US ATLAS Analysis Support & the Facilities

Jim Cochran (ISU) & Rik Yoshida (ANL)

Outline

What are users doing now – most common analysis model

PAT Survey results

Distributed Analysis Job Efficiencies

PAT Plans: Analysis Model Target

DPD Train

DPD Train vs simple Skimming

Other Focus

New Efforts to Improve Performance

Other Issues with Potential Impact on Facilities

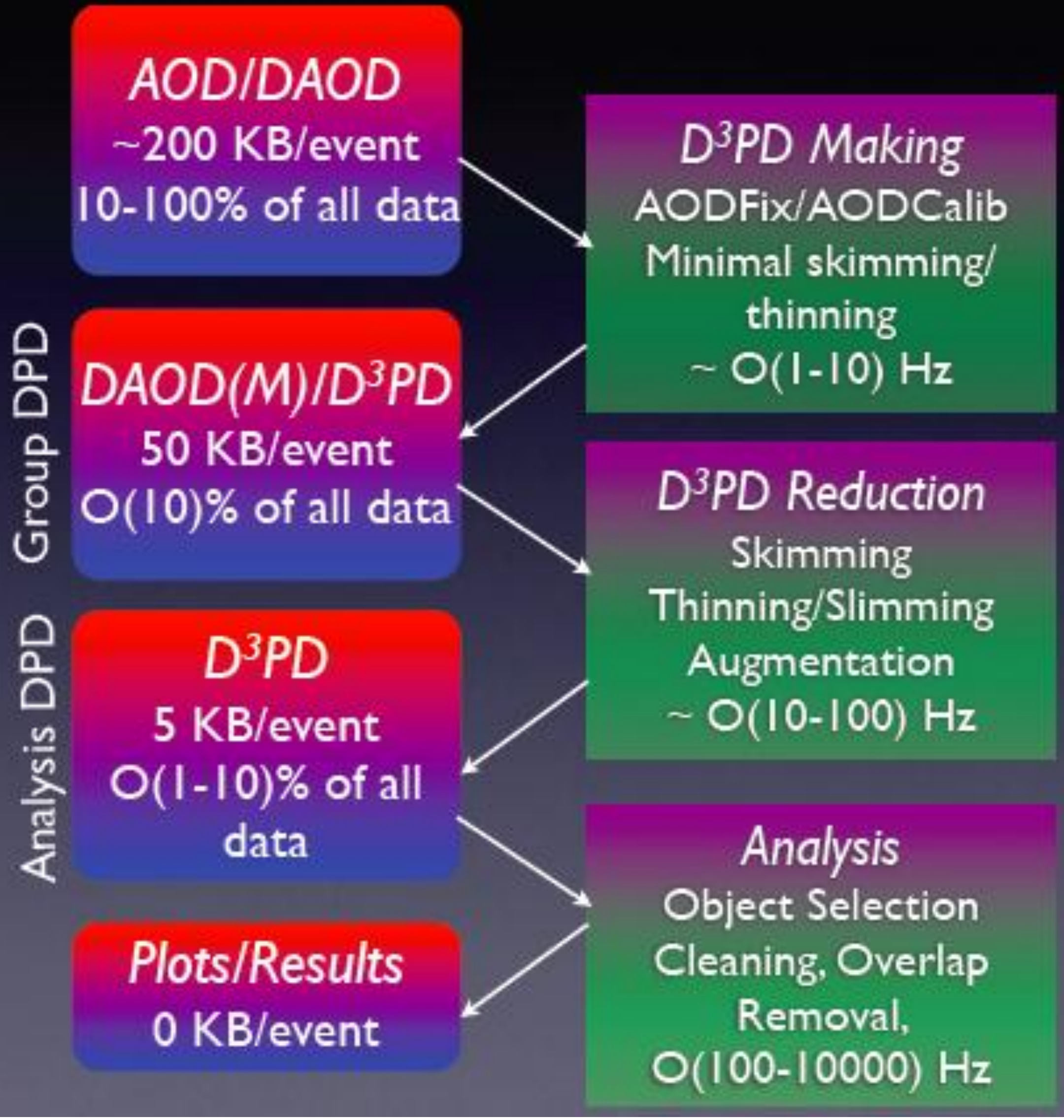
Outlook

Most Common Analysis Model

Data

Task

Organization



Group DPD

Analysis DPD

AOD/DAOD
~200 KB/event
10-100% of all data

D^{3}PD Making
AODFix/AODCalib
Minimal skimming/thinning
~ O(1-10) Hz

DAOD(M)/D^{3}PD
50 KB/event
O(10)% of all data

D^{3}PD Reduction
Skimming
Thinning/Slimming
Augmentation
~ O(10-100) Hz

D^{3}PD
5 KB/event
O(1-10)% of all data

Analysis
Object Selection
Cleaning, Overlap Removal,
O(100-10000) Hz

Plots/Results
0 KB/event

Physics Groups...

- ~10 in ATLAS
- Most use Production System

Physics Sub-Groups...

- Nominally ~5-10 per Physics Group
- Many use the GRID

Individual

- Nominally up to ~10 per Physics Sub-group
- Usually on Tier 3 or desktop/laptop

PAT Survey Results

Prior to the PAT Workshop, PAT conducted an extensive survey of active analyzers

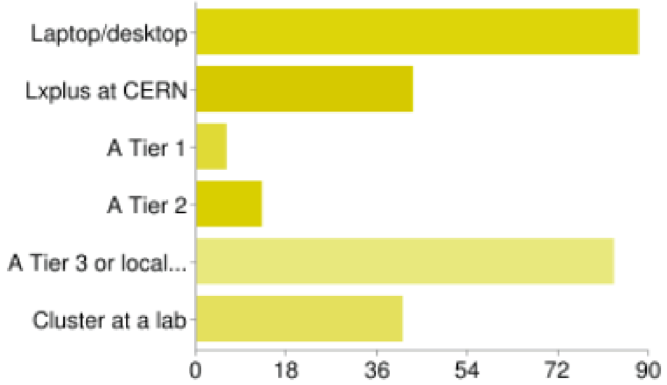
There were 76 questions

They received 274 responses!

What follows is a summary, focused on topics of potential interest to facilties

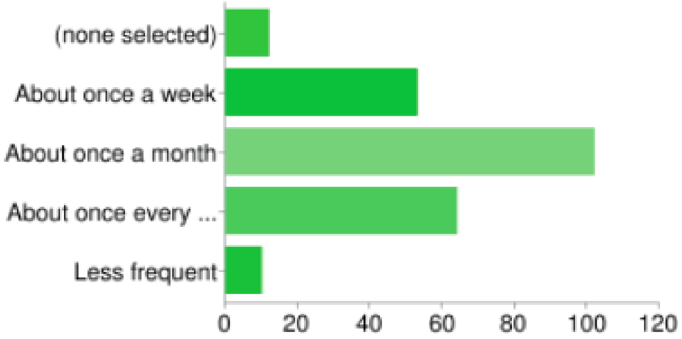
Full survey results are here:
https://indico.cern.ch/getFile.py/access?contribId=21&sessionId=5&resId=0&materialId=0&confId=149202

4. What is your primary computing resources for data analysis (i.e. where you login)?



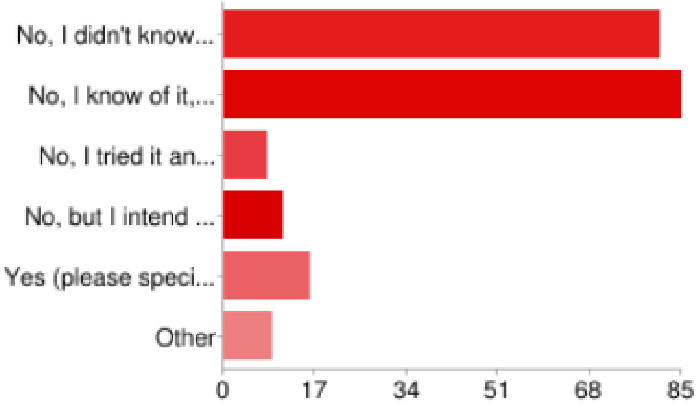
Laptop/desktop	**88**	32%
Lxplus at CERN	**43**	16%
A Tier 1	**6**	2%
A Tier 2	**13**	5%
A Tier 3 or local cluster	**83**	30%
Cluster at a lab	**41**	15%

17. How often do you update your analysis to include new data?



(none selected)	**12**	4%
About once a week	**53**	19%
About once a month	**102**	37%
About once every few months	**64**	23%
Less frequent	**10**	4%

34. If you use D3PDs for physics analysis, have you used the provided D3PDReader (https://twiki.cern.ch/twiki/bin/view/AtlasProtected/D3PDMakerReader)?

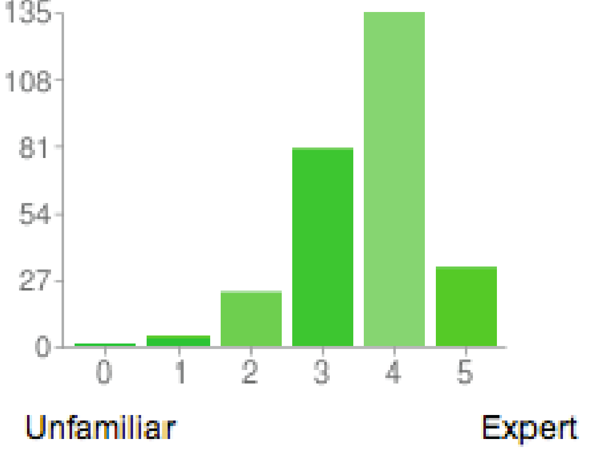


No, I didn't know this exists	**81**	40%
No, I know of it, but haven't tried it yet	**85**	42%
No, I tried it and didn't like it (please specify why not in the "Other" box)	**8**	4%
No, but I intend to do so	**11**	5%
Yes (please specify in the "Other" box, if you have comments)	**16**	8%
Other	**9**	4%

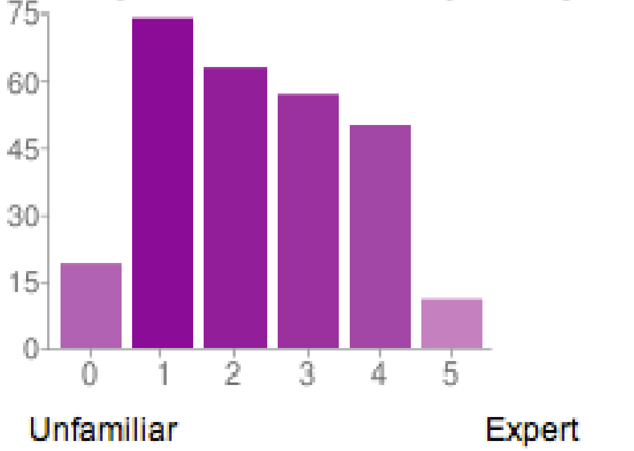
People may select more than one checkbox, so percentages may add up to more than 100%.

