ADAPTIVE SWITCHED DIGITAL SERVICES IN CABLE NETWORKS

Michael Adams
Terayon Communication Systems

Abstract

Switched Digital Broadcast (SDB) is a new technology that has generated considerable interest because it promises better utilization of HFC bandwidth combined with the potential to offer the customer a virtually unlimited number of broadcast (linear) channels. However, progress on implementation has been slow, and the focus has been on vertical solutions that do not inter-work with other services such as digital ad insertion, VOD, and network-based PVR. Moreover, there is no clear revenue model for SDB; it is a technology rather than a service.

SDB promises the greatest increase in bandwidth efficiency when it is applied to the 'long-tail' content which is rarely viewed. However, it is difficult to predict which channels will represent this group and how real-world events dynamically change the popularity of a given channel. What is needed is an intelligent system that can adapt to instantaneous changes in channel viewing behavior.

This paper describes an approach to providing Switched Digital Services that combines the video digital processing required for ad insertion (both into broadcast and on-demand programming) with the ability to switch video in unicast or multicast modes. Therefore future digital switched services must seamlessly include support for advanced features such as targeted Ad Insertion, Fast Channel Change, VOD, and Network PVR.

INTRODUCTION

Switched digital services (often called switched broadcast or switched digital broadcast) are conceptually very simple; *"Only transmit those channels that are being watched"*.

In a cable system of any significant size every broadcast channel will have at least one viewer. However, deployment of Video on Demand (VOD) and High Speed Data (HSD) services into cable systems has naturally led to the creation of "service groups". Essentially, the cable system is segmented into a large number of HFC segments, each with its own dedicated bandwidth. By doing this, narrowcast services can be supported at high penetrations by re-using the same spectrum over and over. This is often termed "spatial re-use".

Within these service groups, which are typically between 500 and 2,000 homes-passed, the broadcast channels have a much lower probability of being watched. Trials conducted by Time Warner Cable in Austin, Texas [Brooks] indicate that, even at peak times, only 45 out of 170 channels will be simultaneously viewed in a service group that serves 1000 homes. The other 125 channels are just occupying spectrum and are completely unused. If 12 standard definition programs are carried in each 256-QAM channel, the total spectrum required for 170 programs is 84 MHz (14 QAM channels). And 74% of this spectrum is not being actually used – a tremendous waste of resources in a 750 MHz plant – in fact, about 9% of the total downstream capacity!

Switched Digital Services Algorithm

So how does *"Only transmit those channels that are being watched"* actually work in practice?

1. When a set-top 'tunes' to a switched program:

    - If that program is currently being transmitted in the service group, the set-top selects it as normal.

    - If that program is not currently being transmitted to the service group, the set-top signals to request it. The system then starts transmission of that program, and acknowledges the request. Then the set-top selects it.

2. When the set-top tunes away, the set-top signals to release the program.

3. When all set-tops within a service group tune away from a program, transmission can be stopped (if necessary) to recover channel capacity.

Benefits of Switched Digital Services

The biggest advantage is that new linear programming services can be added without limit to a switched digital services line-up. This "virtually infinite" linear programming line-up allows many niche programs to be added. This will be an important addition to a cable operator's ability to respond to competition from satellite and telco providers.

Another significant advantage is to reclaim bandwidth for other services, such as VOD, High-Speed Data, VoIP or Digital Simulcast.

In addition, there are several other advantages that can be gained by deploying switched digital services including targeted advertising, fast channel change, and integration with VOD services.

Service Group Size

In practice, a block of QAM channels is assigned to each service group, and the following "VOD math" is used to calculate the optimal service group size as follows. First we need to calculate how many digital set-tops can be supported simultaneously:

$$\#\text{Digital Set-tops} = \frac{\#\text{QAM Channels} \times 10}{\text{Peak Usage}}$$

Let's plug some numbers into this equation; if a block of 4 QAM channels per service group is chosen, then each service group receives a maximum of 10 constant bit rate VOD streams, each encoded at 3.75 Mbps. If the peak usage is 7.5%, we have a service group size of 533 digital set-tops. Taking into account the digital set-top penetration gets us to the number of homes passed in the service group, as follows:

$$\text{Service Group Size} = \frac{\#\text{Digital Set-tops}}{\text{Digital Set-top Penetration}}$$

Let's take an example, where the digital set-top penetration is 50% of homes passed. In this case, we have a service group size of 1,066 homes passed.

