Archiving storage for Scientific Research Data

DAVID YU

10TH **A**NNUAL **G**LOBAL **IT** **E**XECUTIVE **S**UMMIT

About Brookhaven National Laboratory (BNL)

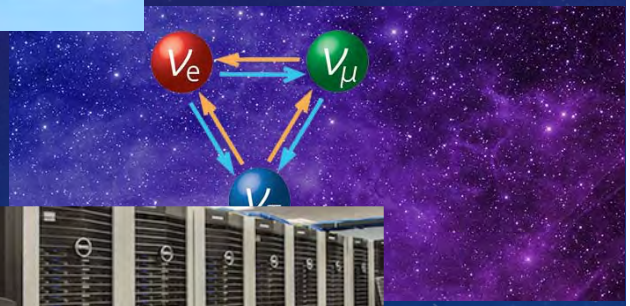
About Brookhaven National Laboratory (BNL)

Energy & Photon Sciences
Center for Functional Nanomaterials
Chemistry
Condensed Matter Physics & Materials Science
National Synchrotron Light Source II
Sustainable Energy Technologies

Environment, Biology, Nuclear Science & Nonproliferation
Biology
Environmental and Climate Sciences Department
Nuclear Science and Technology
Nonproliferation and National Security

Nuclear & Particle Physics
Collider-Accelerator
Instrumentation Physics
Superconducting Magnet
RIKEN BNL Research Center

Computational Sciences
Computer Science and Mathematics
BNL Scientific Data and Computing
Center Center for Data-Driven Discovery
Computational Science Laboratory



Facilities at Brookhaven National Laboratory

Research Facilities

- Relativistic Heavy Ion Collider (RHIC)
- National Synchrotron Light Source II
- Center for Functional Nanomaterials (CFN)
- NASA Space Radiation Laboratory
- Scientific Data and Computing Center
- And more …

BROOKHAVEN NATIONAL LABORATORY | Scientific Data and Computing Center

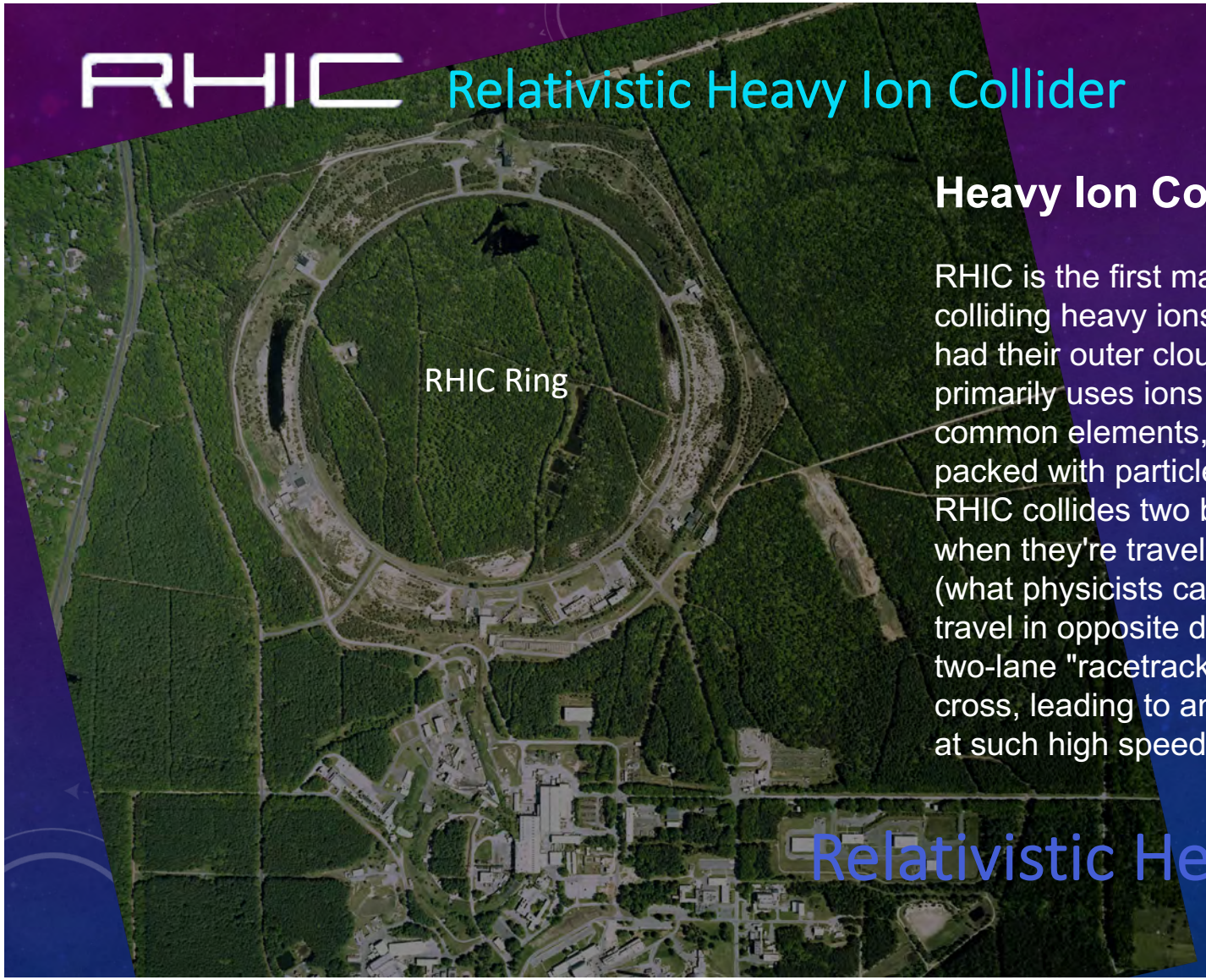
Provides computing services for **RHIC**

- Tier-0 Data storage for all experiment data.
- Primary center for data processing.

Serves as LHC **ATLAS** Tier-1 for the US

- Secondary data storage for fraction of data (~23%).
- Primary US site for data storage, processing and distribution.

Serves as **Belle-2** Tier-1 (New)



Heavy Ion Collisions

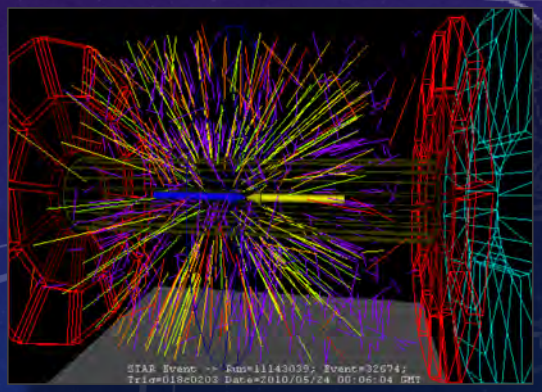
RHIC is the first machine in the world capable of colliding heavy ions, which are atoms which have had their outer cloud of electrons removed. RHIC primarily uses ions of gold, one of the heaviest common elements, because its nucleus is densely packed with particles.

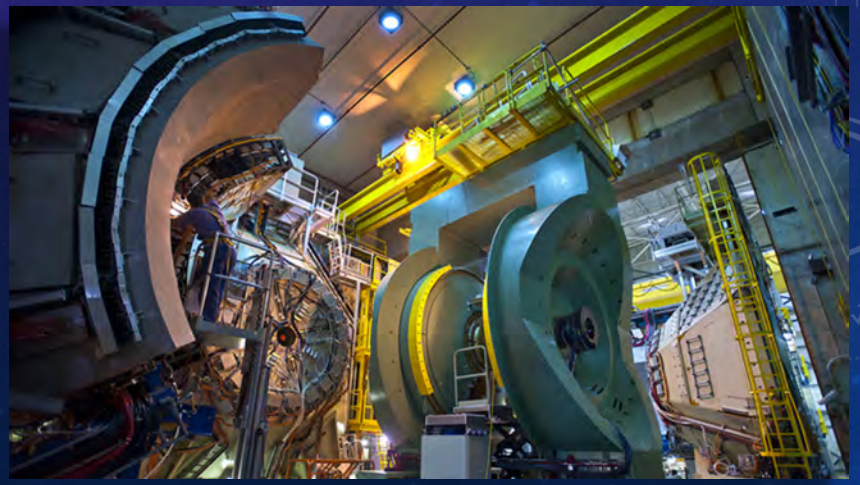
RHIC collides two beams of gold ions head-on when they're traveling at nearly the speed of light (what physicists call relativistic speeds). The beams travel in opposite directions around RHIC's 2.4-mile, two-lane "racetrack." At six intersections, the lanes cross, leading to an intersection. When ions collide at such high speeds fascinating things happen.

