Collaboration Systems Design: Lessons Learned and Challenges

Atul Prakash

Department of EECS
University of Michigan, Ann Arbor
aprakash@eecs.umich.edu

My background

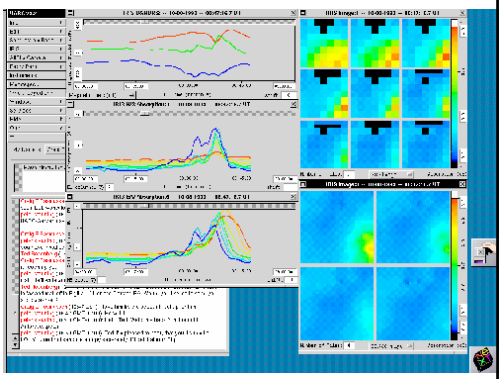
- Infrastructure for CSCW systems
  - DistEdit toolkit for group editors [CSCW'90]
  - Undo in group editors [CSCW'92, ACM TOCHI Dec'94]
  - DECAF: real-time object-oriented groupware [IEEE TC'98]
  - Multimedia session capture and replay [Manohar&Prakash'96]
  - Fault-tolerance in collaboration [SRDS'98, ICDCS'99]
  - Modular approach to building CSCW systems [CSCW'00]
- Participated in the design of space science collaboratories at Michigan: UARC and SPARC
- More recent work: Security policy management and group security [McDaniel&Prakash'99, McDaniel&Prakash'02]
- Goal: To share with you some experiences and observations

Presentation Outline

- **History of Collaboratory systems at Michigan**
- Some observations and challenges: my top 10 list

Upper Atmospheric Research Collaboratory – 1993-96

- Quick-view Sondrestrom data
- Group chat
- DistView Toolkit: Window replication on multiple desktops (CSCW'94)
- CBE: Collaboratory Builder's Environment (CSCW'96)
- Objective-C, NextStep
- Good proof-of-concept
- Many innovations



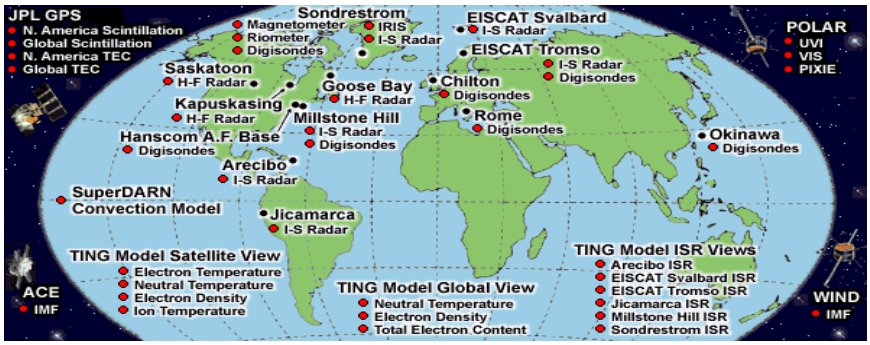
UARC -- 1997

- Increase community size and leverage the web.
  - Get away from NextStep: Java-based via Netscape browser
- Many research contributions
  - Used CBE architecture philosophy (CSCW'96)
  - Started running into data format issues
  - QoS-sensitive data distribution research (Litiu&Prakash'98)
  - Fault-tolerant servers, clients research (Litiu&Prakash'98)
  - Increased support for asynchronous users (Manohar's thesis, 98)
    - Persistent chat and other collaboration tools
    - Session capture and replay research
- Some loss of functionality:
  - images and files could previously be easily exchanged via NextStep chat via cut-and-paste into the Chat window.
- http://www.eecs.umich.edu/distview
- Images: Link to various snapshots

Space Physics and Aeronomy Research Collaboratory, SPARC (1998-2002)

