

Применение гибридного моделирования
для задач масштабирования
распределённых
информационных систем
на примере ИАС «ИСТИНА»

Зензинов Андрей Александрович
Абанкин Олег Алексеевич

МГУ имени М.В.Ломоносова

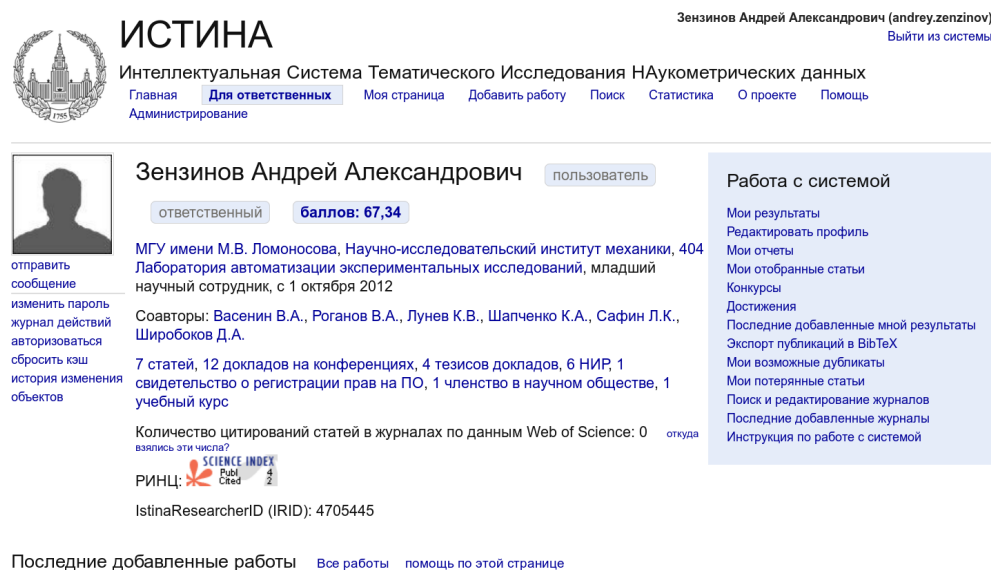
АПСПИ 2017
15 ноября 2017 г.

План доклада

- Введение
- Масштабирование ИС
- Модель распределённой ИС
- Параметры качества сервиса
- Эксперименты

ИАС «ИСТИНА»

Информационно-аналитическая система поддержки принятия решений в области науки и образования на основе наукометрических данных



ИСТИНА
Интеллектуальная Система Тематического Исследования Наукометрических данных

Зенин Андрей Александрович (andrey.zeninov) [Выйти из системы](#)

[Главная](#) [Для ответственных](#) [Моя страница](#) [Добавить работу](#) [Поиск](#) [Статистика](#) [О проекте](#) [Помощь](#)
[Администрирование](#)

Зенин Андрей Александрович пользователь



ответственный **баллов: 67,34**

МГУ имени М.В. Ломоносова, Научно-исследовательский институт механики, 404
Лаборатория автоматизации экспериментальных исследований, младший научный сотрудник, с 1 октября 2012

Соавторы: Васенин В.А., Роганов В.А., Лунев К.В., Шапченко К.А., Сафин Л.К., Ширококов Д.А.

7 статей, 12 докладов на конференциях, 4 тезисов докладов, 6 НИР, 1 свидетельство о регистрации прав на ПО, 1 членство в научном обществе, 1 учебный курс

Количество цитирований статей в журналах по данным Web of Science: 0 откуда
взялись эти числа?

РИНЦ:  

IstinaResearcherID (IRID): 4705445

[отправить сообщение](#)
[изменить пароль](#)
[журнал действий](#)
[авторизоваться](#)
[сбросить кэш](#)
[история изменений](#)
[объектов](#)

Работа с системой

- Мои результаты
- Редактировать профиль
- Мои отчеты
- Мои отобранные статьи
- Конкурсы
- Достижения
- Последние добавленные мной результаты
- Экспорт публикаций в BibTeX
- Мои возможные дубликаты
- Мои потерянные статьи
- Поиск и редактирование журналов
- Последние добавленные журналы
- Инструкция по работе с системой

Последние добавленные работы [Все работы](#) [помощь по этой странице](#)

- МГУ имени М.В.Ломоносова
- Пилотный проект с ФАНО по внедрению системы в институты (Москва, СПб, Ижевск, Иркутск)
- Новые требования к доступности Системы

