SHORT DISTANCE STRUCTURE OF NUCLEI - MINING THE WEALTH OF EXISTING JLAB DATA

Larry Weinstein
Old Dominion University, Norfolk, VA
And a cast of thousands

Collaboration:

Spokespeople:
L.B. Weinstein, S.E. Kuhn, M. Strikman, M. Sargsian

Data Mining Scientist: G. Gavalian

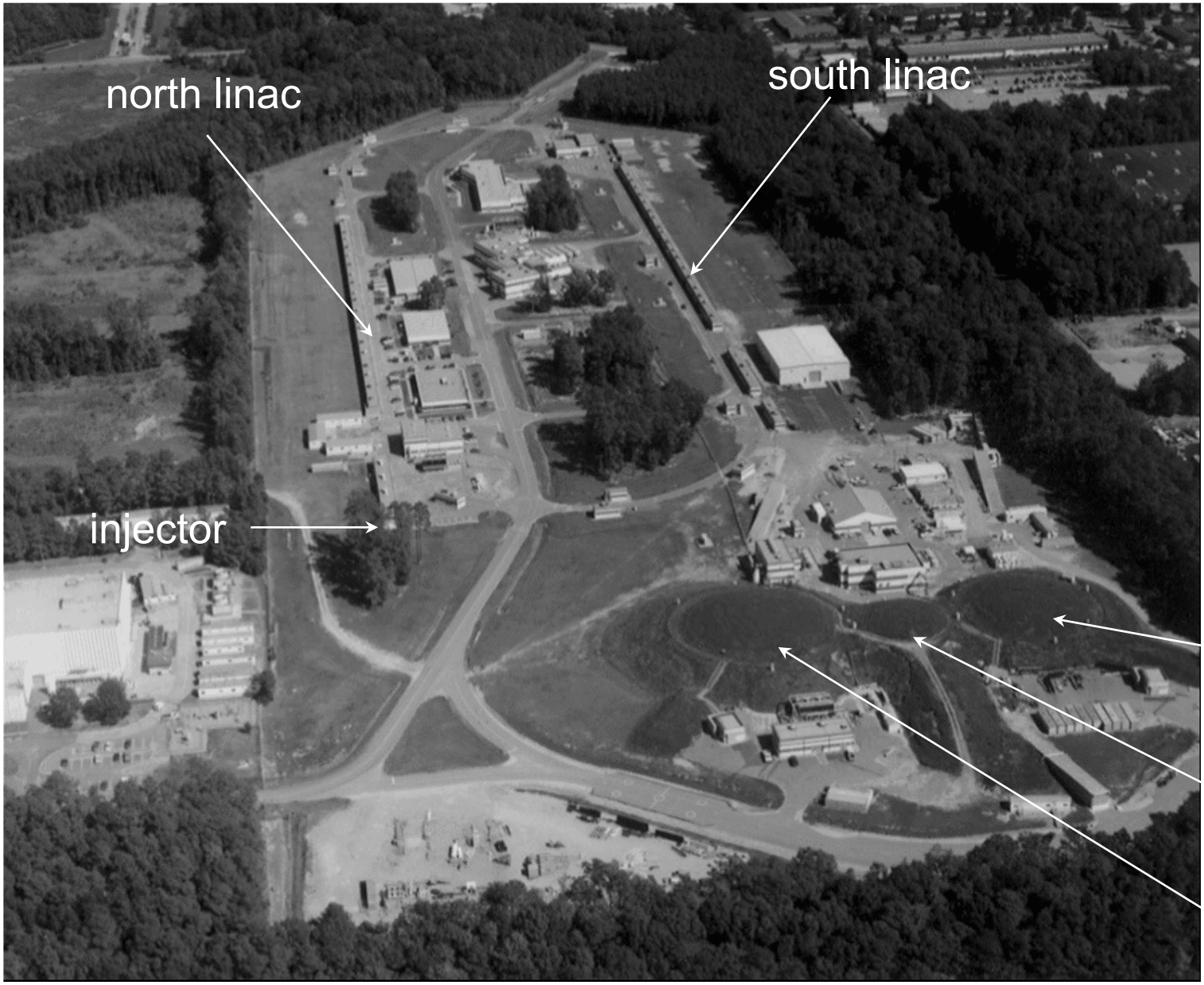
Institutions:
ODU, Penn State, Florida International, Tel Aviv, Glasgow, William and Mary, Edinburgh, Ohio State, New Hampshire, MIT, Richmond, George Washington, Your Name Here

