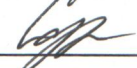


Утверждаю
Руководитель
ЦКП "Центр данных ДВО РАН"

к.т.н. А.А. Сорокин



"1" 07 2017 г.

**Методика измерения производительности подсистем
вычислительного кластера**

Версия 1

Хабаровск, 2017 г.

Местоположение документа

Этот документ редактируется и хранится в электронной форме. Бумажные копии действительны только в течение дня их вывода на печать и после подписания документа. В случае возникновения вопросов относительно актуальности документа - обращайтесь к его владельцу.

Перечень изменений

Дата внесения последних изменений:	Дата следующего пересмотра
------------------------------------	----------------------------

Версия	Дата	Описание внесенных изменений	Изменения отмечены
1.0.0	17.04.2017	Первая версия документа	-
1.0.1	18.04.2017	Добавлены результаты прогона теста LINPAK и результаты тестирования сети Infiniband.	-
1.0.2	23.04.2017	Добавлены выводы	-
1.0.3	30.05.2017	Исправлены расчеты на бустовый режим	

Оглавление

Местоположение документа.....	2
Перечень изменений.....	2
Описание кластера	4
Задачи тестирования и инструменты.....	5
Инструменты тестирования.....	6
Общие замечания	7
Тест на доступность нод кластера с mgmt узла.	8
Тест работоспособности Infiniband соединения.	9
Определение характеристик работы Infiniband сети.....	10
Определение характеристик работы кластера при помощи HPL.....	12
ВЫВОДЫ	16
Приложение 1. Характеристики работы Infiniband сети.	17

Описание кластера

В состав вычислительного кластера входит следующее оборудование:

- **1 сервер Sitonica PW22LC/IBM Power Systems S822LC 8335-GTB** - управляющий сервер / сервер хранения данных;
- **4 сервера Sitonica PW22LC/IBM Power Systems S822LC 8335-GTB** - узлы кластера;
- **2 x 10-core 3.492/4.023 GHz POWER8 Processor module** на каждом сервере, всего 100 POWER8 ядер на кластер;
- **8 x 32 GB DDR4 Memory** на каждом сервере, всего 1280Gb RAM на кластер;
- **5 x 2 Air-Cooled NVIDIA Tesla P100 GPU**, всего 10шт на кластер;
- Вычислительная сеть Infiniband построенная на основе коммутатора **Sitonica 8288-E36/Mellanox SB7700-ES2F**;

программное обеспечение (ПО):

- CentOS Linux release 7.3 LE;
 - ПО управления кластером – xCAT2 2.13.2;
 - ПО мониторинга кластера – Ganglia 3.7.1;
 - IBM Spectrum MPI 10.1.0.3;
 - Планировщик заданий PBS Pro.
-

Задачи тестирования и инструменты

Основными задачами данного тестирования являются:

1. Проверка работоспособности оборудования: серверов, систем хранения данных, высокопроизводительной сети Infiniband.
2. Определение реальных характеристик производительности для:
 1. Высокоскоростной сети Infiniband.
 2. Вычислительного кластера.
3. Сравнение производительности кластера с теоретическим расчетом.

Для выполнения данных задач используются:

1. Стандартные утилиты из состава ОС.
 2. Стандартные системные тесты для определения производительности подсистем Infiniband, MPI.
 3. Вспомогательные скрипты для тестирования.
 4. Тест High Performance Linpack (HPL) — промышленный стандарт для определения производительности кластеров (в т.ч. и с распределенной памятью). В настоящем тестировании используется скомпилированная версия HPL от NVIDIA, оптимизированная для работы на GPU NVIDIA GP100.
-

Инструменты тестирования

В качестве вспомогательных инструментов в тестировании использовался скрипт clsh:

```
#!/bin/sh
for i in `cat nodes`
do
    echo --- $i ---
    ssh $i "$1"
    echo -----
done
```

Для корректной работы скрипта необходимо наличие файла nodes в директории запуска. Этот файл должен содержать список вычислительных узлов кластера. В данном тестировании файл nodes выглядит следующим образом:

```
jupiter1
jupiter2
jupiter3
jupiter4
jupiter5
```

