

# Virtualization for Future Internet

2009.08.26

Korea University

Chuck Yoo

(hxy@os.korea.ac.kr)

# Why Virtualization

- Internet today
  - Pro and con
- Your “wonderful” research results
  - Mostly with simulation
  - Deployment impossible
- Virtualization
  - Aim to set you free
  - Thus, center piece to future internet

# Life before Virtualization

- Testbeds
  - Production: Internet2
  - Research: DETER
    - Mainly incremental changes
- Overlay
  - PlanetLab
    - Centered around specific problem to solve
    - No interaction between overlays
  - NO major architectural innovation

# Contents

- Virtualization introduction
- Related work in Virtualization
  - Server Virtualization
  - Embedded System Virtualization
  - Network Virtualization
- Future Virtualization for Future Internet
  - OS Virtualization
  - Trusted Network
  - E2E Virtualized Transport
  - Adaptive Video Streaming
- Summary & Reference

# What is Virtualization?

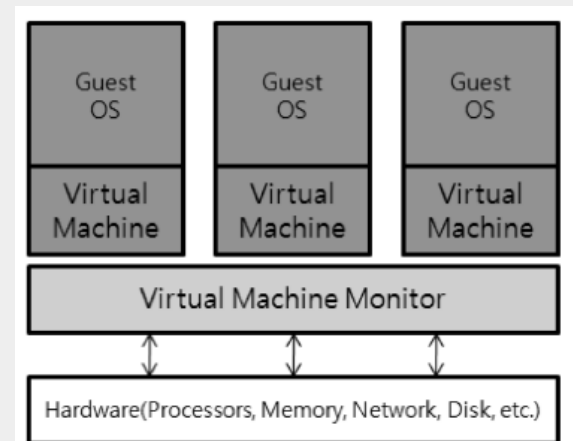
- General meaning
  - An abstraction of resources that provides a logical rather than an actual physical incarnation of those resources
  - Such resources are: CPU, storage, network, etc.
- Introduced to fully utilize mainframes in 1960s
  - Could run several projects on a single mainframe
- Has been extended to several areas
  - System Virtualization → Consumer Electronics
  - Storage Virtualization → Cloud Computing
  - Network Virtualization → Future Internet

# Virtualization Features

- Isolation
  - Provide reliability and fault-tolerance by fully separating virtual systems
- Decoupling
  - Provide portability and mobility by minimizing dependency between software and hardware
- Consolidation
  - Reduce hardware costs by consolidating multiple similar machines into one machine

# Virtualization Architecture

- Virtual Machine (VM)
  - Self-contained software execution environment
  - Guest OS runs over a virtual machine
- Virtual Machine Monitor (VMM)
  - Supervises virtual machines
  - Interacts with physical hardware



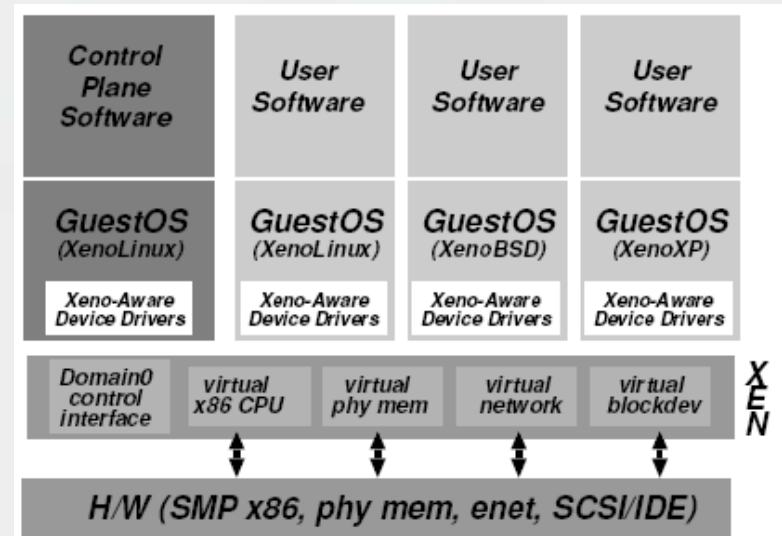
# Related work in Virtualization

- Server Virtualization
- Embedded System Virtualization
- Network Virtualization

