

THE USE OF BROADCAST INFRASTRUCTURES FOR ON-DEMAND SERVICES

Kelly L. Dempski
Accenture Technology Labs
161 N. Clark St.
Chicago, IL 60601 USA
Kelly.L.Dempski@accenture.com

Abstract: Most on-demand architectures focus on delivering services through new delivery mechanisms such as broadband networking or digital cable. Very few focus on increasing the capabilities of the endpoints. This paper proposes the use of inexpensive consumer electronics devices such as PVRs to create a storage based on-demand solution. Such a system would be capable of utilizing currently available broadcast networks to deliver content, alleviating the need for new and costly real time streaming architectures. This paper examines the feasibility of such a solution and suggests additional services that could be delivered through a storage-based architecture.

Keywords: Personal Video Recorder (PVR), Video On-Demand, Personalized content.

1. INTRODUCTION

Traditionally, nearly every form of transmitted content has been delivered to consumers first through a broadcast mechanism. The reason for this is simple. Broadcasting is the easiest and least expensive way to deliver the maximum amount of content to the maximum number of people. The downside is that broadcasting by definition doesn't allow much room for personalization. Broadcasters have been forced to adopt solutions such as pay-per-view (PPV) in approximate an on-demand experience through a broadcast architecture.

However, the advent of digital networking solutions has presented broadcasters with an alternate delivery mechanism. Networks allow packets to be directed at specific endpoints and are therefore a natural medium for personalized or directed content. As the technology has improved, content providers have begun to build solutions such as video-on-demand (VOD) based on these networks. One could argue that this is this obvious choice given the features of IP networks. However, the purpose of this paper is to

show how other technologies might provide better, cheaper solutions that leverage existing and nearly ubiquitous broadcast infrastructures.

We propose that digital video recorders and related technologies can be used to create on-demand, personalized services that deliver high quality media streams through existing broadcast networks. Time shifting, coupled with intelligent scheduling, can deliver an on-demand experience without necessitating on-demand delivery. The concepts described can be applied to variety of broadcast technologies, from analog cable to multicast IP video. The majority of this paper discusses issues as they relate to video distribution. However, it is important to realize that similar principles could be applied to the distribution of any form of content where some form of broadcast infrastructure is already in place.

2. CURRENT ARCHITECTURES

At a high level, most broadcast architectures are modeled after the idea that a single head-end transmitter feeds a large number of consumers through a broadcast medium. Content is scheduled and queued based on popularity and programming guidelines. As stated above, broadcast architectures create very little opportunity for personalization. At best, personalized service is provided through different channel offerings, scheduling, and in some cases, pay-per-view.

At the other end of the spectrum, IP networks are capable of feeding individual users through point-to-point connections through an IP network. This creates the opportunity for personalized service if the service and the head-end are designed to respond to individual requests. The drawback to this architecture is that the overall system performance is heavily dependent on network bandwidth and head-end performance. For many broadband customers, bandwidth is still very limited. Providers typically deal with this by offering lower quality streams. Although compression quality has increased dramatically, in most cases streaming digital video is still lower quality than the analog streams that have been available for the last 50 years.

Finally, an intermediate solution exists. Some providers have mitigated performance issues by placing server stations nearer to clusters of users. This improves overall system performance by distributing the load. This can improve overall system performance, but quality is still constrained by the bandwidth between the consumer and the nearest station. In many cases, the connection into a single home is still relatively thin. Very few subscribers

are capable of receiving IP streams with a quality equal to that of a traditional broadcast transmission.

For the sake of argument, we will label each of these solutions as examples of “transmission-based architectures”. The last two examples try to improve on the first through significant changes to the transmission infrastructure and minimal changes to the endpoints. In most cases, the endpoints are set-top-boxes that are extremely thin client computers that offer minimal user interface functionality and a network address. These approaches are all based on the assumption that it is extremely expensive to roll out powerful end points. Historically, this has been true. However, the remainder of this paper reexamines this issue and proposes that more powerful set top boxes may in fact offer a much better solution for delivering personalized on demand content.

3. STORAGE-BASED ARCHITECTURES

When talking about extending the functionality of set top boxes, improvements can be made on many dimensions. One could talk about user interface improvements, better transaction security, etc. For the purposes of this discussion, we will limit our scope to a view of the set top box as a storage device with rudimentary UI features. If we assume that a large population of users will eventually own set top boxes with storage capabilities, we can envision “storage based architectures” as shown in Figure 1 below.

