

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
3 March 2005 (03.03.2005)

PCT

(10) International Publication Number
WO 2005/020556 A2

- (51) International Patent Classification⁷: **H04N**
- (21) International Application Number:
PCT/US2004/022575
- (22) International Filing Date: 14 July 2004 (14.07.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
10/639,006 11 August 2003 (11.08.2003) US
- (71) Applicant (for all designated States except US): **nCUBE CORPORATION** [US/US]; 1825 NW 167th Place, Beaverton, OR 97006 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **MATARESE, Joseph** [US/US]; 2830 NW Westover Road, Portland, OR 97210 (US). **POOLE, Andrew** [US/US]; 322 South Taft Court, Louisville, CO 80027 (US). **SCOTT PINSON, George** [US/US]; 14613 NW Wendy Lane, Portland, OR 97229 (US).
- (74) Agent: **WESTBERG, Derek, J.**; Law Offices of Derek J. Westberg, Two N. Second Street, Suite 1390, San Jose, CA 95113 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)
- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)
- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE,

[Continued on next page]

(54) Title: OPTIMAL PROVISIONING AND MANAGEMENT OF BANDWIDTH IN A VIDEO-ON DEMAND SERVICES ARCHITECTURE

(57) Abstract: A method for provisioning a Video-On-Demand server with enough bandwidth to stream videos to subscribers such that the probability of the Video-On-Demand server not having enough bandwidth to stream video to a subscriber is less than a set value.

WO 2005/020556 A2



IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*

Published:

- *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

OPTIMAL PROVISIONING AND MANAGEMENT OF BANDWIDTH IN A VIDEO-ON-DEMAND SERVICES ARCHITECTURE

FIELD OF THE INVENTION

The present invention relates generally to the delivery of Video On-Demand (VOD) services and more particularly, to the provisioning and monitoring of the delivery of the VOD services.

BACKGROUND OF THE INVENTION

In recent years, the media industry has expanded its horizons beyond traditional analog technologies. Audio, photographs, and even feature films are now being recorded or converted into digital formats. With this conversion to digital formats comes the ability of the media industry to provide Video-On-Demand (VOD) services to subscribers. VOD is similar to a traditional video rental service except that the video is delivered directly to the subscriber through a media delivery network such as a cable network.

In a conventional VOD delivery system, VOD services are provided to the subscribers via a delivery network. The subscribers are divided into service groups. A fixed bandwidth for delivery of VOD services is associated with each service group. The bandwidth the VOD server has available to provide VOD services is greater than or equal to the sum of the bandwidth of the service groups. The VOD server is provisioned with this bandwidth to ensure that a subscriber is never denied service because the VOD server does not have enough bandwidth. This increases the cost of the VOD server.

Therefore, what is needed is a method for optimally provisioning and managing the bandwidth of the VOD server.

BRIEF SUMMARY OF THE INVENTION

The invention is a technique for optimal provisioning and management of bandwidth in a video-on-demand services architecture. In one embodiment of the present invention a VOD system, includes a VOD server and a delivery network. The VOD server generates VOD services for subscribers while the delivery network is coupled to the VOD server for delivering the VOD services to the subscribers. The bandwidth of the VOD server to provide VOD services to the subscribers is less than the bandwidth of the delivery network to provide the VOD services to the subscribers. In

another aspect of the invention, the VOD server has flexibility in selecting a path in the delivery network by which a VOD service is provided. The delivery network can also be reprovisioned based on monitored usage of the delivery network.

BRIEF DESCRIPTION OF THE DRAWINGS

5 FIG. 1 is a block diagram of a VOD server and a distribution network in which the present invention may be implemented;

 FIG. 2 is a block diagram of the VOD server, a distribution network and a switched network in which the present invention may be implemented;

 FIG. 3 is a graph showing an exemplary daily peak demand of a service group in
10 the system of FIGS. 1 or 2;

 FIG. 4 is a graph showing an exemplary daily peak demand of a module in the system of FIGS. 1 or 2;

 FIG. 5 is a graph showing the bandwidth utilization of multiple service groups at one instant time;

15 FIGS. 6A-C is a group of graphs showing the time dependence of bandwidth utilization of multiple service groups; and

 FIG. 6D is a graph showing the time dependence of bandwidth utilization of the VOD server.

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT OF THE

20 INVENTION

 The present invention relates to the provisioning and monitoring of a VOD server 102 and an associated delivery system 100, as shown in FIGS. 1 and 2.

