Engineering Research Center for
Collaborative Adaptive Sensing of the Atmosphere

Collaborative P2P Systems for Distributed Data Fusion & Beyond

H. M. N. Dilum Bandara and Anura P. Jayasumana
Colorado State University

Dilum Bandara
dilumb@engr.colostate.edu


University of Massachusetts Amherst


University of Oklahoma

Colorado State University


University of Puerto Rico Mayaguez

CASA is primarily supported by the Engineering Research Centers Program of the National Science Foundation under NSF award number 0313747.

Research Goal

Data/resource search solutions for resource intensive collaborative P2P systems

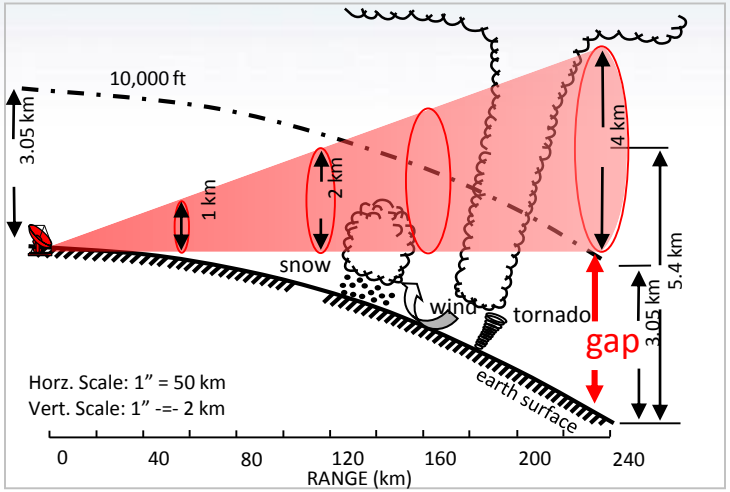
- Resources – radars, CPU, storage, bandwidth, algorithms
- Resource intensive – processing, storage, bandwidth
- Collaborative – thinking beyond current P2P

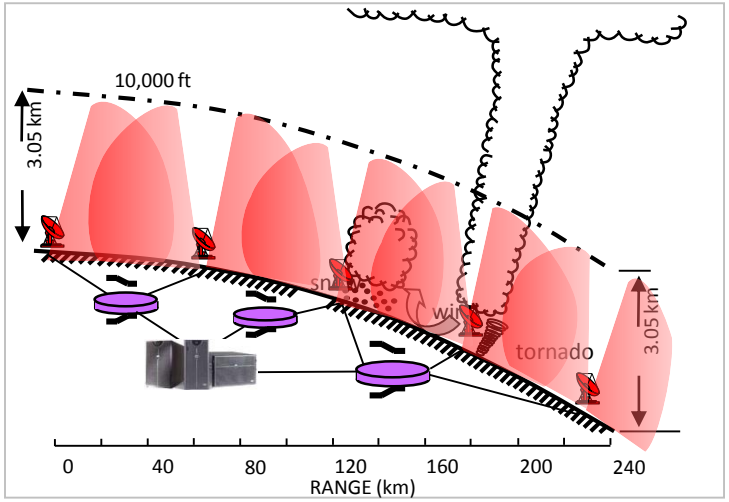
Outline

- CASA
  - Radar networking
- Collaborative P2P systems
  - Distributed data fusion
- Research focus
- Potential applications

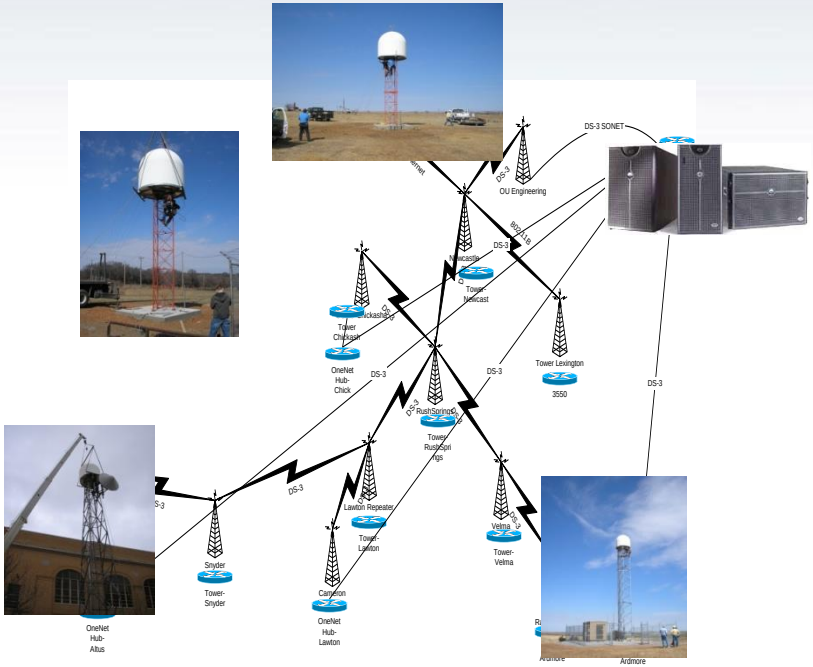
Collaborative Adaptive Sensing of the Atmosphere (CASA)

