Issues in Next Generation Streaming Server Design

Youjip Won

yjwon@ece.hanyang.ac.kr

Distributed Multimedia Computing Lab.

Hanyang University

What is Streaming?

On-line Playback of multimedia data

- Remote Playback vs. Local Playback
- Bi-directional(ITV) vs. Uni-directional(VoD)
- Unicast vs. Multicast
- Streaming Requirement
  - Excessive Space Requirement
    - ATSC(19.2MBits/sec) movie of 110 minutes: 15 GBytes
  - Excessive Bandwidth Requirement
    - ATSC: about 19.2 Mbits/sec, MPEG1/MPEG4: approx. 300Kbits/sec, MPEG2: about 10 Mbits/sec

Markets in Multimedia Streaming

- Worldwide growth of digital STBs(Feb. 8, 2001- MRG, Inc.)
  - > $11.5 B in annual sales in 2004, over 140 M units by 2004.
  - Growth of related digital services > $11 B (annually) by 2004.
  - in aggregated new revenues > $54 B by 2004.
  - service revenues
    - Electronic Program Guides (EPG)
    - Personal Video Recorders (PVRs)
    - Video-on-Demand (VOD)
    - Interactive TV (ITV)
    - Pay per View (PPV)
- Companies
  - NDS, Sony, TiVo, Motorola, Microsoft, Pace, Sarnoff, DirecTV, EchoStar, Hughes, Philips, Broadcom, Intel, National Semiconductor, Liberate, OpenTV, nCube, AOLTV, WebTV, Scientific Atlanta, Thomson, CacheVision, NBC, Wink, RespondTV and many others.

Components of Streaming: from Technology Aspect

System

File System, RTOS
Storage Subsystem
Authoring

L2: IMT-2000, Ethernet
L3: IP
L4: RTP/RTCP/RTSP

Network

MPEG-1/2/4
Sorenson Video
H.323
Etc.

Compression

Components of Streaming: from Developer's Aspect

Server

File System, RTOS, Storage Subsystem, Encoder, Network Transport

Player, Network Transport, Decoder, Presentation(BIFS, SMIL, DMIF)

Client

Encoder, Decoder

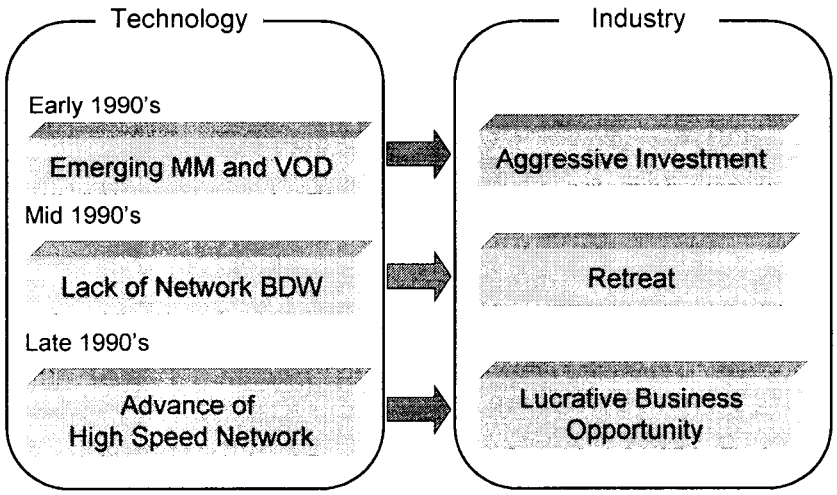
Presentation Tool(BIFS, SMIL)

