#### RCDAQ and EIC and all of that

#### Martin Purschke, BNL

A bit of the RCDAQ history Dedicated EIC test beams and other data-taking campaigns High points of RCDAQ Meta data Actual EPIC detector readout

# History of RCDAQ

- From 2006 or thereabouts we had quite a number of smaller-scale data taking campaigns, R&D for future detectors, but also our then-medical imaging R&D... RCDAQ was "born"
- A large number of commercial or otherwise widely used devices were implemented (today about 120 "active" ones).
- In 2012 we ramped up several EIC-themed measurements of various detectors, GEMs, calorimeters, TOF, RICH, ... all of which used RCDAQ
- RCDAC's first test beam at FermiLab's FTBF was in 2013 (some calorimeter modules)
- 2014 the EIC FLYSUB consortium at the FTBF, sort of the Super Bowl of data taking with 5 setups – BNL (Woody/Purschke), SBU (Hemmick/Dehmelt), UVa (Kondo), Yale (Majka/Smirnov), FIT(Hohlmann)
- 2015 sPHENIX adopts RCDAQ as its DAQ technology, first sPHENIX test beam w/ RCDAQ in 2016 (outer HCal and Emcal, different W calos)
- Also a parade of "field-unrelated" users, UTexas, MPI Munich, NIKHEF, Weizmann, ... ~15 users and also contributors (usually plugins, in a minute)

#### EIC-themed test beams and other datasets

The RCDAQ system has been a pillar of EIC-themed data taking for R&D, test beams etc since 2013 – eRD1, eRD6, LDRDs, ...

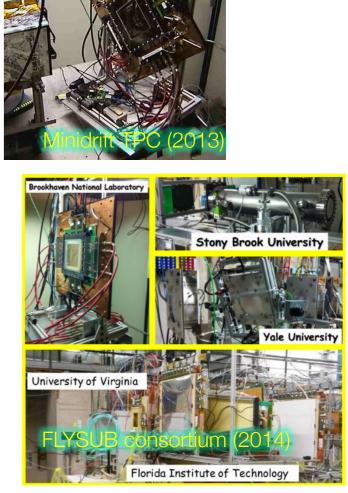
Estimated 25 active RCDAQ installations in the EIC orbit + ~30 elsewhere

Usual entry by ease-of-use for standard devices (DRS, SRS, CAEN, ...) and support for fully automated measurement campaigns

From Markus Diefenthaler at the Users meeting:

- **Detector Simulations** 
  - Validation of Geant4: Make test-beam setup and results available.

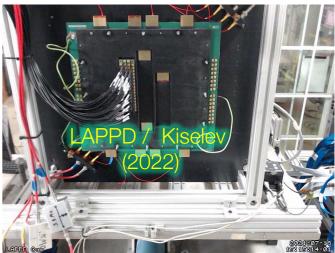
Not sure what prompted that remark... I cannot see how this could be a problem











# Some RCDAQ Concepts

- Modularity
- Binary payload format agnostic
- Different event types
- Set of tools to inspect / display / manipulate files
- Online monitoring support
- Electronic Logbook support
- OS integration
- Interface to community analysis tools (these days: root and 3rd-party frameworks)
- It's gotta be fun to use!

That's quite a list...

#### Data Formats in general...

One of the trickiest parts when developing a new application is defining a data format

It can take up easily half of the overall effort – think of Microsoft dreaming up the format to store this very PowerPoint presentation you are seeing in a file. We used to have ppt, now we have pptx – mostly due to limitations in the original format design

A good data format takes design skills, experience, but also the test of time

The tested format usually comes with an already existing toolset to deal with data in the format, and examples – nothing is better than a working example

Case in point: Parts of the RCDAQ's native format have their roots at the CERN-SPS, and the Bevalac Plastic Ball experiment in the 80's – that's a solid "test of time"

#### Modularity and Extensibility

No one can foresee and predict requirements of a data format 20 years into the future.

Must be able to grow, and be extensible

The way I like to look at this:

FedEx (and UPS) cannot possibly know how to ship every possible item under the sun

But they know how to ship a limited set of box formats and types, and assorted weight parameters Feederate Tube Brance B

Whatever fits into those boxes can be shipped

During transport, they only look at the label on the box, not at what's inside

We will see a surprisingly large number of similarities with that approach in a minute

#### **Events and Packets**

Deliberately lightweight and minimalistic packet and event header structures

No assumption on what "each structure" has in common – you are almost always wrong

```
typedef struct evt_data
{
    unsigned int evt_length;
    int evt_type;
    int evt_sequence;
    int run_number;
    int date;
        8 32bit fields
    int time; 8 32bit fields
    int reserved[2]; =32 bytes
    int data[];
} *evtdata_ptr;
```

```
typedef struct packet_data
{
    unsigned int packet_length;
    short packet_id;
    short packet_type;
    short packet_hitformat;
    short packet_padding;
    short reserved[2]; 4 32bit fields
    int data; =16 bytes
} *packetdata_ptr;
```

Max length for both Event and Packet 16GBytes (length is in units of int's)

#### "Binary payload agnostic" – what is that?

Most of the "devices" we read out provide their data in some pre-made (and usually quite good) compact binary format already. Usually done in some FPGA.

(Think DAM/FELIX.)

All you want to do is to grab the blob of data, stick it into a packet, put a label (packet header) on that says what's in it, done.

That is literally all we do to the data

From that point forward, the DAQ does not care. The "FedEx" approach – they ship boxes, we ship *packets*.

More generally: Usually we store data from our readout devices, but we must be able to store literally *anything* in our data stream.

Want to store an Excel spreadsheet? A text file? A jpeg image? No problem.

If you think "why would one want to do that!", just wait a few minutes.