Общие замечания

В состав тестируемого кластера входят следующие узлы

Имя узла	Роль	Примечания
jupiter1	Узел управления	Используется для запуска всех тестов и сборки программ и других административных задач. Далее в тексте обозначается как MGMT .
jupiter2-jupiter5	Вычислительные узлы	

Для запуска тестов и компиляции программ служит сервер **mgmt**. Тестирование производится от пользователя `skovylov`, без специальных привилегий. Следует учитывать, что при отрицательном результате тестирования следует прекратить выполнение процедуры тестирования, устранить неполадки. Если отрицательный результат теста был вызван не ошибкой процедуры (например, неправильный ввод команды), а неполадками в кластере, то после устранения неполадок следует повторить процедуру тестирования сначала.

План тестирования:

1. Контроль работоспособности и корректной настройки компонент.
 1. Тест на доступность нод кластера с `mgmt` узла.
 2. Тест работоспособности Infiniband соединения.
2. Определение характеристик работы Infiniband сети.
3. Определение характеристик работы кластера при помощи HPL.

Тест на доступность нод кластера с mgmt узла.

```
[root@mgmt ~]# ./clsh date
```

Со всех узлов кластера должно быть корректно получено текущее значение времени.

Тест работоспособности Infiniband соединения.

Тест запускается от пользователя root из домашней директории. Для теста необходим файл nodes_ib0 со следующим содержанием:

```
192.168.50.1
192.168.50..2
192.168.50..3
192.168.50..4
192.168.50..5
```

Действия для проверки:

```
[root@mgmt ~]# ibv_devices
```

```
[root@mgmt ~]# ./clsh ibv_devices
```

```
[root@mgmt ~]# for i in `cat nodes_ib0`; do echo $i; ping -c3 $i | grep loss; done
```

Результат:

Выдача состояний IB карт и пинга по IroIB для всех узлов.

Выводы: На всех нодах настроен и работоспособен IB интерфейс. Настроен IroIB протокол.

Определение характеристик работы Infiniband сети.

Тест запускается от пользователя root из домашней директории. Для тестирования используются утилиты `ib_send_bw` и `ib_send_lat` из состава ОС. Для автоматизации тестирования используются скрипты `ib_bench_bw.sh` и `ib_bench_lat.sh`.

`ib_bench_bw.sh`

```
#!/bin/sh
RESULT=ib_bw_test_results
echo IB BW send tests results > $RESULT
for i in `cat nodes_ib0`
do
    ssh $i "ib_send_bw -a -F 1>/dev/null 2>/dev/null&"
    sleep 1
    echo ---$i--- >> $RESULT
    ib_send_bw -a -F $i >> $RESULT
    echo -----
done
```

`ib_bench_lat.sh`

```
#!/bin/sh
RESULT=ib_lat_test_results
echo IB LAT send tests results > $RESULT
for i in `cat nodes_ib0`
do
    ssh $i "ib_send_lat -a -F 1>/dev/null 2>/dev/null&"
    sleep 1
    echo ---$i--- >> $RESULT
    ib_send_lat -a -F $i >> $RESULT
    echo -----
done
```

Для корректной работы скрипты требуют наличия в директории запуска файла `nodes_ib0` со списком IP адресов IB интерфейсов.

Действия для проверки:

```
[root@mgmt ~]./ib_bench_lat.sh
```

```
[root@mgmt ~]/ib_bench_bw.sh
```

Результат: Данные производительности записываются в файлы `ib_bw_test_results` и `ib_lat_test_results`. Содержание файлов приведено в Приложении1.

Определение характеристик работы кластера при помощи HPL.

Для получения экспериментального показателя производительности кластера использовался скомпилированный тест HPL от NVIDIA.

Он запускался с использованием IBM Spectrum MPI.

