

# Оптимизация скорости моделирования многоядерных микропроцессоров с применением нейронных сетей

Выполнил: Александр Голенев, 816 гр.  
alex.golenev@gmail.com

Научный руководитель: д.т.н. Дроздов А. Ю.

Факультет радиотехники и кибернетики  
Кафедра информатики и вычислительной техники

МФТИ, 2012

- Создать моделирующий программный комплекс, использующий нейронные сети для ускорения процесса моделирования
- Сравнить скорости процессов моделирования с применением нейронных сетей и без него
- Рассчитать ошибку, добавляемую применением нейронных сетей к результатам моделирования

# Поиск оптимальной конфигурации многоядерного процессора

## Проблемы

- Конфигурации образуют большое параметрическое множество
- Существующие симуляторы не обладают достаточной производительностью для возможности оценки всех конфигураций
- Неинтуитивные зависимости между параметрами конфигурации

# Поиск оптимальной конфигурации многоядерного процессора

## Возможное решение

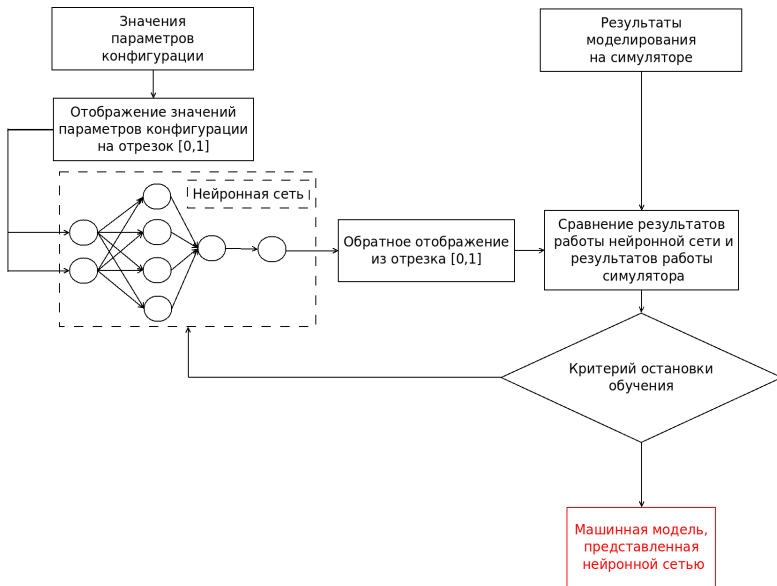
- Создать выборку результатов, на основе симуляции небольшого подмножества конфигураций
- Обучить машинную модель с помощью полученной выборки
- В дальнейшем использовать только машинную модель, не тратя время на симуляцию

- Линейная регрессия - не выявляет нелинейных зависимостей
- Аппроксимация полиномами - даёт плохие предсказания вне обучающей выборки
- Радиальные базисные функции - точность сильно зависит от выбора конкретной функции
- Нейронные сети

Мы выбрали нейронные сети, потому что они:

- Не требуют начальных знаний зависимостей между параметрами
- Выявляют любые нелинейные зависимости
- Работают с зашумлёнными данными
- Хорошо исследованы и широко применяются в различных областях знаний

