

Technical Standards for Delivery of TV Programmes to



Technical Standards for Delivery of TV Programmes to SVT

The SVT Standards include:

- Technical Specifications, i.e. the technical production methods which must be used, and the parameters which all material must meet to be acceptable by SVT.
- Picture and Sound Quality requirements, which also form a binding obligation on producers of material. Assessment of quality is by nature subjective, and is highly dependent on the nature of the programme. Some of the Quality Requirements are expressed in relative terms (“reasonable”, “not excessive” etc.), and it will be necessary to make a judgement as to whether the quality expectations of the intended audience will be fulfilled, and whether SVT will feel that value for money has been achieved.
- Delivery Requirements, which specify the form and layout of the programme material.

Every programme submitted for transmission must satisfy a Quality Control (QC) process specified by SVT. Any programme failing the QC process on tape or file may be rejected and returned to the supplier for repair.

Technical Responsibility

General Responsibility

SVT’s department for Technical development has the responsibility for this document. Questions, remarks and thoughts are appreciated and could be directed to our mailbox:

Mail to: techspec@svt.se

This document does not cover any specific requirements for delivery of programmes to co-producers.

File and Tape programme delivery

The Duty Engineering Managers (DEMS) in SVT Media Centres are the main point of contact for technical enquiries affecting immediate (defined as “on the day”) delivery.

Duty Engineering Managers (DEMS):

Stockholm: tel. +46 8 784 69 14 / +46 8 784 76 80

Göteborg: tel +46 31 83 71 46 / +46 31 83 72 50

Malmö: tel +46 40 22 71 57 / +46 40 22 72 00

Umeå: tel +46 90 17 50 06 / +46 90 17 51 40

For all other enquiries please contact your commissioner or consult the SVT B2B web site:

www.svtb2b.se

File Naming Requirements

Programme files delivered to SVT must be named as follows:

A prefix identifying a programme for delivery to SVT (PG-), a programme-id (seven numbers, 0-9) identifying the programme (series), a hyphen, and an episode number (three numbers, 0-9) followed by a version-identifier, followed by a hyphen and the programme title in upper case (one or more characters, A-Z).

Any spaces in the programme title must be replaced with underscore “_”. The maximum length of the title is 238 characters.

Filename extensions must be in lowercase. Umlaut characters (e.g. Å, Ä and Ö) are *not* permitted.

Contact SVT for programme specific identifier information.

Requested Programme File Name Format	Example Filename	Notes
PG-1234567-001A-PROGRAMME_TITLE.xxx	PG-1300828-006A-SASKADETLATA.mxf	xxx = mxf, mov, wav, doc, xls etc.

Live Programme Delivery

Circuit Bookings and Circuit Handling

The SVT point of delivery is currently the Master Control Room (MCR) in the “TV House”, Oxenstiernsgatan 34 in Stockholm. However as new buildings come on-line programmes may be required to deliver to other locations. Unless otherwise agreed, it should be assumed MCR is the required point of delivery.

The production company is responsible for the costs of all communications and for ensuring all circuits are booked from the source to the point of delivery. Bookings of circuits should be directed to SVT’s contribution booking office at SVT News Media Centre.

Media Centre News

Telephone: +46 8 784 20 20

Mail: kontribution@svt.se

Hours: Every day, around the clock

The production company should also inform the MCR of the full route preferably two weeks in advance, or as soon as a circuit route has been decided before the programme production date.

Master Control Room (MCR)

Telephone: +46 8 784 75 50 / +46 8 784 75 51

Hours: 06:00 – 23:30 Monday to Friday

06:00 – 23:30 Saturday, Sunday and Public Holidays

For deliveries outside these hours MCR should be informed 28 days in advance.

Facilities for down linking satellite circuits to MCR in Stockholm are available at commercial rates. SVT also has permanent circuits to MCR.

Play-out Point of Contact

The Duty Engineering Managers (DEMS) in MCR are the main operational point of contact working closely with the playout editors. Technical enquiries affecting immediate delivery should be made to:

Duty Engineering Managers (DEMS), tel. +46 8 784 75 50 / +46 8 784 75 51

SVT Resilience Levels

Resilience levels will be discussed with each live programme during the commissioning stage. However, as a guide, programmes will need to meet the following:

OPTION A: Programmes with a major Swedish and International public interest

Main and Backup links are geographically and electrically separate as far as possible. Where satellite is used, there are two separately located uplink trucks, powered differently from each other. The source should be able to continue in some form, during a break down or power loss affecting some facilities. The programme producers must discuss the contingencies for alternate content should all facilities be affected.

OPTION B: Live Primetime Programmes (e.g. Saturday early evening)

Main and Backup links and power supplies have some commonality, which creates an accepted single-point-of-failure risk. Reserve links could be lower quality or even SD. A reserve power supply should be available; however it only need be capable of supplying the minimum infrastructure to keep some of the programme on air. The programme producers should discuss alternate content should all facilities be affected.

OPTION C: All other programmes including live links into studio programmes

A single link, with no backup is acceptable. There only need be minimal alternate power supply and mobile phones can be used as alternate communication devices. SVT may provide alternate content in the event of programme failure. The Production is responsible for filling any loss of link into a live studio programme to meet the scheduled duration.

Contents

General Responsibility.....	2
File Naming Requirements.....	2
Circuit Bookings and Circuit Handling.....	3
Play-out Point of Contact.....	3
SVT Resilience Levels.....	3
1 General Quality Requirements	8
1.1 Picture Quality.....	8
1.2 Sound Quality.....	8
1.3 Access for People with Disabilities.....	8
2 Technical Requirements – Video	9
2.1 High Definition Format.....	9
2.1.1 Origination.....	9
2.1.2 Post-production.....	9
2.1.3 Film motion or “film effect”.....	10
2.1.4 Field dominance.....	10
2.2 Video Line-Up.....	10
2.3 Video Levels and Gamut (illegal signals).....	10
2.3.1 Measuring signal levels.....	10
2.3.2 Tolerance of out of gamut signals.....	10
2.4 “Blanking”.....	10
2.5 Aspect Ratio.....	11
2.5.1 “Cinemascope ratio” letterbox.....	11
2.5.2 Floating images.....	11
2.5.3 “Pillarboxed” HD material.....	11
2.6 Archive Material.....	11
2.6.1 General quality – archive.....	11
2.6.2 Aspect ratio – archive.....	11
2.6.3 Safe areas – archive.....	12
2.7 Use of Non-HD material.....	12
2.7.1 Non-HD material.....	12
2.7.2 Up-converted SD video material.....	12
2.8 Film for High Definition Acquisition.....	12
2.9 Photosensitive Epilepsy (PSE).....	12
2.10 Safe Areas for Captions.....	13
2.10.1 14:9 Caption Safe Area.....	13
2.10.2 4:3 Caption Safe Area.....	13
2.10.3 16:9 Caption Safe Area.....	13
2.11 Standards Conversion.....	13
3 Technical Requirements – Audio.....	14
3.1 Stereo Audio Requirements.....	14
3.1.1 Stereo line-up tones.....	14
3.1.2 Stereo audio levels and measurement (loudness or volume).....	14
3.1.3 Stereo phase.....	15
3.2 Surround Sound Requirements.....	15
3.2.1 Surround line-up tones.....	16
3.2.2 Surround audio levels and measurement (loudness or volume).....	16
3.2.3 Surround phase.....	16
3.2.4 Audio metadata.....	16
3.3 Sound to Vision Synchronisation.....	16
3.3.1 Audio / Video sync markers.....	16
3.4 Dynamic Range.....	17

4	File Delivery Requirements	18
4.1	File format	18
4.2	Video codec	18
4.3	Image format	19
4.4	Audio	19
4.4.1	Track allocations	19
4.5	Programme Layout / Format	20
4.5.1	Start and end	20
4.5.2	The Ident Clock or Slate	21
4.6	3D Delivery	21
4.7	Closed captions (Subtitles)	21
4.8	Timecode	21
4.9	Audio only files	21
4.10	SD Files (Legacy programmes only)	21
4.10.1	File format	22
4.10.2	Video codec	22
4.10.3	Image format	22
4.10.4	Audio essence	22
4.11	Metadata	22
4.11.1	Filenames	22
4.11.2	SVT Metadata application	22
4.11.3	Delivery Requirements in MXF	22
4.11.4	Delivery requirements in XML	22
4.11.5	Required Metadata	23
4.12	File Delivery Options	26
4.12.1	Hard Disk	26
4.12.2	LTO Tape	26
4.12.3	File transfer over Leased line / circuit	26
4.12.4	File transfer over a secure managed network service	26
4.12.5	File Transfer using sFTP	26
5	Tape Delivery Requirements	27
5.1	Videotape recording	27
5.1.1	Tape format	27
5.1.2	"i" and "psf" Flags	27
5.1.3	Time-code	27
5.2	Programme Layout / Format	27
5.2.1	Start and end	28
5.2.2	Programmes longer than a single tape	28
5.2.3	Compilation tapes	28
5.2.4	Ad breaks	28
5.2.5	The Ident Clock	28
5.3	Paperwork	28
5.4	Audio Track layout	29
6	Live Delivery	30
6.1	Responsibilities of the Production	30
6.2	Definitions	30
6.3	Link Specifications	31
6.3.1	Uncompressed via Optical Fibre:	31
6.3.2	Compressed via Optical Fibre:	32
6.3.3	Compressed via Satellite Link:	32
6.3.4	Compressed via Microwave Point-to-Point Link:	33
6.4	Picture Quality & Bit Rates (concatenation issues)	33
6.5	Standards Conversion	33
6.6	Standard Definition	33
6.7	Audio	34
6.7.1	Stereo Audio	34
6.7.2	Multichannel Audio (when required)	34

6.7.3	Audio Track Allocation	34
6.7.4	AV Synchronisation.....	35
6.7.5	AV Synchronisation – Radio Link Cameras.....	35
6.8	TOD timecode.....	35
6.9	Cue and Communication	35
6.10	Photosensitive Epilepsy (PSE) and Live Programmes	36
6.11	Ancillary Data (VANC)	36
6.12	UPS & Generator Provision	36
7	3D (Interim for Live and Tape only)	37
8	Appendix 1. Encoders used for Live Delivery – Additional Requirements	38
8.1	MPEG-2 Encoders – additional requirements	38
8.2	MPEG-4 Encoders – additional requirements	38
9	Appendix 2. Audio Track Layout for Live Delivery	39
10	Appendix 3. Dolby E.....	41
11	Appendix 4. Naming of Audio Formats when using multiple mono WAV files	42
12	Appendix 5. Parameters for satellite contribution	44
12.1	Contribution for HD (informative)	44
12.2	Contribution for SD (informative).....	45

1 General Quality Requirements

1.1 Picture Quality

The picture must be well lit and reasonably but not artificially sharp.

The picture must be free of excessive noise, grain and digital compression artefacts.

The picture must be free of excessive flare, reflections, lens dirt, markings and obstructions (e.g. lens hood), and lens aberrations.

Movement must appear reasonably smooth and continuous, and must not give rise to distortions or break-up to moving objects, or cause large changes in resolution.

The picture must be free of excessive black crushing and highlight compression. Hard clipping of highlights (e.g. by legalisers) must not cause visible artefacts on screen.

There must be no noticeable horizontal or vertical aliasing, i.e. jagged lines, field or frame rate fluctuations in fine detail.

Colour rendition, especially skin tones, must be consistent throughout, and a realistic representation of the scene portrayed unless it is altered as an editorially essential visual effect.

The picture must be stable and continuous – i.e. no jumps, movements, shifts in level or position.

There must be no visible contouring / artefacts caused by digital processing. Quantisation noise must not be apparent.

There must be no noticeable spurious signals or artefacts e.g. streaking, ringing, smear, echoes, overshoots, moiré, hum, cross-talk etc.

Note – EBU R 118 is used to assess the suitability of cameras for HD use.

1.2 Sound Quality

Sound must be recorded with appropriately placed microphones, giving minimum background noise and without peak distortion.

The audio must be free of spurious signals such as clicks, noise, hum and any analogue distortion.

The audio must be reasonably continuous and smoothly mixed and edited.

Audio, whether multichannel or stereophonic, must reflect the spatial properties of the picture (left/right and front/rear).

