Journal of Physics: Conference Series

PAPER • OPEN ACCESS

The role of dedicated data computing centers in the age of cloud computing

To cite this article: Costin Caramarcu *et al* 2017 *J. Phys.: Conf. Ser.* **898** 082009

The role of dedicated data computing centers in the age of cloud computing

Costin Caramarcu, Christopher Hollowell, William Strecker-Kellogg, Antonio Wong and Alexandr Zaytsev

Brookhaven National Laboratory, Department of Physics, P.O. Box 5000, Upton, NY 11973, USA
E-mail: tony@bnl.gov

Abstract. Brookhaven National Laboratory (BNL) anticipates significant growth in scientific programs with large computing and data storage needs in the near future and has recently re-organized support for scientific computing to meet these needs. A key component is the enhanced role of the RHIC-ATLAS Computing Facility (RACF) in support of high-throughput and high-performance computing (HTC and HPC) at BNL. This presentation discusses the evolving role of the RACF at BNL, in light of its growing portfolio of responsibilities and its increasing integration with cloud (academic and for-profit) computing activities. We also discuss BNL's plan to build a new computing center to support the new responsibilities of the RACF and present a summary of the cost benefit analysis done, including the types of computing activities that benefit most from a local data center vs. cloud computing. This analysis is partly based on an updated cost comparison of Amazon EC2 computing services and the RACF, which was originally conducted in 2012.

1. Introduction

The RHIC and ATLAS US Tier 1 Computing Facility (RACF) at Brookhaven National Laboratory (BNL) supports the computational needs of the RHIC experiments, the U.S. collaborators in the ATLAS experiment at CERN and the other activities undertaken by the Physics Department at BNL. It has operated continuously since the mid 1990's, serving a geographically diverse, worldwide HEPN community. The RACF is made up of a multi-silo robotic tape storage (currently at 75PB) facility, high-availability disk storage systems (currently at 45 PB) and an integrated Linux-based worker node cluster with over 61,000 compute cores.

In the Fall of 2015, the Computational Science Initiative (CSI) was launched to consolidate HPC activities at BNL. The Scientific Data and Computing Center (SDCC) is the computing facility arm of CSI, and it was formed by leveraging the RACF staff for general computing infrastructure services, with additional staff hired to operate the new HPC-centric resources.

The main data center has ~15,000 ft^2 (1,400 m^2) of space and 2.3 MW of UPS-backed power available. It is primarily assigned to RHIC and ATLAS and nearly full. See figure 1. A satellite data center with ~3,000 ft^2 and ~600 KW of UPS-backed power is similarly full and cannot support HPC-centric projects, leaving little infrastructure to support current and future programs, such as the National Synchrotron Light Source (NSLS-II), Center for Functional Nanomaterials (CFN) and Computational Science.

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 082009 doi:10.1088/1742-6596/898/8/082009

This paper discusses BNL's approach to address this problem. Based on its experience with Amazon Web Services (AWS), BNL compared the cost of cloud computing and storage services with the cost of a new, on-site data center.

Figure 1: Existing data center is full.

2. Current trends in the BNL data center

The SDCC supports dedicated HTC resources for RHIC, ATLAS and other closely-related activities and also newly-available resources in support of HPC-related activities at BNL (see table 1).

Table 1: SDCC current resources.

Cluster	Servers	Cores
HTC (RHIC/ATLAS)	2.2k	61k
Institutional Cluster (IC)	108	3888
Knights Landing Cluster (KNL)	144	9216
Others	176	1856

The recent evolution of the infrastructure requirements (space, power and cooling) contains both good and bad news. While the cost per HS06 [1] is falling (see figure 2), increasing power usage to meet growing computing requirements (see figure 3) limits future growth and the ability to support new scientific programs in the existing BNL data center.

In order to meet the known resource commitments and to accommodate future requirements, BNL is proposing a 25,000 ft^2 (2,320 m^2) usable data center floor space with 6 MW of UPS-backed power capacity (2.4 MW on day one). The new data center would be a shared facility for HTC and HPC resources and conform to US Department of Energy (DOE) requirement for a PUE between 1.2 and 1.4. This can be accomplished with hot-aisle containment and natural air-cooled facility supplemented by redundant chillers.

