



(86) **Date de dépôt PCT/PCT Filing Date:** 2013/02/07
(87) **Date publication PCT/PCT Publication Date:** 2013/10/03
(85) **Entrée phase nationale/National Entry:** 2014/09/05
(86) **N° demande PCT/PCT Application No.:** US 2013/025180
(87) **N° publication PCT/PCT Publication No.:** 2013/148003
(30) **Priorités/Priorities:** 2012/03/26 (US13/429,656);
2012/08/06 (US13/567,681)

(51) **Cl.Int./Int.Cl. H04L 29/06** (2006.01),
H04N 21/845 (2011.01)
(71) **Demandeur/Applicant:**
CACTI ACQUISITION LLC, US
(72) **Inventeur/Inventor:**
MCGOWAN, ALBERT JOHN, US
(74) **Agent:** SMART & BIGGAR

(54) **Titre : SELECTION DYNAMIQUE DE PISTES AUDIO POUR UNE DIFFUSION DE CONTENU MULTIMEDIA EN FLUX CONTINU**
(54) **Title: DYNAMIC AUDIO TRACK SELECTION FOR MEDIA STREAMING**

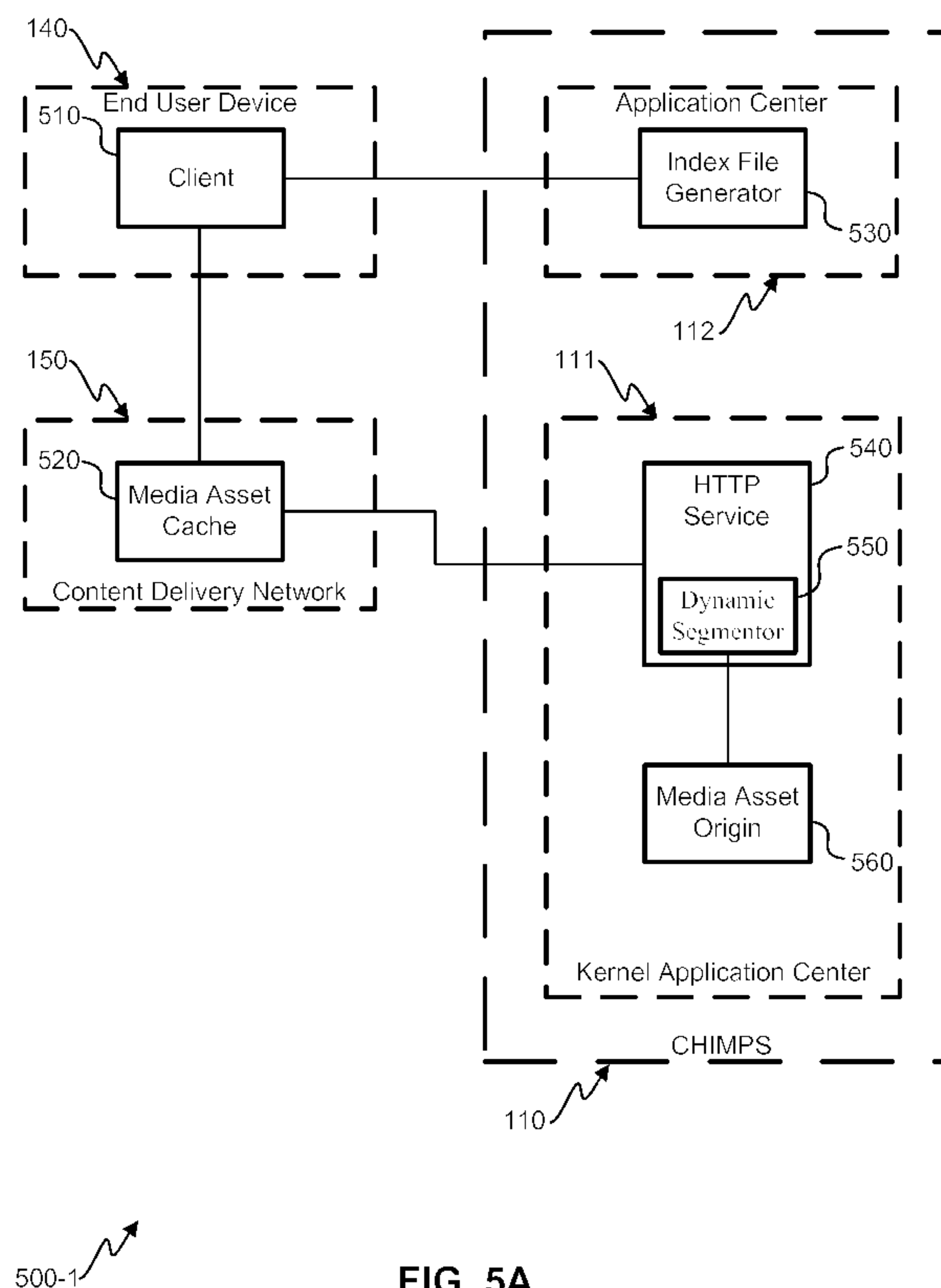


FIG. 5A

(57) **Abrégé/Abstract:**

Systems and methods for providing index files for dynamic audio track selection for media streaming can include receiving requests for media file segments (e.g., "chunks") with a certain audio track and responding to the requests by dynamically generating the corresponding media file segment with the requested audio track. The generated media file segment can then be provided to the requesting entity.



WO 2013/148003 A1

Published:

— *with international search report (Art. 21(3))*

DYNAMIC AUDIO TRACK SELECTION FOR MEDIA STREAMING

BACKGROUND OF THE INVENTION

[0001] The delivery of media over networks such as the Internet can be accomplished in many ways, including progressive downloading or streaming. Streaming is often preferred to progressive downloading because it offers additional features such as content protection and the ability to seek to undownloaded parts of a media file. The process of preparing a media file for streaming typically involves “chunking” the file, or dividing files up into smaller segments for delivery. Information including where chunks may be accessed can be stored in an index file. This index file can be delivered to a client, such as a media player application, for use in streaming.

[0002] The processes of chunking and indexing files for streaming present challenges to a content delivery network or content provider desiring to host media files for streaming. For example, a significant amount of preprocessing is required to prepare media for streaming. Media content may be encoded into several different files to accommodate several different sub-streams. Each of these files typically are chunked, stored, and indexed before the media content is available for streaming. This preprocessing leaves little flexibility for dynamically selecting an audio track during streaming and can require a large amount of storage space to store the preprocessed chunks of media.

BRIEF SUMMARY OF THE INVENTION

[0003] Systems and methods are disclosed provide a technical solution for enabling customized media file delivery by dynamically creating media segments, or chunks, having an appropriate audio track. The generated media file segment can then be provided to the requesting entity. Because media file segments are created dynamically, and an appropriate audio track is included, the systems and methods provided herein can reduce the required bandwidth for streaming a media file by removing unused audio tracks.

[0004] An example server providing audio track selection for media streaming via a data network, according to the description, includes an interface for communicating with the data network, a memory configured to store a media file, and a dynamic segmentor coupled with the memory and the interface. The dynamic segmentor is configured to obtain, via the

interface, a request for a first segment of the media file, determine, based on the request for the first segment of the media file, a starting point and an ending point of the requested first segment of the media file, and retrieve a first portion of the media file from the memory, the first portion of the media file having a video component and at least one audio track. The
5 dynamic segmentor is further configured to generate, without transcoding the first portion of the media file, the requested first segment of the media file such that the requested first segment of the media file has a single audio track, and send, via the interface, the requested first segment of the media file.

[0005] An example method of providing audio track selection for media streaming via a
10 data network, according to the description, includes obtaining a request for a first segment of a media file, determining, based on the request for the first segment of the media file, a starting point and an ending point of the requested first segment of the media file, and retrieving at least a portion of the media file, the at least a portion of the media file having a video component and at least one audio track. The method further includes generating,
15 without transcoding the at least a portion of the media file, the requested first segment of the media file such that the requested first segment of the media file has a single audio track, and sending providing the requested first segment of the media file.

[0006] Another example server providing audio track selection for media streaming via a data network, according to the description, includes means for obtaining a request for a first
20 segment of a media file, means for determining, based on the request for the first segment of the media file, a starting point and an ending point of the requested first segment of the media file, and means for retrieving at least a portion of the media file, the at least a portion of the media file having a video component and at least one audio track. The server further includes means for generating, without transcoding the at least a portion of the media file, the
25 requested first segment of the media file such that the requested first segment of the media file has a single audio track, and means for sending providing the requested first segment of the media file.

30 BRIEF DESCRIPTION OF THE DRAWINGS

[0007] The present disclosure is described in conjunction with the appended figures:

[0008] FIG. 1 illustrates a block diagram of a media servicing system.

[0009] FIG. 2A illustrates a block diagram of an embodiment of a kernel application center connected with application centers.

[0010] FIG. 2B illustrates a block diagram of an alternative embodiment of a kernel application center.

5 [0011] FIG. 3 illustrates a block diagram of an embodiment of an application center.

[0012] FIG. 4 illustrates a block diagram of processes and objects utilized by a cloud-hosted integrated multi-node pipelining system for media ingestion.

[0013] FIG. 5A illustrates a simplified block diagram of an embodiment of a system configured to provide dynamic indexing and chunking for media streaming.

10 [0014] FIG. 5B illustrates a simplified block diagram of another embodiment of a system configured to provide dynamic indexing and chunking for media streaming.

[0015] FIG. 5C illustrates a simplified block diagram of another embodiment of a system configured to provide dynamic indexing and chunking for media streaming, utilizing a redirector.

15 [0016] FIG. 6 illustrates a simplified flowchart of an embodiment of a method for implementing a dynamic index for media streaming.

[0017] FIG. 7 illustrates a simplified flowchart of an embodiment of a method for dynamically chunking a media file for streaming.

20 [0018] FIG. 8 illustrates a simplified swim lane flowchart describing the interaction of components in a system configured to provide dynamic indexing and chunking for media streaming, according to one embodiment.

[0019] FIGS. 9A-9C are illustrations of embodiments of how media file chunks can be generated to provide for dynamic audio track selection for media streaming.

25 [0020] FIG. 10 is a flow diagram illustrating a method for providing dynamic audio track selection for media streaming, according to one embodiment.

[0021] In the appended figures, similar components and/or features may have the same reference label. Further, various components of the same type may be distinguished by following the reference label by a dash and a second label that distinguishes among the similar components. If only the first reference label is used in the specification, the

description is applicable to any one of the similar components having the same first reference label irrespective of the second reference label.

DETAILED DESCRIPTION OF THE INVENTION

5 [0022] The ensuing description provides preferred exemplary embodiment(s) only, and is not intended to limit the scope, applicability or configuration of the disclosure. Rather, the ensuing description of the preferred exemplary embodiment(s) will provide those skilled in the art with an enabling description for implementing a preferred exemplary embodiment. It is understood that various changes may be made in the function and arrangement of elements
10 without departing from the spirit and scope as set forth in the appended claims.

[0023] The increased availability of media content over data communications networks such as the Internet has mirrored the increased bandwidth for these networks. Because media has recently taken a more prominent role in data communications, the distribution of media and the data associated with such distribution has become increasingly important, particularly
15 to media content providers. Media streaming has become a widely-used method of media distribution, but the preprocessing associated with streaming can be burdensome. Certain protocols, including forms of Hypertext Transfer Protocol (HTTP) streaming, require chunking and storing media assets, and generating a corresponding index files. These requirements can deprive a content provider of the ability to dynamically insert additional
20 media such as advertisements into a media stream, and can consume a large amount of storage space to store chunks of media for a media asset, including chunks for any alternative sub-streams (e.g., streams with alternative bitrates, captions, alternative languages, etc.). Certain systems and methods can be utilized, however, to introduce the desired functionality back into the system.

25 [0024] A traditional approach to preprocessing media for streaming involves chunking and storing media assets, then creating corresponding index files to indicate where chunks may be located to download for streaming. Streaming protocols often provide for frequently updating an index file for instances where the corresponding media is frequently updated, such as during live streaming. Thus, an index file does not need to contain all chunks for a
30 requested media asset. In addition, because media files are frequently stored in a format that requires little additional processing to chunk, the chunks can be created in real time, during

the streaming of a media file. The systems and methods disclosed herein take advantage of these features to enable dynamic index file creation and dynamic media file chunking.

[0025] For instance, rather than preprocess media assets for streaming by chunking and indexing all files with relevant sub-streams prior to streaming the media, a server can dynamically create and update an index file during streaming. The dynamically-created index file can contain information regarding a next chunk of media in the various available sub-streams. The next chunk of media may not be cached at a location specified in the index file, in which case a chunk may be dynamically created by pulling all or part of the media file of interest from a media file origin, chunking it, and making it available for download. The chunk also may be cached, thereby eliminating the need to create the chunk again if it is requested at some later time.

[0026] Because the chunks are created during streaming, a content provider and/or media distributor can have more information and control during the streaming process. Rather than generate single index file for a given media asset, an instance of the index file generator may be created at the beginning of the media streaming to provide individualized media content to a particular end user and unique information regarding the streaming session to a content provider. The file index generator can vary the length of each chunk by, for example, indicating starting and ending points in the index file. Thus, the file index generator may determine a uniform chunk length for a media asset, varying the length of the chunks for different media assets, or the file index generator may adjust the length of the chunks within a single media asset. The index file generator can further insert additional media, such as an advertisement, at any time during the streaming by specifying the location of the additional media in the index file. The determination to insert advertisements can be based on any information, including data collected during the streaming session.

[0027] As the index file generator receives requests for and generates index files, it can further gather data regarding the streaming session for reporting to a content provider. Content providers often rely on beaconing data collected from media player applications to determine when an end user stops, plays, pauses, skips, etc., the streaming media content. Such information can be vital in determining the value of the media.