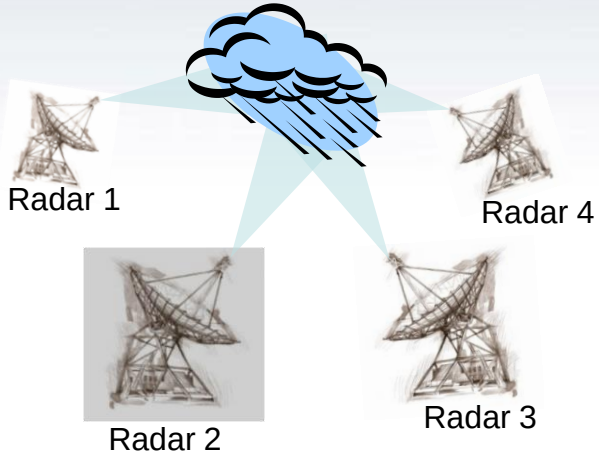
- Concept
  - A network of small radars instead of one large radar
  - Sense lower 3 Km of atmosphere
  - Collaborating & adapting radars
    - Improved sensing, detection, & prediction
- CASA goal
  - Improve warning time & forecast accuracy for hazardous weather

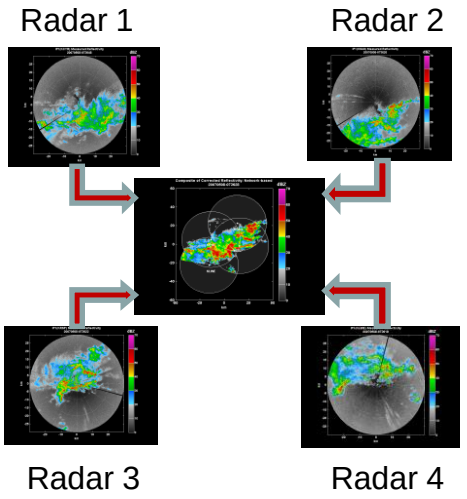




CASA Oklahoma Test-bed

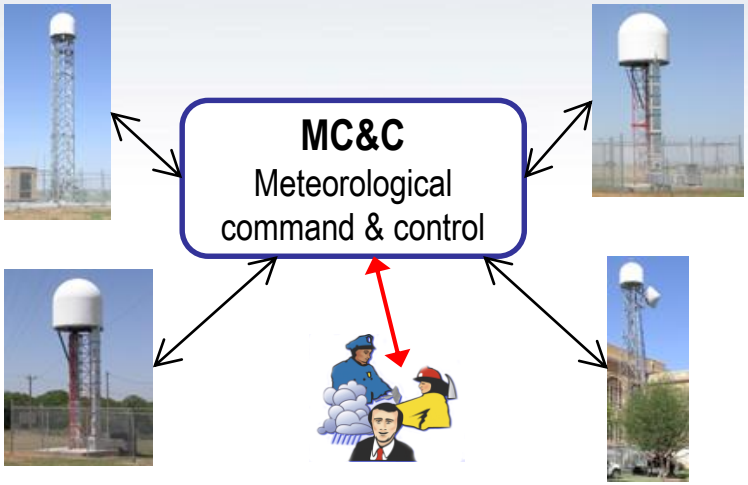


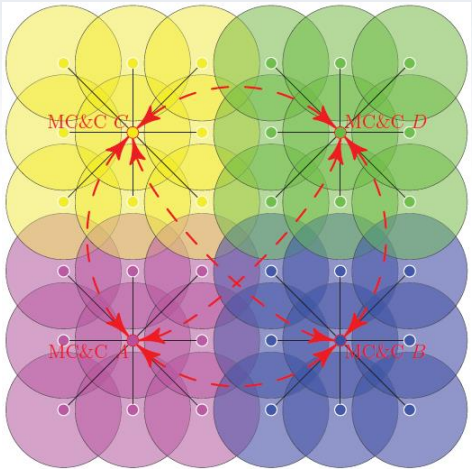




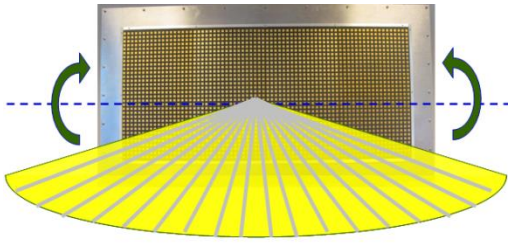
