Hall A Software & Analysis for the 12 GeV Era

Ole Hansen

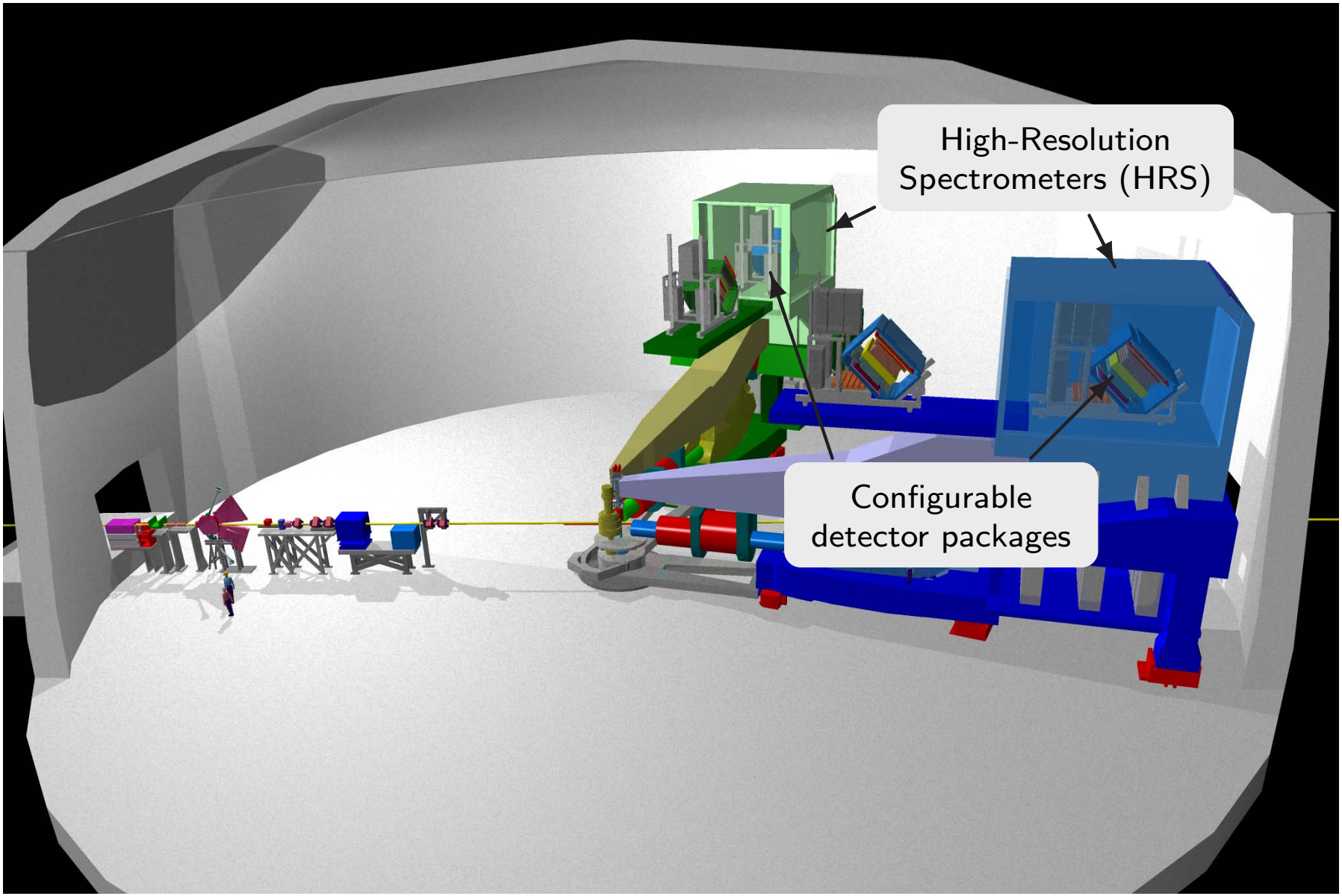
Jefferson Lab

JLab 12 GeV Software Review
Morning Session
June 7, 2012

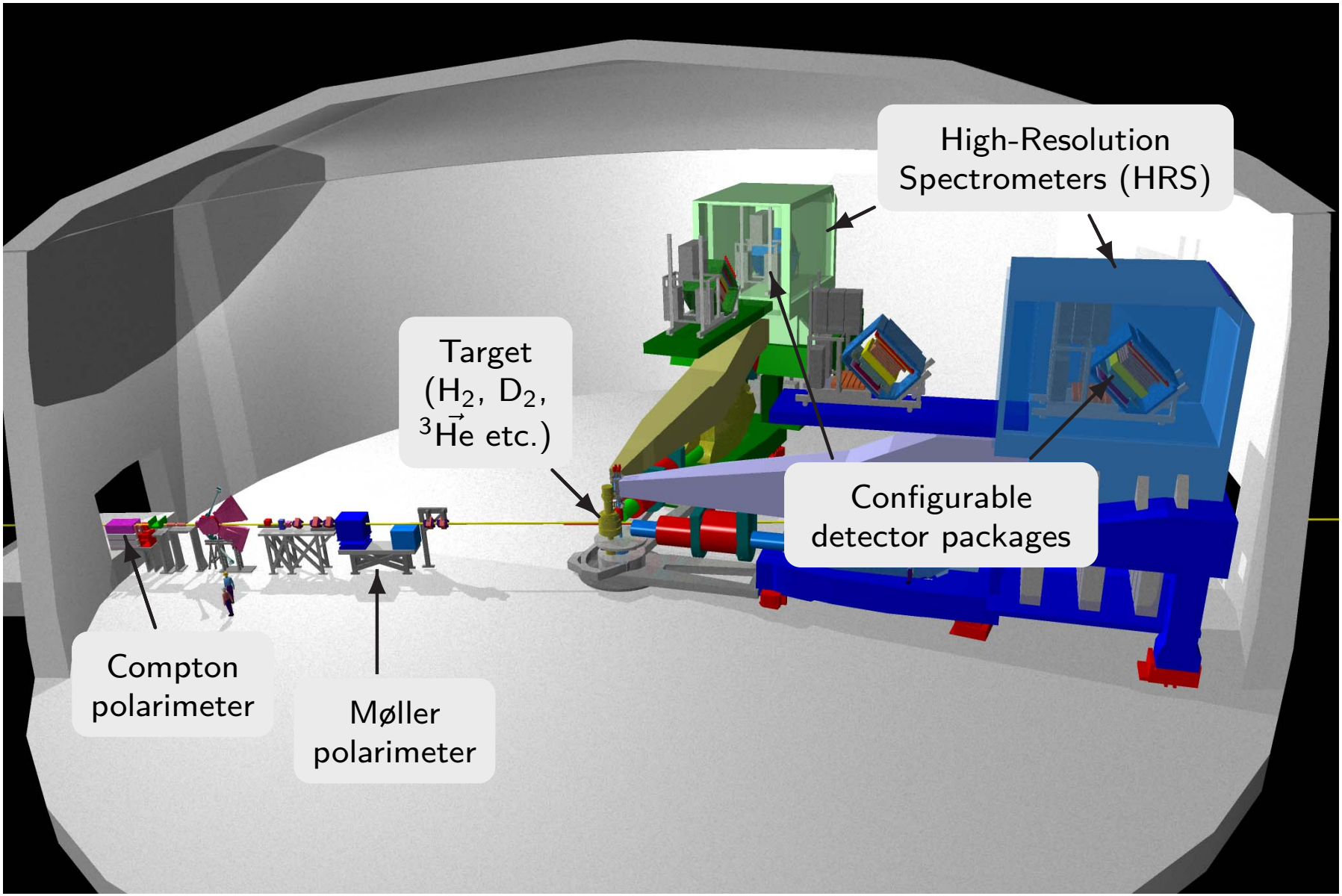
Jefferson Lab
Thomas Jefferson National Accelerator Facility

