Intro & STAR DBs Overview

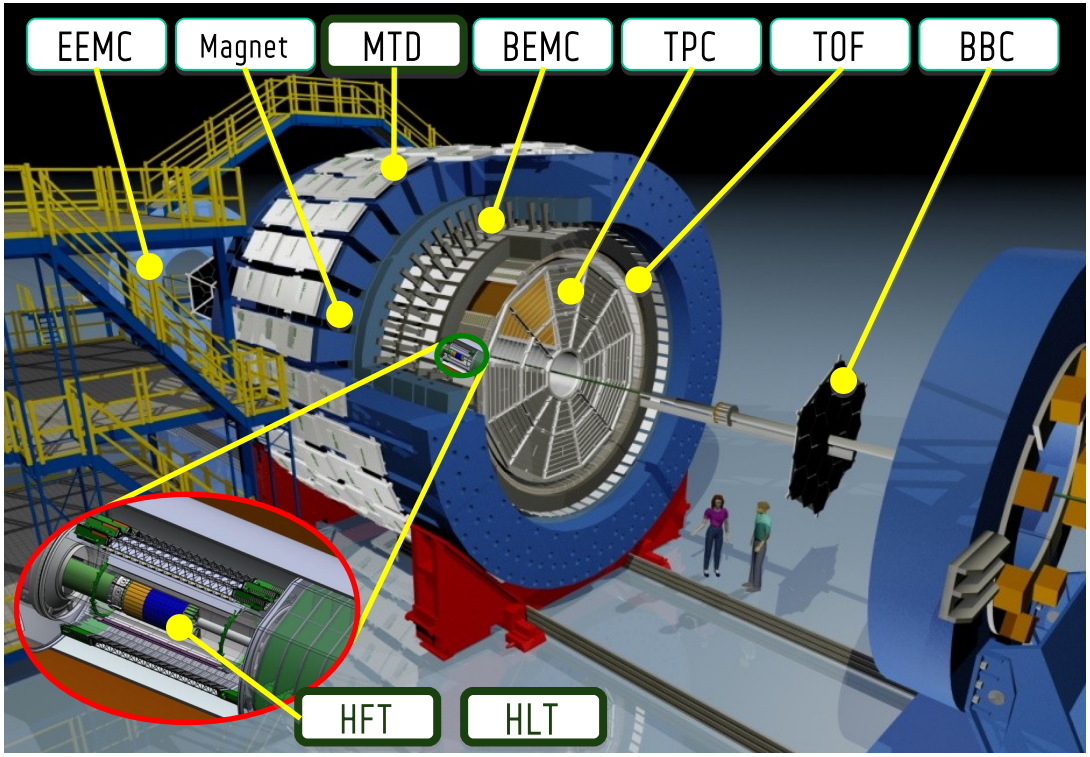
Dmitry Arkhipkin

NPPS group meeting

2019-06-05

BROOKHAVEN
NATIONAL LABORATORY

STAR @ RHIC



Solenoidal Tracker at RHIC

STAR detector:

- Operating since 1999 (till 2025)
- 22 subsystems and growing
- Detector Control System is EPICS-based, having over 60k process variables
- Data taking rate: ~2kHz (started at 1Hz!)
- Colliding species: AuAu, CuCu, pp

STAR Collaboration:

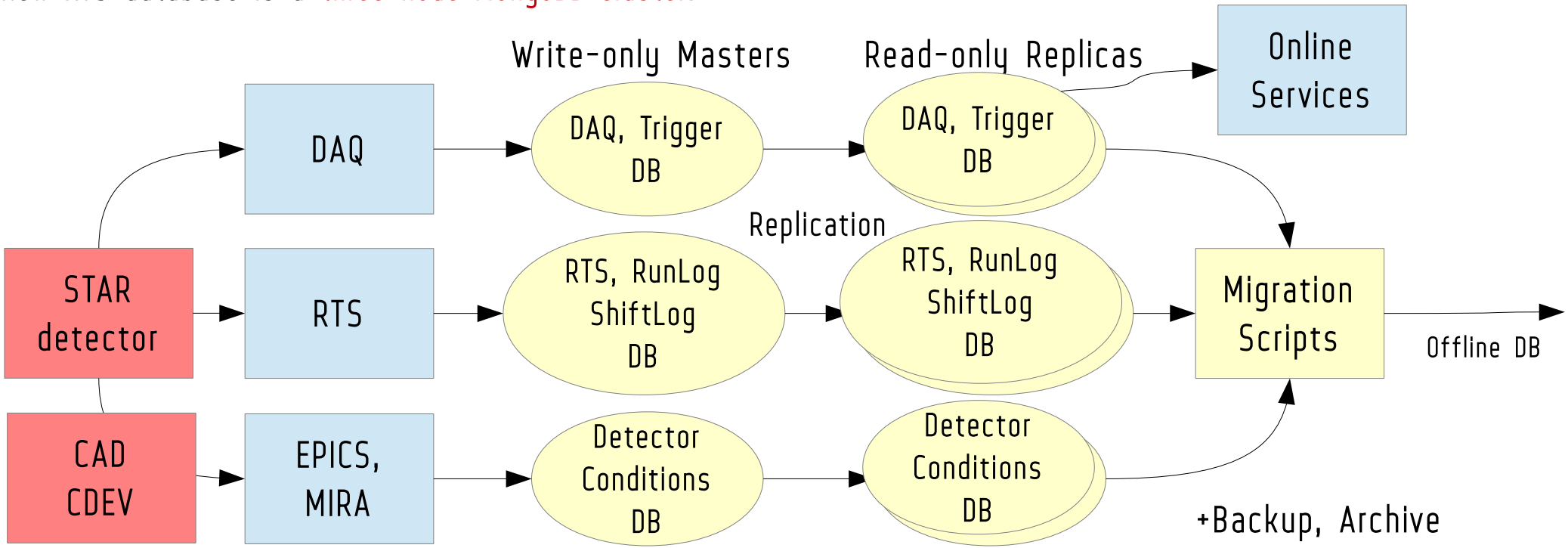
69 institutions from 14 countries, with a total of ~680 collaborators.

My current responsibilities

- STAR Databases
  - Online: Conditions, RTS, RunLog, Shifts
  - Offline: Calibrations, Geometry (+API)
  - FileCatalog
  - Software Infrastructure
- STAR Services
  - MIRA: SCADA Framework
  - SKM: SSH Key Management
  - PhoneBook: Collaboration Record Keeping
  - Shift Signup & Accounting
  - Experiment's RunLog
  - Online Event Display
- Misc Tools and Interfaces
  - DB Interfaces: Monitor, Browser, Explorer
  - Author tools: author lists (LaTeX, Inspire)
  - Online Service Aggregator
  - jobStat: nightly tests UI
  - dbPlots: Conditions DB archive viewer
  - dbSlice: db readiness checker
  - talkstats / simstats
  - Online-to-Offline data migration scripts & monitoring tools
  - Drupal modules: STAR papers, meeting, conference etc

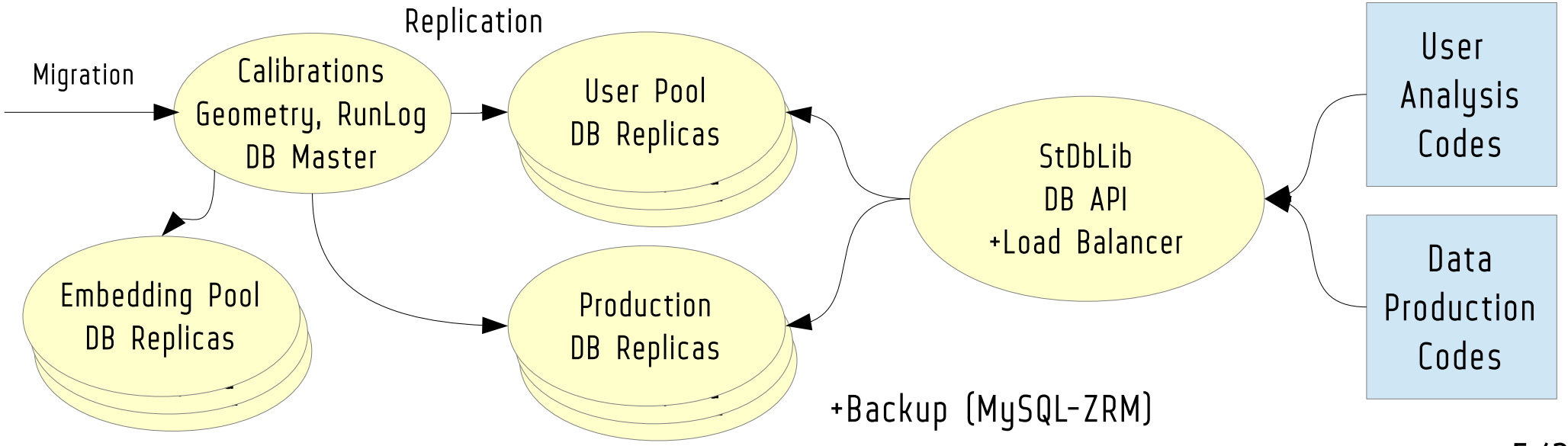
Online Databases

"Online" Databases: used during data taking, optimized for fast writes, not fully structured. MySQL: two master servers containing four independent db instances, four slave servers. Each replica contains all online databases. New RTS database is a three-node MongoDB cluster.



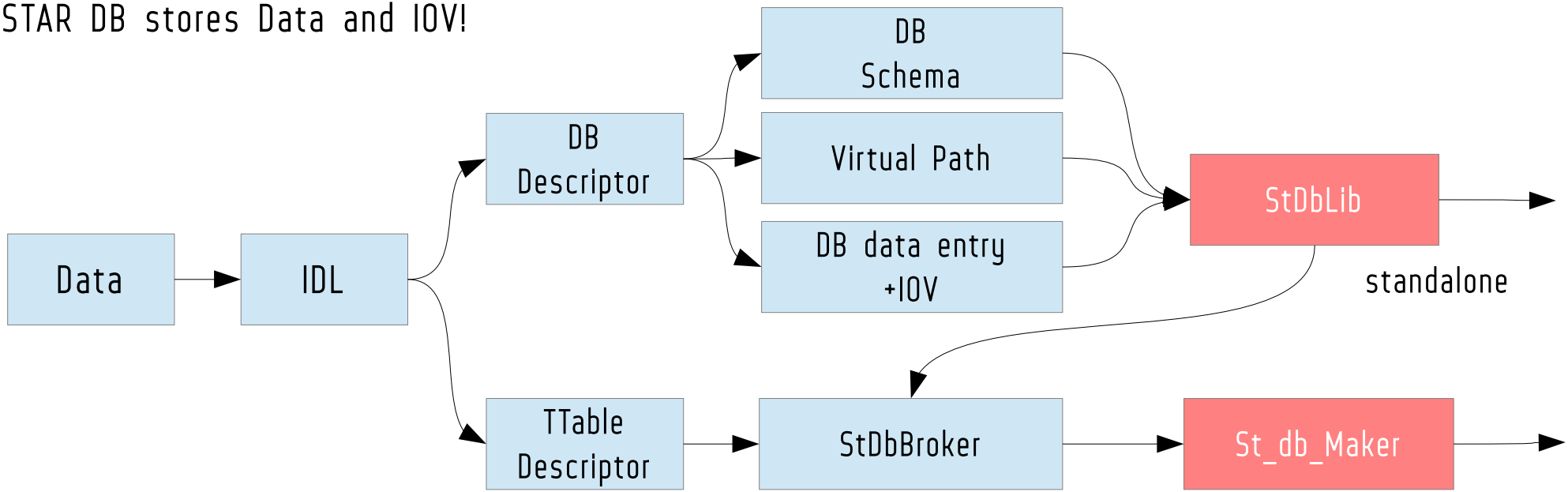
Offline Databases

"Offline" Databases: Calibrations and Geometry Databases used during data production. Highly structured and optimized for fast reads. Replicated setup: single MySQL master, 15 MySQL slaves (three groups). Load Balancer is built into the client DB API (StDbLib, cpp). Database is not a file lookup service but data distribution service (+descriptors). Highly optimized for performance: peak load of 150k queries per second was handled without interruptions. Routine average load is ~20k queries per second.



