

Audio Processing for DTV: Solving **LOUD** Commercials and Other Problems in 5.1

Tim Carroll

Linear Acoustic Inc.

Tim@LinearAcoustic.com

What We Will Cover

- Basics of Dolby Digital (AC-3)
- Dolby Digital (AC-3) Decoders
- Metadata: Essentials & Problems
- Loudness Problems (Real World)
- One possible solution
- Q&A

Dolby Digital Quick Review

- One to 5.1 channels of audio
- Highly efficient: 56-640 kbps (5.1 is 448 kbps)
- Part of ATSC, DVB, and OpenCable specifications
- Includes basic provisions for loudness matching:
 - Wideband dynamic range control (dynrng, compr)
 - Data-controlled attenuator in consumer decoders (dialnorm)
- Advanced features require accurate *Metadata*

Metadata

- Data about the audio
- MUST be in time with audio
- Describes number of channels, loudness, how to downmix, etc...
- 20-30 parameters per program

Essential Metadata – Never Leave Home Without These:

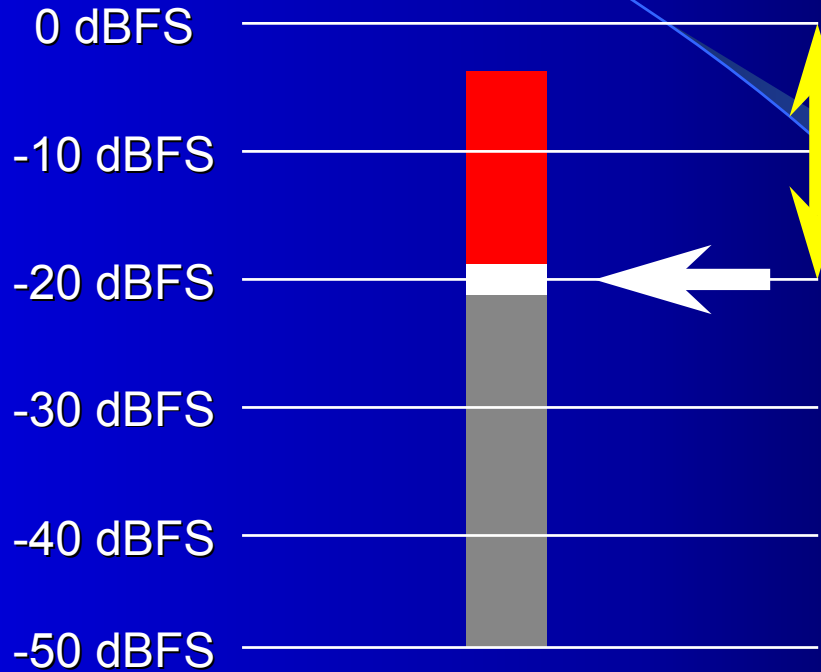
- Dialogue Level
 - Either make the value match the program, -or- make the program match a preset value
- Dynamic Range
 - Pick a preset (Film Standard is good)
 - Make sure RF Overmod is OFF!
- Audio Coding Mode (2/0, 3/2L)
 - Beware of the switching effects!

Dialogue Level (dialnorm)

- A measure of the *Dialog Loudness* of a program.
- Measured using an LAeq meter (long-term A-weighted loudness equivalent), can use a Dolby DP570 or better is LM100.
- Value is unique to each program, However:
 - A single value *can* be chosen and audio can be adjusted to match this value.

