

## **ОТЗЫВ**

**Члена 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»**

Представленная на соискание ученой степени доктора технических наук диссертация Тхуреин Киав Лина определяется тем, что широкое применение систем поддержки принятия решений для государственного управления тормозится отсутствием эффективного инструментария для управления распределенными информационными системами государства и средств консолидации разнородных баз данных в различных ведомствах. Фактически, до недавнего времени такого инструментария не существовало вообще, и даже в развитых странах объединение данных различных ведомств сталкивалось и с техническими и с юридическими проблемами. В диссертации рассматривается технология консолидации баз данных для системной интеграции гетерогенных комплексов, решающих задачи проведения массивных вычислений и поддержки принятия решений.

Исходя из этого определяется основная **цель работы** – она посвящена анализу различных программных систем и программного обеспечения для консолидации серверов и гигантских объемов данных. Задачи создания оптимальных вычислительных архитектур, предназначенных для хранения и обработки информации сверхбольших объемов, наличествующих в настоящее время во многих областях государственного управления, и, следовательно, требующих вычислительных средств сверхвысокой производительности, стоят перед современными передовыми научными коллективами и представляют

значительную научную проблему. Научная идея работы основана на использовании модели распределенных СУБД, Грид технологий и концепции метокомпьютинга для консолидации данных и построения 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.

При решении этих проблем соискателем получены **новые научные результаты**, совокупность которых позволяет получить инструментарий для создания архитектуры и программного аппаратного решения суперкомпьютерного центра поддержки государственного управления республики Мьянма.

Диссертационная работа состоит из Введения, пять Глав, Заключение, Списка литературы и Приложений. Она содержит 249 машинописную страницу, 3 таблиц, 45 рисунков и 187 источников в списке литературы.

**Во Введении** обосновывается актуальность и важность темы диссертационного исследования, формулируются: предмет исследования, цели работы, решаемые задачи, постановка задачи, методы исследования, практическая ценность, а также положения, выносимые на защиту, и обосновывается научная новизна.

**В первой главе** рассмотрены ключевые принципы правительство Мьянмы по развитию информационно-коммуникационных технологий (ИКТ), из которых ключевую роль играет создание суперкомпьютерного центра для государственного управления, субъектов экономики, науки и образования. Его основная задача предоставление качественных информационных и вычислительных услуг и планирования социальных и экономических процессов и 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, которое должно отвечать задачам Ресурсного центра. Обеспечение состоит из собственно информационного массива (интеграции данных, баз данных, знаний и САПР), систем управления базами данных и интеллектуальных систем, из расчётного комплекса систем моделирования, виртуального полигона в составе систем обеспечения проектной и конструкторской деятельности (САПР), а также непосредственно корпоративных систем управления. Основная цель данной главы состояла в разработке модели и анализе инструментария для суперкомпьютерного центра. Основной упор делается на Информационно-вычислительных системах с интеллектуальной поддержкой для решения сложных задач с возможностью достаточно произвольно изменять размерность и содержание модели.

**В третьей главе** осуществлена теоретический анализ и экспериментальное 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 главе описываются применяемые алгоритмы распределенных вычислительных сред, технология консолидации и Больших данных, различные аппаратные средства и архитектура программного обеспечения используются для распределенного вычисления система. В данной главе рассмотрена вычислительная среда, которая может быть распределённой как из-за больших размеров, так и из-за неоднородного состава компонент, и позволяет пользователям совместное использование распределённых ресурсов на основе специфического промежуточного программного обеспечения и эти

среды могут быть разными по промежуточному программному обеспечению и по подходу к организации доступа к ресурсам.

