Ongoing investigations around Spark and bringing code to the data

André Schaaff, François-Xavier Pineau, Thomas Boch Centre de Données astronomiques de Strasbourg

Corentin Sanchez<sup>1</sup>, Paul Trehiou<sup>1</sup>, Jérôme Desroziers<sup>2</sup> <sup>1</sup>Université de technologie de Belfort-Montbéliard, <sup>2</sup>Telecom Nancy



#### IVOA, Santiago, 27-29/10/2017







H2020-Astronomy ESFRI and Research Infrastructure Cluster (Grant Agreement number: 653477).



### Apache Spark & X-Match

### • Reminder, status, next steps

### Reminder: Apache Spark

- "Apache Spark is a cluster computing platform designed to be fast and general purpose"
- Extension of the MapReduce model to support more types of computations (interactive queries, stream processing, etc.) offering APIs for Scala, Java, Python, R,...
- Computations in memory (as much as possible, otherwise pilling to the



## Reminder (2): Use case & data

- Cross-matching of large source catalogues:
  - 2MASS<sup>1</sup>, 470,992,970
  - SDSS<sup>2</sup> DR9, 469,053,874
    - Full sky (<u>all</u> the sources), cone or HEALPix cell

Fuzzy join between 2 tables



1 - 10       F10.6       deg       RAdeg       (ra) Right ascension (Jule 11, 10, 10, 10, 10, 10, 10, 10, 10, 10,		Byte	s Format	t Units	Label	Explar	nation	S		-			a de la constante de la constan En especta de la constante de la	
1 - 10       F10.6       deg       RAdeg       (ra) Right ascension (J         12 - 21       F10.6       deg       DEdeg       (dec) Declination (J200         23 - 26       F4.2       arcsec       errMaj       error ellipse         28 - 31       F4.2       arcsec       errMin       error ellipse         33 - 35       I3       deg       errPA       [0,180] (err_ang) Position angle of error         Image: State St														
28- 31       F4.2       arcsec       error       error       ellipse         33- 35       I3       deg       errPA       (0,180)       (err_ang)       Position angle of error         Ver       ************************************		1- 1 12- 2 23- 2	0 F10.0 1 F10.0 6 F4.2	6 <u>deg</u> 6 <u>deg</u> <u>arcsec</u>	RAdeg DEdeg errMa	(ra) F (dec) j (err_m	tight Decli naj) S	ascens nation emi-ma	ion (J (J200 jor axi					
33- 35 I3 deg errPa       [0,180] (err_ang) Position angle of error         veit       veit       (0,180] (err_ang) Position angle of error         veit       (0,180] (err_ang) Position angle of error         veit       (error ellipse         (error ellipse)       (error ellipse)         (error ellipse)       (error ellipse)         (error ellipse)       (errang) Position angle of error         (error ellipse)       (errang)         (error ellipse)       (error ellipse)		28- 3	1 F4.2	arcsec	errMi	error n (err n	ellip min) S	se emi-mi	nor axi	s of	pos		San Sug	20
335 - 35 - 13       Geg       OFFRA       [0,1801] (BFF_alig) POSITION and Te OF OFFRA			E T2			error	ellip	se	Desiti					
• view		33- 3	5 13	deg	errpa	[0,180	)] (er	r_ang)	POSICI	on a	ngre		rior	
Image: Control in the stand of the stan	📍 🖣 Vizie	ieR ×	+											
Note         Yange	<ul> <li>i vizier.u-strasbg.fr/viz-</li> </ul>	-bin/VizieR-4				C Q Rec	hercher:		+ 1		i 🛡 -	4 9	4 + + =	Ξ
