

## Объектная модель задач в Grid-системе обработки спутниковых данных

Шелестов А.Ю.

Институт космических исследований НАНУ-НКАУ  
email: inform@ikd.kiev.ua

### **Abstract**

*Shelestov A.Yu. Object Model of Jobs in Grid System for Satellite Data Processing. In this paper a problem of Grid system decomposition is solved through the construction of object model. Unified Modelling Language (UML) is used as a formalization tool. This approach is motivated by the complexity of the system and the need for simulation model design.*

### **Введение**

В последние годы проблемы разработки Grid-систем привлекают все более пристальное внимание ученых и специалистов в Украине и во всем мире. В частности, анализ современного состояния и мировых тенденций развития подобных распределенных систем обработки информации приводится в [1] и [2]. Развитию Grid-систем в Украине посвящены, в частности, работы [3], [4] и [5].

Поскольку Grid-система является сложной многомерной системой, обеспечивающей решение сложных задач за счет эффективного использования вычислительных и информационных ресурсов, к ее разработке следует подходить с позиций системного анализа. Особенно это касается Grid-систем экологического мониторинга, связанных с использованием спутниковых данных и наземных измерений, результатов их обработки и моделирования в различных масштабах. Такая система не просто является вычислительной Grid-средой или системой управления данными, а представляет собой виртуальную организацию, решающую вычислительно сложные задачи с использованием данных из географически распределенных архивов с заданной надежностью (или качеством обслуживания QoS) и необходимым уровнем безопасности. Примерами такого класса систем являются международные инициативы GMES (Global Monitoring for Environment and Security) [6], GEOSS (Global Earth Observation System of Systems) [7] и WAG (Wide Area Grid) [1], в которых участвует Украина. В процессе анализа таких систем в первую очередь возникает задача декомпозиции системы и ее структурно-функционального анализа с целью последующей системной оптимизации и

эффективного управления [8].

В данной статье задача декомпозиции Grid-системы обработки спутниковых данных решается с помощью построения ее объектной модели. В качестве средства формализации используется унифицированный язык моделирования UML (Unified Modeling Language). Такой способ моделирования обусловлен сложностью аналитического описания структуры и функциональности системы, а также необходимостью последующего построения имитационной модели.

### ***Свойства Grid-систем обработки спутниковых данных***

Grid-система представляет собой объединение вычислительных, информационных и других ресурсов, принадлежащих к разным административным доменам. Традиционно такие системы предназначались для решения вычислительно сложных задач, требующих использования высокопроизводительных кластеров и суперкомпьютеров, или задач, связанных с многократным выполнением однотипных функций для больших массивов входной информации. К числу таких систем относятся проекты EGEE (<http://www.eu-egee.org>) для решения задач физики высоких энергий [9] и гравитационных волн [10], астрономии [11] или биоинформатики [12]. Такие Grid-системы получили название вычислительных.

Однако в последнее время все больше внимания уделяется развитию Grid-систем экологического мониторинга на основе обработки спутниковых данных. Обзор наиболее значительных проектов в этой области приводится в [2].

Grid-система мониторинга или поддержки принятия решений на основе наблюдения Земли из космоса может использоваться не только для организации высокопроизводительных вычислений, но и как инфраструктура для управления большими объемами поступающих данных [13], которые необходимо обрабатывать, каталогизировать и помещать в соответствующие архивы. Например, инструмент GOME, размещенный на спутнике Envisat и предоставляющий информацию о содержании озона на разных высотах атмосферы Земли, генерирует в общей сложности около 400 терабайт данных в год [14]. Система распространения спутниковых данных EUMETCast, являющаяся составной частью глобальной сети GEONETCast системы GEOSS [15], позволяет получать более 50 терабайт в год, содержащих как необработанные данные, так и сгенерированные продукты. При этом, как правило, обработка спутниковых данных выполняется не одним приложением с «монолитным» кодом, а несколькими, в общем случае распределенными, приложениями. Т.е. такой процесс можно рассматривать как сложный (комплексный) процесс выполнения (workflow), состоящий из множества подзадач, связанных с геометрической и радиометрической коррекцией

данных, фильтрацией, перепроектированием, созданием композитов изображений, классификацией, тематической обработкой, пост-обработкой, визуализацией и т.д. [16]. Например, калибровка и создание мозаики из 80 радарных изображений, полученных прибором ASAR спутника Envisat, занял 3 дня на 10 рабочих узлах системы Earth Science GRID on Demand, разрабатываемой Европейским космическим агентством и Европейским институтом космических исследований.