General Software Skills

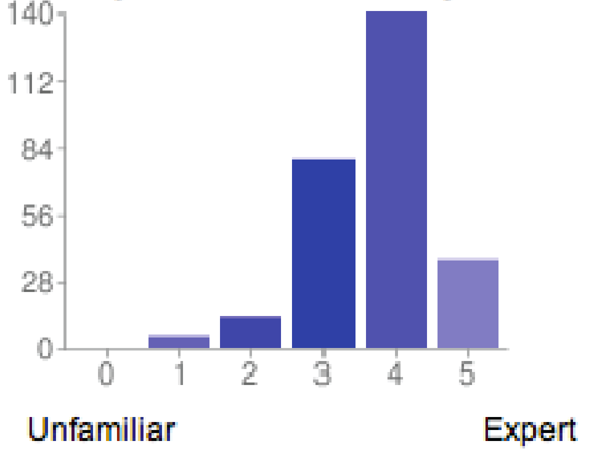
5. Rate your level of familiarity with C++



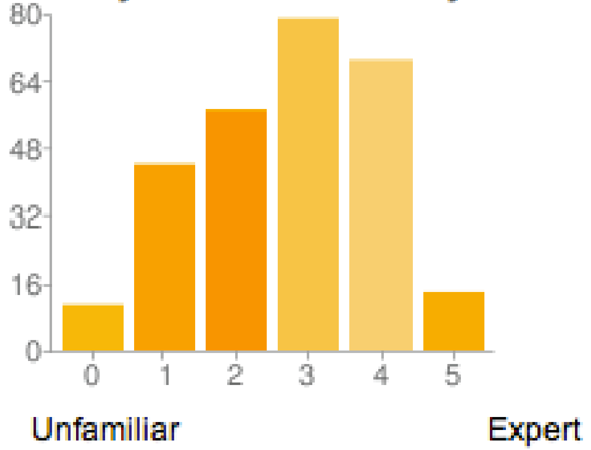
6. Rate your level of familiarity with Pytho



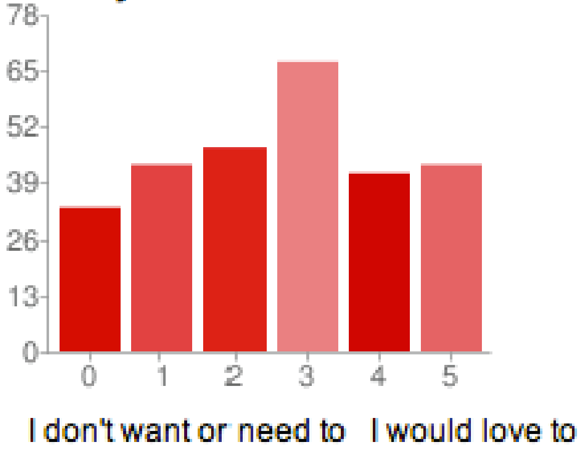
7. Rate your level of familiarity with ROO



8. Rate your level of familiarity with Athena

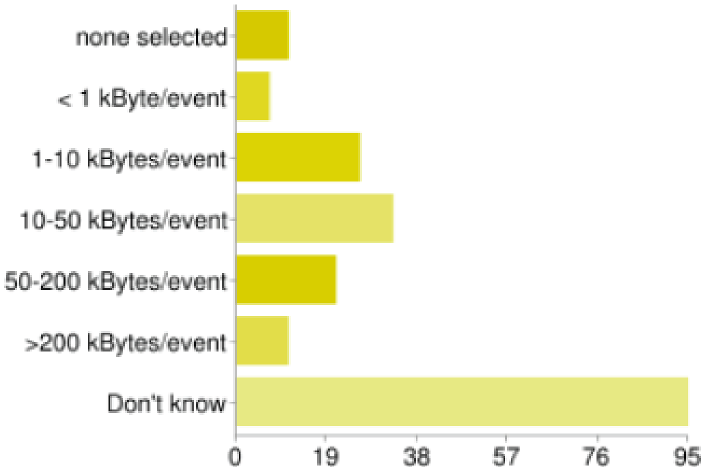


9. Rate your desire to learn more about Athena



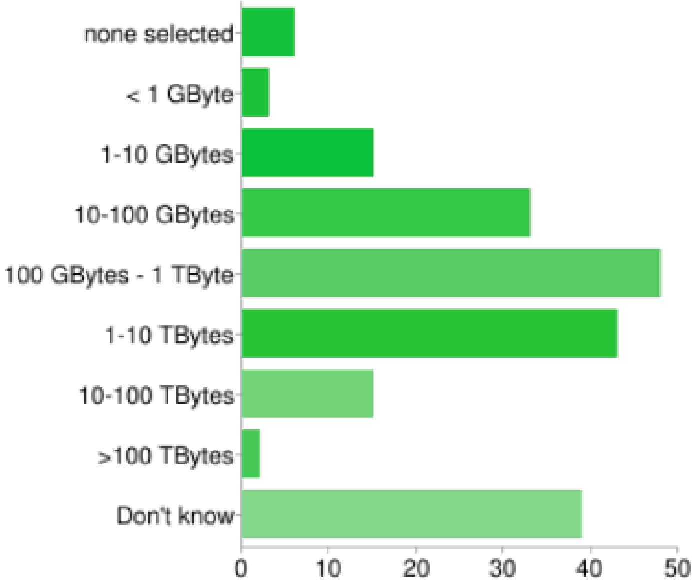
- Quite a few people expressed their wish to learn python (and also C++)
- Need for teaching software and software design in ATLAS

29. For the data D3PD/ntuple files you most frequently use for your physics analysis, what is their approximate per-event size?



none selected	**11**	4%
< 1 kByte/event	**7**	3%
1-10 kBytes/event	**26**	9%
10-50 kBytes/event	**33**	12%
50-200 kBytes/event	**21**	8%
>200 kBytes/event	**11**	4%
Don't know	**95**	35%

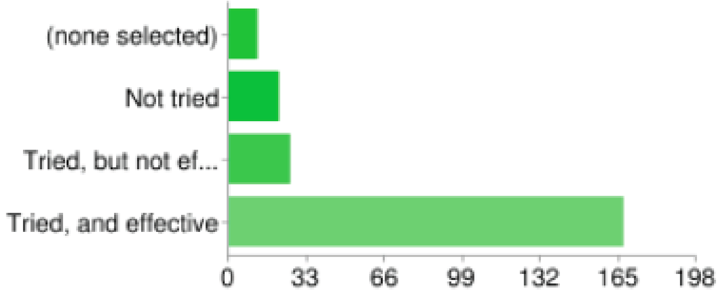
30. For the D3PD/ntuple files you most frequently use for your physics analysis, what is their approximate total size (including all data and MC that you regularly process)?



none selected	**6**	2%
< 1 GByte	**3**	1%
1-10 GBytes	**15**	5%
10-100 GBytes	**33**	12%
100 GBytes - 1 TByte	**48**	18%
1-10 TBytes	**43**	16%
10-100 TBytes	**15**	5%
>100 TBytes	**2**	1%
Don't know	**39**	14%

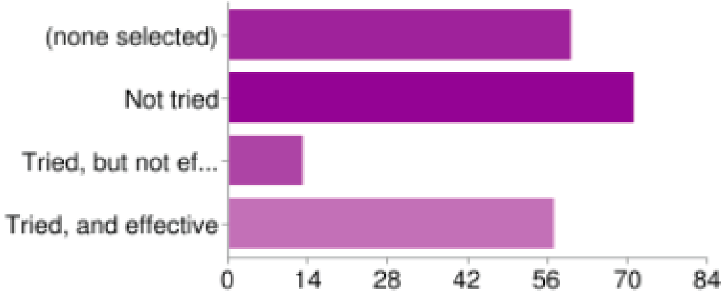
User efforts at speeding up analysis

16. Please select any of the following methods for speeding up analysis that you have tried and your conclusion: - Compiling Macros



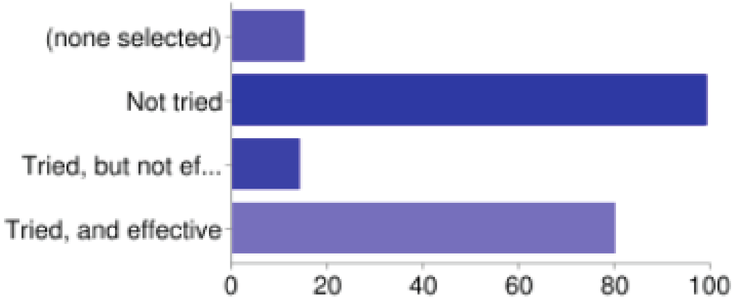
(none selected)	**12**	4%
Not tried	**21**	8%
Tried, but not effective	**26**	9%
Tried, and effective	**167**	61%

16. Please select any of the following methods for speeding up analysis that you have tried and your conclusion: - Moving from Python to C++



(none selected)	**60**	22%
Not tried	**71**	26%
Tried, but not effective	**13**	5%
Tried, and effective	**57**	21%

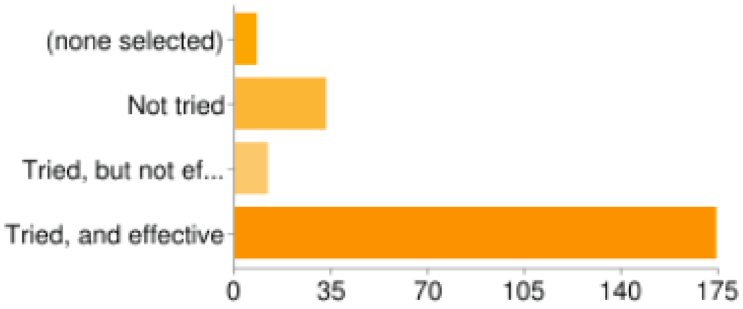
16. Please select any of the following methods for speeding up analysis that you have tried and your conclusion: - Profiling analysis code and optimizing



(none selected)	**15**	5%
Not tried	**99**	36%
Tried, but not effective	**14**	5%
Tried, and effective	**80**	29%

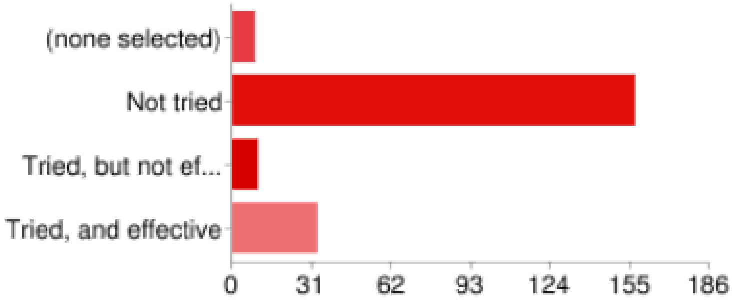
User efforts at speeding up analysis

16. Please select any of the following methods for speeding up analysis that you have tried and your conclusion: - Activating select branches of TTrees



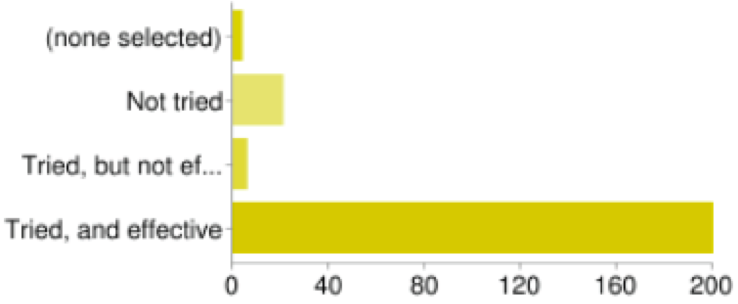
(none selected)	**8**	3%
Not tried	**33**	12%
Tried, but not effective	**12**	4%
Tried, and effective	**174**	64%

16. Please select any of the following methods for speeding up analysis that you have tried and your conclusion: - Optimizing ROOT i/o (eg using TTreeCache or optimizing split level)



(none selected)	**9**	3%
Not tried	**157**	57%
Tried, but not effective	**10**	4%
Tried, and effective	**33**	12%

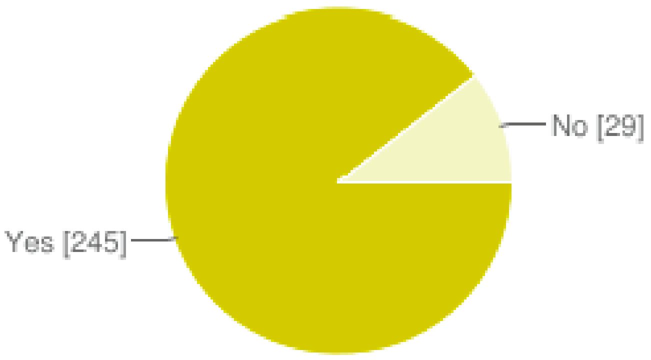
16. Please select any of the following methods for speeding up analysis that you have tried and your conclusion: - Creating intermediate (typically smaller) ntuples