# Server Virtualization

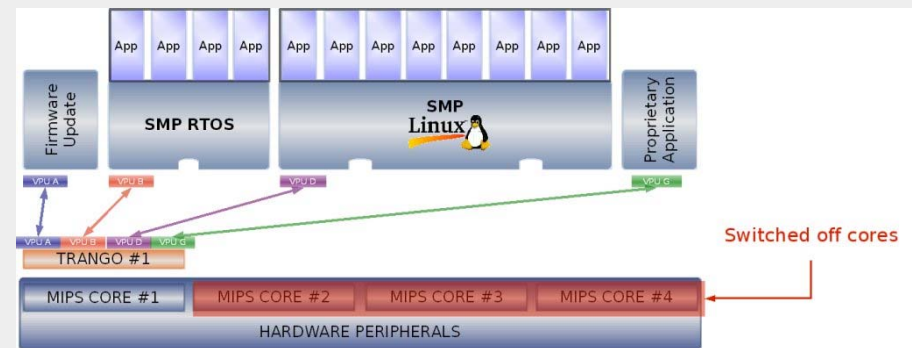
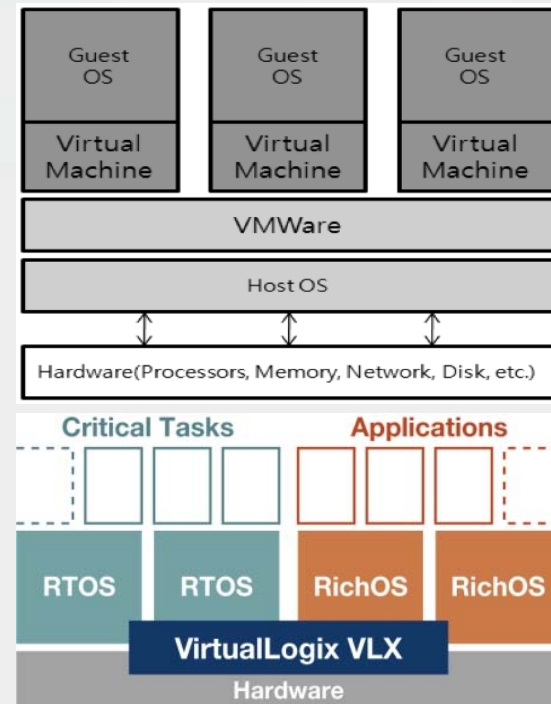
- Xen – Hypervisor

- University of Cambridge
- Para-virtualization: modify partially guest OS to improve Xen performance
- Split-driver model: device drivers are at a special domain (Domain0) to improve system reliability



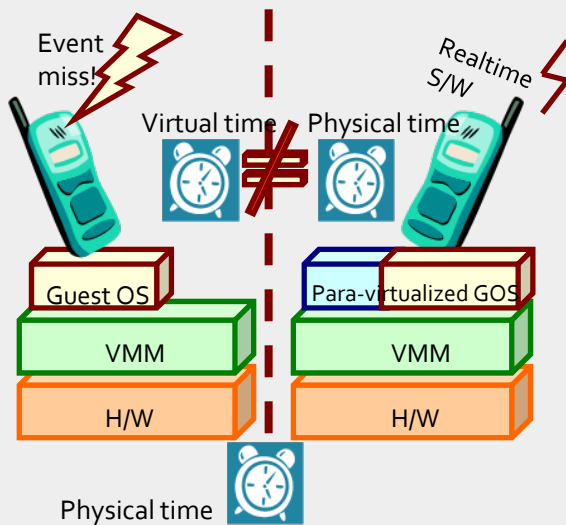
# Server Virtualization Products

- Commercial products
  - VMware
    - Operates on a host OS
  - VirtualLogix
    - Real-time support virtualization platform
  - Trango
    - Multi-core virtualization platform
    - Supports SMP, ARM MPCore processor



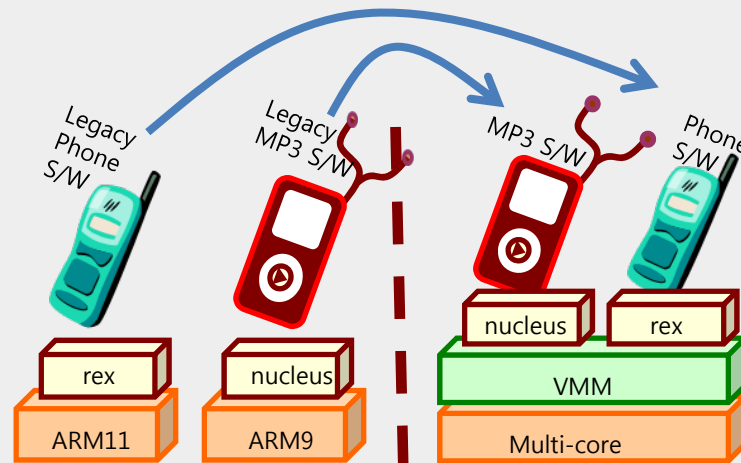
# Embedded System Virtualization

- Currently hot research topic in virtualization
  - New requirements: 1) real-time support, 2) multi-core support, 3) power efficiency, etc.

