Pangeo community platform and its use at CNES

BiDS 2019, Guillaume Eynard-Bontemps, CNES/Pangeo







Mission

To cultivate an ecosystem in which the next generation of open-source analysis tools for the geosciences can be developed, distributed, and sustained.



Problems

- Data volume crisis in (geo)sciences
- Software multiplication, non reproducibility
- Many copies of the same datasets
- Local vs HPC vs Cloud
- Technology gap: industry vs academia

Goals/vision

- Foster collaboration around the open source Scientific Python ecosystem:
 - open and collaborative development
 - Welcoming and inclusive culture
- Support the development with domain-specific (geo)science and transverse packages
- Improve scalability of these tools to handle gigabytes to petabytescale datasets

Pangeo ecosystem



- Set of tools that will facilitate science at all scales
- Platform agnostic
- The core of the Pangeo ecosystem includes:
 - Xarray (data-model and toolkit for working with N-dimensional labeled arrays)
 - **Dask** (parallel computing)
 - Jupyter (interactive computing)
- Extensible: Series of 3rd party packages that build on top of core libraries
- Flexible: Individual components may be swapped in/out



Examples of 3rd party packages in the Pangeo Ecosystem:

- Data discovery
- Regridding and GIS
- Vector calculus
- Signal processing
- Thermodynamics



BUILD YOUR OWN PANGEO

Storage Formats	HJF	OPeNDAP	Cloud Optimized COG/Zarr/Parquet/etc.
ND-Arrays	NumPy	DASK	More coming
Data Models	xarray	S Iris	$\begin{array}{c} pandas \\ y_i t = \beta' x_{it} + \mu_i + \epsilon_{it} \\ \hline \bullet \\ \bullet \\$
Processing Mode	Jupyter Interactive		Serverless
Compute Platform	HPC	Cloud Google Cloud Platform	Local

Pangeo vs state of the art

VS



Mature

Robust

JVM/Python

Query optimized

Collections & Dataframes

Python overhead

For big tabular data

Hadoop/Cloud/HPC

DASK

Less Mature Pretty strong Python only Science optimized

Collections, DF, Arrays, Futures...

Python only

For science data

Hadoop/Cloud/HPC



Pangeo public cloud deployment



hub.pangeo.io

pangeo.binder.io

JupyterHub/BinderHub running on the Google Cloud

- Kubernetes for both Jupyter and Daskdistributed
 - Dask-kubernetes
- Exploring/evaluating:
 - Cloud storage
 - User environment customization
 - Data discovery

0/1

0/1

0/1

0/1

0/1

1/8

1/8

0/3

0/3

anmax-partia

nanmin-nanmi

nanmax-aggre.

nanmax-nanma

- Kubernetes Helm-chart (github.com/pangeo-data/helm-chart)
- CI/CD with Hubploy and CircleCI
- Deployments exist on AWS and Azure.

CNES Datacenter overview

HPC (HAL)

- 300Tflops
- 380 batch servers / 8400cores
- 4 interactive servers pre/post processing w/ GPU
- 6,2 PB GPFS / 200TB burst buffer/ 50GBs bandwidth
- Low latency network
- GPGPU Nvidia Volta V100

HPC DRSF (Ktulu)

- 20 Tflops
- 2 interactive servers pre/post processing w/ GPU
- 24 servers / 576 cores
- 120TB GPFS
- Low latency network

HPC usecases in CNES



Two main kinds of processing

Numerical simulation (HPC)

- Upstream phase, R&D
- Highly optimized technics
- Fine grain parallelism

Trends : multiscale, multiphysics





- Downstream phase, operation
- Sensors data \rightarrow scientific data
- Coarse grain parallelism

Trends : data volume explosion



Pangeo at CNES

- JupyterHub and notebooks for • interactive computing
 - Hub on a VM with qsub access •
 - Batchspawner, Wrapspawner •
- dask.distributed: parallel workers across many HPC nodes
- Xarray for computational toolkit and I/O

