June 2017 Harvard IT Summit

June 2017 Harvard IT Summit

- Keynote by Nicco Mele, Director Shorenstein Center on Media, Politics, and Public Policy

• 3D printing is awesome!

- He has been printing lego blocks, t-rex's, even crocs!
- Then, someone released a blueprint to print a functional handgun

• Thesis technology moves "power" of the institution into the hands of users

• Challenging "at scale"

• People in power often taken by surprise by change

- World order fragile to technological disruption

• Examples of power moving into hands of the users

- Personal computer
- Apple
- Microsoft Windows
- Internet
- Smartphones

• It becomes about:

- The individual
- Intimate
- Intention I search for a white mini-van and google knows that I want to buy a mini-van

• Internet's effect on presidential elections

- Gary Hart proved he might be able to beat Mondale in the primary in 1984, but took too long to get checks/donations to fund his campaign
- In 2008, Barack Obama proved he might be able to beat Clinton in the primary and had a website to immediately get donations and fund his campaign
- Tea Party

- Donald Trump and tweeting unconstrained by norms and institutions

• Major milestones in US history mediated by news/journalism

- JFK shot

- Watergate

- Berlin Wall

- 911

- Common experience created by journalism

• Compare these experiences to Bin Laden getting killed

- Very first announcement a tweet from Keith Urbhan, goes viral, changes White House strategy on announcement

- Happened w/out news media

• Journalism is collapsing and fake news is winning

• What do we do about it?

• Re-think our relationship

- Authority and expertise not taken at face value

- Self-funded solutions to municipal problems normally covered by taxes

• Speed and intensity of response and mutual support after Boston Marathon bombing

- Donations

- Google spreadsheet for temporary housing

• Internet also tried to find the 2 bombers

- When people get angry, they are going to use the Internet as an outlet for their anger

• Demand technical literacy from our leaders and each other

• Fight "nerd disease" technical elitism

- "genius bar"

- Pebkac

• Demand accountability One

- 10:50am Concurrent Session 1 OpenNebula at FAS RC Sever 102

• Presented by:

- Justin Riley, John noss, Wess Dillingham
- Dr Ignacio Llorente Project Director at OpenNebula and visiting scholar at Harvard

• FASRC VM's are provisioned similar to HUIT ticket request, request reviewe,d if viable, deploy it, make sure it's up and running

• Why Private Cloud?

- Workload requirements
- Performance, latency
- Security, confidentiality, privacy (isolated infrastructure)
- Leverage existing infrastructrue and know-how
- Storage, networking, config mgmt
- Cost analysis
- Variability of demand (x10 cheaper storage)

• The old system based on KVM and GlusterFS

• New System

- OpenNebula + KVM + Ceph Cluster
- Multi-datacenter hypervisor and storage
- Automated VM scheduling
- Image based OS installs

• Why OpenNebula?

- Light and simple
- Robust production ready, highly scalable
- Flexible
- Powerful innovative functionality for private/hybrid cloud and DC virtualization
- 1 week for OpenNebula versus months to really get OpenStack working

• Why Ceph?

- Killer storage solution for private cloud
- Software defined storage
- Acquired and maintained by RedHat
- File, block, and object store
- Can backend into S3 and Swift
- Can scale to exabyte

- RC primarily uses it for block storage for VM hosting storage

- Architecture Review of the full stack

  - They're using OpenNebula to also start provisioning VM's in AWS and Azure
  - Single dashboard for mgmt

- Integration and Mgmt

  - Puppet
  - Using Puppet to do operations and configuration mgmt for OpenNebula
  - Can spin up multiple clusters w/ their puppet scripts
  - Backups
  - Production Ceph Cluster daily disk snapshots
  - Backup Ceph Cluster
  - Deltas between today and yesterday transferred from prod ceph cluster to backup ceph cluster
  - More CPU intensive work done on the backup cluster
  - QCOW2 Deep Backups
  - Export RBD devices from backup cluster to qcow2 file formate and keep in separate filesystem in separate datacenter
  - Monitoring
  - Ceph dashboard
  - Web dashboard Grafana
  - Nagios alerting
  - VM Leaderboard which VM's are hammering disk?
  - OpenNebula hooks/plugins
  - IP address check to ensure
  - OneDNS
  - Provide dynamic DNS resolution for VM's on OpenNebula
  - Dynamic DNS generation for all VM's based on sanitized name
  - Automatic forward and reverse records per VM

- Operations

  - They use an internally developed tool called Cangallo to generate clean VM images similar to dockerfile for containers
  - Simple YAML file format

- They use a single git repo for all our base and derived images

• NESE will be based on Ceph

- 11:45am Lunch
- 1:00pm Concurrent Session 2 JupyterHub on AWS – Demba and Farras Sadek

• Bridging the gaps between EE and CS education

- EE data set and data collection
- CS Data processing and IoT

• Goal integration of theory and computation and design of seamless coding interface

• Jupyter Notebooks on AWS

- Minimize lead time to get a useful programming environment up and running a "sandbox"

• JupyterNotebook

- A web application for coding, documentation, simulation

• JupyterHub

- A way to give a Notebook ervier to each person or group of people (PaaS)
- Users can log in w/out any setup and start usign it

• Hub functions

- Authentication, authorization,

Manages AWS resources

- 2:20pm Concurrent Session 3 Video Delivery and Retention at Harvard

• Video at HBS Kaltura based

- Use cases both academic (course videos, etc.) and administrative (marketing, external relations, initiatives)
- Academic Solution Design
- KMC Media Console
- Classroom Delivery KMS (MediaSpace)
- Playlist
- LMS
- CCTL KMS

- Sharepoint
- Administrative Solution Design
- KMC
- KMS
- Sharepoint
- Websites
- Links
- Upfront metadata design and video migration
- Canvas and Kaltura were rolled out at the same time to ease migration disruption
- Extensive time spent w/ faculty on tool requirements gathering and training

- Video Retention and Metadata Recommended Guidelines
  - Video storage workgroup
  - Tasked w/ developing a centrally managed service to manage, distribute, and retain Harvard's video content
  - Defining standards and guidelines metadata and retention in particular
  - Retention and Metadata Sub-groups
  - Recommendations to serve as basis for wider retention and metadata efforts moving forward

- Questions:
  - For donation of lecture capture, Intellectual property owner can choose to donate, which bypasses the default retention policy. What about department head? (lectures that only belatedly recognized for historical significance)
  - For metadata and digital asset management, how does versioning play into asset mgmt?

- Video Metadata Recommendations minimum required fields
  - Title
  - Description
  - Ingest data
  - Creation date
  - GRS Category
  - Retention Requirements
  - Organizational Owner
  - Usage Class

- Filename/path
- Security Level

#learning/conferences