Relativistic Heavy Ion Collider

RHIC Relativistic Heavy Ion Collider

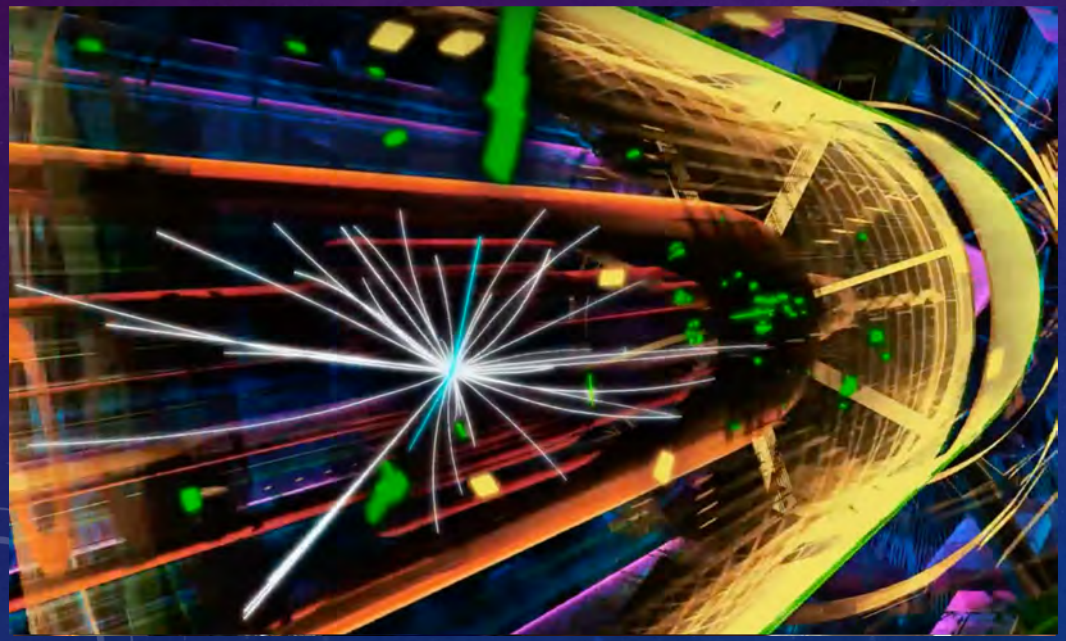
STAR and PHENIX, two large detectors located around the 2.4-mile-circumference accelerator.





The Large Hadron Collider (LHC)

The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelerator on the planet. It first started up on 10 September 2008, and remains the latest addition to CERN's accelerator complex. The LHC consists of a 27-kilometre ring of superconducting magnets with a number of accelerating structures to boost the energy of the particles along the way.



CERN, the European Organization for Nuclear Research, founded in 1954 to unite nations through science. Physicists and engineers are probing the fundamental structure of the universe. They use the world's largest and most complex scientific instruments to study the basic constituents of matter – the fundamental particles.

CERN is derived from the acronym for the French **"Conseil Européen pour la Recherche Nucléaire"**, (European Council for Nuclear Research)

A System of Tiers

- Tier 0 is the CERN Data Centre. All of the data from the LHC passes through this central hub
- Tier 1 consists of 13 computer centers large enough to store LHC data. They provide round-the-clock support for the Grid, and are responsible for storing a proportional share of raw and reconstructed data

Country	Tier 1 Grid site
Canada	TRIUMF
Germany	KIT
Spain	PIC
France	IN2P3
Italy	INFN
Nordic countries	Nordic Datagrid Facility
Netherlands	NIKHEF / SARA
Republic of Korea	GSDC at KISTI
Russian Federation	RRC-KI and JINR
Taipei	ASGC
United Kingdom	GridPP
US	Fermilab-CMS
US	BNL ATLAS

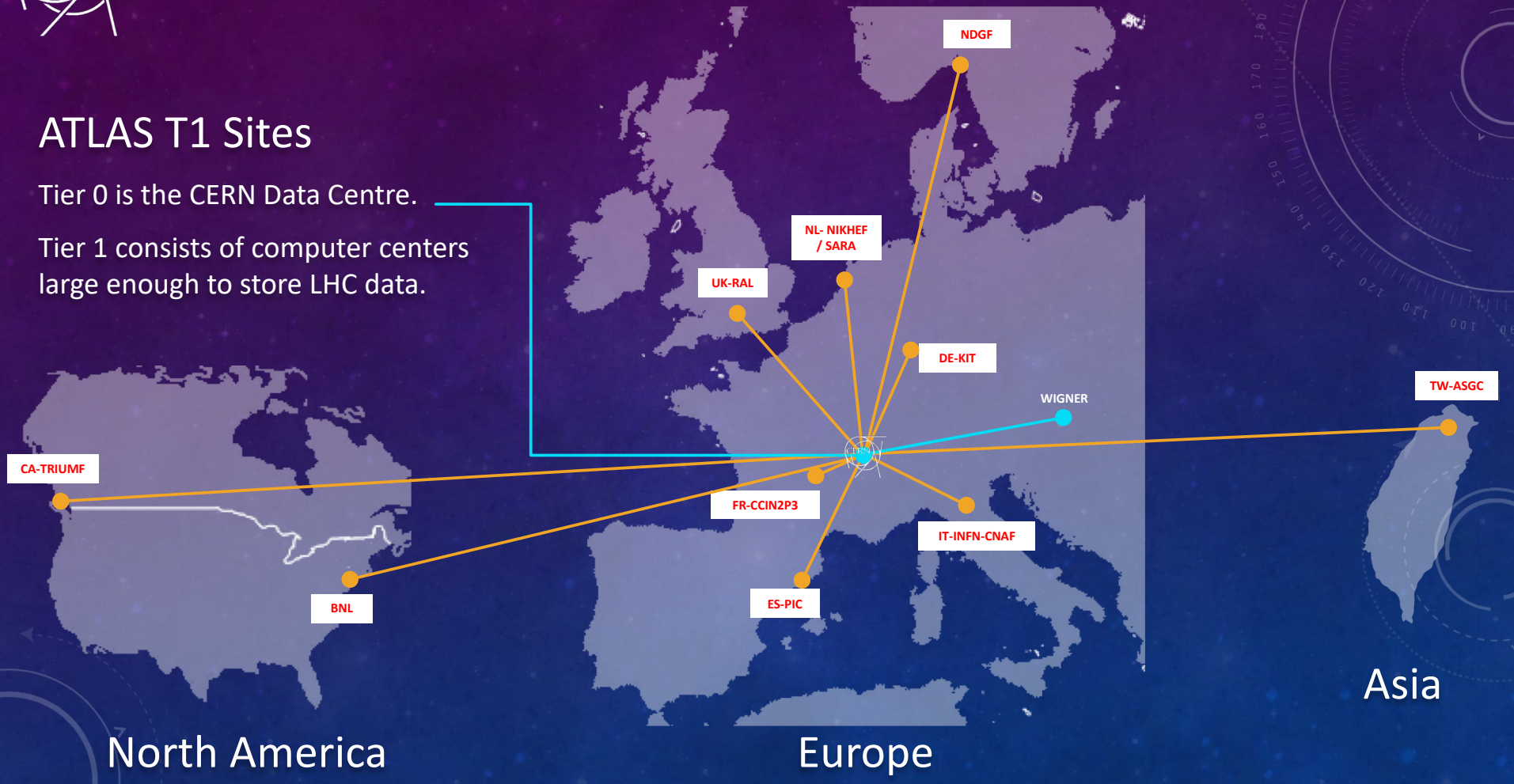
https://home.cern/about/computing/grid-system-tiers

