

OceanStore: An Architecture for Global-Scale Persistent Storage

Presented By David Deschenes
January 28, 2003

A True Data Utility

- Utility services provided by a confederation of companies
- Service providers buy and sell capacity among themselves
- Consumers pay a monthly fee for access to the utility

Potential Applications

- Transparent data synchronization across devices
- Groupware and personal information management tools
- Distributed design tools
- Large digital libraries and repositories for scientific data

System Goals

- Uniform and highly-available access
- Support of nomadic data
- Ability to survive major disasters and regional outages
- Ability to be constructed from an untrusted infrastructure

System Requirements

- Connectivity must be provided to client devices
- Information must be kept secure from theft and denial-of-service
- Information must be extremely durable
- Archiving should be automatic and reliable
- Information must be divorced from location
- Servers must be geographically distributed

System Design

- Servers are linked together to form a mesh
- Fundamental unit of storage is a persistent object
- An object is identified by a GUID (hash of the object's name and its owner's key)
- Objects are replicated on multiple servers, and referenced by pointers propagated throughout the server mesh
- Users interact in sessions, or sequences of read and write operations

System Features

- Objects may be located quickly using a probabilistic search or reliably using a slower hierarchical method
- Updates are secure and are propagated to replicas for consistency
- Introspective data collection is used for future optimization
- Provides multiple APIs to ease porting existing applications

Importance of this Paper

- Introduces the concept of a cooperative utility for global-scale persistent storage
- Discusses interesting approaches to addressing the problems of data availability and survivability in the face of disaster
- Investigates the uses of introspection for optimization of a self-organizing system