Таким образом, в силу своей сложности, Grid-системы мониторинга и поддержки принятия решений на основе спутниковых данных представляют особый интерес с точки зрения системного анализа. На исследование специфики подобных систем направлен международный проект DEGREE [17], инициированный в рамках программы EGEE-II. Как следует из опыта авторов данной статьи и отчетов по проекту DEGREE, системы экологического и геофизического мониторинга на основе спутниковых данных характеризуются набором серьезных требований, которые можно сформулировать следующим образом.

— Требования к безопасности. Многие данные спутниковых наблюдений и специализированное программное обеспечение для их обработки распространяются согласно различным лицензионным соглашениям (в том числе, коммерческим), поэтому строгое соблюдение политики безопасности является важнейшим условием работы подобных систем.

— Требования к надежности. Системы мониторинга должны функционировать в бесперебойном режиме, поэтому непреложным требованием является обеспечение соответствующего уровня обслуживания (QoS — Quality of Service), например, предоставление результатов мониторинга или прогнозов развития ситуации к заданному времени.

— Требования к стандартизации. Безусловным требованием работы систем мониторинга и поддержки принятия решений является использование сертифицированных методик обработки данных и верифицированных моделей. Не менее жесткие требования по стандартизации предъявляются и к используемым данным и протоколам их обмена: зачастую решение одной задачи предполагает применение данных разной природы, полученных с различным временным и пространственным разрешением.

Эти требования определяют чрезвычайно высокую гибкость процесса выполнения задачи (workflow) и необходимость эффективного управления ресурсами на физическом и логическом уровне. Важнейшей задачей разработки таких систем является выполнение их структурно-функционального анализа на основе моделирования.

### ***Необходимость моделирования Grid-систем***

Необходимость моделирования работы Grid-систем, особенно связанных с обработкой спутниковых данных, определяется следующими целями.

Поскольку создание Grid-инфраструктуры требует существенных капиталовложений, моделирование работы системы на этапе ее проектирования позволит разработать оптимальную структуру распределенной системы с учетом решаемых ею задач.

Не менее важную роль играет моделирование работы уже существующей системы. Построение модели нагрузки системы позволяет выявить узкие места этой системы, оценить существующие резервы и возможность добавления в систему новых задач, а также спланировать направление и темпы развития системы.

И наконец, моделирование работы системы является неотъемлемой составляющей процесса прогнозирования производительности и разработки механизмов планирования решения задач в системе на основе эффективного управления ресурсами. Под управлением ресурсами понимается процесс, состоящий из следующих этапов: определение параметров вычислительных ресурсов, удовлетворяющих потребностям задачи, поиск соответствующих ресурсов, резервирование, планирование выполнения задач и мониторинг. Это, в свою очередь, приводит к необходимости разработки алгоритмов управления ресурсами и планирования их использования, что в свою очередь, требует многократного проведения повторяемых и управляемых экспериментов. Выполнение разных сценариев экспериментов может потребовать наличия сотен пользователей с наперед заданным поведением. Проведение таких экспериментов в реальной системе практически невозможно, в первую очередь, в силу распределенной и гетерогенной природы Grid-систем. Рассмотрим эти факторы более подробно.

— Вычислительные ресурсы Grid-системы принадлежат разным владельцам, что усложняет управление экспериментом.

— Свойства и потребности пользователей и ресурсов непостоянны во времени, что делает невозможным повторное проведение идентичного эксперимента.

— Обычно размеры существующих Grid-полигонов недостаточны для исследования сложных алгоритмов поведения системы.

— Создание работоспособной Grid-инфраструктуры для потребностей конкретного эксперимента является длительным и ресурсоемким процессом, требующим значительных капиталовложений.

Поэтому моделирование является единственно возможным подходом к исследованию Grid-инфраструктуры и анализу поведения как системы в целом, так и ее отдельных ресурсов и задач (или пользователей).

### ***Подходы к моделированию Grid-систем***

Для моделирования различных аспектов Grid-систем применяются разные подходы, среди которых преобладают попытки построения аналитических или статистических моделей [18]. Как правило, свойства элементов и

систем удается эффективно представить в аналитической форме с учетом предположений о независимости одних параметров от других, линейности некоторых зависимостей, мгновенности переходов между состояниями и т.п. Если эти предположения соответствуют действительности, модель хорошо описывает зависимости между характеристиками и параметрами системы. Однако во многих случаях принимаемые допущения приводят к существенным отличиям модели от реального объекта и существенным ошибкам в описании систем.

