Old Wine in New Bottles
Key Security Technologies in Data Center Networks for Cloud Computing

Jun Li
Tsinghua University

With many help from my students Yaxuan Qi, Fei He, Baohua Yang
Outline

- Cloud Labels the “New Bottles”
  - Security Concerns of Cloud
  - DCN is Key to Cloud Performance
  - Topology and Algorithm Matters
- Cloud DCN inside the New Bottle
- “Old Wine” with a New Taste
Frost & Sullivan’s Survey

What are the key concerns of ‘Cloud Computing’? – Top 6 Responses

- Security concerns (i.e. data security) - 55.3%
- Lesser control/ownership - 31.0%
- Reliability issues (i.e. performance, time) - 16.3%
- Over-reliance on vendors - 10.7%
- Lack of options in the market - 7.7%
- Lack of support (internal/external) - 7.3%
IDC’s Survey

Q: Rate the challenges/issues ascribed to the 'cloud'/on-demand model
(1=not significant, 5=very significant)

- Security: 74.6%
- Performance: 63.1%
- Availability: 63.1%
- Hard to integrate with in-house IT: 61.1%
- Not enough ability to customize: 55.8%
- Worried on-demand will cost more: 50.4%
- Bringing back in-house may be difficult: 50.0%
- Regulatory requirements prohibit cloud: 49.2%
- Not enough major suppliers yet: 44.3%

Source: IDC Enterprise Panel, August 2008, n=244
Gov Going Cloud with Reservations

- Governments are using cloud computing to reduce IT costs and increase capabilities
  - US government GSA (General Services Administration) now offers a portal for cloud computing services
  - By 2014, over $1 billion of the US federal IT budget would be devoted to cloud computing

- Governments have serious hurdles to overcome
  - Public perception of the secure processing of citizens’ personal information in cloud computing infrastructures
  - Legal and regulatory obstacles which prevent many eGov applications from moving to cloud

Ref. [1], etc.
The Facts about Cloud DCN

- Cloud Data Centers are the big “new bottles”

- Typical data center lifecycle is 10 to 15 years
- Data center instance: Costs in billion range with >100K servers
- By 2010, >15M of servers installed in US, and >$45B to power and cool servers
- 10GE (40, 100 GE later) fiber from servers to ToR, and then to switch fabric
- Traffic volume between servers to entering/leaving data center is 4:1
- Majority of flows are small and each machine has <10 flows in >50% of time (only <5% time with >80 flows)
Network Perspective of Cloud Computing

Performance(cloud computing) = \sum \{\text{performance(network connections)}\} + \sum \{\text{performance(IT resources)}\}

Cloud networking is centered around cloud data centers:

- **Resource networking inside DCN**
  - **Network of resources**: servers and storages
  - Connects the IT elements together to create data centers

- **Access networking to DCN**
  - **Network of clients**: direct and pervasive connect
  - Allows users to access the applications running in those data centers

- **Federation networking between DCN**
  - **Network of clouds**: private and public clouds

Most of challenges are within Data Center Networking (DCN)
Impacts of DCN Virtualization

- **Topology-independent Service Assignment**
  - Decouple service assignment in DC from network topology to support evolutorial service deployment

- **Location-independent Server Addressing**
  - Decouple server’s location in DC from its address to enable seamless server migration

- **Topology-independent Policy Enforcement**
  - Decouple security policy in DC from network topology by flow-based redirection to provide borderless security

- **Location-independent Private Access**
  - Decouple confidential access in DC from physical clustering by Virtual Private Network (VPN) to create virtual private cloud (VPC)
The Two “L”s in Cloud Networking

- The two specific variables likely to determine success in data center networking are the "two 'L's"

- Loss
  - All network protocols have to protect against data loss through retransmission of corrupted information as it takes time, and loss of an information packet is particularly critical with storage protocols because of the risk of creating a corrupted file or leaving a storage device in a bad operating state

- Latency
  - Latency is a special problem in data center and storage networks because it accumulates quickly across the tens of millions of operations involved
Performance Focuses for DCN

- **Topology: Flat and Reliable**
  - Latency accumulates in networks largely in proportion to the number of interfaces a packet transits, and each switch that handles packets poses a risk of loss, in addition to contributing to the total delay → It’s better to reduce the number of intermediate interfaces, and that means reducing the number of switches.
  - A few very large switches will provide better performance than several layers of smaller ones, but concentrating switching into a few devices could increase failure risk too → It’s important for the switches to have the highest possible MTBF and also that the components be redundant and support automatic failover.

- **Algorithm: Fast and Scalable**
  - Flow identification, distribution, and screening is now employed almost everywhere, from TOR to fabric, and from generic processor to dedicated chips → Acceleration is required at algorithmic level to adapt the evolution of hardware platforms and the trend of flow-based switching.

