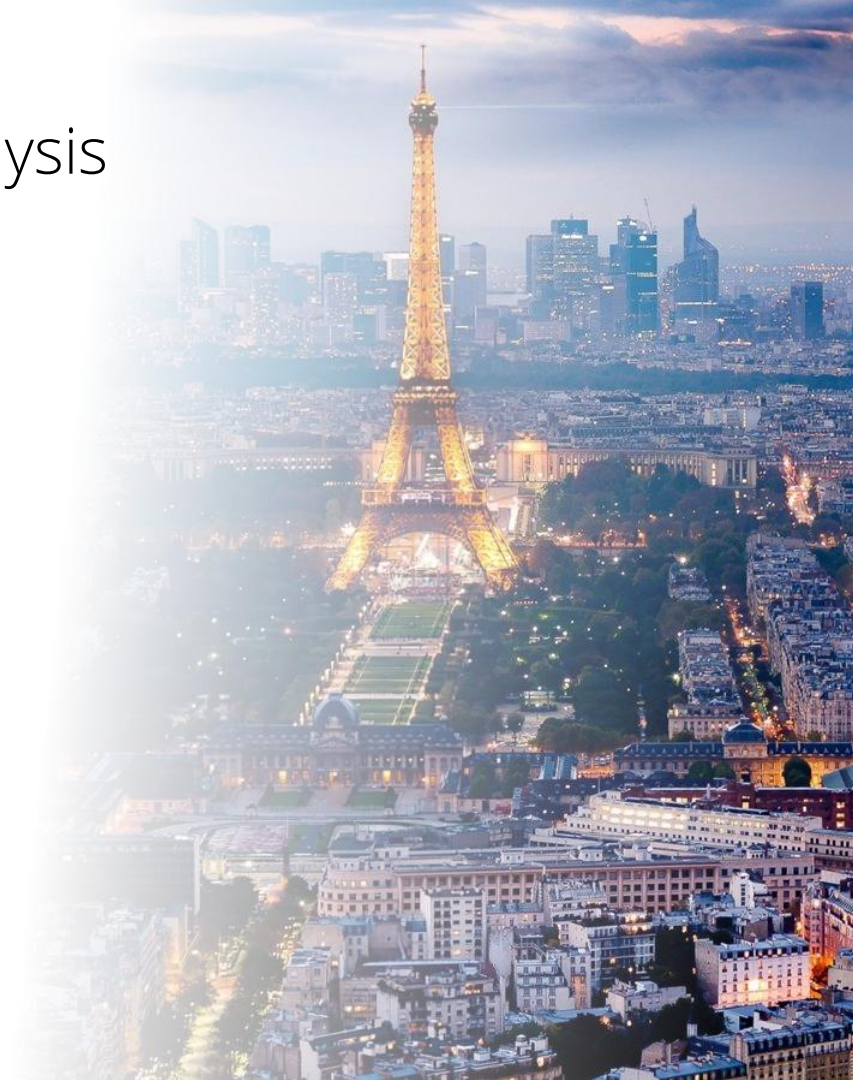
Prototype for a CTA Real-Time-Analysis

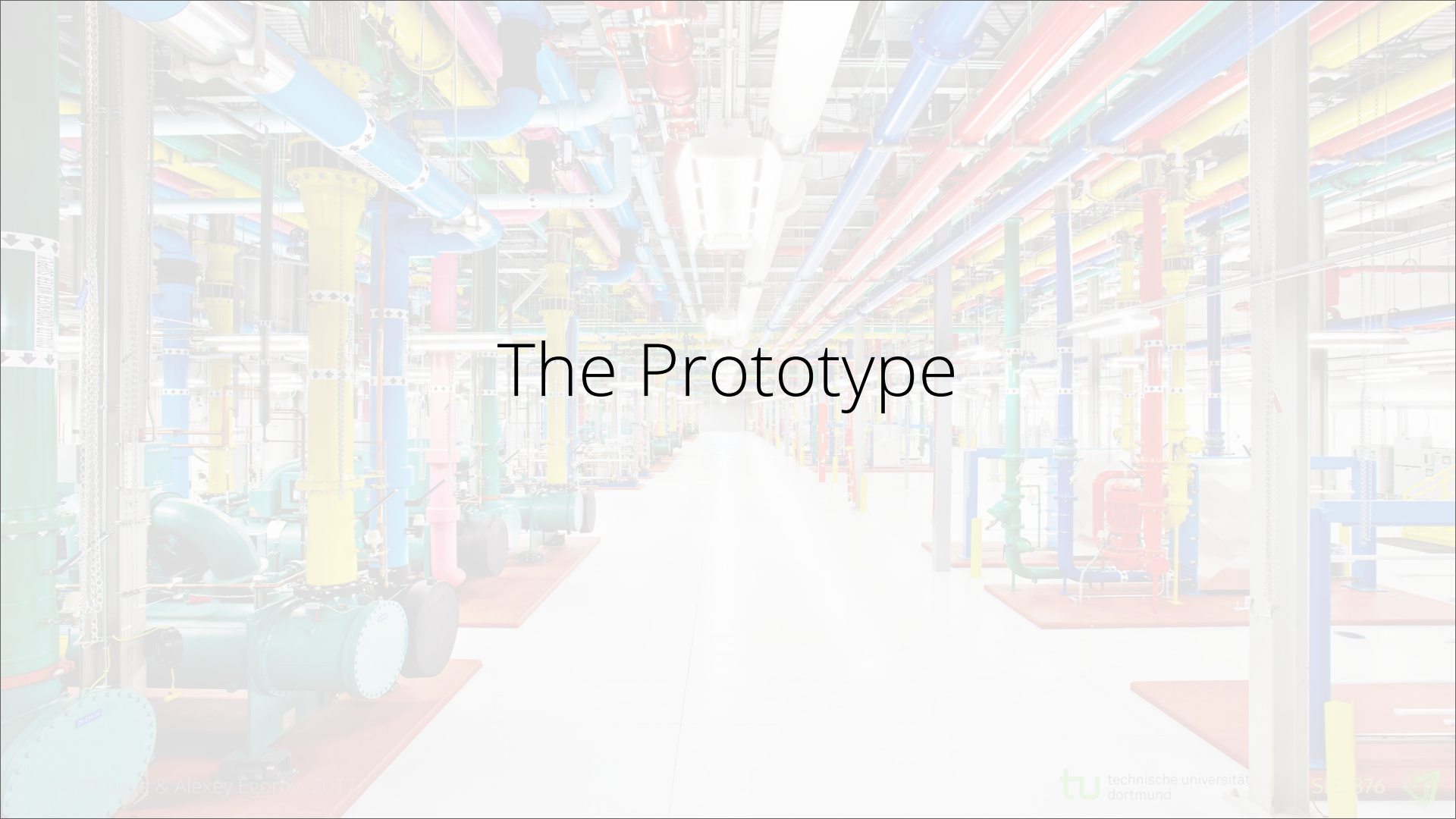
Thierry Stolarczyk
Jeremy Decock
Karl Kosack
Tino Michael
Jens Buß
Max Nöthe
Wolfgang Rhode
Christian Bockermann
Katharina Morik

Kai Brügge & Alexey Egorov 2017



Contents

1. RTA Prototype
2. Analysis Results
3. Runtime Performance



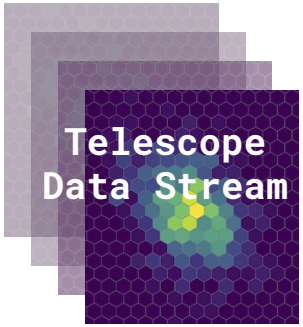
The Prototype

Steps performed by the RTA prototype:

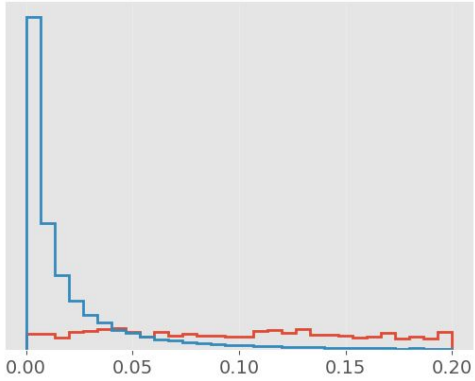
1. Read Images
2. Perform Image Cleaning
3. Extract Hillas Parameters
4. Reconstruct Direction
5. Random Forest Energy Regression
6. Random Forest Gamma/Hadron Classification
7. Write Results

Simulated Type	Width	Length
gamma	50	153

model persisted to file or database



Type	Energy	Ra	Dec
hadron	60	4.9	24.3



Goal for the RTA prototype:

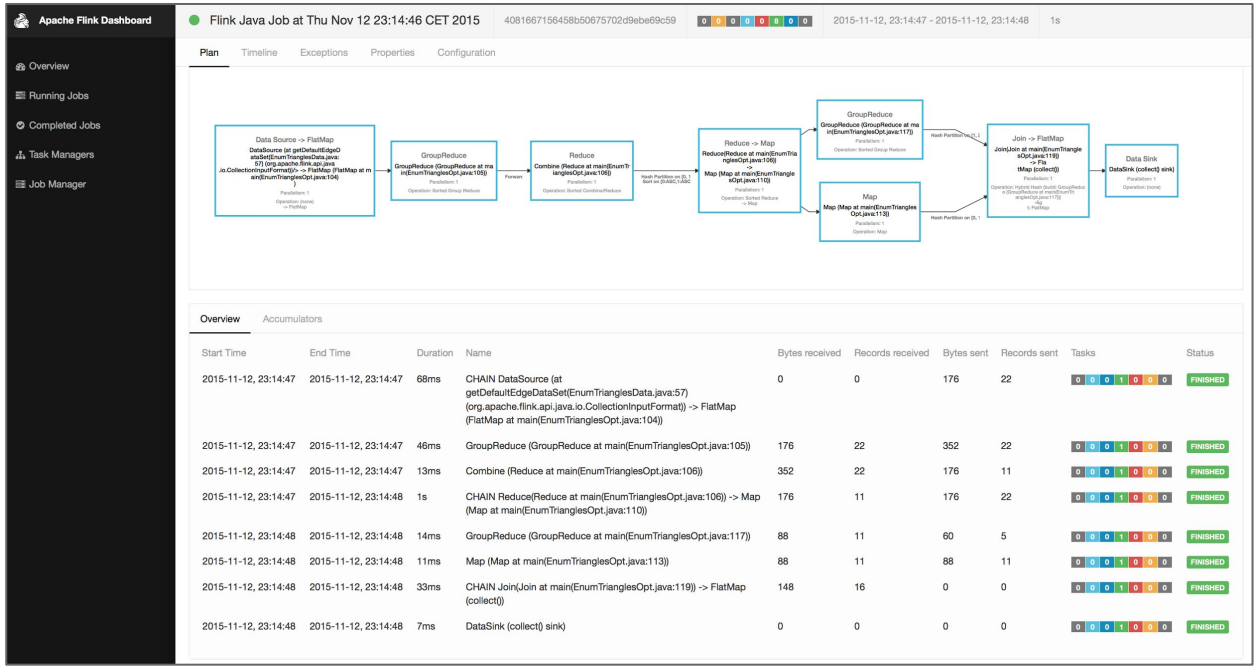
Leverage Big-Data frameworks to distribute streamed computing to several computers.

Wikipedia on Apache Flink:

Flink provides a high-throughput, low-latency streaming engine[7] as well as support for event-time processing and state management. Flink applications are fault-tolerant in the event of machine failure and support exactly-once semantics.[8] Programs can be written in Java, Scala,[9] Python,[10] and SQL[11] and are automatically compiled and optimized[12] into dataflow programs that are executed in a cluster or cloud environment.[13]

In simpler terms:

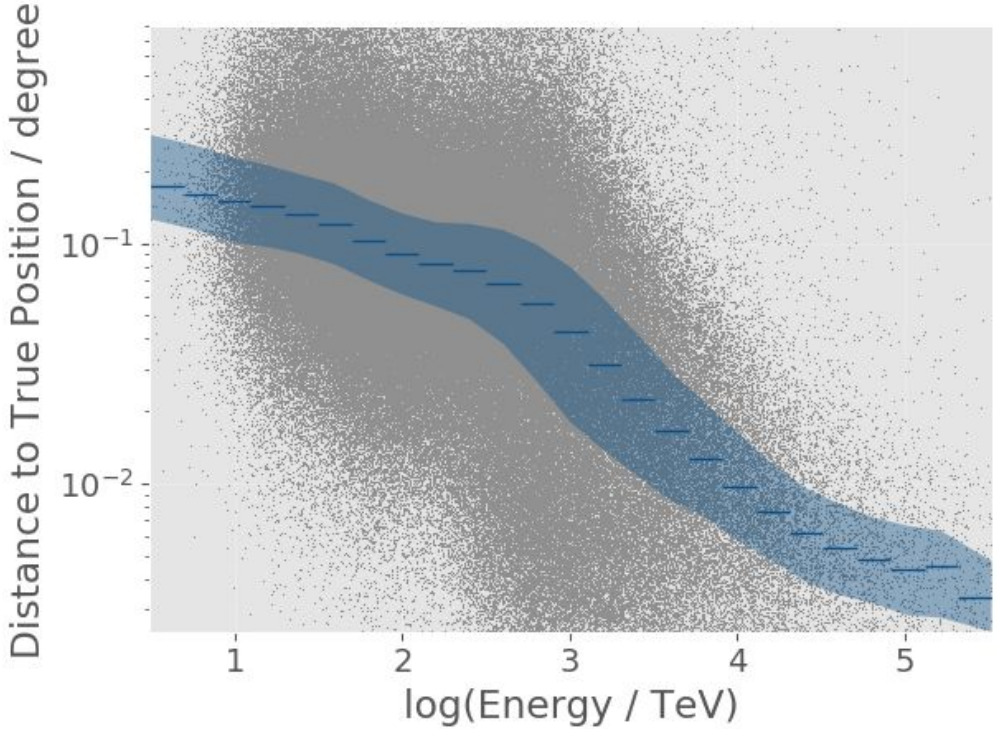
Apache Flink is nice for distributing streamed computations.