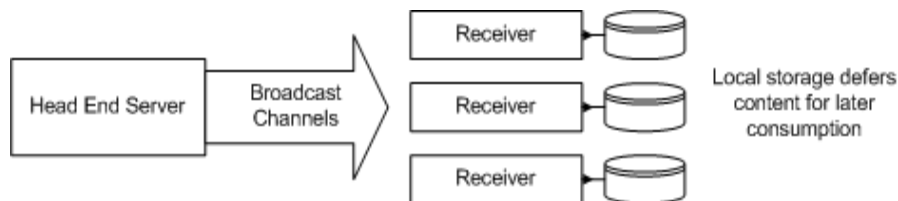


Figure 1. A storage based architecture.

In Figure 1, content is delivered to each consumer through a broadcast medium, where it might or might not be stored at the end point. Such a system is in no way novel; this has been possible since the advent of the VCR. Personal video recorders (PVRs) such as TiVo [1] extend the functionality of VCRs through richer interfaces, better functionality, and digital storage. Novelty (and more importantly, consumer value) could be delivered as a function of how the content provider leverages the capabilities

of a PVR. Some content providers such as AT&T and DirecTV are offering PVRs to their subscribers [2], but they aren't using the technology to deliver new services. Yet, the power to do so is at their fingertips.

Although adoption of PVR technology has been slow, PVR users are already using them as "on-demand" devices. Users will typically set their PVRs to record a number of programs. From those selections, the PVR service will make further guesses about what the viewer might also want to see. When the user turns on the TV, they are given a complete list of the programming currently available on the PVR. Many users are thereby freed from the notions of "channel" or "schedule". They know that they can watch the most recent episode of "Friends", but they don't necessarily care how it was delivered to them.

One could argue that such a system is not "on-demand", or that it is at best "near video on demand" (NVOD). For many, "on-demand" is synonymous with "delivery on demand". However, it is much more interesting to talk about consumption on-demand. For many PVR users, consumption of television programming is on-demand even though delivery might have happened much earlier. This is an important but often forgotten point. Transmission based architectures are built on the premise that delivery and consumption are temporally linked, but this need not be true in many cases.

Currently, the functionality of PVRs is enjoyed by a relatively small number of television viewers, but it does provide a glimpse of what is possible with a storage-based architecture. For the PVR user, the broadcast network is simply a transmission mechanism that feeds their set top device. By virtue of storage and a small amount of processing, the PVR presents them with an enhanced viewing experience. This is done with no change to the broadcast infrastructure and in most cases the customer has happily purchased their own PVR equipment with little or no subsidies from the content provider.

We suggest that content providers should take notice of this paradigm shift and use storage-based architectures as the basis for new delivery services. At the most basic level, a storage-based architecture is one that assumes that a traditional broadcast network supplies client devices and that these devices have storage capability and a small amount of decision-making power. Based on this, we can create on demand services with minimal changes to the existing infrastructure.

4. EXISTING INFRASTRUCTURES

Existing broadcast infrastructures come in many forms. Terrestrial cable providers supply analog and/or digital cable. Additionally, many cable providers offer broadband access or at least a thin backchannel. Satellite providers deliver several audio and video channels, and in some cases they also offer broadband access. Regardless of the actual technology used, there are two features that are common to most systems.

The first feature is a wide range of channels. A single provider might offer hundreds of channels. These channels might carry continuous feeds of audio or video, or they might deliver pay-per-view or special event programming. Some channels might be “dark” at times. The large number of channels increases the chances that, at any given moment, there is something to capture the interest of any single user. However, it is also possible that, at any given moment, large numbers of channels are completely ignored by the user base. Current systems offer a wide range of content, but the signal to noise ratio can be quite low.

Another feature is the fact that these infrastructures represent a large investment for the providers. Ideally, any further value they can derive from the existing system will further offset the cost of the investment. In some cases, a considerable portion of that investment comes from the addition of broadband services. In many cases, the broadband system might be adequate for web browsing, but not for the delivery of broadcast television quality media. The best way to maximize the value of the existing investment would be to use the existing media channels for media delivery and the broadband channels for service selection, transactions, and control.

5. EXISTING AND FUTURE SET TOP CAPABILITIES

As stated above, the value of PVRs is their ability to easily record and catalog content on a digital medium. Their weakness, from the point of view of the content provider, is that they are largely a double blind system. Aside from predetermined schedule information, the PVR has very little knowledge of the current state of the content provider. Preempted programming or last minute scheduling changes can disrupt recording schedules. Also, the network provider has no knowledge about the use of the content once it has been delivered to the PVR. The content provider might

get aggregate usage statistics [3] from the PVR service provider, but there is no individual real time content usage feedback.

