TRUSTED CI
THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

Trusted CI: The NSF Cybersecurity Center of Excellence
American Museum of Natural History (AMNH)
Engagement Report

June 18, 2019
For Public Distribution
Version 1.0

Terry Fleury[1], Jeannette Dopheide[2], John Zage[3]

[1] Project Lead, tfleury@illinois.edu
[2] Project collaborator, jdopheid@illinois.edu
[3] Project collaborator, jzage@illinois.edu

About Trusted CI

Trusted CI is funded by NSF's Office of Advanced Cyberinfrastructure as the NSF Cybersecurity Center of Excellence (CCoE). In this role, Trusted CI provides the NSF community a coherent understanding of cybersecurity's role in producing trustworthy science and the information and know-how required to achieve and maintain an effective cybersecurity program. Trusted CI achieves this mission through a combination of one-on-one engagements with NSF projects, training and best practices disseminated to the community through webinars, and the annual community-building NSF Cybersecurity Summit for Large Facilities and Cyberinfrastructure.

Acknowledgments

Trusted CI's engagements are inherently collaborative. The authors wish to thank the American Museum of Natural History (AMNH) team, including Michael Benedetto[4] and Ben Virgilio[5], for the collaborative effort that made this document possible. AMNH is supported by the National Science Foundation under Grant Number OAC-1827153[6].

This document is a product of Trusted CI. Trusted CI is supported by the National Science Foundation under Grant Number ACI-1547272[7]. For more information about Trusted CI please visit https://trustedci.org . Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Using & Citing This Work

Cite this work using the following information:
http://hdl.handle.net/2142/105406

[4] AMNH Director of IT and Deputy CIO, mbenedetto@amnh.org
[5] AMNH Senior Information Security Engineer, bvirgilio@amnh.org
[6] https://www.nsf.gov/awardsearch/showAward?AWD_ID=1827153
[7] https://www.nsf.gov/awardsearch/showAward?AWD_ID=1547272

Table of Contents

Executive Summary

The American Museum of Natural History (AMNH) conducts research and education activities spanning multiple branches of science. Scientific collaborations require high network capacity among scientific instruments, collaborators, and researchers. The National Science Foundation (NSF) awarded AMNH funds to develop and install a Science DMZ to enable high speed transfer of large datasets. Connections were deployed regionally via NYSERnet and nationally via Internet2. Additionally, AMNH's ADFS identity management system was federated with InCommon to give researchers access to Globus data transfer nodes (DTNs).

Trusted CI's engagement with AMNH initially focused on developing an information security program tailored to the new Science DMZ. This effort started by reviewing existing AMNH policies and procedures which might apply to the Science DMZ. After this initial examination, it was decided that the accelerated timeline for installation and configuration of both the Science DMZ and the ADFS federation with InCommon left little time for refinement of the few security policy documents deemed to be lacking. Instead, effort was focused on fine-tuning system configuration for the Science DMZ by consulting outside expertise from ESnet.

AMNH intends to document the processes of installation and configuration of their Science DMZ and the federation of their ADFS identity management system with InCommon. This documentation will give other similarly sized institutions a good starting point for installation of a Science DMZ or ADFS integration with InCommon.

1 Overview

1.1 Introduction

This document describes the Trusted CI - AMNH engagement which occurred January 2019 to June 2019. The goals of the engagement included creating a cybersecurity program tailored to the new Science DMZ using Trusted CI's Guide to Developing Cybersecurity Programs; addressing security concerns of the Science DMZ network by consulting with additional Science DMZ experts; assisting with the federation of AMNH's ADFS identity management system with InCommon; and drafting documentation about AMNH's Science DMZ installation for use by other institutions.

1.2 Background

AMNH is home to over 200 scientists who conduct scientific research in diverse fields such as astrophysics, biology, anthropology, geosciences, and genomics. Scientific collaborations require increasing network capacity among scientific instruments, collaborators, and researchers. Through the National Science Foundation's Campus Cyberinfrastructure (CC*) program, AMNH was awarded funds for a major data network upgrade that makes scientific data flows a priority. Improvements include high-speed "science-access" switches for research departments, a new Science DMZ complete with data transfer nodes (DTNs) implementing high-speed transfer via Globus, network performance monitoring with perfSONAR, connections with regional (NYSERNet) and national (Internet2) high-speed networks, deployment of federated login with InCommon, and education and training for scientists and the broader research and education community. The networking improvements to AMNH directly support research activities. AMNH's ability to move large data sets quickly between its campus and other sites across the nation and throughout the world is critical to the success of the Museum's research program.

2 Engagement Process

2.1 Determine Engagement Scope

Michael Benedetto applied for a Trusted CI engagement for AMNH in October 2018. He expressed interest in Trusted CI reviewing current plans and policies for securing the Science DMZ, and providing guidance in AMNH's effort in federating with InCommon. The experience setting up the Science DMZ would be documented by AMNH to create a "playbook" for use by other institutions.

