

# Design and Evaluation of a P2P IPTV System for Heterogeneous Networks

Meng-Ting Lu, Jui-Chieh Wu, Kuan-Jen Peng, Polly Huang, *Member, IEEE*, Jason J. Yao, and Homer H. Chen, *Fellow, IEEE*

**Abstract**—NTUStreaming is an overlay P2P-based IPTV system that integrates innovations in both overlay networking and video coding for optimal user experience. The system consists of three key components: partnership formation, robust video coding, and video segment request scheduling. For partnership formation, a graph construction mechanism TYPHOON based on epidemic algorithms is developed to reduce disconnect time and isolated peers. For robust video coding, a multiple description coding (MDC) scheme with spatial-temporal hybrid interpolation (STHI) is proposed to adjust streaming traffic according to the bandwidth and device capability of each peer. For request scheduling, an optimization algorithm is developed by taking the available bandwidth and the video segment type into account. Experimental results show that NTUStreaming is able to deliver optimal video quality in lossy and dynamic networking environments.

**Index Terms**—Error recovery, IPTV, multiple description coding, overlay network, peer-to-peer (P2P), video broadcast, video streaming.

## I. INTRODUCTION

WITH the rapid advances in multimedia entertainment, broadband networks, and semiconductor manufacturing technologies, high-speed multimedia communication is coming of age. Video broadcasting over the Internet, i.e., IPTV, is one of the most promising multimedia entertainment applications on the rise. The key to a successful IPTV system lies in the quality of services (QoS). For content providers with enough financial power and high sensitivity to the copyright issue, a conventional

distributed network consisting of servers and caches may be a preferred solution for content delivery. However, the recent success of P2P IPTV services, such as PPStream [1], PPLive [29] and CoolStreaming [12], is an evidence of the fact that the digital content world is evolving. Web 2.0 allows the general public to create and share contents of their own. In this new content world, an economic P2P approach to content distribution is attractive. According to a recent measurement study, Huang *et al.* [34] shows that over 80% of the users on the Internet have download capacity more than 1500 kb/s and upload capacity more than 384 kb/s, which is quite enough for P2P streaming. Our objective is to develop a scalable and robust P2P IPTV system for quality video broadcasting over the Internet.

NTUStreaming, the proposed IPTV system, takes the following design features into considerations: 1) an overlay peer-to-peer networking approach, with which the system can exploit the resources, such as storage and bandwidth, available among the peer users for better scalability; 2) an MDC approach, with which heterogeneous peers can flexibly subscribe to streams of different resolutions and enjoy the best quality their available resources permit; and 3) a bandwidth-aware scheduling algorithm, where the schedule of the video segment transmissions is based on the resolution of video segments and network bandwidth available. The architecture consists of three major technical components: partnership formation, multiple description coding, and video segment request scheduling. The partnership formation component is responsible for finding the peers to serve video segments. The multiple description coding determines the encoding and playback strategy for better video quality. The video segment request scheduling algorithm manages how and where to request the video segments. NTUStreaming is one of the pioneering IPTV systems that integrate innovations in both overlay networking and video coding for optimal user experience.

The partnership formation component is essential for the overlay P2P-based video broadcasting. In the P2P network paradigm, a video stream is divided into segments, which are delivered from peers to peers. More specifically, a user chooses a number of peers to form a partner group, and it receives video segments from these partners. A user may also be chosen by other peers in the network to be their partners. How the partner and consumer relationship is established impacts the efficiency and stability of the network. In fact, load balancing and stability of the partnership in the presence of peer dynamics are the main design considerations in the partnership formation of NTUStreaming. Experimental results show that, relative

Manuscript received November 15, 2006; revised July 27, 2007. This work was supported in part by grants from the Intel Corporation, Institute for Information Industry, and the National Science Council of Taiwan under Contracts NSC 94-2219-E-002-016, NSC 94-2219-E-002-012, NSC 94-2220-E-002-027, and NSC 94-2725-E-002-006-PAE. The associate editor coordinating the review of this manuscript an approving it for publication was Dr. Jin Li.

M.-T. Lu and K.-J. Peng are with the Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan, R.O.C.

J.-C. Wu is with the Graduate Institute of Computer Science and Information Engineering, National Taiwan University, Taipei 10617, Taiwan, R.O.C.

P. Huang is with the Department of Electrical Engineering, Graduate Institute of Communication Engineering, and Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei 10617, Taiwan, R.O.C. (e-mail: phuang@cc.ee.ntu.edu.tw).

J. J. Yao is with the Department of Electrical Engineering and Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan, R.O.C.

H. H. Chen is with the Department of Electrical Engineering, Graduate Institute of Communication Engineering, and Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei 10617, Taiwan, R.O.C. (e-mail: homer@cc.ee.ntu.edu.tw).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMM.2007.907456

to the state of the art [2], the partnership formation scheme of NTUStreaming is able to enhance the stability of the P2P partnership without compromising the system load balance.

MDC is the component that impacts the user perception of the video. Built on top of the multiple state video coding [3]–[5], our proposed MDC scheme adds two streams of lower-resolution pictures to enhance scalability and error resilience. When a description is missing, we introduce a hybrid interpolation scheme, called Spatial-Temporal Hybrid Interpolation (STHI) [6], which exploits both spatial and temporal information to improve the video quality.

The networking and video coding components are further integrated through a video segment request scheduling mechanism. The video segment request scheduling in the state-of-the-art CoolStreaming is heuristic based. A policy that requests the rarest segment first could be inadequate for heterogeneous networks. We formulate the segment scheduling component as an optimization problem, which allows each node in a heterogeneous network to request segments of particular spatial or temporal resolutions based on the available bandwidth. As a result, the system is able to better utilize available resources and achieve better video quality.

The rest of this paper is organized as follows. Section II describes related work, and Section III presents the partnership formation of the overlay network. Section IV explains MDC-STHI and the integration of the whole system. Section V discusses the comparison between MDC-STHI and layered video coding. The bandwidth-aware data scheduling mechanism is given in Section VI. Section VII presents experimental results and discussions, and Section VIII concludes this paper.

## II. RELATED WORK

The distribution of video streams over IP is a quasi-reliable multicast problem in nature. Solutions at the network layer have long existed. Multicast routing protocols such as DVMRP [7] and PIM-SM [8] are widely implemented in commercial routers. Reliable multicast solutions at the network layer such as SRM [9] have also been proposed since 1997. The common problem of the network-layer solutions is the difficulty in deployment because every router in between the senders and receivers needs to run the required mechanisms for the multicast video broadcasting to work. To achieve scalability through adjustment in the application layer, several designs have been proposed recently, including Application-level Multicast [10], ZigZag [11], CoolStreaming [12], SplitStream [13] and AnySee [14]. All of them exploit the deployment flexibility at the application level. Among the existing systems, CoolStreaming is closest to ours but different in the design of partnership formation, video coding, and video segment scheduling [15].

MDC has received only little attention in P2P streaming. Padmanabhan *et al.* [16] introduced data redundancy to CoopNet using MDC. Zink *et al.* [17] discussed the feasibility of MDC for P2P networks and compared it to hierarchically layered video coding. Khan *et al.* [18] compared the performance of MDC over P2P networks with that of content delivery networks (CDNs). Conventional MDC schemes are seldom used in practice because they are too complex and usually use nonstandard

TABLE I  
NOMENCLATURE

Partner Peer	Peer of a node that serves data to the node
Consuming Peer	Peer of a node that receives data from the node
Contact Node	The node that is first connected to by a newly joined peer
PartialView	The list of partner peers for a node
InView	The list of consuming peers for a node

video codecs. On the other hand, our scheme is built on top of multiple state video coding, which is easy to implement using standard codecs as basic components.

Layered video coding (LVC), such as scalable video coding (SVC) [23] and MPEG-4 fine grained scalability (FGS) profile [32], which encodes videos into streams with temporal, spatial, and SNR scalability, is adopted in some P2P streaming systems [19], [20], [30]. The scalability of LVC is especially suitable for heterogeneous networks. However, the major drawback of the layered coding approach is its high complexity and lack of error resilience. On the contrary, the simple MDC-STHI scheme provides both scalability and error resilience.