Authoring Tool

History of Multimedia Technology

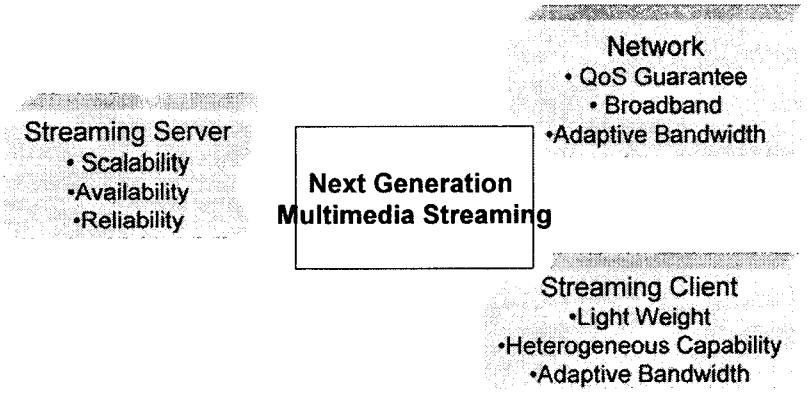
- Early 90's: Local Playback of Multimedia Contents(OS issues)
  - File System
  - Resource Allocation
  - Call Admission Control in single address space
- Mid 90's: Remote Playback of Multimedia Contents(NW Issues)
  - RTP/RTCP/RTSP
  - Synchronization
  - Bandwidth Guarantee on Network
  - Smoothing
- Late 90's: Contents Management and Mobile Issues
  - Image Processing(Transcoding)
  - Pattern Matching(DB Search)
  - Authoring Tool
  - Multimedia in Mobile Environment
  - Heterogeneous Support(Scalable Encoding, Layered transmission)

Market Trends

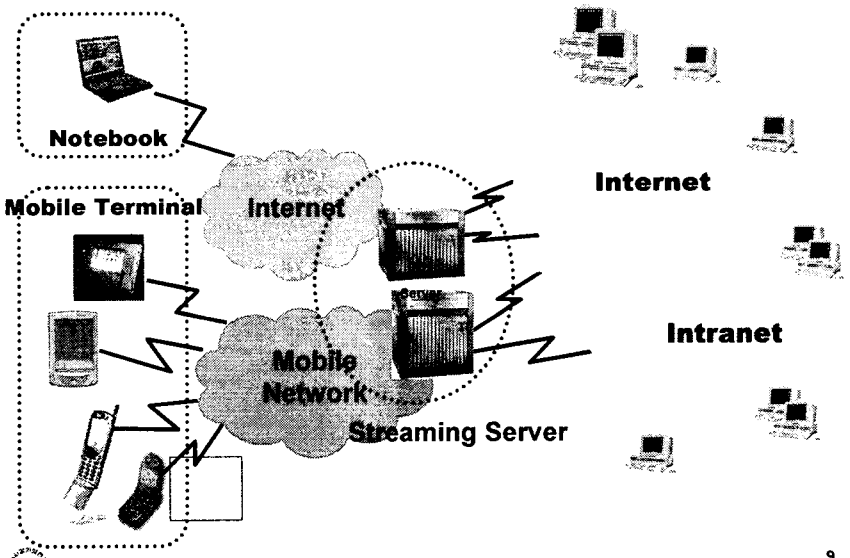


Next Generation Streaming Support

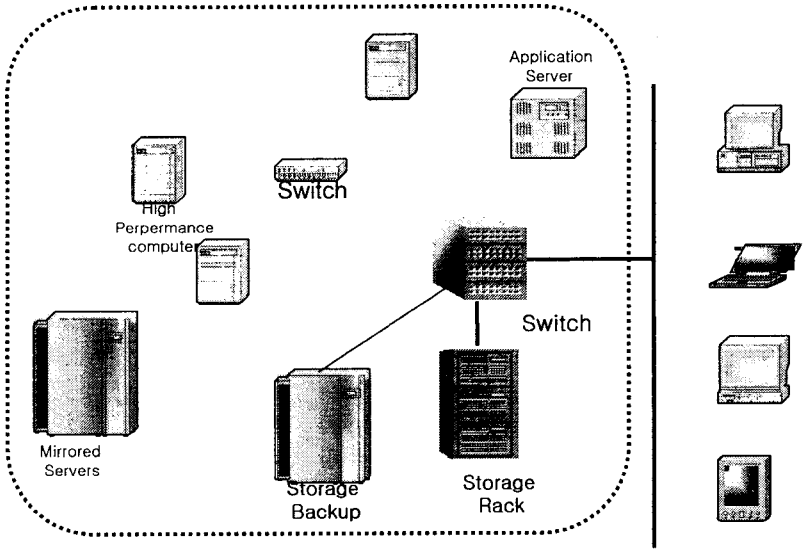
- Support for Heterogeneous Terminal
  Few tens of Mhz – Ghz terminal
- Support for Heterogeneous Network
  Few tens of Kbits/sec – Several tens of Mbits/sec
- Support for RASUM



