Synchronizing multi-media streams

Kishan Shenoi

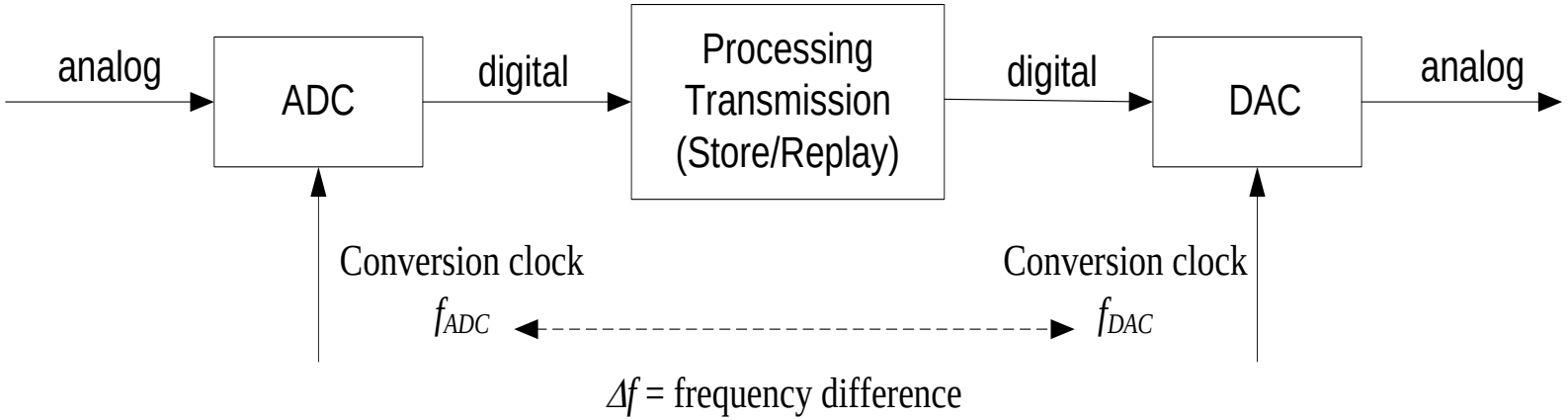
CTO, Qulsar

WSTS – 2013, San Jose, April 16-18, 2013

Outline of Presentation

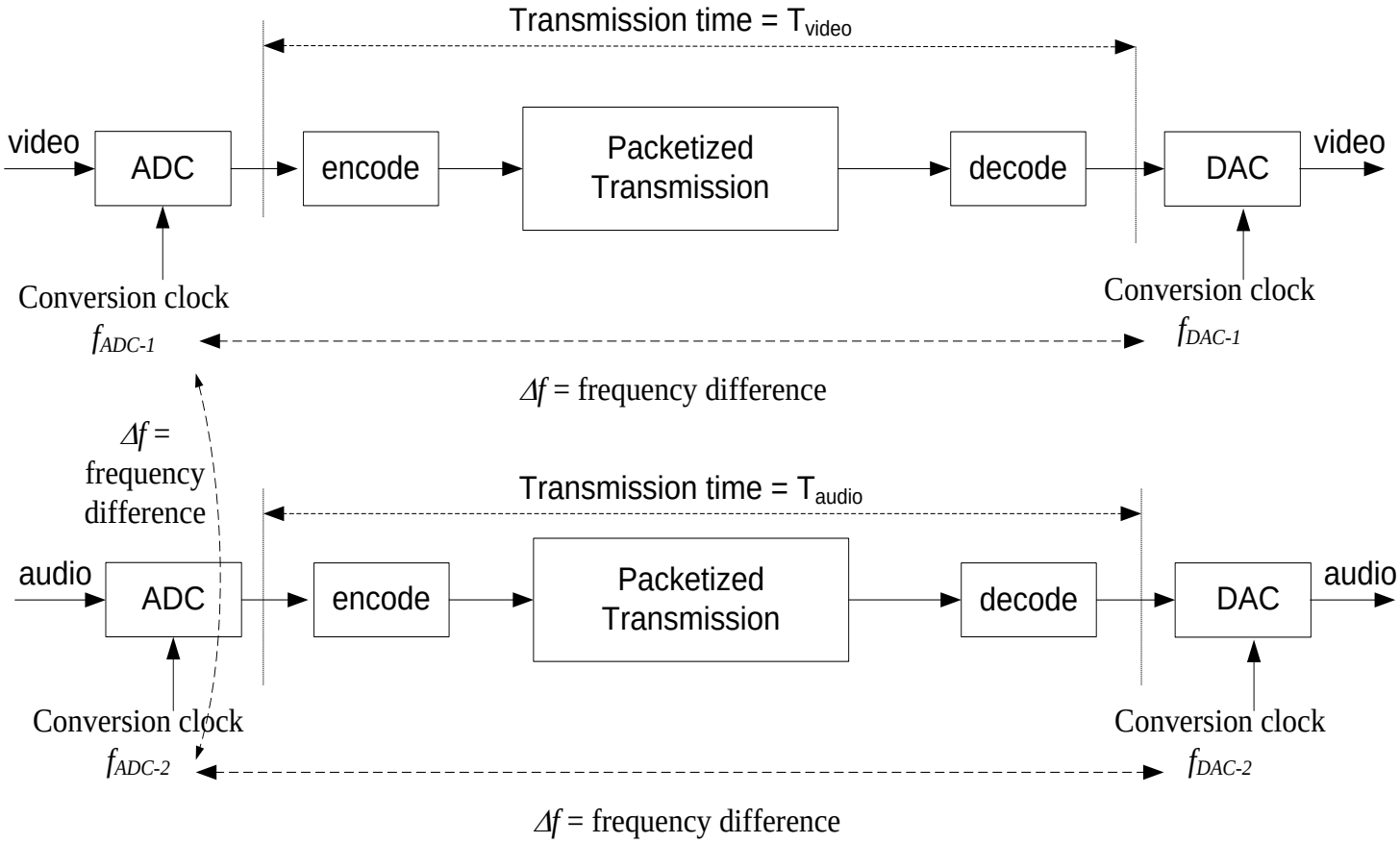
- Fundamental need for synchronization
- Alignment of multiple streams
- Conventional approach
- Time alignment in multi-media
- Using time-stamps for alignment
- Concluding remarks

