Traffic Shaping, Bandwidth Allocation, and Quality Assessment for MPEG Video Distribution over Broadband Networks

Steven Gringeri, Khaled Shuaib, Roman Egorov, Arianne Lewis, Bhumip Khasnabish, and Bert Basch
GTE Laboratories Incorporated

Abstract

This article provides an overview of residential video delivery systems and presents the applications, benefits, and challenges of using VBR MPEG video encoding in broadband video distribution networks. The network resources required to transmit stored variable-rate MPEG can be reduced by properly analyzing and smoothing the video stream before transmission. A scheduling technique is presented which selects a traffic contract for a pre-encoded MPEG video stream with the criteria of minimizing network resources and maintaining video quality. Several effective bandwidth metrics are discussed and used to model the potential savings in network resources for the shaped streams.

Telecommunications providers have started to offer broadband services such as broadcast video, high-speed Internet access, and videoconferencing. Broadband network architectures such as fiber to the curb (FTTC) coupled with digital subscriber line (xDSL) modem technologies [1, 2] are extending the fiber plant closer to the end user and facilitating the migration of narrowband networks to broadband networks. These architectures use asynchronous transfer mode (ATM) technology as the underlying transport protocol. ATM uses small fixed-size (53 bytes) cells to allow multiplexing of various services such as voice, video, and data with guaranteed cell rate, cell loss, and cell delay variation parameters. These capabilities make ATM well suited for real-time applications such as multimedia applications.

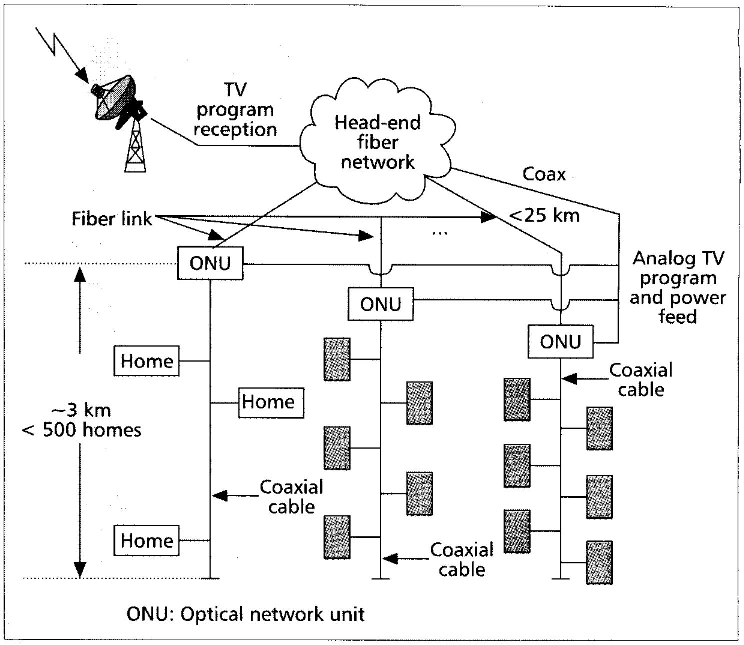
The Moving Picture Experts Group (MPEG) standards [3, 4] have been widely adopted for providing digital video services. The MPEG standards do not prescribe the encoding process; instead, they specify the data input format for the decoder (the syntax) as well as detailed specifications for interpreting this data (the decoding semantics). An encoder must follow a set of steps known as the encoding process to compress video data. This process is not standardized, and may vary from application to application depending on a particular application's requirements and complexity limitations. This allows the encoding process to be optimized for an application's bandwidth and quality requirements. Today, commercially available encoders support a wide range of transmission and storage applications using both constant bit rate (CBR) and variable bit rate (VBR) encoding.

Transmission of real-time video streams is resource-intensive even when the video is compressed using sophisticated algorithms like MPEG. As digital video systems are deployed, efficient utilization of storage and network resources will be needed to reduce costs and increase revenues. Variable-rate MPEG encoding has the potential to save bandwidth and increase channel capacity while reducing system cost and server storage requirements [5–7]. Compressed digital video such as MPEG is inherently variable-rate since its complexity and motion content affect the encoding bit rate required to maintain picture quality. ATM networks have the flexibility to support both constant- and variable-rate services, and can provide statistical gains when using variable-rate MPEG. However, uncontrolled burstiness will lead to inefficient use of network resources by occasionally requiring excessively high processing, storage (buffering), and transmission capacity from the network. These requirements motivated the early implementers and providers of digital video services to use a CBR channel for delivering real-time video to customers. Digital video encoder manufacturers, accordingly, implemented additional rate-control and buffering in the encoder to generate a

CBR stream for transmission and distribution applications. The fullness of the rate-control buffer dynamically controls the quantization resolution so that the number of bits generated per picture satisfies the bit rate constraints of the video stream. These rate adaptation methods not only lead to variable video quality, but also may result in poor utilization of bandwidth since the rate must be selected to accommodate the most complex scenes, and the encoder may need to use stuffing bits to maintain the CBR.

