**PHOTON AND NEUTRON** 

#### DATA ANALYSIS AS A SERVICE



#### Andy Götz on behalf of

#### Photon and Neutron Data Community (PanData)

Talk at e-IRG meeting @ Bratislava 15/11/2016

Who we are

**User community** 

**Example experiments** 

What our users need

Data analysis as a Service

How to get there with e-infrastructures

Conclusion



#### PHOTON AND NEUTRON SOURCES + COMMUNITY

- 10 photon sources
- 7 neutron sources
- Represents
   ~40 000
   users/year
- PaN is not CERN



#### PHOTON AND NEUTRON SOURCES



#### **HOW DOES A SYNCHROTRON WORK**







#### PAN USER COMMUNITY

- Counted users between 2012 + 2014: see http://pan-data.eu/Users2012e-Resultsfor details
- Total of 38338 users of which 14347 new each year



#### PAN USER COMMUNITY

- Counted users between 2012 + 2014: see http://pan-data.eu/Users2012e-Resultsfor details
- Total of 38338 users of which 14347 new each year



#### PAN USERS IN THE WORLD



The European Synchrotron

#### PAN USERS COUNTRY DISTRIBUTION





#### DATA ANALYSIS AT PAN INSTITUTES

#### Two quite different types of data analysis tasks :

- Analysis of experimental data (~3/4)
  - processing time / dataset typically short (seconds)
  - but large amounts of data (10 TB / day)
  - **difficult** to move data off site e.g. cloud (so far)
- Theoretical calculations and modeling (~1/4)
  - small amounts of input+output data (a few GB max)
  - but long processing times (80 cores for 5 days not uncommon)
  - **not difficult** to move to cloud

#### → theoretical calculations and modeling are candidates for cloud



#### **EXAMPLE - ESRF DATA PRODUCTION IN 2016**

• 499 experiments produced :

### 4 PetaBytes in 2016

- Data volume is growing each year
- Next generation experiments can
   produce 10 PB in 1 week



#### **EXAMPLE – DIAMOND DATA SINCE 2007**





#### MAIN DATA CHALLENGES FOR PAN COMMUNITY

- **1. Data management and policy**
- **2.** Medium to Huge and increasing data volumes
- **3.** Helping (new) users to analyse their data
- 4. Developing Data Analysis as a Service
- **5. Software development resources**
- 6. Integrating private and public clouds
- 7. PAN community not well known
- 8. How the e-infrastructures can help

Conclusion



#### **MULTIPLE DATA BOTTLENECKS**

- **1.** Acquiring data from detectors to storage (Gigabytes/s)
- 2. Reducing data online fast enough (Gigabytes/min)
- → **3. Analysing data fast enough (Terabytes/day)**
- $\rightarrow$  4. Getting results to users efficiently (Giga to Terabytes)
- → 5. Scientific + infrastructure software bottleneck ...



credit: http://arc.umich.edu/

#### PARALLEL ALGORITHMS + DEVELOPERS NEEDED

- **PyFAI radial integration**
- **PyMCA** multi-variate analysis
- PyNX ptychography
- PyHST tomography
- **XPCS** coherent diffraction correlation
- List keeps on growing ...
- Parallel programming on GPUs / CPUs with MPI / OpenCL / OpenMP requires specialised skills



#### **USE CASE - PYFAI**

- PyFAI radial integration code (pyfai.readthedocs.org)
- Hosted on GitHub: https://github.com/pyFAI/pyFAI
  - 23000 lines of Python (plus 5000 for the test)
  - 8000 lines of Cython which are converted into ... C (about half a million lines)
  - 5000 lines of OpenCL kernels
- 4 years of almost constant development speedup, new detectors, ...







#### **USE CASE – XRAY DIFFRACTION TOMOGRAPHY**



Probing the structure of heterogeneous diluted materials by diffraction tomography Pierre Bleuet, Eléonore Welcomme, Eric Dooryhée, Jean Susini, Jean-Louis Hodeau & Philippe Walter Nature Materials 7, 468 - 472 (2008) doi:10.1038/nmat2168

The European Synchrotron

#### **USE CASE – XRAY DIFFRACTION TOMOGRAPHY**



#### **USE CASE – XRAY DIFFRACTION TOMOGRAPHY**

- Data rate = 250 Hz
- Experiment = 1 week
- Sample volume = 1000<sup>3</sup>
- Data volume = 10 PB
- Have to reduce data online to < 10 TB</li>
- Use PyFAI accelerated on GPU to reduce data in realtime a factor of 1000





#### • Main issues are

- How to develop maintainable software
- How to avoid scientific software developers wasting time developing non-scientific software e.g. gui code, input/output, standard algorithms
- How to get developers to adopt best practices