 As shown in FIG. 1, the delivery system 100 connects one or more VOD servers 102 to one or more service groups 108a-f such that the VOD servers 102 provide the
25 service groups 108a-f with content in the form of streaming digital video and generally also audio, though either audio or video may be provided. Preferably a single VOD server 102 is provided that is scaleable, however, multiple servers may be provided, for example, if the servers are not scaleable. Each service group 108a-f is a group of subscribers that all share the same bandwidth in the delivery network 100. Each service
30 group 108a-f may comprise, for example, one hundred to over a thousand subscribers.

 The topology of the delivery network 100 allows the VOD server 102 to deliver unique content to each service group 108. Thus, the VOD server 102 tailors the content

to the desires of the subscriber. The subscriber may request that the VOD server 102 provide a particular video program, and may request that the VOD server 102 perform trick functions such as: pause, rewind, fast forward, slow play, slow rewind, or some other function on the video.

5 An interface between the delivery network 100 and the VOD server 102 may be provided by one or more modules 104a-c such as: a 10 Gigabit Ethernet Module, a Gigabit Ethernet Module, a 10/100 BaseT Ethernet module, an ATM OC-3/STM-1 module, an ATM OC-12/STM-4 module, an DVB-ASI module, an QAM 64 module, an QAM 256 module, or any module that allows the VOD server 102 to communicate to the
10 subscribers over the delivery network 100. In a switched network, Ethernet or ATM is preferably used. Higher speed interfaces may be utilized in the invention without changing the nature of the invention. The bandwidth of the VOD server 102 may be proportional to the number of modules 104a-c with which the VOD server 102 has been provisioned.

15 Each module 104 in the VOD server 102 may be connected to one or more service gateways 106 via network links (e.g. Gigabit Ethernet links). In addition, each service gateway 106 may be connected to one or more modules 104. For example, a connection may be provided between a service gateway 106b and a module 104a and between the service gateway 106b and the module 104b as in FIG. 1. The connection
20 between a service gateway and a module may be direct or through one or more intermediaries. The intermediaries may include distribution hubs, switches or other devices that provide routing or transmission functionality in the delivery network 100. Each of the service gateways 106a-d may split the VOD services being sent by the VOD servers 102 into one or more service groups 108. A service gateway 106a-d may deliver
25 video streams to each of several service groups (e.g. 108a-c) via a delivery channel such as one or more Quadrature Amplitude Modulation (QAM) channels and may deliver different video streams to each service group. A delivery channel is a mechanism for carrying one or more video streams and may include one or more QAM channels, a Digital Subscriber Line (DSL) or a wireless access point. In the following description of
30 the invention the delivery channels are described as QAM channels. However, it will be apparent that any delivery channel may be used in place of a QAM channel.

A QAM channel may contain multiple video streams. Where each QAM channel has an upper limit (QAM channel bandwidth) on the number streams of digital video that it can deliver (e.g. 10). Thus, each service group 108a-f has an upper limit (a service

group bandwidth) on the number of digital streams that may be delivered to it. A service gateway 106a-d may use any delivery mechanism that allows the VOD server 102 to deliver the VOD services to subscribers in the service group 104 such as by sending MPEG-2 Single Program Transport Streams over IP (Internet Protocol) / Ethernet.

5 Each service gateway 106a-d has a service gateway bandwidth that is a sum of the service group bandwidth of the service groups 108a-f or portions of the service groups connected to that particular service gateway 106a-d. Each service group may be connected to one or more service gateways. Whenever a subscriber in a particular service group 108a-f requests a VOD service from the VOD server 102 and the service
10 group bandwidth that would be delivered to that particular service group 108a-f would be exceeded, then the VOD server 102 must deny the subscriber's request. To maintain a reasonable level of service, the probability of a request being denied, while greater than zero is low, typically between 0.1 and 1%. The lower the probability of service denial, the better the resulting subscriber experience.

15 The delivery network 100 may provide multiple paths between the VOD server 102 and the subscriber. Each of the paths may be divided into multiple links. Each of the paths may share some of the links among themselves. A management module 110 may monitor the state of each of the links either through the VOD server 102 or by interacting directly with the device providing the link.

20 In one aspect, the invention provides a method for connecting a set of the modules 104 to a set of the service groups 108, such that no one service group 108 is connected to only one of the modules 104. Bandwidth on the VOD server 102 may be assigned via any of the modules 104 comprising the VOD server to each service group 108 as needed. This is more efficient than conventional techniques as it allows the VOD
25 server to be provisioned with fewer modules. During the VOD service's busiest hours of day, demand for bandwidth in a service group may experience rapid variation, in particular "peaks" or "spikes", that exhaust the bandwidth capacity of the QAM channels feeding the service group as shown in FIGS. 6A-C. In FIGS. 6A-D the horizontal axis is time and the vertical axis is bandwidth utilization. Therefore number of QAM channels
30 and associated bandwidth feeding a service group must be sized appropriately large to accommodate these demand spikes. Because the VOD server aggregates bandwidth demand over all service groups, the bandwidth at the VOD server will have a much smaller relative variation owing to statistical averaging as shown in FIG. 6D. In other words, the decrease in variation is a consequence of summing up the demand of all the

subscribers to the VOD server 102 as opposed to summing up the demand of the subscribers in one service group 108. The present invention takes advantage of the smaller variation present at the server to provision the VOD server with fewer modules 104 than with the conventional method, thereby reducing the cost of the VOD server.