VBR encoding can achieve improved coding efficiency by better matching the encoding rate to the video complexity. Variable-rate encoding is currently used in storage applications such as digital versatile discs (DVDs) to achieve significant storage savings. Using VBR MPEG, it should be possible to achieve substantial savings in networking resources while maintaining video quality, if the burstiness of the video can be controlled using some prespecified constraints. In addition, it is possible to offer a more constant video quality than can be provided using constant-rate encoding. Savings in storage resources such as server capacity allow more movies or video sequences to be stored using the same capacity. Similarly, savings in networking resources imply that the same amount of switching, buffering, and transmission capacity can be used to deliver more video content to the customer while maintaining the application-level quality of service (QoS) requirements. Either source coding or output shaping or a combination of the two can be used to adapt MPEG video streams for transmission over VBR ATM channels. To achieve such savings, it is imperative that we control the burstiness and bitrate variability of a single stream or a set of multiplexed streams in order to make the stream(s) adhere to a traffic contract while maintaining the desired video quality. The bandwidth saving for transmission applications may be smaller than for storage applications due to additional restrictions the network places on the video's burstiness.

This article addresses issues in encoding and distributing variable-rate MPEG video over broadband networks. The primary objective is to realize bandwidth savings while maintaining video quality. The issues related to shaping MPEG video to control its burstiness and produce streams that are more suitable for transmission over switched ATM networks are discussed. An overview of residential video delivery systems is also included to provide the architectural information about these systems relevant to implementing variable-rate MPEG. The relative quality of the various CBR and VBR video streams is compared using both subjective viewing and quantitative measurements. The goal of the comparison is to ensure that video quality is being maintained for variable-rate encoding. An analysis was conducted to determine the required ATM network resources and traffic contracts for various video categories (action movie, sports, talking head, etc.) when MPEG video is transported over an ATM network using real-time VBR service. Potential savings are analyzed for single streams and groups of streams to determine whether variable rate MPEG streams have traffic contracts consistent with the capabilities of today's ATM switches. The bandwidth savings are quantified using several effective bandwidth metrics [8–11]. Our results to date are based primarily on shaping existing streams before they enter the network to reduce their burstiness and thus their utilization of network resources. This technique is well suited to MPEG streams that have already been encoded, although encoding variable-rate streams specifically for transmission applications could further reduce the required network resources. These rate control techniques for encoding "ATM-friendly" streams have been discussed extensively in the literature [5, 6, 12–16] but to date are not implemented in real-time transmission systems. Our results indicate that using variable-rate MPEG in transmission applications such as residential video distribution systems has the potential to decrease bandwidth requirements and improve channel capacity.



■ Figure 1. *The architecture of hybrid fiber coax networks.*

An Overview of Residential Video Distribution Systems

Traditional video service providers deliver analog video to residential customers using cable TV networks or direct broadcast satellite (DBS) technology, mainly for entertainment purposes. Digital representation of video opens up the possibility of computerized processing of multimedia information, allowing consumers to archive, index, and retrieve programs in a content-based manner. Digitally encoded information can be made more resilient to degradation during transmission, distribution, and storage, giving the consumer a better picture and sound quality. Digital transmission of audio-visual information also enables multimedia communication over the same network that supports voice and data communication using xDSL technologies [1, 2, 17]. In the next two paragraphs, related issues from the point of view of network and transport architectures are discussed. Details of the various architectures can be found in [2]. We then briefly describe the dominant transmission protocols followed by some recommendations on the usage of the various architectures and protocols for video delivery.

