



**Институт программных систем  
Российской академии наук**

**Разработка  
высокопроизводительных  
клиент-серверных приложений  
для работы  
с данными дистанционного  
зондирования Земли**

Московский А. А., Первин А.Ю.

Институт Программных Систем РАН  
**Москва, ИКИ РАН, 14 ноября 2006**



# Тезисы доклада

- ❑ Кластерная архитектура - эффективное средство повышения производительности при обработке данных ДЗЗ.
- ❑ Технологии «клиент-сервер» позволяют создавать мощные параллельные программные средства с интуитивно понятным интерфейсом.
- ❑ Шаблоны параллельного программирования



# План доклада

- Институт Программных Систем
- «Платформа» для создания приложений обработки данных ДЗЗ
- Среда распараллеливания вычислений
- Тестовое приложение: классификация мультиспектральных изображений



# ИПС РАН, Переславль-Залесский





## Флагман: «СКИФ К-1000»



- ❑ **Пиковая производительность: 2,5 Tflops**
  - ❑ **Linpack-производительность: 2,0 Tflops**
  - ❑ **КПД=80.1 %**
- 
- ❑ **Ноябрь 2004: Наиболее мощная машина на территории СССР**
  - ❑ **Ноябрь 2004: № 98 в Top500**

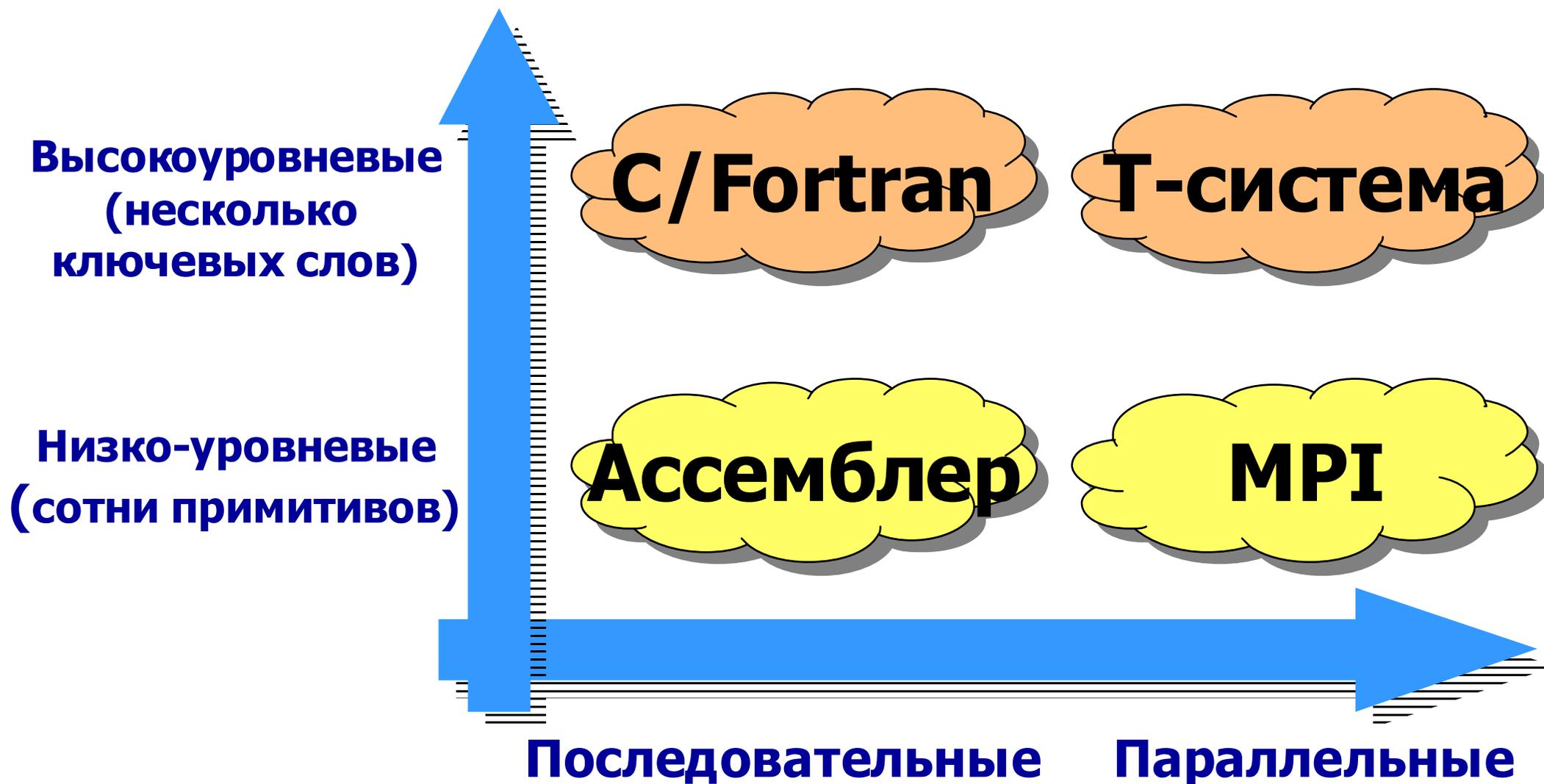


# Параллельные программы

- ❑ Кластеры – много компьютеров, соединённых вместе.
- ❑ Проблема: как запрограммировать?
- ❑ Решение: использование высокоуровневых средств программирования:
  - ★ Т-система (результат программы «СКИФ»)



# Сравнение: Т-Система и МРІ





## T-Подход

- ❑ Высокоуровневое программирование
- ❑ Автоматическое динамическое распараллеливание программ
- ❑ Сочетание функциональной и императивной парадигм (и ООП)
- ❑ T++ язык: неготовые значения, «чистые» T-функции, всего 7 ключевых слов.
- ❑ Реализации для C/C++, Fortran, Рефал, C#
- ❑ Работает на SMP, кластерах и мета-кластерах



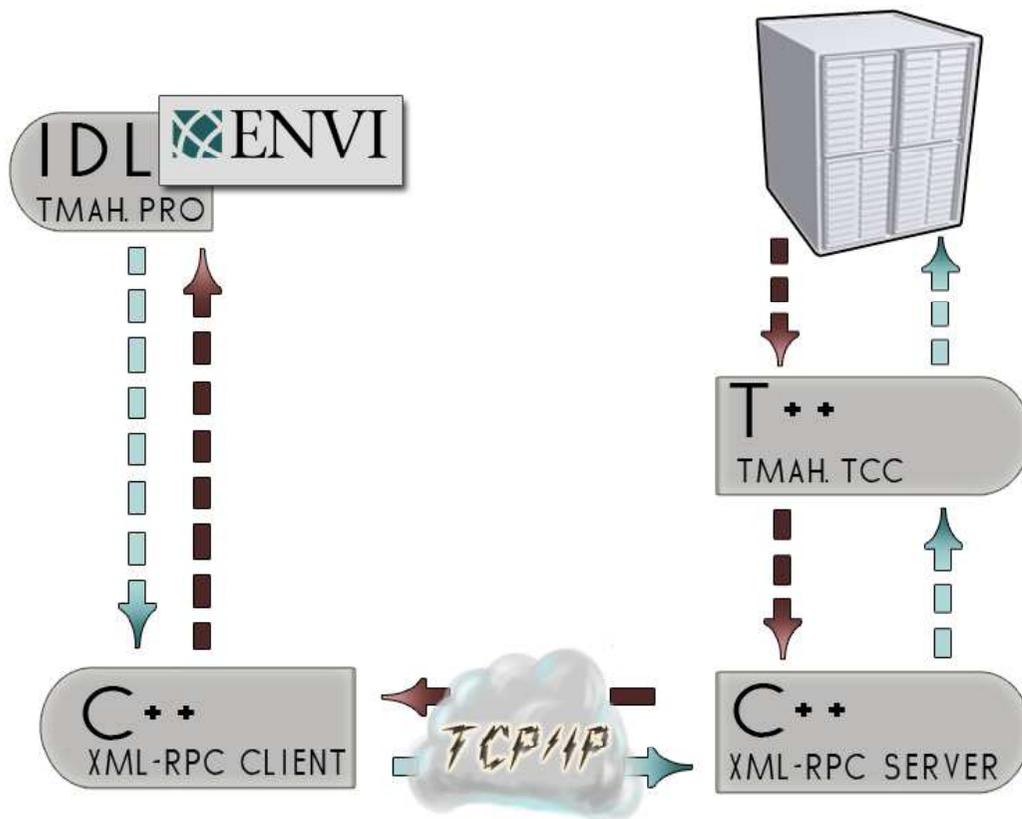
**Институт программных систем  
Российской академии наук**

