Institut für Betriebssysteme und Rechnerverbund

Networked Music Performance (NMP)

Xiaoyuan Gu, Matthias Dick, Ulf Noyer and Lars Wolf
Institute of Operating Systems & Computer Networks
Technical University Braunschweig

IBR

Agenda

- Why Networked Music Performance
- Related work
- What is NMP
- Design considerations
- Evaluation
- Conclusions & future work

Motivation

- IT has penetrated into nearly every aspect of the work and life of human beings
- The market of networked entertainment is growing
- The usage of Internet as music databases has been well established and exploited
- Emerging interests in exploring the nature of Internet for new paradigms of networked music
- Emerging applications: networked collaborative composition, networked conducting, and distributed musical performance.

Our focus!

Limitations of Tradition

- Requires physical presence of the musicians
- Not an easy task to find a common timeslot
- Time and costs on traveling
- Find a player of the desired level
- Different versions of the sheet music

➔ *A basic need to improve the way of music performance for sakes of flexibility, economy, efficiency, productivity and creativity.*

Definition

- A concept of rehearsals/concerts via networks with acceptable audio quality
- **Bandwidth-demanding**: Mono PCM 0.7Mb/s, and up to 27.6 Mbps for high definition multi-channel natural audio
- **Highly-delay sensitive**: 120ms E2E delay upper-bound for real-time interactive apps, 20ms desired for music for professionals.
- **Strict requirement on audio stream synchronization:** clocks of PCs, latencies from sound device, NIC, and rhythm adjustment etc.

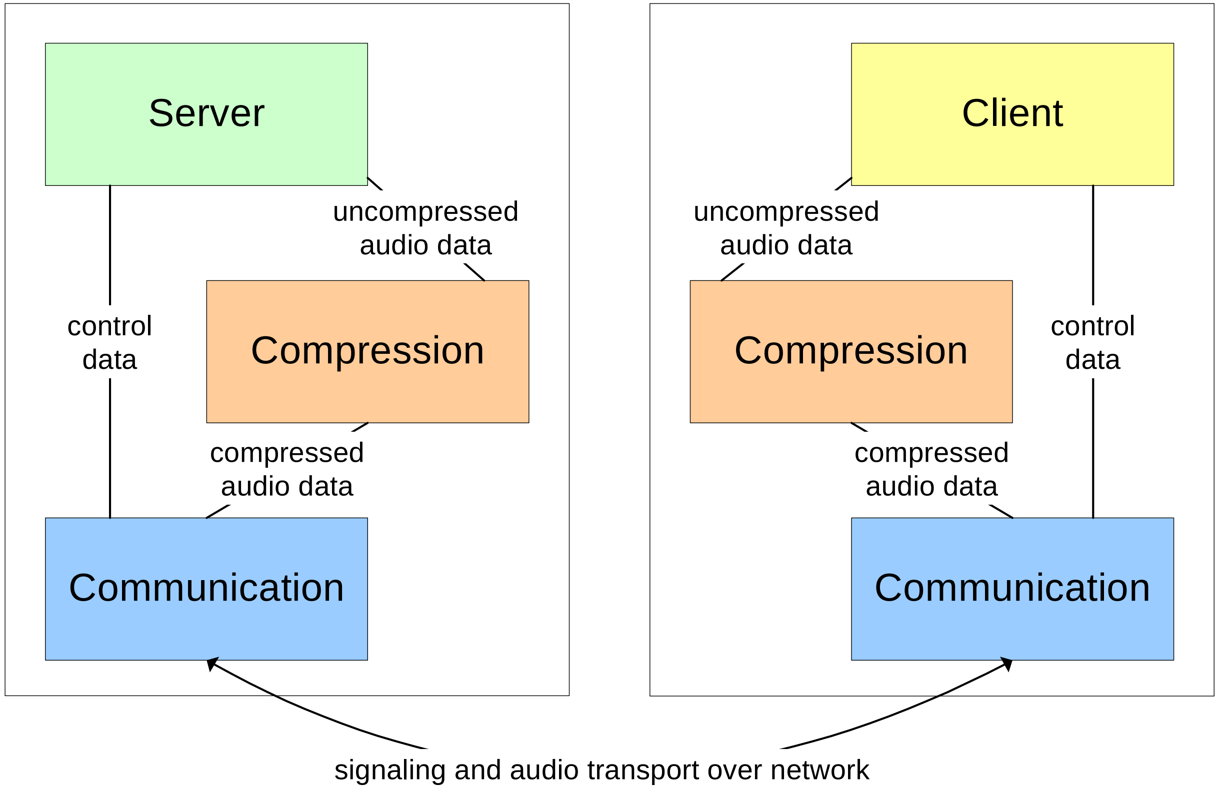
Related Work

- Only a few studies on this topic present in literature
- Earlier work was focused on mono audio transport over ATM Network
- Some experiments on using MIDI to convey synthtic audio
- The Master Class approaches
- Xu and Copperstock‘s work
- The SoundWire Project
- The Conductor-driven Scheme

