The Role of Dedicated Data Computing Centers in the Age of Cloud Computing

CHEP 2016 – San Francisco

Tony Wong
Brookhaven National Laboratory
Computing @ BNL (2)

• Existing data center space mostly devoted to RHIC and ATLAS is nearly full
  – 15,000 ft² (~1,400 m²), including new expansion space built in 2009
  – ~2.3 MW of UPS power
• HPC-centric existing space (~2,500 ft² and ~500 kW of UPS power) not sufficient
• Little space and power left for expansion to support new programs at BNL
  – Center for Functional Nanomaterials (CFN)
  – Computational Science
  – Others
Cost per HS06 is trending down
Fast-rising computing requirements

![Graph showing the increase in logical cores from 2013 to 2017 (est). The categories are Other, HPC, ATLAS, and RHIC. The number of logical cores increases each year, with the most significant increases seen in the HPC category.]
New Data Center Profile

• 25,000 ft$^2$ (~2,320 m$^2$) of usable space
• 2.4 MW of UPS power on day 1 (expand up to 6 MW in future)
• PUE of 1.2 to 1.4 (mandated by DOE)
• Shared facility for ATLAS, RHIC, CSI, Photon Science
• Natural air-cooled supplemented by redundant chillers
• Hot-aisle containment
External Considerations

- DOE mandate to prioritize “Cloud” as alternative to building new data centers
- Current budgetary realities and program requirements have compelled the HENP community to evaluate off-site alternatives, independent of DOE mandate
- Commercial providers (Amazon, Google) offer increasingly price-competitive cloud services
- Virtual (non-profit) organizations (ie, OSG) are harnessing the compute power of non-dedicated (HTC and HPC) resources
Alternative Analysis for CD-1

• Four scenarios considered
  1. Do nothing
  2. Utilize existing BNL facilities
     a. Renovate current data center
     b. Re-purpose another building
  3. Build new facility
  4. Use cloud resources

• Compare two most cost-effective solutions (options 2b and 4) on a hypothetical 3-yr deployment and operations scenario

• Several assumptions made to simplify calculations
  – Local hosting (power, cooling, staff, etc) costs remain constant
  – Future requirements do not deviate from forecast estimates
  – Tape storage (capital and operations) not included—even though it is essential component of archival storage at RACF
Cloud vs. in-House cost evolution

- BNL presentation at CHEP 2013 in Amsterdam ([http://iopscience.iop.org/1742-6596/513/6/062053](http://iopscience.iop.org/1742-6596/513/6/062053))
  - Computing
    - $0.013/hr (m1.medium spot instance)
    - $0.02/hr (RACF)
    - $0.12/hr (m1.medium on-demand instance)
  - Storage
    - $0.05/GB/month
- Current AWS costs (as of July 2016)
  - Computing
    - $0.017/hr (c4.large spot instance)
    - $0.015/hr (RACF)
    - $0.105/hr (c4.large on-demand instance)
  - Storage
    - $0.0275/GB/month
- Note: switched to c4.large instance to match current requirements
Summary

• In-house cost-competitive with cloud resources
  – True over past ~4 years – confident it will hold true over 25-yr lifetime of data center
  – Irreducible cost of hardware makes up ~70% of Total Cost of Ownership – hard floor to any further competitive gains at BNL or elsewhere

• Access to cloud resources still important
  – Upcoming HEP computing/storage requirements cannot be met without “external” contributions
  – In-house competitiveness depends on volatile factors (cost of electrical power, infrastructure support, etc) and cannot be taken for granted as enduring advantages
  – Motivates the development of mechanisms and models for cost-effective access, such as event server to use AWS spot pricing