## III. PARTNERSHIP FORMATION

The partnership formation component determines for each user a list of partner peers that can serve video segments to the user. Because the peers may leave or join the overlay network at any time, it is necessary to select substitute partners for reliability. A balanced and robust partnership formation policy achieves both load balance and stability. In this section, we survey a state-of-the-art partnership formation mechanism SCAMP and then propose TYPHOON to improve the system stability while maintaining load balance. The terms used in this section are defined in Table I.

### A. SCAMP

SCAMP (Scalable Membership protocol) [2], which is used in CoolStreaming, is a peer-to-peer partnership formation protocol operating in an unstructured and fully decentralized manner. Each node under this protocol searches for potential neighbors to connect with by a gossip-based messaging mechanism [21], [22]. Each node maintains a list of partner peers, PartialView, and a list of consuming peers, InView. To avoid staleness, every entry in the PartialView has a finite lifetime, and each node sends subscription messages periodically to reset the expiration time. These messages are also used to update the partner list as nodes join and leave the system dynamically.

The process of a new node joining the system is illustrated in Fig. 1. When a peer node 7 joins the system for the first time, it adds the contact node 1 in its PartialView. The contact node then advertises the identifier of this newly joined node to all its partner peers and to a predetermined number of additional peers randomly selected from its PartialView. When receiving a subscription message, the peer adds the newly joined node to its PartialView with a probability of  $1/(1+M)$ , where  $M$  is the PartialView size. If the newly joined node is not added to the PartialView, the subscription message is randomly forwarded to another node in the PartialView.

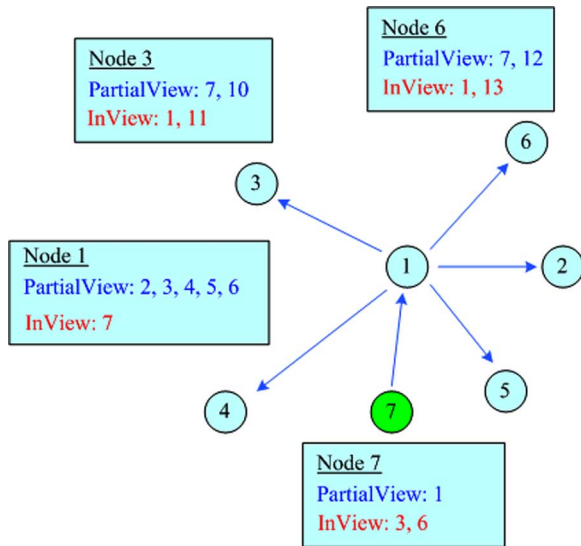


Fig. 1. Illustration of partnership formation.

In SCAMP, subscription forwarding and partner selection processes are purely random. Although this achieves good load balancing, our experiments show that the network becomes unstable in dynamic environments, resulting in disconnection and video interruption.

### B. TYPHOON

The robustness of traditional gossip-based protocols like SCAMP heavily relies on the success of message forwarding. When a peer leaves the overlay network, it should inform all peers in the InView and PartialView of its departure. If a peer leaves the SCAMP network without informing other peers, all peers in its PartialView can notice the partnership change after a heart-bit cycle. However, the peers in its InView can not get any signal of its departure. This kind of asymmetry results in network instability and frequent message losses. To solve this problem, we propose TYPHOON—an improved SCAMP protocol—based on the concept of preferential message forwarding and epidemic algorithms [21].

TYPHOON consists of two major parts: preferential random forwarding (PRF) and preferential random selection (PRS). The first part, PRF, allows preferential random forwarding of the subscription messages. When a node re-subscribes itself, partners with larger partnership are favored and have a higher probability to be chosen. For the balance of the partnership, subsequent forwarding of the subscription messages is sent to partners with smaller partnership. The second part, PRS, allows preferential selection of partners. In SCAMP and PRF, a node is selected as a partner with the probability  $1/(1+M)$ . In PRS, a peer with small partnership size selects peers with higher probability until reaching the lower bound of the partnership size. Similarly, if the partnership size of a peer is larger than an upper bound, the peer simply replaces one of its partners with the newly selected partner and forwards the subscription messages of the replaced partner to others.

TYPHOON is designed as a general architecture that works equally well regardless of the size of membership. Compared to SCAMP, TYPHOON is more resilient to heavy churns that

may break down or disconnect loosely connected small overlay networks. Unlike SCAMP, TYPHOON maintains a moderate membership size for each peer. This helps keep the traffic between peers manageable. As the number of partners of each peer grows, the message processing traffic becomes heavier and can potentially go beyond the limit for, for example, handheld devices in a heterogeneous network that have very limited resource. For applications considered in our work, this is certainly undesirable.

Compared to SCAMP, TYPHOON is more resilient to heavy churns that may break down or disconnect loosely connected small overlay networks. TYPHOON and SCAMP both periodically search for new partners to add and time out the old partners. To make the search scalable, the subscription messages they send and the partner list they record are both local. The disconnection in the overlay network under peer dynamics is the problem that rises from such local discovery of new partners. A peer may break away to add certain partners randomly or with preference to available bandwidth and content. Not knowing whether the added partners are well connected to the rest of the network, a small island can be formed and becomes disconnected from the major crowd. The subsequent searches for new partners are limited to peers on the island because none of the peers have any information of the major crowd from whom they get disconnected by chance. In simpler words, the disconnection of overlay network does not mean ‘peers physically breaks from the network’. Rather, it means the search network for potential partners clustered so much that islands are formed.

Peers on the disconnected island may need to spend several swarming cycles to figure out the fact that they are isolated and some additional time to re-contact the bootstrap servers to get back on the network. In delay-sensitive video streaming systems, such outage of video streams may cause a significant amount of disturbance to multiple users. The best way to solve the isolation problem is to prevent it from happening. This is why TYPHOON, considering the connectivity of the candidate partners, can effectively mitigate the peer isolation problem and, in return, provides better overlay network stability.

### IV. MULTIPLE DESCRIPTION CODING

MDC is a critical component directly affecting user satisfaction. In a P2P network, each peer downloads data from and uploads data to multiple peers, but the peers many join or leave the network at any time. MDC is especially useful for such a dynamic network with path diversity because it allows a receiver to obtain the baseline quality as long as one supplier survives. In our system, we decompose a video stream in the temporal dimension for diversity gains, and encode the odd frames based on the previous odd frames and the even frames based on the previous even frames. The two sub-streams are transmitted separately and played back with synchronization at the receiver. Such a design is more fault tolerant than conventional video coding because when one sub-stream is lost due to network congestion, the decoder can still independently display the remaining sub-stream with controlled degradation. This method would lower the coding efficiency a little because the temporal relationship is not fully exploited. As shown in Fig. 2, multiple state video coding, using MPEG4 for both descriptions, loses

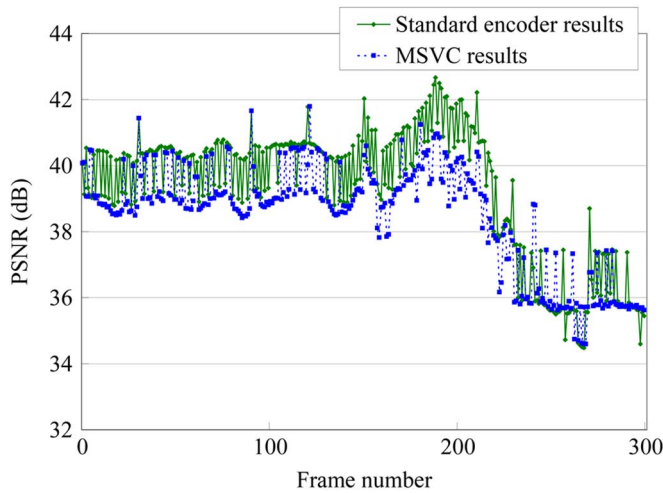


Fig. 2. Comparison between standard encoder results and MSVC results.

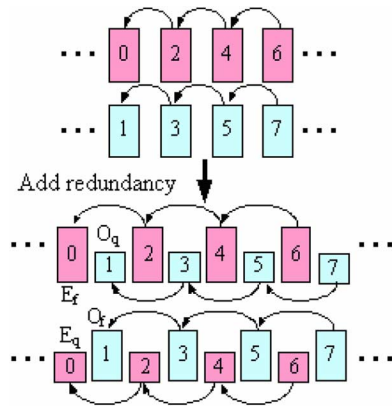


Fig. 3. Video streams in MDC-STHI.

about 1 dB in PSNR in comparison to normal MPEG4 video at the same total bit rate.