 $\leftarrow \rightarrow$ Q Les plu

jup Filo 5 ÷

- New tool for deploying dask clusters on HPC: dask-jobqueue
 - Start a cluster from a notebook
 - Interactive (or not) distributed computing
 - Auto scaling capabilities

		In [4]:	cluster										
		and a second sec	PBSCI	uster									
			Workers 2 Cores 24 Memory 120.00 GB		Manual Scaling Adaptive Scaling								
					Minimun	n 0 🖭	Maximun	n 100 🖭	Adapt				
			Dashboard: /u	C Sabboard: /user/eynardbg/proxy/8787/status						,			
	τ¢.	Status	Workers	Tasks	System	Profile	Graph	Info			×		
	CPU Us	e (%)											
	P	- 11 AD											
	Memory	y Use (%)											
es plus visitės 🗖 🛛	Débuter avec Firefox 🛅 CNES 🛅 Pe	150 🔘 0.PAGE 1	D'ACCUEIL AC.	Technique	n, C) Ξ	← → C ^a Q Les plus visit	tés 🛄 Débuter	evec Firefox	45 Ell Perso O	0.PAGE D'ACCUEIL	🖸 Tec	hnique 🖽 IRT	
es plus visités 🖬 (lupyter da Edit Vice + 🔉 🎗 🕅	Débuter avec Firefox ≧CNES È Pe ask_local_vs_cluster_exa r Insart Cell Kernol b ↑ ↓ N Run ■ C ₩	eno 🕲 DPAGE I amples Widgots H	ND MCCUELAC.	Technique	III III E III t Control Panel I dask_rt @	 ← → Q^a Q Les plus visit Ø Stat Bytes stored: 	 	evec Firefox 🛗 Cl evec Firefox 🛗 Cl ers Tasks	IS Perso O System Ta	OPAGE D'ACCUBIL Profile Gi sks Processing	AC- ETTec	₩ ₩ hnique IRT fo	
es plus visités 🖬 (jupyter da Edit Vion + 😹 😢 🗓	Debuter avec Firetox CNES Pre ask_local_vs_cluster_exa r Insert Coll Kernol b + + H Run C H	mo 😧 DPAGE 1 amples Widgots H Code dt.ypes=("ne	Np i i i i i i i i i i i i i i i i i i i	Technique	IT Control Panel I Gask at	← → Q ^a Q Les plus visit Ø Les plus stored: Bytes stored: 0,0	02 tes ■ Déouter tus Work 44.90 GB	ers Tasks	System	Profile G	AC. The raph In	hnique [2] IRT	
Les plus visités de la jupyter de jupyter de la Visité de la Visité de la Visité de la companya	Debuter avec firefox CAUS Pre- ask_local_vs_cluster_exa r inser: Cell Kernol b • • • H Run C Pre- decoupe les donnees, puis calcule re on affiche que la head, on ne v	eno DAGE amples Widgots H Code dt.ypes=(*ne e une moyerne a pas calculer t	ND PACUELAC.	Technique III I Logour Trusted : int, *x*	IR CD = IFT I Control Panel I dask, st @ : float, 1 > xt y.	← → C Q Les plus visit Ø Stal Bytes stored: R ₂ Task Stree	CD tes Débuter tus Work 44.90 GB 57 _{20 469} ¹ 7 am	ers Tasks	ES EN Prio Consultation of the second	Auser open	AC. The Technology of the AC.	hnique 🛅 IRT fo	0
Les plus visités III (JUPYter da Edit Visit + 3c 2 II On re Comm In [6]: df[[Out[6]:	Debuter avec firefox ■ CNES ■ Re ask_local_vs_cluster_exa r insert Coll Kernol b ◆ ◆ H Run ■ C → deecupe les donnees, puis calout re on affiche que la head, on ne v ***, 'y']].resaasple ('lv')	eno DPAGE amples Widgots H Code d ypes=("nac e une moyenne a pas calculer t .mean () .hea	DPACCUELAC.	Technique	IN CD = IFT I Control Parel I dask, rt @ : float, * * Xy.	← → C [*] Q tespin visk Ø tespin visk Ø Stal Bytes stored: b ₀ Task Streit	CD teis CD Débuter tus Work 44.90 GB 57 _{20 Ats} 7 am	vec Firefox ⊞ C ers Tasks	ES ≧ Perio System Ta ¹ 20 C _{EB} 0	Anase processing	AC_ The rephonent		