Most of the traditional residential video distribution systems usually deliver one-way analog video using either full coaxial cable (CATV) or hybrid fiber-coax (HFC) networks [1, 2]. A typical HFC network is shown in Fig. 1. These networks offer only limited capability for real-time interaction with the service provider's facilities for content selection. For wireless delivery of video, cost-effective small dish antennae for receiving DBS signals have recently appeared on the market. DBS networks use geostationary satellites which orbit the

Technology	Frequency band (up/down, GHz)	Bandwidth (MHz)	Channel capacity (each analog TV channel is 6 MHz)	Area coverage
Direct broadcast satellite (DBS)	Ku-band 11–18	500 downstream (11.7–12.2 GHz) 500 upstream (14.0–14.5 GHz)	150 digital channels	1000s of km radius, using ~50 cm. home dish antenna
Local multipoint distribution system (LMDS)	27.5–29.5	850 downstream (27.5–28.35 GHz) 150 upstream (29.2–29.35 GHz)	42 analog channels	~5 km radius, using ~30 cm. home dish antenna
Multichannel multipoint distribution system (MMDS)	2.15–2.70	198 downstream 4 upstream	33 analog channels (120 or more digital Channels)	~50 km radius, using a ~60 cm home dish antenna

■ Table 1. *Video distribution using wireless transmission technologies.*

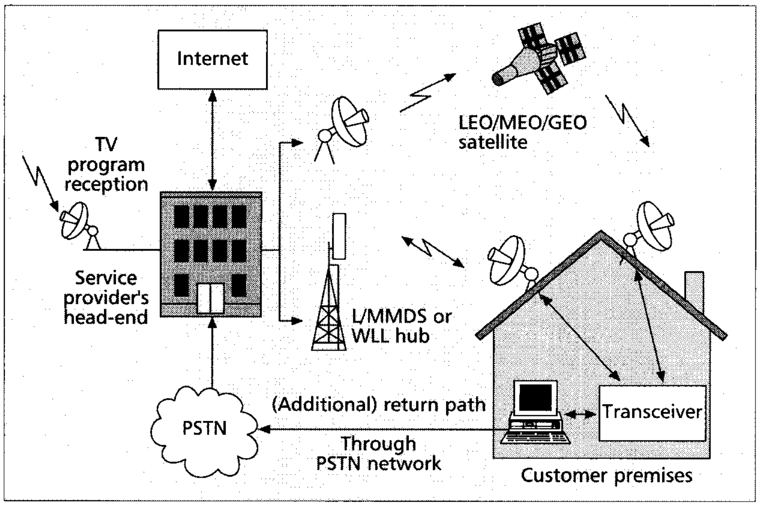
Earth 23,500 mi above the surface. Although they were originally intended for delivery of analog/digital video to customer premises, it is possible to utilize these systems for interactive delivery of digital video and Internet information. Figure 2 presents a high-level description of such architecture. To support interactivity, this architecture uses the public switched telephone network (PSTN) as the return path from the customer premises. All of the above systems are based on shared media architectures. These architectures exhibit a number of problems, including maintaining secure and reliable operation as well as providing customized services for individual users. In addition, these architectures cannot easily support the necessary infrastructure for delivering integrated telephone, data, and video services.

Traditional CATV networks are designed to offer low-cost unidirectional bandwidth from 550 MHz to 1 GHz to provide analog TV services over coaxial cable. Tree-and-branch-type architectures with cascaded unidirectional amplifiers to improve signal strength and reliability are commonly used in such networks. Existing CATV networks are being upgraded to an HFC architecture, as shown in Fig. 1. This architecture uses an all-optical backbone network to interconnect the head-ends. From the head-end, fiber links run to regional distribution centers which provide opto-electric conversion. Coaxial cable is used in the local feeder loop from the distribution center to the home. Traditional analog broadcast services occupy the spectrum between 50 MHz and 550 MHz for about 80 channels, where each channel uses 6 MHz of bandwidth. The 6 MHz analog channel can carry digital information using different modulation techniques. For example, with 256 quadrature amplitude modulation (256-QAM) or 16 vestigial sideband (16-VSB) modulation, a 6 MHz channel could deliver approximately 38 Mb/s data (the remaining bits are used for error correction and control) to the home [1, 2]. If a CBR of 6 Mb/s is used for digital video, a maximum of six video streams can be accommodated in one analog channel. Recent research [5, 9] in the area of bandwidth allocation and transmission of video shows how VBR encoding of video can increase the channel multiplexing capability while maintaining the video quality desired by the application.

Another option for video distribution networks is to use wireless transmission to the home. Two possible techniques for offering high-quality digital video distribution and fast Internet access using low-power, high-frequency radio signals over short to medium distance are the local multipoint distribution system (LMDS) and multichannel MDS (MMDS), as described in Table 1 and shown in Fig. 2. Investments in these systems are incremental in the sense that they can grow as new customers are added to the system.

