





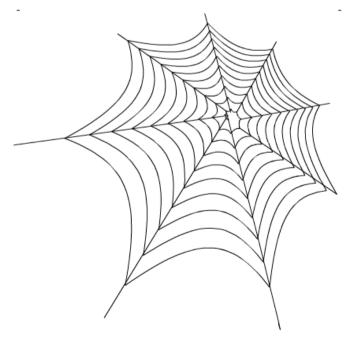
AMI / Rucio / LOCALGROUPDISK/FAX

ATLAS ソフトウェア講習会 2016 河村元 II.Physikalisches Institut, Universität Göttingen

Overview

- AMI (ATLAS メタデータインターフェース網)
 - AMI
 - ATLAS データセット
 - AMI CLI
 - AMI WebUI
- Rucio (ATLAS 分散データ管理システム)
 - 基本概念
- LOCALGROUPDISK
 - ATLAS LOCALGROUPDISK
 - Rucio CLI
- RucioUI (WebUI)
- FAX (Federated ATLAS Xrootd)
 - FAXとは?
 - 基本的な使い方

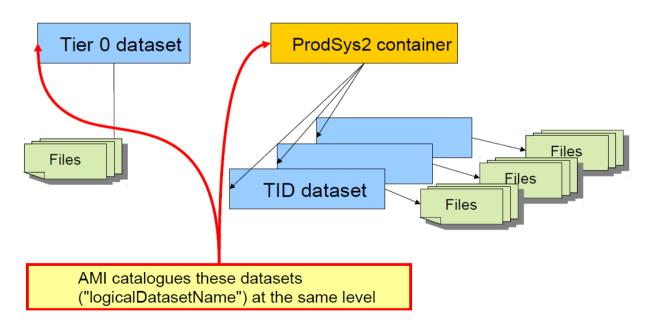
AMI (ATLAS メタデータインターフェース網)



AMI (ATLAS Metadata Interface)

- メタデータ = データのデータ
- ATLASのデータ再構築の過程でそれぞれのステップのAMI tag が作られる
 - ATLAS データセット・メタデータリンク
 - データセットの起源、サイズ
 - ファイル、イベント数
 - ソフトウェア・パラメタ (AMI Tags)
 - MCパラメタ (PDF, generator, cross section, etc)
 - 失われたファイルやLumi blockなど
 - 他のアプリケーションへのリンク(COMA, Rucio)
 - データ ピリオド
 - Luminosity、トリガー、データ取得時期などの情報
 - 特殊インタフェース
 - AMI-Tags
 - ソフトウェア設定パラメタ→ software configuration parameters
 - AMI-Glance
 - データ+出版物
 - · Definition of data periods
 - Definition of physics containers
 - · Reprocessing campaigns
 - Event count comparator

- <u>データセットとは?ファイルの集合体</u>
 - <u>コンテナはデータセットあるいはコンテナの集合体</u>
 - TierO ATLAS DAQの出力ファイル
 - RAWと第一段再構築ファイル
 - ProdSys2のタスクの出力ファイル
 - TID(Task ID) でタグ付けされる
 - MCとリプロセッシング

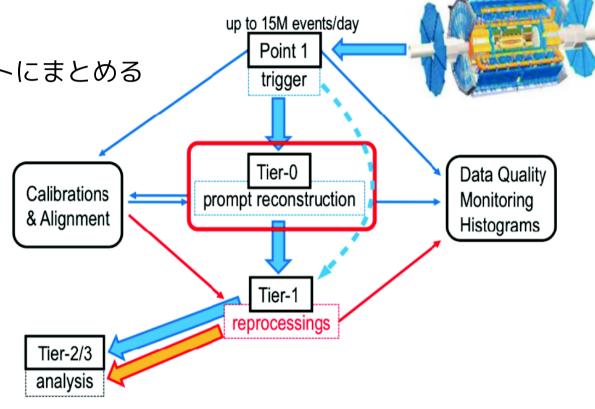


• Tier-0

- RAW ファイルをデータセットにまとめる

- 第一段階のデータ再構築

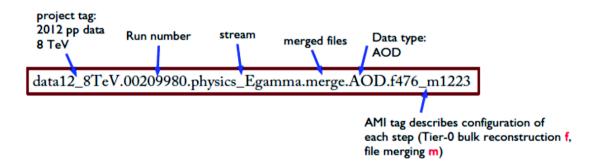
- ProdSys2
 - MCやリプロセッシング
 - PanDA, JEDI, DEFT
- 分散データ管理システム
 - Rucio



