

CGNS/PYTHON & CGNS/HDF5 SPÉCIALISATION POUR LES DONNÉES EN CFD



Motivation

Interopérabilité des simulations CFD (Computational Fluid Dynamics)

- > Calculs longs et coûteux
- > Peu de données de grands volumes
- > Représentation des données adaptée aux besoins de performance

Objectifs

- Capitaliser les outils et les données
- > Reprendre ses calculs et ceux des autres
- > Partager avec des entités quelconques
- > Factoriser les pratiques
- > Déléguer des traitements



Etat des lieux - 1

Des grandeurs physiques et numériques CFD

- > Coordonnées de maillages
- > Connectivités de points et de blocs
- > Champs de valeurs
- > Ingrédients numériques

Des tableaux de scalaires

- > Réels simple et doubles, entiers simples et doubles
- > Peu ou pas de chaînes de caractères
- > Peu ou pas de structures typées

Codes de manipulation et d'échange travaillant sur des tableaux bruts

- > Fortran
- > MPI



Etat des lieux - 2

Pas d'interopérabilité

- > Chaque code a son format
- Mélange langage de description de données vs modèle de données
- > Traducteurs de formats internes en formats internes

Pas d'outils standards

- > Capture des formats propriétaires
- Très bas niveau de manipulation (binaire)

Pas de capitalisation

- > Le savoir faire reste dans les outils
- > Les données sont dupliquées/ traduites avec perte de contexte voire de précision



Initiative CGNS - Organisation

Groupe NASA/ Boeing/ Mc Donnell Douglas

- > Travaux sur des outils CFD et format CFF (1994-1995)
- > Développement du standard (1995-1998)
- > Création du groupe public (1999)

Steering Committee

- Membres votants pour le standard
- Le consensus avant tout
- > Demandes d'extensions hors SC possibles

Un standard désormais incontournable pour la CFD

Airbus ANSYS-ICEM CFD Boeing Commercial Cenaero Colorado State University **HDF** Group Intelligent Light NASA Langley ONERA Pointwise. Inc. Pratt & Whitney SAFRAN Tecplot, Inc. TTC Technologies University of Colorado University of Kansas



Initiative CGNS - Eléments du standard

Le document de référence

> CGNS/SIDS (Standard Interface Data Structure)

Représentation en arbre indépendante de l'implémentation

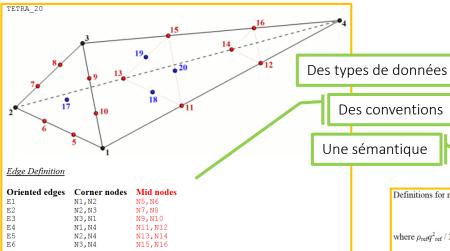
- > CGNS/ADF
- > CGNS/HDF5
- > CGNS/Python

Des exemples

- > Une librairie de manipulation CGNS/MLL
- > Des outils de manipulation, visualisation...



CGNS/SIDS - Le sens des données



Mid-face node Oriented edges

-E3,-E2,-E1

E1, E5, -E4

E2, E6, -E5

E3,E4,-E6

N17

N18

N19

N20

<pre>RigidGridMotion_t := {</pre>	
List(Descriptor_t Descriptor1 DescriptorN) ;	(0)
<pre>RigidGridMotionType_t RigidGridMotionType ;</pre>	(r)
DataArray_t <real, 2,="" 2]="" [physicaldimension,=""> OriginLocation; DataArray_t<real, 1,="" physicaldimension=""> RigidRotationAngle; DataArray_t<real, 1,="" physicaldimension=""> RigidVelocity; DataArray_t<real, 1,="" physicaldimension=""> RigidRotationRate;</real,></real,></real,></real,>	
List(DataArray_t DataArray1 DataArrayN) ;	(0)
DataClass_t DataClass ;	(0)
DimensionalUnits_t DimensionalUnits;	(0)
List(UserDefinedData_t UserDefinedDatal UserDefinedDataN) ; } ;	(0)

Definitions for nondimensional flowfield coefficients follow: the pressure coefficient is defined as,

$$c_p = (p - p_{\text{ref}}) / (\rho_{\text{ref}} q^2_{\text{ref}} / 2)$$

where $\rho_{\text{ref}}q^2_{\text{ref}}/2$ is the dynamic pressure evaluated at some reference condition, and p_{ref} is some reference pressure. The skin friction coefficient is, $c_f = \tau / (\rho_{\text{ref}}q^2_{\text{ref}}/2)$

where τ is the shear stress or skin friction vector. Usually, τ is evaluated at the wall surface.

The data-name identifiers defined for nondimensional governing parameters and flowfield coefficients are listed in the following table.