[0028] Because not all media player applications provide this beaconing data, the data gathered by the index file generator can serve as a substitute for or complement to the beaconing data. For example, if a request is made for a chunk that does not immediately

follow a previously-requested chunk, a skip was made. If the amount of time elapsed between a previous request and a subsequent request exceeds the time for playback of the previously-requested chunk, a pause was made. If a request is not received within a certain time since a prior request, it can be determined that a stop was made.

5 [0029] As illustrated above, the state of a client may be determined from a variety of factors. This can include when the request for the index file is received, when the index file is provided, a length of time to play back the segment of media for streaming, and/or the starting and/or ending point of the segment of media for streaming. The determined state of a client may also be based on whether the request for the index file has been received within a
10 certain amount of time since receipt of a previous request for an index file, whether the segment of media for streaming includes media other than the media file, and more. The state of a client and/or the data from which it was determined, may be used to create reporting data to serve as a substitute or complement to beaconing data from a client media player application. Because the index file generator can determine the length of the chunks, it
15 therefore can determine the frequency of subsequent index file requests and the resolution of the reporting data based on the requests. The index file generator may log the reporting data and/or transmit the reporting data over a network during streaming.

[0030] The determined state of a client may be used by the index file generator and/or other services for various purposes. For example, it may be used in behavioral advertisement
20 targeting and enforcement of session advertisement behavior, adjusting advertisement content and playback based on the behavior of a user as determined by the state of a client. The state of a client further may be used to support resume features on a per client basis, allowing a user to continue playback of a media asset from a point at which the user had previously stopped playback. The state of a client also may be used to support individual encryption
25 keys in an encryption scheme and allow the index file generator to return secure URLs (e.g., time expiring or Internet Protocol (IP) allowed) for chunks to support functions such as payment services.

[0031] Additionally or alternatively, the tasks of generating the index file and providing a location a requested chunk can be split up, thereby enabling the system to determine which
30 chunks are actually requested. For example, a system may be configured to dynamically create an index file having links to one or more redirectors on the system. These redirectors can be configured to issue the location of the chunk, which can be created dynamically. The

redirectors can further determine which chunk is actually requested, thereby enabling, among other things, calculation of Quality of Service (QOS) metrics, an increase the accuracy of reporting data, a decrease the frequency of index file generation if efficient to do so, and the ability to more easily handle keys of an encryption scheme.

5 [0032] While the above embodiments may be implemented in a variety of different systems, some particular embodiments may be implemented as part of a media service system. **FIG. 1** is a block diagram illustrating a media servicing system 100, according to some embodiments of the present invention. The system may deliver media content to the end user device 140 through a network such as the Internet 120. The end user device 140 can
10 be one of any number of devices configured to receive media over the Internet 120, such as a mobile phone, tablet computer, personal computer, portable media device, etc. A media asset provided by a content provider 130 can be processed and indexed by cloud-hosted integrated multi-node pipelining system (CHIMPS) 110, and further stored on content delivery network (CDN) 150. Additionally or alternatively, the CHIMPS 110 may also be adapted to store the
15 media asset.

[0033] The media servicing system further enables a content provider 130 or other entity to gather information regarding user behavior during media playback. For example, a content provider 130 can be provided with data indicating that end users tend to stop watching a video at a certain point in playback, or that users tended to follow links associated with
20 certain advertisements displayed during playback. With this data, a content provider 130 can adjust factors such as media content, advertisement placement and content, etc., to increase revenue associated with the media content and provide the end user device 140 with a more desirable playback experience.

[0034] End user device 140 can request a media asset to stream with a client program
25 executed by the end user device 140. The client program can be, for example, a media player, browser, or other application adapted to request and/or play media assets. In response to a request for a media asset, the CHIMPS 110 can utilize any number of application centers 112 and/or kernel application center(s) 111 to provide the client program with a data object concerning the requested media asset. The data object can include information about the
30 media asset, including where the media asset can be located, such as within the CDN 150 or within the CHIMPS 150 itself. Location information may be provided by Universal Resource Indicator (URI), a Universal Resource Locator (URL) or other indicator. During playback of

the media asset, the CHIMPS 150 can collect data regarding the playback through beaconing provided by a client program executed by the end user device 140 and/or indexing service from within the CHIMPS and/or CDN. The CHIMPS 150 can subsequently provide the data and/or any analytics information derived from the data to the content provider 130.

5 [0035] FIG. 2A is a block diagram illustrating an embodiment of a kernel application 111-1 center connected with application centers from within the CHIMPS 110-1. The kernel application center 111-1 and application centers 112 can be geographically distant and can be connected via the Internet 120, wide area network (WAN), and/or other data communication network. Because application centers can be geographically separated, DNS services (not
10 shown) can be used to allow an end user device 140 to connect to the nearest available application center 112. The kernel application center 111-1 can connect with application centers 112 within the CHIMPS 110-1 through an internal interface 270, thereby enabling the application centers 112 access to the various components within the kernel application center 111-1.

15 [0036] Components within the kernel application center 111-1 can communicate through network 260 such as a local area network (LAN) and can include one or more origin servers 240 and a storage array 230 with which data objects and/or media assets may be stored and distributed. The storage array 230 may also be utilized by services running on processing server(s) 220 and/or transcoding server(s) 250 that may require temporary or long-term
20 storage. Kernel server 210 can utilize processing server(s) 220, transcoding server(s) 250 to provide various functional capabilities to the CHIMPS 110.

[0037] For example, as described in more detail below, the CHIMPS 110-1 can provide transcoding service for media assets provided by a content provider 130 for syndication. Such a service can allow a content provider 130 to upload a media asset to an application
25 center 112, after which the application center 112 would notify the kernel server 210 that the media asset has been uploaded. The kernel server can then notify services running on the processing server(s) 220 of the upload. These services can utilize transcoding server(s) to transcode the media asset, which can then be moved to a CDN and/or stored locally by storage array 230 and origin server(s) 240. Services running on the processing server(s) 220
30 can also update the associated data object stored by the storage array 230 and origin server(s) 240.

[0038] FIG. 2B is a block diagram illustrating an alternative embodiment of a kernel application center 111-2. In addition to the components of the embodiment of FIG. 2A, this embodiment incorporates an application center 112 within the kernel application center 111-2. The application center 112 incorporated within kernel application center 111-2 may be located at or near the other components of the kernel application center 111-2, and can be communicatively connected to the other components via network 260. The incorporated application center 112 can therefore have faster access to kernel application center functionality because it does not need to communicate over long distances. In consideration of this advantage, it will be understood that the CHIMPS 110 can include multiple kernel centers with one or more application centers incorporated therein. Additionally or alternatively, components of the kernel application center may be incorporated into one or more application centers 112 in the CHIMPS 110 to provide quicker access to certain functionality.

[0039] FIG. 3 is a block diagram illustrating an embodiment of an application center 112. The application center 112 can include caching server(s) 330 and a storage array 310 for storing and distributing data objects of media assets requested by end user devices through end user interface 360. Caching server(s) 330 and storage array 310 can also be used to collect, process, and/or store metrics information from beaconing data, media chunk requests, and/or other data sources, including data collected through end user interface 360. The application center can further include ingest server(s) 320 for ingesting uploaded media assets from a content provider 130 through a content provider interface 370. The media assets may be stored on the storage array 310. As with the kernel application center 111, the components of the application center 112 can be communicatively linked through a network 340, such as a LAN. The application center can further include an internal interface 350, providing a communication link from the application center to the rest of the CHIMPS. It is through internal interface 350, for example, that media assets stored on storage array 310 can be made available to a kernel application center 111 for services such as transcoding.

[0040] FIG. 4 is a block diagram 400 of processes and objects utilized by the CHIMPS 110 for media ingestion, according to some embodiments. Although FIG. 4 further indicates the physical systems in which my execute or store these processes and objects, it will be understood that the processes and objects disclosed may be executed or stored on more than one system, including systems not disclosed in FIG. 4. In other words, the processes and

objects shown in FIG. 4 allow for a variety of implementations through one or more of hardware, software, firmware, microcode, etc.

[0041] Media can be ingested into the CHIMPS 110 when a content provider 130 uploads a media asset to ingestion server(s) 410 in an application center 112 by utilizing a client 405.

5 The client 405 can be a stand-alone application or browser based, for example, and can communicate with ingest server(s) 410 through an application programming interface (API) configured for the ingestion of media assets.

[0042] Ingest server(s) 410 can communicate with devices in the kernel application center 111 executing programs such as kernel server 425 and file replication service 430. The

10 kernel server 425 can be configured organize the workflow among services such as transcoding 440 file system manager 435, and other services 445 (e.g., analytics, dynamic API, etc.) Upon a particular event, for example, the kernel server can be configured to notify the relevant services of the event, causing the services to process tasks associated with the event.

15 **[0043]** The file replication service 430, under direction of the kernel server 425, can coordinate the movement of the media assets between services. For example, retrieving the uploaded media asset from the ingest server(s) 410 and storing it on the file archive 450, or retrieving transcoded media assets from transcoding server(s) 460 and storing them in the media asset origin.

20 **[0044]** The data object updater 420 keeps the data object origin 415 up to date in response to any changes in the system. When, for example, a file is uploaded, transcoded, and stored in media asset origin 455, the location and other metadata concerning the transcoded media assets need to be created or updated in the data object origin 415 to ensure an end user device that accesses the object in the data object origin 415 has the correct information regarding the
25 related media asset. Because the data object updater 420 receives updates from the kernel server 425 (which is notified when a transcoded media asset is stored in the media asset origin 455, the system ensures the data objects in the data object origin are constantly up to date.

[0045] The upload of a media asset to the ingest server(s) 410, as described above, can
30 provide an example of how the kernel server 425 may coordinate workflow. For instance, in response to the upload, the ingest server(s) 410 can notify the kernel server 425 that a media

asset has been uploaded. The kernel server 425 informs the file replication service 430 of the uploaded media asset, and the file replication service 430 moves the uploaded media asset into the file archive 450 and notifies the kernel server 425 of the move. In response, the kernel server 425 notifies the file replication service 430, the file system manager 435, and the transcoding master 440 of the move. The file replication service 430 then will know it can delete the uploaded media asset from the ingest server(s) 410, the file system manager 435 will update the file system accordingly, and the transcoding master 440 will notify transcoding service(s) 460 of different transcoding tasks to be performed. The transcoding service(s) 460 can then retrieve the uploaded media asset from the file archive 450 to create transcoded media assets. The transcoding service(s) 460 notify the kernel server 425 once transcoding is complete, and the kernel server 425 relays this information to the file replication service 430. The file replication service 425 then takes the transcoded media assets from the transcoding services 460 and moves them to the media asset origin 455. Once the file replication service 430 notifies the kernel server 425 of the move, the kernel server 425, in turn, notifies the file replication service 430 and the data object updater 420. The data object updater 420 which updates the data object origin 415 accordingly, and the file replication service 430 deletes the transcoded media assets from the transcoding services 460.

[0046] The modular nature of the system enables all tasks associated with an event to be completed quickly. As illustrated in the example above, workflow relating to a particular event, such as a media asset upload, can be spread among the various services simultaneously. Moreover, because the system's modularity enables it to be scaled to accommodate differing hardware capacities, and because the system can be configured to dynamically allocate hardware to different services according to the needs of the system, the speed of completing tasks relating to a particular event can further be increased. For example, a server of the CHIMPS 110 can be configured to dynamically switch its purpose based on external conditions such as load and overall system performance, providing functions such as transcode, upload, metrics collection, application web service, and more, on an as-needed basis.

[0047] Embodiments of such systems may include other systems that manage various requests from end users. For example, a system for dynamic index file generation and media file chunking. Referring to **FIG. 5A**, shows an embodiment of such a system 500-1. Media may be streamed to end user device 140 though a client 510. As mentioned above, the client

510 can be stand-alone media player, a plug-in, a browser, or other application, which can be executed on a personal computer or other electronic device.

[0048] An index file generator 530, as discussed previously, can be a program instantiated for media streaming to a particular client 510. The index file generator 530 can be executed on a server or other computing device within an application center 112 of the CHIMPS 110. Index files generated by the index file generator can include a wide variety of information such as starting, ending, and or run times for media chunks and locations for media chunks. This information can be embedded in a single string of data, such as a URI or a URL. If media includes various sub-streams (e.g., streams with alternative bitrates, captions, alternative languages, etc.) the index file can include data for chunks corresponding to each of the alternative sub-streams, as well as information regarding the bitrate and/or other unique information for each stream. Alternatively or in addition, index files indicating alternative sub-streams may be separate from index files indicating one or more media chunks for streaming.

[0049] It should be understood that the index file can further comprise a wide variety of formats, which can depend on the particular protocol. HTTP streaming may, for example, require index files to comprise one or more of M3U, M3U8, XML, and XML-based formats. Of course, other formats can be used in accordance with relevant streaming protocols.

[0050] **Table 1** illustrates a simplified example of a generated index file in M3U9 format, indicating chunk of media for streaming. The index file in this example provides a URI for a chunk of media. The URI indicates the chunk is to be generated by dynamic segmentor 550, the chunk being 10 seconds long, starting at 9 seconds into the media file and ending 19 seconds into the media file.

Table 1: Example Index File Contents

<pre>#EXTM3U #EXT-X-MEDIA-SEQUENCE:1 #EXT-X-TARGETDURATION:10 #EXTINF:10, http://video.example.com/seg/9/19/seg1.ts</pre>

[0051] Referring again to FIG. 5A, the index file generator 530 can also include an indicator within an index file to indicate whether a chunk of media is to be dynamically created. If, for example, it is determined that a requested media asset has not been chunked

and that the asset will be chunked dynamically, the index file generator can include the indicator in data corresponding to a chunk of media to be created. The indicator, which can be as simple as including the term “/seg/” in a URL, will indicate that a requested chunk of media needs to be generated.

5 [0052] The chunks of media can be generated during media streaming by a dynamic segmentor 550, which can be incorporated into an HTTP service 540. The HTTP service 540, as well as the media asset origin 560 can be located within a kernel application center 111 of the CHIMPS 110 on, for example, a media asset origin server. The system 500-1 can be configured such that the kernel application center 111 provides dynamically-created
10 chunks of media to a CDN 150 for delivery to client 510. The CDN 150 can store the chunks locally in, for example, a media asset cache 520, thereby forgoing the need to dynamically create a chunk again if the same chunk is requested in the future.

