

УДК 004.75.05

О. М. ТАРАСЮК

Национальный аэрокосмический университет им. Н. Е. Жуковского «ХАИ», Украина

ПРОБЛЕМЫ СОЗДАНИЯ ГЕТЕРОГЕННЫХ СИСТЕМ ДЛЯ ГЛОБАЛЬНО-РАСПРЕДЕЛЁННОЙ ОБРАБОТКИ И ХРАНЕНИЯ ИНФОРМАЦИИ

В статье рассмотрены тенденции развития современных информационных технологий и распределенных информационно-вычислительных систем. Революционное увеличение объема хранимых и обрабатываемых данных на фоне отстающего развития информационных технологий привело к появлению системного противоречия между свойствами готовности, оперативности обслуживания и согласованности данных, которые определяют качество современных распределенных информационно-вычислительных систем. Другими актуальными проблемами создания распределенных Интернет-систем является несовершенство технологий беспроводного доступа, а также возрастающая роль уязвимостей программного обеспечения в качестве основного источника угроз для нормального функционирования распределенных компьютерных систем.

Ключевые слова: глобально-распределенные информационно-вычислительные системы, большие данные, гетерогенность абонентов, согласованность данных, готовность, оперативность.

Введение

На сегодняшний день стремительное расширение рынка мобильных устройств и приложений, а также повсеместная реализация концепций «Internet of Things» и «Cyber-Physical Systems» происходит на фоне не менее динамичного развития технологий «Cloud Computing» (Облачных вычислений) и «Big-Data» (Больших данных). С одной стороны происходит повсеместное распространение (проникновение) компьютерных технологий на основе оснащение объектов физического мира встроенными средствами измерения, вычисления и коммуникации в рамках концепции «Ubiquitous Computing». Это приводит к генерации значительных объемов информации, нуждающихся в хранении, обработке и предоставлении одновременного высокоскоростного доступа со стороны множества клиентов.

В то же время имеет место концентрация информационных услуг, вычислительных ресурсов и средств хранения данных, в виде корпоративных центров обработки данных, совокупность которых формирует информационно-вычислительную Cloud-среду [1]. Развитие технологий распределенных вычислений является необходимым условием для обработки стремительно растущего объема данных в информационном пространстве современного общества. Предпосылками такого развития являются:

1) популярность приложений IoT (Internet-of-Things) и социальных медиа-сетей (Facebook, Twitter и др.), которые характеризуются высокой интенсив-

ностью генерации данных значительного объема; например, общее количество пользователей социальной сети Facebook на сегодняшний день превышает 1 миллиард, из которых ежедневную активность проявляют более 1,5 миллиона человек; каждый день объем хранимой информации увеличивается примерно на 500 ТБ, а инфраструктура сети насчитывает более 180000 серверов, размещенных в 5 дата-центрах [2];

2) высокие требования к объемам хранимой информации и скорости их обработки со стороны компаний – Интернет-гигантов, таких как Google, Yahoo, Yandex; например, в 2011 году инфраструктура Goggle насчитывала более 900000 серверов, размещенных в 12 основных и 32 вспомогательных датацентрах [3];

3) появление и активное использование общедоступных сервисов «облачных» вычислений хранения данных, таких как Amazon EC2 и S3;

4) внедрение в практику разработки программного обеспечения (ПО) технологий динамически определяемых данных, поддерживающих возможность быстрой смены схемы данных.

Для описания наборов данных, размер которых превосходит возможности традиционных реляционных баз данных по хранению, управлению и анализу информации в 2008 году был введен термин BigData или «Большие данные» [4]. Отличительными характеристиками «больших данных» является триада VVV (Volume, Velocity, Variety) [5]:

Volume – экстремально большой объем данных,

создающий проблемы для традиционных подходов к их хранению и обработке; большое количество пользователей, одновременно выполняющих запросы на запись и чтение данных;

Variety – неструктурированность данных, отсутствие строгой модели данных; разнообразие (гетерогенность) типов устройств, которые являются источниками и потребителями информации (от смартфонов и беспроводных сенсоров IoT до настольных компьютеров и промышленных серверов);

Velocity – высокая интенсивность генерации данных, а также высокие требования к скорости их чтения и обработки.