- 5 The conventional method would ensure that the bandwidth of the VOD server was greater than or equal to the bandwidth of all the service groups added together.

The VOD server 102 may be connected to a switching network 202 via multiple modules 104, such as Gigabit Ethernet modules, as shown in FIG. 2. Each service gateway 106 in turn is connected to one or more service groups 108, by way of one or
10 more QAM channels though each service group may be connected to multiple gateways. This will give the VOD server 102 the flexibility of not having to assign the QAM channel supplying a service group 108 to a particular module 104 as in the conventional system. The switching network 202 may include an Ethernet network or some other switching network that allows the VOD server 102 to communicate efficiently with
15 multiple service gateways 106.

In contrast to the present invention, a conventional VOD server would ensure that the module, which is connected to one or more service gateways, has a bandwidth that is greater than or equal to the sum of the service gateway bandwidth of each of the service gateways that are connected to it. This is done to ensure that requests for VOD services
20 are not being denied due to lack of module bandwidth. The bandwidth of the VOD server is equal to the sum of the bandwidth of all the modules. Due to the wide variation in demand for bandwidth in each of the service groups, while at any instant in time one or more of the service groups may have used up all their bandwidth, the VOD server may only have used 60 to 70% of its bandwidth.

25 In the present invention, the VOD server 102 is generally provisioned with fewer modules 104 than a conventional system would be. However, if the server 102 is not provisioned with enough modules 104 then a subscriber might be denied service because the VOD server 102 does not have enough bandwidth. Therefore, the provisioning of the VOD server 102 with bandwidth in the form of modules 104, preferably takes into
30 account typical use patterns of subscribers in all the service groups 108. The use patterns may be summed together because the VOD server 102 can assign service to different paths in order to make efficient use of system resources. To ensure that the VOD server 102 is provisioned with enough bandwidth the probability that demand for bandwidth on the VOD server 102 will be greater than some value over a set period may be calculated.

The VOD server is provisioned with a set bandwidth. There is a probability that the demand for bandwidth at the VOD server may exceed the set bandwidth. The probability may be below a set value. The probability may be calculated using past subscriber behavior or some other means.

A probability P_{network} is defined as the probability that a subscriber will be denied service because there is not enough bandwidth in the delivery network 100. A probability function P_{server} is defined as the probability that a subscriber will be denied service because the VOD server 102 does not have enough bandwidth. The probability function P_{server} is a function of the bandwidth of the VOD server 102. In general, as the bandwidth of the VOD server 102 increases the value of the probability function P_{server} decreases. Further, as the bandwidth of the VOD server 102 decreases the value of the probability function P_{server} increases. In one aspect of the invention, the bandwidth of the VOD server 102 is provisioned such that P_{server} does not exceed a set value. Further, the set value may be less than P_{network} , in other words:

15 $P_{\text{network}} > P_{\text{server}}$

In another alternate embodiment, the set value might also be set to 0.1-1% of service requests. A service request might be denied for reasons other than lack of bandwidth on the VOD server 102, such as lack of bandwidth in the delivery network. Provisioning the VOD server 102 with bandwidth above certain point will have a marginal (i.e. insignificant) effect on the probability of service requests being denied as other bandwidth limitations in the delivery system overwhelm any bandwidth limitation that the VOD server 102 might impose. Therefore, in a further alternate embodiment, the set value might be such that a decrease in the set value has no appreciable effect on the probability that a subscriber might be denied service for any reason.

These probabilities may be calculated based on past knowledge of the behavior of the subscribers in order to predict their future behavior. It may also take into account the past behavior of other media viewers not connected to the video server 102, which might have similar use patterns to the subscribers. The probability may also be calculated from a mathematical model of the subscribers demand for bandwidth. Known techniques for implementing mathematical models based on statistical behavior may be applied.

A management module 110 preferably has an ability to monitor and report on all currently active subscriber sessions and server streams. The management module 110

may report on currently active sessions and keep an archive on completed sessions to perform this monitoring.