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 082009 doi:10.1088/1742-6596/898/8/082009

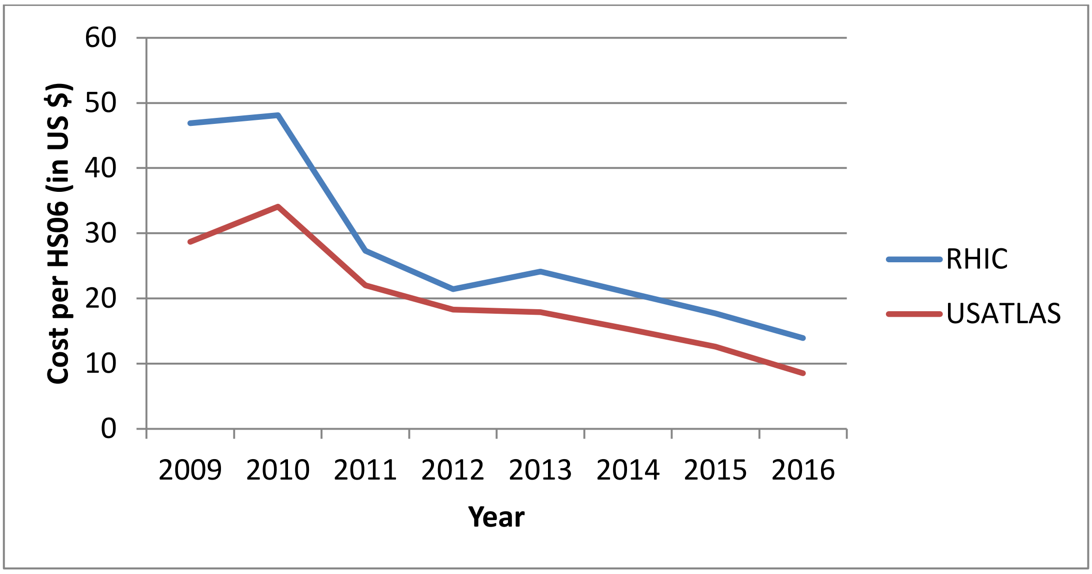

Figure 2: The cost (in US dollars) per HS06 is trending down. RHIC costs are higher due to larger local storage requirements.

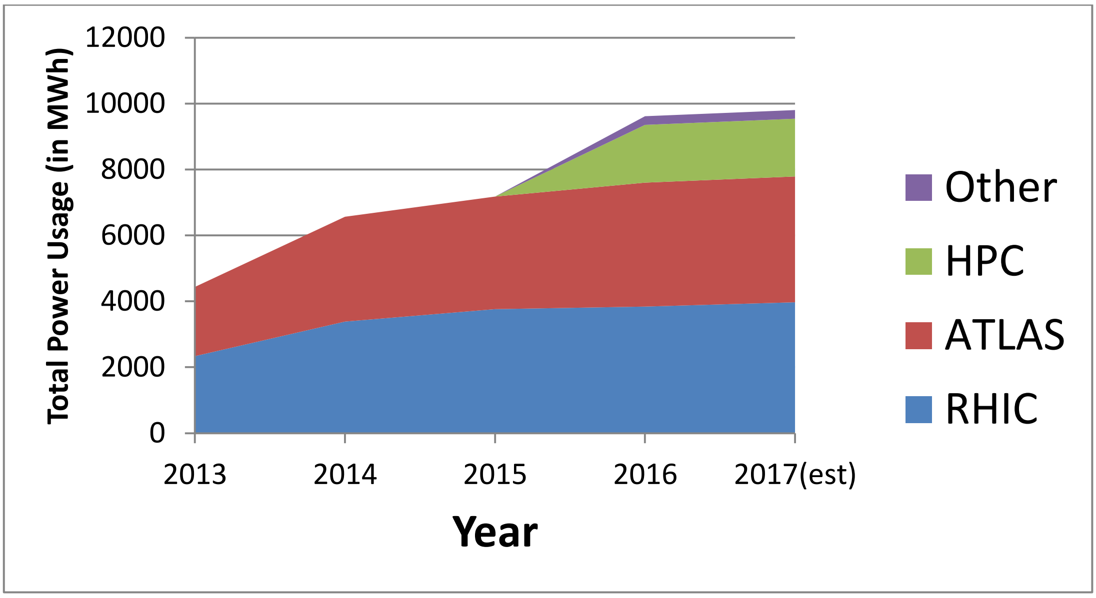

Figure 3: Growing power usage due to ever increasing computing requirements.

