

Суперкомпьютерные дни в России 2017



**Построение коммуникационного профиля
параллельных программ путем анализа
сохраненной копии трафика сети суперкомпьютера**

Градсков Алексей Алексеевич, ВМК МГУ

Стефанов Константин Сергеевич, НИВЦ МГУ

25 сентября 2017

Задача построения коммуникационного профиля

Используемый способ:

Инструментирование (PMPI)

Запуск модифицированной программы

Анализ собранных трасс

Предлагаемый способ:

Нет вмешательства в исходный код приложения

Взаимодействие с коммуникационной сетью

Сбор и обработка сетевых данных

Постановка задачи

Разработать инструмент, позволяющий по копии трафика сети Infiniband, снятой во время выполнения приложения на вычислительной системе, построить коммуникационный профиль данного приложения в виде коммуникационной матрицы.

Алгоритм построения коммуникационной матрицы

- 1. Конфигурация коммуникационной сети для сбора данных.**
- 2. Запуск исследуемого приложения и сбор пакетов Infiniband.**
- 3. Извлечение из пакетов релевантной информации.**
- 4. Построение матрицы.**

Инициализация сбора данных

Выделение узлов slurm

```
sbatch script.sh
```

Получение информации о узлах (srun)

hostname – для формирования списка узлов, где будет запущена задача

LID (ibstat) – для получения списка идентификаторов InfiniBand выделенных узлов

Получение информации о подсети

Определение списка коммутаторов и их портов, к которым подключены выделенные узлы

Сбор данных с сети при выполнении приложения

Распределение узлов

Узлы сброса

Вычислительные узлы (hostfile)