Data-Name Identifiers for Nondimensional Parameters

Data-Name Identifier	Description	Units
Mach	Mach number: $M = q/c$	-
Mach Velocity	Velocity scale (q)	L/T
Mach VelocitySound	Speed of sound scale (c)	L/T
RotatingMach	Mach number relative to rotating frame: $M_r = q_r/c$	-



N10,N9,N8,N7,N6,N5

N5,N6,N13,N14,N12,N11

N7, N8, N15, N16, N14, N13

N9.N10.N11.N12.N16.N15

Face Definition

N1,N3,N2

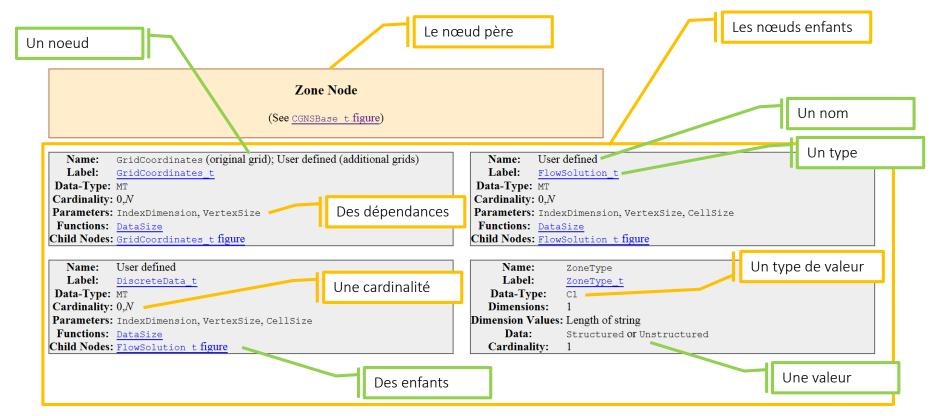
N1, N2, N4

N2,N3,N4

N3,N1,N4

Face Corner nodes Mid-edge nodes

CGNS/SIDS - Correspondance implémentation - 1





CGNS/SIDS - Correspondance implémentation - 2

Prérequis du support d'implémentation

- > Représentation en arbre
- > Nœud avec : nom, valeur, type
- > Référence à un noeud

Implémentation historique

> CGNS/ADF

Implémentations de référence

- > CGNS/HDF5
- > CGNS/python

Implémentations test

- > CGNS/XML
- > CGNS/SQL



Système ouvert - 1

Interfaces publiques

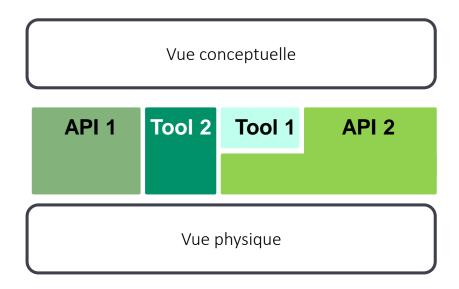
- > Consensus
- > Standard de facto vs Standard de juri
- Outils/Librairies adaptées aux applications

Implémentation de référence

- > Avec: la code fixe la pratique (danger)
- > Sans: standard virtuel
- > Plusieurs: idéal

Principes de base

- > Pas de module/ librairie/ classe spécialisée requise
- > Support de librairies/ services très utilisés faisant déjà l'objet d'un consensus
- > Spécifications publiques établies après discussion





Implémentations de référence

CGNS/HDF5

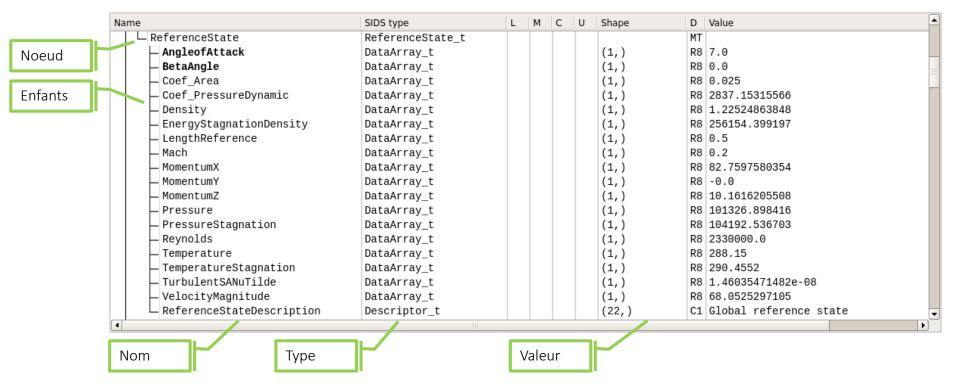
- > Archivage
- > Echange par fichier
- > Spécialisation des groupes/ dataset/ attributes

CGNS/Python

- > Très grande souplesse de manipulation
- > Structure de liste + numpy
- > Intéropérabilité en cours de calcul
- > Echange par buffer mémoire
 - Buffer protocol: mpi4py, numpy...

