

Распределенная система поточковой обработки спутниковой информации

В.А. Кихтенко, В.В. Смирнов, Д.Л. Чубаров

Современные проблемы дистанционного зондирования Земли из космоса

13 ноября 2012

- **Архивная и оперативная обработка данных MODIS**
- **Автоматический режим обработки**
- **Высокая производительность**
- Простота изменения процедуры обработки
- Ведение журнала обработки
- Возможность восстановить всю цепочку генерации каждого продукта

- При наличии вычислительных ресурсов возрастает ценность исходной информации
- Сегодня не так много источников данных, способных породить сотни гигабайт в день. Данные MODIS – один из доступных источников

- 2 спутника — Terra и Aqua
- 2 точки приема — Новосибирск и Хабаровск
- 7-8 витков в светлое время суток
- Архив за 2009, 2010, 2011 и текущий годы
- Суммарный объем архива продуктов ~ 100 ТБ

Основные устройства:

1. Дисковый массив на 150 дисков
+ FC-NFS шлюз
2. Scalable NAS

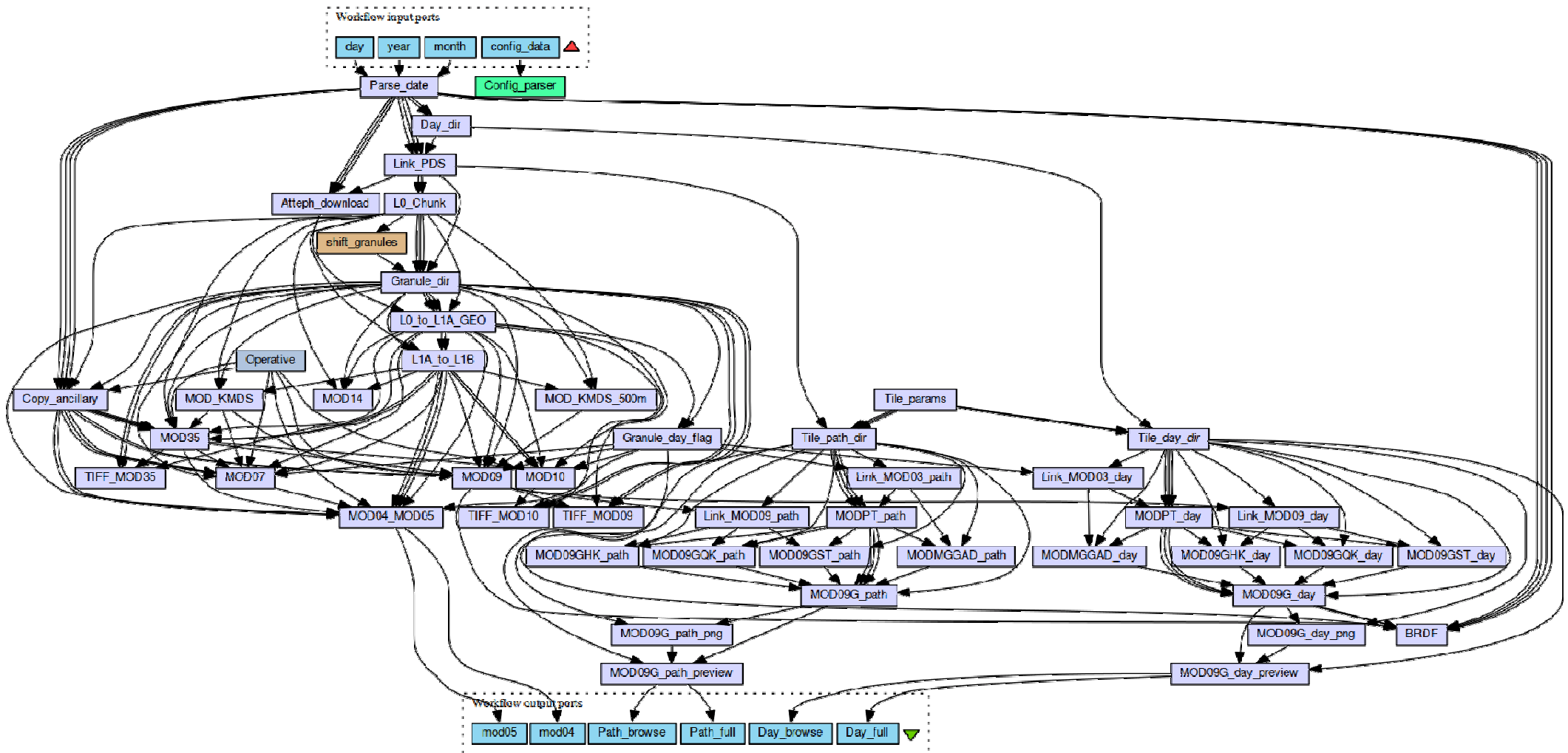
Вспомогательные устройства:

1. Дисковые массивы на 14-16 дисков
2. Шлюзы NFS-iSCSI

Основное вычислительное устройство

- 8 узлов суммарной производительностью около 1 Тфлопс,
2 x Opteron @ 2.66 GHz, ОЗУ 24 GB,
жесткий диск 72 GB
- 4 узла суммарной производительностью 590 Гфлопс,
2 x Xeon @ 3.06 GHz, ОЗУ 24 GB,
жесткий диск 900 GB RAID5 LV

Архивная цепочка MOD09 на основе PGE

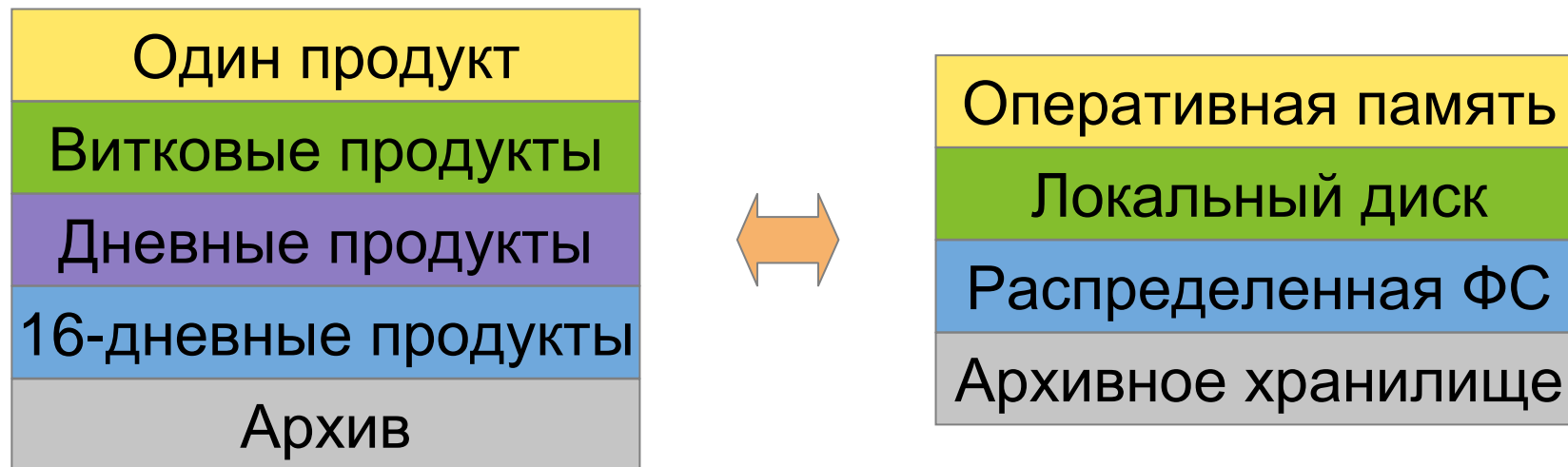


MOD09 – Land Surface Reflectance

Продукт NASA и Univ. of Maryland, E.F. Vermote

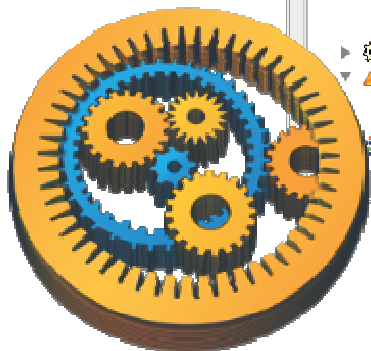
Количество потоков	Время обработки	
	Локальный диск	Сетевое хранилище
1	2:36	3:35
4	0:43	1:31
12	0:32	0:48

1. Создание иерархии уровней хранения
2. Согласование иерархию уровней с процедурой обработки



The screenshot displays the Taverna Workbench 2.2.0 interface. On the left, the **Service panel** is active, showing a search filter for "protein" and a list of matching services under "Local services" and "Biomart". The "Workflow explorer" below it shows a tree view of the workflow components, including "BiomartAndEMBOSSAnalysis", "Workflow input ports", "Workflow output ports", and "Services".

