Part I

AI3 (Asian Internet Interconnection Initiatives)

Title: AI^3 Working Group activity report 2010

Authors (in alphabetical order):
Achmad Basuki, Keio University
Ang Way Chuang, Keio Ubiversity
Teruaki Yokoyama, Cyber University

Date: February 28, 2012

Chapter 1

AI3 Overview

1.1 Introduction

Asian Internet Interconnection Initiatives, or AI^3 (*ei-triple-ai*) [1] in short, was established in 1995, in order to work for the Internet development in Asian region. 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. 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. Our decision and achievements are recorded on papers [2] [3].

Today, the Internet is thought to be a critical and dependable infrastructure. This was also clearly seen when a big earthquake has happened at Tohoku area in Japan on Mar. 11 2011. 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.

This report describes all our activities in 2011.

1.2 Partners

Through our activities, AI^3 has been an international research consortium of 29 organizations in 13 Asian countries as shown in Figure 1.1. 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. Our partners include Japan, Indonesia, the Philippines, Singapore, Vietnam, Malaysia, Thailand, Cambodia, Laos, Nepal, Myanmar, Bangladesh, and Mongolia. We aim to contribute to developing a communications infrastructure as well as human resources in these regions through the project.

Figure 1.1: AI^3 partners

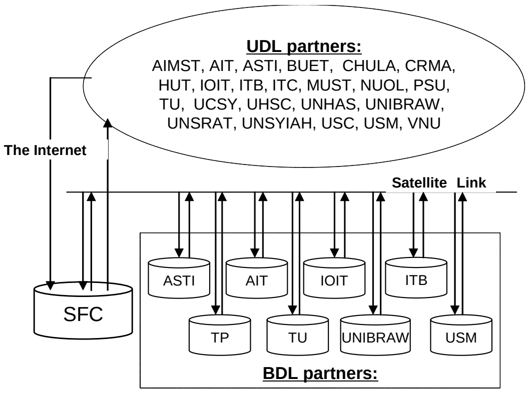

Figure 1.2: AI^3 satellite topology

Figure 1.2 shows the satellite-network topology. The earth station at SFC transmits signals between BDL partners and to UDL partners. The UDL frequency can simultaneously transmit aggregated traffic to many partners. Satellite links are shared by both BDL and UDL signal transmissions. We are developing UDLR technology [4] in this environment.

Chapter 2

Operation

2.1 Virtualization

Recently, we are working on deployment of virtualization technology in NAIST and SFC NOC. Figure 2.1 shows the architecture of the virtualization deployment in NAIST NOC.

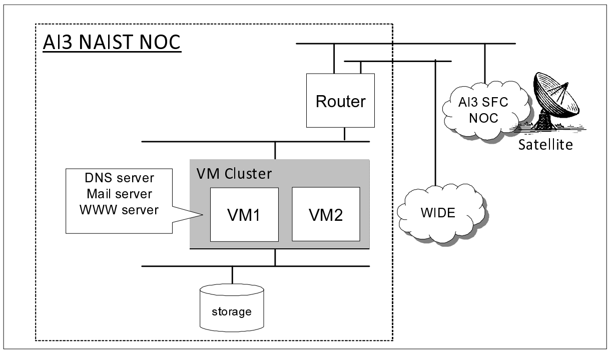

Figure 2.1: Virtualization in NAIST NOC

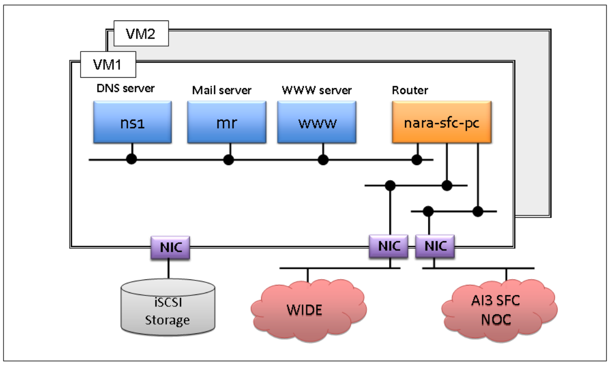

Figure 2.2: Inside of VM

We combine three AI^3 services, i.e., WWW, mail and DNS(ns1) server, and one router into one machine as shown in Fig. 2.2. These services run on top of the linux system as a virtual machine. We maintain redundancy by installing dual VM hosts. Here, we use a Kernel Virtual Machine (KVM) for virtualization platform. We also attached a iSCSI disk system to dual VM (vm1 and vm2) for backup the important configuration files. In Keio SFC NOC, two new servers were also installed as virtualization platforms: one is for SOI Asia workshops, and the other one is for AI3 and SOI Asia services.

