

Эльбрус сегодня

микропроцессоры, вычислительные комплексы, программное обеспечение

**Александр Ким, Владимир Волконский, Федор Груздов,
Юлий Сахин, Сергей Семенихин, Владимир Фельдман
ОАО «ИНЭУМ им. И.С. Брука»**

*III Московский суперкомпьютерный Форум
01.11.2012*

Пути развития суперкомпьютеров

Проект UHPC, DARPA

- Широкий спектр диапазонов производительности
 - Эксафлопные (10^{18}) суперкомпьютеры
 - Петафлопные (10^{15}) массовые компьютеры
 - Терафлопные (10^{12}) встраиваемые системы
- Энергетическая эффективность
 - Производительность МП на ватт улучшить на 2 порядка (до 50 Гфлопс/Вт)
- Программируемость
 - Облегчить разработку ПО (программируемость, совместимость)
 - Поднять надежность и безопасность программ
- Сроки
 - 2010-2014 макетные образцы
 - 2018 опытные образцы



Аппаратура и программное обеспечение должны разрабатываться совместно

Все это было заложено в архитектуру «Эльбрус» еще при создании ПРО

Влияние закона Мура

Закон Мура: удвоение количества транзисторов на кристалле каждые **1,5-2** года за последние 20 лет привел

- технологические нормы уменьшились с 1 мкм. до 22 нм. (**45** раз)
- емкость кристалла увеличилась с 1млн. до 5 млрд. транзисторов (**5000** раз)
- тактовая частота увеличилась с 25 Мгц до 5 Ггц. (**200** раз)

Развитие кремниевой технологии остановило развитие архитектуры микропроцессора, т.к. цикл разработки новой архитектуры составляет намного дольше — **8-10** лет

- реальное архитектурное ускорение процессорного ядра за 20 лет составило чуть более **2,5** раза (от RISC до Суперскаляра)

ЗАДАЧИ

1. Создание параллельной энергоэффективной архитектуры ядра для повышения **однопоточной** производительности.
2. Автоматическое распараллеливание задач **компилятором**

Особенности архитектурной линии «Эльбрус»

Технологии, разработанные в процессе создания МП «Эльбрус»

Высокопроизводительная энергоэффективная параллельная архитектура ядра микропроцессора на базе «широкого командного слова», выполняет до 23 операций за такт

- исполнение планируется оптимизирующим компилятором, число операций за такт (ЧОТ) на пакете SPECcpu2000 больше 3, планируется довести до 6 (Intel Core 2 – 1,1)

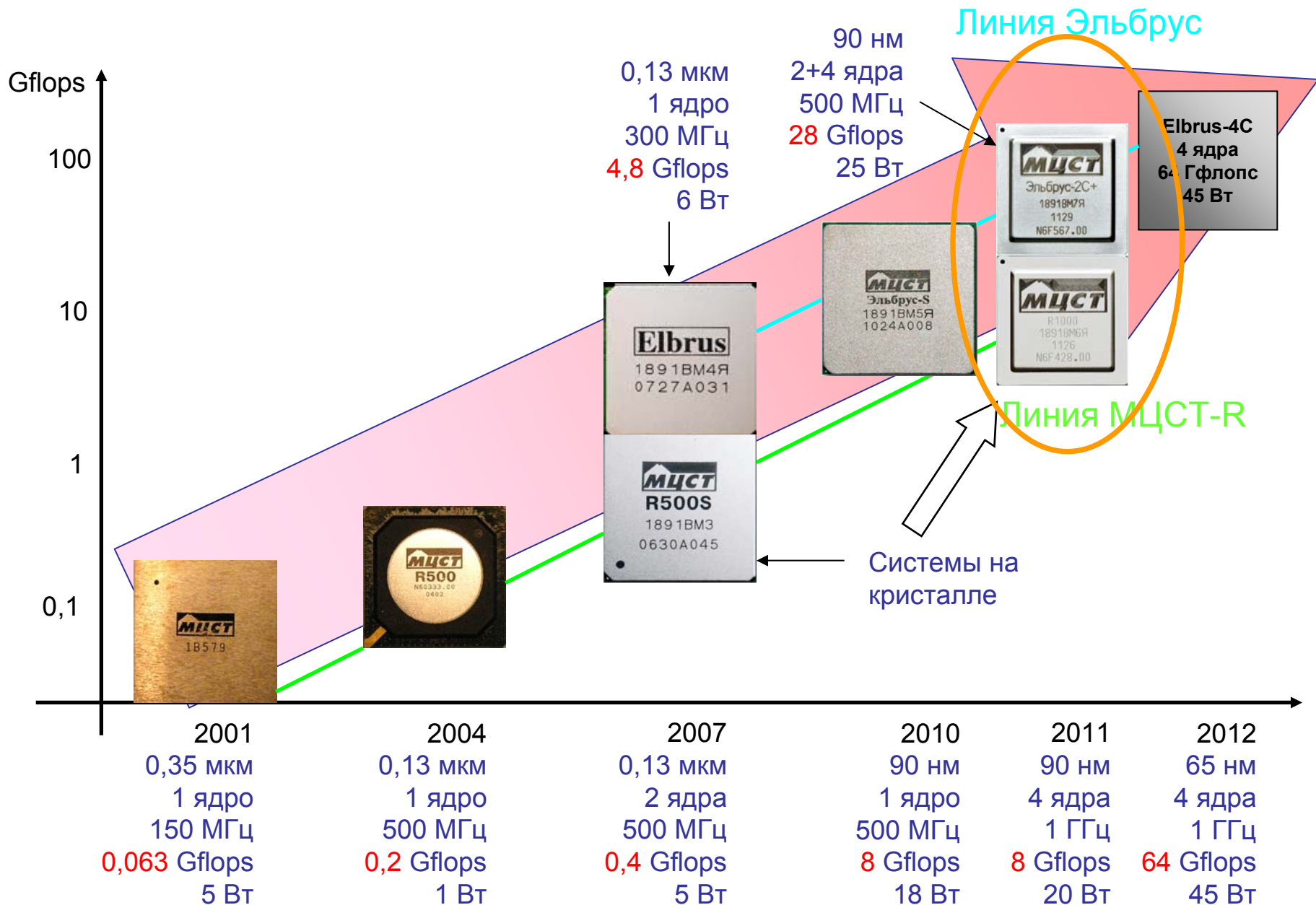
Аппаратно поддерживаемая двоичная трансляция

