World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:6, No:9, 2012

Benchmarking: Performance on ALPS and Formosa Clusters

Chih-Wei Hsieh, Chau-Yi Chou, Sheng-HsiuKuo, Tsung-Che Tsai, I-Chen Wu

Abstract**—**This paper presents the benchmarking results and performance evaluation of differentclustersbuilt atthe National Center for High-Performance Computingin Taiwan. Performance of processor, memory subsystem andinterconnect is a critical factor in the overall performance of high performance computing platforms. The evaluation compares different system architecture and software platforms. Most supercomputer used HPL to benchmark their system performance, in accordance with the requirement of the TOP500 List. In this paper we consider system memory access factors that affect benchmark performance, such as processor and memory performance.We hope these works will provide useful information for future development and construct cluster system.

Keywords**—**Performance Evaluation, Benchmarking and High-Performance Computing

I. Introduction

THE rapid improvement of the microprocessor and memory along with the availability of low-cost and fast interconnection network made it possible for many research groups to put together commodity off-the shelf PCs to build parallel high-performance computers. Having the advantage of delivering high-performance at low-cost, PC Clusters are becoming one of the most important platforms for HPC[6].

The National Center for High-Performance Computing (NCHC), Taiwan has dedicated time and efforts to research, develop, and promote on PC Cluster technologies since 1999. The NCHC is self-madeserial Formosa PC Cluster systems, Formosa, Formosa II[2], Formosa III [3]and Formosa IV HPC Cluster over the past decade. In 2003, Formosa PC Cluster was built and become the fastest computer system in Taiwan then. The Central Processing Unit (CPU) of Formosa II HPC Cluster is AMD Opteron 275 DualCore, extending the IA32 architecture to 64 bits by Intel Extended Memory 64 bit Technology (EM64T).The Formosa II HPC Cluster has achieved1166*Gflops* performance of HPL. It has record #298 in TOP500 lists in Nov. 2004 [8].

The Formosa III HPC Cluster is a 64-bit Beowulf Cluster located in the southern branch of NCHC. It consists of 76 IBMSystem x3550 M3 servers as its compute nodes. This self-made cluster was designed and constructed by the HPC ClusterTeam at NCHC for computational science applications and came online in 2011.

C. W.Hsieh is with the National Center for High-Performance Computing and National Chiao Tung University, Hsinchu, Taiwan (e-mail: david.hsieh@nchc.narl.org.tw).

C-Yi Chouis with the National Center for High-Performance Computing, Hsinchu, Taiwan (e-mail: b00cyc00@nchc.narl.org.tw).

S-H Kuois with the National Center for High-Performance Computing, Hsinchu, Taiwan (e-mail: a00mba00@nchc.narl.org.tw).

T-C Tsaiis with the National Center for High-Performance Computing, Hsinchu, Taiwan (e-mail: 1103911@nchc.narl.org.tw).

I-Chen Wu is with the Department of Computer Science, National Chiao Tung University, Taiwan. (e-mail: icwu@csie.nctu.edu.tw).

Each node has two six-core Intel modelprocessors and 48*GB* of DDR 3 registered ECC SDRAM. All nodes are connected by the 4x QDR InfiniBand high speednetwork and a private subnet with Gigabit Ethernet. An additional four nodes are arranged as login nodes, thus, enablingusers' access easily. The Formosa III missionis virtualization support that different from previous series clusters.

Since the installation of hardware and software, we have tested several popular benchmark programs on the clusters.Performance of processor, memory subsystem and interconnect is a critical factor in the overall performance of a computing system and thus affecting the applications running on it. The LLCbench benchmark suite [4], including cachebench,blasbench and mpbenchare designed to evaluate theperformance of the memory hierarchy, Basic Linear Algebra Subroutines(BLAS) and Message Passing Interface (MPI)[9].

There are also some application programs from different scientific and engineering domains currently running on the clusters for performance evaluation. This paper presents our experience in the performance evaluation of PC clusters based on performance data collected from a broad range of standard benchmark programs and real applications. The results provide some hints about the attainable performance with PC clusters, which is largely application-dependent as we will describe in the following sections.