Based on the concept described above, we propose a new coding scheme named MDC-STHI, which generates four sub-streams as shown in Fig. 3. Sub-streams  $E_f$  and  $O_f$  represent the even and odd streams, respectively, encoded from the original full-size video.  $E_q$  and  $O_q$  denote the down-sampled version encoded at a lower bit rate. In our implementation, a 2 : 1 down sampling is applied in each dimension to the video frame to generate the quarter-size video. While  $E_q$  and  $O_q$  require extra bits to encode, we show in Section VII that the combination of the four sub-streams has many desirable features for video streaming over P2P networks.

Fig. 4 illustrates a possible P2P networking scenario using MDC-STHI. Generally, Internet users with various computational capabilities are connected through networks of different bandwidths, and they are randomly grouped as peers for video streaming in a P2P environment. In Fig. 4, the number of circles indicates roughly the amount of resources at the node. Thus,  $N_3$  has enough bandwidth and computing power to retrieve both  $E_f + O_q$  and  $O_f + E_q$  streams from  $N_1$  and  $N_2$ . In turn,  $N_4$ ,  $N_5$ , and  $N_6$  can receive appropriate combinations of streams from  $N_3$  according to their available resources. For example,  $N_4$  may

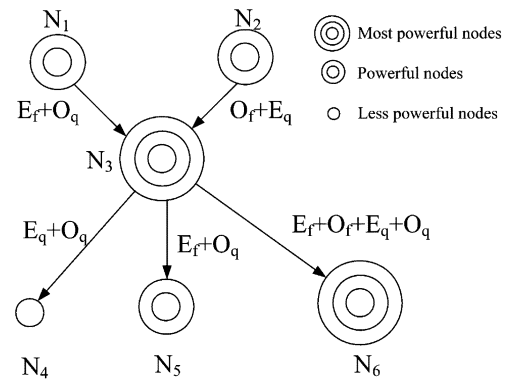


Fig. 4. MDC-STHI over a P2P network.

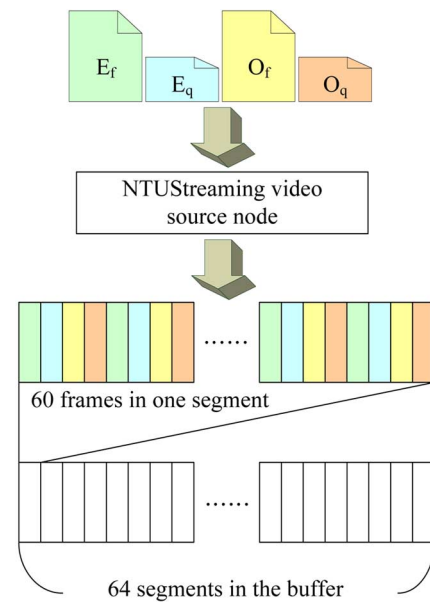


Fig. 5. NTUStreaming video source node.

retrieve only  $E_q$  and  $O_q$  from  $N_3$  because it uses a mobile device with a narrowband wireless connection and a small-size screen that best displays quarter-size videos. On the other hand,  $N_6$  can accommodate all streams from  $N_3$  and may further distribute the video streams intelligently to other peers. Compared to other MDC designs, MDC-STHI offers superior flexibility for P2P video streaming as the four sub-streams are independently encoded and a node may adaptively deliver suitable combinations according to the peers' capabilities.

The combination of  $E_f + O_q$  or  $O_f + E_q$  poses an interesting challenge as it contains two streams of different resolutions. During transmission, the bits from  $O_q$  may piggyback in the same IP packet with  $E_f$  of the previous frame, and *vice versa* for  $O_q$ . This saves the overhead and ensures same time arrival of adjacent frames. An innovative hybrid interpolation scheme is also developed for such combinations, fully explained in [6].

Fig. 5 shows the integration of MDC-STHI with the NTUStreaming video source node. The four independently-encoded MDC-STHI sub-streams are read into the video source node and segmented. Each segment contains 60 frames, 15 frames for

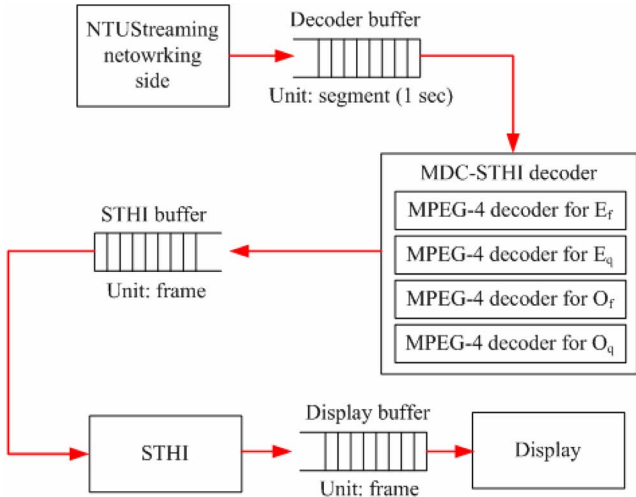


Fig. 6. Video data flow of an NTUStreaming node.

each sub-stream, and there are 64 segments in the buffer, equivalent to 64 s of video data. Fig. 6 shows the video data flow of an NTUStreaming client node. When receiving a segment, the MDC-STHI decoder processes the data using the corresponding MPEG-4 decoder based on the type of the segment. The decoded frames are then put into the STHI buffer. If the node only receives part of the description due to packet loss or bandwidth constraint, the STHI module performs spatiotemporal hybrid interpolation on the lost frames. Finally, the interpolated frames are sent to the display buffer.

V. COMPARISON BETWEEN MDC-STHI AND LVC

For MDC-STHI, because all descriptions are coded independently, a receiving node is able to put together the full resolution video if it receives collectively from its source nodes all descriptions of the full resolution video. In other words, different descriptions of the full video do not have to come from one single source node. If a receiving node does not receive all the descriptions of a full resolution video, it composes a scaled-down (either spatially or temporally) version of the full video. No node is required to receive all descriptions. Note that the system works because a receiving node has the chance to reconstruct a full resolution video even if none of its source peers has the complete full resolution video. This is in fact a novelty of MDC-STHI.

In contrast to MDC-STHI where each description is independently coded and hence missing any description does not cause catastrophic effect, the enhancement layers of LVC are dependent on the base layer. Therefore, the performance of LVC is hinged on successful transmission of the base layer. This limits its flexibility in a dynamic P2P network where peers join and leave at any time.

Another strong distinction between these two approaches is that thin stream is more likely to occur for layered streaming in a dynamic P2P streaming network. Thin stream refers to the case where a low resolution or partial video is received by a peer even if the peer has enough bandwidth for a full video. Because of the coding overhead, layered video coding requires more bits than the regular video coding employed in MDC for each description. Consequently, at a given bandwidth, a P2P network with layered

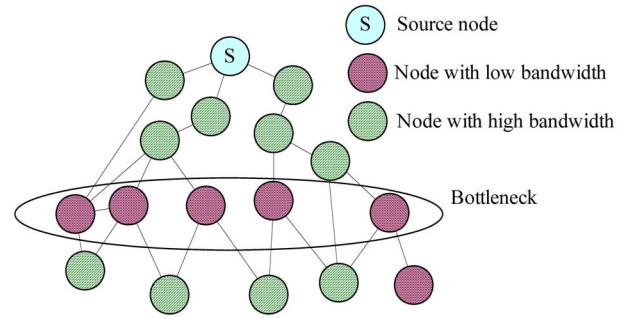


Fig. 7. Topology of a P2P network with bottleneck.