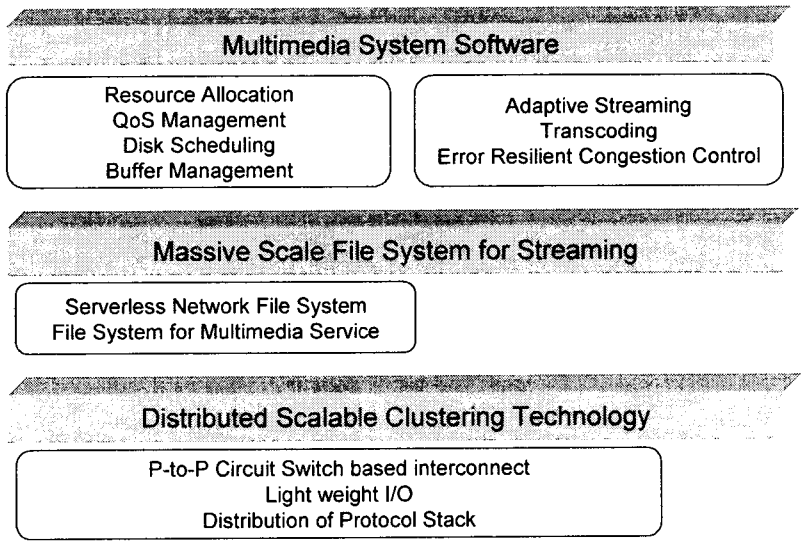
Next Generation Streaming Environment



Distributed Scalable Streaming Server



Component Technology: Server



Multimedia Streaming: Server/Client

- Streaming Server
  - Difficult to provide Bandwidth Guarantee
  - Bursty traffic
  - CPU scheduling: Legacy TS approach is not feasible.
  - File System: Legacy UFS does not fit.
  - How to configure the system to support +1 M users.
- Client
  - Heterogeneous Terminal: PDA, Notebook, Desktop
  - Heterogeneous Network: T1, xDSL, Cable Modem, POT, IMT-2000
  - Wired/Wireless, Static/Dynamic Connection
  - Efficient codec to run on light weight processor(< 33 MHz)

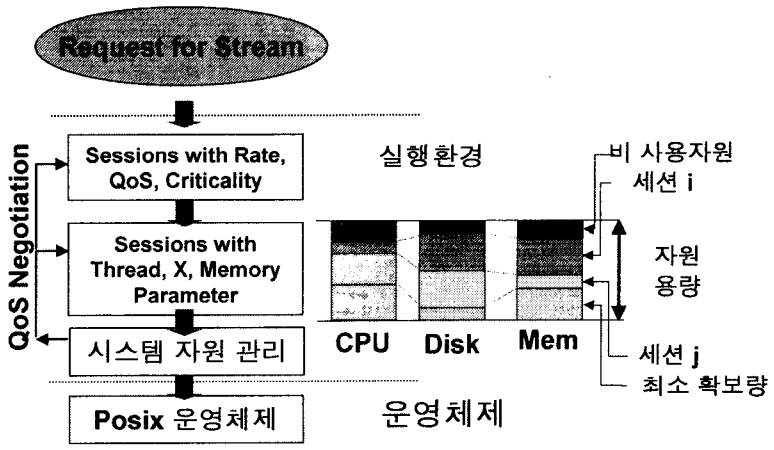
Issues

- **Adaptive stream QoS management** technology for heterogeneous network/client environment
- **Support multiple speed playback**
- **TCP friendly congestion control** mechanism for multimedia streaming
- **Operating Systems Kernel** optimized for streaming

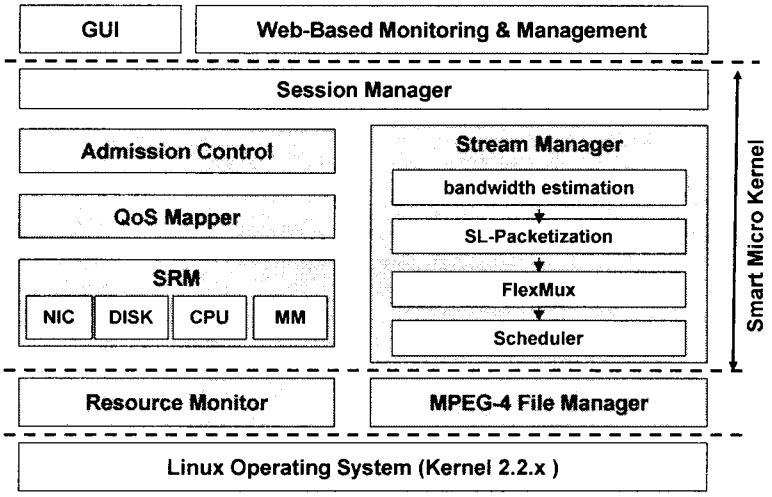
- **Clustered File system technology** optimized for multimedia streaming operation
- **Light weight I/O** technology which can handle hundreds of terabytes data
- **Load Balancing**

