An Adapted Multi-way Tree with Holder-chain for Large Scale P2P VoD System

Pingshan Liu, Guimin Huang, Ya Zhou, Huibing Zhang and Siyun Liu

Abstract. Peer-to-Peer Video-on-demand (P2P VoD) systems have attracted great interest in research community. Many hybrid methods have been proposed to provide high-quality video-on-demand services to support VCR-like operations, which can scale to large population of users. In this paper, we propose a hybrid overlay DONet-AMT, which combines DONet and an adapted multi-way tree. DONet is applied to distribute the video segments efficiently. The adapted multi-way tree consists of two layers: a complete multi-way tree and holder-chain, which can support efficient VCR-like operations. Our scheme can ensure that every video segment has k (typically $k \geq 4$) replicas at least when the number of peers is large enough, and thus guarantees high data availability. Finally, the performance analysis shows that DONet-AMT can provide better VoD services than DONet-VoD.

Keywords: Peer-to-peer • Video-on-demand • adapted multi-way tree • overlay

1 Introduction

With the prevalence of broadband Internet access, VoD services over the Internet have become increasingly popular. The traditional client-server architecture can hardly scale to a large population of users due to the limited service capacity of the servers. Recent years, P2P networks have been shown to be a promising approach to provide high scalability VoD streaming service[1-8].

In the P2P VoD systems, supporting VCR-like operations such as jump forward/backward is a great challenge. The VCR-like operations result in frequent seeking, and thus can impose great impacts on the viewing quality. We can define seeking latency as the interval between the time when a segment is requested and the time when the segment is ready for playback. To achieve a high viewing quality, the seeking latency must be as short as possible. To reduce the seeking latency,

---

Pingshan Liu (✉), Guimin Huang, Ya Zhou, Huibing Zhang, Siyun Liu
Research Center On Data Science and Social Computing,
Guilin University of Electronic Technology, No 1 Jinji Road, Guilin, China
email:liup3@guet.edu.cn

© 2013. The authors - Published by Atlantis Press

each segment has to be fetched prior to its playback time. Thus, the prefetching scheme is required to support the VCR-like operations. For ease of presentation, in the rest of the paper, we use the terms “client”, “peer”, and ”node” interchangeably.

Recently, some studies on P2P VoD systems[8-16] have emphasized the support to VCR-like operations and can handle the random seeks. Most of these studies have applied a hybrid overlay. In this paper, we focus on one kind of hybrid overlays which combine two overlays: a mesh-based overlay and a tree-based overlay. These two overlays have their merits[17]. The tree-based overlays have a simple structure, and support efficient query. The mesh-based overlays yield a higher goodput to the streaming application than the tree-based overlays. So, A hybrid overlay combining a mesh-based overlay and a tree-based overlay to take advantage of their merits is a good solution to support VCR-like operations in large scale P2P VoD systems.

The hybrid overlays of this kind have been explored by some research work[9, 10, 15]. In these hybrid systems, a video is divided into segments. One prefetching unit of DONet-VoD[9] consists of several consecutive segments. The superchunk in SURFNet[15] has the same meaning as the prefetching unit in DONet-VoD. DONet-VoD is a hybrid overlay, which combines a mesh-based overlay DONet[7] and a multi-way tree-based overlay based on BATON*[18]. SDNet[10] is an extension of DONet-VoD, which is more detailed about the underlying mechanism. SURFNet[15] is also a hybrid method, which integrates two networks: a structured network for search and an unstructured network for information exchange. The structured overlay consists of two layers: AVL tree and holder-chain.

