Welcome to the CCoE Webinar Series. Our topic today is SDN and IAM Integration at Duke University. Our host is Jeannette Dopheide.

The meeting will begin shortly. Participants are muted. You may type questions into the chat box during the presentation.

This meeting will be recorded.

The CTSC Webinar Series is supported by National Science Foundation grant #1547272.

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the NSF.
SDN and IDM Integration Duke – The implementation of a Protected Research and Data Network and the road to SDN.

Richard Biever
Charley Kneifel
EVOLUTION OF A PROTECTED NETWORK
Protected Research @ Duke Pre-2012

Highly distributed IT and research environment
Highly distributed IT and research environment

Pre-2012, it was assumed that most protected research data originated from Duke Medicine.
Highly distributed IT and research environment

Pre-2012, it was assumed that most protected research data originated from Duke Medicine.

However, risk assessments and interactions with the IRB showed protected research data across Duke.
The #1 approved way of protecting sensitive data was to keep it unplugged from the network.
The #1 approved way of protecting sensitive data was to keep it unplugged from the network.

But what if you needed to collaborate, OR repurpose the computer?
Evolution of a Protected Network
Evolution of a Protected Network
Evolution of a Protected Network

Faculty, Researchers, Students, Staff

SourceFire IDS

Core Router

SourceFire IPS

Internet Router

Protected Network

Campus Network
Evolution of a Protected Network

Faculty, Researchers, Students, Staff

Internet Router

Campus Network

Core Router

SourceFire IDS

Core Router

SourceFire IPS

Internet Router

Protected Network

SSRI's

Protected Research Data Network

Campus Network
PARTNERSHIPS MATTER
Organization and Administration

Leadership
- Do we have protected data?
- Could a centrally provided service work?
- Partnership between CIO and Vice Provost of Research

Research Partnership (SSRI)
- Who could bridge IT and researchers?
- Who could provide governance support?

OIT’s role
- Focus on development of infrastructure
- Remove the silos
- Research computing focus

ITSO’s role
- Initially – the facilitator/pusher
- Today – policy, risk, monitoring and consultative
SSRI’s Role

Technical and Administrative Support

Data warehousing & honest broker

Grant writing & proposals

Facilitate collaboration
Support Model

OIT
- Technical ownership - Infrastructure (network, servers, storage)
- Functional ownership (relationship with customers; new feature requests)
- Architecture design

ITSO
- Security policies and standards
- Monitoring
- Incident Response
- Architecture design
- Exception requests

SSRI
- Technical ownership – applications and data
- Interface to research community
- Authorization for data access
- Grant writing support

Researcher
- Compliance with DUA
- Responsible for ensuring laptop/desktop security controls are in place

Shared responsibility for the security of the data
Network and Systems Automation

AUTOMATION & VIRTUALIZATION
(SECURELY OF COURSE)
Services/Needs to automate

We started with:

• **Clockworks** – a web application/API to provision VM resources.
• **Locksmith** – a web application to provision SSL certificates.
• **Switchboard** – a web application to provision SDN paths

This led to:

• **Research Toolkits** – Provision services – VMs, Storage, Groups
• **VM Manage** - a pilot service providing access to linux virtual machines (VMs) and Linux Docker containers for software development projects via a web browser.
• **Proconsul** - Linux via Web Browser (Windows, Matlab, Mathematica, …)
Virtual Machine Provision Request

Please complete this form. Please see our VMware pricing page for information on pricing. You can stop at any time and save your work by pressing the ‘Save for Later’ button.

**General Settings**

- **Hostname for the VM**: [Enter hostname]

  To create multiple hosts with the same configuration, list hostnames here, separated with a comma.

- **Associated Service**: [Enter service name]

- **Fund code**: [Enter fund code]

- **Server function**: [Please Select]

- **Location**: [Enter location]

**System Administration**

- **System Administration Options**
  - Off Administration
  - 8/5 Standard C&I OS System Administration
  - 24/7 Standard C&I OS System Administration
  - 8/5 Webhosting C&I System Administration
  - 24/7 Webhosting C&I System Administration

**Operating System**

- **Operating System**
  - [Enter OS version]

**Hosting Options**

- **Hosting level**: [Please Select]

- **Sensitive Data**: [No] [Yes]

  *Visit the IT Security Office web site for data classification information.*

- **Secure project**: [Enter project name]

**System Size**

- **VM Size**: [Custom]

- **RAM**: [Enter GB]

- **CPU**: [Enter CPU(s)]

- **Storage size**: [Enter GB]

**Extra Requirements**

- **Backups Needed**: [No] [Yes]
Containers for research

• How can we provide:
  – Reproducible analysis tool suites
  – Flexible and migratable workloads
  – Custom environments on shared infrastructure
Zero-install consoles

• VNC (virtual network console) and RDP (remote desktop protocol) can run inside modern web browsers
• tablets work …but you may want a real keyboard/mouse
• how is this even possible?
  – HTML5 = live browser-to-server communication
  – fast Javascript in modern web browsers
  – fast networks
About This Service

VM-Manage is a pilot service providing students and instructors with semester-long access to Linux virtual machines (VMs) and Linux Docker containers for software development projects. Typical uses of this service are to develop and test servers for Innovation CoLab projects or coursework. Think of VM-Manage as an alternative to running a server in your dorm room - you have complete control of the server, and you are responsible for server security and backing up your work.

