第 XXIII 部

Asian Internet Interconnection Initiatives

第23部
Asian Internet Interconnection Initiatives

第1章 Introduction

The Internet has become a critical and dependable infrastructure for today's society. This was clearly seen when an earthquake in Taiwan on 26 December 2006 knocked out several undersea cables off Taiwan. Some Asian countries, including Taiwan, Singapore, Thailand, Philippines, and Indonesia, were severely affected. However, satellite connectivities were not affected by this disaster. Thus we can see this as a role of satellite Internet today, in addition to providing connectivities where terrestrial links are still scarce.

Asian Internet Interconnection Initiatives, or AI^3 (*ei-triple-ai*) in short, was established in 1995, in order to work for the Internet development in Asian region[2]. When we started this project, we set some assumptions on what is required to accelerate the deployment process of the Internet: (1) a testbed network as a live demonstration and also as a technical showcase of the Internet technology is required because it always can persuade many people of the potential and possibility for the power of the Internet, (2) research for adapting and localizing the Internet to the region should be conducted simultaneously with the deployment, because the Internet is aiming to be an infrastructure for our society, and (3) human resource development locally in the region is vital for rapid deployment of the Internet because the human resource development process can reproduce more evangelists, supporters and participants for the Internet deployment.

With these assumptions, the AI^3 project decided to start as a research consortium of leading research groups in universities in Asia[188]. Because universities are in charge of human resource development, less restricted to have a testbed network, and a base of research activities, we expect we can find out there many researchers who are working actively on the Internet technologies.

In our 11 years of activities, AI^3 testbed network has been built to connect 24 universities in 12 countries in this region and still expanding. This network has been working on 24/7 basis and turned to be its communication infrastructure for members of this AI^3 project. Recently, we are not only focus on conducting satellite research activity but also conducting IPv6 research activity. In this report, we describe all conducted research activities and research publications in 2007.

第2章 Research and Development

In this chapter, we describe the research activities and achivements in this year.

2.1 Overview of AI^3 Network: Design and Applications of Satellite Network

The main content of the section is published in SIGCOMM 2007 Network Systems for Developping Region Workshop[189].

2.1.1 Summary

Many developing countries do not get benefit from the Internet as the network infrastructure is not yet built due to economic or administrative restrictions of each region. Asian Internet Interconnection Initiatives (AI^3) has been developing a research consortium of 29 partner institutions in 13 countries utilizing Internet infrastructure over a satellite link. The project aims to develop partnerships and human networks to foster researchers within the region to carry out research satellite Internet and to develop the region. As result of more than 10 years of experiments, we have successfully conducted fruitful researches and also achieved to develop a distance learning environment. The research aims to share the operational know-hows and research results of AI^3 activities to construct network and applications supporting for developping region.

2.2 Architecture of Satellite Internet for Asia-wide Digital Communications

The main content of the section is published in Aintec 2007[82].

2.2.1 Introduction

We published the R&D acrivity of AI^3 Working Group that discusses the network architecture of an Asia-wide satellite Internet that considers the situations in developing regions. The design considerations for the architecture are costs, effective use of satellite bandwidth, scalability, and routing strategy when combined with terrestrial links. The architecture adopts one-way shared satellite links to reduce costs, IP multicast to leverage the broadcast nature of satellite links, QoS, audio-video application gateway to adapt to the limited bandwidth of satellite links.

We summarised the characteristics of satellite communication as follows: (1) wide-area coverage, (2) quick installation, (3) independent from terrestrial infrastructure, and (4) broadcast capability. Hence, satellite communication can be leverage to give broadband access to places where terrestrial infrastructure is still underdeveloped, and is a viable option to build a network in developing regions, such as many parts of Asia.

Our focus on this research is sharing the technical aspects of networking using satellite link, that we have experienced through design, installation, operation and R&D activities on the AI^3 satellite Internet.

2.2.2 Satellite Internet Architecture

In this document, we raised 5 requirements for implementation of IP network using satellite communications as shown in Table 2.1.

Figure 2.1 shows the architecture of satellite Internet used by AI^3. There are several technical components for IP communication over satellite. The components of this architecture provide the solutions to meet the requirements discussed above. Table 2.2 summarizes the components and the requirements they will satisfy.

2.2.3 Current Achievement

We discuss that what points are achieved in the architecture. Then we evaluated our achievement qualitatively by analyzing whether each requirement is satisfied or not. Briefly, S1–S4 have been achieved, and S5–S7 are the current challenges, while S8 is under discussion for the future work.