- Multiple high bandwidth streams
- Real-time communication
- Simultaneous observations by multiple radars
- Multi-sensor data fusion
- Heterogeneous infrastructure & end users
- Hostile weather conditions

Large-Scale CASA Deployments

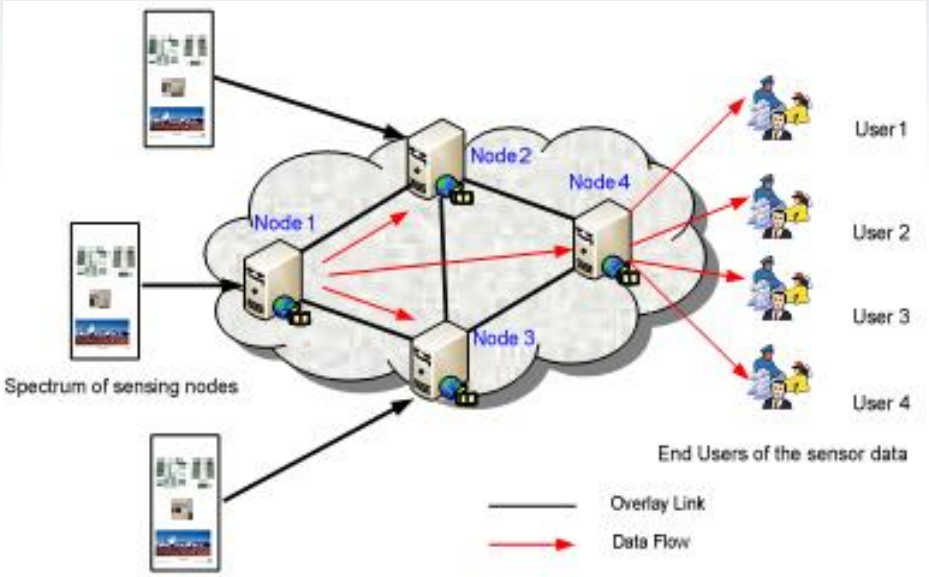


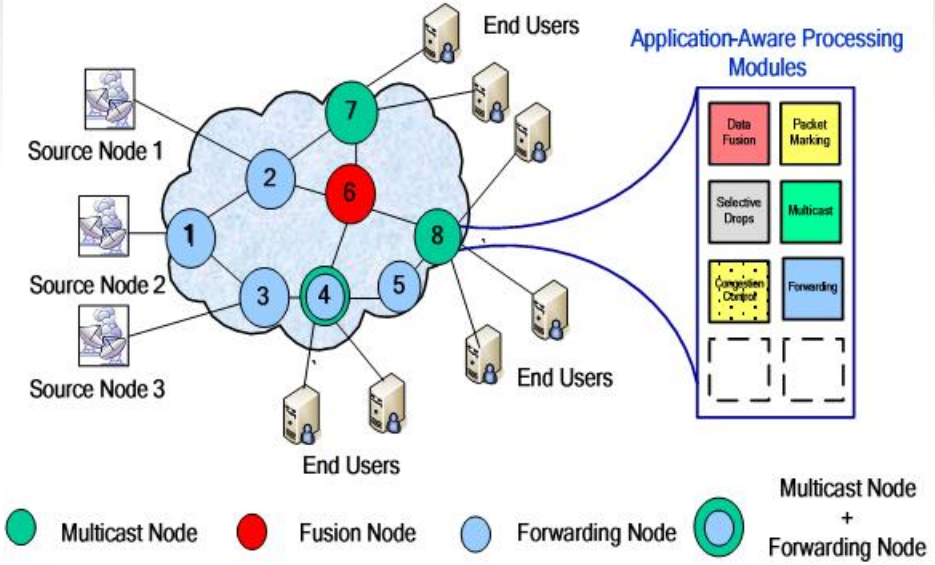


- Large-scale CASA deployments are lot more computation, bandwidth, & storage intensive
- New solid-state radar data rates in Gbps
- Distributed & heterogeneous resources
- Increased resource utilization