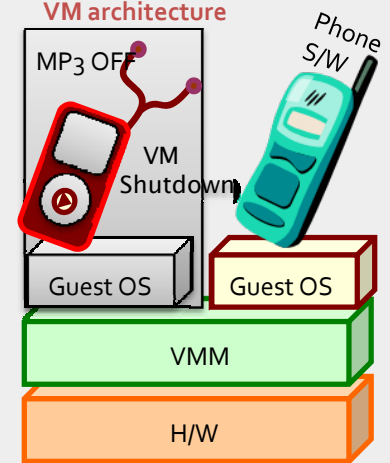


\*GOS=Guest OS

## Identical Programming Model



## Energy Efficient VM architecture

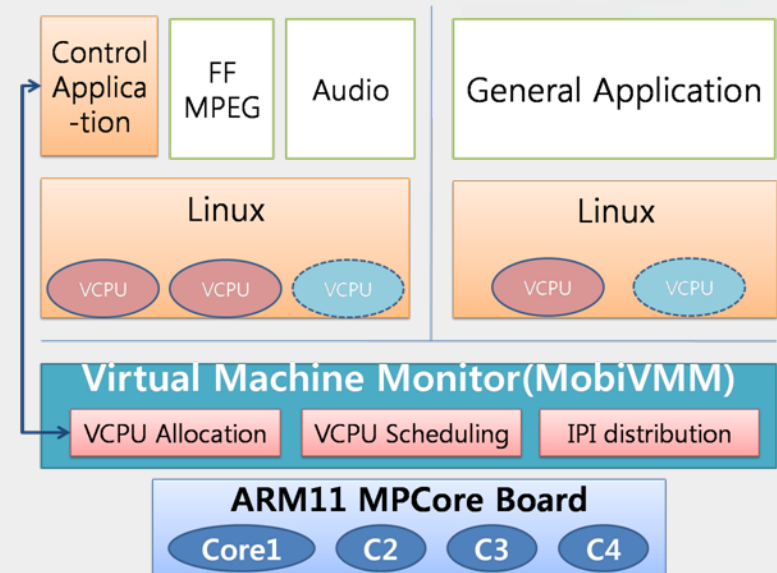


# Embedded System Virtualization Topics

- Xen on ARM
  - Samsung Electronics S/W Lab.
  - Embedded S/W Platform Virtualization
- Real-time Xen
  - Supports real-time for embedded mobile devices based on Xen on ARM
  - Provides real-time scheduling between guest OSes
  - Supports user RTOS
  - Considers power efficiency

# Embedded System Virtualization Example

- MobiVMM (MobiVirt'08)
  - New VMM for mobile phones
  - Supports mobile multi-core processors such as ARM11 MPCore
  - Provides multi-core load balancing by modified credit scheduler



# Embedded System Virtualization

## Korea University status

- MobiVMM
  - Run two guest OSes with MobiVMM over multi-core environments - Demo is available at <http://www.youtube.com/watch?v=SABYIon9vOg>
  - Feb. 2009 Xen Summit presentation
    - [http://www.xen.org/xensummit/xensummit\\_winter\\_2009.html](http://www.xen.org/xensummit/xensummit_winter_2009.html)
- Real-time
  - Para-virtualized uC/OS-II (RTOS) for Xen on ARM contributed to Xen community
    - Source code is available at <http://wiki.xensource.com/xenwiki/XenARM>

# Network Virtualization

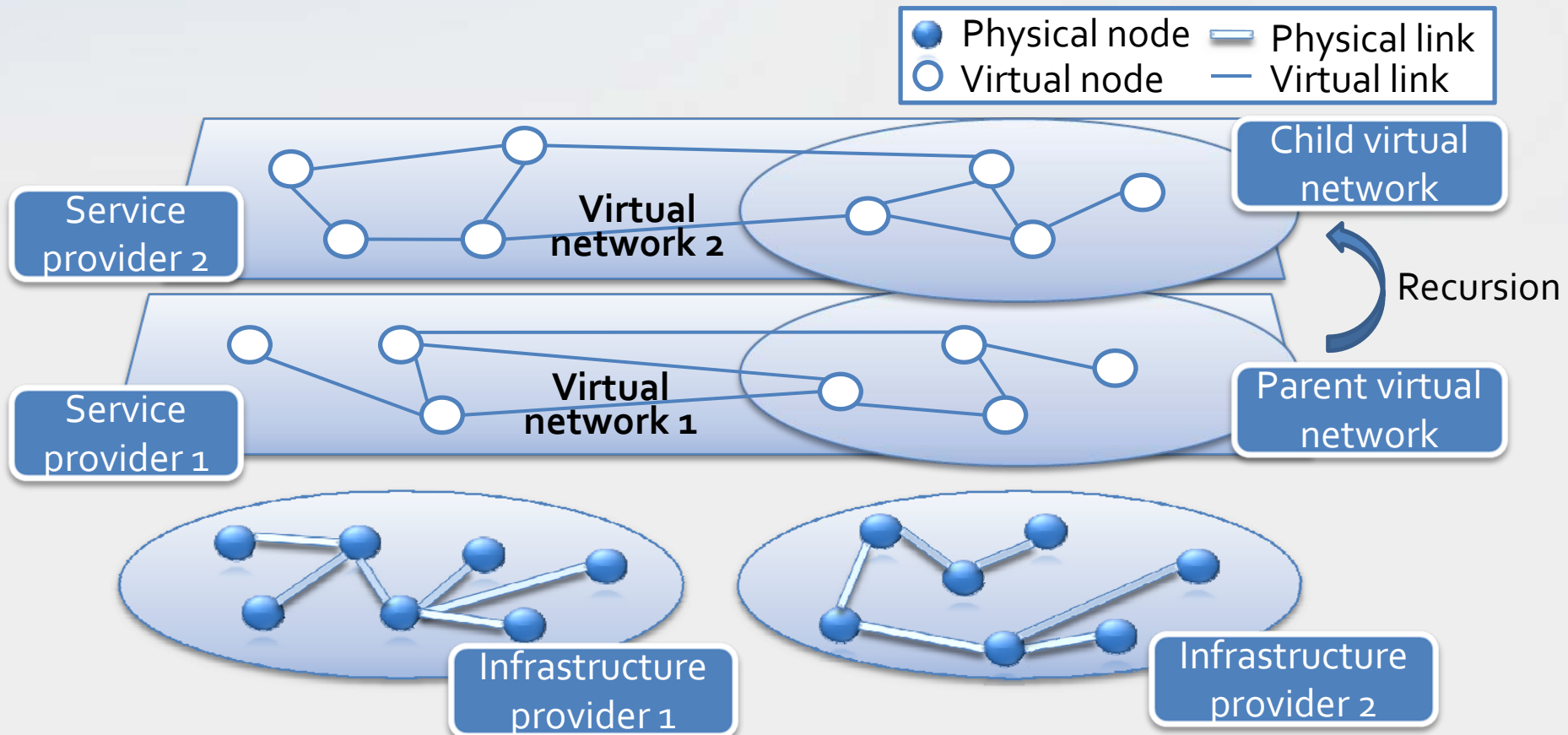
- Network Virtualization is,
  - Defined by decoupling the roles of the traditional Internet service providers into two independent entities
    - Infrastructure providers: manage physical infrastructures
    - Services providers: create virtual networks and offer end-to-end services
  - A tool for evaluating new architectures
  - A fundamental diversifying attribute of the next-generation architecture itself

