# Constructing an IPTV Workload Model

Fernando M.V. Ramos*, Fei Song*†, Pablo Rodriguez‡, Richard Gibbens*, Jon Crowcroft*, Ian H. White*

*University of Cambridge †Beijing Jiaotong University ‡Telefonica Research, Barcelona

## 1. INTRODUCTION

IPTV services are among the fastest growing television services in the world. While in 2005 only two million households had access to this service, this figure jumped to more than 14 million households in 2007.

Incumbent operator's IPTV networks are "walled gardens", well provisioned to guarantee the user experience required by TV viewers. Current networks use static IP multicast within their network domain. The existing architectures have proved successful in broadcasting the hitherto limited number of TV channels, but the IP network was not designed to accommodate a huge number of niche channels with asynchronous viewing patterns. For this reason, understanding [1] and modelling this new type of traffic can offer significant insights into traffic engineering and capacity planning for network providers.

The aim of this work is thus to build the first IPTV workload model. For this purpose we are analysing channel switching logs of a quarter million users from one of the worlds largest Telco-managed IPTV networks.

### 1.1 What this work is, and what it is not

IPTV is sometimes confused with WebTV or P2P-based TV. IPTV in this context means a cable-like TV service offered on top of an IP network (instead of the hitherto predominant cable network or free air broadcast). This is a server-centred architecture, in contrast with a p2p architecture (section 3). On demand-video services are usually offered as a bundle in IPTV service packages, but in this work we are only interested in modelling live TV usage, capturing the way users surf and watch the available TV channels.

## 2. MOTIVATION

IPTV is a brand new service offered on top of IP networks, where users exhibit a behaviour unseen in other IP-based applications. In Figure 1 we illustrate the typical behaviour of an IPTV user.

In this plot one can see the switching events performed by a typical user during a representative day. A switching event occurs when a user selects a new TV channel. Several switching events in a certain time $t$ show that the user is browsing through TV channels. When there is no switching event for a relative amount of time, this means the user is watching a specific channel. There is also the possibility that the user left the channel on, but is away. These "away events", lasting several hours, are quite common because IPTV users very rarely switch off the Set Top Box (STB). As one can see, this type of workload is completely new in an IP network, and therefore we consider it worth to be analysed (as in [1]) and modelled.

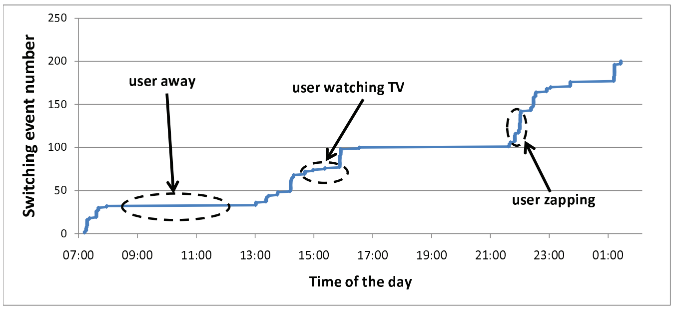


**Figure 1: IPTV switching events (one typical user, one typical day)**

## 3. DESCRIPTION OF AN IPTV NETWORK

Figure 2 illustrates a typical IPTV service "walled gardened" architecture. The TV head-end sources IPTV contents to its routers that are connected to Digital Subscriber Line Access Multiplexers (DSLAMs) through an IP backbone. These DSLAMs aggregate traffic from several users (usually in the order of the hundreds) and connect to the high-speed backbone. The TV channels are distributed from the Headend to the DSLAMs through bandwidth-provisioned, static, multicast trees. The network is only dynamic between DSLAMs and customers, where multicast trees are extended or pruned based on the channel switching signals.