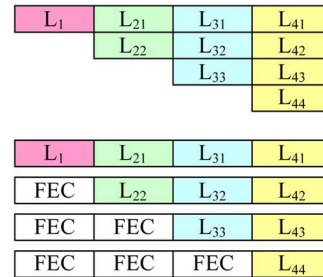


Fig. 8. Illustration of MD-FEC plus SVC with four quality layers.

streaming is more likely to run into the thin stream problem than a P2P network with the MDC scheme. When a network bottleneck occurs that can only accommodate, for example, the base layer, a downstream peer may receive only the base layer even though it has enough bandwidth to receive the enhancement layer as well (see Fig. 7). The complexity of layered coding also poses a great challenge to system implementation in practice. To sum up, LVC has higher video coding efficiency. To avoid the thin stream problem, the peers need to be clustered by the available bandwidths. On the other hand, although MDC consumes more bandwidth, it is resilient to errors and simplifies the mechanism forming the peer-to-peer network. Which approach is more appropriate for a peer-to-peer design at hand depends on the peer and available bandwidth dynamics, as well as the capacity of the underlying physical network. When the peers and network usage are stable and the network capacity is limited, the LVC with clustered peer-to-peer network approach will be better. Otherwise, MDC with the random peer-to-peer network approach appears to be a reasonable design choice.

Although the error resilience and thin stream problem of LVC can be solved by imposing the MD-FEC structure [24], [31] on LVC, the redundancy is high. For MDC-STHI, the redundancy comes from the quarter-size stream, the bit rate of which is normally set to 40% of that of a full-size stream in our simulations. Therefore, its redundancy ratio is  $0.4/1 = 40\%$ . For SVC plus MD-FEC (the layer structure of which is shown in Fig. 8), the redundancy comes from the FEC blocks. We use the MPEG-4 SVC reference software (JSVM) [25] to encode the CIF-size Foreman sequence into four quality layers corresponding to the four layers,  $E_q$ ,  $E_q + O_q$ ,  $E_f$ , and  $E_f + O_f$ , of our scheme. Table II lists the bit rate of each layer. The total bit rate allocated for FEC is  $300 \times 3 \times 150 \times 2 + 150 = 1\,350$  kb/s, whereas the total bit rate of the layered video is  $300 + 300 + 450 + 450 =$

TABLE II  
BIT RATE OF EACH LAYER FOR SVC

Quality layer	L <sub>1</sub>	L <sub>2</sub>	L <sub>3</sub>	L <sub>4</sub>
Bit rate (Kbps)	300	300	450	450

1 500 kb/s. Therefore, the redundancy ratio for this scheme is 1350/1500 = 90%, which is higher than MDC-STHI. Although the redundancy of SVC plus MD-FEC can be further reduced by unequal error protection [26] and selective transmission of only useful layers, the complexity of handling these descriptions also increases, which is not desired for a live streaming system. For these reasons, MDC-STHI is adopted in our system.

VI. SEGMENT SCHEDULING

The segment scheduling policy determines how a peer requests video segments from partners. In our design, the video stream is separated into four sub-streams and thus the decision becomes more complicated. We divide the decision process into two parts: segment type selection and supplier selection.

The selection of segment type is formulated as an optimization problem and solved by linear programming based on information such as segment availability, size, video quality and estimated available bandwidth  $B_a$ . Each peer constantly estimates  $B_a$  by taking a weighted average of the observed transmission rate  $B_{obs}$  and the previously estimated bandwidth. The initial  $B_a$  is set to the bit rate of the video.  $B_{obs}$  is a conservative estimation of the instantaneous available bandwidth and is calculated by dividing the averaged segment size by the averaged interval between the time the request for each segment is made and the time the requested segment is received for each data transmission. Because the transmitted segment traverses both the uplink of the sending peer and the downlink of the receiving peer, both the upload capacity of the sending peer and download capacity of the receiving peer is considered. Although there are more accurate bandwidth estimation methods reported in the literature [27], this simple bandwidth estimation scheme is adopted because most other methods involve sending extra probing packets, which consumes extra bandwidth.

Based on the information exchanged by each peer, a node schedules the transmission of  $K$  available segments at a time by

$$\{x_i^j\} = \arg \max_{x_i^j} \sum_{i=1}^K \sum_{j=1}^L x_i^j U_i^j$$

$$\text{s.t.} \left( T_i = T_{i-1} + 1 - \sum_{j=1}^L \frac{S_i^j}{B_i^j} x_i^j \right) > 0,$$

$$0 \leq \sum_{j=1}^L x_i^j \leq 1, \quad x_i^j \in \{0, 1\}$$

where the symbols used in this optimization problem are all defined in Table III. For the  $K$  segments to be scheduled, our mechanism determines the segment types so that the video quality  $U_i^j$  is maximized subject to the constraint  $T_i > 0$ . The constraint is imposed to avoid the underflow of the video buffer, see Fig. 9. In our system, we actually consider both the download bandwidth of the peer and the upload bandwidth of its partners in  $B_i^j$ . Simply considering the download bandwidth

TABLE III  
DEFINITION OF THE SYMBOLS USED IN SEGMENT SCHEDULING

$U_i^j$	The PSNR value of segment $i$ with type $j$
$x_i^j$	The result of segment scheduling for segment $i$ with type $j$
$P_i^j$	The subset of peers in the PartialView that have segment $i$ with type $j$
$D$	The estimated available bandwidth of the peer
$B_i^j$	$\min(D, \text{the maximum upload bandwidth of a peer in } P_i^j)$
$S_i^j$	The size of segment $i$ with segment type $j$
$T_i$	The size of the video buffer after obtaining segment $i$
$T_0$	The size of the current video buffer
$L$	The number of types

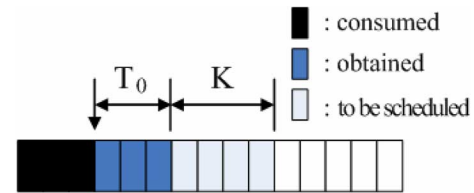


Fig. 9. Status of the buffer before scheduling.

may not work well. A node is only allowed to request one segment type for each segment. The value of  $x_i^j$  is either 0 or 1. If  $x_i^j$  equals 1, segment  $i$  is scheduled to be type  $j$ . After the type of each segment is determined, we use the rarest-first scheduling policy to decide the order of downloads and the suppliers of segments.

VII. SYSTEM EVALUATION

Simulations are conducted to examine the stability and scalability of partnership formation on a large scale network and to verify the error resilience ability of MDC-STHI. Also, the performance of NTUstreaming implemented on a small-scale campus testbed and PlanetLab [33] is evaluated.

A. Partnership Formation Evaluation

The effectiveness of three overlay network construction strategies, SCAMP, SCAMP+PRF, and SCAMP+PRF+PRS, is compared by using packet-level simulations with ns-2 [28], and the results of the three strategies are denoted as SCAMP, PRF, and TYPHOON, respectively. The overlay network size is set to 1000 peers. The physical network topology is not important in the simulation because we do not consider the delay and bandwidth constraints. The only requirement of the physical network is that any two nodes are always connected. The objective of the simulation is to test the robustness of various partnership formation mechanisms, so we only simulate the peer dynamics on the overlay network. Each peer joins and leaves the overlay network with an exponential on-off distribution, and the mean values of both on and off periods equal 250 s. The average number of peers exist in the overlay network is 500. In the initialization phase, all peers join the network in the first 5000 s. The period of subscription advertisement is 128 s, and the simulation time is 75 000 s. The main

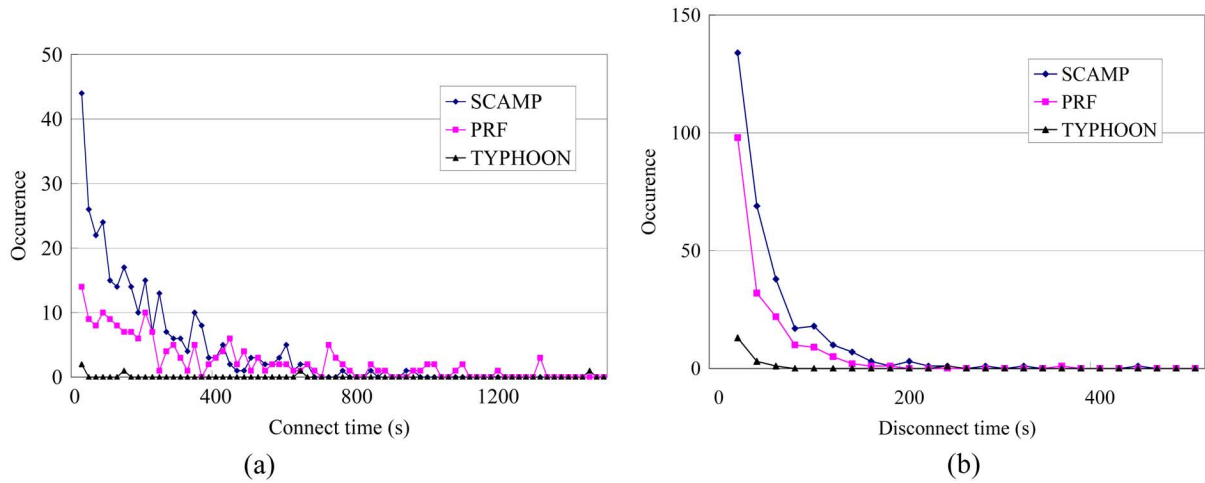


Fig. 10. Distributions of (a) connect time and (b) disconnect time.