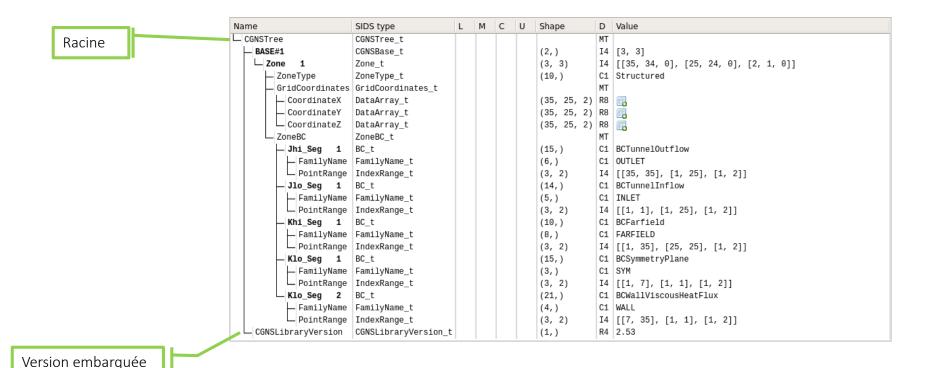


CGNS - Nœud & CGNS/SIDS



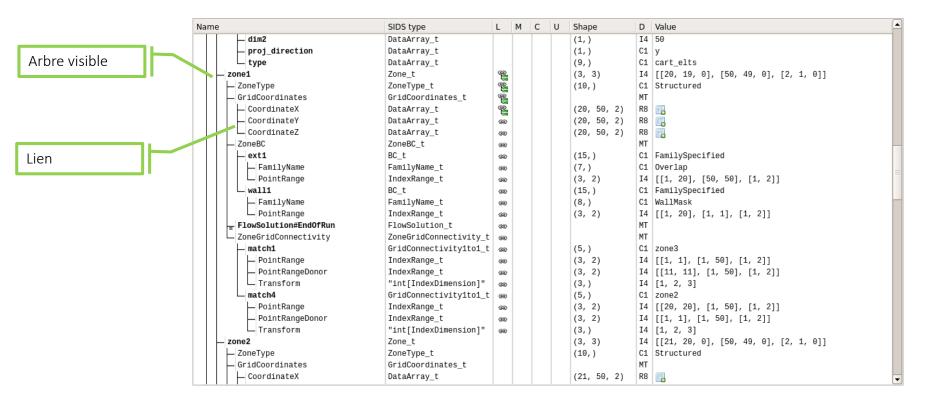


CGNS - Coordonnées





CGNS - Liens



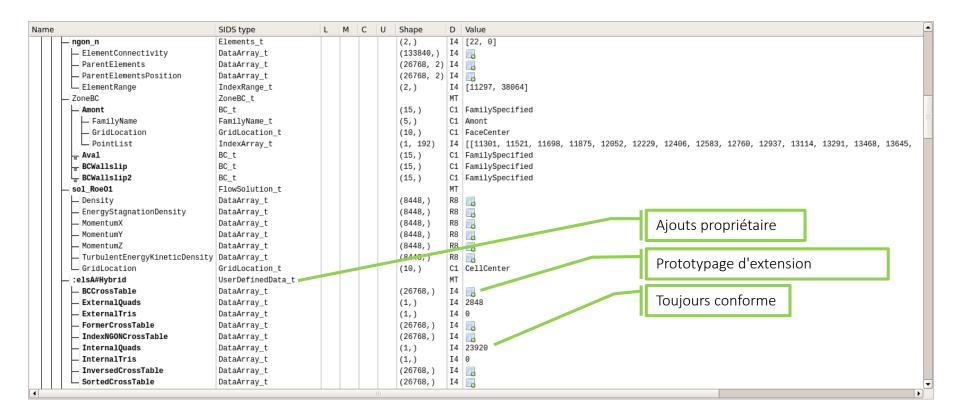


CGNS - Unstructured elements

Autonomie du sous-arbre 7one Une zone pour un processeur (en structuré) D Value Name SIDS type L M С U Shape FamBC FamilyBC_t (14,) C1 BCWallInviscid I4 [[9945, 8448, 0]] b1ku-01 Zone t (1, 3)(11,) FamilyName FamilyName_t C1 Unspecified C1 Unstructured ZoneType ZoneType t (12,) GridCoordinates GridCoordinates t МТ R8 CoordinateX DataArray t (9945,) CoordinateY DataArrav t (9945.) R8 CoordinateZ DataArray t (9945,) R8 — DataClass DataClass t (11.) C1 Dimensional □ DimensionalUnits DimensionalUnits t (160,)C1 Null UserDefined Nu11 Nu11 [14 [7, 0] Amont Elements t (2,)ElementConnectivity DataArray t (768,)I4 [1, 46, 811, 766, 46, 91, 856, 811, 91, 136, 901, 856, 136, 181, 946, 901, 181, 226, 991, 946, 226, (2,) □ ElementRange IndexRange t I4 [8449, 8640] — Aval Elements t (2,)I4 [7, 0] ─ ElementConnectivity DataArrav t (768,)14 [810, 855, 90, 45, 855, 900, 135, 90, 900, 945, 180, 135, 945, 990, 225, 180, 990, 1035, 270, 225, 1 □ ElementRange IndexRange t (2,)I4 [8641, 8832] _ BCWallslip Elements t (2,)I4 [7, 0] __ BCWallslip2 Elements t (2,) I4 [7, 0] — HexElements Elements t (2,)I4 [17, 0] ElementConnectivity DataArray t (67584.) I4 🕞 □ ElementRange I4 [1, 8448] IndexRange t (2,)_ dom-12 Elements t (2,)I4 [7, 0] _ dom-3-split-1 Elements t (2,)I4 [7, 0] (2,) I4 [22, 0] — ngon_n Elements t ─ ElementConnectivity DataArray t (133840,) I4 (26768, 2) 14 — ParentElements DataArray_t (26768, 2) 14 ParentElementsPosition DataArrav t L ElementRange IndexRange_t I4 [11297, 38064]