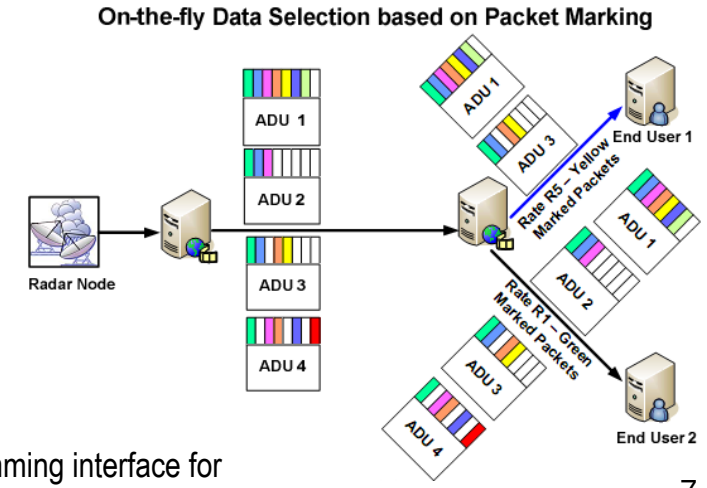


CASA Radar Networking



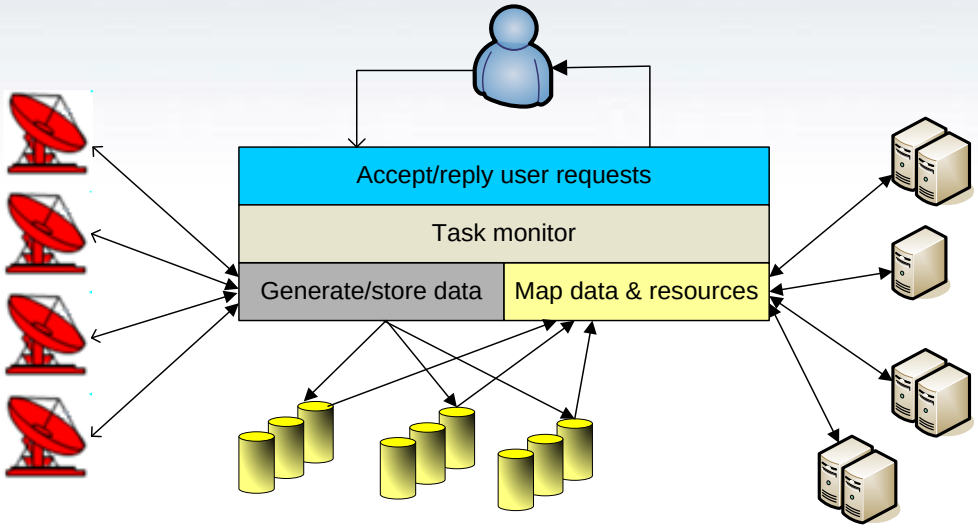


- Application-aware overlay networks
  - Application-aware packet marking
- In-network data fusion
- API for application-aware service deployment
- Data-fusion latency estimation



T. Banka, P. Lee, A. P. Jayasumana, and J. F. Kurose, "An architecture and a programming interface for application-aware data dissemination using overlay networks," COMSWARE 2007, Bangalore, Jan. 2007.

