RECO Status: Goals, Schedule & Effort

Harry Melanson, Fermilab

September 14, 2001

All information shown in this talk is kept up to date on the Algorithms web page:

D0 AT WORK

Algorithms

http://www-d0.fnal.gov/computing/algorithms/

Information is also regularly distributed to *d0algo@fnal.gov*

(To subscribe, see http://listserv.fnal.gov/users.html)

RECO Production Release Schedule

- We build a production release every 3 months (“quarterly”)
- We are currently working on “p10”
- Production releases are *intended* to be stable, well-understood, and timely
- In order to converge on schedule, we go through the following phases for each production release:
  - Major development frozen one month before pxx.00.00 release
  - Bug fixes only allowed two weeks before pxx.00.00 release
  - Certification on “large reco certification samples” for two weeks after pxx.00.00 created, using MC (pxx – 1) and *data* + bug fixes
  - Certification on “moderate reco certification samples” for additional two weeks using MC (pxx) + bug fixes

A first for p10 RECO

RECO Production Release Schedule p10

• **June 23, 2001**: All MAJOR improvements to DØRECO MUST be released in this week's test build. Any major modifications that miss this date must be targeted for next production release.

• **August 6, 2001**: No new functionality will be allowed in this week's test build. Only bug / mistake fixes will be allowed in this build. (and there better not be many...)

• **August 20, 2001**: Production version is created, and formal certification begins.

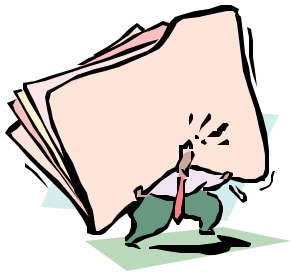
• **Sept. 7, 2001**: Reports on p09 and raw data certification samples at a SPECIAL Algorithms meeting. Declaration of the production worthiness of executable will then be made.

• **Sept. 21, 2001**: Certification on p10 generated samples.

RECO Production Release Schedule
p10 status

- t01.56.00 (same as p10.00.00) created on schedule
- t01.56.00 currently being used to process real data
- p10.03.00 created this week, p10.04.00 scheduled for next Tuesday
- p10.04.00 should fix “muon PDT problem” and “large output size” – if so, will request farms upgrade to this for official real data processing
- Still need to certify using p10 Monte Carlo samples (one new feature is first SMT noise simulation) – **Sept. 21**
- Up-to-date status at:

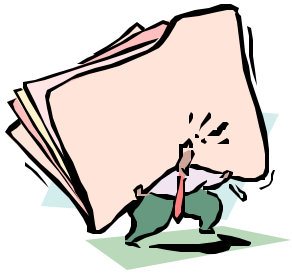
http://www-d0.fnal.gov/computing/algorithms/status/p10.html



Algorithms Group Effort Reporting

In August, I asked all Algorithm / Object ID groups to report:

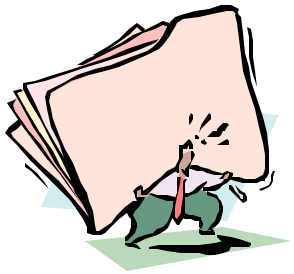
- Who has been working in your group over the *last* three months?
- At what level of a "Full Time Equivalent" (FTE)?
- Who do you anticipate will work within your group for the *next* three months, and at what level?
- What tasks do you have to accomplish over the next three months, for which you have *no one committed*?
- How many FTE's will these projects require?



Algorithms Group Effort Reporting

Group	Active FTE	Projected	Additional Needed	Additional / Needed	Total Required
Align	3.15	3.00	10.30	3.43	13.30
Calib	3.25	3.50	1.25	0.36	4.75
SMT	1.60	1.70	2.50	1.47	4.20
CFT	1.20	0.70	2.00	2.86	2.70
CAL	0.00	0.00	0.00	?	0.00
MUO	12.55	10.15	1.70	0.17	11.85
GTR	4.75	3.30	3.35	1.02	6.65
VTX	2.45	2.70	1.00	0.37	3.70
JET	0.00	0.00	0.00	?	0.00
EMID	11.85	12.25	2.50	0.20	14.75
MUID	2.85	2.95	2.50	0.85	5.45
TAUID	0.00	0.00	0.00	?	0.00
BID	3.05	4.20	1.00	0.24	5.20
JES	0.00	0.00	0.00	?	0.00
L3	5.20	6.30	10.15	1.61	16.45
Total	**51.90**	**50.75**	**38.25**		**89.00**

