Caching Architecture

Tree-structured caching hierarchy, with cache misses in lower levels percolating up through higher levels until the requested document is found

*Primary goal for caching and prefetching:*
Reduce client delay by reducing or removing external latency
Algorithms

- Caching:
  - Passive - without prefetching
  - Active - employs some mechanism to prefetch data

- Prefetching:
  - Local - based on locally available information
  - Server-hinted - based on hints from the web server hosting the document to be prefetched

- Metrics for evaluating caching algorithms:
  - Hit/Miss Rate
  - Response/Access Time
  - Bandwidth Utilization
Caching between NOC and IS

• Advantages:
  • Client access patterns across servers available
  • Size of cache controllable

• Problems:
  • Mechanism required to throttle prefetch traffic when needed

Caching between NOC and HH

• Advantages:
  • Eliminate latency introduced by the satellite
  • Utilize broadcast/multicast characteristics of satellite connections

• Problems:
  • Size of cache uncontrollable
  • Little data available for prefetching decisions
Factors Influencing Cache Performance

- Workload
- Cache size and access characteristics
- Content Aging and Replacement algorithms
- TCP/IP stack performance
- Prefetching lead time