#### • ESRF approach

- Develop a common library which implements common low-level functionality
- Use best practices (unit tests, continuous integration, documentation, ...)
- Python-based, C++ for speed, Open Source, simple licence (MIT), Github
- Hire software engineers to develop the library
- Build scientific applications on top
- Encourage collaborations

#### • Sustainability

The library saves development time + improves maintenance of scientific applications



#### **TOOLKIT APPROACH**





#### SILX TOOLKIT



**Open Source development model works !** 



The European Synchrotron

#### SILX TOOLKIT

- Public project hosted at github
   https://github.com/silx-kit/silx
- Continuous testing

Linux, Windows and MacOSX

- Nightly builds
  - Debian packages
- Weekly meetings
- Quarterly releases
- Code camps before release
- Continuous documentation

http://www.silx.org/doc/silx/

## Conclusion: need a similar effort for web applications



-----

License and copyright information

The European Synchrotron

ESRF

#### **USE CASE – RECIPROCAL SPACE STRAIN SCANNING**

- Kmap strain scanning
- Experiment = 10 minutes
- Sample size = 100x100 um
- Resolution = 100 nm
- Data reduction = 10 hrs
- Data analysis = bottleneck for users i.e. needs specialist

Solution : hire data scientist for 18 months to rewrite scientists application





#### Figure 8

(a) Two-dimensional map distribution of the tilt magnitude (°) and (b) directional representation (arrows) of tilted  $\{001\}$  atomic planes with respect to the sample surface normal.

#### **USE CASE – RECIPROCAL SPACE STRAIN SCANNING**

- Financing = IRT-Nanoelec
- File format = HDF5
- Optimise = legacy code
- Execution = 2 minutes
- Speedup = 300 times !
- Next step = make available remotely as a service
- Long term = new program is based on SILX and therefore easier to maintain

Scientist : « Amazing - I have never seen my data like this ! »

#### **Conclusion:** hire more data scientists to rewrite scientific application + do similar effort for web applications





- Life sciences single crystal diffraction, and (soon) cryo electron microscope
- **Spectroscopy** small to medium sized data sets, lots of modelling (DFT calculations)
- Small Angle Scattering medium data sets but lots of modelling (Molecular Dynamics)
- **Coherent diffraction** large data sets + iterative processing
- Tomography currently the main data producer (<80%) has the highest needs in terms of data processing
- Cultural heritage: archeology, paleontology

#### Fast synchrotron tomography at KIT



- Fast high-throughput synchrotron-based X-ray tomography for material science, biology & paleontology
- ASTOR virtual infrastructure (astor.kit.edu)
  - Virtual analysis infrastructure based on cloud technology
  - Remote access for external users
  - Automated segmentation tools

# gy Contraction of the second sec

Reconstructed anatomy of a fossil beetle

### Example: Digitalization of 30 million-year-old phosphatized insects

- Collections from the Natural History Museums Basel & Stockholm
- Datasets scanned in one week: 1408
- Data created: ~ 79.2 TB (without post-processing!)
  - ∎ raw data: ~36.3 TB
  - reconstructed tomograms: ~ 42.9 TB



Schwermann et al. 2016: Preservation of three-dimensional anatomy in phosphatized fossil arthropods enriches evolutionary inference. *eLife* 5: e12129

- What about Data Policy and PaNs ?
- PAN RIs had no open data policy prior to 2011
- PANDATA-EUROPE (FP7) deliverable produced a generic open data policy (http://wiki.pan-data.eu/imagesGHD/0/08/PaN-data-D2-1.pdf)



- PaN-Data Open Data Policy now accepted
- First adopted by neutron sources :
  - ISIS and ILL
- Then adopted by photon sources :
  - ELETTRA, ESRF, HZB, PSI
- Other sources following :
  - ALBA, DESY, XFEL, ...
- Just in time for Springer + Nature journals



#### **EXAMPLE = ESRF DATA POLICY**

- ESRF is custodian of data and metadata
- ESRF to collect high quality metadata to facilitate reuse of data
- ESRF will keep metadata forever
- ESRF will keep raw (or reduced) data for 10 years
- Data will be registered in a data catalogue (ICAT)
- Data will be published with a Digital Object Indentifier (DOI)
- The experimental team has exclusive access to data during the embargo period (3 years which can be extended on request)
- Data will be made public after the embargo period under licence CC-BY 4.0
- Data Policy will be implemented on all beamlines by 2020



#### • Volume

- 100s of Petabytes for each RI over 10 years
- Cost
  - Hundreds of thousands euros per year
- Metadata
  - Big effort to define metadata (years) but can be shared
- Open access
  - Scientists are encourage by Journals + H2020
- Reusability depends on metadata



#### **PAN-DATA LONG TAIL OF DATA**

Many data sets are medium to small



 $\textbf{IAAS} \rightarrow \textbf{PAAS} \rightarrow \textbf{SAAS} \rightarrow \textbf{DAAS}$ 

• How to go from Infrastructure to Platform to Software to Data Analysis as a Service ?