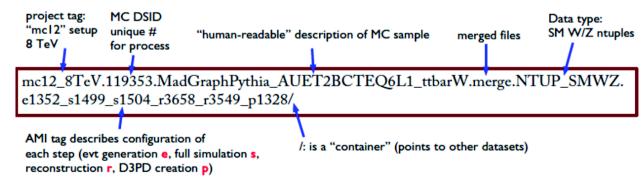
- データ再構築 → 例: x353 = AMI tag
 - Reco_tf.py —AMI=x353 --inputBSFile=tier0_RawData.data
- リプロセッシング・キャンペーン
 - ソフトウェアやフレームワークの大改良後、しばしば RAW から AOD データが再構築される
 - キャンペーン自体は大量の計算能力を使う
 - 結果、複数の異なったバージョンの AOD
 - Derivation framework はその間 AOD から DAOD (Derived AOD) を生成
 - 約1%ほどのサイズのオリジナル AOD
 - 特定のイベントや変数
 - キャリブレーションやアプリ設定などはバージョンごとに潜在的に異なる

- データセットはファイルの集合体
 - 基本的に実験データ (data) と Monte Carlo (mc)

Data:

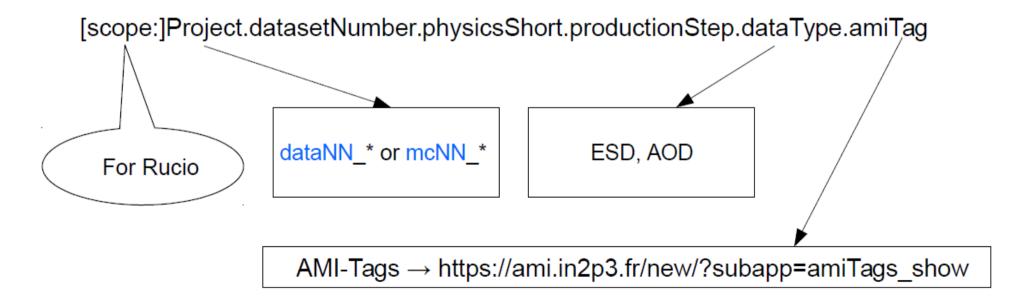


Simulation:



AMI tag と Rucio

[scope:]Project.runNumber.streamType.productionStep.dataType.amiTag



AMI CLI

• pyAMI ドキュメント: https://ami.in2p3.fr/pyAMI/

Loading pyAMI client

Isetup pyami

Displaying metadata

ami show dataset info data13_2p76TeV.00219364.physics_MinBias.merge.NTUP_HI.f519_m1313

logicalDatasetName: data13 2p76TeV.00219364.physics MinBias.merge.NTUP HI.f519 m1313

nFiles : 547

totalEvents : 5771888

totalSize : 2558965331099

runNumber : 219364

period : C4

prodsysStatus : Tier 0 dataType : NTUP_HI beamType : NULL conditionsTag : NULL geometryVersion : NULL