ORIGINAL SIGNAL

Digital Full Scale



DIALNORM VALUE

DIALOG LOUDNESS



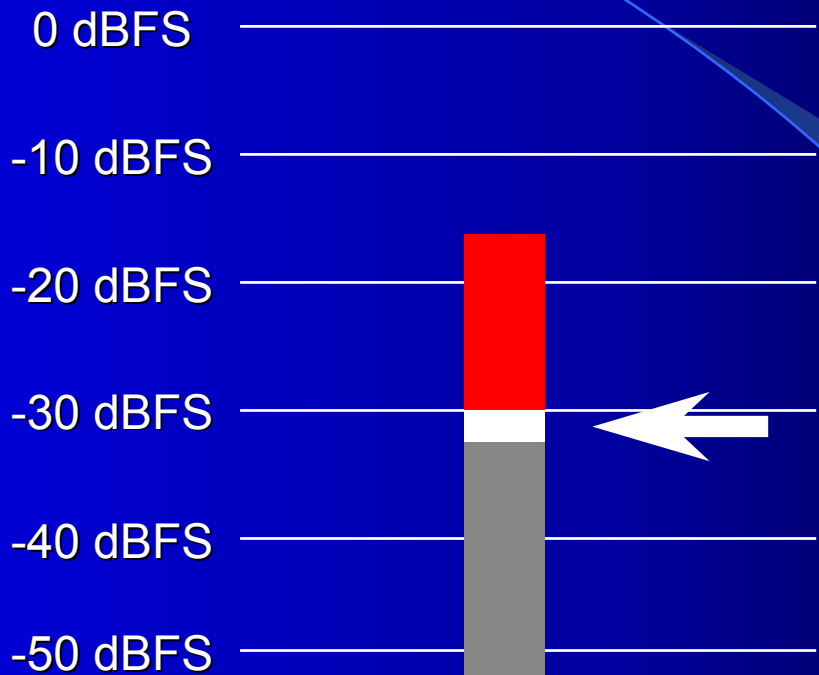
AVERAGE DIALOG



SIGNAL PEAKS

SIGNAL WITH DIALOG NORMALIZATION

Digital Full Scale



PROGRAM LEVEL SHIFTED -11 dB

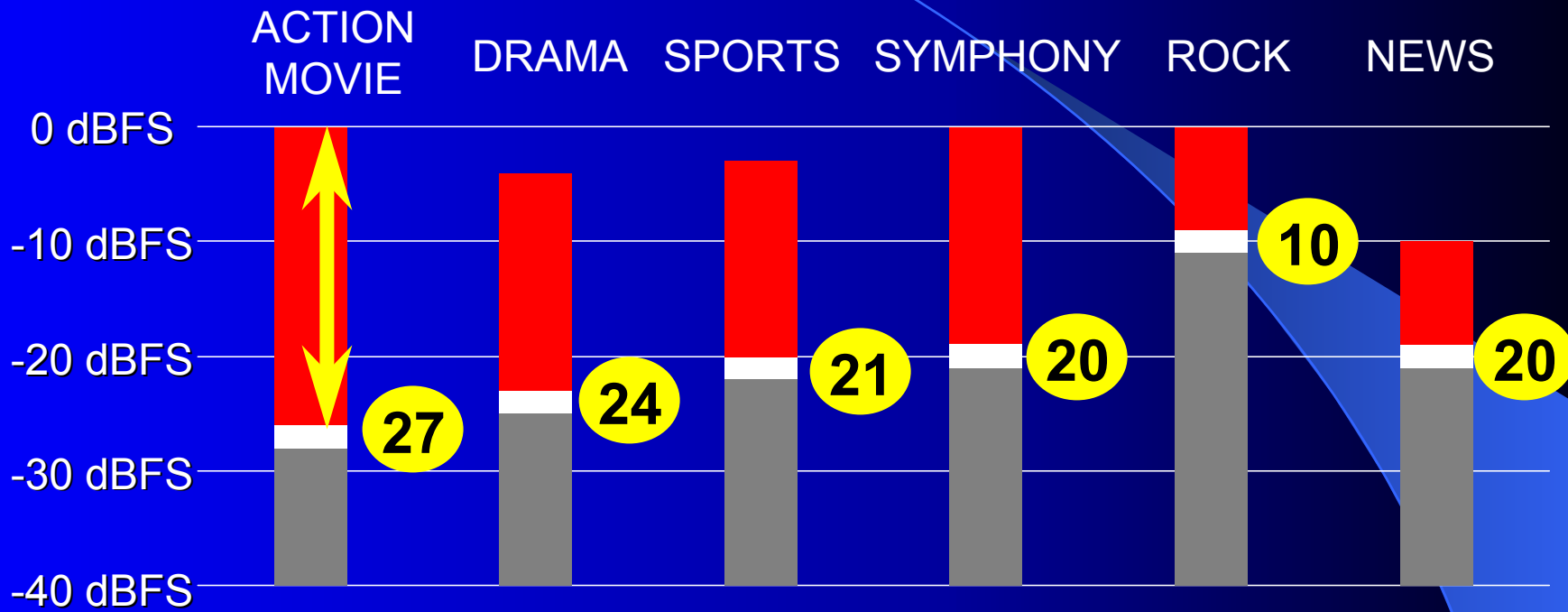
DIALOG LOUDNESS AT -31 dBFS



AVERAGE DIALOG

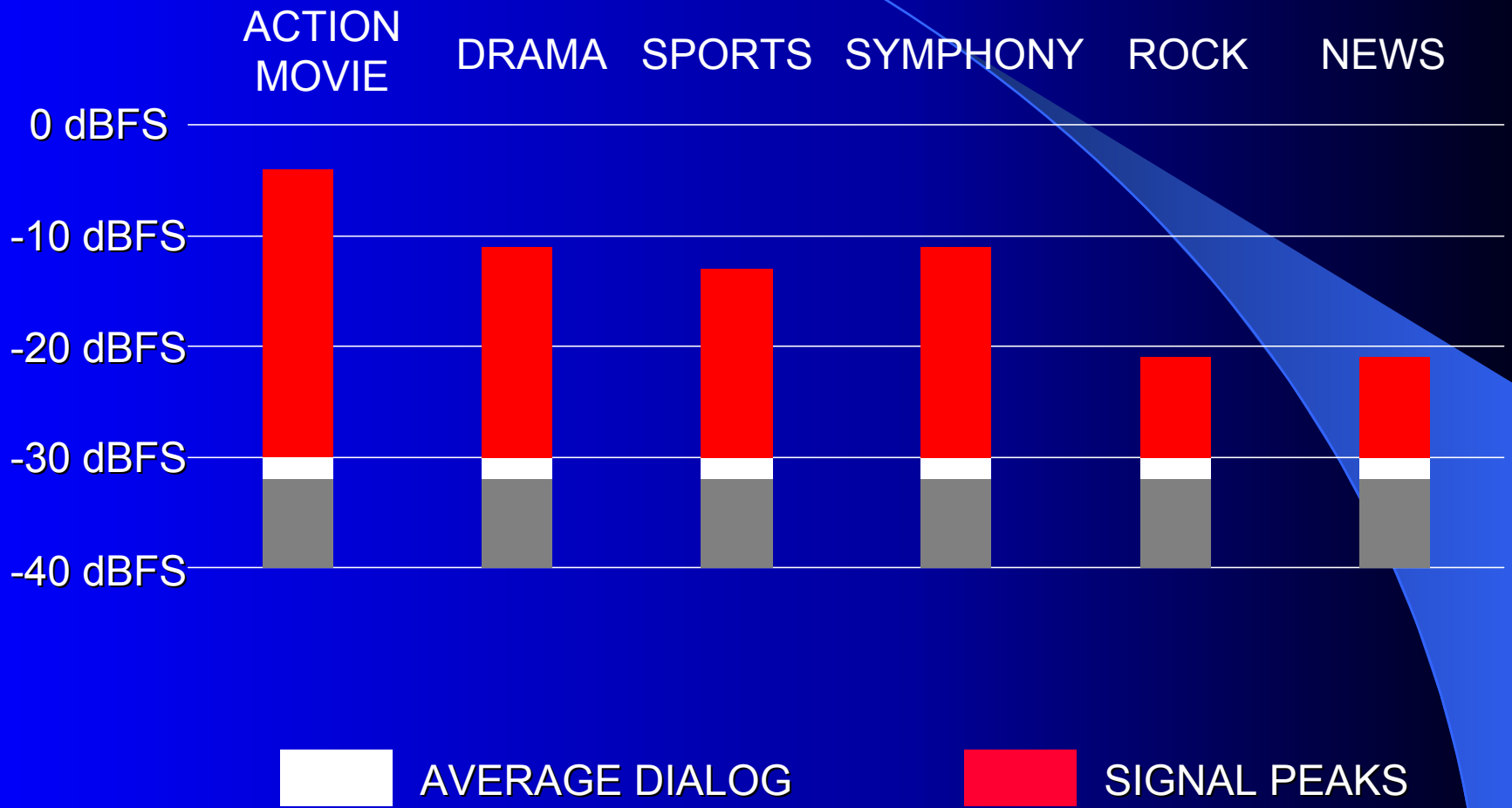


SIGNAL PEAKS



 AVERAGE DIALOG

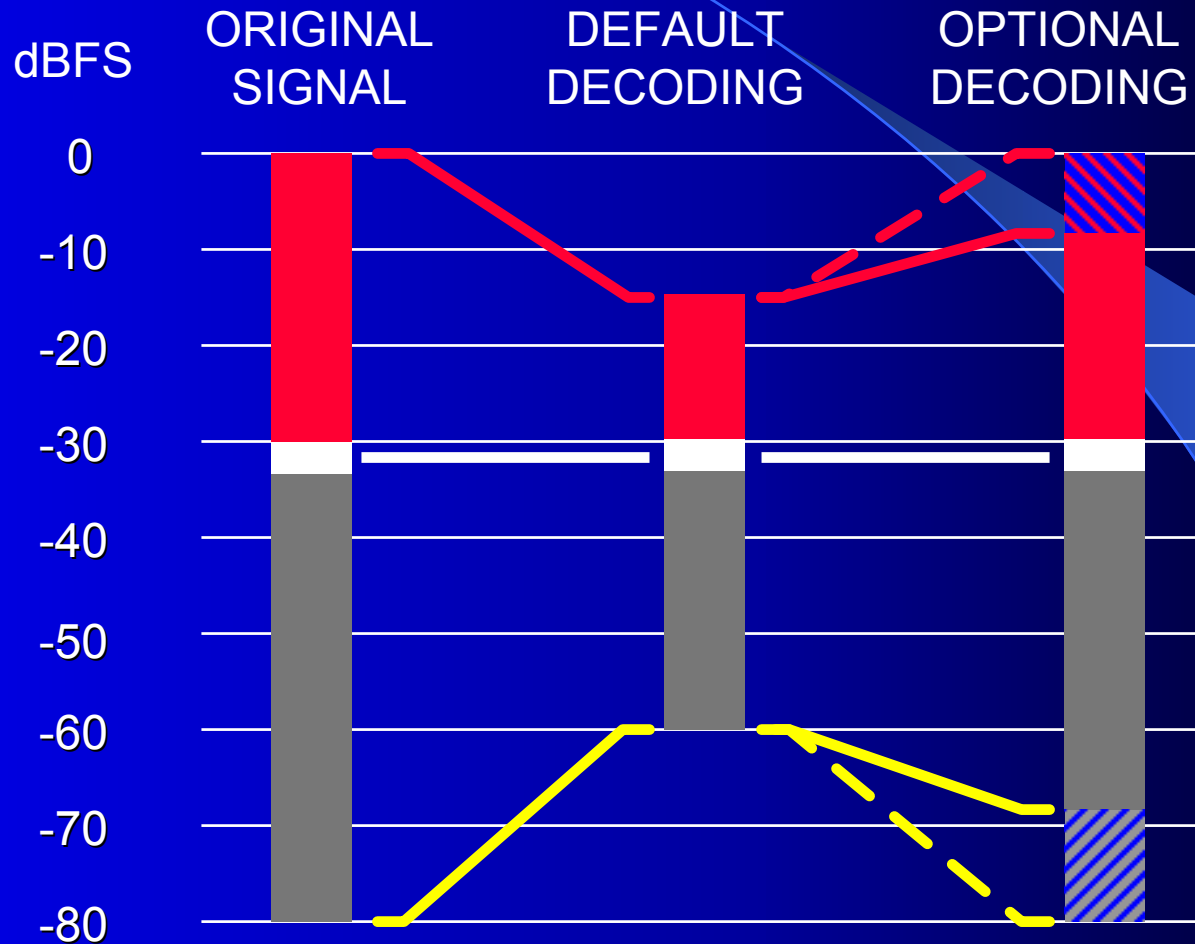
 SIGNAL PEAKS



Dynamic Range

- Control words that allow the *decoder* to modify the dynamic range of the audio:
 - Line Mode (dynrng) = 24dB range
 - RF Mode (compr) = 48dB range