- обеспечивает совместимость с другими платформами на уровне исполняемых кодов

Защищенные вычисления, основанные на контекстной защите памяти на базе тегированной архитектуры

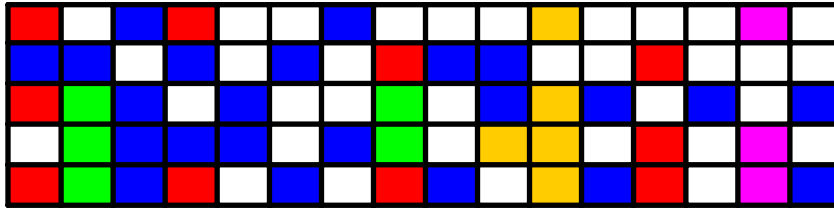
- обеспечивают стойкость к компьютерным вирусам и быструю отладку программ

Микропроцессоры



Параллелизм МП «Эльбрус»

МП «Эльбрус»



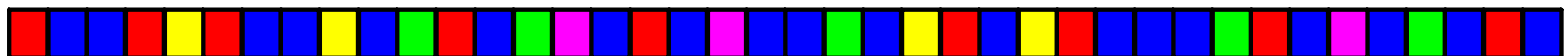
Оптимизирующий компилятор

- анализ зависимостей
- оптимизации
- глобальное планирование
- распределение регистров

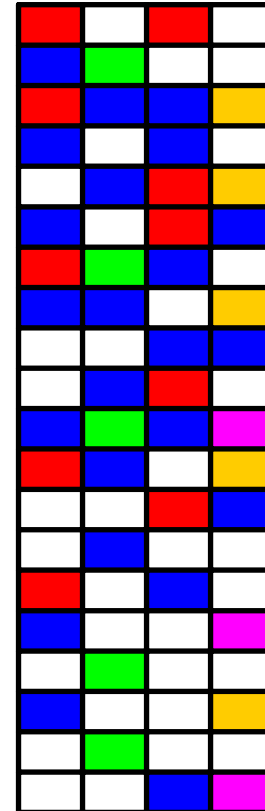
Текст программы

**Больше
параллелизма,
меньше тепла**

Последовательный поток команд после компиляции



Суперскалярный МП



Анализатор
зависимостей,
Перекодировщик
операций,
Планировщик,
Распределитель
регистров

