numerical methods and programming 7, 172–179 (2006)
9. Stepanov, R., Masich, A., Masich, G.: Initiative project "distributed piv". In: Proceedings of Scientific service in the Internet: scalability, parallelism, efficiency. pp. 360–363 (2009)
10. Stepanov, R., Masich, A., Sukhanovsky, A., Schapov, V., Igumnov, A., Masich, G.: Processing the stream of experimental data on the supercomputer. In: Proceedings of Scientific service in the Internet: exaflops future. pp. 168–174 (2011)
11. Tan, K., Song, J., Zhang, Q., Sridharan, M.: A compound tcp approach for high-speed and long distance networks. In: Proc. IEEE INFOCOM. pp. 1–12 (2006)
12. Willert, C., Raffel, M., Kompenhans, J.: Recent applications of particle image velocimetry in large-scale industrial wind tunnels. In: International Congress on Instrumentation in Aerospace Simulation Facilities. pp. 258–266. IEEE (Sep 1997)
13. Willert, C.E., Munson, M.J., Gharib, M.: Real-time particle image velocimetry for closed-loop flow control applications. In: 15th Int Symp on Applications of Laser Techniques to Fluid Mechanics (2010)
14. Yu, H., Leeser, M., Tadmor, G.: Real-time particle image velocimetry for feedback loops using fpga implementation. Journal of Aerospace Computing, Information, and Communication 7, 52–62 (2006)