Both problems could be solved with relatively minor but important changes to the PVR hardware and software. Broadband connections, while often too thin to deliver quality video streams, are more than sufficient to deliver real time scheduling information to storage devices. Devices such as ReplayTV already use broadband to access scheduling information, although it tends to come from a scheduling aggregator instead of the content provider. The broadband connection is also sufficient to send user preferences and requests to the content provider, as well as real time usage statistics.

Endpoints on a storage-based architecture would also need to integrate security and licensing features. This would allow broadcasters to stream premium content to a wide number of consumers in an encrypted form. If a consumer chooses to view the content, the provider can supply a key, decrypt the content, and charge the viewer at the time of consumption. DirecTV [4] and other receivers currently handle encrypted broadcast streams, which are decrypted upon reception. Decryption could be deferred until actual consumption. Integrating it into the PVR functionality would require minimal software changes.

With these two changes, the current PVR becomes a viable endpoint for a storage-based architecture. A relatively low bandwidth bidirectional connection allows the provider to schedule content based on demand, and the security features create a viable platform for settlements and rights management. Together, they allow the PVR to grow into a platform through which new services can take advantage of current broadcast infrastructures.

6. UBIQUITOUS TIME SHIFTING

Currently, PVR users experience the value of “time shifting”, which essentially means that the PVR allows them to break free of the existing broadcast schedule by virtue of the storage capabilities. In a world of relatively few PVRs, this is a small but very important step. In order to illustrate a dramatic change in services, we will fast forward to a time when a critical mass of users has a storage device that incorporates the new features explained above. This critical mass could be the result of a massive rollout by the provider or of a slower migration for individual customers. A fraction of the available channels could be dedicated to feeding the storage-

based network. As the install base grows, so could the number of channels. The following discussion assumes a critical mass of devices receiving signals from hundreds of channels.

When storage becomes ubiquitous, time shifting also becomes ubiquitous and content scheduling quickly loses meaning. People no longer watch shows based on channels and schedules. There is, in effect, no “prime time”. Rather, people watch shows based on their preferences and whims. VOD advocates have long envisioned this type of behavior, but always as the result of on-demand transmission. A storage-based solution supports the behavior with very little change to the transmission mechanism, but the final piece of the puzzle rests in necessary changes to the scheduling process.

When control of consumption rests with the viewer, content scheduling takes a very different form. Because there is no prime time, content can be scheduled based on real time user demand instead of historic patterns. In a sense, scheduling becomes a resource optimization problem where user demands are weighed against available content and network resources. A single content provider might have hundreds of available channels through which content can be transmitted. Ubiquitous time shifting means that content can be streamed at any time in a 24-hour day. For instance, a new release of a movie might be streamed to everyone’s PVR in the middle of the night. From then on, it is available whenever the consumer wants it. This is a new source of value for the content provider. Currently, a large portion of the day has low viewership, and therefore low value. During these times, broadcast channels might carry low quality programming or no programming at all. Time shifting allows for more efficient use of broadcast resources because value is no longer dependent on time of day.

Time shifting creates the need for a content scheduling system that is dynamic, intelligent, and responsive to real time user demand. In a storage-based architecture, the challenge is to deliver a given piece of content *before* the user demands it. This necessitates a scheduling system with some intelligence.

7. INTELLIGENT SCHEDULING

The problem of intelligent scheduling can be broken down into a handful of key pieces. There are network resources (network channels and user storage), atomic pieces of content (episodes, movies, songs, etc.), and user demands. An intelligent scheduling system is one that tries to fulfill the latter

through optimal use of the former. Figure 2 shows a high level diagram of such a system.

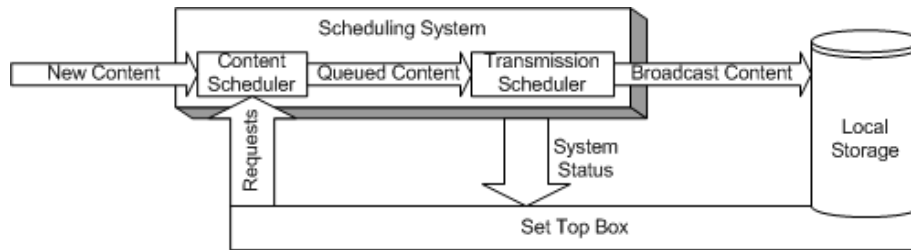


Figure 2. An intelligent scheduling system.