SDS REFERENCE ARCHITECTURE

Figure 1 shows reference architecture for switched digital services.

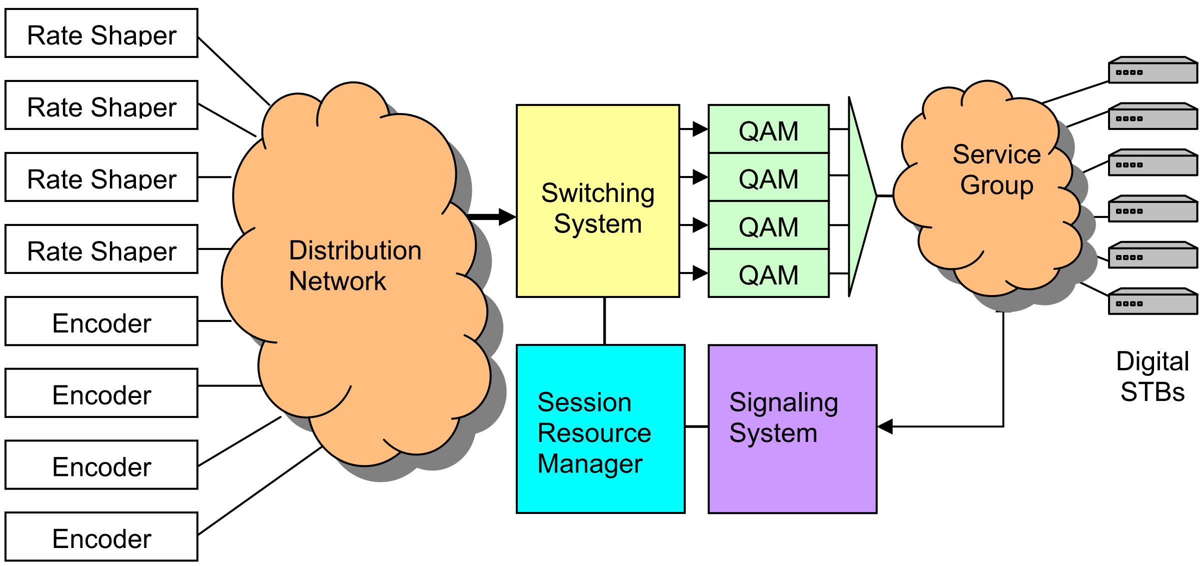

Figure 1. SDS Reference Architecture

The key features of the architecture are:

Constant Bit Rate (CBR) encoders

Before being switched, all programs are converted into MPEG-2 single program transport streams (SPTS) at a constant bit rate of 3.75 Mbps. This is the same format as is used for standard definition VOD content.

There are two ways of doing this:

1. If the channel is already available as a variable bit rate (VBR) program, it may be possible to rate shape it into a CBR program. However, this may not always produce good results. For example, a demanding sports feed such as ESPN, might peak at a significantly higher bit rate, and clamping the bit rate to 3.75 Mbps may cause a significant loss of picture detail. Visually, this will appear as tiling – where all the picture coefficients have been approximated to the point where macroblock boundaries become evident to the human eye.

2. A second option is to re-encode the program from a baseband (uncompressed) signal. This will typically produce a better result because the encoder has more flexibility in coding the stream than the rate shaper. Nevertheless, it can be very challenging to encode certain feeds into a 3.75 Mbps SPTS. In practice this equates to cost – requiring a more sophisticated and expensive encoder.

Despite these drawbacks, 3.75 Mbps was selected in early implementations because it makes the session resource management simpler. At 3.75 Mbps, 10 programs will comfortably fit into a 256-QAM channel.