TABLE IV  
NORMALIZED NETWORK STABILITY STATISTICS

	SCAMP	PRF	TYPHOON
Connect Time (sec)	59180	67160	74320
Disconnect Time (sec)	15820	7840	680
Number of Disconnection (N)	305	181	18
Instability	51700	22100	2580

performance metrics examined here are partnership stability and load balance.

Due to the dynamics caused by peers joining and leaving the network, some peers may become isolated and lose all the information about the remaining part of the overlay network. In such a case, the whole overlay network is in the disconnect state; otherwise it is in the connect state. The state of the overlay network varies with time, and each state may elapse for a period of time. The duration the connect state is defined as connect time, and that of the disconnect state is defined as disconnect time. Fig. 10 shows the distributions of connect time and disconnect time. As we can see, both PRF and TYPHOON outperform SCAMP in preventing frequent and sudden system disconnections [Fig. 10(a)], and fewer disconnections and shorter disconnect time than SCAMP [Fig. 10(b)].

Table IV shows the statistics of the network stability based on the three strategies. Once the network falls into the disconnect state, the number of peers partitioned from the main cluster and the disconnect time are recorded. The instability index is defined as the sum of products of the disconnect time and the number of isolated nodes. It indicates how often the video segment transmission is interrupted. The statistics show that the instability of the P2P partnership for TYPHOON is significantly lower than the other two strategies.

Fig. 11 shows the histograms of the partner size (out-degree) and consumer size (in-degree). For SCAMP, the peak of the out-degree is near 4 and that of the in-degree is near 3. The small in- and out-degrees imply that SCAMP networks may not work well in the presence of network dynamics. For TYPHOON both the peaks of in-degree and out-degree are larger than the other two strategies, implying that the network is more stable.

### B. Comparison Between SCAMP and TYPHOON over Both Static and Dynamic Networks

SCAMP and TYPHOON are further compared by the following experiment to show their vulnerability to network dynamics. The simulation contains two phases. In the first phase, SCAMP and TYPHOON are compared on a static network, which means all peers stay in the network after they join. The average number of active peers is 1000. In the second phase of the experiment, we start to add churns to the network, where all the peers start to join and leave the network according to an on-off model with the mean values of both on and off periods equal to 500 s. The first phase starts from the beginning of the simulation to the 35 000th second, and the second phase starts from the 35 001st second to the end of the simulation. Although the total number of peers in this phase is 1000, the average number of active peers is 500.

The simulation results in Fig. 12 show that the membership size of SCAMP in a static environment is consistently large and stable. However, in a dynamic environment, the overlay network of SCAMP is no longer stable. The in- and out-degrees decrease from 15 to 4. Normally, the in- and out-degrees of SCAMP in a dynamic environment are much smaller than those in a static environment. In contrast, the TYPHOON protocol uses PRF and PRS to prevent the overlay network from heavily weakening. The simulation shows that the membership, though influenced by the on-off model, is still maintained above the lower bound.

### C. Multiple Description Coding Evaluation

To demonstrate the effectiveness of MDC-STHI, four CIF-size ( $360 \times 288$ ) video sequences and one D1-size ( $720 \times 480$ ) video sequence are encoded with only I and P frames and with I frame interval being 30 frames. For CIF-size video sequences, the bit rate for each full-size stream is set to 750 kb/s, and that for each quarter-size stream is set to 300 kb/s. For the D1-size video sequence, the bit rate for each full-size stream is set to 1500 kb/s, and 600 kb/s for each quarter-size stream. When one description stream is lost or deliberately dropped due to bandwidth limitation, only the full-size even frames and quarter-size

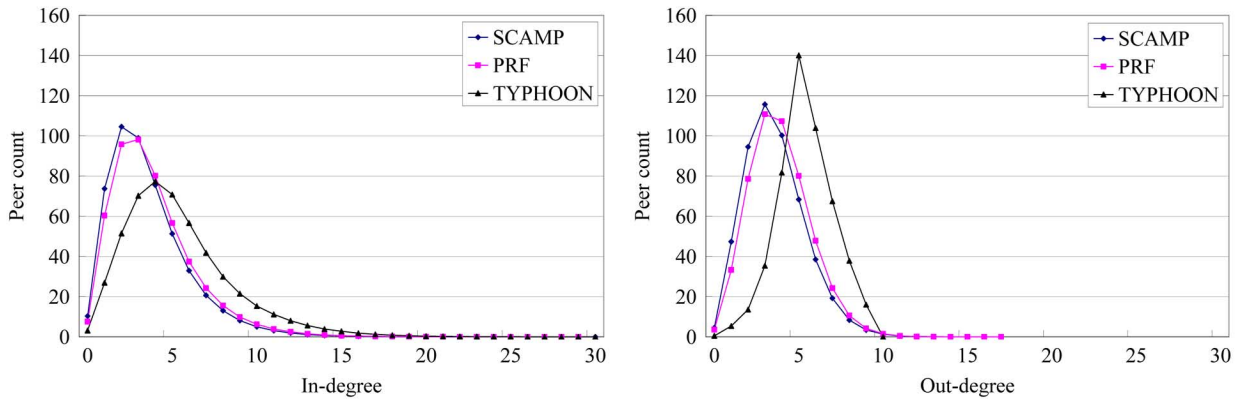


Fig. 11. In-degree and out-degree distributions.

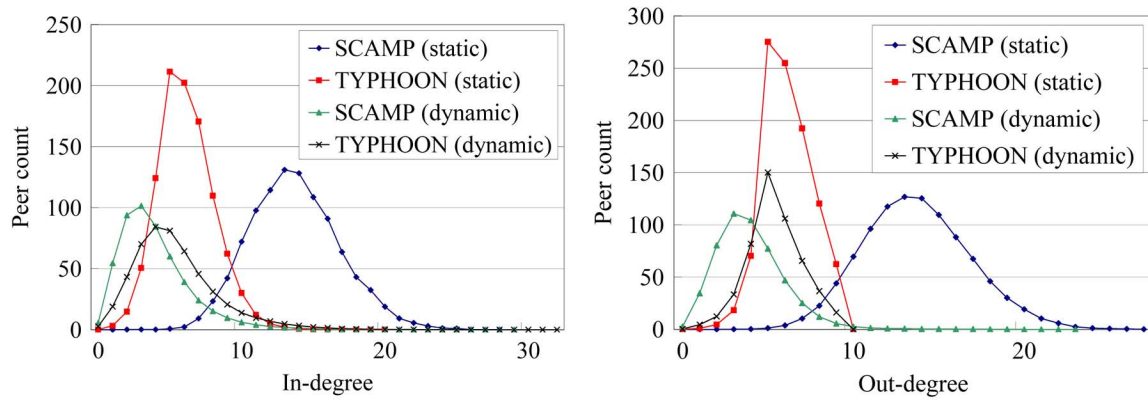


Fig. 12. In-degree and out-degree distributions for both static and dynamic networks.

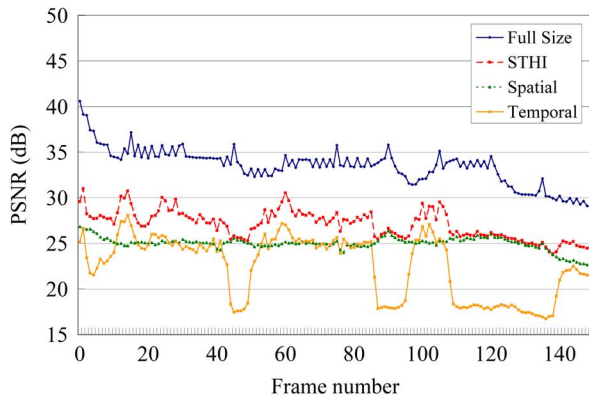


Fig. 13. PSNR values of the CIF-size Stefan video sequence.