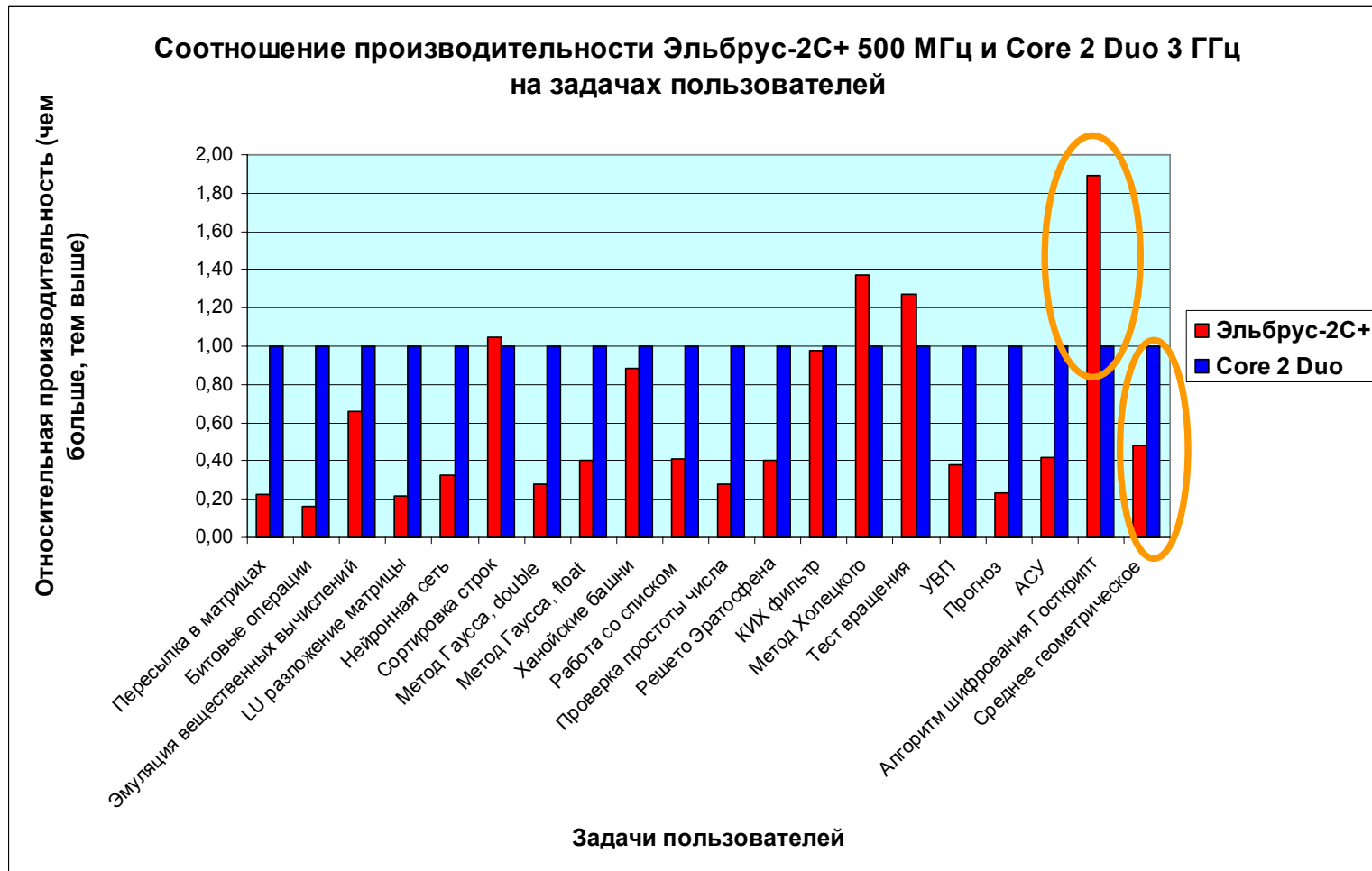
Асинхронная подкачка данных

- Основные свойства
 - Освобождает основной конвейер от операций обращения в память и вычисления адресов
 - Устраняет простои от неготовности данных
 - Не опустошает кэш
 - Увеличивает параллелизм обращений в память
 - Эффективно скрывает время доступа в память
- Применима как для вычислительных задач, так и для задач обработки данных



Значительно сокращает потери от доступа в память

Сравнение на задачах пользователей (ГИ 2011)

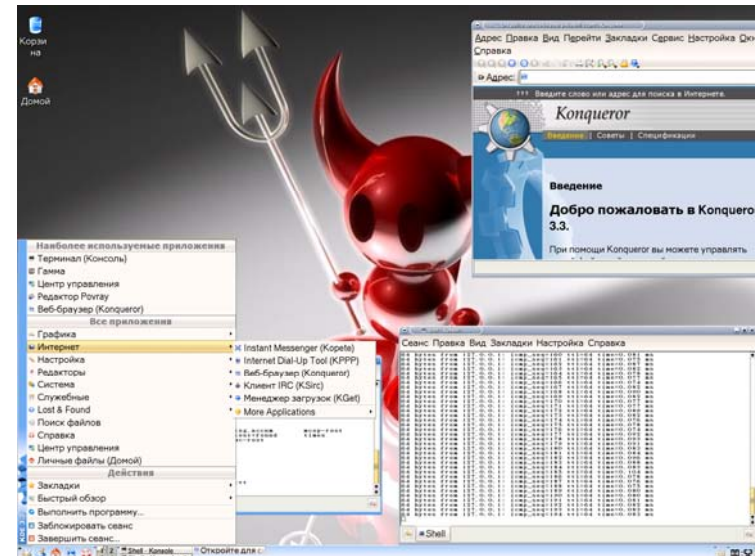
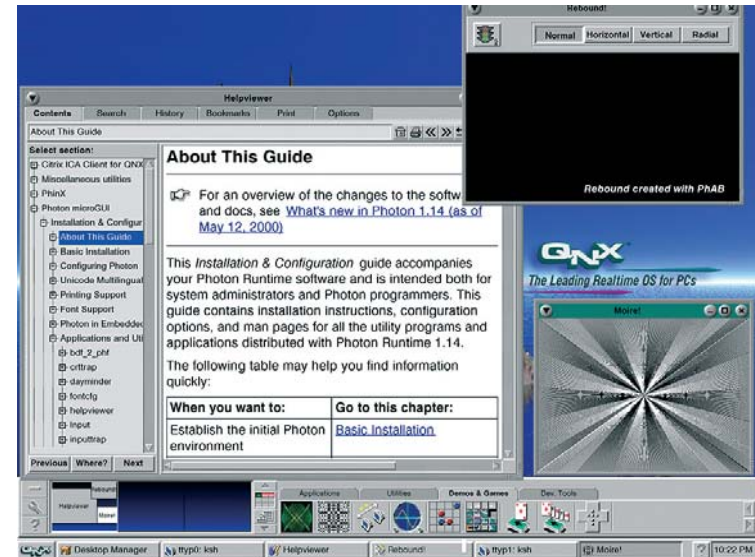


Логическая скорость Эльбруса выше, т.к. при в 6 раз меньшей тактовой частоте абсолютная производительность ниже только в 2 раза