While Trusted CI staff was up to the task of providing guidance on cybersecurity programs and federating with InCommon, Trusted CI reached out to ESnet for their expertise with the Science DMZ model. With these pieces in place, Trusted CI and AMNH agreed to the following engagement goals.

1. Using Trusted CI's Guide to Developing Cybersecurity Programs for NSF Science and Engineering Projects[8], create a cybersecurity program tailored to the new Science DMZ and associated high-speed connections to NYSERNet and Internet2. This process included a survey of AMNH information security policies and procedures currently in use for the campus and enterprise networks, referencing applicable policies so as to prevent duplication of effort.
2. Address security concerns of the Science DMZ network by bringing in additional personnel with Science DMZ expertise. This involved conference calls to give AMNH staff the opportunity to ask detailed questions about current and future specifications of the Science DMZ hardware and software configuration.
3. Assist with federating the current AMNH ADFS identity management system with InCommon. This involved configuring AMNH's ADFS to assert SAML attributes for consumption by external Research and Scholarship[9] (R&S) Service Providers. Trusted CI assisted AMNH with procedural issues for adding AMNH Identity Provider metadata to the InCommon Metadata aggregates, as well as issues of compliance with SIRTFI[10].
4. The culmination of the engagement will feature documentation of the tasks completed during the engagement, as well as a "playbook" for Science DMZs featuring AMNH as an exemplar that other scientific institutions can follow.
5. If time allows, develop training materials for AMNH technology staff to ensure dissemination of cybersecurity "best practices".

[8] https://trusteci.org/guide
[9] https://refeds.org/category/research-and-scholarship
[10] https://refeds.org/sirtfi

2.2 Background Preparation

2.2.1 Science DMZ

Preparation for the engagement included familiarization with Science DMZ core concepts. As the quantity of science data transferred and stored is increasing, the Science DMZ model[11] is becoming widely adopted by the scientific community. The Science DMZ is a dedicated portion of a network, usually located outside the network firewall but behind a high performance router or switch, designed to enable high throughput transfers that are becoming increasingly common with the large datasets being collected as part of scientific experiments.

2.2.2 Security

Security of a Science DMZ is typically approached in one of two ways: narrowly-focused with identification of risks and creation of mitigation strategies for those risks, or broadly-focused with risks mitigated by the implementation of controls. A firewall implementation is an approach to network security typically found in enterprise networks, but does not work well in a Science DMZ due to high latency of large data transfers. Instead a router Access Control List is typically used to provide security, filtering by IP and port number. In addition, network intrusion detection systems, host intrusion detection systems, and blackhole routing can all be used to mitigate attacks. Network Intrusion Detection systems such as Zeek or Snort can be used to trigger blackhole routing as well as to share intrusion data with other institutions.

2.3 Review Cybersecurity Program

Prior to reviewing AMNH cybersecurity policies and procedures documents, Trusted CI prepared a short summary of the Guide to Developing Cybersecurity Programs for NSF Science and Engineering Projects. This summary presented an overview of the guide that summarizes activities involved in creating a cybersecurity program. This summary helped quickly identify an area that needed work in AMNH's cybersecurity program, namely the information asset inventory list.

Trusted CI then performed a review of AMNH's existing set of policies and procedures documents and attempted to map them to those listed in the Master Information Security Policy & Procedures Template[12]. The results of this evaluation can be found in Appendix A.

[11] https://fasterdata.es.net/science-dmz/

[12] https://trustedci.org/guide/docs/MISPP

2.4 Outside Science DMZ Expertise

On February 5th, a conference call was held with ESNet, AMNH, and extended Trusted CI staff to discuss AMNH's proposed Science DMZ configuration. While the majority of the call was dedicated to an overview of the technical aspects of the Science DMZ, the following suggestions were given to AMNH.

- A 40Gb/s DTN to a 10Gb/s campus network can cause bottlenecks in campus networks. That step-down occurs at the Science DMZ core, meaning there will be potential packet loss and congestion occuring on switches. This trickles to other systems that connect via that device. One suggestion is to configure a 10Gb/s DTN with the capability to add more 10Gb/s DTNs when needed. This offers flexibility in balancing user load, and 10Gb/s DTNs can use spinning disks rather than SSDs, resulting in cheaper hardware costs in the longer run.
- IDS (intrusion detection) with blackhole routing is fairly fast (seconds), but is slower than IPS (intrusion prevention) (milliseconds). If the network is doing default "deny", it can cause problems with async traffic, e.g., black holes in the network. ESnet suggested to dedicate resources that would allow all traffic on the /24 subnet, and use host-based security, out-of-line IDS (e.g., Zeek), a signature-based box, and blackhole routing against the router. This needs a well-tuned IDS with a lot of care and feeding. ESnet suggested NCSA's open source blackhole routing software.
- While AMNH is currently using Scrutinizer, ESnet suggested looking at ElastiFlow[13], a bundle addon to Elastic Stack.
- If implementing a default out-flow "allow" scheme, make sure that you have visibility into data so you know what is going with the flow to prevent potential issues with BYOD (bring your own device).
- Get some form of netflow/sflow/etc. monitoring system that gives visibility into network traffic. One suggestion is inMon[14]; another is NetSage[15] which was developed jointly with IU. ESnet offered assistance installing it.
- Identify 1 or 2 pilot users to see how they would utilize the new infrastructure.

