

## Лабораторная работа 1. Исследование пропускной способности каналов связи вычислительных систем с иерархической организацией коммуникационной сети

Имеется распределенная вычислительная система (ВС), укомплектованная  $N$  одинаковыми элементарными машинами (ЭМ). Наиболее распространенные типы ЭМ (NUMA и SMP) современных высокопроизводительных ВС представлены на рисунке 1.

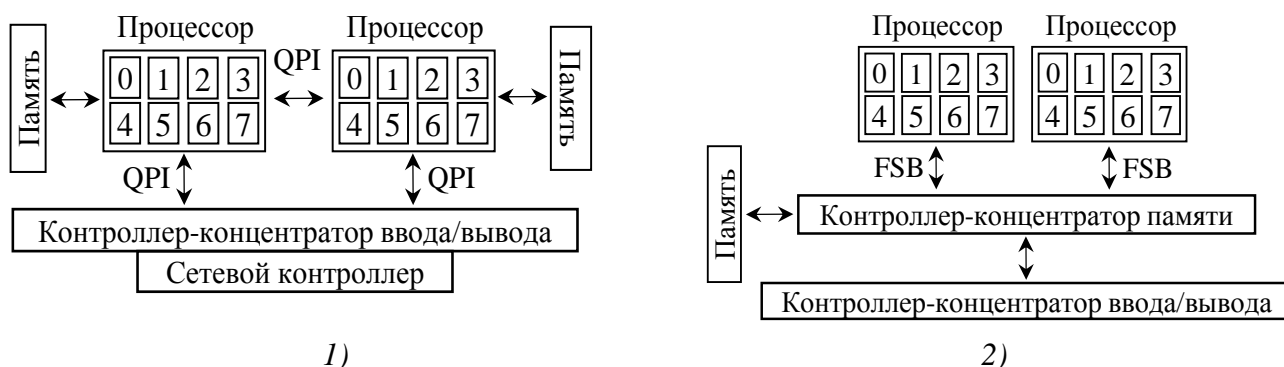


Рисунок 1 – Типы элементарных машин современных ВС:

- 1) NUMA-узел: два 8 ядерных процессора;
- 2) SMP-узел: два 8 ядерных процессора.

На рисунке 1.1 изображена техническая реализация ЭМ NUMA-узла, включающих два 8 ядерных процессора Intel, объединенных шиной Intel QuickPath Interconnect (QPI). На рисунке 1.2 представлен пример технической реализации ЭМ SMP-узла, состоящей из двух 8 ядерных процессоров Intel, объединенных системной шиной Intel Front Side Bus (FSB). Можно выделить три уровня коммуникационной среды, на которых возникает одновременное совместное использование каналов связи: контроллер памяти, внутрисистемная шина, объединяющая процессоры в ЭМ с архитектурой NUMA, и сетевой контроллер.

В рамках лабораторной работы №1 требуется выполнить следующие задания:

**1. Разработать программный комплекс для измерения времени выполнения операции обмена на различных уровнях коммуникационной среды.**

Для определения времени выполнения информационных обменов необходимо разработать тестовую MPI-программу, реализующую вызов каждым процессом операции `MPI_Irecv` и `MPI_Isend`.

Время выполнения операции обмена оценивается путем измерения среднего времени  $n$  выполнений операций `MPI_Irecv` и `MPI_Isend` в цикле. На каждой итерации цикла реализуется ожидание завершения обменов всех ветвей параллельной программы. За время  $t$  выполнения информационного обмена принимается среднее время одного запуска.

## **2. Провести эксперименты на вычислительных кластерах с NUMA/SMP-узлами.**

Кластер с NUMA-узлами укомплектован 6 вычислительными серверами на базе платформы Intel S5520UR. На каждом узле размещено два процессора Intel Xeon E5620, оперативная память – 24 Гбайт (DDR3), сетевой адаптер InfiniBand QDR Mellanox MT26428. Кластер с SMP-узлами укомплектован 18 вычислительными серверами на базе платформы Intel SR2520SAF. На каждом узле размещено два процессора Intel Xeon E5420, оперативная память – 8 Гбайт (DDR2), сетевой адаптер Gigabit Ethernet Intel PRO/1000 EB. Вычислительные кластера функционируют под управлением операционной системы (ОС) GNU/Linux. При компиляции MPI-программы использовать коммуникационные библиотеки MPICH 3.2.1 и MVAPICH2 2.2.

Тестовая программа запускать с разным количеством процессов для передачи информационных сообщений размером  $m$  Мбайт. Каждый процесс привязывать к выделенному процессорному ядру при помощи подсистемы numactl.

## **3. Построить графики зависимости $t(m)$ времени передачи сообщения размером $m$ байт.**

В экспериментах рассмотреть три уровня коммуникационной среды: оперативную память NUMA/SMP узлов, внутрисистемную шину Intel QPI, объединяющую процессоры NUMA-узлов и сеть связи между ЭМ (адаптеры InfiniBand QDR и Gigabit Ethernet).

### **Литература**

1. Хорошевский В. Г. Архитектура вычислительных систем. – М.: МГТУ им. Н.Э. Баумана, 2008. – 520 с.
2. Евреинов Э.В., Хорошевский В.Г. Однородные вычислительные системы. – Новосибирск: Наука, СО РАН, 1978. – 319 с.