TRANSPORT, CONTENT, AND SERVICE IMPLICATIONS ON VOD NETWORK TOPOLOGY

George Kajos, Vice President of Engineering
Conrad Clemson, Sr. VP, Technical Operations
Broadbus Technologies

Abstract

Video on Demand (VOD) is evolving and growing rapidly. As a result, transport, content, and service offerings are changing the fundamental economics and operational efficiency of Video on Demand networks. On Demand services of a few premium movies and HBO On Demand are giving way to the NFL on Demand, Nickelodeon on Demand, Everything on Demand. Content distribution networks comprised of a small array of catchers directly connected to the back end of a small video server farm are evolving into full fledged propagation services and hierarchical storage models. The streaming network has evolved from a network with streaming servers directly connected to ASI ports to a GbE based transport network. Services are expanding from traditional transactional, free, and subscription services to include a new variety of On Demand service offerings. Everything from reality TV, to advertising, to personal ads is becoming available on line. With each of the elements of the On Demand system dissected into its pieces, the paper will put these elements together in a single cohesive view of optimal On Demand network topologies based on the evolution of transport, content and service type.

OVERVIEW

The VOD environment is clearly evolving in multiple simultaneous dimensions. Within this environment, there are several competing architectures for appropriate VOD server deployments. Some architectures propose decentralized VOD server deployments; others propose centralized server deployments, finally some compromise with a hybrid approach.

Initially, content usage data from a small VOD installation is examined. Then, this paper evaluates the effect of content placement on the transport network and the economics of a complete VOD solution as a function of centralizing vs. decentralizing servers. In the face of dramatically decreasing transport costs, the paper identifies the few scenarios in which edge caching may be an effective approach to certain VOD applications. The paper also examines content propagation and replication. Finally, this paper examines how the evolution of new services, some of which may be personalized to a per subscriber basis, should dictate the placement of both the servers and the content.

SERVICES

The last half decade has validated Video on Demand. Hundreds of VOD deployments have occurred featuring MOD, SVOD, and FOD [CED]. A consistent server design point had previously been 500 – 1,000 streams and 1,000 – 2,000 hours of content. Server clustering allowed installations to support 10,000 – 20,000 streams.

Accordingly, these successes have unleashed the potential demand for a wide range of new services. Examples of new services being trialed or conceived include:

- Music on Demand
- Non-linear Live Broadcast
- Network PVR
- Customized content
- HD Content and Widescreen format for all the of above

It is assumed that Music on Demand will not have a major impact on streaming and storage capacity.

HD content is still evolving. The impact of the current format is a 4X multiplier on streaming bandwidth and storage requirements as MPEG 2 HD content is being transmitted at 15 Mbps. The demand for HD is steadily increasing.

MOD Wide Screen versions appear as just another content. They are typically equivalent in size.

Network PVR and Live Broadcast services have the potential to greatly impact capacity requirements. Quantifying the required bandwidth is straightforward. Service offerings and business rules determine temporary storage requirements. For example, a service offering of 100 SD and 10 HD channels has an ingest requirement in excess of 500 Mbps and a content storage requirement, temporary or permanent, of 250 Gbytes/hour. The selection of content to retain and the duration it is to be made available could vary widely based on the business rules of the offering.

SERVER DESIGN

The continuing advance of technology allows this generation of video servers to break the dependency of streaming capacity on disk bandwidth. First generation video servers relied on streaming from hard drives. Consequently, the number of streams served was a one-to-one relationship with the bandwidth available from hard disks. Extremely sophisticated striping and scheduling techniques were employed to drive up stream counts. Moreover, custom trick files were prepared for fast forward and rewind in order to remove the variance from disk access. In general, this meant a trick file for every fast forward/rewind rate or a single fast forward/rewind rate.

A hard drive based server also posed limitations on the amount of content which could be ingested. Updating hard drives with new content reduced streaming bandwidth while disk writes were scheduled. In an environment with just MOD and SVOD services, content propagation could be scheduled at off hours with little adverse affect. However, as we explore live broadcasts and real time propagation, the required inbound content loading bandwidth goes up considerably.