# Network Virtualization

- Virtual Network (VN)
  - **Collection** of *virtual nodes* connected together by a set of *virtual links* to form a *virtual topology*
  - Each virtual network operated and managed by a single service provider
  - Free to implement end-to-end services by deploying *custom packet formats, routing protocols, forwarding mechanisms, management planes*

# Network Virtualization

- Network virtualization environment



# Network Virtualization

- Architecture principles
  - Coexistence
    - Refers to the fact that multiple VNs from different service providers can coexist together
  - Recursion
    - VNs are spawned from another VN creating a VN hierarchy with parent-child relationships
  - Inheritance
    - Child VNs can inherit architectural attributes from their parents

# Network Virtualization

- Design Goals
  - Flexibility
    - Each service provider should be free to implement arbitrary network topology, routing
    - Customized control protocols independent of the underlying physical network
  - Manageability
    - Modularize network management tasks and introduce accountability at every layer of networking
  - Scalability
    - Infrastructure providers must scale to support an increasing number of coexisting VNs

# Network Virtualization

- Design Goals (cont.)
  - Isolation
    - Ensure isolation between coexisting VNs to improve fault tolerance, security and privacy
  - Stability and Convergence
    - Isolation ensures that faults in one VN do not affect other coexisting
  - Programmability
    - To ensure flexibility and manageability, programmability of the network elements is an indispensable requirement

# Network Virtualization

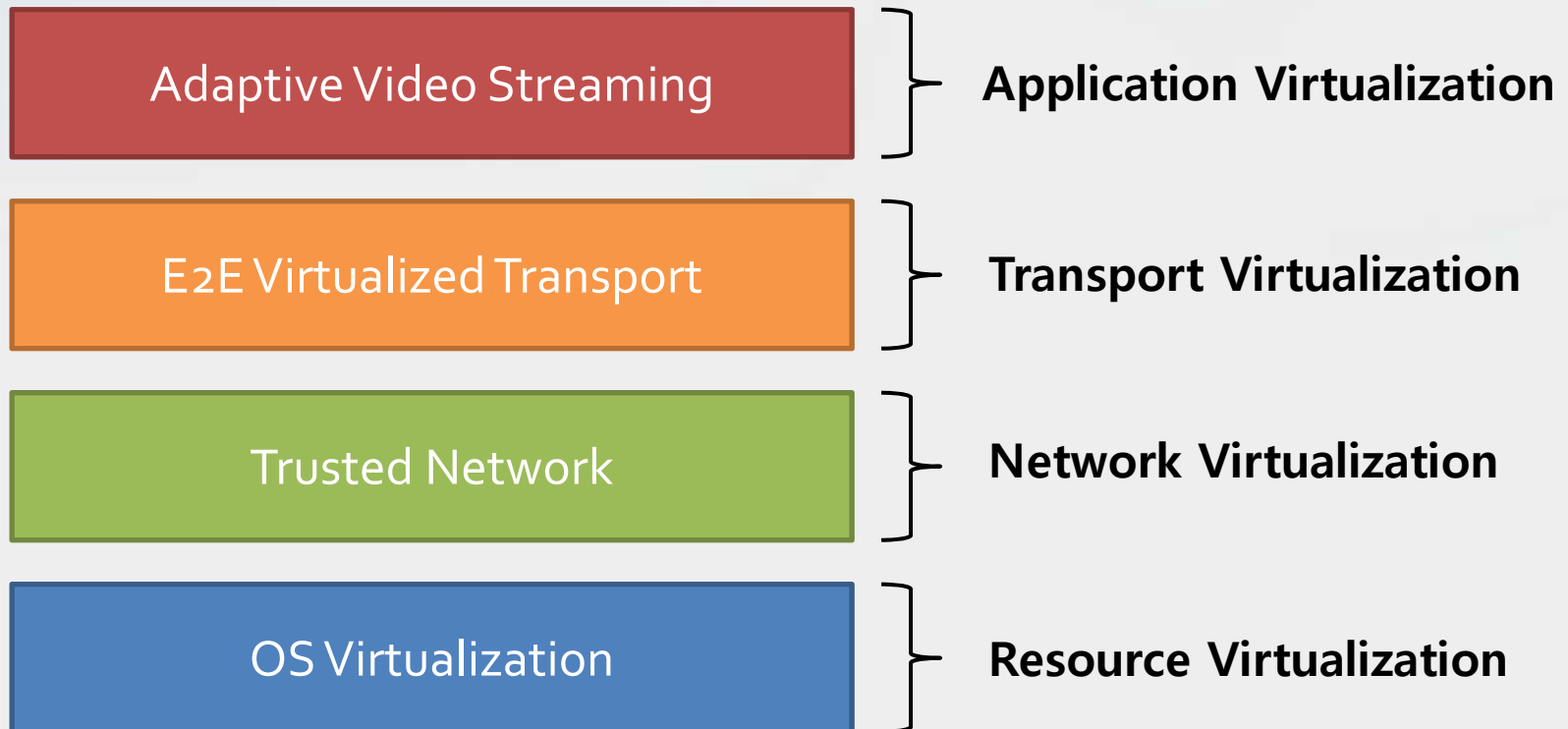
- Virtual Nodes and Virtual Links
  - Virtual Nodes
    - Allow multiple service providers to share the same set of physical resources
    - Implement separate customized control protocols on virtual nodes
  - Virtual Links
    - Links between virtual nodes must also be virtualized
    - Ability to create tunnels over multiple physical links already exists in the context of virtual private networks

