June 2005

MOTOROLA

MPEG-4 Over IP Networks

A New Distribution Paradigm Emerges

WHITE PAPER

INTRODUCTION

The recent standardization of the MPEG-4 video compression format and the rapid proliferation of IP networks will present content distributors with new opportunities and challenges. The MPEG-4 compression standard includes new tools that promise to reduce the bandwidth necessary to distribute "DVD quality" standard definition and high definition content by at least 50%. Tests have demonstrated this and even greater efficiency is likely in the future. In addition, server technology, IP network expansion, and standardized protocols for file movement provide the infrastructure to transport the content efficiently – using high-volume, off-the-shelf IP networking equipment. To round out the technology set necessary to construct viable distribution networks, IP rights management is maturing to the point that studios are ready to entrust their valuable content to this new topology.

New video on demand delivery solutions and broadcast models are already in place and being tested for comparison with current distribution performance. "Greenfield" networks are seriously considering using MPEG-4 and IP network distribution in new deployments because of the simplicity, proven connectivity, and ready availability of a wide variety of reliable networking and high performance equipment.

This paper examines the components of the MPEG-4 standard that make it suitable for this application and presents examples of the potential bandwidth reduction gains. All of the required network components will be presented in an end-to-end architecture to illustrate the practicality of near term deployments. Last, the important techniques associated with IP Rights Management will be discussed, along with in-place facilities for security key generation, key storage and key distribution.

CONVERGENCE

Although "convergence" (PC technology and television technology) has not materialized in a significant way in the living room, it's rapidly being realized in the distribution network. Video compression technology is ready to take a giant leap as MPEG-4 emerges to yield excellent quality at 50% of the bitrate of the latest MPEG-2 advancements. This should facilitate movement of video through some ubiquitous, bandwidth constrained networks such as DSL and facilitate large, inexpensive video storage libraries. As an example of "convergence" leverage, readily available 250 Gigabyte disk drives produced for the PC industry (soon to appear in personal video recorders – PVRs) can store over 250 hours of DVD quality, standard definition MPEG-4 video. As demands on computer networks drive the efficiency of routers and switches, these same devices will be able to manage MPEG-4 video streams without the latency problems of earlier generation networks attempting to route real time MPEG-2 content.

MPEG-4

After ten years of progress with MPEG-2, processing technology has finally advanced to the point where another breakthrough in bandwidth efficiency is on the horizon. While MPEG-2 is now widely deployed enabling sub $70 set top terminals and $15 compression chips, it has reached its practical limits in terms of the bandwidth efficiency improvements that we've seen developed over the past several years.

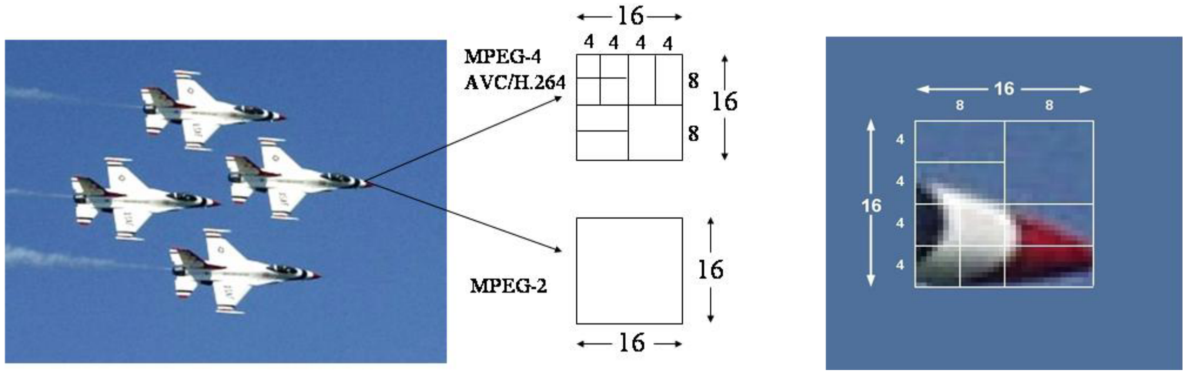
Development of MPEG-4 started shortly after the MPEG-2 standard was published. Work progressed so that by 1999, MPEG-4 ASP (advanced simple profile) was observed to yield an immediate 20% improvement compared to MPEG-2. Four years later, MPEG-4 Part10 (also known as H.264 and/or AVC) was given ISO status and side-by-side comparison tests routinely showed improvements of 50% or higher.