➔***Conclusion:*** *limitations of current work and the demand for further work*

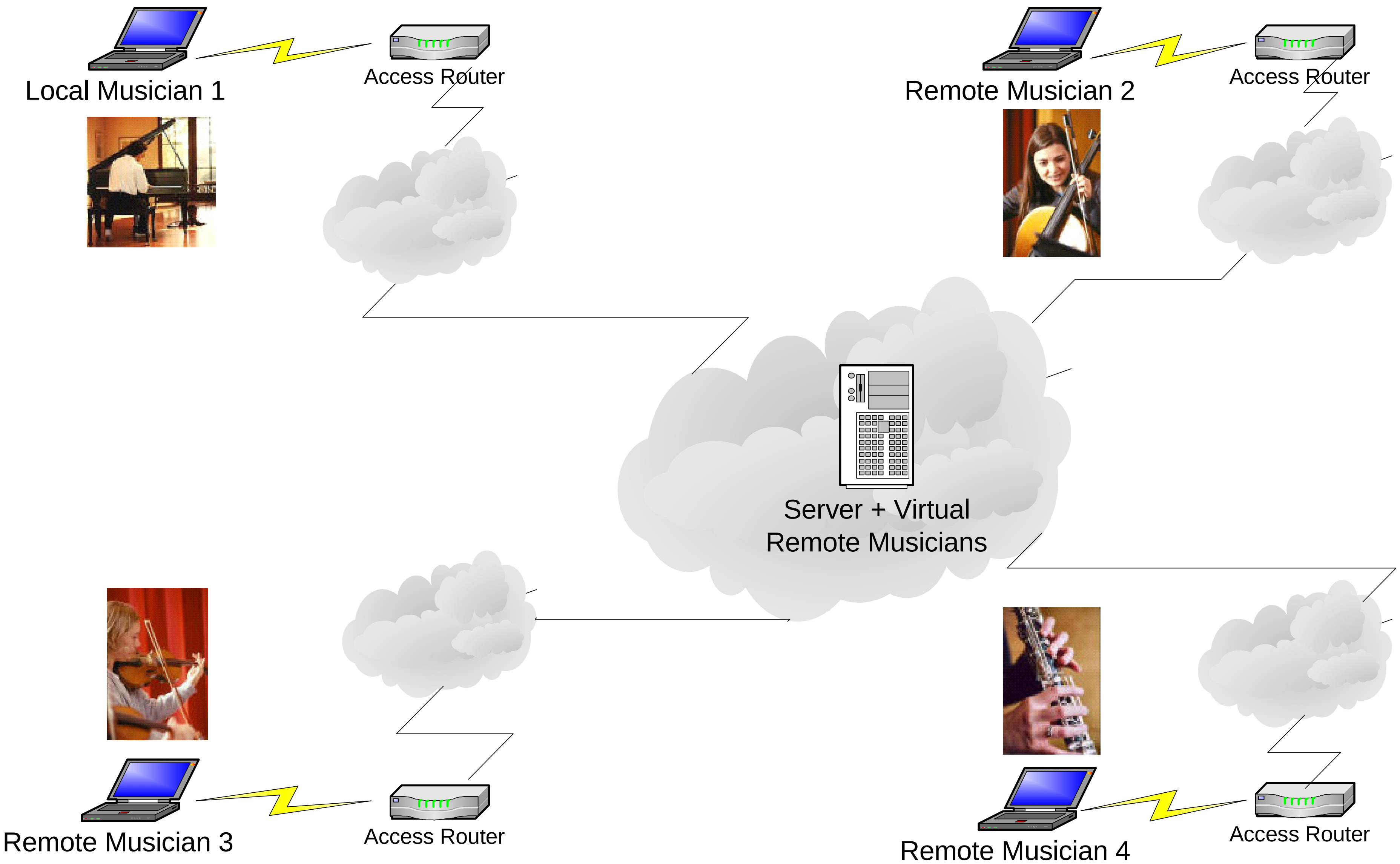
Design: The Architecture

- 4 major components: client, server, compression & communication
- Application boundaries