On the right, the **Workflow diagram** shows a flowchart of the workflow. It starts with the "hsapiensGeneEnsembl" service, which feeds into the "GetUniqueHomolog" service. This service then branches into three parallel services: "getHSapSequence", "getRNorSequence", and "getMMusSequence". These three services feed into the "CreateFasta" service. The output of "CreateFasta" goes through a sequence of services: "seqret", "emma", and "plot". The "plot" service feeds into the "FlattenImageList" service. Finally, the "FlattenImageList" service feeds into a dashed box labeled "Workflow output ports", which contains four output services: "outputPlot", "HSapIDs", "MMusIDs", and "RNoriDs".



- Существует с 2001 года
- Свободное программное обеспечение
- Кросс-платформенные технологии Java
- Несколько десятков разработчиков модулей

Taverna Workbench 2.2.0

Design Results myExperiment BioCatalogue

Workflow runs Remove

Click on a run to see its values
Click on a service in the diagram
to see intermediate values (if available)

lists_iterate 2010-11-29 14:15:15
 BiomartAndFMBOSSAnalysis 2010-11-29 14:13:35
 BiomartAndEMEOSSAnalysis 2010-11-29 14:13:27
 BiomartAndEMEOSSAnalysis 2010-11-29 14:13:09
 Spreadsheet_Import_Example 2010-11-04 17:26:41
 Workflow65 2010-10-27 16:46:05
 Workflow65 2010-10-27 16:43:29
 Workflow65 2010-10-27 16:40:45
 Workflow65 2010-10-27 16:40:09
 Workflow65 2010-10-27 16:39:16
 Workflow65 2010-10-27 16:37:56
 Workflow65 2010-10-27 16:36:32
 Workflow64 2010-10-27 16:35:34
 Workflow65 2010-10-27 16:33:23
 Workflow65 2010-10-27 16:33:05
 Workflow65 2010-10-27 16:32:40
 Workflow65 2010-10-27 16:32:26
 Workflow65 2010-10-27 16:31:20

Graph Progress report

Name	Status	Queued itera...	Iterations done	Iterations w/...	Average time...	First iteration...	Last iteration...
lists_iterate	Running	-	1	-		14:15:49	
boo - ,b	Finished	0	1	0	23 ms	14:15:50	14:15:50
Concat	Running	160	39	0	139 ms	14:15:54	
Concat	Running	0	39	0	31 ms	14:15:54	
Concat	Running	2	36	0	38 ms	14:15:54	
Concat	Running	2	33	0	30 ms	14:15:54	
List_Em	Finished	0	1	0	1.8 s	14:15:51	14:15:53
sdfds -	Finished	0	1	0	176 ms	14:15:50	14:15:50
xxx - X)	Finished	0	1	0	17 ms	14:15:50	14:15:50

Concatenate_two_strings2
 Working output ports
 out

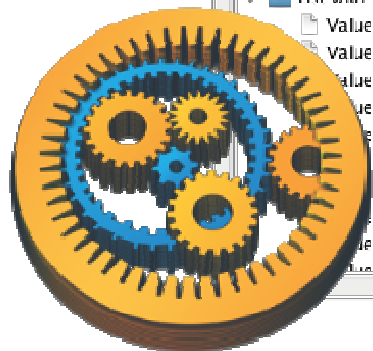
Running Pause Cancel Refresh intermediate values Show workflow results Save all values

Workflow results Save all values

Click in tree to view values

Value type Refresh Save value

- List with 101 values
 - Value 1
 - Value 2
 - Value 3
 - Value 4
 - Value 5
 - Value 6
 - Value 7
 - Value 8
 - Value 9
 - Value 10