In future, it is likely that multiple CBR rates might be supported. For example, 3.75 Mbps for movie channels, 4.5 Mbps for sports channels, and, of course, significantly higher rates for HD channels.

Distribution System

After the SPTS are built at the master headend, they must be distributed to the switching systems. An efficient way to accomplish this is to use IP multicast service. Each switching system will join the multicast group for those channels that need to be delivered to that service group.

Switching System

The output to each service group comes from a switch that essentially has n inputs, where n is the total number of available channels, and m outputs, where m is the maximum number of channels that can be viewed simultaneously.

In some implementations, the switch has been collapsed into the modulation system [Bombelli]. In others, it is a separate piece of equipment.

Modulation System

The modulation system is a bank of QAM modulators. Fortunately, high-density Edge-QAMs, developed for VOD, are ideal for this purpose. The number of QAMs assigned to each service group must be carefully chosen to reduce the probability of blocking to an acceptable level.

Set-top Client

For switched digital services to be deployed, existing set-top boxes must be able to support them. This is achieved by developing a special piece of client software that monitors channel selection, and signals to the system whenever a channel is selected or released.

In early implementations, a dedicated switched digital services client was developed. In future, an integrated client that supports VOD and switched digital services will be required to simplify software integration.

Signaling System

The signaling system allows each set-top to request a program when it is selected by the user. The signaling system that already exists for VOD signaling can be re-used for this purpose. In fact, even the protocol can be re-used with some minor extensions.

Session Resource Manager

The session resource manager (SRM) is responsible for making sure the correct channels are made available in each service group according to the set of signaling requests received over time. In implementations to date, a dedicated SRM has been developed for switched digital services (as opposed to re-using the existing VOD SRM). There are several reasons for this:

1. The switched digital services SRM must understand the concept of shared programs within a service group. In other words, if a second set-top requests a particular channel that is already being transmitted, no action is necessary.

2. The switched digital services SRM must provide an extremely rapid response to the user’s program request. A typical response time of 100 milliseconds is desirable, which is an order of magnitude faster than most VOD SRMs.

3. Implementation simplicity – simply put, it is usually easier to start from scratch to develop an SRM from the ground up than to modify existing code.

IMPLEMENTATION CHALLENGES

Switched digital services are not without significant implementation challenges. Some are similar to VOD and others are unique to switched digital services. Fortunately, there are solutions to each of these challenges and, for clarity that is how we've organized the next sections.

One-way Consumer Electronics agreement

Digital cable-ready receivers built to the one-way "plug and play" agreement [NCTA], cannot signal to request a switched broadcast channel. The only solution is to keep a set of programs in the broadcast tier. Fortunately, not all linear programming needs to be constrained to this tier. In the case of digital simulcast, the same programs are provided in analog, and are therefore available in that format to a cable ready receiver, so these channels can be switched without impact. In addition, any new linear programming could be made available in a switched digital services tier. Obviously, it would be important for customers to be informed that a set-top box would be required to receive them.

The long-term solution is a two-way capable digital cable-ready receiver. Although progress on the definition of a single standard has been slow, several bi-lateral agreements with specific manufacturers have been developed and first units have been deployed. A switched broadcast services application would have to be developed and downloaded to these devices in order to allow them to participate in a switched digital services tier.

Program Blocking

Program blocking will occur at any time when the number of simultaneous program requests exceeds the capacity that has been allocated in a service group. This is the same effect that happens in any switched system, whether it is a telephone system or a VOD system. The difference here is that the customer's expectations have been determined by the behavior of broadcast cable, which doesn't normally block. In light of this what probability of blocking is acceptable? A lot of operators would say "zero!" Unfortunately, engineering the system to ensure that blocking will never occur means building in significant additional headroom.