Interaction begins at the user's set top box. The role of the set top box is to receive content, but it also provides the interface through which the user can request content. Requests can be generated in many ways. The user might request a movie for immediate viewing, as in a true on-demand architecture. They might also request a movie to be delivered to their media library for later viewing much as they would purchase a DVD. In this case, timeliness is less of a factor and the system provides a means of digital distribution. The user can also create persistent requests, such as requests for new episodes of a particular TV series as they become available. Finally, the set top box could include software that creates requests on the user's behalf. Such a software agent could use a variety of techniques to generate requests based on user behavior.

User request objects may have many attributes. An object identifies a desired piece of content, and it may also include a notion of priority, timeliness, or willingness to pay. A request generated by a software agent might be tagged with a lower priority than one that is generated by a real user. These requests might either be persistent or immediate. A persistent request could define standing requests for any new episode of a television show or any new movie from a particular director. An immediate request would be an explicit request for a certain piece of content. In either case, requests are sent to the Content Scheduler, where they are queued, pending available content.

The purpose of the content scheduler is to optimize the distribution of content based on a variety of factors. If an immediate request is received, the content scheduler will first verify that the content is actually available. If so, it will check the general state of the transmission scheduler and queue the content for delivery based on the priority of the request and the availability of network resources. Persistent requests are similar, but queuing is triggered

when new content is available. If new content becomes available, the content scheduler will check for persistent requests and queue accordingly. The content scheduler maintains the queue of content for the transmission scheduler.

The transmission scheduler routes queued content to appropriate network resources. Its role is to optimize the use of the network based on content length, format, and priority. In a homogeneous system, this could be a matter of load balancing between broadcast channels. In a heterogeneous environment, the transmission scheduler could route content to different channel types based on priority and format. Figure 3 shows one such scenario. Low priority requests (where timeliness is not important)are routed through broadcast channels and immediate requests are routed through a smaller collection of digital network resources.

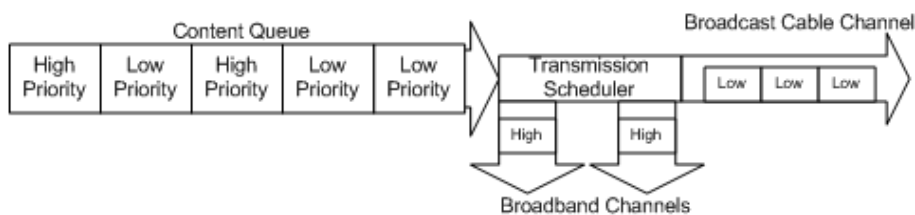


Figure 3. Allocating resources based on request type.

Finally, the scheduling system provides the set top boxes with real time status information about the content queue. The content scheduler provides information about pending content in the queue. The transmission scheduler provides the set top box with channel information needed to access the content. This information allows the set top box to optimize the use of its own resources, both in terms of storage and tuners.

Traditionally, the word tuner refers to a device that is capable of tuning to a particular transmission frequency. We are using the term more loosely as a means of receiving content. One can imagine set top boxes that have a number of traditional tuners that are capable of receiving broadcast content from either cable or satellite channels and a broadband “tuner” that is capable of receiving digital media. The set top box can allocate the use of those tuners based on the information it receives from the scheduling system. In some cases, the number of channels with desired content might outweigh the number of available tuners. In these situations, the set top box will record content based on priority and resubmit requests for lost content. This might not be a detrimental as it might seem.

8. INCREASED EFFICIENCIES AND THE ECONOMIES OF STORAGE

At first glance, the described architecture might seem to be significantly inferior to a true on-demand system, but we believe that this is not actually the case in real world situations. The use of local storage, combined with much better use of network resources, creates a viable platform for “consumption on demand”. To understand why, one must take a deeper look at the way the network is being used.

For the sake of example, we will look at the DirecTV pay-per-view system. DirecTV dedicates 100 of its channels to pay-per-view content [5]. Popular movies might be shown on four or more different channels, allowing a user to wait no longer than one half hour for a two-hour movie. Similar systems are in place to transmit pay-per-view content on cable networks and multicast digital networks. Repeated showings and staggered channels allow many users to view the content when they want to.

The downside of this system is that the transmission network is actually greatly underutilized. For example, four channels are capable of sending 96 hours of content for every 24 hours of time. Instead, they are being used to repeat the same 2 hours of content. Effectively, the channels are only being used at roughly 2% of their actual capacity. In aggregate, The 100 channels of pay-per-view are being used to transmit roughly 60 movies per month. Assuming an average time of two hours per movie, this means that 120 hours of content occupies 72000 hours of broadcast time. In this case, the average utilization is a staggeringly low 0.16%. Stated differently, the same block of channels could be used to deliver roughly 36000 movies a month, which is probably more than the content providers have in their libraries.