One-way shared satellite link is proved to significantly reduce the cost on installation and operation of satellite link. UDL mesh also exhibited the potential to reduce bandwidth consumption while keeping same information speed as P2P mesh. Therefore, R1 is satisfied.

The combination of one-way shared satellite

Table 2.1. Requirements

Item	Description
R1	Cost reduction on installation and operation of satellite communications
R2	Engineering to maximize utilization of radio spectrum
R3	IPv6 to accommodate a large number of connecting nodes
R4	Scalability of network protocols
R5	Routing strategy in the environment where multiple paths are available including a satellite communication

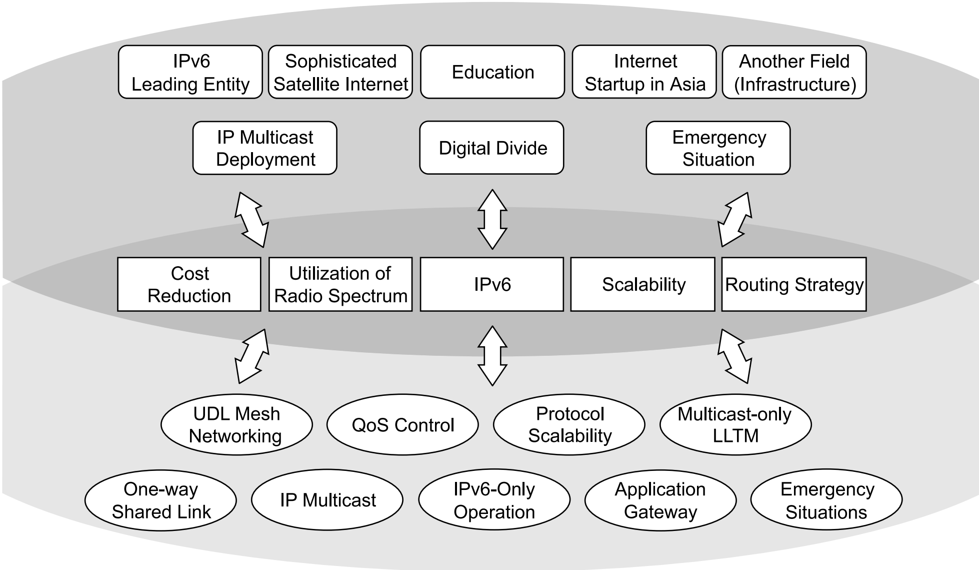

Fig. 2.1. Technology Architecture

Table 2.2. Components of Architecture

Item	Component	Target
S1	One-way shared satellite link	R1, R2
S2	IP Multicast	R2
S3	QoS traffic management	R2
S4	Application gateway for video/audio communication	R2
S5	UDL mesh networking	R1, R2
S6	Transition to IPv6 operation	R3
S7	Scalability of UDL	R1, R4
S8	Sophisticated inter-AS routing using satellite	R5

link and the deployment of IP multicast on the satellite UDL is the biggest contribution for efficient utilization of radio spectrum. QoS traffic management let prioritized traffic and other traffic coexist by guaranteeing or limiting their bandwidth. The video/audio communication between a very high-speed network and our satellite network can be done by rate adaptation and media conversion using application gateway. Given the situation that SOI Asia lecture can be delivered from the global Internet to Asian regions with as good quality as possible with a very limited bandwidth resource, R2 is satisfied.

In addition to OSPFv3 and BGP4+ operation in the AI^3's backbone network, transition to IPv6 operation in receive-only network via the satellite UDL will make our entire network ready for IPv6 in both routing and application. Some AI^3 partner sites have already completed the transition to IPv6 operation, and R3 is satisfied adequately and to be met completely in the near future.

For the scalability issues of satellite network, our achievement is still partial and does not completely satisfy R4 at this moment, There can be many network protocols to work on a large-scale satellite UDL, and they may suffer from limited bandwidth and long delay of the satellite. Hence we still need much effort to satisfy R4.

As a result, the evaluation can be concluded that the architecture described in this document has achieved R1, R2 and R3 adequately. However, this architecture is not completed yet for potential large-scale or advanced usage of this network as set in R4 and R5 requirements.

2.2.4 Conclusion

This document has described the benefits and concerns, and current challenges for architecture of a wide-area satellite Internet based on our experiences and implementations in AI^3 network and SOI Asia.

Although our activity is on-going and needs more evaluation for some parts of the architecture, our architecture can be also applied to establish IP networks on other wireless communications because satellite communication is also a part of them. We are going to continue the operation and deployment of satellite Internet in Asia, research and development on the current and potential challenges on technology to establish a better Asia-wide digital communication infrastructure in the global Internet.