Analysis Results

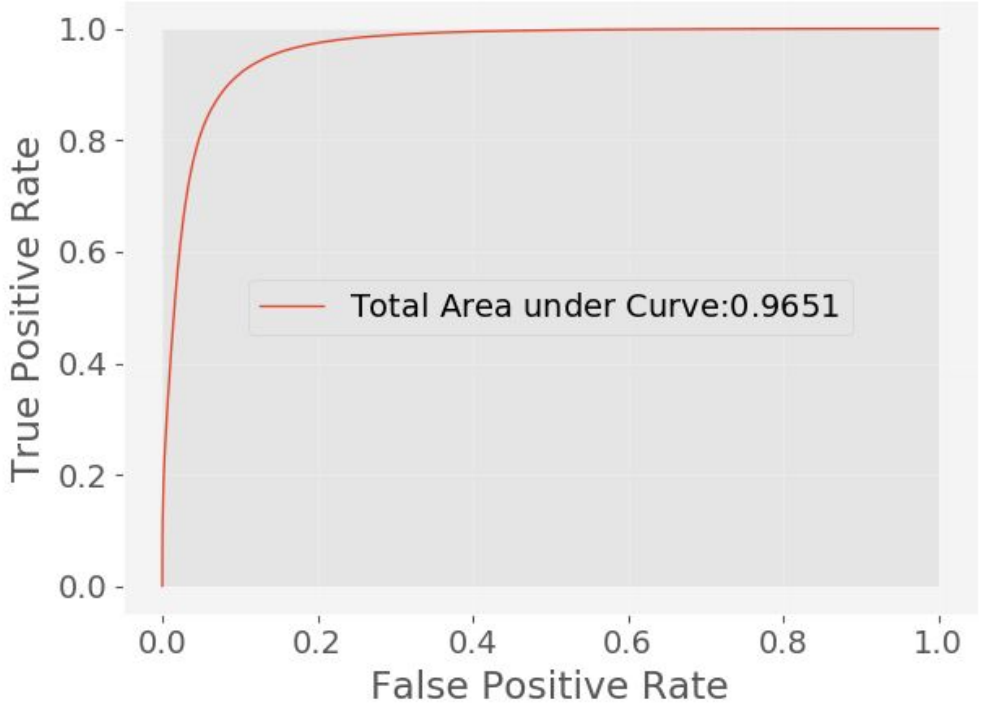
Direction Reconstruction

Implementation of reconstruction algorithm from ctapipe.



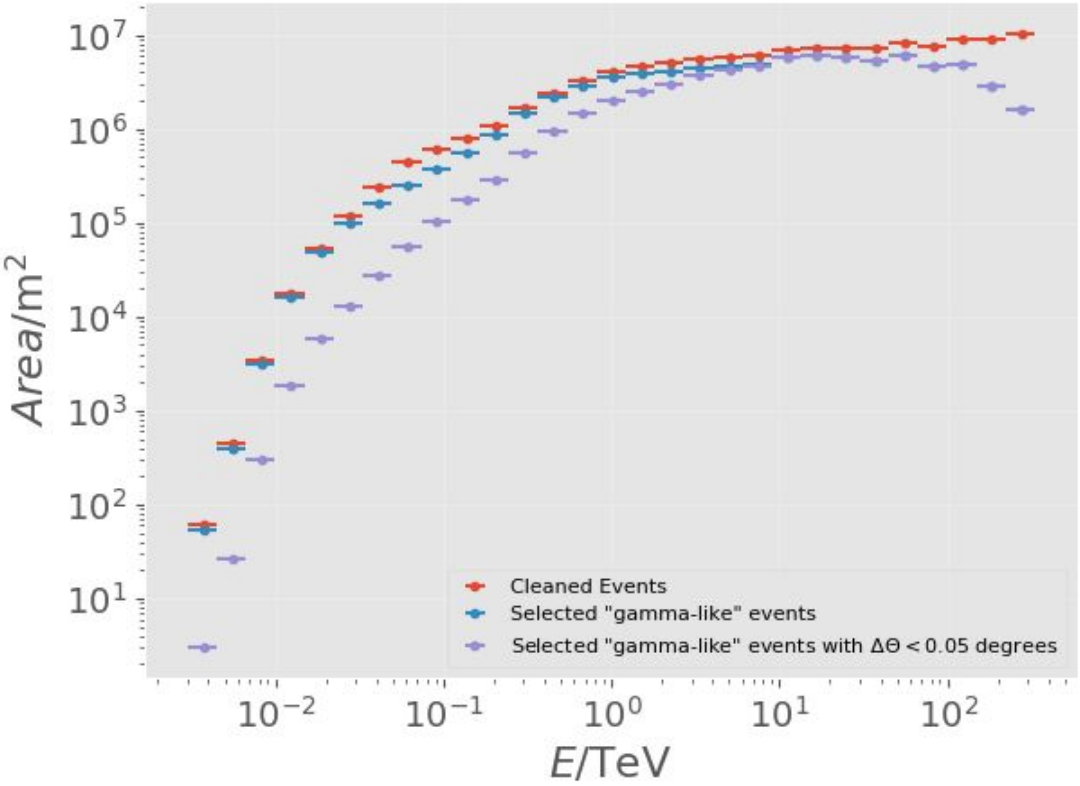
Classification Performance

One single Random Forest for all telescope types.