# Future Virtualization for Future Internet

- Bottom up approach -

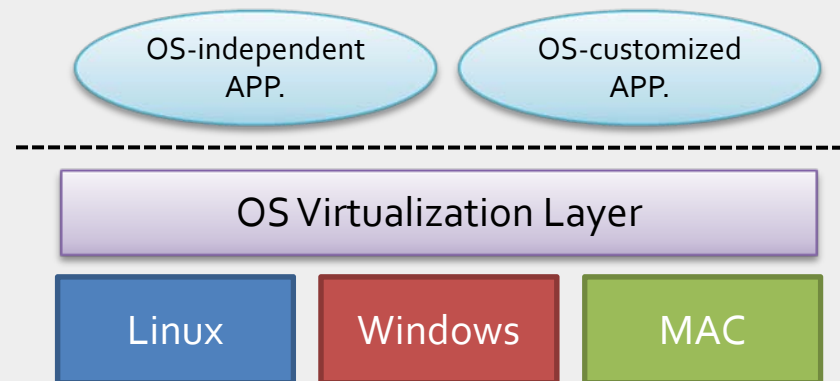
- OS Virtualization
- Trusted Network
- E2E Virtualized Transport
- Adaptive Video Streaming

# Virtualization Stack for Future Internet



# OS Virtualization

- OS Virtualization Layer
  - Supports heterogeneous OSes simultaneously
  - Matches an application's request to corresponding OS
  - Provides two key capabilities to applications
    - OS-independency
    - OS-customization



# OS Virtualization

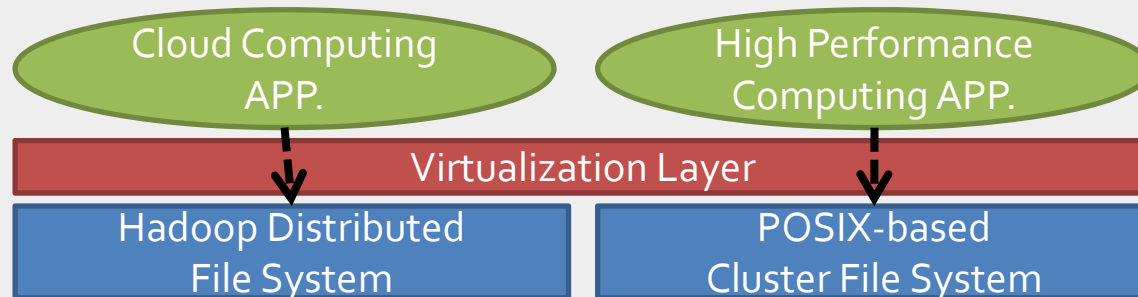
- OS-independency
  - Beyond hardware virtualization, provides OS independent execution environment to applications
  - Also provides language-independency unlike JAVA Virtual Machine
  - Core technology for state migration of applications between heterogeneous OSes

# OS Virtualization

- Seamless state migration between terminal devices
  - Example: VoIP, video streaming services

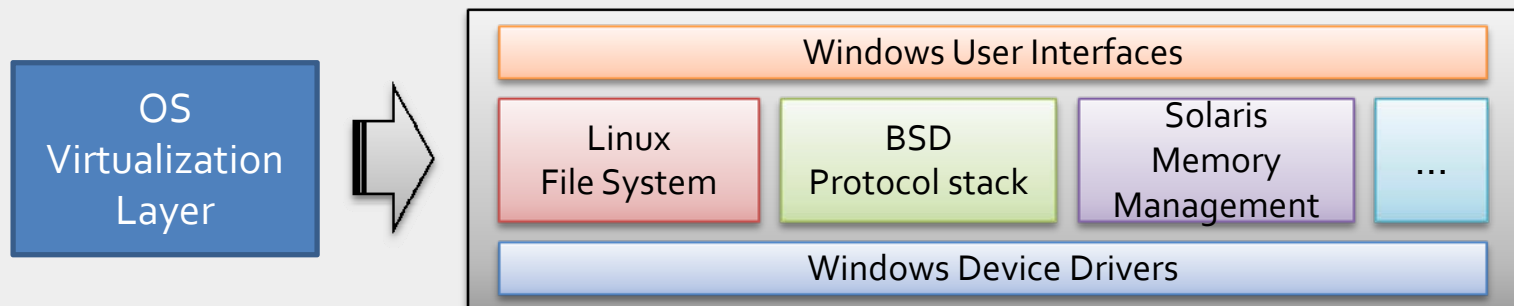


- Portability and availability by OS-independency
  - Example: cloud computing vs. high performance computing