This generation of video servers is designed around two principles:

- Independent scaling of streaming bandwidth and storage capacity
- Real time content ingest and turnaround performance

Server architecture designers must carefully consider the tradeoffs between the following:

- Processor performance
- Disk I/O bandwidth
- Network bandwidth
- Memory bandwidth
- Backplane/Interconnect I/O bandwidth

SERVER PLACEMENT AND NETWORK DESIGN

Extensive research has been conducted and numerous papers have been written on the topic of video server network topology. Most authors describe the approaches as centralized, decentralized, and hybrids of the two. Looking at the next-generation of server capacity and content library sizes, and taking into account both cable HFC and IP video, it is instructive to generalize the following components:

- Video Server Complex – server technology capable of streaming, local content storage, On Demand and scheduled ingest, and session/stream management.
- Transport network – the network supporting streaming to On Demand clients.
- Propagation network – the network supporting On Demand and scheduled propagation to servers.
- Content Storage Server – a generalized library and central repository for the content made available for On Demand services.
- Preparation Server – receivers of live broadcasts which then encoded for propagation to video servers to On Demand clients for play.
- On demand clients – the media decode and display point at the subscriber. Figures 1 and 2 depict centralized and decentralized VOD environments respectively.

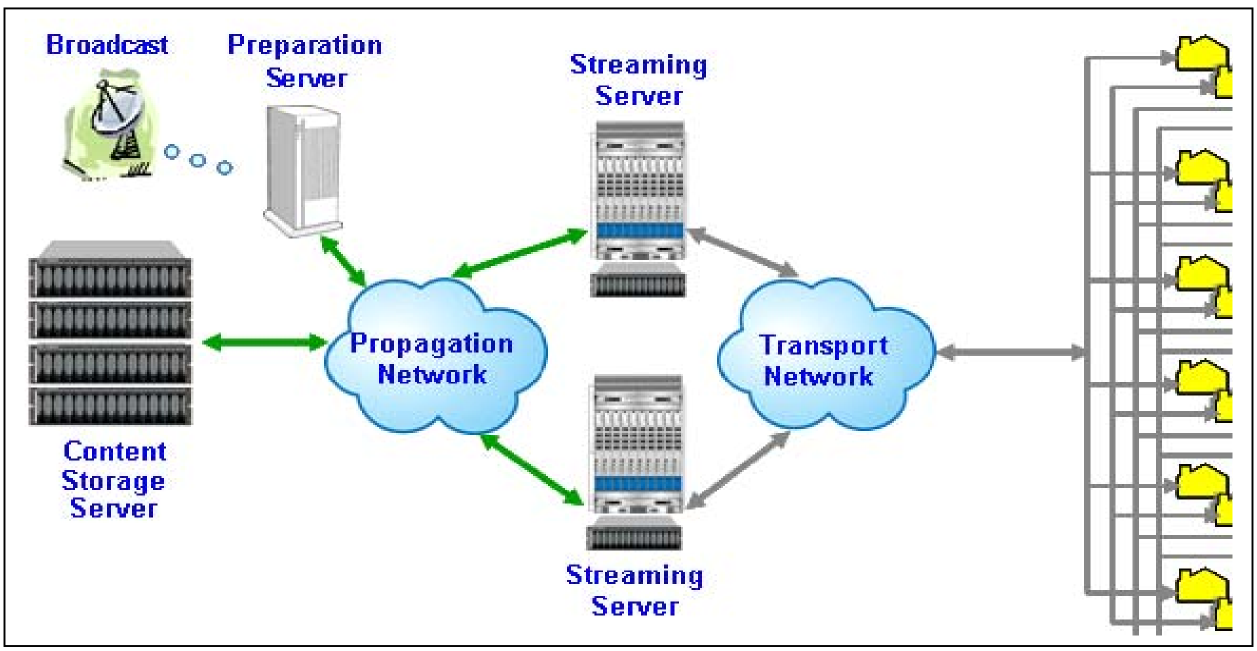

Figure 1: Centralized Server Placement

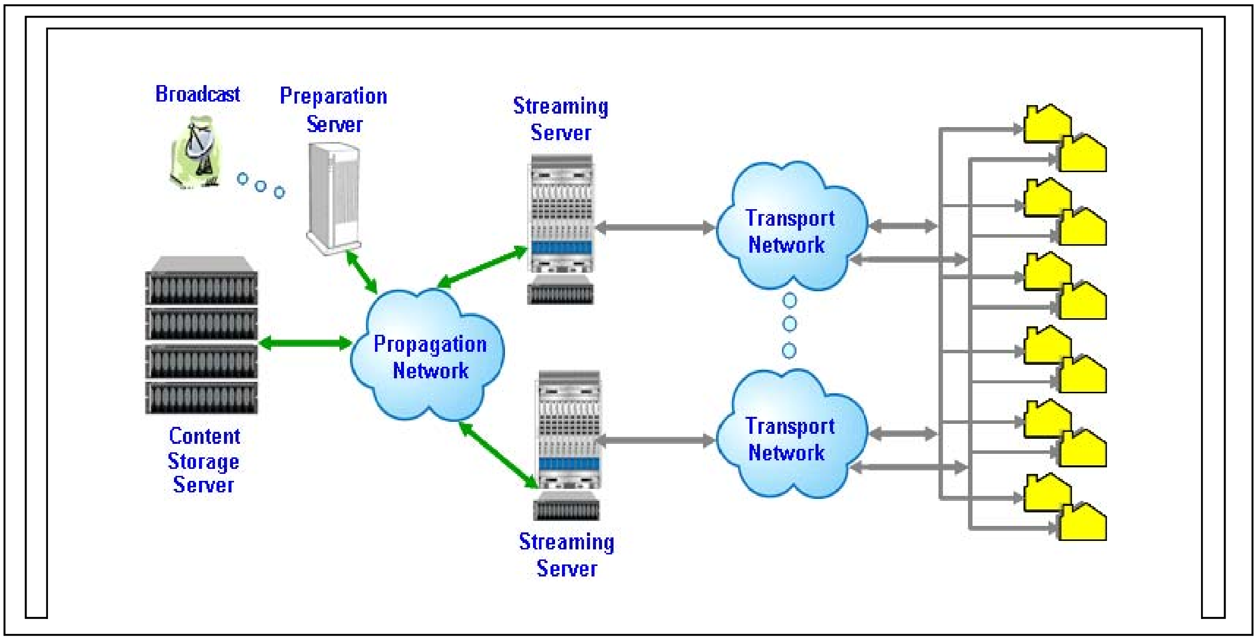

Figure 2: Decentralized Server Placement

The distinctions between these environments follow:

1. The centralized server environment has complete connectivity to the entire population of supported On Demand clients over the centralized transport network.
2. The centralized server environment receives only scheduled content and little or no On Demand content over the propagation network.
3. In the centralized environment, it is possible to load balance across the entire population of clients. In the fully decentralized environment, load balancing is restricted to the partitioned subset of clients.
4. The decentralized server environment has connectivity with a partitioned subset of the population of supported On Demand clients over regional transport networks.