Fundamental need for Synchronization



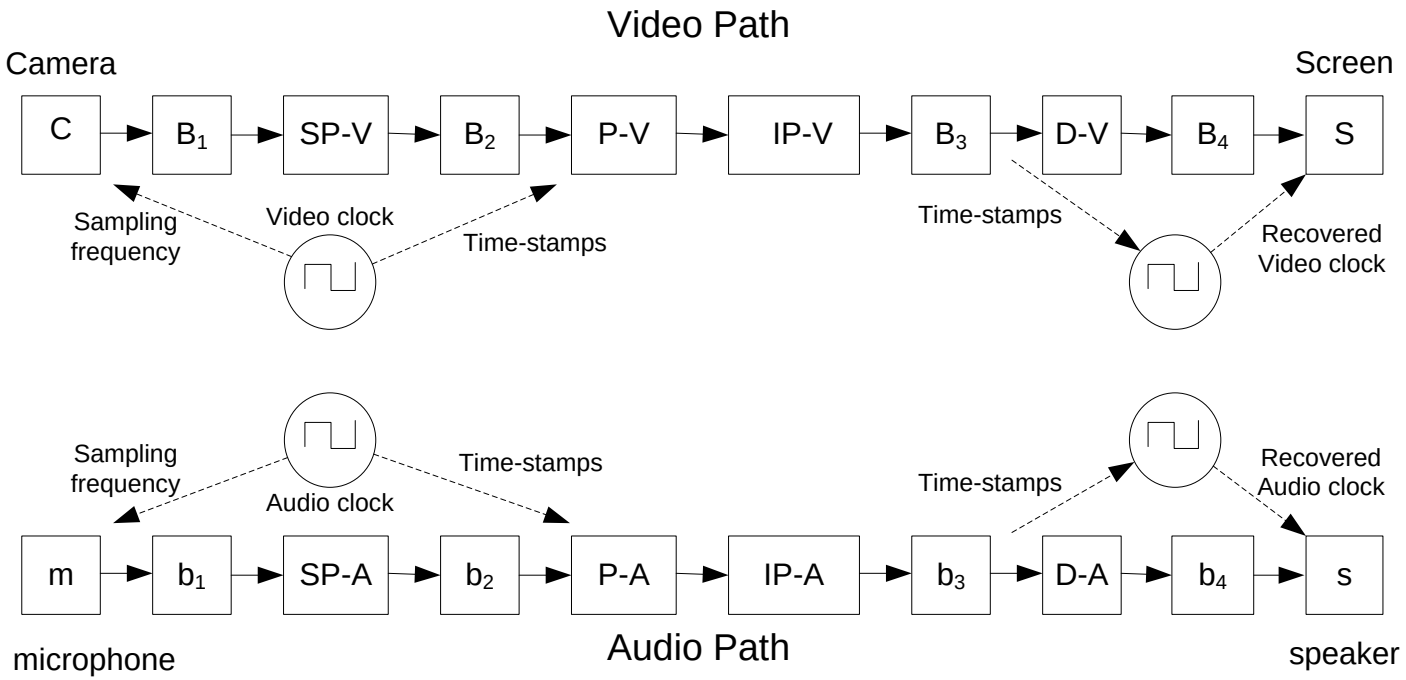
- Real time transmission of audio/video over digital networks requires conversion from analog-to-digital (at source) and digital-to-analog (at destination)
- Impact of frequency difference (Δf):
  - Eventually buffers will overflow/underflow (e.g. slips) (“obvious”)
