Interconnecting International Network Substrates for Networking Experiments

Namgon Kim[1], JongWon Kim[1], Chris Heermann[2], and Ilia Baldine[2]

[1] Networked Media Lab., School of Information and Communications, Gwangju Institute of Science and Technology (GIST), Gwangju, 500-712, Korea
{ngkim,jongwon}@nm.gist.ac.kr
[2] Renaissance Computing Institute (RENCI)
{ckh,ibaldine}@renci.org

Abstract. Large research programs have been launched to address both the development of Future Internet architectures as well as suitable experimental facilities for testing. Recent research activities on experimental facilities try to share resources across organizational boundaries. This paper introduces an international cooperation effort on interconnecting network substrates of two Future Internet testbed projects, FIRST@PC (Future Internet Research on Sustainable Testbed based on PC) in Korea and ORCA-BEN in United States. To build a collaborative research infrastructure available to each other, we first present how to interconnect two network substrates. We then present how we support experiments on the interconnected network substrate and show the demonstration result performed on it.

Keywords: Future Internet, Networking testbed, Interconnection, Federation, FIRST@PC, GENI.

1 Introduction

Research and experimentation on novel network technologies and architectures require new experimental testbeds that combine flexibility, neutrality, a minimum set of constraints for the researchers, reproducibility, and full control of the testing environment. Large research programs have been launched to address both the development of Future Internet architectures as well as suitable experimental facilities for testing [1][2][3][4]. In the testbeds, a control framework controls and manages the resources of underlying substrates. A slice is defined by a set of slivers spanning a set of computing/networking components, plus an associated set of users to access those slivers for the purpose of running an experiment. That is, the control framework supports experimenters to run their experiments on the testbed by giving slices composed of sets of resources of the testbed.

Recently, research activities try to share (i.e. federate) the resources of experimental facilities across organizational boundaries. Federation enables combining infrastructural computing/networking resources and associated services of more

T. Korakis et al. (Eds.): TridentCom 2011, LNICST 90, pp. 116–125, 2012.

than one independently controlled domain in order to enhance the utility of testbeds significantly [7]. For example, GENI (Global Environment for Network Innovations) SFA (Slice Based Facility Architecture) [5] defines a minimal set of interfaces and data types to allow a federation of slice-based substrates to interoperate. The federations among facilities like PlanetLab, Emulab, and OpenFlow have been realized based on it. Network stitching is an associated topic of federation. Typically, it provisions static layer-2 network connections among resource substrates. To make it dynamic, a stitching service is being discussed to visualize global interconnection topology among resources and make requested interconnection topology [6].

In this paper, we introduce our experience on interconnecting network substrates of FIRST@PC (Future Internet Research on Sustainable Testbed based on PC) project in Korea and ORCA-BEN (Open Resource Control Architecture-Breakable Experimental Network) project in the United States. We first interconnect the network substrates of each project. FIRST@PC resources are interconnected by the programmable networking substrate called NetOpen RA (Resource Aggregate). ORCA-BEN resources are interconnected by the network substrate called BEN (Breakable Experimental Network) [10][11]. We provide layer-2 network connectivity between NetOpen RA and BEN by configuring VLANs between the two. Then we perform networking experiments to verify the utilization of shared resources.

2 Network Substrates Interconnection

Generally, a testbed is composed of a number of substrates for computing and networking. There is at least one network substrate that provides network connectivity among other participating substrates. Figure 1 shows an example testbed configuration connecting the substrates of FIRST@PC and ORCA-BEN projects.

The FIRST@PC project is a Future Internet testbed project exploring the possibility of building innovative media-centric Internet applications by composing component services over programmable network substrates. It is composed of two types of RAs, MediaX and NetOpen. The MediaX RA is an aggregation of computing resources focused on producing/processing/consuming media.