Для хранения и обработки больших данных разрабатываются новые технологии и средства, реализующие принципы:

1) горизонтального масштабирования и распределения («шардинга») данных по множеству однотипных серверов;

2) массово-параллельной обработки неопределённо структурированных данных;

3) репликации данных с целью повышения их доступности и скорости обслуживания клиентских запросов.

В то же время, анализ концепции *распределенных вычислений, устойчивых к разделению*, предложенной Эриком Брюером в 2000 году [6] и исследованной в работах [7–9] свидетельствует о том, что современные особенности создания и эксплуатации распределенных информационно-вычислительных систем учтены не в полной мере. Среди них основными являются следующие:

– разнородность клиентов распределенных систем, а также доминирование технологий беспроводной связи при доступе клиентов к распределенным информационно-вычислительным ресурсам;

– фундаментальное противоречие между свойствами готовности, целостности и оперативности, системная взаимосвязь которых может быть выра-

жена не только в качественном, но и в количественном отношении;

– превалирование информационных угроз, обусловленных многочисленными уязвимостями ПО, над риском проявления дефектов аппаратных, программных и сетевых компонентов в качестве основной причины отказов распределенных систем.

Таким образом, целью статьи является анализ актуальных проблем построения информационно-вычислительных систем (ИВС) для глобально-распределённой обработки и хранения информации.

1. Гетерогенность абонентов распределенных ИВС

В современных условиях развития информационных технологий мобильные вычислительные устройства (смартфоны, планшеты и др.) с одной стороны играют роль «генераторов» данных для их хранения в Cloud-среде, а с другой – являются портативными терминальными устройствами, предоставляющими владельцу доступ к вычислительным Cloud-ресурсам и сервисам и предназначенными для отображения результатов удаленных вычислений.

Взаимодействие мобильных клиентов, сервис- и Cloud-провайдеров в этих условиях описывается многоуровневой сервис-ориентированной архитектурой, представленной на рис. 1. Cloud-провайдеры предоставляют IaaS- и PaaS-услуги по размещению интернет-приложений сервис-провайдеров. В свою очередь Интернет-приложения сервис-провайдеров формируют «рынок» информационных услуг, предоставляемых конечным пользователям, значительная часть которых для доступа к этим ресурсам использует мобильные устройства (смартфоны, планшеты, нет- и ноутбуки и др.). Менее распространенным является сценарий, при котором конечные пользователи напрямую взаимодействуют с Cloud-провайдерами для размещения и использования своих приложений и хранения персональных данных.



Рис. 1. Многоуровневая архитектура гетерогенных систем распределенной обработки и хранения информации

Предоставление информационно-вычислительных услуг конечным пользователям, особенно тем из них, кто использует для доступа к услугам мобильные гаджеты, в рамках указанной архитектуры имеет ярко-выраженный энергетический контекст. Мобильные абоненты функционируют в условиях энергетических ограничений, определяемых ёмкостью аккумуляторных батарей. Повышение энергетической эффективности и увеличение времени их автономной работы может быть достигнуто за счет переноса вычислительно-сложных задач с мобильной платформы в Cloud-среду [10]. Целесообразность такого подхода может быть оценена на основе формулирования и решения оптимизационной задачи с учетом [11]:

- 1) требований (ограничений) к оперативности получения результата вычислений;
- 2) разницы в вычислительной мощности мобильной платформы и виртуального Cloud-ресурса, которая определяет выигрыш по времени выполнения задачи;
- 3) энергетических затрат на решение задачи на локальной мобильной платформе и стоимости виртуального Cloud-ресурса;
- 4) энергетических и временных затрат на беспроводную передачу исходных данных и результата вычислений в/из Cloud-среды;

Другой важнейшей проблемой является противоречия между высокой декларируемой скоростью передачи и достаточно умеренной пропускной способностью, которая реально доступна беспроводным абонентам. Одной из основных причин низкой пропускной способности сетей Wi-Fi является неравномерное использование среды передачи абонентами с различной скоростью беспроводного сетевого подключения [12, 13].