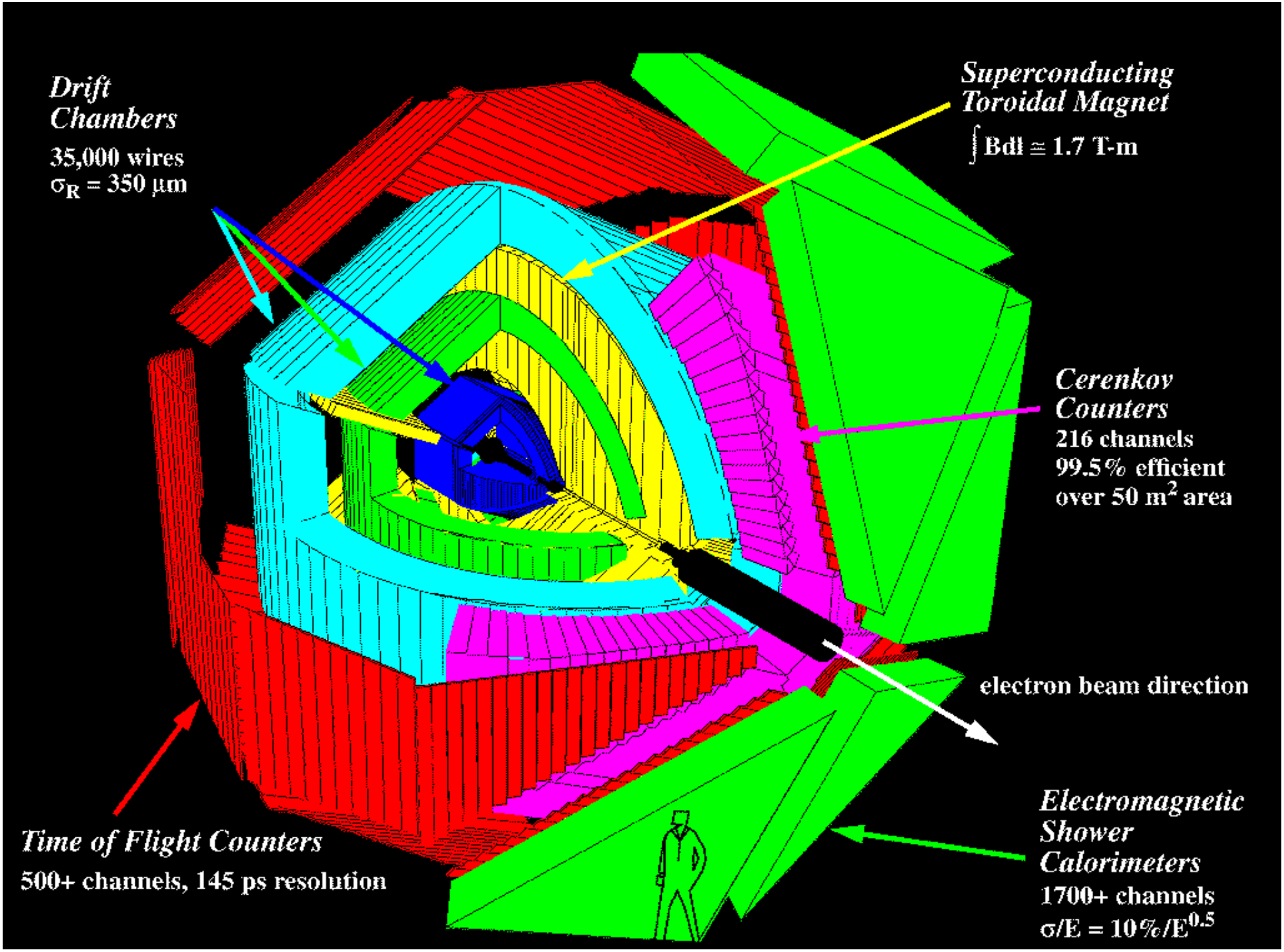
Support:
DOE Grant for two postdocs plus travel 2011-2014
ODU computer support

Why Data Mining?