Digital multimedia services require significantly more bandwidth than traditional voice services. One way to achieve this is to use xDSL modem technologies [1, 2] which enable high-bit-rate data transmission using the existing copper plant (subscriber loop). It is also possible to use this technique for digital video delivery services. Asymmetric DSL (ADSL) can support a bit rate of 6.312 Mb/s (the maximum over 3000 ft is about 9 Mb/s [18]) from the central office (CO) to the user, and around 640 kb/s from the user to the CO over a distance of less than 12,000 ft. Very-high-speed DSL (VDSL) can deliver data at OC-1 rate (51.84 Mb/s) to the home using twisted pair telephone lines over a distance of less than 1000 ft, and the data rate could be 25.92 Mb/s over a distance of 3000 ft. Table 2 shows the maximum ADSL, rate-adaptive DSL (RDSL), and VDSL data rates under ideal line conditions. Data rates depend on a number of factors, including the length of the copper line, its wire gauge, the presence of bridged taps, and cross-coupled interference [17].

Broadband networking architectures such as fiber to the node (FTTN) or FTTC extend fiber from the CO all the way to the home or to some intermediate point such as the node or curb. FTTC or FTTN networks coupled with xDSL for the "last mile" to the customer premises can extend the capabilities of broadband networks (fiber plants) closer to the end user without requiring the initial investments of fiber to the home (FTTH), and thereby can facilitate the migration of narrowband networks to broadband networks [1, 2]. For delivering



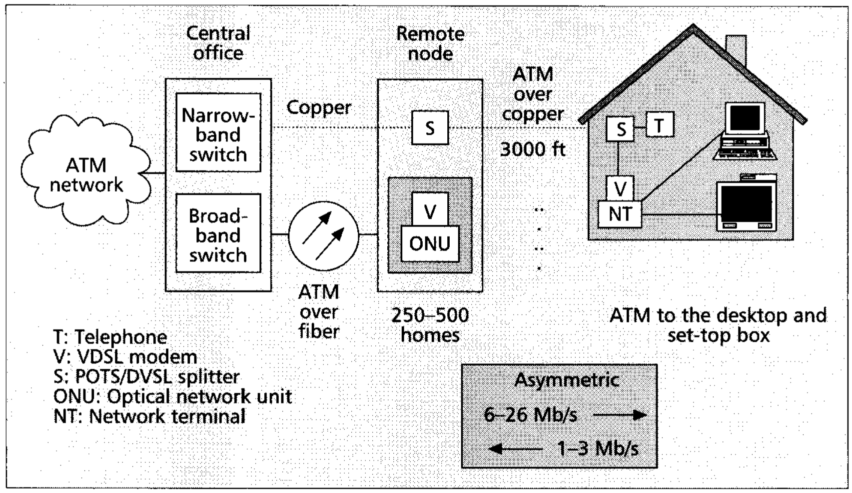
■ Figure 2. *Video delivery using wireless distribution hubs and DBS.*

xDSL	Data rates (downstream/upstream)							
ADSL		9 Mb/s 1 Mb/s		8.448 Mb/s 1 Mb/s	7 Mb/s 1 Mb/s	6.312 Mb/s 640 kb/s	2.048 Mb/s 64 kb/s	1.544 Mb/s 16–64 kb/s
RDSL		12 Mb/s 1 Mb/s		8.448 Mb/s 1 Mb/s	7 Mb/s 1 Mb/s	6 Mb/s 1 Mb/s		1 Mb/s 128 kb/s
VDSL	51.84 Mb/s 2.3 Mb/s	25.92 Mb/s 2.3 Mb/s	12.96 Mb/s 1.6 Mb/s					
Cable length (ft)	1000	3000	4000	5000	9000	12,000	16,000	18,000

■ Table 2. *Theoretical data rates for ADSL, RDSL, and VDSL.*

switched digital video, local exchange carriers are currently investigating FTTC or FTTN architectures (Fig. 3), which can be extended to FTTH (Fig. 4) networks in the future. FTTH is the ultimate wireline architecture for broadband services to the home; the raw bandwidth could be OC-3 (155 Mb/s) or higher.