These significant gains in coding efficiency are due to the development of numerous new standardized coding tools that can be applied as needed to the compression goals set by encoder manufacturers. While it is still very expensive to implement all of the MPEG-4 tools in a real time environment, advanced media processors have recently led to the first commercial encoder products. Fast general purpose processors also enable non real-time MPEG-4 encoding for authoring in the MPEG-4 format.

MPEG-4 TOOLS

While MPEG-4 introduces a large number of new tools, two that provide a significant departure from MPEG-2 techniques and provide a major portion of the coding gains are variable sized macroblocks and multiple reference frames. In MPEG-2, macroblocks are fixed at 16 x 16 pixels, so that picture representations can only be broken down as fine as a single macroblock. With MPEG-4, encoding algorithms can apply a variety of sub-macroblock sizes (from 4 x 4 to 16 x 16 – with rectangular and square regions as options) so that each sub-macroblock can be selected to efficiently enclose an area of the picture that contains varying levels of detail or motion. *Figure 1* illustrates the difference between MPEG-2 and MPEG-4 representation of the same picture area and shows the relative accuracy possible with MPEG-4 macroblock configurations – areas with a greater degree of detail are assigned smaller macroblocks than the broader, homogeneous areas.

FIGURE 1



In addition, motion vectors – which track the movement of macroblocks between frames – can be applied to the sub-macroblocks and more accurately track the motion of picture elements.

For motion estimation, MPEG-2 uses three frame types and dictates strict relationships between these frame types. I frames are standalone frames and completely represent one frame of video. P frames are unidirectional predicting frames, which contain information used to "predict" the movement of macroblocks compared to an I frame. B frames are bi-directional frames which predict the location of macroblocks either before or after the current P frame. (See *Figure 2*)

FIGURE 2

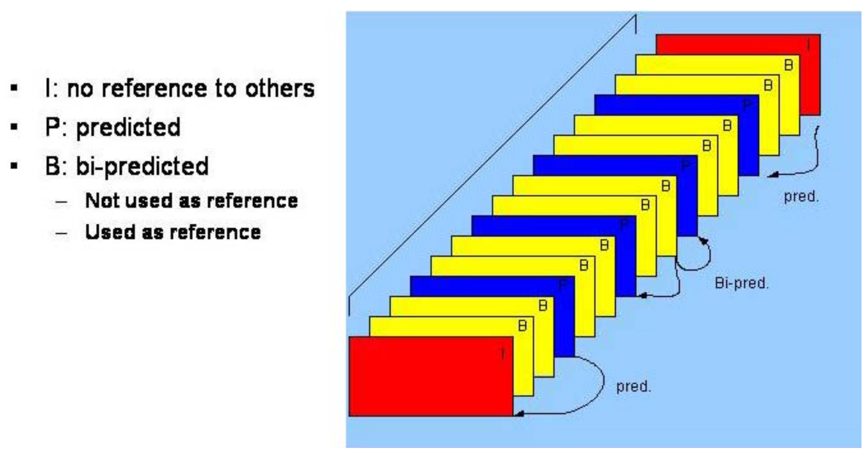
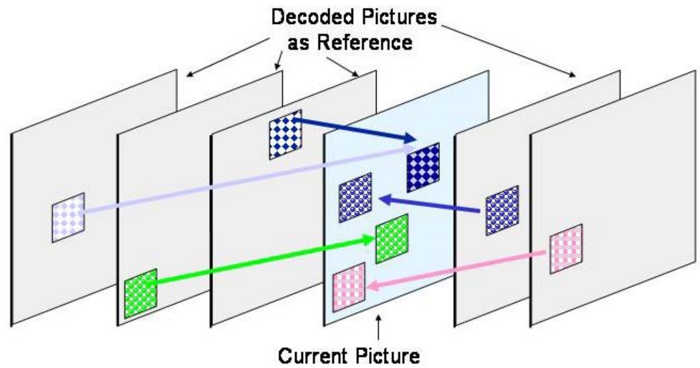


