Topic 20
Parallel I/O and Storage Technology

Rajeev Thakur, Rolf Hempel, Elizabeth Shriver, and Peter Brezany

Topic Chairpersons

Introduction

In recent years, it has become increasingly clear that the overall time to completion of parallel applications may depend to a large extent on the time taken to perform I/O in the program. This is because many parallel applications need to access large amounts of data, and although great advances have been made in the CPU and communication performance of parallel machines, similar advances have not been made in their I/O performance. The densities and capacities of disks have increased significantly, but improvement in performance of individual disks has not followed the same pace. For parallel computers to be truly usable for solving real, large-scale problems, the I/O performance must be scalable and balanced with respect to the CPU and communication performance of the system.

The parallel I/O and storage research community is pursuing research in several different areas in order to solve the problem. Active areas of research include disk arrays, network-attached storage, parallel and distributed file systems, theory and algorithms, compiler and language support for I/O, runtime libraries, reliability and fault tolerance, large-scale scientific data management, database and multimedia I/O, realtime I/O, and tertiary storage. The MPI-IO interface, defined by the MPI Forum as part of the MPI-2 standard, aims to provide a standard, portable API that enables implementations to deliver high I/O performance to parallel applications.

The Parallel I/O Archive at Dartmouth, `http://www.cs.dartmouth.edu/pario` is an excellent resource for further information on the subject. It has a comprehensive bibliography and links to various I/O projects.

Papers in This Track

The Parallel I/O and Storage Technology track at Euro-Par 2000 contains six papers that address different aspects of the I/O problem:

1. "Towards a High-Performance and Robust Implementation of MPI-IO on top of GPFS," by Jean-Pierre Prost, Richard Treumann, Richard Hedges, Alice Koniges, and Alison White describes IBM's implementation of the MPI-IO standard for the GPFS file system.

A. Bode et al. (Eds.): Euro-Par 2000, LNCS 1900, pp. 1251–1252, 2000.

2. "Design and Evaluation of a Compiler-Directed Collective I/O Technique," by Gokhan Memik, Mahmut T. Kandemir, and Alok Choudhary presents a compiler-directed collective I/O approach that detects opportunities for using collective I/O in a program and inserts the appropriate collective I/O calls.
3. "Effective File-I/O Bandwidth Benchmark," by Rolf Rabenseifner and Alice E. Koniges describes a benchmark designed to measure the effective I/O bandwidth achievable by applications on a given parallel machine and file system.
4. "Instant Image: Transitive and Cyclical Snapshots in Distributed Storage Volumes," by Prasenjit Sarkar presents an algorithm for handling snapshots of storage volumes in a distributed storage system.
5. "Scheduling Queries for Tape-Resident Data," by Sachin More and Alok Choudhary investigates issues in optimizing I/O time for a query whose data resides on an automated tertiary storage system containing multiple storage devices.
6. "Logging RAID – An Approach to Fast, Reliable, and Low-Cost Disk Arrays," by Y. Chen, W. Hsu, and H. Young presents a disk-array architecture that uses logging techniques to solve the small-write problem in parity-based disk arrays.