Nullsoft Video (NSV) Format 2.1 Specification

Last Updated: January 24, 2005
Author: Justin Frankel

Table of Contents

1. Introduction

Nullsoft Video (NSV) is a file and bitstream format designed to encapsulate generic audio and video together. NSV is intended to be flexible, efficient, and widely supported. Some key features of NSV are:

- NSV supports nearly any combination of audio and video compressor.
- NSV is a bitstream format, meaning that it has no requirement of a whole file or beginning of file to achieve play back. If you have the middle 1/3 of a 30 minute .NSV file, you can play back approximately 10 minutes of the contents.
- NSV's bitstream format is suitable for streaming; accurate synchronization can be acquired at any point in the stream.
- NSV supports auxiliary data chunks for sending multiple audio tracks, subtitles, or other data.

2. NSV file layout

NSV files consist of as many as two parts. The first part which is optional is the NSV file header (see section 2.1). The second part, which is required, is the NSV bitstream (see section 3).

All multibyte integers are specified in LSB format, i.e. the first byte in the integer is the least significant. i.e. 0xDEADBEEF written to the file would look like EF BE AD DE if the file was viewed in a hex editor. When the bitstream specifies A=4 bits followed by

B=20 bits, the A is the low 4 bits of the first byte, and the high 4 bits of the first byte represent the low 4 bits of B.

2.1. NSV file header format

NSV files may contain a single file header. This header can provide information such as file length (in bytes and milliseconds), a table of contents (which can be used during playback for somewhat accurate seeking of VBR contents), and generic meta data.

The format of the NSV file header is as follows:

Length	Name	Description
4 bytes	nsv_sig	File header signature. Must be: 'N','S','V','f' (0x4E, 0x53, 0x56, 0x66)
4 bytes	header_size	Length of file header. May not be 0 or 0xFFFFFFFF.
4 bytes	file_size	Length of file header combined with length of trailing NSV bitstream data. May not be 0, but may be 0xFFFFFFFF, which specified an unknown bitstream length. If this value is less than **header_size**, then the header is not valid.
4 bytes	file_len_ms	Length of NSV bitstream, in milliseconds. May be any value, with 0xFFFFFFFF signifying unknown total length.
4 bytes	metadata_len	Length of meta data, in bytes. May be 0 to signify no meta data.
4 bytes	toc_alloc	Table of Contents Size Allocated, in entries. May be 0 to signify no TOC.
4 bytes	toc_size	Table of Contents size used, in entries. This value MUST be less than or equal to **toc_alloc**
metadata_len bytes	metadata	Meta data (length specified by **metadata_len**). See section 2.1.1 for information on how meta data is encoded.
toc_alloc*4 bytes	toc	Table of Contents (length specified by **toc_alloc**, multiplied by 4 bytes per entry). See section 2.1.2 for information on how the Table of Contents is encoded.

2.1.1. NSV file header meta data format

The NSV file header can contain additional information about the NSV file, such as title, creator, preferred display aspect ratio, or anything else. This meta data can contain any number of name/value pairs. The meta data stored in the NSV file header is stored as zero or more nonzero bytes.

The meta data is formatted as zero or more of the following structures:

```
[whitespace]NAME=<any nonzero character, C>VALUE<C>[whitespace]
```

The whitespace in the above format is optional. NAME is any sequence of characters not beginning with whitespace, and not containing an equals sign (=). NAME should be followed by an equals sign, and then the next character will be used to delimit VALUE, i.e. it will terminate the value. NAME should not be interpreted as case sensitive, so 'Title' should have the same effect as 'TITLE'. The same NAME can be specified multiple times, but depending on the context in which the value is used, the first instance of NAME will typically be used.

Examples:

```
Title=`Justin's Crazy Movie`
Aspect=|0.83333333|
Framerate="29.97"
```

See Appendix A for a list of currently defined NSV meta data NAMEs.

2.1.2. NSV file header table of contents format

TOC 1.0 depreciated TOC format

This TOC format was the original implementation of TOC in NSV, and is very inaccurate. It is still used as a base for storing the new format, and should be supported in its old format for old content.

The **toc** is an array of X 32-bit unsigned integers. The index of each item in the array reflects the time of that entry. For example, if you have a 60 second NSV, and the **toc_size** is 1024 entries, then the 512th entry in **toc** represents the 30 second mark.

The value of each item in **toc** represents the offset in the NSV bitstream. So to seek to 15 seconds in the above example, you would read the 256th entry from **toc**, add that value to the size of the NSV file header, and seek to that offset in the file. Linear interpolation can be used for seeking to arbitrary times in the file.