Offline Databases: format & API

STAR DB stores Data and IOV!



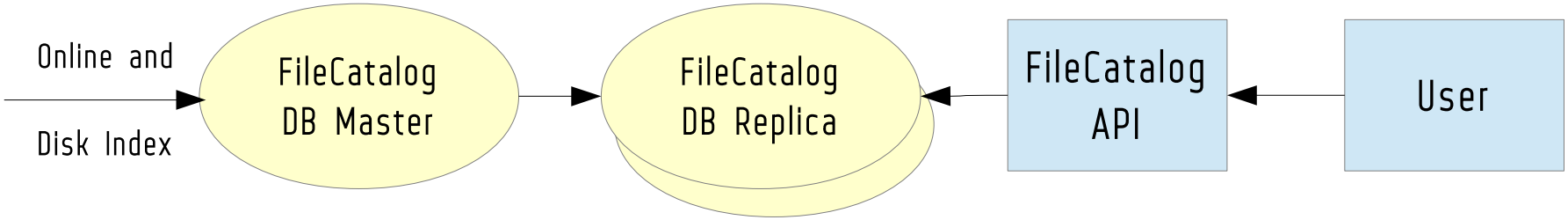
- Every data bit has its own Validity Range
- Data is requested via Event Timestamp + /full/path/to/the/entry
- Three time tags: beginTime, entryTime, deactiveTime
- Complete reproducibility: constrain entryTime and get db state as it was at time X

Offline DB: clusters & clouds

- KEY FEATURES:
  - Easy to maintain: just one service to maintain – MySQL master + N replicas. No separation between file servers and IOV servers. Maintainable by just one person bottom up (online to production).
  - MySQL replication allows near-perfect horizontal scalability, so if performance is a bottleneck, just add more servers to the pool to accommodate for the increased load. Commodity hardware is fine, no need for a super-beefy servers.
  - Client-based load balancing allows simple local LB configuration setups
  - MySQL Query Cache is the only cache, and it is update-aware, no cache expiration time inconsistencies, ~95% efficiency
- CLUSTERS:
  - MySQL is fairly easy to setup (incl. replication), so new cluster setup is not too complicated. Instant replication ensures 100% real-time data propagation across servers;
  - Load is not an issue: add as many db replicas as needed in no time;
- CLOUDS:
  - (from STAR experience) Bring DB server along with your jobs, use it as local server.. One year of STAR db data is ~5GB, no exascale-sized db needed if properly maintained ;)

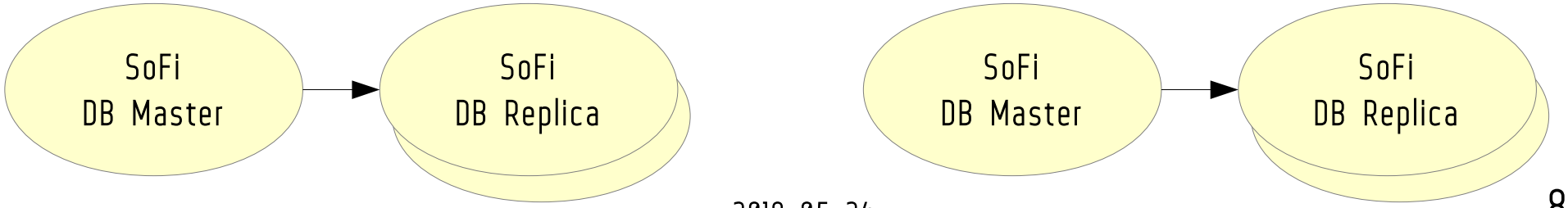
FileCatalog & SoFi databases

"FileCatalog" Databases: contain locations of all BNL-hosted files (HPSS, XROOTD, Distributed Disks) MySQL, one master, two replicas, optimized for frequent updates.



+Backup (MySQL-ZRM)

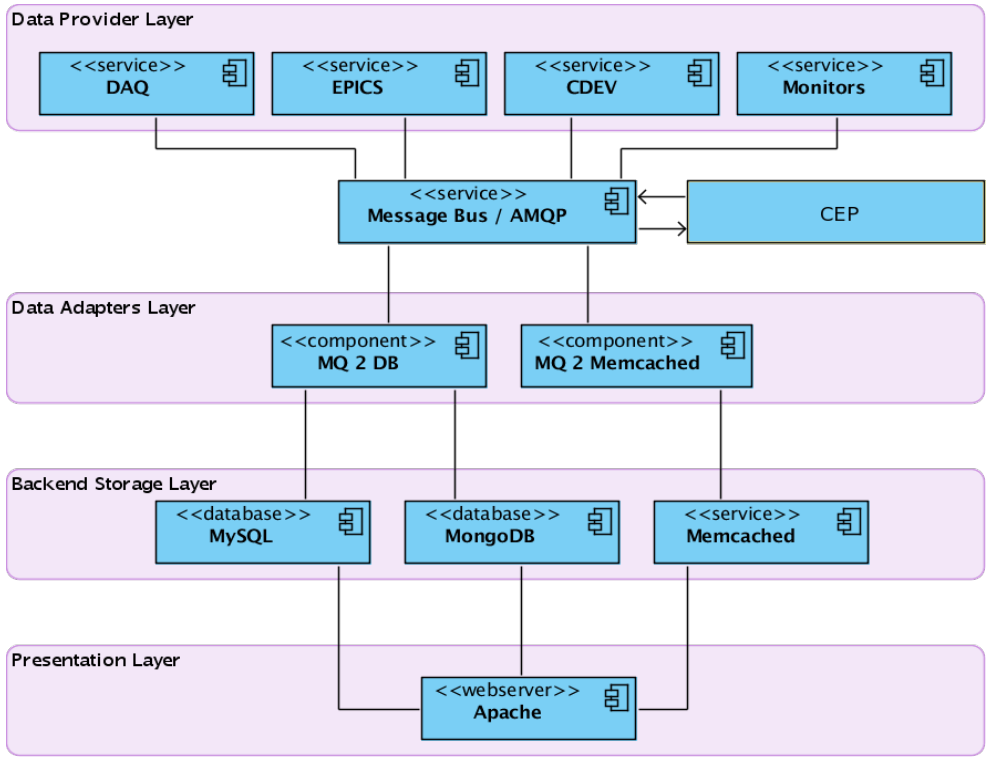
"SoFi" Databases: various Software Infrastructure databases. Loggers, monitoring, web services, SKM, file statistics, user activity stats etc. MySQL, several pairs of "one master, one replica" setups.



MIRA: SCADA Framework

Messaging Interface and Reliable Architecture

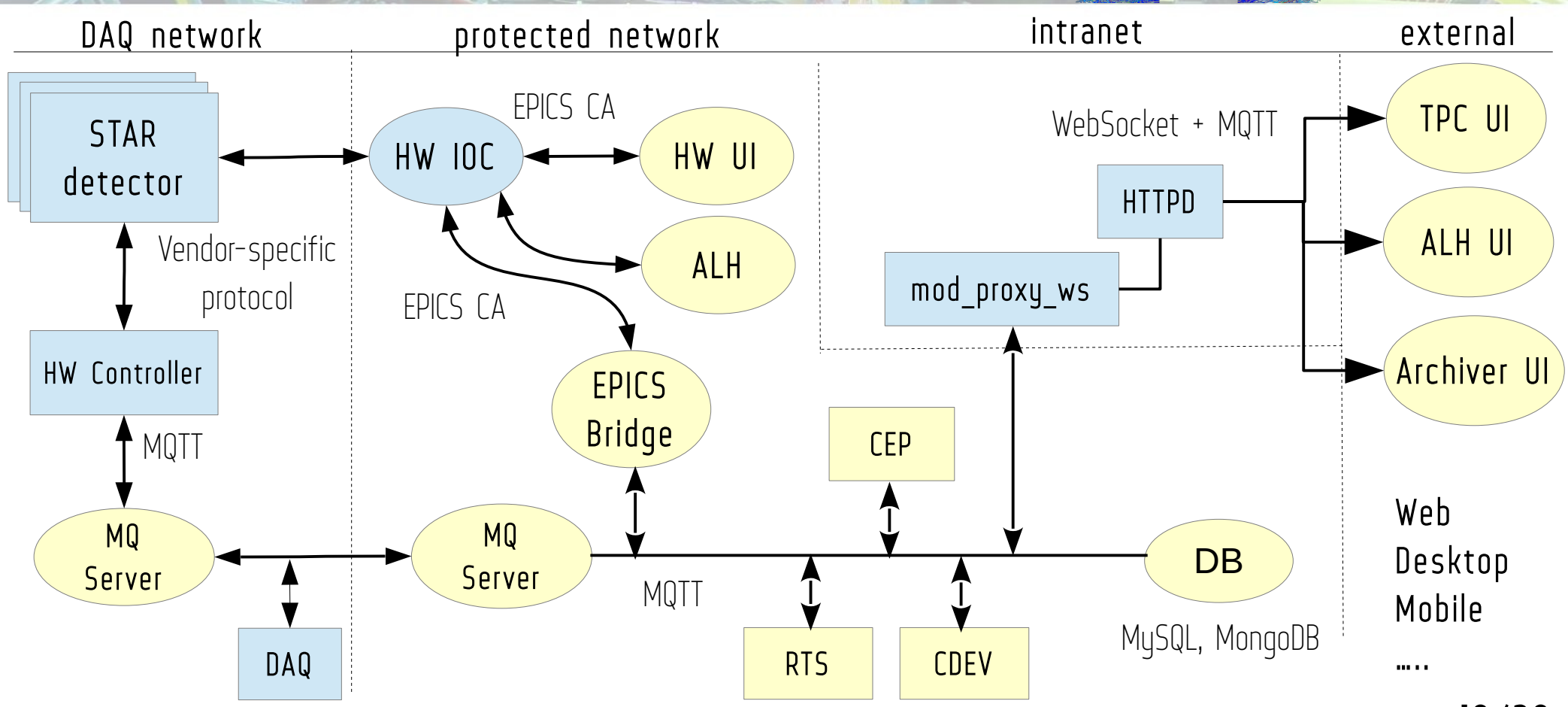
- Features:
  - Scalable architecture
  - Inter-operable, low-overhead protocol
  - Payload-agnostic messaging
  - Quality of Service regulation
- Originally designed to implement better meta-data collection (archiver) and provide basic service messaging bus
- Implemented using Message-Queuing service bus - AMQP, later MQTT
- Supports Complex Event Processing (CEP)
- With time, expanded to the Control System realm and Alarm Handling