On May 14th, there was a second conference call to discuss further updates to the network diagrams, including the specific hardware used for each section of the network. These diagrams, referenced in Appendix B (with confidential information redacted), did not include the DTNs as those were still on order. The suggestions from the previous call on the transfer speeds of the networks (i.e., the step down from 40Gb/s to 10Gb/s) were taken into account by choosing a 10Gb/s DTN. One major issue pointed out by ESnet was that managing BGP preferences in the network can be a large issue. This lead to the discussion of the BGP design document. Another observation by Trusted CI staff was that the network maps are great but are only as useful as they are accurate. Thus, it is prudent to schedule regular quarterly meetings with staff to verify and update them. Finally, ESnet mentioned that storage is

[13] https://github.com/robcowart/elastiflow
[14] https://inmon.com
[15] http://www.netsage.global

becoming an increasing problem in Science DMZs. If local storage solutions are not offered, users of the network will seek outside solutions such as Google or Dropbox.

2.5 InCommon Federation

Trusted CI provided basic guidelines and links for federating the AMNH ADFS identity management system with InCommon. The steps[16] included signing the agreement with InCommon, paying the appropriate fees, and registering executive and administrator contacts. AMNH then verified administrative access to the InCommon Federation Manager[17]. These steps were confirmed to be complete by the end of January 2019.

Trusted CI suggested AMNH review the Baseline Expectations for Trust in Federation[18], and then create publicly accessible links to elements to be included in metadata[19].

In order for ADFS to assert research and scholarship (R&S) SAML attributes, Trusted CI suggested using ADFSToolkit[20]. This proved to be more difficult than available documentation led to believe as InCommon metadata[21] contains several thousand Service Provider (SP) entries. AMNH discovered that Microsoft suggested the use of a separate SQL server to manage the list of SPs[22]. AMNH intends to document this process and offer it to CANARIE for inclusion in their ADFSToolkit documentation[23].

After configuring a test instance of their ADFS server with ADFSToolkit, AMNH registered their Identity Provider (IdP) with InCommon and confirmed that R&S attributes are being asserted by the IdP. This configuration has been deployed to AMNH production ADFS servers as of June 2019.

[16] https://spaces.at.internet2.edu/x/zgKMBw
[17] https://service1.internet2.edu/siteadmin/
[18] https://www.incommon.org/federation/baseline-expectations-for-trust-in-federation/
[19] https://spaces.at.internet2.edu/x/-ZWAAQ
[20] https://github.com/fedtools/adfstoolkit
[21] https://spaces.at.internet2.edu/display/InCFederation/Metadata+Aggregates
[22] https://docs.microsoft.com/en-us/windows-server/identity/ad-fs/design/federation-server-farm-using-sql-server
[23] https://www.canarie.ca/identity/support/fim-tools/

3 Recommendations

3.1 Cybersecurity Policy List Completeness

The review of AMNH's policies and procedures documents led to discovery of the following missing/incomplete components of a full cybersecurity program. These have been prioritized by ease and importance of completion. Links to the Trusted CI Guide templates are provided where applicable.

1. Network Security Policy
2. Access Control Policy[24]
3. Information Asset Inventory List[25]
4. Asset Management Policy[26]
5. Privacy Policy
6. Disaster Recovery Policy[27]
7. Partner and Subcontractor Policy
8. Personel Exit Checklist[28]

Among these missing components, the Information Asset Inventory List is required for the completion of a Risk Assessment, which would be a suitable next step in the development of AMNH's cybersecurity program.

[24] https://trustedci.org/guide/docs/ACP
[25] https://trustedci.org/guide/docs/IAI
[26] https://trustedci.org/guide/docs/AMP
[27] https://trustedci.org/guide/docs/recovery
[28] https://trustedci.org/guide/docs/exitlist

4 Broader Impacts

AMNH is creating a "playbook" for how they selected, installed, and configured components of their Science DMZ. The intent is to offer this documentation to ESnet for hosting on their site[29] for consumption by other institutions of similar size and makeup.

AMNH also intends to document their installation of ADFSToolkit including the additional step of installing a backend SQL server to manage InCommon metadata. This documentation will be provided to the community through CANARIE and announced on the Trusted CI Blog[30] to reach a broad audience.

In addition to this documentation, a brief summary of the Trusted CI Guide to Cybersecurity Programs was created to give AMNH an overview of the purpose of the guide as well as to set expectations for creating a comprehensive set of policies for their cybersecurity program. This document will be made available on the Trusted CI site at a future date.

