

# ПРИМЕНЕНИЕ ПАКЕТА BOOKSIM ДЛЯ МОДЕЛИРОВАНИЯ ТОПОЛОГИЙ ВЫЧИСЛИТЕЛЬНЫХ СЕТЕЙ

*В. Г. Басалов, Д. О. Козлов*

ФГУП «РФЯЦ-ВНИИЭФ», ИТМФ г. Саров Нижегородской обл.

Создание современных высокопроизводительных коммуникационных сетей (КС) ставит перед разработчиками много сложных технических проблем: производительность, расширяемость, надежность, стоимость и потребляемая мощность.

Эти проблемы имеют сильную взаимосвязь, и поэтому решение отдельных проблем не приведет к достижению приемлемого результата.

Топология коммуникационной сети в значительной степени определяет ее производительность и эффективность и, как следствие, всего мультипроцессорного вычислительного комплекса. Существует множество различных топологий, при этом возникает вопрос, какую топологию выбрать?

Практическое сравнение топологий является трудоемким и затратным процессом, поэтому появилась необходимость в таком программном средстве, которое могло бы моделировать работу различных сетевых топологий. Таким программным средством является пакет BOOKSIM [1].

Пакет BOOKSIM разработан в Стэнфордском технологическом университете и представляет собой консольное приложение для моделирования трафика вычислительных сетей. На основе входных параметров Booksim выводит информацию о средней коммуникационной задержке доставки пакета в сети и о среднем количестве транзитных участков. К входным параметрам моделирования относятся: параметры топологии, функция трафика, задающая распределение для генерации номера источника и приемника для пакетов, функция маршрутизации и интенсивность внедрения информационного потока. Средняя коммуникационная задержка измеряется циклами (итерациями) работы программного пакета. На одном цикле происходит: создание пакетов, внедрение их в сеть, продвижение пакетов через выходные каналы коммутаторов, и если пакеты были доставлены до адресата, то происходит их извлечение из сети. Коммуникационной задержкой одного пакета считается количество циклов работы программы, которое прибывал данный пакет в сети, с момента его создания по момент его извлечения. Средняя коммуникационная задержка получается путем вычисления среднего арифметического задержек всех сгенерированных пакетов. По аналогии рассчитывается величина среднего количества транзитных участков

BOOKSIM позволяет моделировать топологии Torus, Mesh, DragonFly, Fat Tree и др. В программном пакете предусмотрено добавление новых возможностей без изменения всего кода программного

пакета. Booksim может применяться для моделирования топологий, разработанных на базе архитектуры InfiniBand и архитектуры отечественной системы межпроцессорных обменов СМПО (разработанной в РФЯЦ-ВНИИЭФ).

С применением данного программного пакета проведено моделирование топологий коммуникационных сетей, реализованных на базе архитектуры InfiniBand: 3D Torus, 3D Mesh, 4D Mesh и 5D Mesh. В ходе каждого эксперимента изменялась интенсивность внедрения пакетов от 5 % до 75 %. Интенсивность внедрения пакетов, это параметр, который влияет на количество пакетов, которое будет создаваться на одном цикле моделирования. Если данная величина равна 50 %, то это обозначает, что на одном цикле моделирования пакетов будет создано ровно половина от количества вычислительных модулей в МВК (мультипроцессорном вычислительном комплексе). Неизменными оставались параметры масштабирования топологии, функции маршрутизации и трафика. Параметры масштабирования подобраны таким образом, что бы сети с данными топологиями объединяли примерно 256 ВМ (вычислительных модулей). Функция маршрутизации для топологий Mesh и Torus представлена алгоритмом Dimension Order Routing [2]. Функция трафика представлена равномерным распределением для каждого генерируемого пакета.

На рис. 1 показаны результаты моделирования средней коммуникационной задержки для топологий 3D Torus, 3D Mesh, 4D Mesh и 5D Mesh в зависимости от загрузки КС МВК.

Исходя из данных представленных на рис. 1 видно, что наихудшую среднюю коммуникационную задержку имеет топология 3D Mesh. А при загрузке сети выше 40 % BOOKSIM вывел сообщение о том, что средняя коммуникационная задержка для этой топологии слишком большая. 3D-тор выигрывает у 3D Mesh, но проигрывает 4D Mesh, а наилучшую коммуникационную задержку показала топология 5D Mesh.