2.2 AI^3 Layer 3 Routing

The March 11th of 2011 earthquake and tsunami brought many difficult situation in Japan. The earthquake and tsunami hit the nuclear power plants operated by Tokyo Electric power company (TEPCO). Right after the March 11, TEPCO rolled out the rolling power outages in their service areas. AI3-SFC NOC is under service of TEPCO, which are also subjected to the rolling blackouts. Against the blackouts, we have decided to change our network configuration.

2.2.1 Before 3/11

The are two AS border routers of AI^3 at Keio University Shonan-Fujisawa Campus (Keio SFC), namely `sfc-gate.ai3.net` and `sfc-gate2.ai3.net`. The AI^3 IPv4 and IPv6 address prefixes are advertised only from these two border routers. AI^3 has two BGP peer neighbours, namely with WIDE and with APAN-JP. Two AS border routers of AI^3 are peered with WIDE AS border router hosted at WIDE-FUJISAWA NOC. The AI^3 border router (`sfc-gate.ai3.net`) is peered with APAN-JP AS border router hosted at TOKYO-XP. The Layer 2 connection between AI3-SFC with TOKYO-XP is provided by WIDE with a VLAN-ID 49. Figure 2.3 shows the general connection diagram of AI^3 networks.

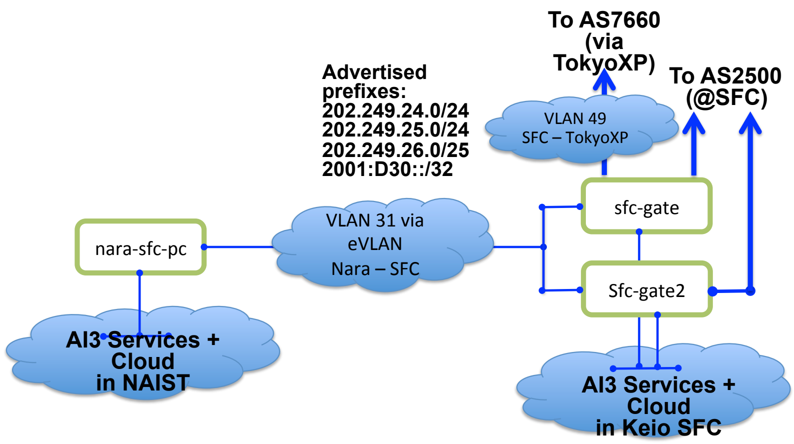

Figure 2.3: AI^3 Layer 3 Routing: Before March 11, 2011. AI^3 has only two AS border routers, where all are located at AI3-SFC . Intra-domain routing run between AI3-SFC and AI3-NAIST via Layer 2 link provided by WIDE with a VLAN ID 49.

2.2.2 After 3/11

The rolling blackouts brought difficult operation of AI^3 networks, especially AI3-SFC. When AI3-SFC was on the rolling blackout times, AI3-NAIST was also unreachable from the Internet. In response to this situation, soon afterward, we set up a temporary AI^3 border router, `nara-wide-pc.ai3.net`, at AI3-NAIST. This `nara-wide-pc.ai3.net` was peered with WIDE AS border router hosted at NARA. The Layer 2 connection between AI3-NAIST to WIDE-NARA is provided by WIDE with a VLAN ID 802. This addition of AI^3 border router at AI3-NAIST brought better network redundancy, especially in time of disaster. Figure 2.4 shows the general network diagram of the Layer 3 routing of AI^3 networks after the March 11 disaster.

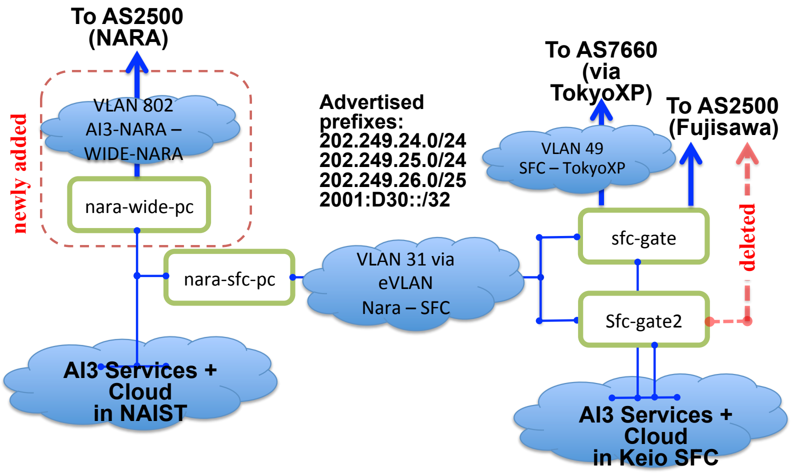

Figure 2.4: AI^3 Layer 3 Routing: After March 11^3, 2011. The BGP peer between `sfc-gate2.ai3.net` and WIDE-FUJISAWA border router was terminated to reduce the power consumption in time of rolling blackouts. The `nara-wide-pc.ai3.net` is newly added border router and linked through a VLAN ID 802 with WIDE border router at WIDE-NARA.