FIGURE 3



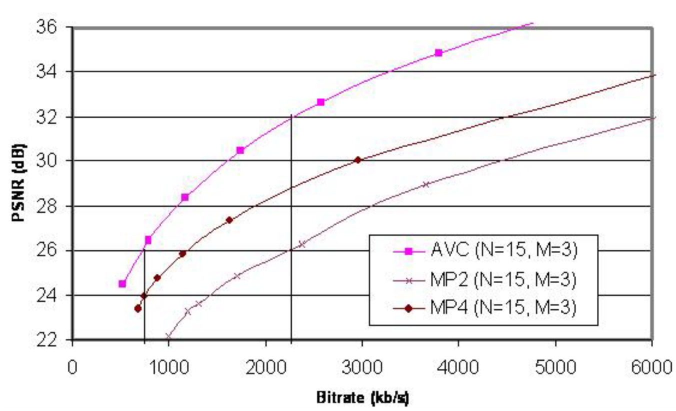
Under this scheme, only two frames are ever considered when building or decoding the current frame. MPEG-4 has now introduced the concept of multiple reference frames – where any frame can reference the data in any other frame – within a five frame sequence. (See *Figure 3*) As an example consider an object (macroblock) that is visible in frame #1, moves behind an object in frame #3, and re-appears in frame #5. Using MPEG-2 techniques, the macroblock would have to be duplicated in frame #5 since reference frames before and after contain no valuable information.

With MPEG-4, the data in frame #1 can be referenced from frame #5, thereby saving the duplication of the object data.

TEST RESULTS

A number of compression tests have been performed using MPEG-2, MPEG-4 ASP, and MPEG-4 part10 (AVC) to validate the coding gains that are likely to be attained using these advanced tools. *Figure 4* is a graph that charts the PSNR (peak signal to noise ratio) – a measure of picture quality – obtained at different bitrates using the

FIGURE 4



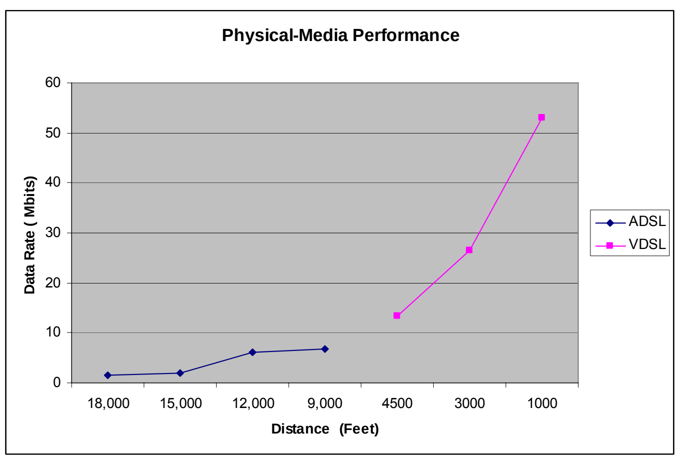
three different coding techniques. The video sequence that was tested is called “mobile and calendar” which is a popular (and difficult) video clip (used throughout the MPEG-2 development process) that contains saturated color and small amounts of movement of every part of the picture. The PSNR that typically is referenced as “good“ video quality is 32 PSNR. Tracing across the 32 PSNR line, the required bitrate to achieve that quality with MPEG-2 coding is 6 megabits per second while only 2.2 Mbps is required to duplicate that quality level with MPEG-4 AVC.