- Get away from Java (not robust enough on browsers): went to HTML
- Several innovations:
  - Shared, editable web pages, with fine-grain access control
  - Many more instruments
  - Many more data types (CDF, Cedar, Jpg..)
  - Simple HTTP-based protocols to collect data from various sources
- Some steps back in CS technology deployed:
  - Images computed at servers
  - Pull-based: screen refresh
  - Fewer display controls at clients
  - No notion of rooms (collaboration workspaces)
- Nevertheless, perceived to be more robust, had more data sources, and higher usage during campaigns (sometimes, less technology can have higher acceptance!)
- Spinoff: Additions to the Windows-to-the-Universe project for K-12 students. Very successful during those years, especially "Ask-the-Scientist" part
- Some parts transitioned to NCAR after end of NSF funding

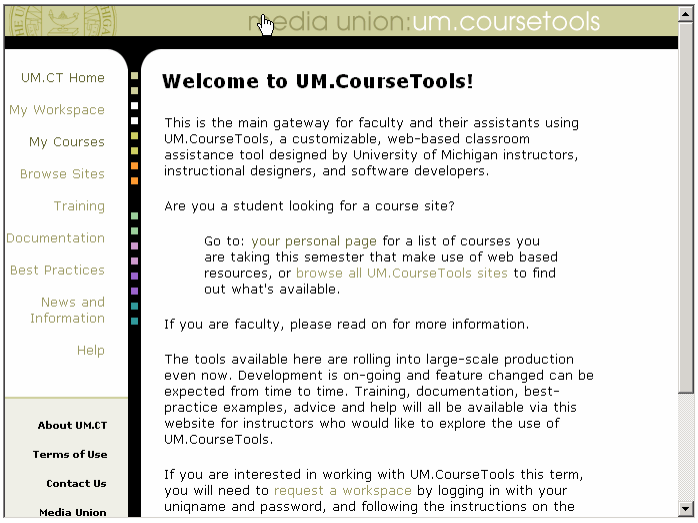
Online instruments



UM.CourseTools (Fall 1999-now)

- Developed by the SPARC team from experience developed in UARC/SPARC
- Focus on helping instructors creating customized web sites for courses, and supporting collaborative activities discussion groups, grading, etc.
- Link: http://coursetools.ummu.umich.edu

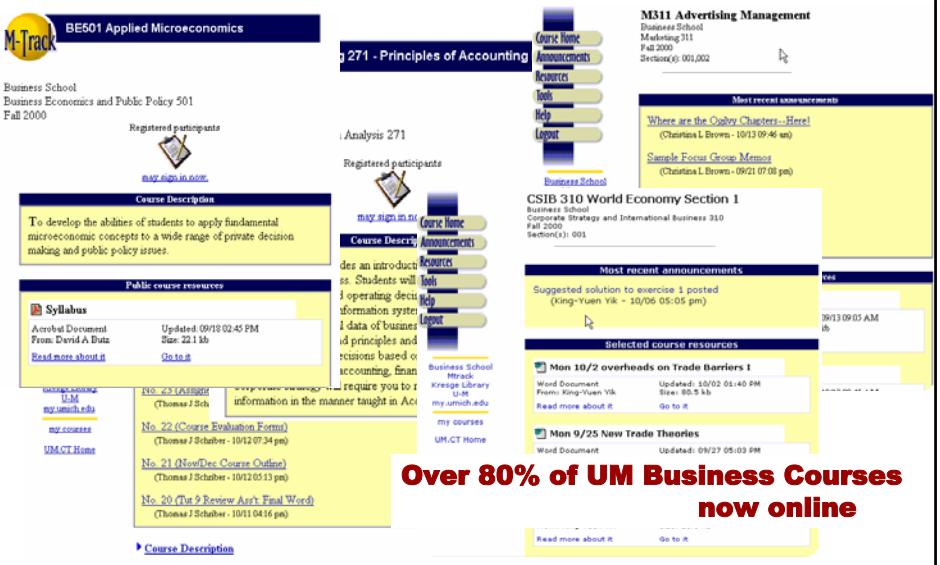
(Source for Coursetools info on following slides: from a presentation by Joseph Hardin)



