The clearinghouse: a decentralized agent for locating named objects in a distributed environment

- Mechanism for uniform naming of objects (users, resources etc.)
- Nice paper to articulate some of the naming concerns
- Built on top of grapevine (email system for data distribution)
  - As a product, this was probably not a good idea (silly window syndrome, distributed deadlock - was sending about 1 message per day :-))
  - Many research papers, including epidemic algorithms, anti-entropy policies came off of this effort
- Relevant to our discussion because naming and location are integral to using a distribute store (of objects)

Analogy to telephone system

- Hierarchical with area codes. Internally only telephone numbers are significant - users however care about intuitive names
- Phone numbers can be inconsistent, numbers are mostly hints (not sure how much of it is true these days, but addressbook entries are typically “hints”)

Naming conventions

- Absolute
- Relative
- Hierarchical
  - Levels of hierarchy
- Aliasing for flexibility
- Client perspective: A single global database, client stub interacts with different clearinghouse servers
- Binding strategies
  - Static
  - Early binding
  - Late binding

Distributed name lookup

- Queries move up the hierarchy and then downward
- Sideways movement possible
- Distributed updates
  - Idea is for one of these updates to win-out. Performace issues in who wins
    - Peer anti-entropy algorithms - when two peers meet, they reduce entroy (differences amongst themselves). Peer processes eventually lead of global consistency
    - Epidemic algorithms - who to perform anti-entropy with. Modeled after spreading diseases. Vary agtressiveness depending on neighborhood activity