streamName : physics MinBias

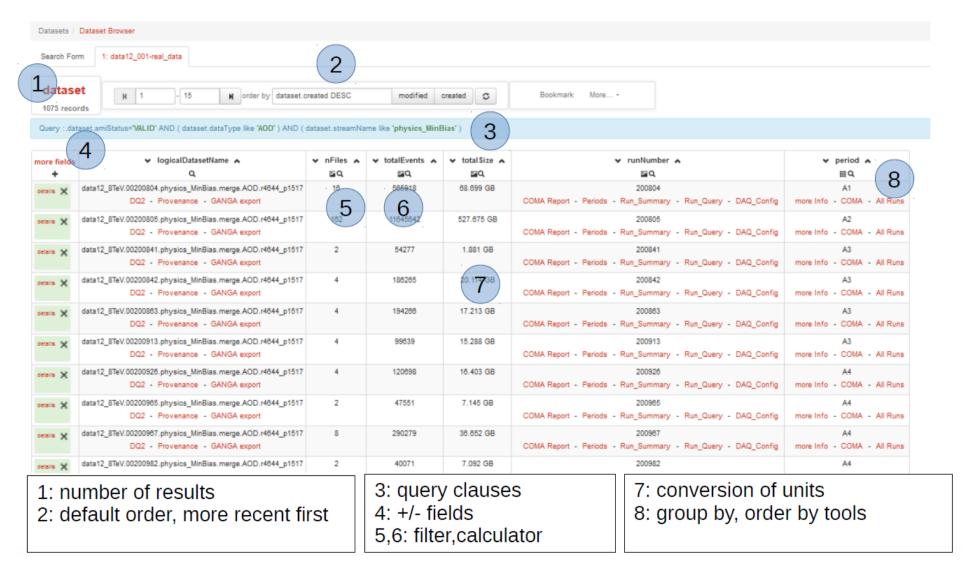
version : f519_m1313 added_comment : NULL

keyword : NULL

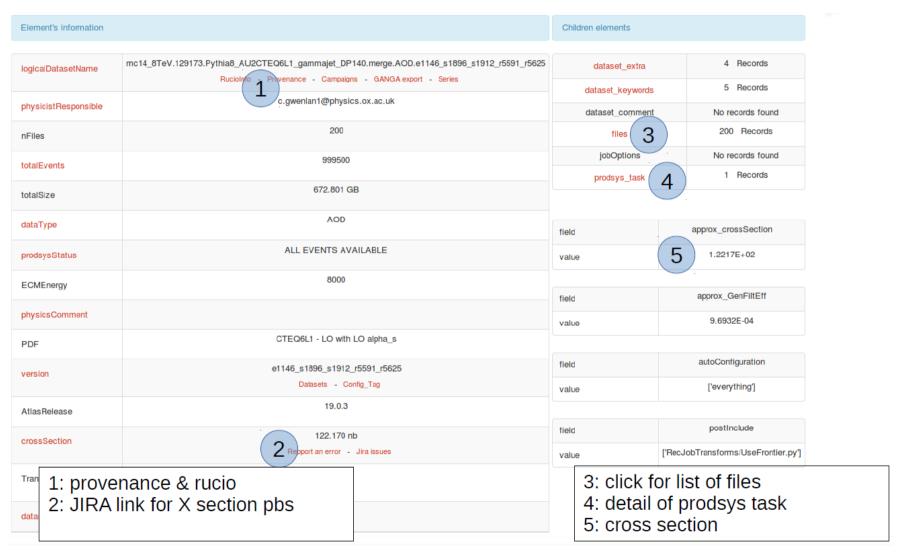
Displaying provenance

ami show dataset prov data13_2p76TeV.00219364.physics_MinBias.merge.NTUP_HI.f519_m1313

AMI WebUI - 1



AMI WebUI - 2



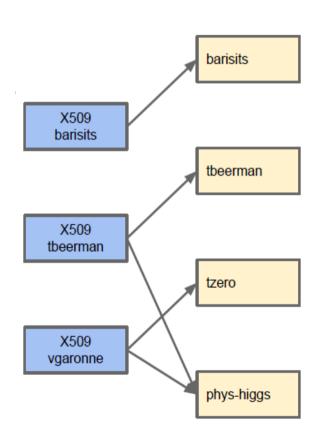
Rucio (ATLAS 分散データ管理システム)



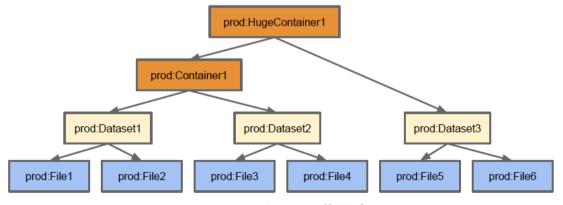
ATLAS ソフトウェア講習会 2016

- ATLAS の分散データ管理システム DQ2 の新しい実装
 - DQ2のサポートは2016年12月で終了
- Grid ジョブ内部でのデータの移動にも使用
- CVMFS 上で Rucio CLI ツールから使用
 - Web インターフェースでも同じような機能を使用 可能(後述)