# Нейронные сети для моделирования микропроцессоров

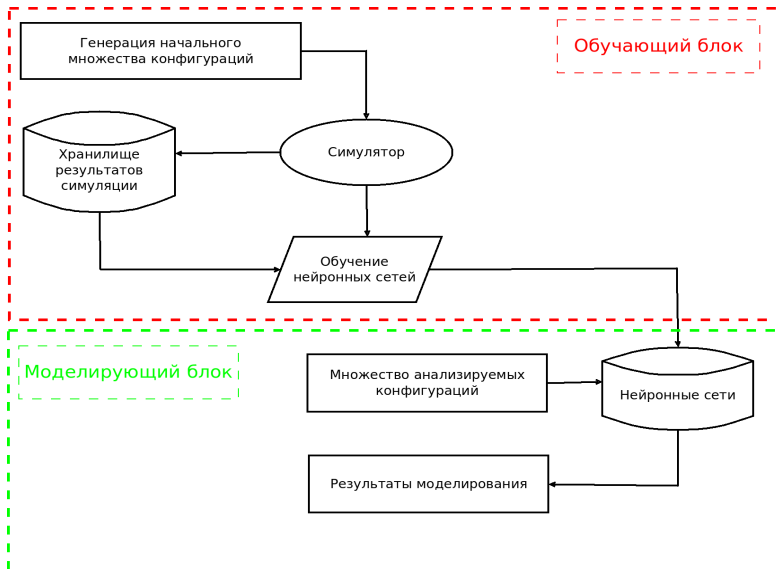


## Конфигурация нейронной сети

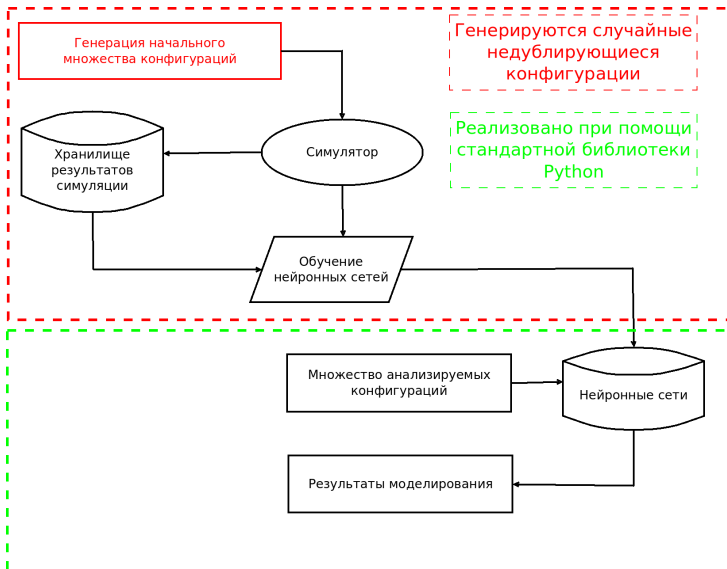
- 1 скрытый слой
- 16 элементов в скрытом слое
- сигмоидная выходная функция
- соединения типа "каждый с каждым" между слоями
- стандартная топология для начального экспериментирования



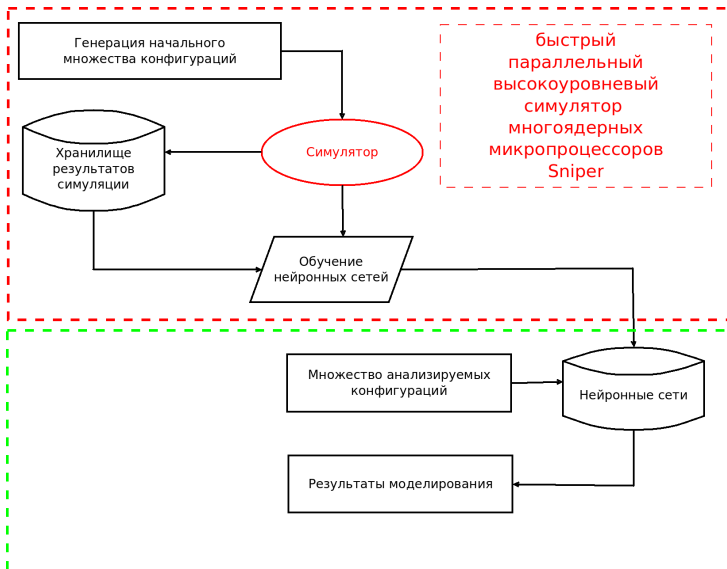
# Моделирующий комплекс с применением нейронных сетей



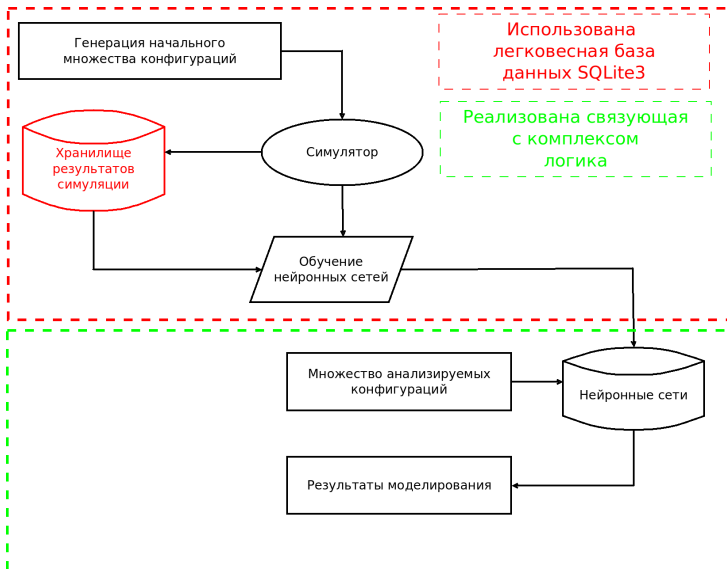
# Моделирующий комплекс с применением нейронных сетей