The Large Hadron Collider (LHC)

ATLAS T1 Sites

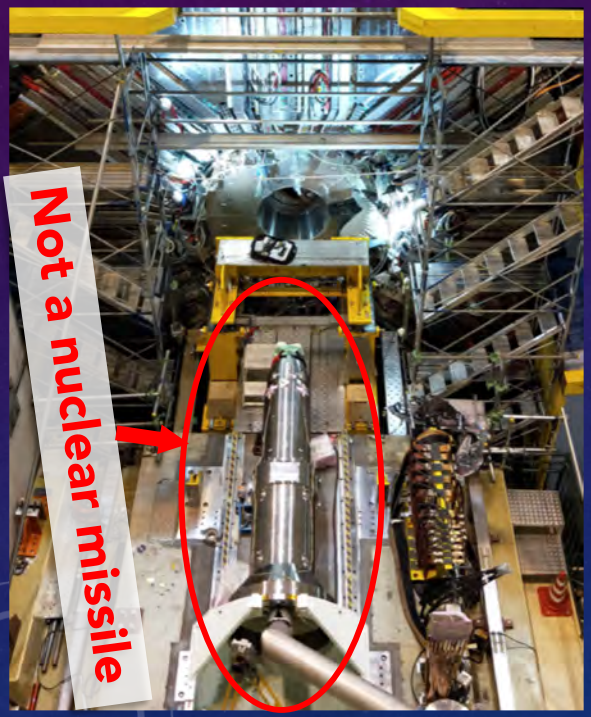
Tier 0 is the CERN Data Centre.

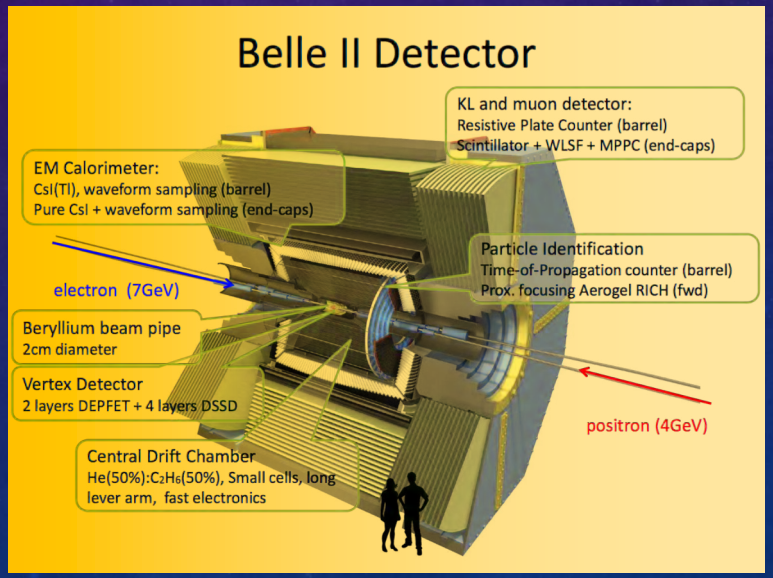
Tier 1 consists of computer centers large enough to store LHC data.

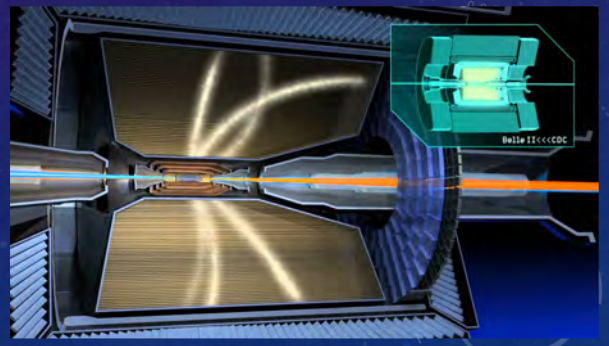


Belle II Experiment

The Belle II experiment is a particle physics experiment designed to study the properties of B mesons (heavy particles containing a bottom quark). Belle II is the successor to the Belle experiment, and is currently being commissioned at the SuperKEKB[1] accelerator complex at KEK in Tsukuba, Ibaraki Prefecture, Japan.







Know More About RHIC, LHC and Belle II Experiment

https://www.youtube.com/watch?v=1cY78LskVjQ

https://youtu.be/nGCrrgXSEOk?t=6s

https://youtu.be/328pw5Taeg0?t=1m27s

https://www.youtube.com/watch?v=pQhbhpU9Wrg

KEK - High Energy Accelerator Research Organization

高エネルギー加速器研究機構

***K**ō **E**nerugī **K**asokuki Kenkyū Kikō*

A pilot mass-storage system for KEK belle experiment

Information | Discussion (0) | Files

Exhibition Objects

Report number	CERN-OBJ-IT-055
Category	Computing and computers
Title	**SONY SD1-1300L**
Description	Cartridge used in the SONY DMS-24 automated tape library system installed at CERN in 1995 and still in use by the NA49 experiment. Tape length is 1300 m with 100 GB storage capacity.
Description (French)	Cartouche utilisée dans le système de bibliothèque de bandes automatiques SONY DMS-24 installé au CERN en 1995 et toujours utilisé par l'expérience NA49. La longueur de bande est de 1300 m avec une capacité de stockage de 100 Go.
Year	1995
Keywords	IT
Physical characteristics	Height: 20 cm Depth: 3,5 cm Length: 36 cm
Availability	On loan
File(s)	

96 GB

© CERN

Total images: 1

CERN-OBJ-IT-055-1

GIF;ICON,JPG,JPG;ICON-180

Many objects of this database may be borrowed: see the loan conditions.

https://cds.cern.ch/record/2273963?ln=en

Data storage medias in 1995

(codenamed **Chicago**)