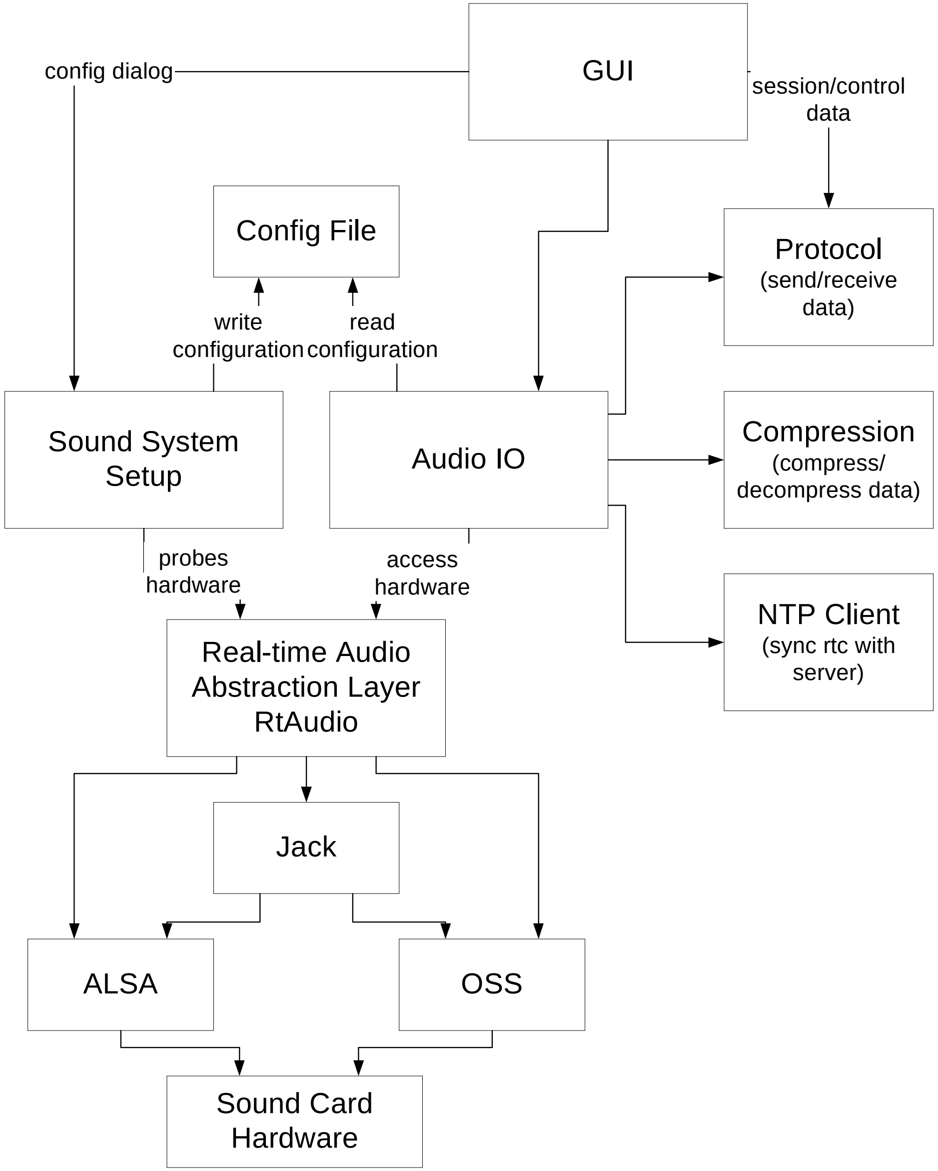


- Two application scenarios: real-time rehearsal & rehearsal on-demand
- Targeted at home Internet users constrained by the last-mile bottleneck link