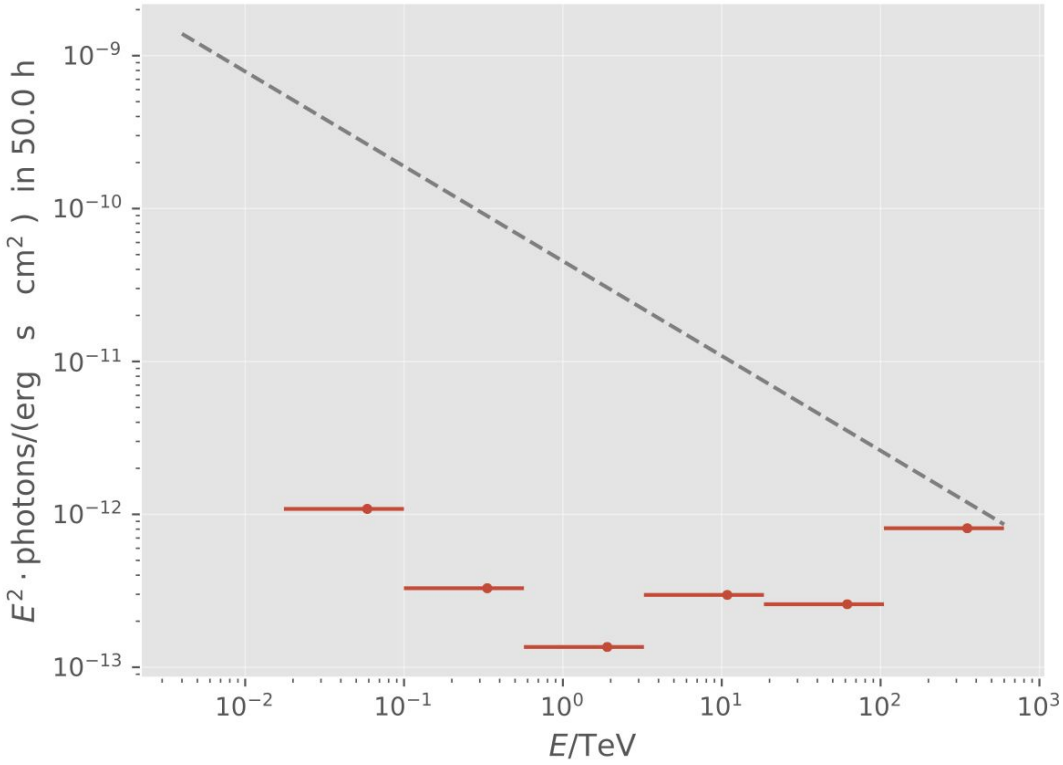
Effective Area

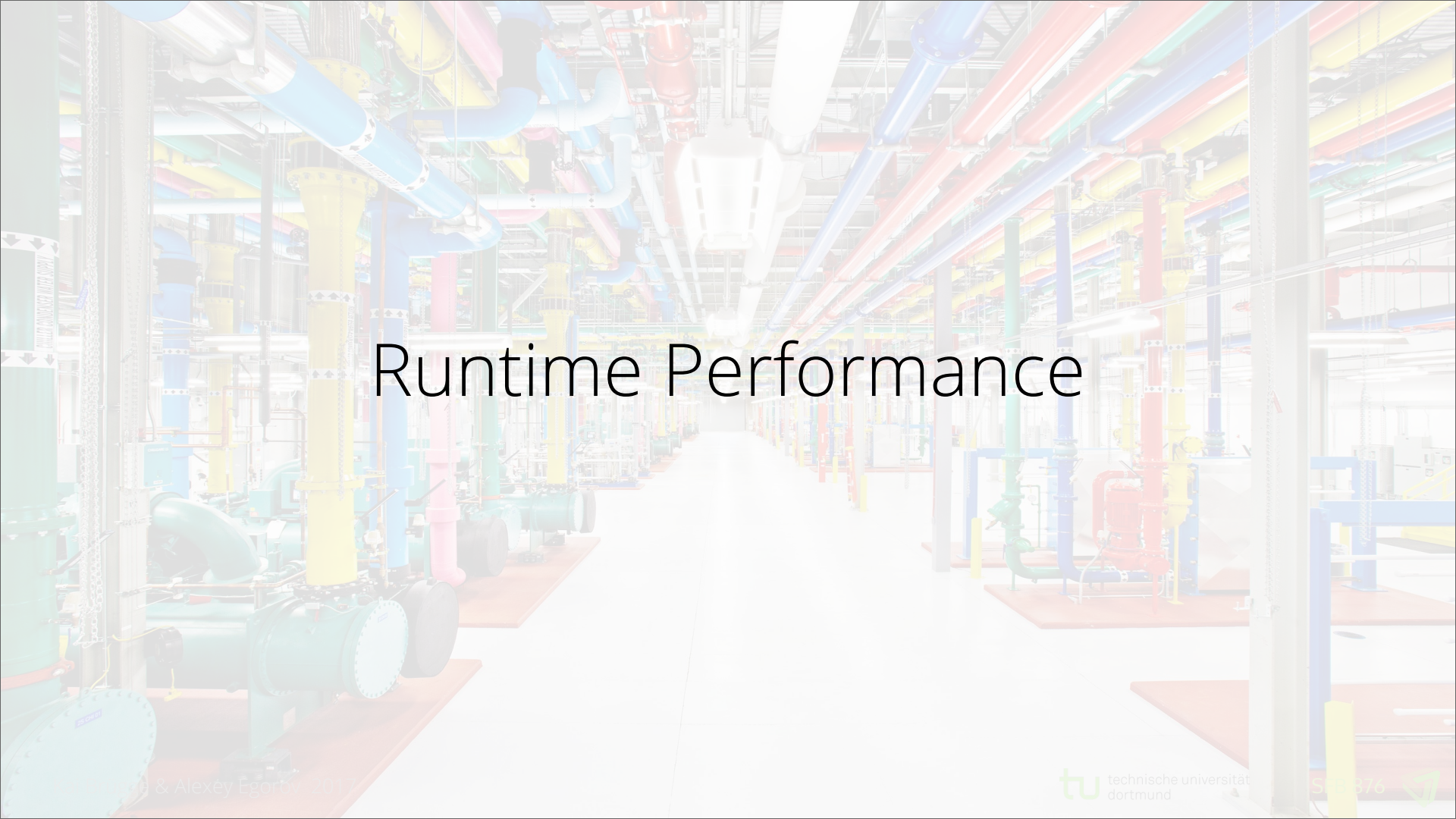
Effective Area for different event selections.