#### SETTING UP PRIVATE CLOUD MODEL



#### PANDAAS

• Photon and Neutron Data Analysis as a Service is a common proposal between all synchrotrons + neutron sources in EU (not funded)





#### PANDAAS GOES AHEAD

- The need for Data Analysis as a Service has not gone away

   instead it has got stronger
- Photon and Neutron sources are working on it with own resources
- ESRF continues to play the rôle of coordinator for PANDAAS
- Organised followup meetings :
  - July 2016 @ ESRF
  - December 2016 @ ESRF
  - ESRF to present PANDAAS @ e-IRG meeting next week
  - Activities reported here: http://pan-data.eu/node/103
- Currently active tasks :
  - Packaging scientific software with Docker containers
  - Setting up and giving users access to private clouds
  - Testing commercial cloud services (AWS, HNSciCloud, ...)



#### **TESTING HPC ON AMAZON CLOUD**

- Amazon provides HPC resources  $\rightarrow$  why not use them ?
- Check out https://aws.amazon.com/hpc/
- AWS <u>cfncluster</u> script allows us to create an HPC cluster in 10 minutes with pre-installed MPI, fast interconnect, batch scheduler, ...





#### **TESTING HPC ON AMAZON CLOUD**



• Results of tests :

Software	Intel Cpu ESRF	Intel Cpu AWS	elapsed	elapsed	time ratio	theory speed	cost AWS
	clock, #cores	clock, #cores	time ESRF	time AWS	ESRF / AWS	AWS / ESRF	
FDMNES	E5-2680	E5-2666 v3	17 h 14 min	24 h 57 min	0.69	0.70	26.73 \$
	2.7 GHz, 16	2.9 GHz, 8	(1034 min)	(1497 min)		(0.5 * 1.07 * 1.3)	(25.0 * 1.069)
FDMNES	as above	E5-2666 v3	as above	20 h 17 min	0.85	0.87	43.40 \$
		2.9 GHz, 10		(1217 min)		(0.625 * 1.07 * 1.3)	(40.6 * 1.069)
Geant4	E5-2680	E6-2670 v2	2 h 05 min	3 h 42 min	0.56	0.60	5.92 \$
	2.7 GHz, 16	2.5 GHz, 8	(125 min)	(222 min)		(0.5 * 0.93 * 1.3)	(3.7 * 1.60)
Geant4	as above	E6-2670 v2	as above	1h 54 min	1.10	1.21	6.08 \$
		2.5 GHz, 16		(114 min)		(1.0 * 0.93 * 1.3)	(1.9 * 3.201)
Quanty	E5-2680	E5-2666 v3	5 h 40 min	7 h 38 min	0.74	0.70	2.03 \$
	2.7 GHz, 4	2.9 GHz, 2	(340 min)	(458 min)		(0.5 * 1.07 * 1.3)	(7.6 * 0.267)
Quanty	E5-2680	E5-2666 v3	4 h 33 min	5 h 32 min	0.82	0.70	2.94 \$
	2.7 GHz, 8	2.9 GHz, 4	(273 min)	(332 min)		(as above)	(5.5 * 0.534)
Quanty	E5-2680	E5-2666 v3	5 h 09 min	5 h 33 min	0.93	0.70	5.99 \$
	2.7 GHz, 16	2.9 GHz, 8	(309 min)	(333 min)		(as above)	(5.6 * 1.069)
Quanty	as above	E5-2666 v3	as above	6 h 26 min	0.80	1.60	13.68 \$
-		2.9 GHz, 18		(386 min)		(1.125 * 1.07 * 1.3)	(6.4 * 2.138)

Remarks:

1) the ESRF computer have 1 thread / core, the AWS comuters have 2 threads / core (i.e. hyperthreading)

2) theory speed: cores(AWS / ESRF) \* clockrate(AWS / ESRF) \* (estimated speedup hyperthreading)

3) the second FDMNES run on AWS used 10 cores of 2 nodes with 8 cores each

4) the nodes used in all Quanty runs at the ESRF have 16 cores, but only part of them were used

5) AWS prices are "on demand prices" for the region "EU Frankfurt", excluding taxes