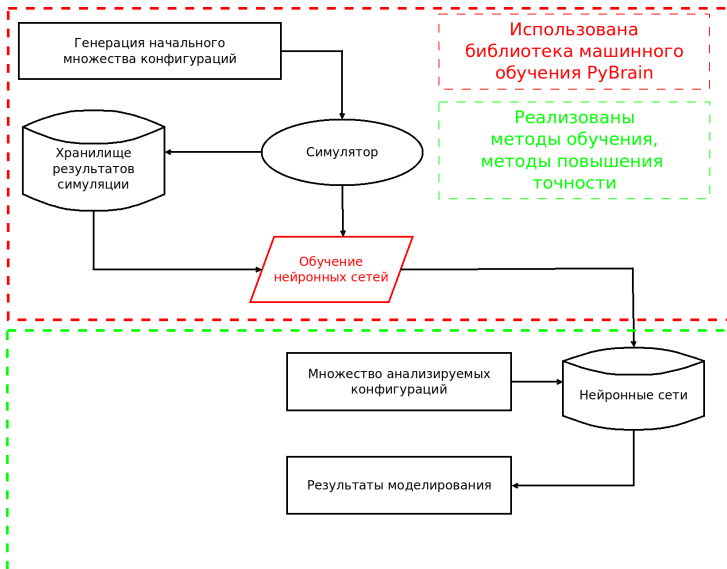
# Моделирующий комплекс с применением нейронных сетей



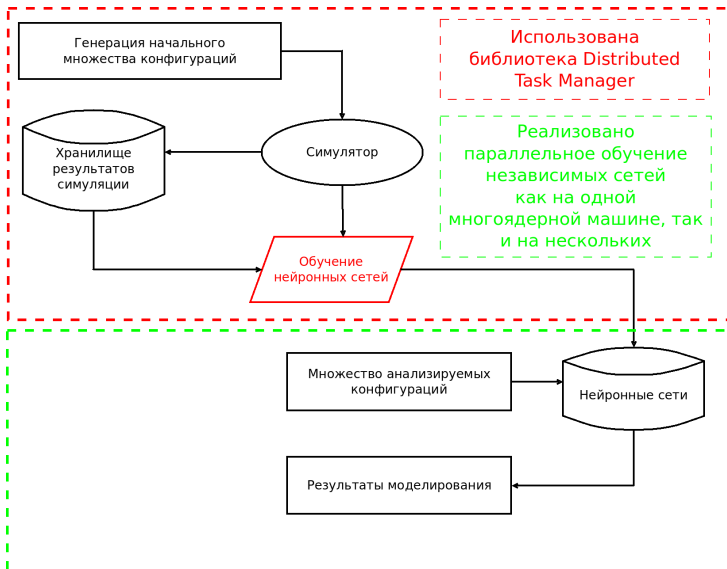
# Моделирующий комплекс с применением нейронных сетей



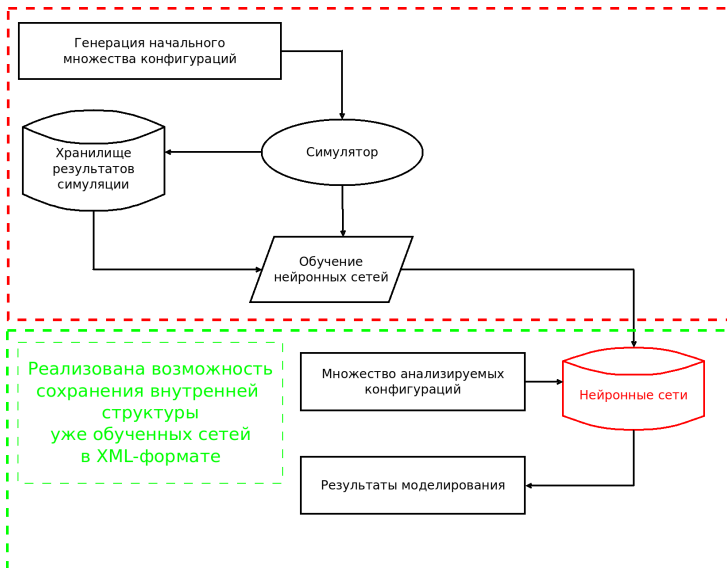
# Моделирующий комплекс с применением нейронных сетей