3. Cost Analysis

In addition to on-site solutions, current budgetary realities, program requirements and a DOE mandate that "cloud" solutions be considered over a new, on-site data center have compelled the HEPN community to evaluate off-site alternatives, such as commercial cloud (such as Google and Amazon) providers, which offer increasingly price-competitive computing (and to a lesser extent, storage) services. For the purpose of this study, the following scenarios were considered:

1) Do nothing
2) Utilize existing BNL buildings
   a. Renovate current data center
   b. Re-purpose another building
3) Build new building
4) Use cloud resources

Option 1 was immediately ruled out. Option 2a entails significant disruptions to current operations, and option 3 (building + 25-yr operational costs) was estimated to cost ~20% more than option 2b. Therefore, options 2b and 4 were compared. Figure 4 is a conceptual design of an existing BNL building re-purposed as a data center to meet the requirements described in section 2. The two options were compared on a hypothetical 3-yr deployment and operational scenario in which cost projections are well understood. Comparisons over a longer time period were discouraged because of uncertainties over cost projections. To simplify the comparison, several assumptions were made: a) local hosting costs (power, cooling and staff) remain constant, b) future computing and storage requirements do not deviate from current forecasts, and c) tape storage (deployment and operations) were not included in the comparison—even though it is an essential component for archival storage at the facility.

3.1 The cloud option

For cloud storage resources, AWS S3 [2] was used to estimate the cost of a 7 PB deployment spread equally (2.33 PB/yr) over a 3-yr period. The estimate does not include the cost of data egress (from S3 back to the internet—current cost varies from $0.05-$0.09/GB). Table 2 (July 2016 prices) summarizes the cost estimate.

Table 2: Cost of S3 storage (July 2016 prices) resources over a 3-yr period.

Year	Cost per PB/month (in US$)	PB/yr	Cost/yr (in US$)	Cumulative cost (in US$)
1	27.5k	2.33	770k	770k
2	27.5k	2.33	1540k	2310k
3	27.5k	2.33	2310k	4620k

For cloud computing resources, the estimate is based on the typical 5,000-core annual acquisition for the ATLAS Tier 1 at BNL. These 5,000 cores are fixed resources over its lifetime, while AWS EC2 [3] resources are upgraded more frequently. For this study, we assume EC2 resources are upgraded annually, and we use prices for the c4.large (spot) instance (July 2016 prices) to calculate the cost of 5,000 core-equivalent to BNL resources (assume a 20% yearly improvement). Table 3 summarizes the cost estimate.

Table 3: Cost of EC2 (July 2016 prices) computing resources over a 3-yr period.

Year	Spot c4.large ($/yr)	Equivalent cores	Cost/yr (in US$)	Cumulative cost (in US$)
1	151.5	3077	466k	466k
2	151.5	2564	389k	855k
3	151.5	2137	324k	1179k

3.2 The on-site option

For on-site storage and computing costs, the estimates are based on actual cost of recent hardware procurements. For storage, the total cost of ownership (TCO) for 7 PB over a 3-yr period is $1300k. TCO includes the capital cost of the equipment and operational costs (staff, power, cooling, etc) of the data center. TCO amortizes the capital cost over the equipment lifespan. TCO for computing depends on the useful lifespan of the worker nodes. The RACF has operated computing equipment with various lifespans. Table 4 shows TCO for two typical cases (3 and 5-yr).

Table 4: Cost of on-site computing (5,000 cores) resources over 3-yr and 5-yr periods. The cost/core is amortized over the lifespan of the computing equipment. Therefore, the longer the lifespan, the lower the cost/core.