The rest of this paper is organized as follows. Section 2 introduces the NCHC cluster system architecture. Section 3 describes the benchmark software. Section 4 shows the performance of different benchmark software. Finally presents conclusions.

II. System Descriptions

In this paper, we evaluatethreeHPCclusters, the NCHC SUN GPU Cluster was built in 2010, the NCHC ALPS was built in 2011 and the NCHC Formosa IV was built in 2011. We show the detail system description as follows:

A. The NCHC SUN GPU Cluster

The role of the GPU accelerator has become more and more important for scientific computing. The GPU has become the world's top driving force behind supercomputer.The NCHC SUN GPU Cluster built in 2010thatwas contains two of Intel X5570inside four cores running at 2.93*GHz*a node sharing 24 GB RAM. As Fig. 1 shows each computing node integrates one GPU device S1070. They are connected together with Mellanox MT26428 ConnectX IB 4x QDR.

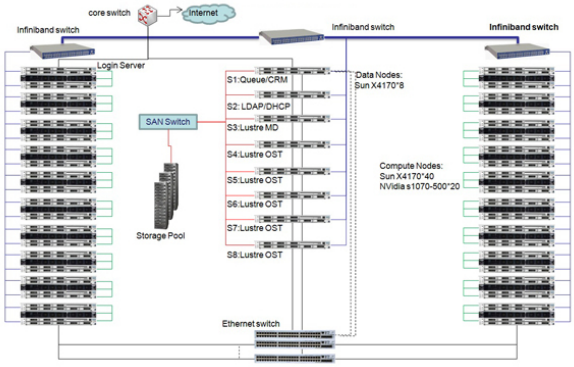


Fig. 1 The NCHC SUN GPU Cluster Hardware configuration

B. The NCHC ALPS Cluster

The NCHC ALPS is short for "Advanced Large-scale Parallel Supercluster" (also known as "御風者" in Chinese). It is a supercomputer that offers an aggregate performance of over 177 *Tflops*. The system uses the AMD Opteron 6100 processors, and has a total of 8 compute clusters, a large memory cluster, and over 25,600 compute cores.

The ALPS Cluster is designed to support a wide range of applications - so many that NCHC could only begin to fathom the jobs that could potentially be run across it. APLS is also designed to function as a test bed for new application design and research; support for open source and academic-specific applications is a must. To offer the most robust level of support, the system was first planned with every angle in mind - not simply raw compute performance, and not necessarily best performance per watt.

The hardware of computing nodes on NCHC ALPS consists of 600 of Acer AR585. They are connected together with Qlogic InfiniBand in 4x QDR(40*Gb/sec*) and the bandwidth throughput of this system achieves 51.8 *Tbps*. In logic point of view, the system comprises eight computing clusters, which consists of four of AMD Opteron 6174 inside 12 cores running at 2.2 *GHz*, that is, 48 cores a node sharing 128 *GB* RAM in four memory-controller non-uniform memory access architecture, and a large memory cluster that includes four of AMD 6136, which comprises eight core running at 2.2 *GHz*, that is, 32 cores a node sharing 256*GB* RAM. There are 25,600 computing cores of AMD Opteron 6100 in this system and the maximal Linpack achieves at 177 *TFlops* and at #42 places in TOP500 list in June 2011[8].

C. The NCHC Formosa IV

The Formosa IV is the NCHC's self-madehigh-performance computing cluster system. The Formosa IV GPU architecture would be improving specific applications performance different the ALPS service general-purpose applications.

The Formosa IV consists two of INTEL X5670 inside six cores running at 2.93*GHz*, that is, 12 cores a node sharing 48 *GB* RAM. Each node inside three the NVIDIA M2070 GPU, that have 515.2*Gflops* performance in a GPU device. The Formosa IV connected together with a Voltaire InfiniBand in 4x QDR (40*Gb/s*).