CGNS - Unstructured BC





CGNS - Solutions

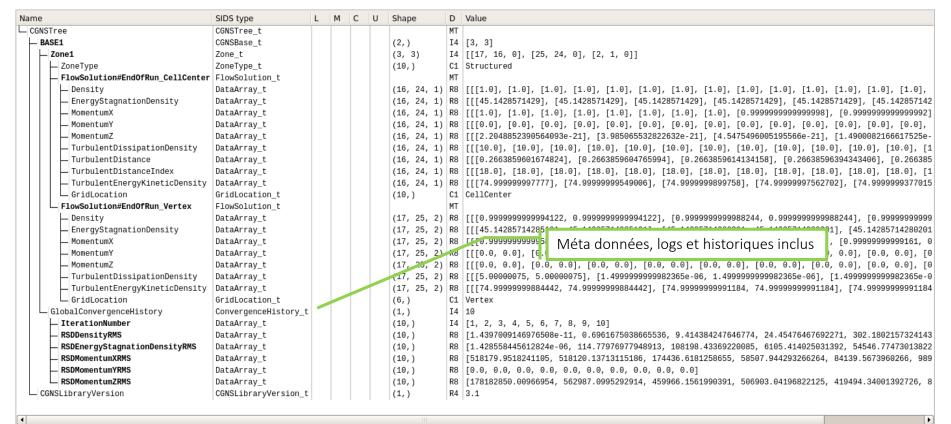
Cohérence du contexte de la simulation

Grand volume de données

-										
Name			SIDS type	L	М	С	U	Shape	D	Value
L	CGNST	ree	CGNSTree_t						MT	
	_ BAS	SE1	CGNSBase_t					(2,)	14	[3, 3]
	L z	Zone1_0	Zone_t					(3, 3)	14	[[19, 18, 0], [25, 24, 0], [2, 1, 0]]
	l 1	ZoneType	ZoneType_t					(10,)	C1	Structured
	1 -	— FlowSolution#EndOfRun_CellCenter	FlowSolution_t						MT	
		— Density	DataArray_t					(18, 24, 1)	R8	[[[0.9998565132575242], [1.000009727031954], [1.0000010145018032], [1.0000001063755095]
		 EnergyStagnationDensity 	DataArray_t					(18, 24, 1)	R8	[[[45.12094121733567], [45.14251457271502], [45.142881514515416], [45.14286894911249],
		— MomentumX	DataArray_t					(18, 24, 1)	R8	[[[0.9390350269087354], [0.9943512054331229], [0.9993435605837907], [0.9999095768346228
		— MomentumY	DataArray_t					(18, 24, 1)	R8	[[[0.0], [0.0], [0.0], [0.0], [0.0], [0.0], [0.0], [0.0], [0.0], [0.0], [0.0], [0.0], [0.0], [0.0], [0.0], [0.0], [0.0]]
		— MomentumZ	DataArray_t					(18, 24, 1)	R8	[[[0.0004699335890680028], [0.00013211426991973708], [1.647843941468899e-05], [2.237100
		 TurbulentDissipationDensity 	DataArray_t					(18, 24, 1)	R8	[[[8855501.218217093], [729437.6436579665], [73621.07245186143], [8553.338059905827], [
		 TurbulentDistance 	DataArray_t					(18, 24, 1)	R8	[[[4.16001675888e-06], [1.3490537268980002e-05], [2.609797064245e-05], [4.50447197898e-
		 TurbulentDistanceIndex 	DataArray_t					(18, 24, 1)	R8	[[[0.0], [0.0]
		 TurbulentEnergyKineticDensity 	DataArray_t					(18, 24, 1)	R8	[[[33.688982572021025], [70.62952752804541], [74.53061025086866], [74.94571750269924],
		└─ GridLocation	GridLocation_t					(10,)	C1	CellCenter
	L	— FlowSolution#EndOfRun_Vertex	FlowSolution_t						MT	
		— Density	DataArray_t					(19, 25, 2)	R8	[[[0.9998565132575242, 0.9998565132575242], [0.999933119602017, 0.999933119602017], [1.
		 EnergyStagnationDensity 	DataArray_t					(19, 25, 2)	R8	[[[44.679984444585244, 44.679984444585244], [45.131727892378265, 45.131727892378265], [
		— MomentumX	DataArray_t					(19, 25, 2)	R8	[[[0.0, 0.0], [0.9666931706558373, 0.9666931706558373], [0.9968473878987397, 0.99684738
		— MomentumY	DataArray_t					(19, 25, 2)	R8	$[[[0.0,\ 0.0],\ [0.0,\ 0.0],\ [0.0,\ 0.0],\ [0.0,\ 0.0],\ [0.0,\ 0.0],\ [0.0,\ 0.0],\ [0.0,\ 0.0],\ [0.0,\ 0.0],\ [0.0,\ 0.0],\ [0.0,\ 0.0],\ [0.0,\ 0.0],\ [0.0,\ 0.0]]]$
		— MomentumZ	DataArray_t					(19, 25, 2)	R8	[[[0.0, 0.0], [0.0003010228302998935, 0.0003010228302998935], [7.429595235502886e-05, 7
		 TurbulentDissipationDensity 	DataArray_t					(19, 25, 2)	R8	[[[232411423.19776723, 232411423.19776723], [4792474.301945026, 4792474.301945026], [40
		 TurbulentEnergyKineticDensity 	DataArray_t					(19, 25, 2)	R8	[[[0.0, 0.0], [52.15927519012747, 52.15927519012747], [72.58007098873153, 72.5800709887
		☐ GridLocation	GridLocation_t					(6,)	C1	Vertex
	L CGN	NSLibraryVersion	CGNSLibraryVersion_t					(1,)	R4	3.1
4								III		