5 Version History

2019-06-18 Version 1.0 - Initial version for AMNH review

[29] http://es.net/science-engagement/knowledge-base/case-studies/science-dmz-case-studies/

[30] https://blog.trustedci.org/

Appendices

Appendix A - Policy and Procedure Document Review

This list of policy and procedure documents recommended by the Trusted CI Guide[31] can be found in the Master Information Security Policy & Procedure Template[32] in Section 5. Existing AMNH policy and procedure documents were reviewed and mapped to this list. A green check (✅) indicates that an AMNH document sufficiently matched. A yellow triangle (⚠) indicates that Trusted CI could not find an adequate match to existing AMNH documents. A red question mark (❓) indicates that additional investigation by AMNH staff is needed to determine the availability of the policy. There may be AMNH internal documents which suffice.

For the policies which lacked a matching AMNH document, a number is given to indicate priority as suggested by Trusted CI.

- ✅ Acceptable Use Policy - A set of rules that a user must agree to follow in order to be provided with access to a network and/or resources. Reduces liability and acts as a reference for enforcement of policy. **This is done in the first section "Acceptable and Secure Use of AMNH Computing Resources" in the "AMNH Data Security Policy (COMBINED)" document, and expanded in section "8 Acceptable Use of AMNH Computing Resources".**
- ⚠ *2* Access Control Policy - Defines the resources being protected and the rules that control access to them.
- ⚠ *4* Asset Management Policy - Requirements for managing capital equipment including: inventory, licensing information, maintenance, and protection of hardware and software assets. Note: this is related to the Information Asset Inventory document.
- ⚠ *6* Disaster Recovery Policy - Contains policies and procedures for dealing with various types of disasters that can affect the organization.
- ⚠ ❓ *8* Personnel Exit Checklist - Form to be completed at the end of employment that addresses revoking access to compute resources, physical spaces, and the return of organizational assets. **This may partially exist in various AMNH policy documents, but it would be useful to have a checklist referencing these existing documents.**
- ✅ Incident Response Procedures - A pre-defined organized approach to addressing and managing a security incident. **This is done in the "Incident Response Policy" document. Note: current policy does not include testing of IR.**
- ⚠ *3* Information Asset Inventory - A document that tracks assets controlled or in use by the organization. The list of assets is often used as the starting point for a Risk Assessment.