- Build on the progress made at JLab on SRC and dynamics of interactions with nuclei at medium Q^2 (see, e.g., most of the experimental talks at this workshop)
- Take advantage of the huge CLAS data set
  - Mostly taken with A(e,e') inclusive trigger
- Take advantage of the otherwise "wasted" time while we upgrade JLab to 12 GeV

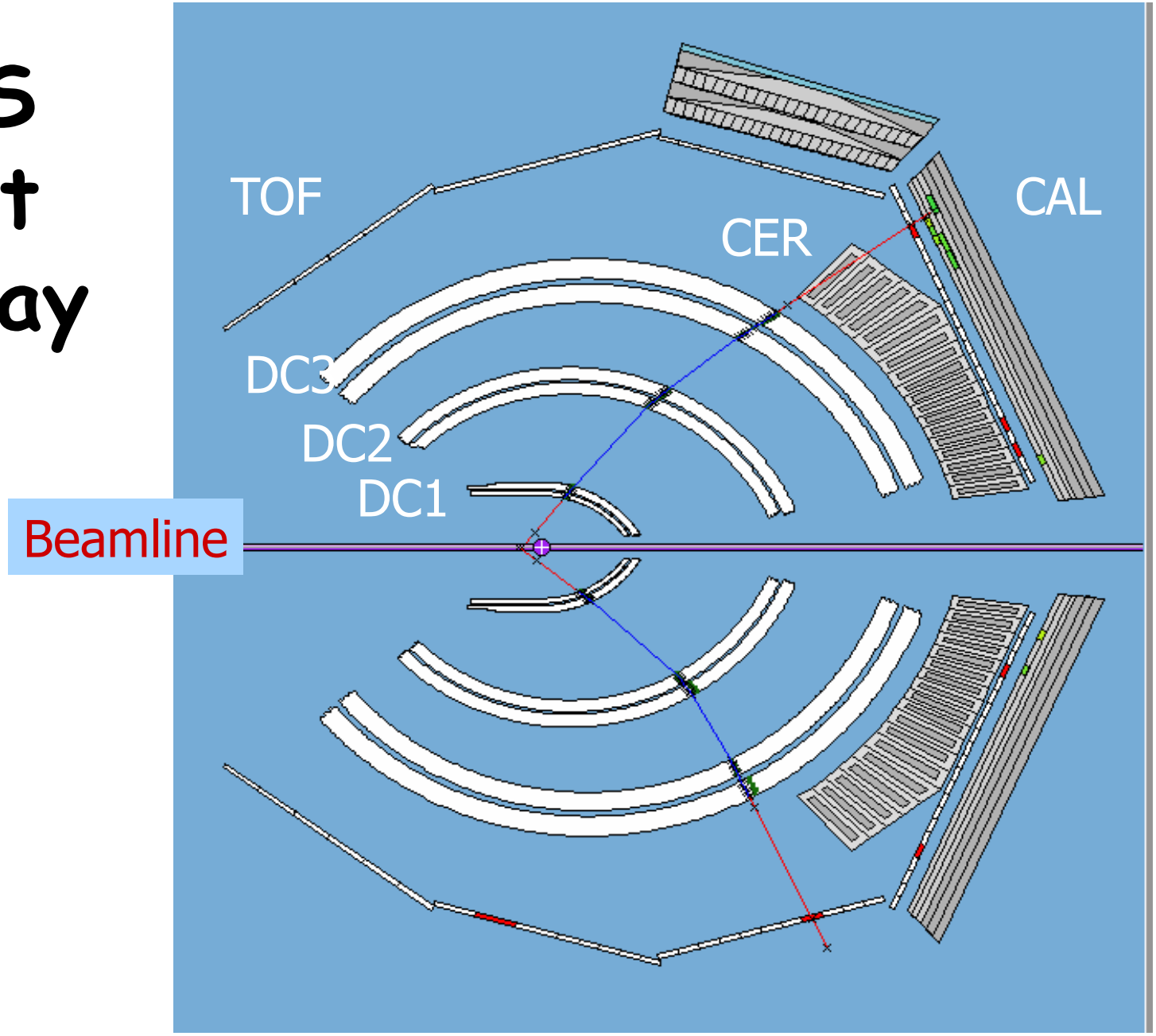