Les plus visités III : jupyter da Edit View + 3x 2) 10 c On re Comm In [6]: df[[Out[6]: time:	Debuter avec firefox ■ CNES ■ Pe ask_local_vs_cluster_exa r insert Cell Kernol b ◆ ◆ H Run ■ C H decoupe les connece, puis calcule ne on affiche que la head, on ne v 'st', 'y']].recasspile ('iu') x y temp	eno DAAE amples Widgots H Code dtypes=(*ne e une moyenne a pas calculer t .mean () .bea	DPACCUEL AC.	Sa Technque III Logout Trusted : Int. *x*	IN CD =	← → C [*] Q Les plus visit d Stati Bytes stored 0,0 0,0 Task Street	Co micis Dobuter trus Work 44.90 GB 5720 Leg	Φ http://juanet.com/picture avec Firefox Cl ers Tasks	IS Pro O	Anase/opensi opAge Draccues Profile G sks Processing	AC_ Tree		
et plus visites and jupyter di Edit View + ac 2 10 On re Comp In [6]: df[(out[6]: time: accord	Obecuter avec firefox CNES Dec ask_local_vs_cluster_exa imset Coll Xernol imset Coll Xernol C H decoupe les sonnees, puis calcut ne on affiche que la head, on ne v ''''', '''']], reclasspie (''I'') x Y x Y Y Y Y calcut de gue la head, on ne v ''''', 'y'']], reclasspie ('I'u'') X Y x Y Y Y Y calcut de gue la head, on text Y Y x Y Y Y	Widgots H Code Code Code cure moyenne a pas calcular f mean () .hea	DPACCUEL AC.	S2 Technque III Logou Truster xstramotros x e	INT CD = FT Control Parent Control Parent	← → C [*] Q tesplus visit Ø Stat Bytes stored: k ₀ Task Street	Co miss Dobuter tus Work 44.90 GB	ers Tasks	In the second se	Under/opmin	AC. □ Tec raph In	hniqor 🛅 IRT to	
es plus visites juppter di E tai Verent + 3c Qi R C On re Cons In [6]: df[[2000- 2000- 2000- 2000-	Cecuter avec firefox ■ CNES ■ Pe ask_local_vs_cluster_exa insert Col Kernol → ◆ ♥ H Run ■ C → cecoupe les donnees, puis calcule ne on affiche que la head, on ne v ***, 'v']).reaample ('14') x y tamp 22.36 0.000455 0.000451 62.31 0.0007/4 0.000453 62.36 0.000453	Widgots H Code Code Code Code cune moyenne a pas calculer + mean () .hea	practure ac.	Technique 1	IN C) = #7	← → C [*] Q tespins visit Ø Stat Bytes stored: Lo Task Stree	CD hes CD Debuter tuss Work 44.90 GB	eres Tests avec Fields Π Cl eres Tasks	13 17 19 19 19 19 19 19 19 19 19 19 19 19 19	Uters/spins	AC. Tree	Minique 12 HT fo	•
es plu violes Las lu violes Est Vione + 3c 2, 10 c On re Corre Corre 2006 2	Obeouter avec firefox CAUS Dec ask_local_vs_cluster_exa insert Coll Xernol insert Coll Xernol C H obeoupe les connees, puis calcut ne on affiche que la head, on ne v * * Y x Y * Y * e206 0.001665 0.00461 * Y x Y * Y * e206 0.001665 0.004651 60.004651 60.004651 62-30 0.002166 0.006453 60.00455 60.00455 62.001655 60.00455	eno O DAGE I amples Widgots H Code Code Code cune moyerne a pas calculer t mean () .hea	alp alp alp alp all all all all all all all all all all	ST ST Technique 11 S Cogoou Trusted : Lint, *x*	IN C = FT Control Pavel i Gask, rt i Gask, rt i float, - i f	C +> C Q Les plus visit Star Bytes stored L ₂ Task Stree	CD hes CD Debuter tuss Work 44.90 GB 51 _{200 Map}	e πηρεζημ svec Fields fill Cl errs Tasks	In the second s	Uters/spans	AC Tree	Anique III fo	
es plu visités Las Universités Las Verse + 32 02 10 10 0 nr e 0	CAUS	ere O DAGE amples Widgots H Code disysteme (*ne pass calculer mean ()hea	In the second se	rechnique Internique In	IN C = FT Control Paral Control Paral Co	C	GD Hels GD Debuter Hutus Work 444.90 GB Stoppe Add Stoppe Addd Stoppe Add Stoppe Add Stoppe Addd Stoppe Addd Stoppe	ers Tasks	System Provide State	Unserverse	AC. The Tee	1 μηματο (1) μηματο (• • •
