

JVOの研究開発

すばる望遠鏡データ解析機能の導入

白崎 裕治 国立天文台

田中 昌宏, 川野元 聡, 本田 敏志, 大石 雅寿, 水本 好彦,
大江 将史 国立天文台

増永 良文 御茶ノ水女子大

安田 直樹 東京大学

石原 康秀, 堤 純平 富士通 株式会社

中本 啓之, 小林 佑介, 坂本 道人 株式会社セック

イントロダクション

- **バーチャル天文台 (JVO) が提供するサービス**
 - **検索サービス**
 - 世界中のデータベース → 一つの巨大なデータベース
 - 試験運用システムとして公開中 (<http://jvo.nao.ac.jp/portal>)
 - **解析サービス**
 - 解析ソフトのインストール不要
 - Web ブラウザがあればどんな環境でも解析可能
 - 統一されたインターフェイスでさまざまな種類のデータを解析
 - 大容量データを使う解析
 - 計算機の低価格化、高速化
 - データと計算資源の集約化
 - グリッドコンピューティング技術

SuprimeCam データ解析機能

- SuprimeCam のデータ量は大きい
 - 2007年3月時点で 6.3 TB
- 基本的にサーベイデータとして利用される
 - 全データを利用したい。
- 多波長データの一部として利用
 - 光赤外の人以外にもX線や電波の人も利用したい。
 - 解析未経験者に対するサイエンスレディーデータの提供。

SuprimeCam データ解析機能

● フラットフレーム作成機能

- オブジェクトフレームのメディアンを計算
- カスタムフラットフレームを作成したい場合に利用
- 利用するデータの観測期間、フィルター、チップ名を指定。そのほか使用データのセレクション条件

● モザイクングデータ作成機能

- 生データのフラット補正後、重ね合わせ処理を行う。
- オブジェクト名とフィルター名を指定(現状)
- 一度モザイク処理されたデータはキャッシュ(予定)
- 任意の期間、任意の領域、利用データのクォリティーを指定(予定)

● 機能の簡略化

- 可能な限り解析を自動化。
- 解析システムが適切なパラメータを自動判断してパイプラインを走らせる
- ユーザは欲しいデータの種類(座標、フィルター、観測期間、クォリティー指標)のみを指定。