## 4. DESCRIPTION OF THE TRACES

This study is based on the characterisation of a large

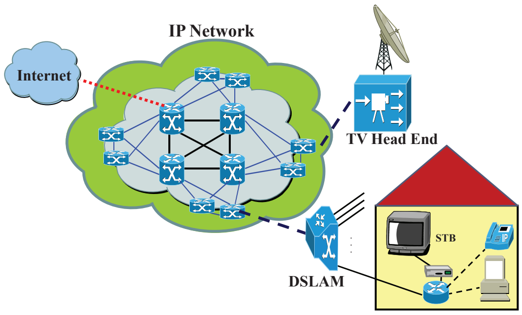


**Figure 2: IPTV network**

dataset of a commercial IPTV service. The logs were collected over a six month period. The traces scale to 250 thousand users, 680 DSLAMs, 150 TV channels, and they cover a whole country. The trace includes all switching events during this period, and each event includes information on its type (user switched to or from a certain channel), the timestamp, the IP addresses of the DSLAM and the STB, and the TV channel.

## 5. QUESTIONS TO ADDRESS

A workload model can be defined by two main components: the arrival rate of requests, and the service time of each request. Finding the right probability distribution for each of this components is thus the main objective of our work.

From Figure 1 one can see a rapid burst of channel selection events - "zapping periods" - followed by an extended period without any channel being selected - "watching/away events". One way forward to model the stream of selection events could be to have a compound model of "burst events" occurring with some inter-arrival distribution (*inter zap block interval*) - perhaps, exponential to form points of a Poisson process - and then for each burst to consist of a random number of selection events (*zap block size*)- say, geometrically distributed.

In order to confirm these assumptions, we propose to address several questions, from which we extract some of the most important ones:

**a)** How many and what channels are browsed during a zapping period? How linear is a zapping period? What is the purpose of zapping?

**b)** For how long does a user watches a channel? What channels do people watch?

**c)** What is the aggregated result per DSLAM? And per region?

## 6. PRELIMINARY RESULTS

The preliminary results show that our simple assumptions were incorrect. The zap block size's empirical data does not fit into a geometric or poisson distribution. However, it seems to fit into a Gamma distribution, with $k$=0.5 and $\theta$=5.8 (Figure 3). The inter-zap block interval does not fit into an exponential distribution, and it may be necessary to fit it with more than one function (in the plot we fit it with three functions, including two gamma with different parameters).

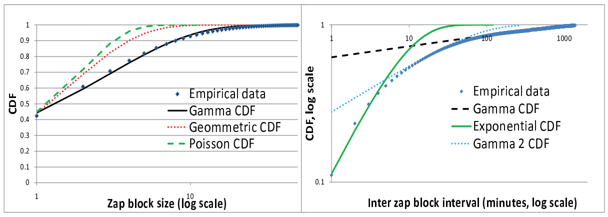


**Figure 3: CDF of zap block size and inter zap block interval (40,000 users, 6 months)**

The majority of channel changes, during zapping events, is sequential[1]. This knowledge of the TV channel that follows in a user's selection can be used to build novel network architectures that may, for instance, reduce the latency experienced by IPTV users. In order to explore this issue further, we want to understand why people zap: is it just to chill out after adverts, or is it mainly as a way of switching channels? For this purpose, our experiment is to calculate the probability of a user returning to the channel he was watching after zapping.

Other type of analysis we are doing is to analyse the aggregate behaviour at the DSLAM level (aggregating hundreds of users) and at the regional level (aggregating dozens of region-specific DSLAMs). Preliminary results show that, at the DSLAM level, only 30% of the channels are watched at any one time. This means the network is wasting more than 70% of its resources by using static IP multicast. We have discovered, however, that for some very popular channels there is **always** at least one household watching the channel, so it makes sense for these to be always available in the DSLAM. But there is a significative percentage of unpopular TV channels that are seldom watched (or even browsed through). The regional level analysis follow the same trend.

All this is work currently in progress, so new results are regularly made publicly available online[1].

## 7. ACKNOWLEDGEMENTS

We would like to thank Telefonica for allowing us to analyse this very rich dataset.

## 8. REFERENCES

[1] M. Cha, P. Rodriguez, J. Crowcroft, S. Moon, and X. Amatriain. Watching television over an IP network. In *Proc. IMC*, 2008.

[1] Check http://www.cl.cam.ac.uk/~fmv21/iptv_workloads