1. Редактируется файл HPL.dat с параметрами запуска LINPACK. Ниже представлен конфигурационный файл для запуска теста

HPL.dat

```
HPLinpack benchmark input file
Innovative Computing Laboratory, University of Tennessee
HPL.out  output file name (if any)
6        device out (6=stdout,7=stderr,file)
1        # of problems sizes (N)
380160   Ns
1        # of NBs
768      NBs
1        PMAP process mapping (0=Row-,1=Column-major)
1        # of process grids (P x Q)
5        Ps
2        Qs
16.0     threshold
1        # of panel fact
2        PFACTs (0=left, 1=Crout, 2=Right)
1        # of recursive stopping criterium
2 8      NBMINs (>= 1)
1        # of panels in recursion
2        NDIVs
1        # of recursive panel fact.
2        RFACTs (0=left, 1=Crout, 2=Right)
1        # of broadcast
2 0 2    BCASTs (0=1rg,1=1rM,2=2rg,3=2rM,4=Lng,5=LnM)
1        # of lookahead depth
0        DEPTHS (>=0)
1        SWAP (0=bin-exch,1=long,2=mix)
192      swapping threshold
1        L1 in (0=transposed,1=no-transposed) form
0        U in (0=transposed,1=no-transposed) form
1        Equilibration (0=no,1=yes)
8        memory alignment in double (> 0)
```

2. В файле запуска run_linpack определяется переменная окружения «CPU_CORES_PER_RANK», равная кол-ву физических ядер, т.е. 10.
 3. В файле запуска run_linpack определяется переменная окружения
-

«GPU_DGEMM_SPLIT» в 1.0.

4. В файле запуска `run_linpack` определяется закрепление ранга MPI процесса за определенным GPU и набором ядер CPU. Следует помнить, что текущее поколение использует SMT8, т.е. физические ядра по факту располагаются со смещением в 8.

```
case ${lrank} in
[0])
    export CUDA_VISIBLE_DEVICES=0;
    numactl --physcpubind=0,8,16,24,32,40,48,56,64,72 --membind=0 \
        $APP
    ;;
[1])
    export CUDA_VISIBLE_DEVICES=1;
    numactl --physcpubind=80,88,96,104,112,120,128,136 --membind=1 \
        $APP
    ;;
esac
```

Действия для проверки:

```
mpirun -n 10 -bind-to none ./run_linpack | tee result_n5.log
```

Результат:

На 5 узлах получили результат 40,39 TFLOPS.

Детали теста приведены ниже:

```
=====
HPLinpack 2.1 -- High-Performance Linpack benchmark -- October 26, 2012
Written by A. Petitet and R. Clint Whaley, Innovative Computing Laboratory, UTK
Modified by Piotr Luszczek, Innovative Computing Laboratory, UTK
Modified by Julien Langou, University of Colorado Denver
=====
```

An explanation of the input/output parameters follows:

TV : Wall time / encoded variant.

N : The order of the coefficient matrix A.
NB : The partitioning blocking factor.
P : The number of process rows.
Q : The number of process columns.
Time : Time in seconds to solve the linear system.
Gflops : Rate of execution for solving the linear system.

The following parameter values will be used:

N : 380160
NB : 768
PMAP : Row-major process mapping
P : 5
Q : 2
PFACT : Crout
NBMIN : 2
NDIV : 2
RFACT : Crout
BCAST : 1ringM
DEPTH : 0
SWAP : Spread-roll (long)
L1 : no-transposed form
U : transposed form
EQUIL : no
ALIGN : 8 double precision words

-
- The matrix A is randomly generated for each test.
 - The following scaled residual check will be computed:
$$\|Ax-b\|_{\infty} / (\text{eps} * (\|x\|_{\infty} * \|A\|_{\infty} + \|b\|_{\infty}) * N)$$
 - The relative machine precision (*eps*) is taken to be 1.110223e-16
 - Computational tests pass if scaled residuals are less than 16.0

gpu_dgemm_split from environment variable 1.000
monitor_gpu from environment variable 1
gpu_temp_warning from environment variable 80
gpu_clock_warning from environment variable 1500
gpu_power_warning from environment variable 280
gpu_pcie_gen_warning from environment variable 3

gpu_pcie_width_warning from environment variable 8

test_system from environment variable 0

displaying Prog:%complete, N:columns, Time:seconds

iGF:instantaneous GF, GF:avg GF, GF_per: process GF

```
=====
```

<i>TV</i>	<i>N</i>	<i>NB</i>	<i>P</i>	<i>Q</i>	<i>Time</i>	<i>Gflops</i>
<i>WR01C2C2</i>	<i>380160</i>	<i>768</i>	<i>5</i>	<i>2</i>	<i>906.89</i>	<i>4.039e+04</i>

```
-----
```

||Ax-b||_oo/(eps(||A||_oo*||x||_oo+||b||_oo)*N)= 0.0000004 PASSED*

```
=====
```

Finished 1 tests with the following results:

- 1 tests completed and passed residual checks,*
- 0 tests completed and failed residual checks,*
- 0 tests skipped because of illegal input values.*

```
-----
```

End of Tests.

```
=====
```

ВЫВОДЫ

Тест сети Infiniband.