However, it is possible to use some other tools that are already in the bag to ameliorate this situation. Statistical multiplexing is commonly used in cable systems to combine multiple variable bit rate (VBR) programs into a constant bit rate (CBR) QAM channel. Each program is processed in real-time to ensure that the instantaneous sum of the bit rates never exceeds the total capacity of the QAM channel. (If it does, even for a few milliseconds, MPEG packets will be dropped and the video will literally fall to pieces, as compressed digital video is extremely vulnerable to packet loss.)

Unfortunately, statistical multiplexing is too expensive to be applied to every service group because there are so many of them. However, if we observe that blocking is unlikely to occur in all service groups simultaneously, we can see that only opportunistic statistical multiplexing is needed. This could be implemented at a realistic price point by sharing the statistical multiplexing resource across a number of service groups. This leads us towards a model where the switching system incorporates a statistical multiplexing resource.

Impact of Channel Surfing

Imagine a user is channel surfing through the switched digital services tier. For each channel change, the system must establish a session for the new channel and tear-down the session for the old channel. Now imagine that a large number of viewers all start to channel surf at the same time (for example, when a popular channel goes to commercial).

The impact of channel surfing can be minimized by transmitting popularly watched channels even when they are not being viewed. (Essentially this is exactly the way the CPU cache works in a computer system.) Caching could be implemented by building a learning algorithm into the SRM. A simple first-in, first-out algorithm will work quite effectively and minimize the number of sessions that have to be established and released. A more sophisticated algorithm could weight channels according to the number of viewers watching them simultaneously.

In addition, a fast channel map indicating which the channels are in the cache can be sent via a broadcast carousel to all set-tops. On a channel change, the set-top client checks to see whether the program is already being transmitted, in which case the set-top can tune to it directly. (Note that it must still signal to request the program because otherwise the SRM would have an inaccurate picture of the system state, and the session carrying the program could be unexpectedly released at some future time.)

However, the signaling system and SRM will still have to keep up with the signaling load, even if the session establishment/release rate is reduced. In addition, we have to be sure that the cache (switched digital services tier size) is reasonably large, otherwise channels will be added and released constantly and we will be back where we started. (In computer science terminology, this behavior is aptly termed "thrashing".)

The best way to build a high-performance signaling system is to distribute it across multiple CPUs, each serving a set of service groups. This also has the nice property of reducing the failure group size. There is also the possibility for building a fault-tolerant system if each set-top client maintains a list of multiple signaling sub-systems. In this arrangement, if a request from the first signaling sub-system doesn't yield a timely response, the client re-tries the request with an alternate signaling sub-system.

Operational Impacts

There is no doubt that the introduction of switched digital services changes the model of a broadcast channel. SDS requires zero non-responders and 24/7 upstream network availability.

However, some additional robustness and graceful degradation can be built into the service:

1. SDS session manager periodically broadcasts tuning information for current state to ensure that the system stays in a consistent state. The set-top client is responsible for monitoring these messages and taking appropriate action. For example, if it is connected to a session but detects that its identifier is not in the reference list for that session, it must re-request the

session to prevent unexpected termination of the session.

2. In the event of a failure to reach the SDS session manager due to a communications path failure or failure of the SDS session manager itself, the STB client may choose to use the last tuning information received for the desired program.

3. In the event of an SDS session manager failure, the sessions can be left untouched. After the SDS session manager recovers, it should be possible for it to recover the system state by polling the set-top clients to retrieve their current state. Thus the impact of an SDS session manager failure does not cause each viewer to have to re-tune their set-top after the outage.

Program Hold Time

How do you know if anyone's watching a program (or if the STB was just left on)? This is analogous to making a phone call to someone, mid-way through the conversation you have to answer the door, and you completely forget about your phone call. Because there is an intelligent party on the other end, they will hang up and the call will be released. But in the case of switched digital services, the system doesn't know whether you are still watching the program or not.

One solution is to monitor the viewer's activity by keeping track of remote control commands. If no remote control commands are received for a preset time-out period, then an on-screen message can be generated to test is the viewer is still there (and paying attention). After a further time-out, the session would be released. Obviously this approach could cause confusion and irritation for some viewers; however it would only become necessary at some threshold where the number of free sessions is becoming almost exhausted, so during periods of lower contention for resources it is less likely this algorithm would have to be executed.