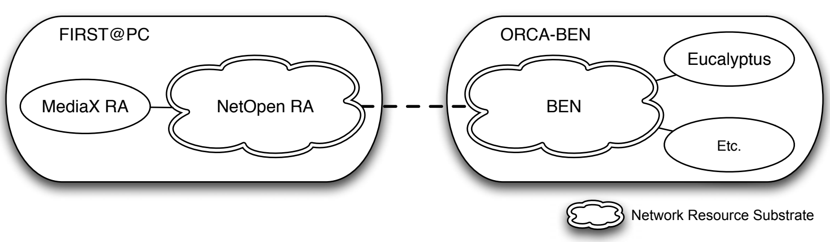


Fig. 1. FIRST@PC and ORCA-BEN

Previous work by the authors in [13] constructed a preliminary prototype of FIRST@PC MediaX RAs, which includes an early version of experiment control tools and PC-based devices for media capturing, processing, delivering, and display. The NetOpen RA is a set of OpenFlow-enabled programmable networking resources connecting MediaX RAs. NetOpen RA is composed of NetOpen nodes supporting flow-based network programmability based on OpenFlow [14]. With NetOpen RA, experimenters can determine how to handle their flows in network.

The ORCA-BEN project is a RENCI-Duke collaboration to build a testbed for GENI (Global Environment for Network Innovation) initiative launched by NSF (National Science Foundation) [12]. ORCA-BEN focuses on building an unified control framework to prototype GENI facilities incorporating RENCI's metro-scale network substrates called BEN with other available substrates. BEN is a multi-layer optical network that supports both IP overlays and dynamic circuit provisioning. It is designed for coarse-grained time-sharing between experiments, enabled by automated fiber switches located at each site. ORCA-BEN also supports other substrates, including Eucalyptus cloud sites and NLR's Sherpa FrameNet service.

As the two projects have partially different research objectives, their testbeds are comprised of different type of substrates as well. For media-centric service composition over programmable network substrates, FIRST@PC provides media-centric resources and OpenFlow-based programmable network substrate. On the other hand, ORCA-BEN aims at building an unified control framework utilizing BEN and has various computing resources connected to BEN. Therefore, by interconnecting the two testbeds and providing the chance of utilizing each other's substrates, FIRST@PC can secure the computing resources of ORCA-BEN and ORCA-BEN can include programmable network substrate as its new substrate.

2.1 Connecting Network Substrates

By providing network connectivity with NetOpen RA and BEN, the two testbeds can be interconnected with each other. The connection between NetOpen RA and BEN should meet the following requirements. First, it should keep layer-2 network connectivity. In Fig. 2, the packet arrived at 'end02' should be the same as the packet sent by 'end01' even it goes through the interconnection. Second, the features of each network substrate should be kept. For the flows between 'end01' and 'end02', it should utilize the flow-based network programmability of NetOpen RA as well as the dynamic circuit provisioning features of BEN. Finally, it should be connected without losing networking performance. That is, the experiments performed on each network substrate should be possible on the interconnected substrate.

Three methods can be considered to provide layer-2 network connectivity. We can use 1) software-based or 2) hardware-accelerated EoIP (Ethernet-over-IP) tunneling solutions to create EoIP tunnel connecting NetOpen RA and BEN. Also we can setup 3) VLANs along the path between NetOpen RA and BEN. The software-based EoIP tunneling solution is the easiest one to use, but we cannot

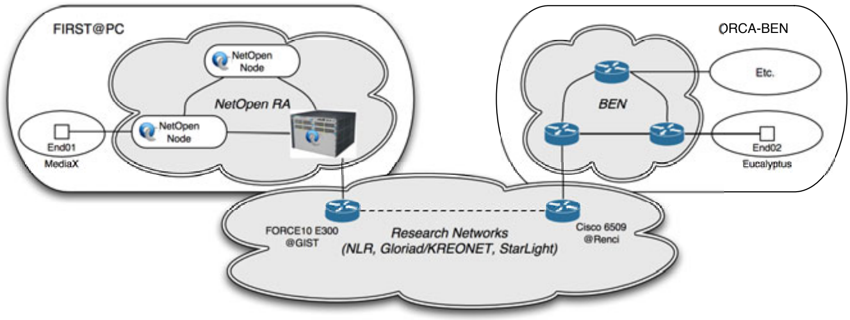


Fig. 2. Interconnecting network substrates

achieve good networking performance. On the other hand, with the hardware-accelerated EoIP tunneling solution, we can get better networking performance. However, both parties should be equipped with the EoIP tunneling solution and it could cost much. Although there is an open source implementation on NetFPGA, its stability has not yet been confirmed. Thus configuring VLAN is the best solution in stability and performance perspective. However, it takes time to configure as it needs to be done manually by the network operators managing the networks between NetOpen RA and BEN.