Although DONet-VoD and SURFNet can combine a mesh-based overlay and a tree-based overlay to support efficient VCR-like operation, both of them cannot guarantee that every prefetching unit has k($k>1$, typically $k \geq 4$) replicas at least. In some cases，some prefetching units may have only one replica, which results in low data availability. If multiple peers request these segments, some peers may not get the desired segments immediately, which results in high seeking latency. To avoid the problem, every prefetching unit should have k replicas at least when the number of peers is large enough, which can guarantee high data availability. Inspired by SURFNet, we proposed a hybrid overlay DONet-AMT, which combines the DONet[7] mesh overlay and an Adapted Multi-way Tree overlay. The DONet mesh overlay is used as the routine overlay, which is responsible for scalable live video streaming and video distribution. The adapted multi-way tree overlay consists of two layers: a complete multi-way tree and holder-chain, which is responsible for supporting efficient VCR-like operations. Moreover, The adapted multi-way tree can make sure that every prefetching unit has k replicas at least. Consequently, high data availability of the system is guaranteed. Refer to the above-mentioned cases, our contributions about P2P VoD as follows: (1) we propose a new hybrid overlay DONet-AMT, which combines the DONet mesh overlay and an Adapted Multi-way Tree overlay. (2) our scheme can guarantee that every prefetching unit of a video has k replicas at least when the number of peers is large enough. Thus, high data availability is guaranteed. Due to high data availability, DONet-AMT can provide better VoD service.

The rest of this paper is organized as follows. In Section 2, we present the design of DONet-AMT model. In Section 3, we evaluate the performance of DONet-AMT. Finally, we conclude the paper in Section 4.

2 DONet-AMT Model

In this Section, we first introduce the architecture of DONet-AMT. Then we describe how to construct the adapted multi-way tree overlay.

2.1 DONet-AMT Architecture

DONet-AMT is a hybrid architecture (Fig. 1), which integrates two networks: an adapted multi-way tree and DONet. The adapted multi-way tree overlay consists of two layers: a complete multi-way tree and holder-chain. The complete multi-way tree is based on BATON*[18] which is balanced, and has a limited size according to the length of a video. The cost of search in BATON* is bounded by $O(\log_m N)$. Holder-chain is a link-list which groups all peers holding the same prefetching unit together and attaches to a node in the multi-way tree which has the same prefetching unit. DONet uses a random gossip algorithm and constructs a mesh overlay across the two layers of the adapted multi-way tree.

In the adapted multi-way tree, the node in the complete multi-way tree not on-

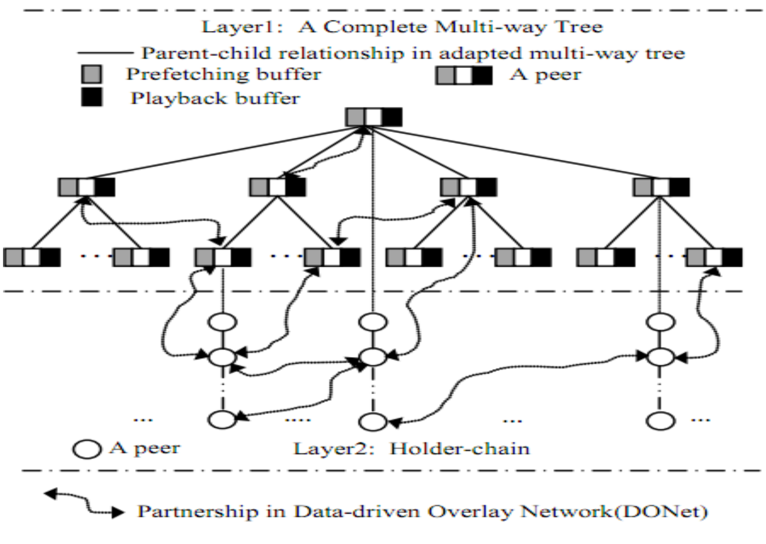

Fig.1 The Architecture of DONet-AMT

ly acts as a node in the multi-way tree, but also acts as the head of the corresponding holder-chain. As far as the head of holder-chain is concerned, we adopt the same mechanism as that of SURFNet. The nodes in the multi-way tree only keep the IP addresses of a subset of peers in the holder-chain, namely sample peers, and the sampling mechanism is also the same as that of SURFNet.