Outline

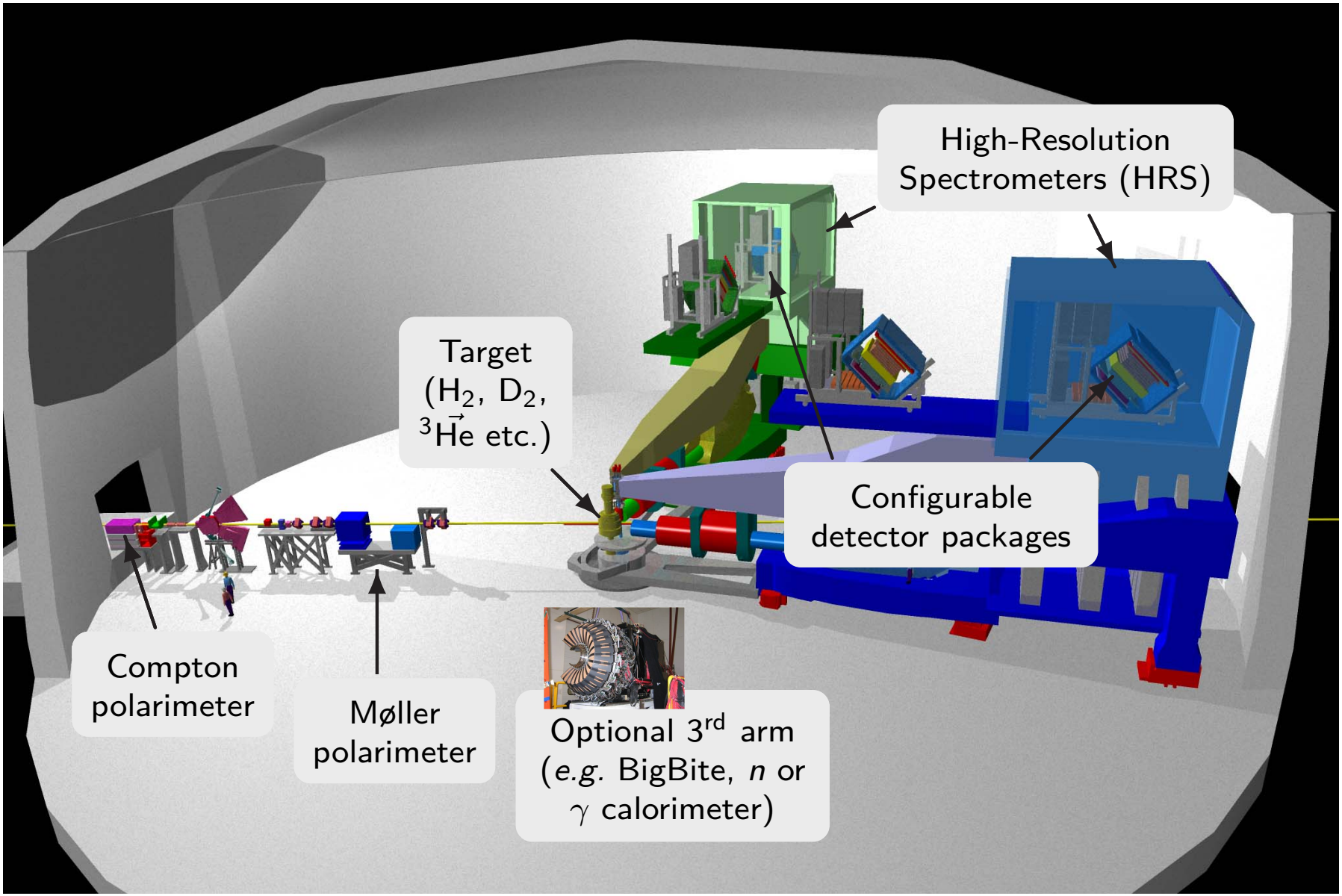
Hall A Experimental Equipment



Hall A Experimental Equipment

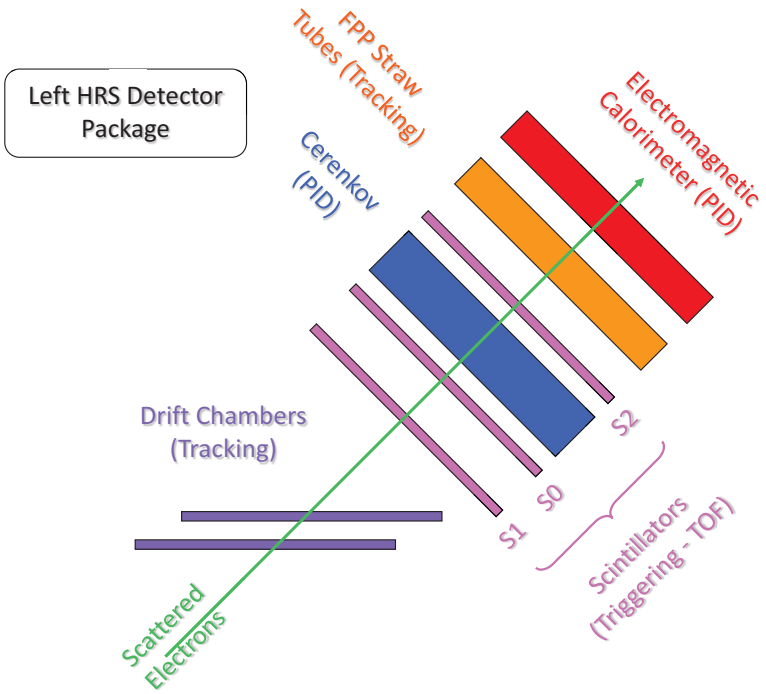


Hall A Experimental Equipment



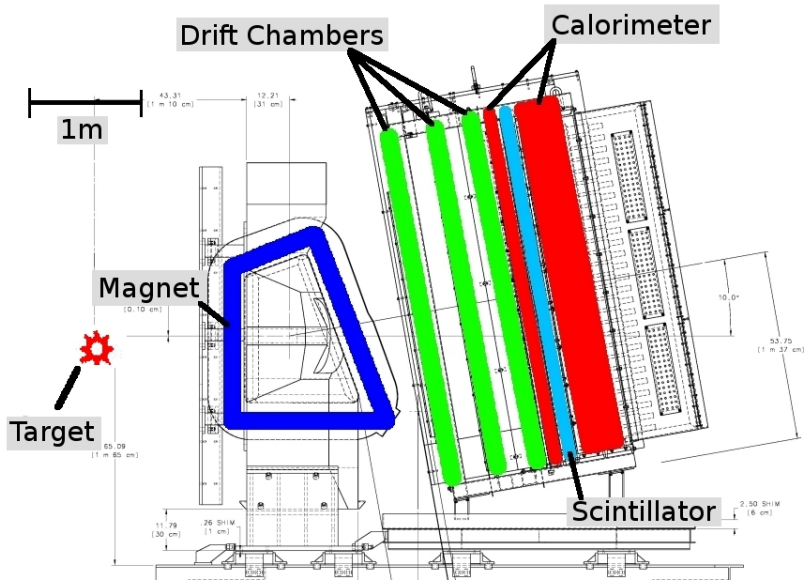
Typical Detector Packages

High-resolution spectrometer (HRS)



- Vertical Drift Chambers (VDCs): 1472 channels
- Focal Plane Polarimeter (FPP): 5112 channels
- Calorimeter(s): 48+96 channels
- Scintillators: 12+12+18 channels
- Cherenkovs: gas & aerogel, 10+26 channels

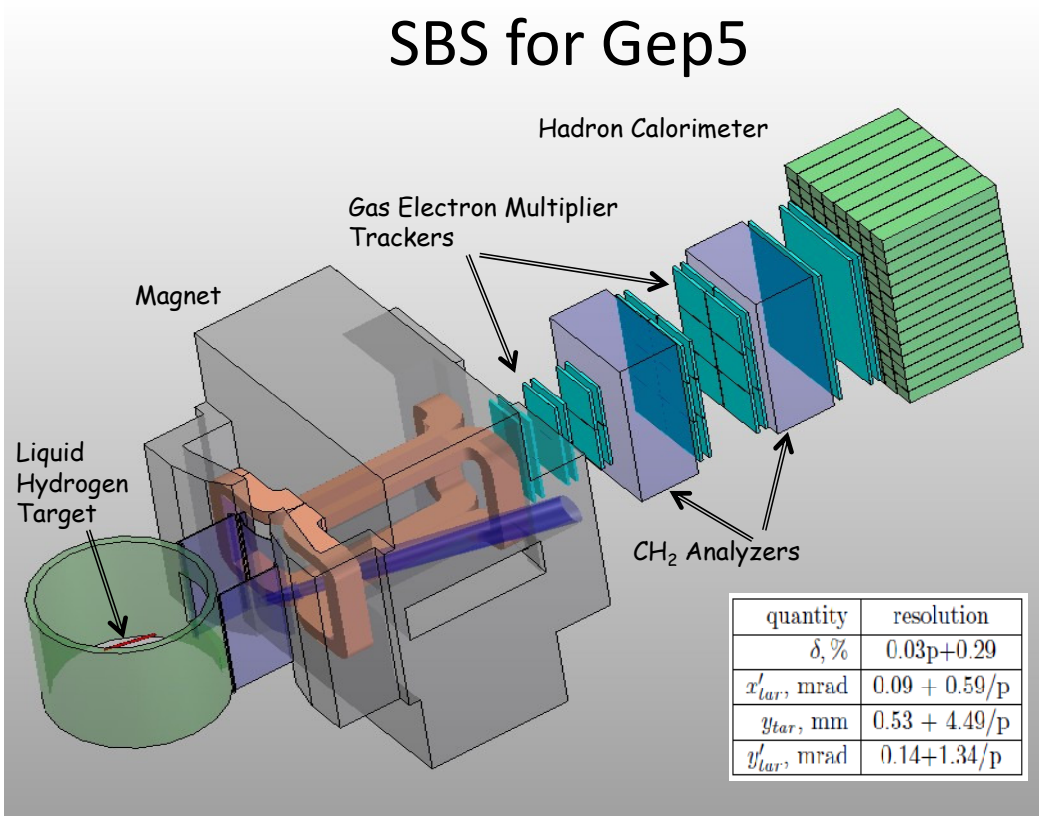
BigBite spectrometer (3rd arm)



- Horizontal Drift Chambers (MWDCs): 3246 chan. total
- Calorimeters: 48+48 channels
- Scintillator: 56 channels

Hall A SuperBigbite Spectrometer (SBS)

B. Wojtsekhowski, G. Cates, *et al.*



- Set of components for flexible medium-acceptance spectrometer configuration
- GEM trackers for high-rate environment (500 kHz/cm^2), up to $\approx$ 100k channels total
- Proposals approved for
  - EM form factor measurements up to $Q^2 \approx 10$ GeV^2
  - SIDIS/Transversity
- Recently fully approved by DOE
- Data taking tentatively starting in 2016