  - Pitch Modification Effect (PME) (analogous to *Doppler*) makes recovered symbol clock ≠ transmit symbol clock (not so “obvious”)
  - Recovered waveform ≠ original waveform (more than just additive noise)

Alignment of multiple streams



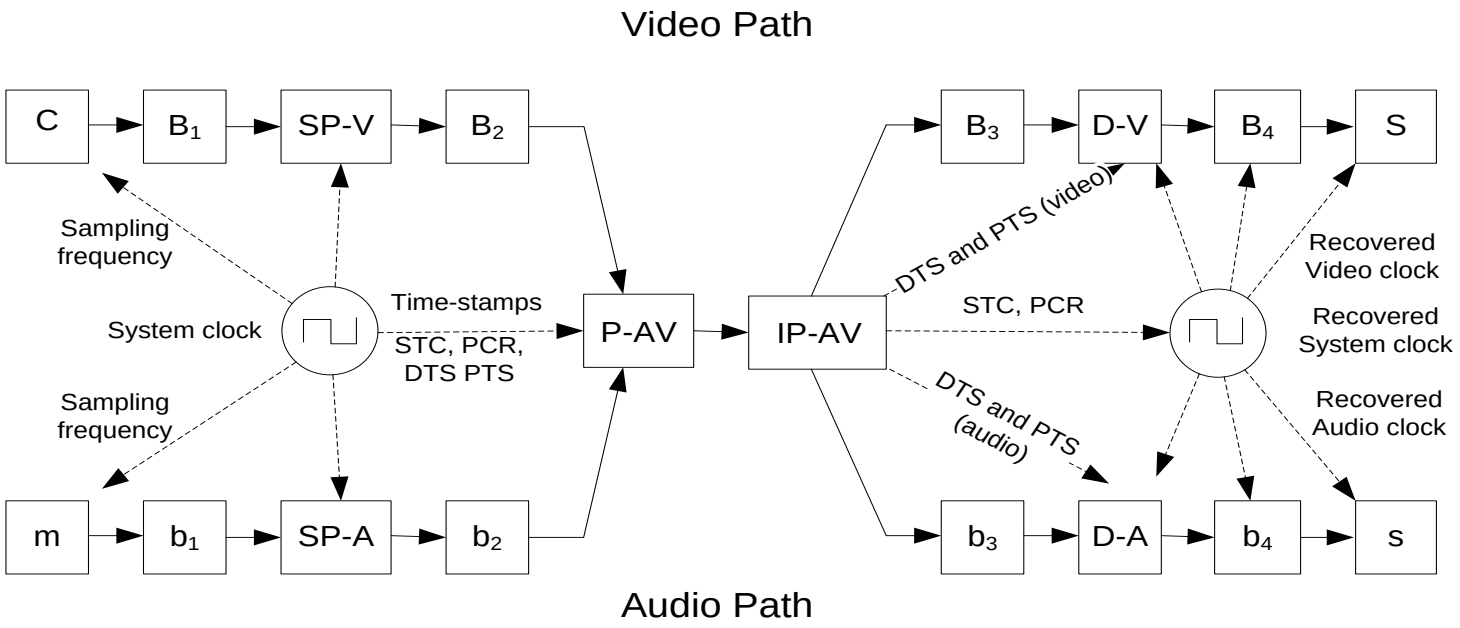
Alignment required in time and frequency between the (multiple) streams