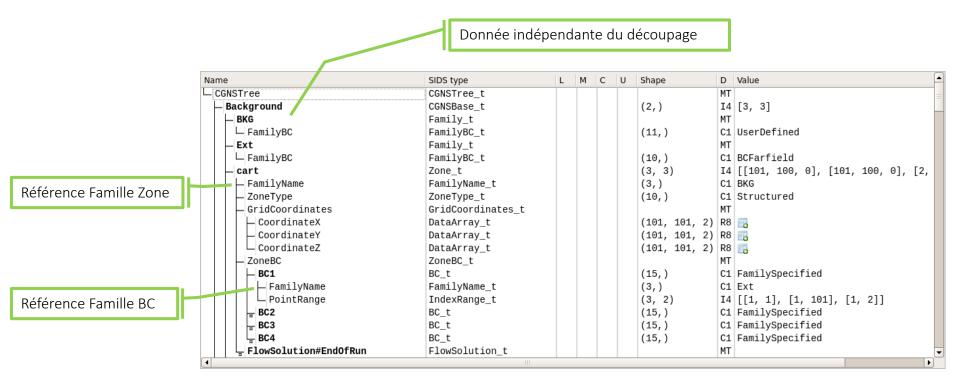


CGNS - Solution convergence





CGNS - Family reference





CGNS/HDF5 - Implémentation de référence pour archivage

Implémentation spécifique

- > Un groupe est un nœud
- Utilisation des attributs pour les champs privés
- > Un nœud a une seule valeur (Dataset)

Support HDF Group

- Support exceptionnel pour les liens puis le parallèle
- > Implication dans le steering committee puis adhésion
- > Pratiques de développement reprises par HDF Group

Avantages de l'implémentation

- > Parfaitement adapté à la simulatin numérique CFD
- > Portage et maintenance sur les plateformes HPC/Compilateurs/Systèmes
- Mapping dataspace idéal pour dissocier modèle de données du modèle physique



CGNS/HDF5 - Liens - 1

Référence à un nœud

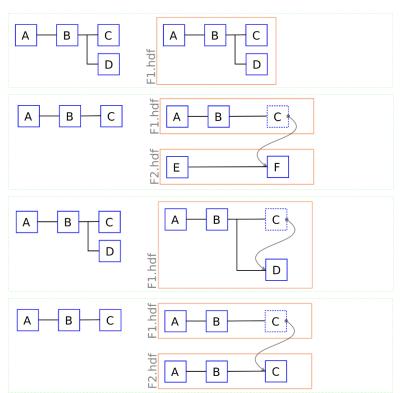
- > Similaire aux liens symbolique file system Unix
- > Référence à un autre nœud, dans le même fichier ou un autre fichier
- > Recherche du fichier avec une liste (ordonnée) de répertoires
- > Le nom du nœud distant est remplacé par le nom local
- > HDF5 se charge de la gestion de façon transparente

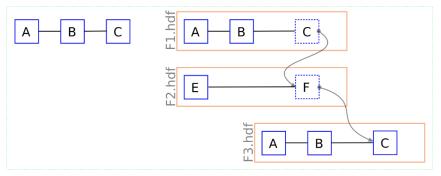
Utilisation

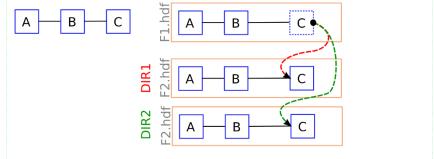
- > Partage de fichiers volumineux (coordonnées, solutions...)
- > Assemblage de sous-arbres (fichiers) en un seul arbre (parallèle, séries...)
- > Création a priori de canevas d'arbre
- Paramètrage local d'arbre