A related problem is that some viewers will manually tune to a channel and set their VCR to record a program on that channel at some later time, leaving the system unattended. In this case it is quite possible that the session would be released either before, or, more irritating still, during the program recording period. The solution is to educate the customer in the use of VCR control functions or to deploy more DVR capable set-tops.

Cost of Switching and Modulation Equipment

From the reference architecture, it is apparent that a lot of additional switching and modulation equipment will be required to deploy switched digital services. Let's take the example of a medium-sized cable system that passes 300,000 homes. Assume that we decide to use a service group size of 1,000 homes-passed, and that 50 sessions per service group are required. For every 10 sessions, we will need to modulate a 256-QAM channel, thus requiring 5 QAM channels per service group. With 300 service groups in the system, a total of 1500 channels of QAM modulation will be required.

Fortunately the solution for this already exists; high-density edge-QAM modulators were developed to provide a low cost modulation for VOD. Given the scale of modulation resources, any viable switched digital services deployment will leverage a highly commoditized edge-QAM market.

Switching cost is less of an issue because a single switch can easily serve a large number of service groups. In our example there are a total of 15,000 streams at 3.75 Mbps, or 56 Gbps of total switching capacity.

In addition, there are a great many advantages to be gained from deploying a narrowcast architecture such as switched digital services. Each session is consumed by a small group of viewers (in some cases only one viewer), and so there is a tremendous potential for customization of the content to that audience. This could be done using a combination of Digital Program Insertion (DPI) and graphical overlay technologies.

ADAPTIVE SWITCHED DIGITAL SERVICES

One of the most difficult challenges in a switched digital services deployment may be choosing which programs can be placed into the switched digital tier. Some programs must stay in the broadcast tier because of programming agreements or because they must be made available to one-way cable ready receivers. Others will be clear candidates for switched digital service tier, such as niche programming.

In the middle is a gray area of channels that could fit into either group according to system dynamics. In an adaptive SDS implementation, the caching algorithm will effectively promote switched channels to broadcast channels if they become heavily used. Moreover, the operator can manually promote them to un-switched mode on a per service group basis if necessary.

Benefits of Adaptive SDS

SDS is an interesting technology, but with significant deployment and operational impacts, operators are looking for additional features and benefits to justify its deployment.

SDS will be deployed first in bandwidth-constrained systems that need to add digital simulcast for competitive reasons but do not have the available spectrum for the additional QAM channels required to carry them. But as a pure bandwidth management tool, SDS is best viewed as a stop-gap measure.
Removing a channel from the analog tier creates tremendous bandwidth saving, but it is a slow process. Nevertheless, each channel removed from the analog tier allows 12 new SD digital channels or 2 HD digital channels.

We can start to see more potential benefits from switched digital services if we observe that, for the first time, the system has a direct way of measuring who is watching broadcast programs and when they are watching them.

Targeted Advertising

This information is of great interest to advertisers because it could support targeted advertising. In turn, targeted advertising promises increased revenues for the cable operator. Of course, special care must be taken to preserve the individual viewer's privacy but this is not incompatible with matching the advertisements to the viewer's interests based on established demographic classification.

How might this work in practice with SDS? When the first viewer tunes to a program, the system would start to insert commercials in that channel according to the demographic classification into which that viewer falls. When a second viewer joins the session, the system might deliver commercials that intersect the two viewer classifications. Alternatively, if the system is lightly loaded, the SDS resource manager could choose to deliver the program over a new session to the second viewer, enhancing the capabilities for targeting commercials.

Fast Channel Change

Fast channel change (FCC) is being offered by some competitive IPTV products as a benefit over conventional digital cable implementations. In FCC, the transition from one channel to another is accomplished more quickly due to two changes to the architecture:

1. The channel change is accomplished by switching in the network, rather than in the set-top, so there is no re-tune latency.