- アカウント (account)
 - 各 Rucio アカウントはユーザー やグループ等にマップ
 - Quota や属性はそれぞれのアカ ウントごとに設定
 - Rucio account は X509 証明書や ケルベロス認証をサポート
 - 一つの証明書は複数のアカウン トにマップ可



- Rucio ネームスペース
 - 3種類の Data Identifiers (DIDs): ファイル,データセット,コンテナ
 - データセット:複数ファイルのセット
 - コンテナ:複数データセットあるいは複数コンテナ
 - ネームスペースはスコープごとに分割可能。ネームスはスコープ 内で一意。 DID はスコープやネームで識別可能



ATLAS ソフトウェア講習会 2016

- Rucio ストレージエレメント (RSE)
 - Rucioストレージエンドポイント,例: CERN-PROD_DATADISK
 - 検索キーも使用可能 (例 tier=2, cloud=FR)
- レプリケーション(複製)要求
 - 例、データセット user.gen:my.dataset を Tokyo-LCG2_ LOCALGROUPDISKへ複製
 - RSE 検索キー: cloud=FR&type=LOCALGROUPDISK
 - Rucioはレプリカ作成の際に最小ディスク容量と最小転送量になるよう最適化

LOCALGROUPDISK



ATLAS LOCALGROUDPDISK

- 一般ユーザが管理可能なディスクエリアは2つある
 - 各サイトはSCRATCHDISK と LOCALGROUPDISKを提供している
- SCRATCHDISK (例:ドイツ Tier1 + Tier2s)
 - FZK-LCG2 SCRATCHDISK
 - DESY-HH_SCRATCHDISK
 - DESY-ZN_SCRATCHDISK
 - LRZ-LMU_SCRATCHDISK
 - WUPPERTALPROD_SCRATCHDISK
 - UNI-FREIBURG_SCRATCHDISK
 - GOEGRID_SCRATCHDISK
- LOCALGROUPDISK (例: DESY-HH と UniGoettingen)
 - DESY-HH_LOCALGROUPDISK
 - GOEGRID LOCALGROUPDISK
 - LOCALGROUPDISK

このエリアはジョブアウトプット等の一時保存用。一定時間後(2週間?)消去される。

このエリアはデータの恒久保存用

ATLAS LOCALGROUDPDISK

- 一般ユーザが管理可能なディスクエリアは2つある
 - 各サイトはSCRATCHDISK と LOCALGROUPDISKを提供している
- SCRATCHDISK (例:ドイツ Tier1 + Tier2s)
 - FZK-LCG2 SCRATCHDISK
 - DESY-HH_SCRATCHDISK
 - DESY-ZN_SCRATCHDISK
 - LRZ-LMU_SCRATCHDISK
 - WUPPERTALPROD_SCRATCHDISK
 - UNI-FREIBURG_SCRATCHDISK
 - GOEGRID_SCRATCHDISK
- LOCALGROUPDISK (例: DESY-HH と UniGoettingen)
 - DESY-HH_LOCALGROUPDISK
 - GOEGRID_LOCALGROUPDISK
 - LOCALGROUPDISK

このエリアはジョブアウトプット等の一時保存用。一定時間後(2週間?)消去される。

このエリアはデータの恒久保存用

日本では <u>TOKYO-LCG2 LOCALGROUPDISK</u> のみ

Rucio CLI

Loading Rucio client Isetup rucio ******************************* Requested: rucio ... Setting up emi 3.17.1-1 v2.sl6 ... Skipping: grid middleware already setup (from UI) Setting up rucio 1.7.3 ... Info: Setting compatibility to slc6 Info: Set RUCIO AUTH TYPE to x509 proxy Info: Set RUCIO ACCOUNT to gkawamur emi: Your proxy has 95h:54m:0s remaining ## In Rucio, check which account you use rucio whoami ## Check all RSEs in Tokyo \$ rucio list-rses | grep TOKYO TOKYO-LCG2 DATADISK TOKYO-LCG2 DET-MUON TOKYO-LCG2 LOCALGROUPDISK TOKYO-LCG2 PERF-JETS TOKYO-LCG2 PERF-MUONS TOKYO-LCG2 PHYS-EXOTICS

