

HEPiX 2018 Highlights

Subjective and incomplete selection


Jakub T. Moscicki, IT-ST meeting

[https://indico.cern.ch/event/676324/timetable/
#20180514.detailed](https://indico.cern.ch/event/676324/timetable/#20180514.detailed)

“Our” disk storage mentions

- **CVMFS** is “everywhere”
- **Xrootd** is “almost everywhere”

Operating a large scale distributed XRootd cache across Caltech and UCSD

Edgar Fajardo Hernandez 

Chamberlin Hall, University of Wisconsin-Madison


11:10 - 11:30

- **EOS** reported by IT-ST, IT-DB, IHEP
- **CERNBox**: BNLBox, IHEPBox and “*original CERNBox*”
- **CEPH** reported by IT-CM, BNL, RAL
 - IT-CM: CephFS in production since Q3 at CERN
- The (majority) rest is: dCache, GPFS, Lustre, ...



“My” Presentation

Next generation of large-scale storage services at CERN

Jakub Moscicki 

Chamberlin Hall, University of Wisconsin-Madison

09:40 - 10:00

“Science Box”



One-Click Demo Deployment

- Single-box installation via [docker-compose](#)
- No configuration required
- Download and run services in 15 minutes

<https://github.com/cernbox/uboxed>

Production-oriented Deployment

- Container orchestration with [Kubernetes](#)
- Scale-out storage and computing
- Tolerant to node failure for high-availability

<https://github.com/cernbox/kuboxed>

The screenshot shows the EOS Community Portal at <https://eos-community.web.cern.ch/>. The main heading is "EOS Open Storage" with a sub-heading "CERN storage technology used at the Large Hadron Collider (LHC)". Below this are three buttons: "EOS.COMMUNITY", "2018 WORKSHOP", and "LATEST". A red box highlights the "EOS.COMMUNITY" button, and a red arrow points from it to a sidebar menu on the right. The sidebar menu has a search bar and a list of items with columns for "Category", "Date", "Status", "Views", and "Activity".

Evolve from CERN Open Source to Community Open Source

The banner features the CS³ logo and the text "Cloud Services for Synchronisation and Sharing" with the dates "28 - 30 January 2019, Roma". Below this, it lists "Previous Workshops" with locations: "Krakow 2018 - Amsterdam 2017 - Zurich 2016 - Geneva 2014". The background is a night view of the Colosseum in Rome.

<http://cs3.infn.it>

THE
WALKING DEAD





OpenAFS Release Team report

Chamberlin Hall, University of Wisconsin-Madison

Mr Michael Meffie



16:25 - 16:45

The OpenAFS Foundation

Chamberlin Hall, University of Wisconsin-Madison



The OpenAFS Foundation, Inc.

Championing OpenAFS technology,
its developers, and its users

Dr Margarete Ziemer



16:45 - 17:05

AURISTOR®

AFS Update: Spring 2018

Chamberlin Hall, University of Wisconsin-Madison



OpenAFS
Technical Support



Migrate to the
AuriStor File System

Mr Jeffrey Altman



10:00 - 10:20

AFS and Linux Containers

Chamberlin Hall, University of Wisconsin-Madison

Mr Jeffrey Altman



11:30 - 11:50



OpenAFS 1.6.x stable series

(5 / year)

- 2018-04-17 1.6.22.3
- 2018-02-01 1.6.22.2
- 2017-12-21 1.6.22.1
- 2017-12-05 1.6.22
- 2017-09-21 1.6.21.1
- 2017-07-04 1.6.21
- 2017-04-13 1.6.20.2
- 2016-12-14 1.6.20.1
- 2016-11-30 1.6.20
- 2016-11-10 1.6.19
- 2016-08-03 1.6.18.3
- 2016-07-18 1.6.18.2
- 2016-06-21 1.6.18.1

OpenAFS 1.8.0

What's new in 1.8.0

- Many internal code cleanups
- Better support for non-DES encryption types
- Pthreaded vserver, ptserver, and most utilities
- Rx atomics and other Rx improvements
- Server log files are no longer truncated on startup
- Remove Linux 2.2 and 2.4 support
- and many more... see NEWS

OpenAFS 1.8.1 upcoming

- Fix for crash when cache by-pass is enabled
- More useful library exports for external tools
- Build fix for broken GNU libtool on Solaris



OpenAFS code contributors (27)

Committed within the last 24 months:

Anders Kaseorg

Andrew Deason

Benjamin Kaduk

Caitlyn Marko

Chaskiel Grundman

Christof Hanke

Damien Diederer

Daria Phoebe Brashear

Dave Botsch

Garrett Wollman

Ian Wienand

Jeffrey Altman

Joe Gorse

Jonathon Weiss

Marcio Barbosa

Mark Vitale

Matt K. Light

Michael Lass

Michael Meffie

Nathaniel Filardo

Neale Ferguson

Pat Riehecky

Sergio Gelato

Seth Forshee

Stephan Wiesand

Tim Creech

Yadav Yadavendra



Secure | https://auristor-my.sharepoint.com/:p/r/person/jaltman_auristor_onmicrosoft_com/_layouts/15/Doc.aspx?sourcedoc=%7B2023ea3e-0ca5-4694-81ec-21c...