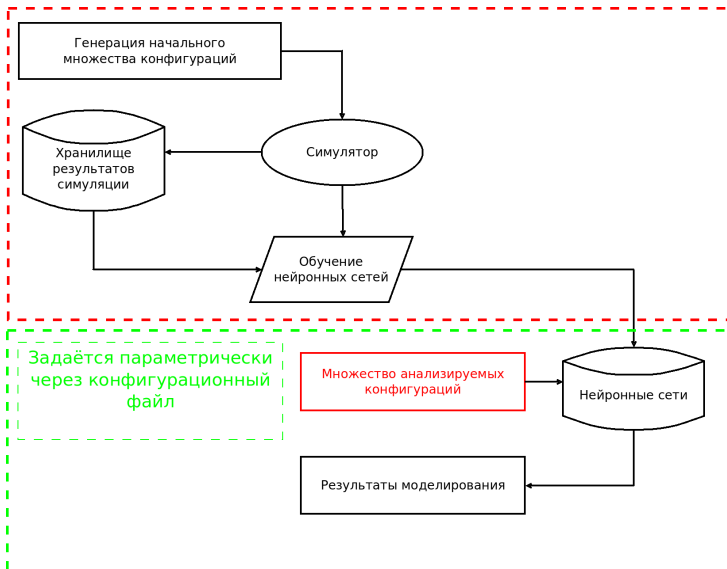
# Моделирующий комплекс с применением нейронных сетей



# Моделирующий комплекс с применением нейронных сетей

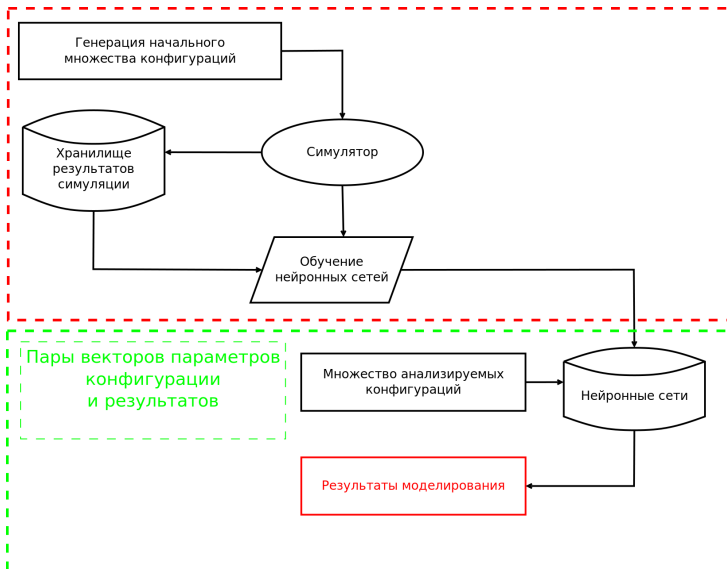


# Моделирующий комплекс с применением нейронных сетей





# Моделирующий комплекс с применением нейронных сетей



- Симулировались приложения из бенчмарка Splash2
- Всего - 2160 конфигураций
- Использовались небольшие наборы входных данных
- Нейронные сети обучались на выборке, составляющей 10% от множества конфигураций
- Все конфигурации были просимулированы

# Моделирование (шаблон моделируемой архитектуры)

- Одинаковые ядра (x86 Nehalem)
- L1, L2 – для каждого ядра
- L3 – общий для 4-ёх ядер
- Ядра соединены общей шиной

## Варьируемые параметры

Название параметра	Значения
Число ядер	{4, 8, 16}
Размер L1-D кэша	{16k, 32k, 64k}
Ассоциативность L1-D кэша	{4, 8}
Размер блока L1-D кэша	{32, 64}
Размер L2 кэша	{64k, 128k, 256k}
Ассоциативность L2 кэша	{4, 8}
Размер блока L2 кэша	{32, 64}
Размер L3 кэша	{1M, 2M, 4M, 8M, 16M}

## Традиционный подход

Симуляция всех конфигураций для каждого из приложений заняла 5 дней

## Подход с применением нейронных сетей

12 часов на получение обучающей выборки + 10 минут на обучение

## Точность

Приложение	ошибка (%)
Splash2.barnes	10.6
Splash2.cholesky	14.0
Splash2.fft	5.9
Splash2.fmm	8.1

Таблица: Средние относительные ошибки предсказаний

- Создан моделирующий комплекс, ускоряющий процесс моделирования с помощью нейронных сетей
- Проведено количественное сравнение скорости и точности традиционного подхода и такового с применением нейронных сетей

- С помощью нейронных сетей можно значительно ускорить процесс моделирования
- Использование нейронных сетей несколько снижает точность, но не настолько, чтобы отказываться от их преимущества применения