An efficient content distribution method using segment metadata annotation in CDN[†]

Jung-Eun Lim[1], O-Hoon Choi[1], Hong-Seok Na[2] and Doo-Kwon Baik[1*]

[1] Dept. of Computer Science & Engineering, Korea University, Seoul, Korea
[2] Dept. of Computer & Information Science, Korea Digital University, Seoul, Korea
{jelim, ohchoi, hsna99, baikdookwon}@gmail.com

Abstract. Recently, user who uses a multimedia files demands a high capacity file via internet. However, it is difficult to guarantee QoS about providing high capacity files by internet because that has an inconstant bandwidth. For guaranteeing the QoS, CDN (Content Delivery Network) is generally used as a method of contents delivery service. Based on CDN, we propose Content Distribution and Management using Metadata (CDM) system which can provide advanced transmission and searching method. For enhancing a transmission rate, CDM system supports segment unit based transmission method that can be possible parallel transmission. Also, we propose a distribution method through substance based search. CDM system can primarily distribute the files which includes content that user wants. Information about substance of file is stored in metadata through metadata management interface, and then the information is inserted to the high capacity file itself. When user requests a high capacity file that user wants, CDM system doesn't transmit all of applicable contents but firstly transmits a particular part of applicable files. Therefore a proposed CDM system can guarantee QoS for user.

Keywords: CDN, QoS, Metadata, Content Distribution, Searching

1 Introduction

Internet provides a relatively low quality streaming service than contents transmission method via existing broadcasting and storage, because internet has inconstant bandwidth. Now, CDN is used to satisfying a request about high quality streaming service of user contents use. CDN solved network bandwidth problem as decentralization contents distribution server [1,2]. CDN transmits contents as decentralizing multi servers in close location to user that internet bandwidth is guaranteed. But since CDN transmits high capacity contents file as file unit, it can't support firstly transmission via substance based search [3,4]. Transmission method using CDN has two problems.

- No method of searching based on substance
- Increasing of file transmission time

To use CDN effectively as solving above problems, we need a method for distributing contents files effectively. For this, this paper proposes Content Distribution and Management (CDM) system using metadata, it can apply in CDN. Roles of CDM system divides into contents distribution using file segment transmission method and management using metadata for substance search of contents [5].

2   Contents distribution and management (CDM) system using metadata

2.1   CDM system architecture

CDM system is installed into CDN staging server. Figure 1 shows architecture of CDM system. CDM consists of 4 modules and 3 interfaces to distributing and managing contents efficiently. Detail descriptions of Modules are below.
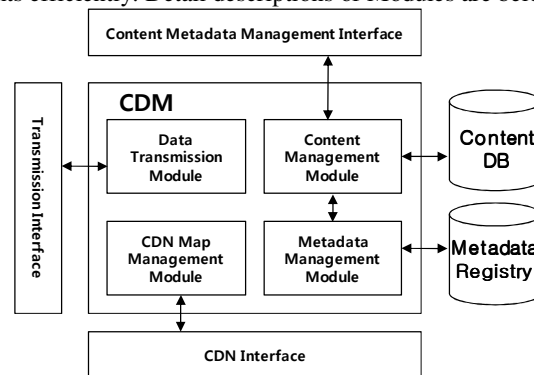

Fig. 1. CDM system architecture

Contents Management Module: It divides contents into low capacity contents as inserting metadata to contents metadata received by contents provider.

Data Transmission Module: It performs transmission error check, redundancy check occurred when file unit transmission and segment unit transmission.

Metadata Management Module: It manages metadata stored in metadata registry. It provides proper metadata to contents management module according to contents type.

CDN Map Management Module: This module enables CDM system to use map in CDM system. The map is distribution path of contents managed to CDN. CDM system transmits divided low capacity contents file according to distribution path defined in map. The divided low capacity contents file is manufactured via content management module, data transmission module, metadata management module.

CDN Interface: It is interface to use CDM system without change about staging server of CDN

Content Metadata Management Interface: It is interface for manager that is used to

divide high capacity contents file logically as low capacity contents files.
Transmission Interface: It is interface that transmission method can select about high capacity contents file. It supports file unit transmission and segment unit transmission as file transmission method.
Content DB: It is database for storing and managing contents provided from contents provider. Metadata inserted by content management module and related information are also stored to Content DB.
Metadata Registry: It is registry that provides standard metadata.