TOKYO-LCG2_PHYS-HIGGS TOKYO-LCG2_PHYS-SUSY

TOKYO-LCG2_SCRATCHDISK TOKYO-LCG2_TRIG-DAQ

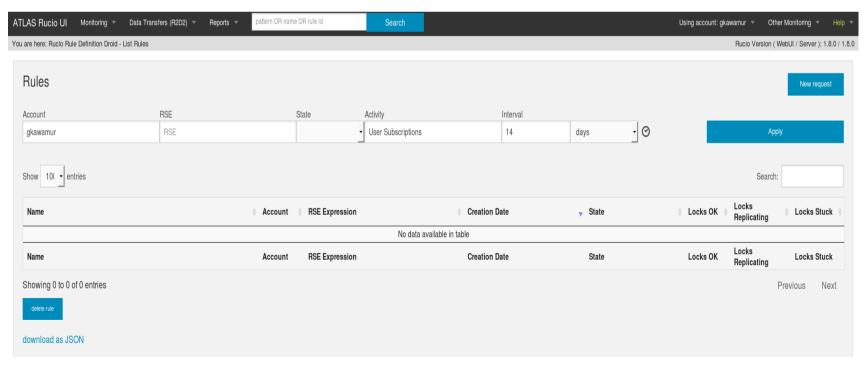
RucioUI (WebUI)



RucioUI

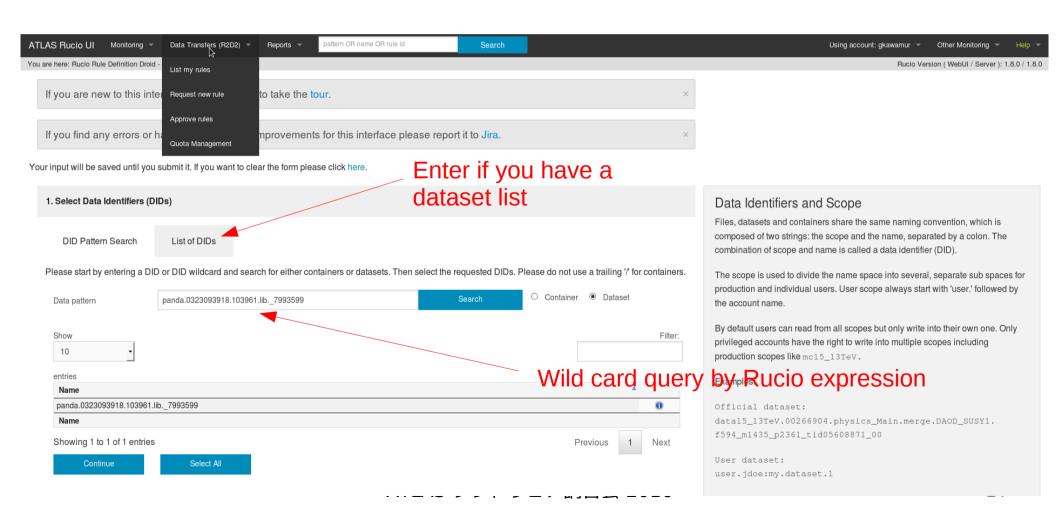
- データを RSE の間で転送、 Rucio 複製ルールの使用
- RucioUI ツール: Rucio Rule Definition Droid (R2D2) https://rucio-ui.cern.ch/r2d2
- ブラウザに X509 ユーザー証明書が必要