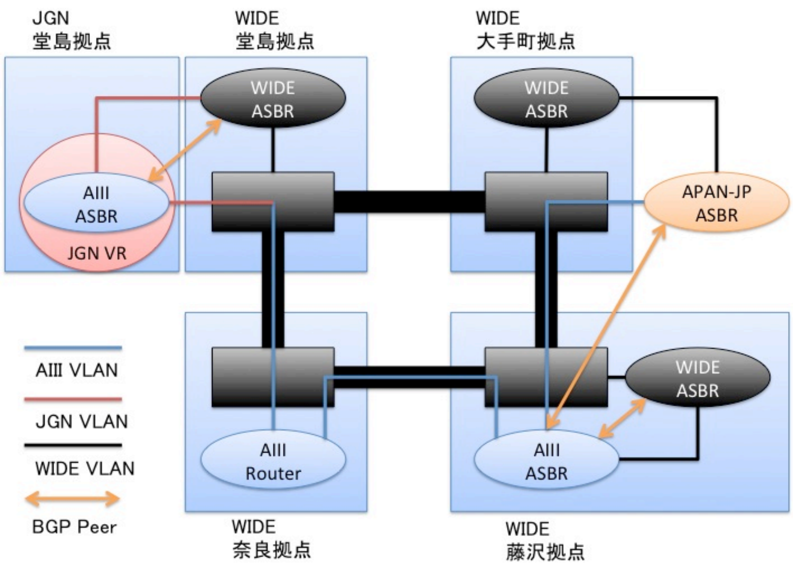

Figure 2.5: AI^3 Layer 3 Routing: A research proposal submitted to JGN-X of NICT. The objective of this proposal is to improve the network redundancy of AI^3 networks.

Considering the usefulness of network redundancy operation at AI3-NAIST, we proposed a research proposal to JGN-X of NICT to extend the network redundancy for better operation in time of possible future disaster or other future purposes. Under a research project titled "Network Redundancy for Disaster using JGN-X Virtual Router", an MoU between Keio University Shonan-Fujisawa Campus and National Institute of Information and Communications Technology (NICT) has been signed in September 2011. Figure 2.5 shows the proposed network configuration.

2.2.3 Current Status

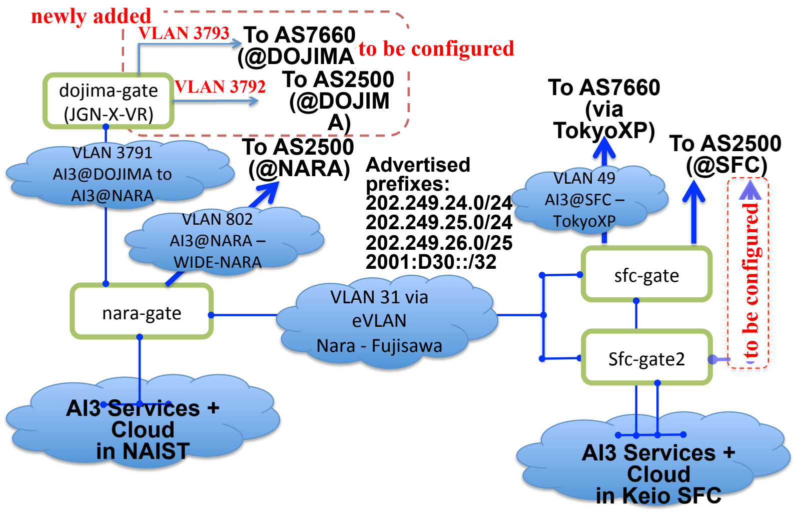

Figure 2.6: AI^3 Layer 3 Routing: Current Status.

The `nara-wide-pc.ai3.net` is a temporary border router running on top of a virtual machine (VM) of a PC server at AI3-NAIST. To improve the operation performance, we dedicated a new powerful PC as a permanent border router at AI3-NAIST to replace the VM. This new router also replaced the very old PC router `nara-sfc-pc.ai3.net` that interconnected AI3-SFC and AI3-NAIST. The new router is `nara-gate.ai3.net`. In this new router, we enabled multicast routing (it was absence in the old router of AI3-NAIST) and prepared it for the interconnection with the newly added AI^3 border router at JGN-X located in Dojima. Figure 2.6 shows the current status (as of this writing) AI^3 interconnection.