Conventional Approach



- RTP time-stamps are based on a “count” of samples
- Additional step required to translate “count” to “time”
- Frequency offset between video/audio clocks can introduce QoE impairments

Timing Alignment in Multimedia



- Frequency offset (wander) between audio and video sampling results in loss of lip-sync – use System Clock for both
- Frequency offset (wander) between send-side and receive-side system clock results in freeze (video), breaks (audio), and possible loss of lip-sync (align System Clock)

Using Time-stamps for Alignment

- Emulate a constant delay:
  - Generate a “creation” time-stamp C when a block of digital samples are collected from the A/D
  - Predetermine a suitable delay X
  - Convert block to analog at time (C+X)
- Time-stamps for audio and video are struck using a common System Clock
- System Clock at source and destination are synchronized
- Synchronization best achieved using:
  - Common PTP Grandmaster
  - Common GNSS (GPS)

Concluding Remarks

- Using time-stamps linked to a common clock provides the following benefits:
  - Alignment of audio in frequency
  - Alignment of video in frequency
  - Alignment between multiple streams (audio and video)
  - Jitter buffer action to absorb network PDV
  - Prescribed delay
  - Audio and video sources do not have to be in same device (or geographic location)

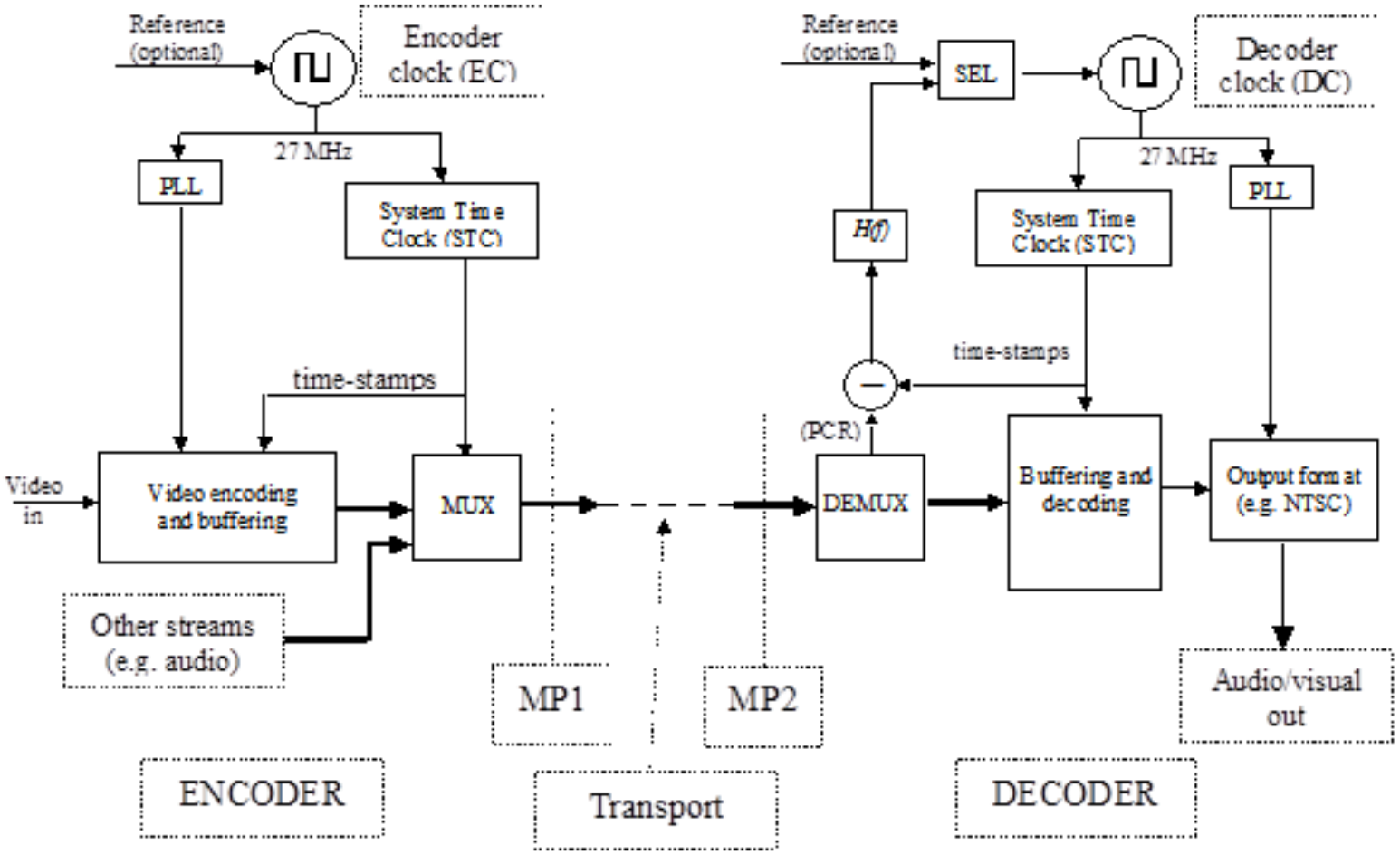
Thank You!

Questions?