Jefferson Lab Site



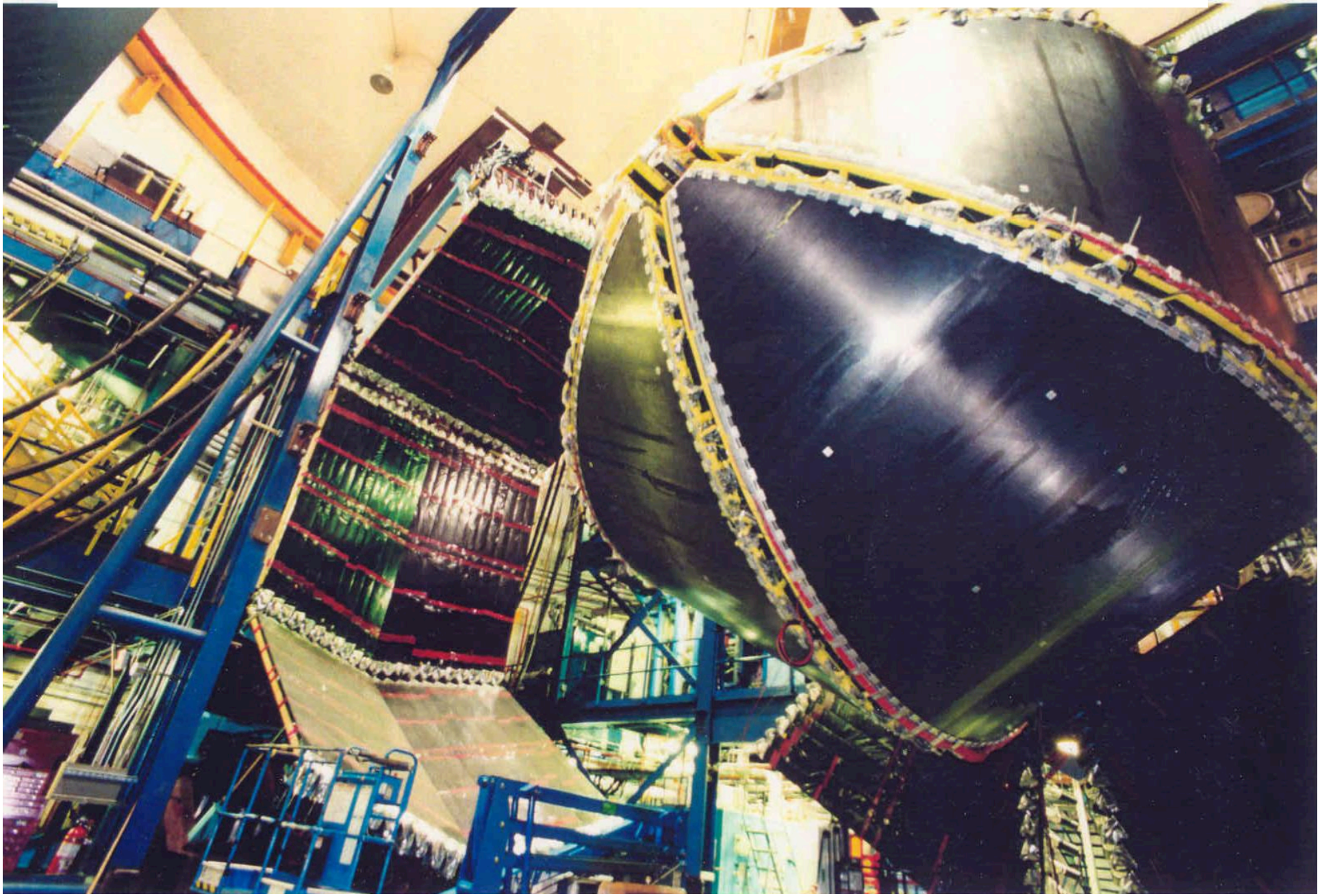

Drift Chambers
35,000 wires
σ_R = 350 μm
Superconducting Toroidal Magnet
$\int Bdl \cong 1.7$ T-m
Cerenkov Counters
216 channels
99.5% efficient
over 50 m^2 area
electron beam direction
Time of Flight Counters
500+ channels, 145 ps resolution
Electromagnetic Shower Calorimeters
1700+ channels
$\sigma/E = 10\%/E^{0.5}$