VM-Manage VMs are well-suited to acting as backend servers for mobile applications or public-facing web services because they have public IP addresses and are available 24x7. However, since this is a software development sandbox, you should not run production or high traffic servers on VM-Manage machines.

Get your own VM

Students and instructors can reserve Virtual Machines for a semester.

Click here to get your own VM.

Service Info and FAQs

What kind of VMs are available?

How do I stop and restart my VM?

How do I get started?

Where is the documentation for my VM?

How can I install additional software on my VM?

I totally messed up my VM and it no longer works - what can I do?

How should I back up my work?

I will be teaching a course and want to be sure VMs are available for my students, what should I do?
Research Toolkits and RAPID VMs

- Manage research group memberships
  - Who has access to what
- Manage/Provision VMs using RAPID
  - On research network (private IPs) or on public IPs
- Manage storage allocations (and access)

- RAPID – Acronym – Research Accelerating Preconfigured Individual Dynamic Virtual Machines (RAPID VMs)
Technologies Applied – Proconsul

- Powershell
- MMC
- Active Directory
Technologies Applied – Proconsul

- Powershell
- MMC
- Active Directory
Technologies Applied – Proconsul

- **Powershell**
- **MMC**
- **Active Directory**
- **NoVNC**
- **Browser access With Shib+MFA**

Random user, random password for the session.
AUTHENTICATION AND AUTHORIZATION
Authentication and Authorization Flow

1. Remote User -> Duke Website Local SP
3. Remote Shib IDP -> Duke Website Local SP
5. Duke Shib IDP -> LDAP/Grouper
6. Duke Shib IDP -> LDAP/Grouper
7. Duke Shib IDP -> Duke Website Local SP
8. Duke Website Local SP -> Services via VNC/RDP in Browser
Technologies Applied – Federated Research

Campus Network

Protected

Network

Researcher

Duke VM Manage
Duke CLOCKWORKS
Duke ANTIKYTHERA

Duke UNIVERSITY

Campus Network
Technologies Applied – Federated Research

- Protected Network
- Technologies
- Applied
- Federated
- Research

- Identified Data
- Deidentified Data
- Dockerized App
- Full VM
- Batch Process

Campus Network
Technologies Applied – Federated Research

Campus Network

Protected Network

Technologies Applied – Federated Research

Browser access With Shib+MFA

Full VM

Dockerized App

Protected Network

Identified Data

Deidentified Data

Batch Process

Campus Network

Duke UNIVERSITY
Technologies Applied – Federated Research

- Campus Network
- Protected Network
- Federated Research
- Browser access With Shib+MFA
- Students
- Researchers
- Collaborators
- Full VM
- Dockerized App
- Identified Data
- Deidentified Data
- Protected Network
- Batch Process

Duke University

Campus Network
Technologies Applied – Federated Research

- Browser access With Shib+MFA
- Federated IDP
- Full VM
- Dockerized App
- Identified Data
- Deidentified Data
- Protected Network
- Batch Process
Technologies Applied – Federated Research

- Campus Network
- Protected Network
- Federated Research
- Browser access with Shib+MFA
- Students
- Researchers
- Collaborators
- Federated IDP
- Full VM
- Dockerized App
- Identified Data
- Deidentified Data
- Protected Network

Duke University

VM Manage
CLOCKWORKS
ANTIKYThera

Browser access with Shib+MFA
SOFTWARE DEFINED NETWORKING AS A RESEARCH TOOL
Why Implement an SDN architecture?

• Traditional networks can inhibit transfers:
  – firewalls
  – intrusion prevention systems
  – backups/data transfers
  – Netflix/Twitch.tv

• SDN is designed for automated configuration

• Self-service configurable bypass network

• Researchers may need access to national backbones via Science DMZ (e.g. Open Science Grid)
SDN at Duke

Goal: How do we more efficiently move large data sets around the network?