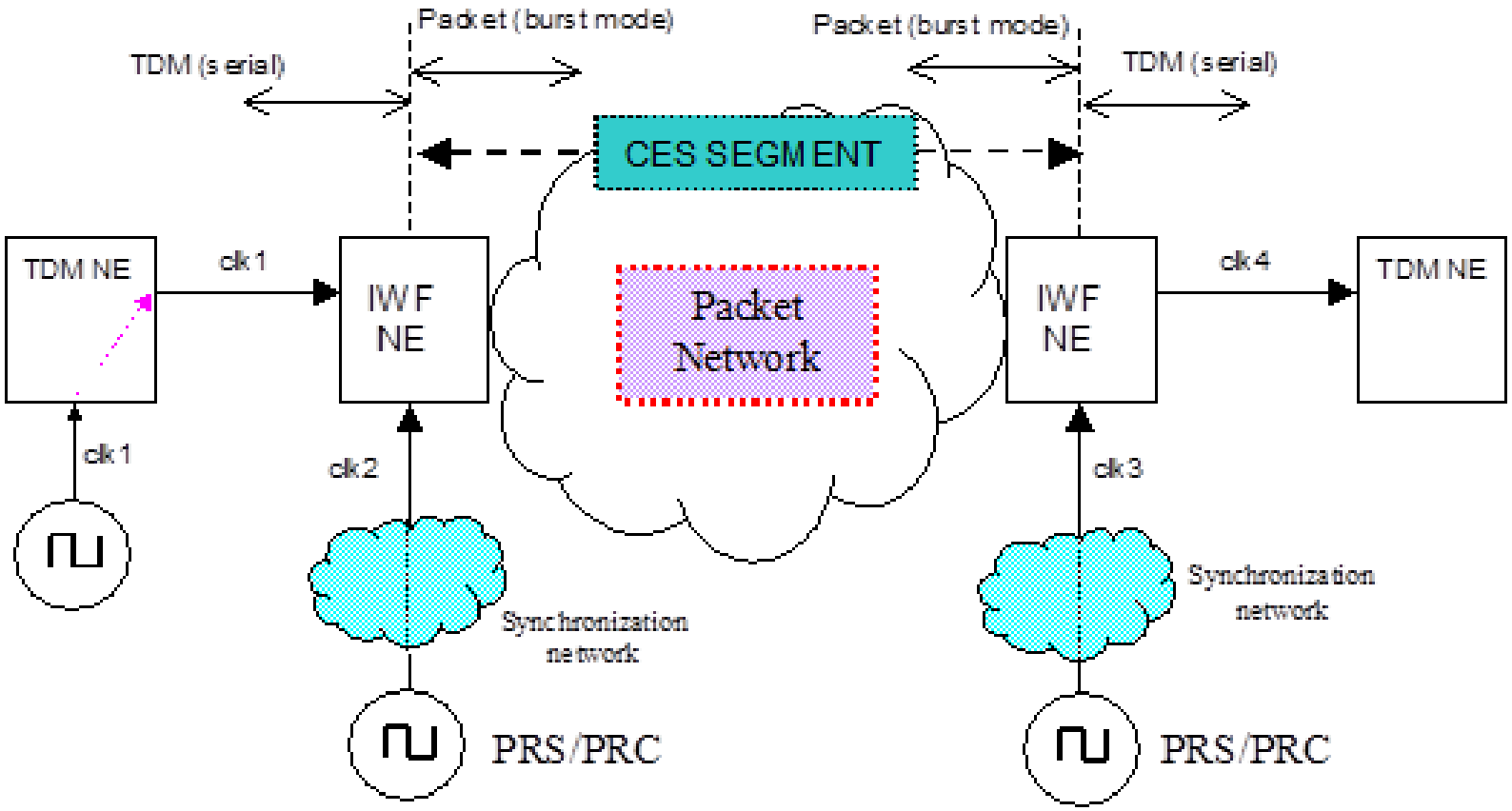
Kishan Shenoi
Title: CTO, Qulsar, LLC
Email: kshenoi@qulsar.com
www.qulsar.com