CGNS/HDF5 - Liens - 2





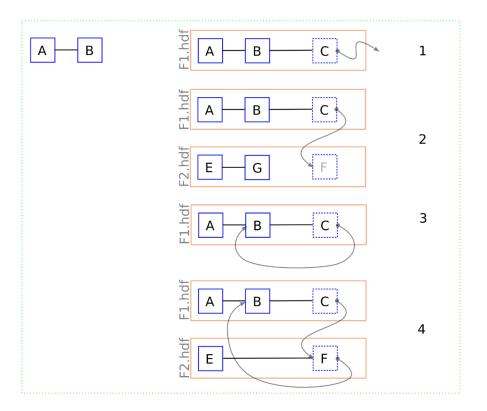




CGNS/HDF5 - Liens - Fautes

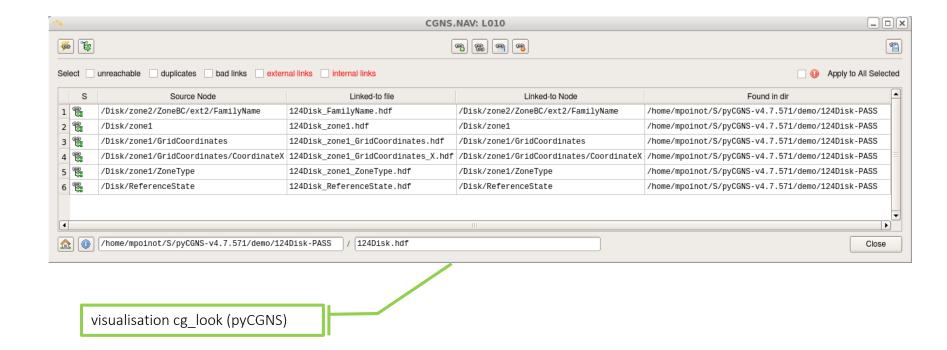
Responsabilité de l'application

- 1. Fichier destination absent
- 2. Nœud destination absent
- 3. Boucle interne
- 4. Boucle externe



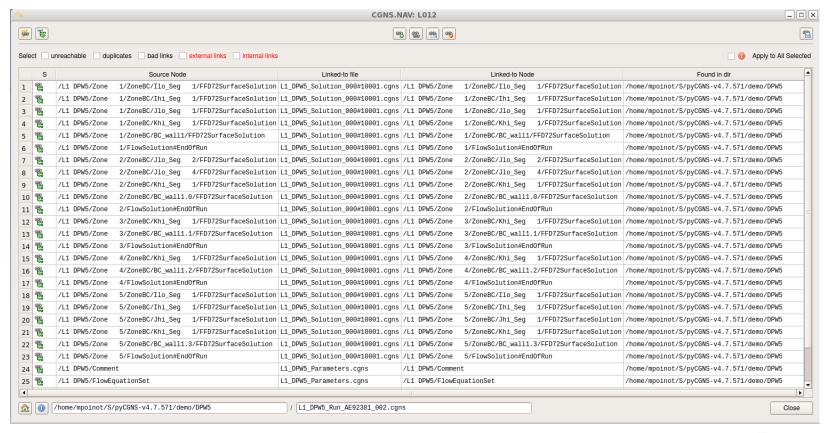


CGNS - Liens



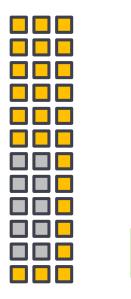


CGNS - Localisation des liens





CGNS/HDF5 - Hyperslab - Zones partielles

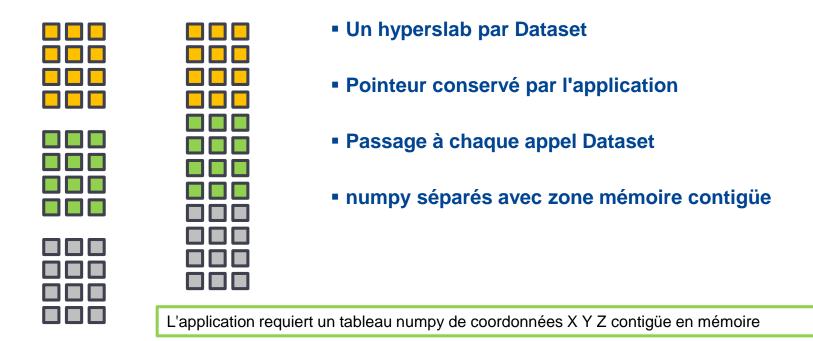


- Lecture de un ou plusieurs Datasets
- Même pointeur mémoire
- Extraction d'une partie des données

N processus en parallèle se partagent de façon exclusive un tableau de coordonnées X Y Z Chaque processeur lit la partie dont il a besoin avec un calcul d'index basé sur son rank

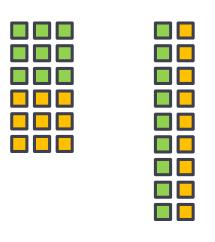


CGNS/HDF5 - Hyperslab - Zones contigües





CGNS/HDF5 - Hyperslab - Zones entrelacées

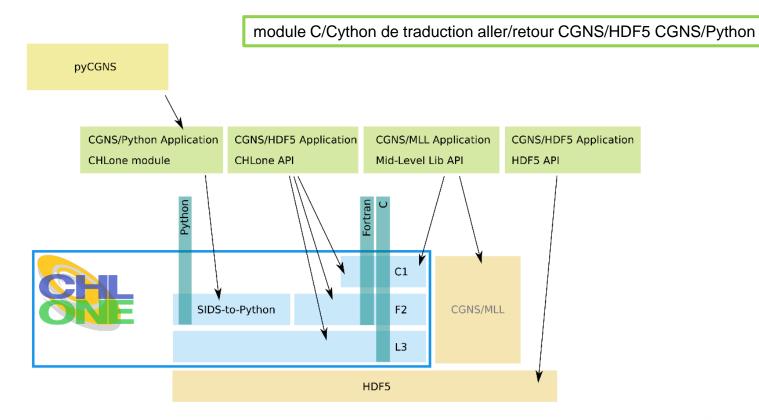


- Un ou plusieurs Datasets
- Même pointeur mémoire
- Entrelacé

L'application requiert un tableau numpy de coordonnées X Y Z entrelacées en mémoire



