Advanced Operating Systems Lecture notes

Dr. Clifford Neuman
Dr. Tatyana Ryutov
Dr. Dongho Kim
University of Southern California
Information Sciences Institute

CSci555: Advanced Operating Systems

Lecture 14 – Contemporary Topics
3 December 2004

Dr. Dongho Kim
University of Southern California
Information Sciences Institute

Administrative

- Class evaluations before today's break
- Instructions for submitting paper on the course web page
  - _Hardcopy – no electronic version
  - _12/5 - no penalty through 5PM 12/12
- Final exam Thursday December 11 at 11AM
  - Details will be sent to class
  - Exam is comprehensive
  - Answer the questions asked
  - Bring paper with name and ID#
  - Separate sheet per question

Today's Lecture

- Advances in Perspective
- USC's Computing Environment
- Internet Search Techniques
- Securing today's systems
- Ubiquitous computing
- Sensor Networks
- Grid Computing
- Peer to Peer
- The Semantic Web
- Current work at ISI

Advances in Perspective

- Operating Systems
  - Virtual systems
  - Ubiquitous applications
- Distributed Systems
  - "System" expands
- Ubiquitous Computing
  - Virtual systems
  - "System" turns inward and contracts, while reach of the system expands.
- Disintermediation leads to reintermediation
  - Agents are the new intermediaries

USC's Computing Environment

- Several NFS File Servers
  - Accessed by Sun's in lab
  - Samba and other file "gateways" supported
- NIS used for login authentication
  - But users registered with Kerberos when they sign up or change passwords
  - Kerberos used for back-end data access through web interfaces
- DNS and LDAP both supported
- Mail service tied to directory
- More bandwidth than most other universities
- Separate network for administrative use

Internet Search Techniques

- Issues
  - How much of the net to index
    - _How much detail
    - _How to select
  - Relevance of results
    - _Ranking results – avoiding spam
    - _Context for searching
      - –Transitive indexing
- Scaling the search engines

Internet Search Techniques - Google

- Data Distribution
  - Racks and racks of servers running Linux – key data is replicated
    - _Some for indices
    - _Some for storing cached data
  - Query distributed based on load
  - Many machines used to for single query
- Page rank
  - When match found, ranking by number and quality of links to the page.

Securing Today's Systems

- Security technologies are well understood
  - Software bugs and configuration errors are the dominant vulnerabilities
  - Policy is not well understood
- Denial of service
  - Is the main kind of attack that we don't know how to prevent
  - Physical DOS attacks resisted through redundancy.
  - Online DOS attacks require a way to distinguish legitimate traffic from attacks, and this is hard to do.

Ubiquitous computing

- According to Mark Weiser at Xerox:
  - Transparent computing is the ultimate goal
  - Computers should disappear into the background
  - Computation becomes part of the environment

Ubiquitous Computing

- Computing everywhere
  - Desktop, Laptop, Palmtop
  - Cars, Cell phones
  - Shoes, Clothing, Walls (paper / paint)
- Connectivity everywhere
  - Broadband
  - Wireless
- Mobile everywhere
  - Users move around
  - Disposable devices

Ubiquitous Computing

- Structure
  - Resource and service discovery critical
  - User location an issue
  - Interface discovery
  - Disconnected operation
  - Ad-hoc organization
- Security
  - Small devices with limited power
  - Intermittent connectivity
- Agents
- Sensor Networks

Grid Computing

- Federated system
  - No single controlling authority
- Scheduling
  - Processors, bandwidth and other resources
  - Scheduling already discussed in lectures
- Policy is an important issue
  - Reliability, security, of who can use, and what one is willing to use

Grid Computing: Systems and Apps

- Systems
  - Globus toolkit
    - GRAM, GSI, MDS, GASS, HBM, Nexus, globus_io
  - Legion
  - Condor
  - Related but not grid – CORBA, DCOM, DCE
- Applications
  - Seti at home
  - Smart instruments
  - Teraflop desktops
  - Distributed supercomputing

What's different about Peer-to-Peer

- Non peer to peer environment
  - Client-Server (bipartite) trust model
    - Server's trusted, clients aren't
      - This was never a good trust model anyway
    - Goal of security is to protect the servers
      - And the clients data on the servers
  - Servers are more available than clients
    - When you can't contact a server it is more likely to be a problem on the client's side
  - Server side security policy
  - Client side software configuration

What's different about Peer-to-Peer

- Peer-to-Peer Assumptions
  - Many servers are clients
    - Not more trusted than other users
      - Need policy to tell us which can be trusted
      - Policy will affect selection of servers (configuration)
    - Certificates and credentials help the client decide the extent to which a "server" should be trusted.
  - Trust issues are similar to those in administratively decentralized distributed systems
    - But may have even less trust than in another organization's servers.
  - Trust issues extend beyond traditional security
    - Reliability, service guarantees, recourse for failure

Policy in Peer-to-Peer networking

- Policies associated with many entities
  - "Server" policies on access to local machine
    - Which client can access a peer
  - Application object policies associated with stored objects or running processes
    - Control access to the objects
    - Often set when process or object is created on "peer".
  - Client side policies used to select servers
    - And to set object policies on "server"
    - The policies may be combine with the "server" policies on the node.

