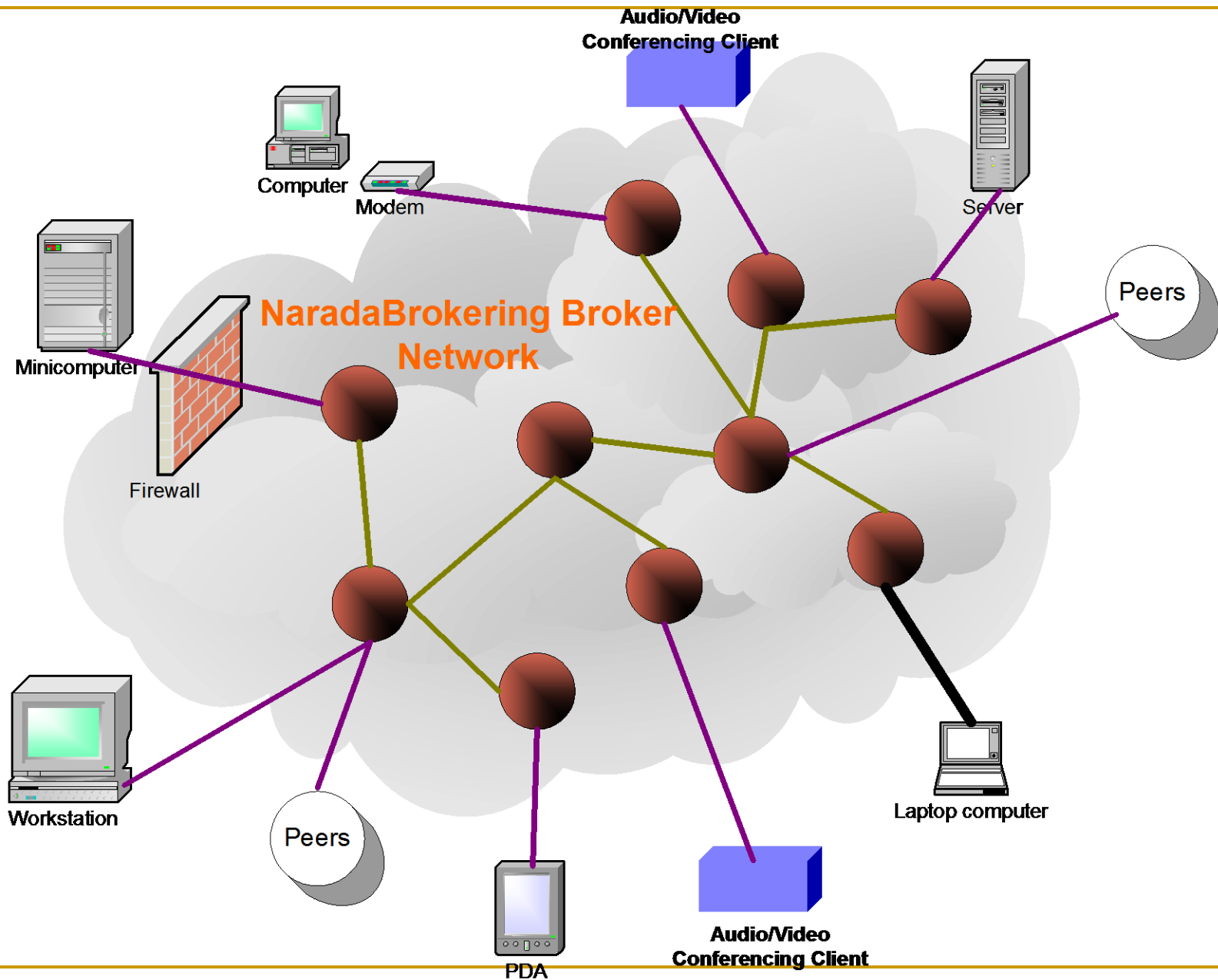# NaradaBrokering: Managing data distribution in distributed systems

Shrideep Pallickara

Community Grids Lab

Indiana University

Data Distribution Problem

- Data dissemination
  - Registration of interest
  - Receipt of data when we need it
- Radio, TV, Cell-phones
- Routing data in distributed systems.
- What, where, when, why and how.

NaradaBrokering Broker Network

Audio/Video Conferencing Client

Computer

Modem

Server

Peers

Minicomputer

Firewall

Workstation

Peers

PDA

Audio/Video Conferencing Client

Laptop computer

VLab Workshop: July 20-23 2005   **http://www.naradabrokering.org**

Problems unique to our situation

- Rate and size of data is not fixed.
- Data Selectivity
- Data Volumes
- Producer-Consumer roles
  - Clear de-coupling of roles.
- Best-effort is not enough!

Data Distribution in Distributed Systems: Characteristics (**I**)

- I don't know who will be sending me the data, but I know what I want and don't.

- Send me everything that I am interested in.

- Don't send me anything that I am not interested in.

- Don't reveal/send my data to someone else.

Data Distribution in Distributed Systems: Characteristics (**II**)

- Send me what I want, in a timely manner, when they occur

- Make sure data is secure

- Manage and interoperate with multiple data formats.

- Global scope of dissemination

- Producer's ability to restrict who receives data

System Constraints

- Large number of clients/entities.
- System is in constant flux.
- Very high data rates
- Voluminous data sizes
- High selectivity
- Susceptibility to hardware/software failures
- Guaranteed delivery
- Preserve data ordering

Packaging data

- Data is packaged within events (akin to a mailing envelope)

- Events encapsulate data and include descriptors related to routing, transport protocols, timestamps, content synopsis etc.

- Events thus have expressive power at multiple levels.

Additional NaradaBrokering Services

- Replay and Recording
- Buffering and time-spacing
- Discovery
- Timing (synchronized clocks on nodes)
- Fragmenting/Coalescing large payloads
- Compressing/Decompressing data
- HPSearch Project: Easy deployment of large broker networks.

NaradaBrokering Application domains

- Distance education

- Real-time audio/video conferencing

- Grid applications

- GIS settings

- Web Service infrastructures

- Visualization services

Visualization Services

- Substrate discovers services that are capable of performing requested visualization task(s)

- Can cope with multiple replicas of visualization servers

- Substrate preferentially deploys replica that is most under-utilized

  - Facilitates dynamic, real-time load-balancing

- Can sustain failures to replicas

Final Comments

- Open source technology
- Deployed in wide variety of settings
- Multiple entry points to avail of capabilities and services offered by the substrate
- Additional Info
  http://www.naradabrokering.org/