If the **toc_alloc** is greater than **toc_size**, then there are filler entries following the **toc**.

TOC 2.0 extended TOC format

The original TOC format was very lacking in that it didn't provide enough information for precise seeking. So there is an updated TOC 2.0 format that is mostly backwards compatible.

If **toc_alloc** is greater than 2***toc_size**, and **toc**[**toc_size**] is equal to ('T','O','C','2'), then the TOC is an extended TOC.

The extended TOC format is similar to the old TOC format, with the addition of time information for each point in the file. Each **toc**[x] specifies the offset of a keyframe in the bitstream portion of the file (the size of the NSV file header should be added to get the offset in the file), and **toc**[x+**toc_size**+1] specifies the number of frames (from the start of the file) that the keyframe is. This allows precise seeking.

Note that when using the TOC 2.0, no interpolation should be used, since the TOC entries specify exact keyframe locations.

2.1.3. NSV HTTP meta data header extensions

For utility, meta data can be set in the format of HTTP header extensions, such as

```
x-nsv-title:NewTitle
x-nsv-aspect:0.83333
```

would override the 'TITLE' and 'ASPECT' fields of the NSV file header (or if no NSV file header was found, it will set these fields). See Appendix A for more information on NSV meta data names.

3. NSV bitstream format

The NSV bitstream consists of a sequence of NSV frames, which can contain synchronization frames and nonsync frames. An NSV bitstream MUST contain at least on synchronization frame to be valid, but does not have to contain nonsync frames.

The two types of NSV bitstream frames begin differently, but both ultimately contain the same payload.

3.1. NSV synchronization frame header

Synchronization frames provide an easy to detect frame signature as well as crucial information describing the contents of the video. Typically encoders should set Synchronization frames to be on or right before video keyframes (if applicable). Whenever the decoder needs to resynchronize with an NSV bitstream, it will look for a synchronization frame.

Length	Name	Description
(The first 136 bits of synchronization frames should not change from frame to frame):		
32 bits	sync_hdr	Must be: 'N','S','V','s' (0x4E,0x53,0x56,0x73)
32 bits	vidfmt	video format (see Appendix C for valid formats)
32 bits	audfmt	audio format (see Appendix C for valid formats)
16 bits	width	width of video frame (0 = invalid)
16 bits	height	height of video frame (0 = invalid)
8 bits	framerate_idx	framerate index (0 = invalid) (See Appendix B for information on how framerates are encoded into this byte)
(The next 16 bits can change):		
16 bits	syncoffs	(treat as signed) a/v sync offset (number of milliseconds ahead of the video the audio is at this frame)
? bytes		NSV frame payload (see section 3.3)

3.2. NSV nonsync frame header

Nonsync frames are frames that simply encapsulate more frame data, but do not provide much additional information. Nonsync frames are provided to lower bandwidth requirements.

Length	Name	Description
16 bits	nosync_hdr	Must be: 0xEF, 0xBE

? bytes		NSV frame payload (see section 3.3)

3.3. NSV frame payload structure

The NSV frame payload immediately follows one of the NSV synchronization frame headers or the NSV nonsync frame headers. The payload contains information on lengths for each of the data types, as well as the data itself for the current frame.

Length	Name	Description
4 bits	num_aux	number of auxiliary data chunks present
20 bits	aux_plus_video_len	combined video and auxiliary data length (if this is greater than 524288 + num_aux*(32768+6), the frame should be deemed invalid)
16 bits	audio_len	audio data length (maximum length: 32768 - if the value is greater than this, the frame is not valid)
(auxiliary data) repeat **num_aux** times:		
16 bits	aux_chunk_len	length of data for this aux chunk (if this is greater than 32768, the entire frame should be deemed invalid)
32 bits	aux_chunk_type	32 bit type of the auxiliary data. (See Appendix C for information on valid types).
aux_chunk_len bytes	aux_chunk_data	auxiliary chunk data
(end of auxiliary data) end repeat		
? bytes	video_data	video frame data, length = **aux_plus_video_len** - total auxiliary data chunks
audio_len bytes	audio_data	audio data

The audio and video data packeted in the frame payload should represent approximately one frame's worth of data. The audio data may be sent ahead of or behind the video, as needed.

3.3.1. Auxiliary data chunks

Auxiliary data chunks can be used to send subtitles, pan information for 16:9 to 4:3 pulldown, additional audio tracks, or just about anything else. There can be up to 15 auxiliary data chunks per frame, and each chunk can be anywhere from 0 bytes to 32768 bytes. Each chunk has a 4 byte type. For a list of the currently defined auxiliary types and how their data should be interpreted, see Appendix C.