- The control word is actually generated in the Dolby Digital encoder.
- Values are also part of *metadata*, but are set via “Profiles”

Dynamic Range Control

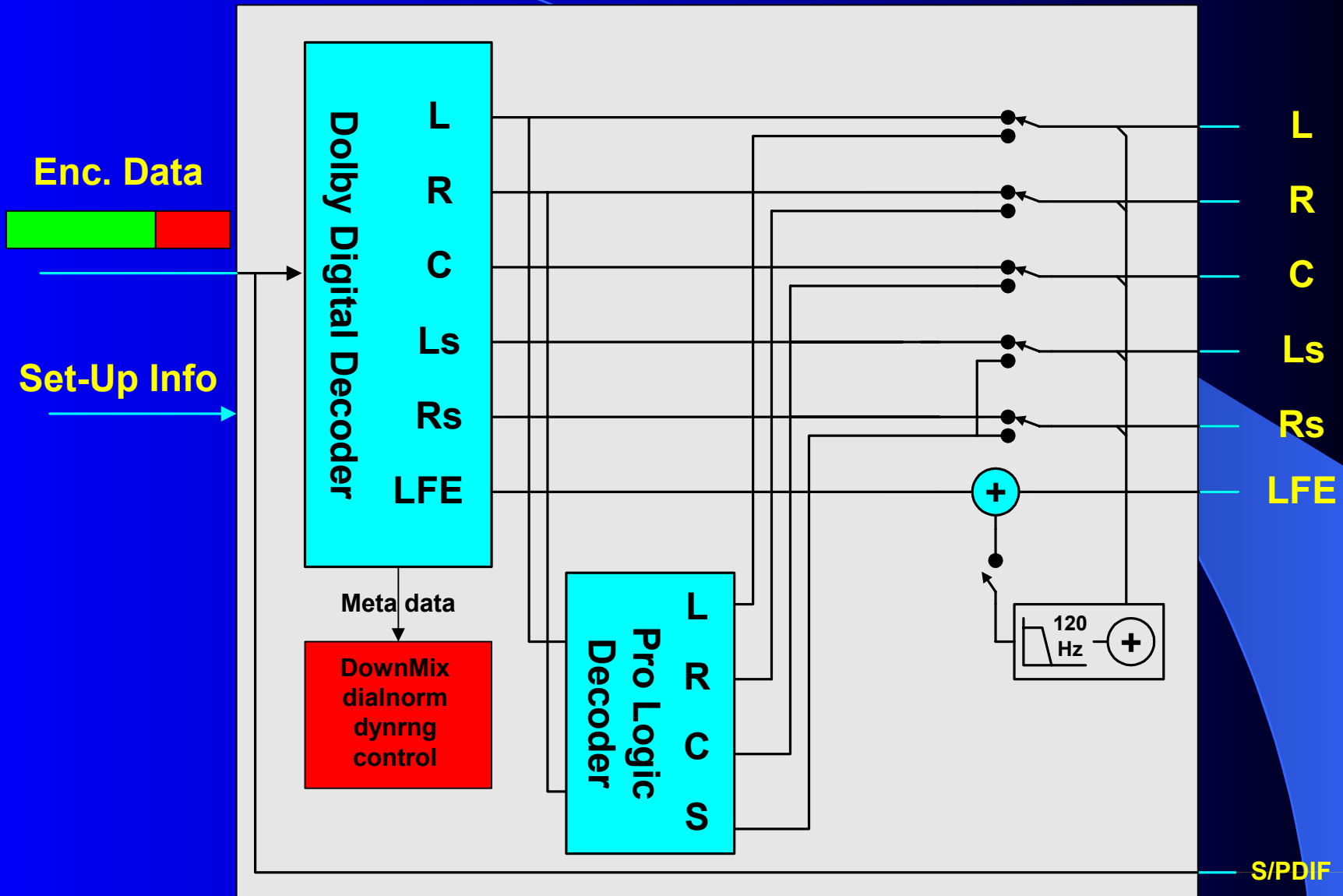


Audio Coding Mode

- Notation is simple: Front/Rear + L (LFE)
- 5.1 = 3/2L
- Stereo or surround = 2/0
- Switching modes causes changes in consumer decoders

Guts of a Decoder

- All two channel products must make AC-3 bitstream available (S/PDIF or TOSLINK)
- Bass management implementation varies, but is usually present in some form:
 - Front speakers large/small
 - Surrounds large/small
 - Subwoofer present?