Sensitivity

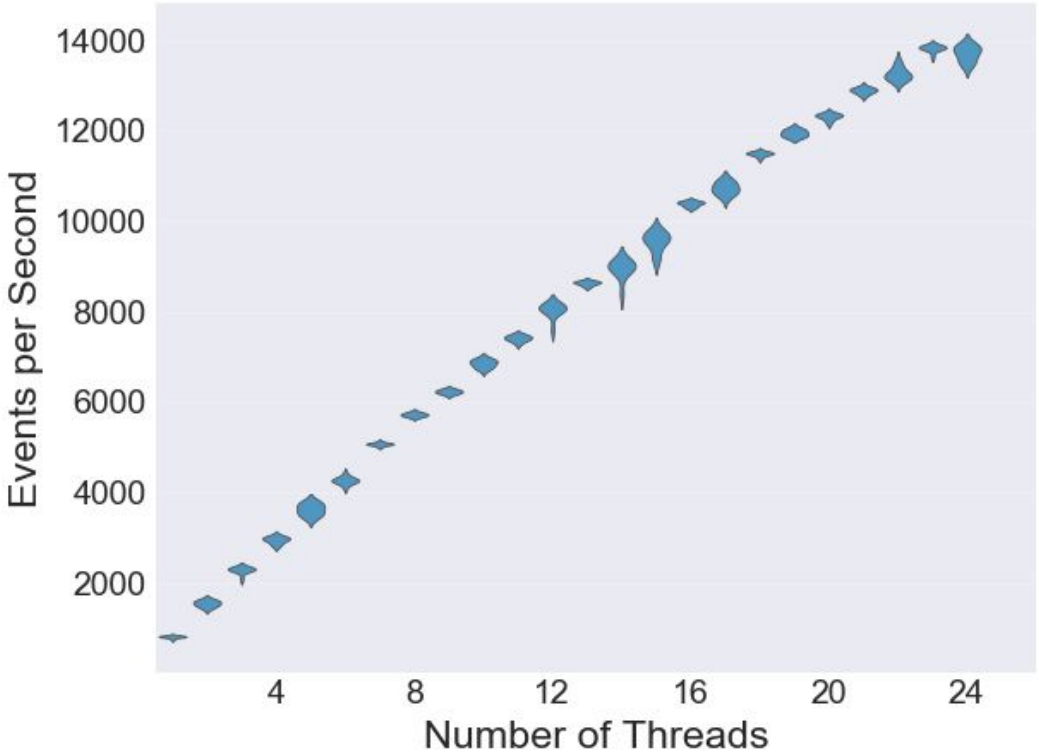
Preliminary sensitivity for the RTA prototype.



