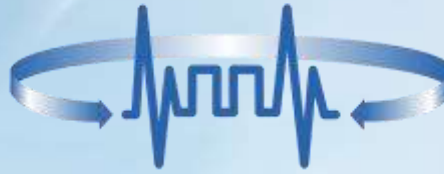


РОССИЙСКАЯ АКАДЕМИЯ НАУК



ИНСТИТУТ ПРОБЛЕМ ПЕРЕДАЧИ ИНФОРМАЦИИ  
имени А.А. Харкевича

DATADVANCE

# Иерархическая модель потока данных с автоматическим управлением файлами для инженерных и научных приложений

Алексей М. Назаренко (докладчик)  
Александр А. Прохоров

# План

- ⑩ Задача оптимизации
- ⑩ Требования к системе
- ⑩ Модель потока работ
- ⑩ Сохранение результатов вычислений



# **Задача ОПТИМИЗАЦИИ**



# Мотивация

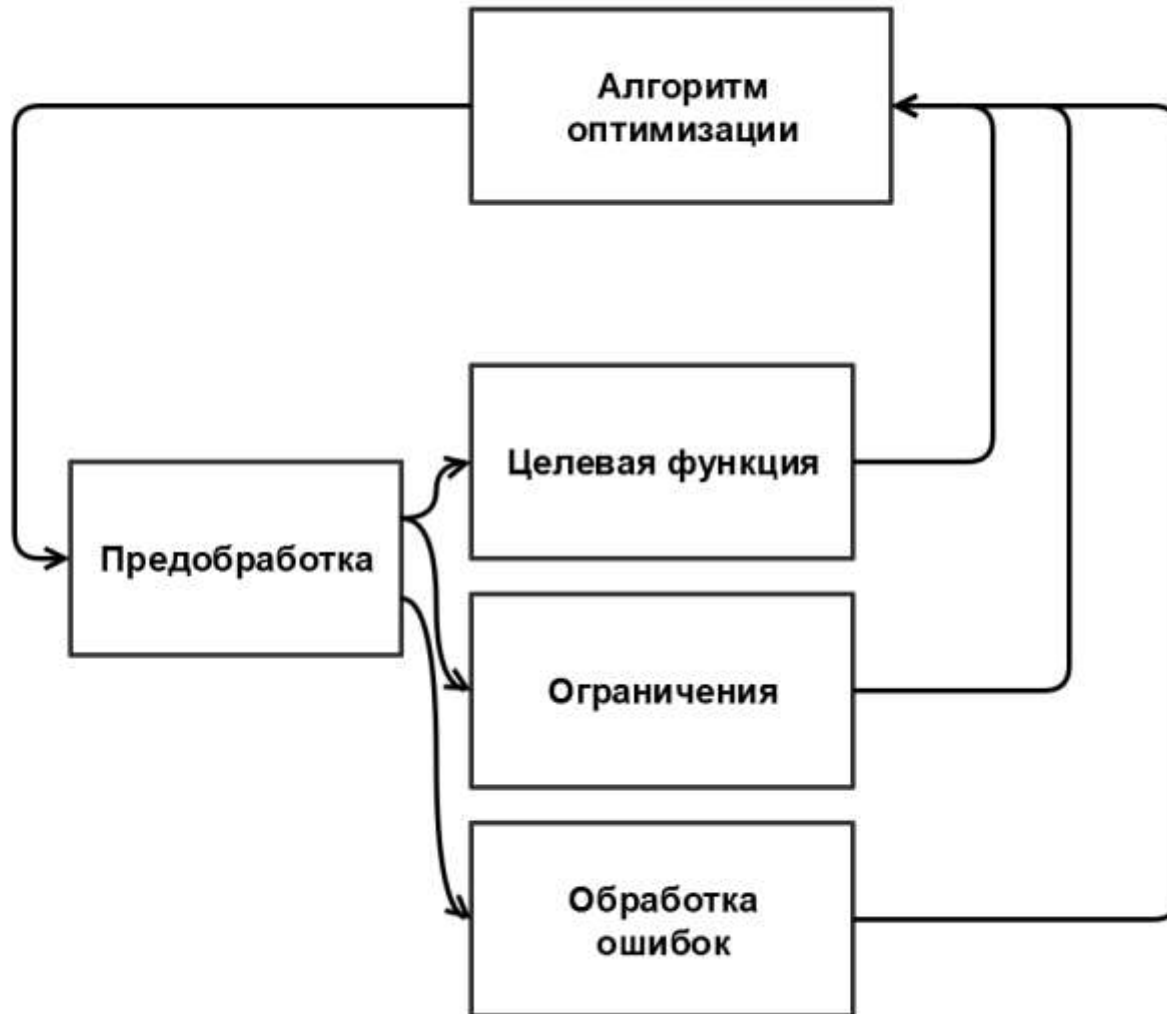
---

- ⑩ Задача оптимизации черного ящика рассматривается как показательный пример
- ⑩ Оптимизация является одной из задач, требующей полной автоматизации вычислений
- ⑩ Устройство черного ящика часто естественно представляется в виде потока работ



# Общая схема

- Поток работ – набор задач с заданной последовательностью запуска



# Инженерный пример



# Требования к системе



# Функциональные требования

- ⑩ Возможность представления циклов и вложенных циклов
- ⑩ Возможность представления ветвлений
- ⑩ Параллельное выполнение независимых вычислений
- ⑩ Поддержка распределенных вычислений
- ⑩ Сохранение истории вычислений и всех использованных файлов





# Дополнительные требования

- ⑩ Естественное соответствие потока работ и неформальной схемы решения задачи
- ⑩ Возможность сопоставить вычисления и использованные или созданные файлы
- ⑩ Минимизация включения в схему технических элементов, не имеющих предметного смысла