Est Voor Est Voor Est Voor * 30 (10) Corr Corr Corr Corr Corr Corr 2006 2	Chebuter avec firefox ■ CNES ■ Pe ask_local_vs_cluster_exa r msart Col Kernot b ◆ ◆ H Run ■ C H decoupe les donnees, puis calcule ne on affiche que la head, on ne v 'st', 'y']].recample ('lu') x y temp 22.06 0.001665 0.000451 62.26 0.001665 0.000451 62.27 0.00166 0.000453 62.27 0.0016 0.000453 62.27 0.0016 0.000453 62.27 0.0016 0.000453 62.27 0.0016 0.00045 62.27 0	ero O DAGE 1 amples Widgots 19 Code Co	do do menti atr. 10 por semano des p ar tout le classet. d() ut donc calouler su porte ()	r tout le datas	IN C = FT Control Parallel dask.rt • : float, t • dask.rt • state of the state of th	C - S C Les pius visit C Les pius visit Portes stored Reg Tank Stree 10	s 15	erec Friedca III of erec Friedca III of erec Tasks	to be the second secon	oracle practure of a second seco	AC_ Tree	10 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
ites plux visités III jupyter dia a Eat Verente a Eat Verente c Control Control Dott (6): III Control Inn (6): diff Control Dott (6): IIII Control Inn (1): IIIII IIII	Consert review CAUS IP Reask_local_vs_cluster_exa resert Col Kernel consert Col Kernel consert consert Col Kernel consert consert Col Kernel consert co	ero O DAGE amples Widgots H Code dt ypes= (*ne a pas calculer t .mean () . bea	ap ap ap ap ar 1 atr, 1 atr, 1 atr por semanc de p ar tout le classet. d () ut donc calculer su pute () .pics ()	r tout le datas	IN C = FT	C C	s 199	exec Fuetos El σ eres Tasks	System System Secan o 176 cm o	Undergrand	AC_ Tree	10 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	* O
te plu viole	Cheloter avec firefox CAUS Pre- ask_local_vs_cluster_exa r insert Col Kennet b + H Run E C H decoupe les donneles, puis calcule re on affiche que la head, on ne v ***, 'y']].resample ('io') * * *** * *** * **** * *	me O DAGE 19	by the the the second secon	r tout le datas	IN C = #7	C de pur via C tes pur via Pytes stored by tes Task Stree 55 Progress - tes 15 Progress - tes 15 Progress - tes 15 15 15 15 15 15 15 15 15 15	GD Hels C Debuter Hels C Debuter 44.90 GB Store Man Store Man	emory: 326, proc	¹ ¹ ¹ ¹ ¹ ² ² ² ² ² ² ² ²	Viele/generation	AC Tree	Maria (1997) 160 197 (2000) 197 (2000)	
ter plu visites III jupyter da a E.31 Visit a E.31 Visit a E.31 Visit c C C On re Correr C In [61: df[[Correr 2000 2000 2000 2000 2000 2000 Ion [7]: temper 2000	Debute avec firefor ■ CNES ■ Pe ask_local_vs_cluster_exa r insar: Col Kennel b ◆ ◆ H Run ■ C → decoupe les donnels, puis calcute ne on affiche que la head, on ne v 's', 'y']].resample('lu') x y temp 22.0 000160 0.00083 62.20 0.00160 0.00083 62.21 0.000160 0.00083 62.22 0.00086 0.00083 62.27 0.00086 0.00085 62.28 0.00086 0.00085	ere O DAGE amples Widgots H Code Code cure moyerne a pas calculer t .mean () . com	to to to to to to to to to to	r tout le datas	IN C = #77 #7 Control Plant	C de pour visit C tes pour visit C tes pour visit De Stat De S	GD Dester tes GD Dester 44.90 GB 51 ₂₀ 4 ₆₀ 5 st ₂₀ 4 st	eres Friedox III of eres Tasks	System System System 1 Seque 1 Secue S	Undergrand	AC The second se		