Для решения данной проблемы предлагаются различные подходы, основанные на предоставлении приоритетного доступа к среде передачи высокоскоростным абонентам, а также регулирование размера сетевых кадров пропорционально скорости передачи абонентов [14, 15].

2. Противоречие между свойствами готовности, целостности и оперативности

Следствием увеличения масштаба современных распределенных систем является увеличение вероятности отказов и сбоев. Например, корпорация Google в своем отчете [16] отмечает в среднем около тысячи отказов, происходящих в год в типовом дата-центре, объединяющем 1800 компьютеров. При этом среднее время между отказами составляет всего 8,76 часов. Кроме серьезных отказов в глобально-

распределенных системах может возникать множество менее значительных сбоев, обусловленных не только сбоями аппаратных и программных средств, но и нестабильностью сетевых задержек при передаче сообщений между удаленными компонентами.

Высокая вероятность отказа компонентов распределенных систем стимулирует применение различных методов обеспечения отказоустойчивости, среди которых наиболее популярным является резервирование и репликация. Кроме того, это позволяет дополнительно решить проблему производительности за счёт распределения пользовательских запросов между несколькими репликами, что позволяет сократить время ожидания обработки в очереди. Глобальная географическая распределенность пользователей обуславливает и глобальную распределенность реплик – ближе к обслуживаемым клиентам для сокращения временных задержек.

Однако, поскольку современные распределенные информационно-вычислительные системы должны поддерживать не только операции чтения данных, но и их изменение и запись новых, возникает необходимость их синхронизации для обеспечения согласованности (consistency).

В этом случае глобальная распределенность реплик и увеличение их количества, учитывая неопределенность временных характеристик глобальной сети Интернет, имеет обратный эффект – т.е. приводит к существенному увеличению времени синхронизации, а, следовательно, к увеличению времени выполнения операций записи и продолжительной блокировке обновляемых данных для некоторых операций чтения. В случае отказа отдельных реплик, синхронизация вообще не сможет успешно завершиться. Если же завершить транзакцию не дожидаясь завершения операции синхронизации, известной как двухэтапное квитирование, то распределенная система не сможет гарантировать строгую согласованность данных.

Указанные противоречия были положены в основу, так называемой, CAP-теоремы [6], предложенной Эриком Брюером в 2000 году, в которой утверждается, что в распределенных информационно-вычислительных системах невозможно одновременно обеспечить выполнение трех свойств: согласованности данных (Consistency), готовности (Availability) и устойчивости к разделению (Partition Tolerance). Позднее CAP-теорема была признана Гилбертом С. и Линчем Н. частным случаем более общего противоречия между свойствами согласованности и готовности, характерного для ненадежных распределенных систем, допускающих отказы компонентов или же нарушение связности между ними [7]. По нашему мнению важным дополнением к системе указанных противоречий является учет опера-

тивності (Responsiveness) и своевременности обслуживания (Timeliness). В этом случае такой параметр, как установленное время ожидания или таймаут (timeout) определяет границу между своевременностью обслуживания и неготовностью системы. В соответствии с моделью надежности сервис-ориентированных систем [17], увеличение таймаута приводит, с одной стороны, к увеличению готовности системы, а с другой – к снижению оперативности обслуживания (т.е. увеличению времени ожидания ответа от системы, которое, в свою очередь, является вероятностной величиной).

Другими важными параметрами, определяющими взаимосвязь между свойствами готовности, согласованности и оперативности являются общее количество реплик в системе и требуемый уровень согласованности (см. рис. 2).