[0053] In sum, the system for dynamic index file generation and media asset chunking 500-1 can, after receiving a request for an index file from a client 510, dynamically generate an
15 index file with an index file generator 530. The index file can, among other things, indicate where a next chunk of media may be located. A client can then request the chunk from the location indicated by the index file, which can comprise a media asset cache 520 in a CDN 150. If the chunk is not found in the media asset cache 520, the cache miss can redirect the request to a segmentor 550 of an HTTP service 540, which can dynamically generate the
20 requested chunk of media by accessing the corresponding media asset in the media asset origin 560. The requested media chunk can then be provided to the CDN 150 for storage in the media asset cache 520 and delivery to the client 510. If the same chunk is requested at a later point in time, the CDN 150 can deliver the chunk from the media asset cache 520, thereby forgoing the need to redirect the request to the segmentor 550 to regenerate the
25 chunk.

[0054] **FIG. 5B** illustrates an alternative embodiment 500-2 of a system for dynamic index file generation and media file chunking. Rather than utilize a CDN, this embodiment 500-2 includes a media caching server within an application center 112 of the CHIMPS 110. The media caching server can receive chunk requests from and provide the corresponding chunks
30 to a client. It will be understood that such a media caching server(s) or similar device(s) can be located anywhere within the CHIMPS and/or in a system(s) communicatively linked to the CHIMPS.

[0055] FIG. 5C illustrates an embodiment 500-3 of a system for index file generation used in conjunction with a redirector 590. As discussed above, an index file generator 530 may create an index file having one or more URIs or other location information directing the client 510 to one or more redirectors 590. Redirector 590, can then provide the URI of the requested chunk with a redirect, such as a redirect under HTTP status code 302. The URI can be located on a CDN 520 or other location (such as a media caching server 570) and/or dynamically created by the dynamic segmentor 550. It will be understood that there can be any number of redirectors 590, which can be at any location, including locations of the CHIMPS 110 such as the application center 112 (as shown in FIG. 5C) or the kernel application center 111. It also will be understood that the URI or other location information provided by redirector 590 can be generated by or provided to the redirector 590 in any number of ways. The URI can be generated, for example, based on the request received by redirector 590 from client 510. Finally, it will be understood that the CHIMPS 110 can be configured to dynamically implement any combination of embodiments 500-1, 500-2, and 500-3, further choosing whether to utilize one or more redirectors based on factors such as, for example, a detected type of client 510.

[0056] Embodiments utilizing one or more redirectors can have several advantages. For example, and not by way of limitation, if a certain client were implemented in such a way that it “reads ahead” to request chunks, it could result in incorrect reporting data. Thus, it would be advantageous to determine which chunk is actually requested by the client. Additionally or alternatively, where chunks are available in various sub-streams with different bitrates, determining the actual requested chunk can be useful in calculating Quality of Service (QOS) metrics. Furthermore, it there may be scenarios in which it is more efficient to create larger index files having many chunks comprising large segments of media, reducing the number of index files required to stream a media asset, and thereby reducing the processing requirements to create the index files. If encryption is used having, for example, a rotating key or a per client key encryption scheme in which a valid key might change during playback of a media asset, it also may be advantageous to incorporate redirector(s) for handling legacy keys for some period of time.

[0057] FIG. 6 illustrates a simplified flowchart of an embodiment of a method 600 for implementing a dynamic index for media streaming. The method 600, which can be executed by the index file generator 530, begins at block 610, where a request for an index file is

received from a client 510. According to media streaming protocols contemplated by this embodiment, if data regarding a chunk of media is not provided in an initial index file, the client will continue to request or refresh an index file until it reaches an indicator in the index file that signals the end of a stream. Thus, the method can be assured to receive more than
5 one request for an index file from a client provided that an initial index file does not include an indicator signaling the end of the stream.

[0058] At block 615, the method 600 additionally provides for receiving input from an advertising service. According to some embodiments, this input could be the availability of an advertisement, and can be provided by a service inside or outside the CHIMPS. In other
10 embodiments, the input could come from a service that factors in any of a variety of factors to indicate that a specific advertisement or type of advertisement should be shown. Or that any advertisement should be shown.

[0059] At block 620, a determination is made whether to include an advertisement in the next chunk. According to some embodiments, this determination can be made with or
15 without input from an advertisement service. It should be known that this determination can include the factors used by an advertisement service to provide the input of block 615.

Whether the determination includes input from an advertisement service of block 615 or not, the determination can still include factors such as information about an end user collected before or during streaming of the media. This can include behavior of the end user during
20 streaming of the media (as determined, for example, by machine-based logic through beaconing data and/or requested chunks of media provided by a client 510). Factors can also include information regarding the media asset used for streaming (such as type of content or preferred points within the media for an advertisement), preference(s) and/or selection(s) of an end user, when a previous advertisement was shown, time of day, and more. It can further
25 include information regarding the source of a media asset, such as who created and/or provided the asset for viewing by an end user. It will be understood that other embodiments contemplate include secondary media, other than advertisements into the media stream in this manner. Moreover, the secondary media and/or advertisement can be of any length and also may be chunked. Thus, it may be determined that the next chunk includes all, or a select
30 portion, of an advertisement of any specific length.

[0060] An index file is created based on the request as well as the determination of whether media, such as an advertisement, should be streamed, indicated by block 625. As discussed

above, the index file can assume a variety of formats and include any amount of information regarding a next chunk of media for streaming. For example, HTTP streaming can utilize index files having the URLs of available chunks of media. Information can be embedded in these URLs to indicate a location to download the corresponding chunk of media, starting
5 point and/or ending point of a chunk of media, an indicator to indicate whether the chunk is to be dynamically created by a segmentor 550, a location of an advertisement to be streamed, and more. This information is included in the index file and sent to the client at block 630.

[0061] At block 635, reporting data can be created based on information included in the index file. As previously discussed, information included in an index file and/or index file
10 request can indicate the behavior of an end user device 140 during streaming, such as a pause, stop, skip, play, etc. of the media. According to some embodiments, this information can be extracted from requests for an index file and/or providing the requested index file. The information can be gathered in addition to or as a substitute for beaconing data provided by a client 510. Moreover, if beaconing data is provided, the creation of reporting data may be
15 omitted altogether.

[0062] Reporting data can include any amount of information regarding end user behavior, as indicated through index file requests and/or provided index files. This can include a particular action and when it was performed. Additionally or alternatively, the data may be kept in a more rudimentary form, depending on the application or embodiment, indicating the
20 data included in an index file request and/or an index file. This reporting data may be stored in a log file for reporting after streaming and/or transmitted during streaming to a relevant service that collects such metrics.

[0063] As indicated by block 640, the reporting data may be sent to a metrics collector for analytics. A metrics collector, according to certain embodiments, may be an application
25 executed by a server from within the application center 112 in which the index file generator 530 is executed, or it may be executed elsewhere, such as in a kernel application center 111 or in a system outside the CHIMPS 110. Depending on the form of the reporting data, the metrics collector can further process and store the information.

[0064] **FIG. 7** illustrates a simplified flowchart of an embodiment of a method for
30 dynamically chunking a media file for streaming 700. This method can be employed by a variety of systems and/or programs. For example, it may be executed by a dynamic

segmentor 550 of an HTTP service 540 running on a server located in a kernel application center 111 of the CHIMPS 110.

[0065] The method 700 can begin at block 710, when a request for a chunk of media is received from a CDN 150. As discussed above, this request may be made in response to a cache miss at the CDN 150 and/or because an indicator was included in the request for the chunk of media that the chunk was to be created dynamically. As discussed herein, if the CDN 150 has the requested chunk cached from a prior request, the CDN 150 can provide the requested chunk and preclude the need to send the request to a dynamic segmentor 550 to generate the chunk. It should be understood that the request may come from sources other than a CDN 150 according to alternative embodiments. One such source includes the media caching server 570 of embodiment 500-2, as shown in FIG. 5B.

[0066] The starting and ending points of a requested chunk of media are then determined at block 715. This information can be included directly in the request or derived from the request, a previous request, and/or other sources. At block 720, the information, as well as information identifying the requested chunk of media, can be used to retrieve all or part of the relevant media asset from a media asset origin 560. The retrieved portion will include at least the relevant media from the starting point to the ending point of the requested chunk of media.

[0067] At block 725, the requested media chunk is generated by converting the relevant portion of the media asset into a deliverable chunk. The media asset, as stored in the media asset origin, may not be chunked; it may be stored in its entirety as a media file (or group of alternative files corresponding to alternative sub-streams). Generating the chunk therefore can require determining the starting and ending points from the retrieved portion of the media asset and converting the resulting segment of media into a deliverable chunk.

[0068] Although the generation of the deliverable chunk may involve transcoding, it may not. The media asset can be stored in a format where transcoding may not be needed, thereby reducing the processing requirements for creating chunks of media during streaming. For example, media assets may be stored such as H.264 or MPEG-4 video format and/or AAC, HE-AAC, or MP3 audio format. According to some streaming protocols, such as some forms of HTTP streaming, chunks of media in these formats would not need transcoding before being wrapped in an MPEG-2 transport stream container format. Instead, such a conversion essentially would require the addition of metadata to create the streaming format from the

format of the stored media asset. In other words, generating a deliverable chunk of media may only require identifying the stored media asset, extracting the relevant segment of the media from the media asset, and adding certain metadata in accordance with a container format. This process requires little processing power and can be easily performed on the fly during streaming. Once the deliverable chunk of media is generated, it is sent to the CDN 150 or other requesting entity, at block 730.

[0069] FIG. 8 illustrates a simplified swim lane flowchart describing the interaction of components in a system configured to provide dynamic indexing and chunking for media streaming, according to one embodiment. In this embodiment, a client can send a request for an index file 805, the request received by an index file generator 810. A particular request may be made to initiate the streaming of a media asset, or it may be during streaming. Depending on the streaming protocol, the request may be made while a client plays a chunk of media previously downloaded during streaming.

[0070] The index file generator generates an index file to indicate the next chunk of media 815. As described above, this chunk may include an advertisement, and the index file can include any amount of information about a chunk of media, including information regarding alternative sub-streams for streaming. The dynamic index file generator can include information regarding existing chunks of media, and, when used in conjunction with a dynamic segmentor may also include information regarding chunks of media that may need to be created. As detailed above, if a chunk of media is to be generated dynamically, the index file generator may indicate this by including an indicator in the generated index file, such as in a URL for one or more chunks described within the index file. Once the index file is generated, the index file generator sends the index file 820, which is received by the client 825.

[0071] Alternative embodiments may provide for the generation of index files containing more than a next chunk of media. For example, an index file generator may generate an index file containing information regarding several chunks of media, in which case the chunks of media can be dynamically generated by a dynamic segmentor when requested by the client. The determination of whether to include information regarding more than a next chunk of media can include factors such as whether the index generator is generating reporting data, the desired frequency of such reporting data, and more.

[0072] Using information contained in the index file, the client can then request the next chunk of media 830, and this request can be received by a CDN 835. The CDN then checks to see if the chunk is already stored in the cache 840. If so, the CDN can provide the requested chunk to the client, blocks 845 and 850. The requested chunk may be found in a
5 CDN's cache if the chunk was created and stored in the CDN during preprocessing or if the chunk was dynamically created and stored in the CDN from an earlier request.

[0073] If the chunk is not found on the CDN, the chunk can be requested of the dynamic segmentor, which receives the request 855 and retrieves the corresponding media asset from an asset origin server 860. As discussed above, the entirety of the relevant media asset does
10 not need to be retrieved as long as at least the portion containing the relevant segment for the requested chunk is retrieved. It will be understood that alternative embodiments can provide for the media asset being stored in a variety of locations accessible, directly or indirectly, to the dynamic segmentor.

[0074] The dynamic segmentor can then generate the requested chunk by converting the
15 retrieved media into a deliverable chunk. That is, the dynamic segmentor converts the retrieved media into an acceptable format for streaming, which can vary depending on the streaming protocol utilized. The dynamic segmentor can then return the requested chunk 870 to the CDN, which can cache the chunk and return it to the client 875. Once the chunk is received by the client 880, the client can play the chunk to an end user.

[0075] The ability to dynamically generate chunks can enable the CHIMPS 150 to provide
20 dynamic audio track selection for media streaming. Traditional chunking methods typically require media file chunks to be created and stored beforehand, in which case separate chunks must be created and stored for each audio track (causing storage inefficiencies by storing chunks with each audio track separately) or each chunk must include a plurality of audio
25 tracks (causing bandwidth inefficiencies by transmitting chunks with audio tracks that will not be used). The dynamic segmentor's ability to create chunks of a media file in real time, however, can enable the CHIMPS 150 to avoid these inefficiencies by creating requested chunks on demand, having only the requested audio track.

[0076] FIGS. 9A-9C illustrate some examples of how audio and video components of a
30 media file can be stored and combined to provide for dynamic audio track selection for media streaming. In the embodiment shown in **FIG. 9A**, for example, the video and audio components of a media asset are stored in separate video 910 and audio chunks 920. The

dynamic segmentor can create media file chunks 930 on demand by selecting, for a requested media asset, the corresponding video 910 and audio 920 chunks, and combining the video and audio into one or more deliverable media file chunks 930 having a the requested audio track.