APPLICATIONS

Constrained bandwidth and storage applications will be the first to benefit from the coding gains of MPEG-4. One example of this type of application is the delivery of video services over broadband connections to the home or within an enterprise. While there may be a great deal of available bandwidth on longhaul networks, last mile IP connections are often limited to 3 – 6 Mbps. Telco and private network operators who depend on DSL (digital subscriber line) technology as their content delivery mechanism are constrained by existing twisted pair infrastructure and distance from their subscribers. The chart in *Figure 5* shows how bandwidth over ADSL (Asymmetric DSL) varied with respect to a subscriber’s distance from a telco central office. From the results of the compression tests, MPEG-4 video coding can help support delivery of multiple streams (for households with multiple TV sets) or single streams to subscribers residing larger distances from the central office. For HDTV, ADSL does not support that format at current bitrates – but soon may be able to carry MPEG-4 coded HDTV signals. Recent tests show excellent picture quality at 7.5 Mbps, so ADSL connections that support 6 – 8 Mbps may be able to carry HDTV to consumers.

Another obvious application is the HDTV DVD player. Where a feature length movie in HDTV MPEG-2 requires approximately 30 Gigabytes of storage (current DVD technology yields about 8 GB of storage) — an MPEG-4 version that only requires 15 GB, combined with laser advancements in DVD technology opens the prospect of affordable HDTV DVDs in the near future.

FIGURE 5



CONTENT DISTRIBUTION

As support for higher network speeds progress and video bandwidth requirements continue to decline, the prospect of content providers delivering their content over in-place, standard IP networks is materializing. As the broadcast domain meets the IP domain, there is a mapping of broadcast and IP constructs. In the IP network space, “multicast” is the term for “broadcast”, “unicast” is the term for “narrowcast”, and Session Description Protocol (SDP) is the equivalent of Program Information (metadata) – to name just a few of the new terms for the IP broadcaster.

One major difference and advantage with IP networks is that they are inherently two-way networks compared to the one-way nature of most broadcast networks (cable networks are two-way today, using a variety of techniques to create a return path to the service provider). Using a two-way network, there is built-in infrastructure for verifying receipt of content and for setting up sessions for users to “ask for” content and to have that content delivered in a secure manner. These techniques can lead to nationwide or worldwide networks where content providers can connect at any point and delivery content to anyone on the network. Using a simple set-top terminal or an application on an advanced personal computer, users can be rapidly added to a network, their identity verified, and a business relationship quickly established.

DISTRIBUTED NETWORKS

Many locales are already connected with high speed fiber and can support a content distribution overlay. *Figures 6 and 7* depict subsets of a content distribution network where content is centrally aggregated and subsequently distributed and staged at a regional level, before delivery to individual subscribers. Inexpensive servers are employed to move and store content at strategic locations, with live content encoded closer to the final consumer. Although most content today is broadcast in real-time, only a small percent is truly “live” broadcast – paving the way for more efficient store-and-forward networks where pre-recorded content is staged at points close to the user and subsequently streamed over a variety of networks.