Эффективная двоичная совместимость с Intel x86, x86-64

- Функциональность
 - Полная совместимость с архитектурой Intel x86 (x86-64 с МП Эльбрус-4С)
 - Прямое исполнение **20+** операционных систем, в том числе: MSDOS, Windows XP, Linux, QNX
 - Прямое исполнение **1000+** самых популярных приложений
 - Исполнение приложений под ОС «Эльбрус» (Linux)
- Производительность – **80%** от нативной
 - Достигается за счет скрытой **системы оптимизирующей двоичной трансляции**
 - Мощная аппаратная поддержка в МП «Эльбрус»
- Лицензионная независимость от Intel

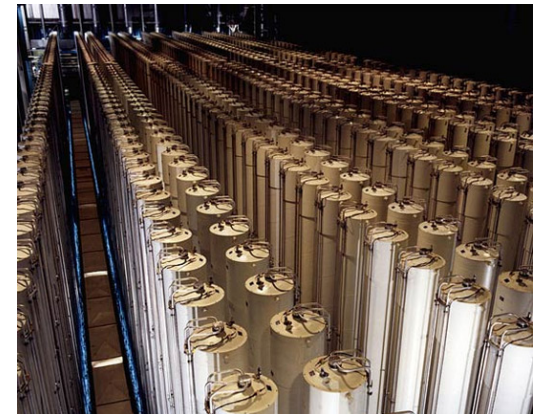
Совместимость по кодам с архитектурой, доминирующей в суперкомпьютерах, серверах, ноутбуках, ноутбуках



Обеспечение безопасности и надежности на базе МП линии «Эльбрус»

Вирусные атаки

- Проект США «Олимпийские игры»
 - Начат при пр-те Буше в 2007 (или раньше)
 - В сотрудничестве с Израилем
 - В мае 2011 пр-т Обама заявил, что киберпространство – зона интересов США
- Червь stuxnet – запущен в 2009, обнаружен в 2010 г.
 - Атака на иранский ядерный центр с целью скрытного вывода из строя центрифуг по обогащению урана
- Червь duqu – запущен раньше или одновременно со stuxnet, обнаружен в 2011
 - Атака на различные научные центры с целью получения технической информации
- Червь wiper (чистильщик) – обнаружен в 2012
 - Атака на компьютерные системы министерства нефтяной промышленности Ирана, привела к падению продаж нефти
- Червь Flame – запущен в декабре 2006, обнаружен в конце мая 2012
 - Размер 20 мегабайт, точные цели пока не известны, обнаружен после обнаружения Wiper
- Червь Gauss – обнаружен в мае 2012
 - Цель – получение любой информации с зараженных компьютеров, из семейства Flame



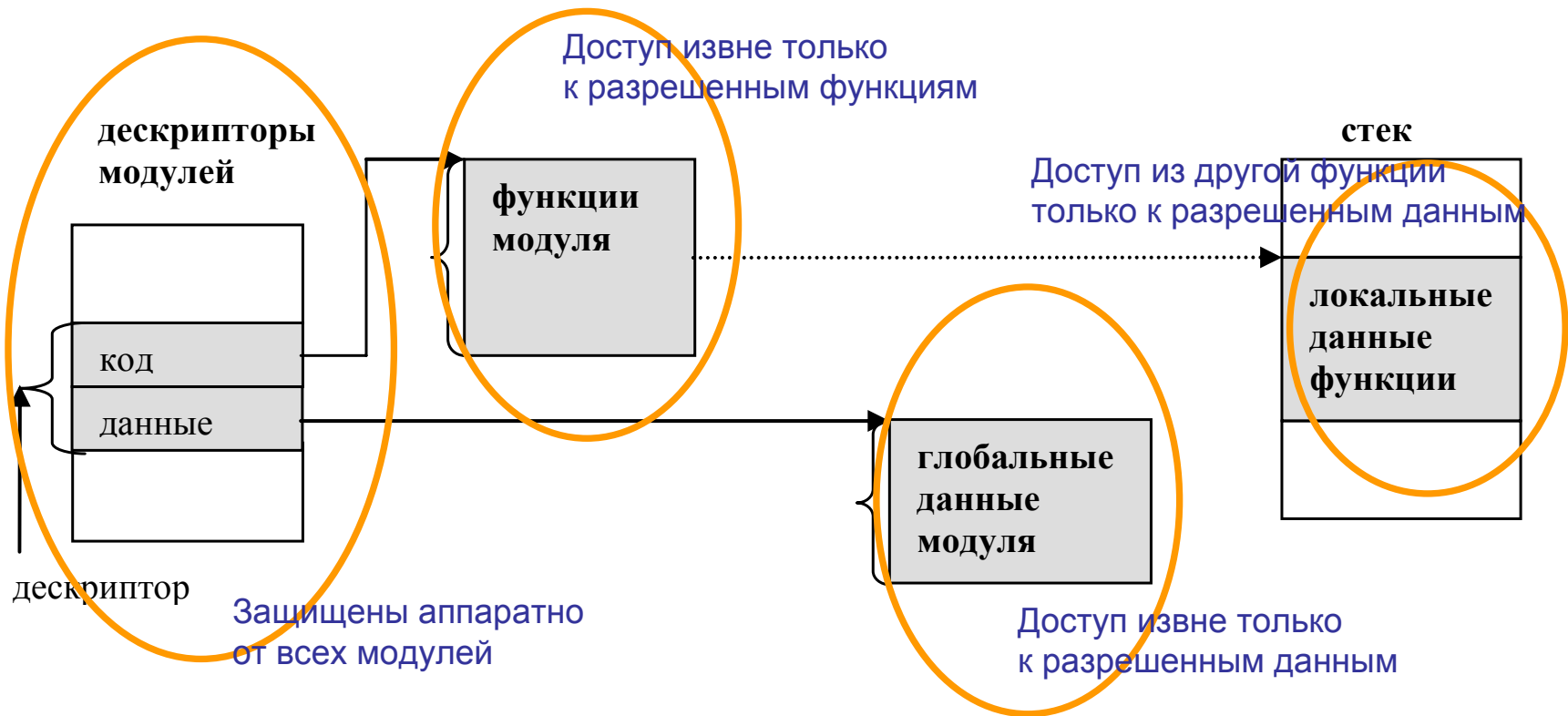
Это кибероружие, все черви тщательно маскируются, заматают следы, принадлежат одному семейству и имеют общие части, но разные цели и разные способы проникновения