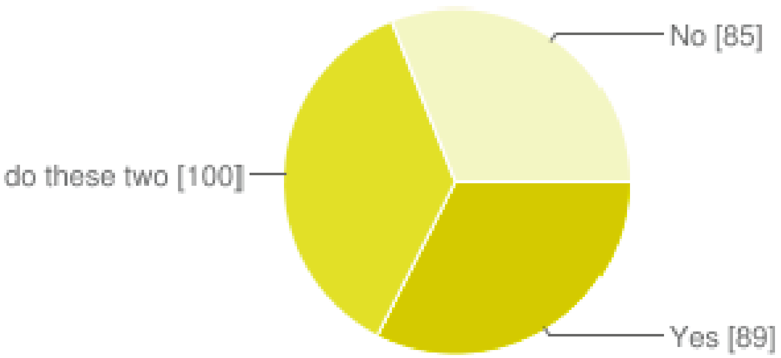
(none selected)	**4**	1%
Not tried	**21**	8%
Tried, but not effective	**6**	2%
Tried, and effective	**200**	73%

Do you do physics and/or performance studies

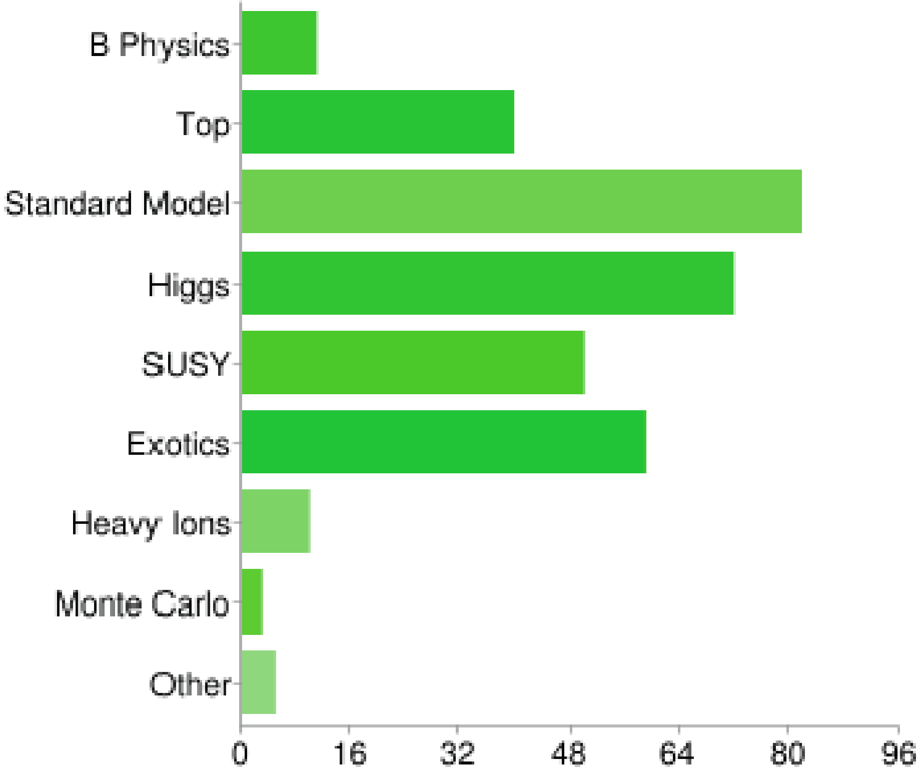
11. Do you do physics analysis ?



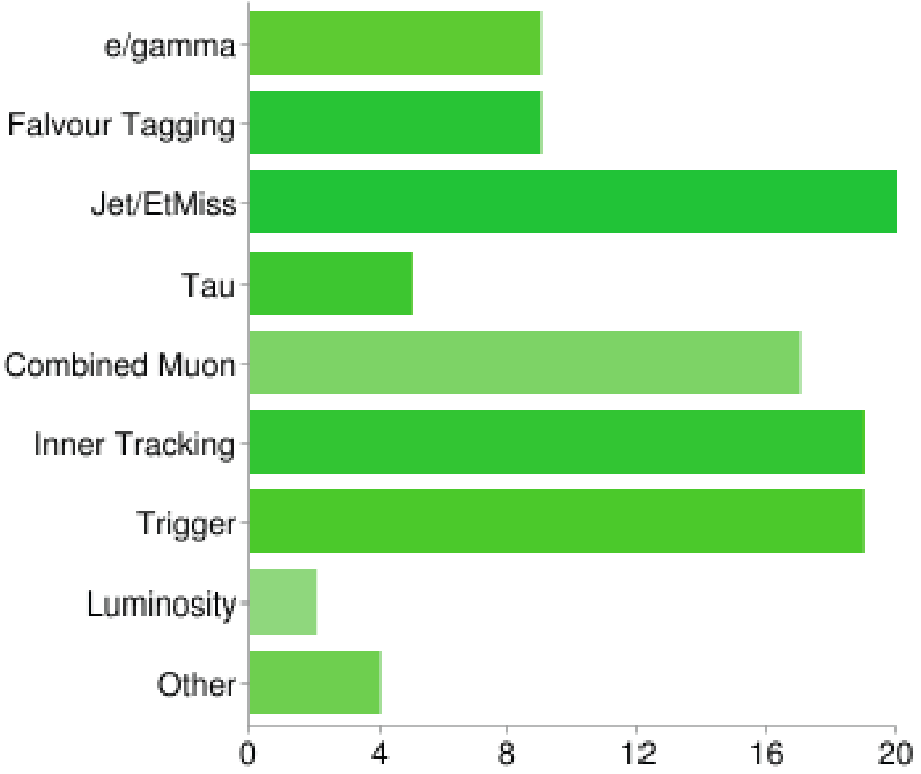
36. Do you do performance studies?



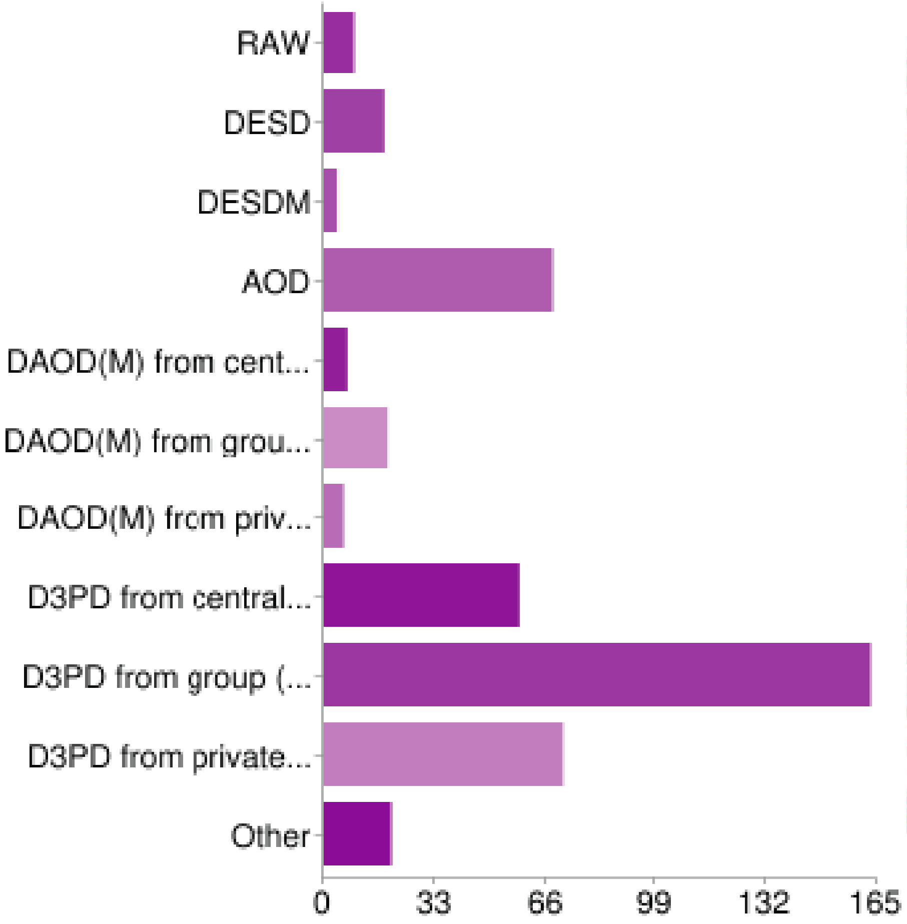
12. In which PHYSICS group do you work mostly (multiple ansv

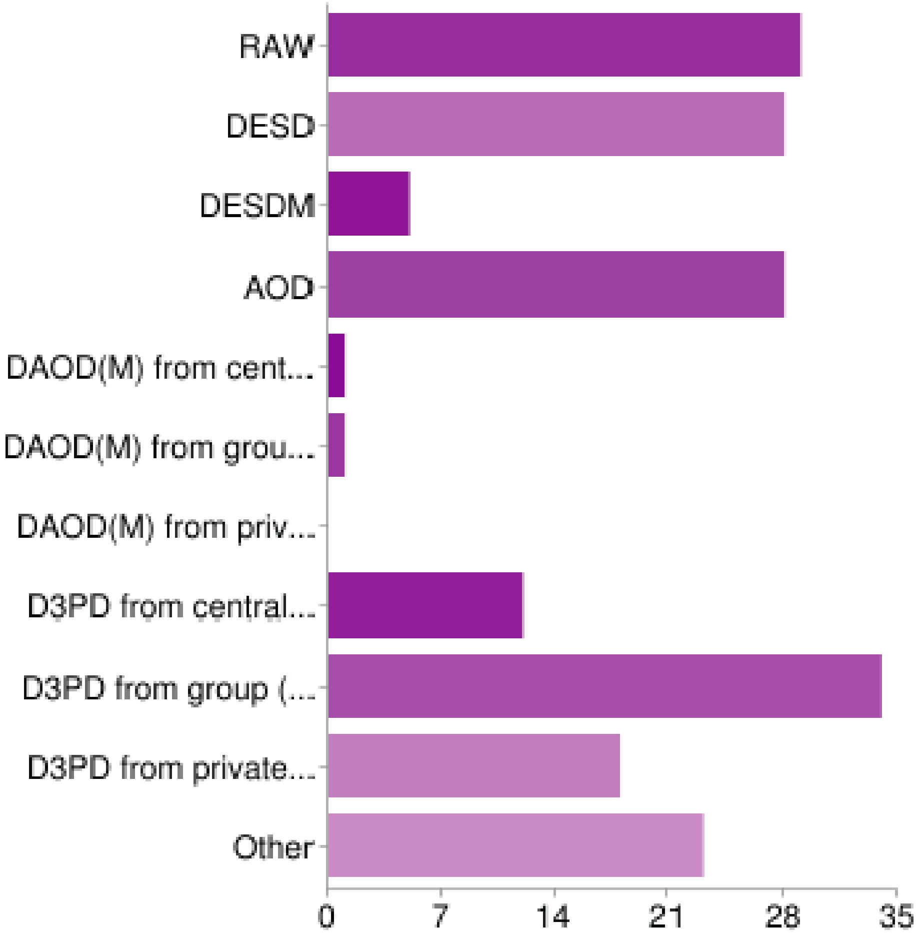


37. In which PERFORMANCE group do you work mostly (multip



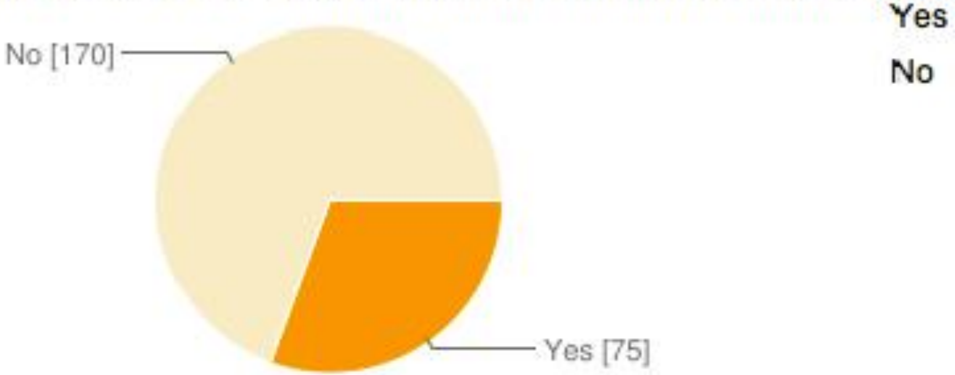
Data Formats


13. What data formats do you use for your physics analysis (

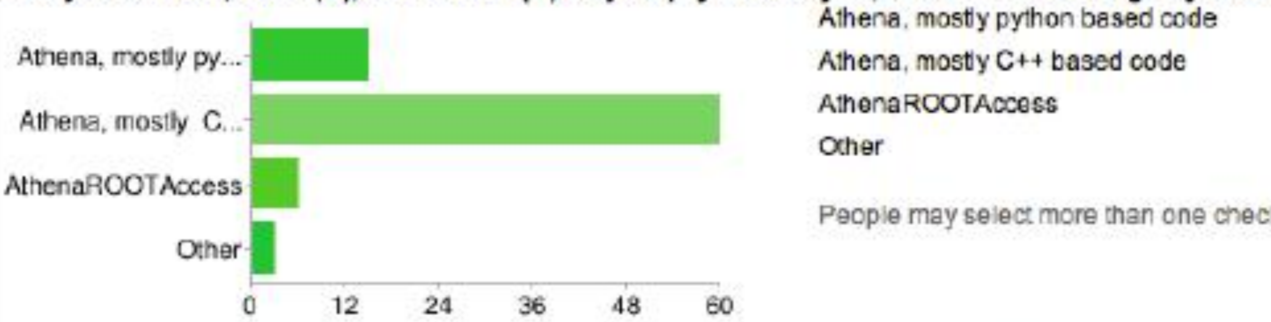

38. What data formats do you use for your performance studies

File usage for Physics Analysis (POOL)