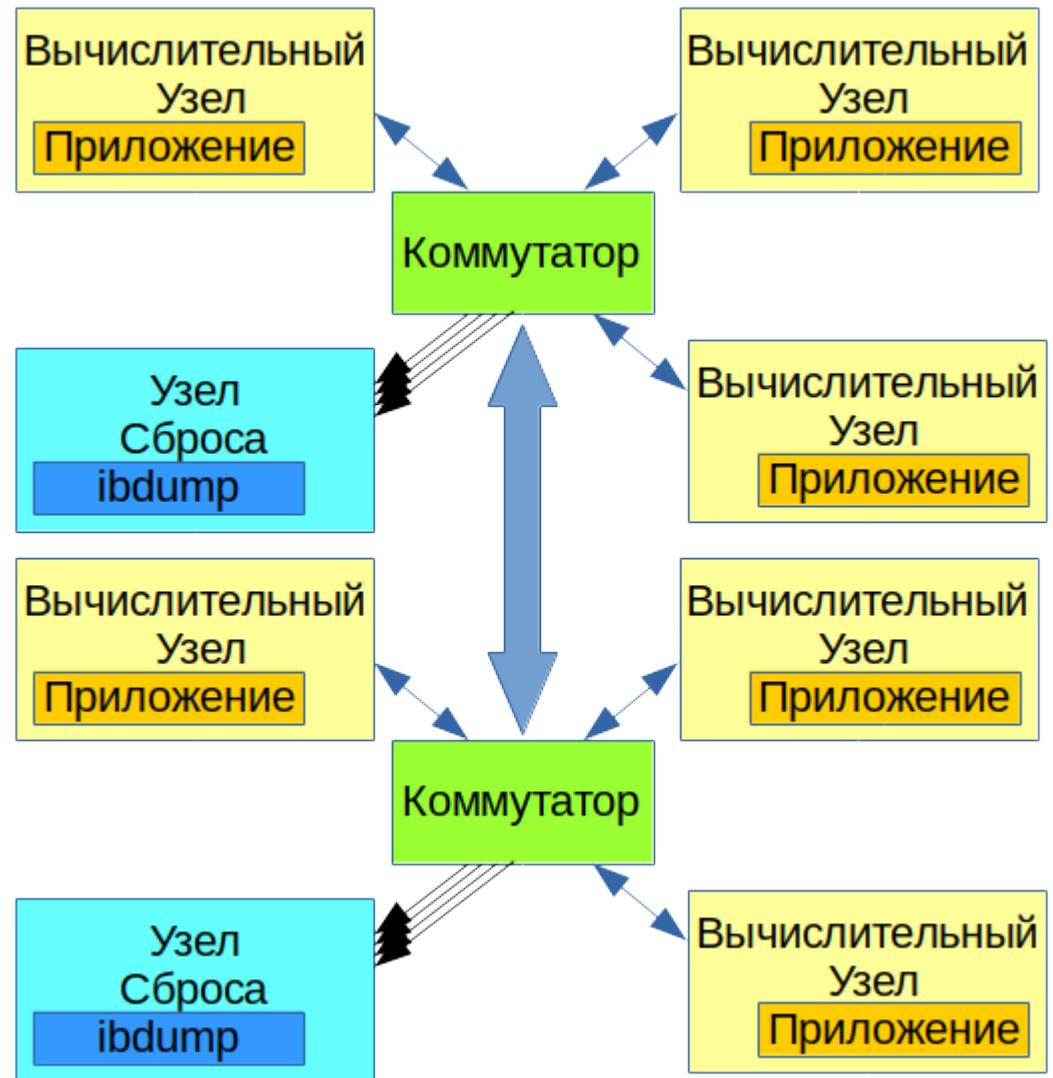
Настройка для сбора трафика

Дублирование трафика (ibmirror)

Запуск ibdump на узлах сброса (ssh)

Запуск приложения

```
mpirun -hostfile hostfile ./a.out
```



Завершение сбора данных

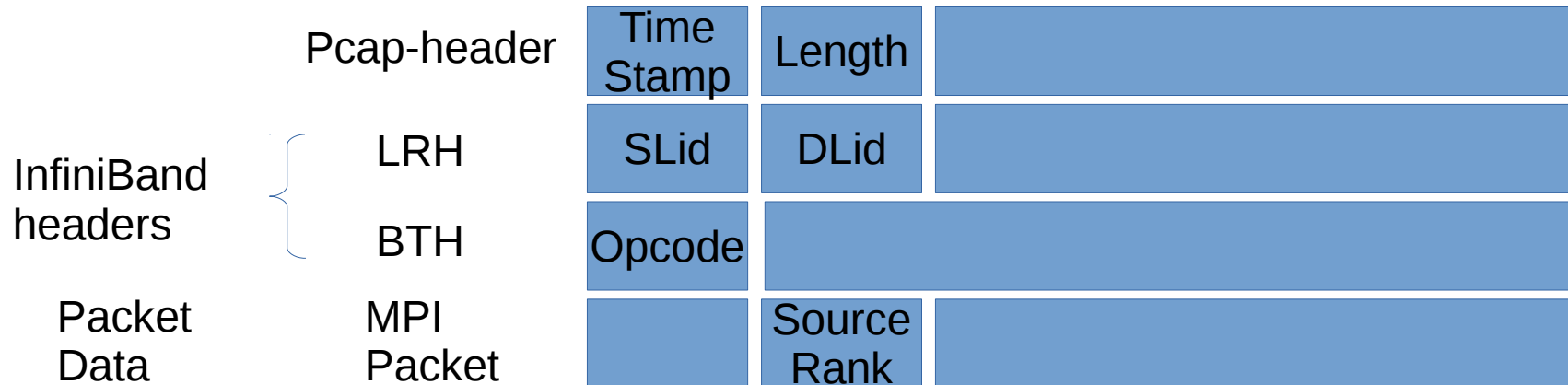
Остановка приложения и очистка

Остановка ibdump

Сброс конфигурации коммутаторов

Сбор и слияние трасс (mergescap)

Получение информации из копии трафика



Timestamp	Slid[rank] -> Dlid	Size
1476371973.641989	4 [3] -> 8	2074 bytes
1476371973.642835	8 [4] -> 4	2074 bytes
1476371973.642895	8 [4] -> 4	2074 bytes
1476371973.644601	8 [0] -> 4	2074 bytes