5 [0077] The URL (or other type of request) received by the dynamic segmentor can include information indicating a requested audio track. According to certain embodiments, this audio track information can be inserted into the URL by the client and/or index file generator. For example, a client media player application may have a user interface that enables a user to select an audio track from a list of available audio tracks. Audio tracks can include different
10 languages, audio channels (e.g., stereo, surround sound, etc.), and/or other options the like. The client can then include, in a URL requesting a chunk for media playback, an indication of the selected audio track. (In alternative embodiments, the media player can indicate the audio track selection in a request for an index file, in which case an index file generator can include an indicator of selected audio track in the URLs included in the index file.) The URL is then
15 provided to the dynamic segmentor, which retrieves the video chunks 910 and the audio chunks 920 corresponding to the selected audio track, and creates media file chunks 930 having the selected audio track. Because the dynamic segmentor can create the media file chunks 930 in real time, this process can accommodate audio track selection changes during playback of the media file. Furthermore, although URLs are used in the examples provided
20 herein, requests of other types can be used additionally or alternatively.

[0078] FIG. 9B illustrates another embodiment that includes combining a video file 940 with an audio file (having the selected audio track) into a media file. Although the video chunks 910 and audio chunks 920 of FIG. 9A can be stored in separate files as well, FIG. 9B illustrates how a video file 940 and audio files 950 can be combined in a media file 960,
25 which is then chunked into the media file chunks 930. Thus, the video file 940 and audio files 950 can include larger portions (e.g., all video/audio content) of the requested media file. Although FIGS. 9A and 9B illustrate multiple audio files 950 and audio chunks 920, it will be understood that some embodiments may include a single audio file/chunk having multiple audio tracks, or multiple audio files/chunks having one or more audio tracks.

30 [0079] FIG. 9C illustrates an embodiment utilizing a master file 970. In this embodiment, the master file 970 includes both video 980 and audio components (tracks) 990. In this embodiment, the dynamic segmentor can, after receiving a request for a chunk, retrieve and

chunk the master file 970. In instances where the master file has a plurality of audio tracks 990, the dynamic segmentor can combine the requested audio track with the video component 980 while ignoring the other audio tracks. Thus, the resulting media file chunks 930 can contain only the requested audio track.

5 [0080] FIG. 10 is a flow diagram illustrating a method 1000 for providing dynamic audio track selection for media streaming, according to one embodiment. The some or all of the method 1000 can be performed by any of a variety of systems, such as a dynamic segmentor of the CHIMPS 150 discussed previously, or other computer and/or processing system. Moreover, the method 1000 can be utilized to provide chunks and/or other segments of a
10 media file for media streaming.

[0081] At block 1010, a request is received for a segment (e.g., chunk) of the media file. As indicated previously, the request can come from a media player client, CDN, or another requesting entity, in the form of a URL or other resource indicator. The request can include information regarding a requested segment (e.g., start time, end time, audio track, etc.). This
15 enables a starting point and ending point of the requested segment of the media file to be determined, at block 1020.

[0082] At block 1030, a portion of the media file is retrieved from memory. This can be done in several ways, as shown in FIGS. 9A-9C. For example, the portion of the media file can include separate video and audio chunks and/or files. For embodiments in which audio
20 tracks are stored in different files, a requested audio track can be identified and only the file with the requested audio track may be retrieved. Alternatively, the portion of the media file can include a master file with both video and audio components. To be clear, the portion of the media file can also include video and/or audio components of all or a large portion of the entire media file, in which case multiple segments can be created from the same media file
25 portion retrieved from memory. Alternatively, as shown in FIG. 9B, the portion of the media file can comprise a media file with a single audio track, which can be generated and stored previously or generated after upon receipt of the request for the media file segment.

[0083] At block 1040, the requested segment of the media file with a single audio track is generated. For embodiments in which the portion of the media file retrieved at block 1030
30 includes more than one audio track, generating the requested segment of the media file may include selecting the requested audio track and/or ignoring other audio tracks. In other embodiments, generating the requested segment of the media file may simply involve

chunking a media file or combining audio and video chunks. Once the requested segment of the media file is generated, it can be provided at block 1050.

[0084] It should be appreciated that the specific method 1000 illustrated in FIG. 10 shows an example of providing dynamic audio track selection for media streaming. Alternative
5 embodiments may include alterations to the embodiments shown. For example, alternative embodiments may include including more than one audio track onto the requested segment of the media file. Furthermore, additional features may be added, removed, or combined depending on the particular applications. One of ordinary skill in the art would recognize many variations, modifications, and alternatives.

[0085] It should be noted that the methods, systems, and devices discussed above are intended merely to be examples. It must be stressed that various embodiments may omit, substitute, or add various procedures or components as appropriate. For instance, it should be appreciated that, in alternative embodiments, the methods may be performed in an order
10 different from that described, and that various steps may be added, omitted, or combined.

Also, features described with respect to certain embodiments may be combined in various other embodiments. Different aspects and elements of the embodiments may be combined in a similar manner. Also, it should be emphasized that technology evolves and, thus, many of the elements are examples and should not be interpreted to limit the scope of the invention.
15

[0086] Specific details are given in the description to provide a thorough understanding of the embodiments. However, it will be understood by one of ordinary skill in the art that the
20 embodiments may be practiced without these specific details. For example, well-known circuits, processes, algorithms, structures, and techniques have been shown without unnecessary detail in order to avoid obscuring the embodiments. This description provides example embodiments only, and is not intended to limit the scope, applicability, or
25 configuration of the invention. Rather, the preceding description of the embodiments will provide those skilled in the art with an enabling description for implementing embodiments of the invention. Various changes may be made in the function and arrangement of elements without departing from the spirit and scope of the invention.

[0087] Also, it is noted that the embodiments may be described as a process which is
30 depicted as a flow diagram or block diagram. Although each may describe the operations as a sequential process, many of the operations can be performed in parallel or concurrently. In addition, the order of the operations may be rearranged. A process may have additional steps

not included in the figure. Furthermore, embodiments of the methods may be implemented by hardware, software, firmware, middleware, microcode, hardware description languages, or any combination thereof. When implemented in software, firmware, middleware, or microcode, the program code or code segments to perform the necessary tasks may be stored
5 in a computer-readable medium such as a storage medium. Processors may perform the necessary tasks.

[0088] Having described several embodiments, it will be recognized by those of skill in the art that various modifications, alternative constructions, and equivalents may be used without departing from the spirit of the invention. For example, the above elements may merely be a
10 component of a larger system, wherein other rules may take precedence over or otherwise modify the application of the invention. Also, a number of steps may be undertaken before, during, or after the above elements are considered. Accordingly, the above description should not be taken as limiting the scope of the invention.

WHAT IS CLAIMED IS:

1. A server providing audio track selection for media streaming via a data network, the server comprising:

an interface for communicating with the data network;

a memory configured to store a media file; and

a dynamic segmentor coupled with the memory and the interface, the dynamic segmentor configured to:

obtain, via the interface, a request for a first segment of the media file;

determine, based on the request for the first segment of the media file, a starting point and an ending point of the requested first segment of the media file;

retrieve a first portion of the media file from the memory, the first portion of the media file having a video component and at least one audio track;

generate, without transcoding the first portion of the media file, the requested first segment of the media file such that the requested first segment of the media file has a single audio track; and

send, via the interface, the requested first segment of the media file.

2. The server recited in claim 1 wherein the dynamic segmentor is configured to generate the requested first segment of the media file during playback of the media file.

3. The server recited in claim 1 wherein the dynamic segmentor is configured to:

obtain, via the interface, a request for a second segment of the media file;

determine, based on the request for the second segment of the media file, a starting point and an ending point of the requested second segment of the media file;

retrieve a second portion of the media file from the memory;

generate, without transcoding, the requested second segment of the media file such that the requested second segment of the media file has a single audio track different than the single audio track of the requested first segment of the media file; and

send, via the interface, the requested second segment of the media file.

4. The server recited in claim 1, wherein the server is configured to store the video component and the at least one audio track in separate files.

5. The server recited in claim 1, wherein the dynamic segmentor is configured to:

generate a file comprising the media file with the single audio track; and
store the generated file in the memory;

wherein the first portion of the media file comprises the generated file.

6. The server recited in claim 5, wherein generating the file comprises combining audio from a first file and video from a second file.

7. The server recited in claim 1, wherein:
the at least one audio track comprises two or more audio tracks; and
the dynamic segmentor is configured to cause the server to select the single audio track from the two or more audio tracks based on information received in the request for the first segment of the media file.

8. A method of providing audio track selection for media streaming via a data network, the method comprising:

obtaining a request for a first segment of a media file;
determining, based on the request for the first segment of the media file, a starting point and an ending point of the requested first segment of the media file;

retrieving at least a portion of the media file, the at least a portion of the media file having a video component and at least one audio track;

generating, without transcoding the at least a portion of the media file, the requested first segment of the media file such that the requested first segment of the media file has a single audio track; and

sending providing the requested first segment of the media file.

9. The method recited in claim 8, further comprising:
obtaining a request for a second segment of the media file;
determining, based on the request for the second segment of the media file, a starting point and an ending point of the requested second segment of the media file;

generating the requested second segment of the media file, wherein the requested second segment of the media file has a single audio track different than single the audio track of the requested first segment of the media file; and
providing the requested second segment of the media file.

10. The method recited in claim 8, wherein retrieving the at least a portion of the media file includes retrieving the video component and the at least one audio track from separate files.

11. The method recited in claim 8, further comprising:
generating a file comprising the media file with the single audio track; and
storing the generated file;
wherein the at least a portion of the media file comprises the generated file.

12. The method recited in claim 11 wherein generating the file comprises combining audio from a first file and video from a second file.

13. The method recited in claim 8 wherein the at least one audio track comprises a plurality of audio tracks, the method further comprising selecting the single audio track from the plurality of audio tracks based on information received in the request for the first segment of the media file.

14. A server providing audio track selection for media streaming via a data network, the server comprising:

means for obtaining a request for a first segment of a media file;
means for determining, based on the request for the first segment of the media file, a starting point and an ending point of the requested first segment of the media file;

means for retrieving at least a portion of the media file, the at least a portion of the media file having a video component and at least one audio track;

means for generating, without transcoding the at least a portion of the media file, the requested first segment of the media file such that the requested first segment of the media file has a single audio track; and

means for sending providing the requested first segment of the media file.

15. The server recited in claim 14 wherein the at least one audio track comprises a plurality of audio tracks, the server further comprising means for selecting the single audio track from the plurality of audio tracks based on information received in the request for the first segment of the media file.