Audibility has to be ensured at all times, regardless of the consumer equipment used, such as a small mono television set or a home cinema. Even viewers with slightly impaired hearing should be able to hear what is said. Particular care needs to be taken when adding background music and sound effects to passages with dialogue.

Audio levels must be appropriate to the scene portrayed and dynamic range must not be excessive. They must be suitable for the whole range of domestic listening situations.

Stereo audio must be appropriately balanced and free from phase differences which cause audible cancellation in mono, and in a multichannel case, in stereo and mono downmix.

The audio must not show dynamic and/or frequency response artefacts as a result of the action of noise reduction or low bit rate coding systems.

1.3 Access for People with Disabilities

SVT programming is subject to the provisions of the Radio Act, to terms set out in the charter between SVT and the state. The Radio Act requires service providers to take positive steps to make their services accessible to people with disabilities.

Programme suppliers are therefore required to consider the needs of people with hearing or visual impairments while generating captions, subtitles and graphics, using voiceovers, and while mixing sound.

For further information, please refer to the project manager of the specific SVT programme.

2 Technical Requirements – Video

NOTE – This section is applicable to all types of deliveries. Requirements which are different for file, tape and live are covered in separate sections 4, 5 and 6.

2.1 High Definition Format

SVT is a "50p HD broadcaster" – 720p/50 for the time being, in addition simulcasting down-converted 576i/25 (SDTV).

Note: Independently of SVT's current transmission format 720p/50, SVT requires delivery of programmes in each programme's native production video format. The native production video format should preferably be chosen according to SVT's priority order below. This concept is chosen by SVT to be able to convert and rerun programmes in optimum quality when the transmission format eventually is changed to 1080p/50.

50 Hz motion portrayal material delivered for SVT transmission must be:

- 1280 x 720 pixels in an aspect ratio of 16:9
- 50 frames per second, progressive scan – now known as 720p/50
- colour sub-sampled at a ratio of 4:2:2
(Fully specified in ITU-R BT.1847)

Or (optionally, for 50 Hz motion portrayal):

- 1920 x 1080 pixels in an aspect ratio of 16:9
- 25 frames per second (50 fields) interlaced – now known as 1080i/25
- colour sub-sampled at a ratio of 4:2:2
(Fully specified in ITU-R BT.709-5 Part 2)

Or (optionally, for 50 Hz motion portrayal standard definition legacy programmes, by agreement with SVT):

- 720 x 576 pixels with the centred 702 x 576 pixels in an aspect ratio of 16:9 or 4:3
- 25 frames per second (50 fields) interlaced – now known as 576i/25
- colour sub-sampled at a ratio of 4:2:2
(Fully specified in ITU-R BT.601-7)

25 Hz motion portrayal material delivered for SVT transmission must be:

- 1920 x 1080 pixels in an aspect ratio of 16:9
- 25 frames per second, progressive scan (or progressive scan segmented frame) – now known as 1080p/25 (or 1080psf/25)
- colour sub-sampled at a ratio of 4:2:2
(Fully specified in ITU-R BT.709-5 Part 2)

2.1.1 Origination

50 Hz motion portrayal material should preferably be originated with progressive scan (or, optionally, with interlaced scan).

50 Hz and 25 Hz motion portrayal may be mixed within a 50 Hz motion portrayal programme if it is required for editorial reasons or the nature of the programme requires material from varied sources.

2.1.2 Post-production

Electronically generated moving graphics and effects (such as rollers, DVE moves, wipes, fades and dissolves) must be generated and added with 50 Hz motion portrayal to prevent unacceptable judder in 50 Hz motion portrayal programmes.

Regarding programmes delivered in 1080p/25 (or 1080psf/25), all moving graphics and effects must be generated and added with 25 Hz motion portrayal.

2.1.3 Film motion or "film effect"

It is not acceptable to shoot with 50 Hz motion portrayal and add a film motion effect in post-production. Most High Definition cameras can capture in either 720p/50 or 1080p/25. Where film motion is a requirement, progressive capture in 25 Hz motion portrayal (25p or 25psf) is the only acceptable method.

2.1.4 Field dominance

Cuts in interlaced scanned material must happen on frame boundaries (i.e. between field 2 and field 1). Motion on *psf* (Progressive scan Segmented Frame) material must always occur between field 2 and field 1 (i.e. field 1 dominance).

Note: It is preferred to shoot with progressive scan for 50 Hz motion portrayal (e.g. 720p/50), hence also possible to shoot material at 1080p/50. When this is done, the correct Two-Frame Marker phasing (SMPTE ST 2051-2010) must be maintained (to not introduce cuts between field 1 and field 2 if subsequently converting to interlace scanned video formats – e.g. 576i/25 or 1080i/25).

2.2 Video Line-Up

Programme video levels must be accurately related to their associated line-up signals. Video line-up must be colour bars of the type known as EBU 75 % (100/0/75/0), or optionally 100 % (100/0/100/0), and filling the 16:9 raster. SMPTE pattern bars are not acceptable. For required durations, see the relevant Delivery Requirements for File (section 4.4.1) or Tape (section 5.4).

2.3 Video Levels and Gamut (illegal signals)

High Definition digital signals will be assessed according to the recommendations ITU-R BT.1847 (720p) and ITU-R BT.709-5 Part 2 (1080i and 1080p/1080psf).

Video levels must be received within the specified limits so that the programme material can be used without adjustment. Any signal outside the specified limits is described as a gamut error.

2.3.1 Measuring signal levels

Digital video levels are usually measured with a device which displays a trace like a traditional waveform monitor. This gives readings in mV (emulating an analogue signal), or as a percentage of the allowable levels.

The limits of signal levels are defined by reference to a nominal black level and a nominal white level. Black level comprises R, G and B all at zero (or 0 % or 0 mV) and white level is all three components at 100 % or 700 mV.

In a picture signal, each component is allowed to range between 0 and 100 % (or 0 mV and 700 mV). This equates to digital sample levels 16 and 235 (8-bit systems) or 64 and 940 (10-bit systems).

2.3.2 Tolerance of out of gamut signals

In practice it is difficult to avoid generating signals slightly outside this range, and it is considered reasonable to allow a small tolerance, which has been defined as follows under EBU R 103:

- RGB components must be between -5 % and 105 % (-35 mV and 735 mV)
- and
- Luminance (Y) must be between -1 % and 103 % (-7 mV and 721 mV)

Slight transient overshoots and undershoots may be filtered out before measuring, and an error will only be registered where the out of gamut signals total at least **1 % of picture area**. Many monitoring devices are designed to detect errors to this specification.

2.4 "Blanking"

HD images must fill the active picture area (1280 x 720 pixels and 1920 x 1080 pixels). No "blanking errors" are permitted on new, up-converted, or archive material.

However, a two pixel tolerance will be permitted during CG or complex overlay sequences where key signals, graphic overlays or other effects do not fully cover the background image. Where animated key signals or overlays cause

moving highlights at the edge of the active image it is preferable to blank these pixels completely. A note of the timecodes and reasons for these errors should accompany the delivered programme.

2.5 Aspect Ratio

All high definition programmes (except as below) must be delivered in 16:9 Widescreen. This means that the active picture must fill 16:9 screens vertically and horizontally without geometric distortion.

2.5.1 "Cinemascope ratio" letterbox

Programmes may be delivered with an active picture in cinema ratios (e.g. 2.39:1 or 1.85:1), centred vertically between black bars in a 16:9 frame, filling the width of the frame according to SMPTE RP 199-2004, and with no geometric distortion. Film aspect ratio examples:

Film Aspect Ratio	ITU-R BT.1847 (1280x720)		ITU-R BT.709 (1920x1080)	
	h	v	h	v
1.85:1	1280	692	1920	1038
2.39:1	1280	536	1920	803

2.5.2 Floating images

Short sequences of images surrounded by black borders, (floating images), may be used for artistic effect. Note however, that widescreen consumer TV sets operating in Auto Zoom / Auto mode often interpret large black borders at the top and bottom of the screen as letterbox, so are likely to enlarge the picture. The resulting unpredictable zooming can be annoying for the viewer and undermine the artistic intent. If used, the black space around floating images must be consistent across sequences of images.

2.5.3 "Pillarboxed" HD material

Some "pillarboxed" material is acceptable where it has been acquired on a medium that has the capability to be transferred to a legitimate HD resolution, for example, 35mm film shot using 4 perf at an aspect ratio narrower than 16:9. The pictures must be centrally framed in a 16:9 raster according to SMPTE RP 199-2004 with no geometrical distortion. Film Aspect Ratio examples:

Film Aspect Ratio	ITU-R BT.1847 (1280x720)		ITU-R BT.709 (1920x1080)	
	h	v	h	v
1.33:1	960	720	1440	1080
1.66:1	1195	720	1793	1080

2.6 Archive Material

Archive material must meet all the requirements in this document, including those for up-converted SD video where relevant, except for the following:

2.6.1 General quality – archive

Archive material must be taken from the best available source, and any improvement or restoration work which could reasonably be expected must be done (for example grading, dropout repair or audio equalisation.)

2.6.2 Aspect ratio – archive

Archive material should be zoomed to fill the 16:9 raster where possible without compromising the image quality or composition, otherwise it may be presented in a pillar-box format, which:

- may be of an intermediate ratio between 4:3 and 16:9, but must be of consistent width across sequences,
- must be centrally framed in the 16:9 raster according to SMPTE RP 199-2004,
- must show no geometrical distortion,
- must have clean and sharp pillarbox edges (i.e. any video or film edge artefacts may need to be blanked),

- must be black outside the active picture.

Note however, that consumer TV sets operating in Auto Zoom / Auto mode may enlarge the picture to fill the screen horizontally. The resulting unpredictable zooming can be annoying for the viewer and undermine the artistic intent.

2.6.3 Safe areas – archive

Any captions or text already in the archive material should be kept within the caption safe area if possible, but if not, should be noted in the accompanying documents.

2.7 Use of Non-HD material

Some high definition programmes will contain some material from standard definition originals, and sources which are not considered to meet HD broadcast standards, such as domestic camcorders. This material is all called "non-HD" in this document.

To maintain a high standard and meet audience expectations the amount of non-HD material is limited to **25 %** of the programme's total duration. Non-HD material must not be used for large uninterrupted sections of the programme, unless agreed by SVT. This includes archive material.

2.7.1 Non-HD material

Material acquired using the following methods or formats is considered to be below the high definition standard and will therefore be treated as non-HD:

- HDV from all manufactures
- Most cameras with image sensors under 1/2"
- Frame based (intra-frame) recording formats below 100 Mb/s
- Inter-frame based (Long GOP) recording formats below 50 Mb/s – e.g. AVCHD \leq 28 Mb/s
- Film not meeting the requirement for HD in section 2.8 below

2.7.2 Up-converted SD video material

Particular care must be taken to deliver the best possible quality of up-converted material. In general, standard definition pictures must look no worse than the original after being up converted, post processed and down converted. Only high quality up-conversion processes will achieve this.

Standard definition video contains a half-line at top and bottom on alternate fields. This must be removed on up-conversion to HD, or it will be visible flickering at top and bottom of the HD frame.

Any VITC or switching signals visible at the top of SD material must be removed.

Any line blanking from SD signals must not appear in the HD conversion.

For these reasons it is necessary that all SD material is zoomed in by a small amount on up-conversion.

2.8 Film for High Definition Acquisition

Super16 film is *not* considered to be high definition no matter what processing or transfer systems are used.

The following 35 mm film types and stock are acceptable for high definition acquisition:

- 3 perf – any exposure index although an exposure index of 250 or less is preferred.
- 2 perf – only if daylight stock with an exposure index of 250 or less is used

To avoid causing problems with high definition transmission encoding, film should be well exposed and not forced more than one stop.

2.9 Photosensitive Epilepsy (PSE)

Flickering or intermittent lights and certain types of repetitive visual patterns can cause serious problems for viewers who are prone to photosensitive epilepsy. Children & teenagers are particularly vulnerable.

SVT is, for the time being, *not* subject to prevent photosensitive epilepsy.

2.10 Safe Areas for Captions

Captions and credits must be clear and legible and must be within the safe area specified.

All font sizes must be legible as HD and also after downconversion for the SD viewer.

