

УДК 004.77

ПОДХОД К ОБЕСПЕЧЕНИЮ ИНТЕРОПЕРАБЕЛЬНОСТИ В ВЫСОКОПРОИЗВОДИТЕЛЬНОЙ СРЕДЕ НА ПРИМЕРЕ E-SCIENCE

А. А. Башлыкова^{1,2}, А. А. Каменщиков¹, А. Я. Олейников¹, Т. Д. Широбокова¹

¹ ИРЭ им. В.А. Котельникова РАН, 125009, Москва, ул. Моховая 11-7

² МИРЭА - Российский технологический университет,
119454, Москва, просп. Вернадского, д. 78

Статья поступила в редакцию 24 октября 2019 г.

Аннотация. Проведена адаптация единого подхода, зафиксированного в ГОСТ Р55062-2012 для случая высокопроизводительной среды, а именно применения ГРИД и облаков на примере электронной науки (e-science). Приведены концепции интероперабельности, модели интероперабельности, а также архитектуры интероперабельности для ГРИД и облаков, соответственно. Подчёркивается разница в реализации единого подхода для ГРИД и облаков, которая состоит в способе предоставления вычислительных мощностей. Модель интероперабельности для ГРИД зафиксирована в ГОСТ Р 55768-2013. Модель облачных вычислений направлена, в первую очередь, на обеспечение интероперабельности между облачными сервисами на уровне приложения, платформы и инфраструктуры. Согласно моделям интероперабельности, в профиль интероперабельности должны войти стандарты и документы технического, семантического и организационного уровня.

Ключевые слова: ГРИД, облака, интероперабельность, стандарты, электронная наука, e-science.

Abstract. The unified approach, fixed in GOST R55062-2012 for the case of a high-performance environment was adapted (GRID and clouds first of all) on the example of e-science. The concepts of interoperability, interoperability models, and interoperability architectures for GRID and clouds, respectively, are presented. The difference in the implementation of a unified approach for GRID and clouds is highlighted. The key difference between GRID and cloud concepts in the way of providing computing power. The interoperability model for grid is fixed in the GOST

R 55768-2013. The cloud computing model is primarily aimed to ensuring interoperability between cloud services at the application, platform and infrastructure levels. The interoperability profile should include standards of a technical, semantic and organizational level in accordance with the models of interoperability.

Key words: GRID, clouds, interoperability, standards, e-science.

Введение

В настоящее время при решении проблемы электронной науки (e-science) используются разнородные вычислительные платформы, и образуется распределенная высокопроизводительная гетерогенная среда [1,2]. В данной среде возникает проблема взаимодействия разнородных платформ, получившая название «проблемы интероперабельности». Интероперабельность должна достигаться на основе использования наборов ИКТ-стандартов (профилей), включающих стандарты ГРИД-технологий (в ряде случаев, далее - ГРИД), стандарты облачных технологий (в ряде случаев, далее - облака). Достижение интероперабельности — сложная комплексная научно-техническая и организационно-методическая проблема, имеющая как фундаментальные, так и прикладные аспекты, не получившая окончательного решения во всем мире. Отметим, что задача развития принципов интероперабельности, стандартов и технологий открытых информационных систем включена в Программу фундаментальных научных исследований государственных академий наук на период 2013-2020 гг. (см. п.34). Чем выше уровень гетерогенности среды, тем актуальнее достижение интероперабельности, а высокопроизводительная среда является именно такой и может быть отнесена к классу сверхсложных систем (система систем — System of Systems(SoS)) [3].

Данная статья является продолжением статьи [4] и подготовлена в рамках проекта «Исследование проблемы интероперабельности в высокопроизводительной защищенной среде», выполняемого в рамках программы №2 фундаментальных исследований РАН по приоритетным направлениям на 2018-2020 г. «Механизмы обеспечения отказоустойчивости и

высоконадежных вычислений», а именно, статья посвящена некоторым важным аспектам обеспечения интероперабельности в высокопроизводительной среде. Проблему интероперабельности следует решать совместно с проблемой информационной безопасности, что еще усложняет проблему [4].