RucioUI: DIDs の選択 - 1

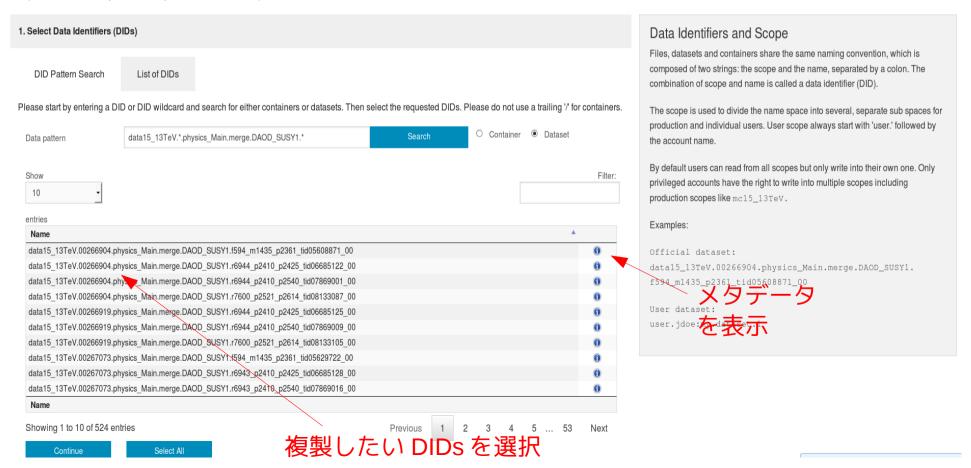
• Data transfers (R2D2) → 新しいルールの定義



RucioUI: DIDs の選択 - 2

- 例、ワイルドカード
 - "data15_13TeV.*.physics_Main.merge.DAOD_SUSY1.*"

Your input will be saved until you submit it. If you want to clear the form please click here.



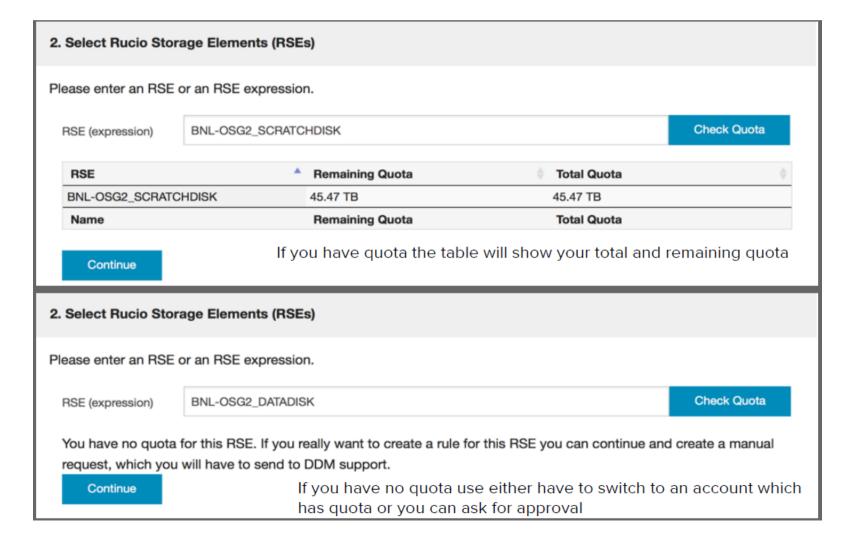
RucioUI: DIDs の選択 - 3

メタデータ

data15_13TeV.00266904.physics_Main.merge.DAOD	_SUSY1.f594_m1435_p2361_tid05608871_00
accessed_at	Mon, 14 Sep 2015 04:34:53 UTC
account	panda
availability	AVAILABLE
closed_at	Mon, 24 Aug 2015 14:33:30 UTC
created_at	Sun, 07 Jun 2015 10:09:59 UTC
datatype	DAOD_SUSY1
did_type	DATASET
events	10593
filesize	237.94 MB
hidden	false
is_open	false
length	7
monotonic	false
name	data15_13TeV.00266904.physics_Main.merge.DAOD_SUSY1.f594_m1435_p2361_tid05608871_00
obsolete	false
prod_step	merge
project	data15_13TeV
run_number	266904
scope	data15_13TeV
stream_name	physics_Main
suppressed	true
task_id	5608871
updated_at	Mon, 14 Sep 2015 06:35:06 UTC
version	f594_m1435_p2361