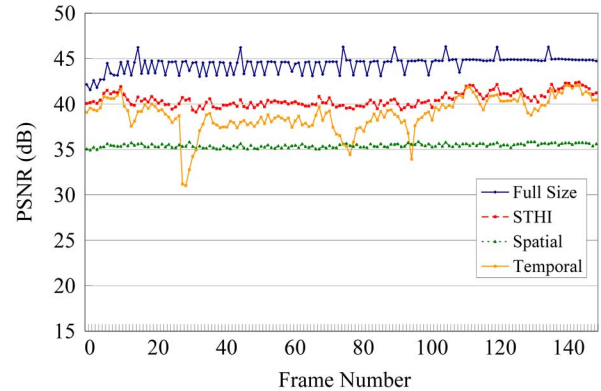


Fig. 14. PSNR values of the CIF-size Mother and Daughter video sequence.

odd frames are received. In this case,  $E_f$  and  $O_q$  are interpolated to produce a full-size video sequence.

The PSNR values of the interpolated odd frames are shown in Figs. 13–16. The red curve represents hybrid interpolation, and the green and orange ones, respectively, represent spatial and temporal interpolation. These figures show that the spatial interpolation scheme performs better for videos with fast motion while the temporal interpolation scheme performs better for videos with slow motion. Furthermore, STHI outperforms both the spatial and temporal interpolation schemes. The frames obtained from different interpolation schemes for the CIF-size

Foreman video sequence are shown in Figs. 17–20. The circled areas clearly show the difference in picture quality between these schemes. Fig. 18, which is generated by STHI, shows better quality than Figs. 19 and 20.

#### D. Evaluation on a Small-Scale Testbed

To examine the video quality based on the proposed partnership formation, video coding, and scheduling mechanisms, the system is implemented on a campus testbed. Two sets of experiments are performed. The first set focuses on the impact of partnership formation protocol on video quality, and the second compares the bandwidth-aware segment scheduling



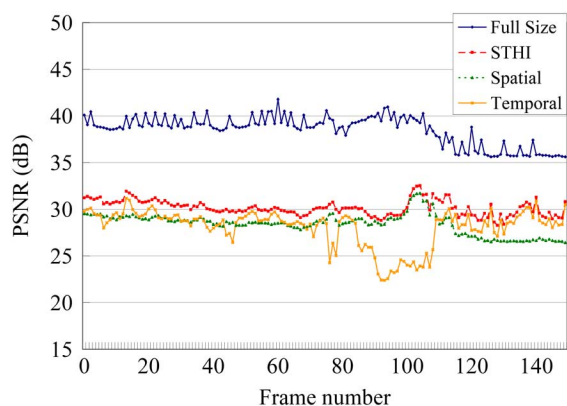


Fig. 15. PSNR values of the CIF-size Foreman video sequence.

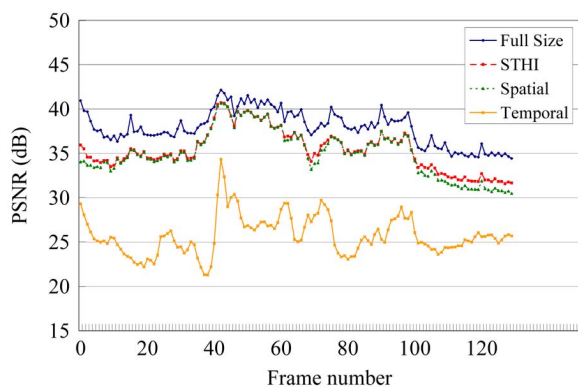


Fig. 16. PSNR values of the D1-size Football sequence.



Fig. 17. The 33rd frame of the full-size odd stream.

scheme with the rarest-first scheme. In these experiments, twelve computers in the campus are used.

*Impact of Partnership Formation:* In this set of experiments, SCAMP and TYPHOON separately distribute the CIF-size Foreman video sequence at 400 kb/s. The bandwidth available for each node is significantly higher than 400 kb/s, so segment loss is due to segment request failure rather than bandwidth limitation. For each protocol, the experiments are repeated three times to collect average PSNR values, and the simulation time is 1800 s for each trial. Since the experimental results are consistent over time, Fig. 21 only shows the data averaged over twelve nodes and three trials for a period of 500 s. As shown in



Fig. 18. The 33rd frame of the weighted interpolated odd stream.

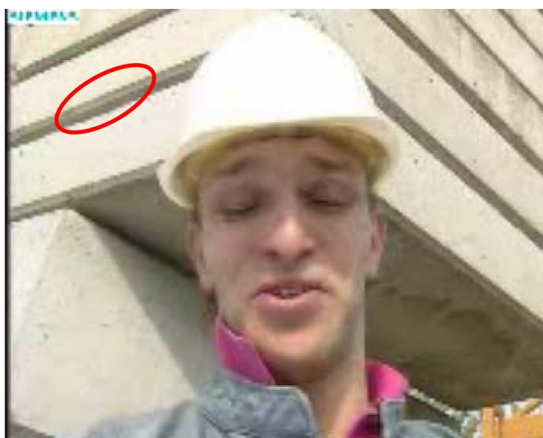


Fig. 19. The 33rd frame of the spatially interpolated odd stream.



Fig. 20. The 121st frame of the temporally interpolated odd stream.

Fig. 21, the video quality of SCAMP drops occasionally due to segment request failure. In contrast, video quality drops seldom occur for TYPHOON.

*Impact of Segment Scheduling:* In this set of experiments, MDC-STHI with bandwidth-aware scheduling and MPEG-4 with rarest-first scheduling is compared. The partnership formation protocol used is TYPHOON.

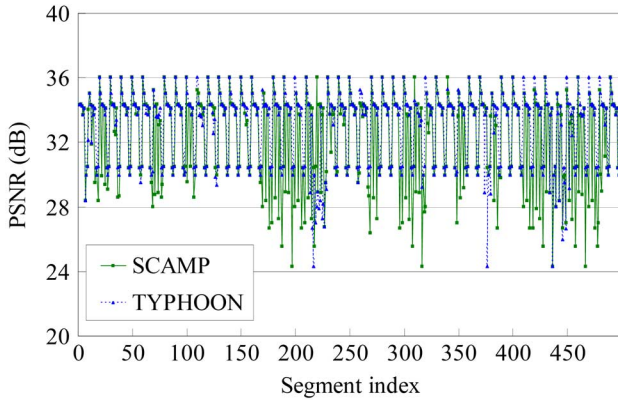


Fig. 21. PSNR values for different partnership protocols.

For MDC-STHI with bandwidth-aware scheduling, the CIF-size Foreman video sequence is encoded into four sub-streams:  $E_f, E_q, O_f$  and  $O_q$ . The bit rates of  $E_f$  and  $O_f$  are set to 750 kb/s and the bit rates of  $E_q$  and  $O_q$  are set to 300 kb/s. Three segment types are defined: “ $E_f E_q + O_f O_q$ ”, “ $E_f + O_q$ ” and “ $O_q$ .” Segments of different types are different in size, redundancy, and PSNR value. Nodes having  $E_f E_q + O_f O_q$  segments can serve nodes requesting  $E_f + O_q$  or  $O_q$  segments, and nodes having  $E_f + O_q$  segments can serve nodes requesting  $O_q$  segments. In the experiment, ten nodes with fixed bandwidth 3.0 Mbps always request segments with type “ $E_f E_q + O_f O_q$ ”. In addition, we set up two special nodes A and B to demonstrate the performance of MDC-STHI under the bandwidth constraint. Node A is configured with varying bandwidth ranging from 0.8 Mbps to 2.0 Mbps and the average bandwidth is 1.5 Mbps. Node B is also configured with varying bandwidth ranging from 0.8 Mbps to 1.2 Mbps and average bandwidth 1.0 Mbps. All the twelve nodes follow the bandwidth-aware segment scheduling mechanism. The computation time of the optimized scheduling for  $K = 10$  and  $L = 3$  is at most 20 ms. This guarantees that the optimal schedule can be obtained in real time.