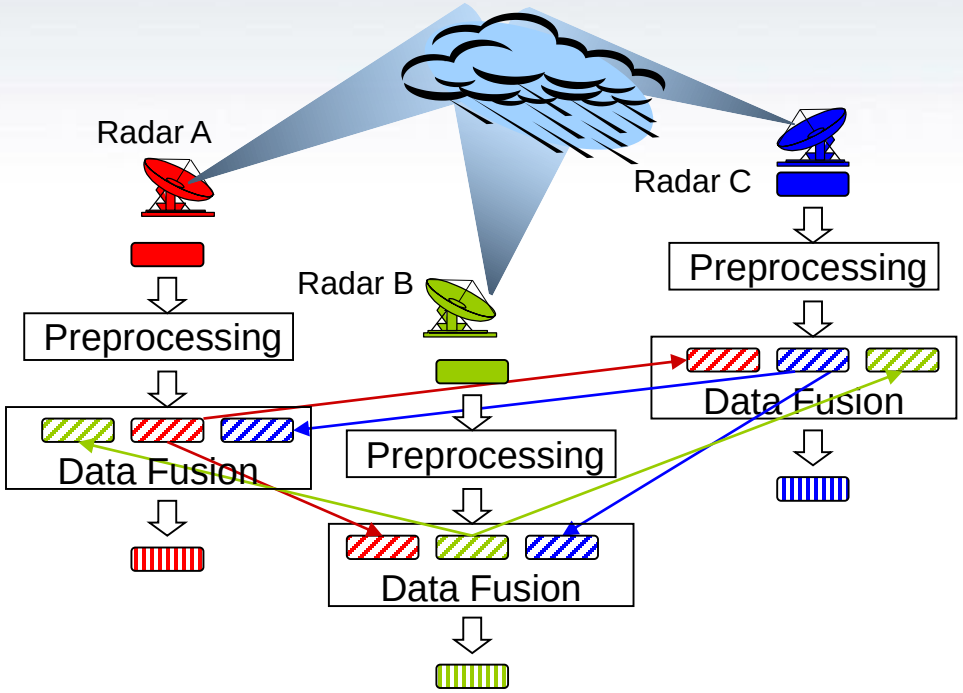
Collaborative P2P for CASA Data Fusion

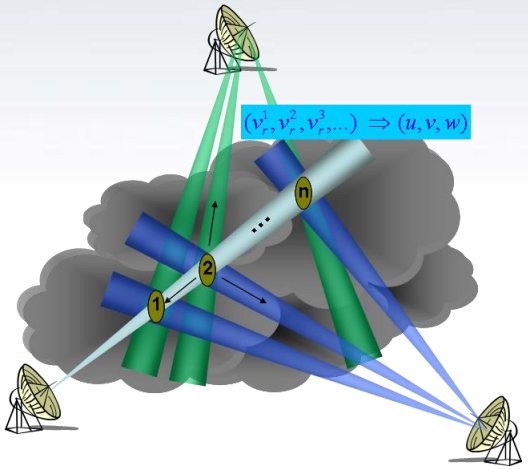


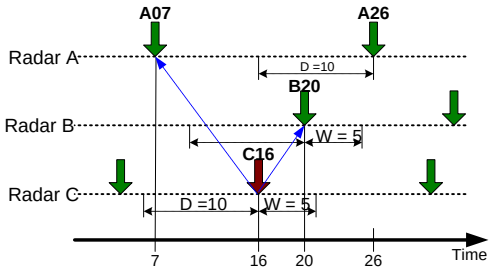
Resources
- Radars
- CPU & memory
- Storage
- Bandwidth
- Algorithms

- *How likely is it to snow in Fort Collins within next 6 hours?*
- Distributed data generation, dissemination, resource aggregation, & fusion
- Push & pull data
- Peers with divers capabilities performing different tasks
- Engage in a greater task that cannot be accomplished by individual peers