Иногда для удобного размещения МВК в машинном зале применяется топология Torus, в которой количество ВМ в одном измерении гораздо больше, чем в других измерениях, к примеру, размерность тора  $16 \times 2 \times 2$ . На рис. 2 показаны результаты моделирования средней коммуникационной задержки для топологий 3D Torus с разной размерностью и одинаковым количеством вычислительных модулей в зависимости от загрузки КС МВК.

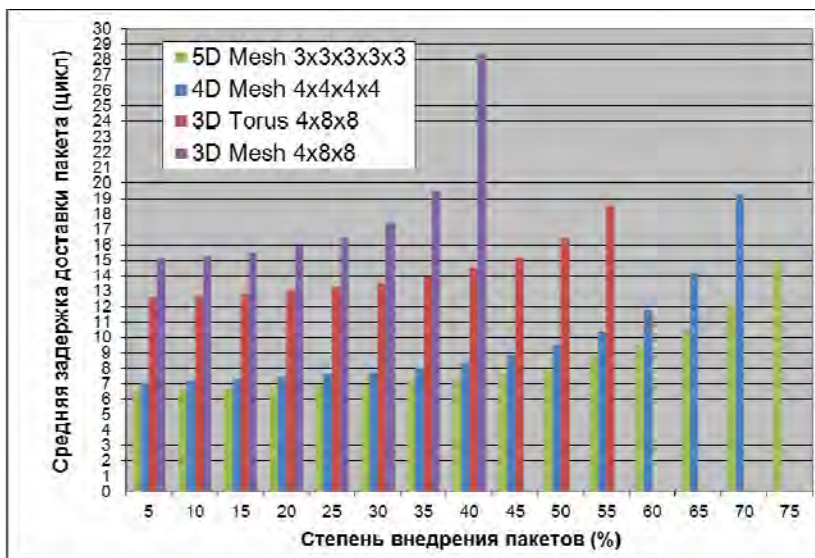


Рис. 1. Гистограмма зависимости средней коммуникационной задержки от степени внедрения пакетов для топологий 3D Torus, 3D Mesh, 4D Mesh и 5D Mesh

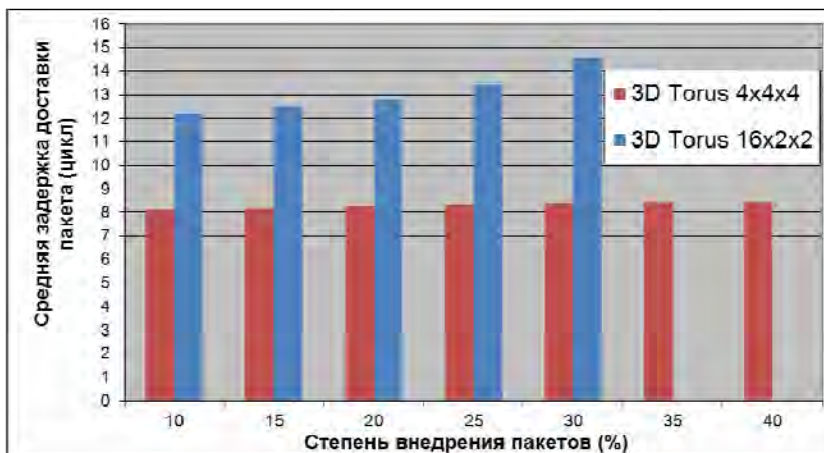


Рис. 2. Гистограмма зависимости средней коммуникационной задержки от степени внедрения пакетов для топологий 3D Torus  $4 \times 4 \times 4$  и 3D Torus  $16 \times 2 \times 2$

Исходя из данных приведенных на рис. 2, можно сделать вывод, что данный подход организации расположения ВМ МВК, когда в одном измерении тора больше вычислительных узлов, чем в других, не эффективен, ввиду того что при высокоинтенсивном трафике в сети получаются очень большие коммуникационные задержки.

Оптимальное расположение ВМ МВК для топологии Torus является, когда в каждом измерении тора находится одинаковое количество ВМ, тогда система работает стабильнее и выдерживает более интенсивный трафик.

Еще BOOKSIM применяется для моделирования топологий коммуникационных сетей: 4D-тор (используемой при создании высокоскоростной сети «Ангара», разработанной в ОАО «НИЦЭВТ» [3]), МультиТор и KNS (используемых при создании высокоскоростной сети на базе архитектуры СМПО). Для поддержки топологий МультиТор и KNS в программный пакет BOOKSIM были реализованы и добавлены

новые модули топологий и алгоритмы маршрутизации к ним.