The virtual router (VR) hosted at JGN-X `dojima-gate.ai3.net` will have BGP peers with WIDE and APAN-JP that are hosted in Dojima NOC, as shown in Figure 2.6. This VR acts as a new AI^3 border router operated from Dojima to support network redundancy. JGN-X VLAN ID 3791 has been extended to WIDE-NARA up to AI3-NAIST. JGN-X has also provided two more VLAN ID 3792 and 3793 for links to other networks. These Layer 2 connection has been activated since mid of 2011. We have already activated the Layer 3 routing of `dojima-gate.ai3.net` with `nara-gate.ai3.net` since December of 2011. We are now in process of coordinating the BGP peers between AI3-DOJIMA to WIDE and APAN-JP at Dojima.

2.2.4 Expected Operation from 2012 onward

We have been experiencing more stable power supply from mid of 2011 and there are no more rolling blackouts from TEPCO so far. We try to improve the AI^3 network redundancy as well as better operation of AI^3 networks. The expected operation of AI^3 Layer 3 routing from 2012 onward would be seen as what we described in Figure 2.7. In this new configuration, we will also resume the BGP peer between `sfc-gate2.ai3.net` with WIDE-FUJISAWA.

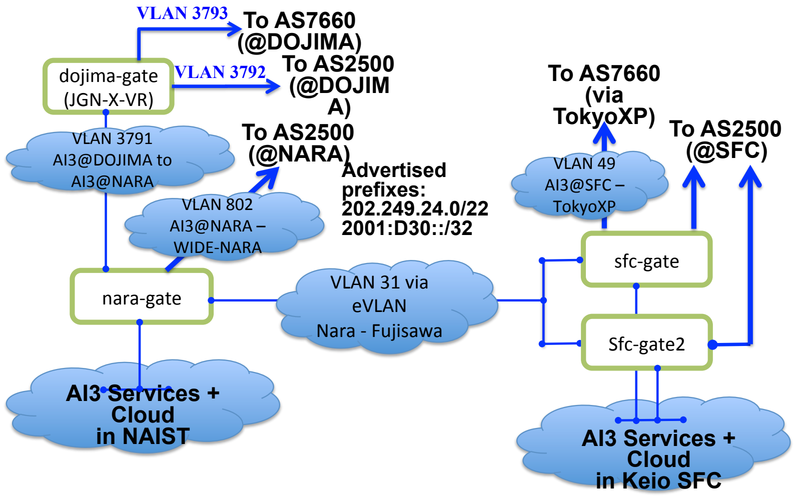

Figure 2.7: AI^3 Layer 3 Routing: Expected Operation in 2012 and onward.

The full operation of Layer 3 routing of AI^3 networks will be served by four AS border routers: 1) `sfc-gate.ai3.net`, 2) `sfc-gate2.ai3.net`, 3) `nara-gate.ai3.net`, and 4) `dojima-gate.ai3.net`. AI^3 networks will have BGP peers with WIDE at three location (at WIDE-FUJISAWA, WIDE-NARA, and WIDE-DOJIMA) and with APAN-JP at two location (TOKYO-XP and DOJIMA).

2.3 UDP Tunnelling for Link Layer Tunneling Mechanism (LLTM)

The original Unidirectional Lightweight Encapsulation (ULE) receiver uses Generic Routing Encapsulation (GRE) as the tunneling protocol to carry traffic back to ULE feed. However, using GRE imposes a severe limitation. Some ISPs block GRE traffic. This problem was reported by some of the AI3 partner sites. For these sites, a request had to be made to their ISP to allow GRE traffic to Shonan Fujisawa Campus (SFC) ULE Feed to pass through. Consequently, the IP address of ULE Feed at SFC has to be 202.249.25.98 as these partner sites have issue forwarding GRE traffic to any address other than 202.249.25.98. The same issue happened to the newest SOI Asia partner site, National University of East Timor (UNTL). For UNTL's case, GRE traffic cannot be forwarded through their ISP. As such, we need to add support for UDP tunneling in addition to supporting GRE tunneling for existing partner sites.

To enable UDP tunneling, the software at the feed and receiver site had to be modified. Figure 2.8 is the design of the software at the feed to support for both UDP and GRE tunneling. Feed requires an additional dummy network interface `tap1` to generate Dynamic Tunnel Configuration Protocol (DTCP) [5] Hello packet containing UDP tunnel information. This dummy network interface is required because existing ULE receivers identify a feed via MAC address. Using the same network