## Everything in RCDAQ is a shell command

One of the most important features. Any command is no different from "Is -I" or "cat"

That makes everything inherently scriptable, and you have the full use of the shell's capabilities for if-then constructs, error handling, loops, automation, cron scheduling, and a myriad of other ways to interact with the system

Nothing beats the shell in flexibility and parsing capabilities (of course you also have GUIs)

You can type in a full RCDAQ configuration on your terminal interactively, command by command (although you virtually always want to write a script to do that)

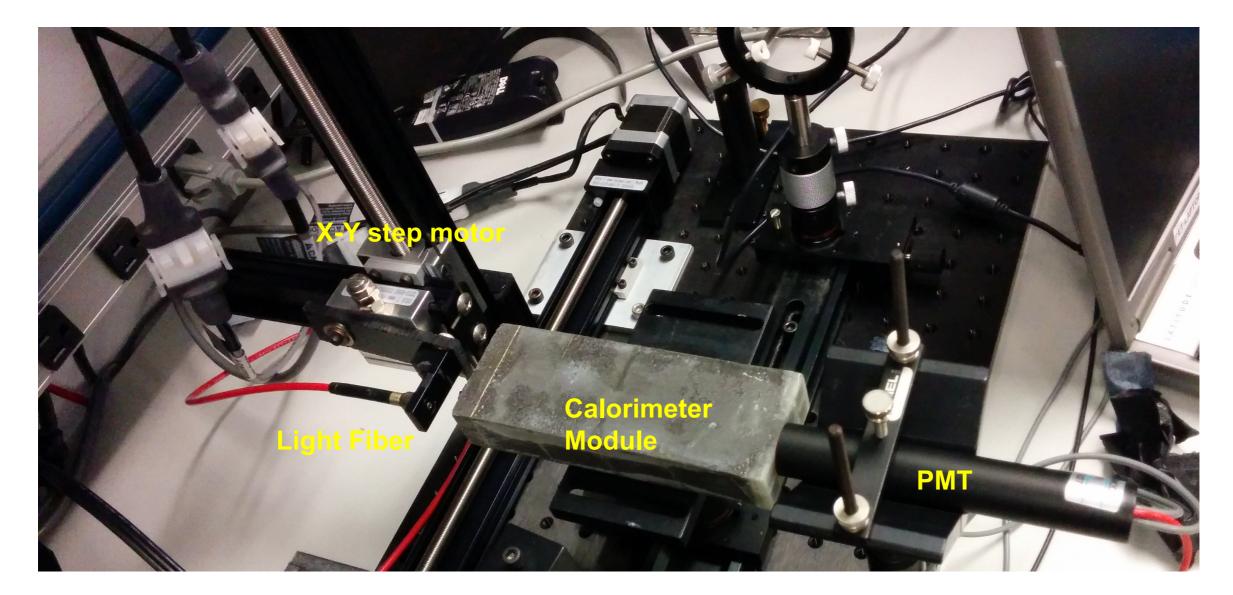
This is quite different from "my DAQ supports scripts"!

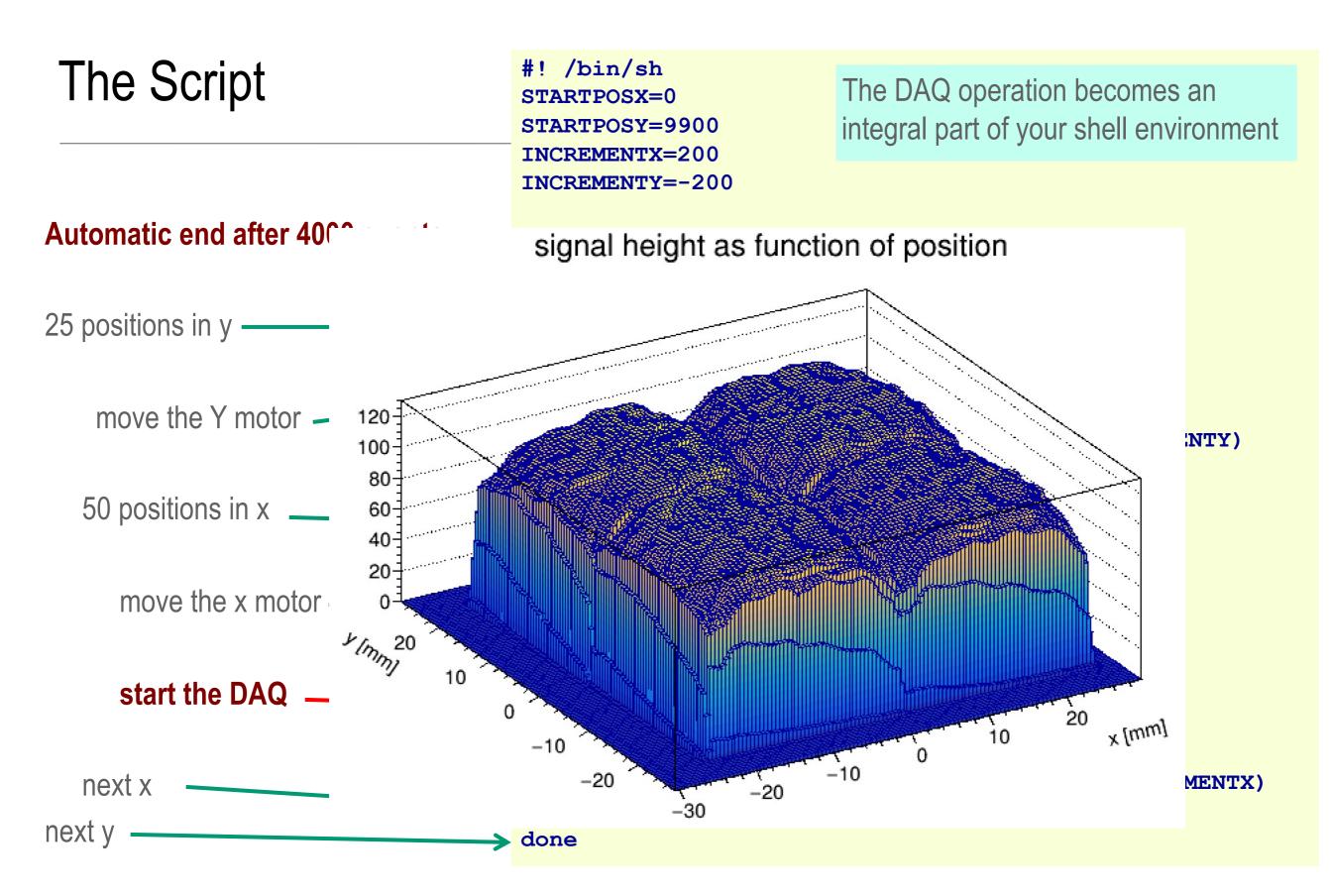
I do not want to be trapped within the limited command set of any application!