Consumer Decoders

- All 5.1 channel decoders contain a Pro Logic (and usually a PLII) decoder:
 - Decodes legacy 2-channel material to fill all the speakers
 - May cause a quick code re-boot (mute) during switch, varies depending on method of changing metadata:
 - Single value change (acmod) is best
 - Switching entire metadata stream on VI works
 - Switching Dolby Digital encoder preset is a disaster (i.e. how the audio guys finally got back at the video guys)

Metadata Problems

- Creation – Ideally generated at the time audio program is created, but matches that audio only:
 - If audio levels are changed (i.e. in master control), metadata MUST be re-created
- Distribution – How do you carry this data?
 - If you are distributing baseband, plan for a 115kbps synchronous RS485 layer
 - Dolby E includes metadata path

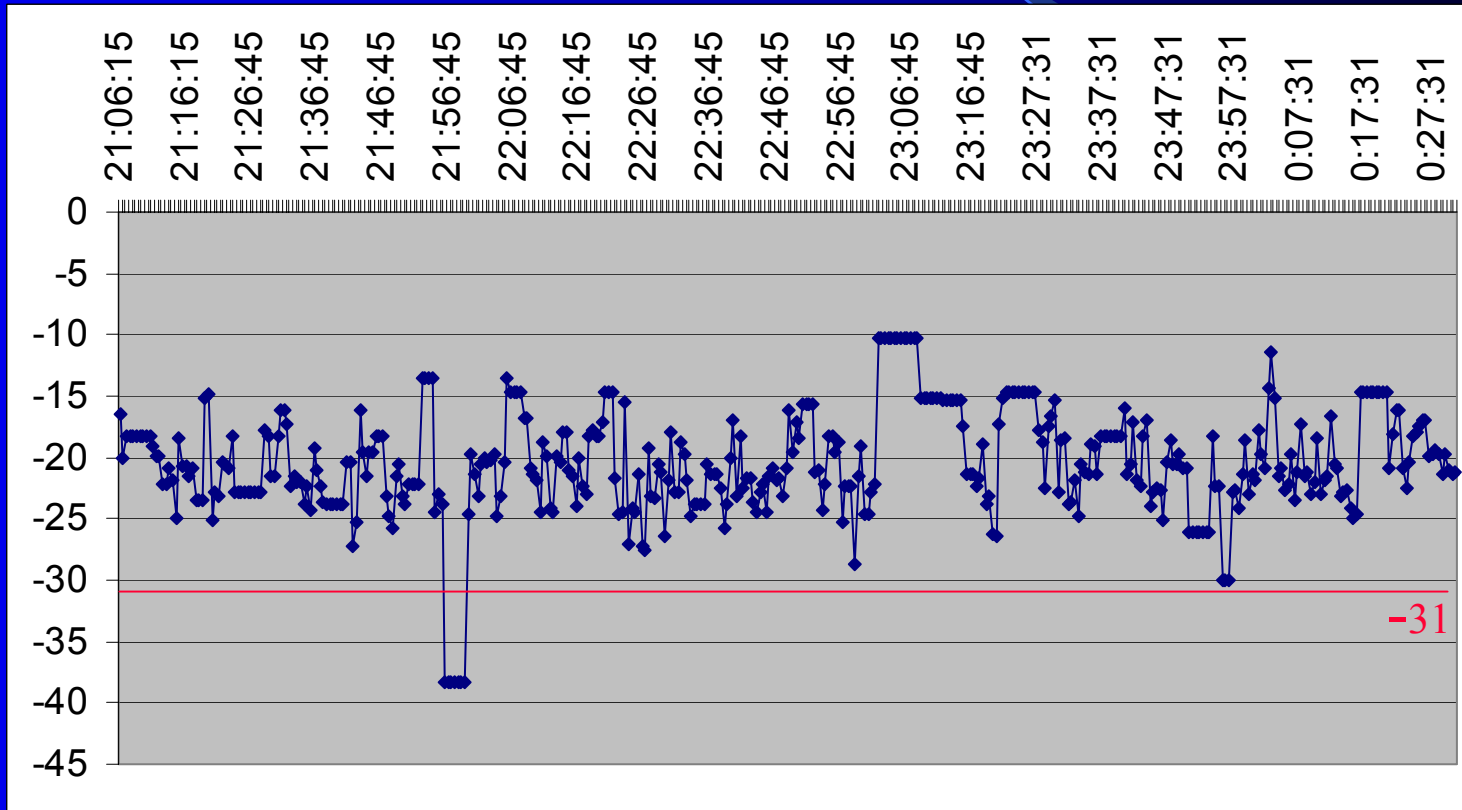
Metadata Problems

- Interruption – What happens if metadata fails?
- “Tricking” the system – What happens if metadata is wrong? (a.k.a. “Trust me, my dialogue level really is -31 even though my audio is $\frac{1}{2}$ LSB from FS...)
- 2004 Grammy Awards broadcast was a good example of both situations