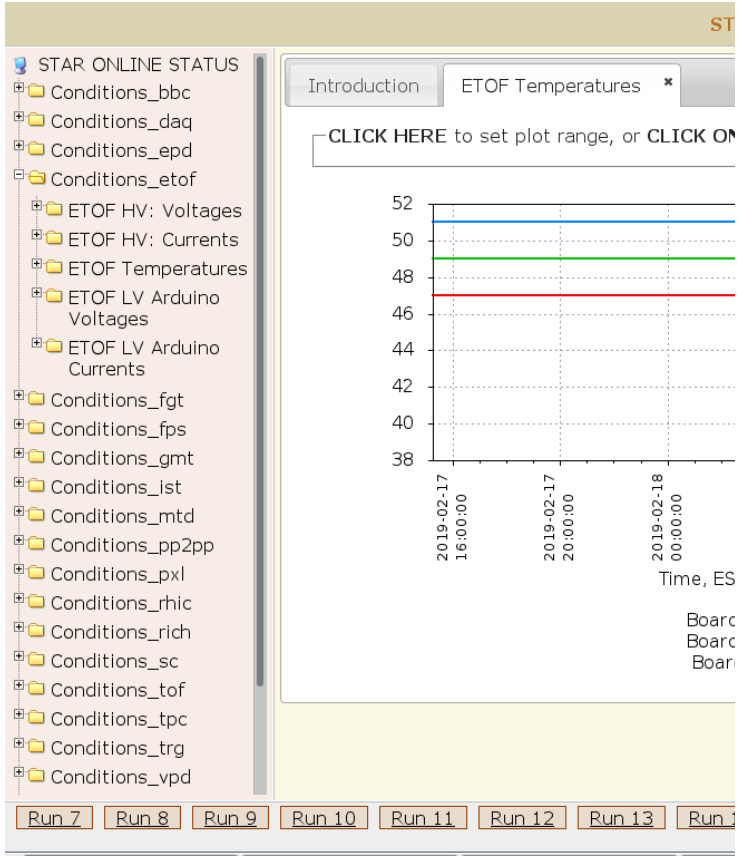
MIRA: basic components overview

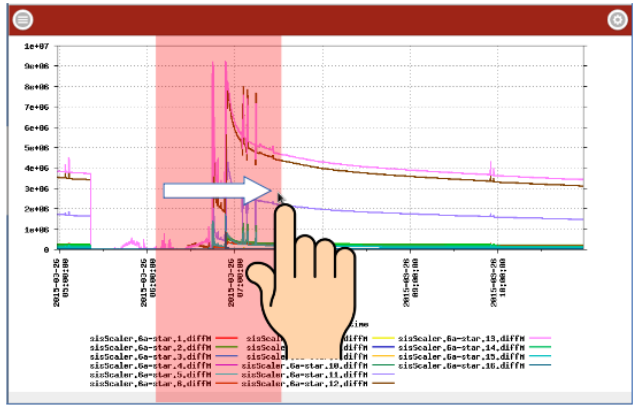
D Arkhipkin and J Lauret 2015 J. Phys.: Conf. Ser. 608 012036

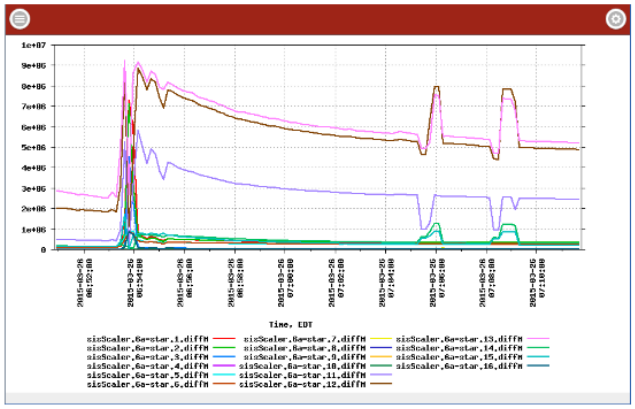
MIRA: Scada Framework

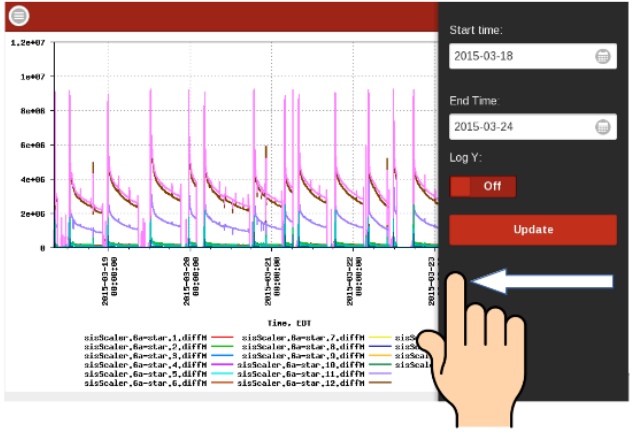


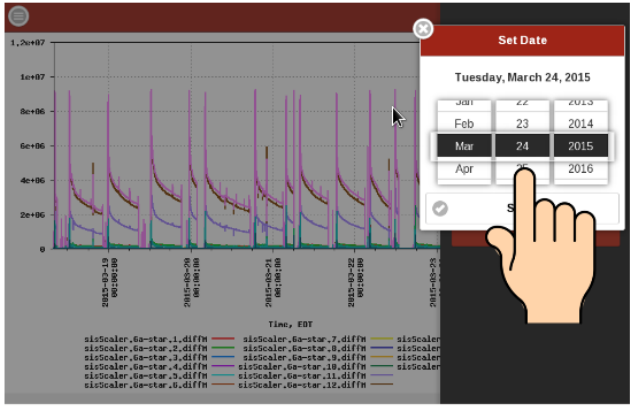
MIRA: Archive Viewer







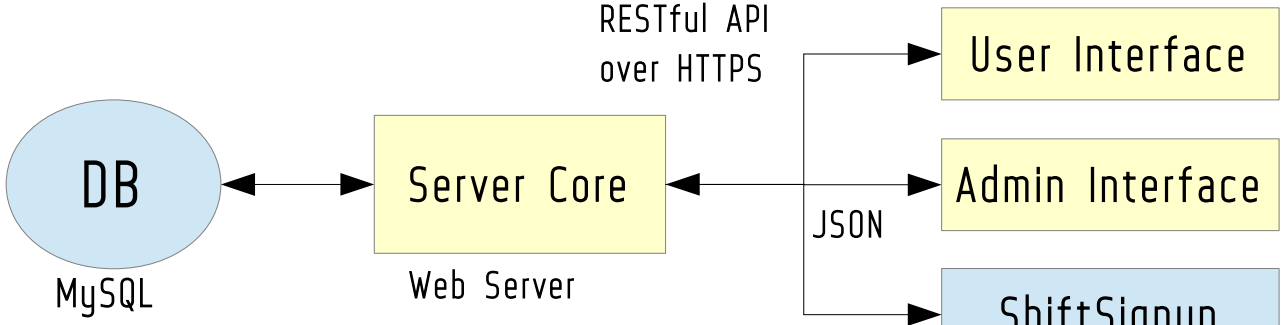




Experiment's PhoneBook

https://www.star.bnl.gov/pnb/client/

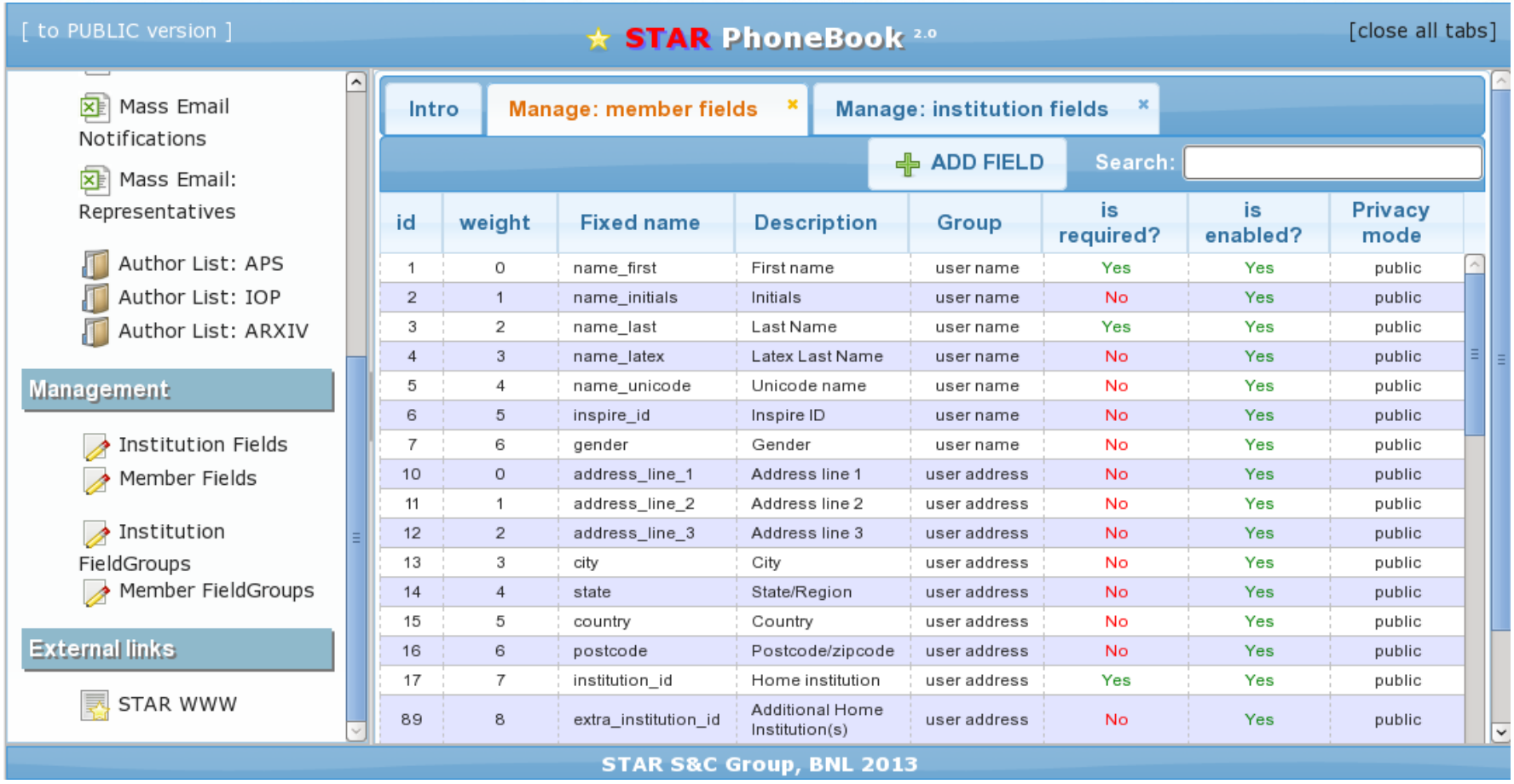
- MySQL database backend (EAV model, schema-free) which has detailed historical information on every member of STAR collaboration, back to y1998. New fields could be configured on a fly without any interruption of service or database schema updates
- Modern user interface, which is more than just interface. Its HTML5 frontend is a client app, written in JavaScript
- Server core, exposing RESTful API (single source of data) for all possible clients: PhoneBook, ShiftSignup, Disk Space allocators etc..