Image: State of the s	Portal Simb	ad VizieR /	Aladin X-Mato	h Other Help									á	
Concerning     C					VizieR	Result Page							Send to	VO tools
True       Bit Control       <		Show the target fr	orm		VIZICII	nesult rage							•	ž
and COSpecting       and Cospecting       2005/Clill284_CC       ReadMarting         Dial Social       2005/Clill284_CC       ReadMarting       2005/Clill284_CC       ReadMarting         Dial Social       2005/Clill284_CC       2005/Clill284_CC       ReadMarting       2005/Clill284_CC       ReadMarting         Dial Social       2005/Clill284_CC       2005/Clill284_CC       2005/Clill284_CC       ReadMarting       2005/Clill284_CC       ReadMarting         Modify Guerrer       2005/Clill284_CC       2005/Clill284_CC <td< td=""><td>arch Criteria</td><td><ul> <li>Show constraint in</li> </ul></td><td>nformation</td><td>outed by ViziaP or</td><td>nd any not part</td><td>of the original data</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></td<>	arch Criteria	<ul> <li>Show constraint in</li> </ul>	nformation	outed by ViziaP or	nd any not part	of the original data								
Numeric Back         ZAMAGONI To Control Sources (Cust): 2003)         ZOUSCALERADA LS Monthering (Cust): 2003)         ZOUSCALERADA LS Monthering (Cust): 2003)           Data and the state of the Call Sources (Cust): 2003)         Call Sources (Cust): 2003)         ZOUSCALERADA LS Monthering (Cust): 2003)         ZOUSCALERADA LS Monthering (Cust): 2003)           Data and the custope of the Call Sources (Cust): 2003)         Call Sources (Cust): 2003         ZOUSCALERADA LS Monthering (Cust): 2003)         ZOUSCALERADA LS Monthering (Cust): 2003           Zada         Call Sources (Cust): 2003         Example (Cust): 2003         ZOUSCALERADA LS Monthering (Cust): 2003         ZOUSCALERADA LS Monthering (Cust): 2003           Stat Andrein Lie         Custope (Cust): 2003         Example (Cust): 2003         ZOUSCALERADA LS Monthering (Cust): 2003         ZOUSCALERADA LS MONTHERING (Cust): 2003         ZOUSCALERADA LS MONTHERING (Cust): 2003           Monthly Cust;         Monthly Cust;         ZouSCALERADA LS MONTHERING (Cust): 2003         ZouSCALERADA LS MON	e in CDSportal	The 5 columns in	i color ale com	Juted by Viziek, at	iu ale noi puri	oj ine originai aaia.								
Construit         Construit <thconstruit< th=""> <thconstruit< th=""> <thc< td=""><td>ywords Back</td><td>II/246/out Post annotation</td><td>2MASS All-S</td><td>ky Catalog of Poin</td><td>t Sources (Cutr 70 002 070 com</td><td><u>i+ 2003)</u> mar. Planca asknowi</td><td>lodge the u</td><td>rogo of the 1</td><td>2003yCa</td><td>£ 22460</td><td>C Read?</td><td>Me+ftp</td><td></td><td></td></thc<></thconstruit<></thconstruit<>	ywords Back	II/246/out Post annotation	2MASS All-S	ky Catalog of Poin	t Sources (Cutr 70 002 070 com	<u>i+ 2003)</u> mar. Planca asknowi	lodge the u	rogo of the 1	2003yCa	£ 22460	C Read?	Me+ftp		
240     Comparing a start of the start of th	bles	1 ber terenennen	2MASS Pages	Note that the m	0,992,970 sour	d correspond to low	quality res	age of the 2	mits or very poc	r photor	(; see also metry)	me		-