Application Scenarios

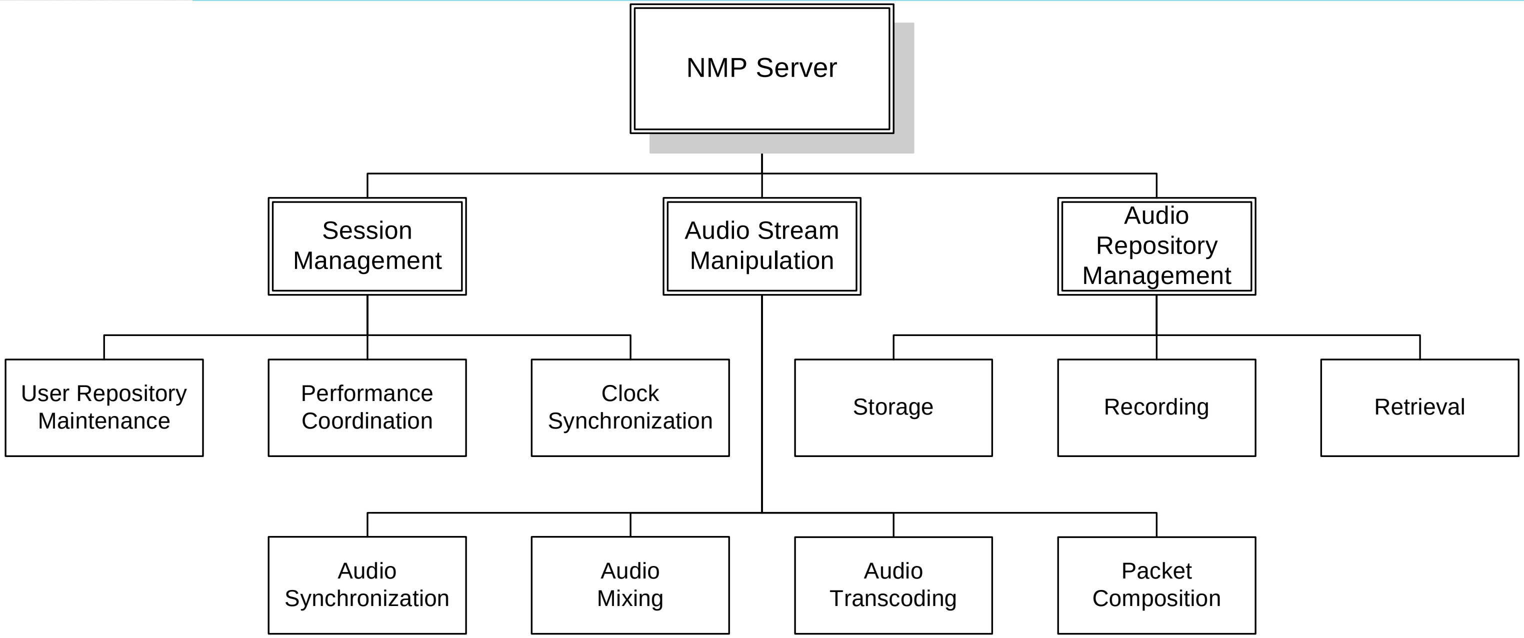


The Client



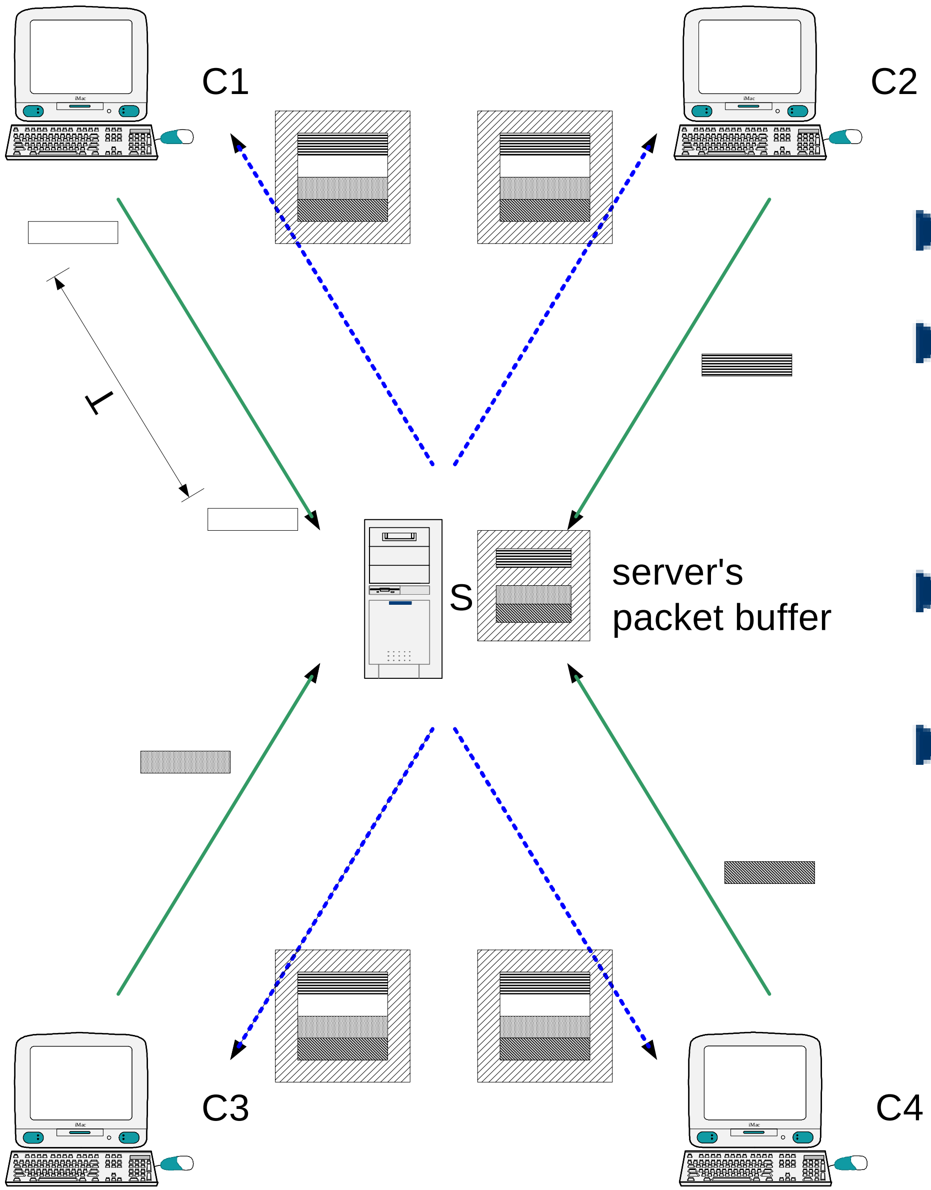
- The user interface to NMP
- Service configuration: *instrument selection and tuning, music piece determination, partner selection, rhythm control, starting-point signaling etc.*
- Sound Card Manipulation: *duplex is a must, allow controlling the related latency due to buffering*
- Clock Synchronization: *all client clocks are in sync with that of the server using NTP*