1. Интероперабельность. Основные понятия. Методика обеспечения

Согласно общепринятому определению: «Интероперабельность - способность двух или более информационных систем или компонентов к обмену информацией и к использованию информации, полученной в результате обмена» (ГОСТ Р 55062-2012, ИСО/МЭК/ИЕЕЕ 24765:2010) [5].

В основе достижения интероперабельности лежит использование согласованных наборов стандартов ИКТ-технологий – профилей. В ГОСТ Р 55062-2012 приведены эталонная модель интероперабельности (Рис.1) и методика обеспечения интероперабельности (Рис. 2).

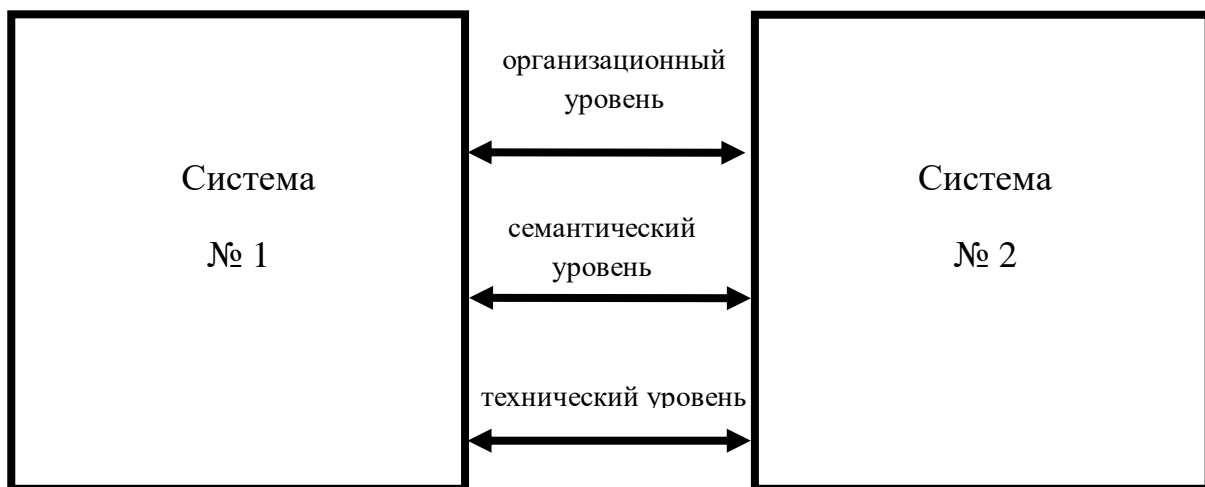


Рис. 1. Эталонная модель интероперабельности.

Здесь следует подчеркнуть, что использование ИКТ-стандартов обеспечивает только нижний, т.н. «технический» уровень интероперабельности. Полная интероперабельность может быть достигнута на более высоких уровнях - семантическом и организационном.

Ранее предложенный нами подход мы применяли к разным областям применения, а именно в здравоохранении [6], образовании [7], военном деле [8], электронным библиотекам [9], в том числе частично для научных

исследований (e-science) [10,11], а в настоящей работе данный вопрос прорабатывается более детально с учетом того, что основные компоненты, для которых актуальна интероперабельность - ГРИД и облака [4].