# Модель потока работ

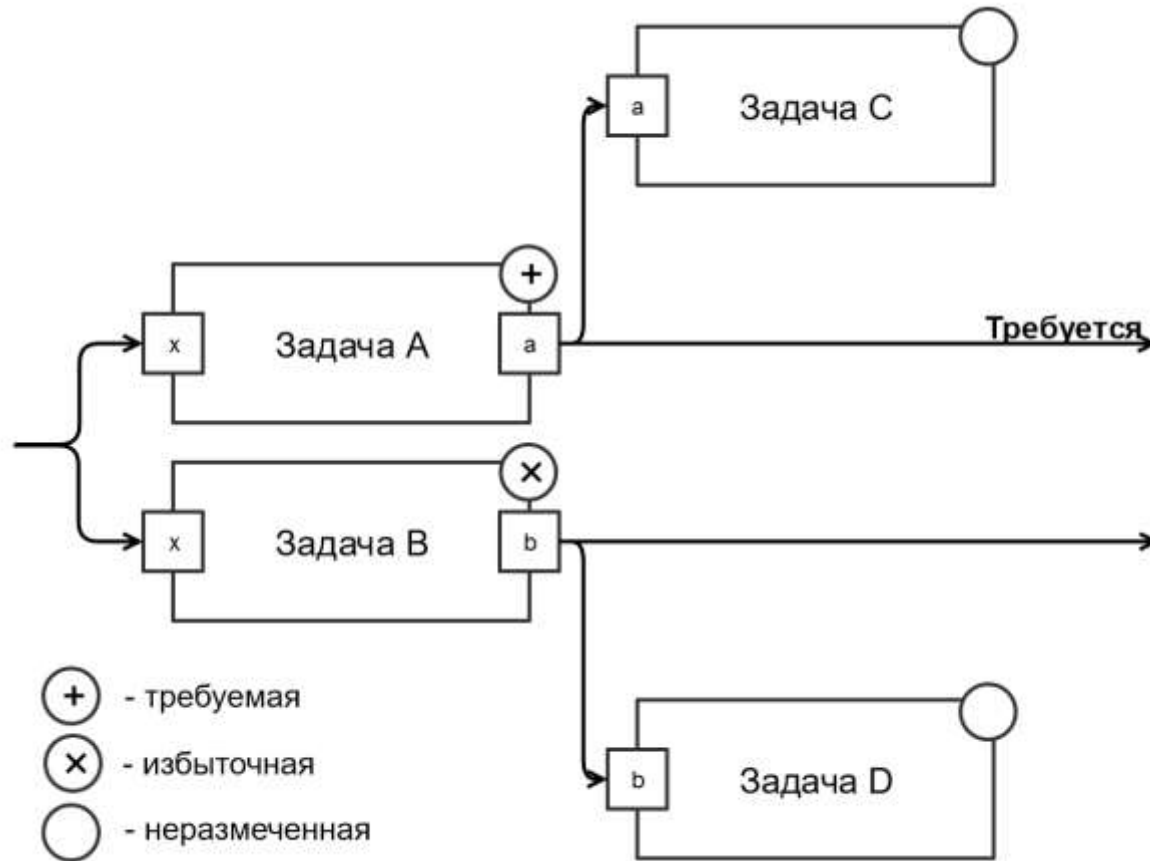


# Определения

- ⑩ Поток работ состоит из задач
- ⑩ Каждая задача имеет набор входов и выходов
- ⑩ От выхода к входу можно провести связь



# Базовая модель



- Модель потока данных
- Каждая задача требует все входы и выписывает все выходы
- Неподключенные входы/выходами становятся параметрами и результатами потока работ
- Циклы и множественные входящие связи не допускаются

# Иерархия

- ⑩ Задачи могут быть составными
- ⑩ Дочерние задачи образуют внутренний поток работ, соответствующий базовой модели
- ⑩ Внутренний поток работ может быть запущен произвольное число раз с различными параметрами и запрашивая различные выходы



# Задача оптимизации

- Решение демонстрационной задачи с помощью предлагаемой модели
- Простое соответствие логических блоков и задач в потоке работ
- Обработка некорректных моделей может быть легко реализована с помощью составных задач («ветвление»)



# Сохранение результатов вычислений



# Необходимость работы с файлами

- ⑩ Многие из задач прямо или косвенно представляют запуск внешних программ
- ⑩ Обмен данными часто происходит через файлы, что без специальной обработки нарушает модель потока данных
- ⑩ Файлы, полученные во время выполнения потока работ, часто представляют ценность и должны быть сохранены





# Известные подходы

- ⑩ Два основных подхода
- ⑩ Все файлы доступны всем задачам (используется общая файловая система)
  - Нарушает модель потока данных, требует введения дополнительной синхронизации в поток работ
  - При распределенном запуске появляется необходимость использования сетевых файловых систем и передачи большого объема лишних данных по сети
- ⑩ Передавать каждый файл через входы/выходы задач