This site uses cookies for analytics, personalized content and ads. By continuing to browse this site, you agree to this use.

PowerPoint Online - Auristor-HEPiX-Spring-2018-AFS-Containers - Read-Only

PowerPoint ribbon with tabs: FILE, HOME, INSERT, DESIGN, TRANSITIONS, ANIMATIONS, REVIEW, VIEW. Home tab includes: Undo, Paste, Cut, Copy, Format Painter, Delete, New Slide, Duplicate Slide, Hide Slide, Font, Paragraph, Shapes, Arrange, Quick Styles, Drawing, Editing.

READ ONLY You don't have permission to edit this file.

Slide 1: AURISTOR YOUR DATA · YOUR FILE SYSTEM. THE AFS NAMESPACE AND CONTAINERS. A SECURE DISTRIBUTED FILESYSTEM APPROACH TO SECURELY PROVIDE PERSISTENCE TO LINUX CONTAINERS. LINUX AF_RRPC AND KERNEL AFS BY DAVID HOWELLS (DHOWELLS@REDHAT.COM). DEMOS BY MARC DIONNE (MARC.DIONNE@AURISTOR.COM). PRESENTED BY JEFFREY ALTMAN (JALTMAN@AURISTOR.COM).

Slide 2: CREDITS. The demonstration you are about to see is the product of David Howells, Marc Dionne, the rest of the Auristor team, and all those who led the way at IBM, Google, Akamai, University, and Twitter. David Howells began development of Linux file systems in October 2007 (Linux 2.6.23). Today, AFS is a single architecture for implementation on systems with diverse underlying hardware that can be used to implement various services. Copy and file cache are more transparently managed by many Linux based and non-Linux systems. AFS provides access to the file namespace through the kernel's file system interface (VFS).

Slide 3: SOFTWARE DEPLOYMENT USING JAFS. The file namespace has long been used for distribution of software and configuration. All infrastructure and storage services are implemented for hundreds of other distributed systems. (Auristor also provides a range of other software deployment and distribution services.)

Large slide content: AURISTOR YOUR DATA · YOUR FILE SYSTEM. THE AFS NAMESPACE AND CONTAINERS. A SECURE DISTRIBUTED FILESYSTEM APPROACH TO SECURELY PROVIDE PERSISTENCE TO LINUX CONTAINERS. LINUX AF_RRPC AND KERNEL AFS BY DAVID HOWELLS (DHOWELLS@REDHAT.COM). DEMOS BY MARC DIONNE (MARC.DIONNE@AURISTOR.COM). PRESENTED BY JEFFREY ALTMAN (JALTMAN@AURISTOR.COM).



Preparing for RHEL8 - python

RHEL8 will be python3 by default

The expectation is that python 2.7 will be provided as a software collection by upstream.

Folks who depend on python 2 should really begin the migration to python3.