Clients: cpp, php, js, python

PhoneBook UI

https://www.star.bnl.gov/pnb/client

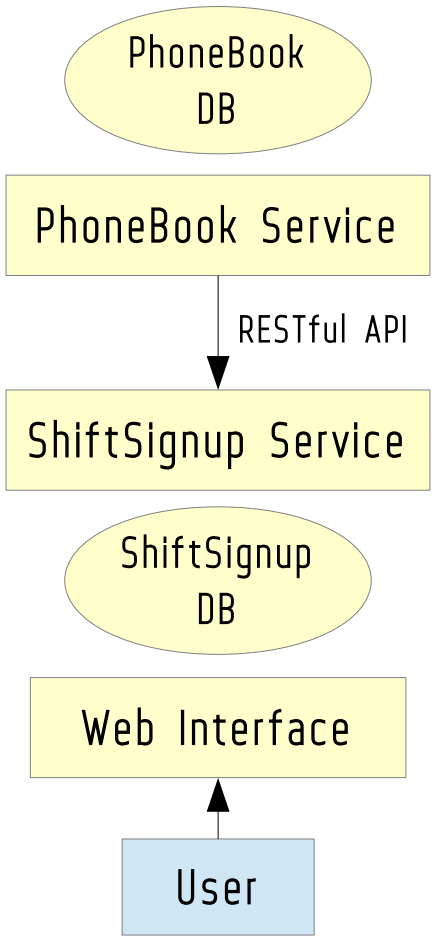


id	weight	Fixed name	Description	Group	is required?	is enabled?	Privacy mode
1	0	name_first	First name	user name	Yes	Yes	public
2	1	name_initials	Initials	user name	No	Yes	public
3	2	name_last	Last Name	user name	Yes	Yes	public
4	3	name_latex	Latex Last Name	user name	No	Yes	public
5	4	name_unicode	Unicode name	user name	No	Yes	public
6	5	inspire_id	Inspire ID	user name	No	Yes	public
7	6	gender	Gender	user name	No	Yes	public
10	0	address_line_1	Address line 1	user address	No	Yes	public
11	1	address_line_2	Address line 2	user address	No	Yes	public
12	2	address_line_3	Address line 3	user address	No	Yes	public
13	3	city	City	user address	No	Yes	public
14	4	state	State/Region	user address	No	Yes	public
15	5	country	Country	user address	No	Yes	public
16	6	postcode	Postcode/zipcode	user address	No	Yes	public
17	7	institution_id	Home institution	user address	Yes	Yes	public
89	8	extra_institution_id	Additional Home Institution(s)	user address	No	Yes	public

Shift Signup & Accounting

https://online.star.bnl.gov/ShiftSignup/

- Features:
  - Highly-configurable Shift Signup and Accounting tool. Integrated with STAR phonebook. Provides detailed overview of STAR shift crews and Online QA shifts, contains expert list.
- Administrative Features:
  - Semi-automatic shift dues calculation per STAR institution for each RHIC Run. Manual override for shift assignments. Variety of summary tables.
- Accounting Features:
  - Automatic checks for BNL mandatory shifter trainings, statistics of shift dues per institution, special shift dues calculations for experts



Shift Signup & Accounting UI

https://online.star.bnl.gov/ShiftSignup/

STAR SHIFTS

Shift Signup
Run 19

(1) To Signup: First select your Institution and Name, then choose a signup sheet
(2) To view schedules: choose a signup sheet

--- Institutions ---

------ People ------

Compact menu

Sign-up Sheets: Experiment Operations | Offline QA | OR | Shift Table: reduced view

Print | PDF | Institutions | Graphs | Controls

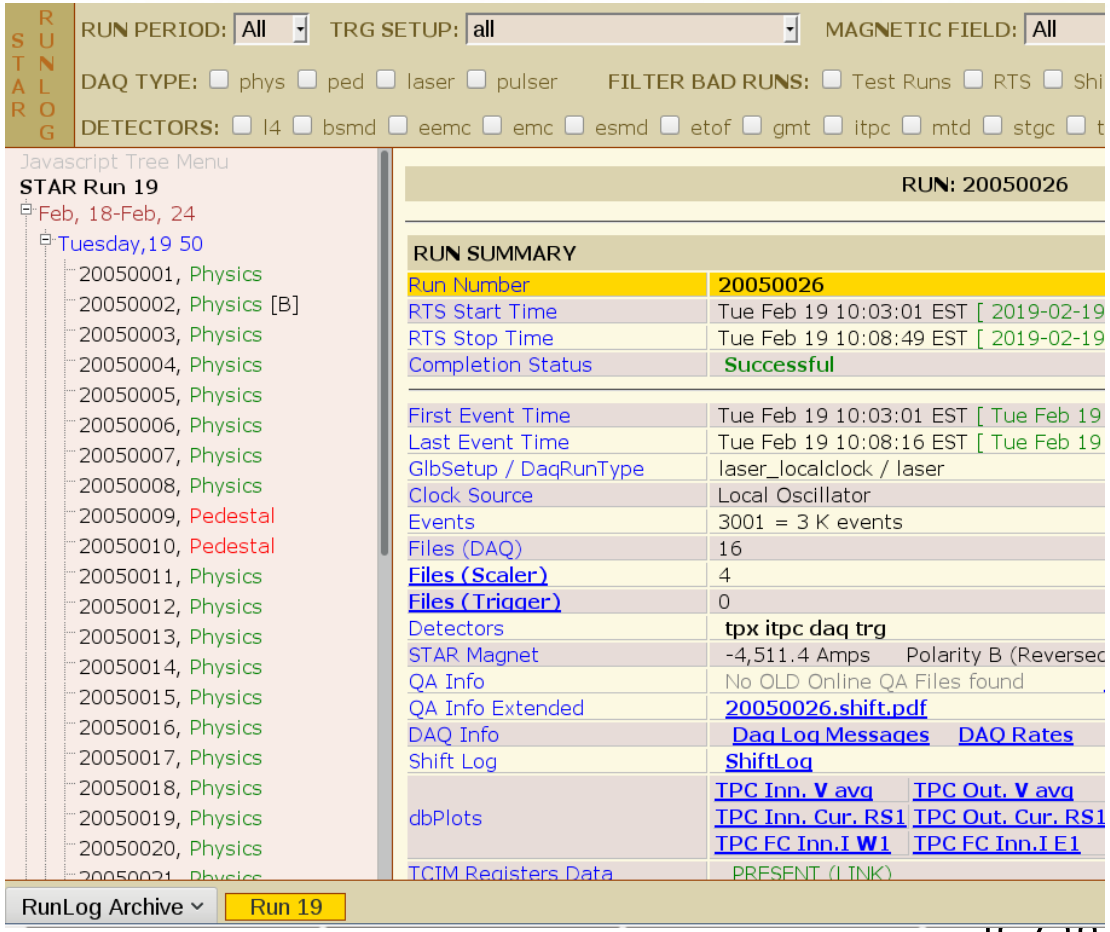
Feb 26th - Mar 5th	Daniel Cebra University of California - Davis	0:30-7:30	Alexander Jentsch University of Texas at Austin	Jan Vanek Nuclear Physics Institute, The Czech Academy of Sciences	Lukas Kramarik Czech Technical University in Prague	Maria Stefaniak Warsaw University of Technology		
		7:30-16:30	Zilong Chang Brookhaven National Laboratory	Joel Mazer Rutgers University	Todd Kinghorn University of California - Davis	Xiaohai Jin Shanghai Institute of Applied Physics		
		16:30-0:30	David Kapukchyan University of California - Riverside	Dingwei Zhang Central China Normal University	Zhenzhen Yang Central China Normal University	Yue Liang Kent State University		

Week	Period Coord.	Shift	Shift Leader	Detector Opr.	Detector Opr.	Shift Crew	Leader Trainee	Det.Opr.Trainee
Mar 5th - Mar 12th	Daniel Cebra University of California - Davis	0:30-7:30		Bogdan Pawlik AGH University of Science and Technology	Susumu Sato University of Tsukuba	MIna Hatakeyama University of Tsukuba		Kosuke Okubo University of Tsukuba
		7:30-16:30	Shuai Yang Brookhaven National Laboratory	Joel Mazer Rutgers University	Declan Keane Kent State University	Xiaohai Jin Shanghai Institute of Applied Physics	Kenneth Barish University of California - Riverside	Audrey Francisco Yale University
		16:30-0:30	David Kapukchyan University of California - Riverside	Dingwei Zhang Central China Normal University	Anjali Attri Panjab University	Zhenzhen Yang Central China Normal University		

Experiment's RunLog

https://online.star.bnl.gov/RunLog/

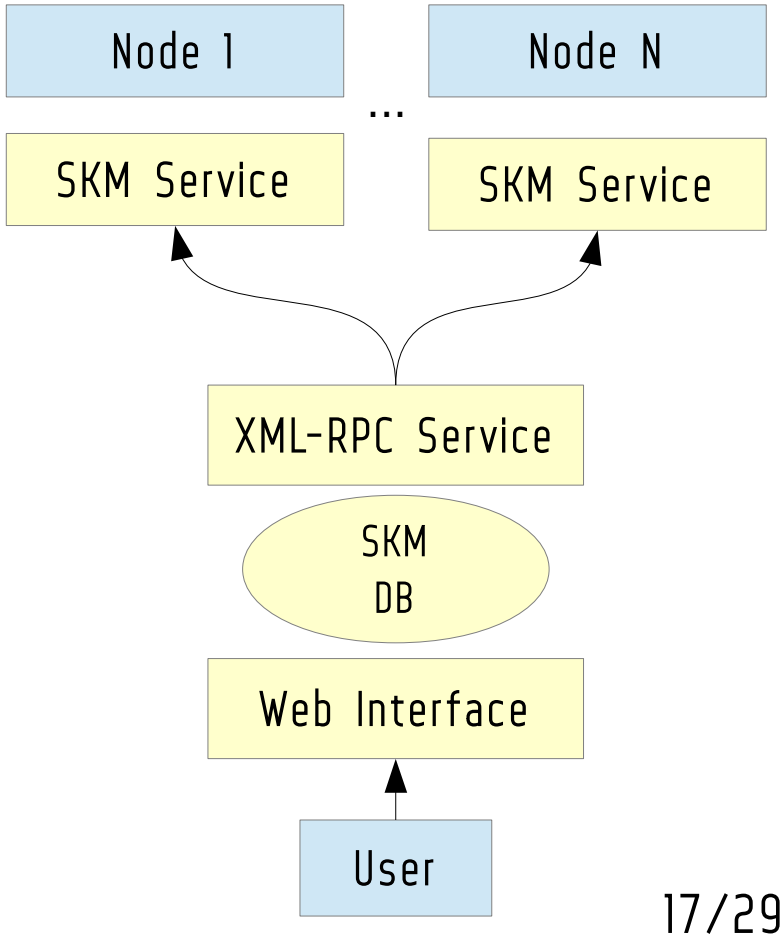
- Features:
  - Extensive web interface for all STAR runs, taken during RHIC data taking Runs.
  - Provides run statistics (time, events, triggers, files etc) filtering, monitoring logs, conditions overview and other information
  - Collects and organizes information from a variety of sources: Run-Time System, DAQ, Conditions, Slow Controls etc;
  - Composed of a ~dozen services, three database instances and a web interface.
  - Archived annually, to provide historical records for past Runs
  - Web interface was fully re-written from scratch in 2010 as Model-View-Controller application



SSH Key Management

https://www.star.bnl.gov/starkeyw/

- Features:
  - Completely automatic SSH key management across mid-sized Linux cluster (online domain).
  - Allows to satisfy CyberSecurity requirements for sensitive domain access.
  - Enables user fingerprinting via personal SSH keys.
  - Eliminates the need for password-protected shared accounts (aka sticky-note passwords)
- Administrative Features:
  - User, Host, Public Key or Public-Private Key management.
  - Assign user keys to accounts, enable/disable offending users or hosts, receive notifications of new requests, approve requests.