The management module 110 preferably maintains detailed information of the nodes and links of the delivery network topology. Such topological information may include an inventory of the various components of the VOD server 102 such as the modules 104, the service gateways 106, and the service groups 108 and the connectivity between them (i.e. the connection provided by the communication links). The VOD server 102 may create logical groups of bandwidth, where a logical group of bandwidth may be defined as: one or more QAM channels, one or more service groups, one or more service gateways or one or more modules. For example, a table mapping the delivery network topology may contain a row for each said logical group in the delivery network 100. Columns of such a table may contain the bandwidth associated with the logical group as well as the identification of a physical entity like a service gateway, Gigabit Ethernet link or QAM channel associated with the logical group. Each row may also contain information regarding every connection associated with each said logical group. The delivery network topology could alternatively be stored as a graph, a linked listed, a hierarchical database, or some other data structure.

The management module 110 may provide an operator of the delivery network with the ability to modify the delivery network topology table as needed to reflect additions, deletions or other changes of the VOD server 102 resources, the service gateways 106 or the service groups 108. Such modifications may be made during scheduled maintenance windows or while the system is live to subscribers.

The management module 110 may monitor the bandwidth utilization and availability across all resources in the delivery network. In particular, the management module 110 may track and record bandwidth utilization on the VOD server 102, the modules 104, the service gateways 106, the QAM channels, the service groups 108a-f and any other logical groups of bandwidth, and any link between any node in the delivery network. The information on bandwidth utilization may be obtained by, polling devices in real time or by knowing the network topology and knowing what is being transmitted. Archived data on this bandwidth utilization may be viewed subsequently for system planning and trend analysis. Examples of reports showing said bandwidth utilization are shown in FIG. 3 and FIG. 4. Some or all of the features described herein as being performed by the management module 110 could be performed by other elements of the system such as the VOD server 102. Similarly, features described herein

as being performed by the VOD server 102 may be performed by the management module 110.

FIG. 3 is a chart showing the daily peak demand for bandwidth allocated to the service group 108. FIG. 4 is a chart showing the daily peak demand for bandwidth allocated to the Module 104, which is equivalent to, the daily maximum of the sum of the demand in eleven service groups 108. FIG. 3 shows the daily peak demand for bandwidth in a single service group 108, which may have a bandwidth of 150 megabits per second (Mbps), while typical daily peak utilization may vary from sixty Mbps to 100 Mbps. While FIG. 4 shows the daily peak utilization for the module 104 serving eleven such service groups varying from fifteen Mbps to eighty-six Mbps. A VOD server 102 provisioned to service subscribers as in the prior art would ensure that the Module 108 providing service to eleven service groups each with a bandwidth of forty Mbps would have a bandwidth greater than four hundred and forty Mbps as shown by the line 402 in FIG. 4. In contrast, a VOD server 102 provisioned to service a delivery network such as in FIG. 2 could have far fewer modules and more modules may be added later as demand increases. In one aspect of the present invention one module 104 is not assigned to provide VOD services to a small subset of the service groups linked to the VOD server 102 but can provide VOD service to all or a large subset of the service groups that are linked to the VOD server 102.

A single service group will have wide temporal variation in demand for bandwidth relative to its bandwidth capacity as shown in FIGS 6A-E. In contrast, the temporal variation of bandwidth demand relative to aggregate bandwidth capacity decreases when the demand for bandwidth associated with a number of the service groups is summed together as shown in FIG. 6F. The present invention takes advantage of this decrease in variation to provision a VOD server with less bandwidth than if the VOD server was provisioned in the conventional manner.

The delivery network topology information preferably allows the VOD server 102 to choose a video delivery path by way of any of a combination of links between a VOD server module 104 and the a subscriber device in a service group, each time the subscriber makes a request. The VOD server 102 may choose a path based upon the bandwidth utilization of the path or links or a priority level of the path or links. If a failure is detected on a particular link, such as a QAM channel in the delivery network, the VOD server 102 may choose to avoid this path for some period of time. For example, the VOD server 102 may de-prioritize that particular link and send an alert to

the operator of the delivery network via an appropriate protocol such as Simple Network Management Protocol and/or through a Graphical User Interface and/or through some other means that allows the VOD server 102 to contact the operator of the delivery network. A VOD server 102 might detect a failure in the particular link by noting a high error rate, or high demand for the retransmission of data. The transmission method might include error detection and correction code that would help in the detection of a failed link. The VOD server 102 may avoid using the de-prioritized link again until: the operator intervenes to resolve the underlying failure and then explicitly re-prioritizes the link, or bandwidth is completely consumed on the other links feeding the service group.

10 The present invention has been described in terms of specific embodiments incorporating details to facilitate the understanding of the principles of construction and operation of the invention. Such reference herein to specific embodiments and details thereof is not intended to limit the scope of the claims appended hereto. It will be apparent to those skilled in the art that modifications may be made in the embodiment
15 chosen for illustration without departing from the spirit and scope of the invention.