As shell commands, the DAQ is fully integrated into your existing work environment.

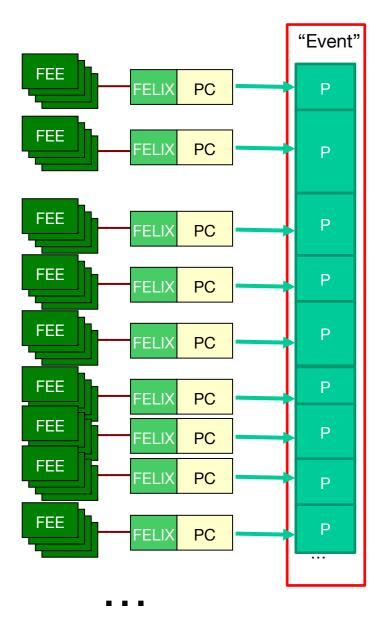
#### Measurements on autopilot

You want to run measurements where you step through some values of a parameter completely on autopilot Here: Move a light fiber with 2 step motors, take a run for each position w/ 4000 events  $50 \times 25 = 1250$  positions (you really want to automate that) Let it run overnight, come back in the morning, look at the data





#### Event / Streaming Data Structures



Each Front-End Card generally contributes what we call a "Packet" to the overall event structures

A Packet ID uniquely identifies the detector component / front-end card where it comes from

A hitformat field identifies the format of the data, und ultimately selects the decoding algorithm

You interact with a standard set of APIs to access the data

We can change/improve the binary format and assign a new hitformat for a packet at any time

Insulation of offline software from changes in the online system

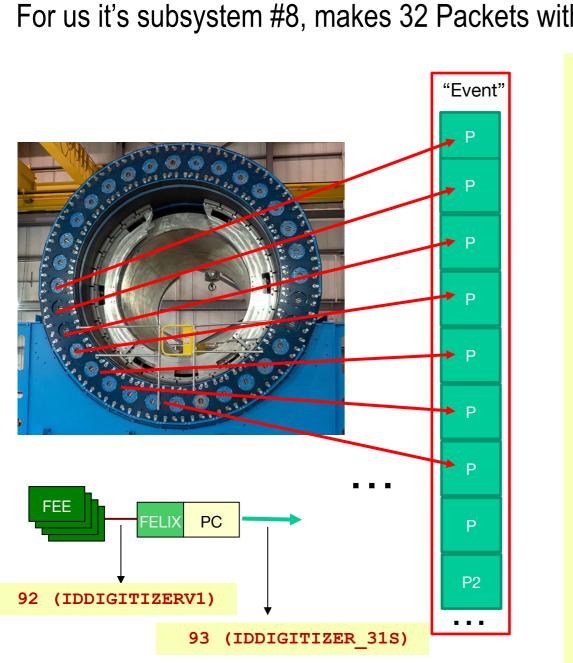
API delivers the data independent of internal encoding

Very rough number: 1200-2500 packets collectively

In case of a triggered DAQ, such an event structure and the packets therein would correspond to the data from one crossing

#### Example: Full EPIC Outer HCal, Real Events

That's one of the detectors that will survive into Detector 1



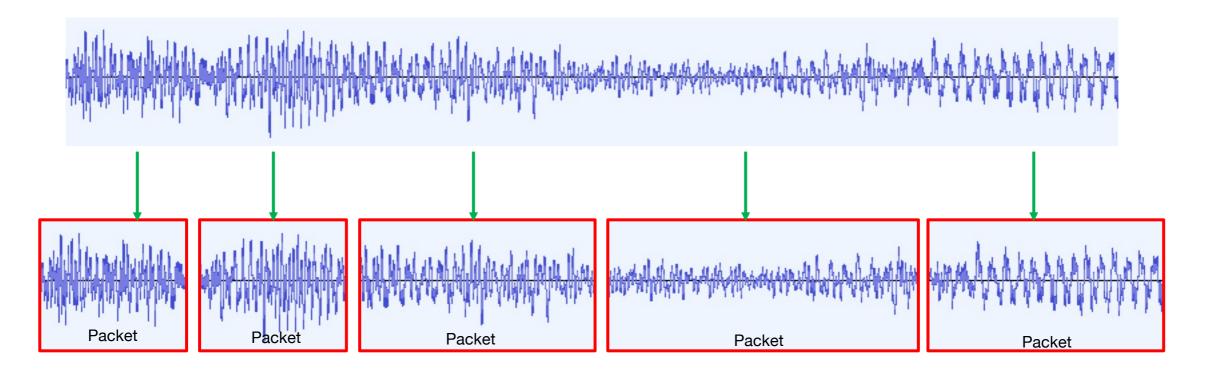
Different formats from the front-end or after the full chain – no change in analysis code!