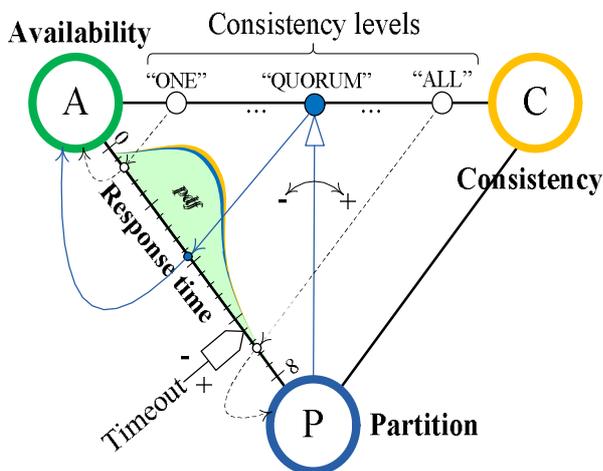


Рис. 2. Взаимосвязь между свойствами согласованности, готовности и оперативности

В современных распределенных нереляционных базах данных, например, Cassandra, уровень согласованности задается количеством реплик от общего их числа, после получения ответа (результата чтения или же подтверждения о выполнении записи) от которых транзакция считается успешно завершенной. Учитывая, что пользователю может быть не известно общее количество реплик, наибольшее распространение получили три уровня согласованности:

ONE – ожидается результат чтения или же подтверждение об успешной записи от одной (наиболее оперативной) из реплик;

QUORUM – ожидается результат чтения или же подтверждение об успешной записи от большинства реплик ($\text{общее_количество_реplik} / 2 + 1$);

ALL – ожидается результат чтения или же подтверждение об успешной записи от всех реплик.

Чем больше общее количество реплик и чем выше требуемый уровень согласованности, тем, в

среднем, дольше оказывается время ожидания результата чтения или же подтверждение об успешной записи. Как следствие увеличивается вероятность того, что ответ от системы будет получен после истечения установленного таймаута, т.е. система окажется неготовой предоставить требуемое обслуживание за ожидаемое время (см. рис. 2).

Таким образом, теория создания глобально-распределенных информационно-вычислительных систем требует разработки новых моделей, которые бы позволили установить количественную взаимосвязь между свойствами согласованности, готовности и оперативности, а также предложить разработчикам таких систем математически-обоснованный инструментарий для нахождения компромисса между этими свойствами с учетом предъявляемых со стороны пользователей требований.

3. Уязвимости системного программного обеспечения

На сегодняшний день общепризнанным является тот факт, что наибольшую угрозу нормальному функционированию компьютерных систем несут не дефекты программного или аппаратного обеспечения, а программные уязвимости, которые используются злоумышленниками для нарушения работоспособности систем, целостности и конфиденциальности хранимой и обрабатываемой информации. Об этом, в частности, свидетельствует и возрастающее количество исправлений безопасности, выпускаемых компаниями-разработчиками для своих продуктов. Например, количество исправлений безопасности (Security Updates), выпускаемых компанией Microsoft, фактически сравнялось по количеству функциональных исправлений (Тщт-Security Updates), а в некоторые периоды времени и превышает их [18].

Уязвимости обнаруживаются как в прикладном программном обеспечении, так и в системных компонентах и операционных системах. При этом последние являются, как правило, наиболее многочисленными и критическими. Исследование баз данных уязвимостей показывает [19], что в операционных системах одновременно может присутствовать несколько десятков известных уязвимостей, для которых вендорами еще не выпущены исправления. При этом время, которое компании-разработчики затрачивают на выпуск исправлений, может варьироваться от нескольких недель, до нескольких месяцев.

В этих условиях наибольшую актуальность приобретают модели оценки уязвимости компьютерных систем, учитывающие количество и критичность уязвимостей программных компонентов, а также методы проактивной защиты. Один из таких

методов, использующий принципы диверсности системного программного окружения и его динамической реконфигурации для минимизации активных уязвимостей предложен в [19].

Суть метода заключается в постоянном мониторинге количества уязвимостей и степени их критичности в группе диверсных операционных систем и системных программных компонентов (серверах приложений и баз данных, динамических библиотеках и т.п.). Основываясь на результатах мониторинга выполняется динамическая реконфигурация системного программного окружения таким образом, что наиболее уязвимые системные компоненты будут автоматически заменены на аналогичные по функциональному назначению, но менее уязвимые, т.е. имеющие меньшее количество активных критических уязвимостей.

Заключение

Современная тенденция развития информационных технологий такова, что наибольшее распространение получают гетерогенные ИВС, включающие как мобильные устройства, так и мощные центры обработки данных, распределенные по всему Интернет-пространству. Основной проблемой построения таких систем является системное противоречие между свойствами согласованности данных, готовности и оперативности обслуживания. Не менее важными проблемами остаются низкая производительность беспроводных компьютерных сетей на фоне стремительного увеличения количества мобильных клиентов ИВС, а также превращение уязвимостей программных средств в основной вид угроз для нормального функционирования компьютерных систем.