CLAIMS

What is claimed is:

- 1 1. A Video-On-Demand system including a server for generating VOD
2 services for subscribers and a delivery network coupled to the server for delivering the
3 VOD services to the subscribers, wherein a bandwidth of the server for providing Video-
4 On-Demand services to the subscribers is less than a bandwidth of the delivery network
5 for providing the Video-On-Demand services to the subscribers, and wherein the system
6 flexibly selects a path in the delivery network by which a requested VOD service is
7 provided.
- 1 2. The system according to claim 1, wherein the links in the paths in the
2 delivery network are monitored.
- 1 3. The system according to claim 2, , wherein when a failure in a particular
2 path between the subscriber and the server is detected an alert is sent to an operator and
3 the particular path is de-prioritized such that one or more alternative paths are selected
4 between the subscriber and the server.
- 1 4. The system according to claim 3, wherein the particular path remains
2 de-prioritized and the alternative paths are selected until the bandwidth in all the
3 alternative paths between the server and the subscriber is consumed or the operator
4 intervenes to resolve the underlying failure in the particular path and re-prioritizes that
5 particular data path.
- 1 5. The system according to claim 3, wherein the status of the particular path
2 is checked and the particular path is re-prioritized if the underlying failure in the
3 particular path is corrected.
- 1 6. The system according to claim 2, wherein utilization of the links and the
2 server is monitored.
- 1 7. The system according to claim 6, wherein the system ensures that the
2 utilization of a link does not exceed the bandwidth of the link.
- 1 8. The system according to claim 7, wherein the system denies service to a
2 subscriber to ensure that the utilization of the link does not exceed the bandwidth of the
3 link.

1 9. The system according to claim 7, wherein the system assigns a service to
2 a different link to ensure that the utilization of the link does not exceed the bandwidth of
3 the link.

1 10. The system according to claim 6, wherein the system uses a past record of
2 the utilization of the links and the server to compute a probability that the demand for
3 Video-On-Demand services will exceed the bandwidth of an individual link to deliver
4 said services or a probability that the demand for the services will exceed the bandwidth
5 of the server to deliver said services.

1 11. The system according to claim 10, wherein the server is provisioned with
2 enough bandwidth to ensure that the probability that demand for the Video-On-Demand
3 services exceeds bandwidth of the server to deliver said services does not exceed a set
4 value.

1 12. The system according to claim 11, wherein the set value is greater than
2 zero.

1 13. The system according to claim 11, wherein the set value is such that the
2 probability that the demand from the subscriber is denied for any reason does not
3 significantly decrease if the set value decreases.

1 14. The system according to claim 10, wherein the probability is further
2 determined from a past record of behavior of viewers that are not subscribers of the
3 Video-On-Demand services.

1 15. The system according to claim 10, wherein the probability is further
2 determined from a mathematical model of the subscribers behavior.

1 16. The system according to claim 10, wherein the probability is further
2 determined by forecasting future behavior of the subscribers.

1 17. The system according to claim 1, wherein the system monitors the
2 utilization of elements of the delivery network; and assigns paths in the delivery network
3 to deliver the Video-On-Demand services in response to the monitoring.

1 18. The system according to claim 17, wherein the server reprovisions the
2 delivery network in response to a historic record of the monitoring.

1 19. The system according to claim 17, wherein the system updates its
2 knowledge of topology and available bandwidth of the delivery network in response to
3 additions, deletions or other changes of the elements of the delivery network that are
4 subject to monitoring.

1 20. The system according to claim 17, wherein monitoring of the links further
2 includes monitoring of failures in the delivery of the Video-On-Demand services to the
3 subscribers.

1 21. A Video-On-Demand system including a server for generating VOD
2 services for subscribers and a delivery network coupled to the server for delivering the
3 VOD services to the subscribers, wherein a bandwidth of the server for providing Video-
4 On-Demand services to the subscribers is less than a bandwidth of the delivery network
5 for providing the Video-On-Demand services to the subscribers, and wherein the server
6 is provisioned with enough bandwidth such that the probability that the server will not
7 have enough bandwidth to deliver the services to the subscribers is less than a set value.

1 22. The system according to claim 21, wherein the system flexibly selects a
2 path in the delivery network that connects the server to the subscribers.

1 23. The system according to claim 22, wherein the system monitors links in
2 the paths in the delivery network.

1 24. The system according to claim 23, wherein when a failure in a particular
2 path between the subscriber and the server is detected an alert is sent to an operator and
3 the particular path is de-prioritized such that one or more alternative paths are selected
4 between the subscriber and the server wherein other subscribers share the same path
5 between themselves and the server.