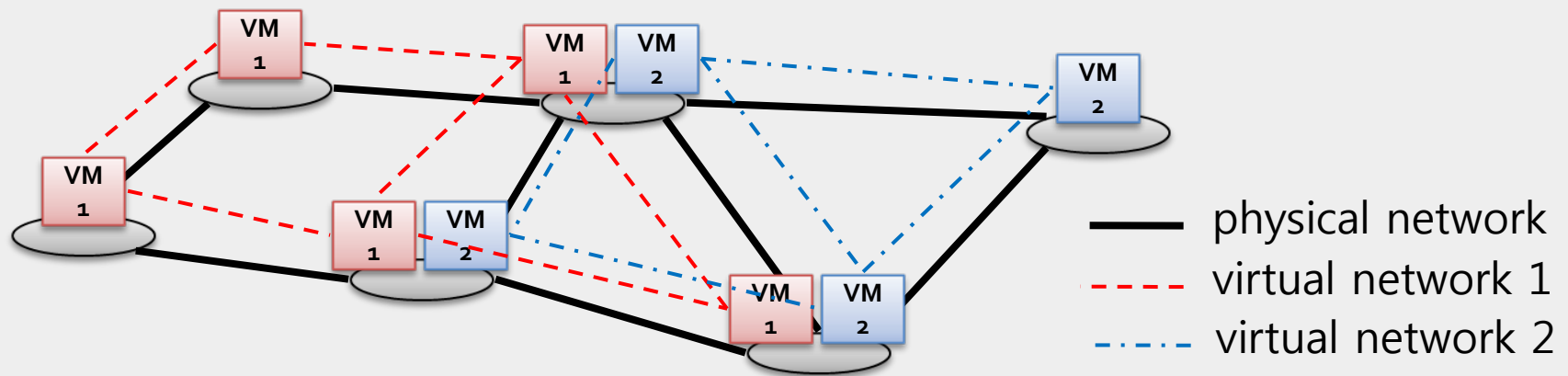
# OS Virtualization

- OS-customization
  - Provides common interfaces for each sub-system (e.g. file system, network stack)
  - Provides a new OS view by customizing functionalities of heterogeneous OSes



# Trusted Network

- Network Virtualization
  - Has widely developed in Future Internet
  - Can create several virtual networks over one physical network by system virtualization



# Trusted Network

- Extension to trusted network
  - Should minimize interference between virtual networks
    - There are shared resources between VMs such as CPU
    - Need an advanced VMM scheduler for fairness, load balancing, etc.
    - Considered partially in ShadowNet (USENIX'09)
  - Should support QoS for real-time services
    - Real-time support in VMM is still challenging issue

# E2E Virtualized Transport

- New requirements for transport layer
  - High network utilization
  - Support for multi-path networking
  - Integration of various transport protocols
- New virtualization concepts
  - Flow Virtualization
  - Protocol Virtualization

# E2E Virtualized Transport

- Flow Virtualization
  - One *physical* session, multiple *virtual* flows
  - Utilizes several heterogeneous network interfaces simultaneously (e.g. WLAN, Wibro, HSDPA, etc.)
  - Obtains more bandwidth in access networks
  - Related issues
    - Dynamic NIC allocation for each connection
    - Flow control for each virtual flow
    - Data aggregation for each connection at receiver side

# E2E Virtualized Transport

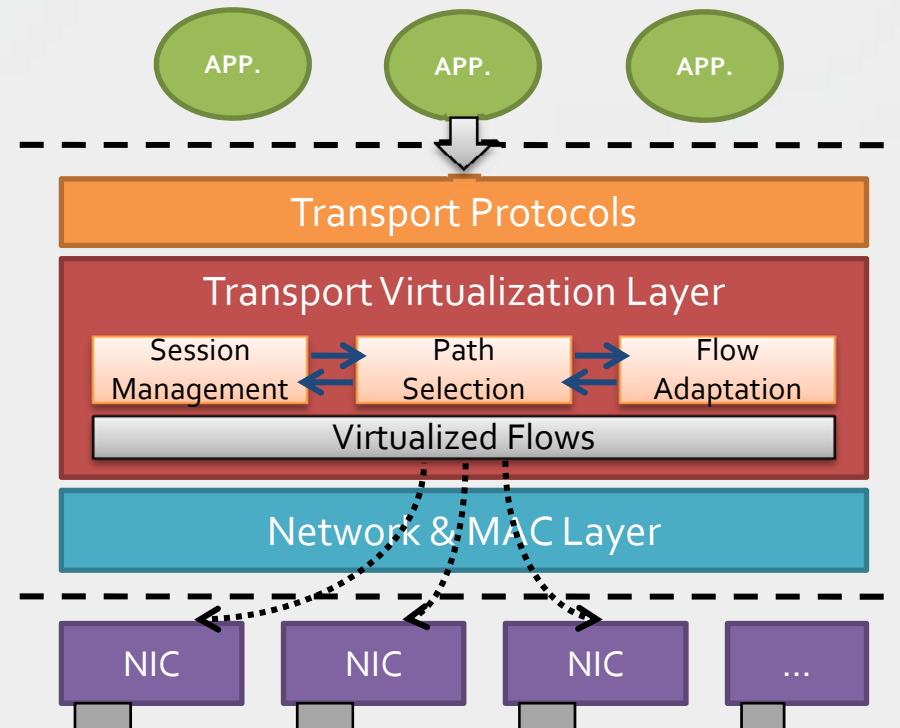
- Protocol Virtualization
  - One *physical* protocol, multiple *virtual* transmission algorithms
  - Utilizes various transmission algorithm that is optimal in a given network environment
  - Provides adaptability against network dynamics

