



Discovery Communications, LLC

## **Global Deliverables: Audio Requirements**

November 2, 2007  
Mark P. Edmondson

## Section 3: Audio Requirements

---

last updated November 1, 2007

### Technical specifications

#### 3.1.a. Audio Phase

All stereo content must be fully mono compatible. When the left and right stereo channels are summed to mono, there should be no phase cancellation, comb filtering, or other discernible change in fidelity. Any use of stereo-enhancement software should be employed in a manner that enhances the stereo experience, without compromising mono playback in any way.

#### 3.1.b Track coherence

All tracks must be time-coherent, aligned to within ½ frame (approx.15 milliseconds) of the full mix.

### 3.2. Audio Levels

Audio Levels are evaluated by means of three measurements: peak program, signal average, and dialog loudness. Program audio must adhere to the following standards for all three measurements:

#### 3.2.a. Peak Program

Peak, or transient, program levels are evaluated using a digital True Peak (PPM) meter with a 0ms rise response. On full mix and mix minus tracks, peak program levels may not rise above -10 dBFS at any point. Please note that PPM's with slower responses will not produce results that satisfy Discovery standards, and should not be used.

#### 3.2.b. Signal Average

Signal average is evaluated using the ANSI C16.5-1942 VU ballistic, with a 300 millisecond per 20dB broadband rise response.

To ensure that mixes occupy a dynamic range best suited for broadcast, RMS/VU levels for narrative speech should maintain the space between -20 and -32 dBFS. This is only a baseline reference; while it is not possible to gauge an exact VU measurement of speech within a full mix, this range should be the target during the mix process. Speech producing RMS/VU readings above or below this range is likely to transmit poorly, and is therefore discouraged.

Within a full mix, RMS/VU levels may not rise above -17 dBFS at any point.

*Note: Some software-based VU meters use a slower (≈500ms) integration time, and consequently produce measurements several decibels lower than 300ms VU meters. Use of such meters will result in measurements that do not match Discovery spec. Please take care, if using a software meter, to ensure that it has a 300ms VU integration time.*

#### 3.2.c. Program Dialog

Program dialog is analyzed using a Dolby LM100 Broadcast Loudness Meter, which is able to measure the *perceived loudness* of human speech discretely within a fully mixed program. Discovery utilizes the Dolby Universal Remote application in obtaining measurements. For a comprehensive tutorial of Discovery's application of the LM100, please refer to the Vendor Guide.

When measured on an LM100 set to measure “stereo” (or “all”, depending on content), with Dialog Intelligence enabled, short term dialog levels should reside between -22 and -32, and infinite term levels should reside between -26 and -28.

When measuring a single track (monaural) mix on the LM100, the reported dialog level will be 3dB lower than if the same signal were to be duplicated out to two tracks and measured as a stereo signal. Consequently, the dialog level standards for monaural mixes are 3dB lower than their stereo equivalent.

### LM100 MEASUREMENT CONTROL SETTINGS

Program Type	Channel Select	Dialogue Intelligence
SD, Stereo	STEREO	ON
HD, 5.1+2	ALL	ON

#### 3.2.d. Dynamic Compression

Programs should be mixed with a dynamic range that is comfortable to experience for sustained periods of time; one that is neither as dynamic as those mixes for DVD/theater playback, nor heavily hyped or compressed.

For more information regarding Discovery’s policy on dynamic compression, please refer to the Vendor Guide.

### GENERAL AUDIO LEVEL GUIDELINES

Track Type	Max PPM 0ms TRUE PEAK	Max VU 300ms/20dB	LM100 Dialog Level SHORT TERM	LM100 Dialog Level INFINITE TERM
Full mix (stereo) SD / NTSC	≤ -10 dBFS	≤ -17 dBFS	Between -22 and -32	-26 to -28
Full Mix (single-track mono) SD / NTSC	≤ -10 dBFS	≤ -17 dBFS	Between -25 and -35	-29 to -31
Full Mix (stereo) SD / PAL	≤ -10 dBFS	≤ -15 dBFS	Between -20 and -30	-24 to -26
Full Mix (single-track mono) SD / PAL	≤ -10 dBFS	≤ -15 dBFS	Between -23 and -33	-27 to -29
5.1 mix HD	≤ -3 dBFS	UNRESTRICTED	Between -22 and -32	-26 to -28
Element Tracks (Music, Dialogue, Narration, FX)	≤ -3 dBFS	UNRESTRICTED	UNRESTRICTED or N/A	UNRESTRICTED or N/A

#### 3.3 Lip Sync

There should be no detectable margin between the timing of visual action and its associated sound.