1/13

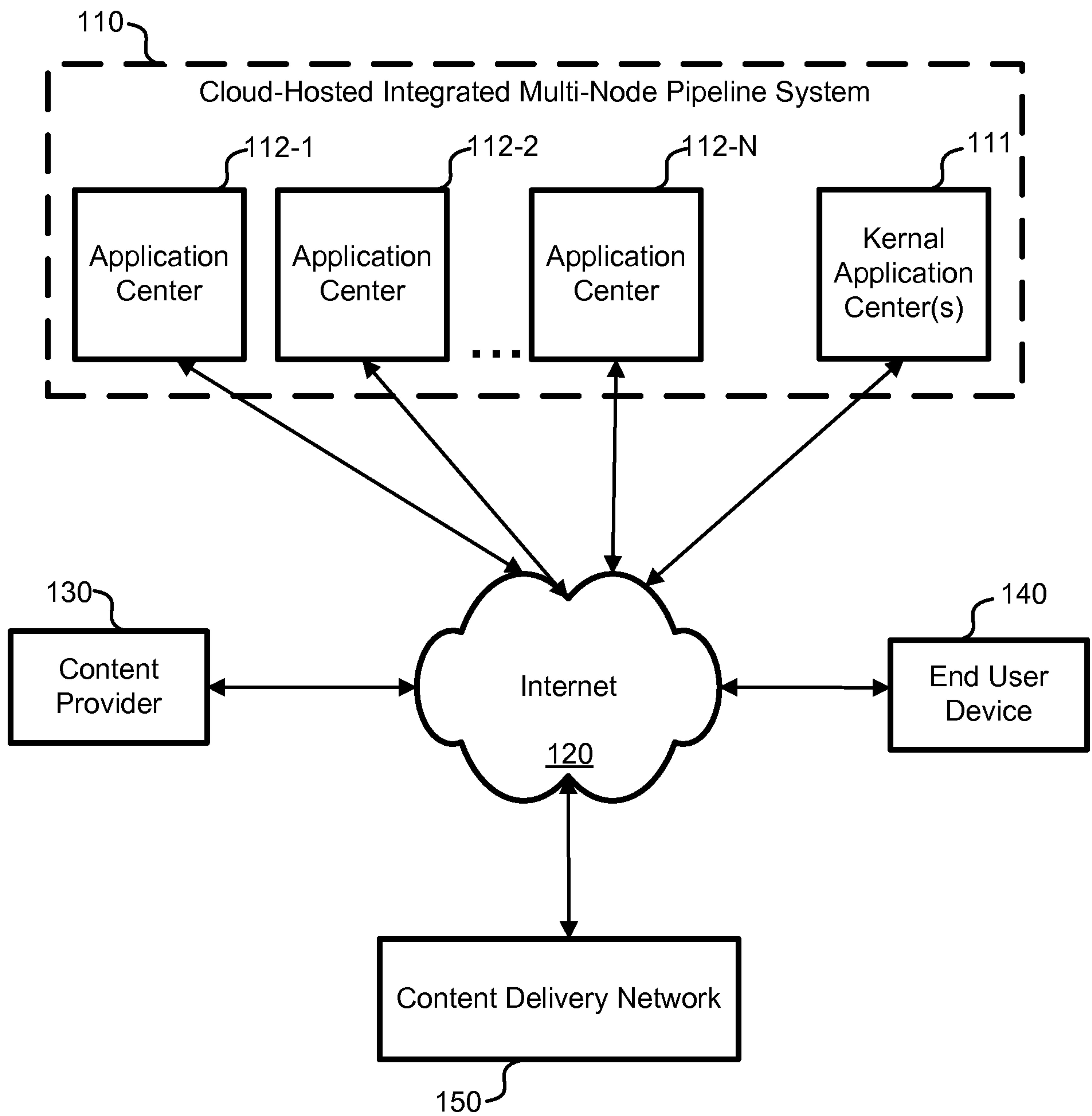


FIG. 1

2/13

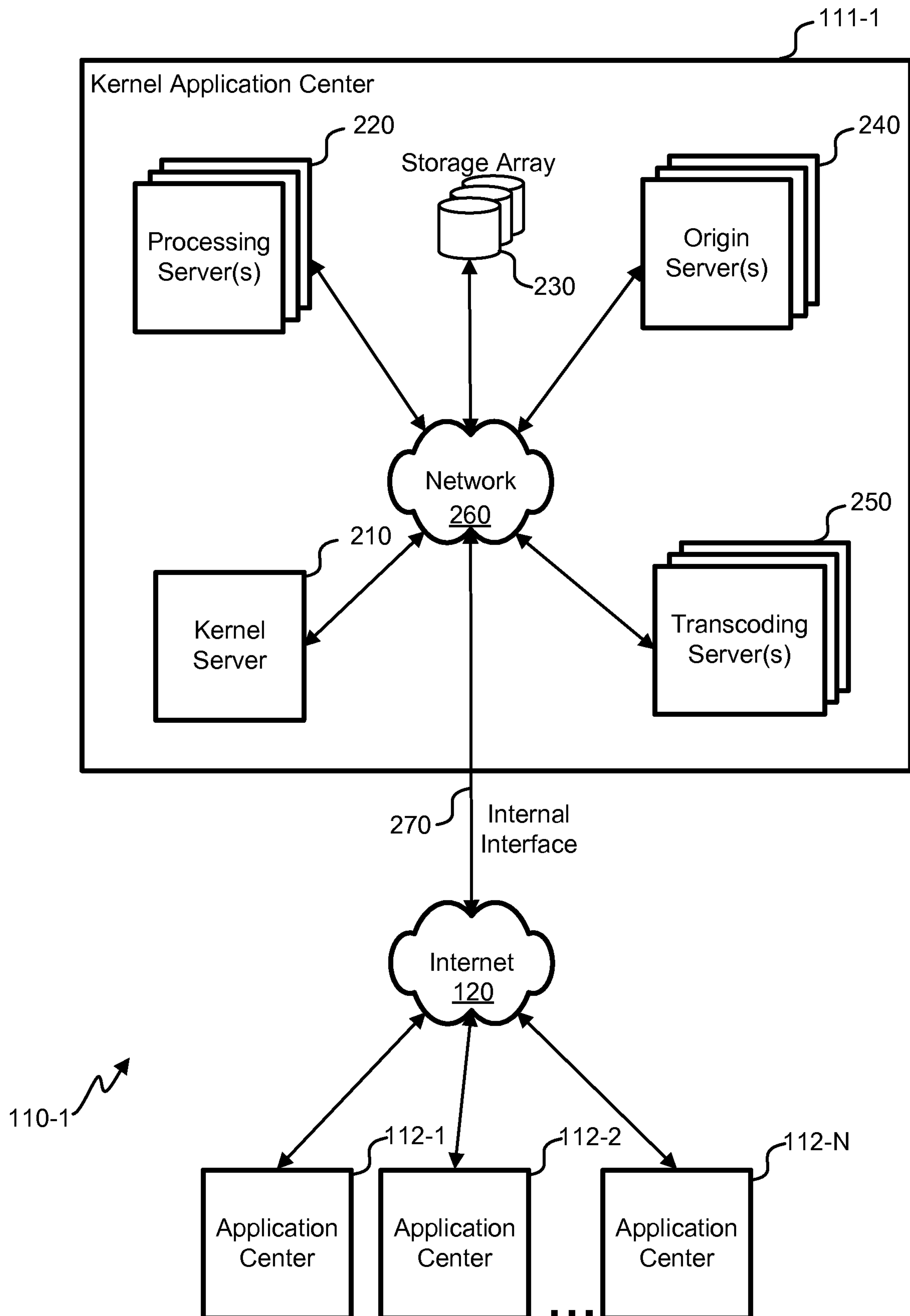
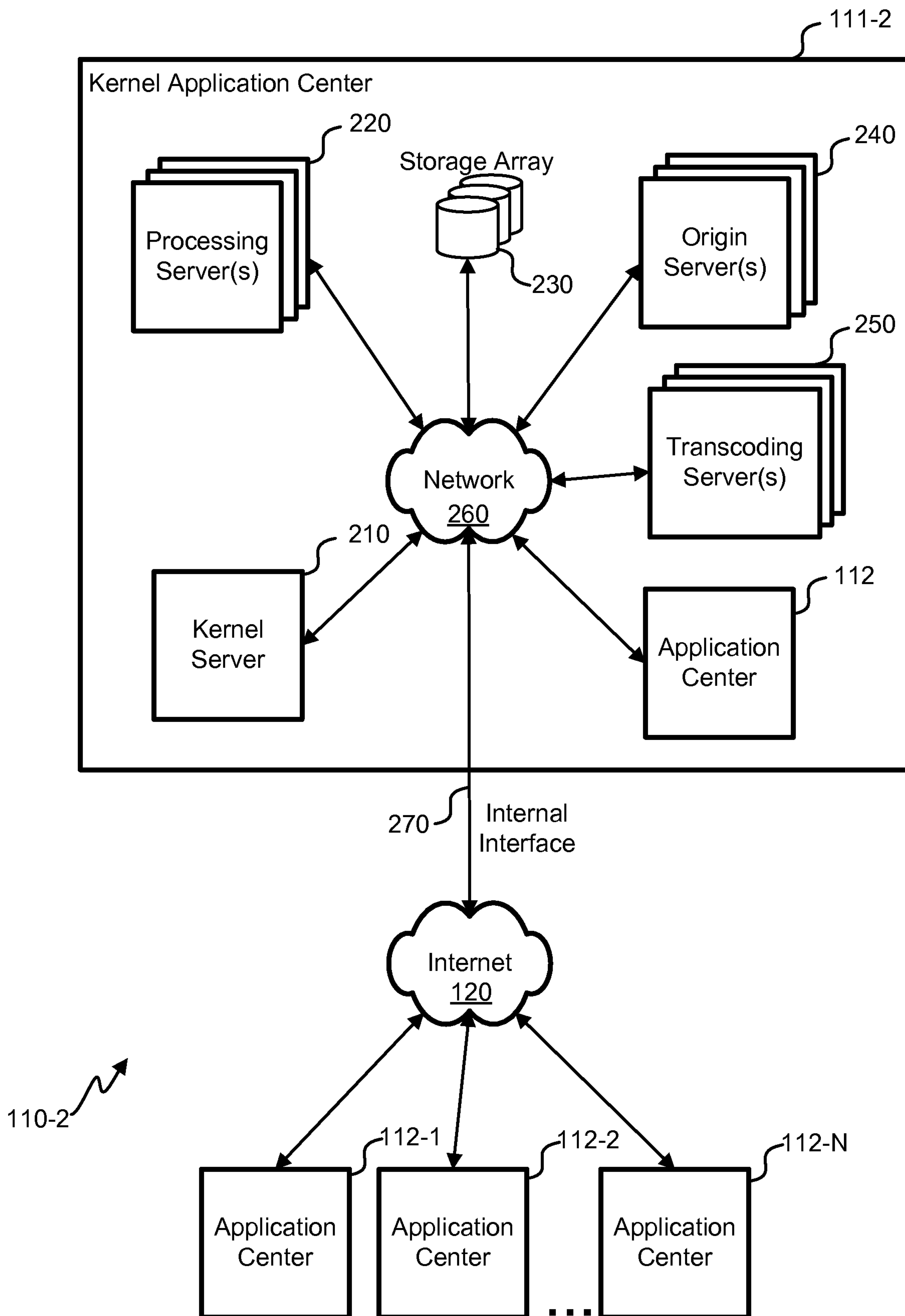
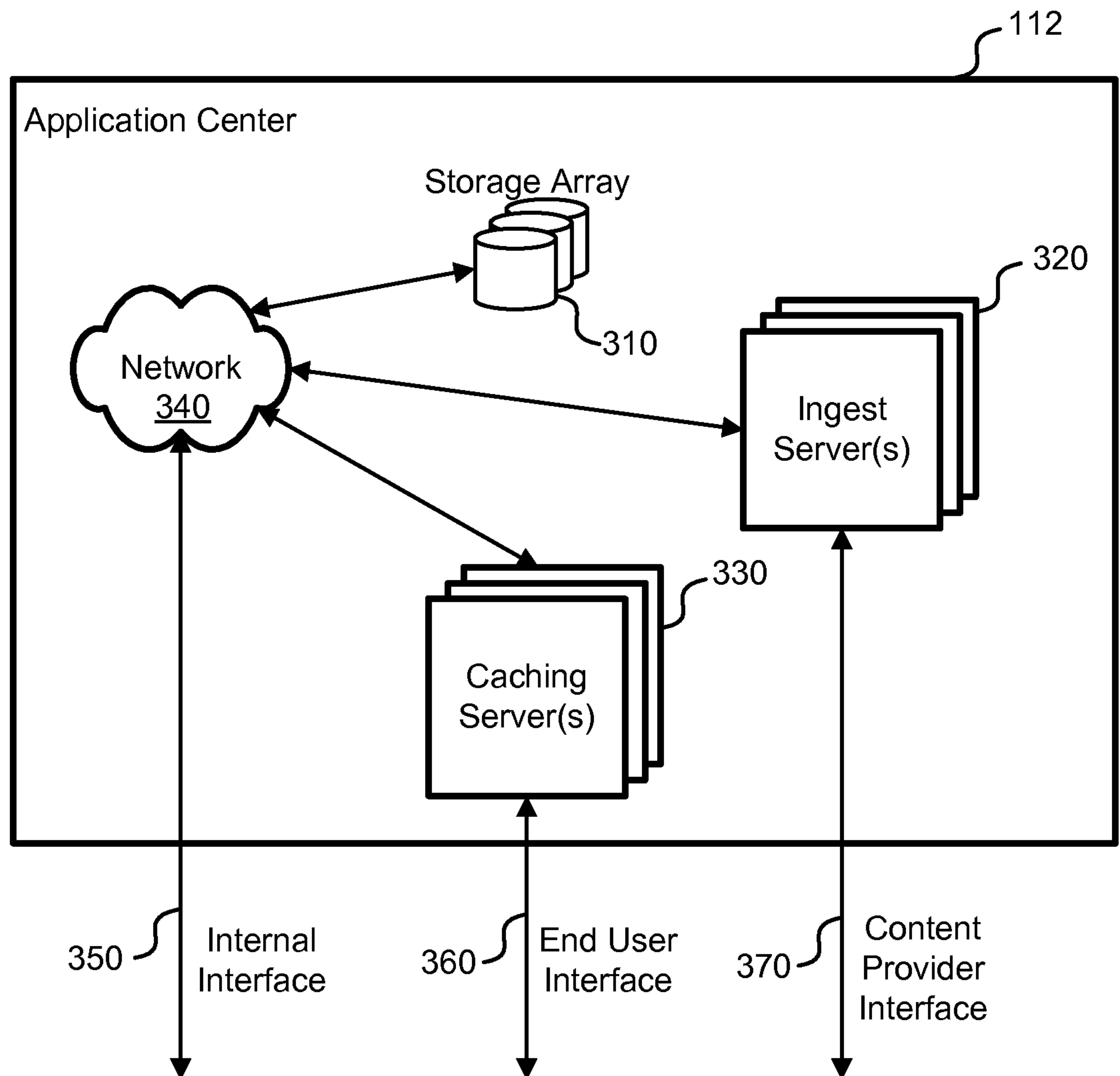