/[	th IDs 8	3001	- 8032			/						
	\$ dlist	oHCal	-0000010	0-0000.evt								
	Packet	8001	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8002	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8003	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8004	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8005	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8006	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8007	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8008	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8009	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8010	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8011	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8012	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8013	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8014	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8015	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8016	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8017	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8018	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8019	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8020	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8021	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8022	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8023	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8024	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8025	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8026	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8027	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8028	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8029	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8030		(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8031	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						
	Packet	8032	1000 -1	(sPHENIX Packet)	92	(IDDIGITIZERV1)						

Hitformat

#### Streaming Readout and Packets

For streaming data, the "Packet" paradigm changes its meaning a bit It becomes like a packet in the Voice-Over-IP sense - VoIP is chopping an audio waveform into conveniently-sized chunks to transfer through a network



We are chopping the streaming detector data into conveniently-sized packets for storage Here: Streaming sPHENIX TPC data (entire sPHENIX tracking system streams!)

```
$ dlist rcdaq-00002343-0000.evt -i
-- Event 2 Run: 2343 length: 5242872 type: 2 (Streaming Data) 1550500750
Packet 3001 5242864 -1 (sPHENIX Packet) 99 (IDTPCFEEV2)
$
```

#### Streaming readout, here we come!

Past the Front-end, the readout is **completely oblivious to the readout mode** It doesn't care how the front-end + FELIX arrived at the decision to send up the data. Triggered or streaming, from the readout perspective they look 100% the same Our FELIX drivers adhere to POSIX standards (here: our INTT):

```
FD_SET(_intt_fd, &read_flags);
sel = select(_intt_fd+1, &read_flags, NULL, NULL, &timeout);
if (FD_ISSET(_intt_fd, &read_flags))
        {
            // data available . . .
```

```
sevt->packet_length = SEVTHEADERLENGTH;
sevt->packet_id = ipacket;
sevt->packet_type=2;
sevt->packet_decoding = IDINTT_v1;
sevt->reserved[0] = 0;
sevt->reserved[1] = 0;
uint16_t *dest = (uint16_t *) &sevt->data;
int ret = read(_intt_fd, dest, _length);
sevt->packet_padding = ret%2 ;
sevt->packet_length += (ret + sevt->packet_padding);
return sevt->packet length;
```

So we grab the data when they come

#### RCDAQ doesn't know "readout"...

#### Huh?

Out of the box, RCDAQ doesn't know how to read out any kind of hardware. Nada.

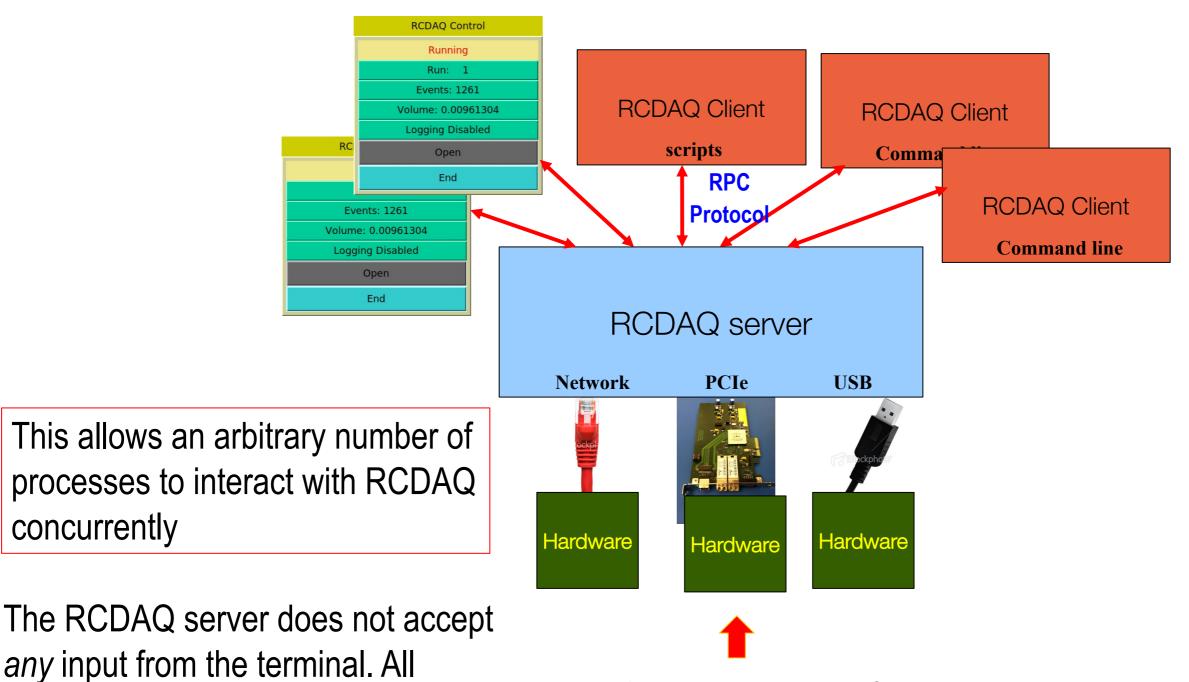
RCDAQ is "taught" about any hardware device by way of a plugin

Many readout routines wouldn't co-exist to begin with, no monolithic binary possible

<pre>\$ daq_status -11 Stopped Logging disabled Filerule: rcdaq-%08d-%04d.evt Buffer Sizes: 32832 KB adaptive buffering: 15 s Web control Port: 8899 Elog: not defined defined Run Types: (none) No Plugins loaded</pre>	This makes RCDAQ modular and also distributable Several plugins contain licensed (all FELIX firmware) or EULA-restricted code (CAEN), etc But we can make RCDAQ and all other plugins available!					
<pre>\$ rcdaq_client load librcdaqp1 \$ daq_status -11 Stopped Logging disabled Filerule: rcdaq-%08d-%04 Buffer Sizes: 32832 KB a Web control Port: 8899 Elog: not defined  defined Run Types: (non List of loaded Plugins: - INTT Plugin, provides -</pre>	<pre>Stopped Logging disabled Filerule: rcdaq-%08d-%04d.evt Buffer Sizes: 32832 KB adaptive buffering: 15 s Web control Port: 8899 Elog: not defined  defined Run Types: (none) List of loaded Plugins:</pre>					