Demo : Dask and dask-jobqueue basic example





Demo : Multi temporal NDVI with Xarray

	A https://jupyterhub.sis.cnes.fr/user/eynardbg/notebooks/git/rt-nouvelles-technos-distrib/notebook/sentinel-ndvi-visu-multitemporal.	pynb#	▣ … ☺ ☆	⊥ m/ @ =	ð
🌡 Débuter av	ec Firefox 🗎 CNES 🗎 Perso 🔀 0.PAGE D'ACCUEIL AC 🛅 Technique 🛅 IRT	_			X ser
	CJUPYTEY sentinel-ndvi-visu-multitemporal Last Checkpoint: il y a 28 minutes (unsaved changes)	÷	Logout Control Panel		
	File Edit View Insert Cell Kernel Widgets Help	Truste	ed Pangeo 2019-02 O		
	🖹 🕂 🔀 🔁 🛧 🔸 M Run 🔳 C 🗰 Markdown 🔍				
	datasets = []				
	for tiff in tif files:		akowave		
	try: print('loading'_tiff)		aneways.		
	ds = create_multiband_dataset(tiff)				
	datasets.append(ds) except Exception as e:	•	Interactiv	e processino	n over ten
	<pre>print('ERROR loading, skipping acquisition!') print(e)</pre>				
	20180215-105300-871 122 73170.1 D V1-5/SENTINEL28 20180215-105300-871 122 73170.1 D V1-5 FOF B	-	series wit	n Jupyter/Da	ask/xarra
	loading zip:///datalake/S2-L2A-THEIA/31TCJ/2018/02/25/SENTINEL2B_20180225-105018-458_L2A_T31TCJ	<mark>_</mark>			
	20180225-105018-458_L2A_T31TCJ_D_V1-5/SENTINEL2B_20180225-105018-458_L2A_T31TCJ_D_V1-5_FRE_B loading zip:///datalake/S2-L2A-THEIA/31TCJ/2018/02/27/SENTINEL2A_20180227-104236-560_L2A_T31TCJ		Intoractiv		h
	20180227-104236-560 L2A T31TCJ D V1-5/SENTINEL2A 20180227-104236-560 L2A T31TCJ D V1-5 FRE B		meraciiv		
	20180302-105023-464 L2A T31TCJ D V1-5/SENTINEL2A 20180302-105023-464 L2A T31TCJ D V1-5 FRE B				
	loading zip:///datalake/S2-L2A-THEIA/3ITCJ/2018/03/07/SENTINEL2B_20180307-105016-460_L2A_T3ITCJ 20180307-105016-460_L2A_T3ITCJ_D_V1-5/SENTINEL2B_20180307-105016-460_L2A_T3ITCJ_D_V1-5_FRE_B	-	Recompl	ite unon use	r Action
	loading zip:///datalake/S2-L2A-THEIA/31TCJ/2018/03/09/SENTINEL2A_20180309-104022-069_L2A_T31TCJ 20180309-104022-069_L2A_T31TCJ_D_V1-5/SENTINEL2A_20180309-104022-069_L2A_T31TCJ_D_V1-5_FRE_B	_	Rooompe		
	loading zip:///datalake/S2-L2A-THEIA/31TCJ/2018/03/14/SENTINEL2B_20180314-104014-461_L2A_T31TCJ	<mark>u</mark>			
	20180314-104014-461_L2A_T317C3_D_V1=//SEMTINEL2B_20180314-104014-461_L2A_T317C3_D_V1=7_FRE_B loading zip:///datalake/S2-L2A-THEIA/317CJ/2018/03/15/SEMTINEL2A_20180315-110609-689_L2A_T317C3		Statistica	I analysis	