1 25. The system according to claim 24, wherein the particular path remains
2 de-prioritized and the alternative paths are selected until the bandwidth in all the
3 alternative paths between the server and the subscriber is consumed or the operator
4 intervenes to resolve the underlying failure in the particular path and re-prioritizes that
5 particular data path.

1 26. The system according to claim 24, wherein the status of the particular
2 path is checked and the particular path is re-prioritized if the underlying failure in the
3 particular path is corrected.

1 27. The system according to claim 23, wherein utilization of the links and the
2 server is monitored.

1 28. The system according to claim 27, wherein the system ensures that the
2 utilization of a link does not exceed the bandwidth of the link.

1 29. The system according to claim 28, wherein the system denies service to a
2 subscriber to ensure that the utilization of the link does not exceed the bandwidth of the
3 link.

1 30. The system according to claim 28, wherein the system assigns a service to
2 a different link to ensure that the utilization of the link does not exceed the bandwidth of
3 the link.

1 31. The system according to claim 27, wherein the system uses a past record
2 of the utilization of the links and the server to compute the probability that the demand
3 for Video-On-Demand services will exceed the bandwidth of an individual link to
4 deliver said services.

1 32. The system according to claim 21, wherein the set value is greater than
2 zero.

1 33. The system according to claim 21, wherein the set value is such that the
2 probability that the demand from the subscriber is denied for any reason does not
3 significantly decrease if the set value decreases.

1 34. The system according to claim 21, wherein the probability is further
2 determined from a past record of behavior of viewers that are not subscribers of the
3 Video-On-Demand services.

1 35. The system according to claim 21, wherein the probability is further
2 determined from a theoretical model of the subscribers behavior.

1 36. The system according to claim 21, wherein the probability is further
2 determined by forecasting future behavior of the subscribers.

1 37. The system according to claim 21, wherein the system monitors the
2 utilization of elements of the delivery network; and assigns paths in the delivery network
3 to deliver the Video-On-Demand services in response to the monitoring.

1 38. The system according to claim 37, wherein the server reprovisions the
2 delivery network in response to a historic record of the monitoring.

1 39. The system according to claim 37, wherein the system updates its
2 knowledge of topology and available bandwidth of the delivery network in response to
3 additions, deletions or other changes of the elements of the delivery network that are
4 subject to monitoring.

1 40. The system according to claim 37, wherein monitoring of the links further
2 includes monitoring of failures in the delivery of the Video-On-Demand services to the
3 subscribers.

1 41. A method for monitoring utilization in a delivery network between a
2 server providing Video-On-Demand services and subscribers, comprising: monitoring
3 the utilization of elements of the delivery network; and assigning paths in the delivery
4 network to deliver the Video-On-Demand services in response to the monitoring.

1 42. The method according to claim 41, further comprising reprovisioning the
2 delivery network in response to a historic record of the monitoring.

1 43. The method according to claim 41, further comprising updating the
2 knowledge of topology and available bandwidth of the delivery network in response to
3 additions, deletions or other changes of the elements of the delivery network that are
4 subject to monitoring.

1 44. The method according to claim 41, wherein the monitoring of the links
2 further includes monitoring of failures in the delivery of the Video-On-Demand services
3 to the subscribers.