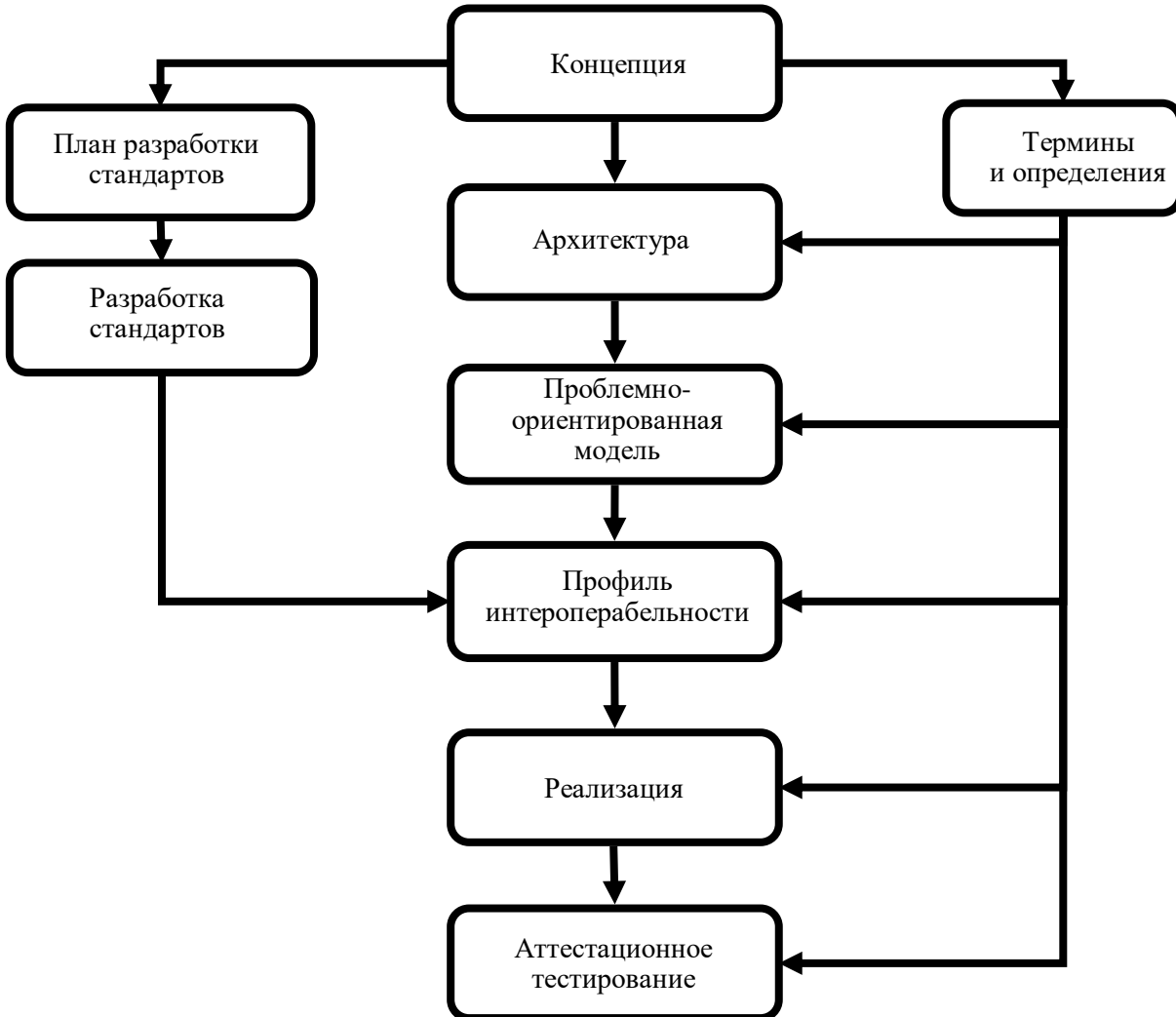


Рис. 2. Блок-схема методики обеспечения интероперабельности.

2. Концепция интероперабельности в e-science

В [2] утверждается, что в XXI веке большая часть огромного объема научных данных, собранных новыми приборами, наряду с результатами компьютерного моделирования, вероятно, навсегда будет в общедоступном, контролируемом состоянии для целей дальнейшего анализа. Этот анализ в свою очередь приведет к разработке многих новых теорий. Существование 4-ой парадигмы состоит в интенсивном применении самых перспективных ИКТ, применении распределенной высокопроизводительной среды, куда входят

суперкомпьютеры, кластеры, GRID-системы, системы облачных вычислений и конечные персональные ЭВМ пользователей [1,2].

В [12] утверждается, что из всех компонентов высокопроизводительной среды проблема интероперабельности для суперкомпьютеров и кластеров проявляется в меньшей степени, т.к. суперкомпьютеры и кластеры чаще представляют собой готовое решение, позволяющее пользователям получать услуги от каждого конкретного кластера или суперкомпьютера. Проблема интероперабельности наиболее актуальна для ГРИД и облаков, поэтому далее в статье мы будем рассматривать применение единого подхода [5] к облакам и ГРИД, как это сделали ранее в предыдущей статье на эту тему [4]. При этом, где это возможно, выделим общие черты. Однако, следует отметить, что вследствие использования суперкомпьютеров для развертывания облачных сервисов в скором времени может возникнуть необходимость сравнения не только облаков и ГРИД, но также добавить в сравнение суперкомпьютеры [12].

