

# БЕЛОРУССКИЕ КЛАСТЕРЫ СЕМЕЙСТВА «СКИФ»: СОСТОЯНИЕ И ПЕРСПЕКТИВНЫЕ НАПРАВЛЕНИЯ РАЗВИТИЯ

В.В. Анищенко, А.М. Криштофик, Н.Н. Парамонов, О.П. Чиж

Приводятся практические результаты создания в Республике Беларусь моделей суперкомпьютеров семейства «СКИФ» Ряда 4.

Суперкомпьютерное направление «СКИФ» развивается в рамках программ Союзного государства (рисунок 1).

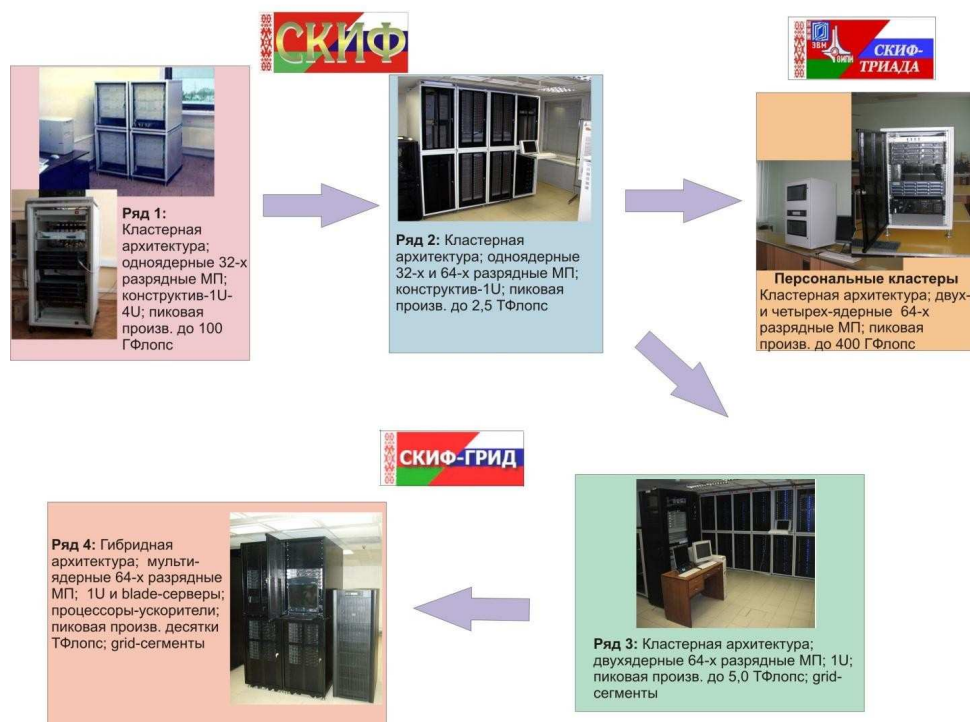


Рис. 1. Развитие суперкомпьютерного направления «СКИФ» в РБ

Важнейший практический результат выполнения программы «СКИФ» (2000-2004 гг.) – выпуск образцов кластерных конфигураций семейства «СКИФ» Ряда 1 и Ряда 2 с пиковой производительностью в диапазоне от десятков миллиардов до нескольких триллионов операций в секунду.

В рамках программы Союзного государства «СКИФ-ГРИД» (2007-2010 гг.) создана очередная генерация моделей семейства «СКИФ» – суперкомпьютерные конфигурации «СКИФ» Ряда 3 и Ряда 4. Примерами конфигураций Ряда 3 является кластер «СКИФ К-1000М» (пиковая производительность 5,0 Тфлопс), созданный в Республике Беларусь.

При создании в РБ кластерных конфигураций «СКИФ» Ряда-4 учитывались следующие концептуальные положения:

1. Диапазон производительности – терафлопсный.
2. Архитектура – гибридная метакластерная с однородными вычислительными узлами или кластерная с гибридными вычислительными узлами с сопроцессорами.
3. Вычислительные узлы – на базе классических мультиядерных процессоров с архитектурой x86-64 и/или специализированных процессоров (сoproцессоров) типа Cell, графических процессоров-ускорителей (GPU) и др.
4. Базовые конструктивно-технологические решения:
  - форм-факторы – 1U, 2U, blade-серверы;
  - стойки – 19” высотой до 44 U;
  - система охлаждения – воздушная.

Ниже приведены основные технические параметры белорусских кластеров «СКИФ ОИПИ», «СКИФ-GPU» и «СКИФ-ГРИД» семейства «СКИФ» Ряда 4, созданных с учетом изложенных выше принципов. Опытные образцы этих кластеров установлены и эксплуатируются в суперкомпьютерном центре ОИПИ НАН Беларуси.