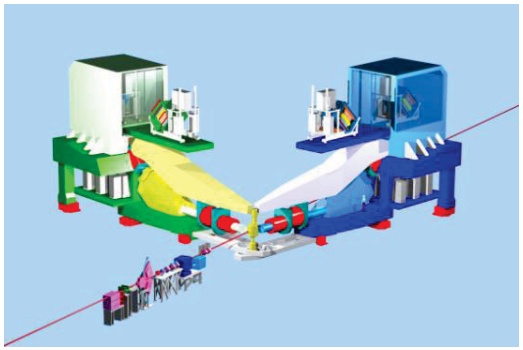
Physics Topics in Hall A

1. Transverse Hadron Structure
   - Nucleon electromagnetic form factors $G_{E/M}^{n/p}$
2. Longitudinal Hadron Structure
   - g_1 structure function at high x
   - F_2^n/F_2^p and d/u ratios
   - Deuteron tensor structure function b_1
3. 3D Hadron Structure
   - Generalized Parton Distributions (GPDs)
   - Transverse Momentum Distributions (TMDs)
4. Hadrons and Cold Nuclear Matter
   - Short-range correlations
   - Hypernuclei
5. Low-energy tests of the Standard Model and Fundamental Symmetries
   - Ultra-precise measurement of $\sin^2\theta_W$
   - Dark photon search
   - Neutron skin radius of heavy nuclei
   - Axial vector quark couplings
   - Charge symmetry violation

Traditional Hall A Physics Analysis "Mode of Operation"

- Each experiment typically employs
  - Base equipment (*e.g.* HRSs, targets) in varying configurations
  - Add-on equipment, both existing and new (*e.g.* RICH, photon calorimeter, neutron detector, etc.)
  - Experiment-specific electronics and trigger setup
- Requires highly modular analysis software. Framework with modules for all "6 GeV" standard components exists (see later in this talk)
- Calibration and analysis software for add-on equipment usually provided by users (members of the experiment's collaboration)
  - Effort typically relatively small
  - Users routinely have best expertise with new or specialized equipment
- Users encouraged, but not required, to write code compatible with the standard Hall A analysis framework. Hall A software experts provide support with code integration, if needed.

Timeline of Early 12 GeV Hall A Experiments (R. Michaels)



Preliminary

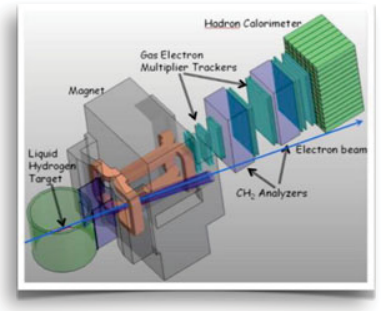
16 mo. Shutdown — **12 GeV Commissioning** — **Early Experiments** — **SuperBigBite**

Dark Photon Search (APEX) and/or **Neutron Skin Radius Meas. (PREX)**

Neutron Spin Structure (A1n)

Deeply Virtual Compton Scatt. (DVCS)