2.3 IPv6-only Operation in UDL RO Sites

AI^3 network started the migration of RO sites into IPv6-only operation environment on July 7, 2007. The objective of running IPv6-only operation is to gain a complete experience in operating and using IPv6. We believe that the experiences from this operation will be beneficial to the Internet community.

The environment of the IPv6-only operation is depicted in Figure 2.2. We introduced the components shown in bold font to enable the IPv6-only operation:

1. IPv6-enabled Squid proxy server
2. NAT-PT (Network Address Translation–Protocol Translation)
3. totd (trick or treat daemon)
4. Private IPv4 address

2.3.1 IPv6-enabled Squid

We run a patched Squid 2.6 proxy server to make it IPv6-enabled for the IPv6-only operation. A host on the RO network is configured to use the IPv6-enabled Squid on its network and the Squid proxy server is configured to use the IPv6-enabled Squid proxy server in SFC as its parent (Figure 2.3). The default access method to the Internet from the IPv6-only networks is via the Squid proxy server. We have confirmed that the following applications can run using the proxy server:

- Internet Explorer 6
- Mozilla Firefox 1.5
- MSN Messenger
- Yahoo! Messenger
- KVIrc

Applications that cannot use Squid proxy server connect to the Internet using the NAT-PT mechanism, which is explained later. Figure 2.4 shows the incoming and outgoing traffic of the IPv6-enabled Squid server located in SFC.

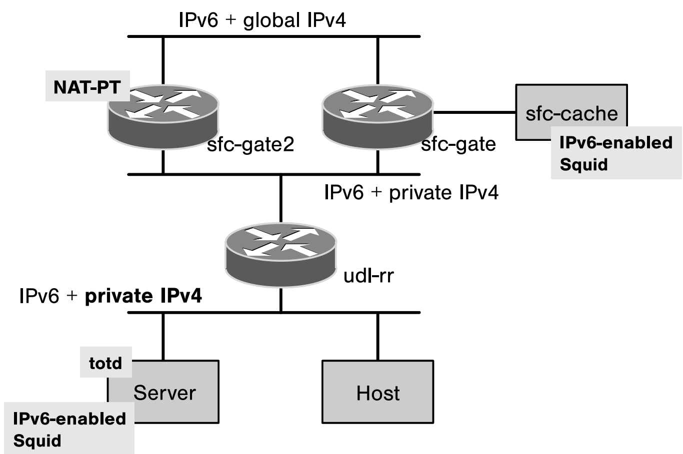

Fig. 2.2. IPv6-only operation in AI^3 network.

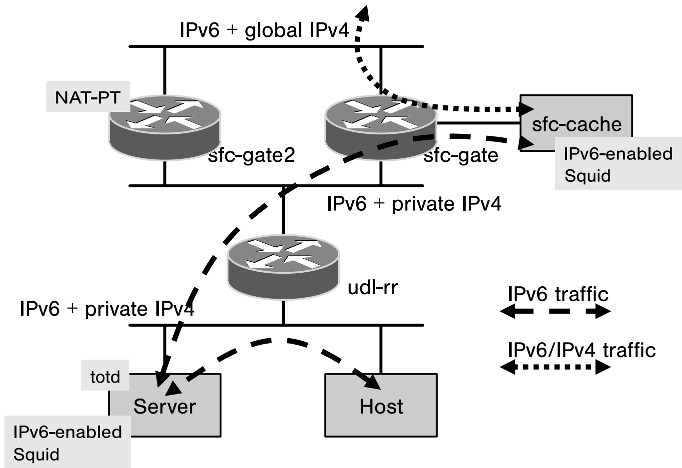

Fig. 2.3. Accessing the Internet via IPv6-enabled Squid proxy server.