The Server



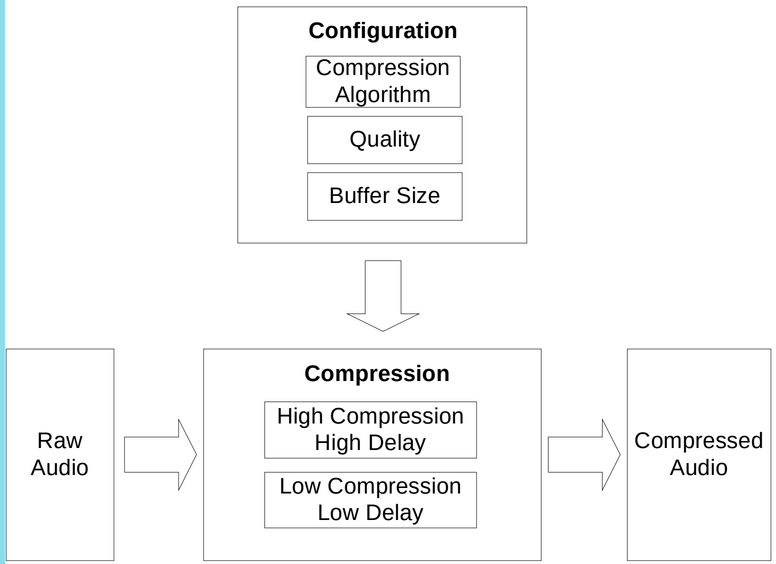
- A point of centralized control
- Session Management: *user repository maintenance, performance coordination, clock synchronization*
- Audio Stream Manipulation: *audio stream synchronization, audio mixing, transcoding, packet composition*
- Audio Repository Management: *storage of the performance examples for either emulation of the remote musicians or playback of the live performance*

The Communication



- A packet-switched paradigm
- Hybrid delivery mechanism:
  Unicast clients' mono audio to server and multicast the multi-channel audio from server back to clients
- Audio Data Transport:
  RTP over UDP, standard compatible
- Session Control:
  Proprietary session management protocol over TCP. Support for session management, performance coordination. Message object serialization.

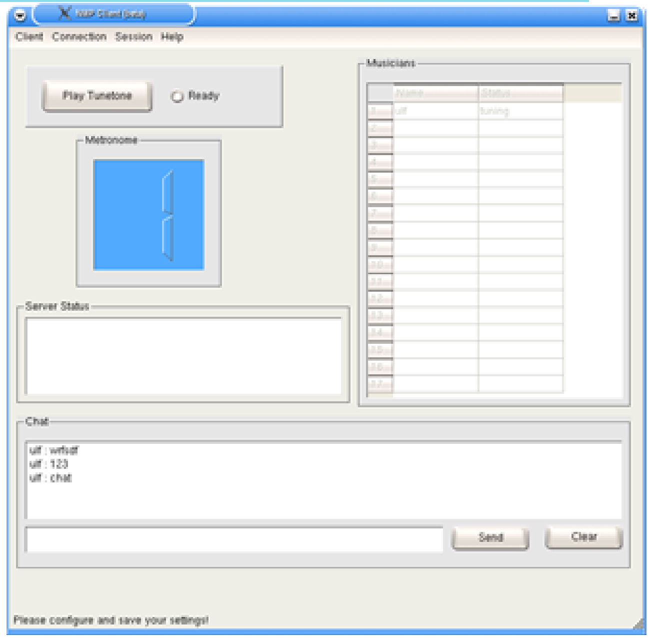
The Compression



- A trade-off between bandwidth efficiency and latency
- MP3 and MPEG-4 AAC not for real-time due to buffering
- ADPCM is the selected codec for real-time rehearsal scenario
- The usefulness of NMP is decided by E2E delay
- Buffer size has a significant impact on delay
- Compression as Library : *Flexibility in choosing the optimal codec, configure the desired algorithm, quality, and buffer size.*