В связи с этим актуальной является разработка концепции гетерогенных систем для глобально-распределённой обработки и хранения информации, которая бы позволила динамически оптимизировать структуру и параметры функционирования таких систем для достижения необходимого компромисса между ключевыми свойствами в условиях вероятностной неопределенности надежностных и временных характеристик компонентов, а также Интернет-среды.

Литература

1. Бец, М. ЦОД нуждаются в оптимизации глобальной сети [Электронный ресурс] / М. Бец // Журнал сетевых решений/LAN. – 2012. – № 1. – Режим доступа: <http://www.osp.ru/lan/2012/01/13012478/> – 11.04.2016.
2. Коваленко, К. Распределенная сеть дата-

центров Facebook [Текст] / К. Коваленко // ЦОДы.РФ. – 2013. – № 2. – С. 56–60.

3. Архитектура Google 2011 [Электронный ресурс]. – Режим доступа: <https://www.insight-it.ru/highload/2011/arkhitektura-google-2011/> – 11.04.2016.

4. Черняк, Л. Большие Данные – новая теория и практика [Электронный ресурс] / Л. Черняк // Открытые системы. – 2011. – №10. – Режим доступа: <http://www.osp.ru/os/2011/10/13010990/> – 11.04.2016.

5. Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data [Электронный ресурс] // Gartner, 2011. – Режим доступа: <http://www.gartner.com/newsroom/id/1731916> – 11.04.2016.

6. Brewer, E. Towards Robust Distributed Systems [Текст] / E. Brewer // 19th Ann. ACM Symposium on Principles of Distributed Computing: conference proceedings. – Portland (USA), 2000. – P. 7–8.

7. Gilbert, S. Brewer's Conjecture and the Feasibility of Consistent, Available, Partition-Tolerant Web Services [Текст] / S. Gilbert, N. Lynch // ACM SIGACT News. – 2002. – Vol. 33(2). – P. 51–59.

8. Abadi, D. J. Consistency Tradeoffs in Modern Distributed Database System Design [Текст] / D. J. Abadi // IEEE Computer. – 2012. – № 2. – P. 37–42.

9. The Impact of Consistency on System Latency in Fault Tolerant Internet Computing [Text] / O. Tarasyuk, A. Gorbenko, A. Romanovsky, et al. // Distributed Applications and Interoperable Systems, LNCS 9038 / Eds.: A. Bessani, S. Bouchenak. – Berlin, Heidelberg (Germany) : Springer-Verlag, 2011. – P. 179–192.

10. Lawton, G. Cloud streaming brings video to mobile devices [Текст] / G. Lawton // Computer. – 2012. – № 2. – P. 14–16.

11. Зеленая ИТ-инженерия. Т. 2 : Системы, индустрия, социум [Текст] / О. М. Тарасюк, А. В. Горбенко, В. С. Харченко и др. ; под ред. В. С. Харченко. – Харьков : Нац. аэрокосм. ун-т "ХАИ", 2014. – 687 с.

12. Impact of Legacy Devices on 802.11n Networks [Text]. – Airmagnet Inc., 2008. – 15 p.

13. Throughput estimation with regard to airtime consumption unfairness in mixed data rate Wi-Fi networks [Text] / O. Tarasyuk, A. M. Abdul-Hadi, A. Gorbenko, V. Kharchenko // Communications. – 2014. – № 1. – P. 84–89.

14. Тарасюк, О. М. Способы решения проблемы неравномерного распределения времени удержания среды передачи в беспроводных сетях [Текст] / О. М. Тарасюк // Радіоелектронні і комп'ютерні системи. – 2014. – № 6 (70). – С. 79–84.

15. Tarasyuk, O. Performance Enhancement in Multirate Wi-Fi Networks [Текст] / O. Tarasyuk, A. Gorbenko // Journal of Information, Control and

Management Systems. – 2014. – Vol. 12, № 2. – P. 141–152.