Execution Environment

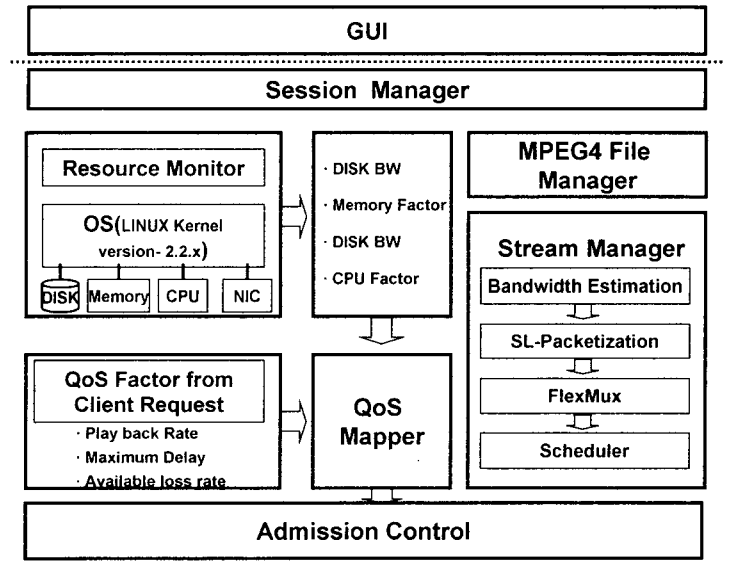
- 분산 실행환경



Server Architecture



SMART Architecture

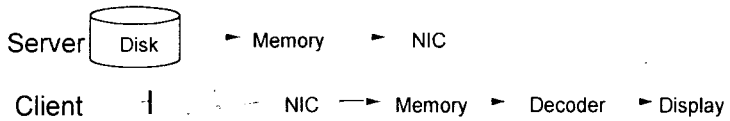


Protocols used in SMART

- **RTP/RTCP(RFC 1889, RFC 1890)**
  - A Transport Protocol for Real-Time Applications.
  - allow monitoring of the data delivery in a manner scalable to large multicast networks
  - provide minimal control and identification functionality.
- **RTSP**(RFC2326)
  - application-level protocol for control over the delivery of data with real-time properties.
  - RTSP provides an extensible framework to enable controlled, on-demand delivery of MPEG 4 data in SMART.
  - RTSP provides a means for choosing delivery mechanisms based upon RTP.
- **SDP**(RFC2327)
  - session announcement, session invitation, and session initiation.
  - SDP conveys sufficient information to discover and participate in a multimedia session in SMART.