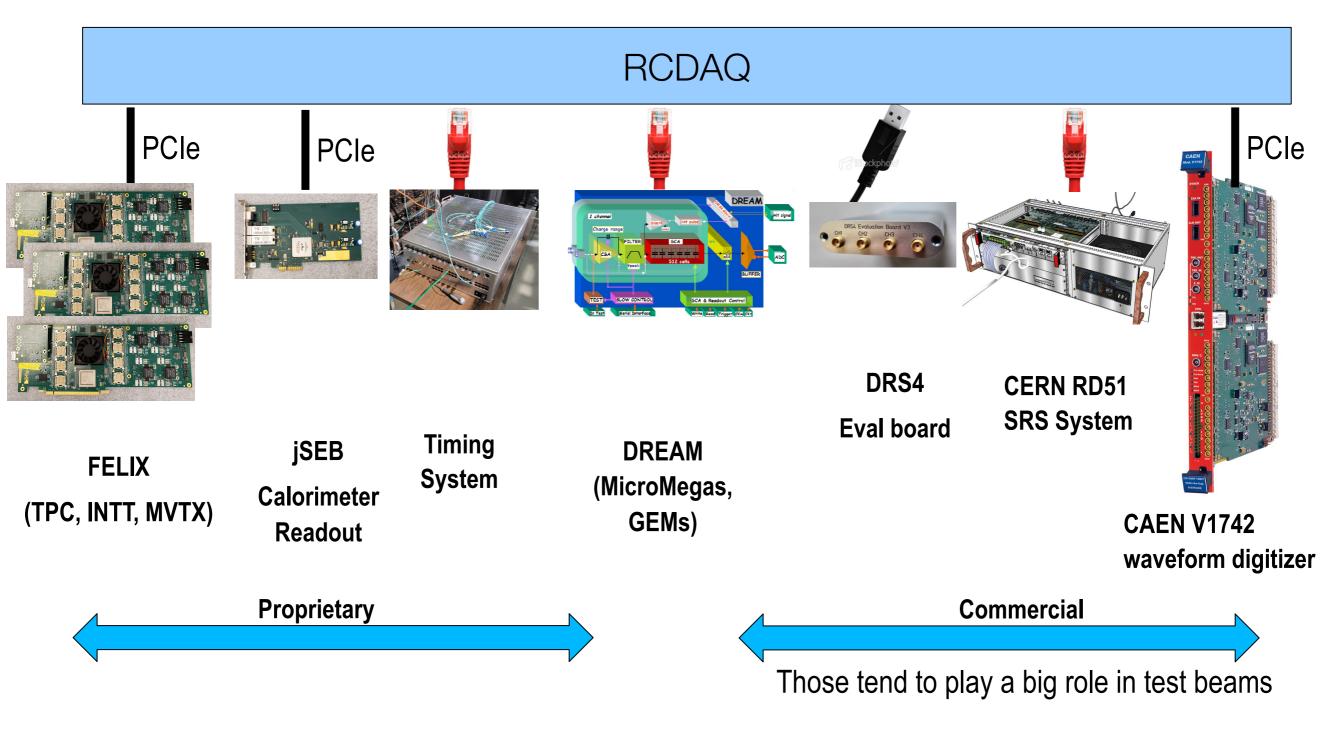
#### The RCDAQ client-server concept

interaction is through the clients.



"The 3 main pathways into a PC"

#### A selection of devices implemented in RCDAQ



There are many more not shown...

#### Where does RCDAC run?

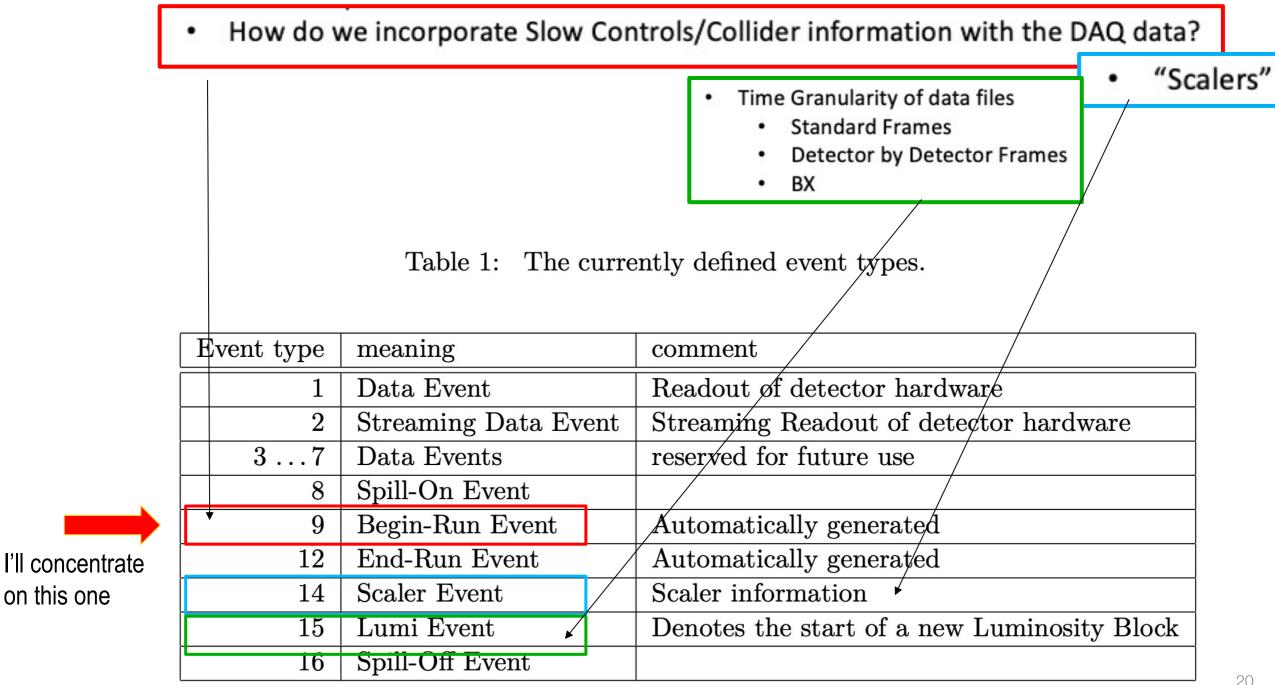
- Pretty much any Linux flavor. RHEL, CentOS, Debian, Ubuntu, Arch, ... and Raspian!
- Yes, it can run on a Raspberry Pi (no PCIe devices, of course)
- Not really just a gimmick, we have run weeks-long cosmics measurements where you don't want to tie up a higher-end PC, and you get like 250Hz event rate with the DRS4
- Runs on PCs, laptop, and of course what sPHENIX has right now in Bldg 1008...