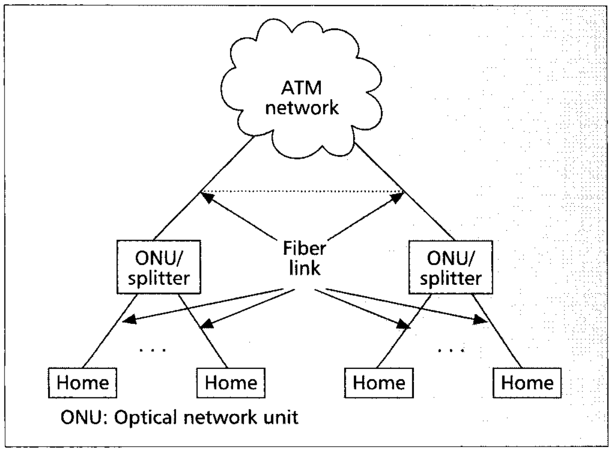
As video distribution systems change from analog to digital, or new digital systems are deployed, efficient utilization of storage and network resources will be needed to improve the capacity and cost effectiveness of these systems. One way to achieve this would be for video broadcast systems to support variable-rate MPEG encoding and transmission in addition to the traditional CBR. Issues related to the encoding and transmission of MPEG video at constant and variable bit rates are discussed next.



■ Figure 3. *Video delivery over ATM/VDSL using FTTN/FTTC networks.*

Constant and Variable Bit Rate MPEG Video Encoding

Through a combination of spatial and temporal compression techniques, analog video can be significantly compressed while preserving image quality using the MPEG [3, 4, 19] compression standards. The spatial compression techniques are the same as those used by JPEG and include discrete cosine transform (DCT), quantization, and entropy coding. The temporal compression techniques rely on block-based motion compensation to reduce the temporal redundancy. In MPEG, three main picture types are defined: intra-coded (I-picture), predictively coded (P-picture), and bidirectionally coded (B-picture). I-pictures are self-contained and are coded without reference to other pictures. They are used as reference frames during the encoding of other picture types and are coded with only moderate compression. P-pictures are coded more efficiently using motion-compensated prediction from a past I- or P-picture. B-pictures provide the highest degree of compression, and require both past and future reference pictures for motion compensation. Several frames are grouped together in a pattern to form a group of pictures (GOP).



■ Figure 4. *The architecture of FTTH networks.*

Digital video is bursty in nature [20], and the burstiness depends on the frequency of changes in the background and the movement of objects in the foreground. Without rate control the output bitstream of a video encoder will be VBR since it depends on the complexity of the scene, degree of motion, and frequency of scene changes. Digital video encoder manufacturers accordingly implemented additional rate control in the encoder to generate a CBR stream for transmission and distribution applications. The rate control algorithm of the MPEG-2 encoder dictates the quantization resolution, which directly controls the number of bits used to encode a picture. By using an output buffer to smooth the bit rate variation, the encoder can maintain a CBR at its output somewhat independent of the number of bits needed to code individual pictures. The rate control algorithm uses the fullness of the buffer as a limit on how large particular frames can be without causing overflow (i.e., it uses source rate control to limit the number of bits entering the output buffer). This is achieved by using coarse quantization to generate a smaller number of bits per picture when the buffer is almost full, or selecting a finer quantization when the buffer is almost empty. These rate adaptation methods not only lead to variable video quality, but occasionally may cause poor utilization of bandwidth because the encoder may need to use stuffing bits to maintain the CBR.

VBR encoding of MPEG video can be achieved using single-pass (real-time) or multipass techniques. In both cases the available options for producing a VBR video stream include source rate control in the MPEG-2 encoder, shaping the output bitstream of the encoder, and rate control which uses feedback from the network. Source rate control, output shaping, or a combination of both will be required to limit the burstiness of a video stream. Network-feedback-based rate control will be needed if the network has to provide a guaranteed QoS that requires some knowledge of the traffic sources' behavior. The video quality depends not only on the rate control and smoothing method selected, but also on the video characteristics and category (sports, news, movie) of video being encoded. The challenge is to find the optimal rate shaping and smoothing strategy that maintains acceptable video quality without putting an excessive burden on the network.