Кластер «СКИФ ОИПИ». Опытный образец суперкомпьютера «СКИФ ОИПИ» (рисунок 2) создан в

ОИПИ НАН Беларуси в 2009 году. Кластер, являющийся «первенцем» конфигураций «СКИФ» Ряда 4, представляет собой метаcluster, состоящий из двух кластеров, - кластер на основе blade-технологий (blade-кластер) и кластер на основе Cell-технологий (Cell-кластер).



Рис. 2. Опытный образец суперкомпьютера «СКИФ ОИПИ»

В Cell-кластере суперкомпьютера «СКИФ ОИПИ» реализовано 8 двухпроцессорных ВУ на базе процессоров PowerXCell 8i 3.2Ghz. Пиковая производительность Cell-кластера:  $R_{peak} = 1638$  Гфлопс (double);  $R_{peak} = 3276$  Гфлопс (single).

Вычислительные узлы (ВУ) blade-кластера реализованы на базе двухпроцессорных серверов T-blades российской компании «Т-Платформы» с использованием 4-х ядерных процессоров Intel XEON E5472 3.0GHz. Общее количество ВУ-50, общее количество вычислительных ядер – 400. Пиковая производительность blade-кластера суперкомпьютера «СКИФ ОИПИ»: 5,0 Тфлопс.

Реальная производительность опытного образца суперкомпьютера «СКИФ ОИПИ» на тесте Linpack для 52 узлов (включая два управляющих узла) на blade-серверах составляет 4027 Гфлопс при показателе эффективности кластера  $U_{eff} = 80,67\%$ . Кластер «СКИФ ОИПИ» включен в 14 редакцию (от 29.03.2011) списка Top50 самых мощных суперкомпьютеров СНГ под номером 47.

Кластер «СКИФ-GPU». Опытный образец кластера «СКИФ-GPU» (рисунок 3) создан в 2010 году ОИПИ НАН Беларуси совместно с ОАО «НИИЭВМ». ВУ и другие модули кластера «СКИФ-GPU» располагаются в двух 19” стойках высотой 44 U собственной разработки (рисунок 3). В кластере «СКИФ-GPU» 34 вычислительных и 2 управляющих узла. Каждый ВУ кластера содержит два 4-х ядерных процессора Nehalem Intel Xeon X5570 2,93 GHz с архитектурой x86-64 и графический процессор NVIDIA GeForce GTX 295.



Рис. 3. Опытный образец кластера «СКИФ-GPU»

Пиковая производительность кластера «СКИФ-GPU» для графических процессоров при выполнении вычислений с плавающей запятой одинарной точности – не менее 60,0 Тфлопс, двойной точности – не менее 5,0 Тфлопс. Реальная производительность опытного образца суперкомпьютера «СКИФ-GPU» на тесте Linpack на 36 узлах без графических процессоров: 3025 Гфлопс при показателе эффективности кластера  $U_{eff} = 89,62\%$  от пиковой 3375.36 Гфлопс. По результатам тестирования кластер «СКИФ-GPU» был включен в 13 редакцию списка Top50 от 21.09.2010 г. под номером 45.

Кластер «СКИФ-ГРИД». Опытный образец кластера «СКИФ-GPU» (рисунок 4) создан в 2010 году ОИПИ НАН Беларуси совместно с ОАО «НИИЭВМ». В кластере «СКИФ-ГРИД» используются перспективные технические решения на базе 12-ядерных процессоров AMD Opteron с архитектурой x86-64 и графических процессоров-ускорителей. Вычислительные узлы и другие модули кластера «СКИФ-ГРИД» располагаются в одной 19” стойке высотой 44 U собственной разработки. Производительность кластера «СКИФ-ГРИД» является самой высокой в семействе белорусских моделей «СКИФ». Пиковая производительность кластера «СКИФ-ГРИД» (36 вычислительных узлов и 2 управляющих узла) без учёта ускорения с помощью графических процессоров составляет 8,0 Тфлопс при показателе эффективности на тесте Linpack 82,15%. Кластер «СКИФ-ГРИД» включен в 14 редакцию списка Top50 от 29.03.2011 под номером 40.