На всех узлах корректно работает сеть Infiniband. Показатели производительности соответствуют ожидаемым и одинаковы для всех узлов в пределах статистической погрешности.

Минимальная задержка при передаче сообщений малого размера ~1.3мкс. Скорость работы на больших сообщениях 11.4Gbps.

Ожидаемый результат выполнения LINPACK:

Теоретическое значение производительности кластера рассчитывается следующим образом:

Для центрального процессора расчет производится по следующей формуле:

$$5 \times (2 \times (\text{CPU_FREQ} \times 8 \times \#_CORES)),$$

где

CPU_FREQ – это базовая частота ядра процессора (3,492 ГГц)

#_CORES – кол-во ядер процессора

Для графического процессора расчет производится по следующей формуле:

$$5 \times (2 \times (\text{G_CLOCKS} \times 2 \times 1792)),$$

где

G_CLOCKS – частоты GPU базовая (1,48 ГГц)

Исходя из выше обозначенных формул следует, что:

1. Производительность CPU составит:
 $5 \times (2 \times (3.492 \times 8 \times 10)) = 2,794 \text{ TFLOPs}$
2. Производительность GPU составит:
 $5 \times (2 \times (1,48 \times 2 \times 1792)) = 53,043 \text{ TFLOPs}$
3. Суммарная производительность составит
 $R_{\text{peak}} = 2,794 + 53,043 = 55,84 \text{ TFLOPs}$

Важно понимать, что теоретическая производительность рассчитывается исходя из базовых частот счетных модулей (CPU и GPU), т.к. повышенные частоты (буст) имеют краткосрочный режим работы, который не приемлем для теста linpack.

Реальная производительность, достигнутая в тесте, составляет **40,39 TFLOPs**, что составляет **~72,33%** от теоретической пиковой.

Приложение 1. Характеристики работы Infiniband сети.

ib_bw_test_results

IB BW send tests results

---192.168.50.1---

Send BW Test

Dual-port : OFF Device : mlx5_0
Number of qps : 1 Transport type : IB
Connection type : RC Using SRQ : OFF
TX depth : 128
CQ Moderation : 100
Mtu : 4096[B]
Link type : IB
Max inline data : 0[B]
rdma_cm QPs : OFF
Data ex. method : Ethernet

local address: LID 0x03 QPN 0x0336 PSN 0x14080c

remote address: LID 0x03 QPN 0x0335 PSN 0x42e11f

#bytes	#iterations	BW peak[MB/sec]	BW average[MB/sec]	MsgRate[Mpps]
2	1000	7.75	7.42	3.891687
4	1000	16.14	15.96	4.184026
8	1000	33.97	33.56	4.399258
16	1000	71.02	70.01	4.588421
32	1000	146.03	144.21	4.725317
64	1000	303.40	291.14	4.770037
128	1000	606.79	594.64	4.871293

256	1000	1201.92	1186.90	4.861543
512	1000	2380.96	2364.81	4.843131
1024	1000	4761.94	4695.79	4.808493
2048	1000	8771.90	8751.32	4.480677
4096	1000	10695.13	10644.92	2.725100
8192	1000	11235.89	11219.47	1.436092
16384	1000	11331.39	11328.37	0.725016
32768	1000	11379.86	11376.83	0.364059
65536	1000	11444.95	11444.27	0.183108
131072	1000	11461.32	11460.70	0.091686
262144	1000	11473.70	11473.15	0.045893
524288	1000	11473.22	11473.08	0.022946
1048576	1000	11476.51	11476.49	0.011476
2097152	1000	11475.49	11475.44	0.005738
4194304	1000	11470.41	11470.36	0.002868
8388608	1000	11464.64	11464.64	0.001433

