

Масштабируемые по количеству ядер и периферийных интерфейсов промышленные вычислительные системы

И.Н. Бычков^{1,2}, Ю.Х. Халиуллин²

¹Московский физико-технический институт (государственный университет)

²ОАО «Институт электронных управляющих машин им. И.С. Брука»

На данный момент на рынке высокопроизводительных серверов широко представлены многопроцессорные решения с неоднородным доступом к оперативной памяти NUMA (Non-Uniform memory access). Процессоры для реализации таких решений имеют межпроцессорные каналы повышенного быстродействия для соединения друг с другом [1,2]. Пример топологии многопроцессорной системы с межпроцессорными каналами представлен на рис. 1.

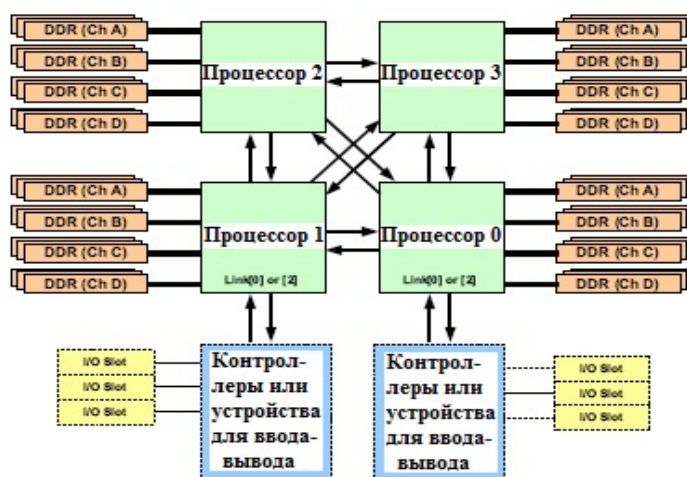


Рис. 1. Топология на основе процессоров Opteron или Эльбрус-8С.

Традиционными являются многопроцессорные решения это на основе одной двухпроцессорной или четырехпроцессорной материнской платы с установкой в корпус (chassis). В последнее время наблюдается рост популярности серверных решений на основе модулей или лезвий (blade servers), устанавливаемых в специализированные корпуса или крейты [3]. В условиях выпуска небольших серий промышленных систем повышенной надежности предлагается реализовать многопроцессорную систему с применением нескольких вычислительных лезвий, обеспечивающих масштабирование по количеству ядер и периферийных интерфейсов.

В настоящее время широко представлены вычислительные лезвия двух типов: первые с набором периферийных интерфейсов и без каналов для построения многопроцессорной системы; модули второго типа с каналами для построения многопроцессорной системы и без набора периферийных интерфейсов. Лезвия первого типа для промышленных систем предлагаются различными производителями телекоммуникационного оборудования и могут иметь в своем составе несколько процессоров [4]. Лезвия же второго типа для воспроизводительных вычислений предлагаются несколькими производителями суперкомпьютеров [5].

Особенность предлагаемой многопроцессорной системы заключается в наличии нескольких одинаковых вычислительных лезвий, которые не относятся к известным типам двум типам лезвий. Вычислительным лезвием является компьютер на плате с выводом межпроцессорных каналов. В составе лезвия имеется процессор, оперативная память, контроллер периферийных интерфейсов (южный мост), энергонезависимая память с программой начального старта (boot), жесткий диск, соединители и разъемы пользователя. Несколько таких лезвий образуют вычислительную систему на общей памяти. Этим предлагаемая многопроцессорная система отличается от распространенных систем с одним вычислительным модулем в корпусе, соответствующих стандартам VPX (VITA 46, VITA 48 выпуска 2007 - 2013 годов), Advanced®TCA (PICMG – 3.x выпуска 2008 - 2012 годов). Как и существующие подобные системы в крейте предлагаемая многопроцессорная система может содержать съемные периферийные модули, модуль управления и модули питания. Съемные модули подключаются к соединителям коммутационной панели из состава крейта.

