

## ОТЗЫВ

**Председателя диссертационного совета о диссертации Тхуреин Киав Лина на тему: «Система консолидации данных и распределенных вычислений для поддержки информатизации Союза Мьянма», представленной на соискание ученой степени доктора технических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей»**

Актуальность, представленной на соискание ученой степени доктора технических наук диссертации Тхуреин Киав Лина, определяется тем, что широкое применение концепции автоматизации государственного управления тормозится отсутствием эффективного инструментария для управления распределенными информационными системами государства и средств консолидации разнородных баз данных в различных ведомствах. В диссертации разработана технология баз данных для консолидации и системной интеграции гетерогенных комплексов, решающих задачи проведения научных вычислений и поддержки принятия решений.

**В диссертационной работе** решены задачи создания оптимальных вычислительных архитектур, предназначенных для хранения и обработки информации сверхбольших объемов, характерных в настоящее время для многих областей науки и практики, и, следовательно, требующих вычислительных средств сверхвысокой производительности. Эти задачи давно стоят перед современными передовыми научными коллективами и представляют значительную научную проблему. Работа основана на использовании облачных моделей, методов и алгоритмов проектирования для консолидации данных и анализа программ и программных систем, позволяющих эффективно работать в многоплатформенных средах с неоднородными данными, и таким образом решать задачи обработки больших объемов информации для широкого диапазона возможных приложений.

**Научная новизна** диссертации определяется тем, что предложено и создано операционное окружение для работы приложений с интенсивной параллельной и распределенной обработкой информации, разработано специализированное программное обеспечение для создания, консолидации и управления распределенными структурами. Соискателем получены **новые научные результаты**, совокупность которых позволяет констатировать создания инструментария для разработки архитектуры и программно-аппаратного решения суперкомпьютерного центра поддержки государственного управления Союза Мьянма.

**Достоверность** научных результатов и выводов диссертации Тхуреин Киав Лина подтверждается тестированием методики консолидации баз данных в гетерогенном распределенном вычислительном ресурсе, что позволяет моделировать работу приложения с интенсивной параллельной и распределенной обработкой информации, и решением задачи увеличения эффективности NoSQL-баз данных на гетерогенных системах, предназначенных для проведения научных вычислений и поддержки систем принятия решений.

Диссертационная работа состоит из введения, пяти глав, заключения, списка литературы и приложений. Она содержит 249 машинописных страниц, 3 таблиц, 45 рисунков и 187 литературных ссылок.

**Во введении** обосновывается актуальность и важность темы диссертационного исследования, формулируются предмет исследования, цели работы, решаемые задачи, постановка задачи, методы исследования, практическая ценность, обосновывается научная новизна.

**В первой главе** рассмотрены базовые принципы правительства Мьянмы по развитию информационно-коммуникационных технологий, из которых ключевую роль играет создание Ресурсного центра для государственного управления, субъектов экономики, науки и образования. Его основная задача –

предоставление качественных информационных и вычислительных услуг, планирования социальных и экономических процессов, а также в альтернативной организации всесторонней информационно-вычислительной поддержки систем государственного управления Союза Мьянмы. Суперкомпьютерный центр должен быть укомплектован экспертами и специалистами в различных прикладных областях, способных оптимизировать использование вычислительных ресурсов и предлагать новые решения преодоления возникающих проблем.

**Во второй главе** рассмотрено информационно-программное обеспечение Ресурсного центра, отвечающее его задачам. Основная цель работы состояла в разработке модели и анализе инструментария для суперкомпьютерного центра. Основной упор делается на информационно-вычислительных системах с интеллектуальной поддержкой для решения сложных задач и с возможностью достаточно произвольно изменять размерность и содержание модели.

**В третьей главе** дается теоретический анализ и проводится исследование вопросов системной интеграции гетерогенных комплексов для улучшения характеристик их производительности и расширения диапазона приложений, разработка научных методов и алгоритмов организации параллельной и распределенной обработки информации, оптимизация приложений баз данных.

**В четвертой главе** рассмотрены вопросы анализа больших данных в специализированном информационно-ресурсном центре. Под анализом больших данных понимается как анализ массивов данных в рамках возможностей локального компьютера, так и в рамках возможностей систем управления реляционными базами данных, при этом как в первом, так и во втором случае при формировании и статистики, и визуализации возникают определенные трудности, которые заключаются в необходимости обеспечения скоординированной работы компьютерных программ на десятках, сотнях или

даже тысячах серверов. В этой главе также описывается состояние дел в области нереляционных баз данных. С точки зрения производительности, как Cassandra, так и Riak показывают хорошие результаты, а в некоторых ситуациях достигают лучших результатов, чем решения MySQL.