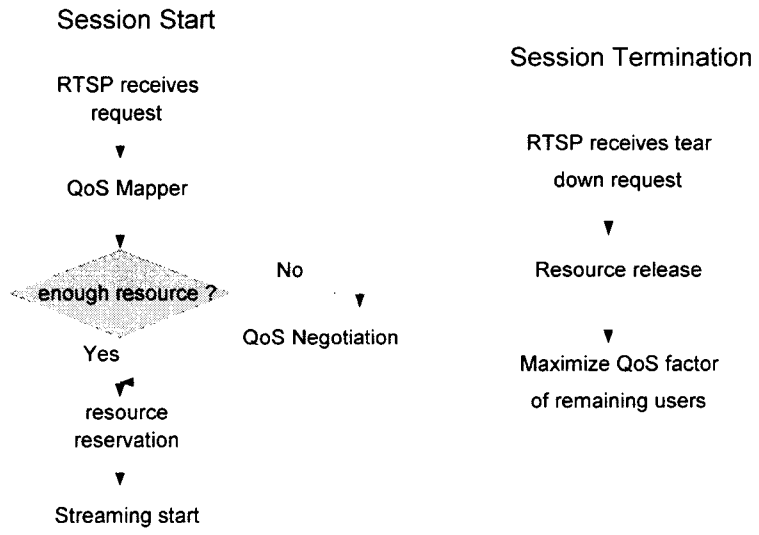
SMART System Components

- Session



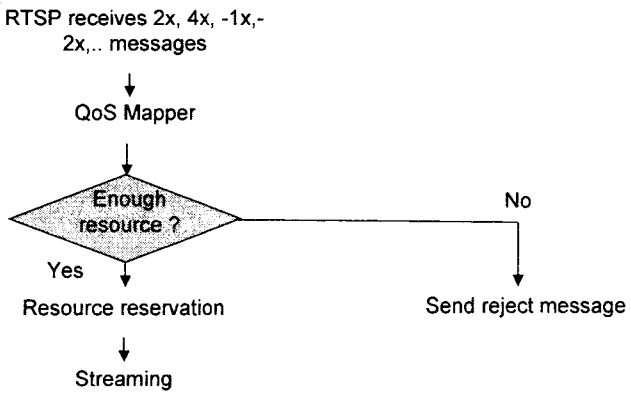
- System Resource Manager
  - resource allocation/scheduling for QoS guarantee
  - Component: CPU, Disk, Memory, NIC
- QoS Mapper
  - Map QoS metric to system resource metric
  - (25fps, 600*480, 1.5Mbps) → Memory(1.5Mbps) , CPU (5%)
- Call Admission Control
  - request 요청시, QoS factor 가 변경시 (ex, Fast forward, backward,..), 수락/거절 결정

SMART Session Management Algorithm



SMART Session Management Algorithm

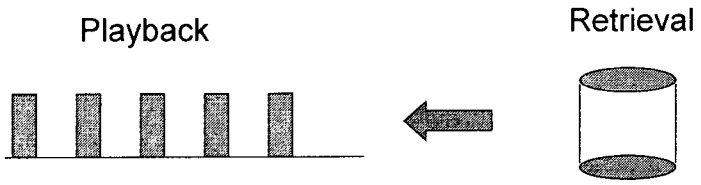
Playback Mode Change



- Thread per session/Thread per Resource
- Issues: Synchronization between threads

File System for Streaming

- Typical Streaming Operation
  - Sequential Read
  - Occasional Fast-Forward, Fast-Backward, Pause
- Characteristics of Streaming Operation
  - Bandwidth Guarantee
  - Minimize Delay variances



File System for Streaming

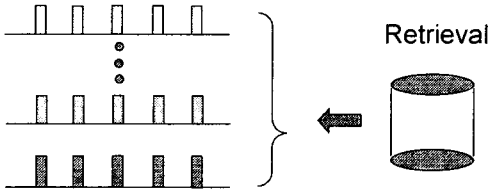
- Minimize Latency
  - Make the file structure flat → Reduce seek overhead
- Minimize Delay Variation
  - File Structure should remain the same with the change in the file size.

Is legacy UFS family OK? Probably Not!

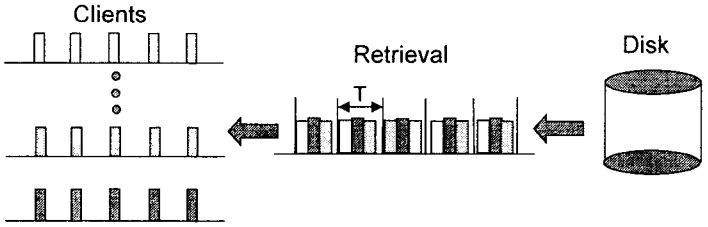
- UFS design philosophy
  - Handling wide variety of file size without loss of disk space
  - Optimized for random I/O
- File System for Multimedia Streaming
  - Minorca(U. of Oslo, Norway), MMFS(SUNY Stony Brook, USA), Presto(U. of Minnesota, USA), HERMES(Hanyang U., Korea)
  - Tigershark(IBM Almaden), Tiger(Microsoft)

Schedulability

- How many streams can we support?