CLAS Event Display

Sectors 1 and 4



CLAS in Maintenance Position



Data Sets with inclusive triggers

Run Period	Beam	Energy (GeV)	Targets
E2a	e	2.2, 4.4	^{3}He, ^{4}He, C, Fe
E2b	e	1.1, 4.5, 4.7	^{3}He, Fe
EG2	e	4.0, 4.7, 5.0	^{2}H, C, Al, Fe, Pb
G8	gamma	<3.1, <4.0	^{2}H, C, Ti, Fe, Pb
E6	e	5.77	^{2}H
EG1a,b	e	1.6 to 5.7	NH_3, ND_3, C
E5	e	2.56, 4.23	^{1}H + ^{2}H
E1e	e	2.04	^{2}H
EG3	gamma	<5.76	^{2}H
E8 (BoNuS)	e	1.1, 2.2, 4.3, 5.4	^{1}H, ^{2}H, ^{4}He

Topics to investigate

- NN Short Range Correlations with A(e,e' pN)
  - A and Q^2 dependence
  - np vs pp
  - Compare NN forward/forward (x>1) and forward/backward (x<1) events
  - Compare real and virtual photons
    - eg: $\gamma n \rightarrow p\pi^-$ plus a backward proton
- Deuteron d(e,e'p)n
  - unpolarized
  - Beam spin asymmetry (A'_{LT}) $d(\vec{e},e'p)n$
  - Beam and target spin asymmetry $\vec{d}(\vec{e},e'p)n$
- Deuteron d(e,e'p)X
  - DIS and EMC

More Topics

- Delta production
  - Backward emitted Deltas
    - Backward hyperons?
  - Forward Delta++ at $x > 1$
  - Delta production on Quasifree neutron in ^{3}He
- Color Transparency
  - Deuteron transverse kinematics
    - Compare pn and pDelta^0 final states
  - $\pi^+\pi^-$ vs ρ production in nuclei
  - S_{11} production in nuclei
- Your idea here!

Procedure

1. Explore channels and data sets (year 1-3)
   1. Make existing data easily available
      - Led by data-mining scientist (Gagik Gavalian)
   2. Identify most promising analyses
      - Guided by theorists with DM theory postdoc
2. Repeat low level analysis ('recook') some data sets (year 2-3)
   1. Maximize statistics and kinematic coverage
   2. Standardize cuts and corrections
   3. Organized by DM scientist
3. Physics analysis (year 1-3)
   1. Entire DM collaboration