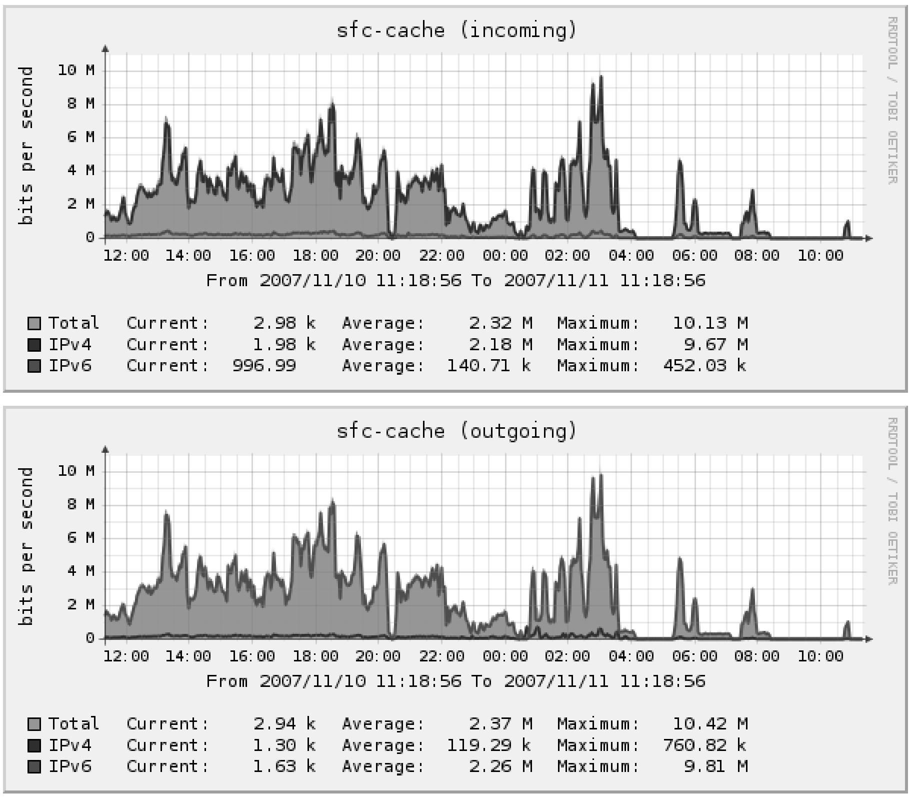

Fig. 2.4. Incoming and outgoing traffic of IPv6-enabled Squid proxy server.

2.3.2 NAT-PT

NAT-PT (Network Address Translation–Protocol Translation), defined in RFC 2766, is a mechanism to enable communications between IPv6 hosts and IPv4 hosts by providing the translation from an IPv6 address to an IPv4 address, vice versa. AI^3 implements NAT-PT by assigining the 32-bit IPv4 address space into a /96 IPv6 prefix and configure sfc-gate2, which is a Cisco 3840 router with IOS 12.4(13r)T, as the NAT-PT router. When an IPv6 host sends a packet to the /96 IPv6 prefix, the traffic flows into sfc-gate2 and sfc-gate2 will perform the necessary translations and then originate an IPv4 packet corresponding to the IPv6 packet. The source address of the IPv4 packet is selected from the IPv4 address pool

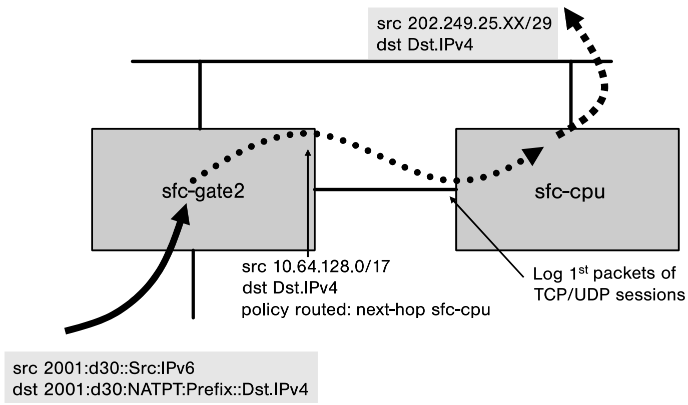

Fig. 2.5. NAT after NAT-PT to scale communication sessions.

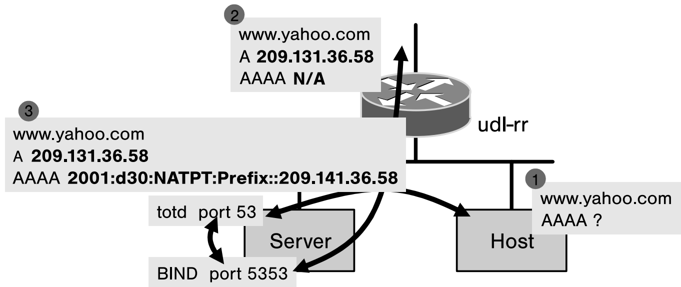

Fig. 2.6. totd translates IPv4 addresses into IPv6 addresses in DNS queries.

dedicated for NAT-PT.