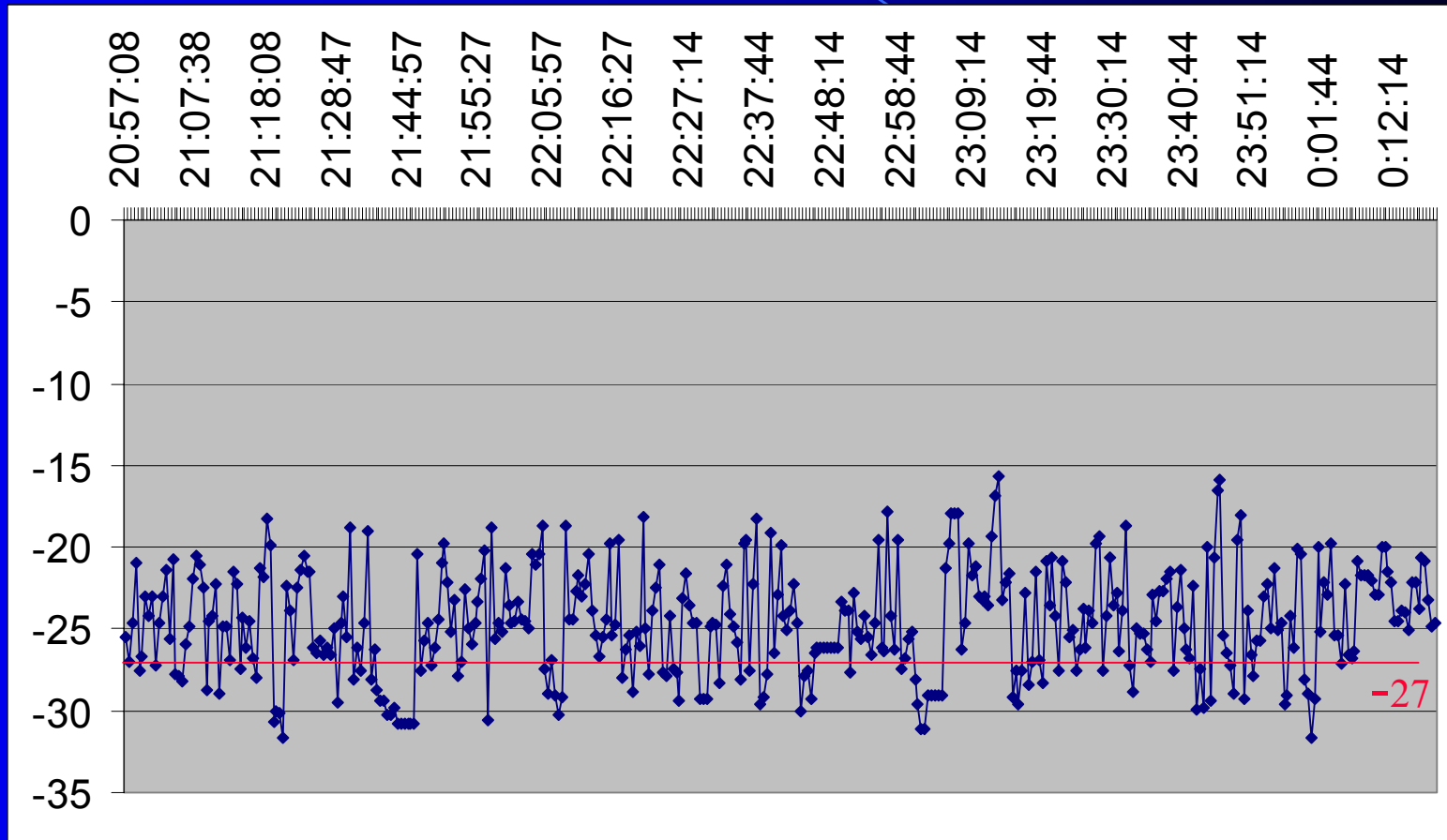
Loudness Issues

- Measurement – Dolby's LM100 makes it easy and accurate, but some of it is still subjective: music vs. dialogue is tough
- Adjusting dialnorm doesn't fix dynamics – matching the dialogue from program to program doesn't take into account the loud explosions

2004 Grammy Awards (NY station)



2004 Academy Awards (different NY Station)



Dynamics Processing is Necessary (but not over processing)

- Dialnorm can center the action, processing can keep it all in check
- As the systems mature, processing can be relaxed (but still present as a safety net)
- With access to metadata, processors can learn to relax automatically

Goals of a DTV Audio Processor

- Handle programs from mono to 5.1 channels
 - Three 2-channel processors does not equal a 5.1 channel processor (as we found out)
- NO clipping – old processors will not work
 - AC-3 is data reduction, clipping ruins efficiency of these systems

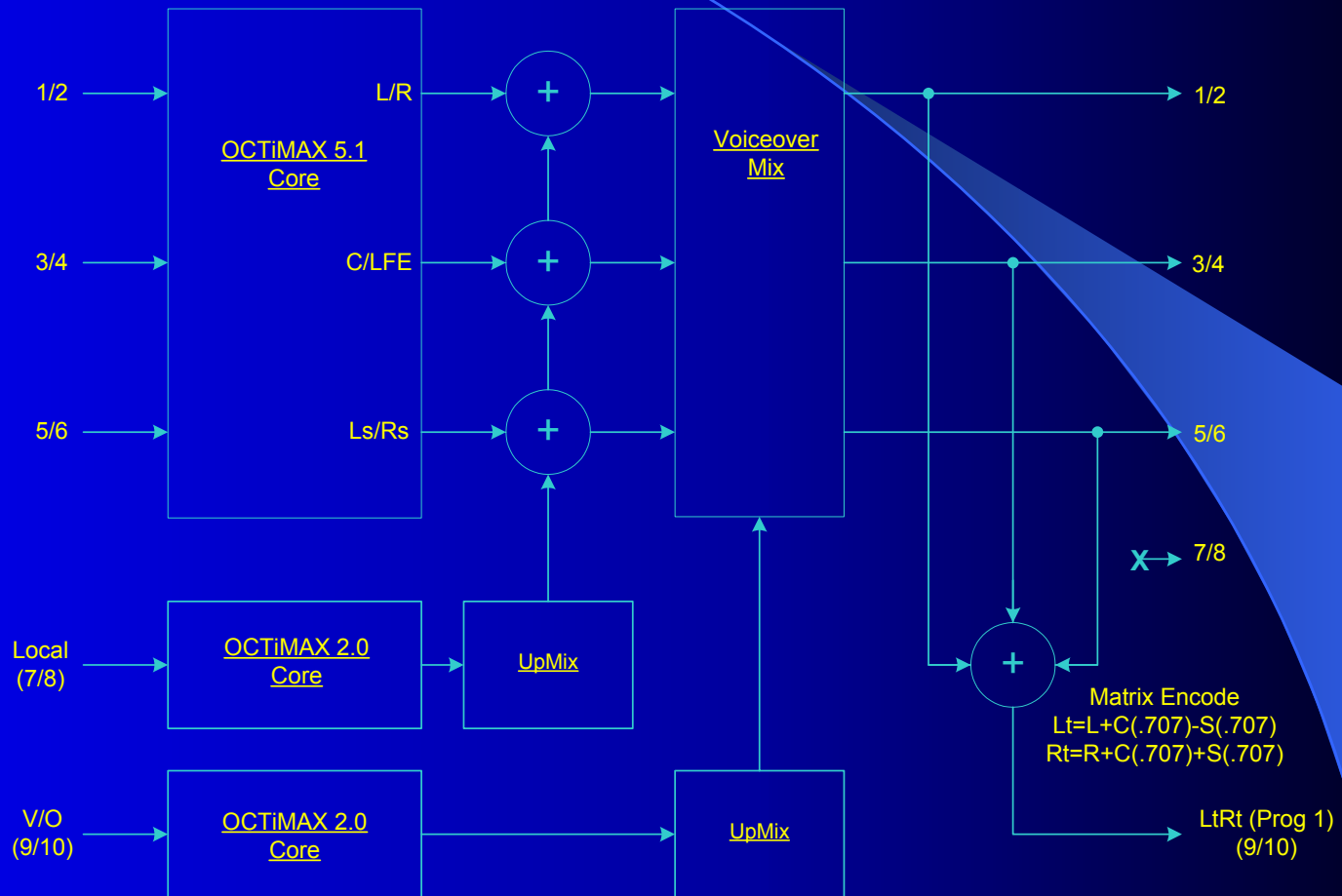
Goals of a DTV Audio Processor

- Must deal with metadata
 - Directly or indirectly, systems must work together
- Local processes:
 - Voiceover of network 5.1 programs
 - Process local audio
- Help integrate legacy 2-channel programs
 - “Upmixing” turns 2-channel into 5.1 channel (with some caveats)

