BBC

R&D White Paper

WHP 087

May 2004

A quantitive comparison of TV-Anytime and DVB-SI

J.C. Newell

Research & Development
BRITISH BROADCASTING CORPORATION

BBC Research & Development
White Paper WHP 087

A quantitive comparison of TV-Anytime and DVB-SI

Chris Newell

Abstract

The DVB-GBS group has recently completed the specification for the "Carriage and signalling of TV-Anytime information in DVB transport streams". One function this specification can support is the signalling of schedule data, mirroring the capabilities of the Event Information Tables defined in the existing DVB-SI specification. Although TV Anytime uses a more complex data structure than DVB-SI there is a potential efficiency gain due to the combined use of Zlib compression and a binary format for XML called BiM. This Technical Note gives a quantitive comparison of the efficiency of the TV Anytime and DVB-SI approach, which may be useful when considering possible roles for the new specification.

Additional key words: EIT, Schedule, metadata

A quantitive comparison of TV-Anytime and DVB-SI

Chris Newell

1   Introduction

The DVB-GBS group has recently completed the specification for the "Carriage and signalling of TV-Anytime information in DVB transport streams". One function this specification can support is the signalling of schedule data, mirroring the capabilities of the Event Information Table (EIT) defined in the existing DVB-SI specification. Although TV-Anytime supports a more complex data structure than DVB-SI there is a potential efficiency gain due to the combined use of Zlib compression and a binary format for XML called BiM. Until recently it was impossible to assess the efficiency of the DVB-GBS approach accurately, due to the lack of a DVB compliant BiM codec. However, a codec has recently become available from Expway allowing this comparison to be made.

2   Approach

To make a meaningful comparison between the two systems, a set of TV-Anytime data was created that matched the capabilities of a typical profile of DVB-SI. The size of the compressed TV-Anytime data was then compared with a calculated figure for an equivalent set of DVB-SI data.

TV-Anytime separates programme information and scheduling information into separate tables, the Program Information Table and the Program Location Table. For this comparison the following fields were supported for every program or event represented in each table:

Program Information Table
CRID
Title (average length: 13 chars)
Synopsis (average length: 144 chars)
Audio/Video Attributes
Genre

Program Location Table
CRID
Published Start Time
Published Duration

To carry an equivalent set of information in DVB-SI, each event in the EIT would have to carry the following descriptors:

- one Short Event Descriptor, carrying a title and synopsis

- two Component Descriptors, signalling the audio and video format

- one Content Descriptor, signalling genre

It is not possible to achieve an exact equivalence between TV-Anytime and DVB-SI, in terms of the information that is carried, but the differences are limited:

- TV-Anytime supports unique programme identifiers (CRIDs) whereas the EIT supports a 16-bit event_id.

- TV-Anytime has a broader set of genres

- TV-Anytime specifies the language code for Title and Synopsis individually, whereas in the EIT a single language code applies to both.

It should be noted that TV-Anytime is capable of supporting a much richer set of metadata when required, beyond the currently defined capabilities of DVB-SI.

3   Compression of TV-Anytime metadata

The DVB compliant codec uses several mechanisms to compress the TV-Anytime metadata:

- The XML structure of TV-Anytime metadata is coded using BiM, an efficient binary format for XML adopted by MPEG-7.

- Time/Date and DVB Locator fields have their own specific codecs.

- Strings are concatenated within each delivery unit (Container) to ensure efficient Zlib compression is achieved in the Object Carousel delivery layer.

The separation of program information from scheduling information in TV-Anytime also provides a useful efficiency gain whenever programs are repeated, since each instance can share a common set of program information.

4   Small data set

To study the efficiency of compression independently from the effect of repeat programs a small TV-Anytime data set was created, representing a single service with 36 unique program events over a 24-hour period. The size of the dataset was measured to be:

TV-Anytime	bytes	Relative size
XML source data	34662	100 %
BiM encoded XML Container	9763	28.2 %
Container after Zlib compression	4089	11.8 %

The size of the metadata container after the BiM encoder and Zlib compression have been applied is 4089 bytes, a compression efficiency of 11.8 %. Note that some additional formatting information (e.g. fragmentIds and version numbers) is added by the codec during the compression process.

The size of an equivalent EIT Schedule service can be calculated, assuming the same average Title and Synopsis lengths (13 and 144 respectively).

DVB-SI - EIT		bytes
Section Header and CRC-32		18
Per event:		
Event header etc:	12	
Short Event Descriptor:	164	
Component descriptors (2):	30	
Content Descriptor:	4	
Total:		210
Total for 8 Sections / 36 Events:		7704

To make a fair comparison between the TV-Anytime approach and DVB-SI an overhead for carriage via the Object Carousel carriage has to be added, which might be 4 %. This gives a figure of 4252 bytes for the TV-Anytime data, compared to a calculated size of 7704 bytes for the equivalent DVB-SI data.

5   Large data set

To explore the efficiency of the TV-Anytime approach for a larger data set, a 7-day schedule was assembled for the six TV services, two pairs of which are time exclusive so that the dataset is equivalent to four 24-hour services. The schedule includes a total of 1171 schedule events, but only 937 unique program descriptions i.e. around 20% of the scheduled programs are repeats of programs broadcast earlier in the week. The size of the data set allowed an Object Carousel to be created, giving an accurate figure for the size of the section data:

TV-Anytime	bytes	Relative size
XML source data	991779	100 %
BiM encoded XML Container	283989	28.6 %
Container after Zlib compression	97860	9.9 %
Object Carousel sections	102813	10.4 %

The size of an equivalent EIT Schedule service can be calculated, assuming the same average Title and Synopsis lengths (14 and 144 respectively). As each EIT Section covers a 3-hour period for each active service, there are 336 Sections in total, an average of 3.5 Events per Section

DVB EIT Schedule		bytes
Section Header and CRC-32		18
Per event:		
Event header etc:	12	
Short Event Descriptor:	165	
Component descriptors (2):	30	
Content Descriptor:	4	
Total for 3.5 Events:		739
Average Section size		757
Total size of 336 Sections		254352

It can be seen that the TV-Anytime approach requires 102813 bytes of section data compared to 254352 bytes for the DVB-SI approach.

6   Conclusions

For the small data set, the TV-Anytime approach requires only 55 % of the data capacity required for the DVB-SI approach. This is reasonable degree of compression, since a significant amount of data capacity (5754 bytes) is taken up by the title and synopsis strings which are not easily compressed.

For the large data set, the TV-Anytime approach requires only 40 % of the data capacity required for the DVB-SI approach, indicating that there can be an appreciable efficiency gain when there is a significant number of repeat events with identical program information.