Масштабирование системы

- Режим «облака» – веб-сервера и СУБД в едином ЦОД
- Веб-сервера в территориально-распределённых организациях, БД — в едином ЦОД
- Веб-сервера и БД в нескольких территориально-распределённых ЦОД

Актуальность моделирования распределённых систем

Масштабирование системы

- Взаимодействие с другими информационными системами
- Появление новых требований
- Изменение внутренней архитектуры системы
- Распределённая и сложная структура
- Необходимо учитывать работу для всех состояний системы (возможные отказы компонентов системы)
- Изменение готовых систем затруднено и влечёт за собой большие издержки

Гибридное моделирование

Гибридный подход предполагает возможность объединения отдельных видов моделирования (натурное, аналитическое, виртуальное, дискретно-событийное)

В зависимости от решаемой задачи возможно достичь необходимого уровня детализации

Такой подход активно используется в научных исследованиях.

Пример: использование сетевого симулятора ns-3 (д-с) с реальными приложениями (вирт.) и реальным оборудованием (натур.) в исследованиях передачи видеопотоков.

Цель работы

- Демонстрация возможности применения гибридного моделирования для оценки масштабируемости распределённых ИС
- Одна из целевых задач – исследование зависимости объективных (время загрузки страницы) и субъективных (качество восприятия сервиса) оценок работы от параметров сетевой инфраструктуры

Распространение на новые организации

- Взаимодействие с различными ИС других организаций
- Географическое распределение приводит к ухудшению канала связи
- Данные о научной деятельности могут быть тесно связаны между организациями

Другие задачи

- Принятие решений по развитию инфраструктуры
- Систематическое тестирование производительности на приближенных к реальным сценариях
- Непрерывная интеграция (автоматический запуск тестов)

Модель распределённой системы

- Модель распределённой системы:
 - Узлы системы
 - Сетевая инфраструктура
 - Программное обеспечение
 - Внешний мир (пользовательская активность, внешние сервисы)

Модель распределённой системы

- Узлы могут быть представлены в нескольких вариантах:
 - Физическая машина (реальное программно-аппаратное обеспечение)
 - Виртуальная машина (реальное программное обеспечение)
 - Логический узел в симуляции (каждый узел имеет своё представление в сети)
 - Моделирование поведения (узлы логически неразличимы)
- Сети могут быть смоделированы:
 - Физическая сеть (реальное сетевое оборудование)
 - Виртуальная сеть (NAT, bridge, изолированная сеть)
 - Симулируемая сеть

Моделирование пользовательской активности

Средства для нагрузочного тестирования веб-приложений:

- Apache Jmeter 3.1
- Простое создание сценариев (запись последовательности запросов через браузер)
- Возможность многопоточного запуска сценариев от нескольких пользователей
- Средства для анализа полученных результатов

Параметры качества сервиса

Субъективные параметры:

- QoE (Quality of Experience) – пользовательское восприятие сервиса
- MOS (Mean Opinion Score) – усреднённая пользовательская оценка по шкале от 1 до 5

Объективный параметр:

- QoS (Quality of Service) – качество обслуживания, опирается на измеримые параметры

Предложены международным комитетом ITU-T (International Telecommunication Union - Telecommunication sector)

Изучение зависимости параметров качества сервиса

1. Daisuke Yamauchi and Yoshihiro Ito A METHOD OF EVALUATING EFFECT OF QOS DEGRADATION ON MULTIDIMENSIONAL QOE OF WEB SERVICE WITH ISO - BASED USABILITY // International Journal of Computer Networks & Communications(IJCNC). – Vol.7, No.1, January 2015
2. ITU-T Nazrul Islam Vijaya John David Elepe On Factors Affecting Web Browsing-QoE Over Time // Master Thesis Electrical Engineering Thesis No: MEE100038. – January 2014
3. S. Canale, F. Delli Priscoli, S. Monaco, L. Palagi, V. Suraci A reinforcement learning approach for QoS/QoE model identification // Control Conference (CCC), 2015 34th Chinese. – July 2015
4. Markus Fiedler, Phuoc Tran-Gia A Generic Quantitative Relationship between Quality of Experience and Quality of Service // IEEE Network (Volume: 24, Issue: 2, March-April 2010). – March 2010