H. Melanson, Sept. 14, 2001

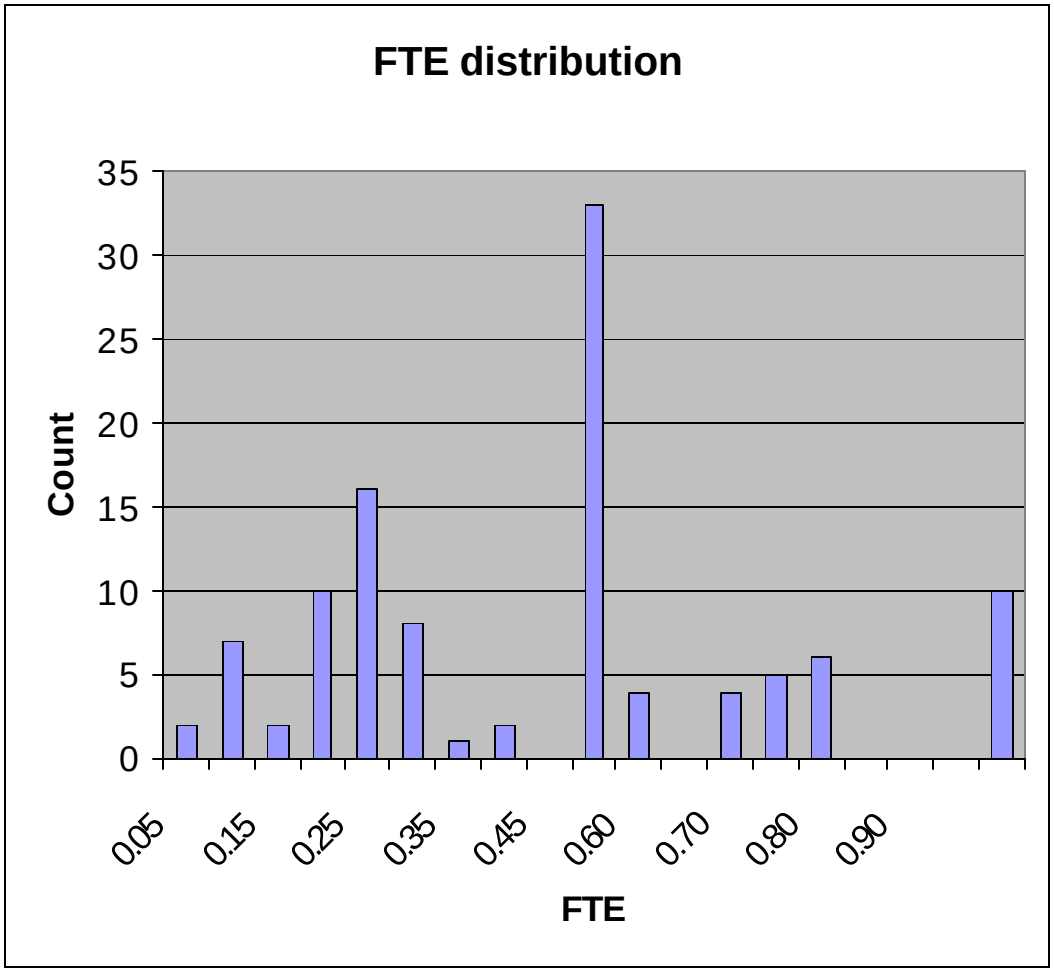


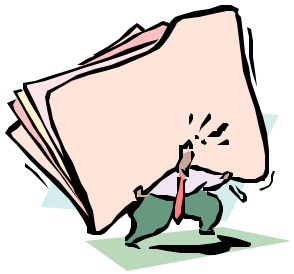
Algorithms Group Effort Reporting

Number of people: 110

Avg. FTE: 0.4

St dev. FTE: 0.3





Algorithms Group Effort Reporting

My personal conclusions from this "data analysis":

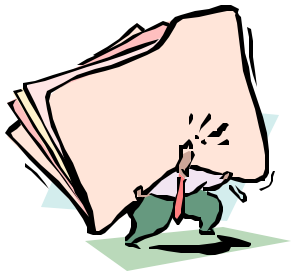
- If we assume the average FTE available is 0.4 for "everyone", we need

38 FTE / 0.4 = **95 NEW** people

- If the currently involved people could be relieved of some of their other duties (46 people with < 0.35 FTE), we could cover some of the outstanding tasks, but I doubt that this would yield much (5 – 10 FTE's???)

- If we don't have enough people, we must prioritize our tasks, and *make sure* we accomplish the most important ones. This will mean not doing everything we want, and may require sacrifices.

H. Melanson, Sept. 14, 2001



Algorithms Group Effort Reporting

My primary recommendations:

- Urge those in the collaboration not already involved to get (deeply) involved.