  - Не всегда известны все файлы, которыми нужно обмениваться; многие форматы данных представляют из себя сложные файловые структуры



# Решение

- ⑩ Каждый запуск каждой задачи происходит в некоторой локальной директории – песочнице
- ⑩ Если задачи соединены в потоке работ, последующая задача может получить доступ на чтение к песочнице предыдущей
- ⑩ Составные задачи назначают песочницы для своих дочерних задач, что позволяет добиться логической группировки вычислений
- ⑩ В случае распределенного выполнения
  - Система выполнения потока работ обеспечивает синхронизацию песочниц
  - В конце выполнения все песочницы образуют единую структуру на головном узле запуска



# Пример структуры песочниц

## ⑩ Запуск 1

- Оптимизация
  - Точка 1
    - Построение геометрии
      - <Входные параметры>
      - <Модель>
    - Расчет массогабаритных характеристик
      - <Входы/выходы>
    - Расчет прочности
      - <Входы/выходы>
      - <График распределения напряжений>
  - Точка 2
    - ....
    - ....
  - ....
  - <оптимальная точка>



# Заключение



# Результаты

---

- ⑩ Разработана модель потока работ для простой постановки задач оптимизации
- ⑩ Результаты запуска задач образуют строгую структуру, удобную для последующего анализа
- ⑩ Предложенная модель на данный момент внедряется в коммерческую систему управления потоками работ



# Контакты

**Назаренко Алексей**  
[alexmnazarenko@gmail.com](mailto:alexmnazarenko@gmail.com)

**ИППИ РАН**  
[www.iitp.ru](http://www.iitp.ru)

**DATADVANCE**  
[www.datadvance.net](http://www.datadvance.net)

*Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 15-29-07043.  
Условия для выполнения работ по проекту предоставлены ООО "ДАТАДВАНС".*