There is one primary caption safe area defined for 16:9 materials at SVT:

- **16:9 graphics safe area for 16:9 presentation (according to EBU R 95-2008).**

Caption Safe Area	Defined as (%)	HD pixels (inclusive). First pixel numbered 1	TV line numbers (inclusive). Numbering as per "Rec1847" and "Rec709"
16:9 Caption safe	80 % of Active Width	129-1152 (of 1-1280) 193-1728 (of 1-1920)	
	90 % of Active Height	37-684 (of 1-720) 55-1026 (of 1-1080)	62-709 (750p) 48-533 (F1) and 611-1096 (F2), (1125i/psf) 96-1067, (1125p)

At the discretion of SVT, programmes such as feature films and some acquisitions may be excluded from this requirement.

2.10.1 14:9 Caption Safe Area

A 14:9 based caption safe area, in combination with a 14:9 based action safe area, is used in the United Kingdom (e.g. by the BBC), but in general *not* used by SVT.

2.10.2 4:3 Caption Safe Area

A 4:3 based caption safe area, in combination with a 4:3 based action safe area, is often used in the United States, but in general *not* used by SVT.

2.10.3 16:9 Caption Safe Area

A 16:9 based caption safe area, in combination with a 16:9 based action safe area, is used by SVT. See the table above and, if one so wish, graphical presentations of "16:9 safe areas for 16:9 presentation" for various scanning raster in EBU R 95-2008, downloadable from <http://tech.ebu.ch/publications/r095>

2.11 Standards Conversion

When standards converted material is included in a programme, e.g. from 59.94 Hz motion portrayal to 50 Hz motion portrayal, Motion Compensation (sometimes known as Motion Predictive or Motion Vector) standards conversion is required.

Currently speed change is the preferred method of changing between 24 fps (including 23.98) and 25 fps standards. Due attention must be given to the audio – i.e. resampling.

Use of non-linear editing platform hardware or software standards conversion is *not* permitted for whole programmes but may be used for short inserts at the discretion of SVT.

3 Technical Requirements – Audio

NOTE – This section is applicable to all types of deliveries. Requirements which are different for file, tape and live are covered in separate sections 4, 5 and 6.

For programmes carrying multichannel audio (>2.0) it is optional to deliver an additional stereo (2.0) mix. SVT transmits a stereo audio stream in conjunction with a multichannel audio stream, but it is made from an in-house mix-down of the multichannel audio, with Audio Metadata applied – i.e. stereo listeners will receive either a mix-down from the surround channels generated in the playout chain or a mix-down generated in their receiver.

For track layout and allocations, see the relevant Delivery Requirements for File (section 4.4.1), Tape (section 5.4.) or Live (section 6.7 and Appendix 2).

For Audio Metadata, see section 3.2.4.

3.1 Stereo Audio Requirements

Stereo tracks, if supplied, must carry sound in the A/B (Left/Right) form.

If mono originated sound is used, it must be recorded as dual mono, so that it may be handled exactly as stereo. It must meet all the stereo standards regarding levels, balance and phase.

3.1.1 Stereo line-up tones

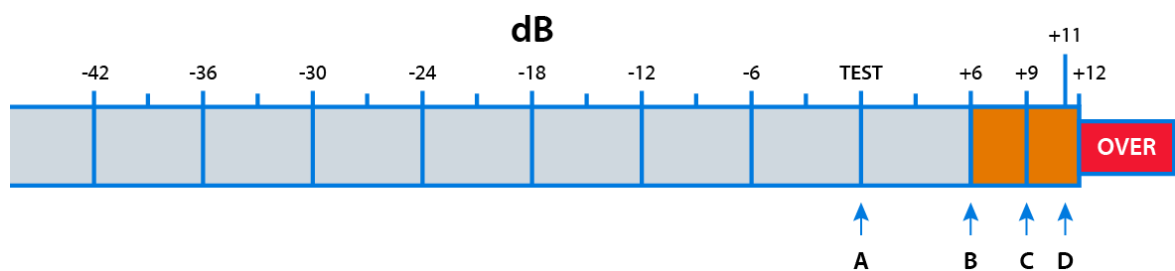
All stereo tracks must use EBU 1 kHz tone (left identification). All tones must be sinusoidal, free of distortion and phase coherent between channels.

Digital Audio Reference level is defined as 18dB below the maximum coding value (-18 dBFS) as per EBU recommended practice R68.

3.1.2 Stereo audio levels and measurement (loudness or volume)

Using the Nordic scale (EBU Tech 3205), in order to achieve uniform programme levels during an entire broadcast, *normal* speech should be levelled between 0 dB and +6 dB.

Short transients may not exceed the reference level “TEST” (0 dBu, corresponding to -18 dBFS) by more than +11 dB, as measured with a peak programme meter with an integration time of 10 ms \pm 2 ms, using the Nordic scale.



A = 0 dB	Reference level (Test tone level)
B = +6 dB	Maximum permitted level for <i>normal</i> speech
C = +9 dB	Maximum permitted level of loud programme parts
D = +11 dB	Maximum permitted level of short transients

Shortly, SVT will adapt to loudness metering according to EBU R 128. In the transition period, both the values of the Nordic scale and the EBU R 128 scale, with the following values, must be met:

- Programme Loudness = -23 LUFS \pm 1 LU,
- Maximum permitted True Peak Level \leq -1 dBTP
- Loudness Range (LRA): *State the measured LRA in the Audio Metadata – section 3.2.4*

For short programmes and trailers (~30 seconds), in addition to the requirements above, both of the following requirements apply (EBU Tech 3343, section 10.1):

- Maximum permitted Short-Term loudness (3 seconds time window) must not exceed -20 LUFS (+3 LU) at any time.
- Maximum permitted Momentary loudness (400 milliseconds time window) must not exceed -15 LUFS (+8 LU) at any time.

Note: All values above must be stated in the Audio Metadata – section 3.2.4.

Note: Programmes that are deliberately mixed with programme loudness out of range will be checked and, if possible, corrected by SVT's internal quality control procedures – otherwise rejected.

3.1.3 Stereo phase

Stereo programme audio must be capable of mixing down to mono without causing any noticeable phase cancellation.

When the right and left channels are mixed down to mono, there shall not be any marked difference in listening level or frequency reproduction. Full audibility is required of the mono downmix. SVT does *not*, during broadcast, guarantee compatibility to matrix-encoded multichannel formats such as Dolby Surround, Dolby Pro Logic or Dolby Pro Logic II.

3.2 Surround Sound Requirements

Surround sound is transmitted as up to 5.1 format, and should be delivered as discrete tracks (i.e. *not* Dolby E, except by agreement with SVT).

Dialogue sound, voice-overs, solo singers and such (voices indispensable to the content of the programme) are placed in the centre channel, enabling viewers with varying degrees of hearing impairment to balance dialogue and sound effects/music in such a way as to maximise audibility.

The control room where the multichannel mix is produced should meet all the requirements of ITU-R BS.775 and Dolby Laboratories' 5.1 Channel Music Production Guidelines. An example, the LFE listening volume in the subwoofer is to be set 10 dB higher than the listening volume of the main channels.

Particular care should also be taken with low-frequency information mixed to the LFE channel compared to the main channels. The LFE channel is *not* part of the downmix to stereo/mono.

SVT allows for one of the following formats to be chosen *during production*, according to the nature of the programme:

- 1.0 – One audio channel (mono) placed in the centre channel. To be used for programmes containing speech only, with no sound effects or music.
- 2.0 – Two audio channels (stereo) placed in the left front and right front channels. May be used for programmes containing only instrumental music/sound effects or for repeats of stereophonic material from archives.
- 3.0 – Three audio channels placed in the left front, centre and right front channels. To be used for programmes containing speech and music / sound effects, e.g. for News. Speech is placed in the centre channel and music/sound effects in the left front and right front channels. 3.0 is comparable to ordinary stereo sound, but with the added benefit of allowing the viewer to strike a balance between dialogue and music/sound effects.
- 3.1 – Similar to 3.0, but enhancing low-frequency sound effects by placing them in the LFE channel, as well. (The LFE channel is not part of down mixing to stereo/mono.)
- 4.0 – Four audio channels placed in the left front and right front channels and in the left surround and right surround channels. May be used for programmes containing only instrumental music/sound effects.
- 4.1 – Similar to 4.0, but enhancing low-frequency sound effects by placing them in the LFE channel, as well. (The LFE channel is not part of down mixing to stereo/mono.)
- 5.0 – Five audio channels placed in the left front, centre and right front channels, and in the left and right surround channels. To be used for programmes where mood and presence are important, and where no enhancement of low-frequency sound is desirable, e.g. musical entertainment and sports.
- 5.1 – Similar to 5.0, but enhancing low-frequency sound effects by placing them in the LFE channel, as well. To be used for documentaries and drama. (The LFE channel is *not* part of down mixing to stereo/mono.)

Audio Format	Left Front	Right Front	Centre Front	LFE	Left Surround	Right Surround
1.0 (1/0/0) mono			X			
2.0 (2/0/0) stereo	X	X				
3.0 (3/0/0)	X	X	X			
3.1 (3/0/1)	X	X	X	X		
4.0 (2/2/0)	X	X			X	X
4.1 (2/2/1)	X	X		X	X	X
5.0 (3/2/0)	X	X	X		X	X
5.1 (3/2/1)	X	X	X	X	X	X

Note: Audio Metadata describing how to mix down the sound to stereo/mono has to be created during production – see section 3.2.4.

3.2.1 Surround line-up tones

All surround tracks must carry BLITS tone, as described in EBU Technical Paper 3304. An audio file of BLITS tone may be downloaded from SVT's website – www.svtb2b.se

3.2.2 Surround audio levels and measurement (loudness or volume)

The same audio level requirements as for Stereo audio (section 3.1.2) apply for Surround sound.

EBU R 128 Loudness measurement must be performed with the same downmix coefficients applied as the ones supplied with the programme delivered to SVT.

SVT reserves the right to check and change the values of supplied Audio Metadata.

3.2.3 Surround phase

Surround audio must be capable of mixdown to stereo and mono without causing any noticeable phase cancellation – i.e. sample accuracy.

For programmes produced with multichannel audio in the 1.0, 3.0, 3.1, 4.0, 4.1, 5.0 and 5.1 formats, the stereo mix (2.0) is automatically created in the Playout Centre at the time of broadcast, using the downmix metadata supplied with the programme – see section 3.2.4.

The resulting stereo mix, in its turn, has to be fully mono compatible – see section 3.1.3.

3.2.4 Audio metadata

Multichannel audio must be delivered together with Audio Metadata.

In the case of file delivery: The spreadsheet (MS Excel format), downloadable from www.svtb2b.se, including both specific audio metadata as well as EBU R 128 metadata, must be filled out and supplied.

In the case of tape delivery: The same spreadsheet as above must be filled out, printed and supplied with the tape.

In the case of live delivery: Use VANC mapping according to SMPTE ST 2020 and Dolby Metadata for Dolby E (see Appendix 3), or at least supply the metadata values in a readable form – e.g. in the spreadsheet mentioned above.

3.3 Sound to Vision Synchronisation

The relative timing of sound to vision should not exhibit any perceptible error. Sound must not lead or lag the vision by more than 5 ms.

3.3.1 Audio / Video sync markers

To assist in maintaining A/V sync through the post-production process, a "sync plop" may be used. If the delivered programme leader contains one, it must meet the following conditions:

- The sync plop must be between timecode 09:59:57:06 and 09:59:57:08 (25 TC).
- The audio plop must be 1 kHz tone on all tracks at -18 dBFS (standard zero level).
- The duration of the vision flash must be 4 frames in 50p video (or 2 frames in 25i video) to allow it to pass through standards conversion successfully.
- The audio plop must be synchronous across all audio PCM audio tracks and with the video flash (within +/- 5 ms).
- If an end sync plop is used it must be no closer than 10 seconds to the end of the programme and comply with the points above.

3.4 Dynamic Range

The dynamic range of sound programme material should be suited to television broadcasts and work for a small mono television set as well as a home cinema.

Listening evaluations guarantee full audibility. Evaluation is performed at a low listening level (57 dBC SPL/channel), using loudspeakers with a linear frequency response.