5. The decentralized server environment receives both scheduled and On Demand content over the propagation network.

The differences in the two environments are chosen to emphasize the trade-off between the cost of transport network bandwidth and content replication.

The network bandwidth/storage trade-off is not the only consideration between centralized and decentralized approaches. Other considerations include:

- Replication of control components at decentralized sites.
- Operational costs of additional decentralized sites.

In reality, most VOD system designs are hybrid approaches. For example, even in a centralized environment, it is unlikely that every server need to have connectivity to every client. Acceptable load balancing is possible with reduced connectivity. In a decentralized environment, for added reliability it would be advantageous to have more than one server capable of reaching each On Demand client. The system could operate in a degraded mode until system repair completes.

CONTENT PROPAGATION

As discussed above, in first generation video servers, content propagation requirements corresponded to the gradual refresh of new MOD, SVOD, and FOD offerings. These could be loaded onto video server complexes with little affect on server performance during low usage periods.

This generation video server must be designed for two new sources for content propagation:

- Live Broadcast
- On Demand propagation.

This paper uses the term "On Demand propagation" to refer to the requirement to move content to a server which has received a purchase request and does not have the required content. The case arises in server environments where not all content is located on every server. In a centralized server environment, a session manager could direct the request to an appropriate server.

However, in a decentralized server environment, clients are partitioned by servers. It is unreasonable to replicate all content at every site in a decentralized server environment. Consequently, this paper defines the case when content must be transferred to a server to grant a client request. Figure 3 demonstrates how On Demand propagation works. The client requests content not available on the local server or server complex. A request is made to a regional propagation server for the required content. A "filler" content is transmitted to the client at the start of the upload. The filler could consist of previews or advertising. When the server buffers enough content, play out begins.