The NAT-PT implementation of this Cisco requires a one-to-one mapping between an IPv6 address and an IPv4 address. It means that the number of IPv6 hosts that can communicate with IPv4 hosts at the same time is limited to the number of IPv4 addresses in the address pool. AI^3 doesn't have a large address space, therefore we use a /17 private IPv4 address space for the NAT-PT IPv4 address pool. The network then can have a large number of NAT-PT-ed communication sessions that is sufficient for AI^3 operation. IPv4 packets resulted from NAT-PT are policy-routed to another router and receiving another NAT process at the next router to make it routable to the Internet (Figure 2.5). This NAT router records the address translations from the private IPv4 address space to the global IPv4 address mainly for security purposes, e.g., when a security incident originated from an IPv6 host happens, we can trace the source using the NAT-PT and NAT translation tables.

2.3.3 totd (trick or treat daemon)

totd is a DNS proxy server that translates IPv4 addresses contained in **A** DNS responses into IPv6 addresses by prepending an IPv6 prefix into the IPv4 address when a host send a **AAAA** query to totd. This software generates the IPv6 address which will then be used in combination with NAT-PT to provide a transparent access mechanism from the IPv6-only network to the Internet. Figure 2.6 shows the totd mechanism running on AI^3 IPv6-only network.

2.3.4 Private IPv4 address

The IPv6-only network still uses private IPv4 address space as the hosts on the network are a mix of FreeBSD, Linux, and Windows machines. Windows machines can use IPv6 for communications, but it still uses IPv4 for the DNS query and responses. The private IPv4 addresses are used for these DNS query and responses and IPv4 packets are not routed to the Internet. IPv4 address may be used as a fallback protocol if there is a problem with the IPv6 network. In such a case, **udl-rr** will route IPv4 traffic to SFC and SFC will perform NAT on IPv4 packets.

2.3.5 Status and Future Plan

The first partner to migrate its UDL RO site to IPv6-only operation is Universiti Sains Malaysia (USM) on July 7, 2007. As of January 7, 2008, nine partners are confirmed to have finished the IPv6-only operation (Table 2.3). University of Health Sciences Cambodia (UHSC) and Asian Institute of Technology (AIT) are not expected to fully migrate to IPv6-only operations, since they are involved in a peroject that uses an application that is not IPv6-enabled. Other partners are expected to finish the migration by March 2008.

The future activities in IPv6-only network are:

- Finishing partner's migration.
- Monitoring IPv6 and NAT-PT generated traffic.
- Advocating the use of IPv6 to National Research and Education Networks in partner countries, and providing IPv6 connectivity to them.

2.4 UDL Mesh Experiment

2.4.1 Network Setup

For this experiment, UDL gateway at USM was either configured as a router or bridge. All of these tests were conducted by configuring UDL gateway at SFC as the default router to Internet since the bandwidth allocation at SFC is much better than that of USM. To configure USM UDL gateway as a router, NAT has to be performed on tap0 interface using netfilter. In total, NAT has to be performed twice for this setup, once on the tap0 interface of USM UDL gateway and eth0 interface of SFC UDL gateway as shown in figure 2.7.

For bridge setup, tap0 and eth0 interfaces of USM UDL gateway were bound to br0 interface as shown in figure 2.8. For this configuration, tap0 and eth0 were only acting as ports

Table 2.3. Partner's IP-only network migration status

Partner	Migration date
Universiti Sains Malaysia	Jul 7, 2007
Hasanuddin University	Oct 25, 2007
Sam Ratulangi University	Oct 25, 2007
Syiah Kuala University	Oct 25, 2007
Brawijaya University	Oct 25, 2007
Advanced Science and Technologi Institute	Oct 25, 2007
Tribhuvan University	Dec 10, 2007
University of Computer Studies, Yangon	Dec 11, 2007
Chulalongkorn University	Dec 10, 2007