2.2   Contents transmission method based on segment

This paper applies segment unit file transmission method to solve the occurred problem when contents of file unit distributes and to re-transmit contents partially when parallel transmission and transmission error occurs. That is a method to transmit high capacity contents as segment unit and to manage transmission information of segment unit. For this, we should get three item of 'file ID' to recognize files according to request, 'Offset' to know start point of file,' File Size' to calculate transmission amount information of file. Figure 2 shows file transmission method of segment unit.
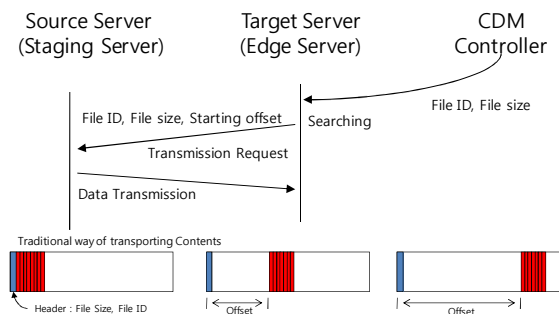
Fig. 2. Procedure of contents transmission based on segment

File transmission method of segment unit can prevent re-transmission of whole file that is pointed out as problem of transmission method of file unit. Server, contents receives, has information about contents existing received. Using this, we can extract a transmission start point of contents file (Offset) and file size of contents file. So, we can prevent duplicate transmission of contents file already transmitted. Also, we can transmit some part of received contents file into another edge server even if server don't receive entire of file as managing contents file transmitted with segment unit.

2.3   Management of contents using metadata

To efficient distribution and management of high capacity contents, we insert substance into contents file, received from contents provider, using metadata then the contents file divides logically as low capacity contents file. For this, this paper

provides content metadata management interface that can use in CDN staging server. We can insert, delete and store proper additional information about contents substance to metadata via interface. Semantically divided contents file according to substance is transmitted only particular part as segment method when user requests. If user requests contents file with particular substance, proposed method can search it using metadata information. So, user receives low capacity file with necessary substance not a whole of high capacity file. Figure 3 shows that logically divided contents file from high capacity files as classifying with 'Sub Chapter' according to substance using metadata. Also it depicts distribution process via substance based search.
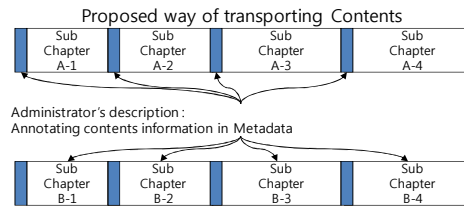

Fig. 3. Contents distribution using metadata

3   Implementation and evaluation

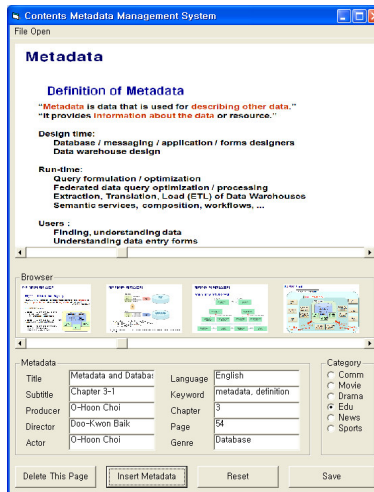## 3.1   Implementation and simulation environment

Fig. 4. Interface for management of contents metadata

For simulation, we classify simulation into distribution time via existing CDN, distribution time using CDM with 4 high capacity contents files. In simulation with existing CDN, we did substance search via keyword inserted by service provides. In

simulation with existing CDN, we did substance search via added keyword through contents metadata management interface. Figure 4 shows a screen shot is inserting proper information to metadata as well as playing contents file via contents metadata management interface that CDM (contents distribution and management) system provides.

4 high capacity contents files stored into staging server are divided into Sub Chapters according to substance, and then we inserted proper information to divided Sub Chapters. Table 1 depicts file size and some of inserted information about low capacity contents files that substances of total 14 have. Used contents files are physically 4 files but, it can be divided into low capacity files of 14 logically.