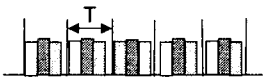


Round Based Disk Scheduling



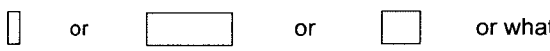
Schedulability

- Problem
  - Given playback rates of sessions
  - And disk parameters
- Determine
  - Retrieval Schedule



Length of round, T

Size of data retrieved in a round



Formulation

- Solve the following two equations

$$T \cdot r_{display} \le n_i \cdot b$$

$$T \ge \sum_{i=1}^{S} \frac{n_i \cdot b}{B_{\max}} + \Theta(S)$$

T: length of a round
n_i: # of blocks/round for stream i } Unknowns!!

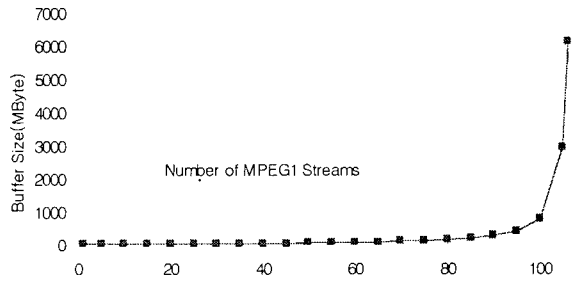
$r_{display}$: playback rate
b: block size
B_{max}: maximum transfer rate

O(s): Disk Overhead
.Governed by Disk Scheduling Algorithm
Data Placement Strategy

Buffer Size Requirement

- Schedulability
  We can admit new stream as long as we can find **n** in below equation.
  n increases very fast.
  O(s): disk scheduling overhead

$$\mathbf{n} \ge \frac{\mathbf{r} \cdot \Theta(s)}{\frac{b}{B_{\max}}\left(B_{\max} - \sum_{i=1}^{s} r_i\right)}$$

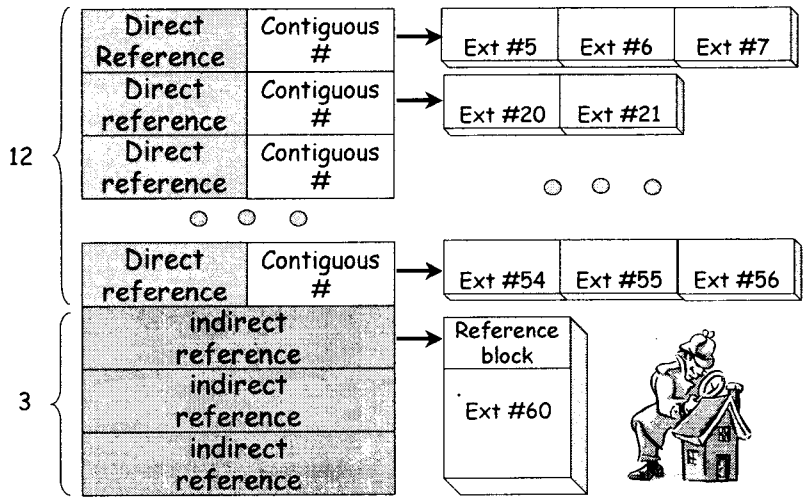


HERMES File System: Synopsis

- Extent based allocation
- Separation of file system meta data and the extents
- Integrated block structure
  Reference pointers + Data Block at the same level
- Objective