Оценка параметров качества сервиса

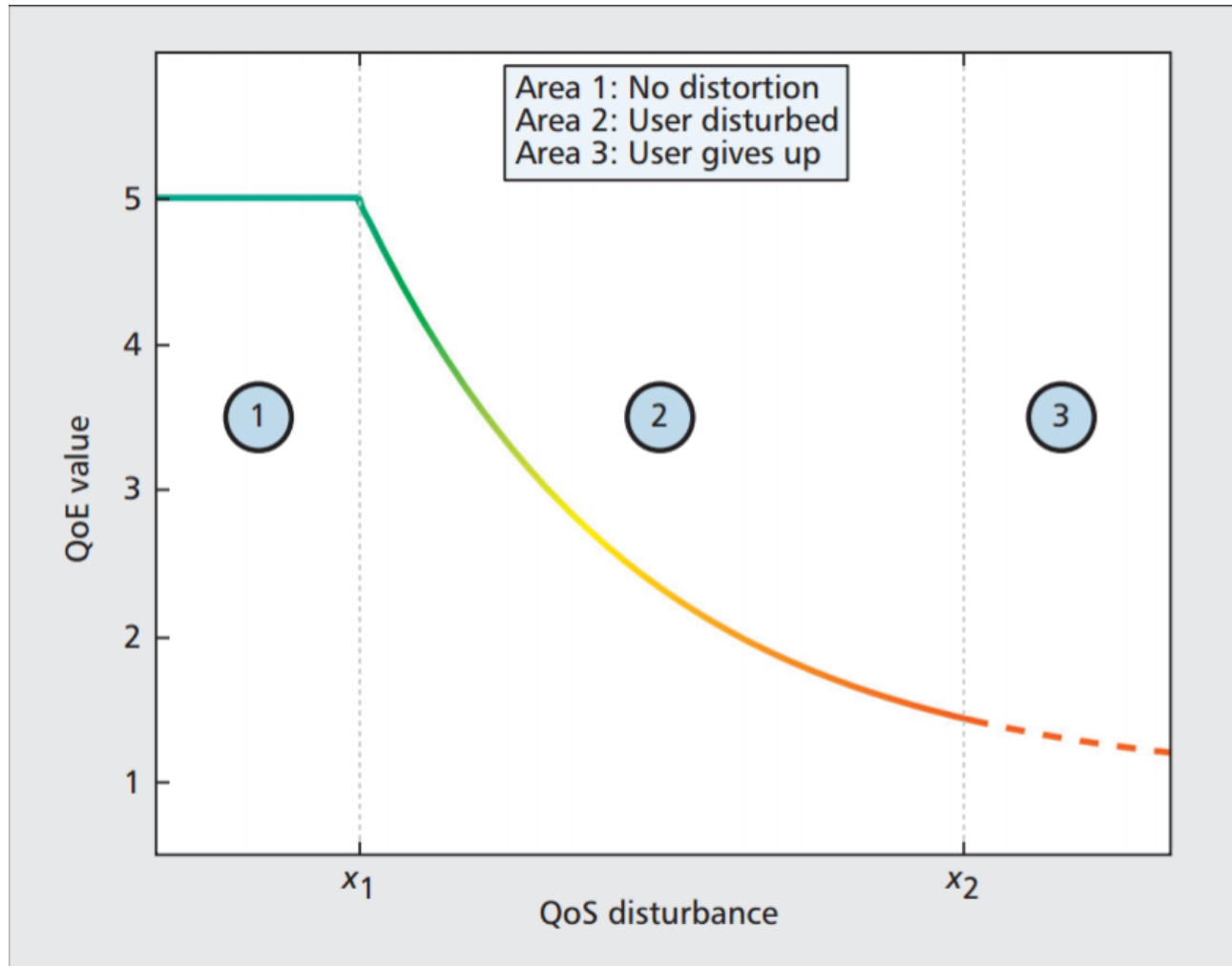
Три категории восприятия скорости сервиса:

- мгновенная загрузка - около 0,1 секунды;
- непрерывная загрузка - около 1 секунды;
- потеря внимания - около 10 секунд.

Jacob Nielsen

Восприятие работы сервиса напрямую зависит от ожиданий пользователя и типа решаемой задачи

Зависимость MOS от QoS



Зависимость MOS от времени сессии

В стандарте ITU-T G.1030 были получены числовые зависимости MOS от времени сессии

Для ожидания порядка десятков секунд зависимость, полученная по формуле ниже, имела корреляцию равную 0,9 с реальными значениями MOS. для более быстрых – 0.72

$$MOS = \frac{4}{\ln(Min / Max)} \cdot \left(\ln(SessionTime) - \ln(Min) \right) + 5$$

Описание эксперимента

Сценарий: поиск человека в системе и открытие его персональной страницы

- Веб-сервер и СУБД на виртуальных машинах
- Сеть моделируется двумя способами: NS-3 и с использованием netem