Appendix A: Common NSV meta data names

The meta data format specified in section 2.1.1 and section 2.1.3 allows for generic name/value pairs to be specified in NSV files and on HTTP served NSV streams. Here is a list of commonly defined meta data items:

Name	Description
Title	Title of file/stream. Should be displayed to the user

URL	URL of relevant information on the file/stream. Useful for attaching homepage information etc.
Creator	Encoder of the content. Useful for putting information regarding what application encoded the material.
Aspect	Preferred aspect ratio of the video. This value describes the relative height of the pixels. For example, to display a 720x480 video in 16:9, the aspect ratio would be 720 / 480 / (16/9), or 0.84375. The default for this value is 1.0, for square pixels.
Framerate	Framerate of the video. Use this when NSV does not support a framerate natively. This should be rare.

Appendix B: Native NSV framerates

The NSV bitstream format can encode to a specific set of framerates (for framerates other than these, you must use the NSV file header or the NSV HTTP header extensions to transmit the framerate). Native framerates are encoded into a single byte, which is stored inside a NSV synchronization frame (see section 3.1). The following psuedocode is used to go from a byte value to the exact framerate:

```
X = byte value
if (X and 0x80 is equal to 0) frame rate = X
otherwise:
  T = (X and 0x7f) shifted right two bits
  if (T is less than 16) S = 1.0 divided by (T plus 1)
  otherwise S = T minus 15.

  if (X and 1) S = S times 1000 divided by 1001

  if ((X and 3) is equal to 3) frame rate = S * 24
  otherwise if ((X and 3) is equal to 2) frame rate = S * 25
  otherwise frame rate = S * 30
```

Basically this formula allows exact representation of all integer framerates from 1 to 127, as well as many multiples and fractions of the standard NTSC and PAL framerates (including drop-frame timecode).

Here is a table of each byte and the framerate associated with that byte:

1=1.0000	2=2.0000	3=3.0000	4=4.0000	5=5.0000
6=6.0000	7=7.0000	8=8.0000	9=9.0000	10=10.0000
11=11.0000	12=12.0000	13=13.0000	14=14.0000	15=15.0000
16=16.0000	17=17.0000	18=18.0000	19=19.0000	20=20.0000
21=21.0000	22=22.0000	23=23.0000	24=24.0000	25=25.0000
26=26.0000	27=27.0000	28=28.0000	29=29.0000	30=30.0000
31=31.0000	32=32.0000	33=33.0000	34=34.0000	35=35.0000
36=36.0000	37=37.0000	38=38.0000	39=39.0000	40=40.0000
41=41.0000	42=42.0000	43=43.0000	44=44.0000	45=45.0000
46=46.0000	47=47.0000	48=48.0000	49=49.0000	50=50.0000

51=51.0000	52=52.0000	53=53.0000	54=54.0000	55=55.0000
56=56.0000	57=57.0000	58=58.0000	59=59.0000	60=60.0000
61=61.0000	62=62.0000	63=63.0000	64=64.0000	65=65.0000
66=66.0000	67=67.0000	68=68.0000	69=69.0000	70=70.0000
71=71.0000	72=72.0000	73=73.0000	74=74.0000	75=75.0000
76=76.0000	77=77.0000	78=78.0000	79=79.0000	80=80.0000
81=81.0000	82=82.0000	83=83.0000	84=84.0000	85=85.0000
86=86.0000	87=87.0000	88=88.0000	89=89.0000	90=90.0000
91=91.0000	92=92.0000	93=93.0000	94=94.0000	95=95.0000
96=96.0000	97=97.0000	98=98.0000	99=99.0000	100=100.0000
101=101.0000	102=102.0000	103=103.0000	104=104.0000	105=105.0000
106=106.0000	107=107.0000	108=108.0000	109=109.0000	110=110.0000
111=111.0000	112=112.0000	113=113.0000	114=114.0000	115=115.0000
116=116.0000	117=117.0000	118=118.0000	119=119.0000	120=120.0000
121=121.0000	122=122.0000	123=123.0000	124=124.0000	125=125.0000
126=126.0000	127=127.0000	128=30.0000	129=29.9700	130=25.0000
131=23.9760	132=15.0000	133=14.9850	134=12.5000	135=11.9880
136=10.0000	137=9.9900	138=8.3333	139=7.9920	140=7.5000
141=7.4925	142=6.2500	143=5.9940	144=6.0000	145=5.9940
146=5.0000	147=4.7952	148=5.0000	149=4.9950	150=4.1667
151=3.9960	152=4.2857	153=4.2814	154=3.5714	155=3.4251
156=3.7500	157=3.7463	158=3.1250	159=2.9970	160=3.3333
161=3.3300	162=2.7778	163=2.6640	164=3.0000	165=2.9970
166=2.5000	167=2.3976	168=2.7273	169=2.7245	170=2.2727
171=2.1796	172=2.5000	173=2.4975	174=2.0833	175=1.9980
176=2.3077	177=2.3054	178=1.9231	179=1.8443	180=2.1429
181=2.1407	182=1.7857	183=1.7126	184=2.0000	185=1.9980
186=1.6667	187=1.5984	188=1.8750	189=1.8731	190=1.5625
191=1.4985	192=30.0000	193=29.9700	194=25.0000	195=23.9760
196=60.0000	197=59.9401	198=50.0000	199=47.9520	200=90.0000
201=89.9101	202=75.0000	203=71.9281	204=120.0000	205=119.8801
206=100.0000	207=95.9041	208=150.0000	209=149.8501	210=125.0000
211=119.8801	212=180.0000	213=179.8202	214=150.0000	215=143.8561
216=210.0000	217=209.7902	218=175.0000	219=167.8322	220=240.0000
221=239.7602	222=200.0000	223=191.8082	224=270.0000	225=269.7303