Ref. [2], etc.
DCN Key Security Technologies

- It all boils down to the basic technologies, and at the core of these innovations is **flow-based network processing**
- Foremost, DCN security topology change leads to new architecture
  - Ethane (SIGCOMM 07)
  - PSwitch (SIGCOMM 08)
  - DIFANE (SIGCOMM 10)
- At the same time, DCN security demands higher performance
  - Packet Classification
  - Pattern Matching
  - Flow Scheduling and Traffic Management
Outline

- Cloud Labels the “New Bottles”
- Cloud DCN inside the New Bottle
  - DCN Architecture is Transforming
  - DCN Security’s Topology Changed
  - Ethane, PSwitch, and DIFANE
- “Old Wine” with a New Taste
Conventional DCN Architecture

Core

Data Center
Layer 3

Aggregation

Layer 2

Access

Key:
- BR = L3 Border Router
- AR = L3 Access Router
- S = L2 Switch
- LB = Load Balancer
- A = Rack of Servers

A Single Layer 2 Domain
Juniper’s Stratus

- Consolidate siloed systems and collapse inefficient tiers into a single fabric for any-to-any connectivity, which is flat, lossless, and delivers very low latency

Ref. [3]
Ethane: Addressing the Protection Problem in Enterprise Networks

Martin Casado
Michael Freedman
Glen Gibb
Lew Glendenning
Dan Boneh
Nick McKeown
Scott Shenker
Gregory Watson

Presented By: Martin Casado
PhD Student in Computer Science,
Stanford University

casado@cs.stanford.edu
http://www.stanford.edu/~casado

NSLab, RIIT, Tsinghua Univ

Ref. [4]
Install Rules on Demand

First packet misses the rules
Buffer and send packet header to the controller

Install rules
Forward

Controller

Ref. [6]
Ethane: High-Level Operation

Ref. [4]
**Ethane: High-Level Operation**

**Network Policy**

“Nick can access Martin using ICQ”

**Secure Binding State**

- **ICQ** → 2525/tcp
- **Host A** →
- **IP 1.2.3.4** →
- **Martin** →
- **Host B** →
- **IP 1.2.3.5** →
- **Nick** →

Ref. [4]
Ethane: High-Level Operation

Network Policy
“Nick can access Martin using ICQ”

Host Authentication
“hi, I’m host A, my password is … can I have an IP address?”

Secure Binding State
ICQ → 2525/tcp
Host A → IP 1.2.3.4
IP 1.2.3.4 → switch3 port 4
Martin →
Host B →
IP 1.2.3.5 → Nick →
Ethane: High-Level Operation

Network Policy

"Nick can access Martin using ICQ"

User Authentication

"hi, I'm martin, my password is"

Secure Binding State

ICQ → 2525/tcp
Host A → IP 1.2.3.4
IP 1.2.3.4 → switch3 port 4
Martin → Host A

Host B →
IP 1.2.3.5 →
Nick →

Ref. [4]
Ethane: High-Level Operation

**Network Policy**
"Nick can access Martin using ICQ"

**Host authenticate** Domain Controller

*hi, I’m host B, my password is …*
*Can I have an IP?*

**Secure Binding State**
- **ICQ** → 2525/tcp
- **Host A** → IP 1.2.3.4
- IP 1.2.3.4 → switch3 port 4
- **Martin** → Host A
- Host B → IP 1.2.3.5
- IP 1.2.3.5 → switch 1 port 2
- **Nick** →
Ethane: High-Level Operation

User authentication

hi, I’m Nick, my password is

Network Policy

“Nick can access Martin using ICQ”

Secure Binding State

ICQ → 2525/tcp
Host A → IP 1.2.3.4
IP 1.2.3.4 → switch3 port 4
Martin → Host A

Host B → IP 1.2.3.5
IP 1.2.3.5 → switch 1 port 2
Nick → Host B
**Ethane: High-Level Operation**

Send tcp SYN packet to host A port 2525

Network Policy

"Nick can access Martin using ICQ"

Secure Binding State

<table>
<thead>
<tr>
<th>Source</th>
<th>Destination</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICQ</td>
<td>2525/tcp</td>
</tr>
<tr>
<td>Host A</td>
<td>IP 1.2.3.4</td>
</tr>
<tr>
<td>IP 1.2.3.4</td>
<td>switch3 port 4</td>
</tr>
<tr>
<td>Martin</td>
<td>Host A</td>
</tr>
<tr>
<td>Host B</td>
<td>IP 1.2.3.5</td>
</tr>
<tr>
<td>IP 1.2.3.5</td>
<td>switch 1 port 2</td>
</tr>
<tr>
<td>Nick</td>
<td>Host B</td>
</tr>
</tbody>
</table>

Domain Controller

- Permission check
- Route computation

Ref. [4]
Ethane: High-Level Operation

- Permission check
- Route computation

Domain Controller

Network Policy
“Nick can access Martin using ICQ”

Secure Binding State
- ICQ → 2525/tcp
- Host A → IP 1.2.3.4
- IP 1.2.3.4 → switch3 port 4
- Martin → Host A