The Implementation

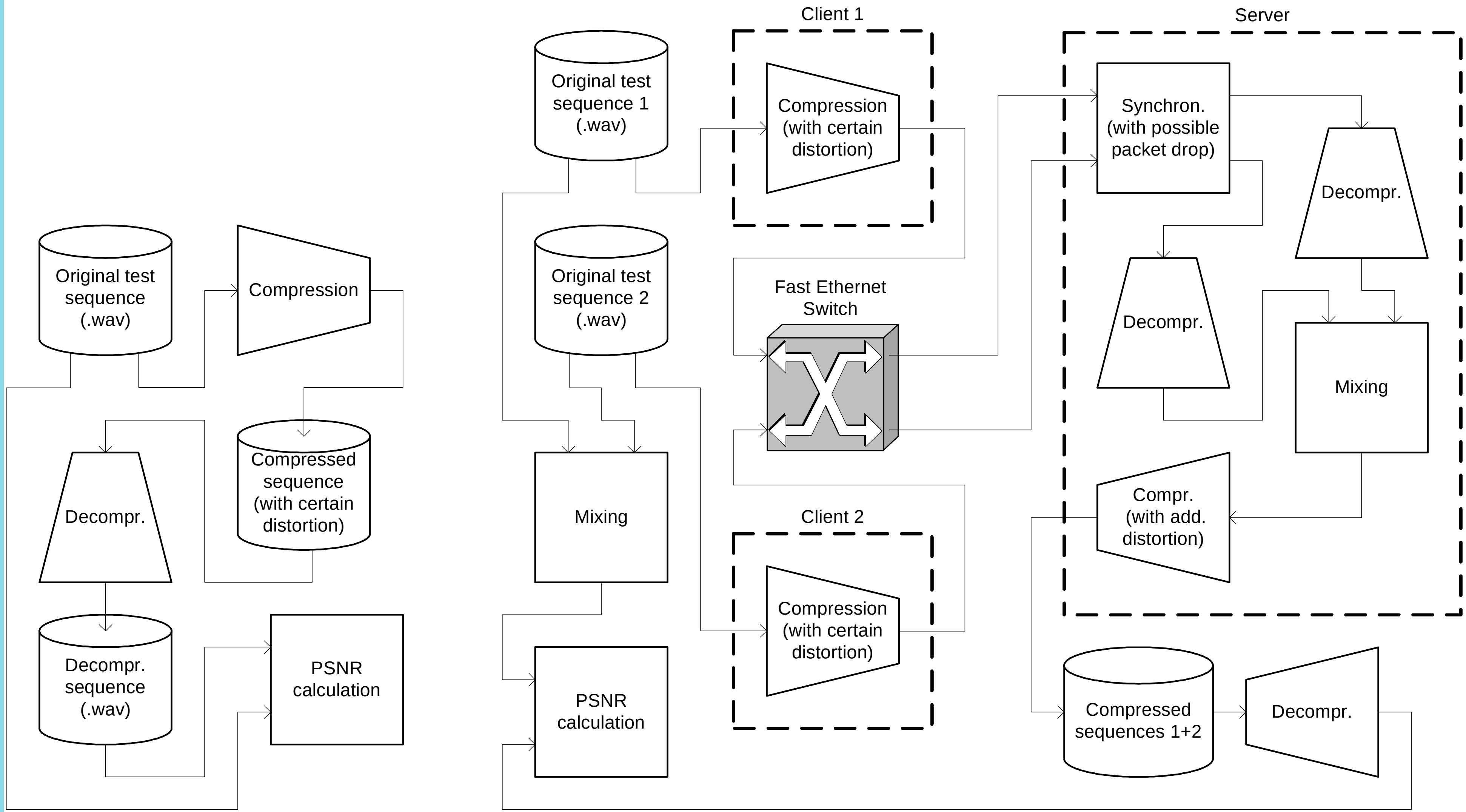
- Test-bed:
  PCs in a controlled LAN
- Operating Systems
  ➔ *Linux: OSS + ALSA*
- Programming Language ➔ C++
- Graphical User Interface ➔ QT
  Envision of cross-platform portability
- Hardware
  PCs, Fast Ethernet switches, multi-channel sound cards and speaker systems, oscilloscope and sweep generator.



Evaluation: Test Procedure

- Targeted codecs: MP3 and ADPCM
- Evaluation categories: object + subjective
- Measurement metrics: PSNR and MOS
- Peak Signal to Noise Ratio: levels 50dB,70dB,90dB
- Mean Opinion Score: 1-5
- Test sequence: publicly available audio samples

