Caching and the Web

- Traditional web: short request-response transactions
- Ideal for some form of network caching
- Do gigabits help or *hurt*?
- How to scale caches in a large and fast network?
Approaches to Caching

- Caching by location
  - At *backbone* nodes
  - At stub domain borders

- Effective use of widely distributed caches is difficult
Caching in an Active Network

- Cache radius — modulo caching
- Trade cache memory for location information
- Co-ordination between active caches — look-around
AN-Sim Environment, Results

- Total cache size constant across methods
- Total cache size increases along $x$-axis
- Each transit-only cache is 25 times larger than each active cache

![Graph showing Round Trip Latency vs Cache Size of Active Caches per Interface]

- Server Dist 0.8, Repeat Prob 0.5, Modulo Radius 3
- Stub connected to Transit
- Transit Only
- Modulo Caching
- Modulo w/ 2 hop Lookaround

Round Trip Latency (No Cache: 13.25)
More Results

- Effect of server distribution and topology

Variation in Server Distribution

Variation in Topology
How do Gigabit Networks Change the Web?

Q.1 What is the Web?

A.1 Reduce congestion?

HTTP or Other Protocols?

A.0 The WEB will expand to fill all available bandwidth.

Not much. Round trip times remain unchanged.
How does the Web Change Gigabit Networks?

A.0 May depend on the type of gigabit network

- Connection-less network (e.g. IP)
- Connection-oriented network (e.g. ATM)

A.1 Makes the network do more connection-setup/teardown processing than other (bulk) applications

- Reduces them to *Megabit* networks