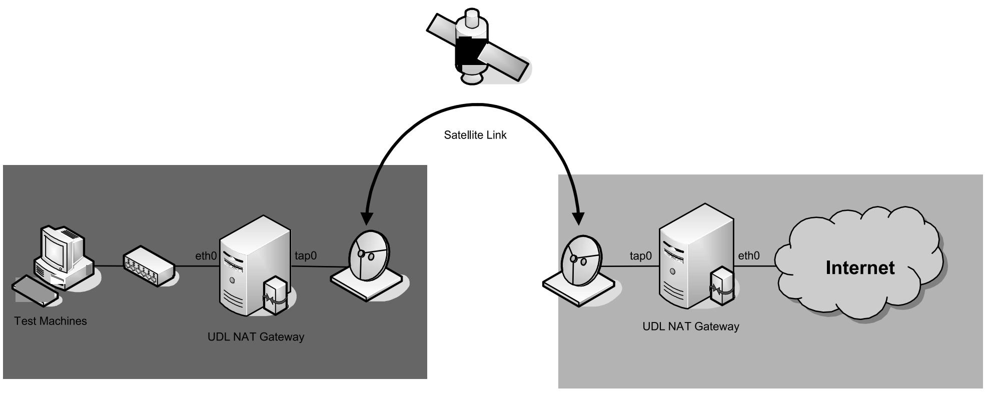

Fig. 2.7. Double NATs configuration. USM UDL gateway acting as a router

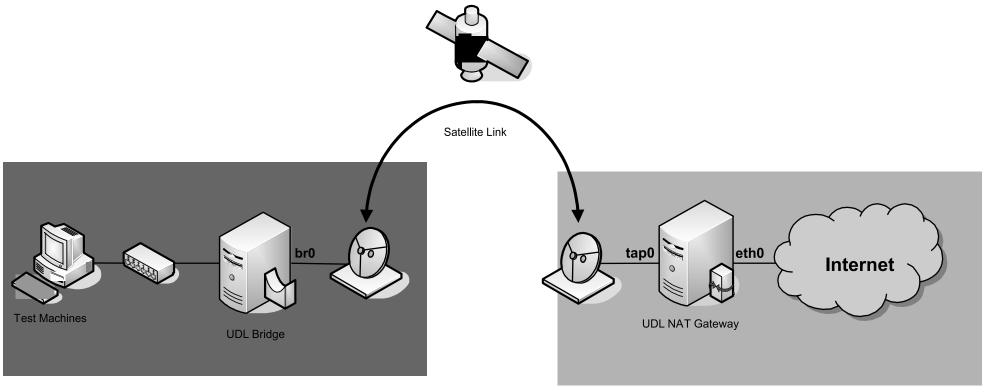

Fig. 2.8. Bridge configuration. USM UDL gateway acting as a bridge to SFC UDL gateway

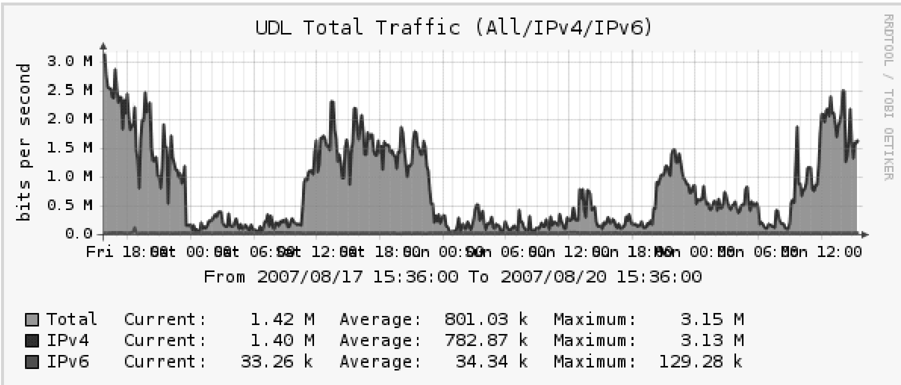

Fig. 2.9. SOI ASIA traffic during UDL mesh experiment

for br0 interface. The setup of UDL gateway at SFC remained unchanged for these 2 configurations. The experiment was conducted using bridge mode for Ultra Lightweight Encapsulation (ULE[44]). Thus, all ULE packets will carry Ethernet header.

The ULE encapsulator at SFC UDL gateway was configured to transmit at 7 Mbps of MPEG2-TS data even though the satellite link can sustain 12.791 Mbps of MPEG2-TS data. This was purposely done to avoid contention of bandwidth resource with SOI ASIA TRAFFIC. The ULE encapsulator at USM UDL gateway was configured to transmit at the rate of 3.507 Mbps.

Figure 2.9 shows that the total traffic for SOI ASIA during the experiment peaked at 3 Mbps and 7 Mbps used by UDL gateway left more than 2 Mbps of safety margin.

2.4.2 Latency

To measure latency of UDL mesh network, two sets of tests were conducted:

- Ping
- Measure Round Trip Time
- Custom tool to measure one way latency of UDL link. For this set of tests, the clocks at both UDL gateways were synchronized using NTP.

Round Trip Time

Average Round Trip Time was taken from 60 continuous ICMP replies. Measurements of RTT were repeated to packing threshold of 0 ms (padding), 3 ms and 40 ms. Again, the tests were conducted from both UDL gateways. SFC to USM shown in table 2.4 below means that ping command was initiated from SFC UDL gateway to USM UDL gateway.