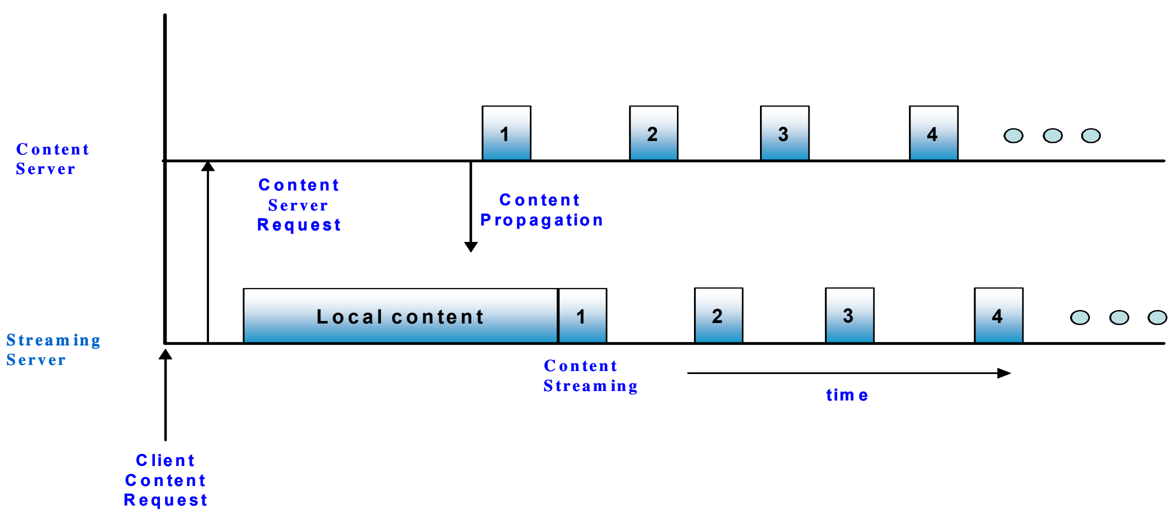

Figure 3: On Demand Content Fulfillment

Live broadcasts and On Demand propagation, move guaranteed quality of service from a server memory or disk subsystem problem to the propagation network. Unlike the dedicated transport network which is most often provisioned for the maximum stream capacity dictated by the HFC QAM capacity or as a percentage (provisioned take rate) of the total pool of clients, the design of the propagation network should include a policy on how to allocate between scheduled, live broadcast and On Demand bandwidth. As discussed earlier, Live Broadcast input bandwidth requirements could range in the 500 Mbps. Figure 4 depicts an example of how the allocation policy could vary during a 24-hour cycle.

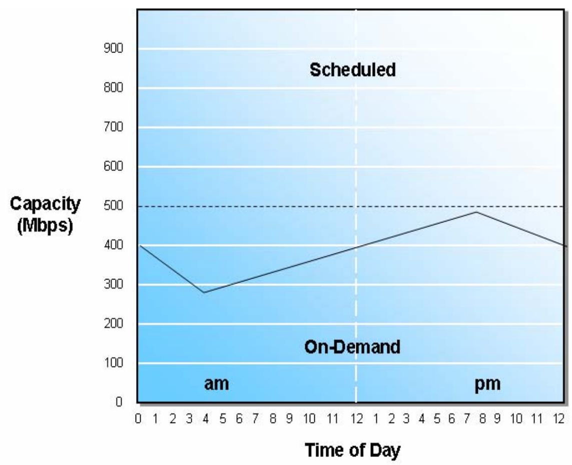

Figure 4: Network Bandwidth Allocation Policy

TRANSPORT AND PROPAGATION NETWORK TECHNOLOGY

The streaming network has evolved from a network with streaming servers directly connected to ASI ports to a GbE based transport network. Ethernet price points have continually fallen due to the ever increasing reach of Internet protocols from traditional LANs to geographically dispersed WANs. While there are alternatives in the WAN, such as Packet over SONET and ATM, as GbE moves to 10Ge, it appears that 10 GbE will become the most cost effective, high bandwidth solution.

Consequently, this paper examines ways to interconnect geographically separated 10Ge pipes. The most straightforward are dark fibers and long reach optics to create point to point links. However, this is an inefficient use of the bandwidth available in the fiber and would only make sense if only a single trunk of 10 Gee is required.

Another approach is wave division multiplexing (WDM). WDM technology allows data from multiple sources to share a single fiber by transmitting on individual wavelengths. WDM interfaces have been incorporated in switches and multiplexers.

Two types of WDM are in use today: Dense WDM and Coarse WDM. DWDM is ideal for high bandwidth, long haul applications. Current DWDM technology can squeeze over 30 channels in C and L optical bands.

CDWM uses lower cost optics and is characterized by wide channel spacing over a wide spectrum. CWDM technology can supply 18 channels from 1270 to 1610 nm.

WDM technology is ideal for accommodating 10 Gee streaming pipes to remote QAM or DSLAM locations

OBSERVED DATA

In this section, VOD and SVOD are examined as a starting point for planning for new services. VOD today is dominated by SD (standard definition) content and is transmitted as MPEG2, 3.8Mbps/stream. It is used for Movies on Demand and "Subscription on Demand" Services.

Figure 5 and 6 depict VOD usage for one day, January 29, 2005, at a relatively small site. The days' totals are 5,133 streams across 703 contents. As one would expect on a Saturday, the maximum number of sessions peaks around 9:30 PM as shown in Figure 6. The distribution shown in Figure 2 plots content usage from most used content (117 plays) to least viewed content (1 play).