It was offers an aggregate performance of over 70.4*Tflops* that #234 in TOP500 lists in Nov. 2011[8].

III. Software Descriptions

A. BLASBench

BLASBench[4] is designed to evaluate the performance of Basic Linear Algebra Subroutines (BLAS). The BLAS is standard programming interface for publishing libraries to perform basic linear algebra operations such as vector and matrix multiplication.The HPL Benchmark[7] was adapted to build upon this DGEMM. HPL employs the LU decomposition to solve a dense $N \times N$ system of linear equation in a floating point workload of$\frac{2}{3}N^3 + 2N^2$. HPL utilizes LU factorization with row partial pivoting to solve a dense linear system while using a two-dimensional block-cyclic data distribution for load balance and scalability. In this paper, we compare DGEMM performance on different machines.

B. CacheBench

Cache is a buffer between the processor and memory.The system cache is designed to achieve fast response of the system performance.As we know size of cache, bandwidth from memory, amount of memory affect performance of HPL benchmark. The CacheBench[4] is designed to evaluate the performance of the memory hierarchy of computer systems. CacheBench specific focus is to performance of possibly multiple levels of cache.It purposes to establish peak computation rate given optimal cache reuse and to verify the effectiveness of high levels of compiler optimization. In this paper, we compare six tests are as follows:

- Cache Read is measure bandwidth of read for varying vector lengths.
- Cache Write is measure bandwidth of write for varying vector lengths.
- Cache Read/Modify/Write is generates read, modify and write bandwidth for varying vector lengths.
- Memset from the C library - provides performance of the C memset() function.
- Memcpy from the C library - provides performance of the C memcpy() function.

IV. Performance Results

In this paper, we are using some of benchmark software to evaluate system performance. The DGEMM is a subroutine in the Basic Linear Algebra Subprograms (BLAS) which performs matrix multiplication that is the multiplication for double precision. We adopt Open64and AMD ACML for NCHC ALPS. The SUN GPU and Formosa IV used INTEL compiler and INTEL MKL library for measure performance.

The cache size and theoretical peak performance is shows in Table I.Fig. 1 shows DGEMM results in *Mflops* that presents the SUN GPU and the Formosa IV performance are unstable before dimension 295, because that illustratea fluctuation in small computation. In this result, we obtainedthe APLS, SUN GPU and Formosa IV DGEMM efficiency has percentage of 93.26, 96.5 and 95.83 respectively.

TABLE I
CACHE SIZE AND PEAK PERFORMANCE OF DIFFERENT CLUSTERS

	Cache	*Peak performance*
ALPS	64K,512K,12MB	8.8 GHz
SUN GPU	32K,256K,8MB	11.72 GHz
Formosa IV	32K, 256K, 12MB	11.72 GHz

The cache read is implemented to measure bandwidth for different vector size, as show in fig.2.In this result, read bandwidth is descending on length of 32*KB* and 256*KB* in the Formosa IV and SUN GPU Cluster both. Besides, read performanceraise on vector size 64*KB* but declining on 256*KB*in the ALPS Cluster.

Cache write presented performance for write bandwidth performance descending on vector length of 24*KB* and 256*KB* in the Formosa IV and SUN GPU Cluster both, as show in fig. 3.Moreover, we gained unsatisfactory resultsfrom read and write test in the ALPS Cluster. Because it architecture is diskless operating system that means load data from remote storage and process data in ram disk.

Cache read-write-modify benchmark is created to measure bandwidth for different vector lengths, as show in fig. 4.In this result, it shows performance decliningon vector size 4*KB*, 32*KB* and 256*KB* in the Formosa IV and SUN GPU Cluster both. The ALPS cluster bandwidth performance is descendon 64*KB* and 512*KB*.