Table 2.4. Average Round Trip Time (milliseconds)

Packing threshold (ms)/Direction	SFC to USM (ms)	USM to SFC (ms)
0	613.447	615.586
3	618.609	626.745
40	701.441	690.598

Table 2.5. Average one way latency (milliseconds)

Packing threshold (ms)/Direction	SFC to USM (ms)	USM to SFC (ms)
0	273.435	333.237
3	285.305	338.312
40	318.304	376.157

One Way Latency

One way latency paints a different picture. While the result of RTT doesn't differ that much no matter which site initiated ping command (for obvious reason). SFC -> USM denotes that UDP traffic was sent from SFC gateway (client) to USM gateway (server) and vice-versa. The tests were conducted using this setting:

- 30 samples of UDP datagram of 36 bytes were sent.
- Each subsequent packet was sent at 100 milliseconds interval

Table 2.5 shows that the average measured latency from SFC to USM is lower. This is probably due to the allocation of bandwidth from SFC to USM is bigger coupled with the behaviour with ASI card. There is an additional latency approximately 60 ms from USM to SFC over 3.5 Mbps link.

2.4.3 Iperf TCP Tests

For the test of TCP performance, the results were selected to exclude the effect introduced by TCP slow start. How these results were derived is discussed below. Since the results were taken with ULE bridged mode, for n byte MTU size selected.

$$\text{TCP payload} = n - 20 - 32$$

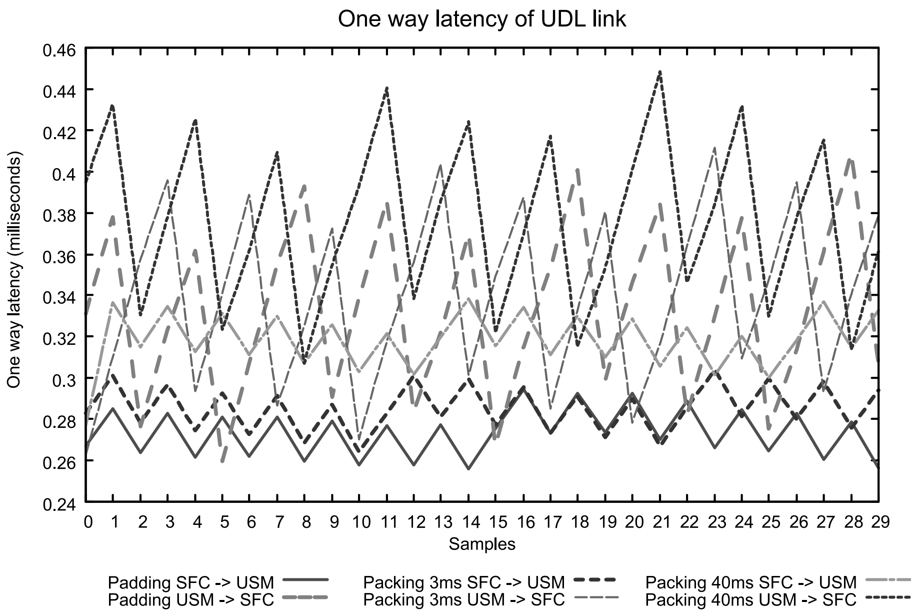

Fig. 2.10. Measured one way latency of UDL link over 30 samples sent at 100 ms interval

Table 2.6. Link Utilization (TCP traffic using default window size)

Packing Threshold (ms)	0	3	40
Normalized Link Utilization (Percentage)	6.08	5.91	5.77

20 is the size of IP header
32 is the size of TCP header
ULE payload $= n + 14$
14 is the size of Ethernet header
Link utilization = TCP throughput reported by iperf $\times (n + 14)/(n - 52)$

The formulae ignores the overhead introduced by ULE header and MPEG2-TS header. The results from figure 2.10 were derived by using 2 Mbytes of transmit and receive TCP buffer at both the client and server.

The final set of results was taken by using default window size of 64 Kbytes found in most operating system.

2.4.4 Iperf UDP Tests

For the checking link performance, we experiments under the different packet size of UDP payloads. ULE payload $= n + 14 + 20 + 8$, 14 is the size of Ethernet header, 20 is the size of IP header, 8 is the size of UDP header, Link utilization = UDP throughput reported by iperf $\times$ (ULE payload)/(UDP payload) Again, the formulae ignores the overhead introduced by ULE header and MPEG2-TS header.

