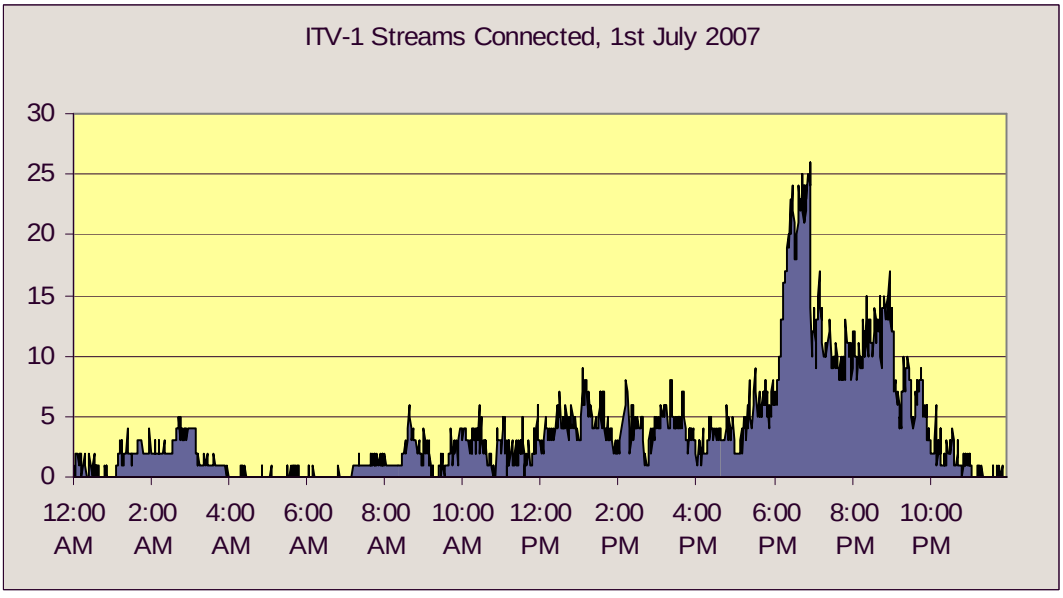
BMWG Stage 1 (Simulcast)

Summary of Methodology

1. Introduction

This document sets out the methodology for preparing the graphs which illustrate the key rights metric agreed by the BMWG for Stage 1 of their work – simulcast streaming. The graphs show the number of simultaneous stream connections active during every minute of a particular period – as in the example shown.



It is anticipated that for continuous reporting of simulcast channels, one graph (with the underlying table of data) per channel, per day is produced. However, for particular one-off events, the schedule provided by the broadcaster – in other words, the length of time over which the event is broadcast – defines the extent of the graph.

2. Data Preparation and Filtering

2.1 Data Format

Ensure that the data provided contains the following fields:

- The URL of the simulcast stream

- IP address of the Internet application (browser or media player) requesting the stream
- Identifier (User-Agent) of the browser or media player requesting the stream
- An indicator of date and time
- A way in which Duration (in seconds) can be derived
- An HTTP-standard status code (e.g. 200, 404, 501) denoting success or failure of the stream request

Important Note:

The two main stream log formats, RealPlayer Helix and Windows Media Player, log different time stamps for streams.

Windows Media Player logs record the time at which the stream **started** and, if configured to do so, a ***Duration*** (x-duration) field in seconds. Thus, the time at which the stream **ended** can be expressed as [Start Time]+[Duration].

```
87.248.198.118 2007-09-01 17:50:39 - /stream1 0 7203 1 200 {00000000-0000-0000-0000-
000000000000} 9.1.1.3814 en-US WMServer/9.1.1.3814 - WMServer.exe 9.1.1.3814
Windows_Server_2003 5.2.0.3790 Pentium 0 0 392456 http TCP - - - 353365484 353365484 72760
72760 0 0 0 0 0 0 0 100 87.248.198.109 scds9ams.llnw.local 1 0 - 2028546
rtspt://83.98.75.138/channel1 http://scds9.ams.llnw.net/stream1 / - - 0
```