That's more than we need in EPIC's year-3

#### Different Event Types for Meta- and other data

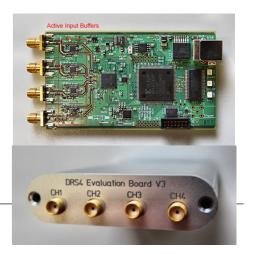
Jeff last week:



# Capturing your environment

- The most-often used event is the begin-run event
- Automatically generated (same as end-run)
- Guaranteed to be first and last, respectively (super-useful in continuous online monitoring)
- Each event type has its own "list of devices to read"
- data events typically read your detectors, begin-run or other types different things
- (some info in the backup, gets too long here)

Let me show you the application of beg-run that is pretty much always used...



# Setting up and reading out a DRS4 Eval board

- \$ rcdaq\_client load librcdaqplugin\_drs.so
- \$ rcdaq\_client create\_device device\_drs -- 1 1001 0x21 -150 negative 140 3
- \$ daq\_open
- \$ daq\_begin
  - # wait a while...
- \$ daq\_end

You can, but of course you wouldn't do it like that

### A Setup Script

Now you got yourself a setup script as I advertised before, call it, say,

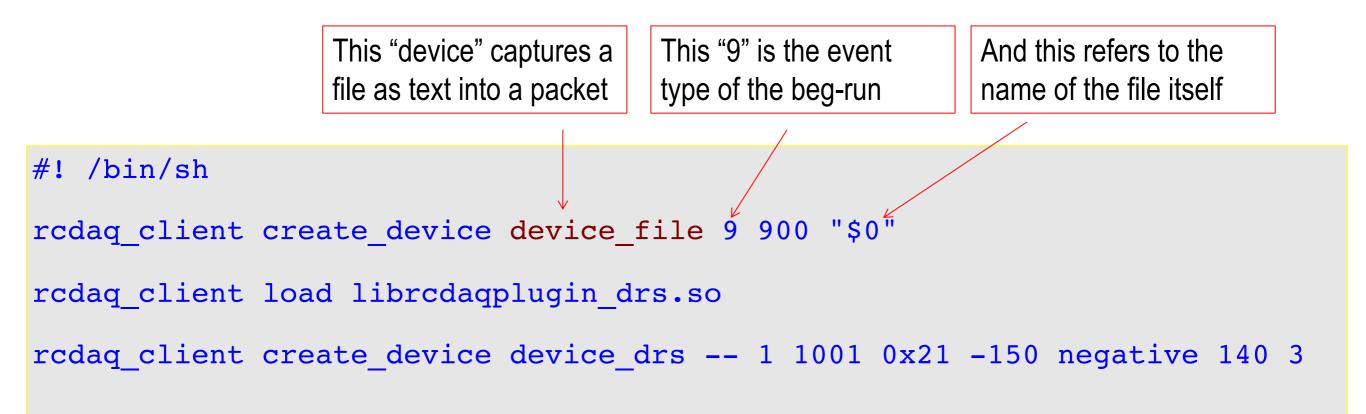
"setup.sh"

#! /bin/sh
rcdaq\_client load librcdaqplugin\_drs.so
rcdaq\_client create\_device device\_drs -- 1 1001 0x21 -150 negative 140 3

Make it executable and you can re-initialize your DAQ each time the same way

#### Capturing the setup script for posterity

We add this very setup script file into our begin-run event for posterity



So this gets added as packet with id 900 in the begin-run

It's not quite right yet - \$0 is usually just "setup.sh", so the server may not be able to find it. We need the name with a full path

# Expanding the \$0 to a full filename

The readlink expands the file to a full filename

```
#! /bin/sh
MYSELF=$(readlink -f $0)
rcdaq_client create_device device_file 9 900 "$MYSELF"
rcdaq_client load librcdaqplugin_drs.so
rcdaq_client create_device device_drs -- 1 1001 0x21 -150 negative 140 3
```

In the same spirit, you can capture anything else as well

Some "pseudo devices": device\_file, device\_filenumbers, device\_file\_delete, device\_filenumbers\_delete, device\_command

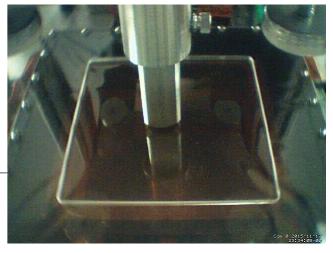
This one is super-powerful

#### Example: Characterizing GEM gain uniformity

Outer script lays a 20x20 or so grid as I've already shown

In begin-run we add a "device\_command" that executes another script

That reaches out to the step motors, gets the x/y positions into a file like this 8377



GEM under a moving X-ray gun with step motors

8031

device\_file captures as ASCII text, device\_filenumbers binary like ADC values (easy programmatic access) We also capture a webcam image for feel-good value (does it move right when it should?)