Рис.4. Опытный образец кластера «СКИФ-ГРИД»

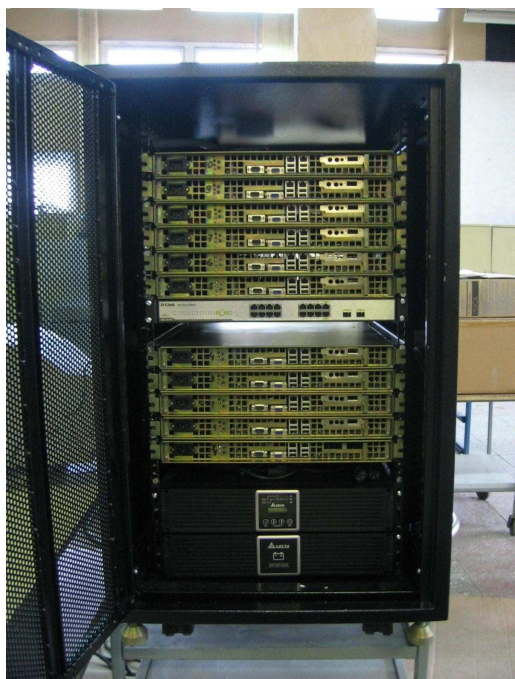
В последние несколько лет довольно широкое распространение для охлаждения серверов получили системы охлаждения (кулеры) на тепловых трубках. В 2010 году ОИПИ НАН Беларуси совместно с ОАО «НИИЭВМ» создан персональный суперкомпьютер с системой охлаждения на отечественных тепловых трубках (кластер «ПСК-СКИФ»). Тепловые трубки используются в кластере для создания более эффективных систем охлаждения, уменьшения энергопотребления и шума, создаваемого вентиляторами воздушно-принудительного охлаждения, за счет уменьшения количества вентиляторов.

Вычислительные узлы и другие модули кластера «ПСК-СКИФ» (рисунок 5 А) смонтированы в 19” стойку собственной разработки высотой 20 U. В кластере обеспечена автономная принудительная воздушная система охлаждения узлов и модулей. Отвод тепла за пределы серверов вычислительных узлов и управляющего узла кластера реализован на базе отечественных серверных тепловых трубок (находятся в серверах; рисунок 5 Б). Серверные тепловые трубки через теплосъемники подключены к отечественным тепловым трубкам стоечной системы охлаждения кластера.

Энергопотребление кластера не более 5,0 кВт. Результаты проведенных экспериментов и тестирования кластера «ПСК-СКИФ» подтвердили работоспособность системы охлаждения на отечественных тепловых трубках (на серверном и на стоечном уровнях).

Узлы кластера (10 вычислительных и один управляющий) выполнены на двухпроцессорных серверах в конструктиве 1U с использованием 6-ядерных процессоров Intel Xeon («Westmer»), RAM 24GB. Объем дисковой памяти управляющего узла - 2 ТВ. Пиковая производительность кластера «ПСК-СКИФ» (10 вычислительных узлов и 1 управляющий) – не менее 1000 Гфлопс. Приемочные испытания кластера «ПСК-СКИФ» проведены в 3 квартале 2010.





А



В

Рис. 5. А) Опытный образец кластера «ПСК-СКИФ» Б) Сервер на отечественных тепловых трубках

Опытные образцы белорусских суперкомпьютерных конфигураций «СКИФ» Ряд 4 являются существенным научно-техническим заделом для практического использования и развития в РБ суперкомпьютерных технологий. При выборе перспективных направлений развития в РБ суперкомпьютеров следующего поколения будут учтены результаты проведенного анализа современного мирового уровня суперкомпьютерных технологий и тенденций развития этого направления.

Исследования суперкомпьютерного рынка [1-3] показывают, что в списках Top500 очевидно подавляющее превосходство суперкомпьютеров с кластерной архитектурой, по-прежнему основной ОС в суперкомпьютерных системах является ОС Linux, а основными сетевыми интерфейсами – GbE и IB. Графические карты сегодня становятся настоящим мейнстримом. В перспективе ядра GPU будут интегрированы с традиционными CPU на одном чипе. Энергосбережение – важнейшая тенденция нынешнего рынка суперкомпьютеров. В последних редакциях списка Top500 появилась графа «Потребление электропитания». Очевидно, что «тера-эра» для суперкомпьютеров уже наступила и приближается «пета-эра». По прогнозам производительность в эксафлопс будет достигнута к 2018 году плюс-минус 3 года. Экзакомпьютеры будут гибридной архитектуры с использованием различных многоядерных ускорителей.

#### ЛИТЕРАТУРА:

1. Суперкомпьютерные конфигурации СКИФ /С.В. Абламейко [и др.] // Мн.: ОИПИ, 2005. – 170 с.
2. Список пятисот наиболее мощных суперкомпьютеров мира Top500 [Электронный ресурс]. – Режим доступа: <http://www.top500.org>.
3. Ежеквартальный журнал «Суперкомпьютеры». ООО «СКР-Медиа», Москва, №№ 1-3 (весна, лето, осень) 2010.