FIG. 2A

3/13

**FIG. 2B**

4/13

**FIG. 3**

5/13

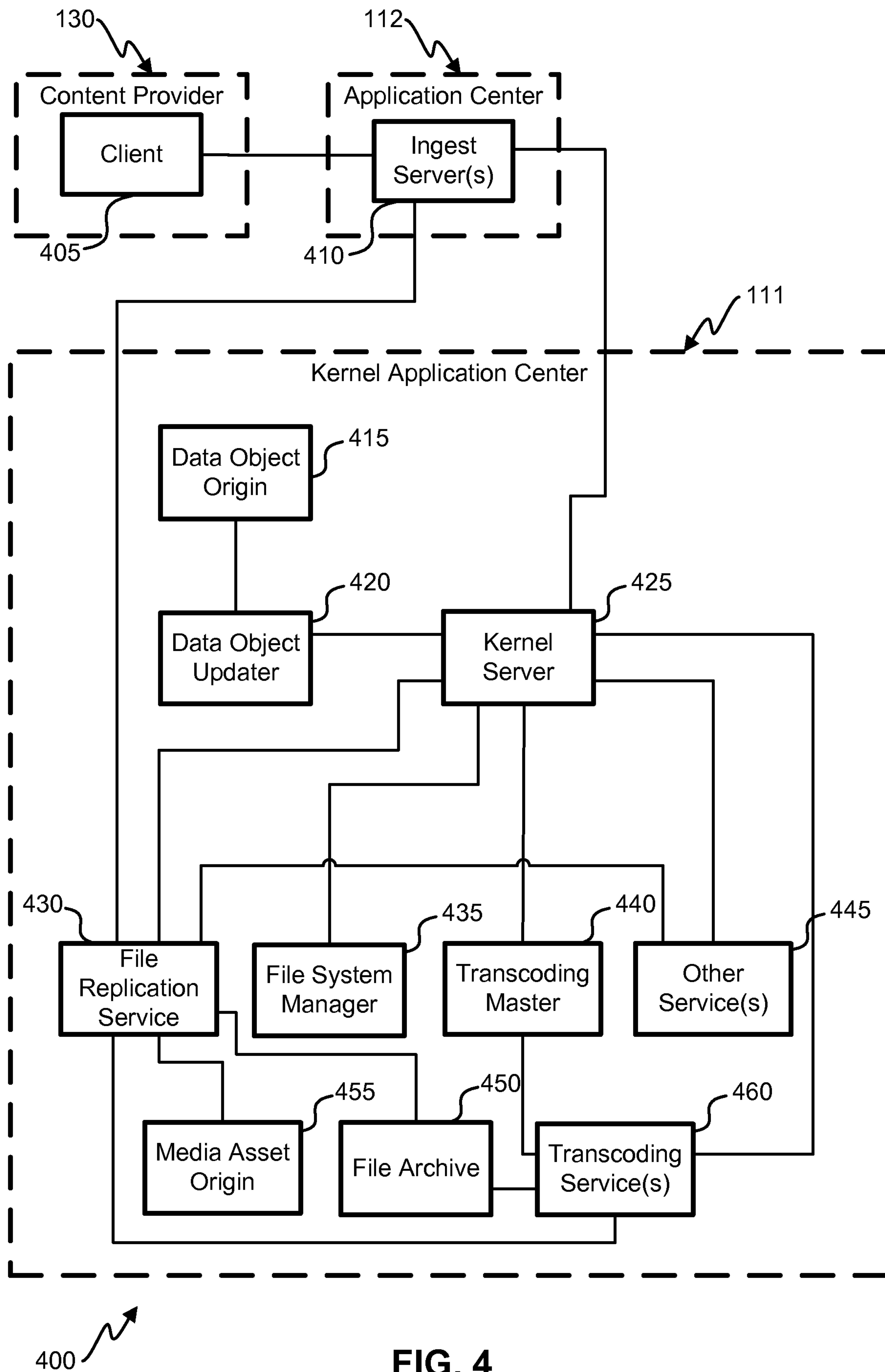
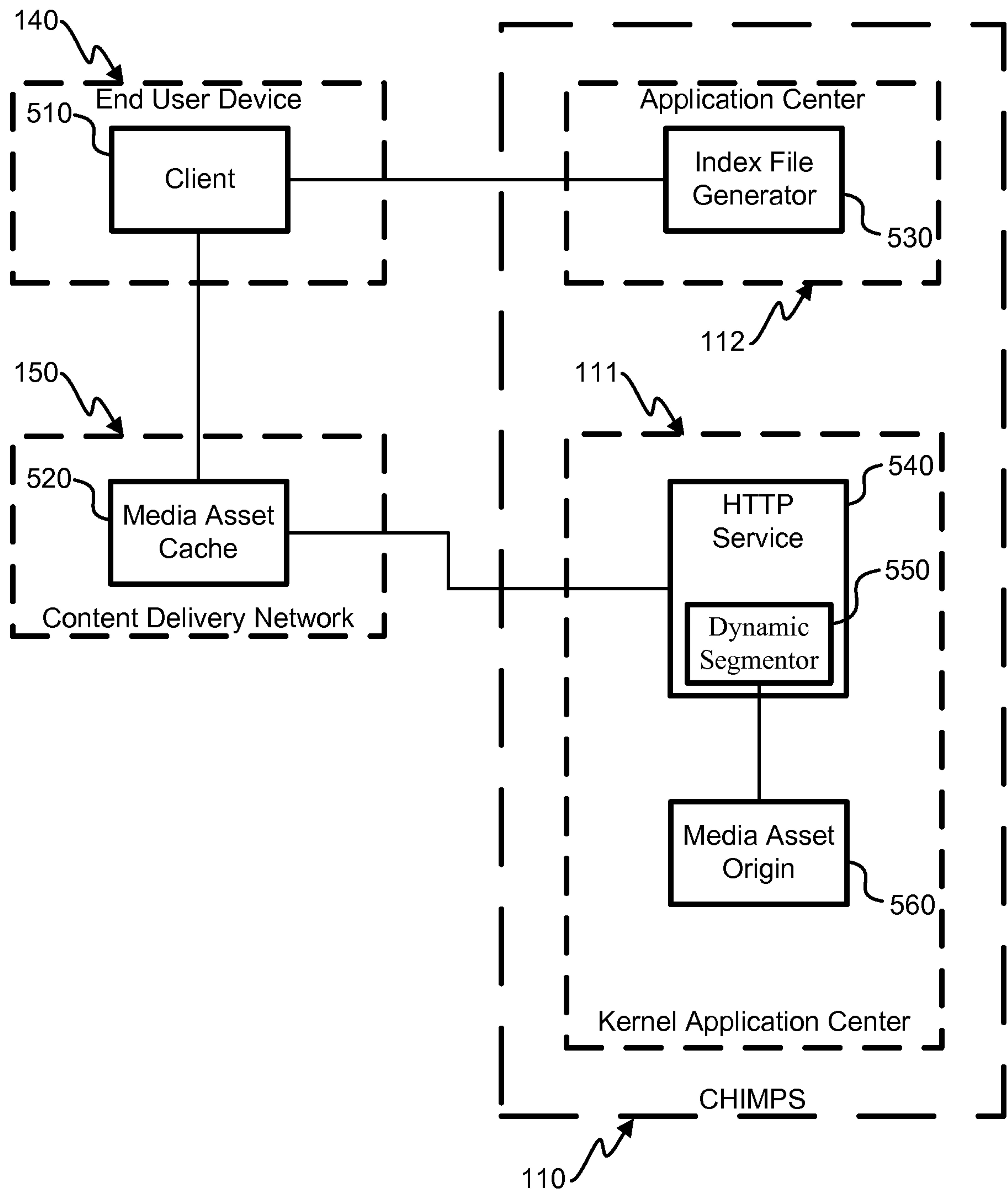