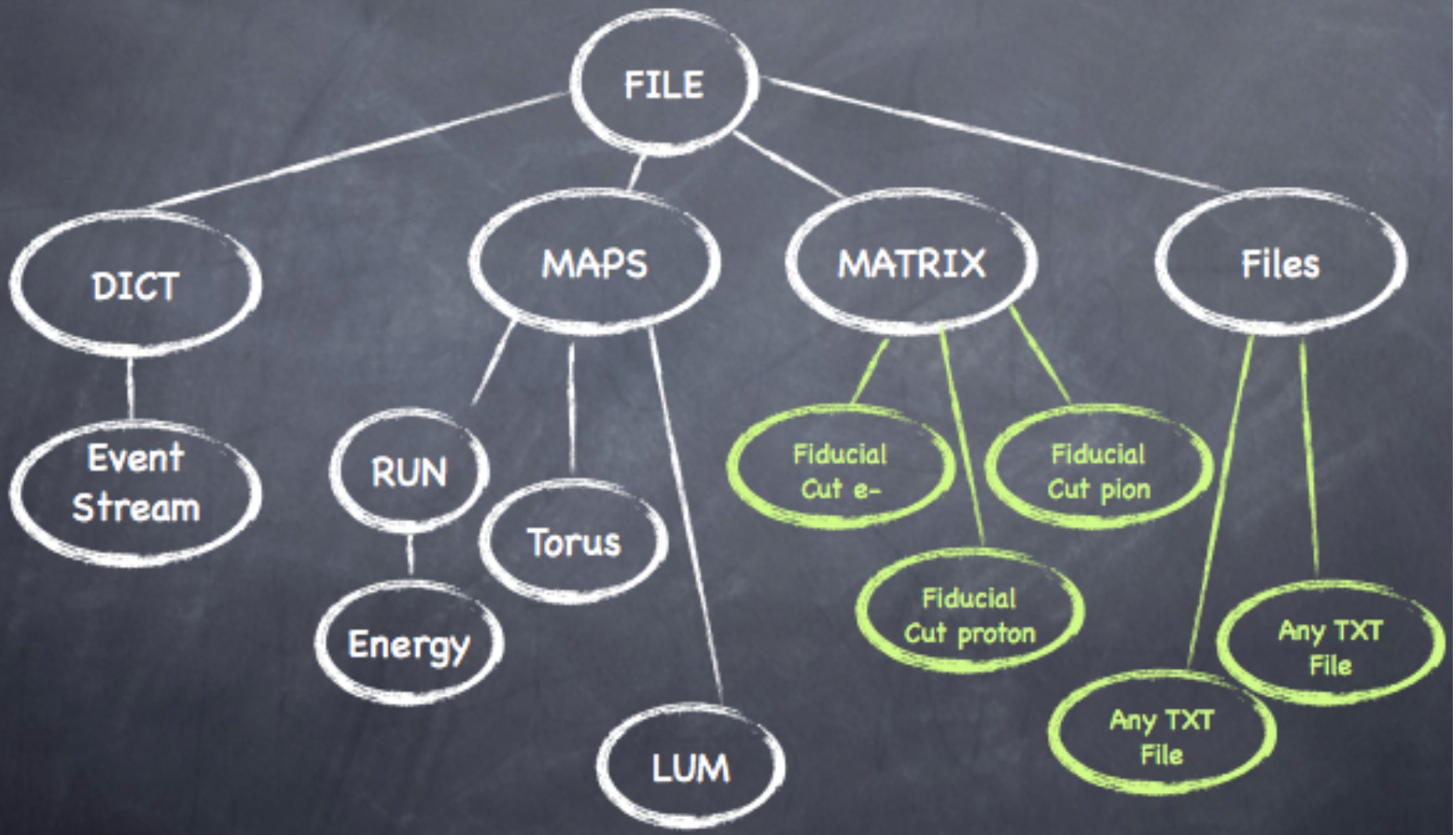
Project Outline

- Collect **all** CLAS nuclear target data in one place with **one** interface
- Provide **easy** access to this data set to participating universities.
- Provide **universal** analysis tools for all data sets (data selection, momentum corrections, fiducial cuts).
- Provide easy framework for combining data from different data sets.
- Provide SOA based multi-process analysis.

Details:

- Processed data stored in a new (HDF5) format
  - Flexible format
  - All needed run information stored in each file
  - Allows random (non-sequential) access
  - Index "interesting events" to process only those
- Data stored at ODU and accessed using CLARA
- A framework was developed around the HDF5-DST's (Data Summary Tapes) for easy (UNIFIED) analysis
  - Python front end
  - No programming, compiling, or linking required
  - Transparent implentation of all standard cuts and corrections
- Output flexibility
  - Event stream
  - Ntuples
  - histograms

Data Structures



Analysis Framework

- Same analysis framework for all data sets with Corrections, Fiducial Cuts and Event Constructor modules loaded automatically based on the run info loaded from the file.
- The data will be indexed. The program will read only events that correspond to selection criteria.
- Parallel processing
- The output can be any of these:
  - Event Stream
  - Ntuples
  - Histograms
- Other analysis tools can be deployed to the server, requires a little advanced programming.