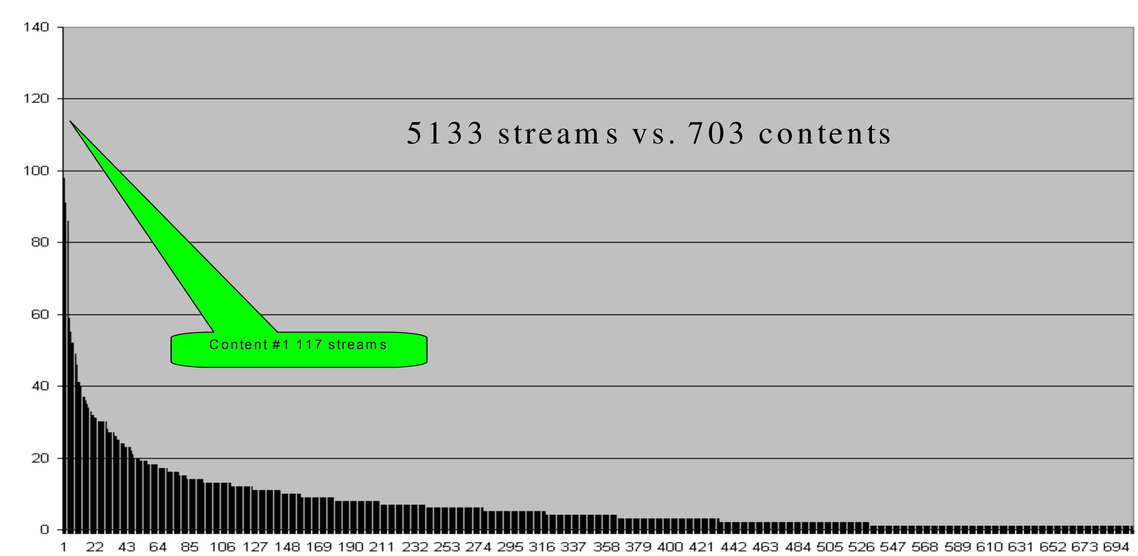

Figure 5: January 29, 2005 – Streams versus Time of Day

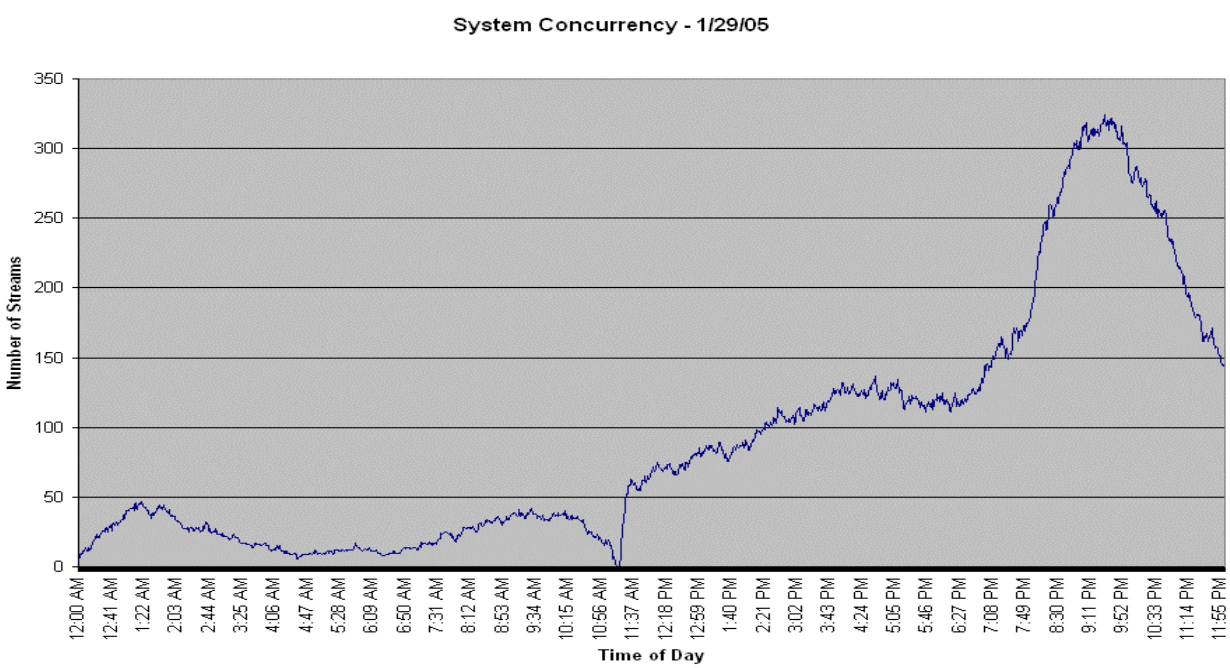

Figure 6: January 25, 2005 – Streams vs. Time of Day

These observations are typical and reported in a number of other works [5]. In fact, the popularity distribution can be fit to Zipf's law, which states that the probability of requesting a program m, where m = 1, 2, 3… out of N movies is : C/m where $C=(1+1/2+1/3+\ldots+1/N)$