Поэтому для моделирования сложных распределенных систем, к которым относятся Grid-системы обработки спутниковых данных, применяют имитационные модели, которые воссоздают процесс функционирования исследуемой системы, исходя из известных свойств ее элементов. Это осуществляется путем объединения элементов в структуру, соответствующую исследуемой системе и имитации функционирования элементов при их взаимодействии. Главным преимуществом метода имитационного моделирования является его универсальность. Имитационное моделирование позволяет исследовать системы произвольной степени сложности и не ограничивает уровень детализации модели. Алгоритмически можно воспроизвести сколь угодно сложные взаимосвязи между элементами системы и процессы их функционирования. При этом имитационная модель является неисчерпаемым источником данных о поведении исследуемой системы — новые эксперименты на модели позволяют получать дополнительную информацию о системе. Благодаря этому достигается детальная оценка характеристик функционирования как системы в целом, так и ее составляющих. Как правило, увеличивая длительность или число экспериментов на моделях, можно достичь высокой точности результатов. Поэтому методы имитационного моделирования широко используются в теоретических исследованиях и при проектировании информационно-вычислительных систем. Имитационные модели позволяют разработчику сформировать представление о свойствах системы и принять обоснованные проектные решения на основе результатов моделирования. Поэтому для моделирования работы создаваемой Grid-системы обработки спутниковых данных будем использовать метод имитационного моделирования.

### ***Классификация задач в Grid-системе обработки спутниковых данных***

При моделировании сложных распределенных систем, в частности Grid-систем, одной из важнейших задач является описание и моделирование задач (task modeling), выполняемых в системе. Данные о задачах обычно служат входными параметрами при описании работы системы в целом и при моделировании ее рабочей нагрузки с целью оценки и

прогнозирования производительности, планирования выполнения задач на ресурсах системы и т.д. Поэтому адекватное описание параметров различных типов задач позволит эффективно моделировать работу системы и ее рабочую нагрузку. Учитывая приведенные выше особенности задач и сложность процессов их выполнения в системах обработки спутниковых данных, рассмотрим основные типы задач в таких системах и построим их объектную модель. Задачи, решаемые в Grid-системах обработки спутниковых данных, будут проанализированы с учетом опыта разработки такой системы в Институте космических исследований НАНУ-НКАУ.

В системах обработки спутниковых данных можно выделить два основных типа элементарных задач: задачи передачи данных и вычислительные задачи, связанные с обработкой данных или запуском моделей. Эти элементарные задачи являются «строительными блоками» сложных процессов выполнения, обеспечивающих комплексное решение задач мониторинга или поддержки принятия решений.

**Задачи передачи данных (Data Transfer Task — DTT).** Задачи, связанные с передачей данных, характеризуются, в первую очередь, объемом передаваемой информации. Этот параметр определяет следующие требования к системе: пропускную способность канала (соединения с Internet или в локальной сети), характеристики накопителей (жестких дисков или магнитных лент), определяющие объем и скорость чтения/записи информации. Рассматривая задачу передачи данных, необходимо учитывать источник получения данных, необходимую частоту решения этой задачи а также требования к процессу передачи.

Моделирование задач, связанных с передачей данных, позволяет обеспечить эффективность управления сетевыми ресурсами системы и определить оптимальные (или минимально необходимые) параметры сети. Задачу передачи данных можно описать следующими параметрами:

- идентификатор задачи;
- частота выполнения (например, по запросу пользователя или другой задачи, регулярно с заданной периодичностью);
- объем входных и выходных данных;
- источник(и) данных (например, Internet, локальная сеть, жесткий диск, магнитные ленты, ...).

**Вычислительные задачи (Computational Task — CT).** Под вычислительной задачей будем понимать элементарный (неделимый) фрагмент программы, выполняющий обработку данных или другие вычисления, например, описывающие одну итерацию работы некоторой модели. Одна элементарная вычислительная задача может выполняться на одном процессоре либо распараллеливаться на несколько процессов, выполняемых на одной или нескольких одно- или многопроцессорных машинах. Таким образом, комплексная задача мониторинга или поддержки

принятия решений может быть описана в виде направленного графа, вершинами которого являются элементарные вычислительные задачи и задачи передачи данных.