解析システムの構成

0. DASサーバは定期的に自身の負荷情報をアップロード

5. データ転送 FTP

~30 コア → 100 コア

DAS

データ解析サーバ

1 Gbps x 11 → 25

MDS

ジョブ管理
サーバ

2. ジョブ実行先の問い合わせ。

JVO Portal



3. ジョブ投入

48 Gbps
→ 96 Gbps

1 Gbps x 5 → 14

SRS

ストレージサーバ

30 TB → 100 TB

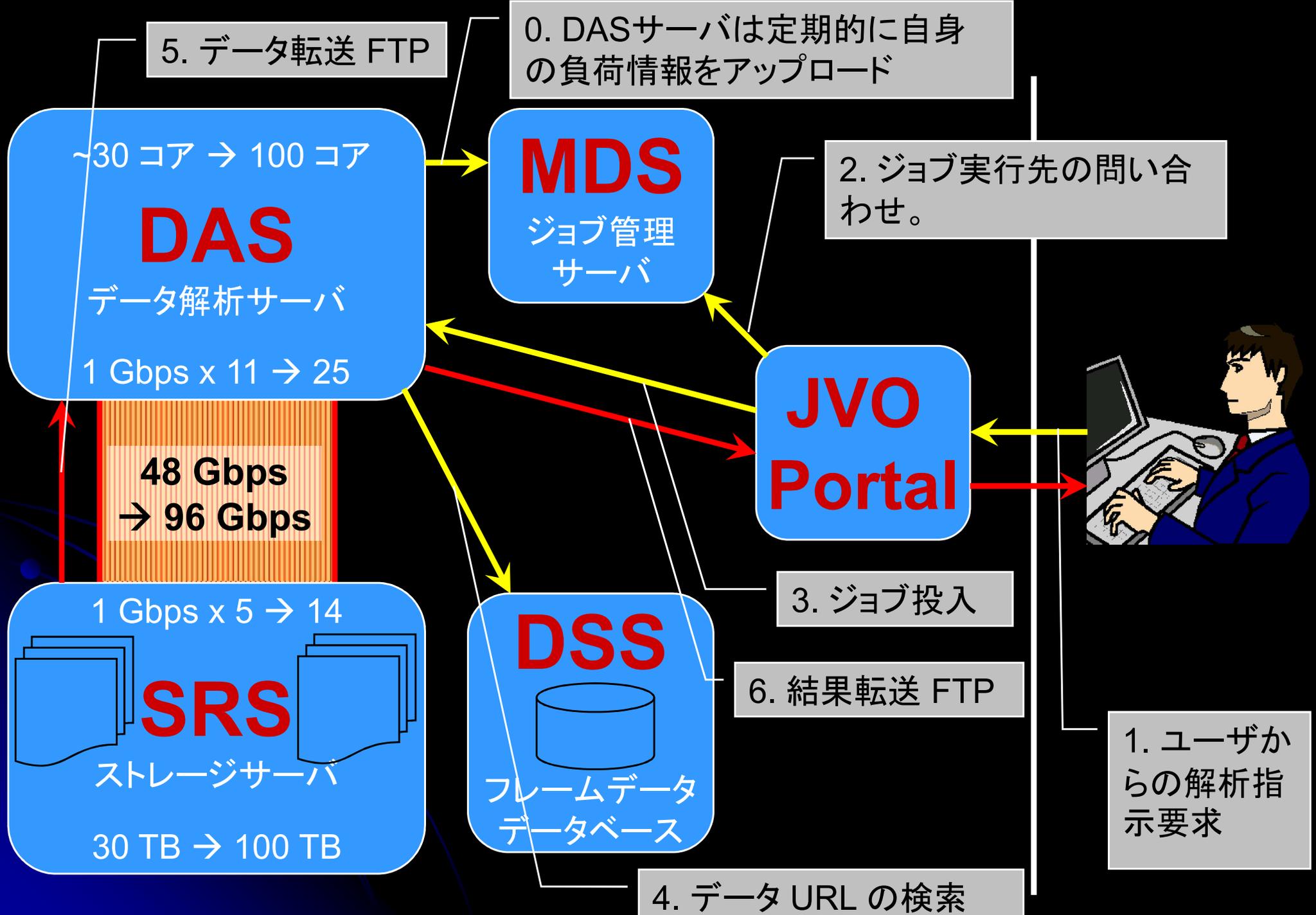
DSS

フレームデータ
データベース

6. 結果転送 FTP

1. ユーザからの解析指示要求

4. データ URL の検索



モザイクサービス GUI

Data Analysis - Mozilla Firefox

ファイル(F) 編集(E) 表示(V) 移動(G) ブックマーク(B) ツール(T) ヘルプ(H)

Mosaic

Status | Registry | Search | Workflow | Result | QSO | DataViewer | Tools | SubaruAnalysis | VOSpace | MDS | Usage | Logout

Target and Filter

Object Name: Filter:

クリック

Google サジェスト的な Object 名, フィルター名の指定

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z CLOSE

クリック

リストから複数選択し 一度に実行

Object	W-J-B	W-J-V	W-C-RC	W-C-IC	W-S-I+	W-S-Z+
A115a	<input type="checkbox"/> 0	<input checked="" type="checkbox"/> 60 S1	<input type="checkbox"/> 0	<input type="checkbox"/> 0	<input type="checkbox"/> 0	<input type="checkbox"/> 0
A115b	<input type="checkbox"/> 0	<input checked="" type="checkbox"/> 60 S1	<input type="checkbox"/> 0	<input type="checkbox"/> 0	<input type="checkbox"/> 0	<input type="checkbox"/> 0
A1401	<input type="checkbox"/> 0	<input checked="" type="checkbox"/> 40 S1	<input type="checkbox"/> 70 S1	<input type="checkbox"/> 0	<input type="checkbox"/> 70 S1	<input type="checkbox"/> 0
A1689	<input type="checkbox"/> 0	<input type="checkbox"/> 40 S1	<input type="checkbox"/> 0	<input type="checkbox"/> 0	<input type="checkbox"/> 60 S1	<input type="checkbox"/> 0
A1835	<input type="checkbox"/> 0	<input type="checkbox"/> 0	<input type="checkbox"/> 0	<input type="checkbox"/> 30 E1	<input type="checkbox"/> 0	<input type="checkbox"/> 0
A183500	<input type="checkbox"/> 0	<input type="checkbox"/> 0	<input type="checkbox"/> 0	<input type="checkbox"/> 50 E1	<input type="checkbox"/> 0	<input type="checkbox"/> 0