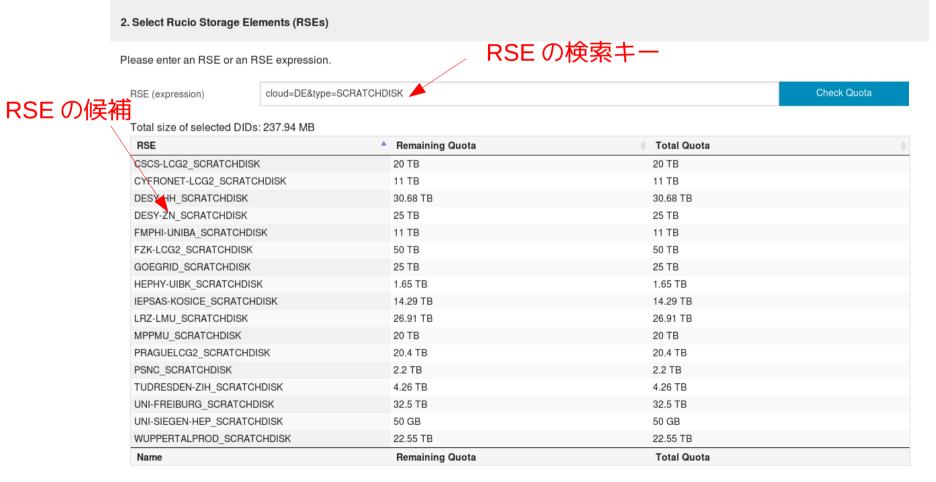
RucioUI: RSE の選択 - 1

• 転送先 RSE を選択



RucioUI: RSE の選択 - 2

- 例: SCRATCHDISK の選択
 - cloud=DE&type=SCRATCHDISK



RucioUI: オプションの選択

3. Options					
Please select/enter your w	anted options and then submit your rule re	equest.			
Grouping ○ All ● Dataset ○ Lifetime (in days). Leave er	None	A grouping definition of how the replica will be distributed			
15	.p.y to minimo monnio				
Copies					
Comment					
For ATLAS-D					
Create sample	Number of files				
Asynchronous Mode	Use if you so randomly	elect files			
Continue					

Rucio UI: 確認

• 複製要求を出す前に最後の確認

4. Summary

This request will create rules for the following DIDs:

DID	^ Co	pies Files	Size	Requested Size
data15_13TeV:data15_13TeV.00266904.physics_Main.merge.DAOD_SUSY1.f594 m1435_p2361_tid05608871_00	- 1	7	237.94 MB	237.94 MB
data15_13TeV:data15_13TeV.00266904.physics_Main.merge.DAOD_SUSY1.r694 p2410_p2425_tid06685122_00	⁴ - 1	14	2.96 GB	2.96 GB
data15_13TeV:data15_13TeV.00266904.physics_Main.merge.DAOD_SUSY1.r694 p2410_p2540_tid07869001_00	⁴ – 1	5	3.4 GB	3.4 GB
data15_13TeV:data15_13TeV.00266904.physics_Main.merge.DAOD_SUSY1.r760 p2521_p2614_tid08133087_00	0_ 1	10	2.62 GB	2.62 GB
Total	4	36	8.59 GB	8.59 GB

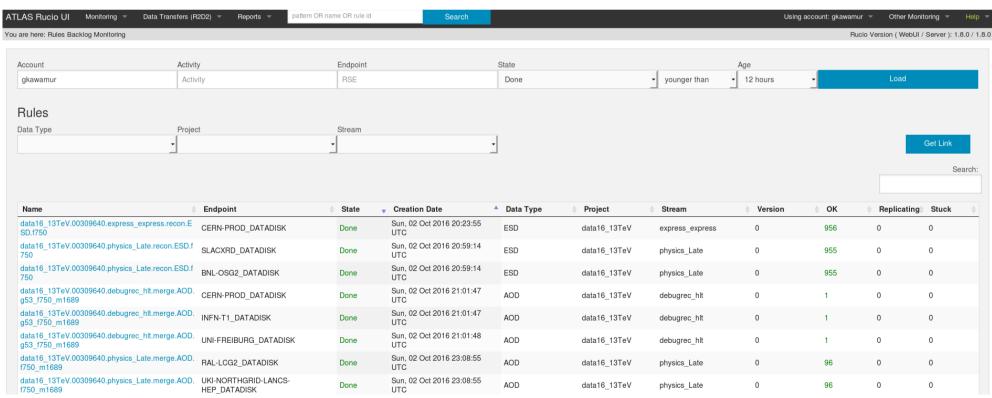
The rules will replicate to one of the following RSEs:

RSE のディスク容量を確認!