Построение коммуникационной матрицы

Фильтрация по вычислительным узлам

LID отправителя и получателя

Группировка пакетов

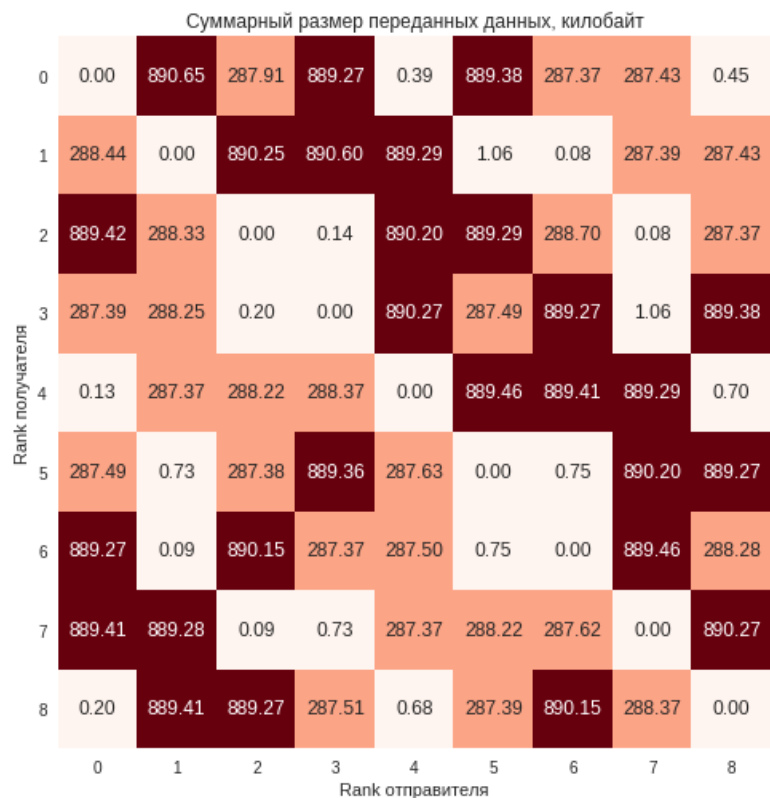
По парам (отправитель, получатель)

Подсчет переданных данных

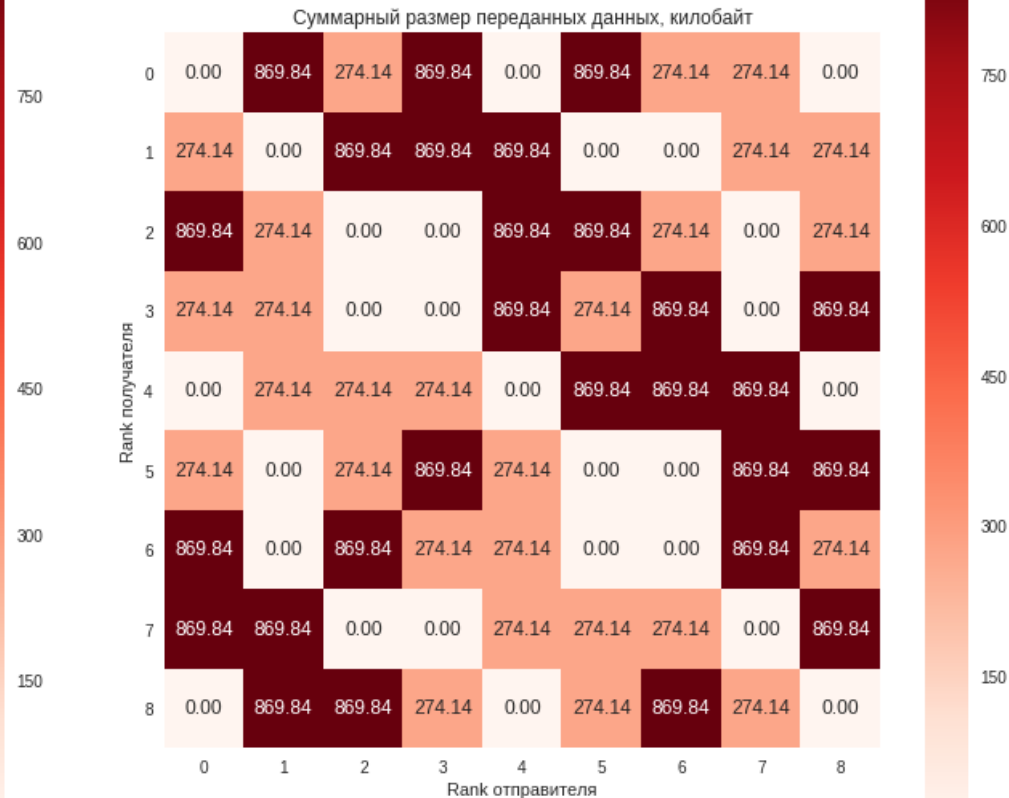
Суммирование размеров пакетов

NAS Parallel Benchmarks BT (Block Tri-diagonal solver)