Bat         Static Labelin-Lite           Choose         BAL2000         DE/22009         BAL2000         DE/22009	246		(470992970 rd	ows)	ugintudos in re	a conceptina to to a	quanty rest	no (apper m	millio or very poo	i photon	licity)			
Start AlabanLate         Start AlabanLate         Bar AlabanLate         Bar AlabanLate         Bar AlabanLate         Composition         Com	but													
Choose         Choose<		Start Alad	InLite	212000 DA 12000	DE 12000	211455	Imag	Umag	Kmag a	Offic	Dfla Bfla	CBa XB	a A 8a	
amaratas         200570 0.22 (4233) +41 (50 5 30 0.649137) +41 25930 0.024335 +11085 0.510 (623 0.051 (627) 0.51 (627) 0.01 (62 - 10) 0.00	Choose	arcmin !	"himis" "d	imis" deg	deg	201000	mag m	s mag u	ag mag mag	s Cong	Kug Dug	, Cug Au	i nug	
andremoda         1000583 (00-22 4003)+41 (00 683409) +41 269583 00 24003+41 (009) 9321         8.614         10.001 00.02 UUE (02 00) 000 2         0         2         0           100058 (00-22 4003)+41 (00 583109) +41 26953 00 24045+41 (000) 1000 000 UUE (02 00) 000 2         0<	astraints	10017100	42 44 337 +41 1	608 53 010 6847?	7 +41 269035	00424433+4116085	945300	52 8 668 0 0	1 8 475 0 05	I FEE	222 111	000	2 0	
Normit 7	ndromeda	20.0568.004	42 44 033 +41 1	6 06 91 010 68346	9 +41 268585	00424403+4116069	9 321	8 614	10 601 0.07	S UITE	002 001	000	2 0	
Monthly Quary         Terms         Top (0.659) (0.24 24.664) +11 (0.922) (0.0586026 +11.269220 0023464+11 (0.922) 92.99)         8.660         10.119 0.056         10.119 0.055         01.019 0.050<	arcmin 2)	3 0.0643 00 /	42 44 558 +41 1	6 10.38 010.68565	7 +41.269550	00424455+4116103	10.773 0.0	59 8.532	8.254	EUU	200 200	c00	2 0	
Interview         State          State         State	Modify Query	4 0.0659 00 /	42 44 646 +41 1	6 09.21 010.68602	6 +41.269226	00424464+4116092	9.299	8.606	10.119 0.05	6 UUE	002 001	00c	2	
And controls       00 07071       00 02 42 4464441       10 10 05 00 0580015       +11 269500 002444444141001       00 120 0000       82.29       UEU (20 050000       00 050000       82.29       UEU (20 050000       00 050000       82.29       UEU (20 050000       00 050000       82.29       UEU (20 050000       82.29       UEU (20 050000       82.29       UEU (20 050000       82.29       UEU (20 0500000       82.29       UEU (20 05000000000000000000000000000000000		5 0.0789 00 4	42 44.032 +41 1	6 10.83 010.68346	5 +41.269676	00424403+4116108	11.507 0.0	56 8.744	8,489	EUU	200 100	c00	0	
Table         Tobic         Tobic <th< td=""><td>terences</td><td>6 0.0791 00 /</td><td>42 44.644 +41 1</td><td>6 10.67 010.68601</td><td>5 +41.269630</td><td>00424464+4116106</td><td>9.399</td><td>9.985 0/</td><td>070 8.429</td><td>UEU</td><td>020 02</td><td>100</td><td></td><td></td></th<>	terences	6 0.0791 00 /	42 44.644 +41 1	6 10.67 010.68601	5 +41.269630	00424464+4116106	9.399	9.985 0/	070 8.429	UEU	020 02	100		
1100 1.100 0.02 42.3383-41 10 0234 010 0.83205 41.1267456 00.023984-11 0221 21.30 0.004         2.26         8.994         AUU 20 100         100           20 1.110 0.02 42 42.334 11 0.0234 010 0.83205 41.1267456 00.024984-11 0005         0.005         9.374         11.500 0.002         100         110         100 <td< td=""><td>max: 50</td><td>Z 0.1008 00 4</td><td>42 44.465 +41 1</td><td>6 01.65 010.68527</td><td>0 +41.267124</td><td>00424446+4116016</td><td>12.070 0.0</td><td>35 9.301</td><td>9.057</td><td>EUU</td><td>20/</td><td>000</td><td>100</td><td>•</td></td<>	max: 50	Z 0.1008 00 4	42 44.465 +41 1	6 01.65 010.68527	0 +41.267124	00424446+4116016	12.070 0.0	35 9.301	9.057	EUU	20/	000	100	•
