



BABAR Status

DPHEP Workshop
June 21st, 2021

C. Cartaro



Outline

- The *BABAR* Collaboration today
- Publications and analysis
- Migration from SLAC
- Open data access



The *BABAR* Collaboration

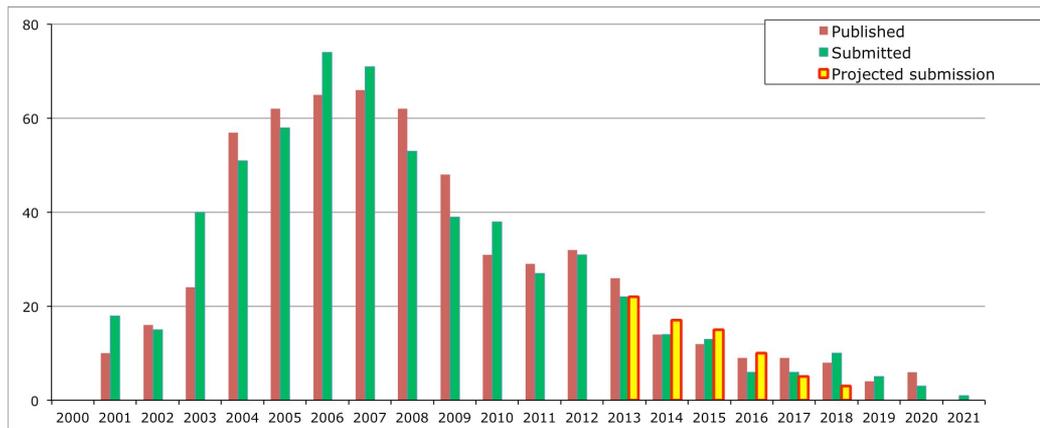


- ~241 Members in the semi-frozen author list
 - Four 2- or 3-days virtual Collaboration Meetings per year
- Includes PhD students in Canada, Germany, Israel, Italy, Russia and the U.S.
 - Postdocs analysing *BABAR* data are funded on other projects
- Continuing to bring in undergraduate students using *BABAR* data
 - Expected this to continue until the data is not superseded by Belle II (~700/fb by July 2022)
 - Y(3S) sample will not be superseded in the foreseeable future
 - Data also useful for detector studies and preparations for large Belle II data set
- Theorists and scientists are welcome to join as Associates to mine the data to test new ideas while working with regular *BABAR* Members taking advantage of our full knowledge of the data and the software
 - True OPEN ACCESS to *BABAR* data
 - The Collaboration also approved to pursue Open Data CERN/DPHEP-style (more on this later)

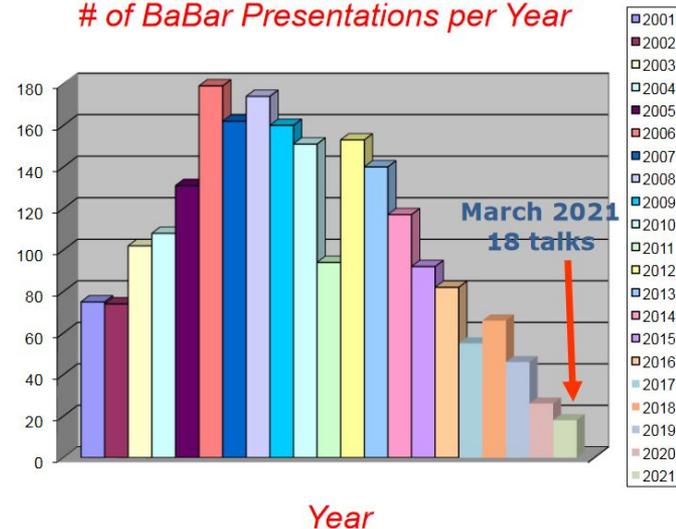


Publications

- 595 papers published or submitted
 - 9 papers published in 2017, 8 in 2018, 4 in 2019, 6 in 2020
 - 3 in the pipeline so far in 2021, few more expected later in 2021
- ~15 analyses active and on track for publication
 - Some are progressing slowly
 - 6 new analyses started last year and expect some more this year
- 25 talks in 2021
 - 7 talks at EPS-HEP, and more already assigned
 - 26 talks given in 2020 (17 cancelled due to COVID-19)
 - Often shared talks (and collaborative analyses) with Belle
- Quality of physics results still excellent



