

Transition from data-archives to knowledge discovery, the O in IVOA

KDD/GWS Session

Saturday October 28 - O'Higgins - 9:00 - 10:30

Speaker	Title	Duration	Materials
Kai Polsterer	Introduction	10'	pdf
André Schaaff	Ongoing investigations around Spark and bringing code to the data	12' + 3'	pdf
Giuliano Taffoni	HPDA (high performance data analysis)	12' + 3'	pdf
Brian Major (for Dave Morris)	Discovering Internal Compute Capabilities	5' + 5'	pdf
François-Xavier Pineau	Prototype of an automated classification service: a use case for KDD?	20' + 5'	pdf
All	Open Discussion	15'	

Transition from

KDD/GWS

Saturday

Speaker

Kai Polsterer

André Schaaff

Giuliano Taffoni

Brian Major (for D)

François-Xavier F

All

□ Bringing the code to the data

- How to allow users to execute code near our data ?
 - Which code ? On which data ?, ...
 - Hardware resources, accounts, security, etc.
- Development of Jupyter Notebooks
 - Submit X-Match jobs to Spark from Python notebooks
 - Ipyaladin, Aladin Lite embedding in a Notebook (see Thomas's talk in Apps)

28/10/2017

IVOA, Santiago, 27-29/10/2017

8

Transition from

KDD/GWS S

Saturday Oct

Speaker

Kai Polsterer

André Schaaff

Giuliano Taffoni

Brian Major (for Dave I

François-Xavier Pineau

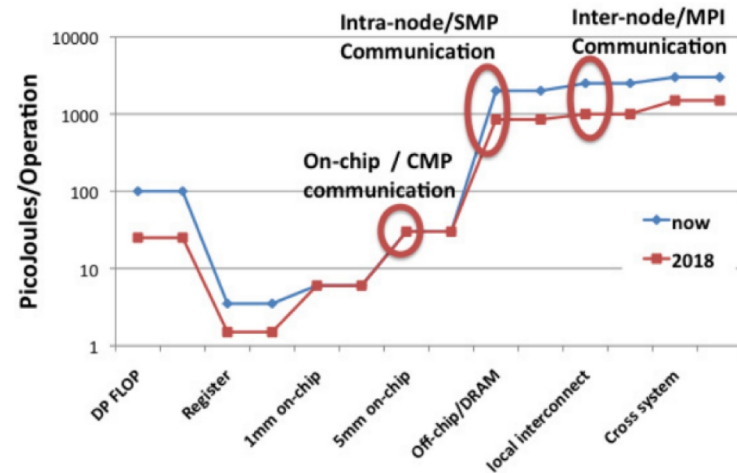
All



Minimize data movement

Move your code close to the data.

