Supporting Privileged Cloud Traffic with SDN (aka VIPlanes)

James Griffioen

Laboratory for Advanced Networking and
Center for Computational Sciences

University of Kentucky
Acknowledgements

• Current/Past Co-PIs
  – Cody Bumgardner
  – Zongming Fei
  – Ken Calvert
  – Vince Kellen

• Staff/Students
  – Lowell Pike
  – Matthew Moseley
  – Jacob Chappell
  – Sergio Rivera P
  – Mami Hayashida
  – Charles Carpenter
  – Yongwook Song
  – Hussamuddin Nasir
  – Andrew Groenewold
  – Peter Oostema

• Funding Agencies

  *This work is supported, in part, by the National Science Foundation under NSF grants OCI-1246332, ACI-1541380 and ACI-1541426.*
Talk Outline

• Problem Statement: Why privileged flows?
• Current Solution: Science DMZs
• A New Approach: An All-Campus DMZ
• Challenges: Technical and Non-Technical
• Prototype: GENI Experiment
• Deployment: Recent Directions
• Demos: GENI and Campus Net
The Problem: Supporting Big Data

• **Large data sets drive many research techniques** today, including (but not limited to) simulation, modeling, data mining, analytics, machine learning, and visualization.

• **These techniques are becoming prevalent across all disciplines** (e.g., biology, bioinformatics/genomics, business, sociology, transportation, engineering/computer science, linguistics, library science, communications)

• The need to **move large data** sets in an efficient and secure manner is now **a daily reality for many researchers**.

• **HPC systems have long dealt with this problem** – but now the user community is much larger.

• **Existing campus networks are not designed to support pervasive big data usage**.
Typical Campus Network

- Internet
- Edge Router
- Firewalls
- Middleboxes
- Campus Core
- Bldg A
- Bldg B
- Bldg C
- HPC

University of Kentucky
Typical Campus Network

- **Internet**
- **Edge Router**
- **Firewalls**
- **Middleboxes**
- **Campus Core**
- **Bldg A**
- **Bldg B**
- **Bldg C**
- **HPC**

Normal Flow Path (100s of Mbps)
Middleboxes

• Provide important services that enforce policy and offer enhanced functionality.

• Example middlebox services include:
  – Network Address Translation (NAT)
  – Traffic Shaping/QoS Enforcement
  – Load Balancing
  – Intrusion Detection/Prevention – and other types of DPI
  – Firewalls
  – VPN
  – (Wireless) Splash Pages
  – Caching Services/Proxies

• Middleboxes are placed strategically throughout the network, not just the edge.

• Because middleboxes operate on packets they pose a bottleneck to network performance.

• Even if the network is upgraded to support high-speed connections, middleboxes can degrade or limit network throughput and/or latency.
Typical Campus Network

- Internet
- Edge Router
- Firewalls
- Middleboxes
- Campus Core
- Bldg A
- Bldg B
- Bldg C
- Normal Flow Path (100s of Mbps)

University of Kentucky
Typical Campus Network

- Internet
- Edge Router
- Firewalls
- Middleboxes
- Campus Core
- Bldg A
- Bldg B
- Bldg C
- Normal Flow Path (100s of Mbps)
- HPC
How do we solve this problem?
Move Nodes Outside the Firewall

- Internet
- Edge Router
- Firewalls
- Middleboxes
- Campus Core
- Bldg A
- Bldg B
- Bldg C
- HPC
Move Nodes Outside the Firewall
Move Nodes Outside the Firewall

Diagram showing the network architecture with:
- Internet
- Edge Router
- Firewalls
- HPC
- Middleboxes
- Campus Core
- Bldg A
- Bldg B
- Bldg C

University of Kentucky logo
Campus Science DMZ

- Internet
- Edge Router
- HPC
- Firewalls
- Middleboxes
- Campus Core
- Bldg A
- Bldg B
- Bldg C

Science DMZ
High Speed Flows (Multiple Gbps)
Science DMZ

(Internal Network Architecture)

Source: https://fasterdata.es.net/assets/dmz-supercomputer.jpg
Standard Science DMZ Solution

- Deploy a *Science DMZ network* connected to the network edge.
- Move HPC machines to the Science DMZ network

**Advantages:**
- Traffic from HPC machines bypass middlebox bottlenecks

**Disadvantages:**
- Science DMZ machines are not protected by middleboxes.
- Campus (middlebox) policy enforcement is not applied to any traffic from Science DMZ machines. Even non-science flows (e.g., Netflix) bypass campus policy enforcement.
- Researchers must decide whether to connect their machines to the Science DMZ or the Campus Network.
Standard Science DMZ Solution

- Deploy a *Science DMZ network* connected to the network edge.
- Move HPC machines to the Science DMZ network

**Advantages:**
- Traffic from HPC machines bypass middlebox bottlenecks

**Disadvantages:**
- Science DMZ machines are not protected by middleboxes.
- Campus (middlebox) policy enforcement is not applied to any traffic from Science DMZ machines. Even non-science flows (e.g., Netflix) bypass campus policy enforcement.
- Researchers must decide whether to connect their machines to the Science DMZ or the Campus Network.
Standard Science DMZ Solution