of BaBar Presentations per Year

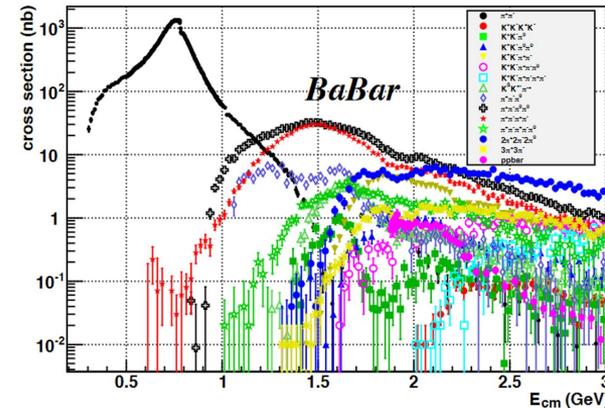
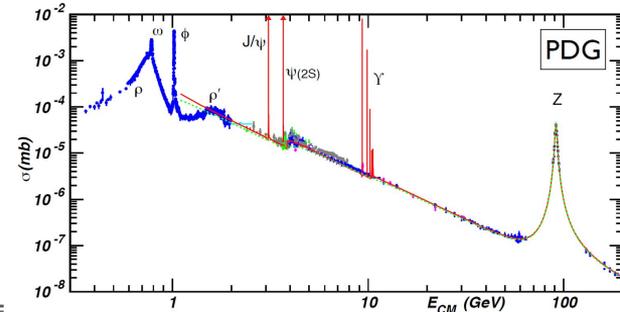




Science

- Dark sector and light dark matter searches
 - Extensive program to search for dark sector signatures, and set stringent limits on their existence
 - Recently set world leading constraints on leptophilic dark scalar
 - More searches ongoing
 - Self-interacting dark matter, dark matter and baryogenesis, heavy neutral lepton, ALP
- Significant contributions to hadron physics in ISR
 - Access mass range from threshold to 4.5 GeV with high resolution (4-15 MeV)
 - Hadronic cross section determination and improved understanding to the muon anomaly dispersion integral
 - Goal: measure all the modes - now looking at 7 hadrons final states
 - High energies well described by perturbative QCD, low energies dominated by $J^{PC} = 1^{--}$ resonances
 - Most precise measurement of $e^+e^- \rightarrow \pi^+\pi^-$, new independent analysis ongoing on the full dataset with lower systematics associated to PID
- Studies of conventional and exotic charmonium
- τ - μ lepton flavour universality in $Y(3S)$
 - Most precise measurement to date
- $B \rightarrow D^* l \nu$ with angular analysis
 - $|V_{cb}|$ tension between inclusive and exclusive studies, NP from FF
- Rare/forbidden semileptonic D_0 decays
 - Better than LHCb for D_0 decays with electrons in final state
 - First observation of $D_0 \rightarrow K^- \pi^+ e^- e^-$

On resonances:	
$\Upsilon(4S)$:	433 fb ⁻¹
$\Upsilon(3S)$:	28 fb ⁻¹
$\Upsilon(2S)$:	14 fb ⁻¹
Off reson./scan:	54 fb ⁻¹
Total:	529 fb ⁻¹