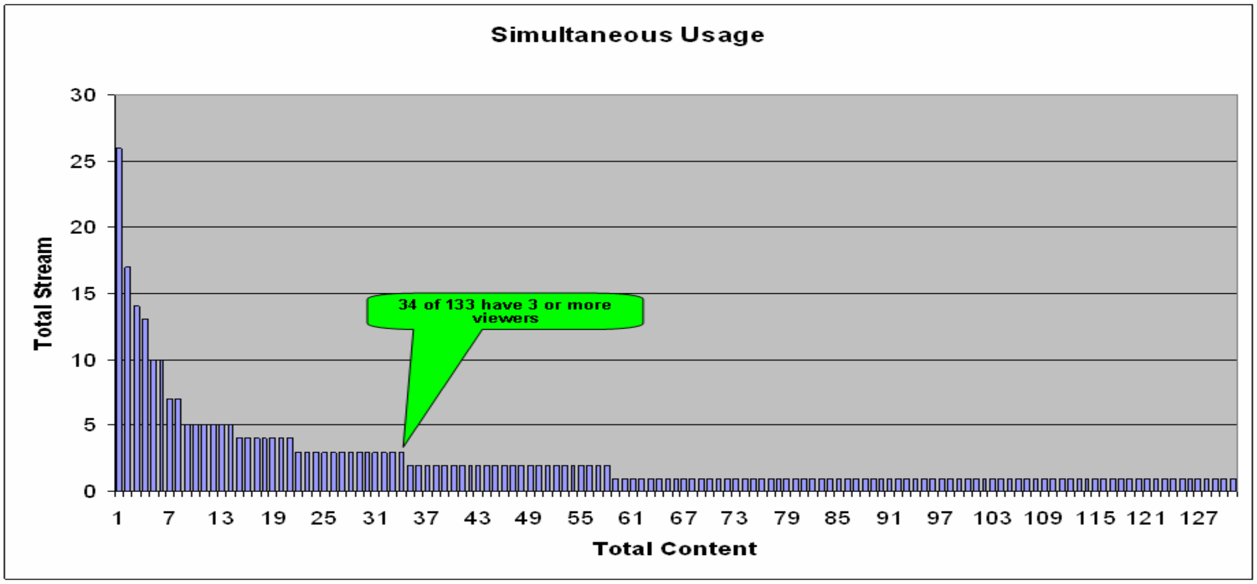

Figure 7: Simultaneous Content Viewing at Peak Usage

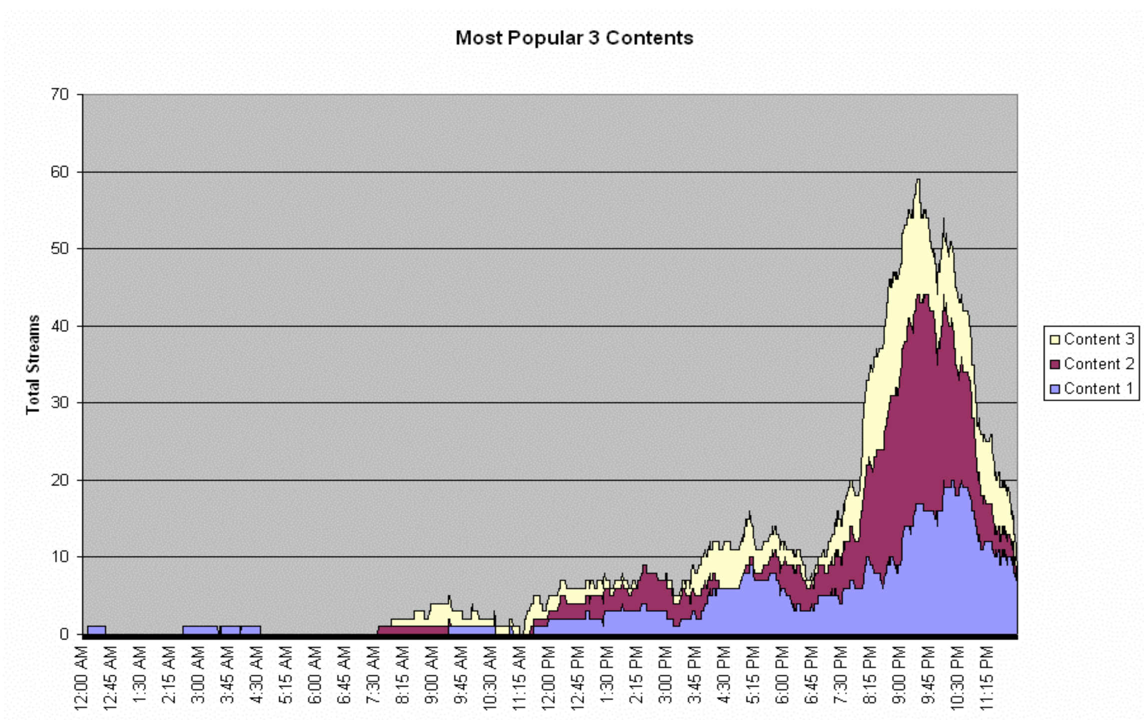

Figure 8: January 29, 2005 – Most Popular 3 Contents

Figures 7 and 8 examine the relationship between simultaneous sessions and content. In Figure 7, it can be seen that 58 of the 133 contents are being viewed by two or more set top boxes. Figure 8, depicts that at the peak of viewing, the top three contents total 57 of the 322 sessions. As video servers are scaled, these relationships will be used to assist with the tradeoff between memory, network, and I/O bandwidth.

Finally, additionally data obtained from discussions with a number of MSO's with mature VOD deployments suggests that actual VOD usage data varies widely from service group to service group. While it is not possible at this time to publicly share specifics of these results, it is generally understood in the industry, that within a system, peak usage may vary from 2% to 14% on a service group by service group basis.

This data proposes some interesting conclusions.

First, real world data validates a Zipf distribution model across multiple deployments. The Zipf curves favor caching architectures, in general, and specifically making caching architectures more effective as the stream count grows. This is the case, because in a Zipf curve, the tail is relatively constant, while the peak of the curve grows dramatically with stream count. As centralization takes place, VOD servers that accommodate extremely high concurrency, such as caching servers, will serve more streams at a lower cost.

Additionally, the tail stays relatively constant. In hybrid architectures, it is generally assumed that less popular titles are streamed from the core to conserve the replication of storage. This data supports that model, but it also points out that a library server which has a caching capability, can perform both tasks.