На рис. 3 приведены результаты моделирования этих топологий размером до 1300 вычислительных модулей, с одинаковым трафиком и с загруженностью сети 50 %.

До 500 узлов лидирует топология MultiTor. После 500 и до 1000 узлов лидирует топология KNS.

Особый интерес вызывает моделирование КС МВК с количеством узлов больше 1000. На рис. 4 приведены результаты моделирования для выше перечисленных топологий с количеством вычислительных модулей от 1300 до 10000, с одинаковым трафиком и с загруженностью сети 1 %.

Хорошо заметно, из данных приведенных на рис. 4, что на этом диапазоне по средней коммуникационной задержке безоговорочно лидирует топология KNS. При этом топология MultiTor начинает проигрывать топологии 4D-тор.

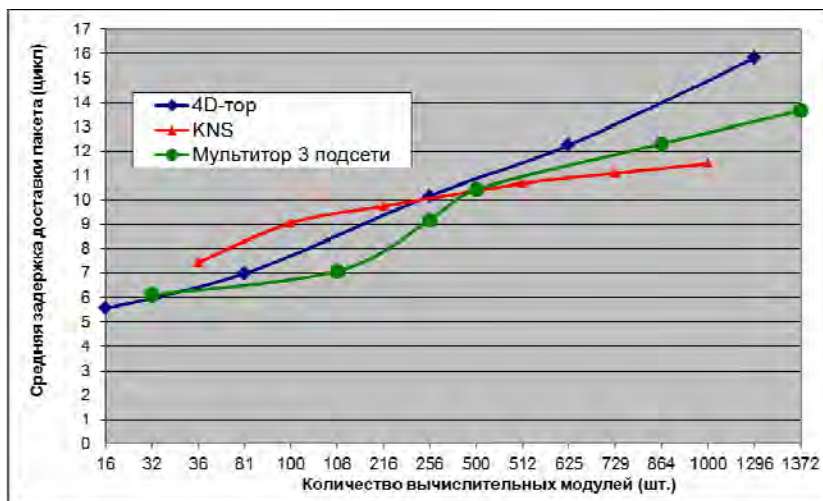


Рис. 3. Графики зависимости средней коммуникационной задержки от количества VM для топологий 4D-тор, МультиТор и KNS

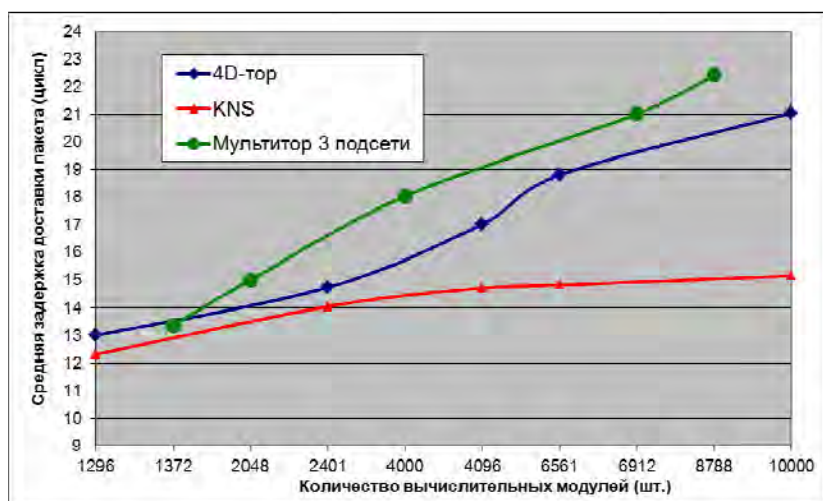


Рис. 4. Графики зависимости средней коммуникационной задержки от количества VM для топологий 4D-тор, МультиТор и KNS

## Литература

1. Главная страница разработчиков программного обеспечения стенфордского технологического университета, программный пакет Booksim [Электронный ресурс]. – Режим доступа: <https://nocs.stanford.edu/cgi-bin/trac.cgi/wiki/Resources/BookSim>.

2. William James Dally, Brian Towles, Principles and Practices of Interconnection Networks, chapter 8.4.2 Dimension-Order Routing in Cube Networks – Morgan Kaufmann, 2004.

3. Жабин И., Макагон Д., Симонов А., Сыромятников Е., Фролов А., Щербак А. Кристалл для «Ангара» // Суперкомпьютеры № 4 (16) 2013. С. 46–49..