Географическое распределение

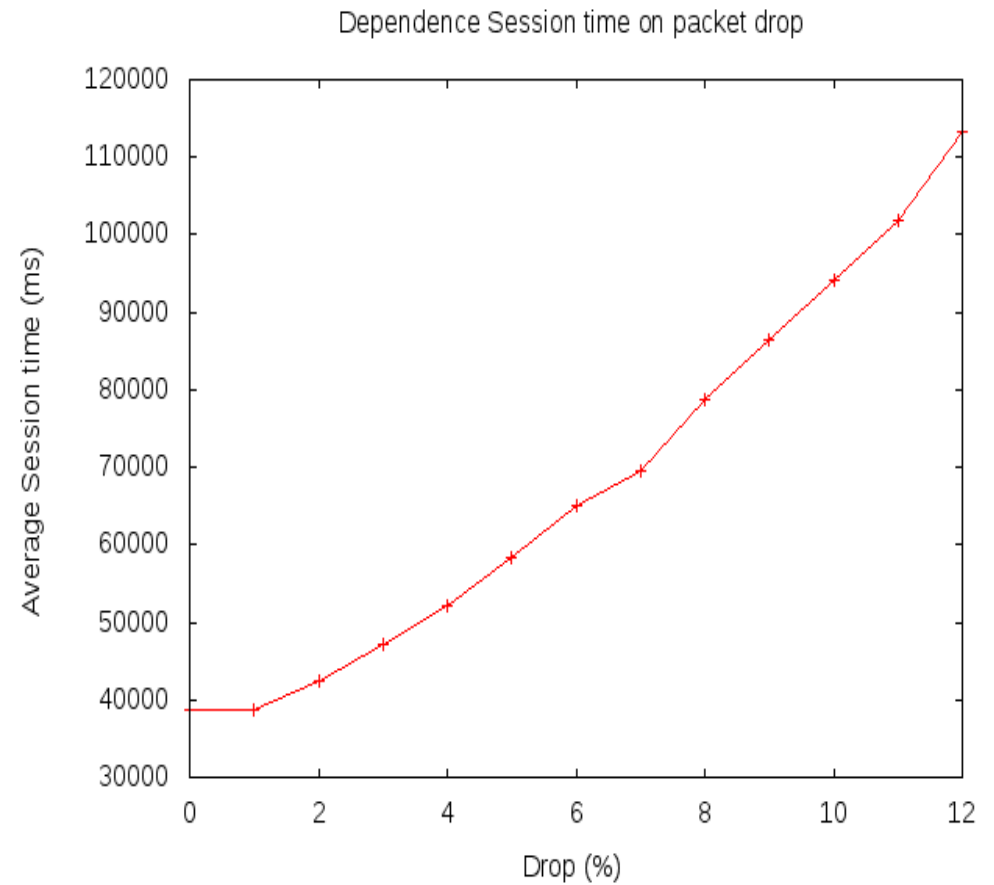
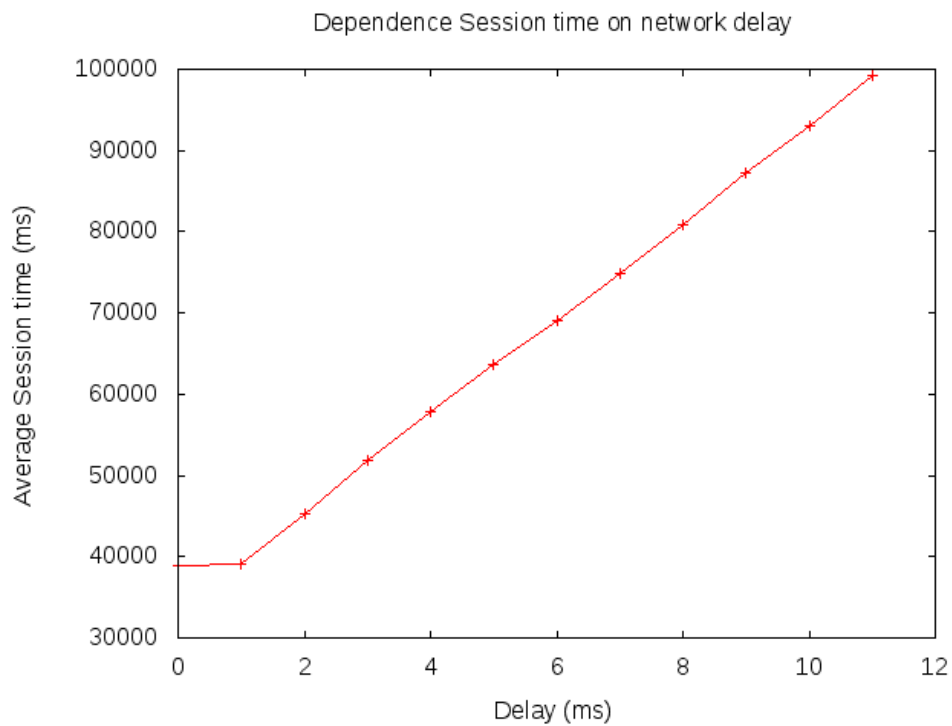
– Варианты распределения:

- Общий дата-центр
- Распределённые веб-сервера
- Распределённые дата-центры

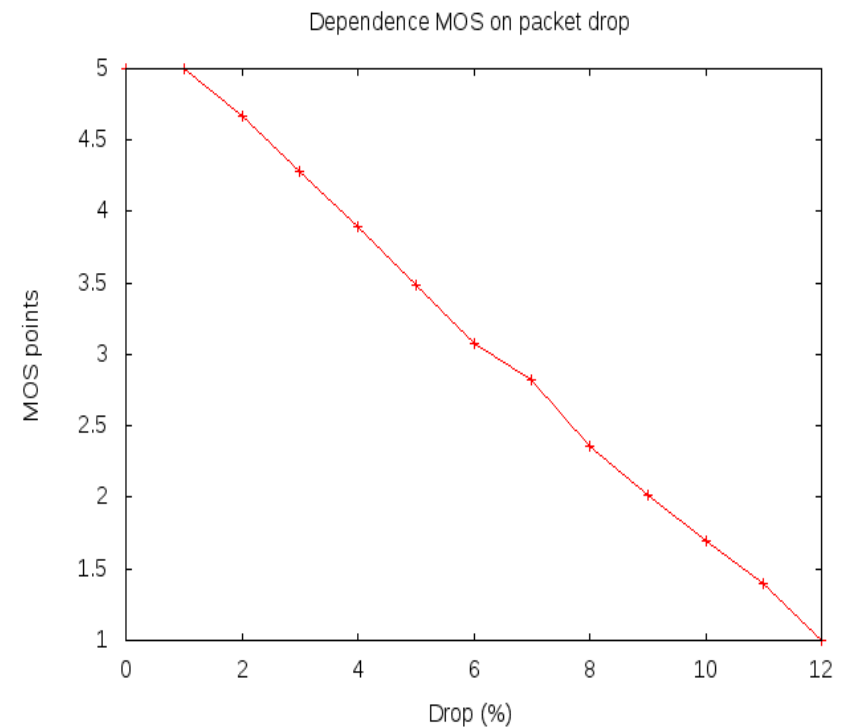
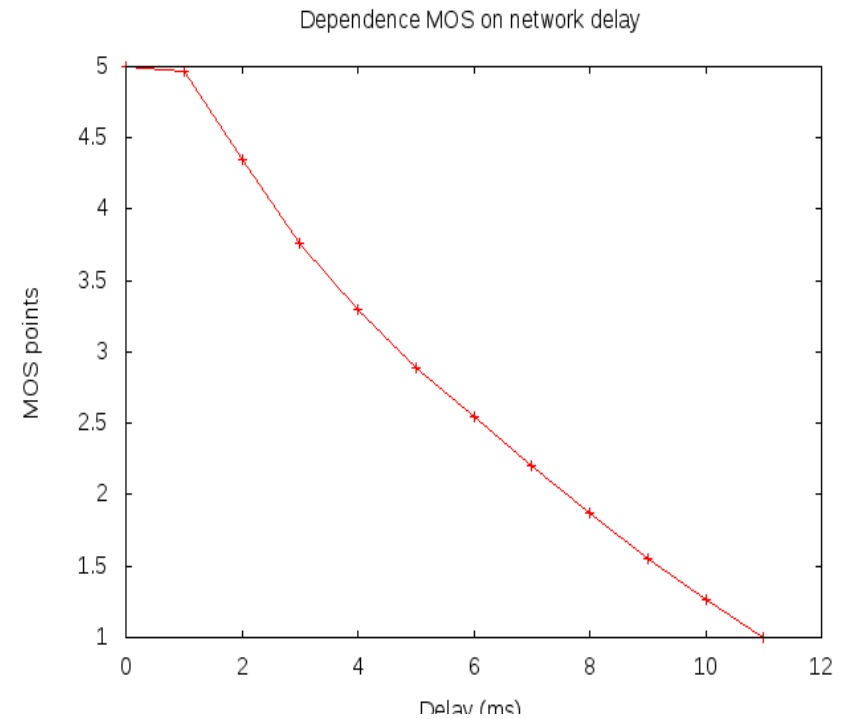
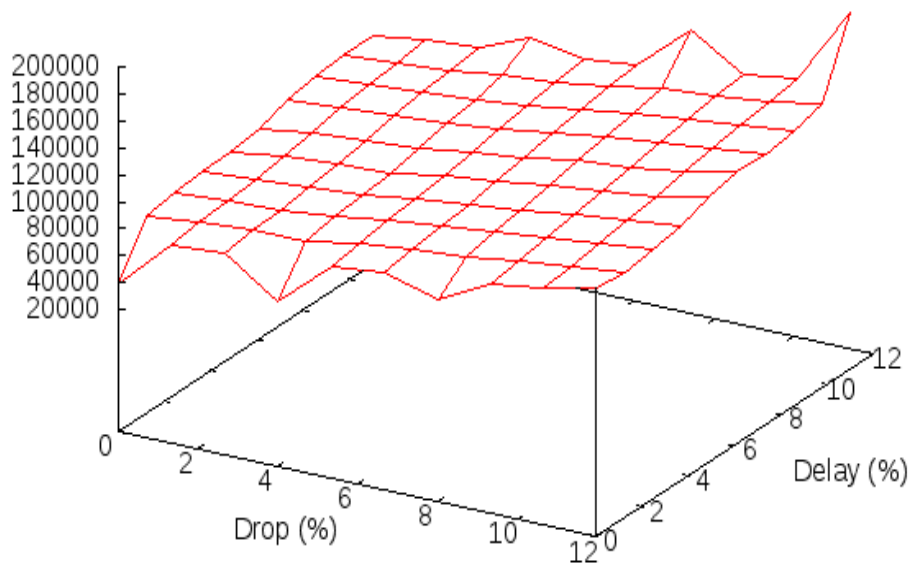
– Приблизительные RTT:

- Между соседними серверами < 0.3 мс
- Физико-технический институт УрО РАН,
г. Ижевск (~ 1000 км), <http://ftiudm.ru>, 10 мс
- Институт динамики систем и теории управления имени
В.М.Матросова СО РАН, г. Иркутск (~ 5000 км), <http://www.icc.irk.ru>,
72 мс
- Дальневосточный Федеральный Университет,
г. Владивосток (~6500 км), <http://dvfu.ru>, 140 мс.

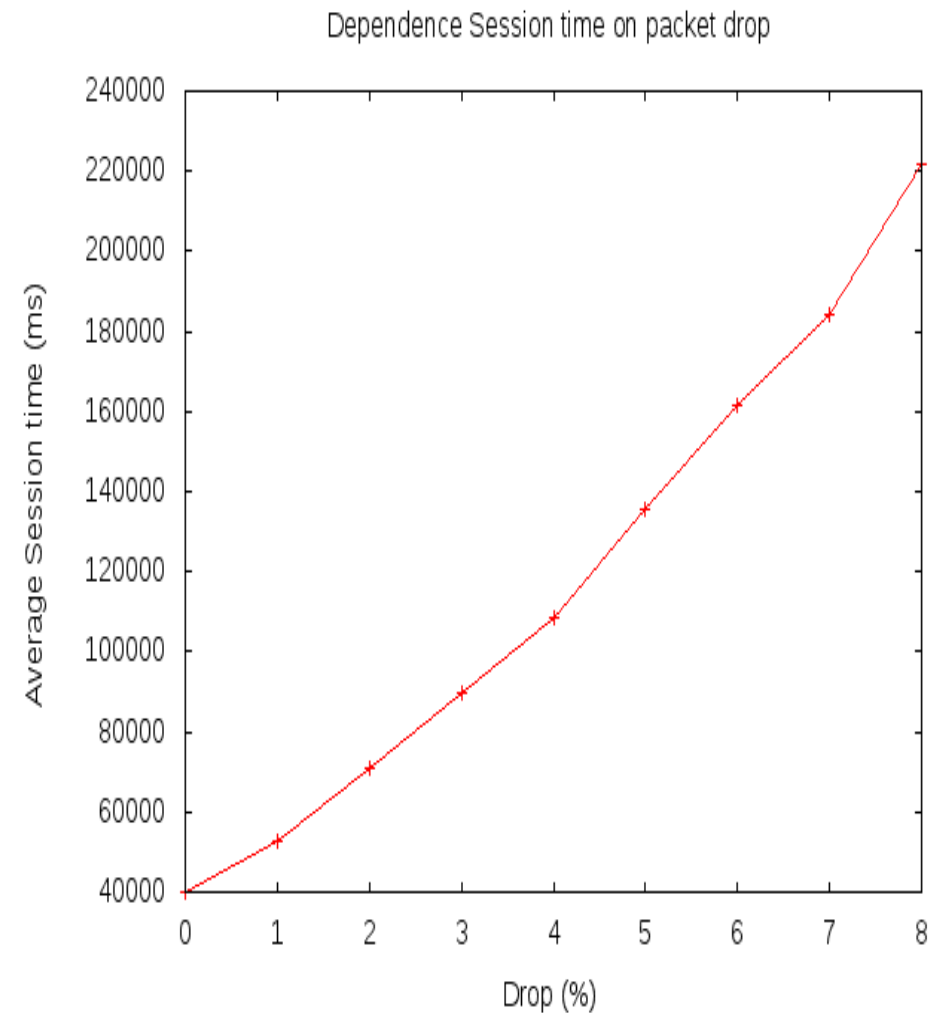
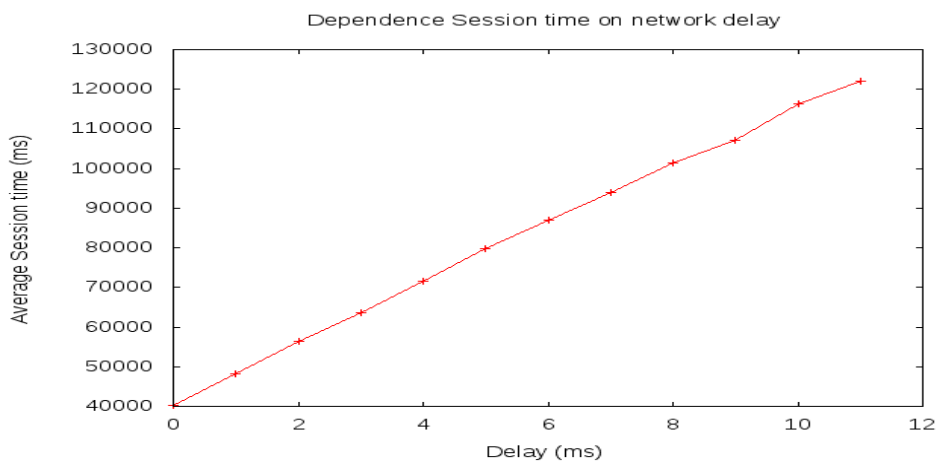
Зависимость времени сессии от сетевых характеристик (NS-3)