Features

- Student view versus instructor view
- Linking with the enrollment database
- Discussions with instructor override
- Online homework submission/grade distribution
- Announcements & Resources can be added and automatically show on the home page
- Instructors can handle exceptions
  - Adding guest students
  - Custom home page view, etc

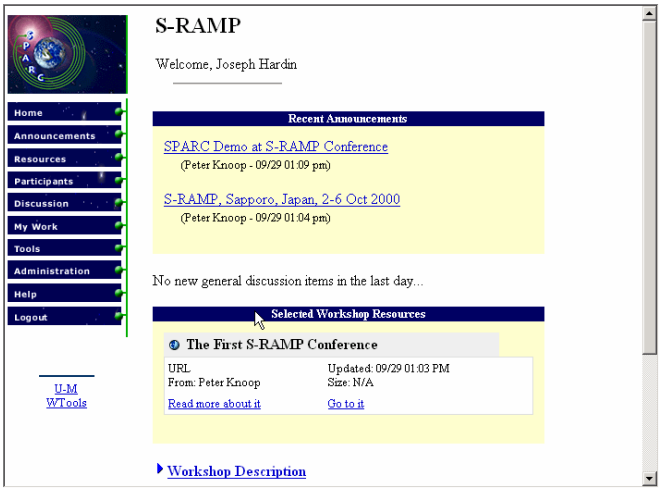
UM.CT Business School Effort - Results



Over 80% of UM Business Courses now online

UM Worktools (followed CourseTools)