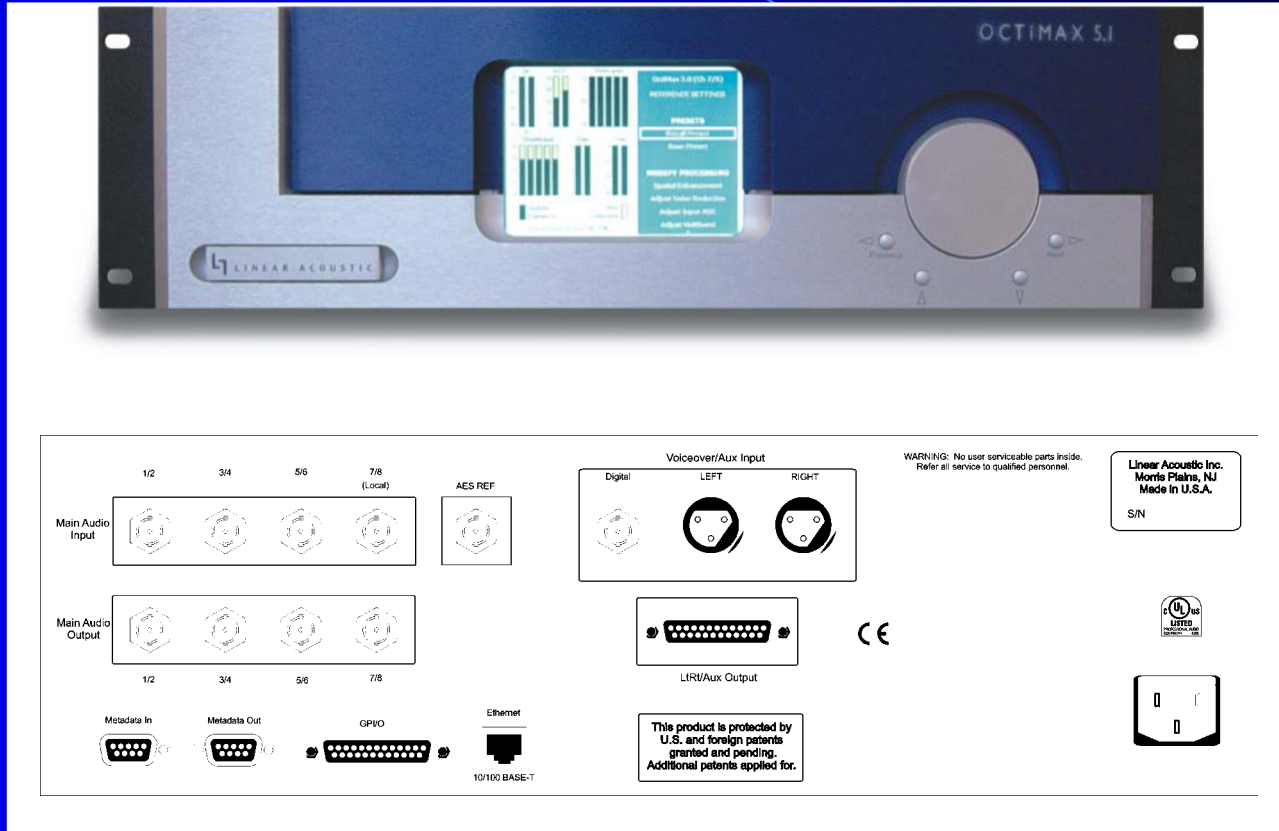
Upmixing

- Not a substitute for a good 5.1 or 6.1 channel mix-
Jerry Springer in pseudo-5.1 is pointless
- MUST be downmix compatible
- Very helpful when trying to maintain dialogue from center channel with no switching artifacts
- Fixes the “2-channel source during 5.1 channel broadcast” problem:
 - If 3/2L mode cannot be changed, all stereo sources come out of Left and Right speakers, no PL or PLII