EM Form Factors at high Q^2



Beam to Hall A

11 GeV

Raw Trigger & Data Rates

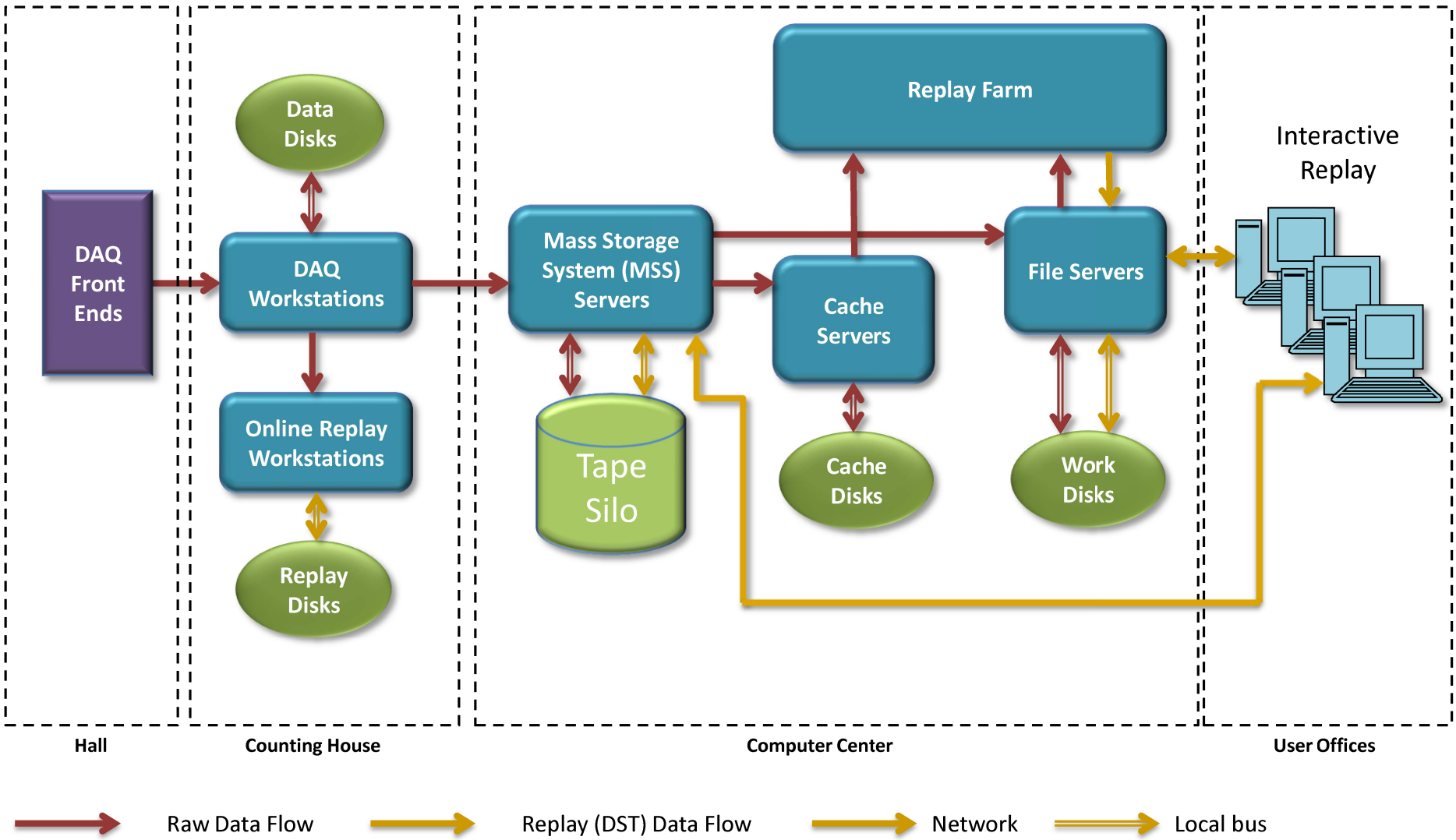
Preliminary schedule & numbers

Experiment	APEX[1]	A_1^n	DVCS	SBS			
				G_E^n	G_M^n	G_E^p	Transv
PAC number 12-	10-009	06-122	06-114	09-016	09-019	07-109	09-018
Config	L+R(CI)	L+R(SA)	L+γCal	BBG+ND	BBG+ND	SBS+BC	SBS+BB