Commax <u>9             1           </u>	All columns	8 0.1014 00 4	42 43.983 +41 1	6 02.84 010.68326	3 +41.267456	00424398+4116028	12.136 0.0	40 9.226	8.994	AUU	20 100	1000		
Submit         Julo 1140 00 42 3372 +41 10 643 30 00.82283 +41 .267925 00.24377+41 1045 12.446 0.061 11.738 0.065 19.75         AA         10           Tree         D35, France         D <tdd< td=""><td>Compute</td><td><u>9</u> 0.1111 00 4</td><td>42 44.203 +41 1</td><td>6 00.99 010.68418</td><td>0 +41.266941</td><td>00424420+4116009</td><td>10.065</td><td>9.374</td><td>11.504 0.05</td><td>2 UU</td><td>2 2</td><td></td><td></td><td>-</td></tdd<>	Compute	<u>9</u> 0.1111 00 4	42 44.203 +41 1	6 00.99 010.68418	0 +41.266941	00424420+4116009	10.065	9.374	11.504 0.05	2 UU	2 2			-
Jones         Juli 10104 00 42 43661+41 10 12.00 010582777 +41.270111 00423864-411023 9977         Dit R39 0.005 (11.390 0.002 UAA           95, France         Juli 20104 00 42 43661+41 10 14.010 0058377 +41.270110 00423864-4110123 9987         Juli 2010 1008 1393 002 UAA         Juli 20104 00 42 4501+4141 161 161.000582773 +41.27059 004244-441410001 25:55 005 9310         9274         Juli 2010 008 13934         UUU 1002 12:55 005 9310         9274         Juli 20104 00 20 12:52 002         2         0           Juli 201028 00 42 44.107+41 16 00.06 0010 083944 +41.26656 0042444+4110010 12:55 0.005 9339         Juli 2010 008 13:930         9391         Juli 2000 002 002 002         2         0           Juli 2010 001 02 42 43157+41 10 10.40 010 058201 3+41 26750 00424385+411014 10.176         Juli 370 005 9527         92010 00 000 02         2         0           Juli 2010 02 42 4379+41 11 031 41 20750 00424385+411014 10.176         Juli 370 000 9421         92010 00 000 02         2         0           Juli 2010 02 42 4379+41 11 031 41 20750 00424385+411014 10.176         Juli 370 000 9412         92010 00 000 02         2         0           Juli 2010 02 42 4379+41 11 10 41 01 003 80064844+41 20741 00424484+4110451 2872 0.060 9413         9178         Juli 200 00 000 02         0         0         0         0         0         0001 000 02         0         0         0         0         0         0         0001 000 02         0	Submit	10 0.1160 00 4	42 43.772 +41 1	6 04.53 010.68238	3 +41.267925	00424377+4116045	12.446 0.0	51 11.753 0.0	063 9.075	AA	.0	-		-
OS. France         Prance         Prancocon	rrors	1 0.1194 00 4	42 43.866 +41 1	6 12.40 010.68277	7 +41.270111	00424386+4116123	9.977	11.683 0.0	056 11.839 0.06	2 UAA	1			
L2[0] 1288 (0) 42 44 147 +41 10 000 60 100 839344 +41 126682 0042414+41 1000 12.555 0.055 9310         9.274         AUUL 1         1         2           L4[0] 326 (0) 42 44 167 +41 10 152 40 100 84907 +41 207080 0042416+4111051 1000 105         9.399         11 1409 0055 ULD 000 000 000 000 000 000 000 000 000 0	DS, France 💿	12 0.1221 00 4	42 44.601 +41 1	6 14.16 010.68583	7 +41.270599	00424460+4116141	9.880	12.051 0.0	368 8.934	UAU				