Предлагаемые лезвия обеспечивают постоянное соотношение количества вычислительных ядер, объема когерентной памяти и пользовательских интерфейсов. Масштабируемость по количеству ядер и периферийных интерфейсов позволяет определить наиболее оптимальные конфигурации крейтовых систем для различных потребителей. При этом возможно резервирование периферийных и вычислительных модулей-лезвий. В случае резервирования лезвий предусмотрен перезапуск системы с меньшим количеством вычислительных ядер, общей когерентной памяти и устройств с подключением к периферийным интерфейсам. Такое резервирование целесообразно для применения в космосе или других агрессивных средах. Поскольку лезвия полностью идентичны, то высока

ремонтпригодность предлагаемой многопроцессорной системы, проста ее настройка и диагностика неисправности в процессе эксплуатации. Пример предлагаемой многопроцессорной системы представлен на рис.2.

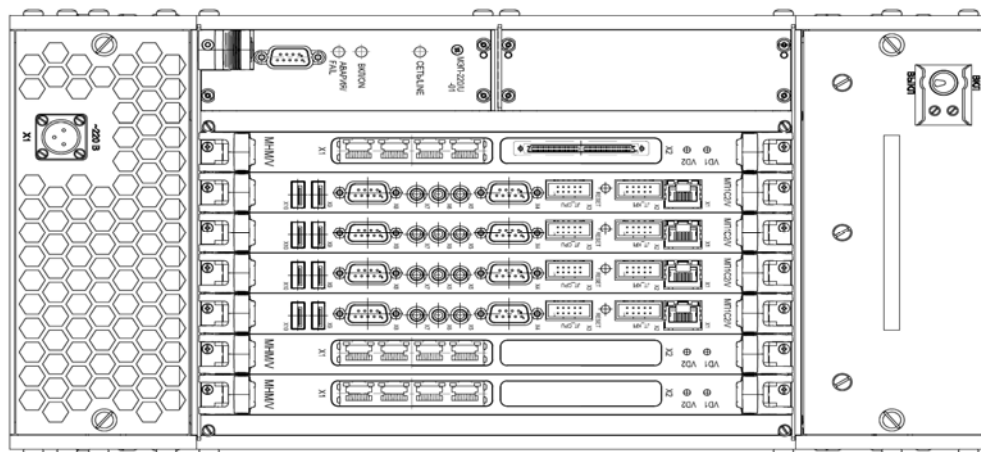


Рис. 2. Вычислительный комплекс «Сивуч-3».

Понижение себестоимости промышленной системы при условии выпуска небольшой серий ограничено. Несмотря на это основным недостатком предлагаемой многопроцессорной системы является повышенная себестоимость. Повышенная себестоимость обусловлена использованием соединителей и приемопередатчиков для передачи данных по каждому межпроцессорному каналу с пропускной способностью более 20ГБ/с, а также возможной избыточностью контроллеров периферийных интерфейсов.

Литература

1. *Tom R. Halfhill*. Sparc M7 tops 10 billion transistors. Oracle's newest 32-core server processor powers bigger iron. - Microprocessors report. September 8, 2014.
2. *Linley Gwennap*. Haswell Crams 18 Cores into Xeon. Mainstream Servers Get More CPUs, 40G Ethernet in Grantley Upgrade. - Microprocessors report. September 22, 2014.
3. *Kuba Stolarski K, Lloyd Cohen, Jed Scaramella et al.* Worldwide Server 2013–2017 Forecast Update. - IDC Forecast Update report. February, 2014.
4. CPCI & VPX Blades & Boards / URL: http://www.advantech.com/products/Master-CPU-Boards/sub_1-2JKLMR.aspx (дата обращения 05.10.2014).
5. Cray® XC™ Series: Adaptive Supercomputing. Aries Interconnect and Dragonfly Topology. / URL: <http://www.cray.com/Products/Computing/XC/Technology.aspx> (дата обращения 05.10.2014).