After all, in this paper, we use both the software-based EoIP solution and the VLAN-based methods. We start with the software-based EoIP solution to check the possibility of this interconnection. At the same time, we requested the network operators of research networks in between NetOpen RA and BEN to configure VLANs. After several months of cooperation, we could successfully have VLAN-based layer-2 network connectivity. Later in Section 4, we show the interconnection result.

3 Support Experiments on the Interconnected Substrates

To perform experiments on the interconnected substrates, substrates should be managed by the control framework of both FIRST@PC and ORCA-BEN.

3.1 Control Frameworks

The control framework manages distributed resources over federated substrates and provides a set of resources to slices, permitted to perform an experiment. For a substrate to be simultaneously used for multiple independent experiments, the control framework schedules and configures computing/networking resources to enable the end-to-end slicing. As shown in Fig. 3, both FIRST@PC and ORCA-BEN provide their own control framework to support flexible configuration and resource isolation in the testbed following the design principles of GENI control framework [1] .

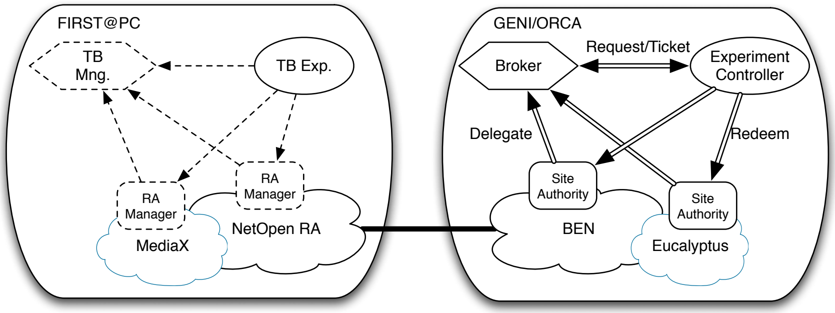


Fig. 3. Control frameworks of FIRST@PC and ORCA-BEN

ORCA-BEN supports experiments by instantiating a network slice hosting an experiment. With the network slice, experimenters manage the network substrate as a resource that can be co-scheduled and co-allocated along with compute and storage resources through *ORCA framework* [8][9] designed at Duke University (Jeff Chase). It has the following structure. Substrate providers are represented by *site authorities*, which delegate their resources to one or more *brokers*. Brokers are containers for resource allocation policies. Experimenter requests for various resources to brokers are satisfied in the form of tickets, which are then presented by the users to different sites in redeem operation. Sites then allocate and configure resources described in tickets. Experimenters can also perform additional configuration actions on the issued slices to further customize the substrates to their needs.

Similar to the ORCA framework, FIRST@PC has the following structure. Substrates are represented as resource aggregates (RAs) and the access and control to RAs are managed by corresponding *RA managers*. A *TB (testbed) management server* is responsible for operating and managing the testbed by involving with slice control, resource management, and resource monitoring. Resource provisioning for an experiment is conducted by the cooperation of RA managers and TB management server. Experimenters can make use of authorized resources of RAs and run media-centric service composition experiments via a *TB experiment control server*. Note that the architecture for FIRST@PC has not completely realized yet and the TB experiment control server takes most of the responsibilities in performing experiments.

3.2 NetOpen RA Supporting the ORCA Framework

For a substrate to be controlled by the ORCA framework, it requires a site authority to manage the substrate. To implement a site authority for NetOpen RA, we should determine the unit resource to be delegated. For example, BEN uses a VLAN tag and Eucalyptus cluster uses a VM (Virtual Machine) for the unit resource. As NetOpen RA is a set of OpenFlow-enabled devices, we should