P2P Collaboration Framework

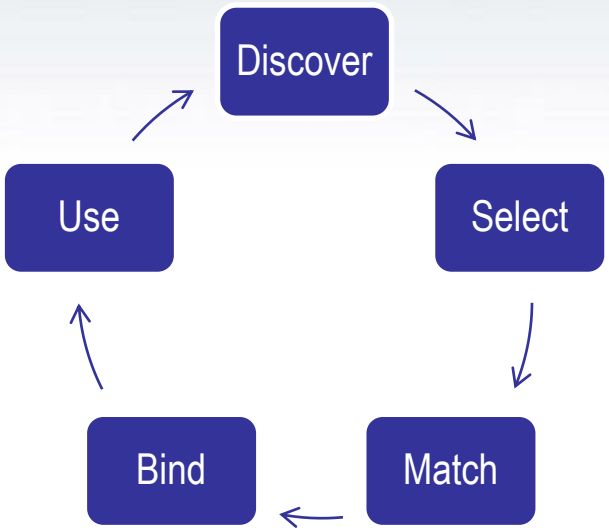


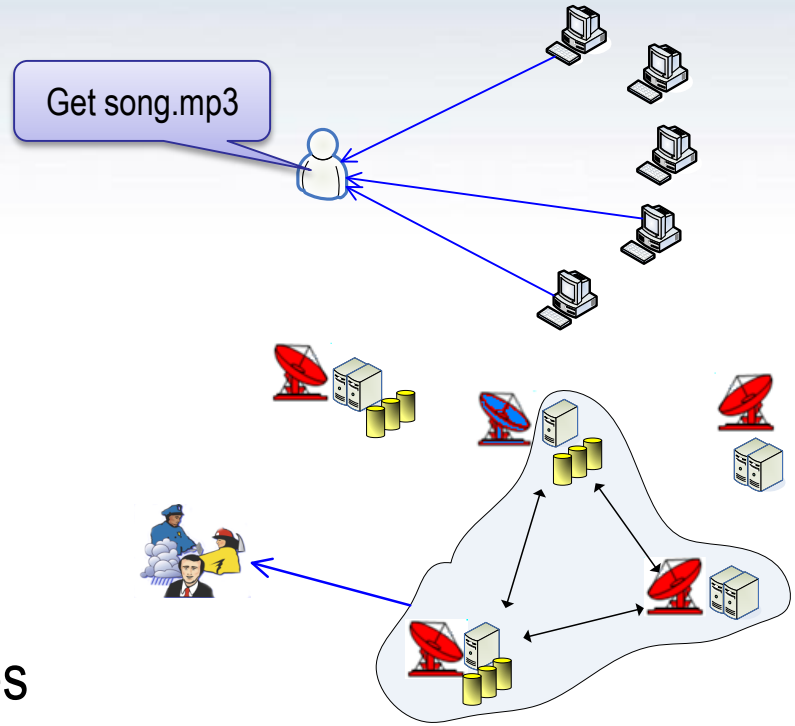




- Distributed data fusion
- Radars depend on each other's data to correct/detect errors
- Locate peers with relevant data
- Locate resources in a timely manner

Resource Aggregation



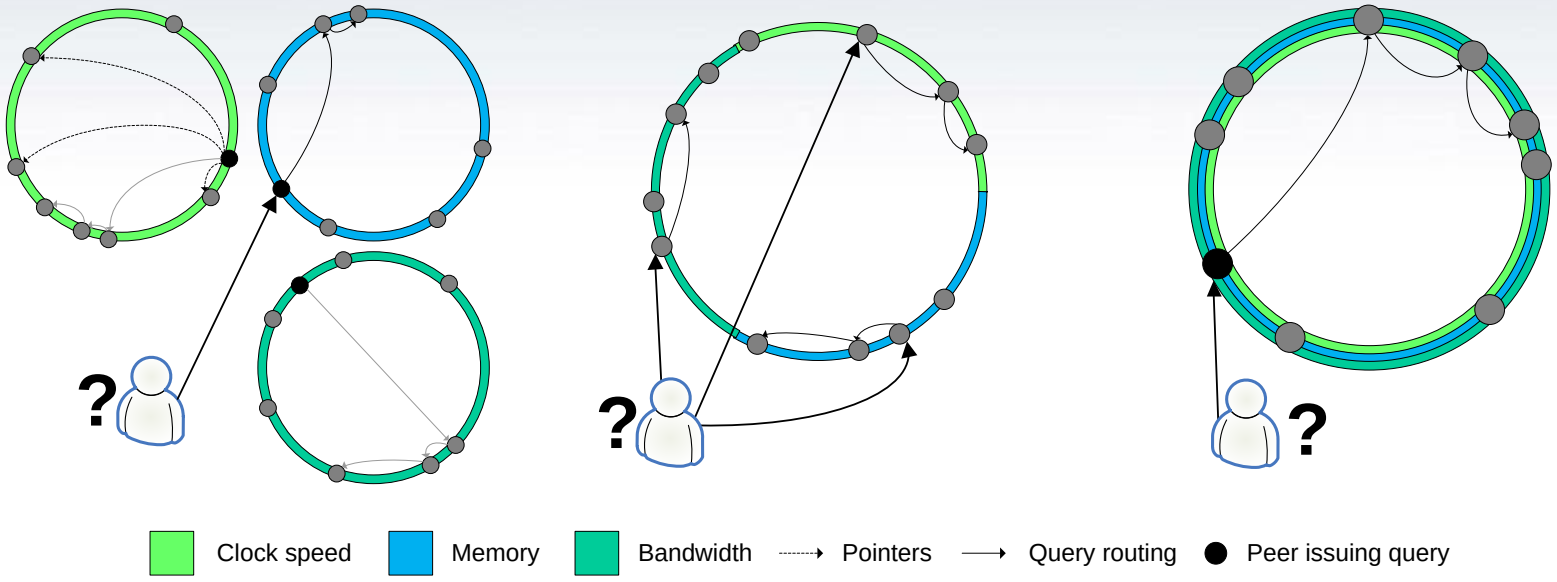


- Finding & aggregating diverse resources
  - Multi-attribute & dynamic resources
    - *Find 2+ Doppler radars covering Fort Collins, 3+ processing nodes with 2+ GHz & 4-6GB memory, 5 TB storage, with 40- ms delay between them & 10 Mbps bandwidth*
  - Not only individual resources, but a group of them working together
  - Resources should relate to each other
  - These phases should work together