The corresponding Loudness Range (LRA) according to EBU R 128 should be stated as Audio Metadata – see section 3.2.4.

4 File Delivery Requirements

SVT's File Delivery Requirements will, in the future, likely be based on the international standard AS-11, for delivery of finished programming as HD (and SD) files – www.amwa.tv/projects/AS-11.shtml

In SVT's opinion, a new version of AMWA AS-11 has to be created (to include 1080p/25 plus Audio Metadata) and published with SVT Shims (subsets) – there will be further announcements, as soon as possible.

Meanwhile, one will below find the File Delivery Requirements of today.

This document covers the requirements for transmission-ready files. There may be additional requirements for programmes intended for further editing, re-versioning or archiving.

There must be only one programme in each file, although a programme may be either soft or hard-parted within that file according to the diagrams below. Only when agreed in advance with SVT, programmes in several parts may be delivered in more than one file, as shown in the third diagram below.

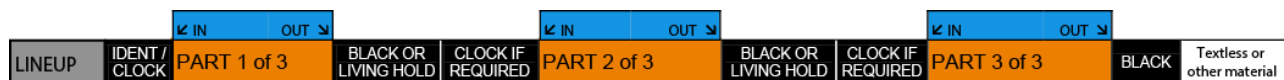
Single part or soft parted programme:

A single part programme will always be played out from start point to end point without interruption. Soft parting is where a programme is provided as a single continuous programme, but the broadcaster may break the transmission of the programme at several points to insert commercials or for other reasons. IN and OUT points for continuous playback only must be included with the delivery metadata; suggested timecodes for breaks should *not* be included.



Hard-parted programme:

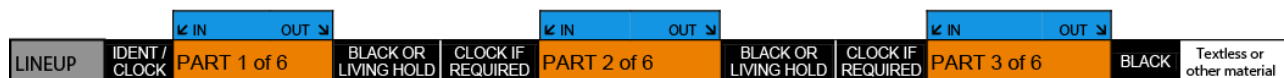
A hard-parted programme is billed and scheduled for transmission as a single entity, but is delivered as a single file containing clearly separated parts between which adverts, trailers etc will be inserted.



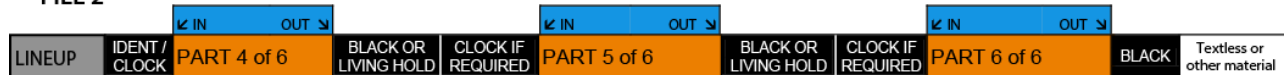
Multi-part programme delivered on multiple files:

Where a programme's delivery must be split over more than one file, e.g. due to editing up to transmission.

FILE 1



FILE 2



4.1 File format

The file format, i.e. the container/wrapper, shall preferably be MXF OP-1a (the layout options for a minimal simple MXF file according to SMPTE 378M, extension “.mxf”). Optionally a Quicktime Movie (extension “.mov”) may be used.

4.2 Video codec

The video essence in the file must, for 50 Hz motion portrayal, be encoded as:

- 720p50, preferably as (Avid) DNxHD 175x (a.k.a. 185x), optionally as (Panasonic) AVC-Intra 100, optionally as (Apple) ProRes 422 (HQ), all 10 bit video codecs, optionally as (Sony) XDCAM MPEG HD422, 8 bit.

- (optionally) 1080i25, preferably as (Avid) DNxHD 185x, optionally as (Panasonic) AVC-Intra 100, optionally as (Apple) ProRes 422 (HQ), all 10 bit video codecs, optionally as (Sony) XDCAM MPEG HD422, 8 bit.
- (optionally and by agreement with SVT) 576i25, preferably as DVCPRO50, optionally as DVCPRO25, both according to SMPTE 314M – see section 4.10 below.

The video essence in the file must, for 25 Hz motion portrayal, be encoded as:

- 1080p25, preferably as (Avid) DNxHD 185x, optionally as (Panasonic) AVC-Intra 100, optionally as (Apple) ProRes 422 (HQ), all 10 bit video codecs, optionally as (Sony) XDCAM MPEG HD422, 8 bit.

4.3 Image format

See section 2, “Technical Requirements – Video”.

4.4 Audio

All audio tracks must be encoded as PCM with a sample rate of 48 kHz at a depth of 24 bits/sample. The audio should preferably be bundled with the video inside the container/wrapper, optionally delivered separately as ”B-WAV” or ”WAV” files, see section 4.9 below.

4.4.1 Track allocations

The files should contain a group of **either 4 or 16 tracks**, with track allocations as on the table below. Unused tracks must carry “48 kHz, 24 bit, PCM digital silence”. The EBU R 48 or R 123 code must be included in the spreadsheet Programme Metadata File, www.svtb2b.se, to identify the track allocations.

EBU Ref code	Programme Type	Audio track numbers															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
R48: 2a	Stereo	St. L	St. R	-	-												
R123: 4b	Stereo with M&E	St. L	St. R	St. M&E L	St. M&E R												
R123: 4c	Stereo with Audio Description	St. L	St. R	St. Aud Desc L	St. Aud Desc R												
R123: 16b	5.1, stereo and M&E	5.1 L	5.1 R	5.1 C	5.1 LFE	5.1 Ls	5.1 Rs	St. L	St. R	5.1 M&E L	5.1 M&E R	5.1 M&E C	5.1 M&E LFE	5.1 M&E Ls	5.1 M&E Rs	St. L	St. R
R123: 16c Opt 2	Stereo, 5.1 and Audio Description	St. L	St. R	St. Aud Desc L	St. Aud Desc R	5.1 L	5.1 R	5.1 C	5.1 LFE	5.1 Ls	5.1 Rs	5.1 M&E L	5.1 M&E R	5.1 M&E C	5.1 M&E LFE	5.1 M&E Ls	5.1 M&E Rs
R123: 16d	5.1 Two languages	5.1 Lang 1 L	5.1 Lang 1 R	5.1 Lang 1 C	5.1 Lang 1 LFE	5.1 Lang 1 Ls	5.1 Lang 1 Rs	-	-	5.1 Lang 2 L	5.1 Lang 2 R	5.1 Lang 2 C	5.1 Lang 2 LFE	5.1 Lang 2 Ls	5.1 Lang 2 Rs	-	-
R123: 16f	Three Languages	St. Lang 1 L	St. Lang 1 R	-	-	St. Lang 2 L	St. Lang 2 R	-	-	St. Lang 3 L	St. Lang 3 R	-	-	-	-	-	-

Note:

- R48:2a, R48:4b, R123:4b, R123:4c, R123:16b, R123:16c must only be used for programmes with single language soundtracks
 - R123:16b is the preferred track allocation for delivery specifically to SVT.
 - Regarding R123:16c – Option 1 is the original 16c allocation. The Option 2 allocation is indirectly indicated when the metadata “Audio Description Present” is true i.e. “yes” – see section [4.11.5](#).
- R123:16d must only be used for programmes with dual language soundtracks.
- R123:16f must only be used for programmes with 3 different language soundtracks.

Any unused audio tracks in the 16 track groups above must contain digital silence and encoded as PCM audio.

For compatibility with stereo systems, any audio generated as mono must be presented as “Dual Mono” on two phase-coherent tracks, and flagged as stereo.

Any additional audio tracks required by SVT must be delivered separately as “B-WAV” or (optionally) “WAV” files, see section 4.9 below.

The naming conventions used in all related documentation and metadata (see section 4.11 below) must match those specified above.

In the case of separate 1-track (mono) “.wav” files, the file names must match those specified above in the cells to indicate the content of the file, e.g. `PG-1234567-001A-PROGRAMME_TITLE-321_ME_LS.wav` (for the left surround channel in a 5.1 (3/2/1) multichannel music and effects mix). See Appendix 4 for naming of all audio format possibilities.

4.5 Programme Layout / Format

All programmes delivered on file must be laid out with elements in the following pattern relative to “25 TC” timecode:

Timecode (25 TC)	Duration (25 TC)	Picture	Sound
09.58.00.00	90"	EBU Bars (100/0/75/0 or 100/0/100/0)	Line-up tone
09.59.30.00	between 27" 00fr and 27" 05fr	Ident Clock or Slate	Silence
09.59.57.06 (optional)	2fr	4 Frames peak white (50p) 2 Frames peak white (25i)	2 Frames tone (50p) 1 Frame tone (25i) (on first video white frame)
09.59.57.06	2" 19fr	Black	Silence
10.00.00.00		Programme	Programme
end of part (multipart programmes)	5"	freeze or "living hold" after end of part	fade or cut to silence by end of part
end of prog	10"	freeze or "living hold"	fade or cut to silence
end of prog + 10" (optional)	2fr	4 Frames peak white (50p) 2 Frames peak white (25i)	2 Frames tone (50p) 1 Frame tone (25i) (on first video white frame)

4.5.1 Start and end

Note that it is usual for sound and vision to be automatically cut to air on transmission, so early vision or sound is not normally required. Vision may fade up from black starting at 10.00.00.00 if desired.

All programmes must end with a fade or cut to silence *before* the intended end point. Any fade out or reverb must be allowed for within the programme duration.

Vision freeze or "living hold" must be held for a further ten seconds (10") after the end point.

Any other programme elements after the end of the programme should not start less than one minute (1') after end of programme.

4.5.2 The Ident Clock or Slate

A countdown clock or slate clearly displaying the following information must precede the start of programme and any subsequent part:

- Programme ID number (the content provider's internal programme number or identifier).
- Programme title (and series number if applicable).
- Episode number (if applicable).
- Episode subtitle (if applicable).
- Version (Pre/post watershed etc. if necessary).
- Part number (if applicable).

No technical information may be included. The clock or slate may display telephone contact numbers for the post-production facility and production company, and may display company branding.

Where a moving clock is used, it must provide a clear countdown of at least 20 seconds, including a hand moving in 1 sec steps (i.e. *not* smooth motion) around a circular clock face. Clocks with only digital countdown are not acceptable.

There must be no audio tone or identification over the clock.

4.6 3D Delivery

SVT does *not* accept stereoscopic 3D programmes for the time being.

4.7 Closed captions (Subtitles)

SVT is interested in any manuscripts used during the production of the programme, or transcripts of any dialogue, for use by SVT's in-house subtitling department when creating subtitles.

Naming of associated material must adhere to the following convention:

PG-<identifier_provided_by_SVT_including_programme_name>-<suffix_describing_content_of_file>.<extension>

Example:

[PG-939393-001A-SASKAETTLATA-Manuscript1.doc](#)

[PG-939393-001A-SASKAETTLATA-Manuscript2.doc](#)

[PG-939393-001A-SASKAETTLATA-Transcript.doc](#)

[PG-939393-001A-SASKAETTLATA-Subtitles.xif](#)

The text describing the content of the file should only contain A-Z, 0-9, i.e. *no* umlaut characters (e.g. Å, Ä and Ö) are allowed.

4.8 Timecode

For content with 50 Hz motion portrayal (progressive and interlace), as well as for content with 25 Hz motion portrayal, a 25 counting timecode (25 TC) must be used.

4.9 Audio only files

Additional audio only files related to a programme, such as Audio Description files, must be supplied as BWF (sometimes called "B-WAV") files, conforming to the specification in EBU-Tech 3285, or (optionally) several 1-track (mono) WAV-files. File duration and timecode must match the principal video file exactly.

4.10 SD Files (Legacy programmes only)

Delivery of standard definition legacy programme files must be by agreement with SVT. Those files must meet the following requirements:

4.10.1 File format

The file format, i.e. the container/wrapper, shall preferably be MXF OP-1a (the layout options for a minimal simple MXF file according to SMPTE 378M, extension “.mxf”). Optionally a Quicktime Movie (extension “.mov”) may be used.

4.10.2 Video codec

The video essence in the file must, for 50 Hz motion portrayal, be encoded as:

- 576i25, preferably as DVCPRO50, optionally as DVCPRO25, both according to SMPTE 314M.

4.10.3 Image format

SD video files must be recorded with a picture area of 702 x 576 pixels, where the 702 pixel wide picture must be centred in the active 720 pixel wide line. The picture information may extend the full width of the 720 pixel wide line, providing the image shape is not distorted.