The reason behind these gross inefficiencies is that current architectures assume that the content must be available for users at unpredictable times. The time shifting capabilities of a PVR remove that constraint. Movies recorded at arbitrary times can be consumed on demand. In fact, there are many users that currently have the capability to create true “*consumption on demand*” systems using their PVRs.

Currently, many DirecTV subscribers have high capacity PVRs with dual tuners. These systems would be capable of recording 48 hours of movies in the first 24 hours of a given month (albeit at great consumer expense, given the current payment structure). These movies would then be available on demand for the rest of the month. Assuming the recording schedule was

prioritized, it would be extremely difficult, if not impossible, to contrive a scenario where content was not yet available at the time a user chose to watch it. These users do exist, and they have acquired the technology without the cost of infrastructure upgrades or equipment subsidies.

Of course, the number of advanced PVR users is still relatively low, and it might require investment from the content providers to build the user base. However, this investment might be substantially less expensive than an infrastructure change. PVRs have value above and beyond their use as part of an infrastructure. If marketed correctly, many users might be more inclined to purchase a PVR than they would another form of set top box. Because storage is an easily configurable variable, service could easily be scaled to different levels for different customers. A basic customer with a 10-hour PVR could store 5 movies at any given time. At the high end, there are 320-hour PVRs capable of storing approximately 160 movies at any time. With proper filtering, these movies could be a subset of a much larger selection. To put this into perspective, 80% of video rental revenues are earned in the first two months of a release [7]. During that two-month period, less than 100 new mainstream videos are released. Therefore, the 320-hour PVR could service 80% of the current rental requests with no filtering whatsoever and still have capacity for other PVR functionality.

The advantages of storage based solutions over IP based solutions become more striking when you factor in the geographical availability of bandwidth. It is very easy for a user to bring a hard drive home, but the same user might not even have the option of subscribing to broadband services. However, nearly all users have access to some form of broadcast content, which is easily recorded. Even if broadband was ubiquitous, there is still a question of quality. This is perhaps most striking in the case of HDTV. A digital HDTV signal could be stored easily, yet the same level of quality would be extremely difficult to deliver through common broadband connections. High quality streams require either high bandwidth or high storage capacity and storage is less expensive for the foreseeable future. Also, consumers will have storage to support other functionality [6].

Greater network efficiency, coupled with the ability to store and time shift content creates a viable platform for on-demand services. The numbers are striking, and become even more so when you factor in the features that are simple byproducts of a storage based system. Some VOD vendors tout features like rewind, pause, and fast forward. A storage-based solution does all of those things trivially. Also, many digital video providers are struggling to deliver video quality that comes close to broadcast quality. In a storage-

based system, quality is dependent on an individual's local compression settings, but broadcast quality is always an option. As HDTV sources become available, they can be easily integrated into a storage-based system and their broadcast nature makes them widely accessible. An IP Stream with the same quality would require significantly increased network bandwidth.

9. CONCLUSION

The purpose of this paper was to outline, at a high level, the value and promise of a storage-based network as it relates to the delivery of several forms of content. For the sake of simplicity, several details about the scheduling process were omitted in favor of a broader discussion of the implications. The intelligence of the scheduler could take many forms, and an actual implementation would be based on the specific business model of the content provider.

We believe that storage-based networks provide both a near and far term solution the problem of delivering personalized content. Overall, the system is more scalable and more capable of providing a broad range of services than a network-based solution. Most importantly, consumers have already begun to adopt both the technology and the behavior patterns. This removes a significant barrier to customer adoption of the PVR as an on-demand solution.

The last significant hurdle is to overcome the mindset of the content providers. Recent changes in technology and cost have presented us with solutions that were unthinkable only a handful of years ago. Content providers should take this time to reexamine the options and make new decisions based on the current state of technology.

REFERENCES

1. TiVo Website (<http://www.tivo.com>)
2. Brown, K., 2001. AT&T, TiVo Record a DVR Deal. *BroadbandWeek*, 11/19/2001.
3. TiVo Press Release (<http://www.tivo.com/5.3.1.1.asp?article=131>)
4. DirecTV website (<http://www.directv.com>)
5. DirecTV Program Guide (<http://www.directv.com/DTVAPP/ProgramGuide>)
6. Dempski, K., 2002. Real Time Television Content Platform. Workshop on Personalization in Future TV 2002.
7. Maddox, G., 2001 The Video Industry (<http://www.afc.gov.au/GTP/wvanalysis.html>)