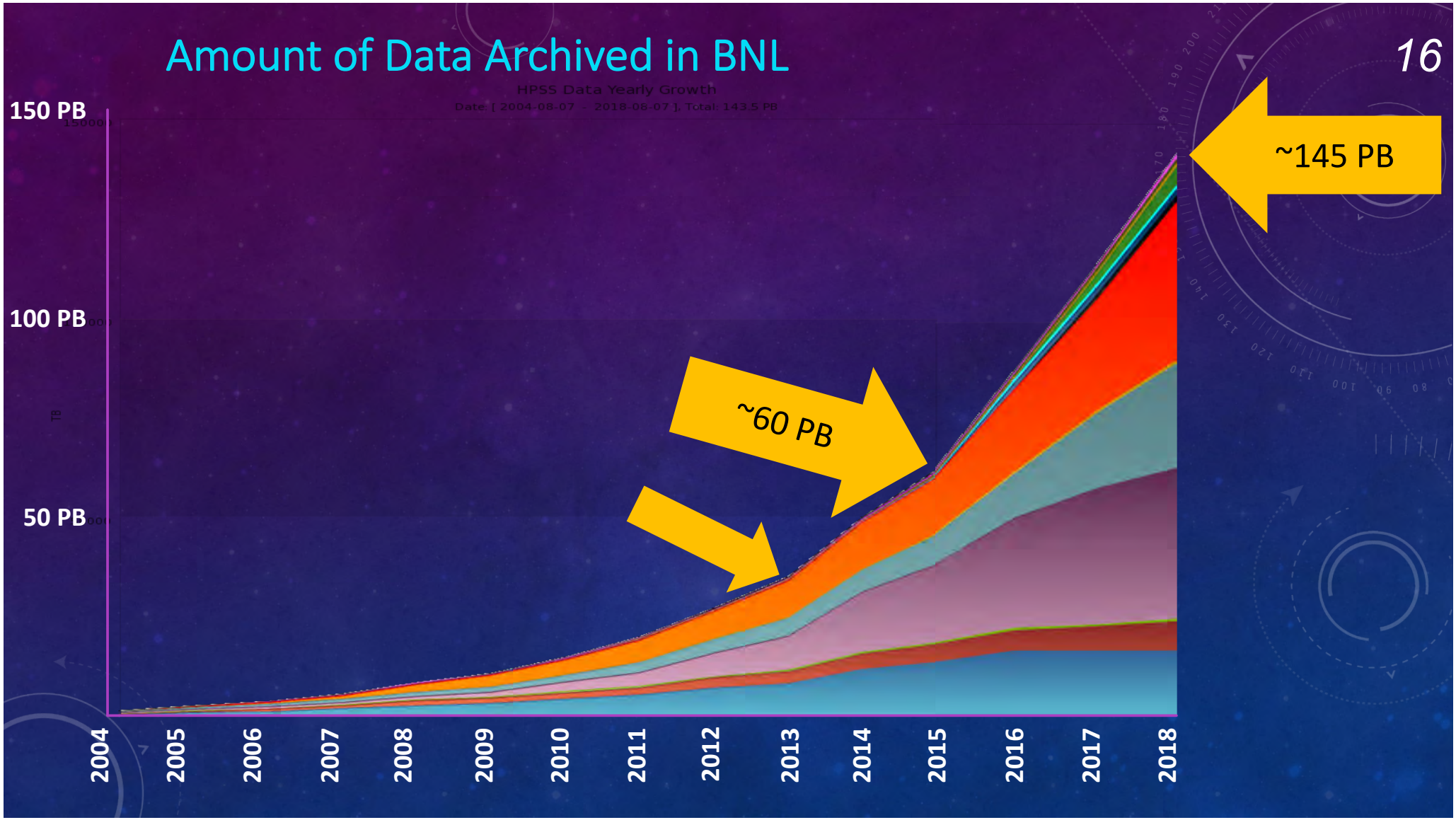
In 2018,
128 GB memory card
$40

Seagate Elite 9 10.8GB Fast SCSI, 10 MB/s

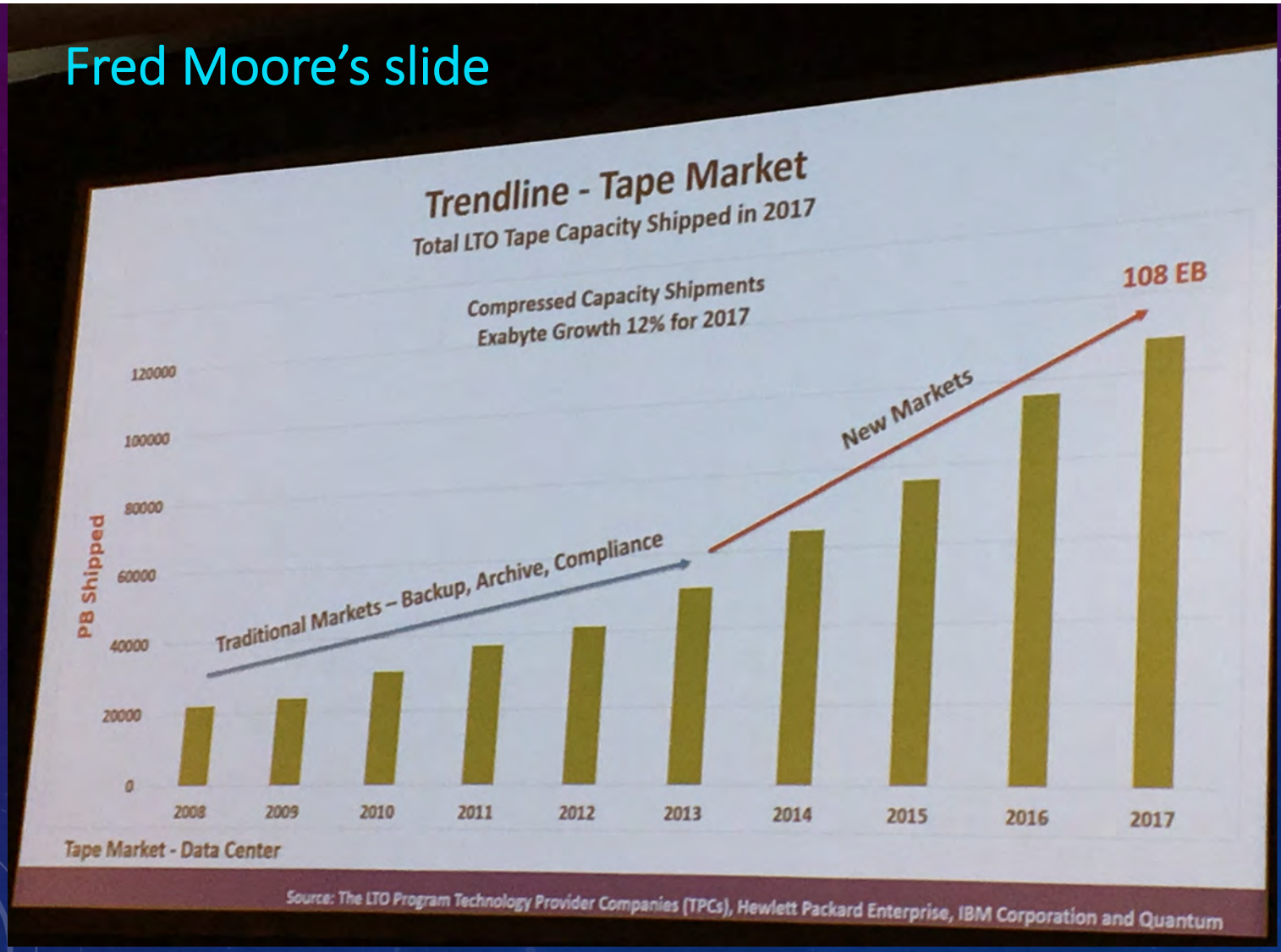
Transfer rate: 32 MB/s

http://www.reactivedata.com/tapedrives/Sony%20DIR1000H.htm

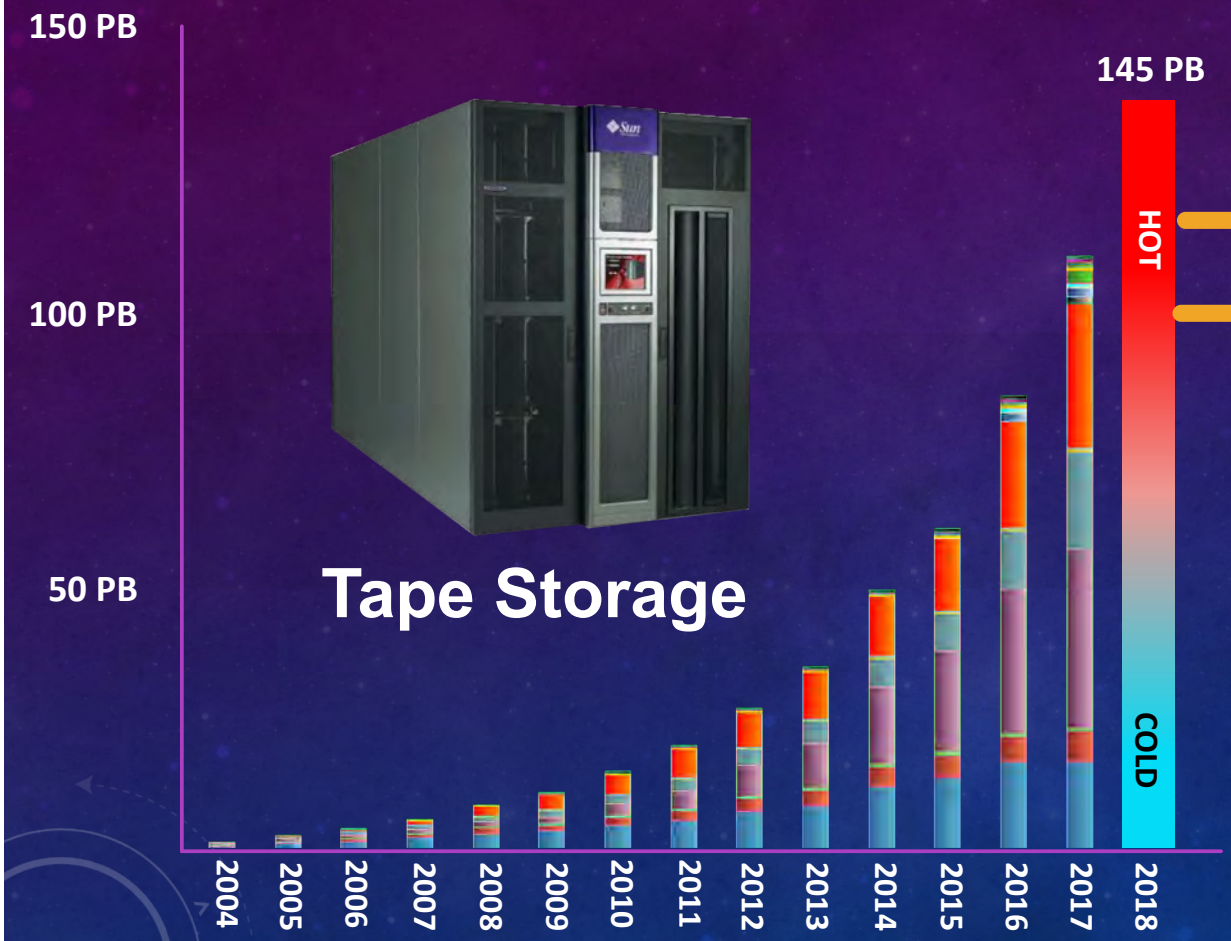
Amount of Data Archived in BNL



Fred Moore's slide



Big Data Storage



Tape Storage

Other Research Institutes

BNL Data Center

150 PB

100

50 PB

2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018

ATLAS

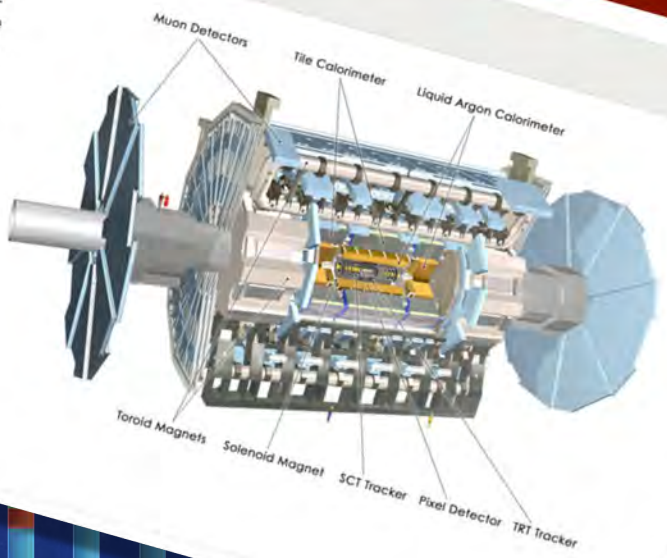
Home People Theses TileCal Trigger Midwest Tier2 Center Physics Links HEP

THE UNIVERSITY OF CHICAGO | Enrico Fermi Institute High Energy Physics

The ATLAS detector studies proton-proton collisions at a center of mass energy of 14 TeV, seven times higher than previous facilities. We have discovered the Higgs boson in 2012 and are constantly searching for new forms of matter such as supersymmetric states. The lighest supersymmetric state is a prime candidate for the dark matter and supersymmetry is an essential ingredient of many attractive theories.