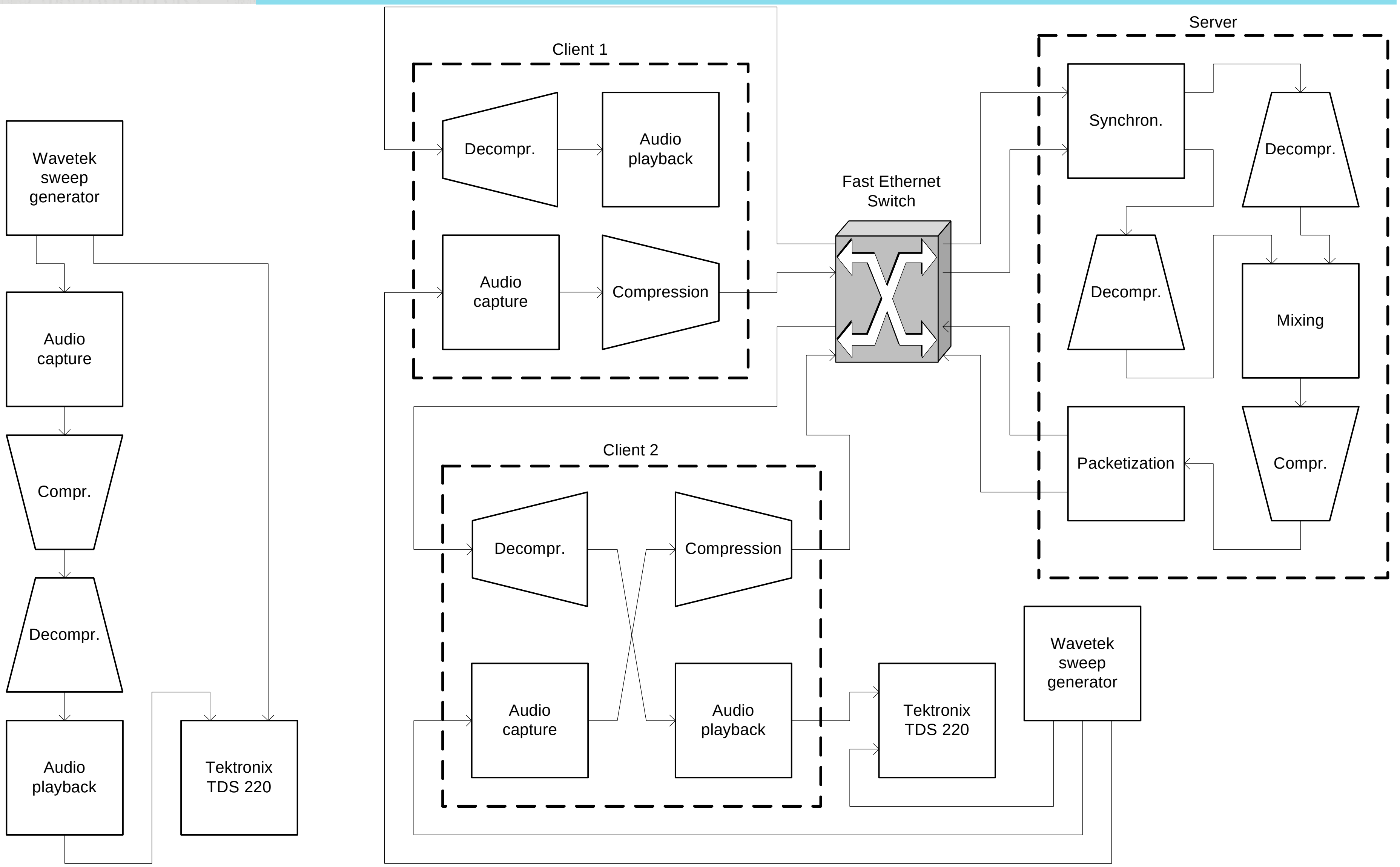
Test Configurations



(a) Test configuration A - distortion due to compression

(b) Test configuration B - end-to-end processing distortion

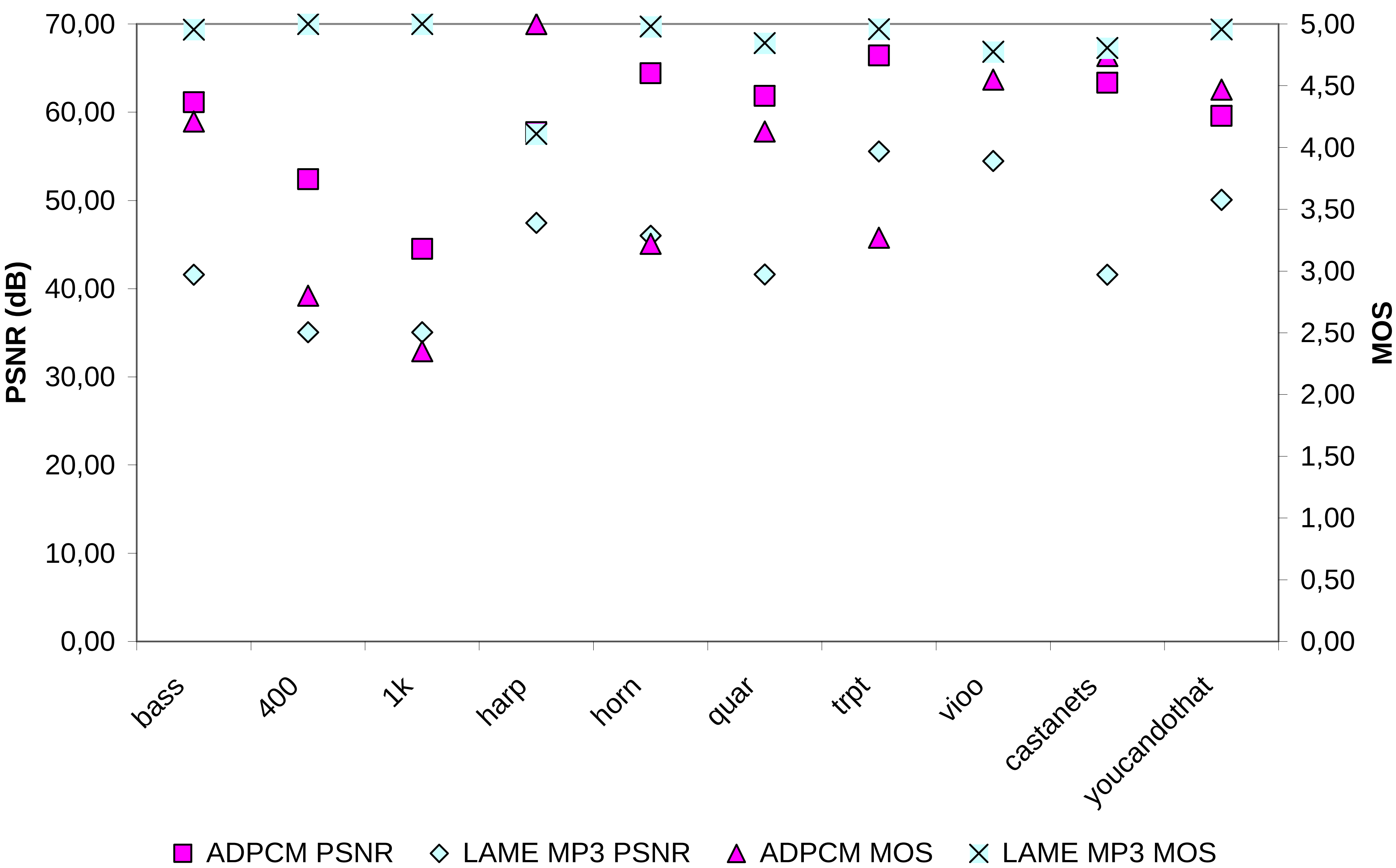
Test Configurations (cont.)