Support For Research Groups



- Link: http://worktools.si.umich.edu

Support for Work Groups

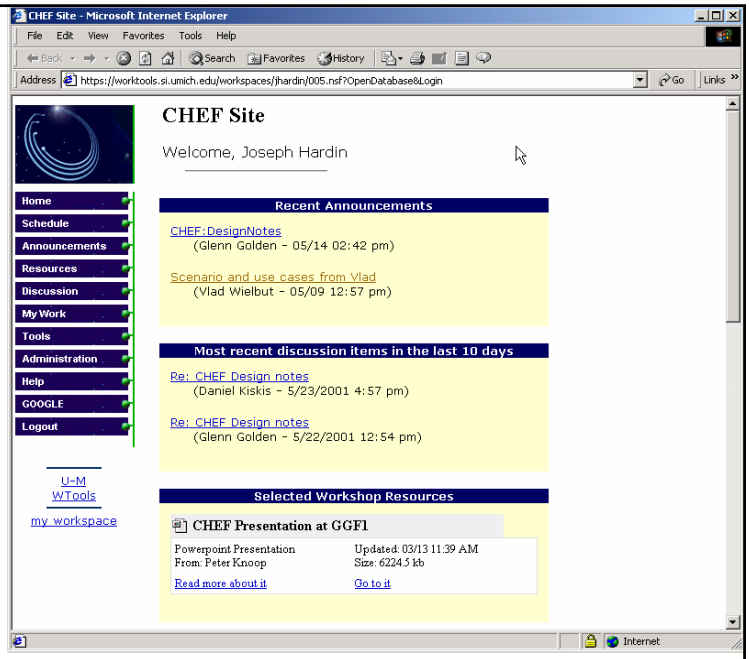
UM.Worktools

Scientific Research Groups

Science Review Teams

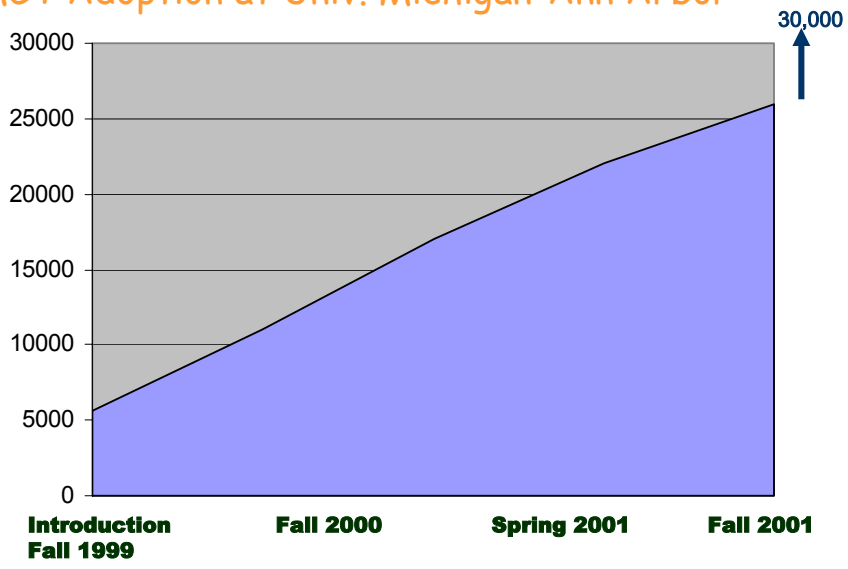
Technology Design Teams

Currently over 3000 users

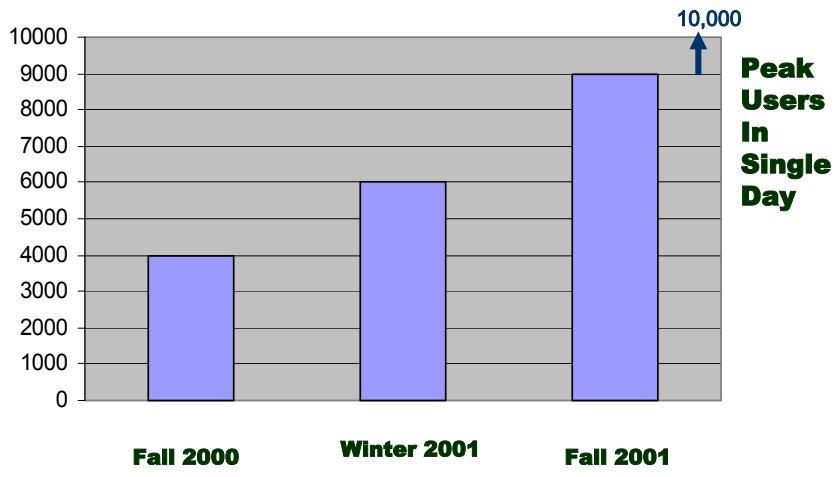


worktools.si.umich.edu

UMCT Adoption at Univ. Michigan-Ann Arbor

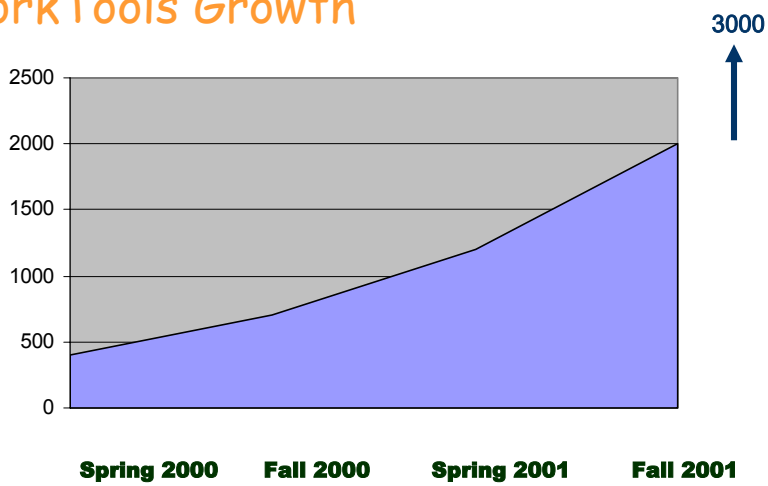


Peak Usage



- 2003: seems firmly entrenched and is considered institutional IT infrastructure

WorkTools Growth



- 2003: Still less pervasive than CourseTools. Increasing use at the college level to support committees: minutes, agenda, and discussions
  - Automatic email notifications of new resources posted
  - Discussions: still often via email in many groups because of concern that not everyone has made or ready to make the switch to regular use