Fig. 5 presents copy areas of memory by using C function memcpy. In this result, the Formosa IV and SUN GPU Cluster memcpy performance decline onvector size 12*KB*and 128*KB*respectively. Nerveless, it affects performance by diskless operating system in the ALPS. The C library provides function memset to initialize regions of memory show in fig.6.In this result, the Formosa IV and SUN GPU Cluster is declining on vector size 24*KB* and 128*KB* and then 64*KB* and 384*KB* in the ALPS.

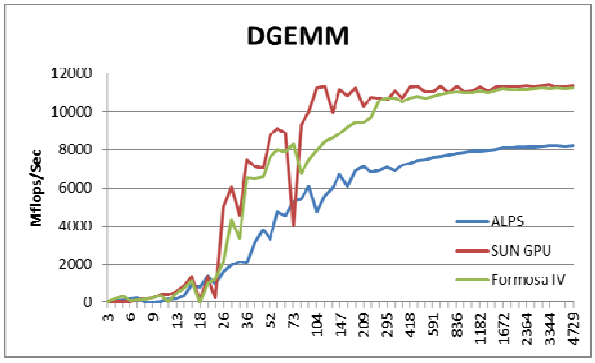
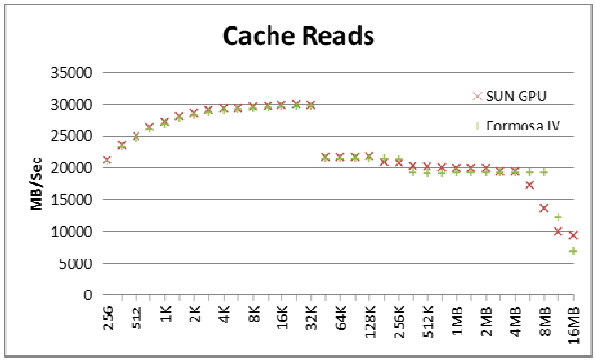
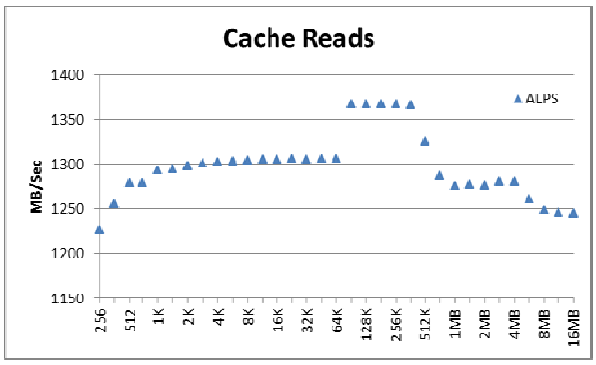


Fig. 1 Performance of Matrix-matrix Multiplication

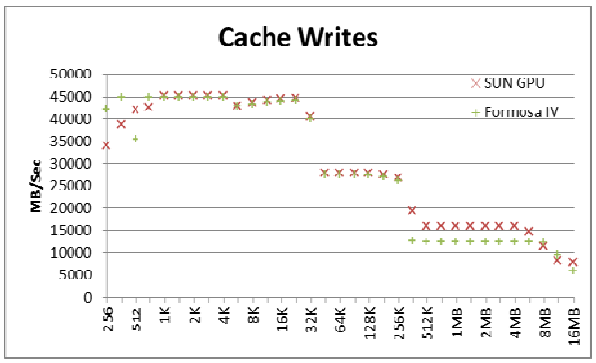


(a)

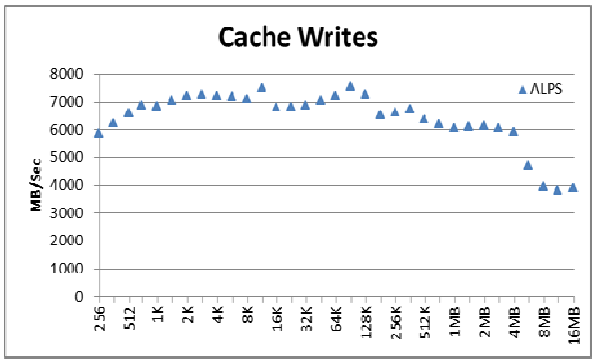


(b)