It should be noted that the results in figure 2.11 had been rounded off before they were normalized. This is due to mistakes of not outputting the results in the range of kbit/s. Thus, the results in figure 2.11 may not be as accurate as it should be. There is no obvious difference in terms of link utilization if packing mode is selected.

2.4.5 Problems Encountered

- Wrong selection of DVB-S receiver card at the initial stage of experiment caused suboptimal result due to enormous amount of corrupted MPEG2-TS frames.
- MTU of tap0 interface can't be changed when tap0 is configured to be as part of bridge interface or else no traffic will go through.

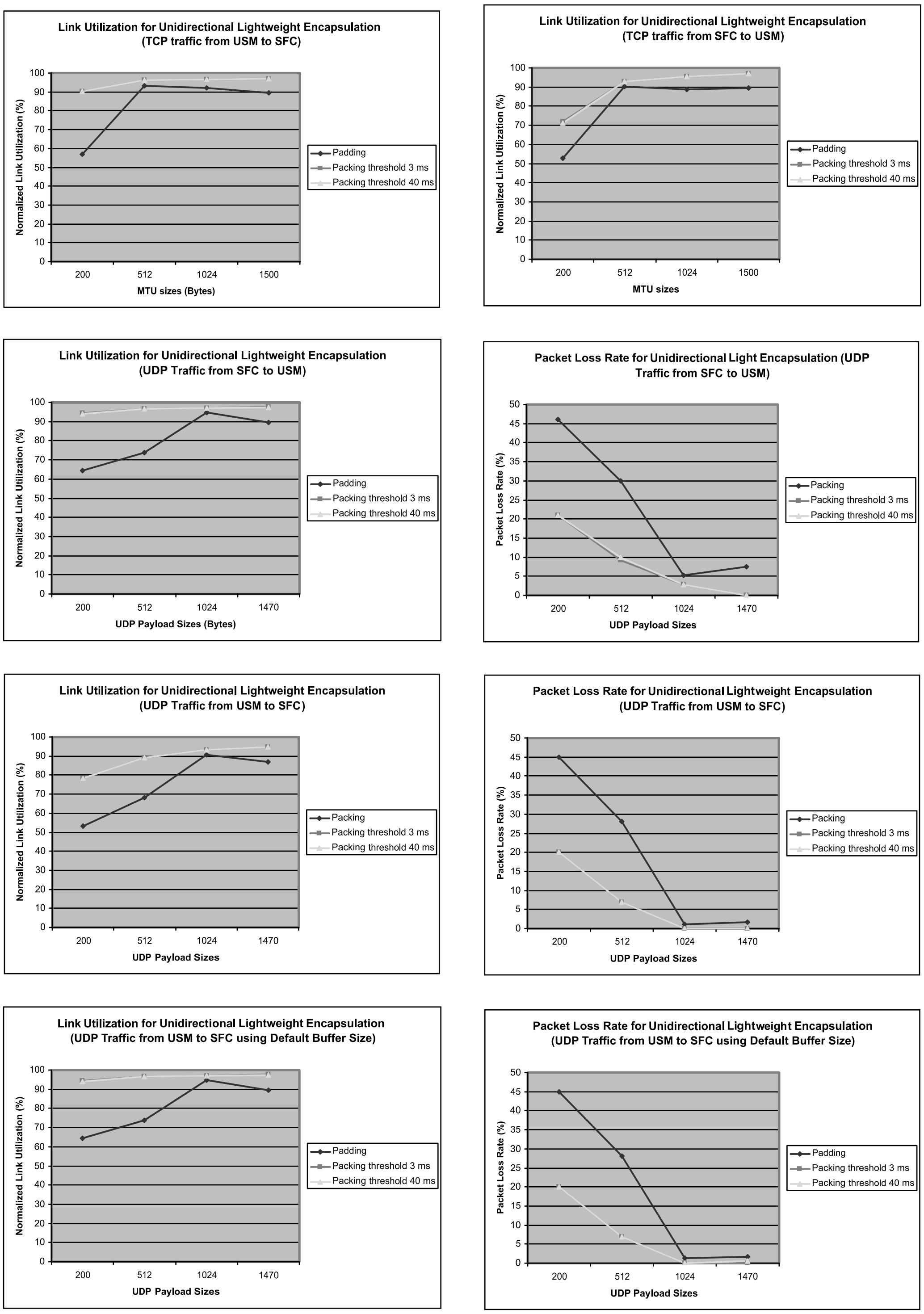

Fig. 2.11. Packet Loss Rate of UDP traffic from SFC under various packing thresholds using default buffer size