Вычислительные задачи характеризуются следующими параметрами:

- идентификатор задачи;
- *частота выполнения* (например, по запросу пользователя или другой задачи, регулярно с заданной периодичностью);
- *сложность задачи* (вычислительная сложность, требуемый объем оперативной памяти, дополнительные требования к программным или аппаратным ресурсам);
- *уровень параллелизма* (возможность распараллеливания задачи):
- параллелизм: по коду или по данным;
- требуемое количество процессоров.

Отметим, что еще одним типом элементарных задач являются задачи управления (поиска в базе данных или выполнения других управляющих операций). Но эти задачи можно отнести к классу вычислительных.

### **Формальное описание задач в Grid-системе обработки спутниковых данных**

Перейдем к формальному описанию приведенных типов задач. Предварительно опишем структуру данных Data. Данные в Grid-системе можно описать следующим кортежем:

$$\text{Data} = \{\text{ID}, \text{V}, \text{DS}, \text{Sec}\}, \quad (1)$$

где ID — идентификатор, V — объем данных (Мб), DS — источник данных, представляющий собой локальный диск, специализированное хранилище данных, ресурс в локальной сети или Internet. Параметр DS может принимать значения из следующего множества:

$$\text{DS} = \{\text{local disc}, \text{Internet}, \text{data storage}, \text{local network}\}.$$

Значение параметра Sec в (1) задает уровень безопасности, необходимый для доступа к данным. Конкретные значения параметров безопасности в данной работе рассматриваться не будут, так как они не влияют на моделирование производительности системы и играют роль только при определении политики безопасности в системе.

**Задачи передачи данных.** Задачу передачи данных опишем следующим кортежем:

$$\text{DTT} = \{\text{ID}, \text{Freq}, \text{I/O: Data}\}, \quad (2)$$

где ID — идентификатор задачи, Freq — частота выполнения задачи, задаваемая одним из элементов

$$\text{Freq} = \{\text{cycle}, \text{request}\}. \quad (3)$$

Если задача выполняется периодически, Freq = cycle, а конкретное значение этого параметра определяет период запуска задачи в (например, в минутах). Если задача выполняется по запросу пользователя, Freq = request, при этом параметры запроса при моделировании должны

задаваться некоторым вероятностным распределением.

Рассмотрим пример описания задачи получения данных.

**Пример 1.** В качестве примера рассмотрим задачу получения данных глобальной модели GFS, используемых для настройки краевых условий при работе метеорологической модели WRF:

$$DTT = \{GFS\_boundary, cycle/360, I/O: 20 \text{ Мб}, Internet\}. \quad (4)$$

Из этого описания видно, что задача запускается периодически раз в 6 часов, объем выходных данных составляет 20 Мб, причем данные поставляются через Internet.

**Вычислительные задачи.** С учетом приведенного выше списка параметров вычислительная задача описывается кортежем:

$$CT = \{ID, Freq, C, Par\}. \quad (5)$$

В (5) приняты следующие обозначения: ID — идентификатор задачи, Freq — частота выполнения, задаваемая согласно (3), C — сложность задачи, описываемая тройкой параметров

$$C = \{CC, Size, Op\}, \quad (6)$$

где CC — вычислительная сложность задачи (требующая отдельного рассмотрения), Size — требуемый объем памяти, Op — дополнительные ограничения (в том числе, требования к программным или аппаратным ресурсам).

Последним элементом кортежа (5) является Par, определяющий возможность распараллеливания задачи и тип параллелизма Parallelism (по коду или по данным), а также требуемое количество процессов NumProcess. Принимая во внимание описание двух основных типов задач, можно построить иерархию наследования, определяющую объектную модель задачи в Grid-системе мониторинга или поддержки принятия решений на основе спутниковых данных. Диаграмма классов этой объектной модели на языке UML приведена на рис. 1.