Fig. 2 Cache Performance of read at (a)SUN GPU cluster and Formosa IV (b) ALPS cluster

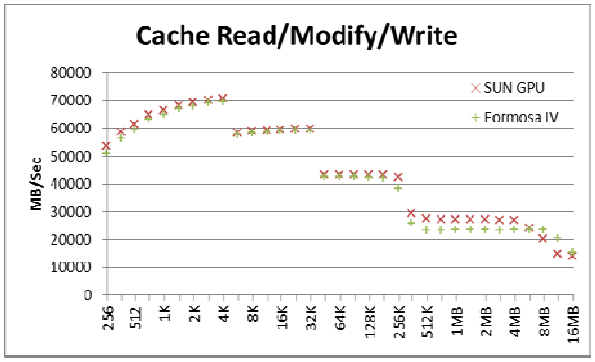


(a)

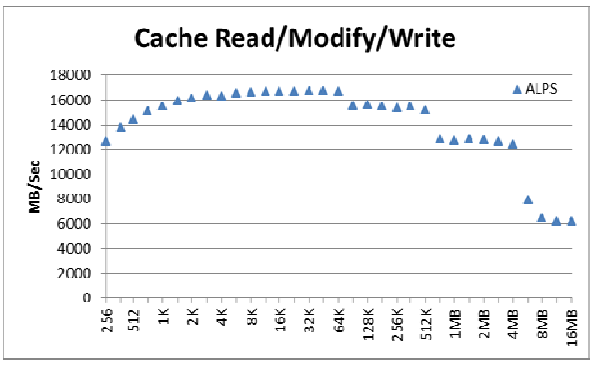


(b)

Fig. 3 Cache Performance of write at (a)SUN GPU cluster and Formosa IV (b) ALPS cluster.

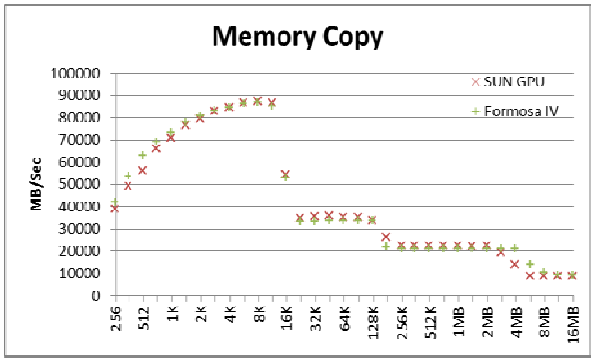


(a)

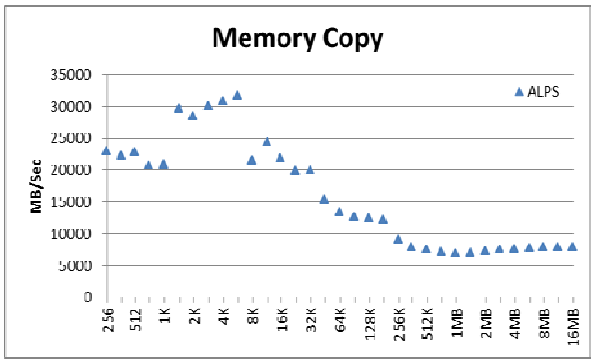


(b)

Fig. 4Cache Performance of read, modify and write at (a)SUN GPU cluster and Formosa IV (b) ALPS cluster

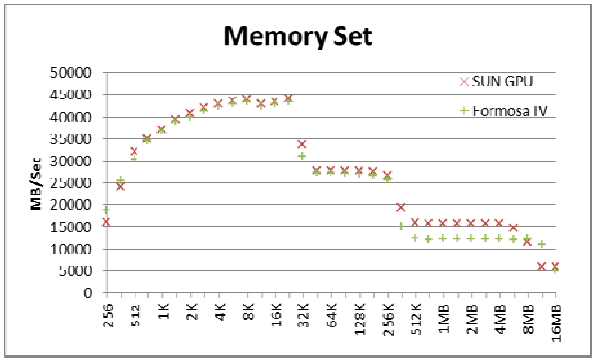


(a)