#### 3.4 Special Circumstances

Any situation in which a content provider is unable to maintain the Discovery-specified volume or loudness ranges and/or standards for quality, in addition to areas of heavy dynamic compression, and/or the intentional use of noise, distortion, static, dropouts or other artifacts for artistic effect, should be clearly explained and noted, with accompanying timecode locations, within the delivered program master materials.

## 3.5. Audio Track Configurations

### 3.5.a. HD Quality Standards

Discovery requires different audio track configurations for HD programs based on their quality (Gold/Silver/Bronze) standard. Programs that adhere to the gold standard, as defined previously in this document, must receive a 5.1 mix along with several secondary elements. Programs that adhere to the silver or bronze standard must meet the standard stereo deliverable.

Primary audio tracks are to be recorded on the videotape master, and secondary tracks must be delivered as Broadcast Wave (.WAV) files on data DVD's. .WAV file specifications and naming conventions are listed in section 3.6.a. A complete listing of all primary and secondary track configurations can be found in section 3.7.

All 5.1 mixes should be delivered in "5.1+2" format (SMPTE standard *320M-1999* for surround track arrangement):

- 1) Left (Lo)
- 2) Right (Ro)
- 3) Center Channel (C)
- 4) Low Frequency Effects (LFE)
- 5) Left Surround (Ls)
- 6) Right Surround (Rs)
- 7) Stereo Full Mix Left or Pro Logic™ Lt (see section 3.5.b)
- 8) Stereo Full Mix Right or Pro Logic™ Rt (see section 3.5.b)

Stereo mixes (tracks 7 and 8) should be acoustically level-matched to their 5.1 counterpart, and consequently, both mixes should measure identically on a Dolby LM100 Loudness Meter.

### 3.5.b. Dolby Pro Logic™/ Dolby Surround

The stereo mix portion of a 5.1 mix may be encoded with the Dolby Pro Logic™ algorithm.

Dolby Pro Logic™ (also referred to as "Dolby Surround") is a system that encodes a two-channel, stereo mix in such a way that it can provide a multichannel surround experience, when decoded with the proper equipment. Dolby Pro Logic™ is transmissible via standard stereo distribution systems, and is commonly supported by many consumer receivers and home theater systems.

The use of Pro Logic™ encoding should be noted within the delivered program master materials.

### 3.5.c. Metadata

Because Discovery level-matches all its programming to an LM100 loudness level of -26 to -28, dialog loudness metadata (DialNORM) should always be set to a value of -27, within Dolby E encoded tracks. All other metadata should be entered at the Dolby default values. For more information on metadata default values, please refer to the Dolby Metadata Guide, available at [www.dolby.com](http://www.dolby.com).

## 3.6. Deliverables

The future of tape, as a means of storing and delivering program content, is limited. Discovery has therefore added digital audio files to its compliment of deliverables. Uncompressed, PCM Broadcast Wave (.WAV) files of certain audio elements now supplement elements delivered on tape.

In addition to providing a first step towards a tapeless future, the system of combined tape + file delivery allows greatly improved freedom in 5.1 mixing, creates a broader range of possibility for reversioning of content, and extends the shelf life of our programming.

### 3.6.a. .WAV file specifications

Broadcast Wave files should be 24 bits in depth, with a sample rate of 48kHz. Files should be named according to their contents, and include the <.WAV> extension: for example <NARRATION.WAV>. Mono files must be properly gain compensated by +3dB: that is, a narration track, metering -20VU when panned to center on a stereo meter, would have a mono equivalent of -17VU. This results in mono stem files that produce LM100 readings identical to their stereo counterpart.

*Note: 16 or 20 bit depth may be used if circumstances, such as workstation limitations or facility design, require doing so. Please note the application of 16 or 20 bit depth within the program master materials.*

*Though popular in the consumer realm, the MP3 format (or any format that utilizes compression) is not acceptable as a Discovery deliverable. Please submit files **only** in uncompressed, PCM Broadcast Wave (.WAV) format.*

### 3.6.b. Reference Tone

Although digital audio allows, in principle, signals to be transmitted and duplicated without the need for analog-style level alignment, reference tone must be present on all audio tracks and files. As its name implies, reference tone establishes the baseline signal level of the content it precedes, and will be utilized as such by Discovery for the foreseeable future.

All NTSC SD and 5.1 HD content, whether tape or .WAV files, must reference a tone of -20 dBFS. All PAL SD content should reference -18dBFS. Although element tracks are subject to fewer technical restrictions than full mix tracks, please take care to ensure that *all tracks* adhere to this baseline. All broadcast (full mix) tracks must reference a 1kHz tone, while non-broadcast tracks (stems) must reference a 500Hz tone (see section 3.7).

Though it may be considered “theoretically correct” to do so, please do not use any reference level other than -20 dBFS for NTSC, and -18 for PAL.