4.10.4 Audio essence

All audio tracks should be encoded as PCM with a sample rate of 48 kHz at a depth of 24 bits/sample, hence delivered *separately* as ”B-WAV” or ”WAV” files, see section 4.9 above. (If interleaved audio is used, according to SMPTE 314M, only 16 bits/sample is obtained).

4.11 Metadata

Metadata is the name for all the information which is not the audio or video essence, but which is required to ensure that contents of the file can be identified correctly, and can be played back or converted in various systems.

The metadata required is specified below and must be delivered as separate Excel-file (the Programme Metadata File), based on a template provided by SVT (downloadable from www.svtb2b.se).

4.11.1 Filenames

For each delivery of one or more files containing video essence for a specific programme, at least one Programme Metadata File containing metadata must also be included. If the metadata differs between episodes, several Programme Metadata Files can be provided. The naming of the Programme Metadata File must adhere to the following convention:

If all episodes are included in the delivery, the programme-id (seven numbers, 0-9) identifies the programme (series), followed by the episode part indicating that this Programme Metadata File provides information about all episodes (NNN).

`PG-3993939-NNNA-SASKADETLATA.xls`

If not all episodes are included in the delivery, the specific episodes must be included in the name, comma-separated:

`PG-3993939-001A,002A,003A-SASKADETLATA.xls`

`PG-3993939-004A-SASKADETLATA.xls`

4.11.2 SVT Metadata application

The template Excel-file containing the required fields can be downloaded from www.svtb2b.se. This file can be edited with any Microsoft Excel-compatible spreadsheet application.

Save the resultant file according to the file convention defined above.

4.11.3 Delivery Requirements in MXF

Not applicable for the time being.

4.11.4 Delivery requirements in XML

Not applicable for the time being.

4.11.5 Required Metadata

Editorial metadata:

Element name	Definition and usage	Typical values
<i>Series Title</i>	The final title of a grouping of publishable assets with sharing identification and branding linked by common characters, subject matter, style or story. This could be a series, serial or themed grouping. 'One-Off' programmes with a single title should give it here as well as in 'Programme Title' below.	Så ska det låta, season 24
<i>Programme Title</i>	The title of a Programme Version for a specific purpose. Note: Programme Title may change between the point of commission/production and final delivery from post production. Note: Must indicate the version where applicable – e.g. pre/post watershed. Note: 'One-Off' programmes with a single title should give it here as well as in 'Series Title' above.	Så ska det låta
<i>Episode Title / Episode No</i>	Final episode name and/or number used to identify an individual episode within a Series. Not used for version information.	Episode 3
<i>Production Number</i>	Provided by SVT! A unique number used to identify an individual Programme Episode and/or Version. Also known as Programme ID or “Vision-ID” internally at SVT.	3393949-001A
<i>Originator</i>	Company responsible for creating the programme.	Production company AB
<i>Copyright Year</i>	Year in which the production was completed.	2013
<i>Distributor</i>	The name of the person or company/companies providing the content. May be a third party for secondary distribution rights.	Providing company
<i>SVT Commissioning Unit</i>	Unit within SVT that commissioned the programme.	SVT ATV, Göteborg

Video metadata:

Element name	Definition and usage	Typical values
<i>Video codec</i>	The video codec employed for the creation of the file.	DNxHD 175x (a.k.a. 185x) (720p/50, 50 Hz motion portrayal)
<i>Picture format</i>	The signal standard (frame resolution and aspect ratio) of the encoded file.	1280x720p/50, 16x9 (720p/50, 50 Hz motion portrayal)
<i>Video comments</i>	The comments which illustrate the subjective quality and any known artefacts or defects (inc. intentional) within the video content discovered during production / post production / or any subsequent technical QC/Review process.	

Audio metadata:

Element name	Definition and usage	Typical values
<i>Audio Track Layout</i>	Code indicating the audio track layout in accordance with EBU R 123 (HD) and R 48 (SD). Assumption is to always have 16 tracks (4 for SD). Unused tracks must contain digital silence. Note: To include valid digital silence is required. Note: This selection directly affects Secondary Language, and Tertiary Language.	R123:16c
<i>Primary Audio Language</i>	Primary audio language used in the Programme Version..	Swedish
<i>Secondary Audio Language</i>	Secondary audio language used in the Programme Version.	Finnish
<i>Tertiary Audio Language</i>	Tertiary audio language used in the Programme Version.	Danish
<i>Audio Comments</i>	The comments which illustrate the subjective quality and any known artefacts or defects (inc. intentional) within the audio content discovered during production / post production / or any subsequent technical QC/Review process.	
<i>Channel Mode</i>	Defines which audio channels are used in the mix	3/2/1
<i>LFE Channel Enable</i>	Defines if Low Frequency Effects channel is used or not	Enabled
<i>Preferred Stereo Downmix</i>	This parameter allows the producer to select either the Lt/Rt or the Lo/Ro downmix in a consumer decoder that has stereo outputs.	Lo/Ro
<i>Centre Downmix Level</i>	The change of level for the centre channel when downmixing to two channel audio.	-3 dB
<i>Surround Downmix Level</i>	The change of level for the surround channels when downmixing to two channel audio.	-3 dB
<i>Loudness Measurement Standard</i>	Defines which loudness measurement standard is used, i.e. EBU R 128, ITU-R BS.1770-3, etc.	EBU R 128
<i>Programme Loudness</i>	The measured integrated programme loudness	-23 LUFS
<i>Max True Peak</i>	The measured max true peak level of audio	-1 dBTP
<i>Loudness Range</i>	The measured Loudness Range of audio	10 LU
<i>Max Short Term loudness (3 s time window)</i>	The measured Maximum Short-Term loudness (3 seconds time window). Relevant only for short programmes and trailers (~30 seconds).	-20 LUFS
<i>Max Momentary Loudness (400 ms time window)</i>	The measured Maximum Momentary loudness (400 milliseconds time window) Relevant only for short programmes and trailers (~30 seconds).	-15 LUFS

Timecode metadata:

Field name	Definition and usage	Typical values
<i>Start of repeating group</i>		
<i>Part Number</i>	Identifier for each hard part no. Note: This should indicate the part number <i>within the</i>	1

	<i>entire programme</i> even if the programme is being delivered as more than one file – e.g. one file might contain parts 3-6 only. (Not required for soft parted material).	
<i>Part Total</i>	Total Programme Parts. Identifier for the total number of parts <i>in the entire programme</i> . (Not required for soft parted material).	2
<i>Part SOM</i>	Part Start of Media (SOM). Timecode for the first frame of the part.	10:00:00:00
<i>Part Duration</i>	Duration of the part as a timecode.	00:30:00:00
<i>End of repeating group</i>		
<i>Total number of parts</i>	This denotes the total no. of hard parts contained <i>within the file</i> . Hard parting is where a programme is delivered as parts separated by black (and or idents) within a file. Note: This is <i>not</i> to be confused with Part Total (Total Programme Parts).	1
<i>Total programme duration</i>	Total duration of programme. For hard parted programmes, this is the total of all part durations.	01:00:00:00

Access services metadata:

Element name	Definition and usage	Typical values
<i>Audio Description Present</i>	This status is to be set if the delivered programme contains any audio description for the visually impaired.	No
<i>Audio Description Type</i>	This describes the format of the audio description employed within the delivered programme. There is two options only – “Control data / Narration” and “AD Mix”.	
<i>Open Captions Present</i>	This status is to be set if the delivered programme contains any visible (in vision) subtitles. Note: This does not include naming or place captions etc.	Yes
<i>Open Captions Type</i>	This is the editorial description of the in vision subtitling employed within the delivered programme. E.g. for the hard of hearing, translation, unclear speech etc.	Translation
<i>Open Captions Language</i>	This describes the primary language of the in vision subtitles.	Swedish
<i>Signing Present</i>	This status is to be set if the delivered programme contains <i>any</i> in vision signing for the hard of hearing.	No

Additional metadata:

Element name	Definition and usage	Typical values
<i>Textless Elements Exist</i>	Indicates whether the delivered file includes any textless elements after the end of the programme.	Yes
<i>Programme Has Text</i>	Indicates whether the main programme has any text or is completely 'clean'.	Yes

<i>Programme Text Language</i>	Primary text language used in the programme.	Swedish
--------------------------------	--	---------

Contact Information metadata:

Element name	Definition and usage	Typical values
<i>Contact Email</i>	The email address of the SPOC (Single Point Of Contact) for the use of the recipient regarding any delivery or technical issues encountered with the delivered file.	contact@yourcompany.com
<i>Contact Telephone No.</i>	The direct telephone number of the SPOC (Single Point Of Contact) for the use of the recipient regarding any delivery or technical issues encountered with the delivered file.	+46 8 784 00 00

4.12 File Delivery Options

The following provides guidance on the delivery options to SVT for File based programmes. As this is intended to offer guidance only, please contact the Post Production Supervisor named in your specific agreement with SVT to discuss further details and options.

4.12.1 Hard Disk

Hard disks must have USB-interface. Both NTFS and Mac OS formatting are acceptable. The physical hard disk case must be clearly marked for easy identification.

Hard disks will normally require some form of content protection, using either hardware or software encryption. In some cases the removable hard disks may be supplied by SVT, but normally need to be provided by the production company.

Note: If using option 1 above, the same delivery address and details should be used as with existing tape deliveries and follow currently processes for delivery confirmation.

4.12.2 LTO Tape

SVT can not handle LTO 5 Data Tape for the time being.

4.12.3 File transfer over Leased line / circuit

For the time being, SVT is only able to open up its firewall and establish UDP connections to *receive* files (not deliver files) via private and secure dedicated data circuits – i.e. standard FTP and 3rd party UDP accelerators such as Aspera, Signiant or File Catalyst. Hence must such a file transfer become initiated by the, to SVT, delivering part. 3rd party accelerators will normally provide delivery reporting.

4.12.4 File transfer over a secure managed network service

SVT has no secure managed network services established for the time being . Projects to resolve this are ongoing, this section will be updated as soon as possible.

4.12.5 File Transfer using sFTP

For the time being, SVT is not using Secure FTP (sFTP). Projects to resolve this are ongoing, this section will be updated as soon as possible.

Note: If using options 3, 4 or 5 above, the method of receipt of media will need to be agreed between SVT and production company.

5 Tape Delivery Requirements

Note that programmes delivered on tape must comply with all the requirements of this document other than those for file or live delivery.

5.1 Videotape recording

5.1.1 Tape format

HDCAM SR is the only format acceptable for HD tape delivery. The recording must be fully compliant with the manufacturer's technical specification thereby ensuring format compatibility.

Tapes must be clean, new stock, in the manufacturer's case, protected by suitable packaging and clearly labelled. Note that flock filled padded envelopes are not suitable since a failure in the packaging can lead to contamination of the tape. All tapes must be supplied with the record lockout "on" and fully rewound. It is recommended to "double rewind" before shipping to ensure an even tape pack. Labels must be fixed to both the cassette case and cassette and must not obscure the spools or obstruct the flap mechanism.

5.1.2 "i" and "psf" Flags

Programmes containing *any* 50 Hz motion portrayal element (e.g. 50 Hz motion graphics), and delivered with 1080i/25 (i.e. not with the preferable 720p/50), must be delivered with flags set to "i" throughout the programme, even if the bulk of the programme has been originated with 25 Hz motion portrayal.

Programmes that for sure contain 25 Hz motion portrayal entirely (including motion graphics, effects etc), and delivered with 1080psf/25, must be delivered with flags set to "psf" throughout the programme.

5.1.3 Time-code

LTC and ancillary timecode (referred to as VITC on HDCAM SR VTRs) must be identical, contiguous and continuous throughout the recording.

It is recommended that assemble edits should not be used between the start of the clock and the end of the programme, as they may introduce LTC discontinuities.