20. Do you YOURSELF use at some point in your analysis chain use pool files (AOD, DESD(M), DAOD(M))?

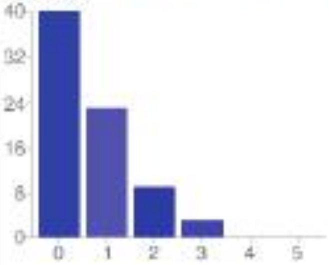


23. If you use AOD, DAOD(M), and/or DESD(M) for your physics analysis, which of the following do you use to analyze them (multiple answers possible)?



People may select more than one checkbox, so percentages may add up to more than 100%.

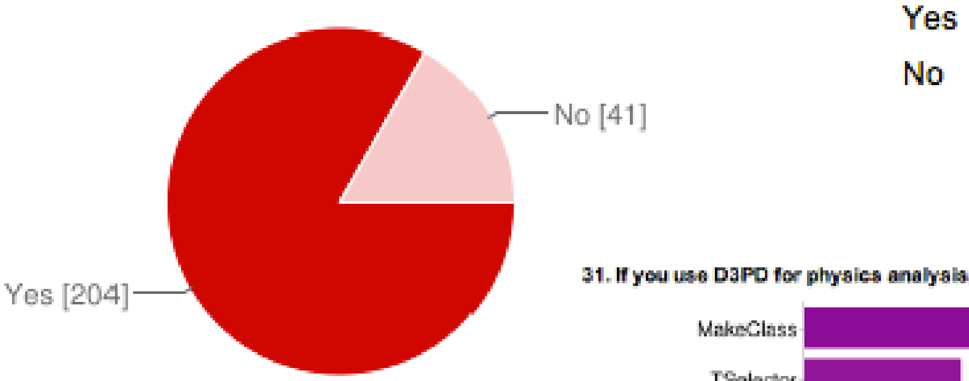
25. How many intermediate pool files do you typically create before making final plots OR dumping a D3PD or ntuple? I you analyze AODs (or centrally produced DESDs) directly, please select 0.



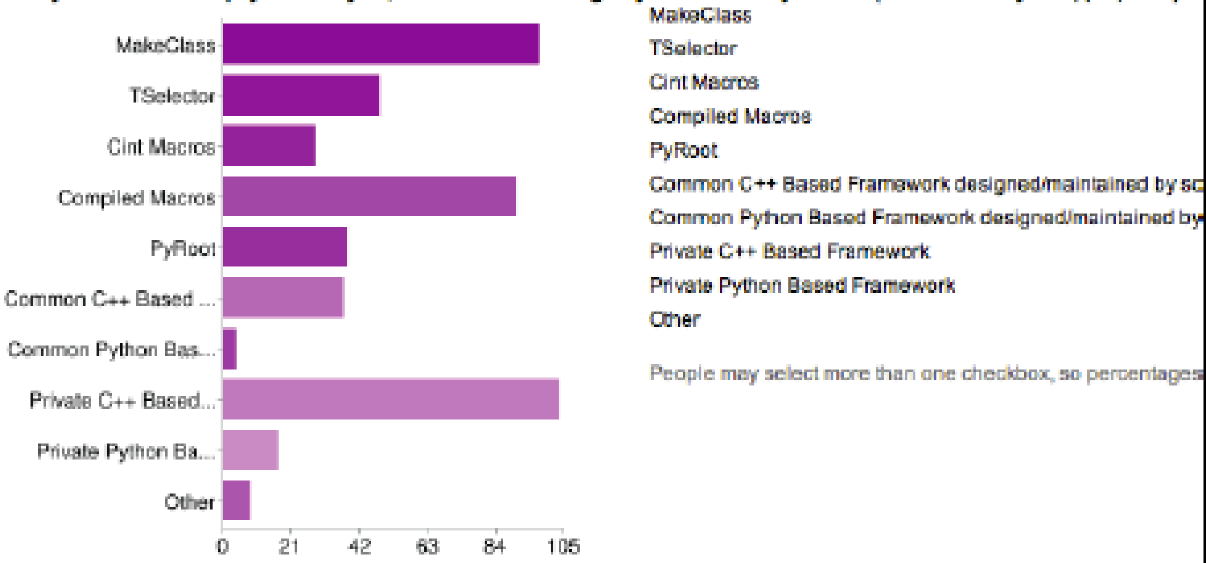
0	40	15%
1	23	8%
2	9	3%
3	3	1%
4	0	0%
5	0	0%

File usage for Physics Analysis (D3PD)

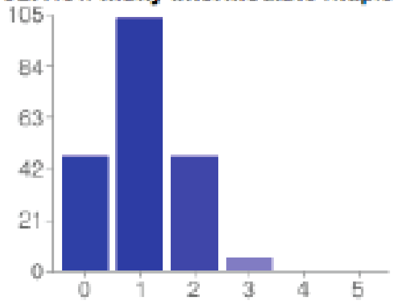
28. Do you YOURSELF use at some point in your analysis chain use D3PDs?



31. If you use D3PD for physics analysis, which of the following do you use to analyze them (select as many as appropriate).



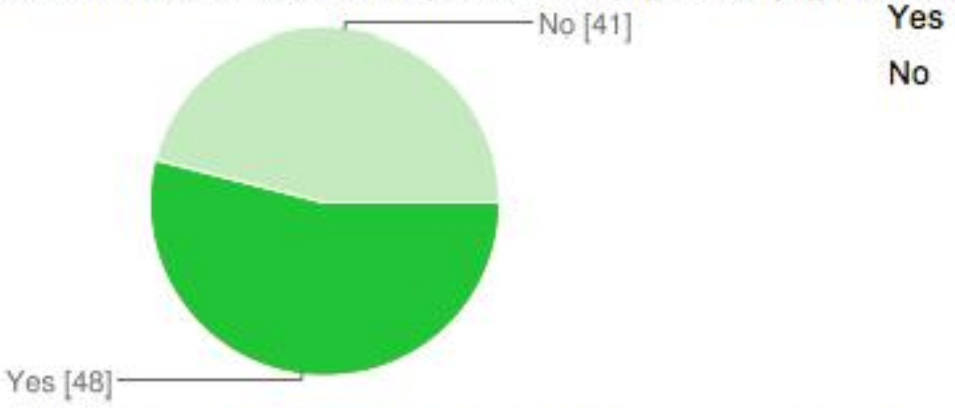
32. How many intermediate ntuple sets do you typically create before making final plots and extracting final numbers? If you analyze D3PDs directly, please select 0.



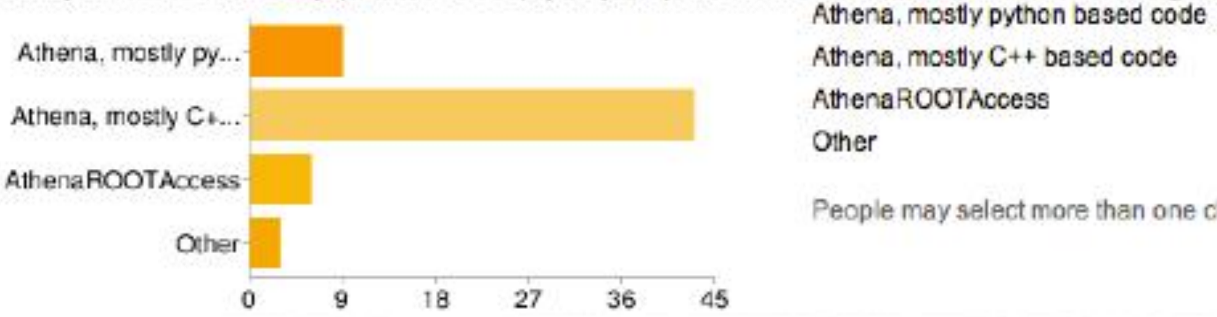
0	47
1	103
2	47
3	5
4	0
5	0

File usage for Performance Studies (POOL)

42. Do you YOURSELF use pool files (AOD, DESD(M), DAOD(M)) at some stage of your performance studies?

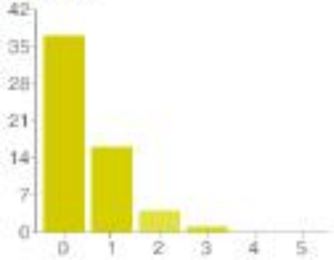


45. If you use AOD, DAOD(M), and/or DESD(M) for your performance studies, which of the following do you use to analyze them (multiple answers possible)?



People may select more than one checkbox, so percentages may add up to more than 100%.

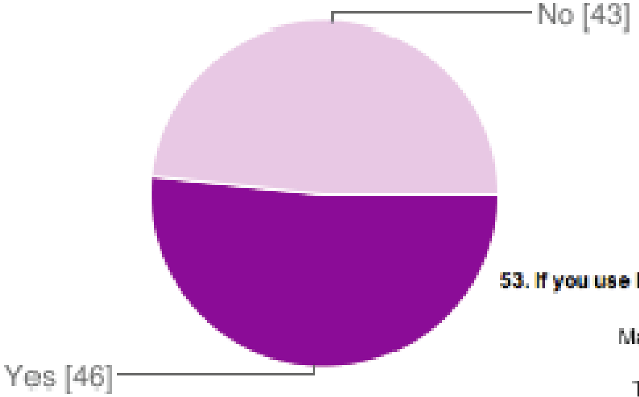
47. How many intermediate pool files do you typically create before making final plots OR dumping a D3PD or ntuple? I you analyze AODs (or centrally produced DESDs) directly, please select 0.



0	37	14%
1	16	6%
2	4	1%
3	1	0%
4	0	0%
5	0	0%

File usage for Performance Studies (D3PD)

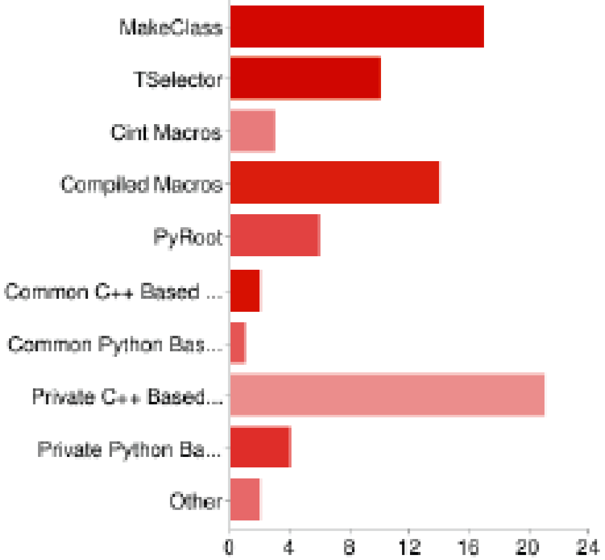
50. Do you YOURSELF use at some point in your performance study chain use D3PDs?