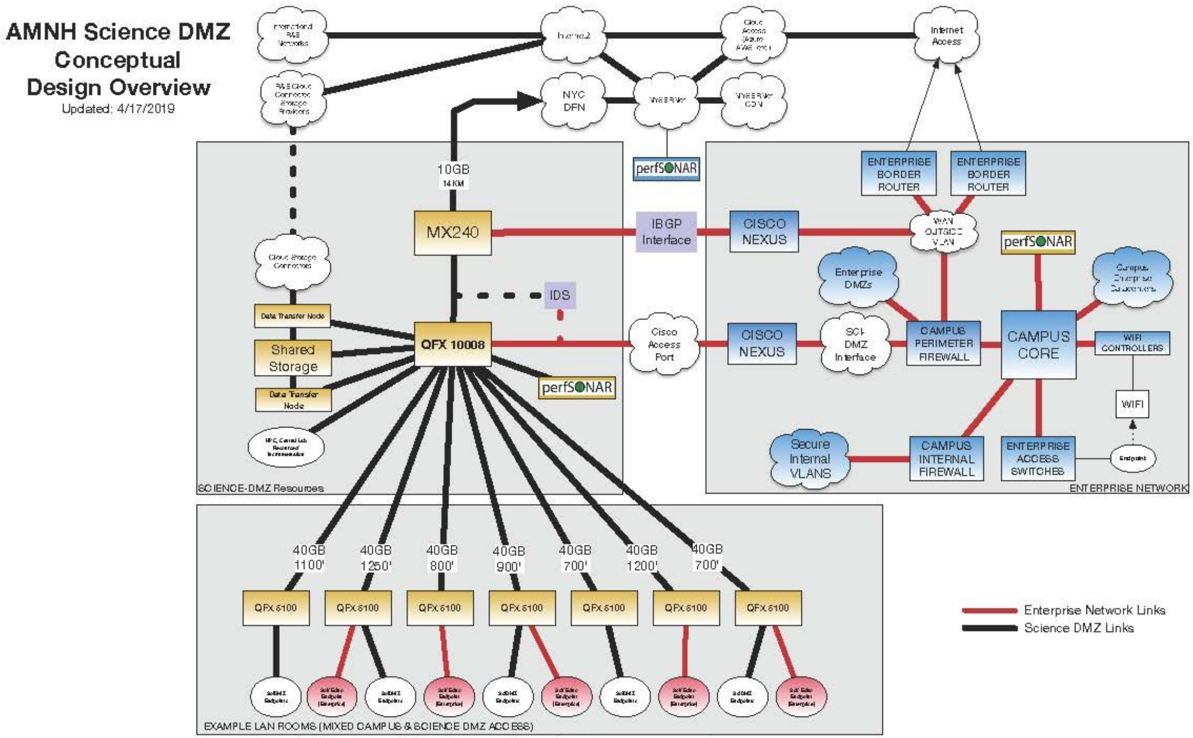
[31] https://trustedci.org/guide

[32] https://trustedci.org/guide/docs/MISPP

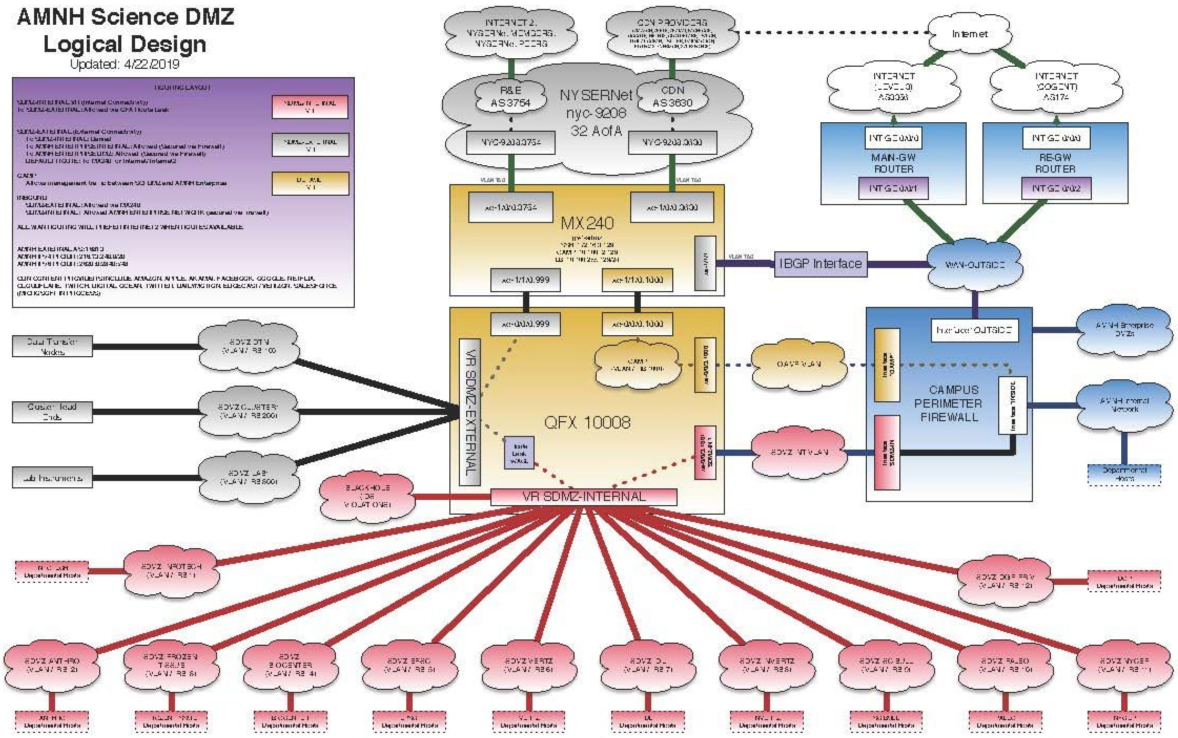
- ✅ Information Classification Policy - Used to ensure consistency in classification and protection of data. **This is done in section "5 Information Classification and Related User Responsibilities" in the "AMNH Data Security Policy (COMBINED)" document.**
- ✅ Mobile Computing Policy - Establish standards for the use of mobile computing and storage devices. **This is done in the "Mobile Devices" section of the "AMNH Data Security Policy (COMBINED)" document.**
- ⚠ *1* Network Security Policy - Outlines the rules for network access, determines how policies are enforced, and lays out some of the basic architecture of the company security/network security environment.
- ⚠ *7* Partner and Subcontractor Policy - An agreement containing a set of rules and expectations to be used between two parties seeking access to the other's network, data, or resources.
- ✅ Password Policy - A set of rules designed to establish security requirements for passwords and password management. **This is done in the "Password Policy" section of the "AMNH Data Security Policy (COMBINED)" document.**
- ✅ Physical [and Environmental] Security Policy - Details measures taken to protect systems, buildings, and related supporting infrastructure against threats associated with their physical environment. **According to Mike, this policy is maintained by Museum Security and Safety Department.**
- ✅❓ *5* Privacy Policy - A statement that discloses the ways a party gathers, uses, discloses, and manages a customer's or client's data. **This is done in section "4 Data Ownership, Access, and Privacy" in the "AMNH Data Security Policy (COMBINED)" document. However, a separate (additional) policy is being created for users accessing Science DMZ / HPC resources.**
- ✅ Remote Access Policy - Outlines and defines acceptable methods of remotely connecting to the internal network. **This is done in section "10 Remote Access to AMNH Data and Networks" in the "AMNH Data Security Policy (COMBINED)" document.**
- ✅ Training and Awareness Policy - Outlines an organization's strategy for educating employees and communicating policies and procedures for working with information technology (IT). **Referenced in section "6 Information Security and Related User Responsibilities" of the "AMNH Data Security Policy (COMBINED)" document: "Details on the AMNH Security Awareness Training Program can be found at http://it.internal.amnh.org/training ."**