16. Dean, J. *Software engineering advice from building largescale distributed systems [Electronic resource] / J. Dean // Stanford CS295 class lecture, 2007.* – Access mode: <http://research.google.com/people/jeff/stanford-295-talk.pdf>. – 11.04.2016.

17. *Dependability of Service-Oriented Computing: Time-Probabilistic Failure Modelling [Text] / A. Gorbenko, O. Tarasyuk, A. Romanovsky, V. Kharchenko // Software Engineering for Resilient Systems: LNCS 7527 / Ed.: P. Avgeriou. – Berlin; Heidelberg : Springer-Verlag, 2012. – P. 121–133.*

18. *Description of Software Update Services and Windows Server Update Services changes in content for 2016 [Electronic resource].* – Access mode: <https://support.microsoft.com/en-us/kb/894199>. – 11.04.2016.

19. *Using Diversity in Cloud-Based Deployment Environment to Avoid Intrusions [Text] / A. Gorbenko, O. Tarasyuk, V. Kharchenko, A. Romanovsky // Software Engineering for Resilient Systems, LNCS 6968 / Ed.: E. Troubitsyna. – Berlin, Heidelberg (Germany): Springer-Verlag, 2011. – P. 145–155.*

References

1. Bets, M. Data TsOD nuzhdayutsya v optimizat-sii global'noy seti [Centres require global networks optimization]. *Zhurnal setevykh resheniy/LAN – Journal of Networking Solution/LAN*, 2012, no. 1. Available at: www.osp.ru/lan/2012/01/13012478/ (accessed 11.04.2016).

2. Kovalenko, K. Raspredeleonnaya set' data-tsentrov Facebook [Distributed network of Facebook data centres], *TsODy.RF – DataCentres.RU*, 2013, no. 2, pp. 56–60.

3. *Arkhitektura Google 2011* [Google Architecture 2011]. Available at: <https://www.insight-it.ru/highload/2011/arkhitektura-google-2011/> (accessed 11.04.2016).

4. Chernyak, L. Bol'shie Dannye – novaya teoriya i praktika [BigData – New theory and practice]. *Otkrytye sistemy – Open Systems*, 2011, no. 10. Available at: <http://www.osp.ru/os/2011/10/13010990/> (accessed 11.04.2016).

5. *Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data.* Available at: <http://www.gartner.com/newsroom/id/1731916> (accessed 11.04.2016).

6. Brewer, E. Towards Robust Distributed Systems. *Proc. 19th Ann. ACM Symp. on Principles of Distributed Computing*, Portland (USA), 2000, pp. 7–8.

7. Gilbert, S., Lynch N. Brewer's Conjecture and the Feasibility of Consistent, Available, Partition-Tolerant Web Services. *ACM SIGACT News*, 2002, vol. 33, no. 2, pp. 51–59.

8. Abadi, D. J. Consistency Tradeoffs in Modern Distributed Database System Design. *IEEE Computer*, 2012, no. №2, pp. 37–42.

9. Tarasyuk, O., Gorbenko, A., Romanovsky, A., Kharchenko, V., Ruban V. The Impact of Consistency on System Latency in Fault Tolerant Internet Computing. *Distributed Applications and Interoperable Systems*, LNCS, vol. 9038, Berlin, Heidelberg, Springer-Verlag Publ., 2011, pp. 179–192.

10. Lawton, G. Cloud streaming brings video to mobile devices. *Computer*, 2012, no. 2, pp. 14–16.

11. Kharchenko, V. (ed.). *Zelenaya IT-inzheneriya. T. 2: Sistemy, industriya, sotsium* [Green IT-engineering, vol. 2. Systems, Industry, Society]. Kharkiv, 2014. 687 p.

12. *Impact of Legacy Devices on 802.11n Network*, Airmagnet Inc., 2008, 15 p.

13. Tarasyuk, O., Abdul-Hadi, A.M., Gorbenko, A., Kharchenko, V. Throughput estimation with regard to airtime consumption unfairness in mixed data rate Wi-Fi networks. *Communications*, 2014, no. 1, pp. 84–89.

