Impact of Cyberinfrastructure Investments on Research at the University of Maryland

September 25, 2018

The opinions expressed in this talk are the speakers, and do not necessary reflect those of UMD or the state of MD.
Notable Events in UMD Networking

1984 – TCP/IP used campus wide for networking (first campus to do so)
1987 – UMD Starts providing Root DNS Services (D Root)
1988 – FIX East Established at College Park (connects NSFnet, Esnet, etc.)
1998 – First Wifi Offered (in McKeldin Library)
1999 – MAX Established
2005 – Wifi expanded to cover most academic buildings
2009 – Campus phones conversion to VOIP started (& major network refresh)
2009 – Wifi Parking meters arrive on campus
2012 – MAX Core network upgraded to 100 Gbps
2015 – Bluecrab HPC Cluster (joint with JHU)
2018 – Second Generation VOIP phone system installed
UMD Campus Network Today Snapshot

Serves over 100 buildings
   695 network closets
   955 Switches and Routers
   Over 85,000 Ethernet ports: This is also power, theoretically up to 1.1MW

Core Networking
   Provided via two primary network hubs (fully redundant)
   Includes 103 firewalls

Wide Area Networking
   two 10 Gps general purpose links to commodity Internet
   one 10 Gps link to Comcast (streaming video to dorms)
   two 100 Gps links to MAX for research traffic (Internet2, Esnet)
   specialty links: Big10 network, State finical network, FIOS
Wireless Utilization - Fall Semester 2018

Wireless Client Authentications

Client Count

Time

Aug 19  Aug 25  Aug 31  Sep 5  Sep 11
There are typically 400+ guests on the network
One Network is Serving Many Needs

Modes of Communication:
- Computer Data
- Voice
- Video

Application Domains
- Instructional (classroom) needs (FERPA)
- Research Computing (NIST 800 171, FISMA)
- Administrative Computing (FERPA, HIPPA)
- Public Facing Web Sites (GDPR, ADA)
- Video Surveillance (Chain of Custody)
- Emergency Calls (E911)
- Credit Card Processing (PCI)
- Entertainment (DMCA)
UMD Data Center Summary

DIT Currently uses six data centers

- Primary Data Center
- Old Secondary Data Center (on campus)
- Research Data Center
- Rivertech Data Center (leased)
- New Secondary Data Center (in a for-profit co-location facility)
- MARC Modular data center (shared interest in with JHU)

Many other data centers exist on campus

- some medium quality raised floors with building UPS
- others small racks near/under students desks or in closets

3.5 KM from core campus
Trends in Computing

Administrative Computing: services are moving to the cloud
  Cloud hosted (AWS) operated by UMD
    Web Site
    Front end to student systems (UMEG)
    Data Analysis (SAS)
  Application As Service (aka you don’t pick your data center)
    LMS (Canvas)
    Email (Google)
    Student Degree Audits (u.achieve)

Research Computing: much slower adoption of cloud
  Investigators will take free cloud service units when offered
  Generally slow to buy cloud services with research dollars
What is Holding Back Cloud for Research Computing?

- Some resistance from funding agencies
- Investigators who want to run computers:
  - Have always done it that way
  - Can employ students & postdocs to run the computers
- Storage and Egress charges from cloud providers
- Purchasing: It’s hard to buy cloud time
- HPC model that separately allocates people/compute time
- Big Issue: Indirect Cost Charging Model
  - Data Center costs are covered in indirect costs
  - Computer purchases do not pay indirect costs
  - Buying cloud time is charged indirect costs
- Networking Challenges to move data
Software as a service is starting to dominate
  AI as service
  Data analytics as a service
  ... as a service

Software as service companies come and go quickly
  It’s a young and rapidly evolving market
  Academic research needs are a small market in this space
  Need to get researchers on/off platforms rapidly (every 3-4 years)

Will still make sense (at least mid term) for local compute resources
  Latency needs
  Costs of storage
  ROI for highly used systems
Public Cloud Research Data Sets are Growing

Genomic Data

Government Data Sets (IRS 990, GOES, U.S. Census American Community Survey)

Social Media data sets that originate in the cloud
New Cyber Infrastructure for Research Support

University of Maryland Campus
Software Defined Network Perimeter
Mid-Atlantic Crossroads Regional Network
Regional Software Defined Science DMZ (SD-SDMZ)
Software Defined Network Perimeter

One Campus Infrastructure
Firewall bypass for approved flows

Possible Future Functions:
- SDN Controller API Driven Service
- Mirror based monitoring
Regional Software Defined Science DMZ (SD-SDMZ)

Hybrid services with turnkey, on-demand, integration of private/public clouds, cyberinfrastructure, scientific resources, and high performance networks

- Sharing SD-SDMZ regionally allows for expertise development
Conclusions

Multiple varied traffic can and should co-exist on one network

Research Data will continue to need to flow freely:
  - Campus to/from near by data centers
  - Campus to/from commercial cloud providers
  - Between cloud providers
  - Between Campuses

Continued cyber Infrastructure investments are required in:
  - Campus Clusters
  - Clouds (both system as service and application as Service)
  - Networks including to support multi-cloud needs
Q&A