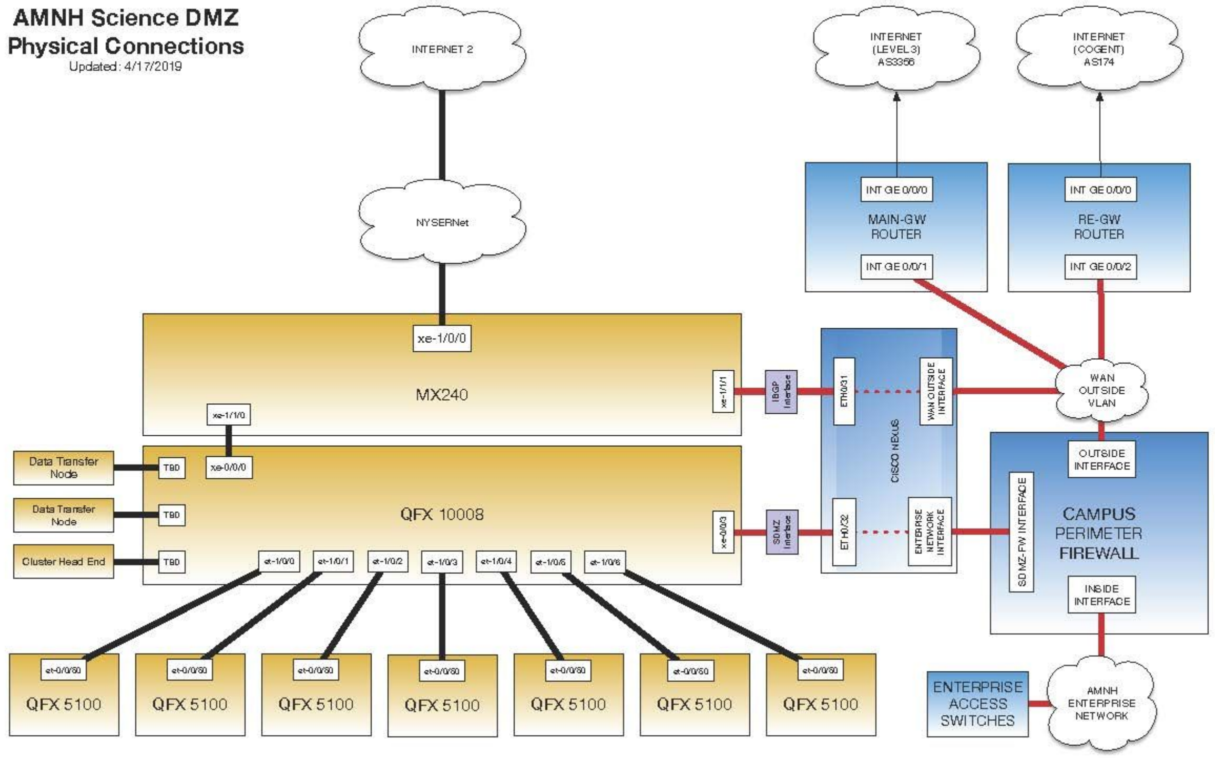
Appendix B - AMNH Science DMZ Network Diagrams