Moving Away from SLAC

- SLAC computing is preparing to support large projects like LCLS II and the Rubin Observatory data facility
- It was decided that supporting *BABAR* servers and the Long Term Data Access enclave(*) was too much effort for an old Collaboration. All hardware, except essential systems, was put in run-to-fail mode and we were given one year to copy the data and move the services outside of SLAC
 - The last day for *BABAR* at SLAC was supposed to be Jan 29th 2021 but we were given until May due to COVID-related delays
- We had to find robust and simple solutions to replace our collaborative services, find a new home for the data, and organize some system to analyze the data and complete the current physics program
- Luckily, *BABAR* has many friends and we found a lot of support

(*) See backup slides



New and Old Homes for the *BABAR* Data

- The dataset at SLAC is still on tape but it is on old T10k-D tapes and they will soon become unreadable
 - Raw data has become inaccessible due old staging software no longer maintained (last full reprocessing was in 2008)
- A second copy of the full data set has always been hosted at CC-IN2P3 (~2PB)
 - At this point CC-IN2P3 is hosting the only existing copy of the raw data
 - MOA with CC-IN2P3 expired, working on a new one: store data until 2025, with the option to renew the agreement until 2030
 - Access to the data limited to few *BABAR* people for data recovery purposes or sync of new data (MC signal production)
- GridKa hosts our federated XRootD redirector, data production databases, and offered to host the data and MC needed by analyses
- CERN offered to host the data for possible use in the Open Data Portal in the long term
 - ***BABAR* is now a CERN experiment!!!**



Collaborative Tools

- Main website moved to University of Victoria together with our main infrastructure
 - Official software releases, repositories, wiki, read-only legacy discussion forums (HyperNews)
- Collaboration forums moved to CERN E-groups
 - Works also with lightweight accounts
- Collaboration Meeting timetables now in CERN Indico (connected to E-Groups)
- Archived/published analysis internal documentation stored in InSpire as private collections, plus local copy at U. of Victoria
- Active membership tracking, on-going analysis tracking (including internal reviews), paper reviews tracking, collaboration talks rehearsals, and anything that requires active tracking is now done through Google Drive tools and documents (free and easy to share and manage)
 - We are not yet 100% sure that we won't lose any data once our Oracle DBs at SLAC will be shut down



Analysis Infrastructure

- A new system is not ready yet
- It will be hosted at U. of Victoria
- Similar to the LTDA
 - Login machines
 - Home dirs and work space on NFS
 - VMs used to run jobs
 - Working on infrastructure
 - User access, sys administration, networking, monitoring, repositories access, batch system, MC production system
- Preparing shipment of a sub-set of LTDA servers from SLAC to U. Victoria to rebuild the system and provide more batch resources and disk space for the analysts
 - It was a long approval process but the program managers at the DOE HQ were very supportive
- Meanwhile, we can download and use a single user VM (BaBar-To-Go) with the last analysis release installed and direct access to data at GridKa through xrootd
 - Originally developed in 2013, SL6, works with qemu and VMWare on any laptop



Open Data Access

- In December 2020, the LHC collaborations signed-off on an Open Data Policy presented to CERN Council and it received huge support.
 - Press release:
<https://www.interactions.org/press-release/cern-announces-new-open-data-policy-support-open-science>
 - [CERN Open Data Policy for the LHC Experiments](#)
- We considered in particular, the section of policy related to ‘Level 3 Data’ is in the “Additional Material” (Level 3 data = reconstructed data useful for particle physicists to undertake new analyses)
- We recognize that BABAR does not have the capacity of providing many resources for implementing L3 data access, and once implemented we may not have the resources for long term support and we’ll have to rely on CERN support
 - We will certainly support and encourage communication on a best-effort basis
- **The vote of the *BABAR* Council:**
 - Yes, the *BABAR* Collaboration want its data available for future analyses through the CERN Open Data Portal
 - When: open some data in 2021, the full dataset to come later.
 - Timing depends on our use of it and Belle II data size



Summary

- The Collaboration is healthy and working as hard and diligently as always and there is still much to do ahead of us
- *BABAR* will join the Open Data Portal
- Support for *BABAR* at SLAC has stopped, machines are cold, tapes dormant
- We had to get creative at times but we got many things going again and we received a lot of support from many institutions
 - CERN, CC-IN2P3, GridKa, U. of Victoria
- The last year has been a bumpy ride for more than one reason... but I think we should be grateful for all the work we were able to do and for the help received
- I have a personal list of “Thank yous” to express but I fear for it to be quite long, involving many colleagues and friends at the institutions above. So...