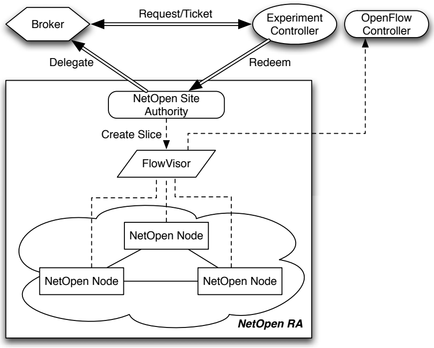


Fig. 4. NetOpen RA supporting the ORCA framework

consider the slice concept of OpenFlow. The OpenFlow switch itself does not support virtualization, but instead it supports the slice concept by FlowVisor [15]. FlowVisor supports flowspace-based virtualization by creating a slice to compose the list of utilized switches. Each slice is connected to an OpenFlow controller and the network works based on the policy implemented in the OpenFlow controller. We simplify the concept of a slice in FlowVisor and use VLAN tag as its unit resource like the BEN site authority. The FlowSpace of a slice is defined by the VLAN tag and all switches are allocated to it. Figure 4 shows the NetOpen site authority interacting with other ORCA entities.

3.3 Eucalyptus Cluster Supporting FIRST@PC Platform Architecture

Until now, the TB experiment control server performs most of operations to support media-centric service composition experiments. It provides a series of operations for the service composition, such as service discovery, matchmaking, placement, stitching, monitoring, and tuning. Based on an experiment description from an experimenter UI, the TB experiment control server communicates with RA managers to perform operations in the experiment description. RA managers support experiments by communicating with *node managers* located in nodes composing the substrate. We apply the RA manager and the node manager to Eucalyptus clusters connected to BEN. A Eucalyptus cluster is configured of a Cloud Controller (CLC), a Cluster Controller (CC), a Node Controller (NC), and a Walrus Storage Controller (WS3). The NC runs on each node and controls the life cycle of instances running on the node. The WS3 contains the machine images and stores snapshots. The CC manages one or more NCs and deploys/manages instances on them. The CLC is the front end to the entire cloud infrastructure. It requests the CC to deploy a VM, then the CC lets the NC to instantiate a VM with specified image in the WS3. Virtual machine images in the WS3 can be considered as services in media-centric service composition.

Operations for the service composition procedure can be realized in Eucalyptus cluster as the processes of deploying VM instances. By applying it to FIRST@PC platform architecture, the CLC and the CC perform the role of the RA manager and the NC works as the node manager.

4 Demonstrations

We had two demonstrations at GEC (GENI Engineering Conference) with the interconnected substrate. At GEC8 on 2010 July, we showed an experiment performed on a slice created by the ORCA framework. At GEC9 on 2010 November, we enhanced the network connection with VLAN and showed the possibility of media service composition experiments on the interconnection.

4.1 Interconnection Using Software-Based EoIP Tunnel

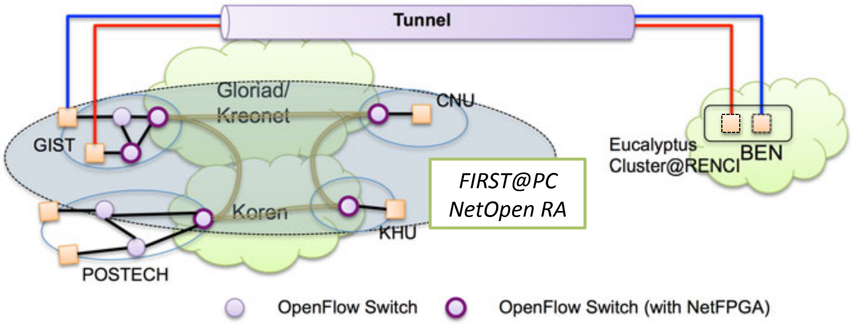
At GEC8, we demonstrated a networking experiment performed on a slice created by the ORCA framework. As it was the first demonstration, we focused on showing the substrate interconnection between FIRST@PC and ORCA-BEN and verifying the interconnection.

Figure 5(a) shows the network configuration between FIRST@PC and ORCA-BEN for the demonstration. VMs in Eucalyptus cluster at RENCI had connections with NetOpen RA through Layer2 tunnels created using OpenVPN ethernet bridge. To verify the interconnection, we showed video streaming between two VMs in different VLANs. We instantiated two VMs in Eucalyptus cluster and assigned VLAN 10 for one VM and VLAN 20 for the other so that they could not have direct connection. NetOpen RA performed VLAN translations for packets from VMs and we could show that each VM can receive the video stream from the other.