226=225.0000	227=215.7842	228=300.0000	229=299.7003	230=250.0000
231=239.7602	232=330.0000	233=329.6703	234=275.0000	235=263.7363
236=360.0000	237=359.6404	238=300.0000	239=287.7123	240=390.0000
241=389.6104	242=325.0000	243=311.6883	244=420.0000	245=419.5804
246=350.0000	247=335.6643	248=450.0000	249=449.5504	250=375.0000
251=359.6404	252=480.0000	253=479.5205	254=400.0000	255=383.6164

Appendix C: Valid NSV audio/video/auxiliary data types

A sequence of 4 bytes. These bytes can be 'a'-'z', 'A'-'Z', '0'-'9', '-', '.', '_', or a space (' ' or 0x20). The first byte must not be a space (0x20).

Example:
"MP3 ", "DivX", "VP31", "TEXT", "S.TL" and "PCM " would all be valid types.

Audio Type	Data description
"NONE"	No audio
"PCM "	PCM audio data. The first byte in the frame is the bits per sample ("bps", typically 16), the second byte in the frame is the number of channels ("nch", typically 1 or 2), and the third and fourth byte in the frame are the low and high bytes respectively of the sample rate ("samplerate", i.e. 0xAC and 0x44 would be 44100 Hz). The remaining data should be PCM samples, and should be aligned to (nch * bps / 8) bytes.
"MP3 "	MP3 compressed audio data.
"AAC "	MPEG-2 AAC compressed audio data.
"VLB "	Dolby MPEG-2 AAC compressed audio data.
"SPX "	OGG speex compressed audio data.

Video Type	Data description
"NONE"	No video
"RGB3"	RGB ordered pixel data, 24 bits per pixel packed.
"YV12"	YV12 planar YUV data. average of 12 bits per pixel.
"VP50"	On2 VP5 compressed video data
"VP60"	On2 VP6 compressed video data
"VP31"	On2 VP3 compressed video data
"VP3 "	On2 VP3 compressed video data (legacy, has U and V planes of output switched. do not use. this will go away soon)
"DivX"	DivX 4 or DivX 5 compressed video data

Auxiliary Type	Data description
"ASYN"	Audio/Video resync. Size should be 0. If the decoder encounters one of these, it should resync the audio and video following this frame. Useful for when concatenating/streaming two files where the audio is not the same granularity as the video. In the future this should be extended to have a 32 bit signed value which would specify how many milliseconds of silence to add/remove
"SUBT"	Subtitle. One or more of: 2 byte size of following data, NUL terminated language string, NUL terminated UTF-8 subtitle string, 2 byte latency (how many frames to wait before showing), 4 byte length (how many frames to show for), 1 byte X position (0 = left, 255 = right), 1 byte Y position (0=top, 1=bottom), 3 bytes of color (R, G, then B), 1 byte of relative font size, and any additional data (which is currently ignored) -- TODO: make more documentation on this format from code
(defined but not used, yet)	
"TIME"	64 bit little endian timestamp (ms)
"FRME"	64 bit little endian timestamp (frames)
"AUXA"	Auxiliary audio track. 1 byte track ID, 4 byte format descriptor, 2 bytes of video sync offset, plus audio data.