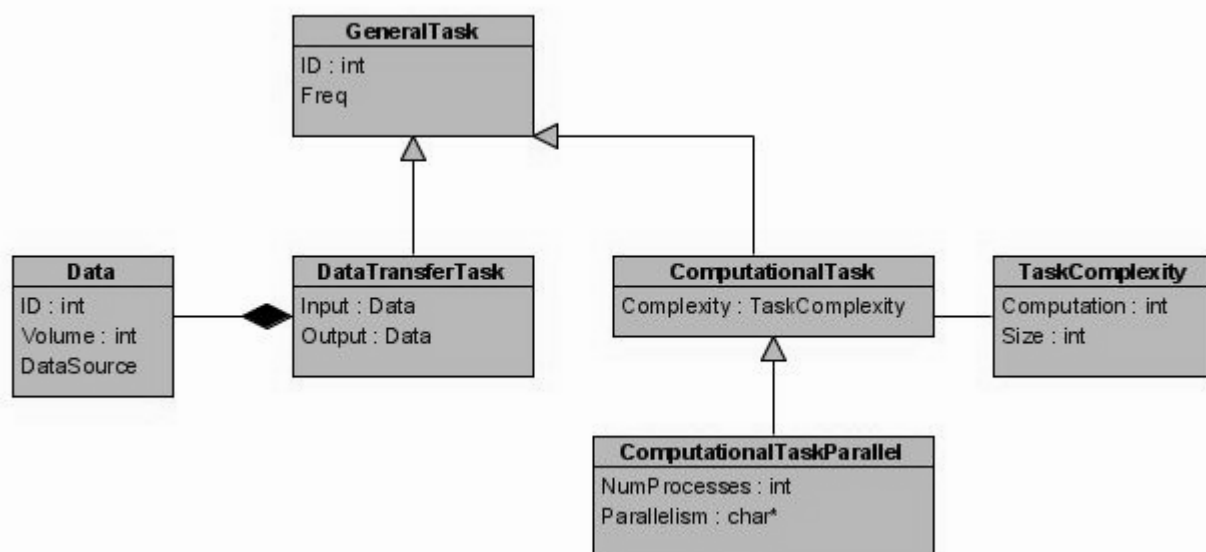


Рисунок 1 – Диаграмма классов объектной модели задачи в Grid-системе обработки спутниковых данных



Без снижения общности модели в ней приняты некоторые упрощения, а именно, уровень безопасности для доступа к данным и дополнительные ограничения вычислительной задачи на диаграмму не вынесены.

### ***Представление комплексных заданий в Grid-системе обработки спутниковых данных***

Формальное представление элементарных задач, полученное в предыдущем разделе, обеспечивает основу для представления комплексных заданий (job) в Grid-системе. Задание (job) представляет собой заданную последовательность (или процесс) выполнения отдельных задач (workflow). Для описания заданий будем использовать ациклический направленный граф. Вершинами такого графа являются элементарные вычислительные задачи или задачи передачи данных, а ребра указывают последовательность выполнения задач. Таким образом, задание определяется следующим набором параметров:

$$\text{Job} = \{S, W, \text{QoS}\},$$

где  $S = \text{ST} \cup \text{DTT}$  — множество элементарных задач (включая вычислительные и задачи передачи данных),  $W \subset S \times S$  — множество пар вершин  $(s_i, s_j)$ , между которыми существует ребро, направленное от вершины  $s_i$  к  $s_j$ , QoS — уровень качества обслуживания. К показателям QoS могут относиться максимальный временной интервал между временем «запуска» задачи и начала выполнения (TTS — time-to-schedule), количество повторных запусков задачи в случае возникновения ошибок и т.д.

### ***Пример построения объектной модели для задачи оценки биоразнообразия на основе спутниковых данных***

В рамках выполнения совместного инновационного проекта Институтом космических исследований НАНУ-НКАУ и Центром аэрокосмических исследований Земли НАНУ был реализован информационный сервис оценки биоразнообразия Причерноморского региона Украины, что соответствует одному из приоритетных направлений развития системы GEOSS [19].

Биоразнообразие ассоциируется с количеством видов как отдельной группы, так и ландшафта в целом. Видовое биоразнообразие характеризуется двумя критериями. Первая составляющая — видовое богатство, т.е. общее количество существующих видов. Второй важный аспект разнообразия — равномерность распределения видов, которое базируется на определении положения вида в структуре доминирования на основе мер значимости [20]. Разработка алгоритмов оценивания и картографирования биологического разнообразия на основе многоспектральных космических снимков и геоинформационных

технологий позволит обеспечить установления особенностей пространственного распределения биоразнообразия на территории Украины.

В процессе вычисления биоразнообразия одной из важнейших задач является обеспечение регулярного приема больших объемов разнородных данных (в разных форматах, проекциях и с разным пространственным разрешением) из различных источников. В таблице 1 приведена информация о получаемых продуктах и источниках данных.