Every peer in the DONet-AMT has both a prefetching buffer and a playback buffer. Videos are divided into uniform segments. The prefetching buffer of every peer has the same size and can buffer one prefetching unit which consists of d(d>1) consecutive segments. The playback buffer is used to cache the video segments for current playing. When a new client enters the system, it may begin to view a video stream on demand. At the same time, it can download in advance some segments using its residual bandwidth and stores them in its prefetching buffer. The prefetched segments are not selected randomly, but are chosen according to the prefetching scheme, which is included in the construction of an adapted multi-way tree and is detailed in section B. After the video segments have been downloaded completely, they will be maintained during the node's lifetime.

In DONet-AMT, DONet is used for efficient video distribution. When the video is being viewed sequentially, each peer gets the multimedia streaming from its gossip partner. When the VCR-like operations occur, the adapted multi-way tree assists the peer to jump to the new playing scene quickly. The adjacent nodes and the attached holder-chain would provide the peer with the multimedia streaming segments from their prefetching buffers. At the same time, the peer seeks for the new partners. After the new partners have been founded, the streaming suppliers are switched to the new partners again.

2.2 Construction of the adapted multi-way tree overlay

Table 1 Notations

Notations	Descriptions
S	The VoD media server
T	The multi-way tree
P	The requested video for playing
Len	Length of P
D	Size of one prefetching unit
R	The root node of T
X	A node in T
HLen	The length of a holder-chain
H	The height of T
m	Fanout of T
N	The number of nodes in the tree T
Level(X)	The level of X in T
Prebuf(X)	Node X's prefetching buffer
Seg(X)	The sequence number of prefetching unit in Prebuf(X)
Parent(X)	Node's parent node in T

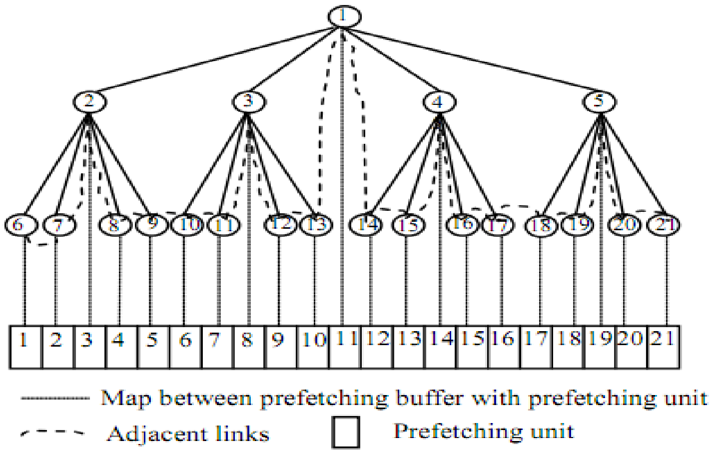

Fig. 2 The First Phase

To be clear, we adopt the notations including those in DONet-VoD[7] and SURFNet[12], which are presented in table 1. The video P has $L = \lceil Len/d \rceil$ prefetching units which are numbered from 1 to L sequentially. The prefetching buffer of node *X* is defined as Prebuf(*X*) with size of *d* pieces of successive video segments. The fanout of the tree T is set as even, denoted as a variable *m*(typically, $m \geq 4$). The multi-way tree T satisfies these formulas and equations:

$$Level(R) = 0; \quad (1)$$

$$Level(X) \leq H, \forall X \in T; \quad (2)$$

$$N = (m^{H+1} - 1)/(m-1); \quad (3)$$

$$H = \lfloor \log_m (m-1)L \rfloor. \quad (4)$$

The construction of an adapted multi-way tree can be divided into three phases. The first phase is to construct a complete multi-way tree T. The second phase is to construct hold-chains and make sure that the length of every holder-chain is no smaller than k. The third phase is to prefetch the prefetching unit randomly, and join the corresponding holder-chain according to the prefetching unit.