Lg(0)         032         042         41.67         013         042         042         02         0           Lg(0)         054         042         0.45         014         01.00         054         015         01         015         055         014         000         015         015         016<		13 0.1288 00 4	42 44.147 +41 1	6 00.06 010.68394	4 +41.266682	00424414+4116000	12.565 0.0	55 9.510	9.274	AUU	2			
120       1358       00       24       353       141       100       140       1370       111       750       527       020       220       1		14 0.1326 00 4	42 44.167 +41 1	6 15.24 010.68402	:9 +41.270901	00424416+4116152	10.063	9.359	11.409 0.05	5 UUA	002 0		2 0	r
Image: 100 are descripted and set of the se		15 0.1358 00 4	42 43.851 +41 1	6 01.40 010.68271	3+41.267056	00424385+4116014	10.176	11.876 0.0	050 9.252	UEU	020 0.			
L2U.1522409 42 et a351+41 10 14 57 (010.86846) et al. ZVV14 00424484+110148 12.872 (20.01) 94/351         9.176         AUU 200 (200 e00 2)         0           Z0.1538 (002 44:87) et al. 10 005 8010.6686673 et al. 26627 0042484+1161005 1050 12.059 (012 e00 2)         20         0         20		16 0.1392 00 4	42 44.979 +41 1	6 03.48 010.68741	4 +41.267632	00424497+4116034	12.371 0.0	56 9.627	9.379	EUU	200 100	CUC	2 0	
L2011329 (M * 4 ***.4) [144] 10 0028 (ULD80905 3+14. 200822 / ULD8484 7+41 10000 [ULD80]         L2008 [ULD8012 (ULD8141 - L20032 (ULD80905 2+14. 2018) - L2004 - L200		12 0.1522 00 4	¥2 44.843 +41 1	0 14.57 010.68684	.0 +41.270714	00424484+4116145	12.872 0.0	9.433	9.178	AUU	200 200	c00	2 0	
220 1000 (00 42 555(4) 11 10 1520 (10008 ft) 141 27002 (0042402411150) 1505 (100 9 (20) 42 (20		18 0.1538 00 4	42 44.8/1 +41 1	6 00.58 010.68696	3 +41.266827	00424487+4116005	10.450	12.094 0.0	33 11.728 0.03	9 UEE	622 021	DCC 2	2 0	
2(0)(297)(0) * 4 × 3 × 60 + 41   10   (2.3 × 10)(0.88800) + 41 / 27(149) (0.0425)(243) + 41 (1.5 ) (2.598) (0.071   7) × (2 + 3.480   AUU 200   100   000   2   0 2(0) (200   0) 4 × 23 × 304 + 41   15 × 200   0) (3.6 × 10 + 41 / 27(12) (0.245) × 401 + 61   (3.253) (0.052   7) × 00   0) (2 + 0)	-	190.1606 004	42 45.027 +41 1	0 13.09 010.68761	.1 +41.270302	00424502+4116130	13.055 0.1	19 9.504	9.246	AUU	200 200	00 2	2 0	
22,0,0005 00 42 43,500 441 10 159,20 010,062,511 441,271721 (0424300441010161 12353) 00,52 19,500 11,662 (0,061 A02) 22 1 (00 2 0 2 0 2 0 2 0 2 0 0 0 0 2 0 4,500 441 15 59,80 010,681043 441 1266632 (042434544115598 10,511 18,61 0,049 UUE (00 2 0 0 1 0 0 2 0 1 0 0 1 0 0 1 0 0 0 0		200.1947 004	42 43 205 +41 1	0 12.34 010.68860	3 +41.2/0149	00424320+4116125	12.090 0.0	1 9.732 52 0.000	9,480	AUU	200 100	000	2 0	
20 0 240 0 4 4 2 10 41 15 2 20 10 6 210 4 1 1 25 20 10 6 210 2 0 1 4 1 2 5 20 10 6 1 1 1 1 2 5 0 1 0 6 1 1 1 2 5 0 1 0 6 1 1 2 5 0 1 0 6 1 0 1 0 1 1 1 5 5 0 1 0 6 1 0 1 0 1 1 1 5 5 0 1 0 6 1 0 1 0 1 1 1 5 5 0 1 0 6 1 0 1 0 1 1 1 5 5 0 1 0 6 1 0 1 0 1 1 1 5 5 0 1 0 6 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0		22 0.2038 00 4	+2 +3.804 +41 1	0 10.20 010.08251	./ #41.2/1/21	00+24380+4116181	12.933 0.0	12 9.900	11.802 0.06	AUA	202 201	000	2 0	
				C CO 00 010 2010A	2 41 766627	0040404614116600	10 511	0.916	11 021 0 04		002 001	000	2 0	