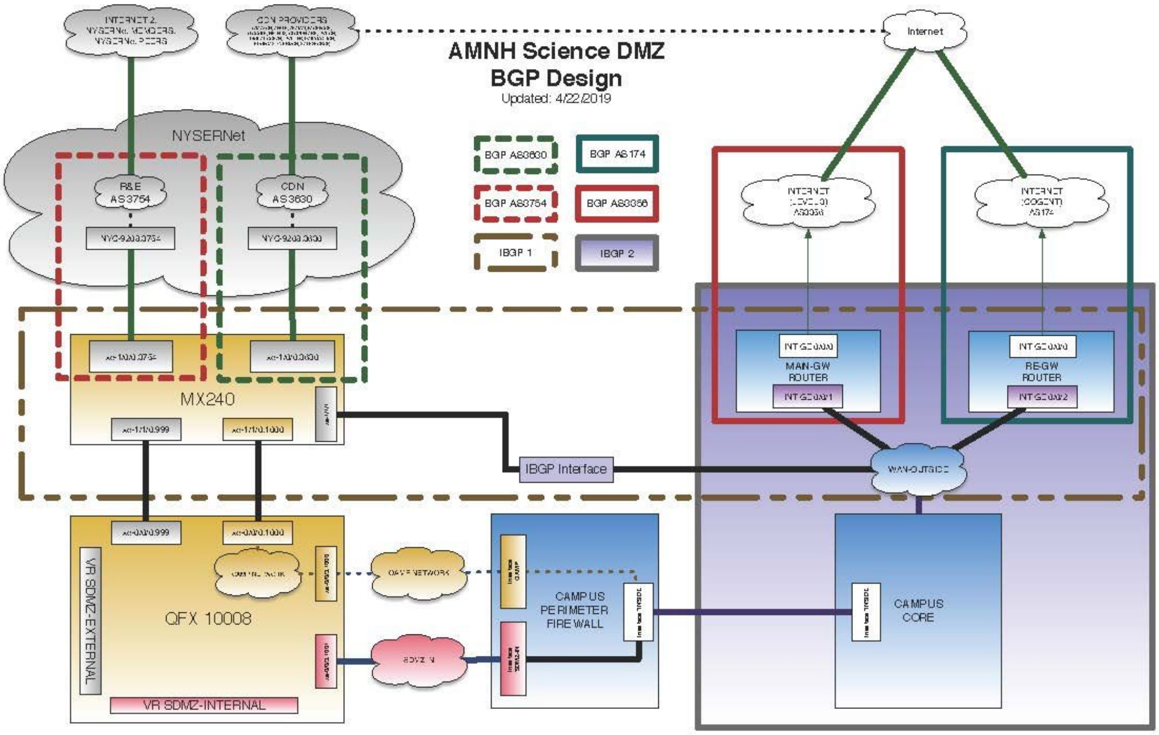
These diagrams illustrate AMNH's proposed Science DMZ and related networks as of May 2019. Note that sensitive information has been removed to make the diagrams suitable for public distribution.

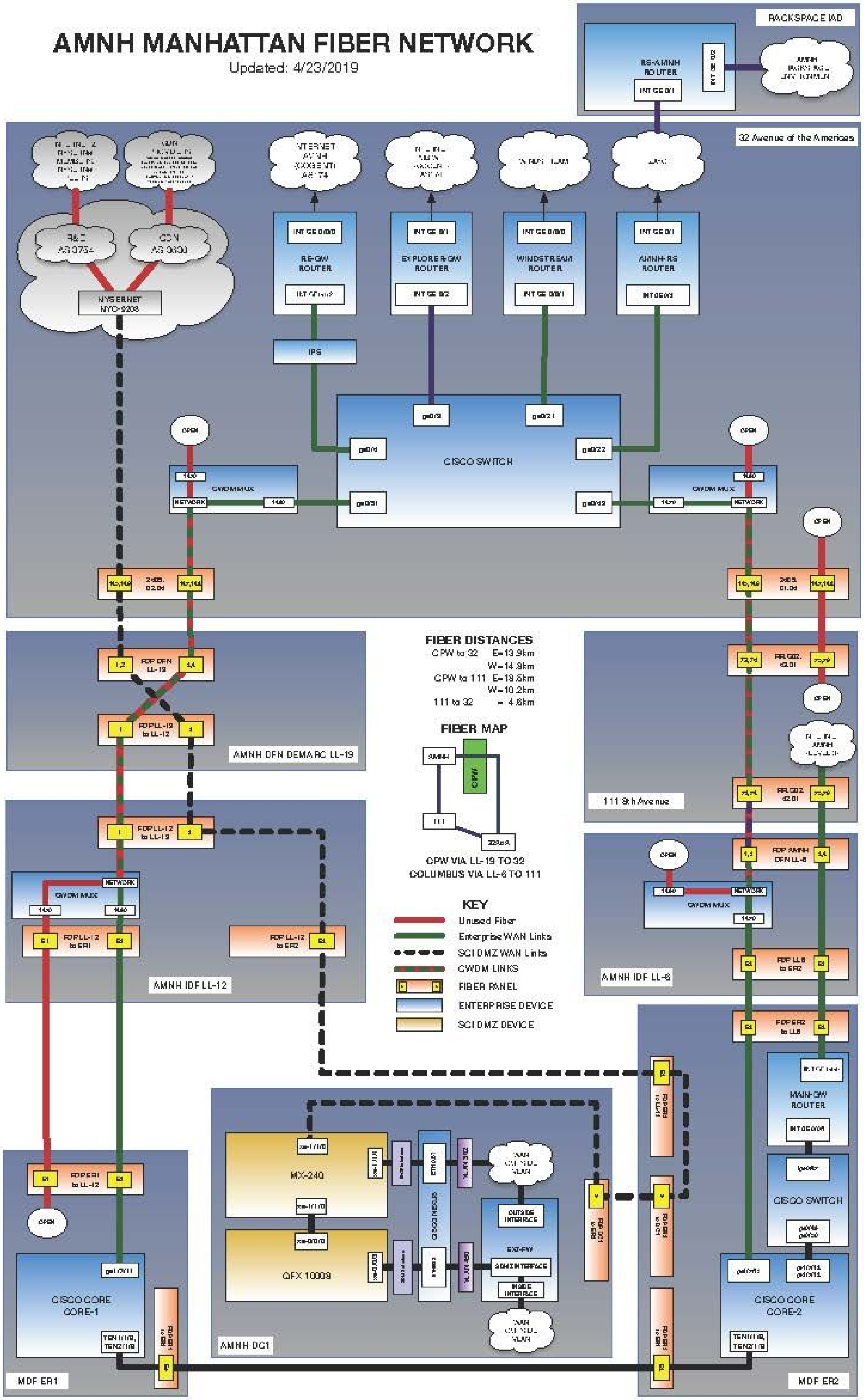


Items shown in blue are existing devices. Items in gold are proposed.