Other Collaboratory Project Examples at Michigan

- Collaboratory on Technology Enhnaced Learning Communities
- Great Lakes Center for AIDS Research
- Network for Earthquake Engineering Simulation (NEESGrid)
- CHEF: Comprehensive Collaborative Framework
  - Combines elements of Worktools and SPARC
  - Integration with the Grid

Presentation Outline

- History of Collaboratory systems at Michigan
- **Lessons learned/observations and some research to watch for impact on collaboration systems in 5-10 years: my top 10 list**

#10: Technology Adaption Challenge is Underestimated

- All technology, but particularly collaboration technology can be particularly challenging to get buy-in into
  - Cultural shift
  - May not fit into existing way of doing things
  - Heterogeneity in client capabilities (Masic'03)
  - Often requires mass adaption to succeed
- Difficulty often underestimated by both CS researchers and enthusiastic domain folks

Adaption Experience

- UARC and SPARC
  - Remains a niche tool -- little evidence of continued use after end of financial support. (Recent transition of technology to NCAR is a good step)
- CourseTools:
  - Developed under SPARC funding as well and finding heavy use at the University across all colleges
- Worktools:
  - Combination of shared file system and discussion group.
  - Finding use as an augmentation to Email by committees and groups for exchanging large documents, archiving minutes, etc. across administrative boundaries and providing access control
- Difficult to predict long-term impact though. Several ideas designed into UARC/SPARC will probably still find wide acceptance
  - chat tools being increasingly seen in web pages,
  - asynchronous messaging support integration likely in instant messaging systems,
  - Web pages could be a great medium of collaboration if editing them can be made a lot easier (e.g., with tablet-based PCs and drag-and-drop features)
  - Earlier work on CSCW at Michigan being cited in patent applications and used to fight granted patents

#9: Light-weight multimedia collaboration

- Most of the collaboration today is still text-based and asynchronous
- Why is multimedia collaboration a challenge?
  - Hetrogeneity among clients:
    - Client hetrogeneity
      - QoS reconcilation among members/infrastructure [open issue]
      - Adapting a session to varying quality levels [Marsic at Rutgers]
  - What if some users have to miss a meeting?
- Promising directions:
  - Mobile multimedia collaboration
    - May start to become promising because of pervasiveness (smarter cell-phones) and increasing b/w.
  - Multimedia asynchronous collaboration
    - Some early indicators of potential demand: exchange of photos/music
    - Meeting minutes
    - Some early work done on multimedia session capture and replay [Manohar&Prakash'96]
    - Key Challenge: summarization and searching.

#8: Understand requirements and show clear benefits to BOTH individuals and groups

- SPARC experience
  - Prior culture in the community on data ownership, sharing, willingness to embrace new technology
  - Tried to create a paradigm shift. Successful with early adaptors, but was difficult to cross the chasm
- CourseTools
  - Clear Benefits for many professors and instructors:
    - Instructors were already creating course home pages. Now they could do it more easily
    - Instructors could protect course material when desired, unlike at normal course web sites
    - Tie-in into University's Kerberos system and course registration database
    - Even HTML-novice instructors could create high-utility course web sites

#7: For success, think BIG. Support MANY communities …

- Instant messaging core is simple
  - Buddy lists or groups
  - Chat
- But it was successful because MANY such groups could be created at a server by users
- Group chat technology existed before IM, but did not get as widely adapted
- Another comparison at Michigan: SPARC (one-community) vs. CourseTools (many communities)
- It would have been interesting to see how SPARC would have evolved if other communities could dynamically create their own collaboratories using the platform