#### **TESTING HPC ON AMAZON CLOUD**



- Conclusions from AWS cloud tests
  - Cloud is adapted for HPC applications
  - AWS HPC offer is easy to use and adapted to our needs
  - Could be used to absorb peak demand for HPC @ ESRF
- Issues still to solve
  - How to optimise best price
  - How to manage access + budget / user
  - How to transfer large volumes of data to cloud
  - How to make it even easier for our non-cloud savvy users e.g. portal
- Hype about cloud is not just hype ...
  - See Viewpoint in Cern Courier « End of steam age of computing » by Eckhard Elsen CERN director for research and computing http://cerncourier.com/cws/article/cern/65819



#### SCIENTIFIC NOTEBOOK PORTAL FOR SCIENCE

- Web noteboooks provide easy access to data + algorithms for users familiar with scripting e.g. Python
- « UGA 2.0 la science augmentée » project submitted to IDEX

	CJUPYTET Lorenz Differential Equations (autosaved)	-
	File Edit View Insert Cell Kernel Help	Python 3 O
	E + ≫ 2 K + ↓ ► ■ C Code + Cell Toolbar: None +	
	Exploring the Lorenz System	
	In this Notebook we explore the Lorenz system of differential equations:	
	$\dot{x} = \sigma(y - x)$	
File Edit View Insert Cell	$\dot{y} = \rho x - y - xz$	
B + ≫ ℓ3 B + ↓ ▶ I	$\dot{z} = -\beta z + xy$	
C jupyter	This is one of the classic systems in non-linear differential equations. It exhibits a range of complex behaviors as the parameters $(\sigma, \beta, \rho)$ are varied, including what are known as <i>chu solutions</i> . The system was originally developed as a simplified mathematical model for atmospheric convection in 1963.	aotic
	<pre>In [7]: interact(Lorenz, N=fixed(10), angle=(0.,360.),</pre>	
Welcome to the	$\sigma = (0.0, 50.0), \beta = (0., 5), \rho = (0.0, 50.0));$	
This Notebook Server way	× angle 308.2	
This Notebook Gerrer Wa	max_time 12	
WARNING	σ10	
Don't rely on this serv	β2.6	
Your server is hosted than	p 28	
Run some Python		
To run the code below:		
1. Click on the cell to se		
2. Press SHIFT+ENTER		
A full tutorial for using the		
In [ ]: <pre>tmatplotlib inline</pre>		
<pre>import pandas as pd import numpy as np</pre>		



- Organise yourselves !
  - PaN-Data collaboration (http://pan-data.eu)
- Elaborate on and drive their e-infrastructure needs
  - PaN-DaaS project (but lack of funds)
- Participate in the innovation of e-Infrastructure services
  - Participate in HNSciCloud, AWS, e-IRG ...
- Contribute to standards and take care of their data
  - Adopting Open Data Policies, Nexus/HDF5

- EOSC is proposed to be a federation of services , but
  - not clear what services will be useful for Photon and Neutron source users
- PaN-Data trends :
  - Large data volumes + Open data policy will mean data stays at the sources in the future
  - Compute power to be at sources i.e. most PaN RI's will set up private clouds
  - Assist users to move data+software to commercial/public cloud where possible
  - Need domain specialists to (sup)port software



#### **PAN-DATA COMMONS IS HAPPENING**

- Common users
- Common technology
- Common techniques
- Common software
- Common (meta)data
- Common catalogues
- Common data policy
- Common data analysis service
- Common e-infrastructure needs