**Решение для обработки данных  
ДЗЗ с использованием  
высокопроизводительных  
кластеров**





# Клиент-серверная архитектура

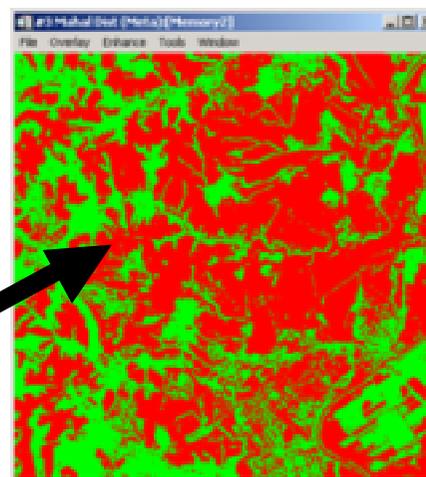
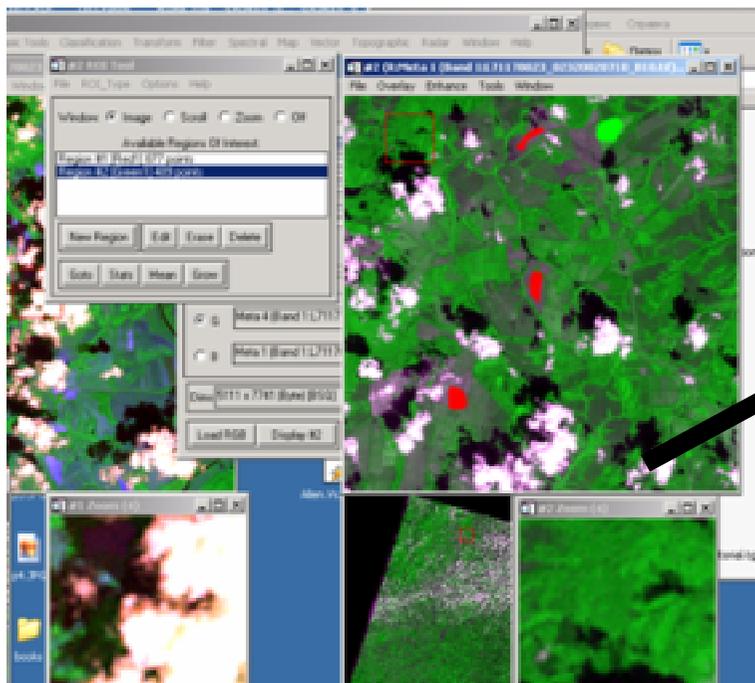




# Задача классификации

- ❑ Пользователь (эксперт) определяет характеристические множества
- ❑ Классификация по расстоянию Махаланобиса:

$$R_j = (X_i - Y_j)^T S^{-1} (X_i - Y_j)$$



Применение:

- Картографирование
- Гидрология
- Решение задач ЧС
- Экологический мониторинг
- Земельный кадастр