Thank you!



Questions?



Backup: LTDA (I)

- The *BABAR* Long Term Data Access (LTDA) project was dedicated to the preservation of the *BABAR* data access and computing environment
 - Provides support of data, code, repositories, databases, documentation, storage, and CPU capacity
- Enclose *BABAR* Analysis Framework into a frozen environment
 - Keep all potential of the Framework including discovery potential
 - Simpler to maintain documentation and provide support
- Minimize the effort needed to maintain the system
 - Validations, upgrades, documentation, hardware
- The *BABAR* LTDA was an integrated cluster of computation and storage resources. The cluster used virtualization technologies to ensure continued operation with future hardware and software platforms, and utilized distributed computation and storage methods for scalability.
- Stability: the cluster was able to run for long periods of time without human intervention and it was able to recover even from complete outages with minimal or no human help.

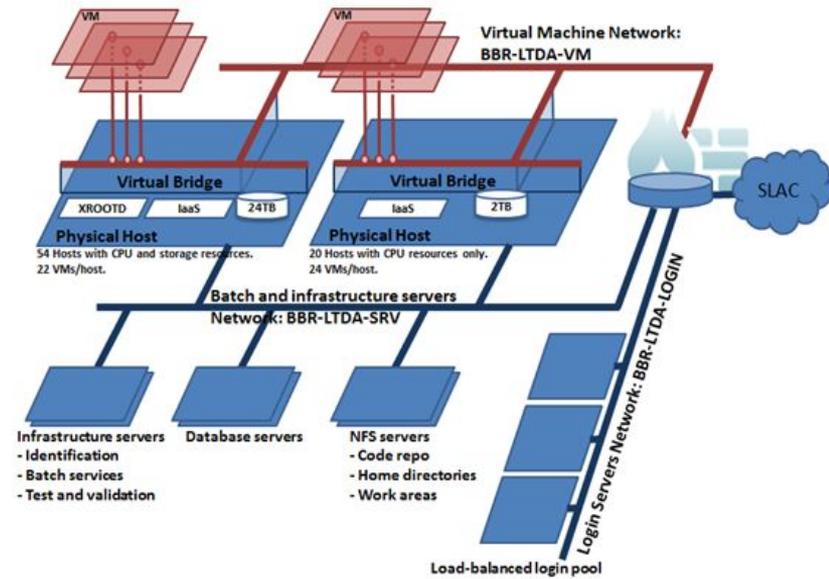


- 9 infrastructure servers
- 75 batch nodes and storage servers
- 2 NFS servers

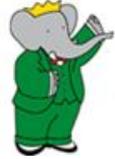


Backup: LTDA (II)