PAC days	34	23	100	58	48	60	64
Schedule		— 2014/2015 —			— 2016/2018 —		
Evt size (kB)	4	2	30	30	20	120	5
Trig rate (kHz)	5	2×10	0.5	2	2	1	5
Data rate (MB/s)	20	40	15	60	40	120	25

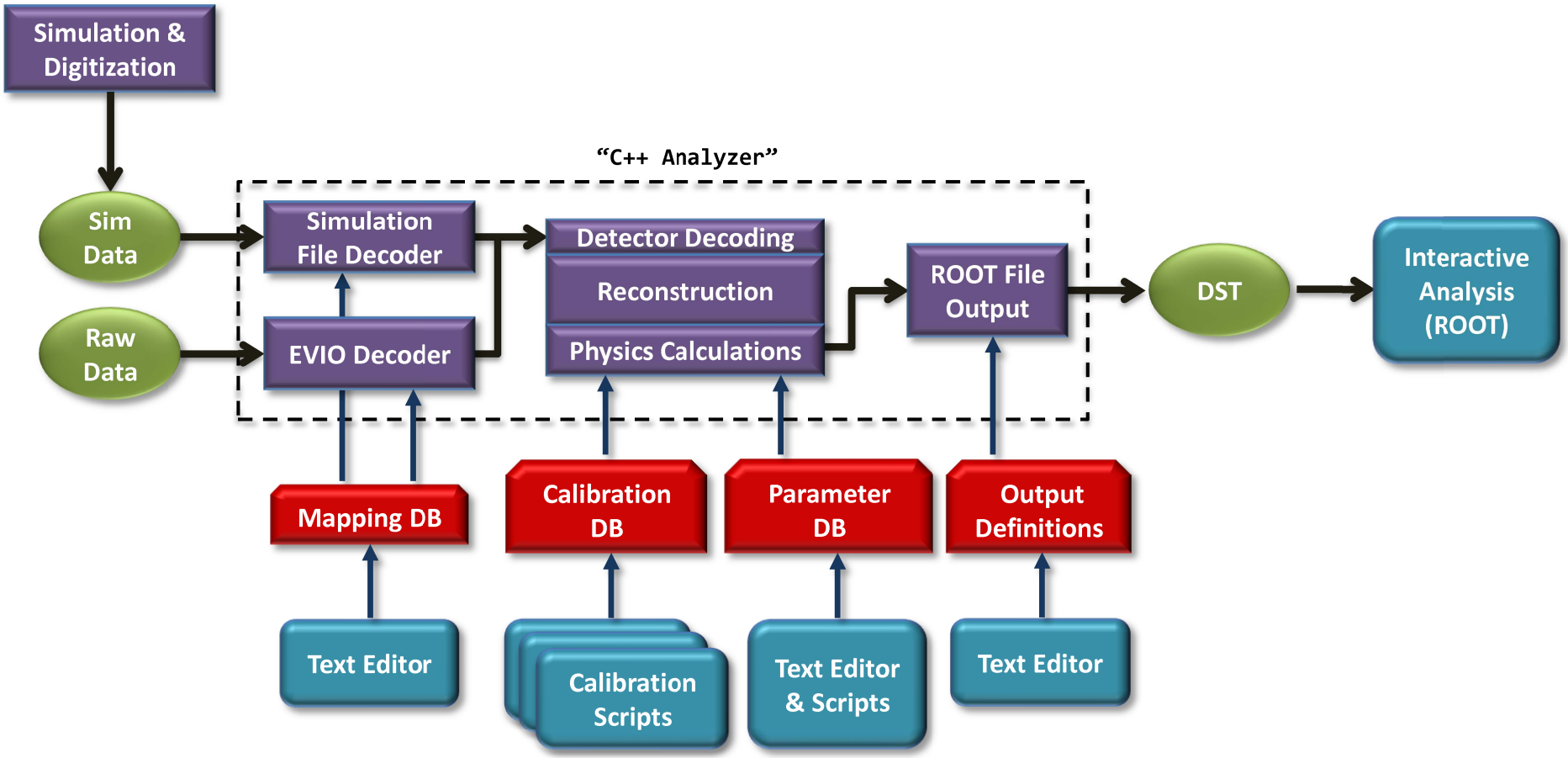
L: L-HRS, R: R-RHS, BB: BigBite, BBG: BB(GEM), ND: neutron det, SBS: SuperBigBite, BC: BigCal, CI: coinc., SA: sing. arm