Предлагаемый метод



ScoreP + Vampir

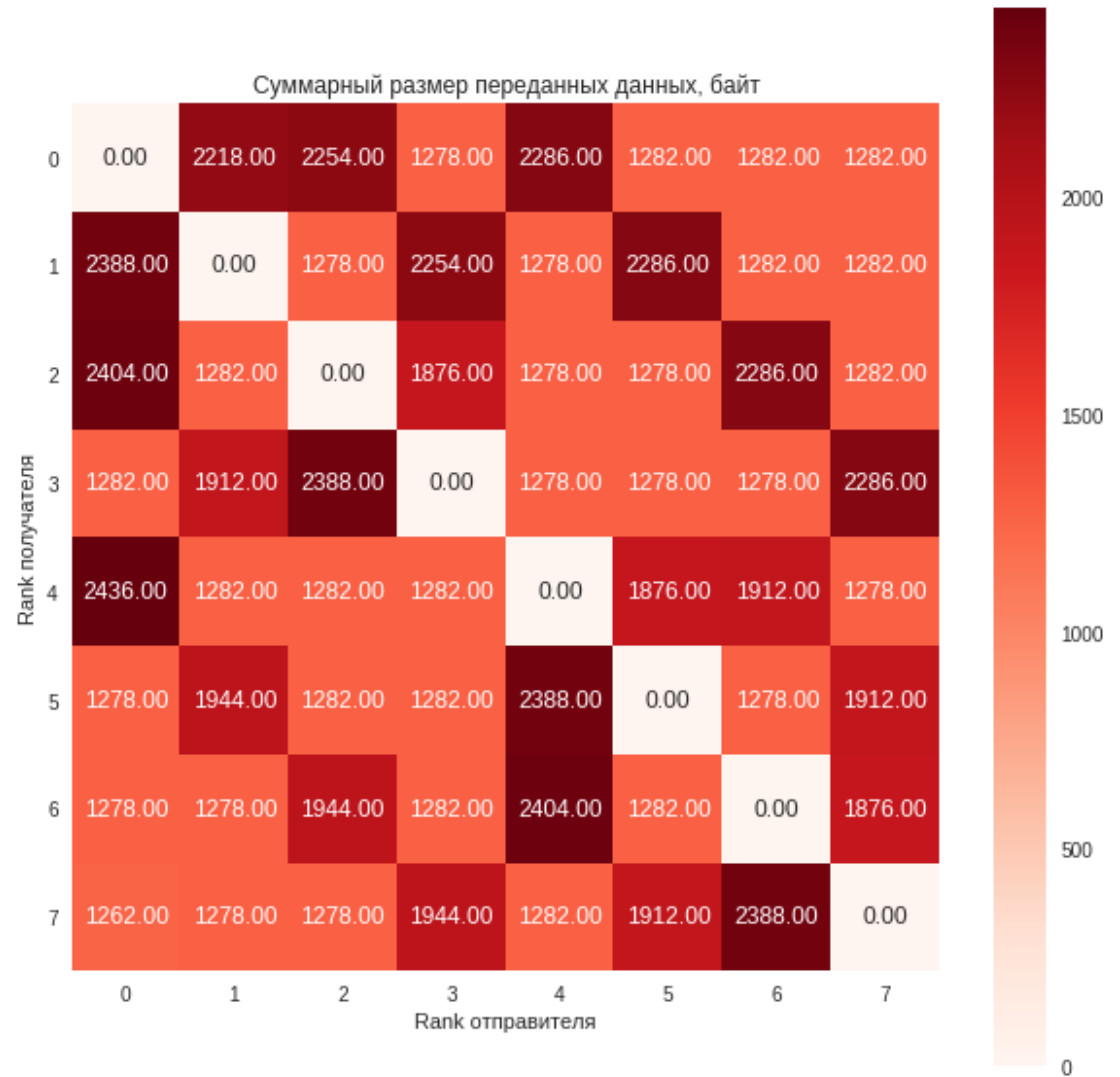


Вычислительных узлов	9
Узлов сброса	3

NAS Parallel Benchmarks

FT (discrete 3D fast Fourier Transform)

Вычислительных узлов	9
Узлов сброса	3



Результаты

Разработан метод построения коммуникационной матрицы, не требующий вмешательства в приложение

Представлена реализация метода для slurm

Проведено тестирование на NAS Parallel Benchmarks

Суперкомпьютерные дни в России 2017



**Построение коммуникационного профиля
параллельных программ путем анализа
сохраненной копии трафика сети суперкомпьютера**

Градсков Алексей Алексеевич, ВМК МГУ

Стефанов Константин Сергеевич, НИВЦ МГУ

Спасибо за внимание!