FIG. 4

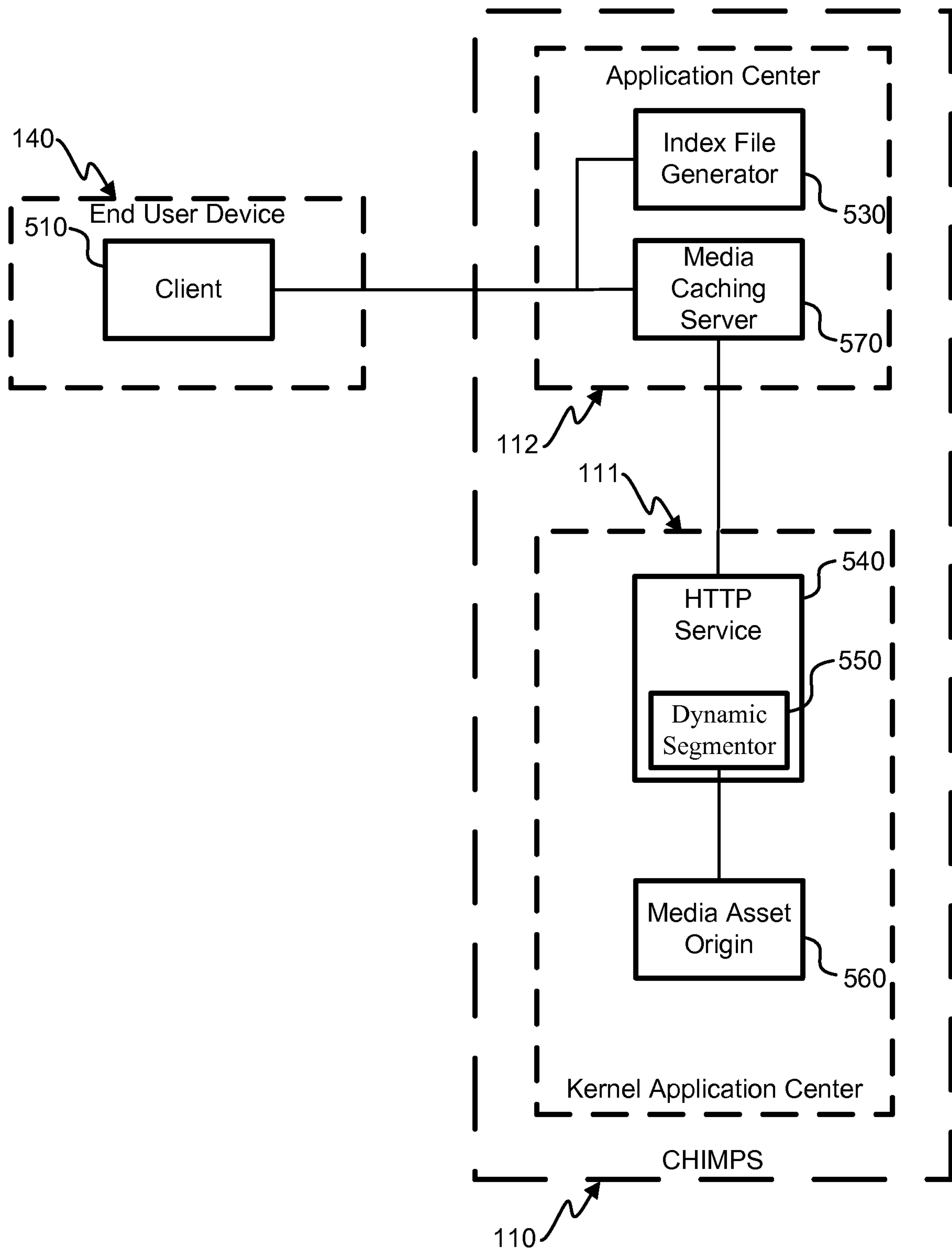
6/13



500-1

FIG. 5A

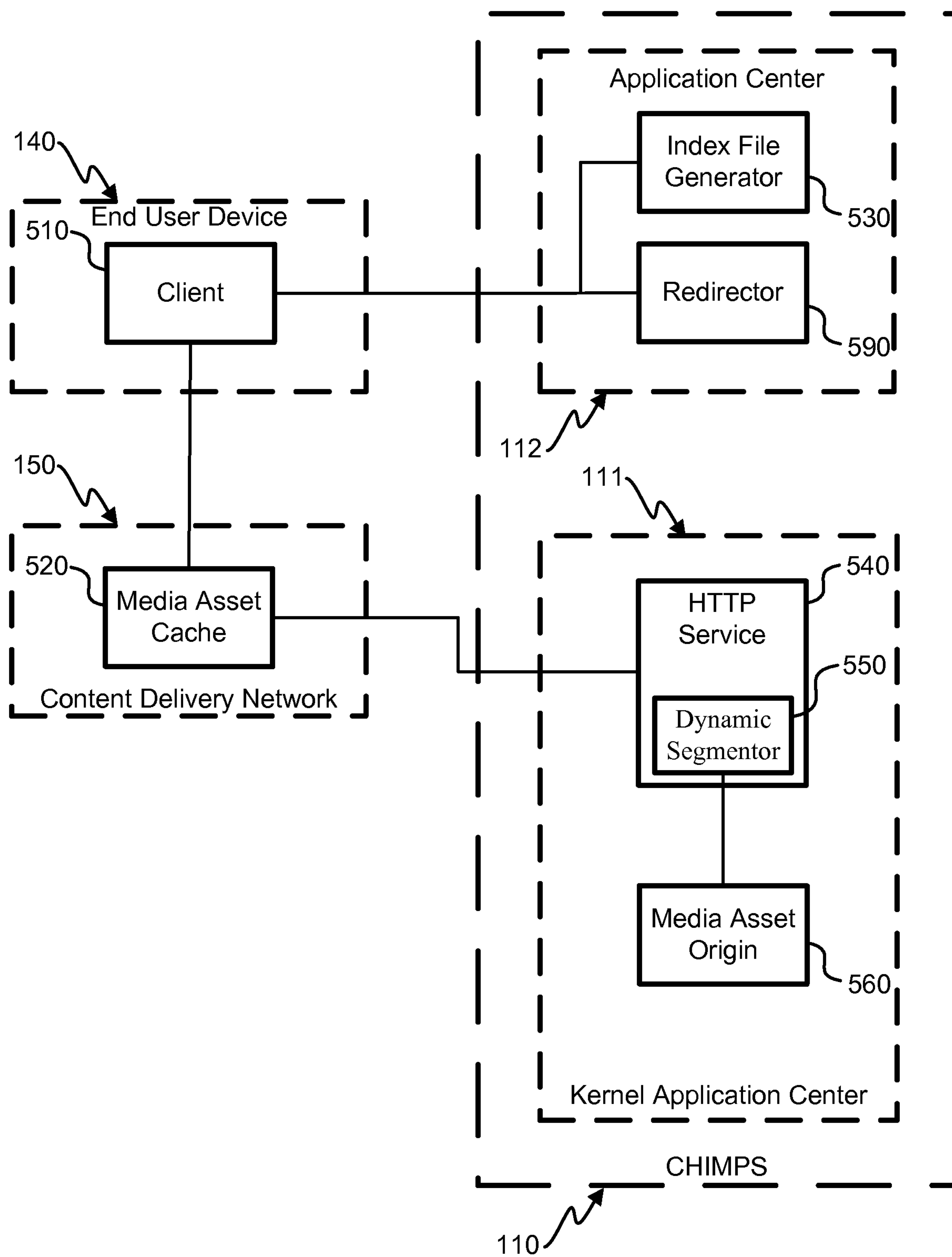
7/13



500-2

FIG. 5B

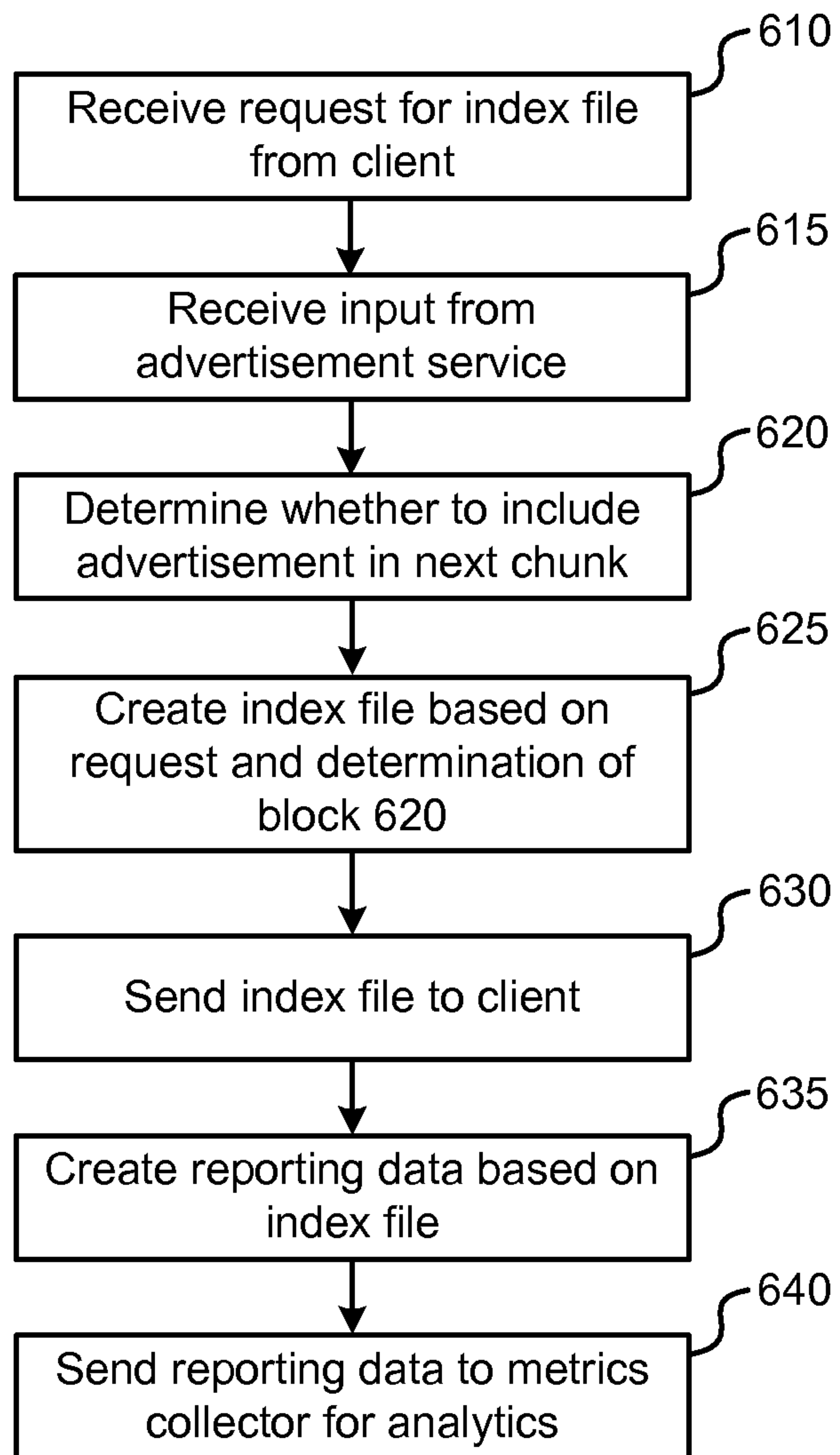
8/13



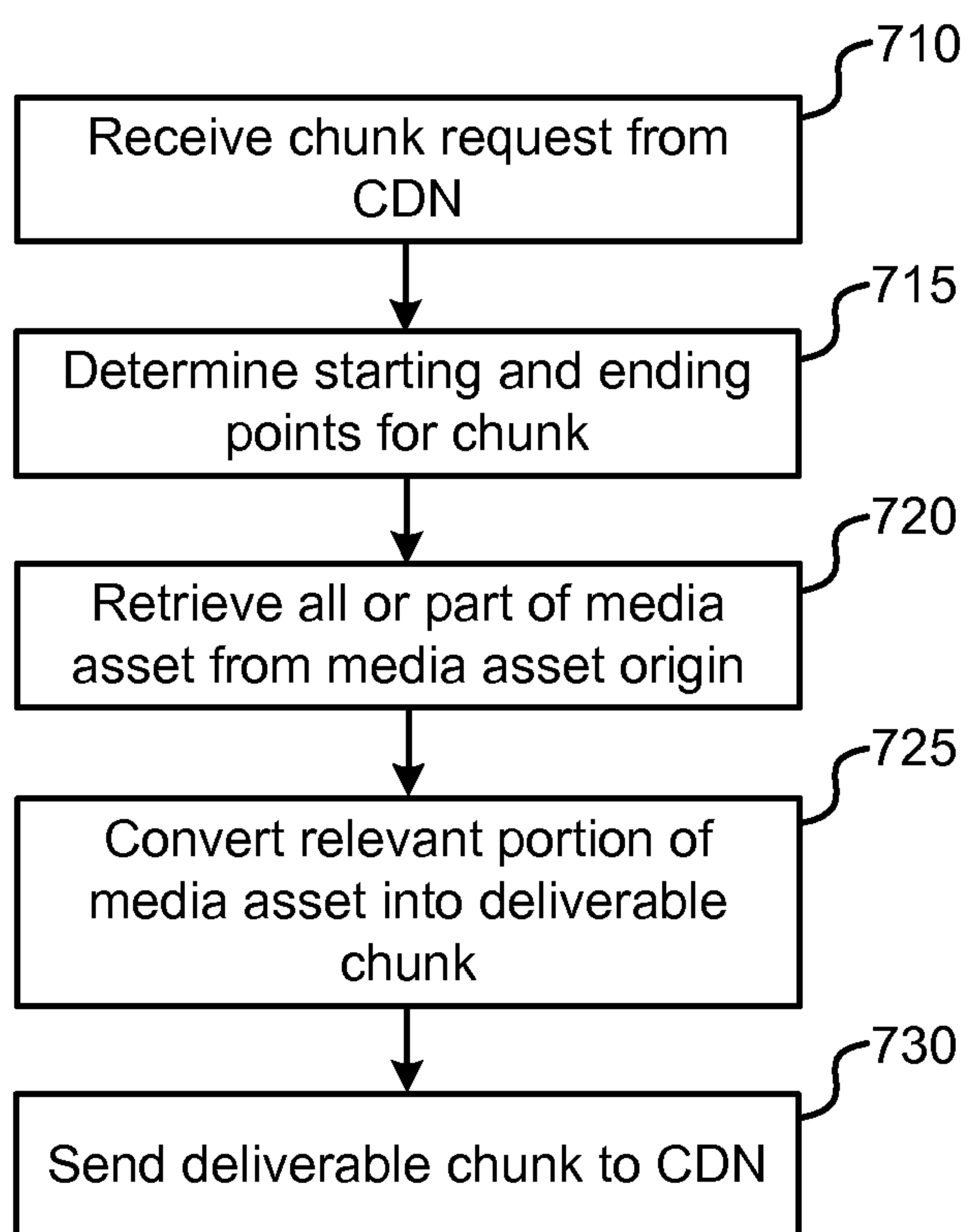
500-3

FIG. 5C

9/13

**FIG. 6**

10/13



700 ↗

FIG. 7

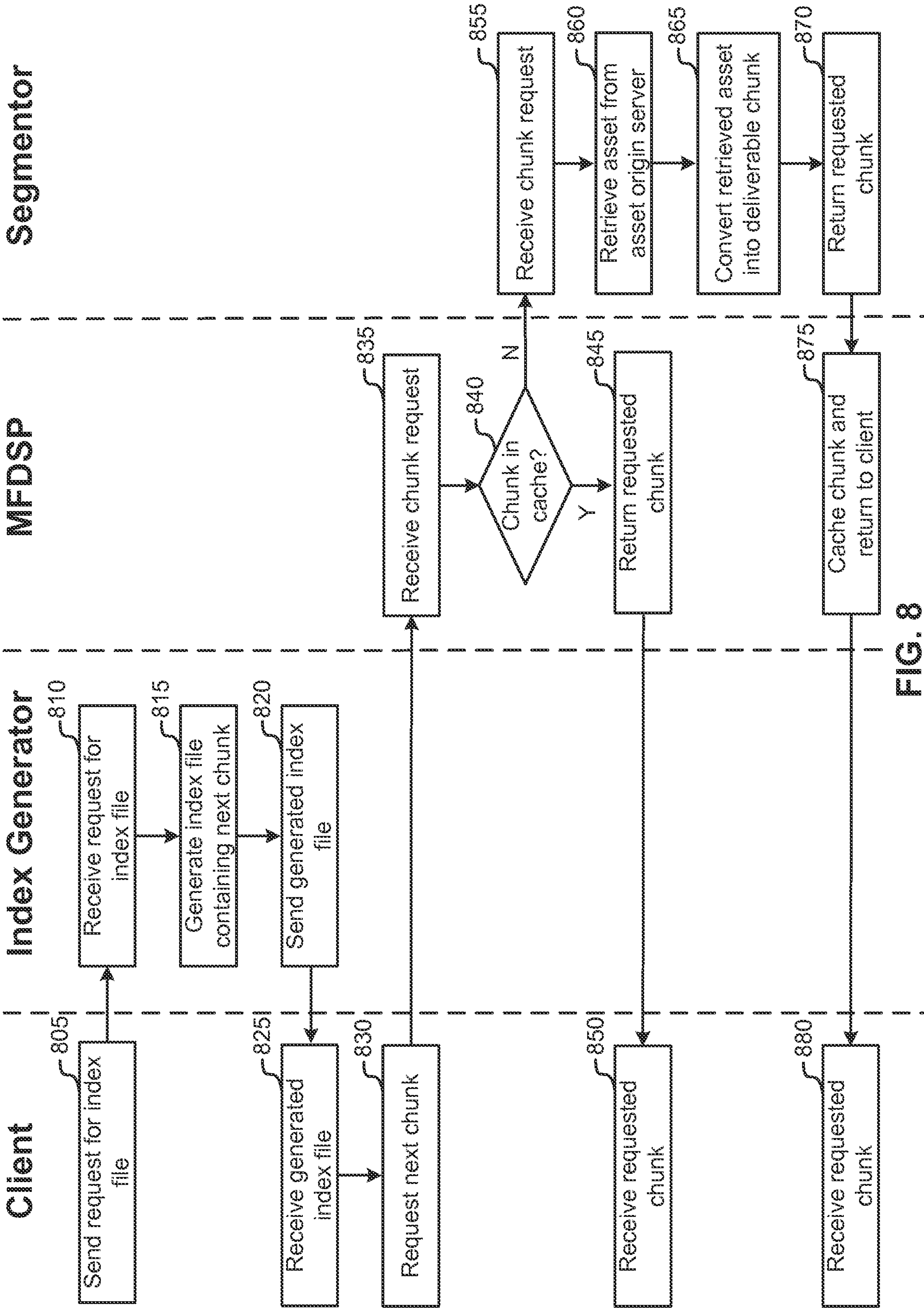