- Deploy a *Science DMZ network* connected to the network edge.
- Move HPC machines to the Science DMZ network

**Advantages:**
- Traffic from HPC machines bypass middlebox bottlenecks

**Disadvantages:**
- Science DMZ machines are not protected by middleboxes.
- Campus (middlebox) policy enforcement is not applied to any traffic from Science DMZ machines. Even non-science flows (e.g., Netflix) bypass campus policy enforcement.
- Researchers must decide whether to connect their machines to the Science DMZ or the Campus Network.
Disadvantages:

- Science DMZ machines are not protected by middleboxes.
- Campus (middlebox) policy enforcement is not applied to any traffic from Science DMZ machines. Even non-science flows (e.g., Netflix) bypass campus policy enforcement.
- Researchers must decide whether to connect their machines to the Science DMZ or the Campus Network.
Deploy a Science DMZ network connected to the network edge.

- Move HPC machines to the Science DMZ network

**Advantages:**
- Traffic from HPC machines bypass middlebox bottlenecks

**Disadvantages:**
- Science DMZ machines are not protected by middleboxes.
- Campus (middlebox) policy enforcement is not applied to any traffic from Science DMZ machines. Even non-science flows (e.g., Netflix) bypass campus policy enforcement.
- Researchers must decide whether to connect their machines to the Science DMZ or the Campus Network.
Deploy a Science DMZ network connected to the network edge.

- **Advantages:**
  - Traffic from HPC machines bypass middlebox bottlenecks.

- **Disadvantages:**
  - Science DMZ machines are not protected by middleboxes.
  - Campus (middlebox) policy enforcement is not applied to any traffic from Science DMZ machines. Even non-science flows (e.g., Netflix) bypass campus policy enforcement.
  - Researchers must decide whether to connect their machines to the Science DMZ or the Campus Network.
Protecting DMZs?

Source: https://fasterdata.es.net/assets/dmz-supercomputer.jpg
Protecting DMZs?

Can we offer any protection?

Source: https://fasterdata.es.net/assets/dmz-supercomputer.jpg
Protecting DMZs?

Yes – for example “No incoming traffic”.

Source: https://fasterdata.es.net/assets/dmz-supercomputer.jpg
Protecting DMZs?

That presents problems (e.g., DTN nodes)

Source: https://fasterdata.es.net/assets/dmz-supercomputer.jpg
Research DMZs
(More than just HPC, and causing additional problems)

Source: https://fasterdata.es.net/assets/dmz-supercomputer.jpg
Research DMZs
(More than just HPC, and causing additional problems)

Source: https://fasterdata.es.net/assets/dmz-supercomputer.jpg
A New Approach:
An All-Campus Science DMZ

- **Observation:** Science DMZs enable “privileged traffic” – traffic that has been pre-approved to by-pass campus middleboxes. (Technically “machines” are pre-approved).

- **Idea:** Use Software Defined Network (SDN) capabilities to intercept approved science flows and route them around performance-limiting middleboxes.

- **Design:**
  - Install SDN-enabled routers/switches in the campus network.
  - Ensure there exists a middlebox free path to the Internet from every end host by deploying an alternative to the campus core (e.g., an SDN core directly connected to the campus edge).
  - Configure a default SDN rule (matching all packets) to forward packets using normal routing (i.e., through the existing campus core and through existing middleboxes).
  - Dynamically install SDN rules (with higher priority) targeting approved flows that route over the alternative SDN core.
Developing an All-Campus Science DMZ
Developing an All-Campus Science DMZ

Diagram showing the network architecture:
- Internet
- Edge Router
- Firewalls
- Middleboxes
- Campus Core
- SDN Switches:
  - Bldg A
  - Bldg B
  - Bldg C
- HPC

Nodes represent different buildings and network components, connected through SDN switches.
Developing an All-Campus Science DMZ
Developing an All-Campus Science DMZ
Developing an All-Campus Science DMZ

Internet

Edge Router

Firewalls

Middleboxes

SDN Core

Campus Core

SDN Switch

Bldg A

Bldg B

Bldg C

HPC
Developing an All-Campus Science DMZ
Developing an All-Campus Science DMZ

- Internet
- Edge Router
- Firewalls
- Middleboxes
- Campus Core
- SDN Core
- SDN Switch
- SDN Switch
- SDN Switch
- Bldg A
- Bldg B
- Bldg C
- HPC

Flow Paths:
- High-speed Flow Path
- Normal Flow Path
Developing an All-Campus Science DMZ

All-Campus Science DMZ
Flows (not machines) join the DMZ.