[1] PREX requirements are negligible compared to APEX

Data Flow


Data Disks
DAQ Front Ends
DAQ Workstations
Online Replay Workstations
Replay Disks
Mass Storage System (MSS) Servers
Tape Silo
Replay Farm
Cache Servers
Cache Disks
File Servers
Work Disks
Interactive Replay
Hall
Counting House
Computer Center
User Offices
Raw Data Flow
Replay (DST) Data Flow
Network
Local bus

Main Reconstruction Software



Hall A "C++ Analyzer"

- C++/ROOT-based framework
- "Straightforward" event-by-event replay
- In production use since 2003
- Highly modular
  - "Everything is a plug-in"
  - Dynamic run-time instantiation of objects representing the configuration and physics of the experiment
  - Dynamic run-time loading of external module libraries
  - Software Development Kit (SDK) available
- Supported on Linux, Mac OS X, Solaris, 32-bit & 64-bit
- Limitations
  - single-threaded $\longrightarrow$ Plan: multi-threaded/multi-process by 2013
  - not distributed $\longrightarrow$ no need anticipated
- Recently begun shared development with Hall C

Additional Software (mostly user-supported)

- Parity-violation experiments (incl. PREX): "Parity Analyzer" (PAN)
  - ROOT/C++-based
  - Independent of standard analyzer
  - Extensively tested in series of prior experiments. Considered ready.
- DVCS
  - Extensive add-on library to standard analyzer
  - Waveform analysis with relatively high CPU requirements
  - Well tested in 6 GeV experiment. Considered essentially ready.
- Polarimeter analysis (Compton, Møller)
  - Standalone code
  - Supported by Hall A staff
  - Generally ROOT/C++-based
  - Minor modifications (*e.g.* support for new front-ends) may be required for 12 GeV. Manpower allocated.
- SBS reconstruction software $\longrightarrow$ next slide

SBS Reconstruction Software

- Pipelined electronics decoder
  - SBS plans extensive use of pipelined electronics
  - Plan to use custom pipelined front-end for APV25 GEM readout
  - Decoder software to be written for both standard and custom modules
- GEM track reconstruction
  - Prototype reconstruction code exists
  - Feasibility for high-rate SBS environment demonstrated for SBS Technical Review, using Monte Carlo input
  - Needs further testing & refinement, especially with real data
- Calorimeter cluster reconstruction
  - Several calorimeter setups proposed for the different SBS experiments (hadron calo, lead-glass calo)
  - "GEp(5)" trigger relies heavily on hadron calorimeter cluster detection
  - At least some cluster analysis software will need to be written. Different algorithms needed for different calorimeters.
- Various user groups (CMU, INFN, UVa, etc.) have taken responsibility for these software projects (*cf.* SBS Program Management Plan)