Python3 is currently available as a software collection for SL6 and SL7

Python 2.7 End of Life is January 1, 2020

Notebooks

- Teraflops of Jupyter: A Notebook Based Analysis Portal at BNL

Teraflops of Jupyter: A Notebook Based Analysis Portal at BNL


Ofer Rind 

Chamberlin Hall, University of Wisconsin-Madison

11:45 - 12:05

- Evolution of the Hadoop and Spark platform for HEP

Evolution of the Hadoop and Spark platform for HEP


Zbigniew Baranowski 

Chamberlin Hall, University of Wisconsin-Madison

15:20 - 15:40

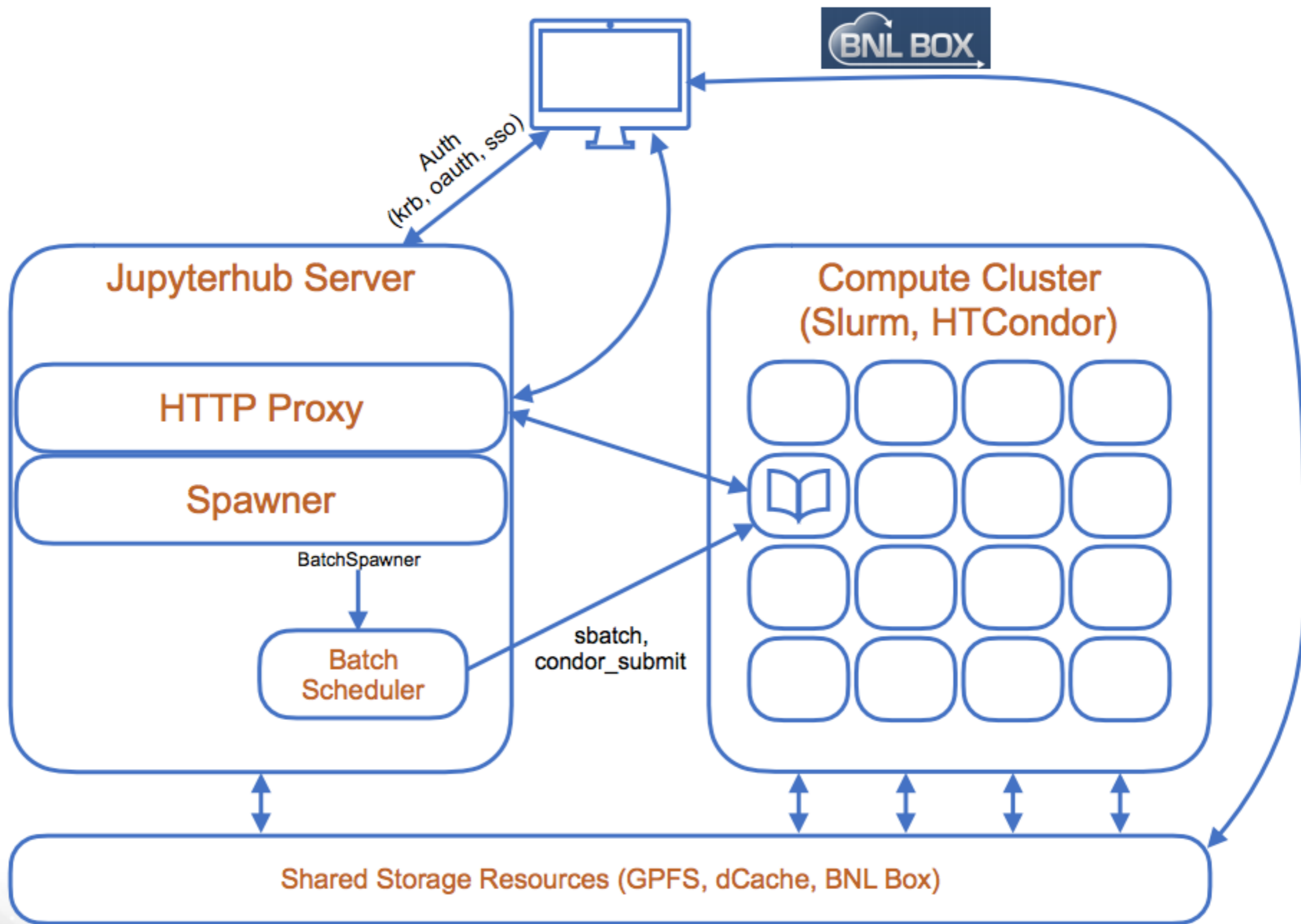
- Data analysis as a service, STFC

Data analysis as a service

Mr James Adams 

Chamberlin Hall, University of Wisconsin-Madison

10:50 - 11:10



What could be gained for CERN?

precision at best 10 – 20 %, but based on conservative assumptions

- $780 * 32 * 0.4 = 9984$ cores
- $>80\%$ can be used: $780 * 32 * 0.8 = 19968$ cores



- Batch farm runs in VMs as well
 - 3% performance overhead, 0% with containers

9000 hypervisors (7K 1yr ago)