In the first phase, a complete multi-way tree is constructed as illustrated in Fig. 2. The height H of the multi-way tree is determined by the length L of the video P according to (4). When T grows fully to the height H, T will form a complete multi-way tree and stop growing. Assuming the node X is the ith child of Parent(X). The map relation between the prefetching buffer with the prefetching unit can be determined by the rules as follows:

If $Level(X) = 0$ **, then**

$$Seg(X) = \lceil L/2 \rceil \quad (5)$$

If $0 < Level(X) \leq H$ ***, there are two cases:***

If $1 \le Level(X) \le \lceil m/2 \rceil$ **, then**

$$Seg(X) = Seg(Parent(X)) - \left\lceil \frac{(m-2i+1)L}{2m^{Level(X)}} \right\rceil \tag{6}$$

If $\lceil m/2 \rceil < i \le m$ **, then**

$$Seg(X) = Seg(Parent(X)) + \left\lceil \frac{(2i-m-1)L}{2m^{Level(X)}} \right\rceil \tag{7}$$

Seg(X) can be computed according to (5),(6),(7). In the case of *Seg(X)*=0, we adjust *Seg(X)*=1. In the case of *Seg(X)*>L, *Seg(X)*=L.

Node *X* can prefetch the corresponding prefetching unit according to the above three rules. When the multi-way tree grows to a complete multi-way tree with a height H, the tree stops growing, and the first phase finishes. Then, the second phase begins to construct holder-chains.

The second phase consists of the following steps:

1) When a new peer *p* is coming, it first queries the root node of the complete multi-way tree, and retrieves the length *HLen* of the holder-chain attaching to the root.

2) If *HLen*< k and there is no a holder-chain attaching to the root, then the peer *p* downloads the prefetching unit from the root with its residual bandwidth. The root becomes the head of the holder-chain, and links the peer *p*. The length *HLen* of the holder-chain increases to be 2.

3) If HLen< k, and there is a holder-chain attaching to the root, then the peer p inserts after the root, and download the prefetching unit from one node of the holder-chain. The length HLen of holder-chain increases one.

4) If HLen≥ k, then the peer p queries the first child of the root, and retrieves the length HLen of the holder-chain attaching to the first child. If the retrieved length HLen< k, then the steps 2),3) executed for the root is applied to the first child. If the retrieved length HLen≥ k, the peer p continues to query the second child of the root, and the process executed for the first child is applied to the second child. The process is iterative until all the nodes of the tree is visited in the order of breadth-first traversal. If all the nodes of the tree have been visited and cannot find a node which satisfy HLen<k, the second phase completes.

In the third phase, the new coming peer will prefetch a prefetching unit randomly. After the sequential number of a prefetching unit is designated by random function, the new coming peer searches the complete multi-way tree and find a node holding the corresponding prefetching unit in the tree. Then the new coming peer inserts after the found node and downloads the prefetching unit from one node of the hold- chain. The length *HLen* of this holder-chain will increase one.

3 Evaluation

In this section, We use two metrics to analyze the performance of DONet-AMT. The two metrics are search latency and jump failure.

3.1 Search Latency

Search Latency is defined as the interval between the time when a seek or jump is performed and the time when a node in the complete multi-way tree having the requested segments is found. Hence, the search latency can be in terms of hop count for locating the node. In DONet-AMT, the search latency is bounded by O(H)+1. In DONet-VoD, the search latency is bounded by O($\log_m (m-1)L$)). In SURFNet, the search latency in the AVL tree is bounded by O(logL). In VMesh, the search latency is bounded by O(logn)(n is the total peers in VMesh). In RINDY, the search latency is O($\log\lceil Len/w \rceil$). When we compute the search latency, we make the computing result be an integer which is no less than the computing result. Assuming the length of a video is Len=7200, including L=720 number of prefetching unit, w=10. A comparison result is illustrated in Fig. 3. From the left side, the first bar denotes the search efficiency of VMesh, the second denotes that of RINDY, and the third denotes that of SURFNet, etc. We can see the search efficiency of DONet-AMT (m=10) is highest.