(a) Test configuration C - single client latency

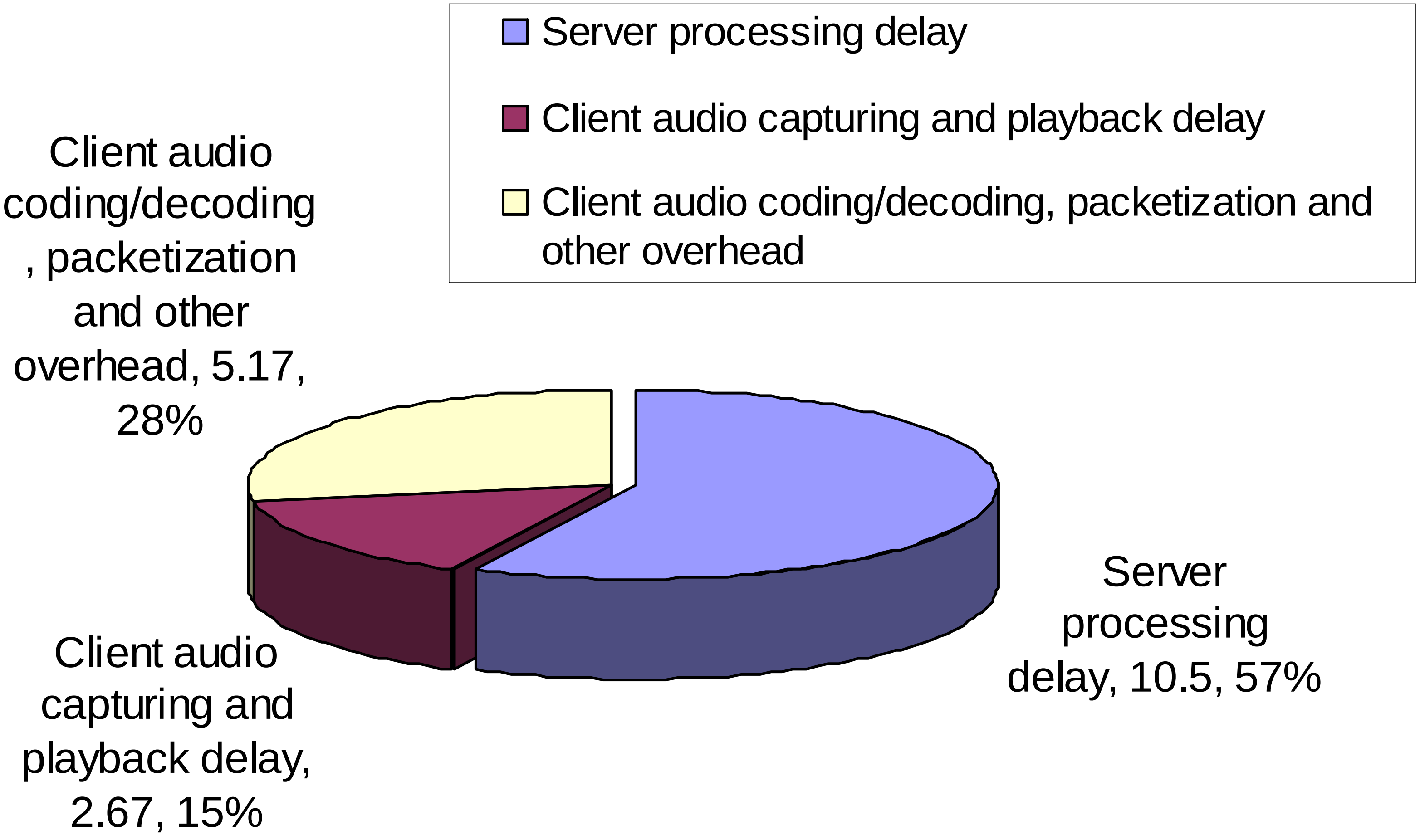
(b) Test configuration D - end-to-end latency

Audio Quality



Latency Distribution

Breakdown of E2E Delay (ms)



Conclusions & Outlook

Conclusions

- The proposed application suffices the real-time constrains and the required audio quality in the LAN.
- Different audio compression schemes and multi-channel audio were supported.
- There exists loose-couplings between MOS and PSNR

Future Work

- To extend the application toward larger scale networks
- To add the support for MPEG-4 AAC
- To consider realistic network conditions
- End-system adaptation schemes and QoS support
- To adopt better object measurement metric like PEAQ