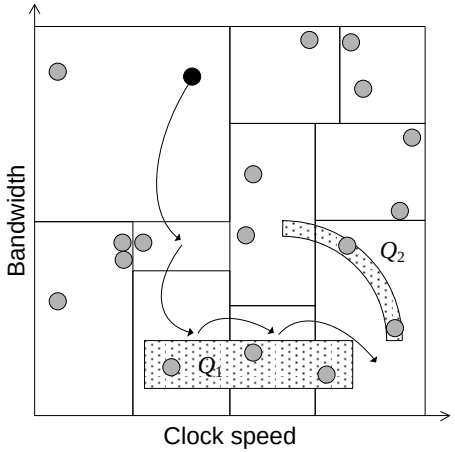
Research Focus

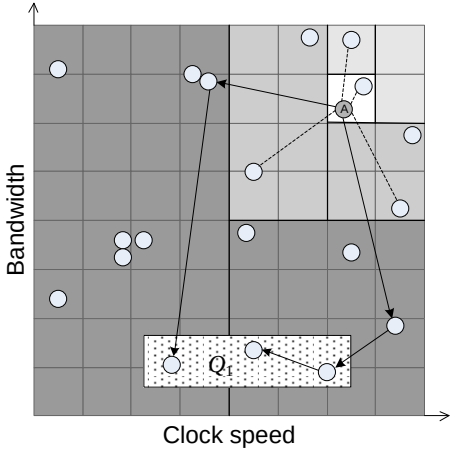
- Building data/resource search solutions for resource intensive collaborative P2P systems
  - CASA requirements
    - Resource intensive, real-time, distributed, multi-sensor data fusion, resource utilization, heterogeneity
  - Specific contributions
    - Supporting push & pull
    - Resource utilization through aggregation
    - Capturing inter-resource relationships
    - Compensation for lack of resources
    - Resource discovery, selection, matching, & binding in a single solution

Current Solution Space



- These solutions are for grid computing
  - Not real-time
  - Low bandwidth





Research Focus (cont.)

- Why current solutions don't work
  - Mainly focus on individual resources
  - Resource discovery, selection, matching, & binding as different phases
  - Unable to compensate for lack of resources
  - Not real time
  - In effective in handling dynamic resources & inter-resource relationships
- Our approach
  - Focus on key attributes
  - Focus on groups of resources
    - Track inter-resource relationships – landmarks
    - Compensation for lack of resources – resource functions
  - Resource discovery, selection, matching, & binding in a single solution
  - Utilize natural growth of Internet backbone
    - Mostly utilize few near by resources

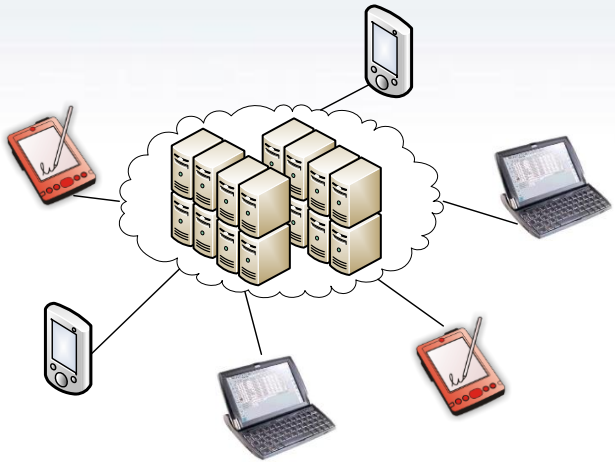
Trends

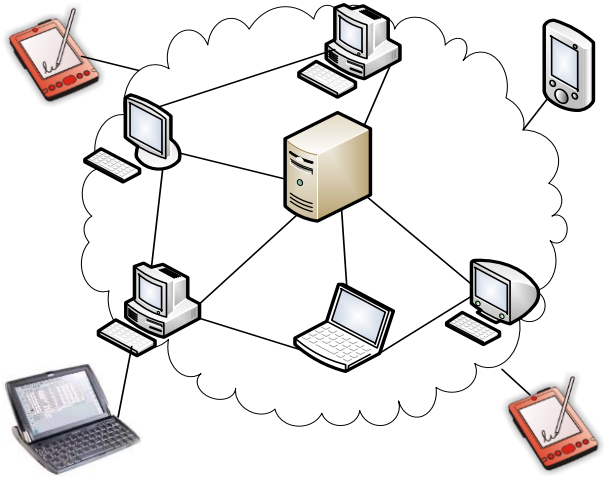
- Resource rich computing devices
  - Households having their own servers 24×7
  - Many resources will be under utilized
- Decreasing communication cost
  - Increasing edge bandwidth
  - Everything will be connected to everything else
  - Real-time on-demand services are norm
- Changing the way we communicate