SSH Key Management

https://www.star.bnl.gov/starkeyw/

SSH PUBLIC KEY MANAGEMENT CONSOLE : STAR

LOGGED AS: **dmitry**, ACCESS: **Admin**
[X] LOGOUT

HOME :: HELP

ADMIN ACCESS : show pending requests :: show pending priv-pub requests :: account and host management :: expired associations :: show hosts :: show host tags :: show users :: show admins :: show logs :: show config parameters :: account scan :: private/public key pair management :: show blacklisted keys

(Re)Upload Public Key

Current state : public key **OK, fingerprint: c7:55:a3:f3:75:5d:07:d2:e8:66:e9:2e:47:f4:06:48**

Public key file in OpenSSH format : Browse... No file selected.

UPLOAD PUBLIC KEY FILE

This will *replace* any previously uploaded key. Key replacement takes up to 10 minutes to propagate to the client nodes.

Contact E-mail

Current E-mail : arkhipkin@bnl.gov

Change E-mail to :

CHANGE E-MAIL

Request for account association

Host: Select Host...

Account: Select Account...

Host Restriction: ex.: *.star.bnl.gov

Reason for this association: ex.: subsystem expert

REQUEST ACCOUNT @ HOST ASSOCIATION FOR YOUR KEY

This is *not* instantaneous -- all associations must be approved by an administrator before becoming active.

Association status list

SSH Key Management

https://www.star.bnl.gov/starkeyw/

List of hosts

Show 10 entries Search:

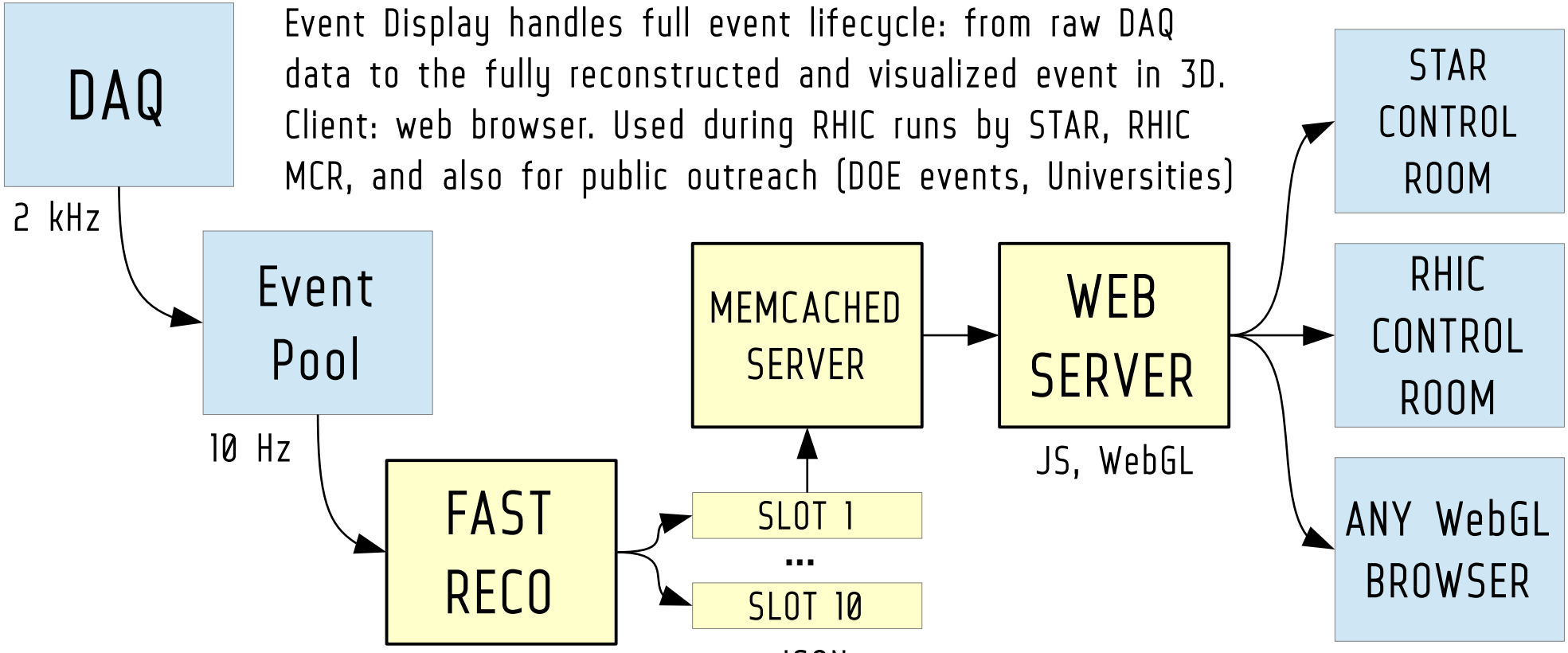
HOST	STATUS	DESCRIPTION	VERSION	ACCOUNTS	USERS	LAST HEARTBEAT	LAST UPDATE	TOGGLE STATUS	DELETE HOST
alh2.starp.bnl.gov (130.199.60.41)	ACTIVE	alh2.starp.bnl.gov, SL6.x, Slow Controls Alarm Handler	2nd gen. / 2.0.5	24	16	2019-02-18 20:08:56	2019-02-15 17:18:31	☐	☐
astaire-run09.starp.bnl.gov (130.199.60.53)	ACTIVE	astaire-run09.starp.bnl.gov, SL 6.x, 64-bit	2nd gen. / 2.0.5	4	3	2019-02-18 20:04:28	2018-09-25 18:52:16	☐	☐
barbados2.starp.bnl.gov (130.199.60.46)	ACTIVE	barbados2.starp.bnl.gov (SL6, x86_64) (Slow Controls)	2nd gen. / 2.0.5	16	9	2019-02-18 20:05:50	2019-02-15 17:15:22	☐	☐
beatrice.starp.bnl.gov (130.199.60.19)	ACTIVE	beatrice (BEMC node), SL6.x, i386	2nd gen. / 2.0.6	5	4	2018-06-19 18:10:01	2019-02-18 20:10:24	☐	☐
bermuda.starp.bnl.gov (130.199.60.55)	ACTIVE	bermuda.starp.bnl.gov, Slow Controls PC in 2C6, SL 7.x, 64-bit	2nd gen. / 2.0.6	22	13	2019-02-18 20:03:20	2019-02-15 17:22:59	☐	☐
blanchett.starp.bnl.gov (130.199.60.133)	ACTIVE	blanchett.starp.bnl.gov	2nd gen. / 2.0.5	12	8	2019-02-18 20:09:47	2018-09-25 18:58:50	☐	☐
burton.starp.bnl.gov (130.199.61.104)	ACTIVE	burton.starp.bnl.gov	2nd gen. / 2.0.6	3	3	2019-02-18 20:08:04	2018-12-20 15:32:19	☐	☐
chaplin-run09.starp.bnl.gov (130.199.60.68)	ACTIVE	chaplin-run09.starp.bnl.gov (SL6; x86_64)	2nd gen. / 2.0.5	4	3	2019-02-18 20:07:03	2018-09-25 18:54:19	☐	☐
daqboot.starp.bnl.gov (130.199.60.214)	ACTIVE	daqboot.starp.bnl.gov (Sc.Linux 7.x, 64-bit)	2nd gen. / 2.0.6	4	3	2019-02-18 20:03:37	2019-02-13 17:42:57	☐	☐
daqman.starp.bnl.gov (130.199.60.86)	ACTIVE	daqman.starp.bnl.gov (SL 6.x, 64-bit)	2nd gen. / 2.0.5	46	40	2019-02-18 20:06:27	2019-01-18 23:03:00	☐	☐
HOST	STATUS	DESCRIPTION	VERSION	ACCOUNTS	USERS	LAST HEARTBEAT	LAST UPDATE	TOGGLE STATUS	DELETE HOST