each for software           cach for software <th>Add new software</th>	Add new software
EPAC CASE OF A CONSTRUCT ON A CONSTR	priori knowledge capabilities. PyHST2 terration synchrotron facilities (10 terabytes defauti filtered backprojection rrove the reconstruction quality or in order to te techniques are based on the total variation
is database can be freely consulted. It gives an overview of software available for neutron and one experiments and their use with respect to instruments at experimental facilities. Tregistering and logging-in new software can be entered and it will appear in the database after construction in the software in the database after construction of the week software in the database after construction with the provide the software in the database after construction with the software in the softwar	priori knowledge capabilities. PyHST2 terration synchrotron facilities (10 terabytes default filtered backprojection rove the reconstruction quality or in order to te techniques are based on the total variation
registering and logging-in new software can be entered and it will appear in the database after deration. Recently updated software PyHST2 World distributed code for high speed tomographic reconstruction with iterative reconstruction and a priori knowledge capabilities. Formerly known as PyHST) has been engineered to sustain the high data flow typical of the third generation synchroton Ladillace er experiment) by adopting a distributed and pipelined architecture. The code implements, beside ad edual tillered backprojection construction, iterative reconstruction techniques with a priori knowledge. The latter are used to improve the reconstruction qualit deuce the required data volume and reach a given quality goal. The implemented a-priori knowledge techniques are based on the enalisation and a new recently found convex functional which is based on overlapping patches. SASTIL MSRI has been written for analyzing and plotting small angle scattering data. It can calculate integral structural parameters liker a rytation, scattering invariant, Prood constant. Furthermore it can fit size distributions together with several form factors including at (diobal fit helps to determine fit parameters unambiguously which by analyzing a single curve would be otherwise strongly correlate (GSULIDI MAX) GSULIDI MAX is high-end software for the visualization and analysis of CT data in combination with the optional add-on module	I priori knowledge capabilities. PyHST2 teration synchrotron facilities (10 terabytes idefault filtered backprojection rove the reconstruction quality or in order to a techniques are based on the total variation
Person Active and the software softwar	priori knowledge capabilities. PyHST2 eration synchrotron facilities (10 terabytes default filtered backprojection rove the reconstruction quality or in order to e techniques are based on the total variation
PyHST2  PyHST2  PyHST4  PyHSt	I priori knowledge capabilities. PyHST2 erration synchrotron facilities (10 terabytes I default filtered backprojection rave the reconstruction quality or in order to a techniques are based on the total variation
PyHST2 bybrid distributed code for high speed tomographic reconstruction with iterative reconstruction and a priori knowledge capabilities. formerly known as PyHST2 has been engineered to sustain the high data flow typical of the third generation synchroton tacilities are experiment) by adopting a distributed and pipelined architecture. The code implements, beside a default filtered backprojection construction, iterative reconstruction techniques with a priori knowledge. The latter are used to improve the reconstruction qualit educe the required data volume and reach a given quality goal. The implemented a-priori knowledge techniques are based on the evailation and a new recently found convex functional which is based on overlapping patches. SASIII SASII has been written for analyzing and plotting small angle scattering data. It can calculate integral structural parameters like ra syration, scattering invariant, Porod constant. Furthermore I can fit size distributions together with several form factors including d is logal if helps to determine fit parameters unambiguously which by analyzing a single curve would be otherwise strongly correlate (Soludio MAX is a high-end software for the visualization and analysis of CT data in combination with the optional add-on module	I priori knowledge capabilities. PyHST2 teration synchrotron facilities (10 terabytes I default filtered backprojection rove the reconstruction quality or in order to a techniques are based on the total variation
ASTIR has been written for analyzing and plotting small angle scattering data. It can calculate integral structural parameters like ra ASTIR has been written for analyzing and plotting small angle scattering data. It can calculate integral structural parameters furdure factors. Additionally an algorithm has been implemented, which allows to simultaneously fit several scattering curves with (global) parameters. This late points in especially important in contrast variation experiments or measurements with polarised n lobal fit helps to determine fit parameters unambiguously which by analyzing a single curve would be otherwise strongly correlate <b>(GStudio MAX)</b> (GStudio MAX is a high-end software for the visualization and analysis of CT data in combination with the optional add-on module	
/GStudio MAX /GStudio MAX is a high-end software for the visualization and analysis of CT data in combination with the optional add-on module	I structural parameters like radius of sveral form factors including different several scattering curves with a common set easurements with polarised neutrons. The be otherwise strongly correlated.
GStudio MAX is a high-end software for the visualization and analysis of CT data in combination with the optional add-on module	
Aeasurement", 'Nominal/Actual Comparison', Porosity/Inclusion Analysis', 'Wall Thickness Analysis', 'Fiber Composite Material Ana mport (with PMI)',	h the optional add-on modules 'Coordinate Fiber Composite Material Analysis' and 'CAD
/ASP	



#### • E-IRG Roadmap

- Good proposals but very high level
- e-infrastructures
  - No interaction so far (except GEANT indirectly),
  - Not sure how scalable e-infrastructures are to cater for PaN-Data community and all the others
  - Easier to get laaS from commercial companies
  - What could e-infrastructures provide that commercial companies don't and which Ris need ? Expert advice + help with software development

## No mention of PaN-Data community in roadmaps / white papers



- PaN-Data community is adopting best practices in data management to create a Data Commons and enable Open Science
- Data analysis of synchrotron and neutron data is a major bottleneck for users (TBs to PBs)
- Providing Data Analysis as a Service (PaNDaaS) is still a strategical issue, work continues very slowly but PaN community needs funds for developing DaaS
- e-infrastructures are not seen as service providers but could be advisors on how to setup e-infrastructures, hybrid clouds and develop SaaS/DaaS