RealPlayer Helix logs record the time at which the stream **ended** as a standard Apache timestamp. The time at which the stream ***started*** is recorded in a field further down the log line, as specified by the logging configuration. So the duration has to be expressed as [End Time]-[Start Time].

```
81.86.224.194 - - [13/Dec/2007:13:59:37 +0000] GET /news/sol_now5a_bb.rm HTTP/1.0 404 187
[WinNT_5.1_6.0.12.1483_play32_GOOG01_en-US_686_axembed] [00000000-0000-0000-0000-
000000000000] [UNKNOWN] 0 0 0 1699243372 1919247457 [0 0 0 0] [13/Dec/2007:13:51:32]
212.58.227.16 0 0
```

2.2 Data Normalisation

Once the start time and end time of each stream has been calculated, these times must be expressed as absolute seconds values – either Unix (epoch) seconds or seconds since the start of the month. In both cases, clearly, the date must be taken into account. This is to cater for differences in log format as noted above and to ensure that the graph can be prepared in a standard manner.

For clarity, these values will be called StartSecs and EndSecs below.

2.3 Filtering

The filtering process should now remove:

- all requests from invalid IP addresses (e.g. those internal to the broadcaster);
- all requests from invalid (robotic) User-Agents;
- all requests for other simulcast channels (or content items, if looking at an individual content item) present in the log;
- all requests with invalid status codes (so not 200-205);
- all requests which do not EITHER:

o have an EndSecs value greater than the absolute seconds value of the first second of the period (typically 00:00:00 on any given day) for which the graph is being produced – so ended within the period, OR

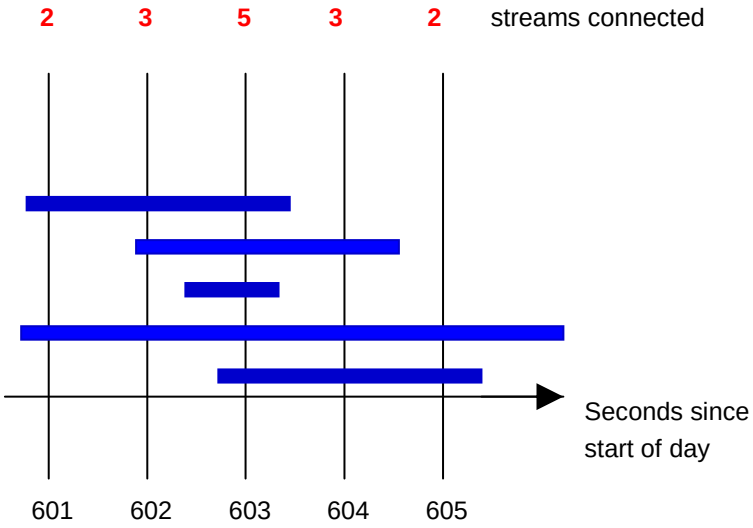
o have a StartSecs value less than the absolute seconds value of the last second of the period (typically 23:59:59 on any given day) for which the graph is being produced – so started within the period.

Following these steps will produce a filtered data set from which the graph can be prepared.

2.4 Graph Preparation

There are 86400 (24 x 60 x 60) seconds in every day. Any logged line of data for which StartSecs is less than a given second and EndSecs is greater counts towards the simultaneous connections for that second. For example, a stream that starts on second 00059 and ends on second 00061, and a stream that starts on second 00023 and ends on second 03201, both count towards the total streams connected in second 60.

Consequently a table will be created which has, as a minimum, 86400 rows (one per second) each showing the total streams connected in that second. This table must then be filtered to show only absolute seconds values within the day or period being measured.



A graph can now be compiled of concurrent connections (y-axis) against seconds (x-axis). To get a minute-by-minute graph, simply extract every 60th record. Analysis and comparison of data produced shows that this continues to give a representative picture of the activity being graphed.