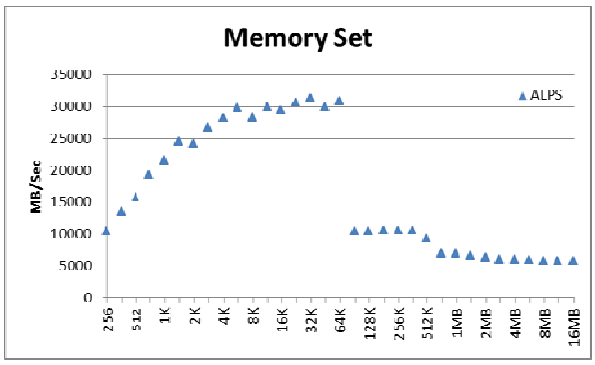


(b)

Fig. 5 Memory Performance of Copy at (a)SUN GPU cluster and Formosa IV (b) ALPS cluster



(a)



(b)

Fig.6 Memory Performance of set at (a)SUN GPU cluster and Formosa IV (b) ALPS cluster

V. Conclusion

This paper presents our experience in benchmarking andperformance evaluation of the NCHC HPC clusters, which can be helpful for those interested for high-performancescientific and engineering computing.Most supercomputer used HPL to benchmark their system performance, in accordancewith the requirement of the TOP500 List. There are various reasons effective system performance results in benchmarking. Such as optimization of operating system, interconnect latency, compiler options, version of the compiler, size of cache, bandwidth from memory, amount of memory and etc. can affect the performance even when the processors are the same.

Although the ALPS HPC cluster has bigger cache size, but as result diskless operating system could be affect performance.On the other hand, L3 cache seems can't be achieved, because t bandwidth performance descending on 4*MB* in the SUN GPU Cluster and declining on 8*MB* in Formosa IV. In this paper shows memory access experiment result in NCHC HPC Clusters. We hope these works will provide usefulinformation for future development and construct cluster system.

References

[1] Chau-Yi Chou, Hsi-Ya Chang, Shuen-Tai Wang, Te-Min Chen, and Chang-Hsing Wu, "Benchmarking and Performance Analysis on a Blade Cluster,"The 3rd Workshop on Grid Technologies and Applications (WoGTA), 2006.

[2] Kuo-Chan Huang, Hsi-Ya Chang,Cherng-YeuShen,Chaur-Yi Chou andShou-Cheng Tcheng, "Benchmarking and performance evaluation of

NCHC PC cluster,"The Fourth International Conference/Exhibition on HighPerformance Computing in Asia-Pacific Region (HPC Asia 2000 Conference), pp. 923-928, Beijing, China, May 14-17, 2000.
[3] Chin-Hung Li, Te-Ming Chen, Ying-Chuan Chen, and Shuen-Tai Wang, "Formosa3: A Cloud-Enabled HPC Cluster in NCHC," International Conference on Electrical, Computer, Electronics and Communication Engineering, 2011.
[4] P.J. Mucci and K.S. London, "Low Level Architectural Characterization Benchmarks for Parallel Computers," Technical Report UT-CS-98-394, Univ. of Tennessee, 1998.
[5] Douglas M. Pase, "Linpack HPL Performance on IBM eServer 326 and xSeries 336 Servers," IBM, 2005.
[6] T. Sterling, D. Becker, D. Savareseet. al., "BEOWULF: A ParallelWorkstation for Scientific Computation," Proc. Of International Conf: On Parallel Processing (ICPP), August 1995.
[7] HPL, Available at: http://www.netlib.org/benchmark/hpl/.
[8] TOP500 Lists, Available at : http://top500.org
[9] The Message Passing Interface (MPI) standard, Available at: http://www.mcs.anl.gov/research/projects/mpi/