**В пятой главе** рассмотрены информационные технологии как объект стандартизации в Союзе Мьянма, цифровое развитие и создания Ресурсного центра в Союзе Мьянма. Республика Союз Мьянма имеет возможность обеспечить эффективное управление различными социально-экономическими условиями через использование системы цифрового «электронного правительства».

Особо следует отметить, что для решения такой задачи применяются специальные подходы, основанные на консолидации платформ, знакомство с которыми соискатель демонстрирует как в обзорной части своей работы, так и в процессе исследований. Отдельные успехи в этом направлении, достигнутые ведущими производителями, к сожалению, упираются в проблему их эффективности только на платформе производителей.

Несомненным достижением автора является то, что при решении такой проблемы основное внимание уделяется возможности объединения ресурсов с сильно отличающейся производительностью и их интеграции в гибридную распределенную среду многопоточных процессоров.

Приведенные в диссертации результаты тестирования разработанного автором программного комплекса, реализующего все описанные в работе алгоритмы, демонстрируют корректность и высокую эффективность разработанных подходов.

**Практическую значимость работы** составляют программные продукты, осуществляющие анализ больших данных в специализированном информационном-ресурсном центре, поддерживающие тестирование в

гетерогенной распределенной вычислительной среде, а также программный комплекс Riak и Cassandra в гетерогенном распределенном вычислительном комплексе.

Изложение в работе является ясным, четким, и последовательным.

Для экспериментальных работ результаты получены на лицензионном программном обеспечении, показана возможность воспроизведения результатов исследования на разных платформах и различных операционных системах. Методики построены на известных проверяемых данных и фактах, согласуются с опубликованными экспериментальными данными по теме диссертации. Автором использованы современные методики хранения и обработки больших объемов данных, представительные результаты экспериментов и подтверждения эффективности предложенных методов, современные подходы и модели для анализа и обработки больших массивов данных.

Среди научных результатов наиболее очевидными и значительными являются:

1. Методы консолидации данных с использованием Распределенной Федеративной СУБД;
2. Применение инструментальной формы вычисления ГРИД (Grid) для работы с распределенными базами данных;
3. Классификация больших данных, основанная на процедуре работы с ними;
4. Проект суперкомпьютерного центра для Правительства Республики Мьянмы;
5. Модель операционной среды, позволяющая симулировать работу приложения с интенсивной параллельной и распределенной обработкой информации;

По диссертационной работе следует сделать следующие замечания:

1) безусловно, основным результатом работы является объединение возможностей мощной распределенной Базы Данных. Однако никаких соображений в пользу того или иного выбора компонент объединенной системы в диссертации не приводится и, хотя правила работы с ней рассмотрены достаточно подробно, вопрос о ее оптимальности остается;

2) наибольший практический интерес в работе имеют эксперименты по масштабированию процесса обработки данных. Получаемые результаты могли бы эффективно использоваться для оптимизации больших систем обработки данных, но для этого количества процессоров и объем передаваемых данных нужно было бы существенно увеличить.

Отмеченные недостатки не влияют на значимость полученных соискателем результатов; рассматриваемая работа представляет собой законченное научное исследование, содержащее решение актуальной задачи, характеризующееся теоретической новизной и практической полезностью.

**Диссертационная работа** оформлена на достаточно хорошем уровне. Основные результаты опубликованы в печати и обсуждались на конференциях различного уровня. Нарушений п. 11 действующего «Порядка присуждения в Санкт-Петербургском государственном университете ученой степени кандидата наук, ученой степени доктора наук» не выявлено.

Диссертационная работа Тхуреин Киав Лина «Система консолидации данных и распределенных вычислений для поддержки информатизации Союза Мьянма» является **законченной, обладающей внутренним единством научно-квалификационной работой**. Считаю, что настоящая работа удовлетворяет всем требованиям действующего «Порядка присуждения в Санкт-Петербургском государственном университете ученой степени кандидата наук, ученой степени доктора наук», предъявляемым к докторским

диссертациям, а её автор – Тхурейн Киав Лин, как достаточно зрелый научный работник, **заслуживает присуждения учёной степени доктора технических наук по специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.**

Председатель диссертационного совета,  
доктор физ.-мат. наук, профессор,  
Санкт-Петербургский государственный  
университет



Тулупьев Александр Львович

04 декабря 2020 года