Showing 1 to 10 of 91 entries ◀ Previous Next ▶

Online Event Display: services

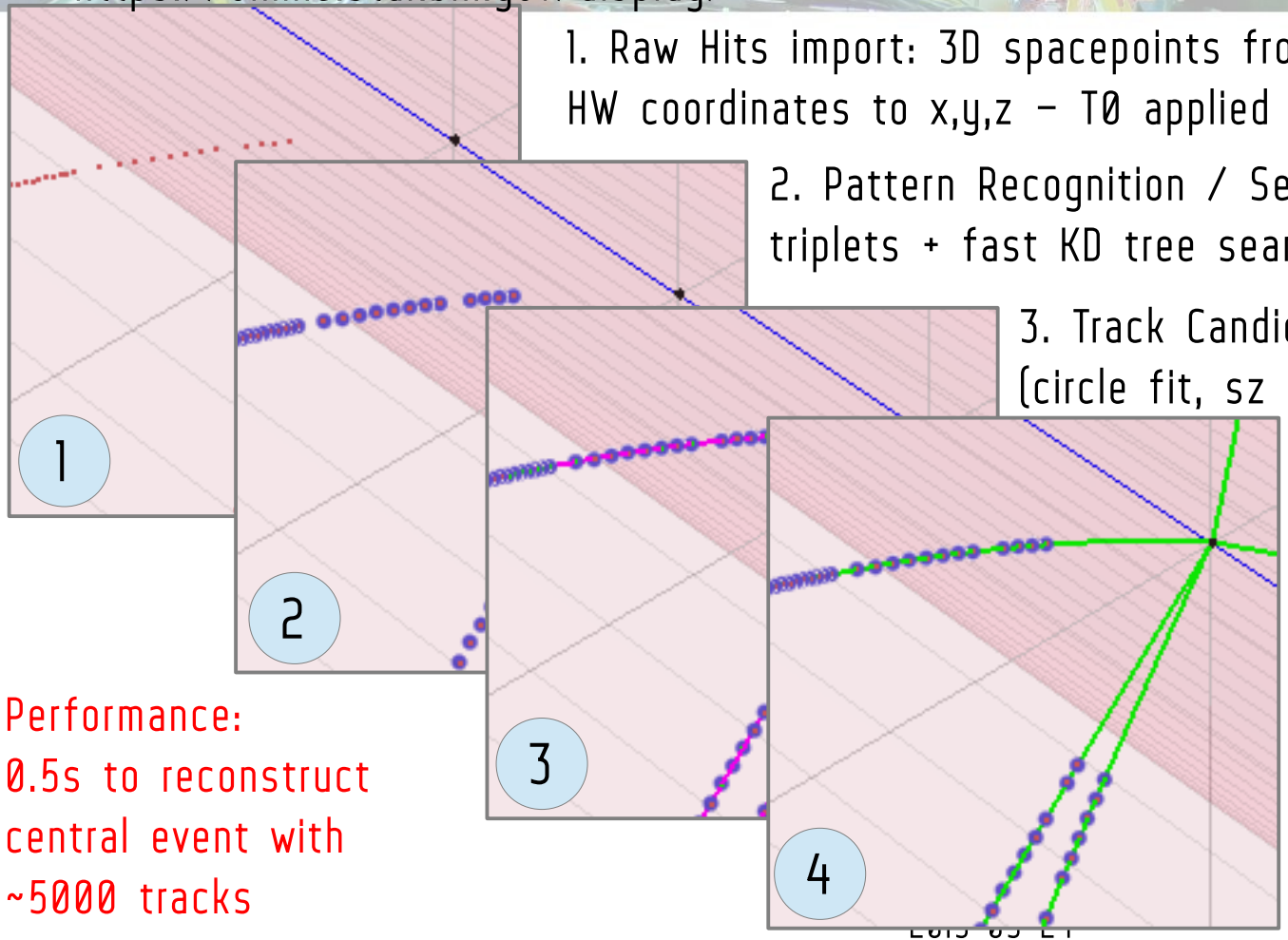
https://online.star.bnl.gov/display/

Event Display handles full event lifecycle: from raw DAQ data to the fully reconstructed and visualized event in 3D. Client: web browser. Used during RHIC runs by STAR, RHIC MCR, and also for public outreach (DOE events, Universities)



Online Event Display: track reco

https://online.star.bnl.gov/display/



1. Raw Hits import: 3D spacepoints from DAQ. Conversion from HW coordinates to x,y,z – T0 applied

2. Pattern Recognition / Seed Finding via triplets + fast KD tree search

3. Track Candidate Following & Fitting (circle fit, sz fit => fully reco'ed momentum

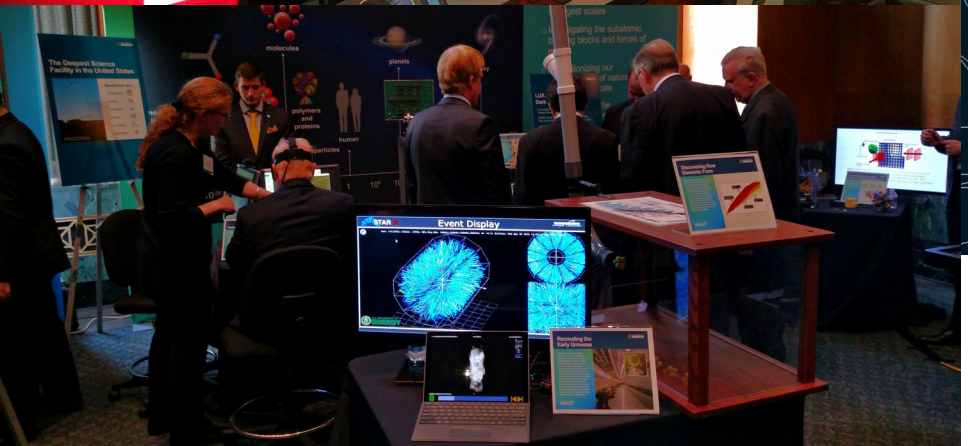
4. Vertex Seed Finding
Centroid found by projecting tracks to DCA(POC) to z-axis

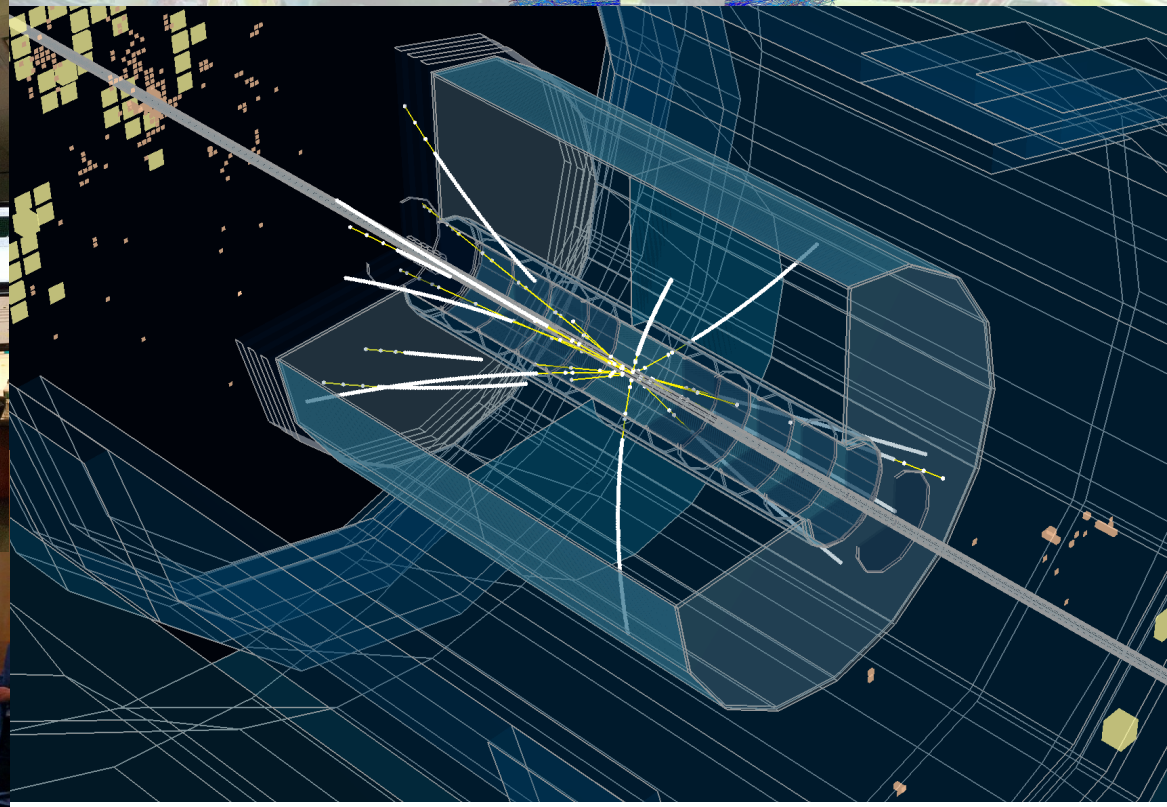
Performance:
0.5s to reconstruct central event with ~5000 tracks

C++11:
- kdfinder.hpp
- nanoflann.hpp

Online Event Display: Web UI v1

https://online.star.bnl.gov/display/





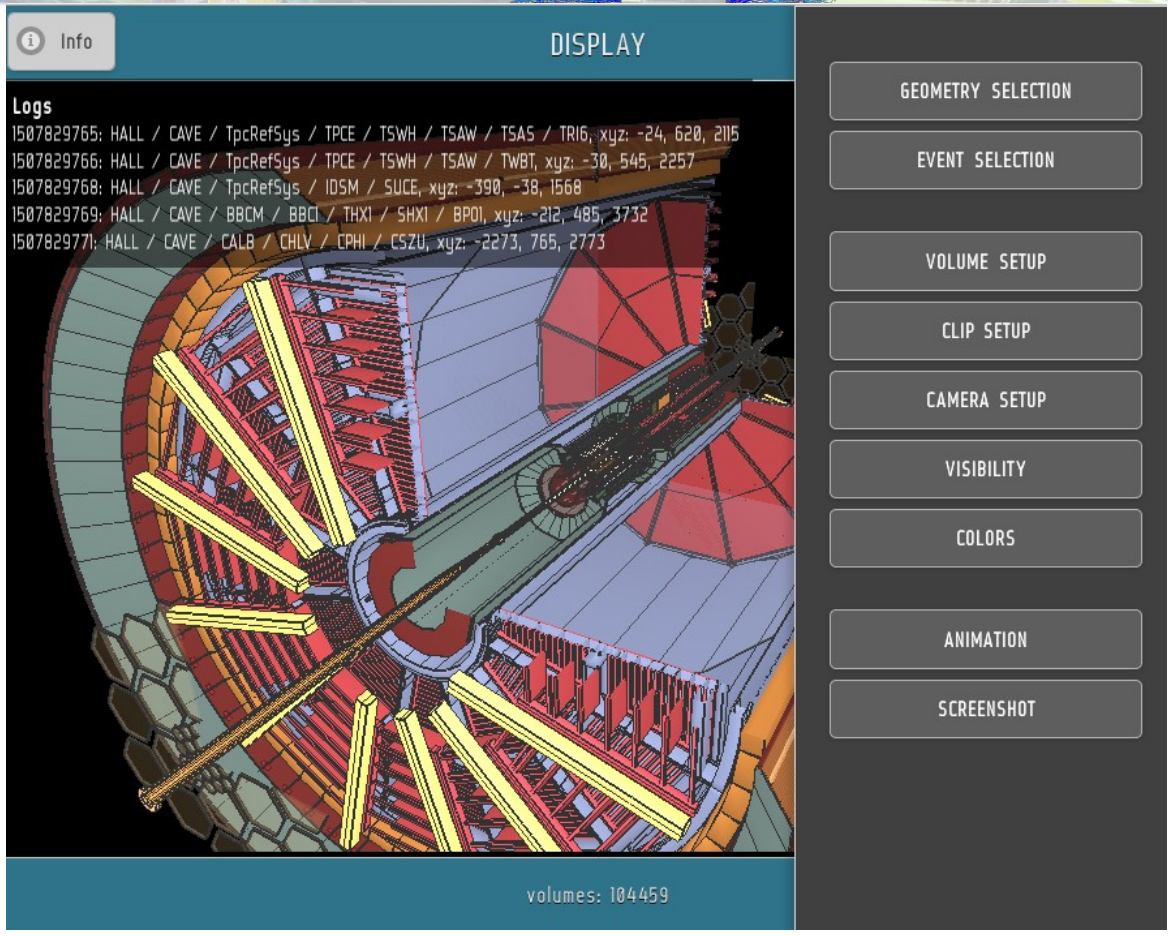
EVD: STAR Control Room, RHIC Control Room