Как stuxnet проникает и получает неограниченные права

- Использует 4 уязвимости «нулевого дня», т.е. ранее никому не известные
 - Заносится с флешки через ошибку в Windows
 - Повышает свой приоритет до уровня операционной системы
 - Использует опасные возможности аппаратуры и ошибки в программах
 - Распространяет себя между компьютерами через уязвимость в программе печати
- Берет под контроль управление промышленным программным обеспечением
 - Подменяет отдельные функции управления
 - Скрытно разрушает оборудование

Червь получает контроль над компьютером из-за ошибок программистов, не обнаруживаемых в коде из-за несовершенства аппаратных архитектур

Реализация межмодульной защиты в архитектуре «Эльбрус»



Каждому модулю соответствует свой дескриптор

Дескриптор модуля хранится при исполнении на аппаратных регистрах, недоступных пользователю

При вызове функции другого модуля происходит смена дескриптора модуля

Обеспечение защищенного исполнения программ

Аппаратная поддержка

- **Контроль целостности данных**
 - Ссылки не могут быть сконструированы из частей
 - Изменение части ссылки приводит к ее разрушению
- **Поддержка работы в определенном контексте**
- Атомарность смены контекста
- **Контроль границ объектов и кода**
- **Защита связующей информации в стеке**
- Контроль над появлением зависших ссылок в стек
- Чистка памяти в стеках
- Контроль прав доступа к типизированным объектам
- Операции преобразования ссылок
- Создание объектов в стеке и ссылок на них
- Создание ссылок на код функций

Программная поддержка

- Операционная система обеспечивает
 - **Создание и уничтожение объектов**
 - **Формирование тегированных ссылок на объекты**
 - Контроль и управление зависшими ссылками
 - Функции для межпроцедурных переходов
 - **Загрузку программ в память для защищенного исполнения**
 - Интерфейсы сборки и отладки
 - Уплотнение данных и чистку мусора
- Компилятор обеспечивает
 - Правильное выравнивание ссылок в памяти и на регистрах
 - Допустимый набор операций при работе со ссылками
 - Поддержку интерфейса связывания
- Средства сборки и связывания обеспечивают
 - Формирование ссылок между модулями при загрузке

Защищенное исполнение: результаты

Задачи	Всего задач	Задач с найденными ошибками
Задачи пользователей	7	4
Пакет SPECint95	8	7
Пакет негативных тестов <i>samate</i> на защищенность	888	874 (14 не выявленных не нарушают защиту)

Типы обнаруживаемых ошибок:

- нарушение границ объекта (переполнение буфера) – 53%
- использование неинициализированных данных – 37%
- использование опасных конструкций языка или опасных отклонений от стандарта языка – 10%

**Эффективный инструмент отладки
и обеспечения надежности программ**

Что дает защищенное исполнение программ

- **Фундаментальные средства** борьбы с любыми уязвимостями
- **Локализацию** сложных программных ошибок без потери эффективности
- **Эффективное исполнение** благодаря параллельному аппаратному контролю
 - Операции не замедляются
- Возможность **создания надежных и безопасных программ** большими коллективами в сжатые сроки

Модули и вычислительные средства на базе микропроцессора «Эльбрус-2С+»

МП Эльбрус-2С+ и платы на нем



- Технология – 90 нм
- Количество транзисторов - 368 млн
- Рабочая частота - 500 МГц
- Производительность - 28 Gflops
- Потребляемая мощность - 25 Вт
- Кэш память - 2,25 МБ,
- Темп обмена с оперативной памятью – 12,8 ГБ/сек
- 3 канала межпроцессорной связи – 3x4 ГБ/сек
- 2 канала Ввода/Вывода – 2x2 ГБ/сек

Плата ПК

- 1 процессор Эльбрус-2С+
- Стандарт ATX
- Интерфейсы: IDE, SATA2, PCI, PCI-Express x8, Gigabit Ethernet, USB2.0, RS232, IEEE1284, Audio
- Группа исполнения 1.1



Плата мини-ПК

- 1 процессор Эльбрус-2С+
- Стандарт mini-ITX (17см*17см)
- Интерфейсы: SATA2, Gigabit Ethernet, PCI-Express x8, USB2.0, RS232, Audio, Video (VGA+DVI)
- Группа исполнения 1.1