Согласно единому подходу к обеспечению интероперабельности [5], концепция интероперабельности должна содержать ряд основных положений, представленных в таблице 1 [10].

Таблица 1 – Основные положения концепции ГРИД и облаков.

ГРИД	Облака
Предполагает объединение групп компьютеров и устройства хранения, позволяющее динамически выделять под определенные задачи необходимые ресурсы по мере появления потребности в них.	Предполагает предоставление по требованию конечным пользователям динамического доступа к услугам (вычислительным ресурсам, приложениям, платформам) через локальную сеть или Интернет.
Интероперабельность в области грид-вычислений означает способность двух или более ГРИД или их узлов обмениваться информацией и использовать эту информацию	Итероперабельность в области облачных вычислений означает способность двух или более облаков и их компонентов к обмену информацией и использованию информации, полученной в результате этого обмена
Цель обеспечения интероперабельности грид-систем – создание единой ГРИД, содержащей множество стандартизованных компонентов, благодаря которым возможно взаимодействие между отдельными частями ГРИД.	Цель обеспечения интероперабельности облаков – создание единой облачной системы, раскрывающей истинный потенциал и преимущества облачных вычислений, заключающихся в возможности обмениваться понятными сообщениями, умении передавать и хранить данные в унифицированном формате, иметь возможность переносить образы виртуальных машин

Ключевая разница в концепциях ГРИД и облаков – это способ предоставления вычислительных мощностей [10]. В случае с ГРИД-вычислениями это - распределенная мощность и ресурсы, предоставляемые на паритетной основе. В случае с облачными вычислениями эта мощность арендуется, и чем ее больше, тем больше приходится платить.

3. Архитектура интероперабельности в e-science

Как следует из предыдущего: в основу архитектуры должны быть положены свойства ГРИД-систем и облачных вычислений. На основе анализа этих свойств формулируются требования к [13]:

- Web-сервисам,
- разделению ресурсов между организациями,
- сервисам данных,
- обеспечению качества обслуживания,
- протоколам безопасности и в соответствии с принятой политикой безопасности,
- стоимости администрирования,
- масштабируемости
- работоспособности
- простоте использования и расширяемости

В таблицах 2 и 3 представлены архитектуры ГРИД-вычислений и облачных вычислений [13].

ГРИД-вычисления сосредоточены на интеграции существующих ресурсов с их аппаратным обеспечением, управлением локальными ресурсами и инфраструктурой безопасности. ГРИД определяют и обеспечивают набор стандартных протоколов, промежуточного программного обеспечения (ППО), инструментов и услуг, построенных на основе этих протоколов. Интероперабельность и безопасность – основа инфраструктуры ГРИД-вычислений.

Таблица 2 – Архитектура ГРИД-вычислений.

<pre> graph TD Application[Application] --> Collective[Collective] Application --> Resource[Resource] Application --> Connectivity[Connectivity] Application --> Fabric[Fabric] </pre>	<p>Прикладной уровень (Application) – содержит любые пользовательские приложения, реализуемые в ГРИД среде через API, и функционирует в средах виртуальной организации (VO)</p> <p>Коллективный уровень (Collective)– перехватывает взаимодействия между наборами ресурсов, службы каталогов, допускает контроль и открытие ресурсов VO</p> <p>Уровень ресурса (Resource)– определяет протоколы для публикации, открытия, согласования, контроля, учета и оплаты совместного использования операций на отдельных ресурсах</p> <p>Уровень связи (Connectivity) - определяет базовую связь и протоколы аутентификации для простых и безопасных сетевых транзакций</p> <p>Структурный уровень (Fabric)– обеспечивает доступ к различным типам ресурсов, таким как: вычисление, хранение, сетевой ресурс, репозиторий кода и т.д.</p>
--	---

Таблица 3 – Архитектура облачных вычислений.