FIGURE 6

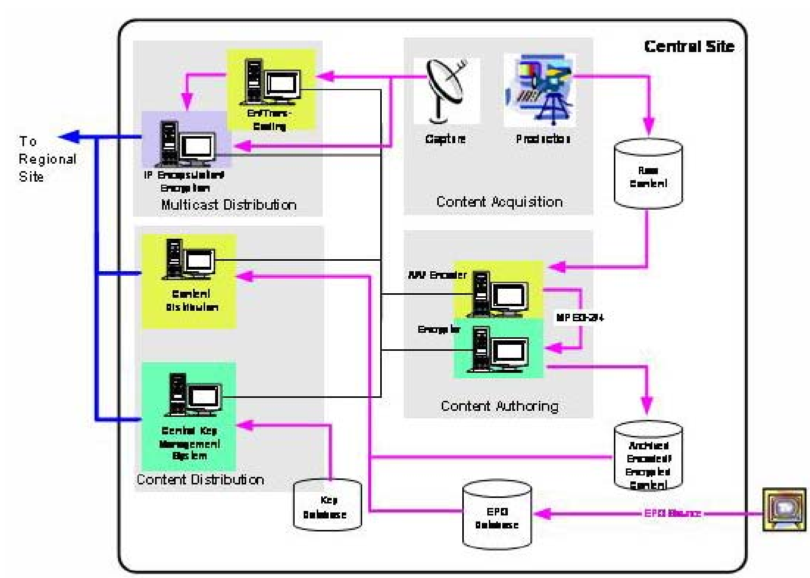
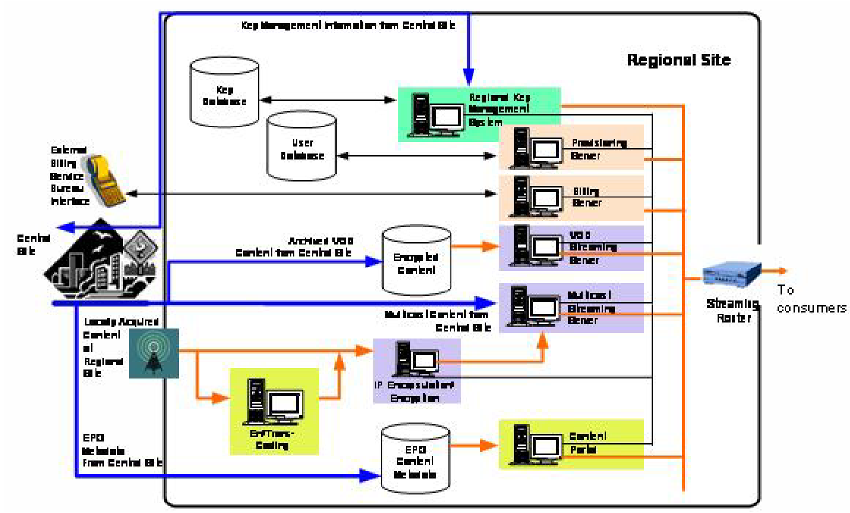


FIGURE 7



IP RIGHTS MANAGEMENT AND CONTENT SECURITY

The ability to protect digital content and ensure the content owners are properly compensated has been one of the hottest topics and biggest challenges in taking advantage of these new distribution technologies. Where content can reside in multiple locations and large numbers of users have IP connections, the topic of digital content and content protection need to be considered as one. The technical challenge presents itself as the requirement to deliver content in a secure manner so that content owners' rights are protected and business models and contracts are enforced while providing end users with a simple and logical method to access content that does not detract from the expected viewing experience. As mentioned earlier, two primary types of content delivery need to be supported; the broadcast model (multicast) and the narrowcast (video-on-demand) model (unicast).

At a high level, there are two operations involved with content security: (1) the encryption of content and association of the content owner's desired rights assignment (number of allowed copies, number of viewing opportunities, duration of viewing windows, etc.); and (2) the distribution of the secure keys necessary to access the content. Two-way networks enable these operations by proving the infrastructure (the physical connections as well as standard protocols) so that individual end-users can be identified, given the access (keys) to desired content, and subsequently be billed for services delivered.