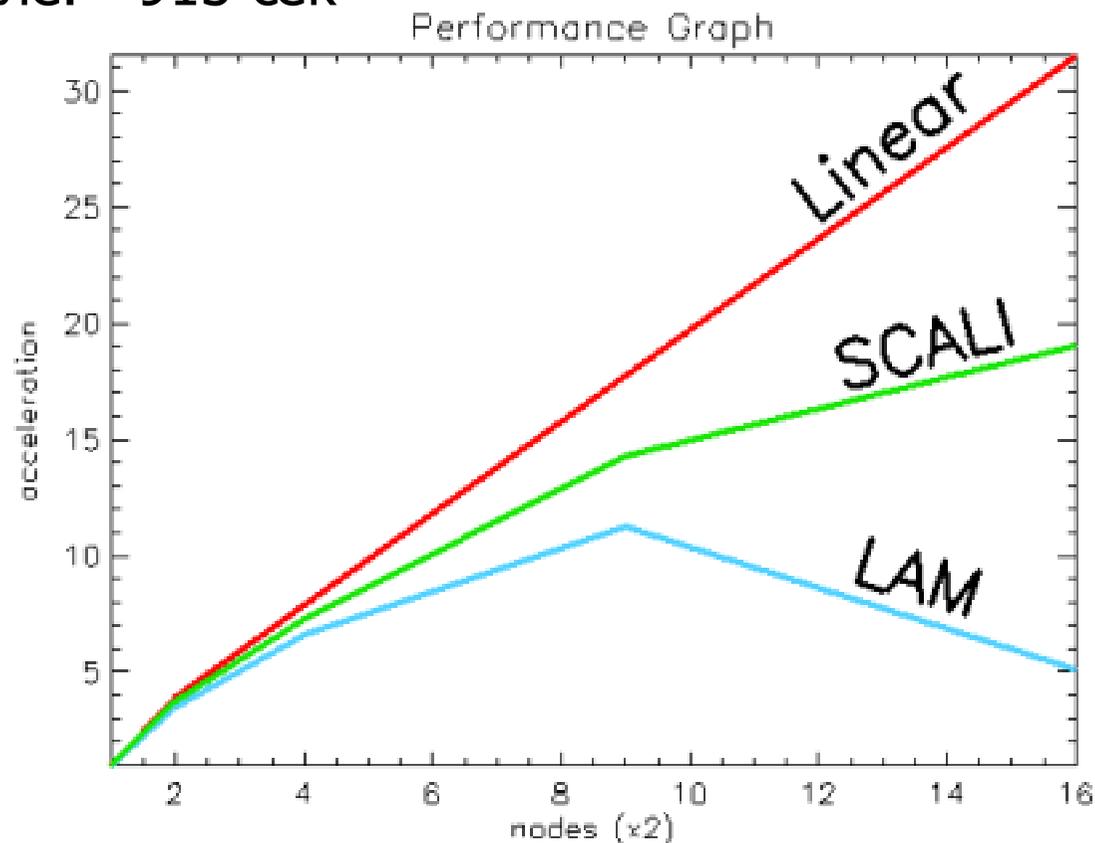
# Производительность

48 классов:

ENVI Mahalanobis Distance: 3420 сек  $\approx$  57 мин

tMahalanobis Distance на 1 узле: 913 сек

	<b>LAM</b>	<b>SCALI</b>
2	256	239
4	136	124
9	80	63
16	176	47





# Шаблоны параллельного программирования

- ❑ Цель – упростить создание параллельных программ
- ❑ Функции высшего порядка (map, reduce)
- ❑ Программирование на T-Системе без T-Системы
- ❑ Реализация – шаблонные функции C++



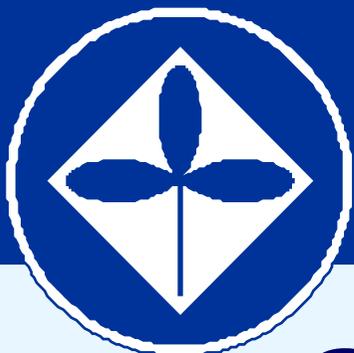
## Дальнейшее развитие

- ❑ Оптимизация для перспективных архитектур (x86-64, Cell и другие)
- ❑ Вычисления в гетерогенных кластерах и грид
- ❑ Программирование современных графических процессоров
- ❑ Новые более эффективные приложения



# Благодарности

- Постановка задачи - Стрыков Александр Иванович, ЦКН
- Разработчики Т-Системы: В.А. Роганов, А.И. Адамович, А.В. Инюхин, В.А. Водомеров, Ю.В. Шевчук, Е.В. Шевчук и другие
- Программы РАН
  - ★ ОИВТС: «Высокопроизводительные вычислительные системы с новыми принципами организации вычислительных процессов»
- РФФИ: грант 05-07-08005-офи\_a



Институт программных систем  
Российской академии наук

**Спасибо за внимание!**

Артем Первин:

[ArtemPervin@botik.ru](mailto:ArtemPervin@botik.ru)