# E2E Virtualized Transport

- Example of protocol virtualization: TCP
  - Utilizes various TCP variants such as CUBIC, Compound, Westwood, Veno, Hybla, etc.
  - Selects a best TCP variant in a given network path
  - Methodology for selection
    - Knowledge-based – performance test
    - Formula-based – throughput prediction
  - Additional requirements
    - Reconfigurable structure such as TCP modules in Linux
    - Network estimation schemes for network detection

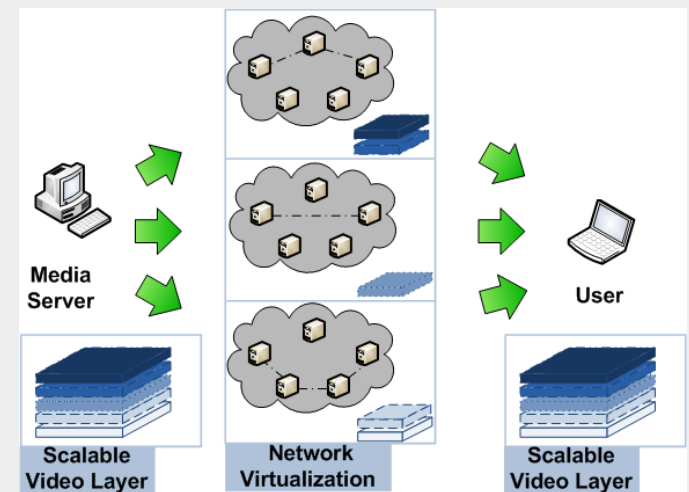
# E2E Virtualized Transport

- Transport Virtualization Layer
  - Supports flow & protocol virtualization
  - Main components
    - Session Management: Keep each connection information
    - Path Selection: Perform Dynamic NIC allocation
    - Flow Adaptation: Select an optimal protocol for each network path



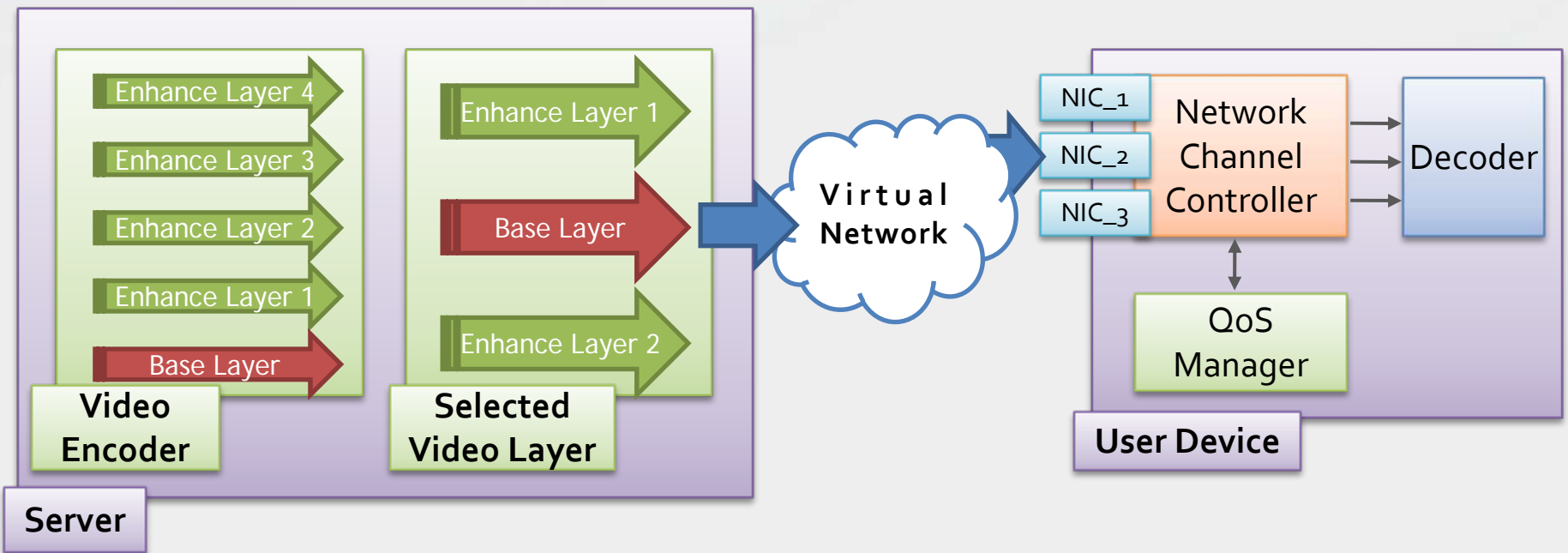
# Adaptive Video Streaming

- Example of Application Virtualization
  - Based on E2E Virtualized Transport
  - Streaming using Scalable Video Codec
    - Support network scalability
    - Adaptive transmission
    - Multi-path transmission
      - Congestion avoidance
      - Load distribution
    - Support real-time decoder