EIC Event Display, sPHENIX Event DIsplay

Event Display Web UI v2

https://www.star.bnl.gov/~dmitry/gide_new/

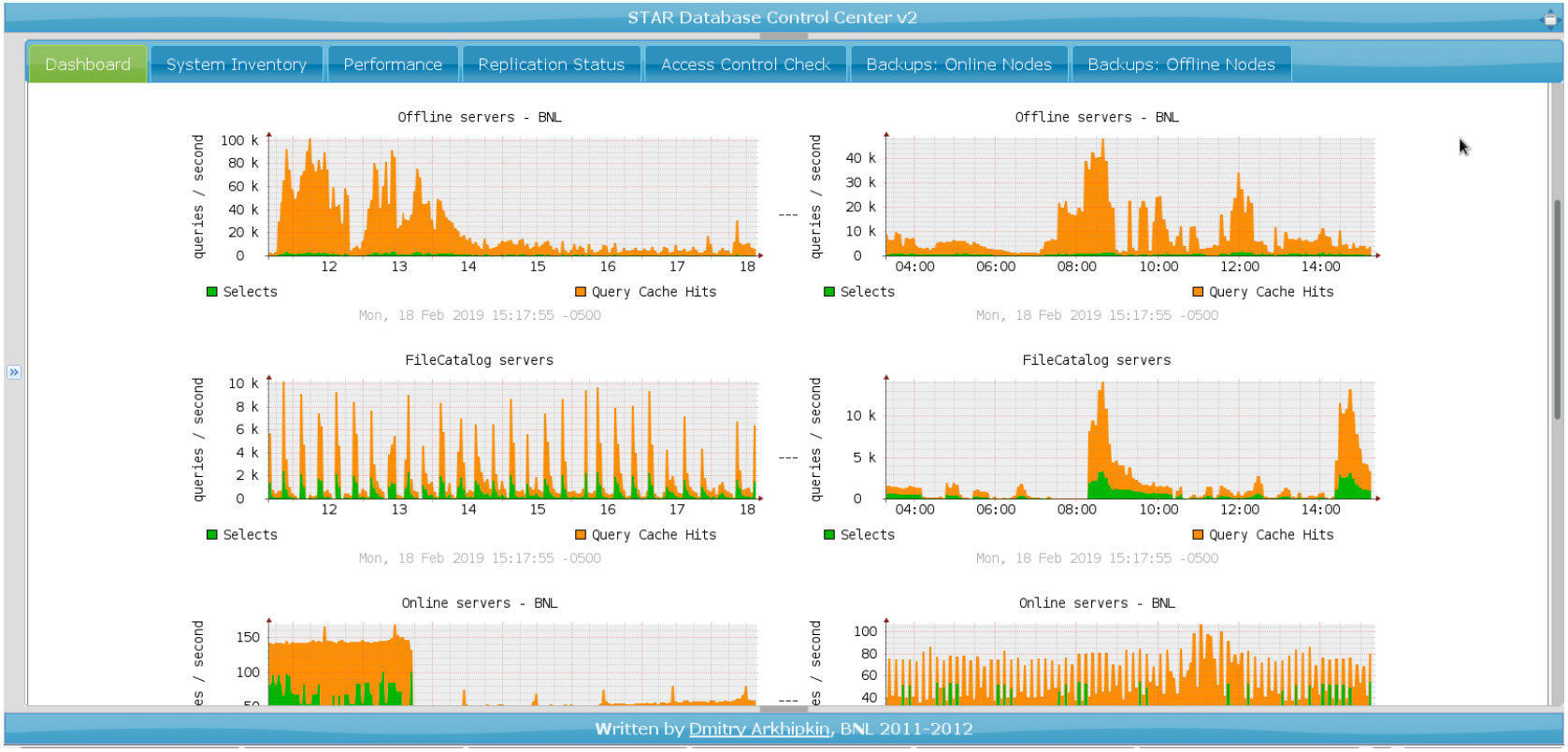
- Geometry Input Format:
  - latest GDML version supported
- Event Input Format: JSON
- Geometry Shapes:
  - 100% coverage of GDML/G4, TGeo, VecGeom
- Interactivity:
  - Subselection of volumes
  - Automatic volume positioning
- Physical Objects
  - Tracks: helix, set of points
  - Hits: 3d points a-la TPC, calorimetric hits
- Extensively used by ITPC experts: debugging!



DB Monitor

https://online.star.bnl.gov/Mon/

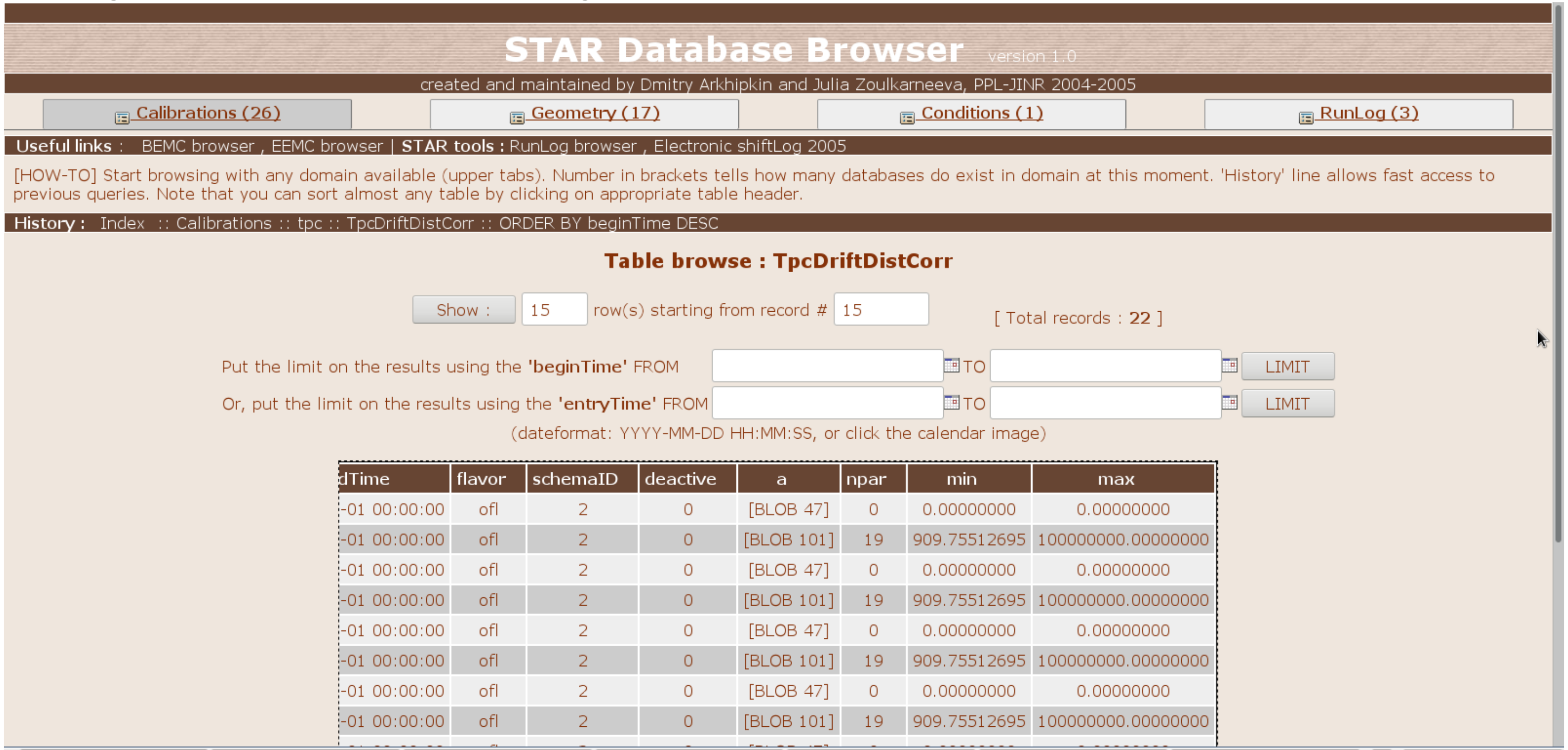
Custom monitoring tool, specialized for large replicated MySQL setups. Monitors all STAR databases, Provides extensive automatic inventory, replication status and performance tuning hints.



DB Browser

https://www.star.bnl.gov/Browser/

Custom database browsing tool. Provides generic database viewer capability, and specialized database viewing for EMC and EEMC subsystems.



STAR Database Browser version 1.0

created and maintained by Dmitry Arkhipkin and Julia Zoulkarneeva, PPL-JINR 2004-2005

Calibrations (26) | Geometry (17) | Conditions (1) | RunLog (3)

Useful links : BEMC browser , EEMC browser | **STAR tools :** RunLog browser , Electronic shiftLog 2005

[HOW-TO] Start browsing with any domain available (upper tabs). Number in brackets tells how many databases do exist in domain at this moment. 'History' line allows fast access to previous queries. Note that you can sort almost any table by clicking on appropriate table header.

History : Index :: Calibrations :: tpc :: TpcDriftDistCorr :: ORDER BY beginTime DESC

Table browse : TpcDriftDistCorr

Show : 15 row(s) starting from record # 15 [Total records : **22**]

Put the limit on the results using the **'beginTime'** FROM TO LIMIT

Or, put the limit on the results using the **'entryTime'** FROM TO LIMIT

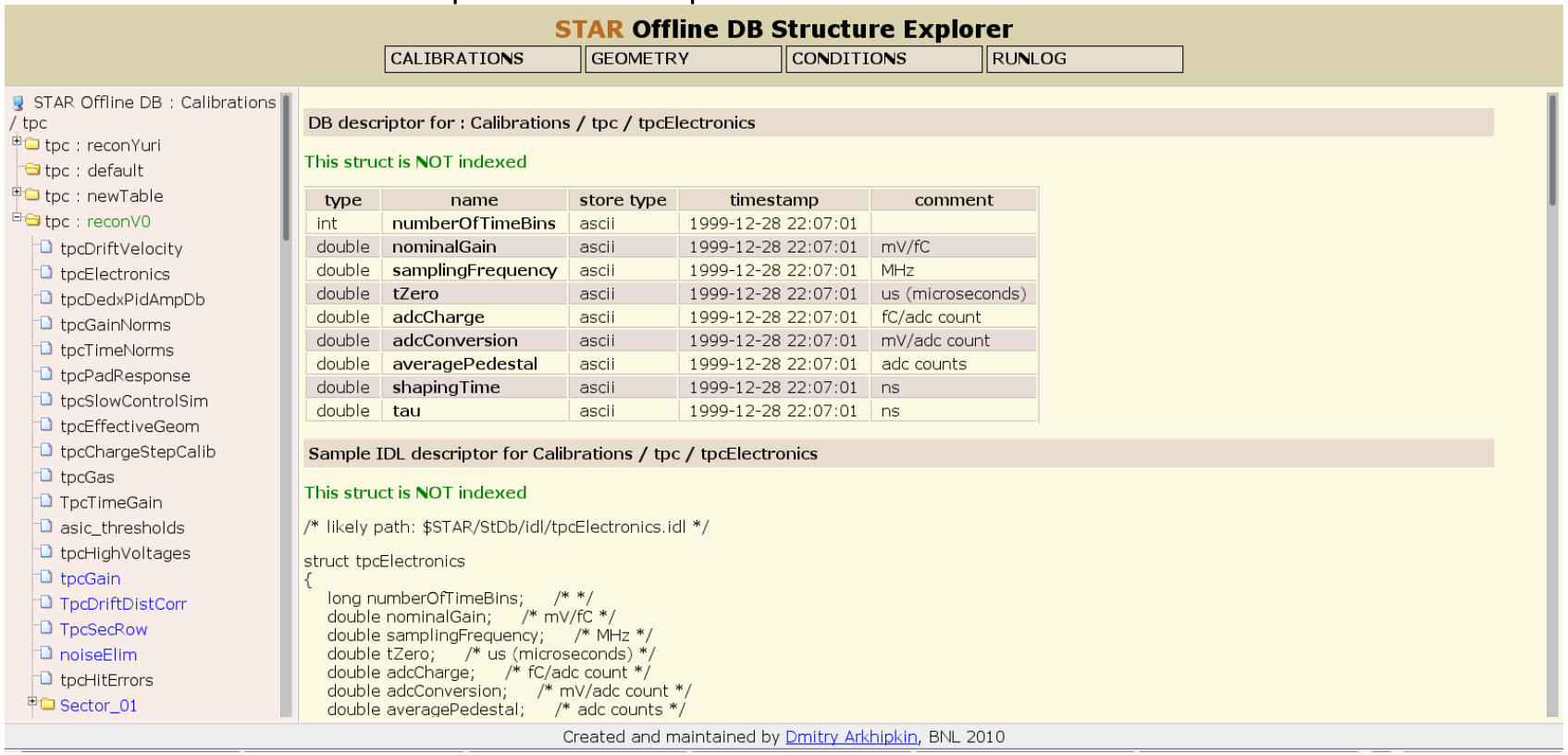
(dateformat: YYYY-MM-DD HH:MM:SS, or click the calendar image)