Table 1. File divided with contents and inserted keyword

File name (*keyword, file size*)	Sub file name of divided file (*keyword,* file size)
Chapter 1 *Introduction* *DB* *320MB*	Chapter 1-1, *(140MB)* *Relational DB, metadata*
	Chapter 1-2, *Multimedia Data, Information Integration (180MB)*
Chapter 2 *ER model* *690MB*	Chapter 2-1, *Entity Set (180MB)*
	Chapter 2-2, *Attribute (150MB)*
	Chapter 2-3, *Relationship (210MB)*
	Chapter 2-4, *ER Diagram (150MB)*
Chapter 3 *Introduction* *Metadata* *540MB*	Chapter 3-1, *Metadata, definition (230MB)*
	Chapter 3-2, *Metamodel (90MB)*
	Chapter 3-3, *Metadata Registry, MDR (120MB)*
	Chapter 3-4, *Interoperability (100MB)*
Chapter 4 *SQL* *560MB*	Chapter 4-1, *Projection (180MB)*
	Chapter 4-2, *Selection (140MB)*
	Chapter 4-3, *Comparison (180MB)*
	Chapter 4-4, *Schema, metadata (160MB)*

Network structure for distribution is composed 2 edge servers with 2 level hierarchies. For performance measurement, we measured distribution time took as distributing same contents file to edge server. Transmission rate are same with 100Mbps. Table 2 shows distribution time according to transmission method. We can recognize that segment unit transmission method is 48% faster than file unit transmission method. Because file unit transmission method that transmits contents file to other edge server after one content file is all transmitting.

Table 2. Kinds and results of measurement test about contents distribution time

Contents size	Distribution time of file unit (sec)	Distribution time of segment unit (sec)
320MB	6.6	3.4
690MB	14.3	7.2
540MB	11.2	5.6
570MB	11.9	5.8

For simulation, we first transmitted 4 low capacity contents files of Chapter 1-1, Chapter 1-3, Chapter 3-3, and Chapter 4-4 searched using substance based search. Selected files sizes are 650MB. In case of CDN, file name is chapter 3 that it has metadata keyword and file size is 540MB. We verified via simulation that substance based search method can firstly distribute several low capacity contents files in a fast time that user's request satisfies.

Table 3. Distribution simulation results of low capacity contents via search based on contents substance

Transmission type	Contents size	Total distribution time (sec)	Related file number
File unit	540MB	11.2	1
Segment unit	650MB	6.6	4

4   Conclusion

CDN (contents distribution and management) is network used for providing high quality contents to user in internet with non-regular bandwidth. This paper proposed CDM (contents distribution and management) system and applied file transmission method of segment unit for solving increase problem of transmission time that existing CDN has. Also, we proposed contents management method using metadata and implemented it as contents metadata management interface for enhancing efficiency of contents distribution and enlarging convenience of contents management. When proposed CDM system applied to CDN, we verified that CDM system can distribute high capacity file with same size 48% faster than existing CDN. Also when contents distribute using distribution method with substance search based on metadata, its transmission speed is 4.8 times faster than CDN. CDM system can enhance QoS of user as firstly providing divided low capacity contents file with substance that user wants.

References

1. Hofmann, Markus and Leland R. Beaumont, "Content Networking: Architecture, Protocols, and Practice.", Morgan Kaufmann Publisher, 2005
2. R. Buyya, M. Pathan and A. Vakali, "Content Delivery Networks", Springer, Germany, 2008.
3. O-hoon Choi, Jung-eun Lim, Dong-hyun Lee, Hong-seok Na and Doo-kwon Baik, "A multimedia contents management system based on a metadata-net in home network", Consumer Electronics, IEEE Transactions on Volume 54, Issue 2, May 2008, Page 468 – 473
4. O-Hoon Choi, Jung-Eun Lim, Doo-Kwon Baik, "MDR-based Framework for Sharing Metadata in Ubiquitous Computing Environment", Lecture Notes in Computer Science (LNCS), vol.3824, pp.858-866, December, 2005
5. ISO/IEC JTC 1/SC 32, "ISO/IEC 11179: Information technology- Specification and standardization of data elements, Part 1~6", 2003