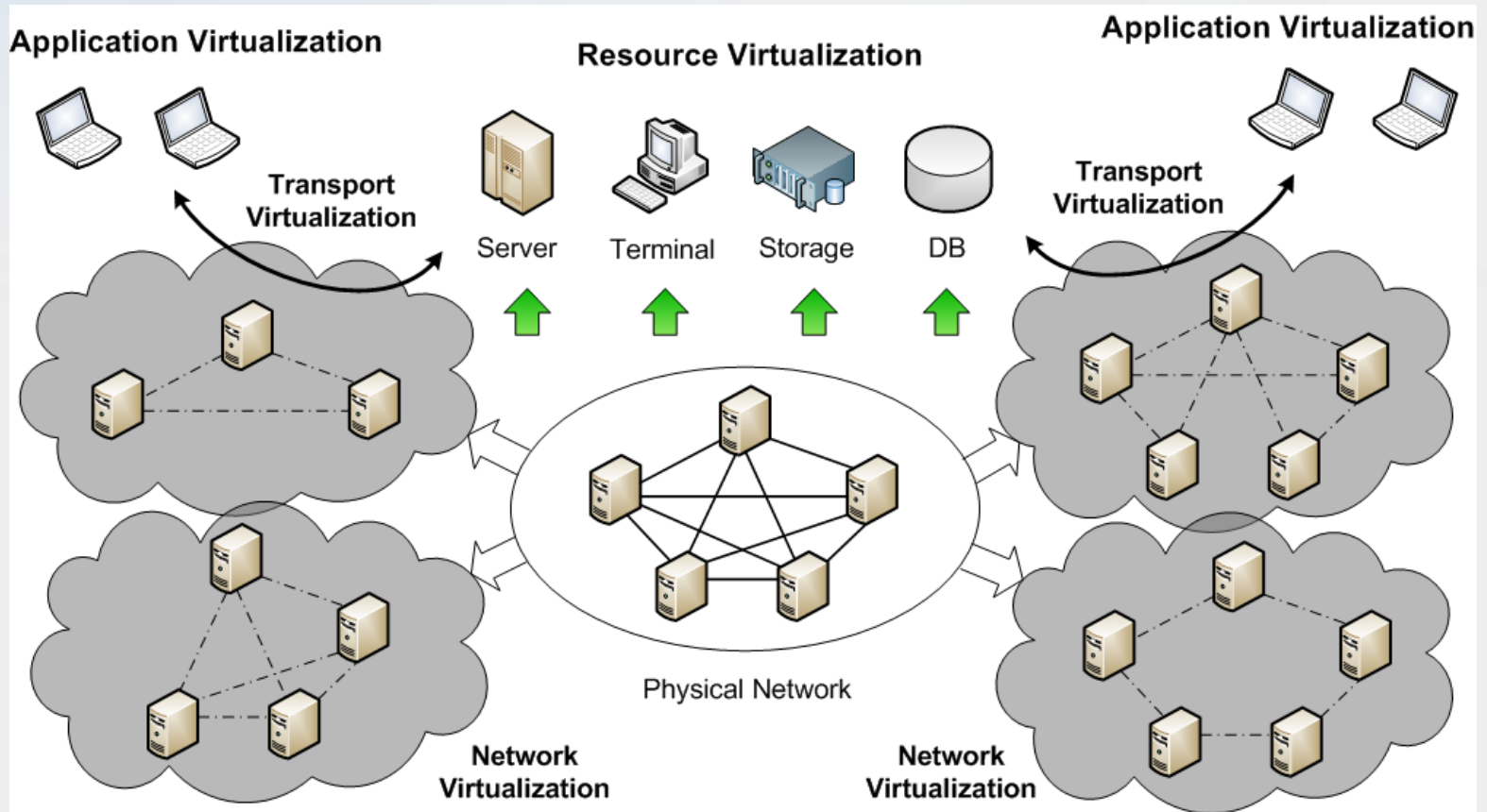


# Adaptive Video Streaming

- Transmits selected video layers through the virtual network



# Overall Architecture



# Summary

- Virtualization technology
  - Starts from system virtualization
  - Has been extended to embedded system, storage, network virtualization, etc.
- For Future Internet
  - Needs further research on fair sharing, real-time, etc.
  - To make a bigger picture:
    - OS/Transport/Application Virtualization to accommodate new requirements

# Reference

- [1] J. Smith and R. Nair, "Virtual Machines: Versatile Platforms for Systems and Processes (The Morgan Kaufmann Series in Computer Architecture and Design)", USA: Morgan Kaufmann Publishers inc., 2005.
- [2] G. J. Popek and R. P. Goldberg, "Formal requirements for virtualizable third generation architectures", Commun. ACM, vol. 17, no. 7. pp. 412-421, 1974.
- [3] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, and A. Warfield, "Xen and the Art of Virtualization", Proc. ACM SOSP, 2003.
- [4] J. Sugerman, G. Venkitachalam, and B.-H. Lim, "Virtualizing I/O devices on vmware workstation's hosted virtual machine monitor", Proc. USENIX ATC, 2002.
- [5] N. M. Mosharaf Kabir Chowdhury, and R. Boutaba, "Network Virtualization: State of the Art and Research Challenges", IEEE Communications Magazine, vol. 47, no. 7, pp. 20-26, 2009.
- [6] X. Chen, Z. M. Mao, and J. V. Merwe, "ShadowNet: A Platform for Rapid and Safe Network Evolution", Proc. USENIX ATC, 2009.

감사합니다.

Operating Systems Lab.

Korea University

(<http://os.korea.ac.kr>)