МП Эльбрус-2С+ и платы на нем



Модуль серверного ВК и АРМ для АСУ

Конструктив – CompactPCI 6U

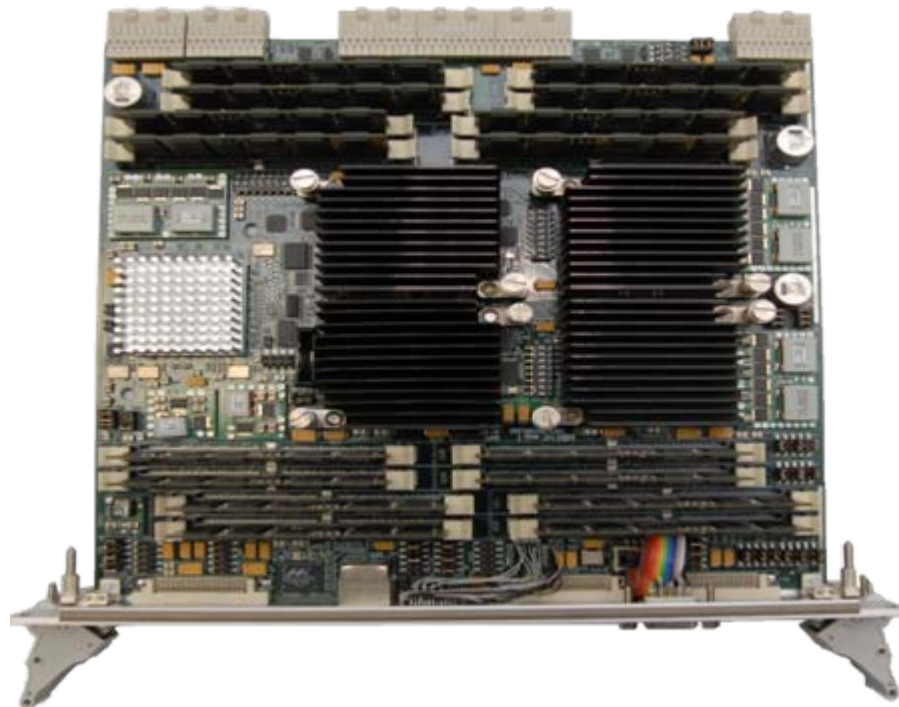
2 процессора

- Производительность 56 Гфлпс

- Оперативная память 8 Гбайт

Группа исполнения 1.1,1.2,1.3

Интерфейсы: Ethernet 3 канала, RS-232 2 канала, RS-422/485 4 канала, USB 4 канала, Audio 3 канала, IEEE 1284, Video RGB 2 канала, IOLink 2 канала, SATA 2 канала



Модуль специального ВК

- 4 процессора Эльбрус-2С+, 4 модуля в одном ВК
- Оперативная память 32 Гбайт
- Интерфейсы: SATA2, Gigabit Ethernet, USB2.0, RS232, IEEE1284, Audio, Video, RDMA
- Группа исполнения 1.1, 1.2, 1.3, 2.1.1, 2.2.1, 2.3.1

Высокопроизводительный вычислительный комплекс

- 16 процессоров Эльбрус-2С+ или МЦСТ-R1000 в одной стойке (расширенный CompactPCI 6U)
- Производительность 0.5 Тфлпс
- Группа исполнения 1.1, 1.3
- Системный модуль (CompactPCI 6U) с интерфейсами USB 2.0, Gigabit Ethernet, Serial ATA, IDE, RS-232, IEEE1284
- Видеокарта 2D/3D*
- ОС Эльбрус

Вычислительный модуль
на базе МП МЦСТ-R1000



Модуль ввода-вывода



Комплекс



Моноблок и ноутбук для широкого применения

Моноблок

- Процессор Эльбрус-2C+
- Экран 21" 1920*1080
- Сенсорный мультитач
- Видеокарта 2D/3D*
- Диски: SATA 3.5" + DVD
- Полный набор интерфейсов (USB 2.0, WiFi, Bluetooth, DVI, Gigabit Ethernet, камера, микрофон)
- Размер 535x415x55(mm)



Ноутбук

- Процессор Эльбрус-2C+
- Экран 14" 1366*768
- Видеокарта 2D/3D*
- Диски: SATA 2.5" + DVD
- Полный набор интерфейсов (USB 2.0, WiFi, Bluetooth, GPS+Glonass, DVI, Gigabit Ethernet, камера, микрофон)
- Размер 341x240x32(mm)
- Стандартный аккумулятор и блок питания