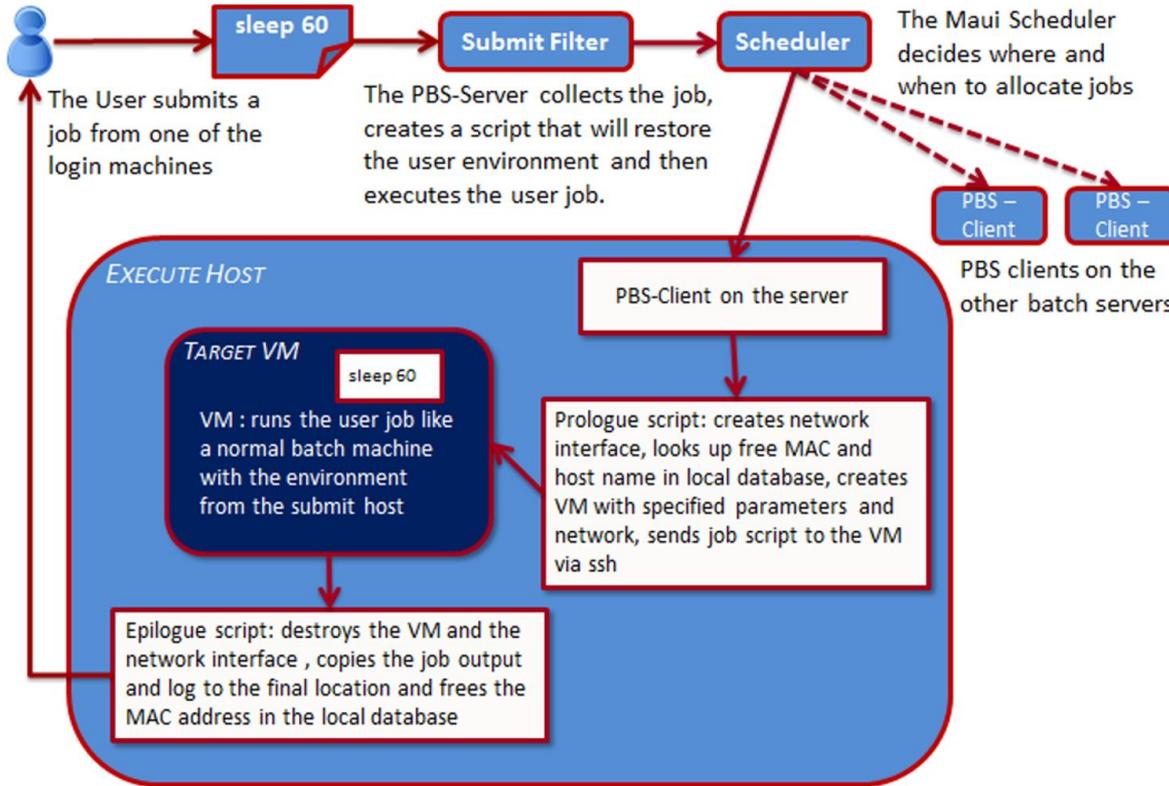
- A simple solution
 - Mix CPU and storage resources on the same node
 - Each machine is a simple building block
 - Standalone capacity
- Scalable and portable design
 - All inclusive design easy to reproduce at any scale in any of our collaborating institutions
- Security and flexibility all in one
 - Stable against disk or node loss
 - Isolation of all back versioned elements
 - Completely on demand
- Facts
 - 1.2 PB XRootD-based storage for physics data
 - 64TB ZFS-based storage for user data
 - 3.5 TB of RAM
 - 1668 job slots
 - SL4, SL5, SL6 platforms available in VMs (qemu+kvm)
- In production from March 2012 to March 2021



- VMs are not allowed to connect to SLAC network or the world. The Login network is protected from the VM network
 - Allow one way ssh from Login to VM network
 - Password less connection via shosts
 - VMs are not allowed to write over the Login network
- Well defined services between VM network and SRV network
 - Infrastructure (DNS, LDAP, NTP), file service (Xrootd, nfs), batch scheduling
 - LDAP is a subset of the SLAC Kerberos list mapped on NFS internal home directories
- Allow SRV and Login networks use SLAC infrastructure



Backup: LTDA (III)



- PBS/Torque was used to manage the batch resources and Maui was the batch scheduler
- PBS Prologues and Epilogues scripts were used to create and destroy the VMs and the needed network environment
 - Home grown system developed by BABAR
 - Create the network interface for the VMs
 - 24 MAC addresses per host and usage status stored in local DB