Yes

No

53. If you use D3PD for performance studies, which of the following do you use to analyze them (select as many as appropriate).



MakeClass

TSelector

Cint Macros

Compiled Macros

PyRoot

Common C++ Based Framework designed/maintained by some

Common Python Based Framework designed/maintained by so

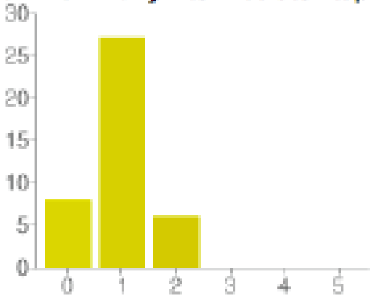
Private C++ Based Framework

Private Python Based Framework

Other

People may select more than one checkbox, so percentages ma

54. How many intermediate ntuple sets do you typically create before making final plots and extracting final numbers? If you analyze D3PDs directly, please select 0.



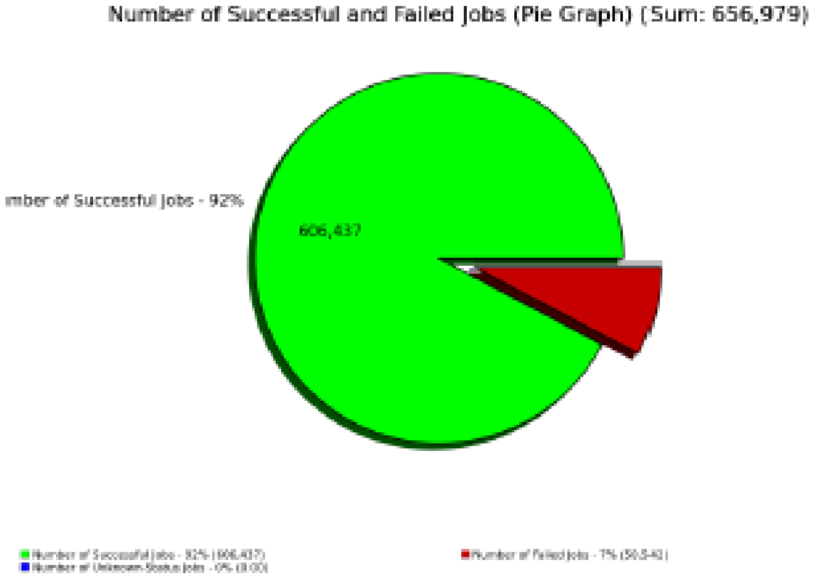
0	8
1	27
2	6
3	0
4	0
5	0

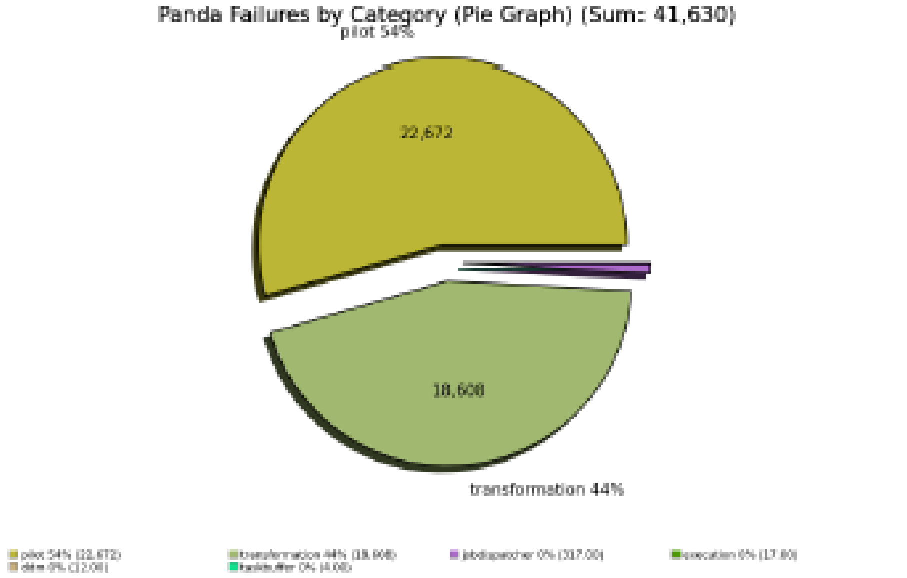
Distributed Analysis Job Efficiencies

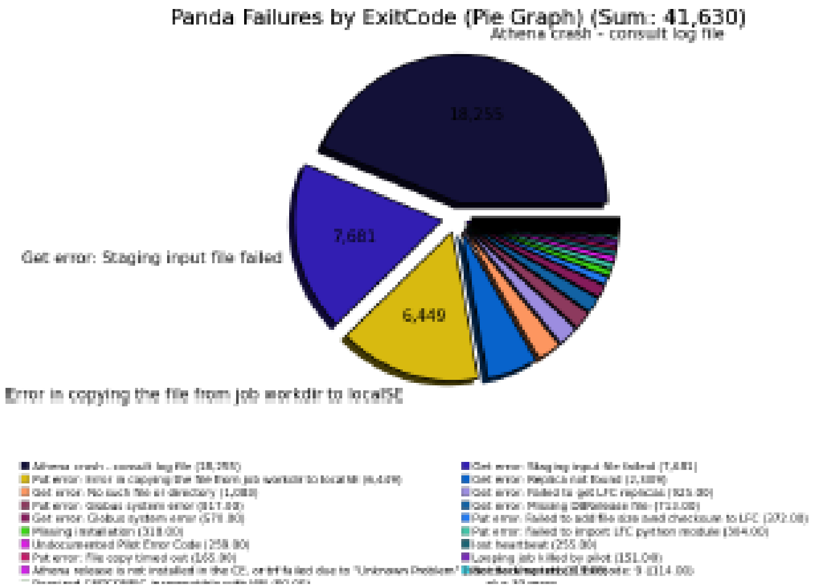
Johannes Elmsheuser and team looked at All Analysis and GangaRobot/HammerCloud jobs in Jun, Jul, Aug, & Sept of 2011