2. The new channel is conditioned such that it starts with an intracoded picture – so that all of the information to display the picture arrives immediately on channel change.

The combination of these techniques can produce a maximum channel change time that is quite fast (i.e. better than analog cable and maybe one fifth that of digital cable; for example, 0.2 seconds versus 1 second is achievable).

In the case where a viewer switches to a SDS program that is not yet being transmitted in the service group, a similar technique could be employed to commence from an intracoded picture, avoiding the delay of waiting for the next I-frame.

In the case where a viewer 'tunes' from one SDS program to another, and is the sole viewer of the program in the service group, the existing session could be re-used and a program switch achieved in the network.

Transition to Network PVR

Some operators are experimenting with PVR services, where the network allows time-shifting of programming using VOD technology. An example of this is Time Warner Cable's "Startover" service [Yahoo!]. The Startover service allows the viewer to watch a particular program from the beginning after selecting it part-way through. To achieve this, the viewer is switched to a VOD session for a delayed version of the program.

A Startover session is actually a type of SDS session, where the program is unicast and the program is streamed from a VOD server rather than a real-time programming source. In fact the SDS session can be re-used when the viewer transitions from a broadcast program, where he or she is the only viewer within the service group, to a Startover session. However, to do this the modulation resource must be shared between Startover (the VOD system) and the rest of the SDS system.

In addition, Startover sessions are another great opportunity to increase the value of advertising by targeting the commercials so they are more likely to be related to the viewer's interests. That same SDS targeting infrastructure can be seamlessly re-used for inserting commercials into a program regardless of whether it is a broadcast or VOD session. In fact, the same SCTE 35 digital cue messages that indicate the ad "avails" can be used to support commercial insertion.

Ad-supported VOD sessions can also re-use the same infrastructure required for SDS. Once a SDS system is deployed it only makes sense to share modulation resources across all switched services, whether they are broadcast programs, Startover sessions, or VOD sessions.

Requirements for SDS systems

Now that we have discussed some of the benefits that can accrue for the deployment of SDS, we can start to develop a set of requirements for SDS systems. In order to use SDS as more than a stop-gap measure for bandwidth management, the system should support the following requirements:

1. Support targeted advertising by incorporating DPI capabilities into the switching system.

2. Support fast channel change by incorporating the ability to synchronously deliver an intracoded picture to the newly selected channel.

3. Seamlessly integrate with Startover by sharing modulation resources with the VOD system.

4. Support opportunistic statistical multiplexing to provide additional, "elastic", headroom when necessary under periods of peak demand.

The common thread of all these requirements is that the switching system is the ideal point in the system to provide the additional functions and benefits we have discussed. This is an interesting conclusion because it flies in the face of the conventional wisdom of some proposed switched broadcast implementations, where the switching system is incorporated into the modulation system in order to minimize cost.

CONCLUSION

In this paper we have summarized the various challenges, solutions, and benefits to deploying switched digital services.

We have argued that enhancements to the switching system can generate significant incremental revenue by supporting targeted advertising in broadcast and VOD programming. In addition, competitive features such as fast channel change could be introduced by adding flexibility into the switching system. Finally, by incorporating opportunistic statistical multiplexing, the bandwidth management properties of the system can be enhanced.

Therefore, in a SDS solution, we argue that the operator should invest additional resources into the switching system in order to realize these benefits.

REFERENCES

1. Brooks, Paul: Time Warner Cable Digital Broadcast Field Trial Results Austin, Texas. SCTE Cable-Tec Expo 2005 Proceedings and Collected Technical Papers.

2. Bombelli, Lorenzo and Brooks, Paul: An Open Architecture for Switched Digital Services in HFC Networks. SCTE Conference on Emerging Technologies 2006 Presentations and Collected Technical Papers.

3. NCTA, Cable MSO-Consumer Electronics Industry Agreement on "Plug & Play" Cable Compatibility and Related Issues. December 19, 2002.

4. Yahoo! Entertainment: Cable Comes to the Rescue of Channel Surfers, Dec 23, 2004.