---192.168.50.2---

Send BW Test

Dual-port : OFF Device : mlx5_0
Number of qps : 1 Transport type : IB
Connection type : RC Using SRQ : OFF
TX depth : 128
CQ Moderation : 100
Mtu : 4096[B]
Link type : IB
Max inline data : 0[B]
rdma_cm QPs : OFF

Data ex. method : Ethernet

local address: LID 0x03 QPN 0x0337 PSN 0x53fd96

remote address: LID 0x04 QPN 0x02cc PSN 0x59d39e

#bytes	#iterations	BW peak[MB/sec]	BW average[MB/sec]	MsgRate[Mpps]
2	1000	6.98	6.61	3.464579
4	1000	13.95	13.55	3.553185
8	1000	30.05	29.25	3.833719
16	1000	64.04	62.57	4.100389
32	1000	143.35	139.90	4.584229
64	1000	300.48	294.28	4.821542
128	1000	631.32	599.99	4.915105
256	1000	1315.79	1281.28	5.248126
512	1000	2717.40	2663.47	5.454782
1024	1000	5681.82	5568.11	5.701747
2048	1000	10204.13	10160.90	5.202380
4096	1000	11560.66	11509.90	2.946533
8192	1000	11695.96	11685.26	1.495714
16384	1000	11764.75	11753.36	0.752215
32768	1000	11790.76	11786.25	0.377160
65536	1000	11799.48	11797.02	0.188752
131072	1000	11812.51	11811.42	0.094491
262144	1000	11815.75	11815.51	0.047262
524288	1000	11816.77	11816.38	0.023633
1048576	1000	11817.73	11817.56	0.011818
2097152	1000	11818.11	11818.06	0.005909
4194304	1000	11819.41	11819.41	0.002955
8388608	1000	11818.05	11818.04	0.001477

---192.168.50.3---

Send BW Test

Dual-port : OFF Device : mlx5_0
Number of qps : 1 Transport type : IB
Connection type : RC Using SRQ : OFF
TX depth : 128
CQ Moderation : 100
Mtu : 4096[B]
Link type : IB
Max inline data : 0[B]
rdma_cm QPs : OFF
Data ex. method : Ethernet

local address: LID 0x03 QPN 0x0338 PSN 0x9d674f
remote address: LID 0x02 QPN 0x00a5 PSN 0x5d45fe

#bytes	#iterations	BW peak[MB/sec]	BW average[MB/sec]	MsgRate[Mpps]
2	1000	6.98	6.59	3.457195
4	1000	13.95	13.53	3.545731
8	1000	30.05	29.15	3.821153
16	1000	64.04	63.27	4.146244
32	1000	131.30	107.69	3.528812
64	1000	300.48	297.97	4.881933
128	1000	625.00	604.32	4.950553
256	1000	1315.79	1311.08	5.370191
512	1000	2747.26	2721.36	5.573341
1024	1000	5681.83	5615.47	5.750240

2048	1000	10204.17	10180.07	5.212194
4096	1000	11560.62	11530.63	2.951841
8192	1000	11695.90	11687.15	1.495955
16384	1000	11764.51	11754.09	0.752262
32768	1000	11790.71	11786.16	0.377157
65536	1000	11799.46	11797.15	0.188754
131072	1000	11812.52	11810.42	0.094483
262144	1000	11814.60	11814.39	0.047258
524288	1000	11817.91	11817.46	0.023635
1048576	1000	11817.65	11817.52	0.011818
2097152	1000	11819.24	11819.20	0.005910
4194304	1000	11818.29	11818.24	0.002955
8388608	1000	11819.53	11819.52	0.001477

---192.168.50.4---

Send BW Test

Dual-port : OFF Device : mlx5_0
Number of qps : 1 Transport type : IB
Connection type : RC Using SRQ : OFF
TX depth : 128
CQ Moderation : 100
Mtu : 4096[B]
Link type : IB
Max inline data : 0[B]
rdma_cm QPs : OFF
Data ex. method : Ethernet