Однопроцессорный модуль



Базовое программное обеспечение

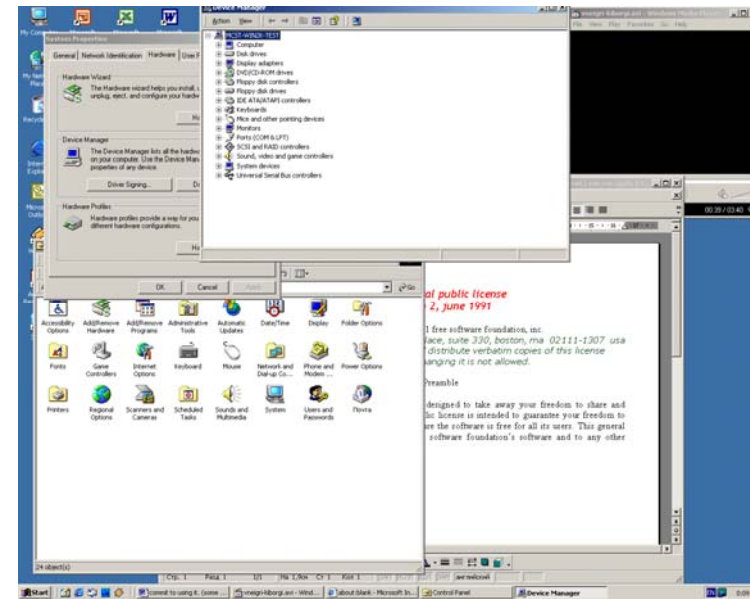
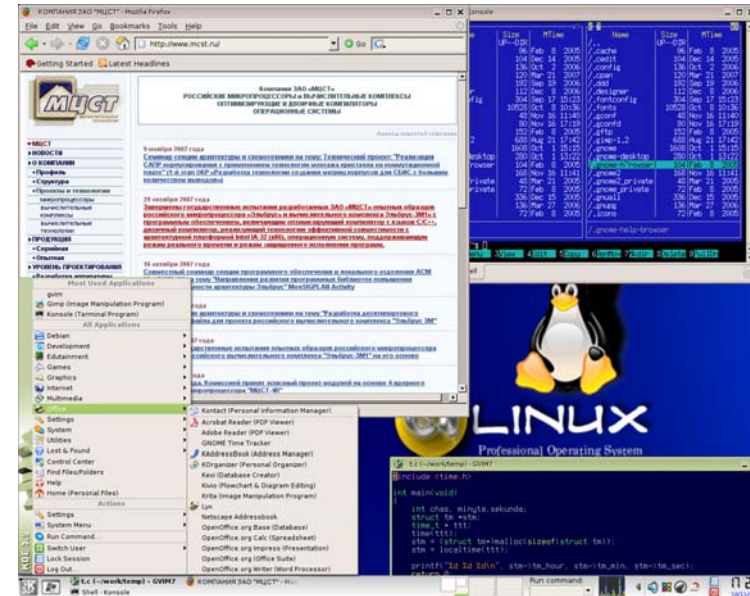
Средства программирования в системе Эльбрус

- Современные средства разработки программ
 - Оптимизирующие компиляторы с языков **C, C++, Фортран**, средства сборки, отладки, профилирования, библиотеки
 - Средства ведения и управления программным проектом
- Средства поддержки пользовательского интерфейса
 - Утилиты, сервисы, библиотеки общего назначения
 - Графическая подсистема, работа с сетью, редакторы текстов, работа с периферийными устройствами
- Возможность использования свободного ПО
 - **совместимость с Гну-компиляторами**
- Высокопроизводительные математические и мультимедийные библиотеки
 - Линейная алгебра, обработка сигналов, обработка изображений, аудио и видео кодеки, графика
- Средства и библиотеки для распараллеливания
 - Библиотека **MPI**, расширения **OpenMP**
 - Автоматическое распараллеливание



Операционная система Эльбрус

- Собственная программа начальной загрузки
- Ядро базируется на ОС Linux
- Обеспечивает стандарты POSIX
- Комплекс средств защиты от НСД **сертифицирован** по 2-му классу, реализует 2-й уровень контроля отсутствия НДВ
- Обеспечивает работу **в режиме реального времени**
 - Разброс времен сокращен до 30 раз по сравнению со стандартным изменением (rt_patch) ОС Linux
- Поддерживает
 - организацию резервирования
 - архитектуру неоднородного когерентного доступа в память NUMA
 - систему **совместимости** для приложений в **кодах Intel x86**
 - эффективное **защищенное исполнение** программ
- Система эффективной совместимости с архитектурой Intel x86