Таблица 1. Исходные продукты и источники данных

Название продукта	Описание	Размер файла	Период обновления
MOD03A2	геолокационные данные	30 МВ	ежедневно
MOD05L2	водосбор	6 МВ	полусутки – день и ночь
MOD11A1	температура поверхности	20 МВ	полусутки – день и ночь
MOD12Q1	типы ландшафта	3 МВ	96 суток и ежегодно
MOD13Q1	вегетационные индексы	30 МВ	16 суток и ежемесячно
MOD15A2	лиственный покров	5 МВ	16 суток, ежемесячно и ежегодно
MOD17A3	продуктивность растительности	7 МВ	8 суток и ежегодно
AE_Land3	влажность грунтов	<1 МВ	Ежемесячно
SRTM DEM	числовая модель рельефа	1,5 Гб	Однократно

Данные о числовой модели рельефа SRTM DEM могут быть получены однократно. Все остальные продукты должны загружаться на регулярной основе. С этой целью был разработан сценарий загрузки данных, который проверяет информацию о новых продуктах, выполняет загрузку данных, помещает данные в хранилище и обновляет сервис индексации данных. Соответствующая диаграмма последовательностей UML для сценария загрузки данных приведена на рис. 2.

На основе предложенного выше подхода к разбиению комплексных заданий на несколько элементарных задач эти сценарии можно представить в виде следующих графовидных структур (рис. 3).

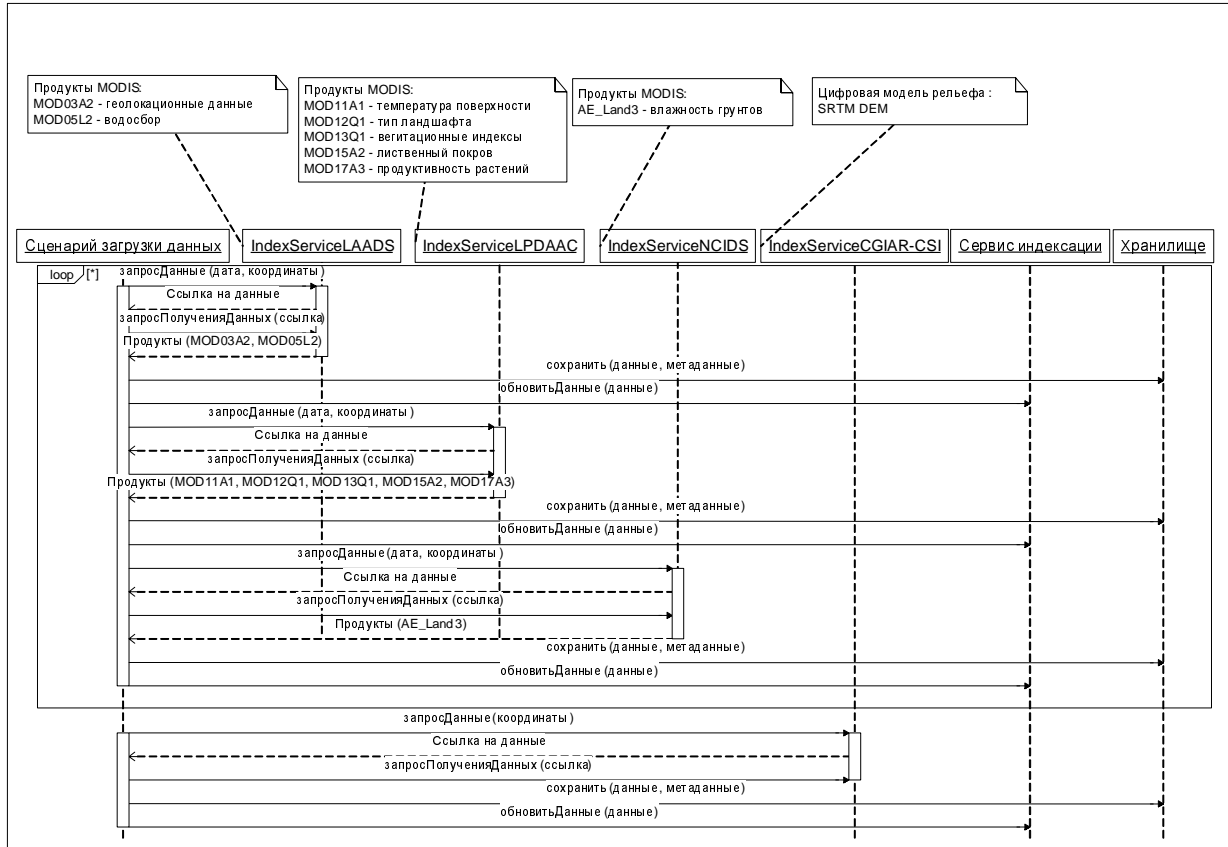


Рисунок 2 – Диаграмма последовательностей для сценария загрузки данных для сервиса оценки биоразнообразия на языке UML