Lifespan	Cost/core (in US$)	Data center charges/yr (in US$)	TCO (in US$)
3-yr	70.7	78k	588k
5-yr	63.6	78k	708k

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 082009 doi:10.1088/1742-6596/898/8/082009

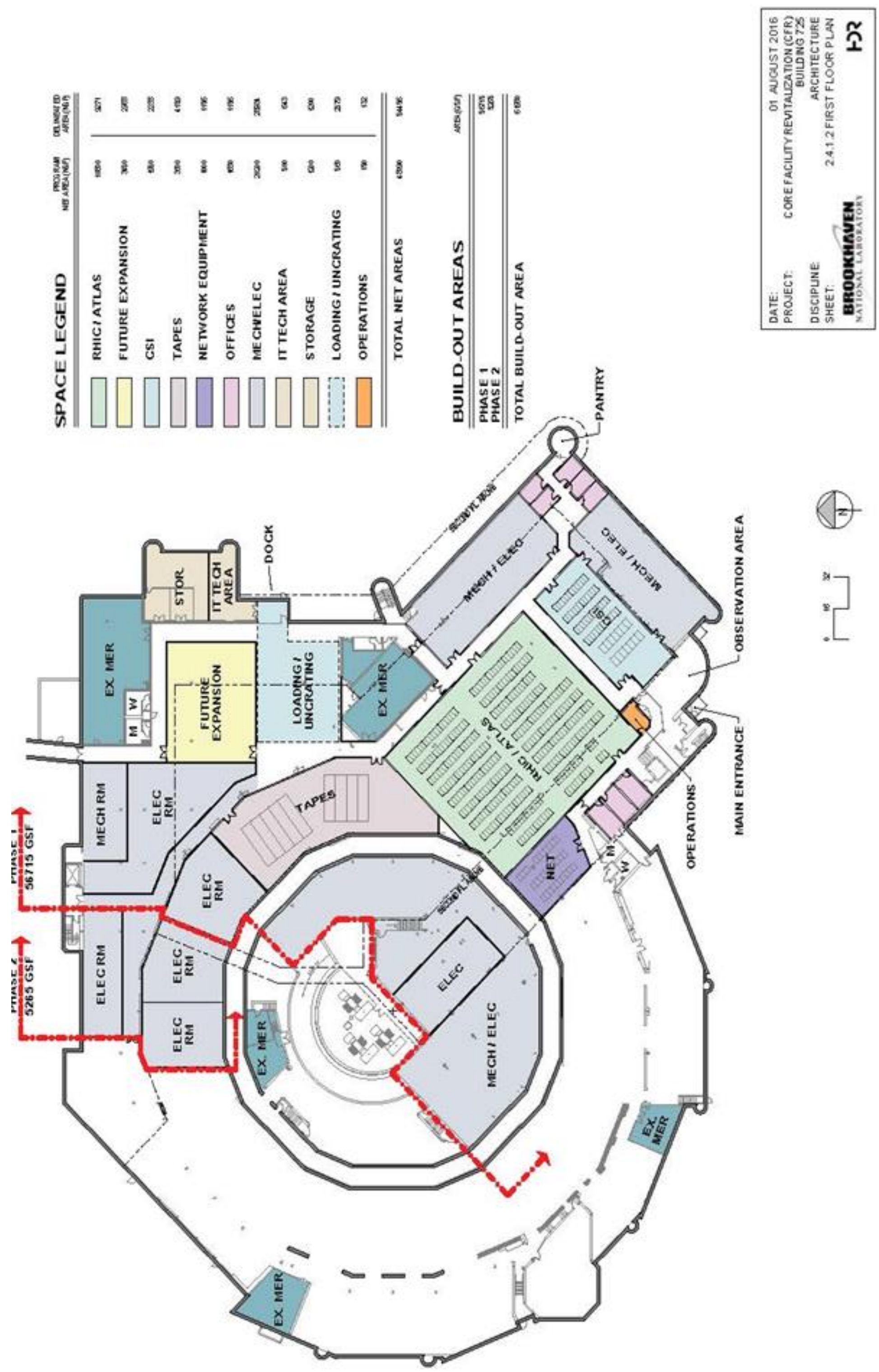

Figure 4: Conceptual design of an existing BNL building re-purposed as a new data center

3.3 Comparative summary and historical cost evolution

The above analysis shows the on-site option is more cost-effective than the cloud (AWS in our case) option. The difference in cost for storage is large ($3.3M), but the difference for computing is significantly smaller ($0.5M - $0.6M, depending on the lifespan). The costs of data movement and replication were not included in the analysis. Using AWS current prices (see subsection 3.1 above) for guidance, one can infer these costs are not insignificant.