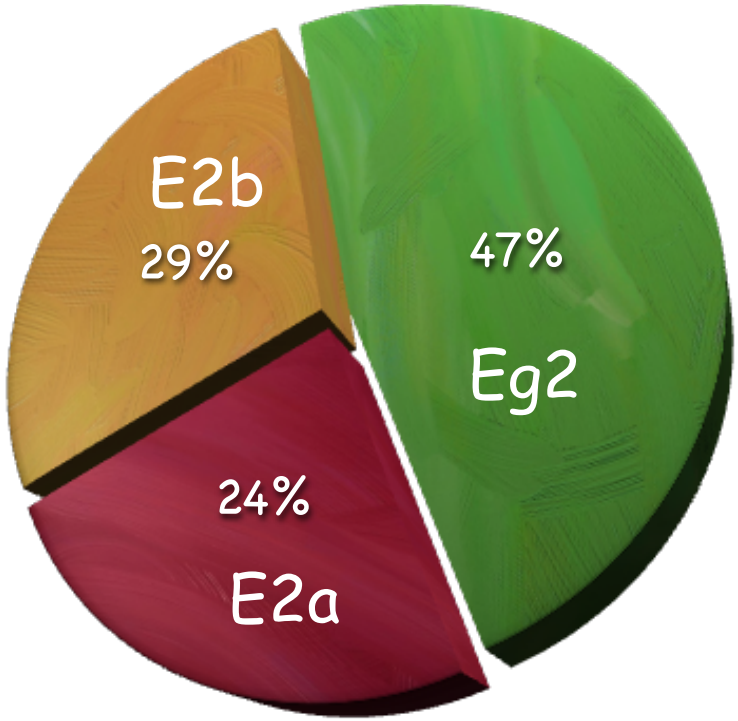
How does it work ? (python)

```
gooApp = gooPyApplication()
gooApp.setInput('e2:e1p1:c12')
//gooApp.setInput('file://myEventStream.hdf5')

gooApp.add('APP::set eventselection (11,2212,-211)')
gooApp.add('APP::set  fiducialcuts true')
gooApp.add('APP::set  momentumcorr  true')
//-----------------------
gooApp.add('NtupleMaker','NT1')
gooApp.conf('NT1:set  eventbank  EVNT')
gooApp.conf('NT1::set ntuplebank NTUP')
gooApp.conf('NT1::variable MxE   VECTMASS([b]+[t]-[11])')
gooApp.conf('NT1::variable MxEp VECTMASS([b]+[t]-[11]-[2212])')
//-------------------------
gooApp.add('HistogramMaker','HST1')
gooApp.conf('HST1::set ntuplebank NTUP')
gooApp.conf('HST1::cut  cMxE  CUT(MxE>2.&&MxE<4.)')
gooApp.conf('HST1::hist  H1MXEP  HIST(100,0.8,1.4,cMxE)')

//gooApp.runLocal() // Run analysis from local file

gooApp.runServer(20) // Run the analysis on the SERVER with 20 jobs
```

Data availability: plans

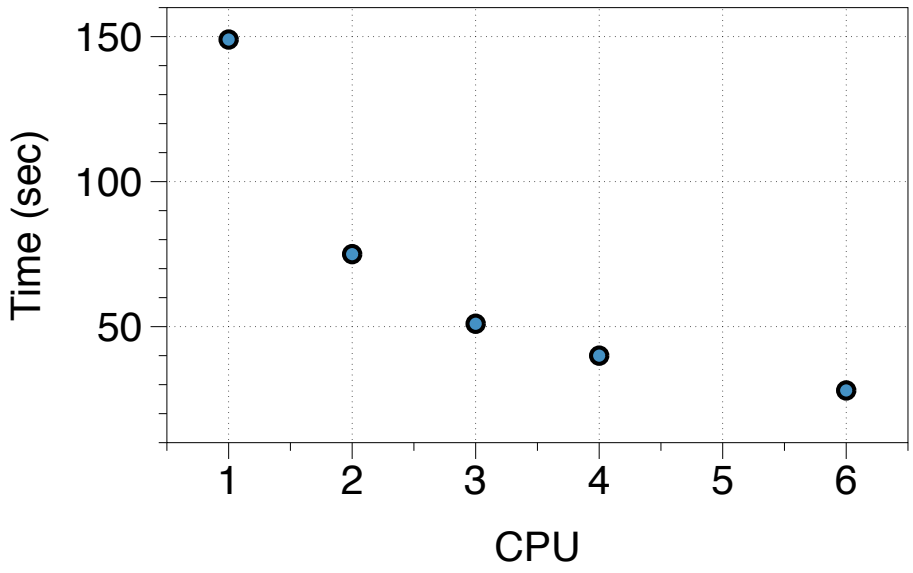
- Will get data from 3 expts within 6 months
- Set up our CLARA service for universal access – tested 11/11/11
- Index the data by event types, target, beam energy and torus current.
  - Event types: (e,e'), (e,e'p), (e,e'pp) …
  - Speed analysis by only reading selected events
- Setup Fast-MC framework for all data sets.

3 TB disk



Successful System Test

11 Nov 2011: Made ntuples at ODU using data stored at JLab



Plan: remote user will make ntuples (or histograms or …) using the ODU farm computers from data stored at ODU.

Summary

- First data mining analyses underway
  - Or Hen (next talk)
  - N. Dashyan: a_2 ratios $A(e,e')/d(e,e')$ at $x > 1$
- Grant approved September 2011
- G. Gavalian organizing data to make analysis easy
- Theory postdoc to help interpret results
- Upcoming workshops
  - Feb 2012 at JLab (CLAS meeting Feb 22-25)
  - Summer 2012 at Tel Aviv?
- Join us!
  - Come analyze some data