CGNS/HDF5 - vers CGNS/Python avec CHLone





CGNS/Python - Implémentation de référence pour les échanges en mémoire

Arbre 100% python

- > Tableaux numpy
- > Très grande simplicité de l'implémentation
- > Liens gérés par l'application

Adapté aux échanges en mémoire

- Les algorithmes de haut niveau sont en python
- > Les spécialisations HPC restent en C/C++/Fortran avec accès direct mémoire
- > La sérialisation permet le passage direct d'arbre en MPI



CGNS/Python - Simple

- Node name
 - > String
- Node value
 - > Numpy array
 - Fortran order
 - I4/I8/R4/R8/C1
 - None
- ◆ Node children
 - > List of nodes
- ◆ Node SIDS-type
 - > String

SIDS-to-Python (CGNS/Python), Release 3.1.2

1.3 Textual representation

It is possible to declare a CGNS/Python node as a textual representation. There is a exemple of a zone connectivity sub-tree with the CGNS/Python in textual mode, a simple PointRange node with two 3D indices:

```
pr=['PointRange',
    numpy.array([[1,25],[1,9],[1,1]],dtype=numpy.int32,order='Fortran'),
    [],
    'IndexRange_t']
```

The PointRange node has no child, the children list is an empty list. The values of the array are initialized with a list, the order of the elements in the list matches the *Fortran* indexing: in that example the first point indices are [1,1,1] and the second point indices are [25,9,1].

The evaluation of this string by the Python interpreter creates a CGNS/Python compliant node as a Python list. Please note the types of this pr node, there are only native Python types (list, string, integer) and *numpy* types or enumerates. You have to have a variable to hold the node or the CGNS sub-tree, if you have no reference to the actually created Python objects these will be unreachable and thus garbaged.



pyCGNS - Modules

APP

Application: outils et exemples d'utilisation des autres modules

WRA

Wrapper en python de la librairie CGNS/MLL

MAP

Mapper: Mapping simple (avec CHLone) de CGNS/HDF5 en CGNS/Python

NAV

Navigator: outil de navigation et d'édition d'arbres CGNS/HDF5 et CGNS/Python

DAT

Database: outils SQL de collecte et de recherche pour arbres CGNS

VAL

Validation: outil de validation d'arbre basé sur des grammaires

PAT

Pattern: librairies de patterns d'arbres et fonctions de manipulation de base



CGNS/Python - MAP

Deux fonctions

> Load/ save

Pilotage par paramètres

- > Lecture/ écriture partielle de sous-arbre (inclusif/ exclusif)
- > Lecture sans données (avec seuil)
- > Profondeur réglable
- > Gestion fine des liens

```
import CGNS.MAP
import CGNS.PAT.cgnsutils as CGU

(tree,links,paths)=CGNS.MAP.load("T0.hdf")

for p in CGU.getAllPaths(tree): print p
```



CGNS/Python - PAT - Fonctions

Utilitaires de manipulation d'arbres

- > Accès par node
- > Accès par path
- > Parcours d'arbres
- Contantes et listes du CGNS/SIDS

CGNS.PAT.cgnsutils.getNodeByPath(tree, path)

Returns the CGNS/Python node with the argument path:

```
zbc=getNodeByPath(T,'/Base/Zone001/ZoneBC')
nchildren=len(childrenNames(zbc))
```

The path is compared as a string, you should provide the exact path if you have a sub-tree or a tree with its *CGNSTree* fake node. The following lines are not equivalent (sic!):

```
zbc=getNodeByPath(T,'/Base/Zone001/ZoneBC')
zbc=getNodeByPath(T,'/CGNSTree/Base/Zone001/ZoneBC')
```

You can change the relative root by giving any sub-node of the complete tree. For example, to get a specific BC node in a zone, you first look for the ZoneBC of the zone and you use the returned node a the new root:

```
zbc=getNodeByPath(T,'/Base/Zone001/ZoneBC')
nbc=getNodeByPath(zbc,'./wall01')
```

- Parameters: tree (node) the target tree to parse
 - path (str) absolute or relative path
- Returns: The CGNS/Python *node* matching the path
 - Returns None if the path is not found
- Remarks: No wildcards allowed (see getPathsByNameFilter() and getPathsByNameFilter())
 - there is no concept of absolute or relative path, the path is always the concatenation of children node names (and then recurse)