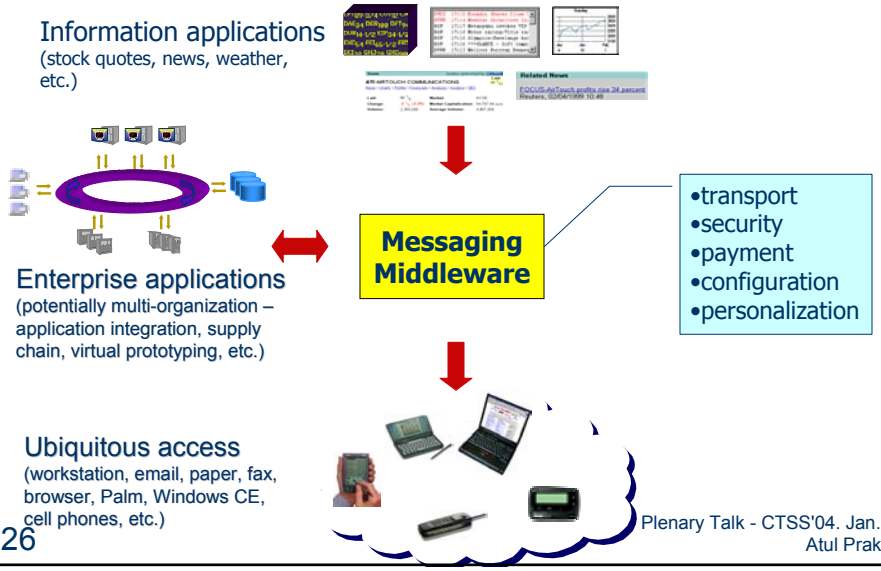
#6: User education and administrative support can be crucial

- CourseTools could have failed or grown much more slowly without one-day training sessions
- Challenge: harder to provide such education remotely, especially for new paradigms to the community
  - E.g., in SPARC, many of the useful features, such as the ability to create your own shared pages, could have found wider set of users if there were mechanisms to provide proactive hands-on training regularly

#5: Distinguishing good data from bad data

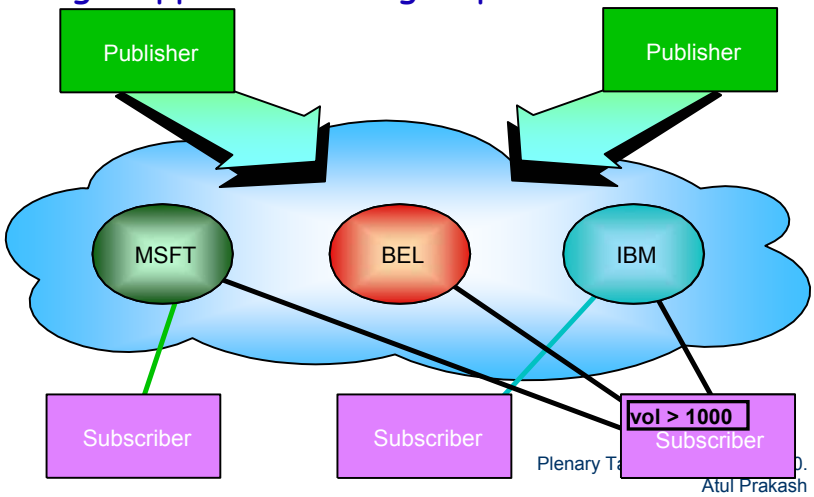
- Currently
  - Systems make it easy to create data for sharing and post it to servers. But…
    - Little incentive to delete the data or maintain it
    - Often, no reliable indicator of ownership. Easy for malicious users to attack systems by contributing corrupted data (e.g., KAAZA)
    - Who should delete the data?
    - Who bears the cost?
- What is good data for one user is useless data for another
  - Finer-grain data delivery paradigms may be become important, E.g., Content-based publish subscribe systems
  - Recommender Systems

