

**0.1. Перышкова Е.Н. Структурно-ориентированный алгоритм формирования подсистем выполнения MPI-программ**

Одним из важнейших архитектурных свойств современных вычислительных систем (ВС) с распределенной памятью является глубокая иерархия средств доступа к оперативной памяти процессорных ядер [1]. Коммуникационные сети большинства ВС списка Top500 имеют как минимум двухуровневую организацию: коммуникационная сеть связи между элементарными машинами (ЭМ): Cray Gemini, IBM PERCS, Fujitsu Tofu, Gigabit Ethernet, InfiniBand; оперативная память, разделяемая процессорными ядрами одной ЭМ.

Если принять во внимание использование коммуникационных сетей на базе составных коммутаторов (например, топология fat tree), а также наличие внутрисистемных шин для объединения процессоров в ЭМ с архитектурой NUMA, то количество уровней в иерархической структуре увеличивается. Конфигурация подсистемы ЭМ и начальное распределение процессов по процессорным ядрам системы в значительной степени определяет время выполнения MPI-программ [2]. В системах управления ресурсами ВС при формировании подсистем ЭМ возникает задача формирования подсистемы из  $P$  процессорных ядер. В ВС на базе многопроцессорных узлов данная задача имеет множество решений [3]. Поэтому практический интерес представляет разработка структурно-ориентированных алгоритмов формирования подсистем ЭМ, учитывающих структуру информационных обменов целевой программы.

Разработан структурно-ориентированный алгоритм формирования подсистем ЭМ, учитывающий падение производительности каналов связи при одновременном использовании канала связи множеством процессов. Эффективность алгоритма исследована на вычислительном кластере с SMP/NUMA-архитектурой вычислительных узлов при выполнении параллельных MPI-программ, использующих шаблон информационных обменов типа All-to-All.

Создана система прогнозирования времени выполнения коллективных операций в стандарте MPI на заданной подсистеме ЭМ по результатам предварительной экспериментальной оценки падения производительности операций MPI-Send/MPI-Recv при одновременном использовании канала связи множеством процессов. Выработаны рекомендации выбора оптимального размещения процессоров параллельной программы с учетом размеров сообщений и разделения каналов связи. Разработанные в данной работе алгоритмические и программные средства расширяют функциональные возможности инструментария параллельного мультипрограммирования.

*Исследования выполнены при поддержке Российского фонда фундаментальных исследований (грант № 20-07-00039).*

**Список литературы**

- [1] Хорошевский В. Г. Распределенные вычислительные системы с программируемой структурой. // Вестник СибГУТИ. 2010. № 2. С. 3–41.
- [2] BALAJI P., GROPP W. MPI on Millions of Cores. // Parallel Processing Letters. 2011. Vol. 21. I. 1. P. 45–60.
- [3] СТЕПАНЕНКО С. А. Мультипроцессорные среды суперЭВМ. Масштабирование эффективности. / М.: Физматлит, 2016. 312 с.