FIGURE 8

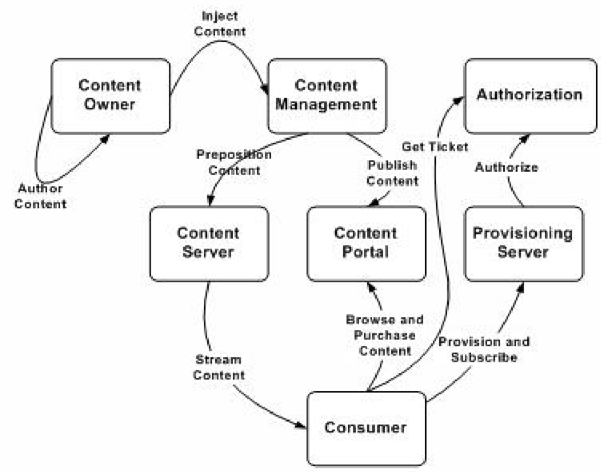
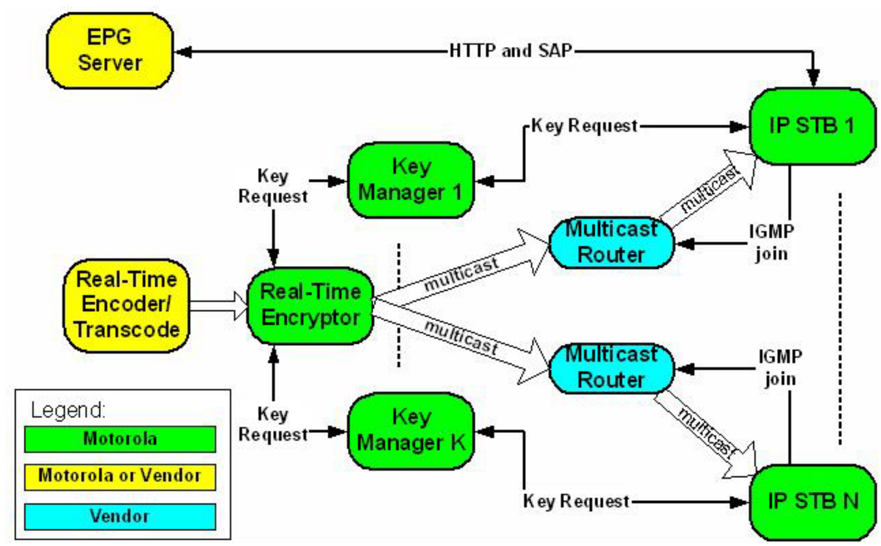


Figure 8 depicts the content transaction functions placed within an IP network and the communications links between the various parties to a transaction. Inexpensive servers allow strategic placement of key storage locations so that transactions can be resolved quickly – with requests and granting of access being fairly transparent to the user. To keep content secure and operate in this environment while preserving the expected viewing experience, techniques have been developed to minimize the level of transactions in certain viewing models. In the case of broadcast content, viewers are accustomed to "surfing" through channels or directly tuning to a desired channel. As users hop back and forth, the network may become overloaded if (in the background) permission request/grant sessions were initiated every time a viewer changed channels. In this situation the security system can permit the user (user's set top terminal) to request channel access authorization in advance so that the necessary keys can be pre-loaded. The network side of the broadcast model uses multicast-enabled routers to set up a multicast group for a broadcast program. This means that one copy of the content is actually streamed, but a list of multicast addresses are generated so that viewers can obtain the address and become part of the multicast group. Once a user has permission and has "joined" the multicast group – the IP packets that make up the multicast are duplicated and directed to the end-user terminal at the last connection between the access network and the user (see *Figure 9*).

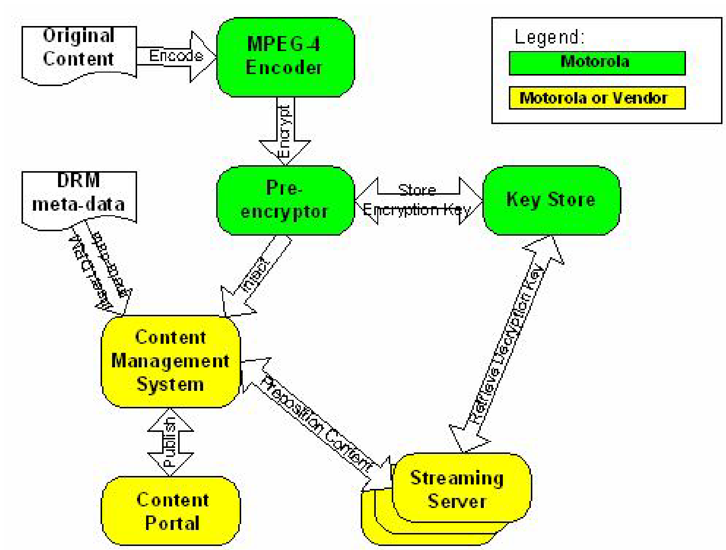
FIGURE 9



VIDEO-ON-DEMAND (VOD)