Back-up slides follow

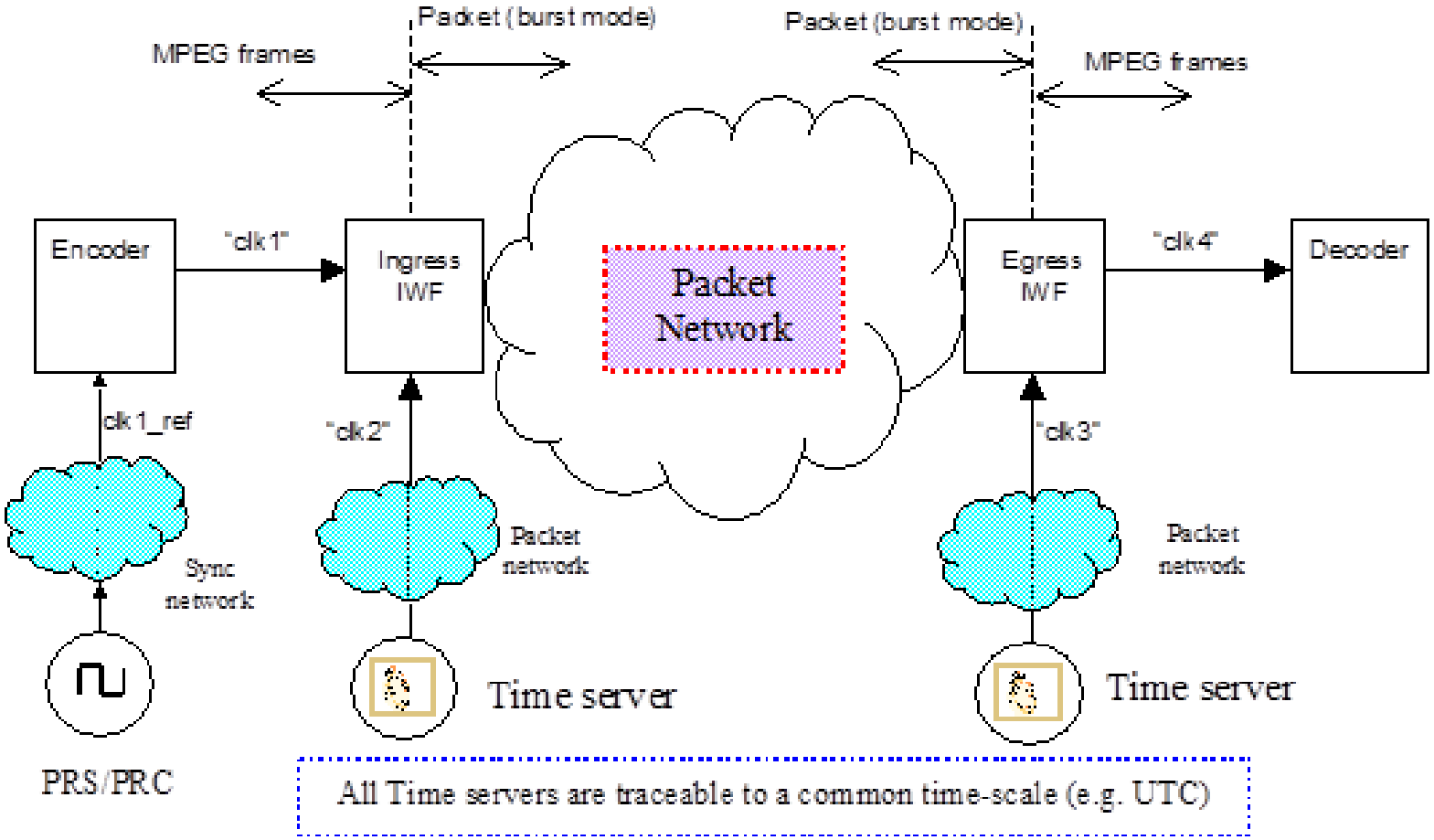
MPEG-2 Timing Model



Robust Circuit Emulation Service



MPEG-2 timing (based on CES)



References

- [1] *MPEG2 Tutorial*, http://www.bretl.com/mpeghtml/MPEGindex.htm.
- [2] *A Guide to MPEG Fundamentals and Protocol Analysis (Including DVB and ATSC)*, An MPEG Tutorial from Tektronix.
- [3] *ATIS IPTV Exploratory Group Report and Recommendation*, Draft Revision: 1.5, June 17, 2005.
- [4] ATIS-0800002, *IPTV Architecture Requirements*.
- [5] ISO/IEC 13818-1, Information Technology — Generic coding of moving pictures and associated audio information — *Systems*
- [6] ITU-T Recommendation H.222.0 {equivalent to [5]}
- [7] ISO/IEC 13818-9, Information Technology — Generic coding of moving pictures and associated audio information — Part 9: *Extension for real-time interface for systems decoders.*
- [8] ETSI TS 102 034 v1.1.1 (2003-5), *Digital Video Broadcast (DVB); Transport of MPEG-2 Based DVB Services over IP Based Networks*.