Calibrations, Data Quality Checks, Prompt Analysis

- Instant replay of raw data from local disks on counting house cluster
- Replay usually in real time $\longrightarrow$ expected to be possible up until high-rate SBS experiments
- Online replay uses the same reconstruction software as offline/farm replay
- Calibrations
  - done on counting house machines or on replay farm, depending on required processing power
  - some standard scripts available
  - custom scripts written for specific experiments by users
- Data Quality Checks via customizable interactive viewer

Simulations

- Typically low volume
- Typically run on user desktops or off-site, but some farm use.
- Existing frameworks
  - SIMA (Hall C's SIMC adapted to Hall A): matrix multiplication for particle transport through spectrometers, no digitization
  - MCEEP: similar to SIMA/SIMC, unmaintained
  - GEMC: Geant4-based Hall B development, adopted by some future large-installation experiments in Hall A (SoLID)
  - SBSsim: Geant4 SBS simulation, maintained by/run at INFN
- Highly specific for each experiment $\rightarrow$ largely user-supported
- Spectrometer-based experiments (*i.e.* early exp'ts through 2015) well simulated with matrix-based codes
- GEMC computing requirements being collected. Modest so far, but expected to increase.

Reconstruction Software Status & Tasks

- Essentially ready for early experiments (though 2015)
- SBS software still largely undeveloped, needed by $\geq$ 2016
- If APEX runs, "must" do before commissioning experiments:
  - VDC track reconstruction for APEX high-rate environment (≈ 1 kHz/cm^2)
    - ★ Code written and successfully tested with test run data
    - ★ Still need to integrate into mainline analyzer (database support etc.)
    - ★ Details this afternoon, if requested
- "Should" do before commissioning, "must" do before SBS
  - Analyzer parallelization
  - Support for pipelined JLab 12GeV DAQ & electronics

Reconstruction Software Status & Tasks Summary

Experiment	Base Software	Required Add-Ons	Required By
APEX	C++ Analyzer	High-rate VDC track reconstruction	mid-2013
PREX	PAN & C++ Analyzer	None	n/a
A_1^n	C++ Analyzer	None	n/a
DVCS	C++ Analyzer	Photon detector analysis (DVCS library)	n/a
SBS program	C++ Analyzer	• Analyzer parallelization • Pipelined electronics decoder • GEM track reconstruction • Calorimeter cluster reconstruction	2016

Red: not yet written Purple: exists, but incomplete and/or not yet fully tested/integrated

Software Management

Project / Subsystem		Responsible / Contact	
		Staff	User
C++ Analyzer development & coordination		O. Hansen	
Front-end decoders		A. Camsonne	
Optics calibrations		D. Higinbotham	
Compton polarimeter		S. Nanda	
Møller polarimeter		J. Gomez	S. Glamazdin (Kharkov)
APEX VDC track reconstruction		O. Hansen	S. Riordan (UMass)
Parity Analyzer		R. Michaels	R. Holmes (Syracuse)
DVCS analysis			C. Muñoz-Camacho (Orsay)
SBS Program			
	Overall coordination	B. Wojtsekhowski	
	GEM track reconstruction		INFN Rome, Carnegie Mellon
	GEM data analysis (calibration)		INFN Catania, UVa
	Hadron calorimeter analysis		Carnegie Mellon
	Coordinate detector analysis		William & Mary

Collaboration Resources

- Annual "Analysis Workshop" in conjunction with collaboration meeting
- Extensive web resources at `http://hallaweb.jlab.org/podd/`
  - Web-based user guide
  - Example scripts
  - ROOT `THtml` reference documentation
  - Software development kit (SDK)
- Hall A Wiki at `https://hallaweb.jlab.org/wiki/`
- Bi-weekly meetings with Hall C

Summary of Hall A Computing Requirements

Anticipated SciComp resource usage assumptions & requirements ▸ spreadsheet

	2013 g2p replay	2014 APEX	2015 A_1^n	2016 DVCS	2017 SBS
Event rate (kHz)	0	5	20	0.5	2
Event size (kB)	0	4	2	30	20
Active data taking (days)	0	17	55	109	109
Raw events per year	0	7e9	9e10	5e9	2e10
Raw data per year (TB)	0	29	189	142	377
Time per event/core (ms)	2	10	2	30	20
Passes through data	1	2	3	3	3
Years to analyze	1	2	2	3	3
Replay duty factor	50%	50%	50%	75%	75%
Output held on work disk	10%	20%	20%	20%	10%
Calculated Totals					
Farm cores (2011 vintage)	6	8	17	33	22
Raw+cooked data to tape (TB)	250	170	410	660	900
Work disk storage (TB)	25	17	44	104	66

Summary & Conclusions

- Hall A has a mature software framework. In production use for over 8 years.
- Hall A reconstruction software essentially ready for early 12 GeV experiments through 2015
- SBS project does require extensive software development. Broad responsibilities for these tasks have been assigned.
- Hall A computing resource requirements through 2017 relatively modest