Minimize Delay and Delay Variances

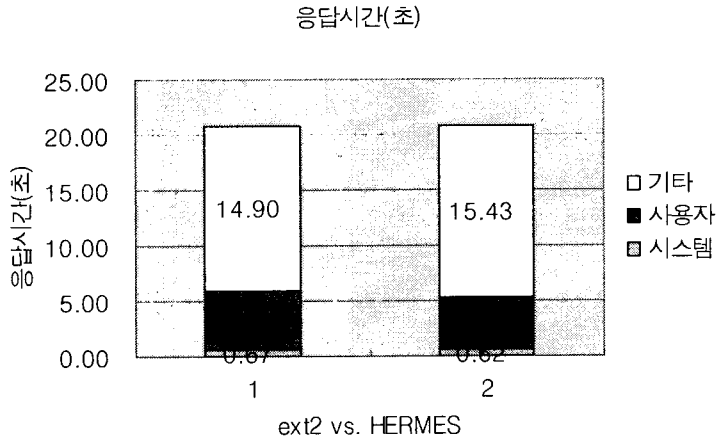
HERMES: Inode Structure



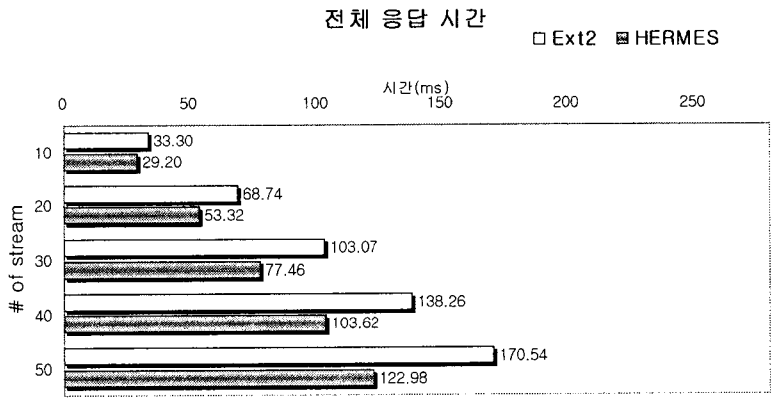
Performance: HERMES vs. EXT2

◆ 파일 읽는 시간

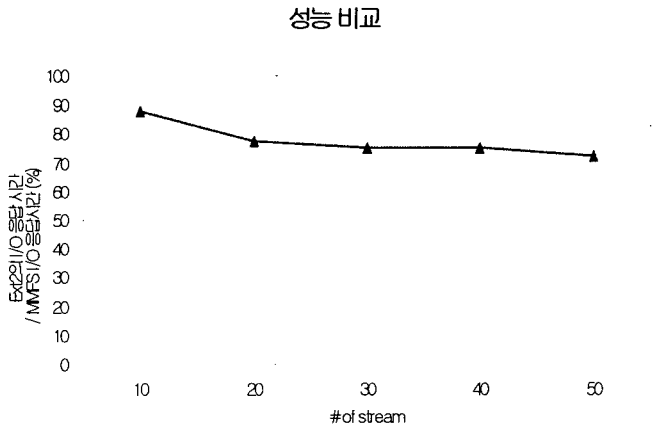
600 Mbyte



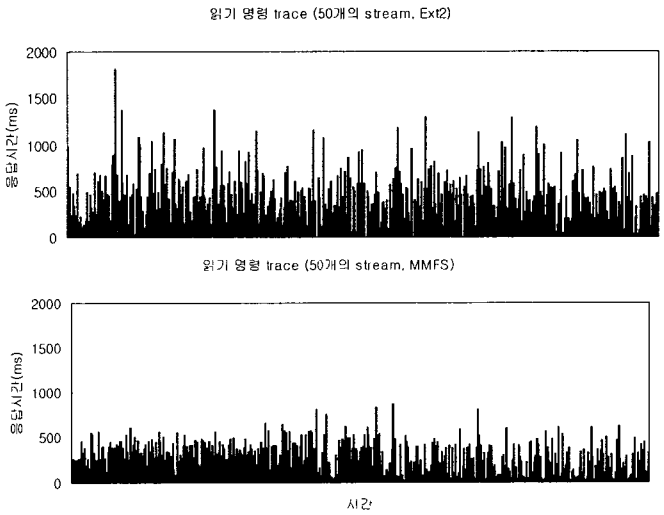
Performance: HERMES vs. EXT2



Performance: HERMES vs. EXT2



Delay Variations



Issues in Massive Scale Cluster Server Design?

How to Support +1 M concurrent session?

Single Gbit NIC supports 100 200Kbits/sec streams.

Server: Architectural O/S issues

- Running entire stack on a general purpose SMP
- No direct disk to NIC transfers
- Bus based architecture
- O/S may cause queues in wrong places
- O/S supported I/O, IPC and synchronization are typically very inefficient