It may be not sufficient



In-memory processing

KDIG / GWS

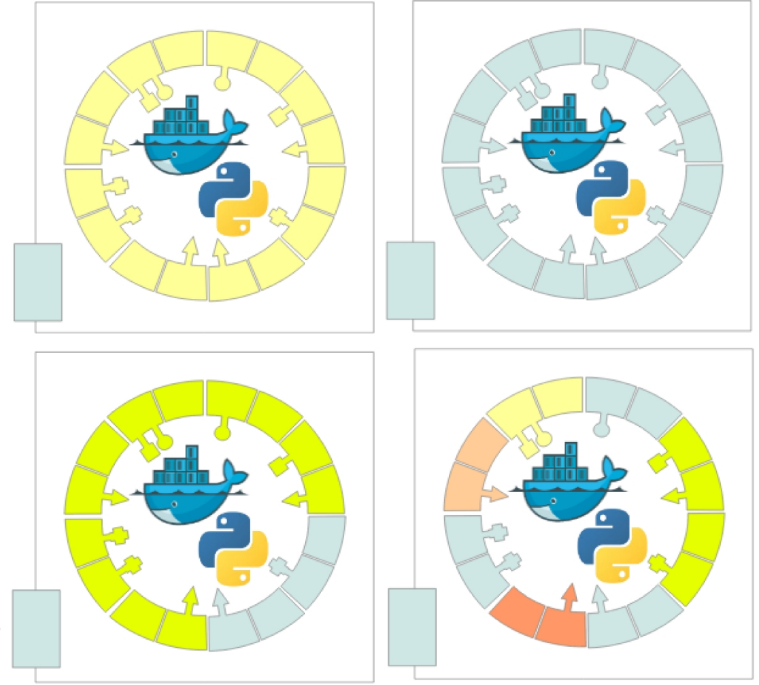


Transition from data

KDD/GWS Sess
Saturday Octobe

- Speaker**
- Kai Polsterer
- André Schaaff
- Giuliano Taffoni
- Brian Major (for Dave Morris)
- François-Xavier Pineau
- All

Standard way of running
the same applications
inside different
data centers



D.Morris
Institute for Astronomy,
Edinburgh University
June 2015



KDIG / GWS



Transition from data

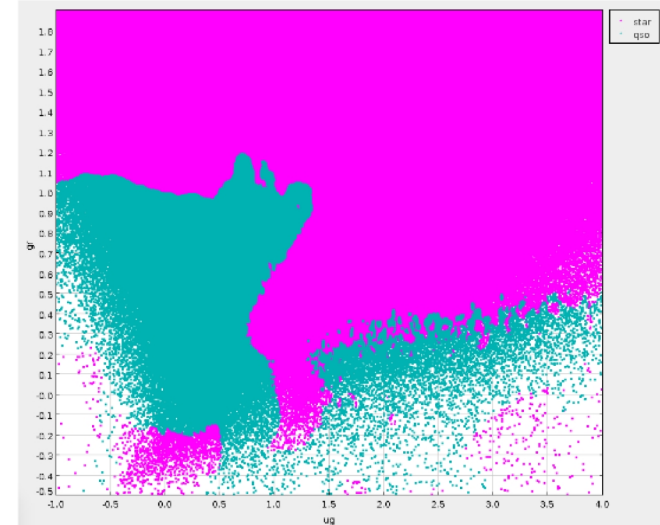
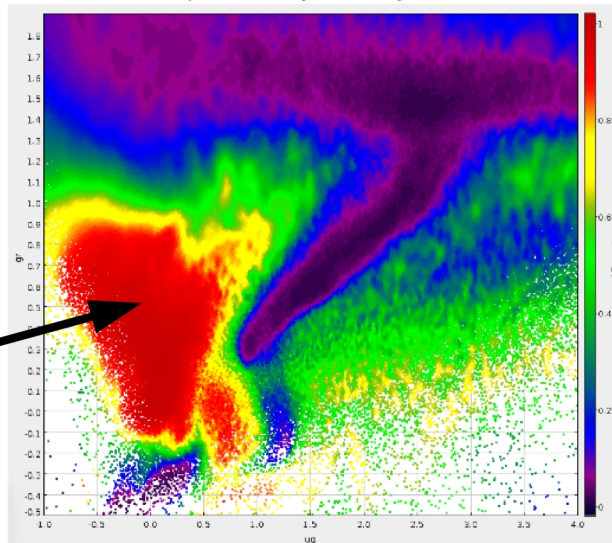
KDC example

KDD/GWS Session
Saturday October 28

Speaker

- Kai Polsterer
- André Schaaff
- Giuliano Taffoni
- Brian Major (for Dave Morrison)
- François-Xavier Pineau
- All

Probability of being a quasar (left) and repartition of sources classes as quasar/star (right).



$$p(qso|\vec{x}) = \frac{p(qso)p(\vec{x}|qso)}{p(qso)p(\vec{x}|qso) + p(star)p(\vec{x}|star)}$$

Cyan: $p(qso|\vec{x}) < 0.5$
Pink: $p(qso|\vec{x}) \geq 0.5$