- Host B → IP 1.2.3.5
- IP 1.2.3.5 → switch 1 port 2
- Nick → Host B

Ref. [4]
A Policy-aware Switching Layer for Data Centers

Dilip Joseph
Arsalan Tavakoli
Ion Stoica

University of California at Berkeley

Ref. [5]
Pre-install Rules in Switches

Packets hit the rules

Pre-install rules

Controller

Forward

NSLab, RIIT, Tsinghua Univ

Ref. [6]
Policy-aware Switching Layer

1. Take middleboxes off-path
2. Separate policy from reachability

HTTP
TCP port = 80

Firewall $\rightarrow$ Load balancer

Existing mechanisms

Policy-aware switching layer

Ref. [5]
- Distributed forwarding
- Loadbalancing middleboxes
- Different policies for different traffic

Ref. [5]
Scalable Flow-Based Networking with DIFANE

Minlan Yu
Princeton University

Joint work with Mike Freedman, Jennifer Rexford and Jia Wang
Packet Redirection and Rule Caching

Ingress Switch

Authority Switch

Egress Switch

First packet

Following packets

Feedback: Cache rules

Redirect

Forward

Hit cached rules and forward

A slightly longer path in the data plane is faster than going through the control plane

Ref. [6]
Outline

- Cloud Labels the “New Bottles”
- Cloud DCN inside the New Bottle
- “Old Wine” with a New Taste
  - Packet Classification
  - Pattern Matching
  - Flow Scheduling
  - Traffic Management
Key Algorithm Development (I)

- Packet Classification
  - Many classical algorithms: HiCuts, RFC, BV, GT, etc.
  - Recent algorithm advancements: HyperSplit (INFOCOM 09), etc.

- State-of-the-art
  - 100 Gbps and beyond throughput
  - Essential for Stateful Inspection
  - Enables rule based search without caching session
  - Makes TOR packet filtering possible
Packet Classification Taxonomy

Prefix Search Algorithms
- Hierarchical Trie
  - Set-pruning Trie
    - Grid-of-Trie

Range Search Algorithms
- Bit Vector
  - ABV

Prefix Search Algorithms
- HiCuts

Range Search Algorithms
- Cross-producting
  - RFC

- HSM

equal-sized space partition

rule-based space partition

NSLab, RIIT, Tsinghua Univ
Key Algorithm Development (II)

- Pattern Matching
  - Many classical algorithms: AC, BM, WM, D²FA, etc.
  - Recent algorithm advancements: Subset-D²FA (TBP), etc.

- State-of-the-art
  - 10Gbps and beyond throughput
  - Essential for Deep Inspection
  - Enables protocol identification at very high speed
  - Makes direct access hacking prevention possible
Pattern Matching Taxonomy

- **RegEx matching**
  - DFA+NFA
  - Hc-FA
  - Hybrid-FA
  - XFA
- **String matching**
  - DFA
  - AC
  - WM
  - BFM
  - RSI
  - Hashing
  - Subset-DFA
  - BloomFilter
- **Prefix-Suffix**
  - Brutal Force
  - BM
  - KMP

**multi-pattern matching algorithms**

**single-pattern matching**

NSLab, RIIT, Tsinghua Univ
Key Algorithm Development (III)

- **Flow Scheduling**
  - Current industry standard: Equal-Cost Multi-Path (ECMP)
  - ECMP drawback: Static + Oblivious to link-utilization

- **Hedera (NSDI 10)**
  - Optimize achievable bisection bandwidth by assigning flows non-conflicting paths (Upto 96% of optimal bisection bandwidth, > 2X better than standard techniques)
  - Uses flow demand estimation + placement heuristics to find good flow-to-core mappings
Key Algorithm Development (IV)

- Traffic Management
  - Previous work mostly on congestion control in end-to-end environment
  - Now approaches start to study global control mechanisms based on flow and protocol identification and measurement

- Besides NOX, recently DCTCP (SIGCOMM 10)
  - TCP-like protocol for DCN to achieve high burst tolerance, low latency, and high throughput with commodity shallow buffered switches
  - Queue length: about 1/10 of that of TCP
Summary

- Cloud DCN is key to overall cloud performance and security
- Cloud DCN requires flat architecture and demands high bandwidth, also to security solutions
- Cloud DCN security has different topology for security deployment, and key algorithms are still the same, but to be accelerated
Summary

Refined *old wine* will have better taste in *new bottles* with modern style!
Reference

2. T. Nolle, Cloud Networking: Inter-networking Data Centers and Connecting Users, Feb. 2010
3. The Cloud-ready Data Center Network, Juniper, May 2010
7. Y. Qi, L. Xu, B. Yang, Y. Xue, & J. Li, Packet Classification Algorithms: From Theory to Practice, THU, Apr. 2009
8. Y. Qi, J. Fong, W. Jiang, B. Xu, J. Li, & V. Prasanna, Multi-dimensional Packet Classification on FPGA: 100 Gbps and Beyond, THU & USC, Dec. 2010