5.2 Programme Layout / Format

All programmes delivered on tape must be laid out with elements in the following pattern relative to "25 TC" timecode:

Timecode (25 TC)	Duration (25 TC)	Picture	Sound
09.58.00.00	90"	EBU Bars (100/0/75/0 or 100/0/100/0)	Line-up tone
09.59.30.00	between 27" 00fr and 27" 05fr	Ident Clock or Slate	Silence
09.59.57.06 (optional)	2fr	4 Frames peak white (50p) 2 Frames peak white (25i)	2 Frames tone (50p) 1 Frame tone (25i) (on first video white frame)
09.59.57.06	2" 19fr	Black	Silence
10.00.00.00		Programme	Programme
end of part (multipart programmes)	5"	freeze or "living hold" after end of part	fade or cut to silence by end of part
end of prog	10"	freeze or "living hold"	fade or cut to silence
end of prog + 10" (optional)	2fr	4 Frames peak white (50p) 2 Frames peak white (25i)	2 Frames tone (50p) 1 Frame tone (25i) (on first video white frame)

5.2.1 Start and end

Note that it is usual for sound and vision to be automatically cut to air on transmission, so early vision or sound is not normally required. Vision may fade up from black starting at 10.00.00.00 if desired.

All programmes must end with a fade or cut to silence *before* the intended end point. Any fade out or reverb must be allowed for within the programme duration.

Vision freeze or "living hold" must be held for a further ten seconds (10") after the end point.

Any other programme elements after the end of the programme should not start less than one minute (1') after end of programme.

5.2.2 Programmes longer than a single tape

If a programme must be delivered on two or more tapes because it is longer than the capacity of a single HDCAM SR tape, the second part must begin at the next whole hour timecode after the end of the first part - e.g. 12:00:00:00 or 13:00:00:00 with appropriate continuous timecode throughout the line-up and clock sequence above.

5.2.3 Compilation tapes

Where SVT has agreed to accept short programmes on a compilation tape, there must be at least 15" of black and silence between the end of one programme and the start of the clock for the following programme – i.e. after the 10" hold.

Each programme must be recorded to begin at a "full minute" – i.e. timecode HH:MM:00:00.

5.2.4 Ad breaks

For hard-parted programmes, each part must be preceded by a countdown clock as below.

There must be at least 15" of black and silence between the end of one part and the start of the clock for the following part – i.e. after the 10" freeze.

Each part must be recorded to begin at a "full minute" – i.e. timecode HH:MM:00:00.

5.2.5 The Ident Clock

A countdown clock clearly displaying the following information must precede the start of programme and any subsequent part:

- Programme ID number (the content provider's internal programme number or identifier).
- Programme title (and series number if applicable).
- Episode number (if applicable).
- Episode subtitle (if applicable).
- Version (Pre/post watershed etc. if necessary).
- Part number (if applicable).

No technical information may be included. This means HD format, tape format, aspect ratio, audio track allocations, safe area etc. Duration should not be included. The clock may display telephone contact numbers for the post-production facility and production company, and may display company branding.

The clock must provide a clear countdown of at least 20 seconds, including a hand moving in 1 sec steps (i.e. *not* smooth motion) around a circular clock face. Clocks with only digital countdown are not acceptable.

There must be no audio tone or identification over the clock.

5.3 Paperwork

Each tape must have the following information on its box and cassette labels and on a VTRR (VideoTape Recording Report) included in its box. The VTRR can preferably be a print out of the spreadsheet Programme Metadata File www.svtb2b.se, containing (but not limited to):

- Programme ID number (the content provider's internal programme number or identifier)

- Programme title (and series number if applicable)
- Episode number (if applicable)
- Episode subtitle (if applicable)

In addition, the VTRR must include:

- Log of tape contents by timecode
- Editor's technical comments
- Audio track allocation (layout)
- Audio Metadata, see section 3.2.4

5.4 Audio Track layout

Audio must be delivered with track layouts as specified in the following table. The chosen track options must be specified in the VTRR (videotape Recording Report).

AES	Track	Format	Content	Track options	Track options
1	1	Digi/SR	Main Stereo L		
	2	Digi/SR	Main Stereo R		
2	3	Digi/SR	M&E Stereo L	2 nd Language L (Digi)	Main Dolby E*
	4	Digi/SR	M&E Stereo R	2 nd Language R (Digi)	
3	5	SR	Main Front L	M&E Dolby E*	
	6	SR	Main Front R		
4	7	SR	Main Centre		
	8	SR	Main LFE		
5	9	SR	Main Surround L		
	10	SR	Main Surround R		
6	11	SR	M&E Front L	2 nd Language L (SR)	AD L (SR Only)
	12	SR	M&E Front R	2 nd Language R (SR)	AD R (SR Only)
7	13		M&E Centre		
	14		M&E LFE		
8	15		M&E Surround L		
	16		M&E Surround R		

Note: Digi equals Digital Betacam. SR equals HDCAM SR

6 Live Delivery

6.1 Responsibilities of the Production

The production should have a technical contact available as far as is possible in advance of the programme, to allow SVT to confirm technical planning, and for dealing with any queries.

There must be a technical contact available at the source during the programme itself and throughout the line-up period.

The technical contact for the programme is responsible for making sure;

- The programme meets the general overall Technical Standards outlined in sections 2 and 3 of this document.
- The cue and communications circuits are adequate and fully operational.
- The video and audio signals are continuous and stable throughout the broadcast period.
- That resilience levels meet SVT's requirements.
- That the signal leaving their site and incoming to SVT can be passed through the playout and transmission chain without the need for further technical intervention unless previously agreed using pre-booked facilities (excludes any synchronisation required at the broadcaster's point of delivery).
- That there is sufficient monitoring in place to confirm the signal quality from the location to the point of delivery.
- That all sources are stable and synchronous at all times.
- That pre-recorded inserts are the same aspect ratio, resolution and match the quality of the live material.
- Line-up signals shall be available at least 30 minutes prior to the programme start time for a single contribution link. Whenever multiple contribution connections are used the line-up signals must be available earlier in order to adjust and secure the link connections. In all cases it is strongly recommended that contact on the day is made well in advance of line-up and all possible links are tested as soon as technically possible.
- Line-up for news contribution may be done using less time for calibration and checking, however, a minimum of at least 10 minutes is necessary.

6.2 Definitions

A **Live programme** is any programme that is not delivered by tape or by file, and requires some form of communications link for delivery.

These programmes will fall under the following sub-categories:

- Live – The programme output from the remote location goes straight to air via the SVT's playout facility.
- Compliance Live – As Live but a short delay exists in the signal path to allow for intervention by SVT for compliance or legal reasons.
- As Live – The programme is produced on-site as if it were live, but the output recorded and played-out at a delayed time (or date) in the schedule. Recording and playout may occur at the production site, or at SVT's playout facility.
- Late Delivery – The programme is produced and edited very close to its scheduled time, and as such, tape or file delivery to SVT's playout facility is not practical. Delivery is via a link or permanent line from another facility.

Point of delivery is the location or building to which the live programme is commissioned to deliver, usually SVT's central routing facility MCR (Master Control Room).

Permanent Link is any dedicated path from the location to the point of delivery that uses land-based circuits that are permanently assigned for use by SVT.

Contribution Link is any path from the location to the point of delivery that is not a dedicated or permanent link, such as a bookable circuit, a satellite feed, or microwave link.

Resilience Level is the level of resilience (back-up) that a live programme is required to have. The level of resilience is a requirement of SVT, and may vary depending upon the production – see page 3.

6.3 Link Specifications

The quality of the link from the remote location to SVT's point of delivery has a major effect on the quality of the programme seen by the audience. The content, genre and workflow requirements of the programme are the primary factors that determine the bandwidth of the link.

For instance, programmes that feed into post production via the link will usually require a higher link specification than programmes that are completed on site, where the link is only used for transmission. The same can apply to programmes that are archived via a link.

Link specifications are always a trade-off between quality, cost and available bandwidth. However the link should never be considered in isolation. So in addition to the type and settings of the link encoder, the use of location radio cameras, and the transmission compression used by SVT, all have to be considered. Where there is any doubt, programme production companies should ask their link provider to speak to SVT's technical contacts.

In all instances, the delivered picture format, specifically for Live Delivery, shall be:

With 50 Hz motion portrayal, preferably:

- 1280 pixels wide x 720 pixels high
- 16:9 Aspect Ratio, Full Frame (no Letterboxing/Pillarboxing)
- 50 frames per second, delivered as 50 progressive frames per second

Or, still with 50 Hz motion portrayal, optionally:

- 1920 pixels wide x 1080 pixels high
- 16:9 Aspect ratio, Full Frame (no Letterboxing/Pillarboxing)
- 25 frames per second, delivered as 50 interlaced fields per second.

The target quality level of compressed programme material is divided into three categories:

- Grade A** Programme material of prime events like concerts, demanding sports and other events where high quality is important – i.e. the main quality target, in line with internationally accepted contribution quality, e.g. in the UK.
- Grade B** (SVT specific). When it is judged that some quality may be sacrificed due to technical or economic reasons.
- Grade C** (SVT specific). For programme material that will not be post-processed and where the content motivates that some quality loss may be acceptable due to technical limitations.

Contribution of programmes in Standard Definition quality (legacy equipment) is described in section 6.6.

Any external reference source at a remote site should be locked to GPS.

Latency of the programme feed must be less than 5 seconds, unless required by exceptional technical considerations.

The *types* of links used for any Live HD programmes shall fall into the following categories:

6.3.1 Uncompressed via Optical Fibre:

1.485 Gbps HD-SDI connection, SMPTE 292M, (often known as 1.5 Gbps HD-SDI). This remains uncompressed along its route to the point of delivery.

Wherever possible, practical, or cost-effective, programmes should use an uncompressed 1.485Gbps HD-SDI connection.

In all instances where the signal can be carried uncompressed, multichannel and/or stereo audio for the programme should be carried as discrete linear PCM (unless Dolby E is requested by SVT).

Uncompressed contribution links fall under category Grade A.

6.3.2 Compressed via Optical Fibre:

Links that provide a 1.5 Gbps HD-SDI connection at the point of delivery, but which use compression/decompression along their route.

Locations with optical Ethernet access should use compression encoders and decoders that allow HD-SDI to be transferred with 10-bit JPEG2000 compression over IP-protocol using MXF encapsulation of compressed data. The circuit must be single-hop or managed in order to guarantee the necessary quality of service.

Locations with access to other single hop fibre connections should use compression codecs that use nominal video bitrates of:

Grade A	JPEG2000	10-bit	140 Mbps
	AVC/H.264	Long GOP 4:2:2, 10-bit	45 Mbps
	MPEG-2	Long GOP, 4:2:2, (8-bit)	60 Mbps

Grade B	JPEG2000	10-bit	110 Mbps
	AVC/H.264	Long GOP 4:2:2, 10-bit	32 Mbps
	MPEG-2	Long GOP, 4:2:2, (8-bit)	42 Mbps

Grade C	JPEG2000	10-bit	80 Mbps
	AVC/H.264	Long GOP 4:2:0, 10- or 8-bit	16 Mbps
	MPEG-2	Long GOP, 4:2:0, (8-bit)	22 Mbps

6.3.3 Compressed via Satellite Link:

Where fibre is not available, links via satellite may be used. The following are permissible and achievable largely by using DVB-S2 modulation schemes. Modulation schemes should be carefully chosen so that the increase in transponder capacity (in MHz) required to deliver the optimal video bitrate (in Mbps) does not come at the cost of a decreased robustness of signal.

Single-hop satellite links should have a nominal video bitrate of:

Grade A	DVB-S2	AVC/H.264	Long GOP 4:2:2	10-bit	45 Mbps
	DVB-S	MPEG-2	Long GOP 4:2:2	(8-bit)	60 Mbps

Grade B	DVB-S2	AVC/H.264	Long GOP 4:2:2	10-bit	32 Mbps
	DVB-S	MPEG-2	Long GOP 4:2:2	(8-bit)	42 Mbps

Grade C	DVB-S2	AVC/H.264	Long GOP 4:2:0	10- or 8-bit	16 Mbps
	DVB-S	MPEG-2	Long GOP 4:2:0	(8-bit)	22 Mbps

Proposed codec settings are listed in Appendix 1 and Appendix 5.

Where the link directly feeds a second compressed link, the signal should not be decoded back to baseband but passed to the second link as a transport stream.

Stereo audio should be carried as MPEG-1 Layer II (stereo) at 384 kbps. Multichannel audio may under certain conditions be carried with phase coherence as multiple MPEG-1 Layer II stereo audio pairs at 384 kbps each, see section 6.7.2, or may have to be carried as Dolby E (see Appendix 2).