P2P File Sharing Issues

- Naming
- Data discovery
- Availability
- Security
  - Encryption
  - Fault tolerance
- Conflict resolution
- Replication

Peer to Peer file sharing

- Napster
  - P2P sharing with central D/S
- Gnutella
  - P2P sharing with distributed D/S
- Servent (SERVer+cliENT)
  - _Bearshare
  - _Gnutella
  - _LimeWire
- Edonkey
  - MFTP: Multisource File Transfer Protocol

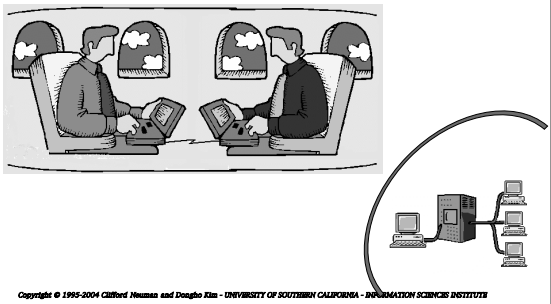
Peer to Peer File Sharing

- Farsite from Microsoft
- OceanStore from UC Berkeley

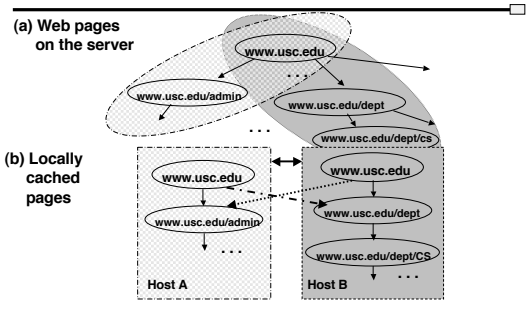
Other Peer to Peer Technologies

- Ad-hoc networking
  - _ Untrusted nodes used to relay messages
  - _ Multiple routes (distributed and replicated)
  - _ Extends range, reduces power, increases aggregate bandwidth.
  - _ Increases latency, management more difficult.
- Sensor networks
  - _ An application of ad-hoc networking
  - _ Add processing/reduction in the network

P2P Reconstruction



P2P Reconstruction



Grouping for Resonstruction

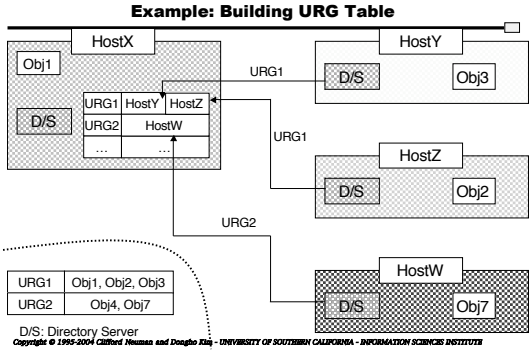
- Finding scattered objects from clients is expensive
  - Keeping location information for individual objects and/or searching for them is expensive

- Group objects and maintain hints about them
  - Reduce:
    - _ The size of database, and
    - _ The required communication
  - by keeping location information only for groups not for individual objects

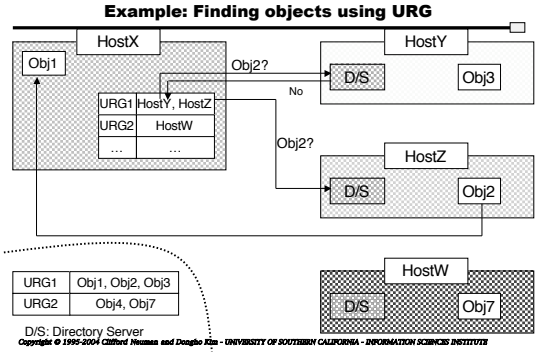
Grouping (continued)

- Group related objects
- Name the group with a URG (Uniform Resource Group name)
- Maintain URG table per host
- The table has the list of hosts that contain objects that have the same URG
  - A host does not necessarily have all the objects with the same URG
- Groups can overlap
  - An object can have multiple URGs

Example: Building URG Table



Example: Finding objects using URG



Semantic Web and XML

- Machine parsed web pages
  - Provides greater structure to data exchanged through web pages.
  - Closure issues apply to the semantics of data.
  - Supports annotation of fields
  - RDF Triples (Object, Attribute, Value)
- XML – Extensible Markup Language
  - Meta tags
  - SGML
  - HTML

Future of OS's

- As we move toward ubiquitous computing and integrated applications, technologies like .net, CORBA, and XML will increase programmatic interactions across protection boundaries
  - Basic technologies are just new names for old technologies, but… unsolved problems…
  - OS Boundaries will blur
    - _ Both TCB boundaries and
    - _ Layer boundaries
  - This enables significant improvment in capability to operate across system boundaries
  - But it creates a more complex policy environment and complicates security issues.

Current OS Research at ISI

- Computer Security
  - Policy and the GAA-API
    - _ Simplifying Policy Specification
    - _ GridSec, Trust Negotiation
    - _ Intrusion detection and response
  - Denial of service detection and countermeasures
    - _ DETER testbed
  - Secure DNS
- File systems
  - Disconnected operation
- Networking: optical, space, active, overlays, simulation, sensor