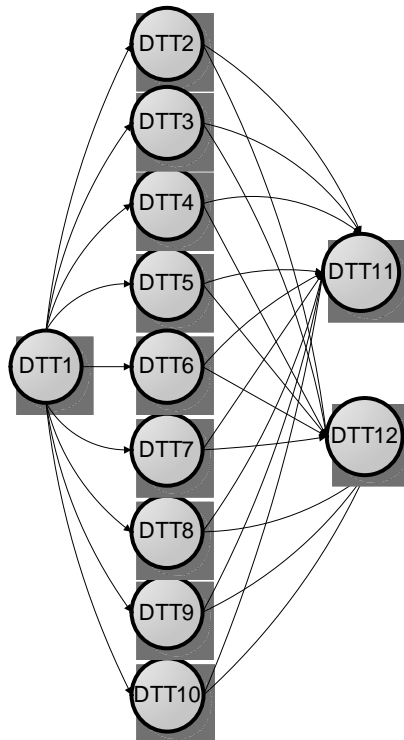


Рисунок 3 – Граф выполнения задания получения данных для сервиса оценки биоразнообразия

Соответствующие обозначения на диаграммах приведены в таблице 2.

Таблица 2. Обозначения для графа выполнения заданий получения данных, обработки данных и визуализации результатов обработки

Название	Описание	Структура задачи
DTT1	Начало загрузки	
DTT2	Загрузка данных (ЗД) MOD03A2	{cycle/24 ч, I/O: 30 Мб, Internet}
DTT3	ЗД MOD05L2	{cycle/12 ч, I/O: 6 Мб, Internet}
DTT4	ЗД MOD11A1	{cycle/12 ч, I/O: 20 Мб, Internet}
DTT5	ЗД MOD12Q1	{cycle/0,5 год, I/O: 3 Мб, Internet}
DTT6	ЗД MOD13Q1	{cycle/16 дн, I/O: 30 Мб, Internet}
DTT7	ЗД MOD15A2	{cycle/16 дн, I/O: 5 Мб, Internet}
DTT8	ЗД MOD17A3	{cycle/8 дн, I/O: 7 Мб, Internet}
DTT9	ЗД AE_Land3	{cycle/1 мес, I/O: 30 Мб, Internet}
DTT10	ЗД SRTM DEM	{request, I/O: 1,5Гб, Internet}
DTT11	Сохранение данных (в хранилище)	
DTT12	Обновление данных (в сервисе индексации)	

Как видно из рис. 2, 3 даже простая задача загрузки данных при решении задач мониторинга окружающей среды описывается нетривиальной объектной моделью. Это связано с использованием разнородных источников информации, что характерно практически для любых задач мониторинга на основе спутниковых данных и данных наземных измерений. Построенная объектная модель обеспечивает возможность имитационного моделирования работы Grid-системы, а также позволяет формировать комплексные задания для выполнения в Grid-системе и задавать процесс их выполнения (workflow) на Grid-портале. Отметим, что представление модели задачи в виде графа (рис. 3) более удобно использовать при планировании выполнения задания в Grid-системе, чем традиционную диаграмму последовательностей UML (рис. 2), поскольку объектная модель в виде графа наглядно иллюстрирует возможность распараллеливания выполнения элементарных задач в рамках комплексного задания.

## **Выводы**

В работе проанализированы требования к Grid-системам мониторинга и поддержки принятия решений на основе спутниковых данных, выделены основные типы задач, решаемых в этих системах, предложены их

формальные описания и построена объектная модель задач в системе обозначений языка UML. Предложенная объектная модель обобщенной задачи является составной частью модели комплексной задачи со сложным процессом выполнения, представляемой в виде направленного графа, и представляет собой удобный инструмент для имитационного моделирования работы Grid-системы [21]. Кроме того, подобные модели позволяют формировать комплексные задания (job) для выполнения в Grid-системе и задавать сам процесс выполнения задания на портале системы. Дальнейшие исследования в этой области будут связаны с построением модели системных ресурсов и расширении программных средств имитационного моделирования распределенных систем путем добавления возможностей моделирования свойств Grid.