Focused on the *network transition bottlenecks* rather than traffic in data center

- architecture & design
- secure the control plane
- authorization for routes
- testing for vulnerabilities
Traditional network switches:
- control functions in local firmware
- packet forwarding rules encoded in local config
- proprietary
SDN @ Duke – Network Bypass

Traditional network switches:
- control functions in local firmware
- packet forwarding rules encoded in local config
- proprietary

SDN switches:
- control functions decoupled from packet forwarding
- controller can view network “as a whole”
- open standards based (Openflow)
user requests network config changes

authorization/approvals

REST configuration commands

SDN controller
(Ryu REST router)

control plane

SDN switch
SDN switch
SDN switch

data plane
Switchboard (Controlling the Controller)

- Simplifies SDN controller/switch configuration and tracks changes
  - who is authorized to enable a bypass/link
  - status of requests
  - update SDN controller based on approved requests
  - rollback/restore SDN controller state
  - audit log of state of network configuration
Traditional Use Cases / Large Science Flows

• Can leverage AL2S – used for transfers to LHCOOne
  – Extensions of VLANs and Remote IP spaces

• Can leverage Regional SDN for collaboration in Research Triangle area

• Can leverage on campus SDN linkages to bypass core/firewalls

• Can leverage less congested 10G connections to the public internet
  – Expansion of capacity can be dedicated to SDSN
Use Case – Data Migrations

• Duke has a protected research data network
  – VMs provisioned
  – Firewalled / ACLs on routing between subnets
  – VPN for general access
  – Jump boxes/SSH/RDP servers

• Bottlenecks on getting data AND code into the network

• Use an SDN path to migrate encrypted traffic into the network
  – Could apply to both external and internal connections
  – One time or potentially recurring use (external protected feeds)
  – Build and deploy docker or singularity containers and automated deployments into protected networks
It’s all about flexibility

SOFTWARE DEFINED SCIENCE NETWORK (SDSN)
What is a Science DMZ?

The term Science DMZ refers to "...a portion of the network, built at or near the campus or laboratory's local network perimeter that is designed such that the equipment, configuration, and security policies are optimized for high-performance scientific applications rather than for general-purpose business systems or 'enterprise' computing."

Security of Model For a Medical Science DMZ

- Router acts as non-stateful packet-filter firewall
- Router manages list of trusted DTNs
- Flows approved by source and destination IP, time, protocol, and application.
- Permissions purged when flow is complete
- IDS (e.g., Bro) monitors for policy infractions and hostile activity
- perfSONAR for performance

SDN to SDSN

• the ability to control or monitor how routes are created
• the ability to control what nodes are added
• the ability to audit routes and traffic flows
• the ability to detect when something malicious enters or exits the network (can be done via SDN flows sent to an IDS)
Architecture overview (phase 1)
Architecture overview (phase 2)

- Change AL2S to Internet link and connect to Edge
- Connect Internet edge to SDN hub
- SDN Hub
- Science DMZ
- Add Data Transfer Node
- DTN Transfer Node 1
- File sharing protocol
- Internet
- Edge-gw1
- Edge-gw2
- IPS/FW
- Campus Core
- Physics (SDN Switch)
- Physics Storage
Architecture overview (phase 2)

Change AL2S to Internet link and connect to Edge

Connect Internet edge to SDN hub

SDN Hub

Add Data Transfer Node

DTN Transfer Node 1

File sharing protocol

Switchboard

Physics (SDN Switch)

SDN Bypass

Bro IDS

Scientific DMZ

Physics Storage

SDN Hub

Campus Core

IPS/FW

Edge-gw1

Edge-gw2

Internet
Architecture overview (final)
Perfsonar – Science DMZ Performance
Perfsonar – Campus Network Path
Conclusions

• We must be able to efficiently move large data sets between internal systems/networks or between organizations.

• How do we accomplish without sacrificing the security of sensitive data

• Interdisciplinary effort between IT (security, network, research compute) and research teams to design a solution that combines:
  – high-throughput transfers
  – detection of security issues
  – authorization for use of network with sensitive data
Using Grants to Drive Innovation

- CC-NIE (OCI-1246042) – Develop and deploy a production SDN network for the purposes of a science DMZ (led to creation of Switchboard).

- EAGER (CNS-1243315) – Deploy SDN to support the Duke ExoGENI project.

- CC*IE (ACI-1440588) – Data at the speed of trust. Federation of external IDs to access Protected Network resources.
What’s Next?

• Research Toolkits –
  – Broader group management
  – More services

• Test federated access with research partners
  – How do we check to see if our partners used MFA?
  – New standards emerging

• Duke Clinical Data (ePHI) Services

• What about security?
  – Logging, logging, logging (and analysis!)
  – Endpoint security – beyond policy and guidance
  – Detection and Response beyond honeypots and log analysis
  – Usability by researchers is VERY IMPORTANT
Want your own?

Public Github Repositories:

NoVNC in Docker:  https://github.com/mccahill/docker-novnc

- This is the generic template for making X Windows Linux environments available via Docker containers to web pages.

RStudio in Docker:  https://github.com/mccahill/docker-rstudio

Switchboard (SDN Controller):  https://github.com/mccahill/switchboard
Thank You!

Questions?

Please take our survey.
About the CTSC Webinar Series

To view presentations, join the discuss mailing list, or submit requests to present, visit:

http://trustedci.org/webinars

The next webinar is April 24th at 11am Eastern
Topic: IT Compliance
Speakers: Susan Ramsey & Anurag Shankar
We thank the National Science Foundation (grant 1547272) for supporting our work.

The views and conclusions contained herein are those of the author and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the NSF.