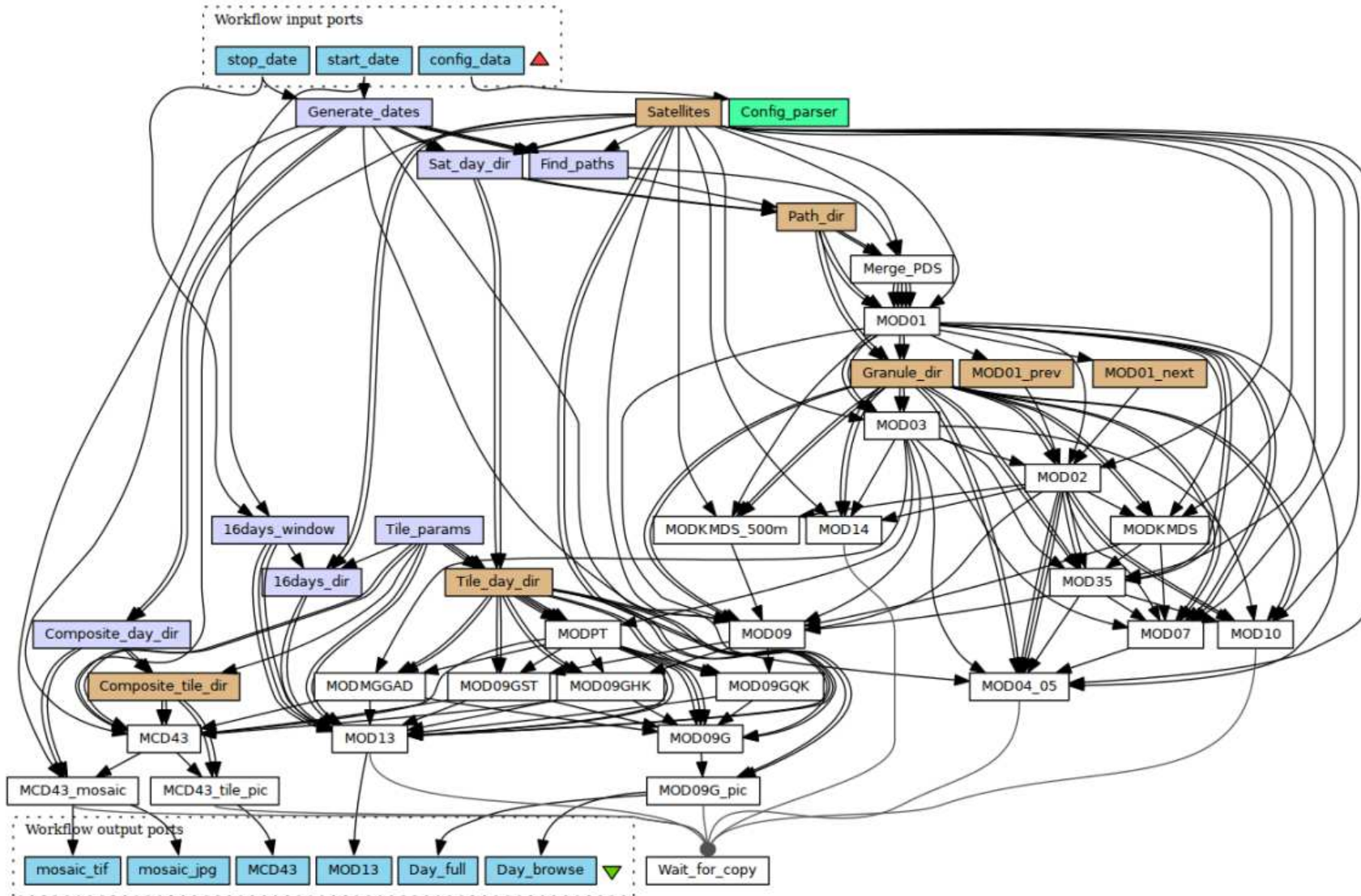


- Строки (in, out)
- Конфигурационные параметры (in)
- Указатели на файлы (in, out)
- Рабочие директории (in)

Задание параметров
процесса обработки

Запуск PGE на
кластере

Запуск BASH
сценария



- PGE запускается на том узле, где доступны все необходимые для этого данные: имеется на локальном диске ИЛИ на сетевой системе хранения
- Копирование происходит асинхронно.
- Специальный блок позволяет при необходимости синхронизировать данные с сетевой системой хранения

- Блок обработки, параметром которого является командный сценарий
- Входные порты (строки) передаются в сценарий как аргументы
- [Оптимизация ввода-вывода] Указатели на файл транслируются в имена файлов

- Запуск заданий для нескольких распределенных кластерных планировщиков: SLURM, Torque, Mesos
- Обработка отчетов планировщика о ходе исполнения
- Установка ограничений на выбор свободного узла, пригодного для запуска задания

Данные Terra + Aqua

Генерация продуктов по всему перечню за несколько дней.

Продуктов за ...	1 день	16 день	30 дней
1 x Xeon node	80 мин	17 ч., 40 мин	
4 x Xeon node	46 мин	7 ч. 40 мин	
1 x Opteron node	110 мин		
8 x Opteron node	85 мин	9 ч. 50 мин	19 часов

Все продукты за 1 год могут быть пересчитаны в течение 10 дней

- Модификации PGE, предложенные А.А. Лагутиным (АлтГУ, ИВТ СО РАН)
- Ожидаемое время генерации витковых продуктов – 20 мин.

- Установка Taverna вместе с дополнительными модулями
- Загрузка сценария обработки
- Запуск задания
- При условии, что есть кластер с совместимым планировщиком и данные в сетевой системе хранения

- Доступ к планировщику только из локальной сети ИВТ СО РАН
- Продукты доступны по FTP, объем за один день – 63 G
- ☺