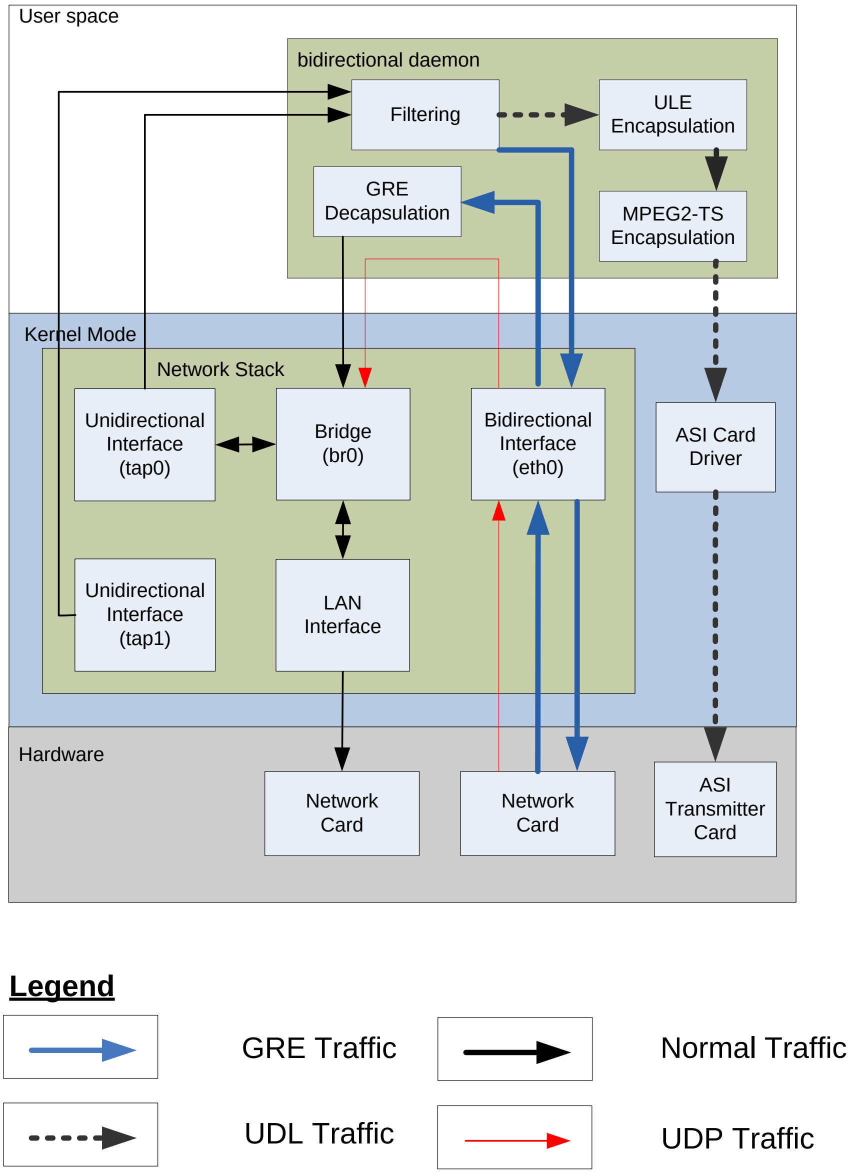

Figure 2.8: Software at the ULE Feed Supporting both UDP and GRE Tunneling

interface to send DTCP Hello packet for UDP and GRE tunnel will confuse existing ULE receivers and may lead to recurring temporary loss of connectivity when UDP tunnel information is received. If, however, a different network interface (thus different MAC) is used to generate DTCP Hello packet for UDP tunnel, existing ULE receiver will discard the information since they can only support GRE tunnel, while new ULE receiver is able to identify new UDP tunnel endpoint.

For the UDP tunnel, we adopt the VXLAN [6]. DTCP Hello doesn't support specifying additional configuration information required to make UDP tunneling work. Therefore, the operator at receiver site has to manually specify UDP port and VXLAN Network Identifier (VNI) required for VXLAN tunnel. For the decapsulation of the tunnel packet, the GRE decapsulation is done in userspace, whereas UDP is automatically decapsulated by the kernel.

To ensure that the new feed wouldn't disrupt the operation of existing ULE receivers, the development and testing was done internally using lab equipments. Once the development phase was over, a quick experiment using a new feed was carried on AI3 satellite network over the weekend to ensure minimal disruption. Existing feed was kept intact in case the experiment fails. Luckily, the experiment completed smoothly. A call was made to partner sites to try out the new ULE receiver software that supports UDP tunneling. Universiti Sains Malaysia participated in the experiment and everything worked as expected. Henceforth, the new system was kept and current AI3 satellite network can be summarized by Figure 2.9 with a mix of GRE and UDP tunneling.