14. Tarasyuk, O. M. Sposoby resheniya problemy neravnomernogo raspredeleniya vremeni uderzhaniya sredy peredachi v besprovodnykh setyakh [Means of solving a problem of uneven distribution of airtime consumption in wireless networks]. *Radioelektronni i komp'uterni sistemi – Radioelectronic and computer systems*, 2014, vol. 6(70), pp. 79–84.

15. Tarasyuk, O., Gorbenko A. Performance Enhancement in Multirate Wi-Fi Networks. *Journal of Information, Control and Management Systems*, 2014, vol. 12, no. 2, pp. 141–152.

16. Dean, J. *Software engineering advice from building largescale distributed systems.* Stanford CS295 class lecture, 2007. Available at: <http://research.google.com/people/jeff/stanford-295-talk.pdf> (accessed 11.04.2016).

17. Gorbenko, A., Tarasyuk, O., Romanovsky, A., Kharchenko V. Dependability of Service-Oriented Computing: Time-Probabilistic Failure Modelling. *Software Engineering for Resilient Systems*, LNCS, vol. 7527, Berlin; Heidelberg, Springer-Verlag Publ., 2012, pp. 121–133.

18. *Description of Software Update Services and Windows Server Update Services changes in content for 2016.* Available at: <https://support.microsoft.com/en-us/kb/894199> (accessed 11.04.2016).

19. Gorbenko, A., Tarasyuk, O., Kharchenko, V., Romanovsky, A. Using Diversity in Cloud-Based Deployment Environment to Avoid Intrusions. *Software Engineering for Resilient Systems*, LNCS, vol. 6968, Berlin, Heidelberg, Springer-Verlag Publ., 2011, pp. 145–155.

Поступила в редакцію 11.04.2016, рассмотрена на редколлегии 14.04.2016

ПРОБЛЕМИ СТВОРЕННЯ ГЕТЕРОГЕННИХ СИСТЕМ ДЛЯ ГЛОБАЛЬНО-РОЗПОДІЛЕНОЇ ОБРОБКИ ТА ЗБЕРІГАННЯ ІНФОРМАЦІЇ

О. М. Тарасюк

У статті розглянуті тенденції розвитку сучасних інформаційних технологій і розподілених інформаційно-обчислювальних систем. Революційне збільшення обсягу даних, що потребують обробки та збереження на тлі відставання розвитку інформаційних технологій призвело до появи системного протиріччя між властивостями готовності, оперативності обслуговування та узгодженості даних, які визначають якість сучасних розподілених інформаційно-обчислювальних систем. Іншими актуальними проблемами створення розподілених Інтернет-систем є недосконалість технологій бездротового доступу, а також зростаюча роль вразливостей програмного забезпечення в якості основного джерела загроз для нормального функціонування розподілених комп'ютерних систем.

Ключові слова: глобально-розподілені інформаційно-обчислювальні системи, великі данні, гетерогенність абонентів, узгодженості даних, готовність, оперативність.

ISSUES OF DEVELOPING HETEROGENEOUS SYSTEMS FOR GLOBALLY-DISTRIBUTED DATA PROCESSING AND STORAGE

О. М. Tarasyuk

The article examines trends in the development of modern information technologies and distributed computing and data storage systems. The revolutionary increase in volume of stored and processed data on the background of lagging development of information technologies has led to the emergence of system contradictions between availability, timeliness and consistency, which determine the quality of the modern distributed systems. Other critical issues of distributed Internet systems are the imperfection of wireless communication technologies, as well as the growing role of software vulnerabilities as the main threats to the normal operation of distributed computer systems.

Keywords: globally-distributed information processing systems, Big Data, users heterogeneity, data consistency, availability, responsiveness.

Тарасюк Ольга Михайловна – канд. техн. наук, доцент, доцент кафедри Комп'ютерних систем и сетей Национального аэрокосмического университета им. Н.Е. Жуковского «ХАИ», Харьков, Украина, e-mail: O.Tarasyuk @ csn.khai.edu.

Tarasyuk Olga – PhD, Docent, Associate Professor with the Department of Computer Systems and Networks, National Aerospace University named after N. Ye. Zhukovsky “Kharkiv Aviation Institute”, Kharkiv, Ukraine, e-mail: O.Tarasyuk @ csn.khai.edu.