This detector is now in operation at CERN's Large Hadron Collider (LHC). Since 2010 it has been operating at a center of mass energy of 7 TeV (2010-2011), 8 TeV (2012) and 13 TeV (2015-present). The Chicago group built much of the readout electronics for the hadron calorimeter (TileCal). The group is also currently involved in preparing upgrades to the trigger system used to identify interesting interactions to be recorded, including a new hardware-based track trigger (FTK), and a new hardware-based calorimeter trigger (gFEX). The group is heavily involved in analysing the data taken by the experiment.

The Chicago group consists of five faculty members together with research personnel and students. Many of the members of the group are in full time residence at CERN, with the rest in Chicago. The physics analysis of the group is supported by a powerful computing facility which includes an ATLAS Tier 3 computing system dedicated to Chicago physicists, and one of the five US ATLAS Tier 2 facilities.



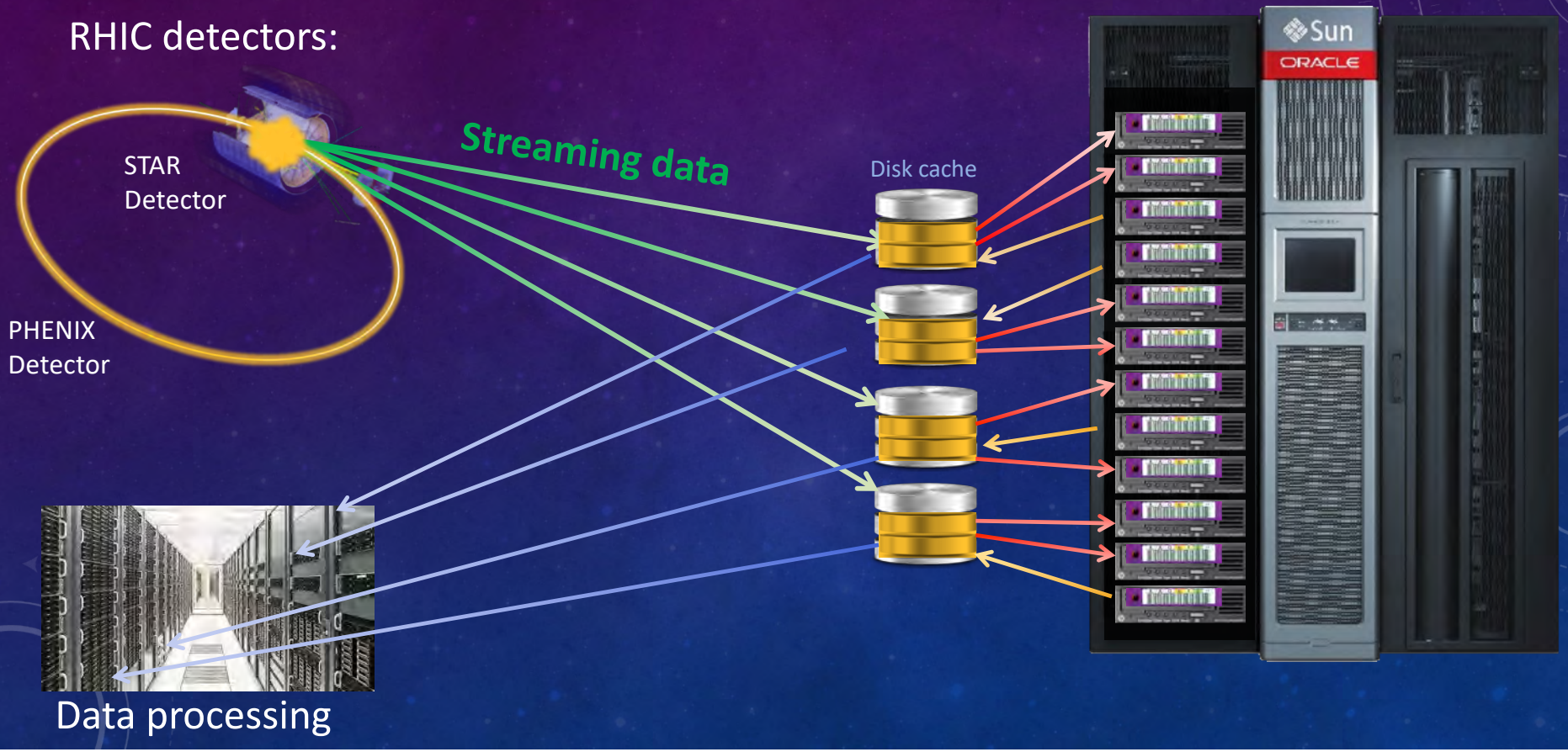
esearch Institutes

BNL Data Center

High Throughput Active Archiving System

RHIC Experiment data directly go to tape storage (primary copy)