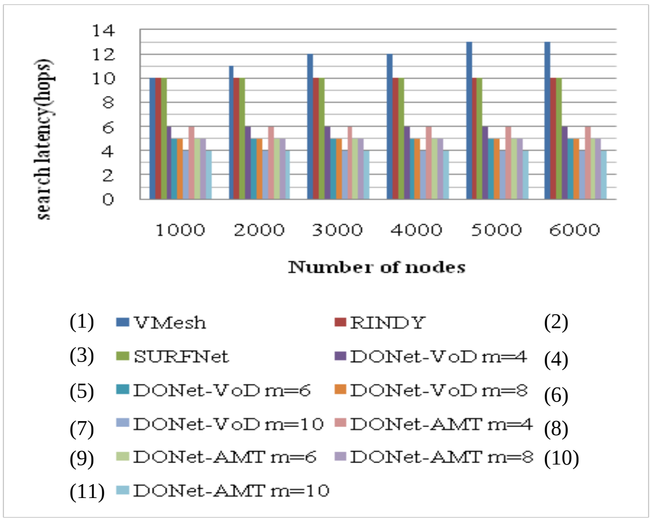

Fig. 3 Search Latency

3.2 Jump Failure

DONet-AMT and DONet-VoD both can provide high success rate for jump because of the unchanged prefetching buffer. However, DONet-AMT can avoid one type of jump failure that can occur in DONet-VoD.

In DONet-VoD, some prefetching units may have only one replica. When one of these prefetching units was requested by multiple peers at the same time, only one peer *p* containing the replica provides service. The peer *p* cannot serve multiple peers at the same time. In this case, the peers that cannot get service in time encounter a jump failure. Hence, these peers cannot download video segments from the prefeching buffer and have to wait for a long time to let DONet find new partners for video segments.

In DONet-AMT, the jump failure described above can be avoided, for every prefeching unit has *k* replicas at least when the number of peers is large enough.

4 Conclusion

This paper has proposed a new hybrid overlay DONet-AMT for large scale P2P VoD systems, which integrates two networks: DONet and an adapted multi-way tree. DONet is used for video distribution. The adapted multi-way tree consists of two layers: a complete multi-way tree and holder-chain. When the number of peers is large enough, the adapted multi-way tree can guarantee every prefetching unit has *k* replicas at least, which ensure high data availability. Compared with DONet-VoD, the seek latency and jump failure of DONet-AMT can be reduced because of high data availability.

Acknowledgments. This work was supported by the Foundation of Guangxi Science Research and Technology Development (No. GuiKeGong1348014-8), and the Foundation of Key Laboratory of Guangxi Wireless Broadband Communication and Signal Processing (No. 11108).

REFERENCES

[1] Guo, Y., Suh, K., Kurose, J., & Towsley, D. (2003). P2Cast: peer-to-peer patching scheme for VoD service. in Proceedings of the 12th international conference on World Wide Web Budapest, Hungary: ACM. doi:10.1145/775152.775195.

[2] Do, T. T., Hua, K. A., & Tantaoui, M. A. (2004). P2VoD: providing fault tolerant video-on-demand streaming in peer-to-peer environment. in Communications, 2004 IEEE International Conference on, pp. 1467-1472 Vol.3.doi: 10.1109/ICC.2004.1312755.

[3] Guo, Y., Mathur, S., Ramaswamy, K., Yu, S. C., & Patel, B. (2007). PONDER: Performance Aware P2P Video-on-Demand Service. in Global Telecommunications Conference, IEEE, pp. 225-230. doi: 10.1109/GLOCOM.2007.50.

[4] Huang, C., Li, J., and Ross, K. W. (2007). Can Internet video-on-demand be profitable?. Computer Communication Review, vol. 37, pp. 133-144. doi: 10.1145/1282427.1282396.

[5] Huang, Y., Fu, T. Z. J., Chiu, D. M., Lui, J. C. S., & Huang, C. (2008). Challenges, design and analysis of a large-scale P2P-VoD system. Computer Communication Review, vol. 38, pp. 375-388. doi: 10.1145/1402946.1403001.