Some server designers would argue that it makes sense to cache the popular titles at the edge to conserve transport costs. While this approach has academic appeal, it does not hold up to real world scrutiny. The problem with this approach is that it assumes a uniform distribution of concurrency in each service group. The data suggests that each service group has its own peak concurrency. These server designers would propose to either provision the entire system to the average concurrency or, worse, the peak concurrency. Provisioning to the average concurrency will result in denial of service to the peak service groups and over provision the low concurrency groups. Provisioning to the peak concurrency will result in massively over provisioning the entire system.

When transport costs are equal to or less expensive than the streaming server costs, as they are today, the only logical way to provision a VOD system is to centralize the architecture. This provides the operator with tremendous economies of scale in the streaming subsystems. At the same time it allows the operator to provision across all service groups without stranding streams at the edge of the network.

NEXT GENERATION VOD ARCHITECTURE

In this section, hypothetical VOD system is created for the purpose of further exploring centralized approaches and decentralized approaches. Assume that an environment is provisioned for a take rate of 300,000 active clients. There are two server capacities available – 3,000 and 15,000 streams. In this exercise, the paper examines the number of servers and the equivalent number of gigabit Ethernet links required at each extreme. One additional consideration is the possibility that a decentralized server will need double the ingest bandwidth of a centralized server to accommodate On Demand propagation.

Table 1: Servers and Transport Links

300,000 streams	**Streams per server**	**Servers**	**GbE Transport Links**	**10GbE Transport Equivalents**	**Propagation GbEs**
Centralized	15,000	20	60	6	20
Decentralized	3,000	100	12		100/(200)

Some general observations can be made:

- More servers, storage, control systems, in decentralized model
- Ability to collapse ten 1GbEs into one 10GbEs in the centralized model
- Potential additional load on the propagation network in distributed model
- Higher cost transport in centralized model
- Greater potential to share storage in centralized model

Figure 9 depicts an architecture which collapses the transport network connectivity.