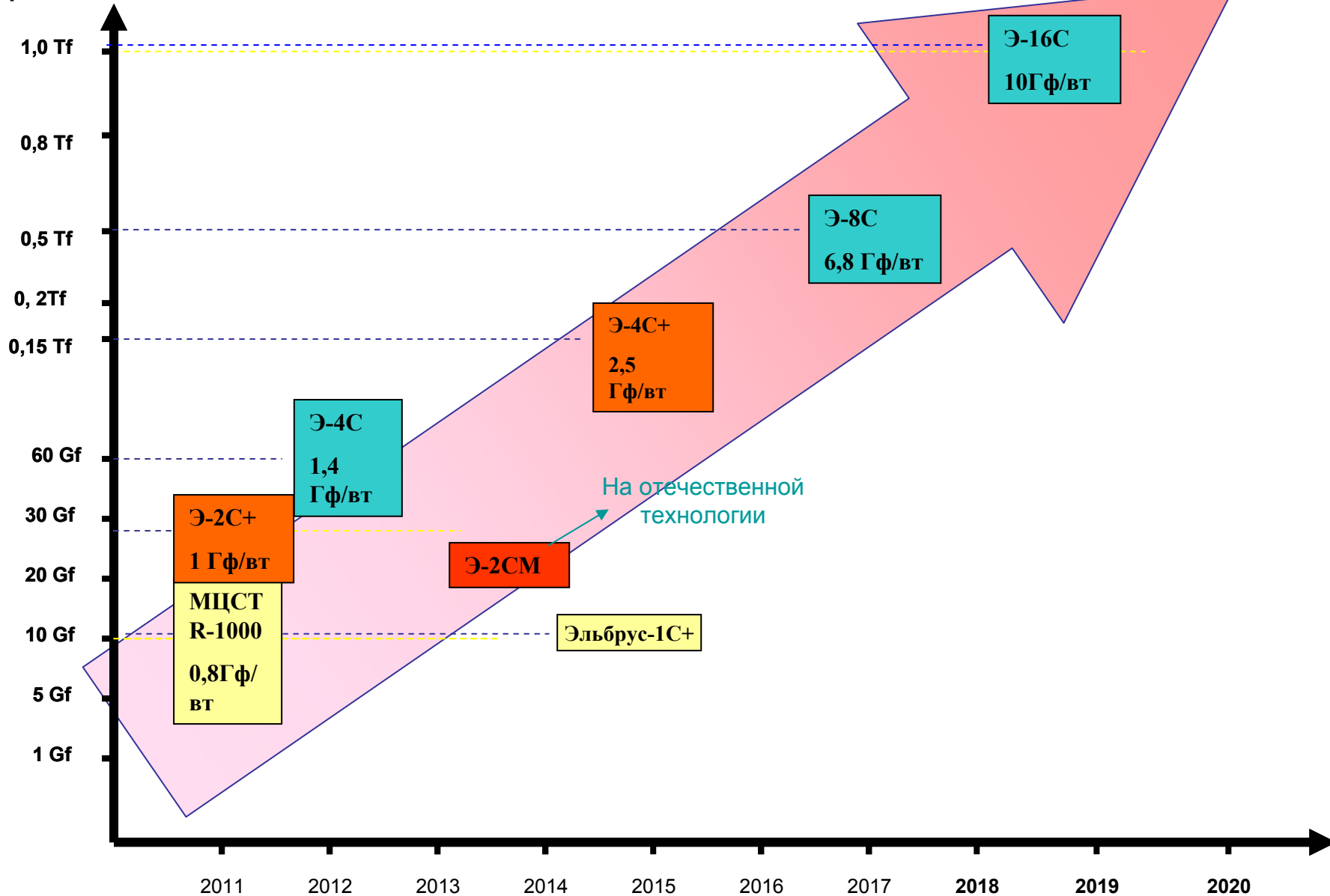
Важнейшие сервисные программы ОС «Эльбрус»

- Управление базами данных (PostgreSQL, SQLite, Линтер)
- Работа с файлами (Gnome Office, Open Office)
- Управление почтовыми сообщениями (sendmail, sylpheed)
- Средства работы с гипертекстовой информацией (Apache, Mozilla Firefox)
- Поддержка растровой и векторной графики (gimp)
- Воспроизведение аудио и видео (xmms, mplayer)
- Средства доступа к удаленному рабочему столу (vnc, rdesktop)
- Поддержка работы со сканером (SANE)

Перспективы развития линии «Эльбрус»

Перспектива развития отечественных микропроцессоров

Производительность



Направления разработки МП

- Разработка МП на технологии 28нм, 22нм, 14нм с переходом на полнозаказную схемотехнику для достижения **тактовой частоты более 3 ГГц**
- Развитие энергоэффективной архитектуры «широкого командного слова» для обеспечения производительности более 32 Гфлпс/ГГц на ядро с достижением **однопоточной** производительности более **100** Гфлпс на ядро и соотношение >10 Гфлпс/Вт, необходимого для создания экзафлопсных систем
- Реализация в одном МП до **16** ядер
- Обеспечение высокой **достоверности и надежности вычислений** введением аппаратного контроля
- Встраивание в МП **высокоскоростных каналов** межпроцессорной и межмашинной связи для создания высокопроизводительных СВТ

Направления разработки серверов и комплексов

- Многопроцессорность на общей памяти
 - до 64 микропроцессоров - 1024 ядер
- Обеспечение суммарного темпа обмена с оперативной памятью 2 Б/флопс
- Разработка конструктивно-технологических решений для создания серверов с производительностью терафлопсного диапазона и комплексов – петафлопсного диапазона
- Разработка системы контроля состояния вычислительного комплекса для поддержания высокой обслуживаемости и систем реконфигураций для поддержания высокой надежности и готовности

Направления разработки ОС

- Разработка алгоритмов распределения памяти и **управления потоками управления с привязкой их к группе процессоров** для работы с определенной физической памятью
- Использование **новых алгоритмов миграции процессов** между процессорами на основе специальной аппаратной поддержки (счетчиков обращения в «чужую» память)
- Возможность использования **копий исполняемых кодов ОС** для различных процессоров
- Возможность **перехода** от систем с общей памятью **к принципам организации принятым в многомашинных комплексах**, но с использованием общей памяти для эффективной организации обмена сообщениями
- Виртуализация

Направления разработки системы программирования

- Двоичная трансляция
 - Реализовать эффективную **многопроцессорность**
 - Обеспечить **совместимость с x86-64**
 - Реализовать **совместимость приложений для x86-64** с помощью двоичного транслятора приложений под управлением ОС «Эльбрус»
- Повысить производительность за счет оптимизирующего компилятора с языков C, C++, Фортран в **1,5-2** раза
- Высокопроизводительные **библиотеки**
 - Линейная алгебра, обработка сигналов, 3D графика и др. – **расширить номенклатуру**
- Эффективная поддержка переносимой **платформы Java**

СПАСИБО за внимание!