- P2P paradigm is a natural fit
  - User driven, distributed, utilize resource-rich edge devices, & encourage sharing

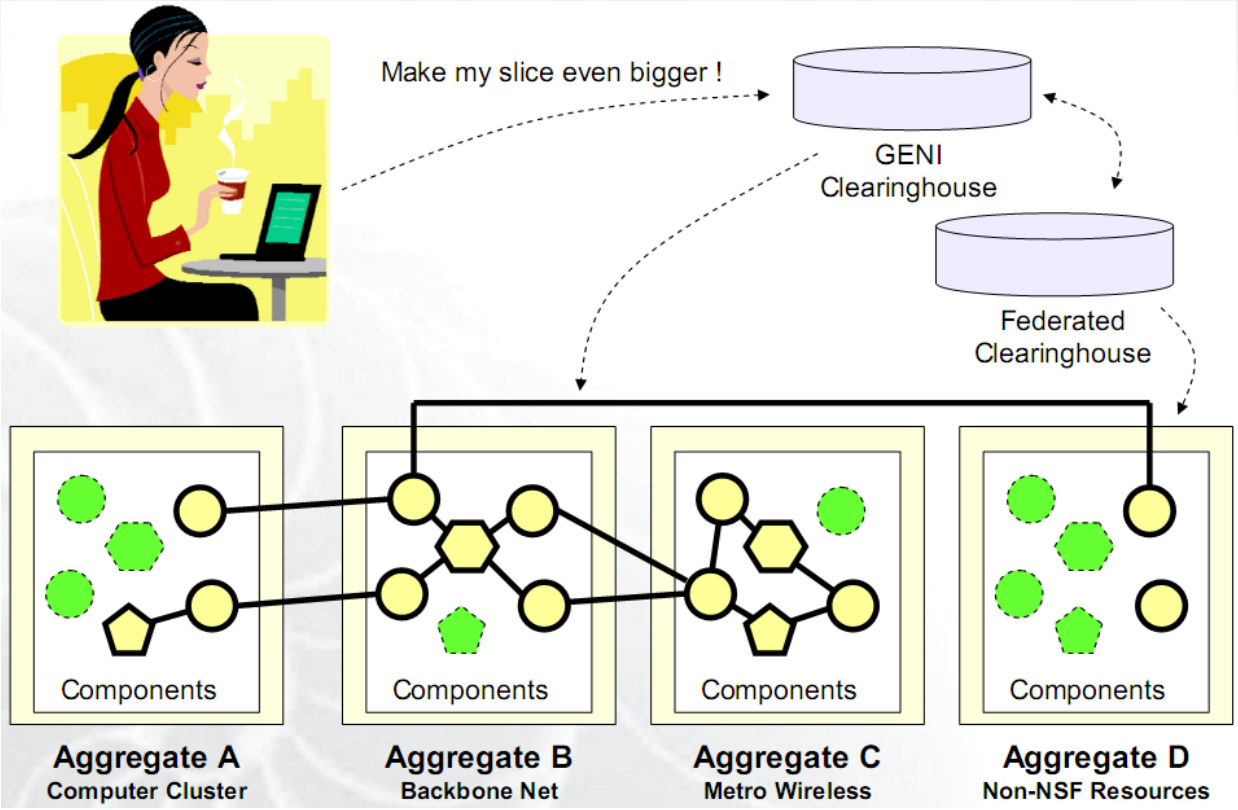
Community Cloud Computing

- Clouds
  - Rapid scale in/out, low start-up cost, pay as you go
  - Centralized data & proprietary applications
- FOSS community says "*it is a trap*"
  - FOSS apps could run in the cloud
  - Where is the datacenter?
- Community is the datacenter
  - Resourceful peers, home servers
  - Users govern themselves & hold data
  - Aggregation of bandwidth at edge
  - Ability to scale in/out
  - Peers could earn points or money





Global Environment for Network Innovations (GENI)



- Understand global networks & their evolving interactions with society
- Innovate at frontiers of network science & engineering
- Transform science of network research & larger world of communications

Questions ?

dilumb@engr.colostate.edu
www.cnrl.colostate.edu