<pre> graph TD Application[Application] --> Platform[Platform] Application --> UnifiedResource[Unified Resource] Application --> Fabric[Fabric] </pre>	<p>Прикладной уровень (Application) – содержит приложения, которые работали бы в облаке</p> <p>Уровень платформы (Platform) – прибавляет набор специализированных инструментов, промежуточного ПО и служб поверх объединенных ресурсов, чтобы обеспечить платформу разработки и/или развертывания (среда web-хостинга, служба планирования)</p> <p>Объединенный уровень (Unified Resource) – содержит ресурсы, которые абстрагировались/инкапсулировались так, чтобы они могли быть представлены верхнему уровню и пользователям как интегрированные ресурсы (компьютер/кластер, логическая файловая система, система баз данных)</p> <p>Структурный уровень (Fabric)– содержит необработанные аппаратные ресурсы, такие как вычислительные ресурсы, ресурсы хранения и сетевые ресурсы;</p>
---	--

Облачные вычисления представляются как источник вычислительных ресурсов или ресурсов хранилища, к которым можно получить доступ через стандартные протоколы и так называемые абстрактные интерфейсы. Облачные вычисления могут быть построены на многих существующих протоколах (Web Services, Web 2.0). Из сказанного следуют различия в представленных выше архитектурах.

4. Модель интероперабельности в e-science

Из отличий в концепциях и архитектурах следуют различные модели интероперабельности для ГРИД-вычислений и облачных вычислений.

4.1 Модель интероперабельности для ГРИД-вычислений

На рисунке 2 представлена модель интероперабельности ГРИД-вычислений, зафиксированная в ГОСТ Р 55768-2013 [14].

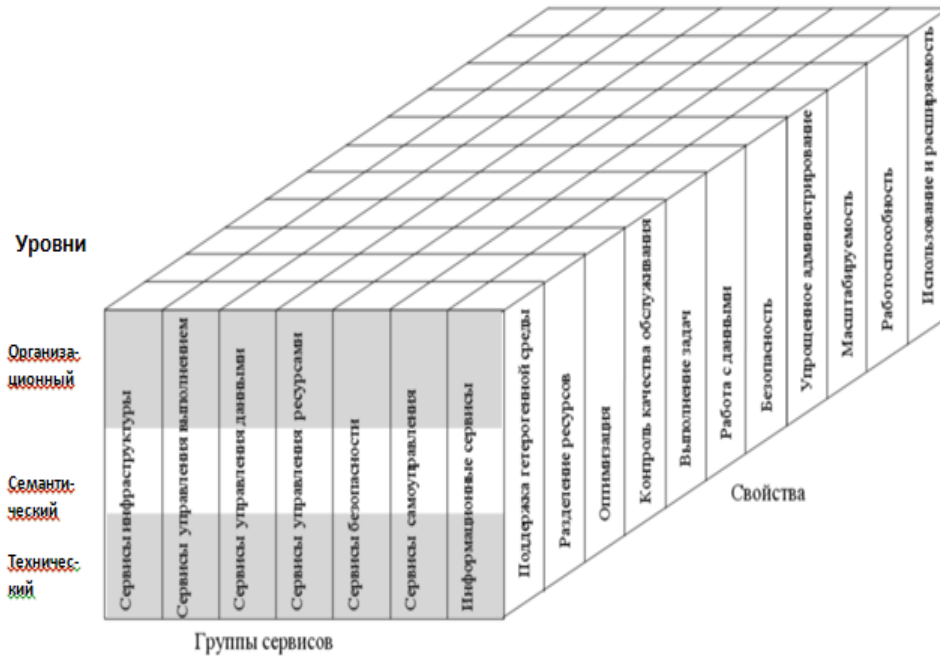


Рис. 3. Модель интероперабельности ГРИД-вычислений.

На рисунке 3 по оси абсцисс перечислены группы сервисов, обеспечивающих реализацию свойств ГРИД-среды. По оси ординат представлены три уровня (технический, семантический и организационный) интероперабельности сервисов. По третьей оси перечислены свойства ГРИД-системы [14,15].

4.2 Модель интероперабельности для облачных вычислений

Модель интероперабельности для e-science на основе облачных вычислений отражает несколько другой подход, чем для ГРИД, и направлена в первую очередь на обеспечение интероперабельности между облачными сервисами на уровне приложения, платформы и инфраструктуры (см. рис.4).

Данная модель представляет собой развитие эталонной модели (см. рис.1) и содержит расширенный набор уровней за счет детализации уровней эталонной модели.