dTime	flavor	schemaID	deactive	a	npar	min	max
-01 00:00:00	ofl	2	0	[BLOB 47]	0	0.00000000	0.00000000
-01 00:00:00	ofl	2	0	[BLOB 101]	19	909.75512695	100000000.00000000
-01 00:00:00	ofl	2	0	[BLOB 47]	0	0.00000000	0.00000000
-01 00:00:00	ofl	2	0	[BLOB 101]	19	909.75512695	100000000.00000000
-01 00:00:00	ofl	2	0	[BLOB 47]	0	0.00000000	0.00000000
-01 00:00:00	ofl	2	0	[BLOB 101]	19	909.75512695	100000000.00000000
-01 00:00:00	ofl	2	0	[BLOB 47]	0	0.00000000	0.00000000
-01 00:00:00	ofl	2	0	[BLOB 101]	19	909.75512695	100000000.00000000

DB Explorer

https://online.star.bnl.gov/dbExplorer/

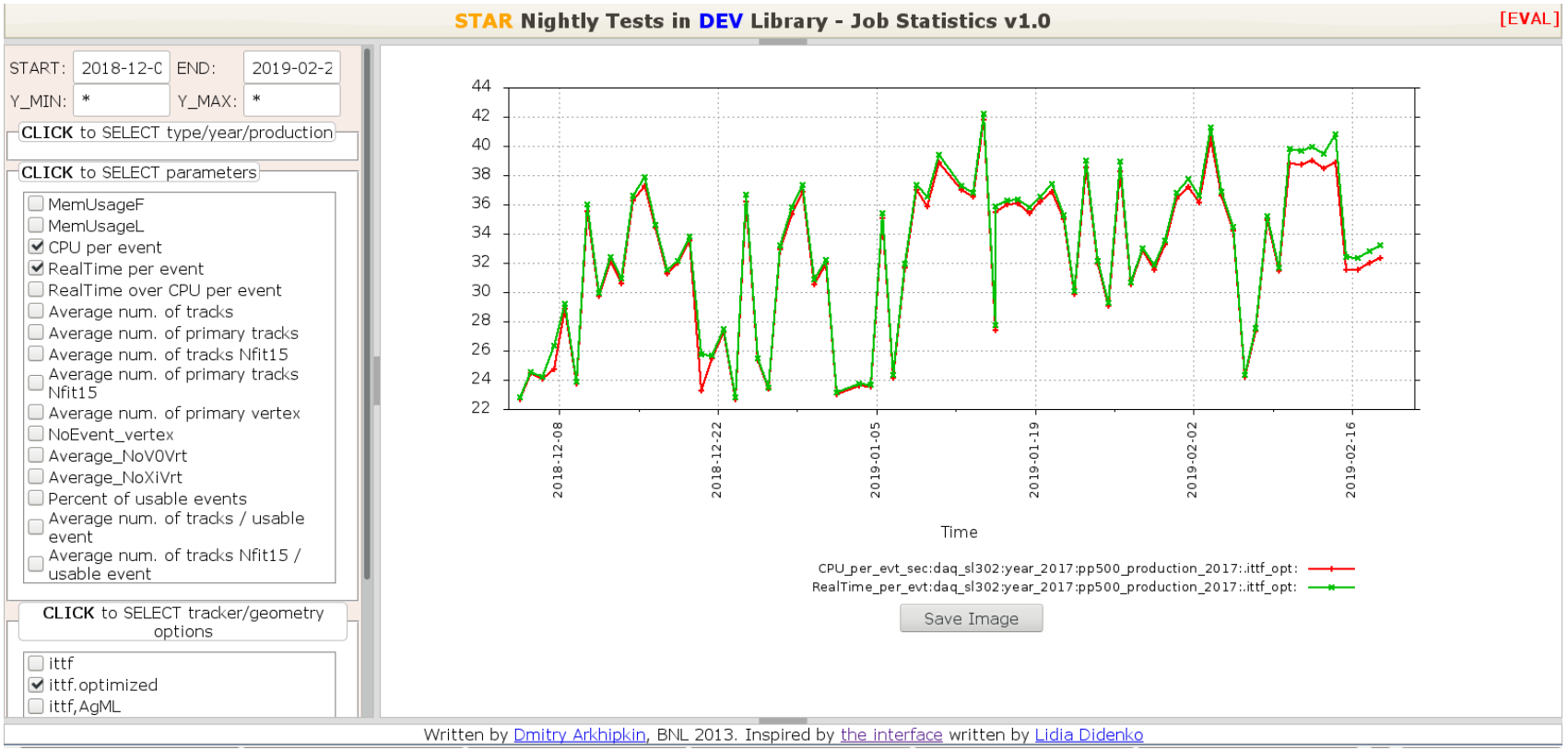
Auto-documentation system for STAR Offline Databases and API. Provides web-based interface for database schema and structure, provides samples for DB read and DB write for each table.



jobStat: nightly tests

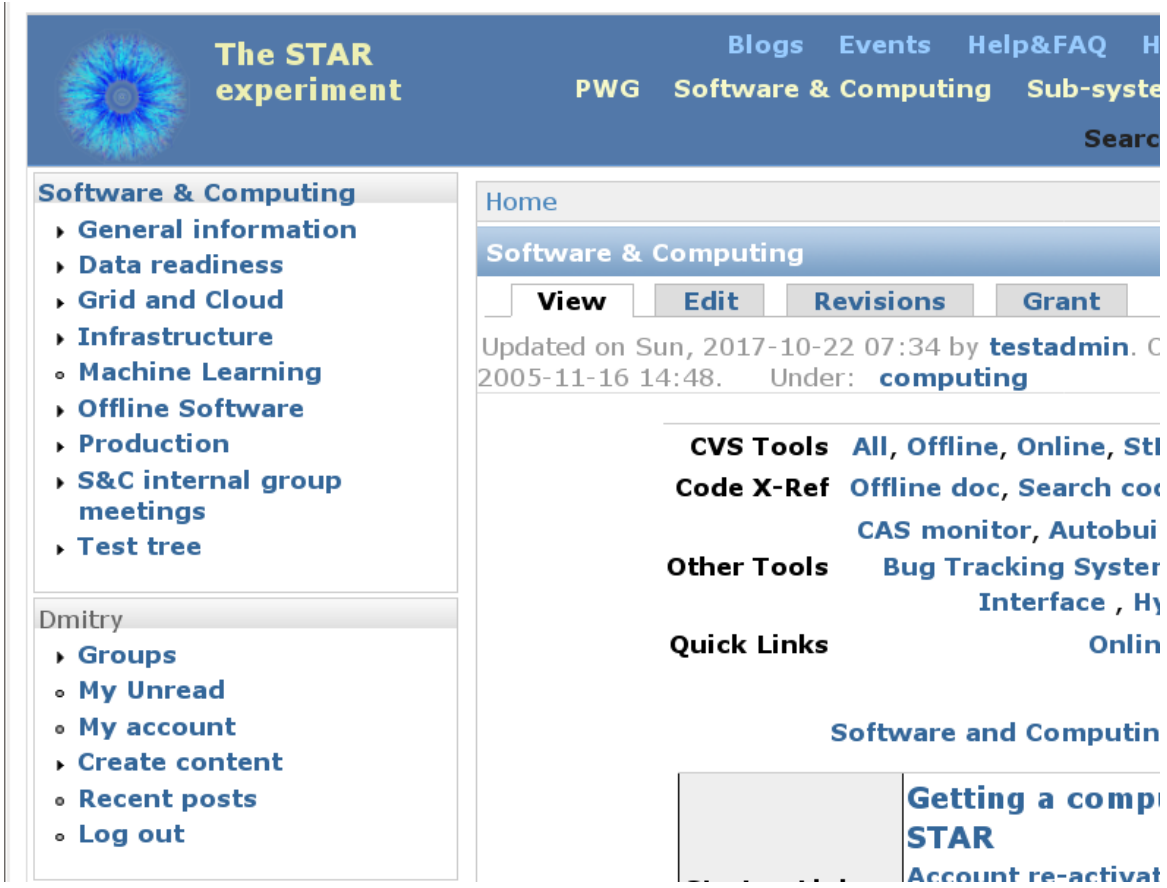
https://online.star.bnl.gov/jobStat/

Web interface to STAR nightly tests system. Provides fast plotting capabilities for all nightly tests.



Drupal Development

- Features:
  - Drupal is modular, easy to extend content management system
  - Provides STAR with web-based document management, blogs, calendar of events, conferences, STAR paper/note archive and many more since 2003.
- Custom modules:
  - STAR conference and meeting
  - STAR publications and notes
  - STAR presentations and thesis
  - STAR simulation requests
  - STAR news and polls



Summary

- Current Duties:
  - All STAR databases – maintenance, support, backups, performance tuning, development for 30+ servers, 50+ MySQL instances, 3 MongoDB instances
- Major RHIC Run Tasks:
  - Online Databases, migration scripts, RunLog service, ShiftSignup service, MIRA services (data collectors), Event Display service
- Major Out-of-Run Tasks:
  - Offline Databases, StDbLib (DB API), FileCatalog databases, DB-related software upgrades, Drupal development and maintenance, R&D development (not mentioned here)
- Commonly-used Languages and Techs:
  - C++, JavaScript, PHP, shell, SQL, RPC, XML, JSON, HTML, CSS etc.