RHIC detectors:



Tape Storage Usage

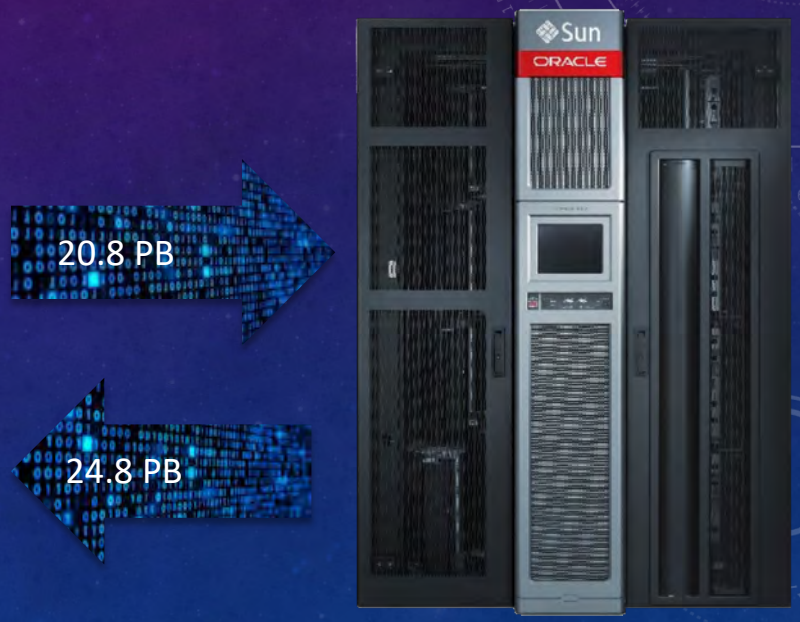
In 2017

Archived to tape:

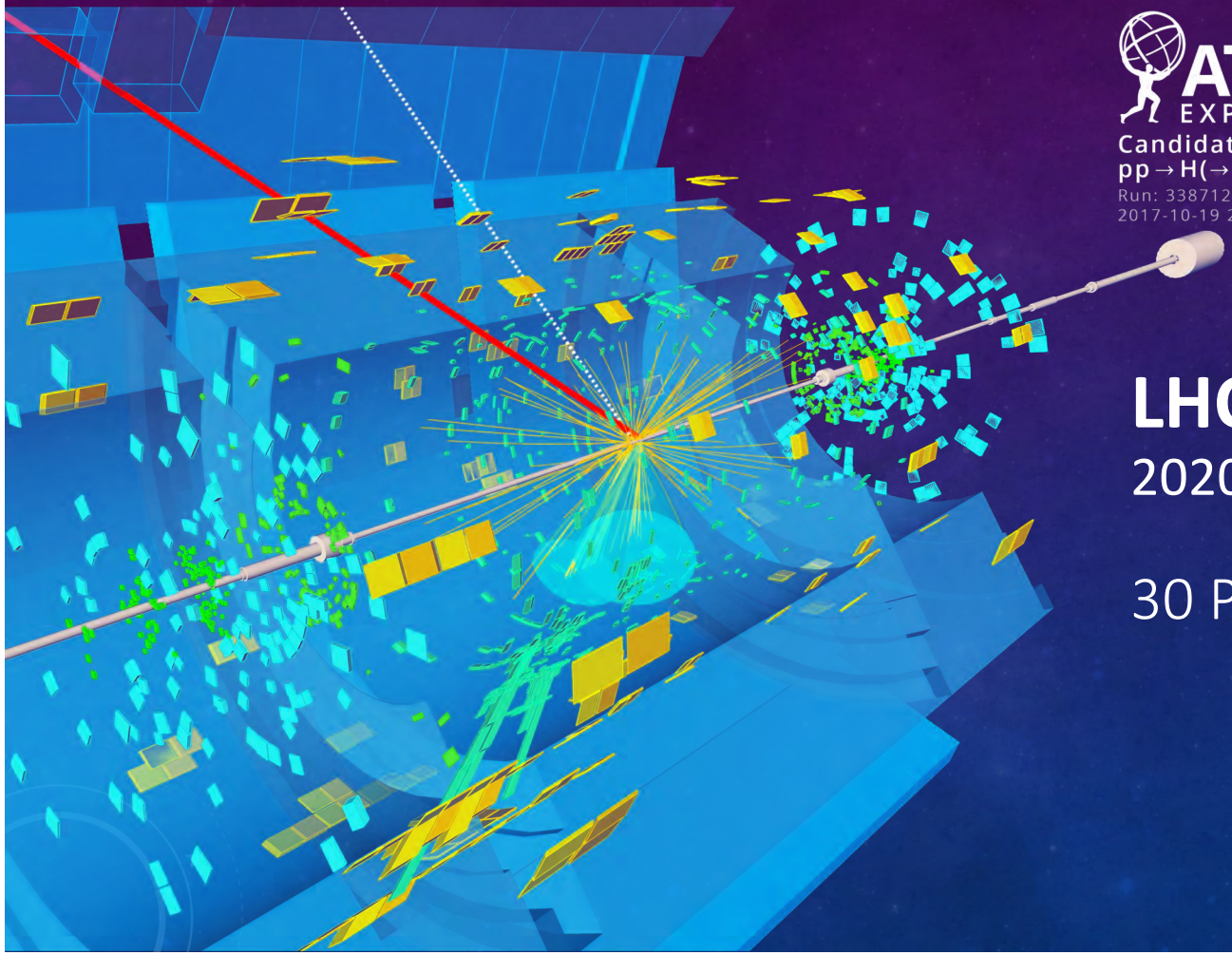
19,412,702 files – Average 53,185 files/day
20.8 PB – Average 58.4 TB / day

Restored from tape:

11,693,141 files - Average 32,036 files/day,
24.8 PB - Average 69.5 TB/day



Future Challenges in Next 10 Years



LHC Run 3

2020, 2021, 2022

30 PB a year?!


High Luminosity Large Hadron Collider
10x more data

sPHENIX:

The next generation heavy ion detector at RHIC

Data acquisition rate at 12 GB/s

LTO-8 is 360 MB/s

Need 35+ LTO-8 drives to keep up the speed!