Семантический уровень, выделенный на рисунке 4, в случае облачных вычислений «расщепляется» на 4 подуровня. Также имеется дополнительная проекция в виде сервисов. Рассмотрим подробнее уровни интероперабельности, выделяемые в данной модели [16].

На техническом уровне между взаимодействующими облачными вычислениями работают протоколы обмена данными. Речь идет об использовании совместимых аппаратно-программных средств, использования телекоммуникационной среды взаимодействия и протоколов связи.

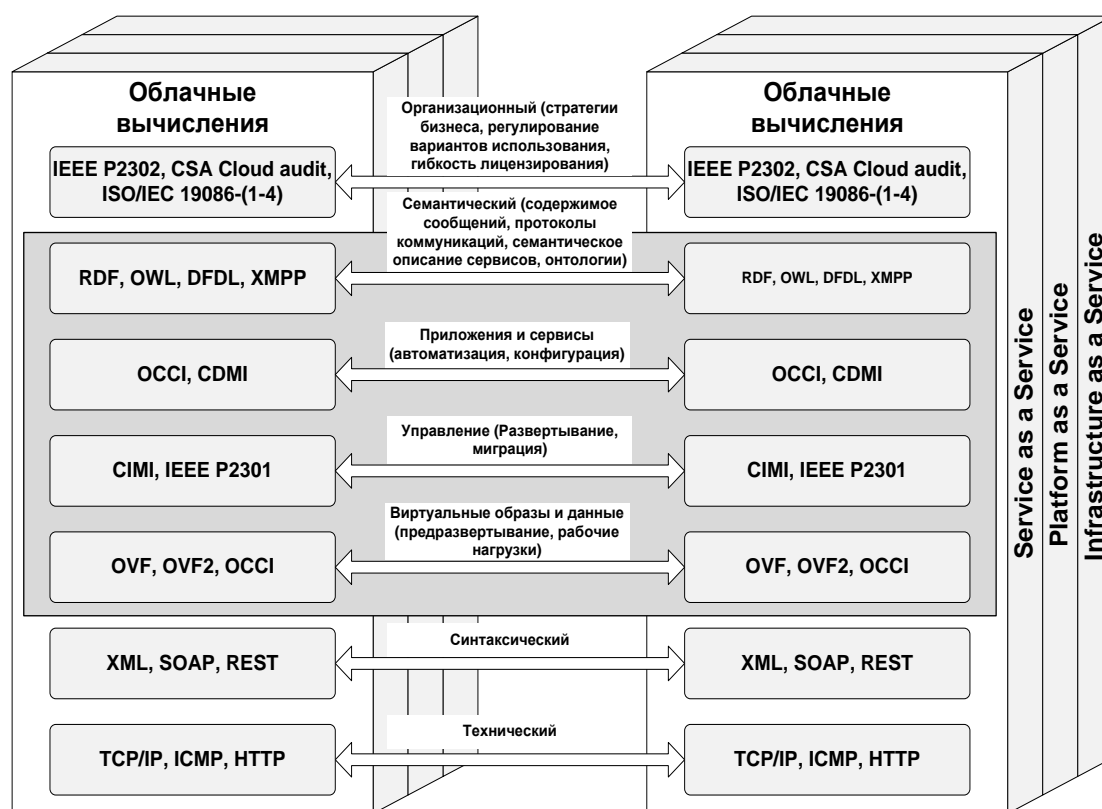


Рис. 4. Модель интероперабельности для облачных вычислений.

Взаимодействие на синтаксическом уровне подразумевает использование единых форматов данных, единую систему кодирования информации.

На уровне виртуальных образов и данных взаимодействие между облачными вычислениями связано с импортом и предоставления общего

доступа к образам виртуальных машин, использования стандартов для доступа к вычислительным мощностям.

Взаимодействие на уровне управления связано с контролированием жизненного цикла приложения, стандартами развертывания приложений на разных облаках [16].

5. Профиль интероперабельности в e-science

Как видим, ГРИД-вычисления и облачные вычисления с точки зрения проблемы интероперабельности имеют много общего. Поэтому изначально можно составить общий профиль интероперабельности, выделив в нем, где это необходимо, отличия для ГРИД -вычислений и облачных вычислений.