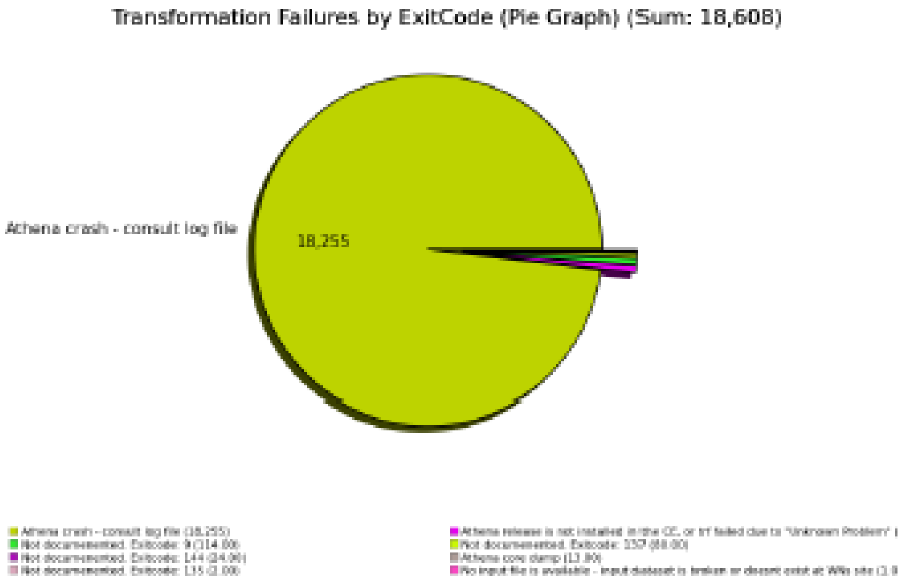
GANGAROBOT SEPTEMBER 2011

All T0-T3



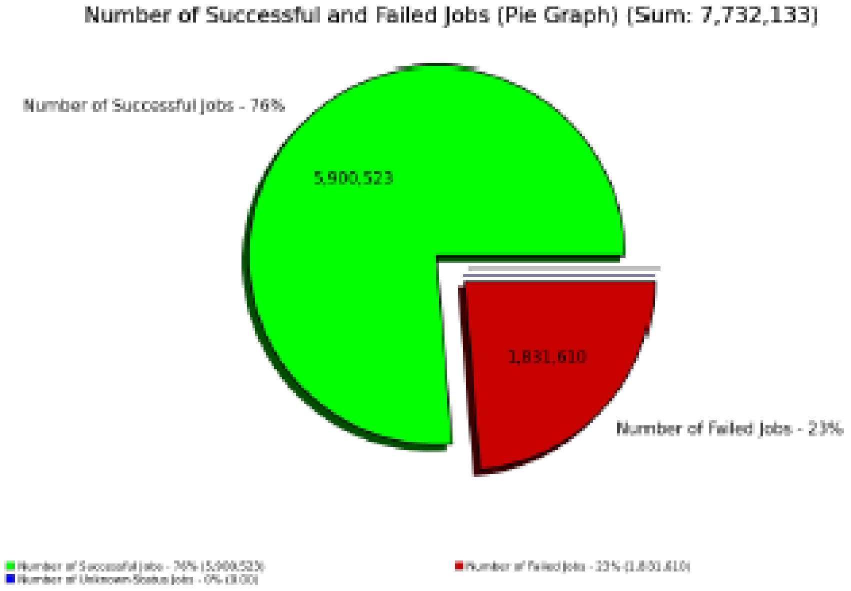




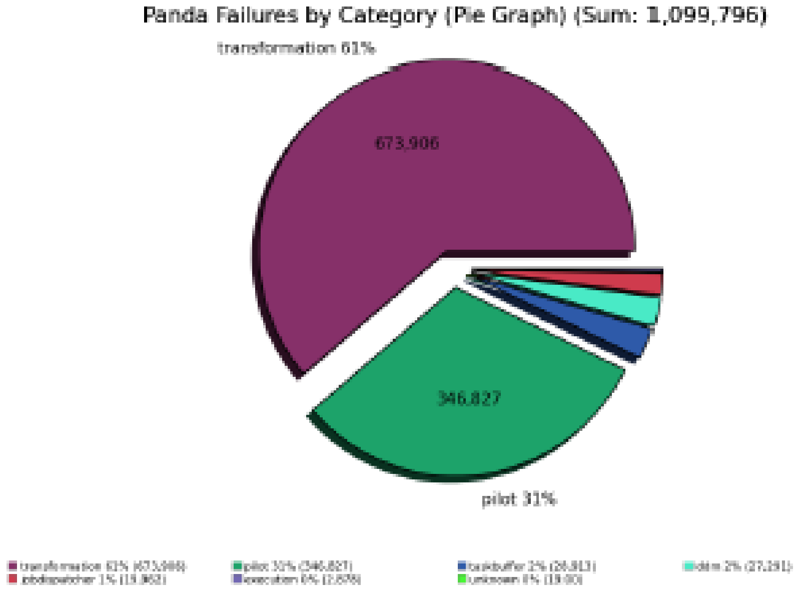


ALL ANALYSIS SEPTEMBER 2011

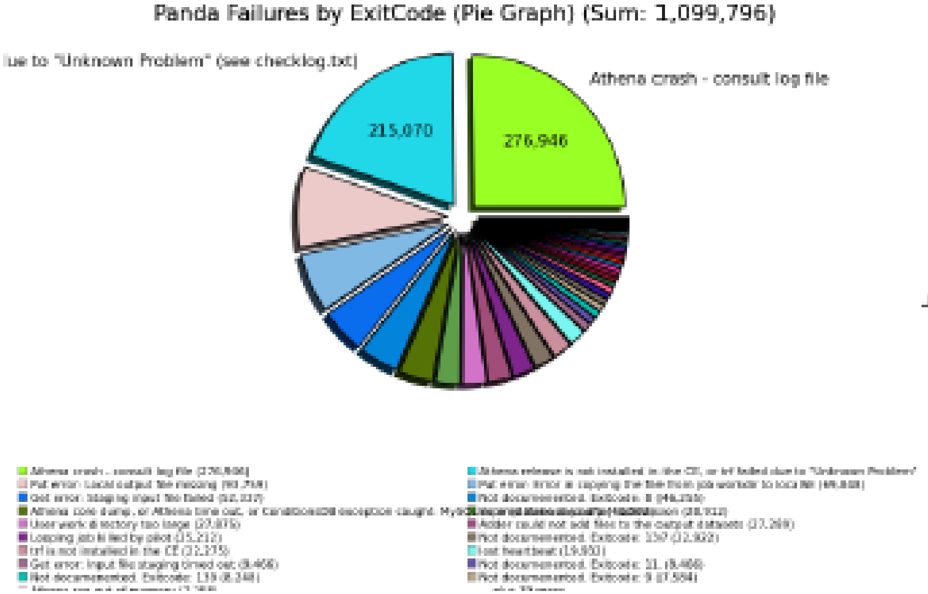
All T0-T3



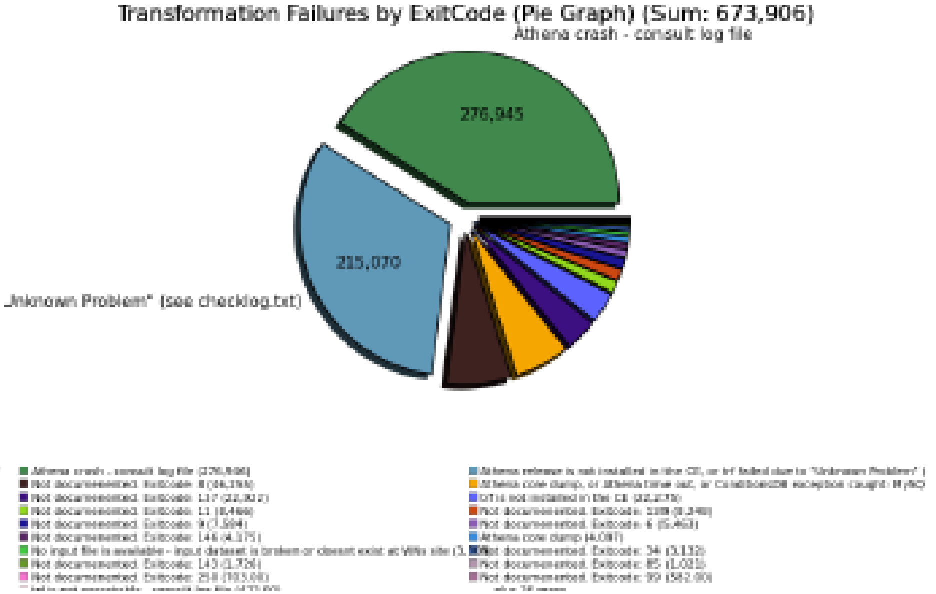
Number of Successful and Failed Jobs (Pie Graph) (Sum: 7,732,133)



Panda Failures by Category (Pie Graph) (Sum: 1,099,796)

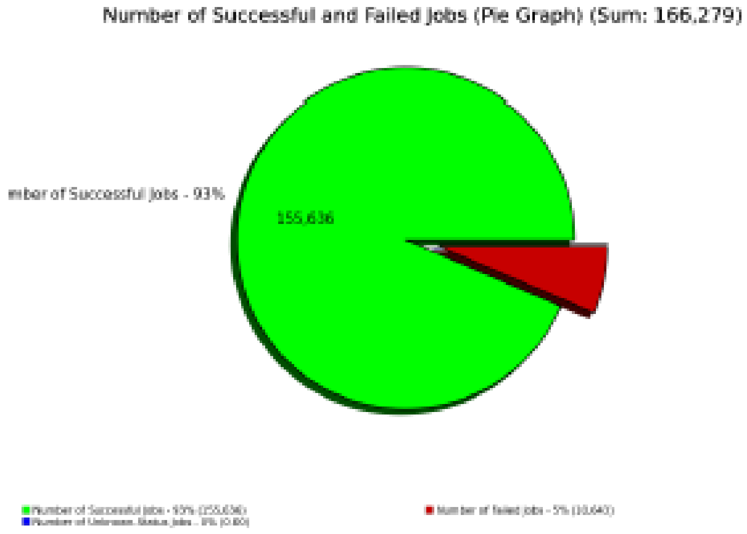


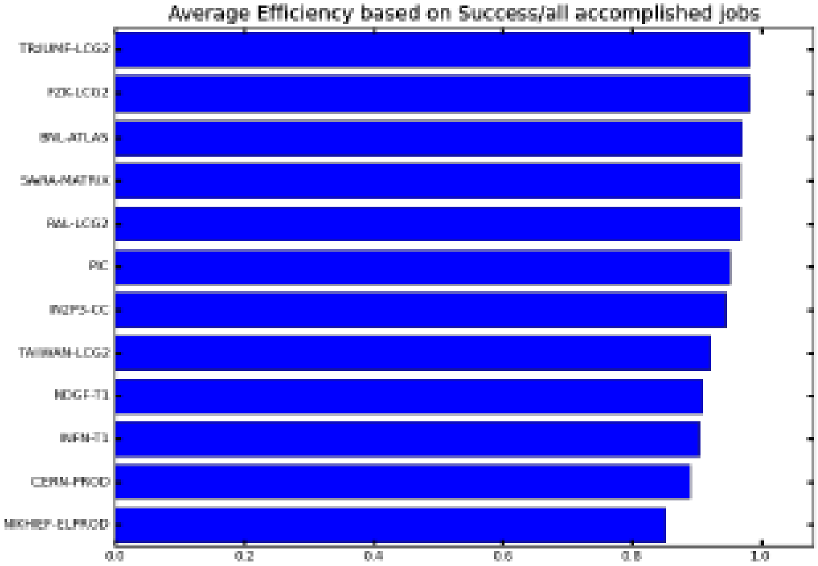
Panda Failures by ExitCode (Pie Graph) (Sum: 1,099,796)

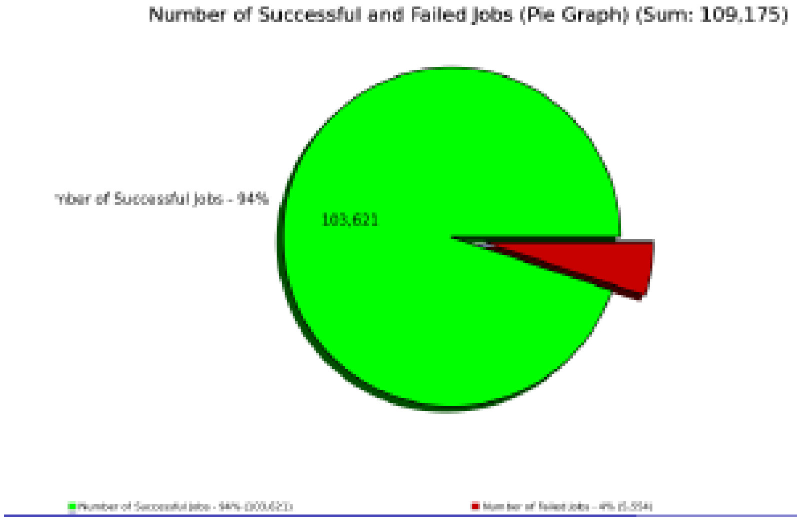


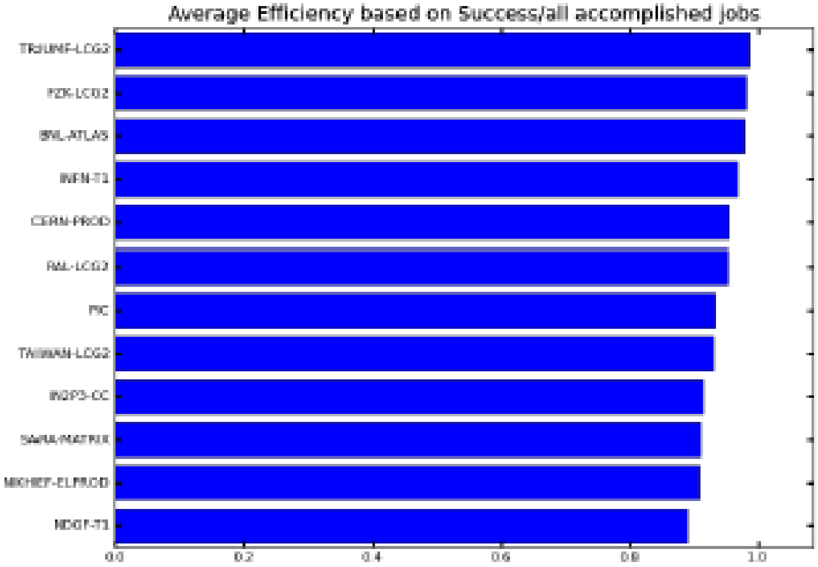
Transformation Failures by ExitCode (Pie Graph) (Sum: 673,906)