Role of Messaging



Content-based Publish-Subscribe Systems

- Example systems: SIENNA, Gryphon
- Challenge: applications to groupware

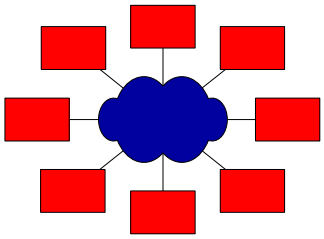


Challenge #4: Security and Privacy

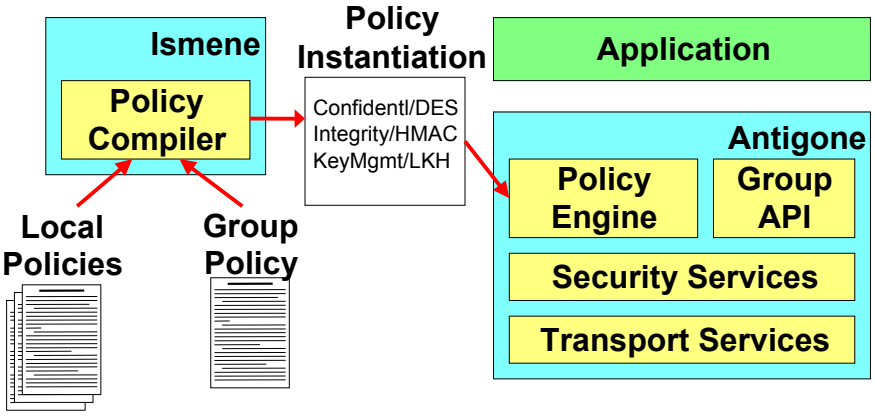
- Security needs to be pushed to the end points
  - Less trust requirements in servers
    - End-to-end confidentiality and source authenticity of data
  - Firewalls at desktops
    - Fewer implications of perimeter firewall breach
- Security needs to be largely transparent
  - PGP or S/MIME is not so good if explicit user action required
  - HTTPS (via SSL) succeeds because users do not usually have to do anything different. Less success likelihood if users had to type "https" explicitly every time they wanted security
- DoS attacks on communities may become an issue as standards develop for building communities
- Support for selective anonymity, with group-authenticity (e.g., anonymous to some users)
- Preventing harassment, tracking of activities, etc.

Challenge #3: Simplify Group Security and QoS Management

- Different organizations have different security policies
- Collaboration across organizations can be a major challenge
- Challenges:
  - Security Policy reconciliation [McDaniel&Prakash'2002]
  - Extension to negotiating QoS issues
  - Handling dynamic policy changes and group membership changes
- A real example:
  - DoD: How does a person manage multiple instant-messaging sessions at different security clearance levels or among different security categories?
  - How are changes in access policies handled dynamically?



The Antigone/Ismene Project at Michigan



Challenge #2: Better modeling and evaluation

- How to make collaboration technology more of a science and engineering discipline?
  - Important contributions tend to be new paradigms of collaboration
  - But little in terms of taking a new paradigm of collaboration and being able to compare it with other paradigms
  - Need better models of collaboration and collaboration infrastructure. Example of previous attempts:
    - 2x2 matrix (distance vs. time)
    - Coordination theory
    - Group editor models
- Little today in terms of benchmarks
- Little systematic testing and evaluation techniques, except often expensive controlled user studies.
- How do we know what the right collaboration system is for an organization?
- How do we configure such a system together rapidly from components?

Challenge #1: New Collaboration paradigms

- Email:
  - Adding presence awareness
  - Adding file sharing capability
- Instant messaging:
  - Adding the ability to exchange attachments
  - Supporting disconnected users
- Peer-to-peer infrastructures (fewer things at servers):
  - Supporting controlled communities
  - Adding conferencing capabilities
  - More security and better QoS/reliability support

Finally, …

- "IT Doesn't Matter": Contrarian view by Jonathan Carr, Harvard Business Review, 5/2003.
- Key point:
  - IT fast becoming a commodity, Increased spending in many areas (replacing hardware every 2 years, for example)
    - May providing little technical edge
    - May make organization less competitive
- IT administrators should focus on reducing technology risks:
  - Security
  - Reliability and availability
  - Wait for robustness, ease-of-use, and adaption of available technologies, rather than quick technology shifts
- Researchers need to also place more emphasis on security, reliability, and finding ways to increase technology adaption, especially for collaborative infrastructure