6.3.4 Compressed via Microwave Point-to-Point Link:

In some locations a point-to-point microwave link may be used as an alternative to satellite links. Microwave links can be used for short hops from the location to a fixed fibre link point or where a satellite up-link has to be remote from the production location. Where microwave links are used to feed a second compressed link, the signal should not be decoded back to baseband but passed to the second link as a transport stream.

The payload on the link should have a nominal video bitrate of:

Grade C	AVC/H.264	Long GOP 4:2:0	10- or 8-bit	16 Mbps
	MPEG-2	Long GOP 4:2:0	(8-bit)	22 Mbps

Please consult SVT if multi-hop microwave links or combination microwave/satellite links are to be used.

Stereo audio should be carried as MPEG-1 Layer II (stereo) at 384 kbps or linear PCM. Multichannel audio may under certain conditions be carried with phase coherence as multiple MPEG-1 Layer II stereo audio pairs at 384 kbps each, see section 6.7.2, or may have to be carried as Dolby E (see Appendix 2).

6.4 Picture Quality & Bit Rates (concatenation issues)

Different devices and contribution links use different compression codecs. A “codec map” must be produced for the broadcaster, listing all codecs (including the bitrates) through which a programme signal passes before arriving at the point of delivery.

This means one must produce a list of all the compression rates used along the route of the signal’s delivery, even if only one encoder/decoder pair is used.

Pictures viewed at the point of delivery should be free from visible compression artefacts when viewed on a broadcast style flat screen 40-inch display at normal viewing distance (3H, where H is the height of the visible screen-size).

To allow further processing (recording, editing etc.) especially if the signal is comprised of any additional contribution links, the highest bitrate possible must be used.

Maintaining as high a bitrate as possible throughout the production and playout process is especially pertinent when considering that the signal has to then undergo further compression and decompression in the transmission chain delivering the final product to the viewer. The aim is to avoid cascade coding artefacts both for picture and audio.

6.5 Standards Conversion

Only very high quality motion compensation (sometimes known as Motion Predictive or Motion Vector) standards converters can be used. Where a programme requiring frame-rate & standards conversion is supplied via a contribution link of less than 100Mbps, the standards conversion must be done *before* the contribution link.

6.6 Standard Definition

In some instances, some contributions may have to be delivered in Standard Definition by exception.

Where Compressed Standard Definition contribution is used, it should be compressed using MPEG-2 or MPEG-4 Long GOP. The GOP structure and encoder setup is the same as the HD requirements in Appendix 1. The payload on the link should have a nominal video bitrate of:

Grade A	AVC/H.264	Long GOP 4:2:2	10-bit	16 Mbps
	MPEG-2	Long GOP 4:2:2	(8-bit)	22 Mbps

Grade B	AVC/H.264	Long GOP 4:2:2	10-bit	11 Mbps
	MPEG-2	Long GOP 4:2:2	(8-bit)	16 Mbps

Grade C	AVC/H.264	Long GOP 4:2:0	10- or 8-bit	6 Mbps
	MPEG-2	Long GOP 4:2:0	(8-bit)	9 Mbps

Proposed codec settings are listed in Appendix 1 and Appendix 5.

Standard Definition video is 702 x 576 pixels, and the 702 pixel-wide picture must be centred in the active 720 pixel-wide line. This leaves 9 pixels to the left and 9 pixels to the right unused. (This is a result of a legacy inherited from PAL analogue TV signals).

The picture information may extend the full width of the 720 pixel wide line, providing the image shape is not distorted.

Stereo audio should be carried as MPEG-1 Layer II (stereo) at 384 kbps. Multichannel audio may under certain conditions be carried with phase coherence as multiple MPEG-1 Layer II stereo audio pairs at 384 kbps each, see section 6.7.2, or may have to be carried as Dolby E (see Appendix 2).

6.7 Audio

6.7.1 Stereo Audio

In all instances where the signal can be carried uncompressed, stereo audio for the programme should be carried as discrete linear PCM. If the signal must be carried in a compressed format, Stereo audio should be carried as MPEG-1 Layer II (stereo) at 384 kbps.

6.7.2 Multichannel Audio (when required)

Multichannel audio (surround sound) should normally be discrete PCM where there is sufficient bandwidth available in the link to the point of delivery.

Dolby E should be used for multi-channel audio when bandwidth is limited or at the request of SVT. Settings for Dolby E encoding are in Appendix 3

Use of the Low Frequency Effect Channel (LFE) channel is optional. Use of the LFE channel should comply with Recommendation ITU-R BS.775.

There should be no sample timing differences between the individual channels of a surround signal.

Each stereo pair or multi-channel group (the 6 audio tracks of a surround sound signal) must be transported in a single SMPTE ST.302 PES in order to maintain the phase relationship between channels.

Audio Metadata must be supplied, see section 3.2.4. Use VANC mapping according to SMPTE ST 2020 and Dolby Metadata for Dolby E – see Appendix 3.

Some encoder/decoder hardware support phase coherence between MPEG-1 Layer II audio pairs. If used, care must be taken to ensure that the phase coherence is preserved. Audio metadata is not carried in such streams and must therefore be supplied by other means, see section 3.2.4.

6.7.3 Audio Track Allocation

It is difficult to prescribe the exact audio track layout for all live programmes. International, host broadcaster, local requirements and link bandwidth may vary the audio layout requirements.

SVT and other broadcasters (e.g. all broadcasters in the UK) are working towards the standard audio layout below, however in many situations this layout is not yet implemented due to legacy requirements.

Please see Appendix 2 for the current requirements of SVT.

AES	Track	PROPOSED Standard
1	1	Main Stereo L
	2	Main Stereo R
2	3	M&E Stereo L
	4	M&E Stereo R
3	5	Main Front L
	6	Main Front R
4	7	Main Centre
	8	Main LFE

5	9	Main Surround L
	10	Main Surround R
6	11	M&E Front L
	12	M&E Front R
7	13	M&E Centre
	14	M&E LFE
8	15	M&E Surround L
	16	M&E Surround R

Note: It is assumed that clean commentary tracks (e.g. Commentary 1 and Commentary 2) are received via additional contribution links.

6.7.4 AV Synchronisation

AV sync should be checked via the audio desk (not just the links truck in the case of OBs) at least once a day.

Any professional AV sync equipment is acceptable but systems that can be used across all audio tracks (up to 16) simultaneously are preferred.

AV sync timing through the audio desk to the point of delivery should be ± 5 ms.

AV sync timing should also be within ± 5 ms from any remote sources into a studio or an OB including non-live inserts from tape or file.

Any external reference source at a remote site should be locked to GPS.

6.7.5 AV Synchronisation – Radio Link Cameras

Radio link cameras are always a compromise between delay and image quality. Image quality should always be the overriding consideration. Where radio and cabled cameras are mixed covering a location with lip-sync, and it is not possible to delay the audio, the radio camera should not have a delay greater than 40 ms compared to the cable cameras. The director is responsible for making sure any visible lip-sync issues are kept to a minimum.

Where all cameras use radio links, the audio must be delayed to match the video. To minimise the issues caused by open talkback and presenter switched talkback the AV sync can be ± 20 ms.

6.8 TOD timecode

Programmes should use local time of day timecode pertinent to the venue and this should be carried in the VANC unless specifically requested by SVT.

6.9 Cue and Communication

A dedicated, stand-alone technical telephone number must be provided and distributed well in advance of the transmission. This should be a fixed landline telephone.

For direct contributions into network transmissions, a feed of the source production talk-back will be required at the playout facility. A dedicated, land-based, “4-wire” circuit offers flexibility and should be considered the minimum requirement.

Talk-back (open or keyed, depending on SVT’s choice) must be offered to playout for the duration of the programme and should be available from thirty minutes before the start of transmission.

It is preferable to arrange instantaneous or low-latency video/audio return or cue paths to sources. Return audio or video cue circuits of the programme may be necessary for programmes that require two-way communication between centres. It is important to consider the latency and reliability of the cue path especially when the programme has live interviews.

Due to the latency of a Digital Terrestrial or Digital Satellite off-air signal (up to 6 seconds), off-air cueing should be considered as a last resort and for contingency purposes only.

It is acceptable to use mobile telephones for communication during the line-up period but during transmission use of mobile phones should be agreed in advance and they should not be relied on as the only means of communication.

6.10 Photosensitive Epilepsy (PSE) and Live Programmes

SVT is, for the time being, *not* subject to prevent photosensitive epilepsy.

6.11 Ancillary Data (VANC)

Where required the following lines must be used for ancillary data

Line	Data	Comments
9	SMPTE ST 2020	Main use: Surround sound metadata
11	SMPTE ST 2016	Main use: AFD
12	OP47	Subtitles if required

These lines should not be used for other data unless agreed by SVT for a specific event or programme.

6.12 UPS & Generator Provision

Unless otherwise agreed, Production companies should ensure OB suppliers, or remote locations, have UPS/Generator provision so the live programme transmission can be maintained in the event of any loss of power at the remote location. Critical systems should always be protected by UPS and if generator power is used it should be a dual system which allows synchronous changeover. This provision should be fully tested prior to transmission to ensure the functionality is fit for purpose.

7 3D (Interim for Live and Tape only)

SVT does *not* accept stereoscopic 3D programmes for the time being.

8 Appendix 1. Encoders used for Live Delivery – Additional Requirements

8.1 MPEG-2 Encoders – additional requirements

- GOP (Group-of-Pictures) should be 30 frames (50p) and 15 frames (25i). This represents a good balance between coding efficiency (requiring long GOPs) and error resilience (requiring short GOPs).
- B-frames should not be used as these are typically coded at a lower quality than I and P frames and will lead to poor picture quality in the home. Note: not all encoders on the market allow B-Frames to be disabled, so please check before accepting the unit.
- GOP structure should be /IPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP/ (50p) and /IPPPPPPPPPPPPPPP/ (25i).
- 4:2:2 colour subsampling should be used to avoid colour smearing when concatenated with the 4:2:0 emission coders used for broadcast transmission.
- **“Intra-DC precision”** should be set to 11 bits. 11 bits are required in the DCT (Discrete Cosine Transform) domain in order to accurately convey an 8 bit video signal. This is not normally a user setting but should be checked with an analyser before accepting the encoder.

8.2 MPEG-4 Encoders – additional requirements

- 10-bit video is preferred. There is no bitrate penalty.
- GOP length should be 30 frames (50p) and 15 frames (25i), in line with MPEG-2.
- Tests suggest AVC/H.264 encoders do not suffer from the poor quality B-frames. Currently B- and hierarchical B-frames are permitted.
- 4:2:2 colour subsampling is preferred, but 4:2:0 may be acceptable whilst encoder technology is developing.

9 Appendix 2. Audio Track Layout for Live Delivery

The tables below give the current audio track layouts in contribution links, where SVT is different from the proposed future 16 track layout – see section 6.7.3.

The abbreviation “M&E” stands for “Music & Effects” (without dialogue) which is also known as “International Sound”.

It is assumed that clean commentary tracks (e.g. Commentary 1 and Commentary 2) are received via additional contribution links.

These tables are subject to change over time.

SVT’s current 4 track layouts

AES	Track	SVT4-A	SVT4-B	SVT4-C	SVT4-D
1	1	Main Stereo L	M&E Stereo L	Main Stereo L	M&E Stereo L
	2	Main Stereo R	M&E Stereo R	Main Stereo R	M&E Stereo R
2	3	M&E Stereo L	-	Main Dolby E multi-ch	M&E Dolby E multi-ch
	4	M&E Stereo R	-		

SVT’s current 8 track layouts

AES	Track	SVT8-A	SVT8-B	SVT8-C	SVT8-D
1	1	Main Stereo L	M&E Stereo L	Main Front L	M&E Front L
	2	Main Stereo R	M&E Stereo R	Main Front R	M&E Front R
2	3	Main Front L	M&E Front L	Main Centre	M&E Centre
	4	Main Front R	M&E Front R	Main LFE	M&E LFE
3	5	Main Front C	M&E Centre	Main Surround L	M&E Surround L
	6	Main LFE	M&E LFE	Main Surround R	M&E Surround R
4	7	Main Surround L	M&E Surround L	Main Stereo L	M&E Stereo L
	8	Main Surround R	M&E Surround R	Main Stereo R	M&E Stereo R

SVT’s current 16 track layouts (legacy)

AES	Track	SVT16-A (PROPOSAL, see section 6.7.3)	SVT16-B
1	1	Main Stereo L	Main Front L
	2	Main Stereo R	Main Front R
2	3	M&E Stereo L	Main Centre
	4	M&E Stereo R	Main LFE
3	5	Main Front L	Main Surround L
	6	Main Front R	Main Surround R
4	7	Main Front C	Main Stereo L
	8	Main LFE	Main Stereo R
5	9	Main Surround L	M&E Stereo L
	10	Main Surround R	M&E Stereo R
6	11	M&E Front L	-
	12	M&E Front R	-

7	13	M&E Front C	-
	14	M&E LFE	-
8	15	M&E Surround L	-
	16	M&E Surround R	-

10 Appendix 3. Dolby E

Dolby DP571 (or equivalent Dolby E encoder).