100+ PB a year

Future Challenges in Next 10+ Years

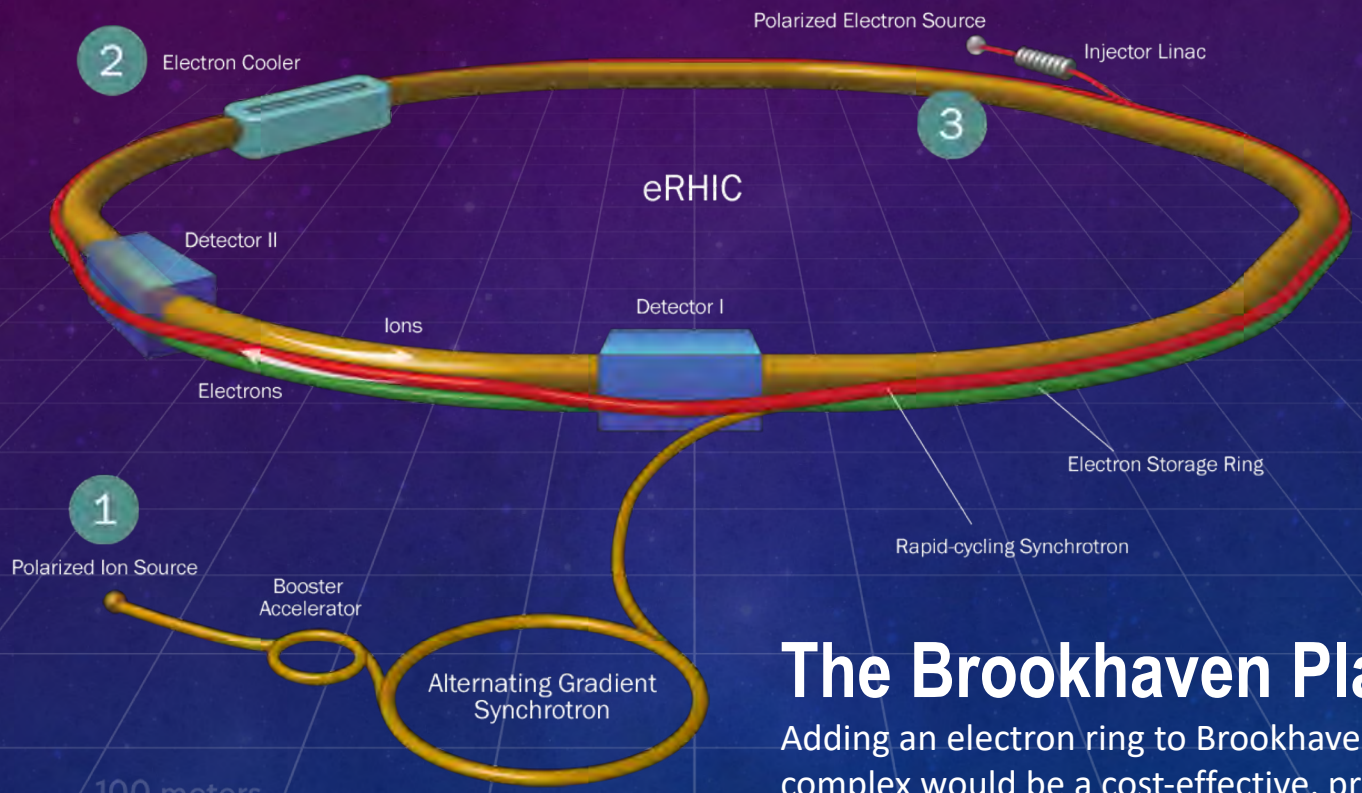
The Electron-Ion Collider

A machine that will unlock the secrets of the strongest force in Nature

The computers and smartphones we use every day depend on what we learned about the atom in the last century. All information technology—and much of our economy today—relies on understanding the electromagnetic force between the atomic nucleus and the electrons that orbit it. The science of that force is well understood but we still know little about the microcosm within the protons and neutrons that make up the atomic nucleus. That's why physicists want to build a new machine—an Electron-Ion Collider, or EIC—to look *inside* the nucleus and its protons and neutrons.

The EIC would be a particle accelerator that collides electrons with protons and nuclei to produce snapshots of those particles' internal structure—like a CT scanner for atoms. The electron beam would reveal the arrangement of the quarks and gluons that make up the protons and neutrons of nuclei. The force that holds quarks together, carried by the gluons, is the strongest force in Nature. An EIC would allow us to study this "strong nuclear force" and the role of gluons in the matter within and all around us.

Future Challenges in Next 10+ Years



The Brookhaven Plan

Adding an electron ring to Brookhaven's existing accelerator complex would be a cost-effective, practical strategy for achieving the scientific goals of an electron-ion collider

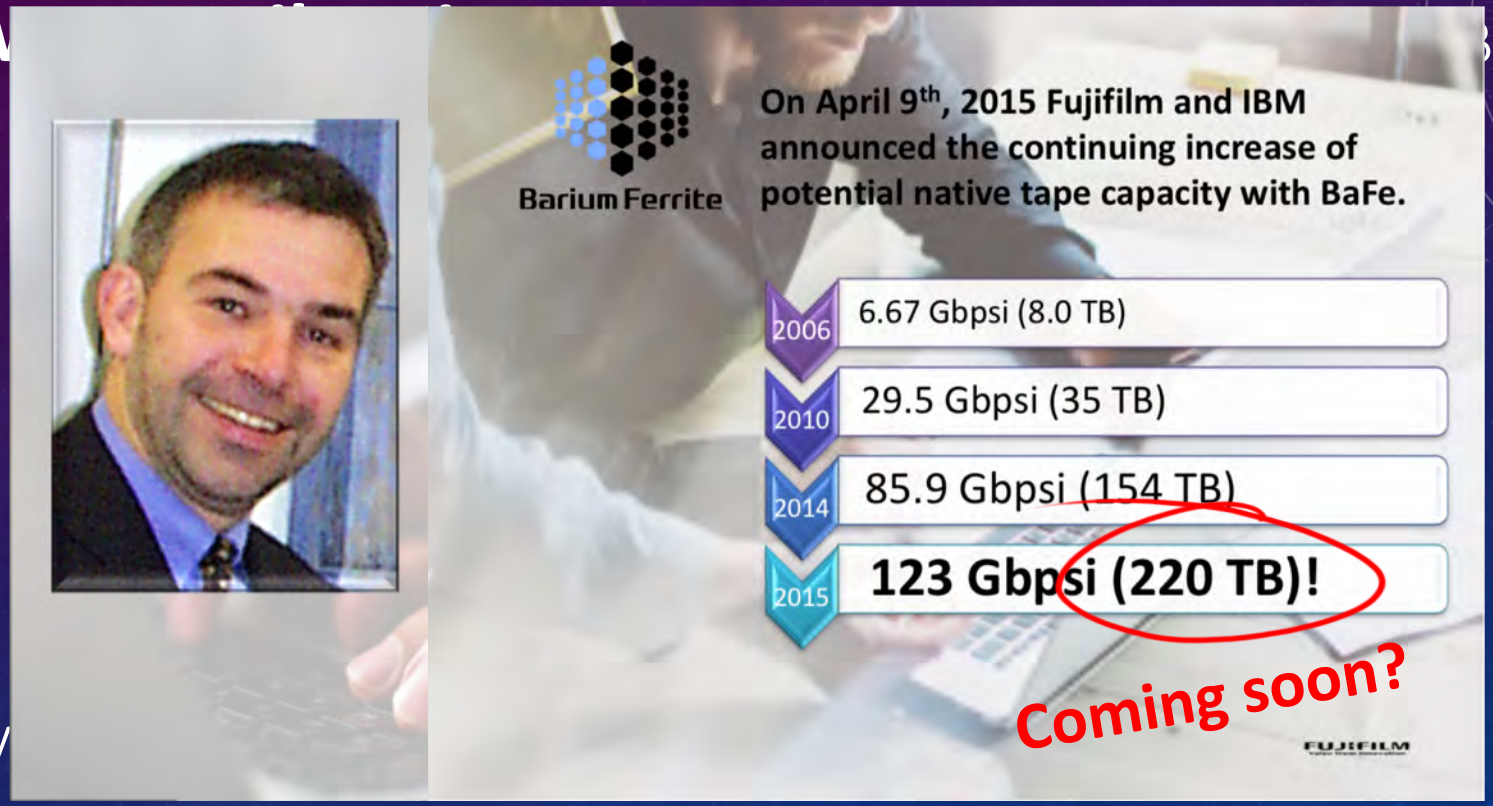
Future Challenges in Next 10 Years

New Tape Libraries Data acquisition volume, max 0.5 EB a year

May need ~100,000 slots in next 10 years

Future Challenges in Next 10 Years

New … a year



May …

Tape Storage Plays a Significant Role in Protecting The Scientific Data

2000 - 2006

4 x STK 9310
9940B Tapes

2005 – now, 9 x SL8500, LTO[3-7]

Are Tapes Really Reliable?