The BNL presentation at CHEP 2013 [4] compared the cost of on-site and AWS 3years ago. In this paper, we redo the calculations with updated costs. In 2013, the m1.medium instance was used as a reference, while in 2016, the reference instance was updated to c4.large to match current requirements. Table 5 summarizes the results.

Table 5: The evolution of AWS and on-site costs (see presentation at CHEP 2013 [4] for calculation details).

Year	AWS spot instance (in US$/hr)	AWS on-demand instance (in US$/hr)	On-site (in US$/hr)	S3 storage (in US$ per GB/month)
2013	0.013	0.12	0.021	0.05
2016	0.017	0.105	0.015	0.028

4. The role of cloud resources

Given the comparative disadvantage, where do cloud resources fit within the BNL computing environment? The answer lies in the computing workflow. To make economic sense, the most cost-effective usage of AWS (and other for-fee cloud resources) is for CPU-intensive, low I/O and latency-tolerant workflows, such as Monte Carlo simulation. Other workflows, such as data analysis campaigns (see figure 5) with time-sensitive, intensive computing and I/O characteristics do not fit the demands for scheduling flexibility and fault tolerance that optimize usage of AWS spot instances.

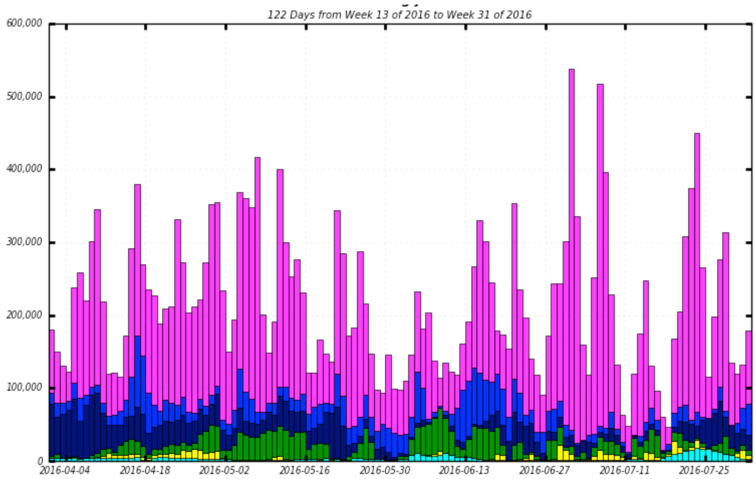

Figure 5: High-priority workflows such as data analysis do not fit the flexible scheduling profile that optimizes usage of spot instances

5. Analysis Summary

BNL's analysis of a hypothetical 3-yr deployment scenario indicates that the TCO of on-site computing and storage resources has a competitive advantage over similar commercial cloud resources. The analysis in 2016 is consistent with a similar analysis done in 2012. Furthermore, historical records at BNL show that the capital cost of re-purposing an existing building (option 2b) is dominated by the significantly larger total TCO (defined as the cumulative cost of IT equipment and related operational costs over the 25-yr expected lifetime of a new data center). Therefore, the competitive advantage is unaffected even if the capital cost of re-purposing an existing building is folded into the TCO of an on-site solution.

The proposal for a new data center is undergoing standard DOE reviews. Critical Decision 0 was (CD-0) granted in the Fall of 2015 on the basis of scientific mission need. A CD-1 review (alternative analysis on lowest cost solution that meets the scientific mission need) was conducted in the Fall of 2016. Preliminary design is expected to occur in 2017, with initial construction in 2018. Initial occupancy is expected sometime in 2021.