local address: LID 0x03 QPN 0x0339 PSN 0x1eceed

remote address: LID 0x05 QPN 0x00a5 PSN 0xe269a1

#bytes	#iterations	BW peak[MB/sec]	BW average[MB/sec]	MsgRate[Mpps]
2	1000	6.98	6.59	3.455956
4	1000	13.95	13.54	3.549990
8	1000	30.05	29.24	3.832885
16	1000	64.04	62.38	4.088233
32	1000	135.87	132.27	4.334160
64	1000	300.48	294.09	4.818435
128	1000	631.31	616.05	5.046663
256	1000	1329.79	1257.93	5.152475
512	1000	2747.25	2687.05	5.503071
1024	1000	5681.82	5547.49	5.680627
2048	1000	10204.17	10150.41	5.197012
4096	1000	11560.66	11530.53	2.951817
8192	1000	11695.87	11684.50	1.495616
16384	1000	11764.77	11753.83	0.752245
32768	1000	11790.65	11786.00	0.377152
65536	1000	11803.87	11803.44	0.188855
131072	1000	11808.11	11807.87	0.094463
262144	1000	11814.72	11814.38	0.047258
524288	1000	11817.95	11817.65	0.023635
1048576	1000	11818.79	11818.68	0.011819
2097152	1000	11819.24	11819.19	0.005910
4194304	1000	11818.42	11818.37	0.002955
8388608	1000	11818.54	11818.51	0.001477

---192.168.50.5---

Send BW Test

Dual-port : OFF Device : mlx5_0
Number of qps : 1 Transport type : IB
Connection type : RC Using SRQ : OFF
TX depth : 128
CQ Moderation : 100
Mtu : 4096[B]
Link type : IB
Max inline data : 0[B]
rdma_cm QPs : OFF
Data ex. method : Ethernet

local address: LID 0x03 QPN 0x033a PSN 0x1b79a3

remote address: LID 0x06 QPN 0x00a5 PSN 0xa3b541

#bytes	#iterations	BW peak[MB/sec]	BW average[MB/sec]	MsgRate[Mpps]
2	1000	6.98	6.60	3.458266
4	1000	15.02	14.59	3.825546
8	1000	32.02	31.22	4.091798
16	1000	67.93	66.33	4.347077
32	1000	143.35	140.03	4.588403
64	1000	315.65	308.63	5.056613
128	1000	657.90	641.89	5.258335
256	1000	1373.63	1308.39	5.359152
512	1000	2873.59	2790.95	5.715864
1024	1000	5813.96	5700.22	5.837030
2048	1000	10204.06	10180.89	5.212617
4096	1000	11560.66	11521.30	2.949453
8192	1000	11695.89	11686.32	1.495849

16384	1000	11764.78	11752.51	0.752160
32768	1000	11790.81	11786.20	0.377159
65536	1000	11803.80	11803.37	0.188854
131072	1000	11810.39	11808.74	0.094470
262144	1000	11814.61	11814.36	0.047257
524288	1000	11817.92	11817.53	0.023635
1048576	1000	11817.64	11817.50	0.011817
2097152	1000	11818.15	11818.07	0.005909
4194304	1000	11818.32	11818.26	0.002955
8388608	1000	11818.43	11818.41	0.001477

ib_lat_test_result:

IB LAT send tests results

---192.168.50.1---

Send Latency Test

Dual-port : OFF Device : mlx5_0
Number of qps : 1 Transport type : IB
Connection type : RC Using SRQ : OFF
TX depth : 1
Mtu : 4096[B]
Link type : IB
Max inline data : 236[B]
rdma_cm QPs : OFF
Data ex. method : Ethernet