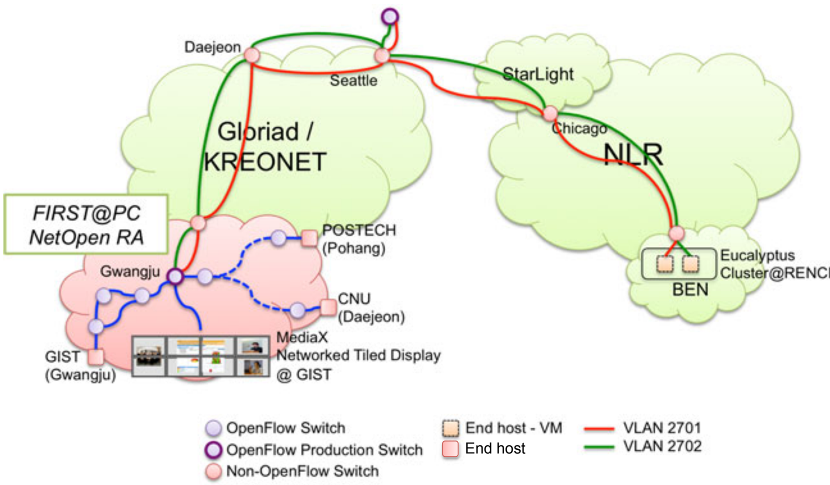
4.2 VLAN-Based Interconnection

At GEC9, we showed the possibility of running both the FIRST@PC experiment and the ORCA-BEN experiment on the interconnection.

Figure 5(b) shows the network configuration for the demonstration. First, we interconnected substrates by configuring layer-2 VLANs between NetOpen RA and BEN. This VLAN configuration was supported by the research networks, Gloriad / KREONET, StarLight and NLR and we could get lower latency and higher throughput than GEC8 demonstration. Next, MediaX RA, such as NeTD (Networked Tiled Display) supporting network-based scalable display, DXT-compression based HD video transmission system, was connected to NetOpen RA for media-centric service composition experiment. Finally, we updated the configuration of NetOpen RA itself. At GEC8, NetOpen RA was configured only with PC-based OpenFlow switches. At GEC9, we added a commercial OpenFlow switch, HP Procurve 5406zl, to NetOpen RA. With the commercial OpenFlow



(a)

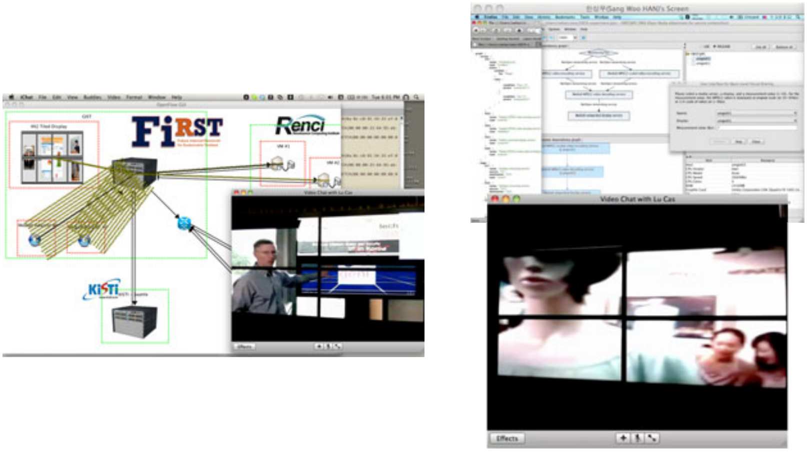


(b)

Fig. 5. Network configurations using: (a) Software-based EoIP tunnel, (b) VLAN

switch, we could support stable network connection for MediaX RA. The OpenFlow protocol version supported by NetOpen RA was also updated to ver.1.0.

The demonstration can be described into two parts. First part is the extension of GEC8 demonstration including MediaX RA. We instantiated two VMs with different VLAN in Eucalyptus cluster using the ORCA framework and each VM transmitted a video stream to the NeTD in MediaX RA and the NeTD visualized the streams. For this demonstration, we used three VLANs, two for each VM and remaining one for MediaX RA. VLAN tags in packets sent by each VM were translated into the VLAN tag of MediaX RA while going through the NetOpen RA. Figure 6 shows screen shots taken at GEC9 demonstration. Figure 6(a) shows flows from two VMs in Eucalyptus cluster go through the HP Procurve switch in NetOpen RA and delivered to the NeTD in MediaX RA. On the right side of the figure, we can see the screen of NeTD showing videos from the VM.