- Focus current efforts on the highest priority tasks. Relieve people of "extra tasks", if any exist. Encourage increasing the existing level of effort towards contributions to "algorithms".

- Set algorithm development goals based on the highest priority issues, and require all groups to reassign manpower as required to meet those goals.

Summary of effort reports:

http://www-d0.fnal.gov/computing/algorithms/reports/august01.html

p11 RECO Goals
(still in Draft form)

http://www-d0.fnal.gov/computing/algorithms/docs/p11goals.html

• **CPU Usage** - RECO is ***required*** to keep up with the maximum data taking rate expected through March, 2002.

• **Memory Usage** - RECO is ***required*** to process any raw data or Monte Carlo file in a physical memory space less than 490 Mb.

• **Output file size** - All groups should review and optimize the size of their persistent chunks.

Algo / ID groups should optimize algorithms, look for bugs, etc.

If any of these goals is not met with the existing algorithms, some functionality in RECO *will be lost*.

p11 RECO Goals
(still in Draft form)

• **Alignment** - All detector groups (*SMT*, *CFT*, *CPS*, *CAL*, *MUO*) ***must*** insure that DØRECO is using valid alignment constants when processing real data. These constants ***must*** be verified with data. Internal and detector-to-detector alignment is required. The Alignment group is charged with coordinating these studies and should validate all such constants.

In order to achieve this goal, tools like tracking and vertexing must be available and understood. In addition, basic understanding of the detector performance is essential.

• **Calibration** - All detector groups (*SMT*, *CFT*, *CPS*, *CAL*, *MUO*) ***must*** insure that all calibration constants available from online calibration procedures are migrated to the offline database, and are used appropriately by DØRECO. The use of these calibration constants ***must*** satisfy the CPU and memory usage requirements listed above. The Calibration group is charged with coordinating these efforts.

p11 RECO Goals
(still in Draft form)

- **Global Tracking** - The global tracking group ***must***

  - optimize the quality of tracks found with real data. This includes SMT and CFT stand alone tracks, as well as global tracks. Tracking over the entire available tracking volume should be studied.

  - work on improving the tracking efficiency for particles in jets.

- **Vertexing** - The vertexing group ***must***

  - optimize primary and secondary vertexing algorithms that run on real data.

  - optimize strategies for reconstrucing K0 particles in real data, and should establish well understood K0 candidate samples.

p11 RECO Goals
(still in Draft form)

- **EM ID** - The EM ID group ***must***
  - optimize electron id algorithms for finding W and Z bosons in the real data, and should establish well understood W and Z candidate samples.
  - develop photon id algorithms to be used on the real data, in conjunction with the needs of the Jet Energy Scale group.

- **MU ID** - The MU ID group ***must***
  - optimize muon id algorithms for finding W and Z bosons in the real data, and should establish well understood W and Z candidate samples.
  - optimize strategies for reconstrucing J/psi particles in real data, and should establish well understood J/psi candidate samples.

- **TAU ID** - The TAU ID group ***must***
  - optimize tau id algorithms for finding W and Z bosons in the real data, and should establish well understood W and Z candidate samples.

p11 RECO Goals (still in Draft form)

- **JET ID** - The JET ID group ***must***
  - optimize jet id algorithms to run on real data.
  - respond to the needs of the JES group.

- **JES** - The JES group ***must***
  - develop the first version of the jet energy scale.

- **BC ID** - The BC ID group ***must***
  - optimize the ability to measure signed impact parameters using real data, and should establish evidence of b jets by showing an excess on the positive side of the distribution.
  - optimize the ability to measure a pt_rel distribution using real data, and should establish evidence of b jets by showing a well understood large pt_rel signal.

- **Thumbnail** - The first complete version of the thumbnail ***must*** be completed.

p11 RECO Goals (still in Draft form)

- Some of these goals might be achievable by just using the current reconstruction program.
- However, I doubt it.
- We definitely need more “users” – I agree more people should “look at the data”.
- But to achieve these goals, I believe we also need “developers” – look at the data, find the problems, and then **do** something about them.
- The Algorithm / ID groups desperately need workers.

http://www-d0.fnal.gov/computing/algorithms/

Summary

- p10.00.00 ("t01.56.00") is running on the farms. It is the first production release officially processing raw data, and was released on schedule.
- An update fixing some major problems is expected next week.
- Complete certification of p10 is expected Sept. 21.
- A schedule and a set of goals for the next release are available.
- The algorithms / ID groups estimate they need an additional 38 FTE's to accomplish all of their tasks for this release.
- We are in desperate need of new people and refocused efforts.
- There are an incredible number of crucial problems to solve.

Now is the time.