GANGAROBOT T0+T1 Aug, Sep 2011

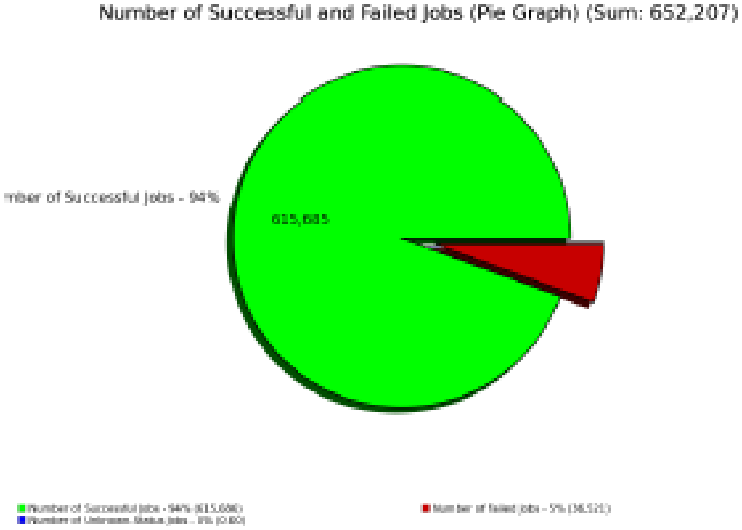


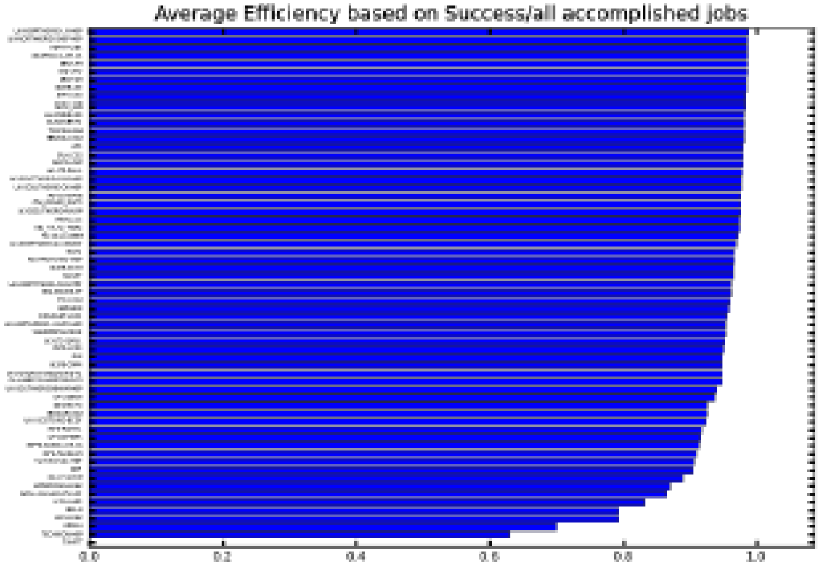


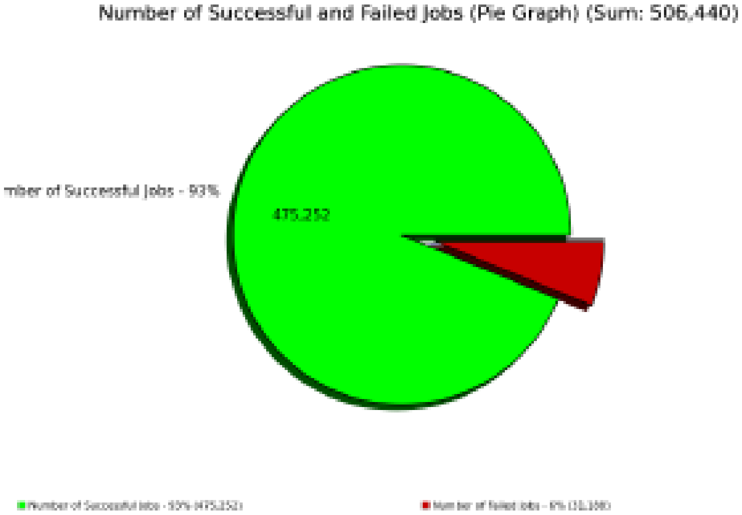


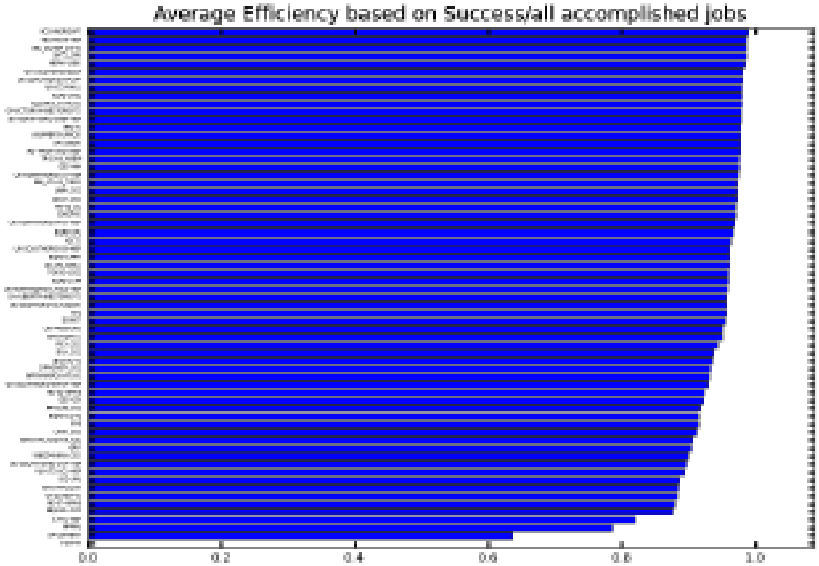


GangaRobot T2 Aug, Sep 2011

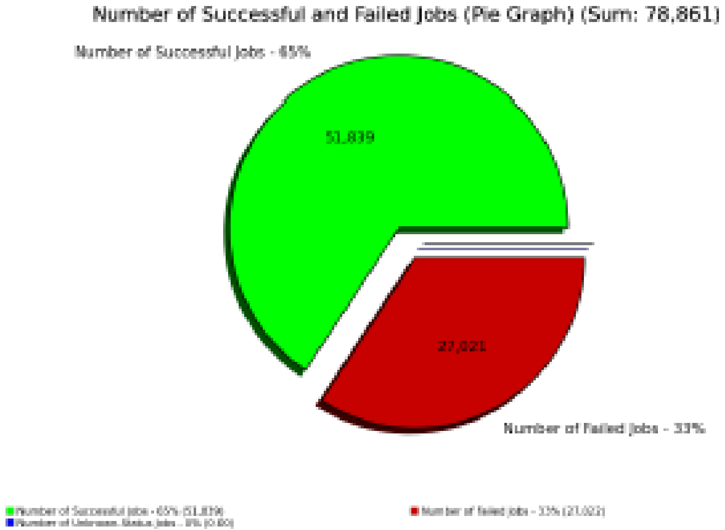


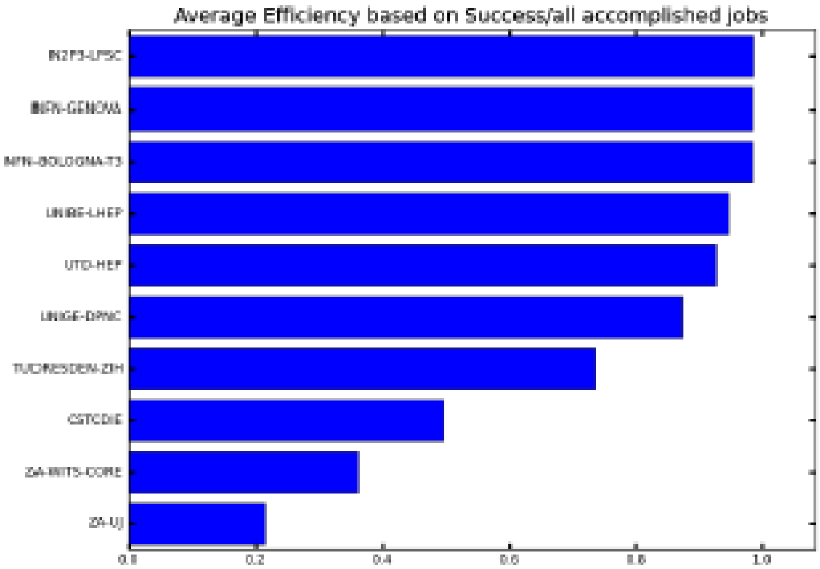


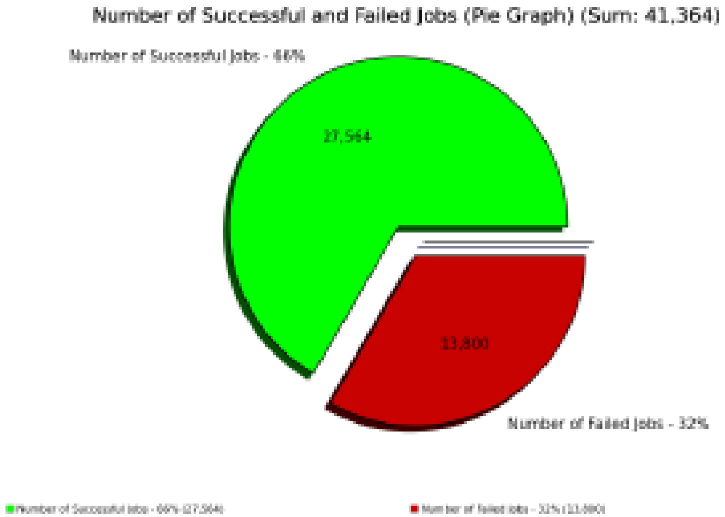


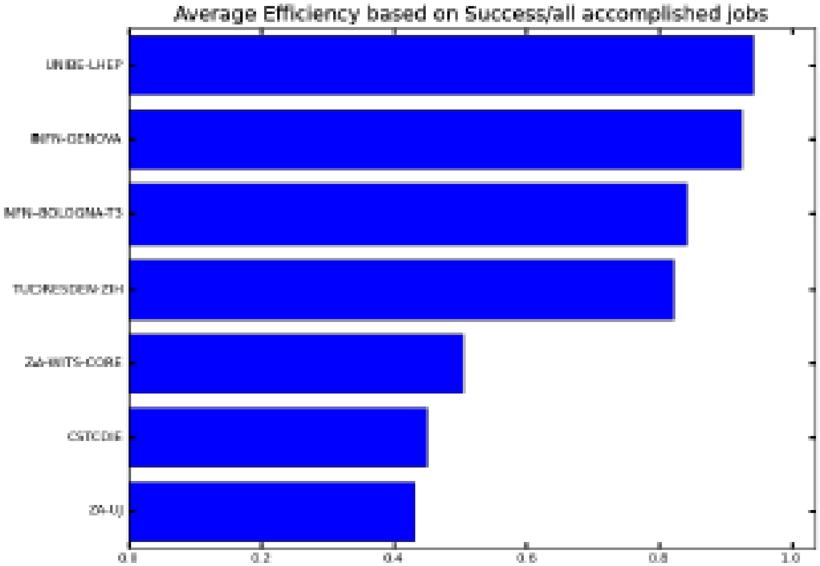


GANGAROBOT T3 Aug, Sep 2011



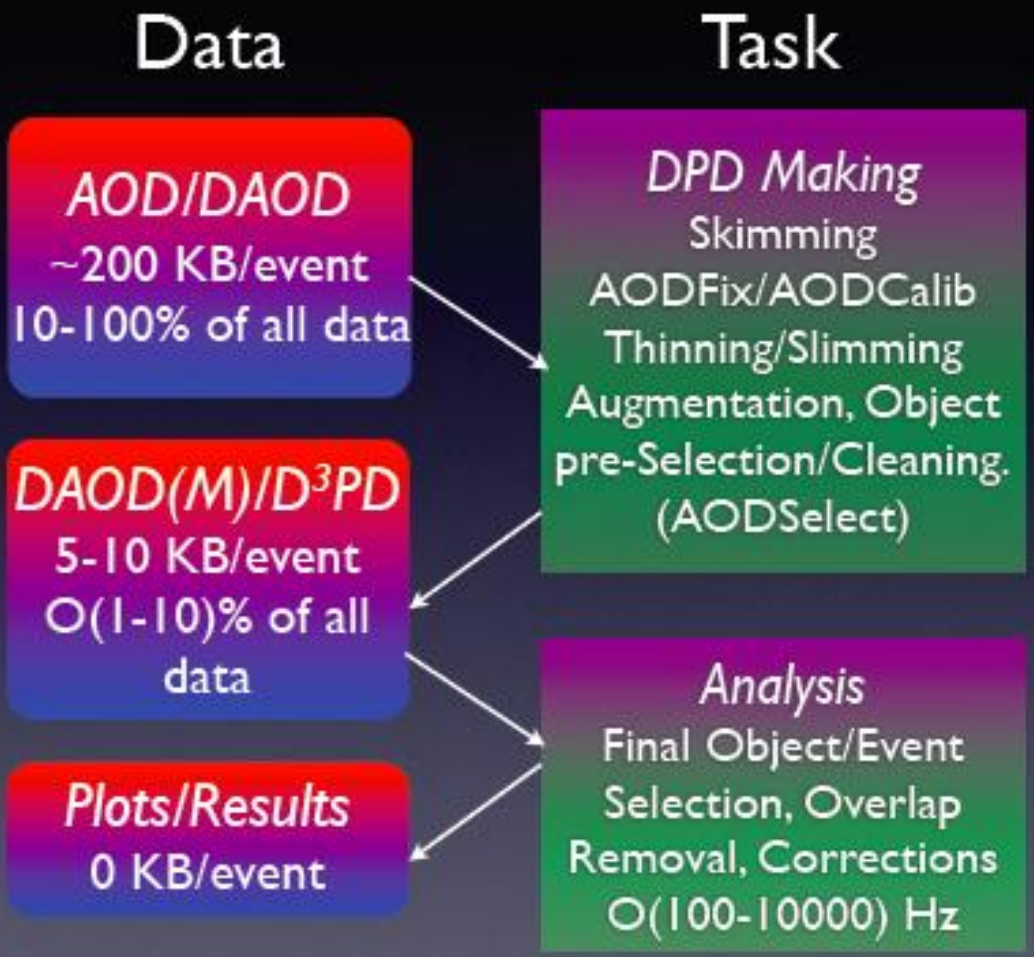


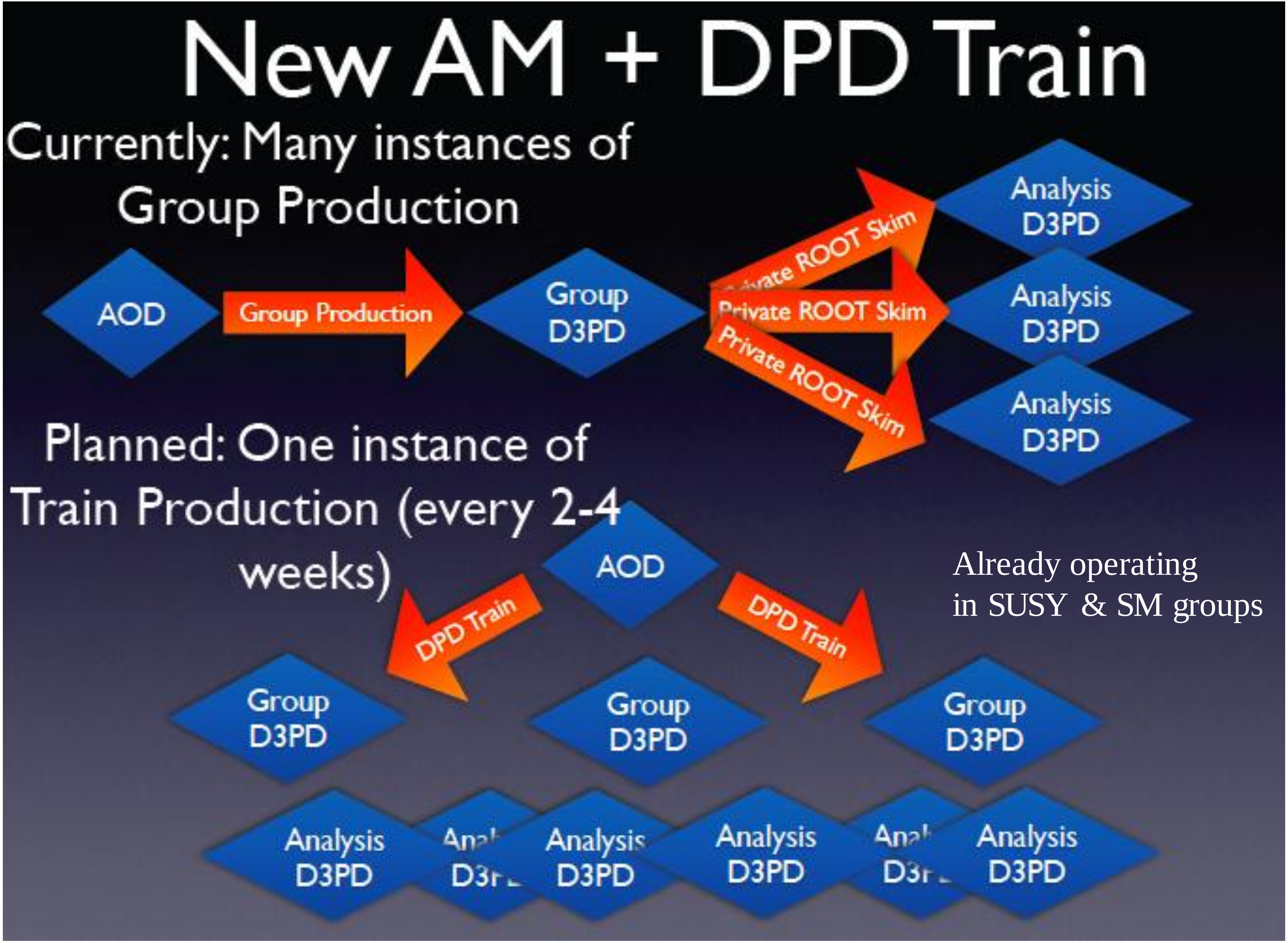




Analysis Model Target

- For most people, the large D3PDs are just an intermediate data-format
    - They are too big to transfer to local resources (10s of TB now, and growing w/ more data). So people skim/thin/slim them on GRID as first step.
- Proposal: Groups create 2 types of DPDs (Mostly D3PDs now, but D2PD also possible):
    - One/few Group DPD (R&D): Large size, meant for detailed studies (eg performance)
        - Follow the analysis model on previous step
    - Many Analysis DPD (Factory): Small size, each finely tuned to needs of one or few specific analyses
        - A goal can be that a graduate student can trivially reproduce recently published result
    - Use the DPD Train to simultaneously build all DPDs every 2-4 weeks
        - Better use of computing and people resources



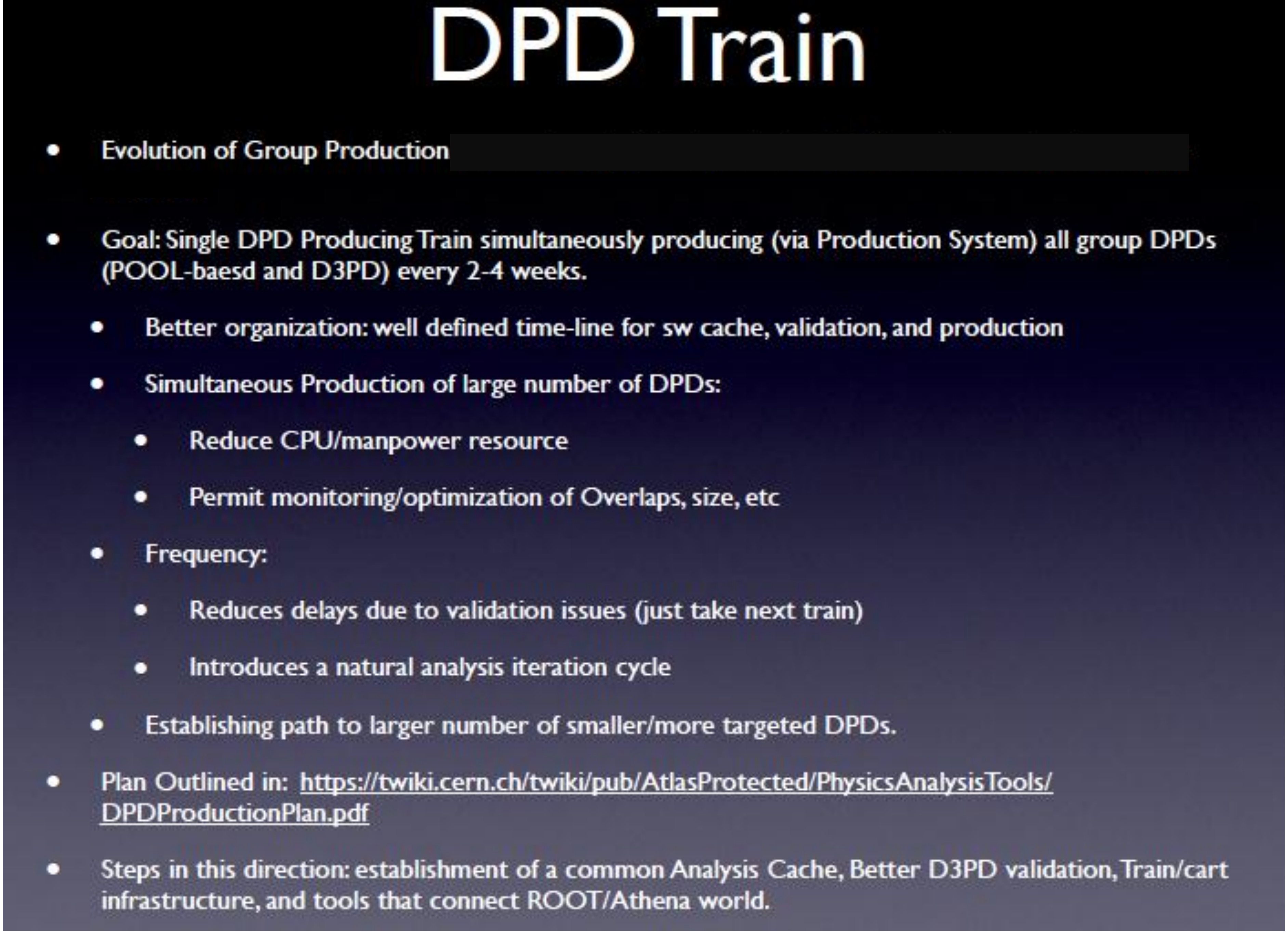


New AM + DPD Train

Currently: Many instances of Group Production

Planned: One instance of Train Production (every 2-4 weeks)

Already operating
in SUSY & SM groups



DPD Train

- Evolution of Group Production [redacted]
- Goal: Single DPD Producing Train simultaneously producing (via Production System) all group DPDs (POOL-baesd and D3PD) every 2-4 weeks.
  - Better organization: well defined time-line for sw cache, validation, and production
  - Simultaneous Production of large number of DPDs:
    - Reduce CPU/manpower resource
    - Permit monitoring/optimization of Overlaps, size, etc
  - Frequency:
    - Reduces delays due to validation issues (just take next train)
    - Introduces a natural analysis iteration cycle
  - Establishing path to larger number of smaller/more targeted DPDs.
- Plan Outlined in: https://twiki.cern.ch/twiki/pub/AtlasProtected/PhysicsAnalysisTools/DPDProductionPlan.pdf
- Steps in this direction: establishment of a common Analysis Cache, Better D3PD validation, Train/cart infrastructure, and tools that connect ROOT/Athena world.

DPD Train vs Skimming

Skimming: user submits grid jobs that reduce a large group n-tuple into a small D3PD

Pros & Cons of DPD Train (as compared to skimming):

- DPD train allows the user to employ all athena features	Pro
- DPD Train should save resources overall (depending on freq)	Pro
- DPD Train requires users to learn Athena	Pro & Con
- Skimming allows user to use their own event selection code DPD Train requires selection to be rewritten using the athena EDM	Con