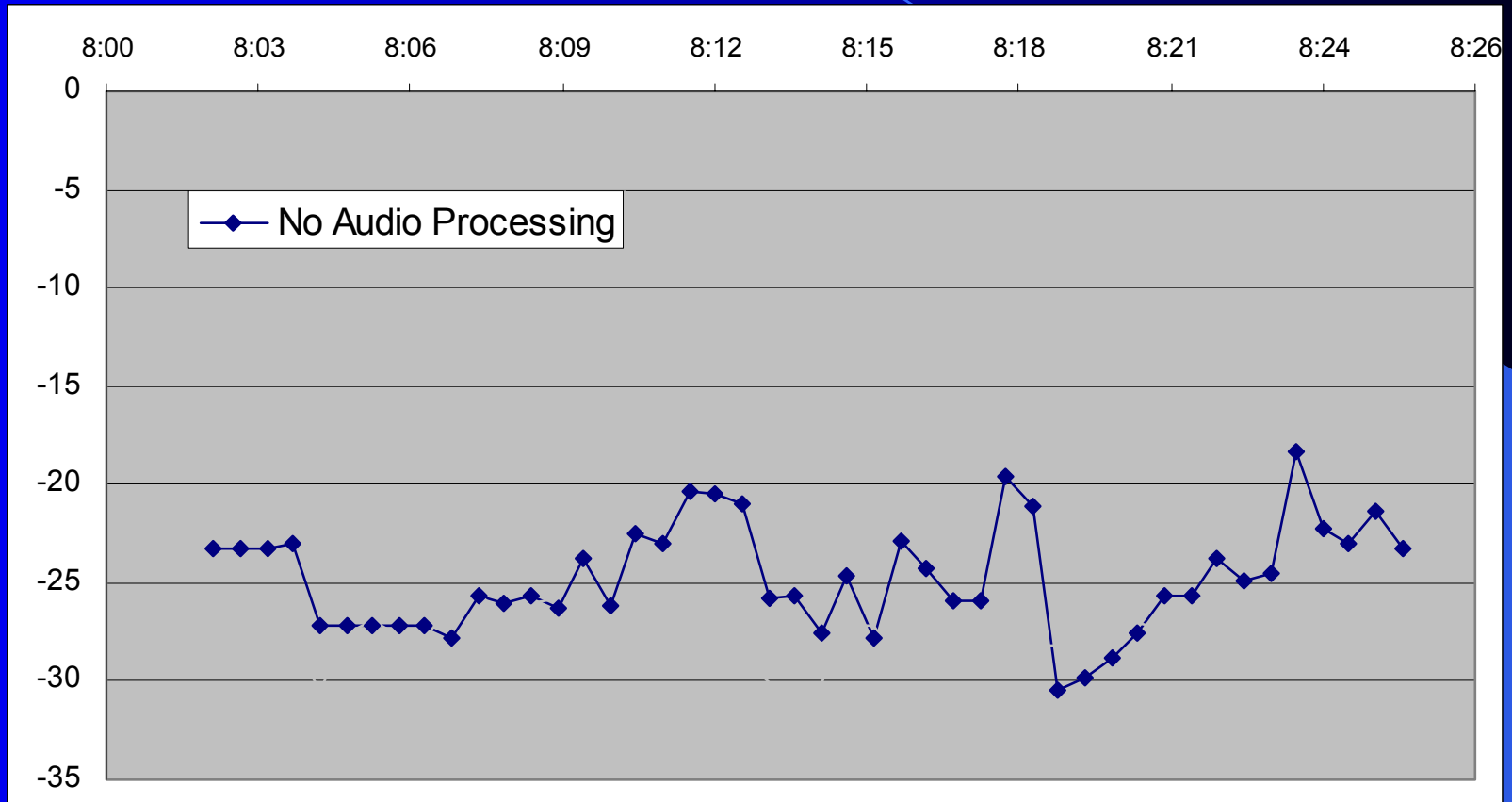
DTV Processor Signal Flow



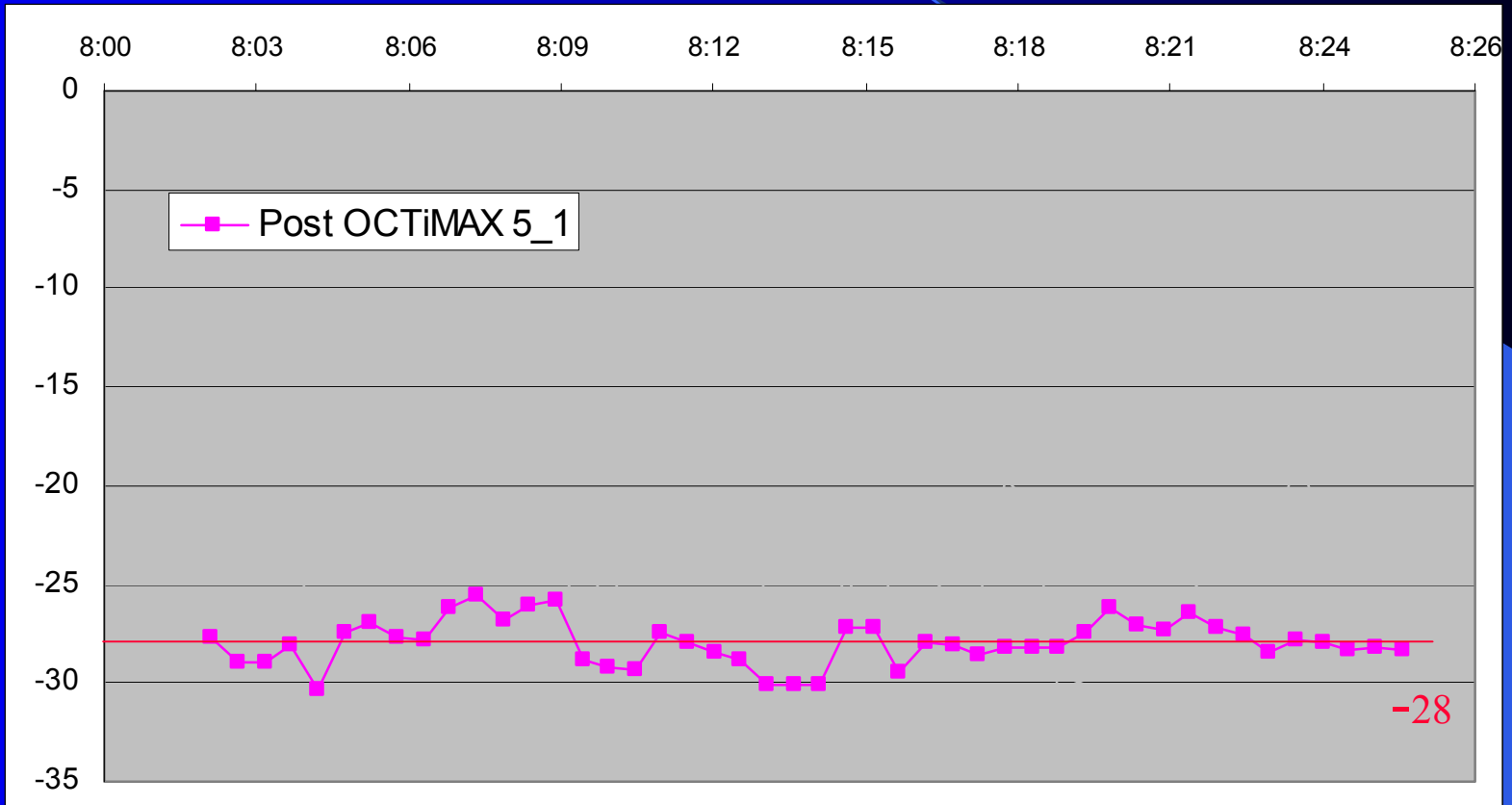
Linear Acoustic OCTiMAX 5.1



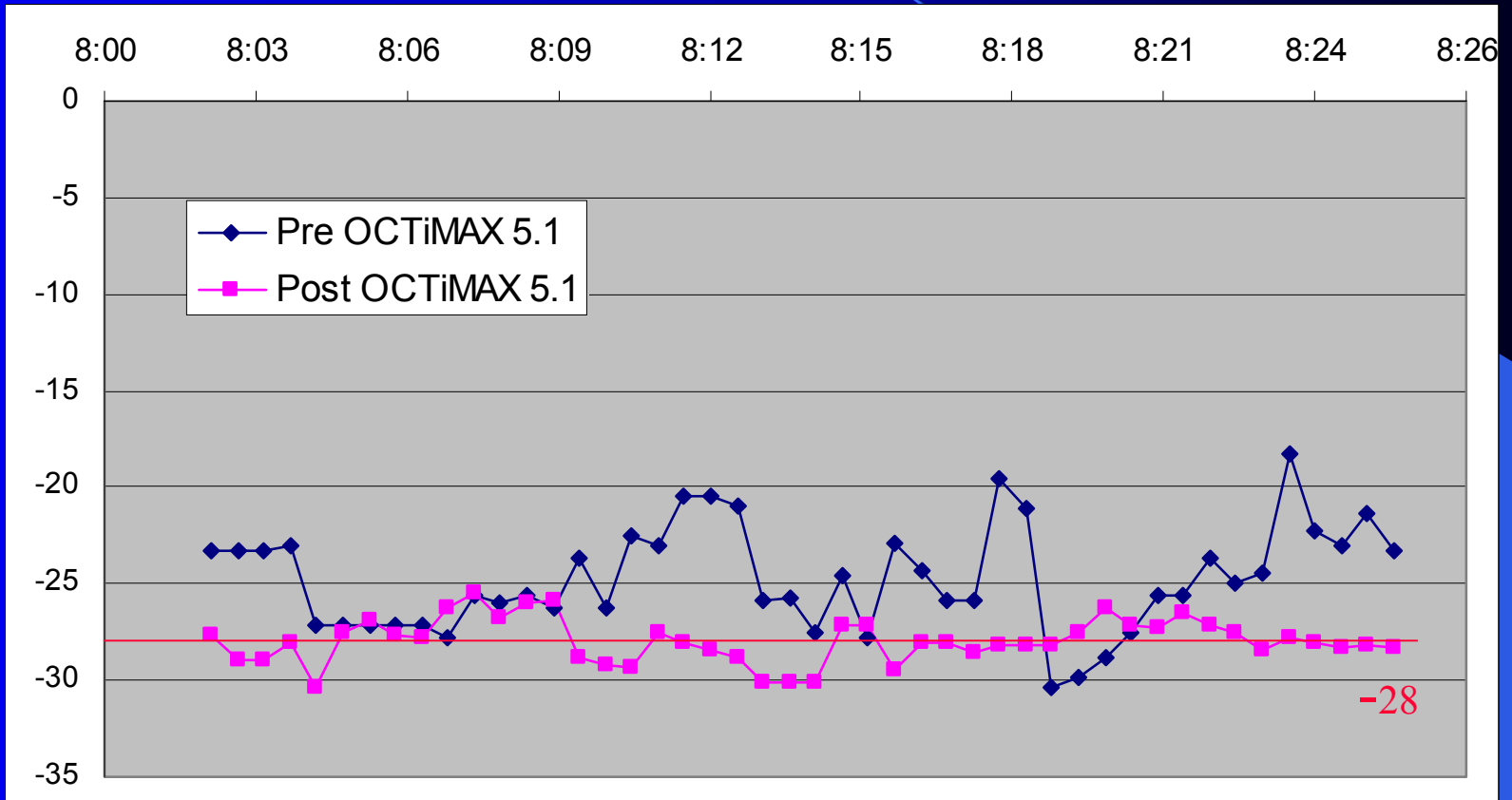
NY DTV Station #1 Revisited



NY DTV Station #1 Revisited



NY DTV Station #1 Revisited

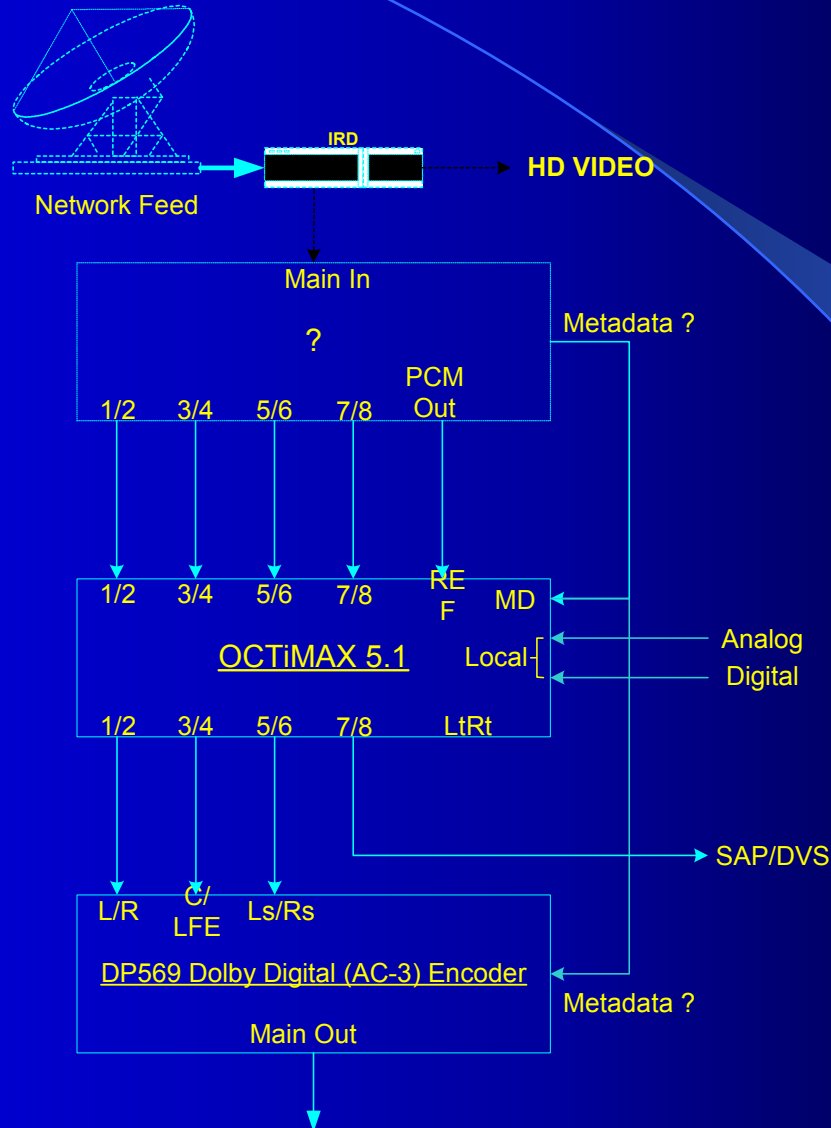


Network Distribution

- ABC – High rate (640 kbps) AC-3
- CBS – Dolby E
- FOX – Dolby E now, transport stream soon*
- NBC – 8 channels via MPEG (built-in)
- PBS – Transport stream (19.39 Mbps)*
- WB, UPN - ? Possibly Dolby E

*Metadata only during network pass-through

Network Distribution



Summary

- DTV audio is complicated – let no one fool you!
- 5.1 channels of audio PLUS metadata
- Dolby Digital (AC-3) solves only the emission part of the puzzle – but how do you feed it?
- Local operations must be accommodated (voiceover, commercial insertion, etc...)
- Audio processing is required for NTSC and is turning out to be a really good idea for DTV

More Information:

- www.LinearAcoustic.com
- www.Dolby.com
- www.ATSC.org

Q & A

Tim Carroll

Linear Acoustic Inc.

Tim@LinearAcoustic.com

www.LinearAcoustic.com

Thanks!

Tim Carroll

Linear Acoustic Inc.

Tim@LinearAcoustic.com

www.LinearAcoustic.com