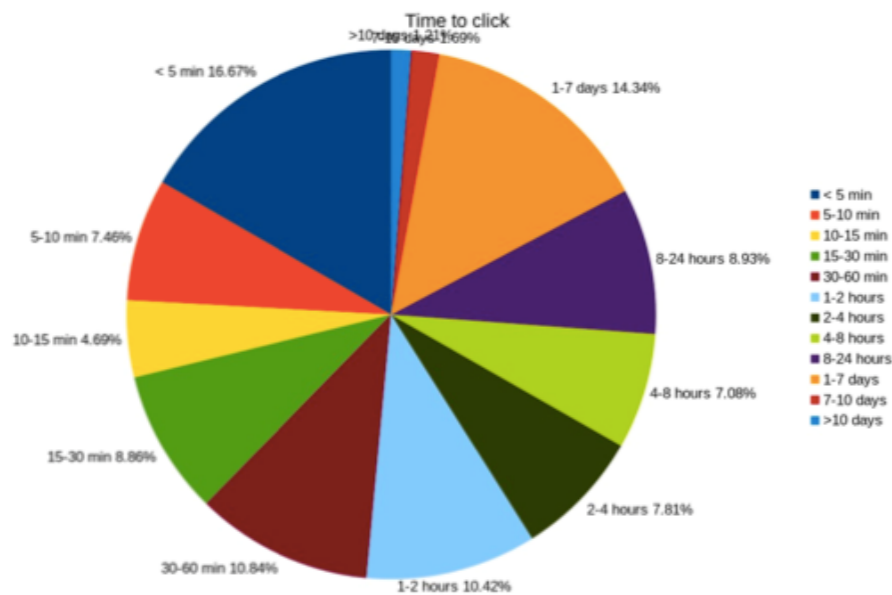
320K cores (210k)

830 TB of RAM (350 TB)

261 container clusters (Magnum) (67)

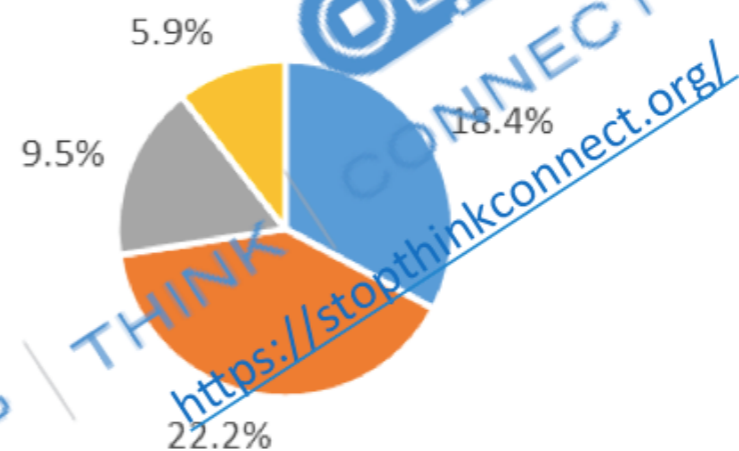
3% of 320K = 10K cores

(2018/4) Watch Your Click (2)



Defense has to be quick...!

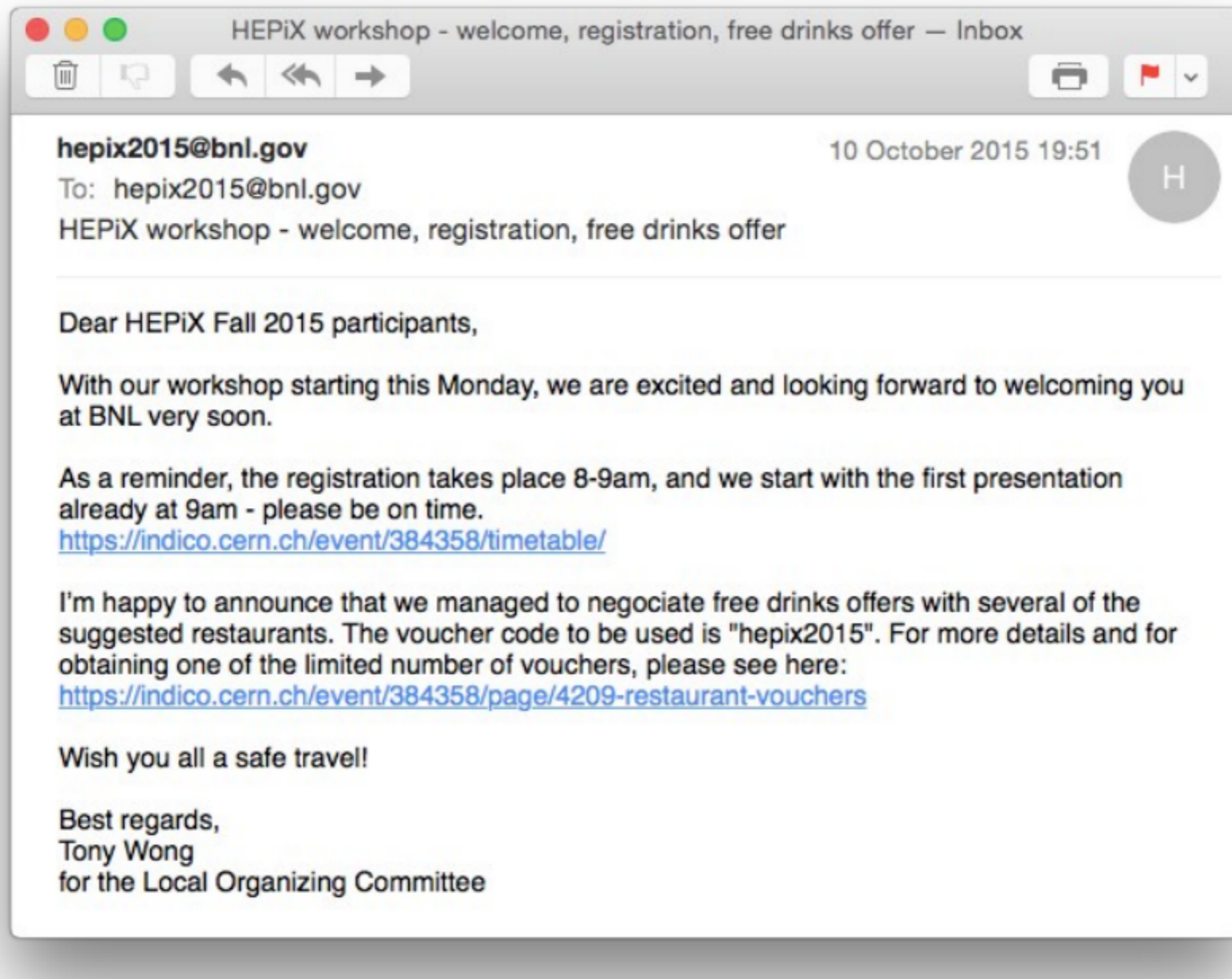
Don't underestimate your people:



Click rate proportional to mail sophistication.
Overall click rate: **14.0% (2018)**
18.7% (2017)
16.5% (2016)

free **B E E R** forged email at HEPiX: ~30% (test)





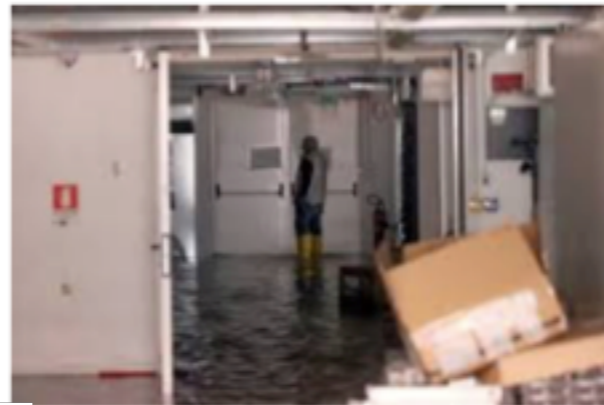
https://security.web.cern.ch/security/reports/presentations/2015/HEPIX_Fall_2015_phishing_campaign.pdf

Disaster recovery

The Tier-1 entrance that morning



- All Tier-1 doors are watertight
- Height of water outside: 50 cm
- Height of water inside: 10 cm (on floating floor) for a total volume of $\sim 500 \text{ m}^3$



Damage to IT equipment: the list

2.2 PB of disk lost (astrophysics data)

- Computing farm
 - ~34 kHS06 are now lost (~14% of the total capacity)
 - No special action taken (replaced)
 - Library and HSM system
 - 1 drive damaged
 - Several non critical components
 - Library recertified in January
 - 4 TSM-HSM servers
 - Tapes
 - 136 tapes damaged
 - Tapes being recovered in lab
- 40 +22 tapes recovered
 - 1 tape partially recovered (LHCb)
 - 3 (CMS) + 3 (LHCb) tapes undergoing second round
 - 6 tapes (CMS) to be recovered
 - Nearly all storage disk systems involved
 - 11 DDN JBODs (2 for CMS)
 - *RAID parity affected*
 - 2 Huawei JBODs (non-LHC experiments)
 - 2 Dell JBODs including controllers
 - 4 disk-servers

System	PB	JBODs	Involved experiments
Huawei	3.4	2	All CSN2 and 3 experiments excepting AMS, Darkside e Virgo
Dell	2.2	2	Darkside and Virgo
DDN 1,2	1.8	4	ATLAS, Alice and LHCb
DDN 8	2.7	2	LHCb
DDN 9	3.8	2	CMS
DDN 10, 11	10	3+2	ATLAS, Alice and AMS
Total	23.9	9	

General Impressions

- Site reports — limited interest (for me) but eat one full day
- CERN talks generally stand out: more content, provide guidance,...
- The event is nice for discussing specific points of interest
 - A bit less exciting for general “discovery” of things
- **Newcomers get a good overview of the community**

Beer & Coffee