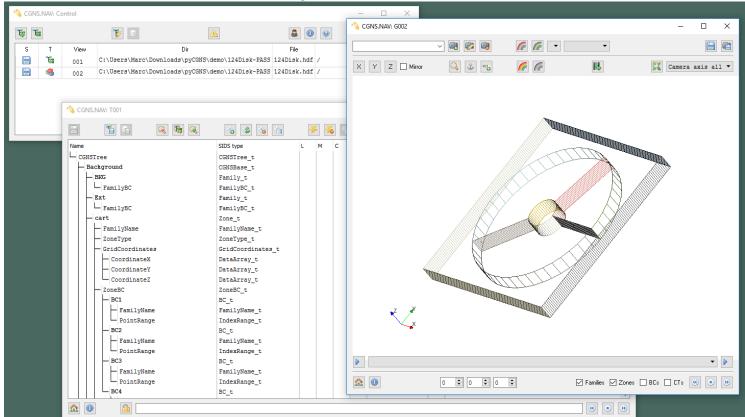
CGNS/Python - PAT - Les données en python



```
import M
                                             import CGNS.PAT.cgnslib as CGL
base[2].append(M.data)
                                             data=CGL.newFamily(None,'{Family}')
                                             CGL.newGeometryReference(data,'{GeometryReference}')
                                             CGL.newRotatingCoordinates(data)
                                             CGL.newUserDefinedData(data,'{UserDefinedData}')
                                             CGL.newDescriptor(data,'{Descriptor}')
def newFamily(parent, name):
                                             CGL.newOrdinal(data)
  if (parent): CU.checkNode(parent)
  CU.checkType(parent,CK.CGNSBase ts,name)
  CU.checkDuplicatedName(parent, name)
  node=CU.newNode(name, None, [], CK.Family ts, parent)
return node
                                             def newNode(name, value, children, type, parent=None, dienow=False):
                                               node=[name, value, [], type]
                                               if isinstance(children, list) and children is not []:
                                               if checkNodeCompliant(children): setAsChild(node,children)
                                               else: node[2]=children
                                               if (dienow): checkNodeCompliant(node, parent, dienow)
                                               if (parent): setAsChild(parent, node)
                                             return node
```



CGNS/Python - NAV - Les GUI faciles de Python

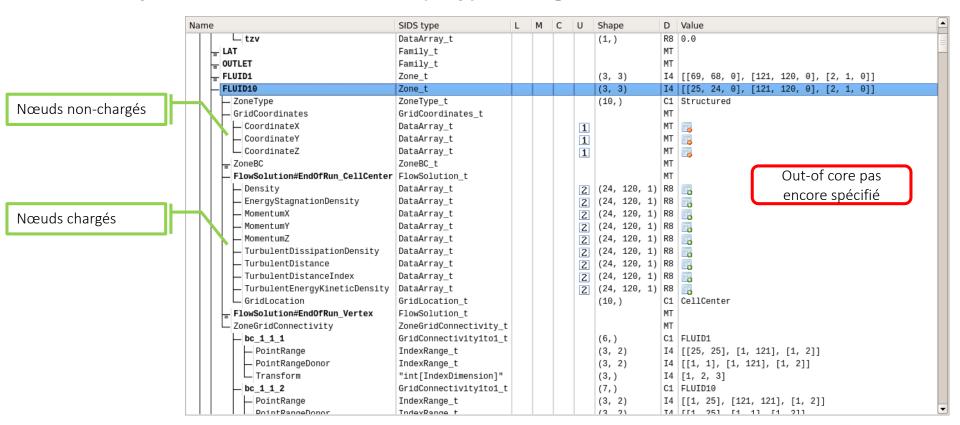


Outils Qt

- > pySide puis pyQt
- > Designer/cython
- 3D avec VTK
- Windows/ Unix



CGNS/Python - NAV - On-demand (lazy) loading





CGNS/Python - VAL

Vérification des arbres

- > Parcours des arbres
- > 3 niveaux de grammaire
 - CGNS/Python implémentation
 - CGNS/SIDS
 - CGNS/user
- > Spécialisation possible par l'utilisateur

Expression des contraintes

- > Impossible hors langage de programmation
- > Fonctions travaillant directement sur l'arbre
- > Requètes 'à-la-XPATH'

```
if (node[0] == CGK. Element Range s):
  eshp=(2,)
  checkData=False
 if (CGU.getShape(node)!=eshp):
    rs=log.push(pth,'s0000.0192',CGU.getShape(node),eshp,node[0])
  elif (CGU.getShape(node)!=shp):
    rs=log.push(pth,'s0000.0192',CGU.getShape(node),shp,node[0])
    checkData=False
  if (ntype not in CGU.getAuthDataTypes(node)):
    rs=log.push(pth,'s0000.0199',ntype,CGU.getAuthDataTypes(node))
  if checkData:
    pr=node[1]
   if (any(pr.flatten()<1)): rs=log.push(pth,'s0000.0206')
      # checking out of range (<1)
      for d in range(self.context[CGK.IndexDimension s][zpth]):
        if (pr[d][0]>pr[d][1]): rs=log.push(pth,'s0000.0207')
```



Conclusion -

- Le modèle de données
- Une approche consensuelle
- Eviter les librairies propriétaires



Prospectives

- Etendre le modèle de données
- Considérer le processus comme donnée
- Gérer une API en mémoire numpy