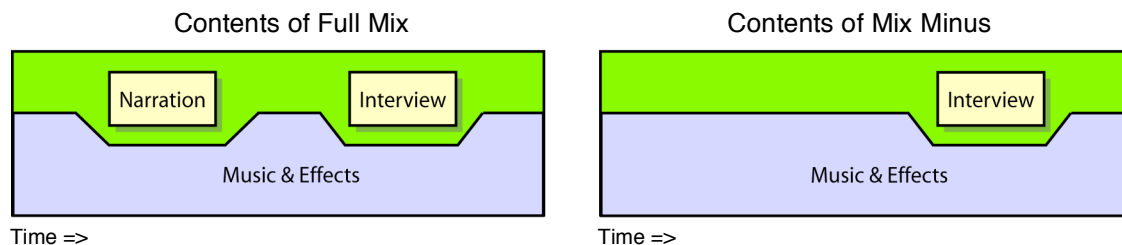
.WAV files must contain reference tone before program start. This tone should align in time with its videotape bars/tone equivalent (usually 1:20 before program start, or 00:58:40:00). .WAV files must also include a “2 pop” sync pulse (a single frame burst of 1k tone), precisely two seconds before program start, to aid in post synchronization.

### 3.6.c. Dips

All source-element tracks or files, a.k.a. “stems” (Music, Effects, M&E, Mix Minus) must be submitted *without dips* for narration or dialog, at *equivalent program volume level*.

For example: a Mix Minus, which contains all elements except narration, should be identical in every way to the full mix, with the exception that during narrated passages, it remains at a constant, or “unity” level (see illustration below).

A Music or FX stem should be at the level it *would be* in the full mix, were no narration present. It must remain at this “unity” level throughout: it should not increase (or decrease) in level, regardless of whether it occurs during narration or dialog.



### 3.7. Track Assignment and Audio Level Specifications

HDCAM-SR: 5.1 Mix							
Track	Contents	Description	PPM	RMS/VU	LM100: Short Term	LM100: Infinite Term	Reference Tone
1	5.1 Mix: Left	Left Channel of Full Mix	≤ -3dBFS	UNRESTRICTED	Between -20 and -32	Between -26 and -28	-20dBFS, 1kHz
2	5.1 Mix: Right	Right Channel of Full Mix	≤ -3dBFS	UNRESTRICTED	Between -20 and -32	Between -26 and -28	-20dBFS, 1kHz
3	5.1 Mix: Center	Center Channel of Full Mix	≤ -3dBFS	UNRESTRICTED	Between -20 and -32	Between -26 and -28	-20dBFS, 1kHz
4	5.1 Mix: LFE (subwoofer)	LFE Channel of Full Mix	≤ -3dBFS	UNRESTRICTED	N/A	N/A	-20dBFS, 60Hz
5	5.1 Mix: Left Surround	Left Surround Channel of Full Mix	≤ -3dBFS	UNRESTRICTED	N/A	N/A	-20dBFS, 1kHz
6	5.1 Mix: Right Surround	Right Surround Channel of Full Mix	≤ -3dBFS	UNRESTRICTED	N/A	N/A	-20dBFS, 1kHz
7	Stereo Mix: Left	Full Mix at equivalent loudness to 5.1 Mix	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 1kHz
8	Stereo Mix: Right	Full Mix at equivalent loudness to 5.1 Mix	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 1kHz
9	Mix Minus Narration & VO: Left	Undipped for Narration or VO	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 500Hz
10	Mix Minus Narration & VO: Right	Undipped for Narration or VO	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 500Hz
11	Music and Effects: Left	Undipped for Narration, VO, or Interviews	≤ -10dBFS	≤ -17 dBFS	N/A	N/A	-20dBFS, 500Hz
12	Music and Effects: Right	Undipped for Narration, VO, or Interviews	≤ -10dBFS	≤ -17 dBFS	N/A	N/A	-20dBFS, 500Hz
48kHz, 24 bit WAV	Stereo Music with "2 Pop"	Undipped stem at equivalent program volume level	≤ -3dBFS	UNRESTRICTED	N/A	N/A	-20dBFS, 500Hz
48kHz, 24 bit WAV	Stereo FX/SOT/BG Dialog with "2 Pop"	Undipped stem at equivalent program volume level	≤ -3dBFS	UNRESTRICTED	N/A	N/A	-20dBFS, 500Hz
48kHz, 24 bit WAV	Mono Interview Dialogue with "2 Pop"	Undipped stem at equivalent program volume level	≤ -3dBFS	UNRESTRICTED	Between -22 and -32	Between -26 and -28	-20dBFS, 500Hz
48kHz, 24 bit WAV	Mono Narration and VO with "2 Pop"	Undipped stem at equivalent program volume level	≤ -3dBFS	UNRESTRICTED	Between -22 and -32	Between -26 and -28	-20dBFS, 500Hz

Notes: Stereo Mix may be Dolby Pro Logic™ encoded (optional). Mono files must be +3dB gain-compensated.