A Remaining Quota	♦ Total Quota ♦
20 TB	20 TB
11 TB	11 TB
30.68 TB	30.68 TB
25 TB	25 TB
11 TB	11 TB
50 TB	50 TB
25 TB	25 TB
1.65 TB	1.65 TB
14.29 TB	14.29 TB
26.91 TB	26.91 TB
20 TB	20 TB
20.4 TB	20.4 TB
2.2 TB	2.2 TB
4.26 TB	4.26 TB
	20 TB 11 TB 30.68 TB 25 TB 11 TB 50 TB 25 TB 1.65 TB 14.29 TB 26.91 TB 20 TB 20.4 TB 2.2 TB

Rucio UI: 転送状況

- 複製(転送)状況の確認をトップページで確認
 - https://rucio-ui.cern.ch/r2d2

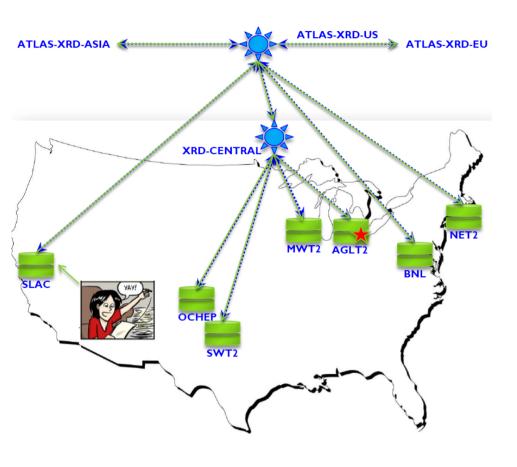


ATLAS ソフトウェア講習会 2016

FAX (Federated ATLAS Xrootd)



FAX とは?



- グローバル・ファイルネーム検索システムによる Xroot protocol による読み込み専用ファイル転送システム
 - X509 (ユーザー)証明書が必要
 - リモート・ファイル・ディスクリプタも 使用可能
 - Root 等からローカルファイルのように直接読み込み可能
- 一意なグローバル・ファイルネームから最近 傍にあるファイルを転送可能
- Grid でも使用可能
 - ローカル・サイトにファイルがなくても 良いのでジョブの成功率が高まる(は ず?)

FAX の基本的な使い方

- データセット名は Rucio と同様。
 - プロトコルとアクセス用のグローバルネームが異なるだけ

Loading FAX client with root

Isetup fax root

Check a dataset

fax-ls user.ivukotic:user.ilijav.HCtest.1

Dataset contains 6 files.

At the moment all of them can be accessed using FAX. in average files have 70 replicas.

Get the dataset

fax-get user.ivukotic:user.ilijav.HCtest.1

Search qLFN

fax-get-gLFNs user.ivukotic:user.ilijav.HCtest.1

root://atlasxrootd-kit-doors.gridka.de:1094//atlas/rucio/user.ivukotic:Details.log.2 root://atlasxrootd-kit-doors.gridka.de:1094//atlas/rucio/user.ivukotic:group.test.hc.AOD.root root://atlasxrootd-kit-doors.gridka.de:1094//atlas/rucio/user.ivukotic:group.test.hc.ESD.root root://atlasxrootd-kit-doors.gridka.de:1094//atlas/rucio/user.ivukotic:group.test.hc.NTUP_SMWZ.root root://atlasxrootd-kit-doors.gridka.de:1094//atlas/rucio/user.ivukotic:group.test.hc.NTUP_SUSY.root root://atlasxrootd-kit-doors.gridka.de:1094//atlas/rucio/user.ivukotic:group.test.hc.NTUP_TRIG.root

For example, copy a file locally

xrdcp root://atlasxrootd-kit-doors.gridka.de:1094//atlas/rucio/user.ivukotic:group.test.hc.AOD.root test.root

リンクと参考文献

- RucioUI
 - https://rucio-ui.cern.ch/
- Rucio Documentation
 - http://rucio.cern.ch/index.html
- ATLAS-D meeting 2015 Rucio Tutorial, Thomas Beermann
- ATLAS-D meeting 2016 Grid/Rucio Tutorial, Gen Kawamura
- The pyAMI
 - https://ami.in2p3.fr/pyAMI/
- ATLAS Software Tutorial, Feb 2016
 - https://indico.cern.ch/event/465378/