Таблица 4 – Организации, связанные с обеспечением интероперабельности ГРИД и облаков.

Организации	Область применения
Open Grid Forum (OGF) Distributed Management Task Force (DMTF) Organization for the Advancement of Structured Information Standards (OASIS)	ГРИД и облака
European Telecommunications Standards Institute (ETSI) Internet Engineering Task Force (IETF) ITU Telecommunication Standardization Sector (ITU-T) Storage Networking Industry Association (SNIA) TeleManagement Forum (TMF)	ГРИД
Open Cloud Consortium (OCC) Institute of Electrical and Electronics Engineers (IEEE) European Telecommunications Standards (ETSI) Open Virtualization Format (OVF) Open Cloud Computing Interface (OCCI) Storage Networking Industry Association (SNIA) Topology and Orchestration Specification for Cloud Applications (TOSCA) Cloud Application Management for Platforms (CAMP) The National Institute of Standards and Technology (NIST) Object Management Group (OMG)	Облака

Профиль [5] представляет собой согласованный набор стандартов, структурированный в терминах модели интероперабельности, который должен обновляться по мере актуализации входящих в него стандартов и может быть издан как отдельный нормативно-технический документ. Согласно моделям интероперабельности, в профиль интероперабельности должны войти

стандарты технического, семантического и организационного уровня. В таблице 4 приведены основные организации, стандарты которых должны войти в профиль интероперабельности.

Разделение стандартов по уровням, в первом приближении может быть сделано для эталонной модели интероперабельности, разделение стандартов по моделям интероперабельности для ГРИД и облаков требует более детальной проработки. Наиболее важные стандарты для ГРИД и облаков представлены в таблице 5 [4].

Таблица 5 – Стандарты, необходимые для с обеспечения интероперабельности ГРИД и облаков.

Стандарты	Область применения
GFD.225 Interoperable Certificate Profile GFD.205 An XACML Attribute and Obligation Profile for Authorization Interoperability in Grids GFD.148 Interoperability Experiences with the OGSA WSRF Basic Profile 1.0 GFD.140 Implementation and Interoperability Experiences with the Job Submission Description Language (JSDL) 1.0	ГРИД
IEEE P2301 Guide for Cloud Portability and Interoperability Profiles IEEE P2302 Standard for Intercloud Interoperability and Federation ISO/IEC DIS 17826:2016 «Information Technology» — Cloud Data Management Interface (CDMI) ISO/IEC DIS 17963:2013 «Web Services for Management» (WS — Management) ISO/IEC DIS 16680:2012 information technology — The Open Group Service Integration Maturity Model (OSIMM) ISO/IEC 19941:2017 Information technology - Cloud computing - Interoperability and portability	Облака

Заключение

Таким образом, на основании изложенного можно сделать следующие выводы:

- Высокопроизводительная среда, как сверхсложная система, имеет высокий уровень гетерогенности, и, следовательно, возникает необходимость решения проблемы интероперабельности.
- Подчеркнута актуальность использование ГРИД и облаков для высокопроизводительной среды для задач e-science.

- Проведено дальнейшее развитие решения проблемы интероперабельности применительно к e-science, на основе использования ГРИД и облаков.

Литература

1. Каменщиков А.А., Олейников А.Я., Проблема интероперабельности при реализации четвертой парадигмы // CEUR Workshop Proceedings 2018, Т.2267, №4(39), С.190-194. URL: <http://ceur-ws.org/Vol-2267/190-194-paper-35.pdf> (дата обращения 07.01.2019).
2. The Fourth Paradigm Data-Intensive Scientific Discovery. Microsoft research Redmond, WA. Available at: <https://www.microsoft.com/en-us/research/publication/fourth-paradigm-data-intensive-scientific-discovery/>. (accessed:17.10.2018).
3. Каменщиков А.А., Олейников А.Я., Широбокова Т.Д. Исследование особенностей проблемы интероперабельности в крупномасштабных информационных системах // Информационные технологии и вычислительные системы – ИТ и ВС. 2018. №3. С.16 -21.
4. Каменщиков А.А., Олейников А.Я., Широбокова Т.Д. Стандарты интероперабельности в высокопроизводительной среде. // Программные системы: теория и приложения. 2018. Т.9. №4(39). С.383–397. URL: http://psta.psir.ru/read/psta2018_4_383-397.pdf (дата обращения 08.12.2018).
5. ГОСТ Р 55062-2012. Системы промышленной автоматизации и их интеграция. Интероперабельность. Основные положения [электронный ресурс] Федеральное агентство по техническому регулированию и метрологии. 2012. URL: <http://protect.gost.ru/document1.aspx?control=31&baseC=6&page=388&month=7&year=2008&search=%D0%93%D0%9E%D0%A1%D0%A2&id=183619> (дата обращения 01.05.2018).
6. Каменщиков А.А., Интероперабельность в области e-health // Информационные технологии и вычислительные системы. 2009. №5. С.67-71.