#### #add the position information

rcdaq\_client create\_device device\_command 9 0 /home/eic/struck/getmotorpositions.sh

rcdaq\_client create\_device device\_file 9 910 /home/eic/struck/positions.txt

rcdaq\_client create\_device device\_filenumbers\_delete 9 911 /home/eic/struck/positions.txt

# add the camera picture

rcdaq\_client create\_device device\_command 9 0 "/home/eic/capture\_picture.sh
 /home/eic/struck/cam\_picture.jpg"

rcdaq\_client create\_device device\_file\_delete 9 940 /home/eic/struck/cam\_picture.jpg

#### Super-easy analysis!

Your throw all 400 files at your analysis

In each data file, the analysis recognizes the begin-run, extracts x, y

Goes through the actual data, determines "gain"

Hits end-run, fills "gain" in the right x/y slot (like a 2D-histogram)

Next file until done. Have a coffee, come back and enjoy the result.

No database access needed! All info contained in the data files.

```
eicdaq2 ~ $ ddump -p 910 -t 9 ZZ48_000001600-0000.evt

8031

8377 

eicdaq2 ~ $ ddump -p 910 -t 9 ZZ48_000001601-0000.evt

8031

eicdaq2 ~ $ ddump -p 910 -t 9 ZZ48_000001602-0000.evt

8031

8409 

eicdaq2 ~ $ ddump -p 910 -t 9 ZZ48_000001603-0000.evt

8409 

eicdaq2 ~ $ ddump -p 910 -t 9 ZZ48_000001603-0000.evt

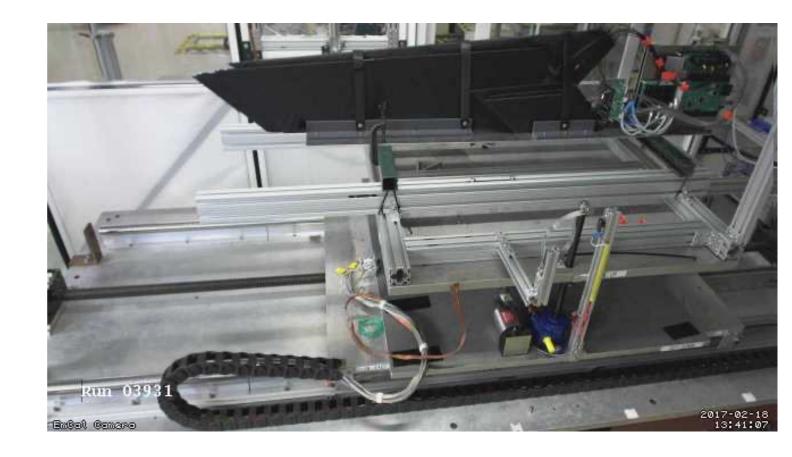
8425
```

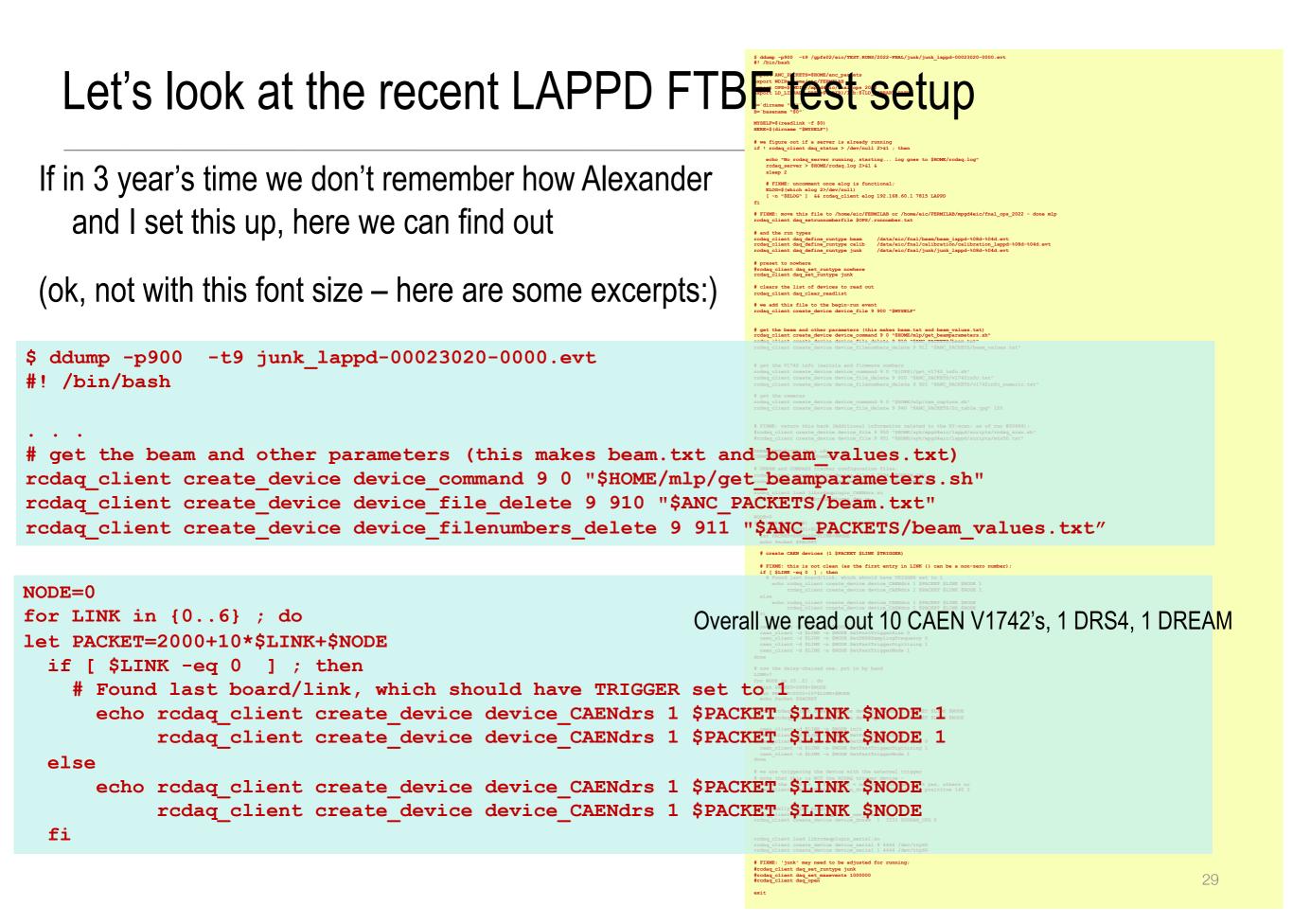
GEM under a moving X-ray gun with step motors