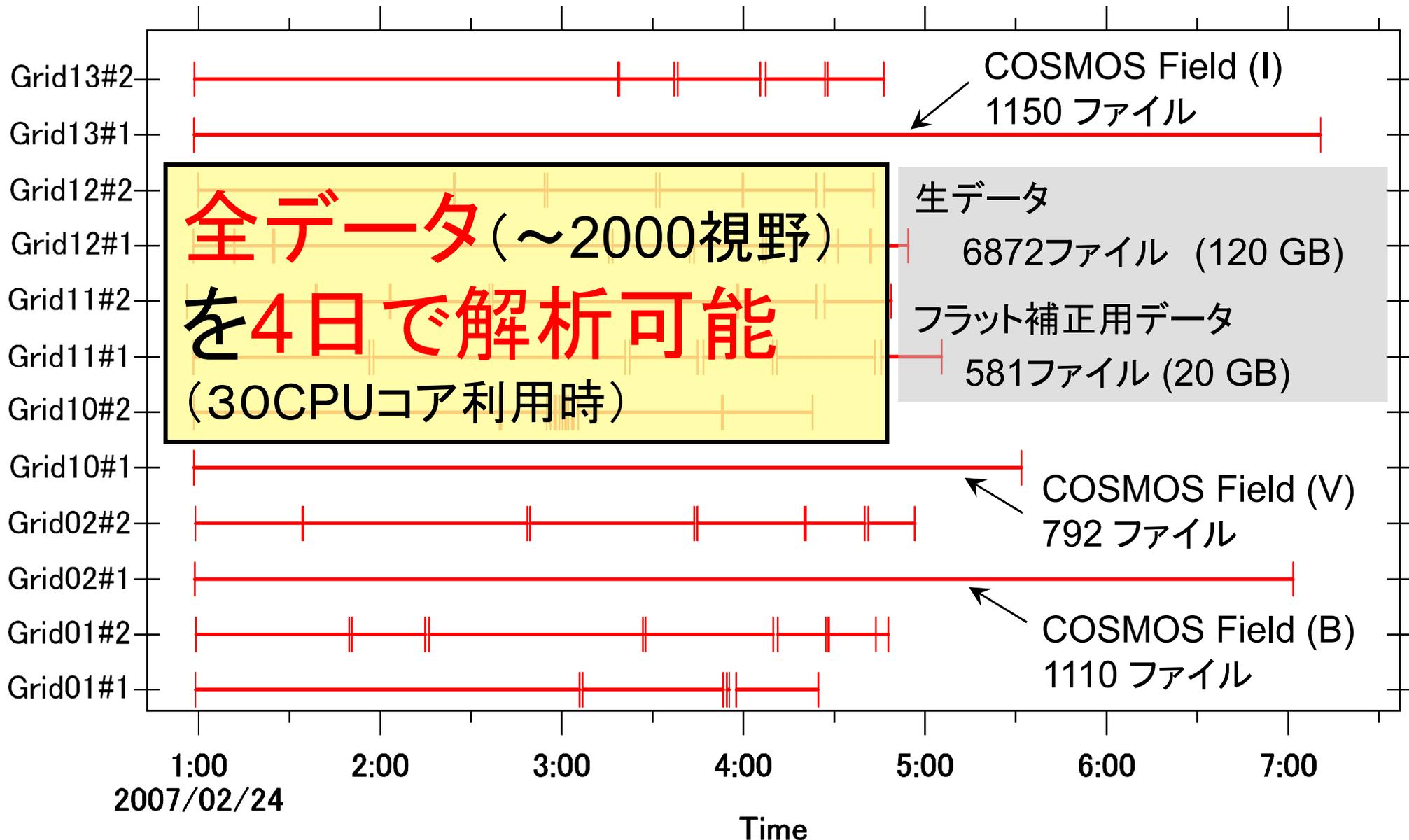
ジョブ実行ステータス画面

Job Status

job #	server id	job id	params	status	reason
0	ivo//jvo/server/grid12	36	mosaic.sh -o APPLES@@deep@@field@@1 -f W-J-B	running	
1	ivo//jvo/server/grid11	22	mosaic.sh -o APPLES@@deep@@field@@1 -f W-S-Z+	running	
2	ivo//jvo/server/grid02	40	mosaic.sh -o ABELL@@2125 -f W-S-Z+	running	
3	ivo//jvo/server/grid10	37	mosaic.sh -o ABELL@@708 -f W-S-I+	running	
4	ivo//jvo/server/grid13	47	mosaic.sh -o ABELL@@521 -f W-S-Z+	running	
5	ivo//jvo/server/grid03	2	mosaic.sh -o Abell@@1185 -f W-J-V	completed	
6	ivo//jvo/server/grid01	40	mosaic.sh -o Abell@@1185 -f W-C-IC	completed	