local address: LID 0x03 QPN 0x033c PSN 0x79b2b3

remote address: LID 0x03 QPN 0x033b PSN 0xa67c99

#bytes	#iterations	t_min[usec]	t_max[usec]	t_typical[usec]	t_avg[usec]	t_stdev[usec]	99% percentile[usec]	99.9% percentile[usec]
2	1000	1.10	4.37	1.29	1.30	0.12	1.51	18.45
4	1000	0.95	2.29	1.04	1.06	0.08	1.31	18.45
8	1000	0.93	5.11	1.02	1.04	0.15	1.27	18.45
16	1000	0.94	1.99	1.02	1.04	0.07	1.26	18.45
32	1000	0.94	2.04	1.03	1.04	0.08	1.26	18.45
64	1000	0.99	3.07	1.08	1.10	0.10	1.37	18.45
128	1000	1.02	3.29	1.12	1.16	0.14	1.68	18.45
256	1000	1.30	2.76	1.37	1.38	0.07	1.58	18.45
512	1000	1.39	2.42	1.47	1.48	0.06	1.66	18.45
1024	1000	1.54	2.55	1.62	1.63	0.06	1.81	18.45
2048	1000	1.81	4.59	1.88	1.89	0.11	2.08	18.45
4096	1000	2.33	3.27	2.39	2.40	0.05	2.57	18.45
8192	1000	2.86	5.23	2.94	2.95	0.10	3.16	18.45
16384	1000	3.92	6.32	3.99	4.00	0.09	4.19	18.45
32768	1000	5.64	8.74	5.75	5.77	0.13	6.00	18.45
65536	1000	8.74	11.22	8.85	8.86	0.12	9.08	18.45
131072	1000	13.51	16.10	14.50	14.46	0.81	15.62	18.45
262144	1000	25.30	28.27	26.03	26.07	0.54	26.91	18.45
524288	1000	46.99	49.76	47.84	47.82	0.53	48.62	18.45
1048576	1000	90.48	93.30	91.36	91.37	0.52	92.17	18.45
2097152	1000	177.59	180.53	178.47	178.48	0.48	179.29	18.45
4194304	1000	351.80	354.91	352.75	352.74	0.40	353.60	18.45
8388608	1000	700.97	703.68	701.57	701.58	0.24	702.18	18.45

---192.168.50.2---

Send Latency Test

Dual-port : OFF Device : mlx5_0
Number of qps : 1 Transport type : IB
Connection type : RC Using SRQ : OFF
TX depth : 1
Mtu : 4096[B]
Link type : IB
Max inline data : 236[B]
rdma_cm QPs : OFF
Data ex. method : Ethernet

local address: LID 0x03 QPN 0x033d PSN 0x658417

remote address: LID 0x04 QPN 0x02cd PSN 0xf64a05

#bytes	#iterations	t_min[usec]	t_max[usec]	t_typical[usec]	t_avg[usec]	t_stdev[usec]	99% percentile[usec]	99.9% percentile[usec]
2	1000	1.19	5.11	1.32	1.33	0.15	1.52	18.45
4	1000	1.05	2.50	1.10	1.11	0.07	1.38	18.45
8	1000	1.03	2.05	1.08	1.09	0.04	1.15	18.45
16	1000	1.02	3.70	1.08	1.08	0.10	1.15	18.45
32	1000	1.01	2.02	1.07	1.07	0.04	1.14	18.45
64	1000	1.08	2.07	1.14	1.14	0.04	1.21	18.45
128	1000	1.11	2.15	1.16	1.17	0.04	1.24	18.45
256	1000	1.43	4.31	1.48	1.49	0.11	1.56	18.45
512	1000	1.54	2.56	1.60	1.60	0.05	1.67	18.45
1024	1000	1.66	3.07	1.72	1.72	0.06	1.79	18.45
2048	1000	1.91	2.93	1.96	1.96	0.04	2.03	18.45

4096	1000	2.35	3.33	2.41	2.42	0.05	2.49	18.45
8192	1000	3.03	5.44	3.09	3.10	0.08	3.22	18.45
16384	1000	4.21	8.32	4.27	4.28	0.20	4.48	18.45
32768	1000	6.15	8.95	6.21	6.22	0.09	6.32	18.45
65536	1000	9.29	12.25	9.37	9.38	0.10	9.52	18.45
131072	1000	14.58	17.43	14.71	14.72	0.11	14.86	18.45
262144	1000	25.16	28.10	25.31	25.31	0.11	25.52	18.45
524288	1000	46.27	49.16	46.47	46.47	0.11	46.64	18.45
1048576	1000	88.63	91.47	88.76	88.77	0.11	88.92	18.45
2097152	1000	173.22	176.05	173.37	173.38	0.11	173.53	18.45
4194304	1000	342.44	345.32	342.60	342.60	0.14	342.79	18.45
8388608	1000	680.86	691.91	681.04	681.05	0.36	681.20	18.45