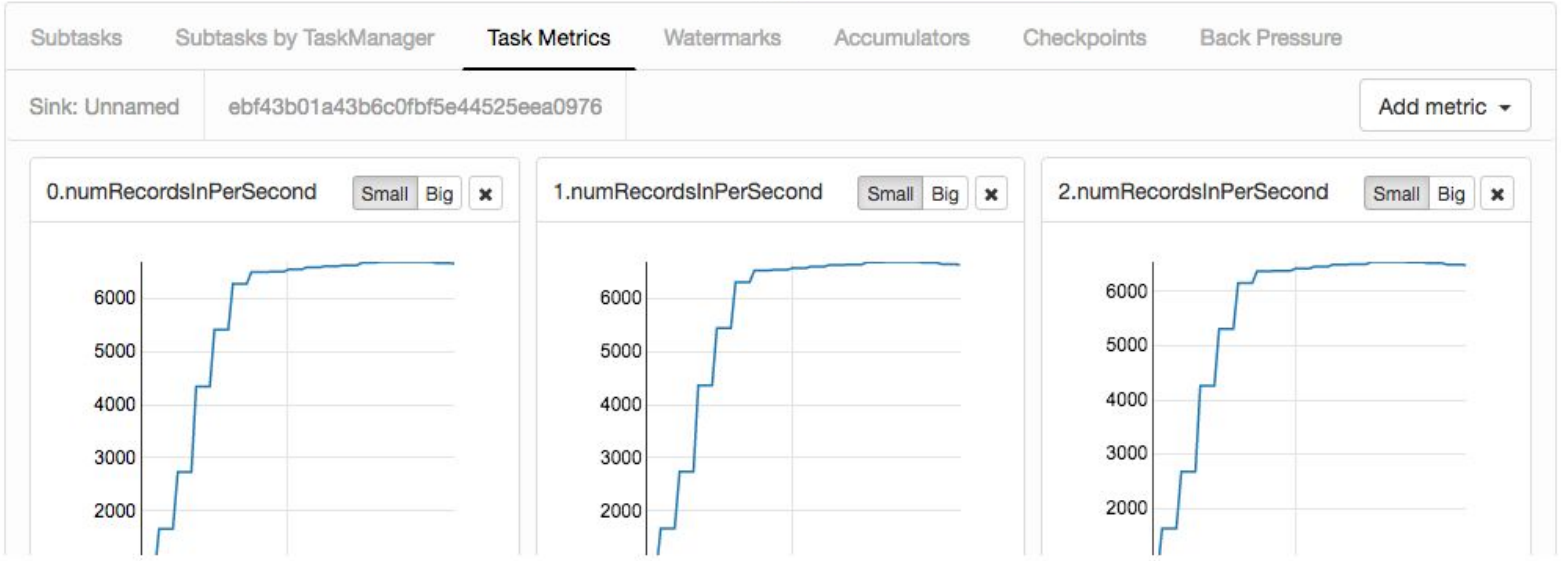


Runtime Performance

Eventrate on a single machine with 24 cores.



~20 000 events/s on two machines.