7	ivo//jvo/server/grid12	37	mosaic.sh -o And@@NE@@Field@@1 -f W-S-I+	running	
8	ivo//jvo/server/grid11	23	mosaic.sh -o And@@NE@@Field@@2 -f W-S-I+	running	
9	ivo//jvo/server/grid02	41	mosaic.sh -o And@@NE@@Field@@3 -f W-S-I+	running	

1 Interval: 100000 sec

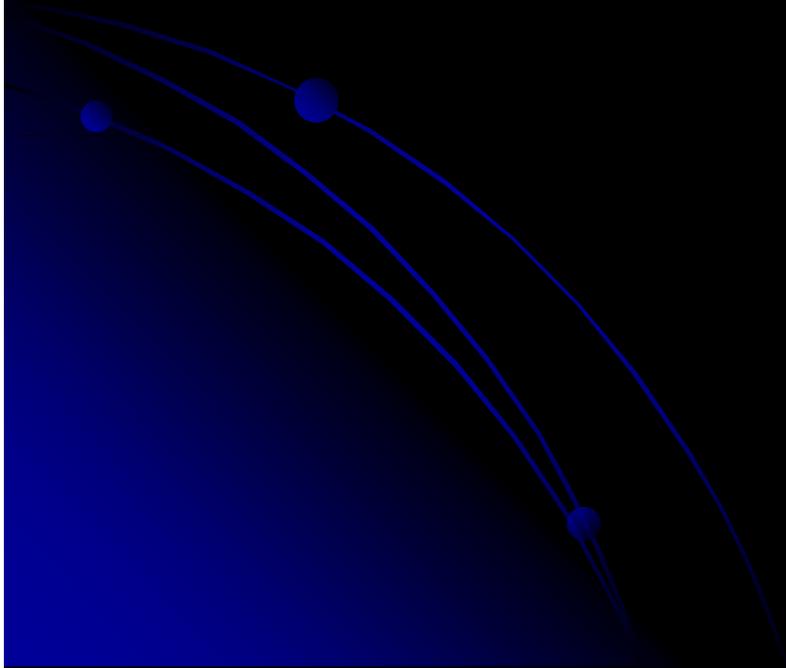
解析実行結果 (58視野/12CPUコア)



まとめと今後の課題

- **大規模データの並列解析システムを開発した**
 - 一晩分のデータ解析
 - 約1日(1CPU) → ~1時間(本システム)
 - 全データ(SuprimeCam 6年分)
 - 約半年(1CPU) → ~1週間程度(本システム)
 - JVO ポータル経由で解析実行可能
- **運用システムへ向けた課題**
 - ジョブ管理の高機能化
 - ユーザ毎の実行時間の割り当て、長時間ジョブ、ジョブの停止、中断、再開...
 - より使いやすいユーザインターフェイスの開発
 - バージョン管理: 過去の解析の再現性

おわり



解析??