Media Types	Cartridges	Average Mounts	Max Mounts
LTO-7	9,057	31.2	736
LTO-6	22,345	52.8	664
LTO-5	15,684	67.3	1,225

We have near zero failures with LTO-6 and LTO-7

Recalled 7,567,301 files from LTO-7

Recalled 18,255,072 files from LTO-6

Most of the IO errors were caused by excessive usage of our tape drives

Are These Numbers Realistic?

LTO-5
140 MB/s

LTO-6
160 MB/s

LTO-7
300 MB/s

Tape is sequential access media

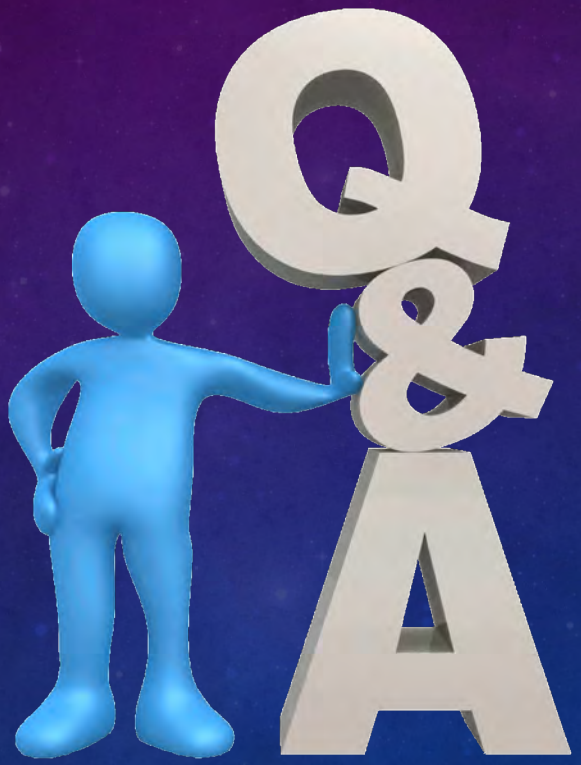
Tape is designed for streaming data

Minimized redundant tape-mounts.

Large files works better.

We have observed files recall at average 132 MB/s/drive for LTO-5 in production. All files were compressed. All overhead (mount and seek) included.

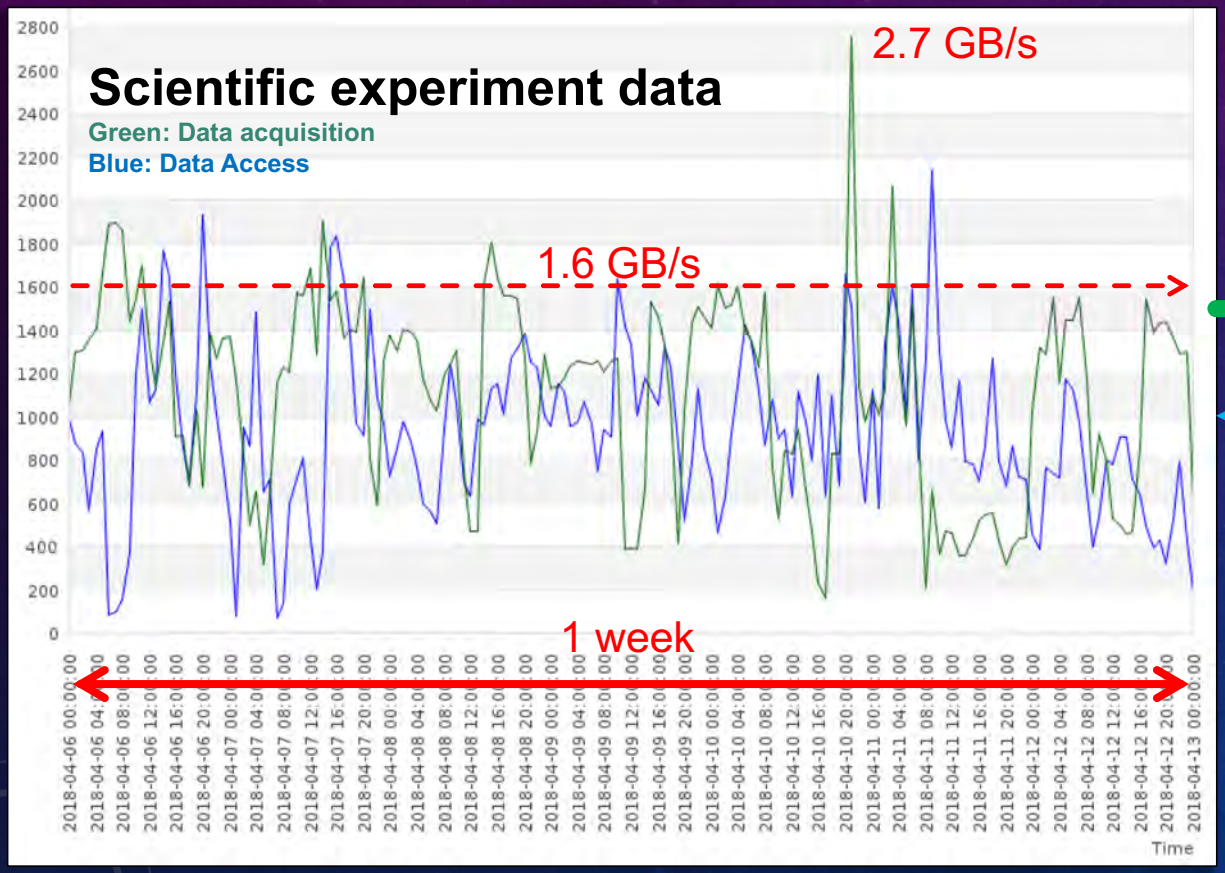
LTO-5 max speed: 140 MB/s

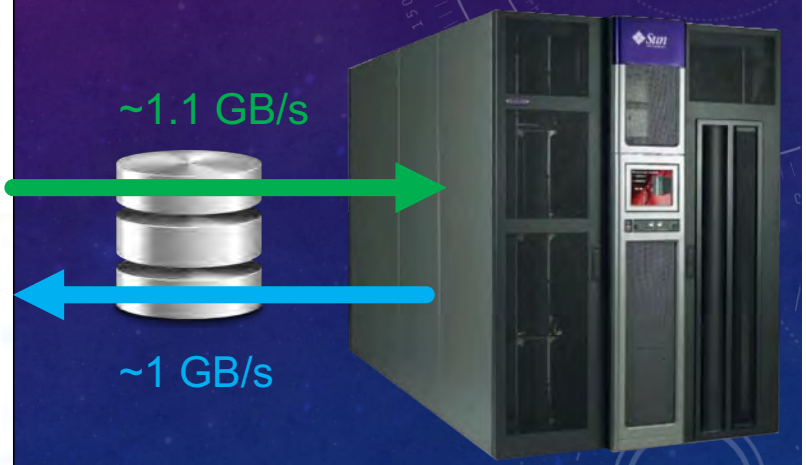

Q&A

THANK YOU!

Supplemental Materials

Are Tapes Fast Enough?





File Recall Performance

42,196 files requested

Average file size:
10 GB

Destination disk limit: <1.5 GB/s

Need to adjust staging drives, to keep @ <1.5 GB/s

15:00 – 21:00 sample: 1.3 GB/s

1.3 / 10 Drives = 132 MB/s/drive

All overheads included (mount/dismount, seek …)