AMNH Science DMZ
Logical Design
Updated: 4/22/2019
INTERNET 2, NYSERNet MEMBERS, NYSERNet PEERS
CDN PROVIDERS
Internet
R&E AS3754
NYSERNet nyc-9208 32 AofA
CDN AS3630
NYC-9208.3754
NYC-9208.3630
INTERNET (LEVEL3) AS3356
INTERNET (COGENT) AS174
MAIN-GW ROUTER
RE-GW ROUTER
MX240
xe-1/0/0.3754
xe-1/0/0.3630
xe-1/1/0.999
xe-1/1/0.1000
VLAN TAG
IBGP Interface
WAN-OUTSIDE
xe-0/0/0.999
xe-0/0/0.1000
VR SDMZ-EXTERNAL
QFX 10008
VR SDMZ-INTERNAL
CAMP VLAN
Interface OUTSIDE
CAMPUS PERIMETER FIREWALL
AMNH Enterprise DMZs
AMNH Internal Network
Departmental Hosts
SDMZ INT VLAN
Data Transfer Nodes
SDMZ DTN
Cluster Head Ends
SDMZ CLUSTER
Lab Instruments
SDMZ LAB
BLACKHOLE (FOR VIOLATIONS)
SDMZ INROTECH
SDMZ DGP PRIV
SDMZ ANTHRO
SDMZ FROZEN TISSUE
SDMZ BIOCENTER
SDMZ ERSC
SDMZ VERTZ
SDMZ IOL
SDMZ INVERTZ
SDMZ SCIBULL
SDMZ PALEO
SDMZ NYCEP
Departmental Hosts


AMNH Science DMZ
Physical Connections
Updated: 4/17/2019
INTERNET 2
NYSERNet
INTERNET
(LEVEL 3)
AS3356
INTERNET
(COGENT)
AS174
INT GE 0/0/0
MAIN-GW
ROUTER
INT GE 0/0/1
INT GE 0/0/0
RE-GW
ROUTER
INT GE 0/0/2
xe-1/0/0
MX240
xe-1/1/0
xe-1/1/1
IBGP
Interface
ETH0/31
WAN OUTSIDE
INTERFACE
CISCO NEXUS
ETH0/32
ENTERPRISE
NETWORK
INTERFACE
SDMZ
Interface
WAN
OUTSIDE
VLAN
OUTSIDE
INTERFACE
SDMZ-FW INTERFACE
CAMPUS
PERIMETER
FIREWALL
INSIDE
INTERFACE
Data Transfer
Node
TBD
xe-0/0/0
Data Transfer
Node
TBD
QFX 10008
Cluster Head End
TBD
xe-0/0/3
et-1/0/0
et-1/0/1
et-1/0/2
et-1/0/3
et-1/0/4
et-1/0/5
et-1/0/6
et-0/0/50
QFX 5100
et-0/0/50
QFX 5100
et-0/0/50
QFX 5100
et-0/0/50
QFX 5100
et-0/0/50
QFX 5100
et-0/0/50
QFX 5100
et-0/0/50
QFX 5100
ENTERPRISE
ACCESS
SWITCHES
AMNH
ENTERPRISE
NETWORK


AMNH Science DMZ
BGP Design
Updated: 4/22/2019
INTERNET 2,
NYSERNet MEMBERS,
NYSERNet PEERS
CDN PROVIDERS
Internet
NYSERNet
R&E
AS3754
CDN
AS3630
NYC-9208.3754
NYC-9208.3630
BGP AS3630
BGP AS174
BGP AS3754
BGP AS3356
IBGP 1
IBGP 2
INTERNET
(LEVEL3)
AS3356
INTERNET
(COGENT)
AS174
xe-1/0/0.3754
xe-1/0/0.3630
MX240
xe-1/1/0.999
xe-1/1/0.1000
INT GE0/0/0
MAIN-GW
ROUTER
INT GE0/0/1
INT GE0/0/0
RE-GW
ROUTER
INT GE0/0/2
IBGP Interface
WAN-OUTSIDE
xe-0/0/0.999
xe-0/0/0.1000
VR SDMZ-EXTERNAL
QFX 10008
VR SDMZ-INTERNAL
OAVE NETWORK
SDMZ-IN
CAMPUS
PERIMETER
FIREWALL
CAMPUS
CORE



AMNH MANHATTAN FIBER NETWORK

Updated: 4/23/2019