Работа выполнена при поддержке проектов УНТЦ №3872 «Разработка эффективных Grid-технологий экологического мониторинга на основе спутниковых данных» и INTAS-CNES-NSAU “Data Fusion Grid Infrastructure” (Ref. Nr 06-1000024-9154).

### **Литература**

1. Шелестов А.Ю., Куссуль Н.Н., Скакун С.В. Grid-технологии в системах мониторинга на основе спутниковых данных // Проблемы управления и информатики. — 2006. — № 1-2. — С. 259-270.
2. Шелестов А.Ю., Кравченко А.Н., Корбаков М.Б., Куссуль Н.Н., Скакун С.В., Рудакова А.И., Ильин Н.И., Тютюнник Л.И. Grid-технология реализации украинского сегмента GEOSS// Загальногалузевий науково-виробничий журнал «Зв’язок» (спеціалізований додаток). - 2006. - С. 106-125.
3. Загородній А.Г., Зінов’єв Г.М., Мартинов Є.С., Свистунов С.Я., Шадура В.М. Грід – нова інформаційно-обчислювальна технологія для науки // Вісник НАН України. — 2005. — № 6. — С. 17-25.
4. Zgurovsky M.Z. Development of educational and research segment of information society in Ukraine // Системні дослідження та інформаційні технології — 2006. — № 1. — С. 7-17.
5. Шелестов А.Ю., Корбаков М.Б., Лобунець А.Г. Реалізація Grid-інфраструктури для розв’язання задач обробки супутникових даних // Проблеми програмування. – 2006. - №2-3. – С. 94-101.
6. Building a European information capacity for environment and security. A contribution to the initial period of the GMES Action Plan (2002-2003) // Office for Official Publications of the European Communities (Luxembourg). - 2004. - 238 p.
7. Global Earth Observation System of Systems (GEOSS), 10-Year Implementation Plan Reference Document // ESA Publication Division, Netherlands, 2005. — 209 p.
8. Згуровский М.З., Панкратова Н.Д. Системный анализ: проблемы,

- методология, применения. — К.: Наукова думка. — 2005. — 743 с.
9. Holtman K. CMS Requirements for the Grid // Proceedings of the International Conference on Computing in High Energy and Nuclear Physics (CHEP2001). — 2001.
  10. Deelman E., Blackburn K., et al. GriPhyN and LIGO, Building a Virtual Data Grid for Gravitational Wave Scientists // Presented at 11th Intl. Symposium on High Performance Distributed Computing. — 2002.
  11. Annis J., Y. Zhao et al. Applying Chimera Virtual Data Concepts to Cluster Finding in the Sloan Sky Survey // Technical Report GriPhyN-2002-05, 2002.
  12. Peltier, S.T., et al. The Telescience Portal for Advanced Tomography Applications // Journal of Parallel and Distributed Computing: Computational Grid. — 2002. — 63(5). — P. 539-550.
  13. Fusco L. Earth Science GRID on Demand // CEOS WGISS-21 GRID Task Team Meeting. — Budapest, May. — 2006.
  14. Fusco L., Goncalves P., Linford J., Fulcoli M., Terracina A., D'Acunzo G. Putting Earth-Observation on the Grid // ESA Bulletin. — 2003. — 114. — P. 86-91.
  15. GEONETCast, [www.earthobservations.org/progress/GEONETCast.html](http://www.earthobservations.org/progress/GEONETCast.html).
  16. Рис У.Г. Основы дистанционного зондирования. — М.: Техносфера, 2006. — 336 с.
  17. Dissemination and Exploitation of GRids in Earth science. — <http://www.eu-degree.eu>.
  18. Попков Ю.С. Макросистемы и Grid-технологии: моделирование динамических стохастических сетей // Проблемы управления. — 2003. — № 3. — С. 10-20.
  19. Куссуль Н.М., Попов М.О., Шелестов А.Ю., Станкевич С.А., Корбаков М.Б., Кравченко О.М., Козлова А.О. Інформаційний сервіс оцінювання біорізноманіття Причорноморського регіону України в контексті розвитку українського сегменту системи GEOSS // Наука та інновація. — 2007. (подано)
  20. Одум Ю. Экология.- Т.2.- Пер. с англ.— М.: Мир, 1986.— 376 с.
  21. Shelestov A.Ju., Kussul N.N., Skakun S.V. Grid-infrastructure simulation// Проблеми програмування. — 2006. - №2-3. — С. 221-230.