High-speed Flow Path

Normal Flow Path

Internet

Edge Router

Firewalls

Middleboxes

SDN Core

SDN Switch

Bldg A

Bldg B

Bldg C

HPC

University of Kentucky
Design Goals

• Switches/Routers must be able to forward packets “normally” by default – i.e., they must support standard routing/forwarding protocols

• Control Software must
  – Discover topology information
  – Discover the placement/role of middleboxes
  – Accept requests to enable “privileged flows”
  – Compute paths that by-pass middleboxes
  – Compute SDN rules and insert them to “pick off” privileged flows
  – Ensure rules remain in place for the duration of the flow
  – Remove rules that are no longer needed
  – Gracefully handle failures
Design Goals - Realized

• Switches/Routers must be able to forward packets “normally” by default – i.e., they must support existing routing/forwarding protocols

• Control Software must
  – Discover topology information
    • Built-in controller support + new module to invoke SNMP
  – Discover the placement/role of middleboxes
    • Via topology discovery and config files
  – Accept requests to enable “privileged flows”
    • VIPlanes server authenticates/authorizes requests
  – Compute paths that by-pass middleboxes
    • Path computation service uses topo info to compute middlebox free paths
  – Compute SDN rules and insert them to “pick off” privileged flows
    • Path computation service and new controller module create and install rules
  – Ensure rules remain in place for the duration of the flow
  – Remove rules that are no longer needed
  – Gracefully handle failures
Challenges (1)

- **OpenFlow Switch Limitations**
  - The OF “normal forwarding” action is often not supported (or not fully supported).
    - OVS (no), HP (yes), Dell (yes?), Cisco (not really)
  - Most Switches are not fully compatible with the OF standards
    - OVS, HP/Dell/Cisco (all miss something)
  - Even when switches support operations defined in the standards, the OF rules/actions may be pushed into software, severely degrading performance
    - OVS (all software), HP (depends on version), Dell/Cisco (?)
  - Switches often only work with certain controllers
    - ODL, Floodlight, HPVAN, RYU largely work with OVS, but are hit/miss HP/Dell/Cisco
Challenges (2)

• **OpenFlow (OF) Controller Limitations**
  – Generic with the expectation of being extended
    • But relatively few extensions exist *(added new modules)*
    • Extension programming environments have caveats
  – Northbound interfaces
    • Have limited functionality
    • Differ from controller to controller *(designed RAPTOR: A REST API TranslaTOR for OpenFlow Controllers)*
  – Largely focused on layer 2
    • *(compute re-write rules that mimic L2/L3 forwarding behavior)*
Rewrite Rules – Mimic Normal Packets
Rewrite Rules – Mimic Normal Packets

Internet

Edge Router

Firewalls

Middleboxes

SDN Core

Campus Core

Layer 2 SDN Path?

SDN Switch

Bldg A

SDN Switch

Bldg B

SDN Switch

Bldg C

HPC

UNIVERSITY OF KENTUCKY
Challenges (3)

• **Campus IT Limitations**
  – Funding to replace campus network switches
  – Staff lack SDN knowledge
  – Single vendor shop
  – Redundancy/Reliability Requirements
  – Simplified Management Requirements
    • e.g., rapid/instant access
  – CAB processes
Flow-level Monitoring

- OpenFlow support flow metrics (packet/byte counts) that enable flow-level monitoring
- Allows Users/IT to monitor performance of specific data transfers
- Can also be used for bandwidth management and debugging purposes

Packet Transferring Speed Graph (per second)
Using GENI to implement A Prototype All-Campus DMZ

- Goal was to emulate the UK campus network topology.
- Created a GENI topology consisting of OpenFlow-enabled switches/routers (OVS nodes) as well as middleboxes and end systems (linux hosts). Used a single host to emulate the Internet.
- Middleboxes were configured to slow down traffic via slow speed links specified in the RSPEC
Emulated UK Network
(GENI Desktop View)
Emulating the UK Network using GENI
Emulated UK Network
(HP VAN Controller View)
Emulated UK Network
(HP VAN Controller View)
A Prototype All-Campus Science DMZ (using GENI)

- Operated the OF switches in “normal” mode (forwarding packets across the campus backbone network as normal). Implementing “normal” required inserting OF rules that caused:
  - L2 nodes to act as normal switches would (e.g., handle ARP, DHCP, etc), and
  - L3 nodes to act as normal routers would (e.g., rules to statically route packets like a router would)
  - All “normal” rules were inserted with low priority (i.e., default behavior)
- Developed code to dynamically install (high priority) SDN rules to intercept approved science flows (used by researchers) and route them over the SDN core.
- Deployed NAT box near the edge to dynamically create alternate source addresses for approved flows, allowing return (incoming) packets to traverse the SDN core.
- Demonstrated significant bandwidth improvement for approved science flows while other (concurrent) flows from the same node took the slow middlebox route – e.g., a scientific secure copy (scp) transfer took the SDN path, while a video transmission took the campus middlebox path.
Demo
Demo: Normal Path Through the Core
Demo: High-Speed Path via SDN
Deploying at the University of Kentucky

- Deployed a new SDN Core network (OpenFlow) and connected it directly to the campus edge router.
- Deployed SDN-enabled switches/routers at the head-end of multiple science buildings.
- Operated the new switches using “normal” mode (forwarding packets across the campus backbone network as normal).
- Installed SDN rules to intercept approved science flows (used by various researchers) and route them over the SDN core.
- Observed an order of magnitude (or more) bandwidth improvement for approved science flows resulting in single flow transfer rates in the multiple Gbps range – all while other (concurrent) flows from the same node took the slow, policy-enforcing, middlebox route.
Thank You

Question?