# Autopilot example: "Tile Mapping" at the Fermi Test Beam Facility

- "Tile mapping" refers to mapping the position-dependent response of a hadronic calorimeter tile.
- About 200 individual positions of the tile relative to the beam you'd go nuts doing all that manually, and you are bound to make mistakes
- The FTBF M2.6 table is controlled via a script that drives to predetermined positions

Same deal as before, positions in the data files, analysis is a snap





#### Forensics

We tend to capture "everything" we can, like a plane's black box... Example:

"It appears that the distributions change for Cherenkov1 at 1,8,12,and 16 GeV compared to the other energies. It seems that the Cherenkov pressures are changed. [...] Any help on understanding this would be appreciated."

Martin: "Look at the info in the data files:"

#### More Forensics (my poster child why this is so useful...)

"There is a strange effect starting in run 2743. There is a higher fraction of showering than before. I cannot see anything changed in the elog."

Look at the cam pictures we automatically captured for each run:

\$ ddump -t 9 -p 940 beam\_00002742-0000.prdf > 2742.jpg \$ ddump -t 9 -p 940 beam\_00002743-0000.prdf > 2743.jpg



#### "Meta Data" Packet list from a recent test beam

	Ad	ditional Pa	ickets	More than 72 environment-capturing						
Begin Run event (type 9) Data Event (type 1)		hitformat	comment							
900	- IDCSTR copy of the setup script for this run			packets (accelerator params, voltages,						
910 1110 IDCSTR beam line info ascii			beam line info ascii	currents, temperatures, pictures,)						
911 1111 ID4EV			beam line info binary (*10000)							
940	-	IDCSTR	picture from our cam of the hcal platform							
941	-	IDCSTR	picture from the facility cam inside the hutch	964	-	IDCSTR	EMCAL_A0 (not changing during run)			
942	-	IDCSTR	picture from the facility cam through the glass roo	968	1068	ID4EVT	EMCAL_KEITHLEY_CURRENT binary			
943	-	IDCSTR	picture from our cam of the Emcal table	969	1069	ID4EVT	EMCAL_KEITHLEY_VOLTAGE binary			
950	1050 IDCSTR HCAL_D0 readback		970	1070	ID4EVT	HCAL_D0 binary				
951	1051	IDCSTR	HCAL_D1 readback	971	1071	ID4EVT	HCAL_D1 binary			
952	1052	IDCSTR	HCAL_10 readback	972	1072	ID4EVT	HCAL_I0 binary			
953	1053	IDCSTR	HCAL_I1 readback	973	1073	ID4EVT	HCAL_I1 binary			
954	1054	IDCSTR	HCAL_T0 readback	974	1074	ID4EVT	HCAL_T0 binary			
955	1055	IDCSTR	HCAL_T1 readback	975	1075	ID4EVT	HCAL_T1 binary			
956	1056	IDCSTR	HCAL_GR0 readback	976	1076	ID4EVT	HCAL_GR0 binary			
957	1057	IDCSTR	HCAL_GR1 readback	977	1077	ID4EVT	HCAL_GR1 binary			
958	1058	IDCSTR	HCAL_KEITHLEY_CURRENT	-	1078	ID4EVT	HCAL_KEITHLEY_CURRENT binary			
959	1059	IDCSTR	HCAL_KEITHLEY_VOLTAGE	-	1079	ID4EVT	HCAL_KEITHLEY_VOLTAGE binary			
960	1060	IDCSTR	EMCAL_D0	980	1080	ID4EVT	EMCAL_D0 binary			
961	1061	IDCSTR	EMCAL_I0	981	1081	ID4EVT	EMCAL_I0 binary			
962	1062	IDCSTR	EMCAL_T0	982	1082	ID4EVT	EMCAL_T0 binary			
963	1063	IDCSTR	EMCAL_GR0	983	1083	ID4EVT	EMCAL_GR0 binary			
		984	-	ID4EVT	EMCAL_A0 binary (not changing during run)					
				988	1088	ID4EVT	EMCAL_KEITHLEY_CURRENT binary			
Captured a	t Captu	red a	igain	989	1089	ID4EVT	EMCAL_KEITHLEY_VOLTAGE binary			
begin-run at spill-off										

## Summary

- Rock-solid, scalable, versatile DAQ system
- Large user base in the EIC community and outside
- As well as many "power users" Kiselev, Azmoun, Stoll, Hemmick, Hohlmann, Tsai, ...
- Superb analysis and online monitoring support (another time)
- Tons of operational experience in sPHENIX, test beams (permanent setup @ the FTBF), dozens of lab setups
- Hard to overstate to importance of having the DAQ + format in hand from day-1 student training, experience
- Early test beam online monitoring and analysis code often grows into the eventual reco code

# That's all for today

Plenty more EPIC-relevant things to talk about maybe at a future meeting

- Current benchmarks
- Timing system integration
- Analysis support / APIs
- Getting Beam Crossing info
- Data inspection tools
- Online monitoring
- CD-2/3a Experience

Thank you!

(I leave you here with a Single-Event Display of an EPIC outer HCal event, real data, cosmics)