Access to cloud resources is and will remain an important issue in scientific computing communities like HEPN. In a flat-funding scenario, near-term computing and storage requirements from existing experiments cannot be met without "external" contributions, such as commercial cloud systems, HPC clusters at Leadership Class Facilities (LCF) and other non-dedicated resources (see figure 6). This is strong motivation to develop and evaluate mechanisms and computing models that leverage cost-effective access to non-dedicated resources, an example of which is the event server approach at ATLAS for AWS [6]. The challenge is to integrate new workflows into the existing computing environment in a cost-effective manner while maintaining dedicated, on-site data centers as the main option for scientific computing activities.

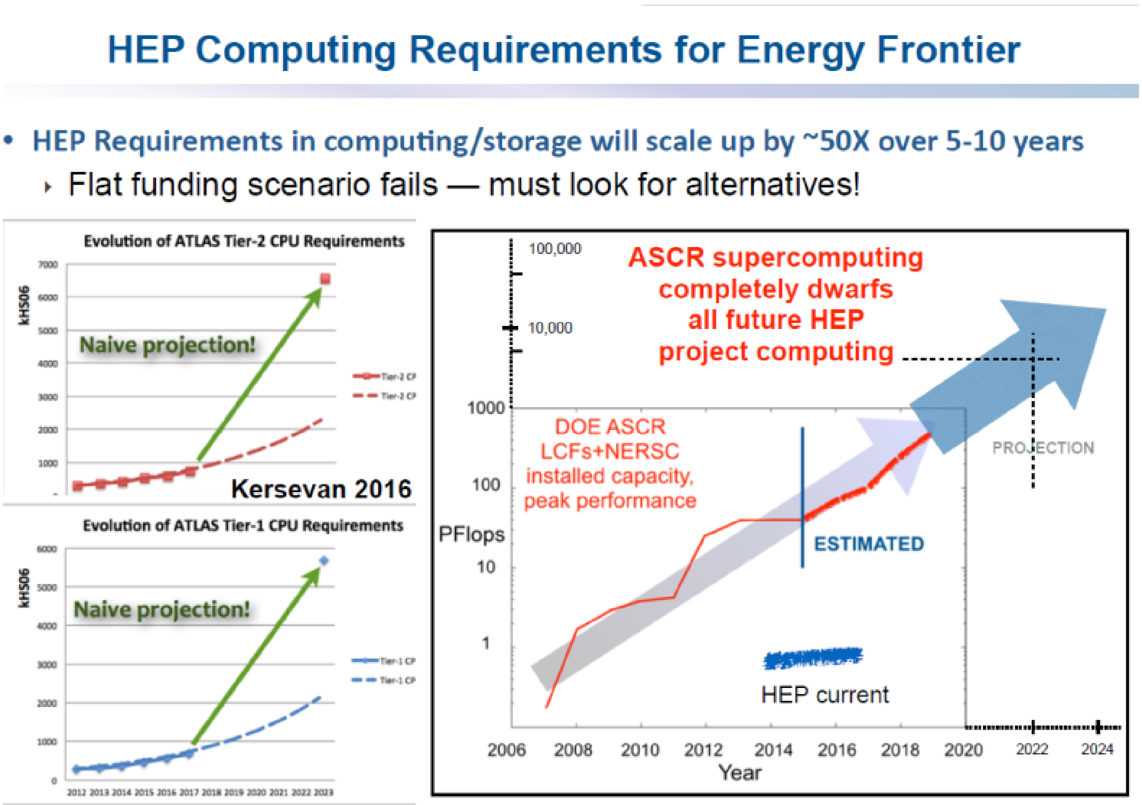

Figure 6: Future deployment of HPC resources are expected to outpace dedicated HEPN resources significantly [5].

References

[1] HS06 is a widely used benchmark within the HEPN community. For more information, see https://www.hepix.org/#section5.
[2] See https://aws.amazon.com/s3/pricing/ for current prices.
[3] See https://aws.amazon.com/ec2/spot/pricing/ for current prices.
[4] See http://iopscience.iop.org/1742-6596/513/6/062053 for published paper proceedings.
[5] See https://www.bnl.gov/nysds16/talks.php (NYSDS 2016).
[6] See https://indico.cern.ch/event/346931/contributions/817790/ (HEPIX Spring 2015).