Appendix D: Licenses

Copyright License

Patent License

America Online believes it may have certain patent rights that would be applicable to any fully compliant implementation of this Specification. In order to encourage the free and unrestricted development of implementations compliant with this Specification, America Online agrees to provide you with the following Patent License.

BY MAKING, USING, SELLING, OR OTHERWISE DISTRIBUTING A LICENSED IMPLEMENTATION, YOU AGREE TO THE TERMS OF THIS PATENT LICENSE.

1. Reciprocal Grants. America Online, Inc., on behalf of itself and its Affiliates, grants to you (“Licensee”) a royalty-free, non-exclusive, non-transferable, non-sublicenseable, worldwide license under its Necessary Claims to make, use, sell, offer to sell, import, and otherwise distribute Licensed Implementations to other Licensed Entities and End Users, subject to a reciprocal grant by you, on behalf of yourself and your Affiliates, to America Online, Inc., its

Affiliates, and all other Licensed Entities of a royalty-free, non-exclusive, non-transferable, non-sublicenseable, worldwide license under your Necessary Claims to make, use, sell, offer to sell, import, and otherwise distribute Licensed Implementations and to authorize such Licensed Implementations to be used by End Users.

2. No Other License. No license or right is granted except as expressly set forth herein.

3. Suspension/Termination. America Online, Inc. reserves the right to suspend or terminate this license if you or your Affiliate asserts that America Online, Inc., any of its Affiliates, or any other Licensed Entity infringes any patent claim as a result of or in connection with the use of this Specification or any Licensed Implementation. You have the right to suspend or terminate your reciprocal grant to any other Licensed Entity that asserts that you, any of your Affiliates, or any other Licensed Entity infringes any patent claim as a result of or in connection with the use of this Specification or any Licensed Implementation.

4. Definitions. “Affiliate” means an entity that is directly or indirectly Controlled by a party to this agreement (“Party”), so long as such Control exists, where “Control” means beneficial ownership of more than fifty percent (50%) of the voting power or equity in an entity, or the direct or indirect right to manage the business affairs of an entity.

“End User” means any ultimate end user of a Licensed Implementation for its own personal use and not to develop, distribute, operate for the benefit of others, or use commercially.

“Licensed Entity” means any person or entity that has agreed to the terms of this Patent License.

“Licensed Implementation” means only those specific portions of an implementation that are fully compliant with this Specification.

“Necessary Claims” means those claims of all patents and patent applications, (other than design patents and design registrations), throughout the world, to which a Party or any Affiliate thereof has the right, at any time, to grant hereunder without such grant or the exercise of rights thereunder resulting in an obligation on the part of such Party or Affiliate to pay royalties or other consideration to any third party (except for payments to Affiliates or employees within the scope of their employment): (i) which are necessarily infringed by an implementation of this Specification where such infringement could not have been avoided by another commercially reasonable noninfringing implementation of this Specification, or (ii) for which infringement is based upon an implementation of any example included in the body of this Specification. Necessary Claims shall not include any claim other than as set forth above even if such claim is contained in the same patent as a Necessary Claim. Moreover, Necessary Claims shall not include claims that read on: (i) the implementation of separate published standards or specifications (including, for example and without limitation, MPEG4, IETF and other standards) which may be referred to, incorporated by reference, or included in this Specification, or (ii) enabling technologies (including, for example and without limitation, operating systems, microprocessors, and software platforms) that may be necessary to make or use any product or portion thereof that complies with this Specification, but which are not expressly set forth in this Specification.

5. Disclaimer of Warranties; Limitation of Liability. THIS SPECIFICATION IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, AND NONINFRINGMENT OF THIRD PARTY INTELLECTUAL PROPERTY RIGHTS ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SPECIFICATION, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

6. Miscellaneous. This License is made under, and shall be construed according to, the laws of the Commonwealth of Virginia, excluding its conflicts of laws principles. Any notice or other communication under this License shall be given in writing, if to America Online, Inc. at 22000 AOL Way, Dulles, VA 20166, and if to Licensee, at such address as America Online, Inc. can reasonably ascertain. In the event that any provision of this License conflicts with the law under which this License is to be construed or if any such provision is held invalid by a court with jurisdiction over the parties to this License, such provision shall be deemed to be restated to reflect the original intentions of the parties in accordance with applicable law, and the remainder of this License shall remain in full force and effect.

End of Document.