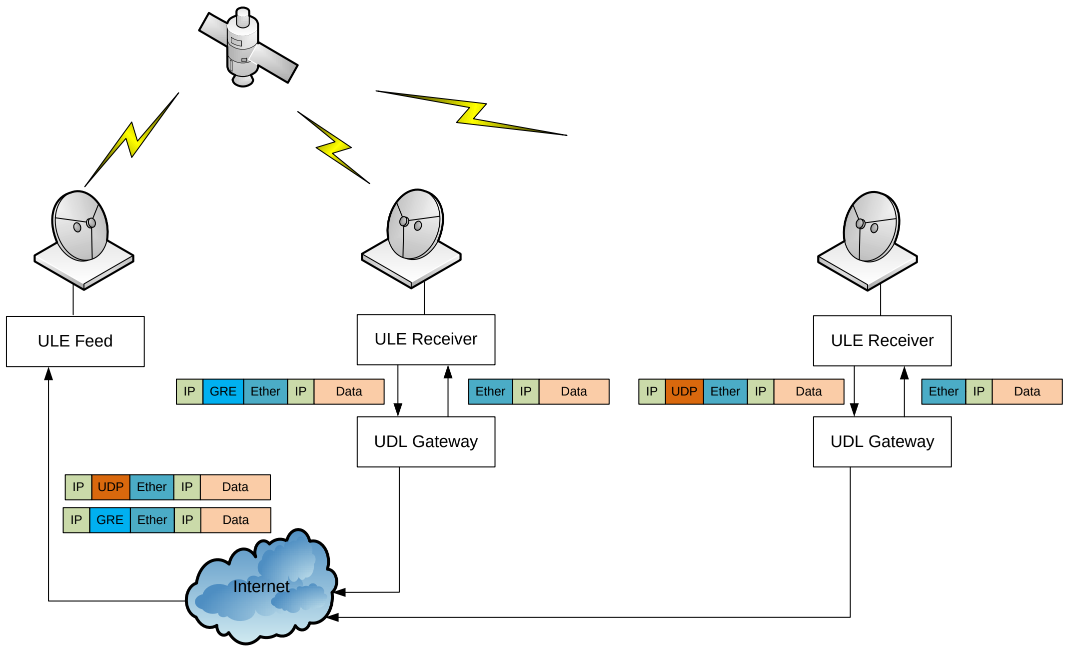

Figure 2.9: Current AI3 ULE Feed Supports for Both UDP Tunneling and GRE Tunneling

As of December 2011, UNSYIAH experienced an issue with bidirectional satellite link. UNSYIAH took this opportunity to try out new ULE receiver software and thus becoming the second partner site to UDP tunneling.

2.4 UDL Migration Experiment

Due to the Great Tohuku Earthquake, the electrical supply by TEPCO was severely limited. Power outages were expected during summer when the peak electrical demand would occur. To keep the AI3 satellite network working during the power outages, it was planned that Universiti Sains Malaysia (USM) would take over the role of transmission hub.

Due to the fact that most of the network facilities needed for AI3 network were still in Japan, the traffic received by USM transmission hub has to be tunnelled back to Japan. The plan was to assume that the essential network facilities at Shonan Fujisawa Campus (SFC) of Keio University would be available even during power outages and to have Nara Institute of Science and Technology (NAIST) acting as the main corresponding site in Japan in case the network at SFC couldn't be sustained by backup power supply. As for the UDL migration experiment, the experiment was limited to SFC and USM.

Apart from that, the ULE feed of AI3 network had to keep the IP address of 202.249.25.98 because the ISPs of some AI3 partner sites only allowed GRE traffic destined to that IP to pass through. If USM were to advertise their own public IP address as the ULE feed tunnel endpoint to partner sites, some sites might lose their connectivities to AI3 network. Since the network prefix for 202.249.25.98 belongs to AI3 network, to assign 202.249.25.98 to USM ULE feed would require that a tunnel to be created between Japan and Malaysia. For that purpose, a GRE tunnel was configured at the router at SFC (sfc-gate.ai3.net) to tunnel LLTM traffic from partner sites to the ULE feed at USM.