7. Рубан К.А. Особенности интероперабельности в системах электронного образования // Информационные технологии и вычислительные системы – ИТ и ВС. 2009. №5. С.72-82.
8. А.А. Башлыкова, А.А. Каменщиков, А.Я. Олейников, О подходах к разработке профилей интероперабельности в военной области // Информационные технологии и вычислительные системы – ИТ и ВС. 2017. №4. С.112 -121.
9. А.А. Башлыкова, А.А. Каменщиков, А.Я. Олейников, Т. Д. Широбокова, И. И. Чусов, Проблема интероперабельности в электронных библиотеках // Журнал радиоэлектроники: электронный журнал. 2017. №12. URL <http://jre.cplire.ru/jre/dec17/12/text.pdf> (дата обращения 14.12.2017).
10. Иванов С.В., Каменщиков А.А., Олейников А.Я., Широбокова Т.Д. Развитие работ по e-science и интероперабельности. Информатизация и связь, 2015. №3. С.5-10. ISSN 2078-8320. URL http://cplire.ru8080/3526/1/%D0%AE%20%D0%92%2080%2030_06.pdf (дата обращения 01.05.2018).
11. Журавлев Е.Е., Олейников А.Я. Интероперабельность в e-science. Информатизация и связь. 2009. №5. С.48-55. ISSN 2078-8320 URL: http://www.isa.ru/jitcs/images/stories/2009/05/48_55.pdf (дата обращения: 01.05.2018).
12. Ali N., Etawi A. A Comparison between Cluster, Grid, and Cloud Computing // International Journal of Computer Applications, 2018, Т.179, №32. URL: <https://www.ijcaonline.org/archives/volume179/number32/etawi-2018-ijca-916732.pdf> (дата обращения: 02.05.2018).
13. Ian F., Yong Z., Ioan R. Cloud Computing and Grid Computing 360-Degree Compared [Электронный ресурс]. URL: <https://arxiv.org/ftp/arxiv/papers/0901/0901.0131.pdf> (дата обращения: 11.09.2019).
14. ГОСТ Р 55768-2013 Информационная технология (ИТ). Модель открытой Грид-системы. Основные положения. 8 ноября 2013 г. URL: <http://docs.cntd.ru/document/1200108057> (дата обращения: 11.09.2019).

15. Журавлев Е.Е., Корниенко В.Н., Олейников А.Я., Широбокова Т.Д. Модель открытой Грид-системы // Журнал радиоэлектроники: электронный журнал. 2012. №12. URL <http://jre.cplire.ru/koi/dec12/3/text.html> (дата обращения 21.05.2019).

16. Иванов С.В., Олейников А.Я. Методика и алгоритм выбора стандартов для профиля интероперабельности в облачных вычислениях. // 7-я Международная конференция «Распределенные вычисления и Grid-технологии в науке и образовании». 4-9 июля 2016. Дубна. ОИЯИ, С.264-270. URL <http://ceur-ws.org/Vol-1787/264-270-paper-44.pdf> (дата обращения 01.05.2019).

Для цитирования:

Башлыкова А.А., Каменщиков А.А., Олейников А.Я., Широбокова Т.Д. Подход к обеспечению интероперабельности в высокопроизводительной среде на примере e-science. Журнал радиоэлектроники [электронный журнал]. 2019. № 11. Режим доступа: <http://jre.cplire.ru/jre/nov19/5/text.pdf>
DOI 10.30898/1684-1719.2019.11.5