Server: Load Management Issues

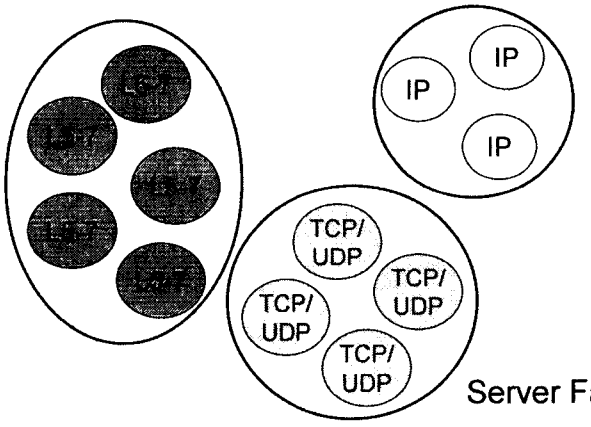
- Symmetric Architecture vs. Layered Architecture
  - Layered Architecture: Easier to manage, configure, engineer, but performance implication is not clear
- Load Distribution for Symmetric Architecture
  - Client Based Approach(Netscape Access)
  - Round Robin DNS(CISCO's LocalDirector, Cisco's Distributed Director)
  - Dispatcher Based Approach(IBM Network Dispatcher)
  - Server Based Approach(Scalable Server WWW)

Server: Content Management Issues

- Large Facilities built as loosely connected clusters of servers
- Significant overheads of contents support
  - NFS mounting → Bottleneck, single point of failure
  - Full Duplication → Consistency management overhead
  - Content Partitioning → Traffic distribution difficult and single point of failure
  - File Cached in every server → Nonscalable
- Content partitioning difficult to handle because
  - Shifting demand phenomenon
  - Heterogeneous servers

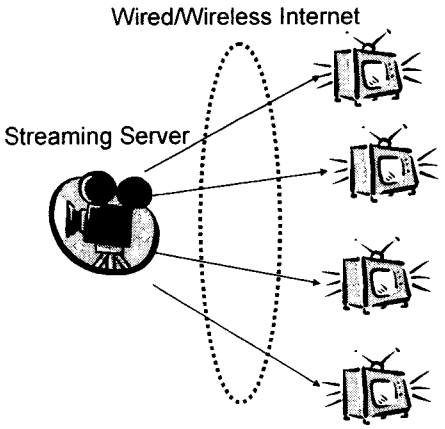
Server: Layered Architecture

- Number of Machines in each layer?
- How to maintain state information across the layer?



Multimedia Streaming: Network

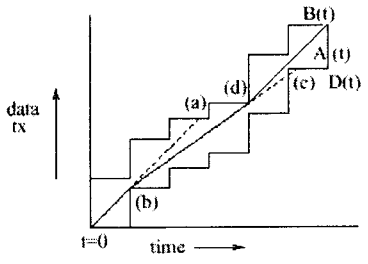
- QoS Guarantee
  - RTP/RTSP/RTCP
  - Diffserv(PHB)
  - Intserv(RSVP)
  - MPLS
- Adaptive Streaming
  - Scalable Encoding
  - TCP friendly congestion control
  - Transcoding
    - Color → Black/White
    - Picture → Text
    - 30 fps → 3 fps
- Mobile Multimedia
  - Smooth handoff
  - Error Resilience



Internet does not guarantee bandwidth!!!

Issues in Network Support for Streaming

- Smoothing
  - Make the bursty video traffic smoother.
  - Server side smoothing(Transmission based on prior knowledge of bandwidth requirement)
  - Client Side Smoothing(Introducing Buffer)
  - Better Congestion Control
  - Improve Loss/jitter situation.
  - More delay is introduced.



(b) - Concave - Rate-decrease.
(d) - Convex - Rate-increase

Fig. 1. Optimal schedule construction

Issues in Network Heterogeniety Support for Streaming

- Network environment gets more diverse.
  - T1, LAN, ADSL, Wireless LAN, 3 G mobile link
- Adaptive Streaming
  - Media Transcoding
  - Source Driven vs. Receive Driven
  - Unicast vs. Multicast
  - Scalable Multimedia Model
- Adjusting the rate
  - Adjust Frame rate → Drop frames.
  - Use hierarchically encoded streams. → Layered Transmission
  - For live broadcast, adjust the encoding rate(e.g. QCIF)

Summary

- Next Generation Multimedia Streaming Technology
  - Massive Scale Support → Clustered Solution
  - Adaptive to Heterogeneous Network
  - Adaptive to Heterogeneous Terminal Capability Presentation Technique
- SMART Server Architecture
- HERMES File System
- Clustered Solution
  - *High Speed Storage Interconnect*
  - Content Partitioning
  - Load Management
- Support for Heterogeniety
  - Adaptive End to End Streaming Transport: Unicast vs. Multicast
  - Scalable Encoding