HDCAM: Dolby E Encoded 5.1 Mix							
Track	Contents	Description	PPM	RMS/VU	LM100: Short Term	LM100: Infinite Term	Reference Tone
1	Full Mix: Left	Full Mix at equivalent loudness to 5.1 Mix	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 1kHz
2	Full Mix: Right	Full Mix at equivalent loudness to 5.1 Mix	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 1kHz
3	Dolby E stream: Channel 1	Encoded 8 channel 5.1+2 (Surround + Stereo) Mix	see "Contents of Dolby E Encoded Tracks" below				
4	Dolby E stream: Channel 2	Encoded 8 channel 5.1+2 (Surround + Stereo) Mix	see "Contents of Dolby E Encoded Tracks" below				

Contents of Dolby E Encoded Tracks							
Track	Contents	Description	PPM	RMS/VU	LM100: Short Term	LM100: Infinite Term	Reference Tone
1	5.1 Mix: Left	Left Channel of Full Mix	≤ -3dBFS	UNRESTRICTED	Between -20 and -32	Between -26 and -28	-20dBFS, 1kHz
2	5.1 Mix: Right	Right Channel of Full Mix	≤ -3dBFS	UNRESTRICTED	Between -20 and -32	Between -26 and -28	-20dBFS, 1kHz
3	5.1 Mix: Center	Center Channel of Full Mix	≤ -3dBFS	UNRESTRICTED	Between -20 and -32	Between -26 and -28	-20dBFS, 1kHz
4	5.1 Mix: LFE (subwoofer)	LFE Channel of Full Mix	≤ -3dBFS	UNRESTRICTED	N/A	N/A	-20dBFS, 60Hz
5	5.1 Mix: Left Surround	Left Surround Channel of Full Mix	≤ -3dBFS	UNRESTRICTED	N/A	N/A	-20dBFS, 1kHz
6	5.1 Mix: Right Surround	Right Surround Channel of Full Mix	≤ -3dBFS	UNRESTRICTED	N/A	N/A	-20dBFS, 1kHz
7	Full Mix: Left	Full Mix at equivalent loudness to 5.1 Mix	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 1kHz
8	Full Mix: Right	Full Mix at equivalent loudness to 5.1 Mix	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 1kHz

Note: Stereo Mix may be Dolby Pro Logic™ encoded (optional)

Digital Betacam (DBC)- Stereo Mix							
Track	Contents	Description	PPM	RMS/VU	LM100: Short Term	LM100: Infinite Term	Reference Tone
1	Full Mix: Left	Full Mix	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 1kHz
2	Full Mix: Right	Full Mix	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 1kHz
3	Mix Minus Narration & VO: Left	Undipped for Narration or VO	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 500Hz
4	Mix Minus Narration & VO: Right	Undipped for Narration or VO	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 500Hz

IMX: Stereo Mix							
Track	Contents	Description	PPM	RMS/VU	LM100: Short Term	LM100: Infinite Term	Reference Tone
1	Full Mix: Left	Full Mix	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 1kHz
2	Full Mix: Right	Full Mix	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 1kHz
3	Mix Minus Narration & VO: Left	Undipped for Narration or VO	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 500Hz
4	Mix Minus Narration & VO: Right	Undipped for Narration or VO	≤ -10dBFS	≤ -17 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 500Hz
5	Stereo Music: Left	Undipped stem at equivalent program volume level	≤ -3dBFS	UNRESTRICTED	N/A	N/A	-20dBFS, 500Hz
6	Stereo Music: Right	Undipped stem at equivalent program volume level	≤ -3dBFS	UNRESTRICTED	N/A	N/A	-20dBFS, 500Hz
7	Mono Sound Effects/SOT/BG dialog	Undipped stem at equivalent program volume level	≤ -3dBFS	UNRESTRICTED	N/A	N/A	-20dBFS, 500Hz
8	Mono Interview & foreground dialog	Undipped stem at equivalent program volume level	≤ -7dBFS	≤ -14 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 500Hz
48kHz, 24 bit WAV	Mono Narration & VO with "2 Pop"	Undipped stem at equivalent program volume level	≤ -7dBFS	≤ -14 dBFS	Between -22 and -32	Between -26 and -28	-20dBFS, 500Hz

Track Assignment and Level Specifications for **5.1 material** delivered on **Tascam DA88 format** should adhere to "5.1+2" format (see section 3.5.a)

Track Assignment and Level Specifications for **stereo material** delivered on **Tascam DA88 format** are identical to the **IMX** tape format (above)