



Parallel Data Laboratory

Department of Electrical and Computer Engineering
School of Computer Science

The PDL

Since 1993, the Parallel Data Laboratory (PDL) has led academic research in storage systems and information infrastructures. With contributions from 12 faculty, 30 students, and 11 staff, the PDL's multifold research agenda includes projects in several areas, including:

Storage Administration

Human administration of storage is a large and growing issue in modern IT infrastructures. PDL's *self-* storage* project explores new storage architectures that integrate automated management functions and simplify the human administrative task; the resulting infrastructures are self-configuring, self-organizing, self-tuning, self-healing, self-managing, etc. Borrowing organizational ideas from corporate structure and technologies from AI and control systems, self-* storage will simplify storage administration, increase system robustness, and simplify system construction.

Data Center Operations

The *Data Center Observatory* (DCO) is a fully-instrumented operational data center at Carnegie Mellon, simultaneously used as a utility and as a research vehicle. Its computation and storage resources are a utility for research activities and real services. Researchers measure all operational aspects, including power/thermal, human administrator time, resource utilization patterns, failures and their consequences, and so on. Technologies and tools are developed and deployed to mitigate operational costs, and measurements quantify their efficacy and guide refinement.

Device-Embedded Security

As self-contained systems, storage devices can execute security functions independent of client OSs. PDL's *self-securing storage* system exploits this fact to help system administrators with detection, diagnosis, and recovery from client system intrusions. From behind the device interface, a device can watch for suspicious behavior, prevent undetected tampering, and provide a history of activity. Storage devices can also perform access control checks. PDL's general *self-securing devices* project also explores embedded security functions for other devices, such as network cards.

Survivable Distributed Storage

PDL's *PASIS* project develops storage architectures that survive failures and malicious compromises of servers. The PASIS data access protocols minimize common-case overheads and allow per-object selection of fault model and encoding scheme, enabling versatile cluster-based storage. The project also explores the associated trade-offs among performance, availability, and security properties to enable automated approaches to configuring and reconfiguring these systems. Automating failure prediction and problem diagnosis increases survivability further.

Database Systems

The PDL has several database system thrusts, including the *Fates storage manager* that transparently matches database access patterns to device-specific features to maximize their efficiency automatically new storage organizations, prefetching techniques for improving memory system per-

formance, and staged query processing architectures for improving robustness under load.

Storage Experimentation Tools

PDL has developed a set of inter-related tools for measuring and experimenting with storage designs. *DIXtrac* is an automated disk-characterization tool that discovers over 100 performance-critical disk parameters. *Disksim* is a flexible and accurate storage system simulator that uses DIXtrac's output to closely predict modern disk performance. The *Memulator* is a timing-accurate storage emulator that makes simulated components look real to existing OSs and applications.

Major Recent Projects

PDL's *Network-Attached Secure Disk* (NASD) project developed the architecture underlying today's object-based storage systems. PDL's *Active Disk* and *Active Storage Networks* projects explored migrating of intelligence into infrastructure devices. PDL researchers explored the impact of *MEMS-based storage* on storage systems.

For further information contact:

Greg Ganger, Director
(412) 268-1297 (phone)
(412) 268-3010 (fax)
ganger@ece.cmu.edu

William Courtright
Executive Director
(412) 268-5485
wcourtright@cmu.edu
Karen Lindenfesler
PDL Business Administrator
(412) 268-6716
karen@ece.cmu.edu