1/4

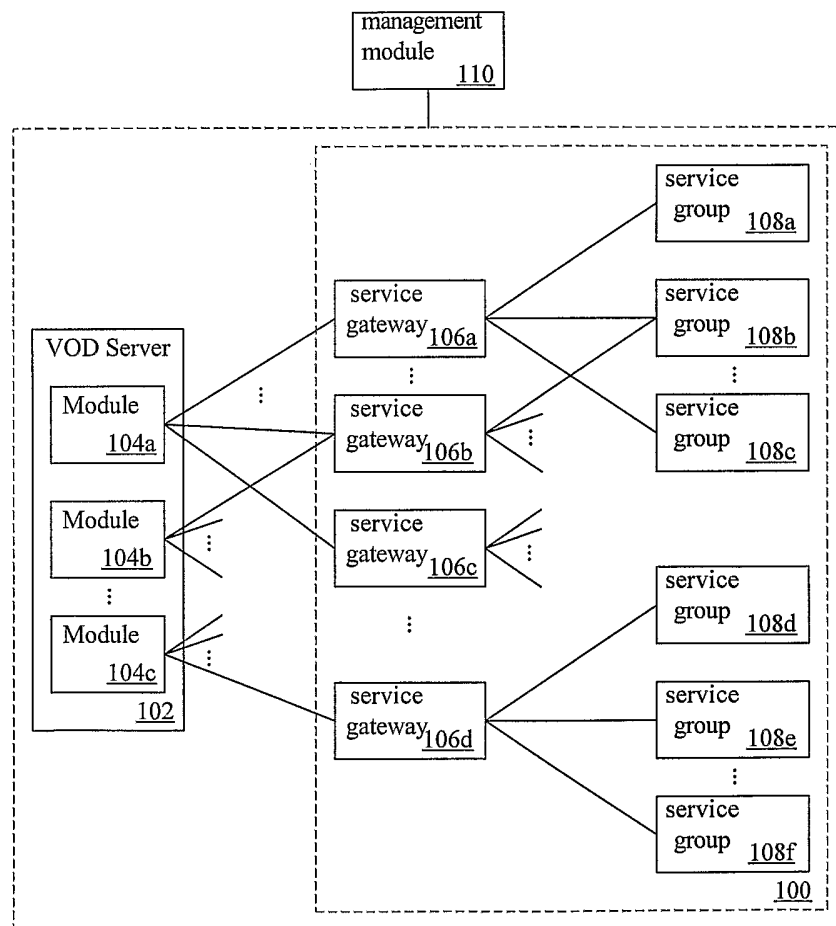


FIG. 1

2/4

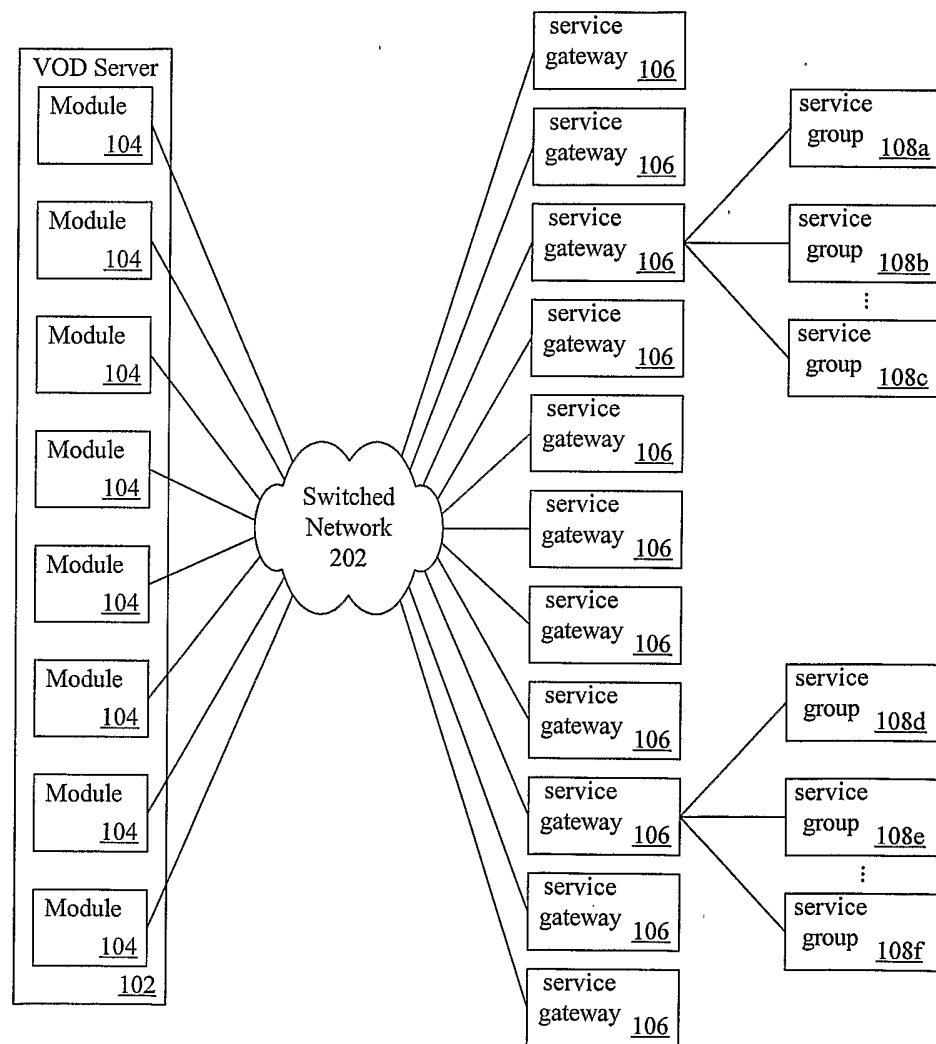


FIG. 2

3/4