- **生データのリダクション** (0次データ→ 1次データ)
 - 観測装置のレスポンスを補正 SuprimeCam フラット処理
 - 複数データの重ね合わせ SuprimeCam モザイクング
- **画像・スペクトルデータに対する測定** (2次データ)
 - 検出、測光 SExtractor (導入済み)
 - スペクトルラインパラメータ JVOSpec/VOSPec(川野元ポスター)
- **科学的知見を得るための処理** (3次データ?)
 - 色 vs 明るさ分布 JVOPlot / VOPlot (導入済み)
 - 赤方偏移 HyperZ (導入済み)
 - クロスマッチ (導入済み)
 - ワークフロー (田中ポスター)

フラットフレーム作成サービス GUI

The screenshot shows a Mozilla Firefox browser window displaying a web application for creating flat calibration frames. The main window is titled "Create Flat Calibration Frames" and contains the following elements:

- From:** 2002-04-06 **To:** 2002-04-15 **W-C-RC** **All**
- exptime:** 100 sec | **max frames:** 999 | **max hum**
- Submit Job** button
- Message:** `action=submitJob&start=2002-04-06&end=2002-04-15&wcr=W-C-RC&time=100&maxFrame=999`

A dropdown menu is open, showing the following options:

- All
- si001s (3)
- si002s (4)
- si005s (2)
- si006s (5)
- w4c5 (8)
- w67c1 (0)
- w6c1 (1)
- w7c3 (0)
- w93c2 (6)

In the background, a table of observation data is visible:

Date						
2002-04-03						
2002-04-04						
2002-04-05						
2002-04-06						
2002-04-07						
2002-04-08						
2002-04-09						
2002-04-10						
2002-04-11						
2002-04-12	0	0	0	0	38	0
2002-04-13	0	0	9	14	0	21
2002-04-14	0	0	9	5	0	26
2002-04-15	0	3	0	0	11	35
2002-04-16	0	0	0	0	21	0

At the bottom of the background window, there is a "Skip: 3 days" control and a "Message:" field with the URL: `action=searchObs&start=2002-4-3&limit=14&expTime=100&maxHum=100`

ジョブ管理情報

Registered Hosts

remove	enable	disable	name	living	enabled	load	numJob	lastTime	ID	address	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	arisa	false	false	0.0	0	2006-07-15 13:26:38	ivo://jvo/server/arisa	192.168.0.4	Intel(R) 2.40GHz
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	grid01	true	true	1.61	2	2006-10-12 14:00:28	ivo://jvo/server/grid01	192.168.0.64	AMD Core F
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	grid02	true	true	1.02	2	2006-10-12 14:00:04	ivo://jvo/server/grid02	192.168.0.66	AMD Core F
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	grid03	true	true	0.56	2	2006-10-12 14:00:16	ivo://jvo/server/grid03	192.168.0.67	AMD Core F
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	grid10	true	true	0.0	0	2006-10-12 14:00:26	ivo://jvo/server/grid10	192.168.0.68	AMD Core F
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	jvo-work01	false	false	0.0	0	2006-07-15 13:26:40	ivo://jvo/server/jvo-work01	192.168.0.3	Intel(R) 3.00GHz
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	jvo-work02	false	false	0.0	0	2006-07-15 13:26:47	ivo://jvo/server/jvo-work02	192.168.0.3	Intel(R) 3.00GHz
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	jvod	true	true	0.0	0	2006-10-12 13:59:35	ivo://jvo/server/jvod	192.168.0.5	Intel(R) 2.80GHz
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	jvoe	true	false	0.0	0	2006-10-12 13:59:48	ivo://jvo/server/jvoe	133.40.212.45	Intel(R) 2.80GHz
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	jvof	true	true	0.76	0	2006-10-12 14:00:08	ivo://jvo/server/jvof	192.168.0.1	Xeon(
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	jvoh	true	false	0.0	0	2006-10-12 14:00:09	ivo://jvo/server/jvoh	192.168.0.7	Dual C Proces
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	jvoi	true	true	1.31	1	2006-10-12 14:00:14	ivo://jvo/server/jvoi	192.168.0.8	Dual C Proces
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	jvoj	true	true	2.15	3	2006-10-12 13:59:52	ivo://jvo/server/jvoj	192.168.0.9	Dual C Proces
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	piglet	false	false	0.69	1	2006-09-11 17:24:39	ivo://jvo/server/piglet	133.40.208.47	AMD 4000+
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	tiger	true	false	0.0	0	2006-10-12 14:00:22	ivo://jvo/server/tiger	192.168.0.65	AMD Core F

Host Name

Heart Beat Status

Load Average

Number of Submitted Job