(a) (b)

Fig. 6. Screen shots of VLAN-based interconnection demo.: (a) NetOpen RA supporting ORCA, (b) Media-centric service composition.

The second part of the demonstration contained media-centric service composition experiments. We showed a media-centric service composition experiment showing two DXT-compressed HD video over the NeTD in MediaX RA. DXT-compressed HD video senders and the NeTD are expressed as component media services. And NetOpen RA also described as networking services providing network connectivity among component media services. Media services in MediaX RA are composed and delivered through the networking service in NetOpen RA. Fig. 6(b) shows the result of service composition experiment.

5 Conclusion

We presented an international work for interconnecting substrates in two Future Internet testbed projects. With the support for each other's control framework, we verified the interconnection by performing experiments in two international demonstrations. In the current interconnection, a control framework can control a resource substrate explicitly supporting it. We will work on the federation of substrates by accommodating GENI AM API which provides general interfaces for a substrate to be used by any control frameworks.

Acknowledgments. This work has been supported by the project (2009-F-050-01), "Development of the core technology and virtualized programmable platform for Future Internet" sponsored by MKE and KCC. We would like to appreciate Dongkyun Kim at Gloriad/KREONET; and Joe Mambretti and Jim Chen at iCAIR. We also thank our collegues Sungho Shin, Namgon Lucas Kim and Sang Woo Han of Networked Media Lab. at GIST; and Anirban Manda, Shu Huang, and Yufeng Xin at RENCI for interesting comments, encouragements, and feedbacks.

References

1. Global Environment for Network Innovation, http://www.geni.net
2. Future Internet Research and Experimentation, http://cordis.europa.eu/fp7/ict/fire/
3. AKARI Architecture Design Project, http://akari-project.nict.go.jp/eng/index2.htm
4. China next generation network, http://www.cernet2.edu.cn/en/bg.htm
5. Peterson, L., et al.: Slice Based Facility Architecture. Princeton (2007)
6. Lehman, T., Yang, X.: Network stitching. In: Proc. of the GEC9 (November 2010)
7. Wahle, S., Magedanz, T., Campowsky, K.: Interoperability in Heterogeneous Resource Federations. In: Magedanz, T., Gavras, A., Thanh, N.H., Chase, J.S. (eds.) TridentCom 2010. LNICST, vol. 46, pp. 35–50. Springer, Heidelberg (2011)
8. Chase, J., et al.: Beyond virtual data centers: Toward an open resource control architecture. In: Proc. of the International Conference on the Virtual Computing Initiative (May 2007)
9. Yumerefendi, A., et al.: Towards an autonomic computing testbed. In: Proc. of the Workshop on Hot Topics in Autonomic Computing (HotAC) (June 2007)
10. BEN: Breakable Experimental Network, https://ben.renci.org/
11. Baldine, I.: Unique optical networking facilities and cross-layer networking. In: Proc. of the IEEE LEOS Summer Topicals Future Global Networks Workshop (July 2009)
12. Baldine, I., et al.: The missing link: Putting the network in networked cloud computing. In: Proc. of the Third International Conference on the Virtual Computing Initiative, ICVCI 2009 (October 2009)
13. Han, S.W., Baek, B.H., Kim, J.: FIRST@PC mediaX: A Service-Oriented Testbed for Realistic Media Networking Experiments. In: Magedanz, T., Gavras, A., Thanh, N.H., Chase, J.S. (eds.) TridentCom 2010. LNICST, vol. 46, pp. 176–185. Springer, Heidelberg (2011)
14. McKeown, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., Shenker, S., Turner, J.: OpenFlow: Enabling innovation in campus networks. ACM SIGCOMM CCR 38(2) (March 2008)
15. Sherwood, R., Gibb, G., Kobayashi, M.: Carving research slices out of your production networks with OpenFlow. In: Proc. of the ACM SIGCOMM Demo. (August 2009)