Credits: http://healpix.jpl.nasa.gov/

### Most significant experiment

- GAIA (1.1 billion sources) vs IGSL3 (1.2 billions)
- 1.6 billion associations in 10 minutes vs 30 minutes for the production X-Match Server\*
- Bottlenecks are
  - network for Loading phase
  - mainly CPU during the X-Match phase (improvable)
  - co-location of data (resolved through a homemade solution)



- 9 nodes (CentOS / OpenStack, remote data storage)
- 24 threads (-> 216 threads), 64GB (-> 576 GB)
- -> possible X-Match of billion sources ...
- \* Rebuild Join not included



### Next steps

- Improvement of the algorithms
- After this phase we will test with Amazon WS
  - AWS offers a wide variety of tools including Apache Spark
  - Comparison with all the previous investigations
- Summary of this work and status during Victoria Interop

### 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 embeding in a Notebook (see Thomas's talk in Apps)

## Spark in a notebook

💭 Jupyter	X-Match.onefile Last Checkpoint: 10/12/2017 (unsaved changes)
File Edit V	iew Insert Cell Kernel Widgets Help Hide Code Trusted 🖋 Python 3 O
* * *	
> Hide/show code	Export to HTML Export to PDF via HTML Export to PDF via Latex
In [2]:	<pre>import pyspark from pyspark.sql import SparkSession import math import healpy class XMatch: definit(self,file1,file2): self.spark = SparkSession\ #Creation of SparkSession and SparkContext if nessessary .builder\ .master("spark://130.79.128.185:7077")\ .appName("X-Match")\ .getOrCreate() self.result=None print("",file1," X ",file2,"") #Load file1 s=loadBinary(self.spark,file1) self.sl=s.flatMap(fDuplicate()) # Duplicate the sources on the border of the pixels print(self.sl.count(),"sources loaded from ",file1," after duplication") #Load file2 self.s2=loadBinary(self.spark,file2) print(self.s2.count(),"sources loaded from ",file2)</pre>
	<pre>def performXMatch(self):     r = self.s1.join(self.s2).filter(filterF) # That's the CrossMatch     print(r.count()," matching sources found")</pre>

## Spark in a notebook (2)

# Switching the previous developments to Python



### ] Ipyaladin

In [2]:

import ipyaladin

aladin = ipyaladin.Aladin(target = 'Cygnus X')
aladin



Aladin Lite integration in notebook and control of displayed target and field of view

In [3]:

aladin.fov = 2
aladin.survey = 'P/allWISE/color'

### Ipyaladin (2)

In [24]:

from astroquery.vizier import Vizier
tgas\_result = Vizier.query\_region('Cygnus X', radius='3 deg', catalog='I/337/tgas')
aladin.add\_table(tgas\_result[0])

aladin.add\_moc\_from\_URL('http://saada.unistra.fr/PNRed/Moc.fits',

{'color': '#009bd6', 'opacity': 0.4})



### Add overlays: Astropy tables and MOCs

### Conclusion

- Exciting experiments, implementation and testing of various technologies, switching between languages (Java, Python, Scala, JS), etc.
- Promising results concerning the X-Match, confident to do better
- Notebooks open the door to a simple way to execute code near the data





H2020-Astronomy ESFRI and Research Infrastructure Cluster (Grant Agreement number: 653477).