MPEG-2 video encoding offers a rich array of possible methods for source rate control. A fundamental parameter is the quantizer scale since it controls the instantaneous bit rate used at the macroblock level. The quantizer scale is used to control the resolution of the DCT coefficients. The quantizer scale can be set globally at the picture layer, and the desired rate can be achieved by varying the scale for different types of pictures. The quantizer scale can also be adjusted at the slice or macroblock layer to provide finer granularity in rate control. The number and length of the DCT coefficients in a macroblock will depend on the complexity of the video coded and the value of the quantization scale selected. Low-frequency coefficients represent the background of a picture, while high-frequency coefficients represent fine picture detail. During high-motion sequences the changes between frames can be large, and therefore more coefficients are required to code the difference. The picture layer controls the display of video frames, and it is possible to control the bit rate variability of video by setting different thresholds for different frame types in the encoder's rate control buffer.

Other methods exist for varying the bit rate during the encoding process. These methods do not provide the same instantaneous control available by changing the quantizer scale, but instead provide the ability to make large adjustments in the rate with the potential of drastically affecting video quality. The most important technique involves adjusting the frame rate to reduce the bit rate requirements during hard-to-encode scenes. This is accomplished by using the repeat field flags provided in the MPEG syntax [3]. For example, if a scene change has been detected and a large number of bits is required to code that next frame, the previous frame can be repeated to lower the bit rate and make room for the large frame in the buffer. Although this technique is easy to implement, the repeat frame can easily be detected in careful viewing of the video: therefore, this technique should only be used when less drastic techniques based on modifying quantizer scale are not sufficient. The video rate can also be adjusted by modifying the video resolution (number of pixels per picture) or the chrominance sampling resolution [19] (e.g., switching from 4:2:2 video format to 4:2:0), although in practice these techniques are seldom used since they lead to large variations in video quality.

VBR MPEG-2 video can be produced using both real-time and offline coding techniques. In the case of real-time VBR encoding, unless the coding and rate control parameters are selected very carefully, the quality of video may vary significantly over a single video session. The nature of real-time coding limits the bit rate optimization process since information about video complexity is limited to the current and past frames, and no information exists about future frames [20–22]. Offline VBR encoding consists of several steps which include video characterization, rate profile generation, and, finally, video coding. For offline encoding the video characterization and rate profile generation steps are performed over the entire video sequence before the encoding process starts, thus optimizing the bit allocation over the entire sequence. This type of multipass encoding can be implemented in real-time systems by using a two-stage pipelined encoding process. In this configuration the first stage of the pipeline characterizes the video, while the second stage encodes the video based on a rate profile which is generated from the first stage. It is possible to look ahead up to several seconds with this approach. This does not give the flexibility of true multipass encoding but can significantly improve both video quality and encoding efficiency since the pipeline allows up to several seconds of future frame statistics to be available before the current frame is encoded.

Video characterization is the video preprocessing or previewing phase. The objective is to determine the frequency of scene changes, the degree of motion of the foreground and background objects, and the coding complexity of the scenes and various objects in the video sequence. This pass helps decide the GOP structure to be used. It also allows the measurement of various parameters that can be used to generate a bit rate profile for a target video quality. It may be possible to use the information from the profile to determine the conditions for virtually open loop or unconstrained VBR video coding. During the rate profile generation pass, the upper and lower bounds on the amount of bits allocated to various types of frames within the GOP are varied. In each profile attempts are made to allocate a large number of bits while coding complex and high-motion segments of a scene, and fewer bits to relatively still scenes so that the target video quality can be maintained. Several such profiles may be generated during this phase. In the video coding pass, the target traffic contract parameters of the channel can be used to determine the GOP structure and to select one of the bit rate profiles generated in the second pass. The quantization scale factor can now be varied to maintain the required bit rate profile while satisfying the long-term bit rate for the application. This may result in highly granular rate control at the subframe level to satisfy the target quality and rate requirements.

Variable-rate encoding is commonly used in storage applications such as DVDs. In this application, the bit utilization is optimized over the entire encoding session to produce the desired peak and average rates for the stream. The encoder needs to maintain a target video quality while not exceeding the specified peak and long-term average bit rates. The source rate control tries to optimize video quality for the DVD characteristics.

The storage requirements for a movie clip encoded using VBR encoding are much less than those using CBR encoding. When clips of similar video quality are compared, the VBR encoding shows significant storage savings. For example, to code a 15-min video clip using CBR encoding at 5 Mb/s, 562 Mbytes of storage are required, whereas recording the same clip using VBR encoding with a peak rate of 6 Mb/s and a mean rate of 3 Mb/s requires only 355 Mbytes of storage. This translates into almost 37 percent savings in disk space for storing clips of very similar video quality (quality assessment of CBR and VBR encoded video clips is presented in the next section). Additional savings can be achieved, depending on the selected

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase®
Smarter legal research.