- DPD Train has longer turn-around time (validation)	Con
- skim requires user to manage his own jobs on the grid for DPD train, production system handles jobs	Pro
- if all users switch to DPD train, would save disk space (eliminating the group D3PDs [size comparable to AODs])	Pro
- DPD train requires more care, one broken piece can derail (thus need centralized approach with coordinator, tests, etc.)	Con

Other (PAT) Focus

Major effort to support/develop a standalone ROOT/D3PD analysis framework but with appropriate links into Athena

See also
Jack's talk

New Efforts to Improve Performance

- There have been previous attempts to compare various analysis approaches:

(1) Akira did a rather exhaustive comparison of the (many) available approaches (carefully documented in [unpublished ?] note by Akira)

(2) Sergey did very careful studies of jobs running at BNL & found much room for improvement (2-3 years ago)

- As we get more data, the stress on the system will certainly increase (dramatically)

- Efforts have begun to improve the framework underlying various user analysis tools
    - D3PDReader
    - Object Selectors for the Performance Tools
    - Event Selectors ? (decision deferred to future meeting)

- Other analysis tools improvements in progress or under discussion
    - central repository of compiled rootcore packages
    - ability to augment D3PDReader objects (method under discussion)
    - modification of D3PDReader so that it can run with python

New Efforts to Improve Performance (continued)

- Would like to build performance benchmarks right into D3PDReader objects

- Furthermore, would like to have a way (at least for grid jobs) to centrally collect these performance measurements
  (would give computing people opportunity to better understand what typical jobs

- Potential performance metrics:
  - number of bytes read
  - TTreePerfStats
  - Maybe also write out which branches were read and how often

- Other (potential) improvements:
  - many new features to SampleHandler (Meta data) [esp ability to use in batch/grid]
  - progress on the development of a PAT framework (probably based on Sframe)
  - a simpler, lighter skimming tool
  - tools for managing jobs and job sequences
  - tools for incorporating systematics and corrections

Other Issues with Potential Impact on Facilities

- Scheme for data distribution to T3s
    - direct output of grid jobs to T3 could decrease total data transfers (?)

- Tag issues
    - keeping tag db up to date can be problematic
    - difficulties reporting trying to use tags on grid
    - tag is more useful for POOL based analyses than D3PD based analyses

Outlook

Active efforts to improve efficiency of analysis and resource use (at all levels)

Built in benchmarking tools will be crucial to further improvements

Success will, however, depend on careful coordination between users, PAT, Physics/Performance Groups, and facilities