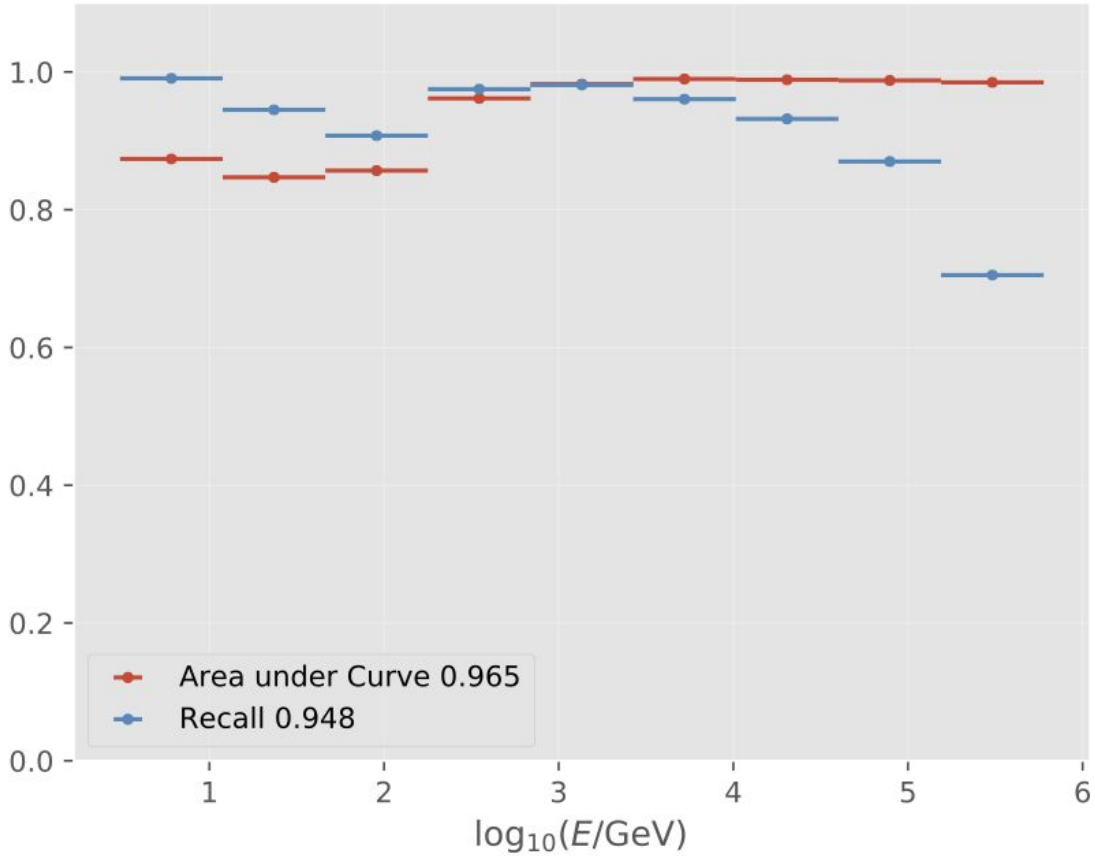
Conclusion

DL3 event lists from images in real time using industry standards for scalable stream computing

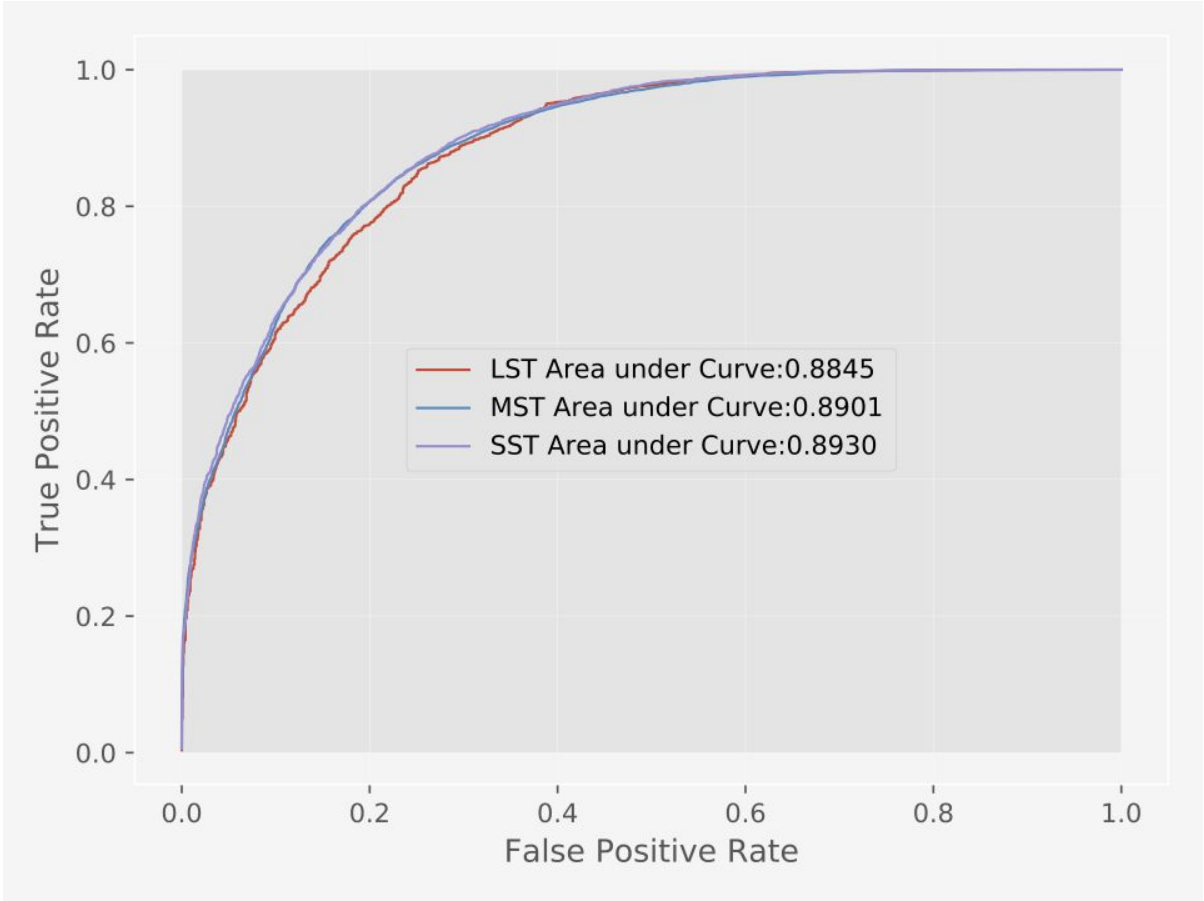
GitHub

`github.com/mackaiver/jayct`

Backup I


1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
4
5
6
$\log_{10}(E/\mathrm{GeV})$
Area under Curve 0.965
Recall 0.948

Backup II


True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
LST Area under Curve:0.8845
MST Area under Curve:0.8901
SST Area under Curve:0.8930