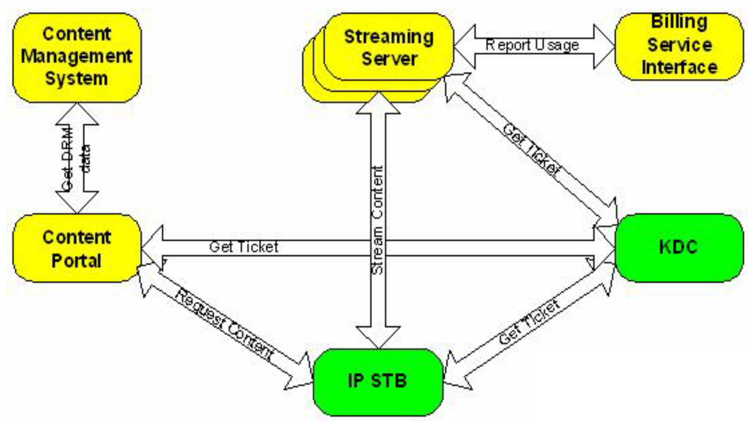
VOD applications take advantage of inexpensive edge servers that aggregate content that is of interest to the specific demographics being server from that location. Using the IP network, a large content archive can be centrally located, with popular content pre-positioned at regional hubs or older content downloaded to regional hubs upon request. (See *Figure 10*) Content owners may choose to author content and store the content at their own facility and distribute to the central hubs of different network/service providers.

FIGURE 10



To protect VOD content, encryption may be applied as the content is streamed to the individual viewer, or can be pre-encrypted and stored in an encrypted state as it moves throughout the network. Once a viewer selects content for viewing, the key exchange transactions begin and the user is issued a "ticket" (i.e. "good for one viewing of Gone with the Wind") if he passes verification and has an account in good standing. (See *Figure 11*)

FIGURE 11



IP SYSTEM ELEMENTS

Figure 12 depicts the various elements and functions of a secure and distributed content distribution system. Key elements include:

Content Server - Content storage and streaming close to the end-user.

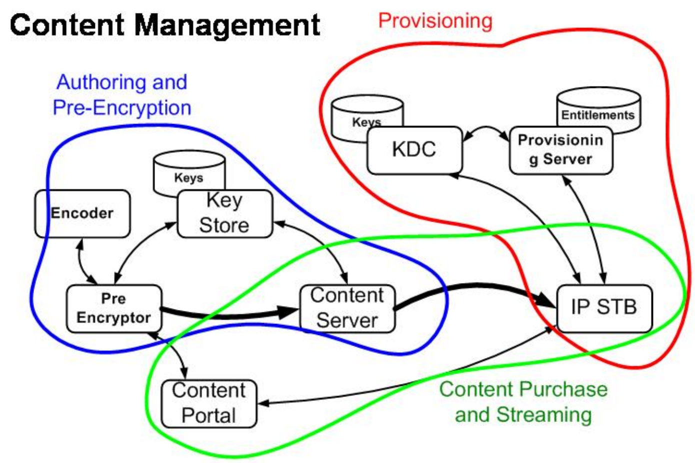
Key Store - A central repository for encryption/decryption keys. Content servers request decryption keys from the Key Store when a user accesses content for the first time.

Key Distribution Center – Authenticates users and issues tickets for users to employ during client-server communications.

Content Portal – Stores descriptive information about the content and the content purchase rules.

Provisioning Server – Adds new users to the system and sets up accounts.

FIGURE 12



CONCLUSION

MPEG-4 technology, IP network-based security and digital rights management, and inexpensive, high performance network components are becoming available in a similar timeframe to enable efficient content distribution over existing IP networks – connecting a large number of content providers with a large number of potential viewers. Two-way IP networks and existing IP protocols support flexible content management scenarios and adaptable, scalable approaches to content security to ensure that the rights of content owners are adequately protected.

Motorola, Inc.
Connected Home Solutions
6450 Sequence Drive
San Diego, CA 92121 U.S.A.

www.motorola.com

5665-0605-0K