[6] Cheng, B., Liu, X. Z., Zhang, Z., Jin, H., Stein, L., & Liao X. F. (2008). Evaluation and optimization of a peer-to-peer video-on-demand system. Journal of Systems Architecture, vol. 54, pp. 651-663. doi: 10.1016/j.sysarc.2007.11.003.

[7] Zhang, X. Y., Liu, J. C., Li, B., & Yum, Y. S. P. (2005). CoolStreaming/DONet: a data-driven overlay network for peer-to-peer live media streaming. in INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE, 2005, pp. 2102-2111 vol. 3. doi: 10.1109/INFCOM.2005.1498486.

[8] Vratonjić, N., Gupta, P., Knežević, N., Kostić, D., & Rowstron, A. (2007). Enabling DVD-like features in P2P video-on-demand systems. in Proceedings of the 2007 workshop on Peer-to-peer streaming and IP-TV Kyoto, Japan: ACM. doi: 10.1145/1326320.1326326.

[9] Xu, C. Q., Muntean, G. M., Fallon, E., & Li, X. G. (2008). DONet-VoD: A hybrid overlay solution for efficient peer-to-peer video on demand services. in Multimedia and Expo, 2008 IEEE International Conference on, pp. 641-644. doi: 10.1109/ICME.2008.4607516.

[10] Xu, C. Q., Muntean, G. M., Fallon, E., & Hanley, A. (2009). Distributed Storage-Assisted Data-Driven Overlay Network for P2P VoD Services. IEEE Transactions on Broadcasting, vol. 55, pp. 1-10. doi: 10.1109/TBC.2008.2006252.

[11] Xu, T. Y., Chen, J. Z., Li, W. Z., Lu, S. L., Guo, Y. , & Hamdi, M. (2009). Supporting VCR-Like Operations in Derivative Tree-Based P2P Streaming Systems. in IEEE International Conference on Communications, 2009. pp. 1-5. doi: 10.1109/ICC.2009.5199502.

[12] Guo, Y., Yu, S. C., Liu, H., Mathur, S., & Ramaswamy, K. (2008). Supporting VCR Operation in a Mesh-Based P2P VoD System. in 5th IEEE Consumer Communications and Networking Conference. pp. 452-457. doi: 10.1109/ccnc08.2007.107.

[13] Cheng, B., Jin, H., & Liao, X. F. (2007). Supporting VCR Functions in P2P VoD Services Using Ring-Assisted Overlays. in IEEE International Conference on Communications, 2007. pp. 1698-1703. doi: 10.1109/ICC.2007.284.

[14] He, Y., & Liu, Y. H. (2009). VOVO: VCR-Oriented Video-on-Demand in Large-Scale Peer-to-Peer Networks. IEEE Transactions on Parallel and Distributed Systems, vol. 20, pp. 528-539. doi: 10.1109/TPDS.2008.102.

[15] Wang, D. Q., & Yeo, C. K. (2009). Superchunk Based Fast Search in P2P-VoD System. in I EEE Global Telecommunications Conference, 2009. pp. 1-6. doi: 10.1109/GLOCOM.2009.5425514.

[16] Yiu, W. P. K., Jin, X., & Chan, S. H. G. (2007). VMesh: Distributed Segment Storage for Peer-to-Peer Interactive Video Streaming. Selected Areas in Communications, IEEE Journal on, vol. 25, pp. 1717-1731. doi: 10.1109/JSAC.2007.071210.

[17] Seibert, J. , Zage, D., Fahmy, S., & Nita-Rotaru, C. (2008). Experimental comparison of peer-to-peer streaming overlays: An application perspective. in 33rd IEEE Conference on Local Computer Networks, 2008. pp. 20-27. doi: 10.1109/LCN.2008.4664147.

[18] Jagadish, H. V., Ooi, B. C., Tan, K. L., Vu, Q. H., & Zhang, R. (2006). Speeding up search in peer-to-peer networks with a multi-way tree structure. in Proceedings of the 2006 ACM SIGMOD international conference on Management of data, Chicago, IL, USA: ACM. doi: 10.1145/1142473.1142475.