- The first DP571 in the chain must be set to "internal" metadata source.
- Metadata from an external source cannot be edited in a DP571 – you have to set the metadata source to internal and rebuild it from scratch if this is required.
- A Dolby E stream uses 75Ω-unbalanced connections – if the loop-through connectors on a DP571 are not feeding another device they should be terminated.
- The most important DP571 setup parameters to check are:
 - Program Config (should be 5.1 + 2 for most things).
 - The Frame Rate. (The factory default is 29.97 fps; SVT's defaults are 25 fps for both 25i and 50p video – there is no 50 fps setting for 50p in DP571. If Dolby E is to be used with 50p video (or 59.94p video), it is subsequently necessary to decode to linear PCM audio before splicing/editing).
 - The Bit Depth (should be 20 bits).
 - The Dolby E Metadata Source (internal or external).
 - Dolby D metadata is enabled (note – if this setting appears in the DP571 menu the unit is running out of date firmware and *MUST* be updated).
- The “Bit Depth” setting does not refer to the number of bits used by the audio signals but the number of bits available in the AES-3 channel used to transmit the Dolby E signal. Six channels of audio can be transmitted down a 16-bit link while 8 channels must be transmitted down a 20-bit link.
- The gain of a Dolby E stream must not be changed – it will break.
- A Dolby E stream suffers one frame of delay on each encode and decode. The SVT standard is to align the stream to be “sync encoded” at each point in the chain.
- For tape delivery the stream must be laid so that the encoded audio is in sync with the video and any stereo audio. The Audio Advance function will be used to move the stream 1 frame ahead of the video before decoding. As Audio Advance applies to all audio tracks, the stereo audio will be delayed by 1 frame to compensate.
- For live or as live programmes the stream must be timed so that the encoded audio is in sync with the video and any stereo audio. It is the responsibility of the area receiving the stream to compensate for the decoding delay.
- The DP571 has a “reversion mode” which tells the unit what Dolby E and Dolby Digital metadata parameters to use in the event of a metadata failure. This should be set to “last used”. This setting guards against temporary metadata failure – if failures occur the Dolby Digital encoding should not be affected.
- In the “Metadata Params” menu there are eight programs where metadata can be changed. The number of active programs depends on the “Program Config” setting. If Program Config is set to the default 5.1+2 value then Program 1 will contain the metadata for the 5.1 stream. Program 2 will set the metadata for the additional stereo pair. If the Program Config is set to 8X1 then Program 1 will set the metadata for channel 1, Program 2 for channel 2, Program 3 for channel 3 etc. It is worth noting that all eight programs remain viewable even if the selected channel configuration does not make use of them.

11 Appendix 4. Naming of Audio Formats when using multiple mono WAV files

Quote from section 4.4.1

“In the case of separate 1-track (mono) “.wav” files, the file names must match those specified above in the cells to indicate the content of the file, e.g. PG-1234567-001A-PROGRAMME_TITLE-321_ME_LS.wav (for the left surround channel in a 5.1 multichannel music and effects mix).”

Short form in the table “Track Allocations”, section 4.4.1	End of the filenaming	Indicates	Comments
St. L	-ST_L	Stereo Left	Main audio
St. R	-ST_R	Stereo Right	Main audio
St. M&E L	-ST_ME_L	Stereo Music & Effects Left	
St. M&E R	-ST_ME_R	Stereo Music & Effects Right	
St. Aud Desc L	-ST_AD_L	Stereo Audio Description Left	
St. Aud Desc R	-ST_AD_R	Stereo Audio Description Right	
5.1 L	-321_L	Multichannel Left (front)	Main audio
5.1 R	-321_R	Multichannel Right (front)	Main audio
5.1 C	-321_C	Multichannel Centre (front)	Main audio
5.1 LFE	-321_LFE	Multichannel Low Frequency Effects	Main audio
5.1 Ls	-321_LS	Multichannel Left Surround	Main audio
5.1 Rs	-321_RS	Multichannel Right Surround	Main audio
5.1 M&E L	-321_ME_L	Multichannel Music & Effects Left (front)	
5.1 M&E R	-321_ME_R	Multichannel Music & Effects Right (front)	
5.1 M&E C	-321_ME_C	Multichannel Music & Effects Centre (front)	
5.1 M&E LFE	-321_ME_LFE	Multichannel Music & Effects Low Frequency Effects	
5.1 M&E Ls	-321_ME_LS	Multichannel Music & Effects Left Surround	
5.1 M&E Rs	-321_ME_RS	Multichannel Music & Effects Right Surround	
5.1 Lang 1 L	-321_LANG1_L	Multichannel Language 1 Left (front)	
5.1 Lang 1 R	-321_LANG1_R	Multichannel Language 1 Right (front)	
5.1 Lang 1 C	-321_LANG1_C	Multichannel Language 1 Centre (front)	
5.1 Lang 1 LFE	-321_LANG1_LFE	Multichannel Language 1 Low Frequency Effects (front)	
5.1 Lang 1 Ls	-321_LANG1_LS	Multichannel Language 1 Left Surround	
5.1 Lang 1 Rs	-321_LANG1_RS	Multichannel Language 1 Right Surround	
5.1 Lang 2 L	-321_LANG2_L	Multichannel Language 2 Left (front)	
5.1 Lang 2 R	-321_LANG2_R	Multichannel Language 2 Right (front)	

5.1 Lang 2 C	-321_LANG2_C	Multichannel Language 2 Centre (front)	
5.1 Lang 2 LFE	-321_LANG2_LFE	Multichannel Language 2 Low Frequency Effects (front)	
5.1 Lang 2 Ls	-321_LANG2_LS	Multichannel Language 2 Left Surround	
5.1 Lang 2 Rs	-321_LANG2_RS	Multichannel Language 2 Right Surround	
St Lang 1 L	-ST_LANG1_L	Stereo Language 1 Left	
St Lang 1 R	-ST_LANG1_R	Stereo Language 1 Right	
St Lang 2 L	-ST_LANG2_L	Stereo Language 2 Left	
St Lang 2 R	-ST_LANG2_R	Stereo Language 2 Right	
St Lang 3 L	-ST_LANG3_L	Stereo Language 3 Left	
St Lang 3 R	-ST_LANG3_R	Stereo Language 3 Right	

Explanation: 5.1 = 3/2/1

When the audio format is not “5.1” in the short form above:

Audio Format – consumer description	Audio Format – professional description	Audio Format – in naming above	Comments
1.0	1/0/0	100	One front channel (centre). No Surround channels. LFE off.
2.0	2/0/0	200	Two front channels. No Surround channels. LFE off.
3.0	3/0/0	300	Three front channels. No Surround channels. LFE off.
3.1	3/0/1	301	Three front channels. No Surround channels. LFE on.
4.0	2/2/0	220	Two front channels. Two Surround channels. LFE off.
4.1	2/2/1	221	Two front channels. Two Surround channels. LFE on.
5.0	3/2/0	320	Three front channels. Two Surround channels. LFE off.
5.1	3/2/1	321	Three front channels. Two Surround channels. LFE on.

12 Appendix 5. Parameters for satellite contribution

12.1 Contribution for HD (informative)

For the number of B-frames in GOP

Allocated Bandwidth	Video Encoding		Video Standard		Sampling		GOP Length		DVB Standard		Modulation		Symbol Rate		Roll Off		FEC		Pilot		TS Bit Rate (188)		Video Bit Rate		Audio Bit rate per 2 ch pair	
	Codec Levels/Profiles	Bandwidth	Video Standard	Video Standard	Sampling	Bit Depth	GOP Length	DVB Standard	Modulation	Symbol Rate	Roll Off	FEC	Pilot	TS Bit Rate (188)	Video Bit Rate	Audio Bit rate per 2 ch pair										
Grade A	H264 L4.1	24 MHz	HD 720p	DVB-S2	4:2:2	10	30	8 PSK	19200	0.25	3/4	on	41808	38500	384											
	H264 L4.1	24 MHz	HD 1080i	DVB-S2	4:2:2	10	15	8 PSK	19200	0.25	3/4	on	41808	38500	384											
	MPEG-2	36 MHz	HD 720p	DVB-S2	4:2:2	8	15	8 PSK	27750	0.25	3/4	on	60416	56217	384											
	MPEG-2	36 MHz	HD 1080i	DVB-S2	4:2:2	8	15	8 PSK	27750	0.25	3/4	on	60416	56217	384											
Grade B	H264 L4.1	18 MHz	HD 720p	DVB-S2	4:2:2	10	30	8 PSK	14400	0.25	3/4	on	31356	27500	384											
	H264 L4.1	18 MHz	HD 1080i	DVB-S2	4:2:2	10	15	8 PSK	14400	0.25	3/4	on	31356	27500	384											
	MPEG-2	24 MHz	HD 720p	DVB-S2	4:2:2	8	15	8 PSK	19200	0.25	3/4	on	41808	38500	384											
	MPEG-2	24 MHz	HD 1080i	DVB-S2	4:2:2	8	15	8 PSK	19200	0.25	3/4	on	41808	38500	384											
Grade C	H264 L4.1	9 MHz	HD 720p	DVB-S2	4:2:0	10/8	30	8 PSK	7200	0.25	3/4	on	15678	13500	384											
	H264 L4.1	9 MHz	HD 1080i	DVB-S2	4:2:0	10/8	15	8 PSK	7200	0.25	3/4	on	15678	13500	384											
	MPEG-2	12 MHz	HD 720p	DVB-S2	4:2:0	8	15	8 PSK	9875	0.25	3/4	on	21503	19500	384											
	MPEG-2	12 MHz	HD 1080i	DVB-S2	4:2:0	8	15	8 PSK	9875	0.25	3/4	on	21503	19500	384											

12.2 Contribution for SD (informative)

For the number of B-frames in GOP-le

Grade	Allocated Bandwidth		Video Encoding		Video Standard		Sampling	Bit Depth	GOP Length	DVB Standard	Modulation	Symbol Rate	Roll Off	FEC	Pilot	TS Bit Rate (188)	Video Bit Rate	Audio bitrate per 2 ch pair
	MHz	MHz	H264 L4.1	MPEG-2	SD 576i	4:2:2	10	12	DVB-S2	8 PSK	7200	0.25	3/4	on	15678	13000	384	
Grade A	9 MHz	18 MHz	MPEG-2	SD 576i	4:2:2	8	12	DVB-S	QPSK	13333	0.25	7/8	on	21503	19500	384		
	6 MHz	9 MHz	H264 L4.1	SD 576i	4:2:2	10	12	DVB-S2	8 PSK	4937	0.25	3/4	on	10750	9000	384		
Grade B	6 MHz	9 MHz	MPEG-2	SD 576i	4:2:2	8	12	DVB-S	QPSK	6111	0.25	3/4	on	15678	13500	384		
	4.5 MHz	6 MHz	H264 L4.1	SD 576i	4:2:0	10/8	12	DVB-S2	8 PSK	3600	0.25	3/4	on	7830	7250	256		
Grade C	4.5 MHz	6 MHz	MPEG-2	SD 576i	4:2:0	8	12	DVB-S	QPSK	4434	0.25	3/4	on	7151	6250	256		
														Multiplexing (kbps)				
													Modulation					
													Codec Levels/Profiles					
													Bandwidth					