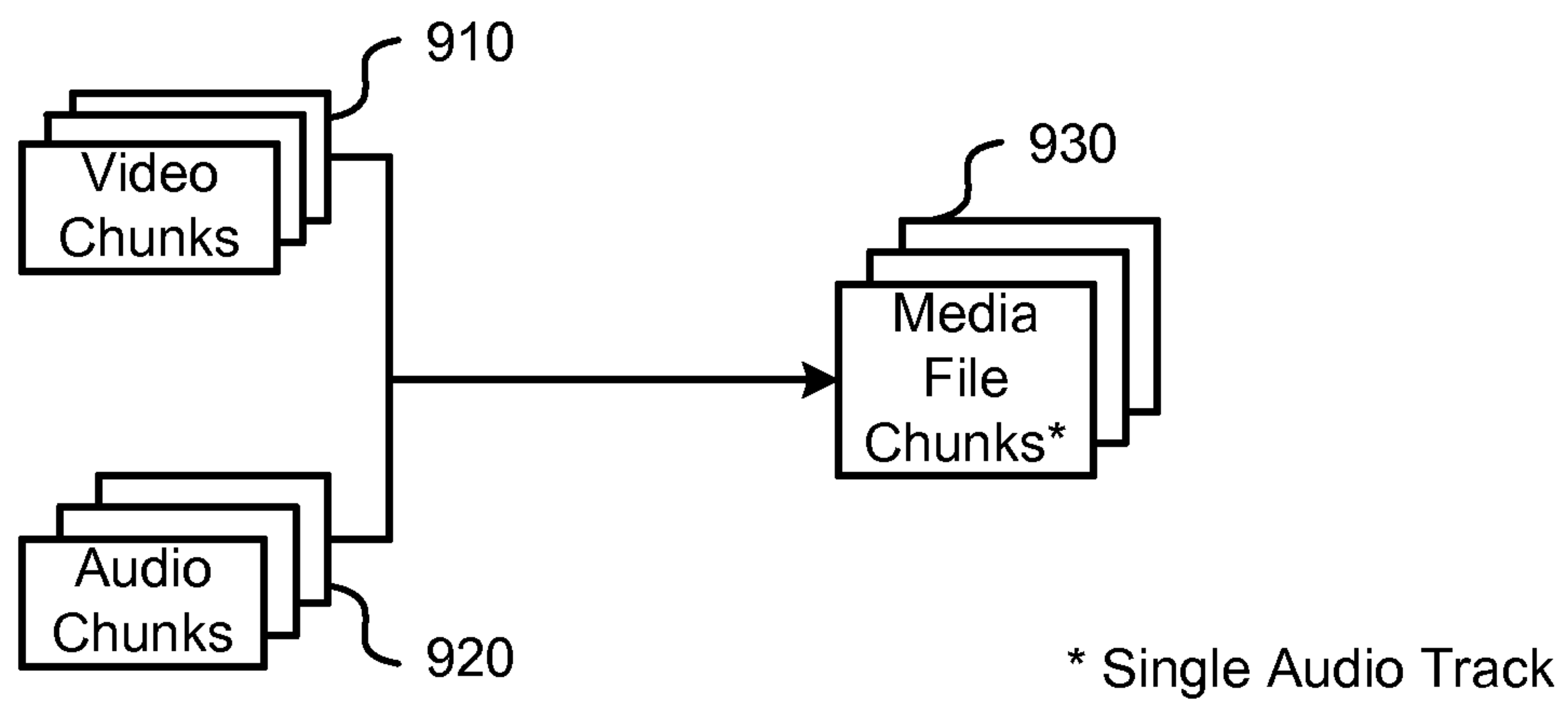
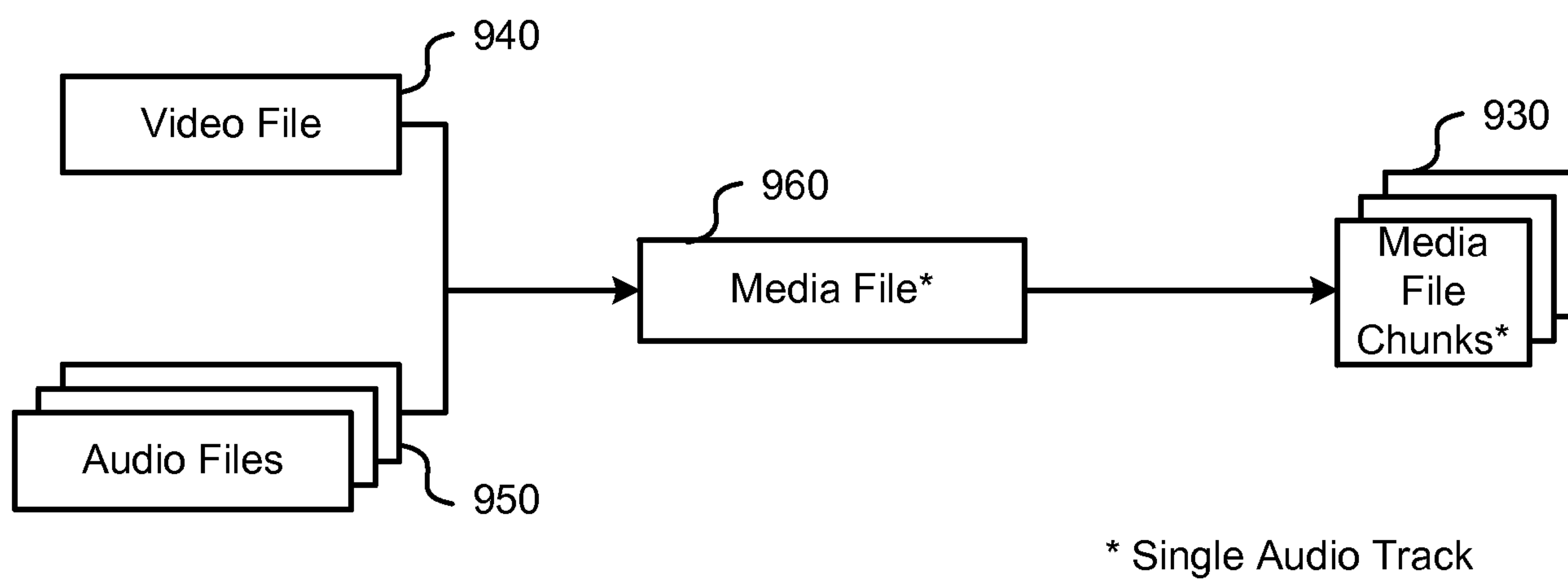
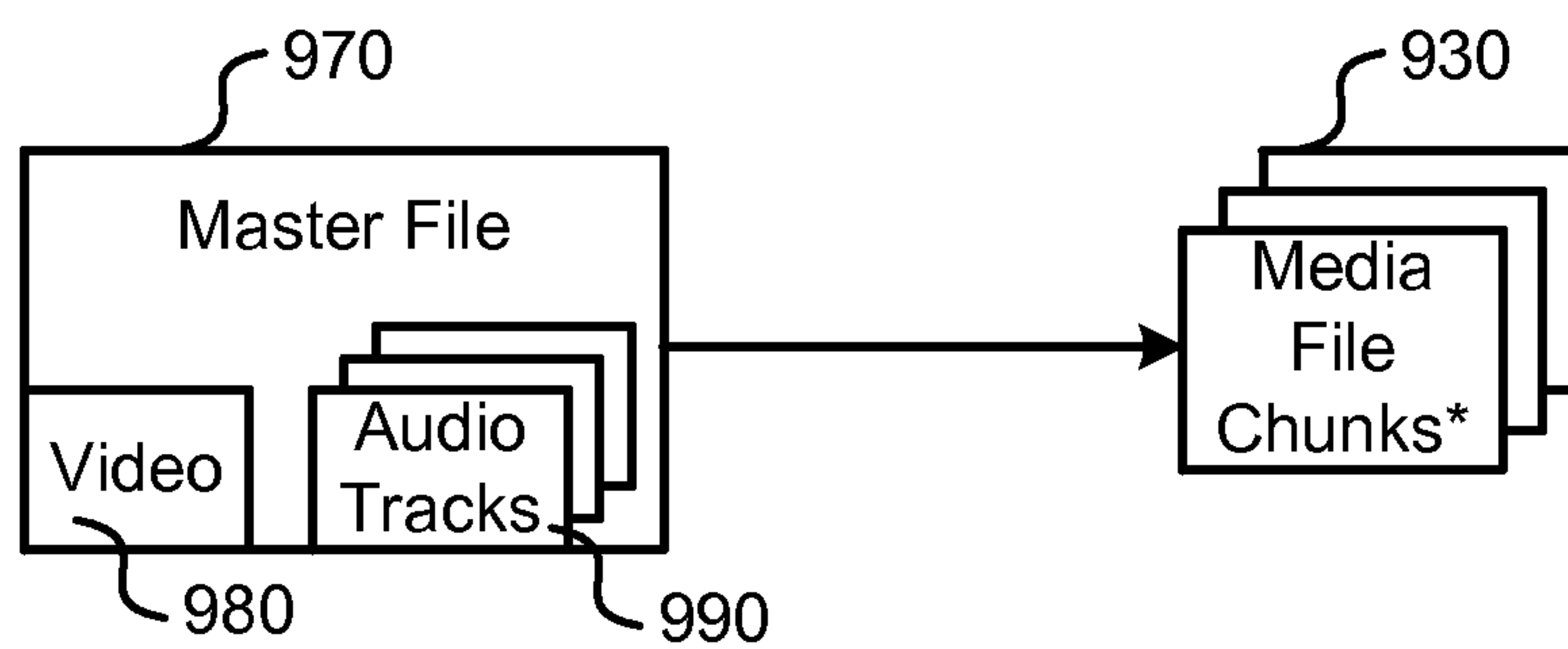
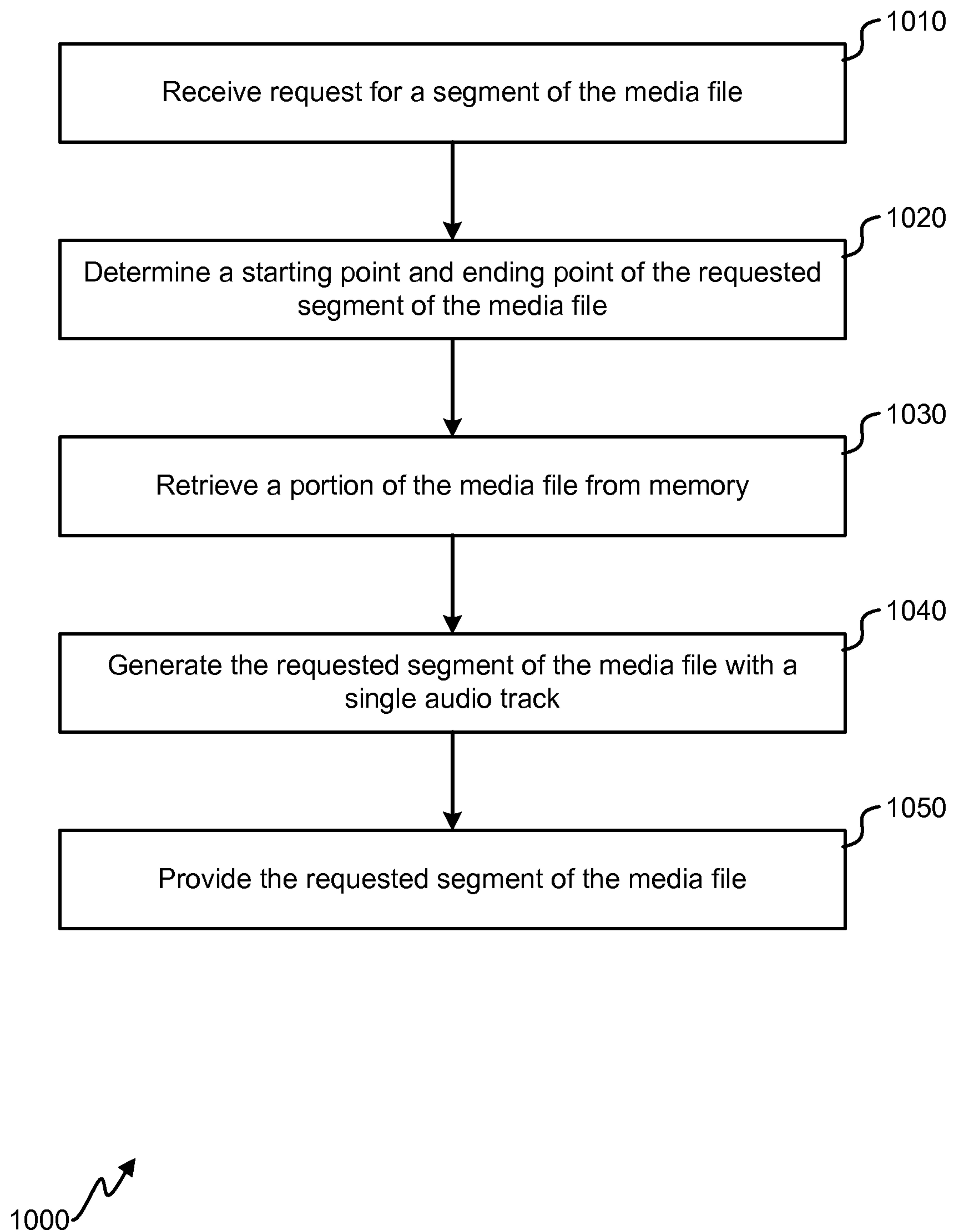


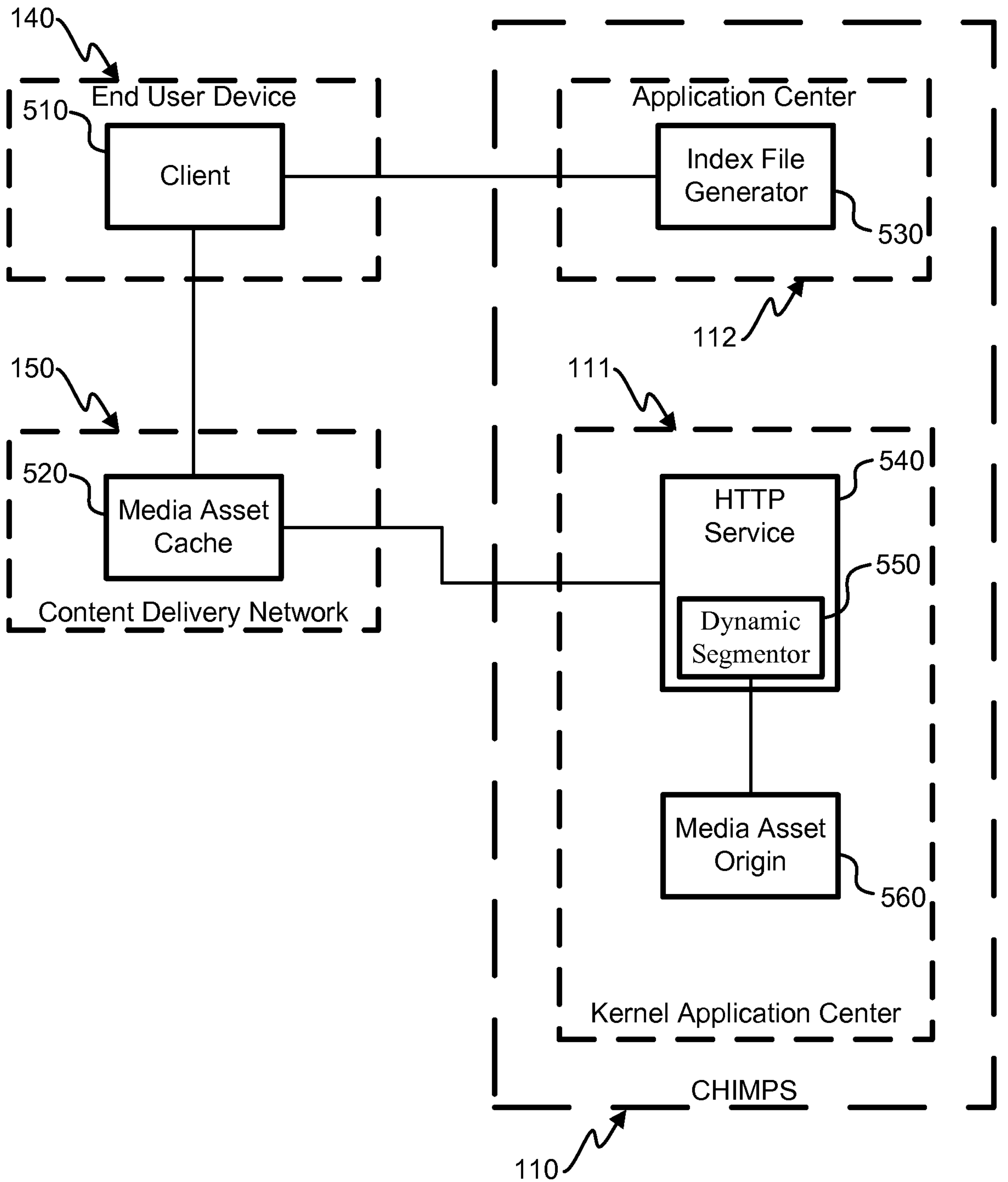
FIG. 8

12/13

**FIG. 9A****FIG. 9B****FIG. 9C**

13/13

**FIG. 10**



500-1

FIG. 5A