	20180315-110609-689 L2A_T31TCJ_D_V1-5/SENTINEL2A_20180315-110609-689_L2A_T31TCJ_D_V1-5_FRE_B loading zip:///datalake/S2-L2A-THELA/31TCJ/2018/03/17/SENTINEL2B_20180317-105355-924_L2A_T31TC;			,	
	20180317-105355-924 L2R T31TCJ D V1-5/SENTINEL2B 20180317-105355-924 L2R T31TCJ D V1-5 FRE B			O is not a d	and intera
	20180319-104019-464_L2A_T31TCJ_D_V1-5/SENTINEL2A_20180319-104019-464_L2A_T31TCJ_D_V1-5_FRE_B	-	JFEGZUL	is not a ge	
			nrocessir	na format	
	In order to concatenate all images in an xarray dataset, we need to get the date for each file. Those dates will be use as a dimension for the xarray dataset.		proceedin	gionnat	
				te to elevial/a	المعالية بالمعا
	<pre>In [23]: list_date = [filename.split('_')[1] for filename in tif_files]</pre>	•	Snare da	ta in cioud/d	Istributed
	<pre>import datetime t = 0</pre>		nrocaecir	a roady for	nat
	for date in list_date:		processii	ig ready ion	Παι
	<pre>list_date[1] = date[:4] + '-' + date[4:6] + '-' + date[6:8] + ' ' + date[9:11] + ':' + date[11: list_date[i] = datetime.datetime.strptime(list_date[i], "%Y=%m=%d %H:%M:%S.%f")</pre>	1.			
	<pre>i += 1 dico = {'date' : list date. 'tiff name' : tif files}</pre>				
	pdf = pd.DataFrame(data=dico)				
	pdi.dtypes			v.	



Conclusions

- Pangeo ecosystem greatily facilitates distributed computing and data analysis at scale
- It changes ways of doing it too
- Non monolithic platform built on top of existing Scientific Python stack and new related packages
- Community is always here to help
- Dask more versatile and easy to use than Spark.

Next steps

- Broaden users and use cases at CNES
- Encourage people to get in touch with Pangeo community
- Work in cooperation with others (Ongoing with Ifremer and CLS on SWOT aval data processing)

Pangeo wesite and discussions:

https://pangeo.io

https://github.com/pangeodata/pangeo/issues

https://medium.com/pangeo

Pangeo Example + Binder:

https://github.com/pangeo-data/pangeoexample-notebooks

http://binder.pangeo.io/v2/gh/pangeodata/pangeo-example-notebooks/master

Dask jobqueue:

https://github.com/dask/dask-jobqueue

Dask simple examples:

https://github.com/dask/dask-examples

My email

Guillaume.Eynard-Bontemps@cnes.fr