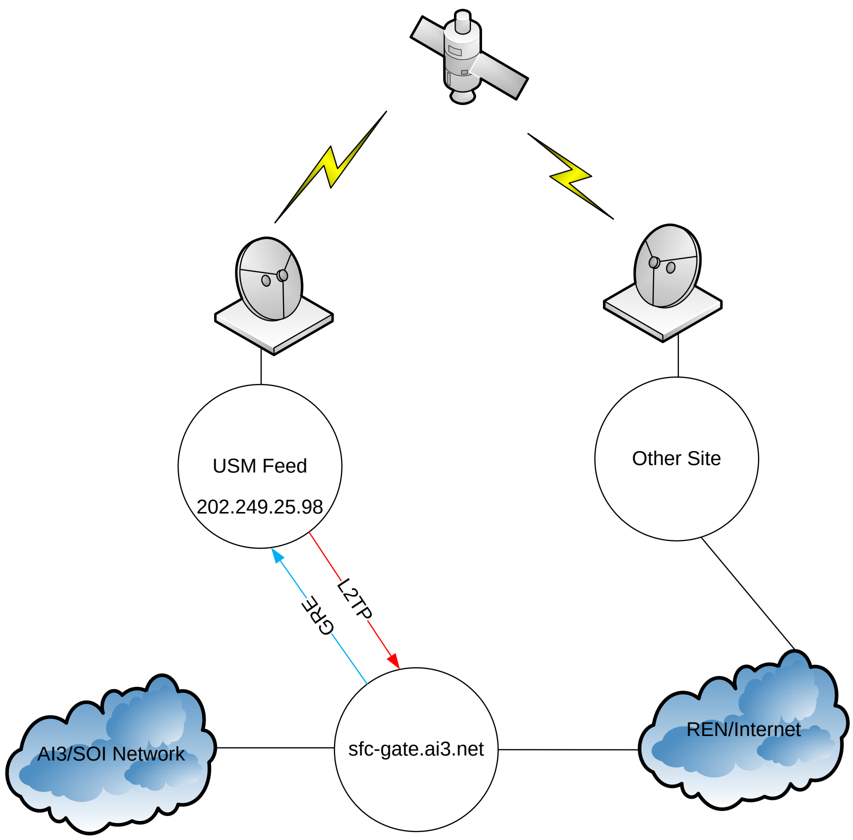

Figure 2.10: Network Topology of UDL Migration Experiment

The decapsulated LLTM traffic had to be sent back to SFC because most of the required network facility required for the operation were still with SFC. The decapsulated LLTM traffic was sent through L2TP tunnel. Figure 2.10 depicts the network topology of the UDL migration experiment. The ULE feed of USM was connected to MYREN network. The experiment was completed succesfully using this configuration. Because of the addition of tunnels, the round trip time of traffic was increased by about 120 milliseconds which was 3 times the latency between MYREN network and SFC. This additional 120 milliseconds was incurred because for every outbound traffic from partner sites, the traffic had to:

- Travel from SFC to USM to channel LLTM traffic to ULE feed.
- Travel from USM to SFC to deliver decapsulated LLTM traffic.
- Travel from SFC to USM because the reply had to be sent from ULE feed at USM.

Due to physical limitation of the equipment at USM, the transmission bitrate from USM was limited to 7168000 bps. To prepare for the eventual migration during the power outage, a script was written to automatically switch the parameters of the tuner of ULE receiver back and forth from USM and SFC. As it happened, there was no power outage during the height of summer.

2.5 Cloud Controllers Usage for AI3/SOI Asia

To take advantage of server consolidation and ease of server migration, AI3 experimented with several cloud controllers. Below is the list of cloud controllers experimented by the AI3 team:

- The earliest deployment of cloud controller was started by NAIST. NAIST used Proxmox as the cloud controller. Several virtual machines were configured and hosted at NAIST cloud controller. The initial setup of NAIST cloud controller consisted of two physical servers. However, Proxmox was unsuitable of wider deployment due to lack of fine grain access control.

- The subsequent effort turned to OpenNebula. The first OpenNebula tried was version 2.2 released on March 2011. While OpenNebula provided great flexibility in control and well documented, however, most of the work must be performed using command line interface as the web administration interface was not complete. Subsequent release of OpenNebula version 3.0 provided major improvement to web administration interface (OpenNebula Sunstone) as shown in Figure 2.11. However, there was a bug that could cause VNC connection to a VM to fail sometimes and the advanced features of cloud controller had to be configured manually by typing the configuration option into web administration interface.

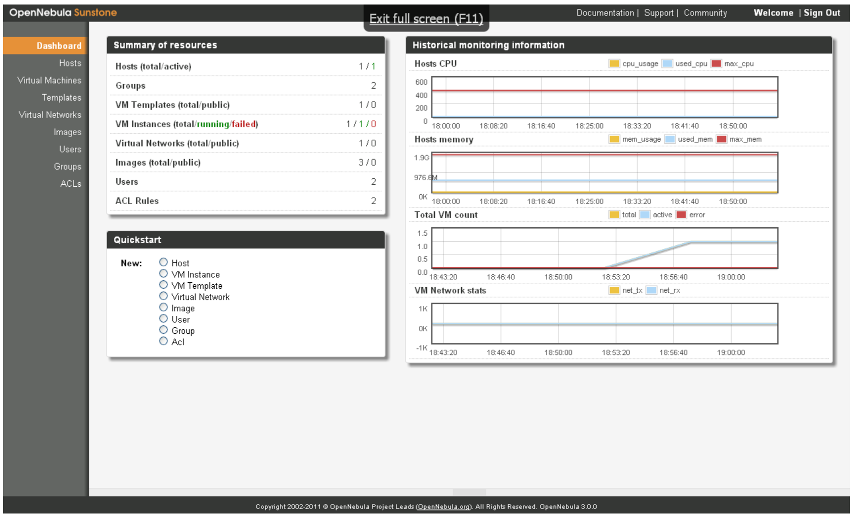

Figure 2.11: OpenNebula Web Administration Interface (Sunstone)