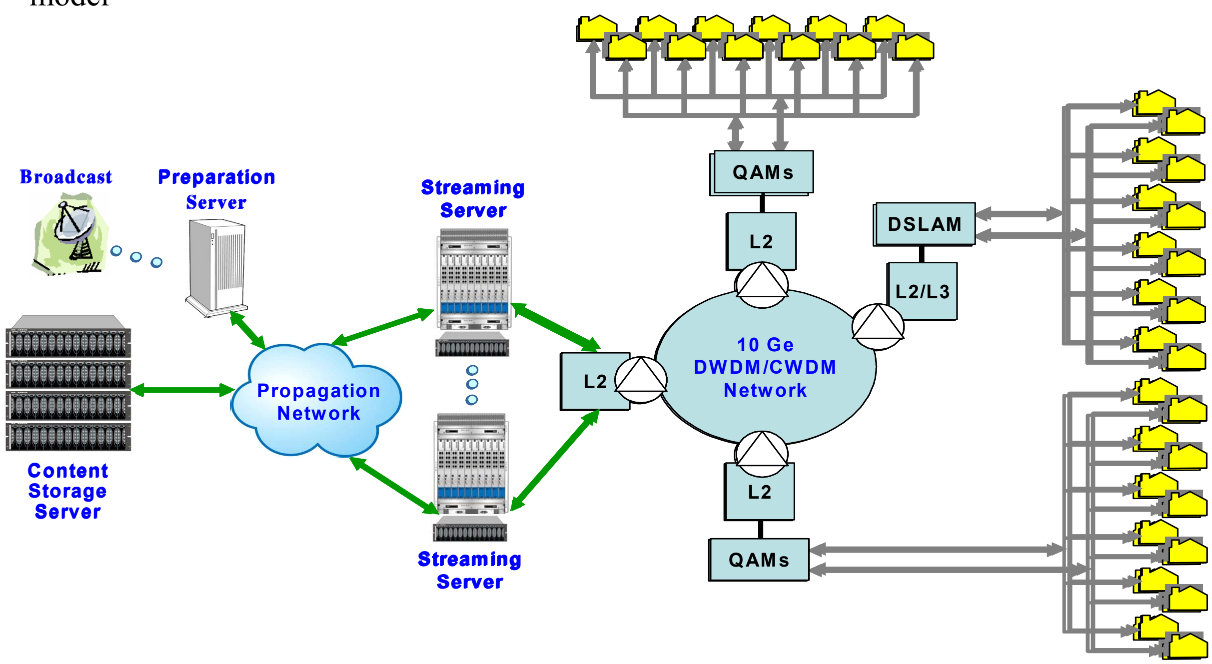

Figure 9: Next-Generation Collapsed VOD Network Architecture

The transport network is represented as a WDM optical network connected through layer 2 switches. Because 10Ge pipes are steered to specific Lambdas and partition the client space, the architecture is not a fully centralized environment. However, most of the advantages of the centralized environment are realized.

SUMMARY

During the previous half decade, many VOD installations were monolithic and self contained for ingest and streaming. Advances in server and transport technology allow new services to be considered for VOD deployments. In this paper, data regarding content popularity during a single day and at peak load was presented from one VOD installation. The next-generation VOD architecture presented in the paper is well suited to meet the scale to the requirements dictated by new services.

ACKNOWLEDGEMENT

The authors wish to acknowledge Michael Kahn for extracting and processing the January 29, 2005 data.

AUTHOR'S CONTACT INFORMATION

George Kajos
VP, Engineering
Broadbus Technologies, Inc.
80 Central Street
Boxborough, MA 01719
gkajos@broadbus.com
Office: 978.264.7906

Conrad Clemson
Sr. VP, Technical Operations
Broadbus Technologies, Inc.
80 Central Street
Boxborough, MA 01719
cclemson@broadbus.com
Office: 978.264.7905

REFERENCES

1. Skarica, Christopher, and Joe Selvage, "Environmentally Hardened optical CWDM Solution for MSO Access Network Capacity Expansion," Cable-Tec Expo, June 15-18, 2004.
2. Kajos, George, "VOD Storage Architecture for Hierarchical Network Content Libraries," Cable-Tec Expo, June 15-18, 2004.
3. Jancola, Mark, "Considerations in Scaling Video-on-Demand Systems," Cable-Tec Expo, June 15-18, 2004.
4. "North American Cable Deployments," CED Magazine, December 2004.
5. Vassiliakis, Constantine, Michael Peterakis, Peter Triantafillou, " Video Placement and Configuration of Distributed Video Servers on Cable TV Networks," Multimedia Systems, 8:92-104, 2000.