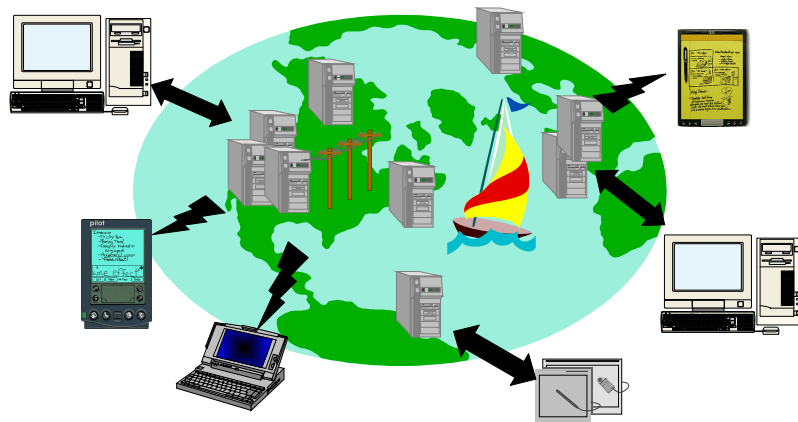


Option 2: The Oceanic Data Utility: Global-Scale Persistent Storage



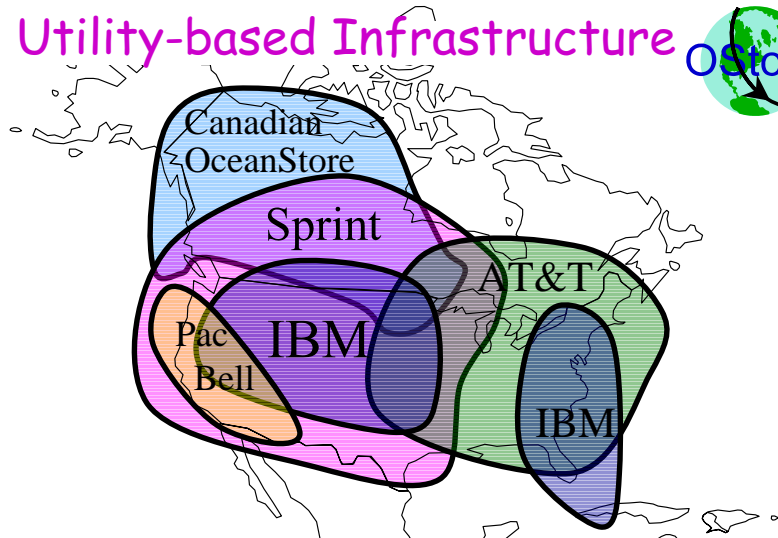
John Kubiawicz

Ubiquitous Devices ⇒ Ubiquitous Storage



- Consumers of data move, change from one device to another, work in cafes, cars, airplanes, the office, etc.
- Properties **REQUIRED** for Endeavour storage substrate:
 - **Strong Security:** data must be encrypted whenever in the infrastructure; resistance to monitoring
 - **Coherence:** too much data for naive users to keep coherent "by hand"
 - **Automatic replica management and optimization:** huge quantities of data cannot be managed manually
 - **Simple and automatic recovery from disasters:** probability of failure increases with size of system
 - **Utility model:** world-scale system requires cooperation across administrative boundaries

Utility-based Infrastructure



- Service provided by confederation of companies
 - Monthly fee paid to one service provider
 - Companies buy and sell capacity from each other

State of the Art?



- Widely deployed systems: NFS, AFS (/DFS)
 - Single "regions" of failure, *caching only at endpoints*
 - ClearText exposed at various levels of system
 - Compromised server ⇒ all data on server compromised
- Mobile computing community: Coda, Ficus, Bayou
 - Small scale, fixed coherence mechanism
 - Not optimized to take advantage of high-bandwidth connections between server components
 - ClearText also exposed at various levels of system
- Web caching community: Inktomi, Akamai
 - Specialized, incremental solutions
 - Caching along client/server path, various bottlenecks
- Database Community:
 - Interfaces not usable by legacy applications
 - ACID update semantics not always appropriate

OceanStore Technologies III: Introspective Monitoring and Optimization

- Requirements:
 - *Reasonable* job on a global-scale optimization problem
 - Take advantage of locality whenever possible
 - Sensitivity to limited storage and bandwidth at endpoints
 - Stability in chaotic environment
- OceanStore Technologies:
 - Introspective Monitoring and analysis of relationships:
 - between different pieces of data
 - between users of a given piece of data
 - Rearrangement of data in response to monitoring:
 - Economic models with analogies to simulated annealing
 - Sub problem of Tacit Information Analysis (option 5)

OceanStore Technologies IV: Rapid Update in an Untrusted Infrastructure

- Requirements:
 - Scalable coherence mechanism which provides performance even though replicas widely separated
 - Operate directly on encrypted data
 - Updates should not reveal info to untrusted servers
- OceanStore Technologies:
 - Operations-based interface using conflict resolution
 - Use of incremental cryptographic techniques: No time to decrypt/update/re-encrypt
 - Use of oblivious function techniques to perform this update (fallback to secure hardware in general case)
 - Use of automatic techniques to verify security protocols

Two-Phase Implementation:

- Year I: Read-Mostly Prototype
 - Construction of data location facility
 - Initial introspective gathering of tacit info and adaptation
 - Initial archival techniques (use of erasure codes)
 - Unix file-system interface under Linux ("legacy apps")
- Year III: Full Prototype
 - Final conflict resolution and encryption techniques
 - More sophisticated tacit info gathering and rearrangement
 - Final object interface and integration with Endeavour applications
 - Wide-scale deployment via NTON and Internet-2