For MPEG-4 with rarest-first scheduling, the CIF-size Foreman video sequence is encoded at 1.5 Mbps. Ten nodes are configured with a fixed bandwidth 3.0 Mbps and two nodes C and D are configured with varying bandwidth. Node C has the same bandwidth setting as node A, and node D has the same bandwidth setting as node B.

The experiments are repeated three times to obtain the average results, and the running time for each trial is greater than 500 s. Since the experimental results are consistent over time, only a period of 100 s is extracted to show the difference between MDC-STHI with coded-aware scheduling and MPEG-4 with rarest-first scheduling. Fig. 22 shows that node A exhibits more stable video quality than node C does. However, the difference is not quite obvious because the bandwidths of node A and C are both sufficient to retrieve the video of 1.5 Mbps. Fig. 23 shows the scenario of nodes with insufficient bandwidth. MDC with bandwidth-aware scheduling allows bandwidth-limited nodes to acquire lower quality frames to preserve a stable and smooth display. This also proves that the NTUS-streaming system adapts to the bandwidth heterogeneity in the network.

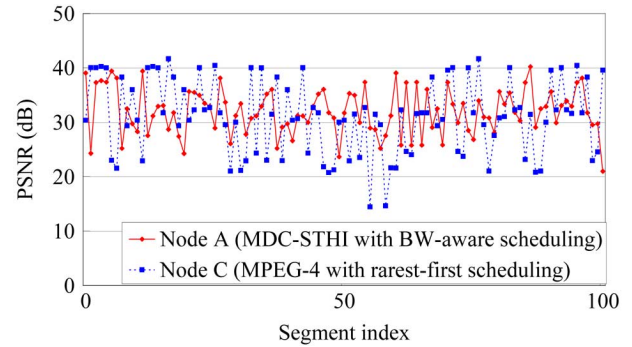


Fig. 22. PSNR values for MDC with bandwidth-aware scheduling and MPEG-4 with rarest-first scheduling. Node A and C both have sufficient bandwidth.

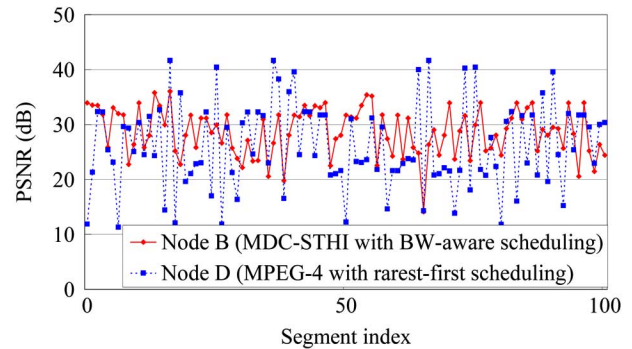


Fig. 23. PSNR values for MDC with bandwidth-aware scheduling and MPEG-4 with rarest-first scheduling. Node B and D both have insufficient bandwidth.

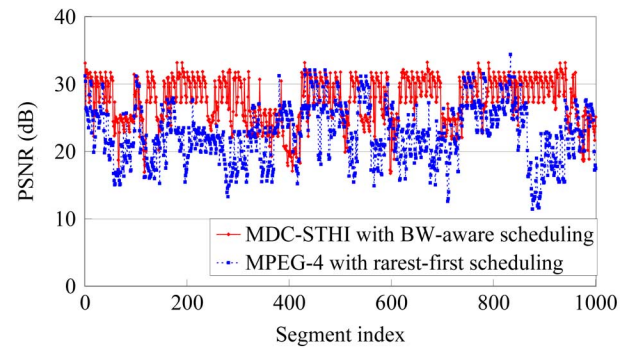


Fig. 24. PSNR values for MDC with bandwidth-aware scheduling and MPEG-4 with rarest-first scheduling on PlanetLab.

### E. Evaluation on PlanetLab

In this experiment, MDC-STHI with bandwidth-aware scheduling and MPEG-4 with rarest-first scheduling are further compared using the PlanetLab testbed. For MPEG-4, the CIF-size Foreman sequence is encoded at 500 kb/s. For MDC-STHI, the bit rates of  $E_f$  and  $O_f$  are set to 200 kb/s, and the bit rates of  $E_q$  and  $O_q$  are set to 50 kb/s. The overlay network consists of 100 nodes, some of which are connected via DSL. Because the performance of MPEG-4 is the same as that of MDC-STHI for nodes with high bandwidth, we only show the average video quality of four DSL nodes with bandwidth ranging from 300 kb/s to 350 kb/s in Fig. 24. The results are similar to the one in the previous section and prove that MDC-STHI with bandwidth-aware scheduling can achieve better quality for bandwidth-limited nodes.

## VIII. CONCLUSION

In summary, we have presented the networking and video coding co-design of NTUStreaming, which consists of three key components: TYPHOON, MDC-STHI, and coded-aware scheduling. TYPHOON makes our system work equally well over both static and dynamic networks. MDC-STHI makes our system resilient to data loss and adaptive to network heterogeneity. Bandwidth-aware scheduling integrates the above two components to provide the best video quality. The experimental results indicate that the combination of P2P streaming with multiple description coding is a promising approach to adaptive provisioning of video broadcast on large-scale heterogeneous networks.

## ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their helpful comments on this paper.

## REFERENCES

- [1] [Online]. Available: <http://www.ppstream.com/PPStream>
- [2] A. J. Ganesh, A. M. Kermarrec, and L. Massoulié, "Peer-to-peer membership management for gossip-based protocols," *IEEE Trans. Comput.*, vol. 52, no. 2, pp. 139–149, Feb. 2003.
- [3] J. G. Apostolopoulos, "Error-resilient video compression via multiple state streams," in *Proc. Int. Workshop Very Low Bitrate Video Coding*, Kyoto, Japan, Oct. 1999, pp. 168–171.
- [4] J. G. Apostolopoulos, "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," in *Proc. Visual Commun. Image Process.*, San Jose, CA, Jan. 2001, pp. 392–409.
- [5] G. Zhang and R. L. Stevenson, "Efficient error recovery for multiple description video coding," in *Proc. Int. Conf. Image Process.*, Singapore, Oct. 2004, pp. 829–832.
- [6] M.-T. Lu, C.-K. Lin, J. Yao, and H. H. Chen, "Multiple description coding with spatial-temporal hybrid interpolation for video streaming in peer-to-peer networks," in *Proc. IEEE Packet Video Workshop*, Hangzhou, Zhejiang, China, Apr. 2006.
- [7] D. Waitzman, C. Patridge, and S. Deering, Distance Vector Multicast Routing Protocol Request for Comments 1075 IETF Network Working Group Nov. 1988.
- [8] D. Estrin, D. Farinacci, A. Helmy, D. Thaler, S. Deering, M. Handley, V. Jacobson, C. Liu, P. Sharma, and L. Wei, "Protocol independent multicast-sparse mode (PIM-SM): Protocol specification," Request for Comments 2362 Jun. 1998, IETF Network Working Group.
- [9] S. Floyd, V. Jacobson, C. Liu, S. McCanne, and L. Zhang, "A reliable multicast framework for light-weight sessions and application level framing," *IEEE/ACM Trans. Netw.*, vol. 5, no. 6, pp. 784–803, Dec. 1997.
- [10] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker, "Application-level multicast using content-addressable networks," in *Proc. 3rd Int. Workshop Networked Group Commun.*, London, U.K., Nov. 2001, pp. 14–29.
- [11] D. A. Tran, K. A. Hua, and T. T. Do, "A peer-to-peer architecture for media streaming (zigzag)," *IEEE J. Select. Areas Commun.*, pp. 121–133, Jan. 2004.
- [12] X. Zhang, J. Liu, B. Li, and T.-S. P. Yum, "DONet/CoolStreaming: A data-driven overlay network for live media streaming," in *Proc. IEEE INFOCOM*, Miami, FL, Mar. 2005, vol. 3, pp. 2102–2111.
- [13] M. Castro, P. Druschel, A.-M. Kermarrec, A. Nandi, A. Rowstron, and A. Singh, "SplitStream: High-Bandwidth multicast in cooperative environments," in *Proc. ACM SOSP03*, New York, Oct. 2003.
- [14] X. Liao, H. Jin, Y. Liu, L. M. Ni, and D. Peng, "AnySee: Peer-To-Peer live streaming," in *Proc. IEEE INFOCOM*, Barcelona, Spain, Apr. 2006.
- [15] J.-C. Wu, K.-J. Peng, M.-T. Lu, C.-K. Lin, Y.-H. Cheng, P. Huang, J. Yao, and H. H. Chen, "HotStreaming: Enabling scalable and quality IPTV services," in *Proc. IPTV Workshop Conjunction 15th Int. World Wide Web Conf.*, Edinburgh, U.K., May 2006.
- [16] V. N. Padmanabhan, H. J. Wang, and P. A. Chou, "Resilient peer-to-peer streaming," in *Proc. IEEE Int. Conf. Network Protocols*, Atlanta, GA, Nov. 2003, pp. 240–247.
- [17] M. Zink and A. Mauthe, "P2P streaming using multiple description coded video," in *Proc. Euromicro Conf.*, Rennes, France, Sep. 2004, pp. 240–247.
- [18] S. Khan, R. Schollmeier, and E. Seimbach, "A performance comparison of multiple description video streaming in peer-to-peer and content delivery networks," in *Proc. IEEE Int. Conf. Multimedia Expo*, Taipei, Taiwan, R.O.C., Jun. 2004, pp. 503–506.
- [19] Y. Cui and K. Nahrstedt, "Layered peer-to-peer streaming," in *Proc. Int. Workshop Network Oper. Syst. Support For Digital Audio Video*, Monterey, CA, Jun. 2003.
- [20] R. Rejaie and A. Ortega, "PALS: Peer-To-Peer adaptive layered streaming," in *Proc. Int. Workshop Network Oper. Syst. Support For Digital Audio Video*, Monterey, CA, Jun. 2003.
- [21] A. J. Demers, D. H. Greene, C. Hauser, W. Irish, and J. Larson, "Epidemic algorithms for replicated database maintenance," in *Proc. 6th Annu. ACM Symp. Principles Distributed Comput. - PODC 87*, PODC 87, Aug. 1987, pp. 1–12.
- [22] P. T. Eugster, R. Guerraoui, S. B. Handurukande, A.-M. Kermarrec, and P. Kouznetsov, "Lightweight probabilistic broadcast," in *Proc. IEEE Int. Conf. Dependable Syst. Networks (DSN2001)*, 2001.
- [23] J.-R. Ohm, "Advances in scalable video coding," *Proc. IEEE*, vol. 93, no. 1, pp. 42–56, Jan. 2005.
- [24] R. Puri and K. Ramchandran, "Multiple description source coding through forward error correction codes," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Asilomar, CA, 1999.
- [25] J. Reichel, H. Schwarz, and M. Wien, "Joint scalable video model JSVM-6," in *Joint Video Team ITU-T VCEG ISO/IEC MPEG, Doc. JVT-S202*, Geneva, Switzerland, Apr. 2006.
- [26] H. Cai, B. Zeng, G. Shen, Z. Xiong, and S. Li, "Error-resilient unequal error protection of fine granularity scalable video bitstreams," *EURASIP J. Appl. Signal Process.*, vol. 2006, 2006, Article ID 45412,11.
- [27] L.-J. Chen, T. Sun, G. Yang, M. Y. Sanadidi, and M. Gerla, "Monitoring access link capacity using TFRC probe," *J. Comput. Commun. Special Issue Monitoring Measurements IP Networks*, vol. 29, no. 10, pp. 1605–1613, Jun. 2006.
- [28] L. Breslau, D. Estrin, K. Fall, S. Floyd, J. Heidemann, A. Helmy, P. Huang, S. McCanne, K. Varadhan, Y. Xu, and H. Yu, "Advances in network simulation," *Computer*, vol. 33, no. 5, pp. 59–67, May. 2000.
- [29] [Online]. Available: <http://www.pplive.com/PPLive>
- [30] Y. Shen, Z. Liu, S. S. Panwar, K. W. Ross, and Y. Wang, "Streaming layered encoded video using peers," in *Proc. IEEE Int. Conf. Multimedia Expo.*, Amsterdam, The Netherlands, Jul. 2005.
- [31] X. Xu, Y. Wang, S. S. Panwar, and K. W. Ross, "A peer-to-peer video-on-demand system using multiple description coding and server diversity," in *Proc. IEEE Int. Conf. Image Process.*, Singapore, Oct. 2005, pp. 1759–1762.
- [32] "Coding of audio-visual objects, part-2 visual, amendment 4: Streaming video profile, ISO/IEC 14496-2/FPDAM4," Jul. 2000.
- [33] [Online]. Available: <http://www.planet-lab.org/PlanetLab>
- [34] C. Huang, J. Li, and K. W. Ross, "Can internet video-on-demand be profitable?," in *Proc. ACM SIGCOMM*, Kyoto, Japan, Aug. 2007.



**Meng-Ting Lu** was born in Peng-Hu, Taiwan, R.O.C. He received the B.S. degree in electrical engineering from National Taiwan University, Taipei, in 2003. He is currently pursuing the Ph.D. degree in the Graduate Institute of Communication Engineering, National Taiwan University. His research interests include video streaming, peer-to-peer streaming, and video coding.



**Jui-Chieh Wu** received the B.S. degree in 2004 and the M.S. degree in 2006 from the Computer Science and Information Engineering Department, National Taiwan University, Taipei, Taiwan, R.O.C.

He is currently a Research Assistant with National Taiwan University. His research interests include overlay network entertainment and sensor networks.



**Kuan-Jen Peng** received the B.S. degree in 2004 from the Computer Science and Information Engineering Department, National Taiwan University, Taipei, Taiwan, R.O.C., and the M.S. degree in 2006 from the Graduate Institute of Communication Engineering, National Taiwan University.

He is currently a Software Engineer in VIA Telecom. His research interests include network planning, video streaming, and overlay network entertainment.



**Polly Huang** (S'99–M'00) received the Ph.D. degree in 1999 and M.S. degree in 1997 in computer science from University of Southern California, and the B.S. degree in 1993 in mathematics from National Taiwan University, Taipei, Taiwan, R.O.C. (NTU).

In February 2003, she joined the Department of Electrical Engineering of National Taiwan University at which she currently holds an Associate Professor position. Prior to joining NTU, she worked as a Postdoctoral Research Scientist at the Computer Engineering and Networks Laboratory (TIK) of the Swiss Federal Institute of Technology (ETH), Zurich, and a Postdoctoral Fellow at the Institute of Pure and Applied Mathematics, UCLA. Her research interest includes sensor networking, multimedia networking, and Internet characterization. She serves currently on the editorial board of the *Journal of Communications and Networks*.

Dr. Huang is a member of ACM.



**Jason J. Yao** received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, R.O.C., and the Ph.D. degree in electrical and computer engineering from University of California, Santa Barbara. He also holds an MBA from Santa Clara University, Santa Clara, CA.

He has worked for AT&T Bell Labs, Fujitsu Network Transport Systems, and Fujitsu Laboratories of America with job functions in advanced research, project management, and strategic planning. He has published more than 20 papers in peer-reviewed

journals and international conferences. His research interests span from digital signal processing, telecommunications, and audio/video systems to Internet traffic engineering and bioinformatics. He is currently a Visiting Scholar at National Taiwan University, focusing on research of multimedia streaming over P2P networks. He is the PI or Co-PI of several research projects sponsored by Fujitsu, Intel, and other government agencies.



**Homer H. Chen** (S'83–M'86–SM'01–F'03) received the Ph.D. degree from the University of Illinois at Urbana-Champaign in electrical and computer engineering.

Since August 2003, he has been with the College of Electrical Engineering and Computer Science, National Taiwan University, where he is Irving T. Ho Chair Professor. Prior to that, he held various R&D management and engineering positions in leading U.S. companies including AT&T Bell Labs, Rockwell Science Center, iVast, and Digital Island over a period of 17 years. He was a U.S. delegate of the ISO and ITU standards committees and contributed to the development of many new interactive multimedia technologies that are now part of the MPEG-4 and JPEG-2000 standards. His research interests lie in the broad area of multimedia processing and communications. He was an Editorial Board Member for *Pattern Recognition* from 1989 to 1999.

Dr. Chen is an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY. He served as Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING from 1992 to 1994 and Guest Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY in 1999.