**В четвертой главе** проанализирована концепция больших данных в специализированном информационно-ресурсном центре. Под анализом больших данных понимается как анализ массивов данных в рамках возможностей многопроцессорного компьютера, так и в рамках возможностей систем управления реляционными базами данных, при этом как в первом, так и во втором случае появляется необходимость обеспечения скоординированной работы компьютерных программ на десятках, сотнях или даже тысячах серверов. В этой главе автор описывает теоретические основы и последние достижения, а также концепции, лежащие в основе технологий, которые требуются для того, чтобы полностью интегрировать в систему нереляционные БД и используемые в `poSQL` СУБД. В основном описание посвящено модели данных и модели запросов, более подробно рассматриваются тестируемые базы данных `Cassandra` и `Riak`. Выводы иллюстрирует применением систем `NoSQL` в реальной ситуации, так что можно увидеть как преимущества, так и связанные с ними недостатки. С точки зрения производительности, как `Cassandra`, так и `Riak` показывают хорошие результаты, а в некоторых ситуациях, достигают лучших результатов, чем решения с `SQLDB`.

**В пятой главе** рассмотрены информационные технологии для внедрения в Союзе Мьянмы и ее цифровое развитие и создание Ресурсного центра в Союзе Мьянмы. В данной главе приведены предложения по насыщению создаваемого вычислительного центра компьютерными системами. К основным компонентам центра относятся вычислительный комплекс, сети и комплекс хранения данных. В данной главе автором показано, что создание подобного суперкомпьютерного Центра потребует решения множества проблем, как правовых, организационных и финансовых, так и технических.

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. Несомненным достижением автора является попытка при решении такой проблемы возможности объединения ресурсов с сильно отличающейся производительностью и интеграции в гибридную распределенную среду многопоточных процессоров.

Соискатель проявил себя как состоявшийся исследователь, решил целый ряд как принципиальных, так и технических проблем и создал программный продукт, который может служить основой для создания распределенных вычислительных комплексов поддержки государственного управления.

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

При этом соискателем получены **новые научные результаты**, совокупность которых позволяет получить инструментарий для создания такого операционного окружения для базы данных и консолидации в распределенной вычислительной среде, который позволяет управлять данными, распределенными в неоднородных вычислительных средах, и создавать виртуальные вычислительные комплексы, адаптированные под обработку данных заданного объема. Среди научных результатов наиболее очевидными и значительными являются:

1. Разработана методика консолидации баз данных в гетерогенном распределенном вычислительном ресурсе.
2. Создана модель операционной среды, позволяющая моделировать работу приложения с интенсивной параллельной и распределенной обработкой информации.

### 3. Создан проект суперкомпьютерного центра для Союза Мьянмы.

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.

**Диссертационная работа** оформлена на достаточно хорошем уровне. Основные результаты опубликованы в печати и обсуждались на конференциях различного уровня.

По диссертационной работе следует сделать следующие замечания:

1. Безусловно, основным результатом работы является объединение ресурсов при использовании 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 об ее оптимальности остается открытым.

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 машины и объем передаваемых данных нужно было бы существенно расширить.

Отмеченные недостатки не влияют на положительную оценку работы, которая представляет собой законченное научное исследование, содержащее

решение актуальной задачи, характеризующееся теоретической новизной и практической полезностью.

Диссертационная работа Тхуреин Киав Лина «СИСТЕМА КОНСОЛИДАЦИИ ДАННЫХ И РАСПРЕДЕЛЕННЫХ ВЫЧИСЛЕНИЙ ДЛЯ ПОДДЕРЖКИ ИНФОРМАТИЗАЦИИ СОЮЗА МЬЯНМА» является **законченной, обладающей внутренним единством научно-квалификационной работой.**

Диссертация Тхуреин Киав Лина на тему: «Система консолидации данных и распределенных вычислений для поддержки информатизации союза Мьянма» соответствует основным требованиям, установленным Приказом от 01.09.2016 № 6821/1 «О порядке присуждения ученых степеней в Санкт-Петербургском государственном университете», соискатель Тхуреин Киав Лин заслуживает присуждения ученой степени доктора технических наук по специальности 05.13.11 — Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей. Пункт 11 указанного Порядка диссертантом не нарушен.

Член диссертационного совета,

доктор физ.-мат. наук,  
профессор, профессор кафедры  
Вычислительной математики  
Санкт-Петербургского  
государственного университета



Бурова Ирина Герасимовна

\_4 декабря\_\_\_\_\_2020 года.