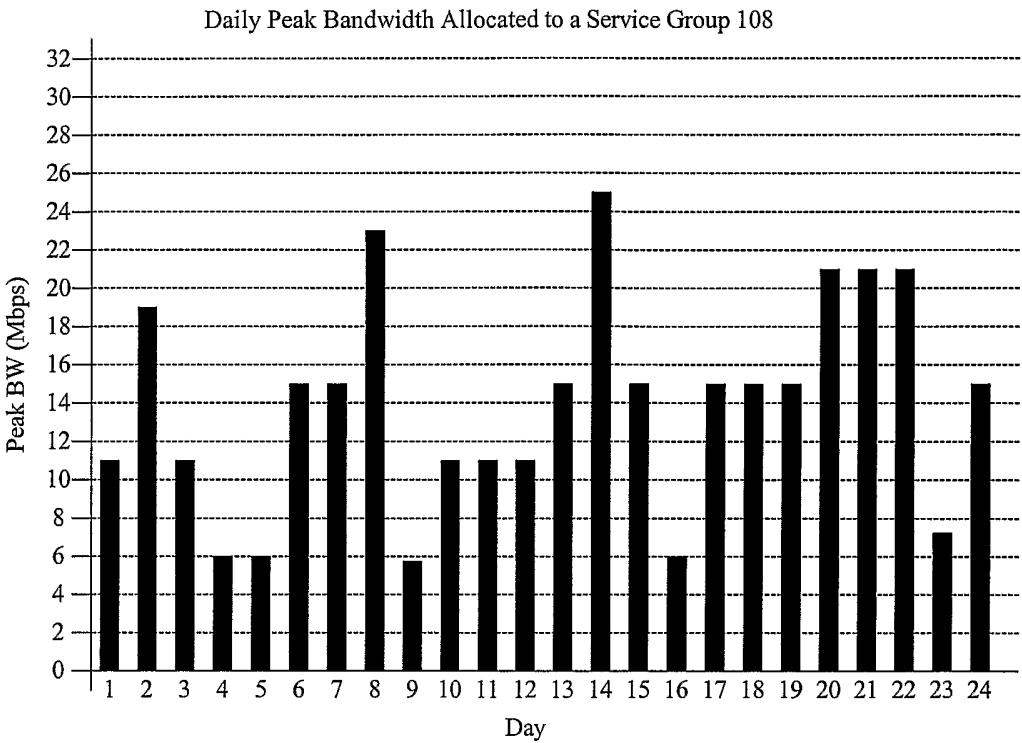


FIG. 3

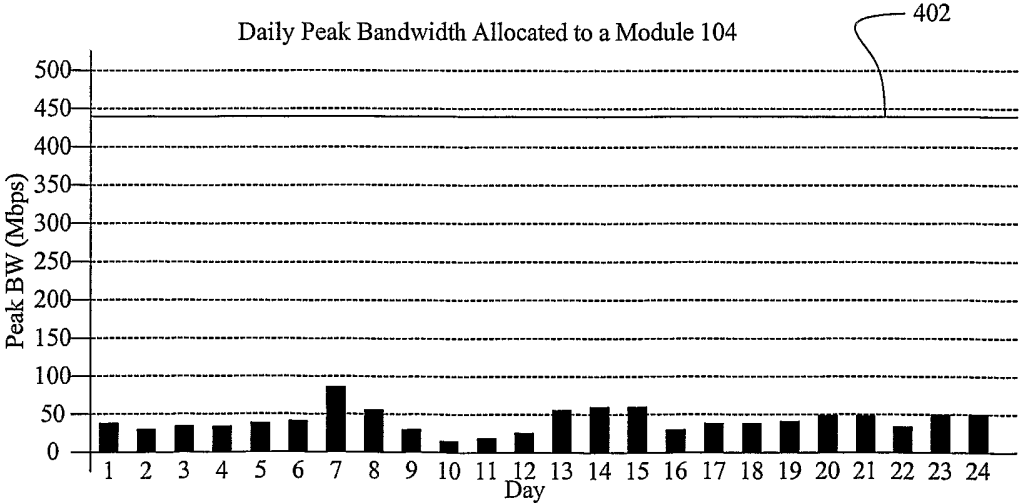


FIG. 4

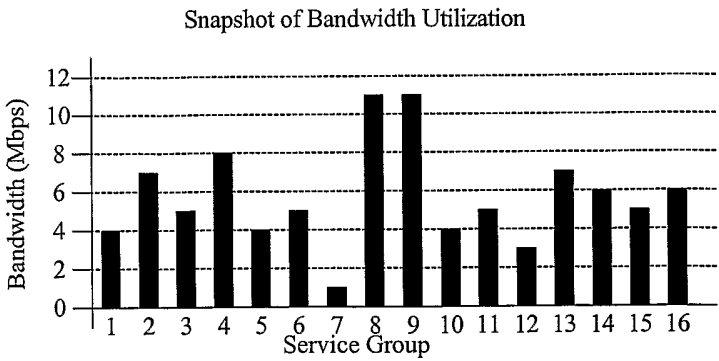


FIG. 5