- Another cloud controller that was tried was the OpenStack project. However, the project was too complicated for our purpose and the project was going through fast pace of change and the documentation did not keep up with code change. Needless to say, OpenStack was not used after initial trial and evaluation

- The last cloud controller that was tried was WIDE Cloud Controller produced by members of the WIDE project. While WCC does not provide all kind of features supported by the OpenStack project, it has sufficient features for SOI Asia deployment. WIDE project kindly permitted the deployment of WCC for SOI Asia. One minor customization made to WCC was to support for multiple network interfaces on a VM. This feature was required by the SOI Asia project. Currently, two physical servers are deployed for the SOI Asia cloud. https://cloud.soi.asia acts as both the cloud controller and a hypervisor, while the other hypervisor is hosted on vm2-sfc.ai3.net. The network setup of the WCC at SOI Asia is shown in Figure 2.12. Both hypervisors has `nat0` bridge interface that is connected through VLAN. Live migration is limited to those VMs that use only `nat0` bridge network.

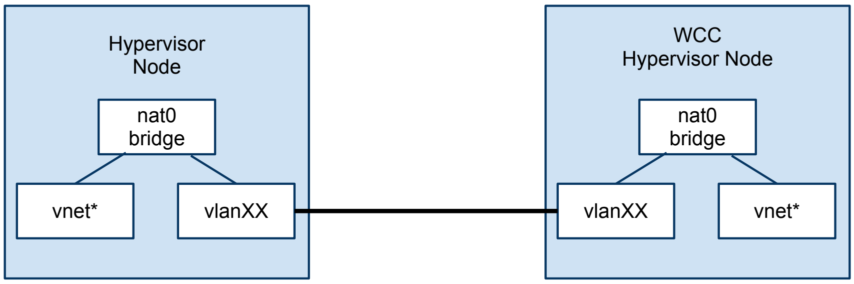

Figure 2.12: WIDE Cloud Controller (WCC) Network Configuration for SOI Asia/AI3

As of this moment, there are only a handful of servers that were migrated to use WCC. A mailing list, cloud@lists.soi.asia, was setup for AI3/SOI Asia partners who are interested to use the cloud controller.

Chapter 3

Meetings

We regularly held meeting twice in a year with partners. The meetings provided opportunities to share operational information and to discuss research topics in future. In this year, we succeeded to held the meeting in Yangon, Myanmar. Myanmar had closed their country to foreign country in long period. However, due to their development of their country and huge pressure from foreign countries, they have intended to open their country. We strongly believe our collaborative activities in past years may have contributed to changing their national strategy.

- Bangkok meeting

  Bangkok meeting was held at the Chulalongkorn University, in Thailand on May 10–12, 2011. In the meeting, we share the information about an influence of great Tohoku earthquake in Japan to AI3 network. We had tried to construct our new network environment with virtualized technology. We discuss the new network architecture. More detailed information is published at the meeting web site [7].

- Yangon meeting

  Yangon meeting was held at the University of Computer Studies, Yangon in Myanmar on Nov 8–10, 2011. We provided a lecture for DNSSEC operation which have been already available in AI3 network. Then we share the information about network operation. Some operators report a status of a trial for network virtualization and call for a request to use the facilities.

  More detailed information is published at the meeting web site [8].

Bibliography

[1] AI3: Asia Internet Interconnection Initiatives. http://www.ai3.net/.

[2] S. Yamaguchi and J. Murai. Asian Internet Interconnection Initiatives. *INET 1996*, Jun. 1996.

[3] : T. Yokoyama, S. Mikawa, J. Takei, K. Cho, S. Yamaguchi, and Jun Murai. Overview of AI3 Network: Design and Applications of Satellite Network. *Sigcomm '07 NSDR Workshop*, Aug. 2007

[4] : E. Duros, W. Dabbous, H. Izumiyama, N. Fjuii, and Y. Zhang. A link layer tunneling mechanism for unidirectional links. *RFC 3077*, Mar. 2001.

[5] A Link-Layer Tunneling Mechanism for Unidirectional Links. http://tools.ietf.org/html/rfc3077.

[6] VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks. http://tools.ietf.org/html/draft-mahalingam-dutt-dcops-vxlan-00.

[7] AI3 Meeting at Bangkok. http://www.eng.chula.ac.th/?q=en/node/2666.

[8] AI3 Meeting at Yangon. http://www.ucsy.edu.mm/ucsy/soi.do.