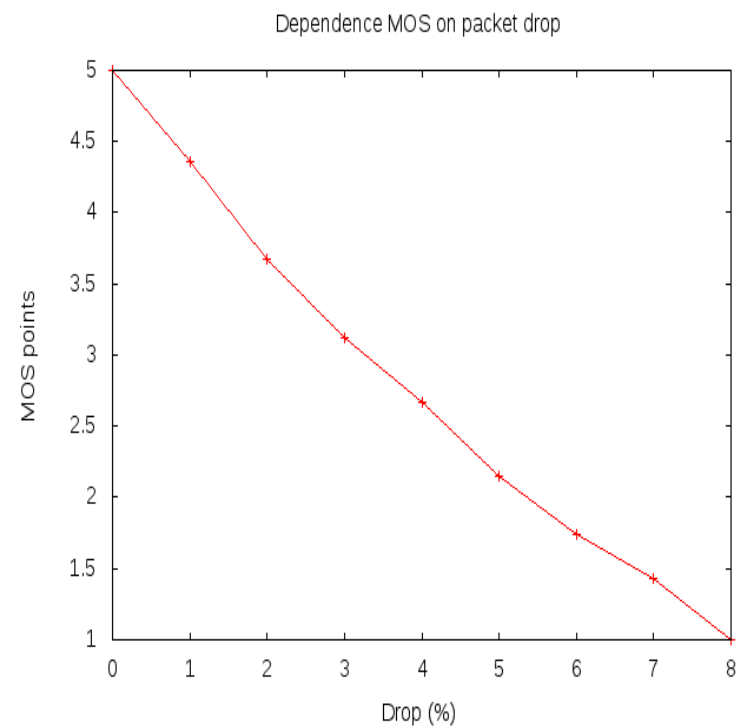
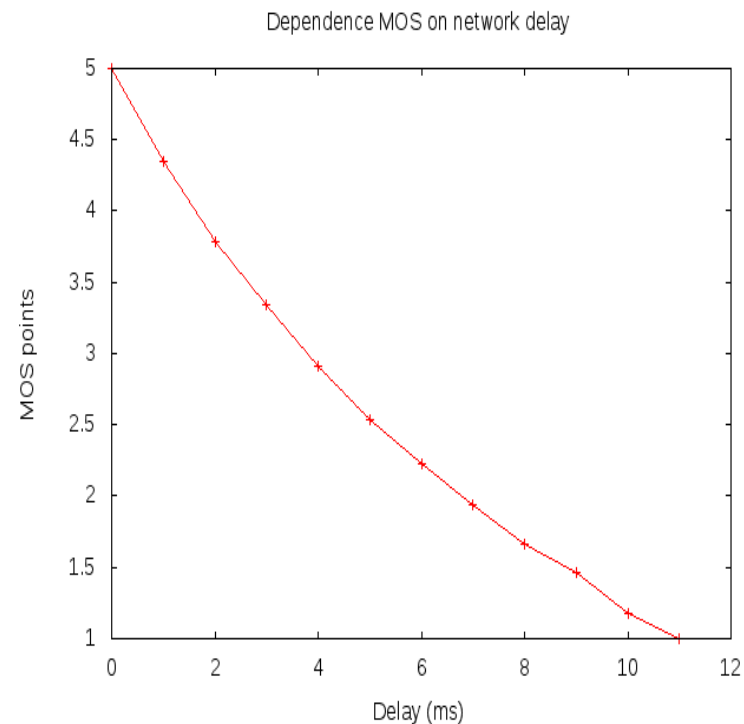
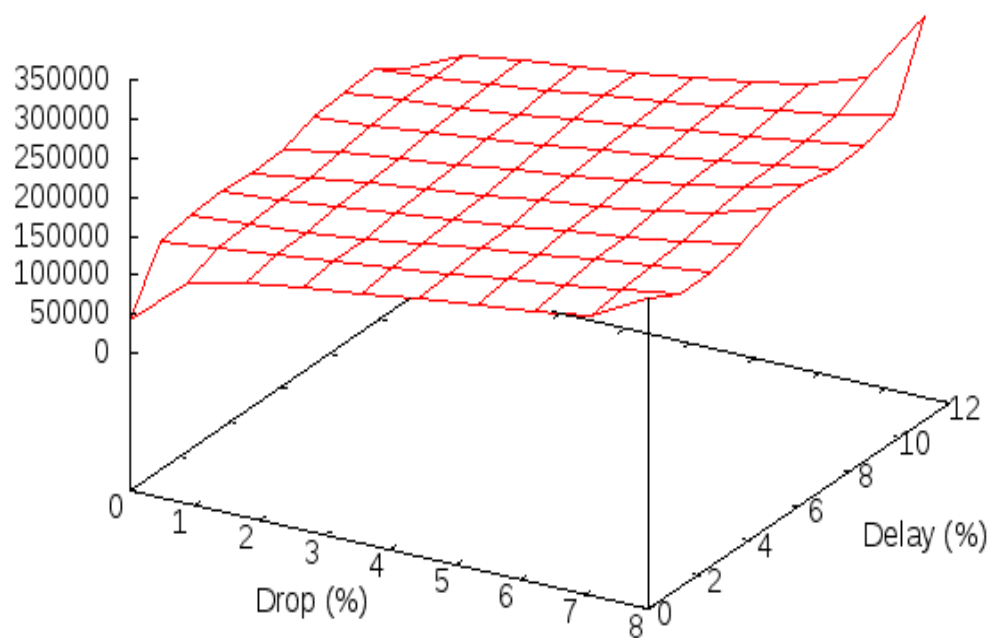
Зависимость MOS от сетевых характеристик



Зависимость времени сессии от сетевых характеристик (netem)



Зависимость MOS от сетевых характеристик



Полученные рекомендации

Рекомендации по требуемым сетевым характеристикам на участке между веб-сервером и сервером БД для достижения MOS ≥ 3 :

- задержка не более 4 мс
- доля потери пакетов не более 3%

Рекомендуется размещать веб-сервера Системы и сервера баз данных, обслуживающих организацию, по возможности, в едином ЦОД. Размещение веб-сервера на значительном удалении от серверов БД приведёт к появлению сильных задержек.

Заключение

- Описан подход к гибриднему моделированию распределённых ИАС
- продемонстрированы возможные применения подобного подхода для решения различных задач по оценке степени их масштабируемости
- рассмотрен пример задачи по предложению рекомендуемых требований к сетевым характеристикам внутренней сетевой инфраструктуры Системы при условии её распределения на новые организации
- получена и проанализирована зависимость времени сессии от таких параметров сети, как задержка и потеря пакетов

Спасибо за внимание!