---192.168.50.3---

Send Latency Test

Dual-port : OFF Device : mlx5_0
Number of qps : 1 Transport type : IB
Connection type : RC Using SRQ : OFF
TX depth : 1
Mtu : 4096[B]
Link type : IB
Max inline data : 236[B]
rdma_cm QPs : OFF
Data ex. method : Ethernet

local address: LID 0x03 QPN 0x033e PSN 0xe0c68e

remote address: LID 0x02 QPN 0x00a6 PSN 0xe8ce0e

#bytes	#iterations	t_min[usec]	t_max[usec]	t_typical[usec]	t_avg[usec]	t_stdev[usec]	99% percentile[usec]	99.9% percentile[usec]
2	1000	1.19	4.23	1.38	1.42	0.20	2.26	18.45
4	1000	1.02	2.35	1.08	1.09	0.07	1.39	18.45
8	1000	1.01	3.82	1.06	1.07	0.10	1.13	18.45
16	1000	1.01	2.82	1.06	1.07	0.07	1.13	18.45
32	1000	1.02	2.01	1.07	1.07	0.04	1.13	18.45
64	1000	1.08	2.60	1.14	1.14	0.06	1.21	18.45
128	1000	1.11	2.98	1.17	1.17	0.07	1.24	18.45
256	1000	1.43	2.84	1.48	1.49	0.05	1.55	18.45
512	1000	1.54	2.53	1.60	1.60	0.04	1.68	18.45
1024	1000	1.65	2.65	1.72	1.72	0.05	1.79	18.45
2048	1000	1.90	3.08	1.96	1.97	0.05	2.04	18.45
4096	1000	2.36	3.36	2.41	2.42	0.04	2.56	18.45
8192	1000	3.05	5.45	3.11	3.12	0.09	3.23	18.45
16384	1000	4.20	6.89	4.28	4.29	0.09	4.41	18.45
32768	1000	6.15	10.77	6.21	6.22	0.17	6.31	18.45
65536	1000	9.24	11.91	9.36	9.37	0.11	9.48	18.45
131072	1000	14.59	22.43	14.72	14.74	0.27	14.88	18.45
262144	1000	25.11	29.36	25.30	25.30	0.17	25.44	18.45
524288	1000	46.36	49.15	46.47	46.48	0.10	46.62	18.45
1048576	1000	88.63	91.48	88.78	88.78	0.10	88.91	18.45
2097152	1000	173.24	176.17	173.38	173.39	0.10	173.52	18.45
4194304	1000	342.46	345.31	342.61	342.62	0.11	342.77	18.45
8388608	1000	680.87	683.76	681.04	681.04	0.11	681.20	18.45

---192.168.50.4---

Send Latency Test

Dual-port : OFF Device : mlx5_0
 Number of qps : 1 Transport type : IB
 Connection type : RC Using SRQ : OFF
 TX depth : 1
 Mtu : 4096[B]
 Link type : IB
 Max inline data : 236[B]
 rdma_cm QPs : OFF
 Data ex. method : Ethernet

local address: LID 0x03 QPN 0x033f PSN 0xd54b8
 remote address: LID 0x05 QPN 0x00a6 PSN 0xd8643

#bytes	#iterations	t_min[usec]	t_max[usec]	t_typical[usec]	t_avg[usec]	t_stdev[usec]	99% percentile[usec]	99.9% percentile[usec]
2	1000	1.28	4.62	1.45	1.50	0.19	2.22	18.45
4	1000	1.07	2.54	1.14	1.16	0.09	1.50	18.45
8	1000	1.06	3.67	1.11	1.12	0.09	1.20	18.45
16	1000	1.04	3.11	1.10	1.11	0.07	1.20	18.45
32	1000	1.04	2.08	1.10	1.10	0.05	1.20	18.45
64	1000	1.11	2.69	1.17	1.17	0.07	1.27	18.45
128	1000	1.14	2.19	1.19	1.20	0.05	1.30	18.45
256	1000	1.49	2.90	1.54	1.55	0.06	1.68	18.45
512	1000	1.58	2.56	1.63	1.64	0.05	1.75	18.45
1024	1000	1.70	2.66	1.75	1.76	0.05	1.86	18.45
2048	1000	1.94	2.91	2.00	2.01	0.05	2.10	18.45
4096	1000	2.38	3.36	2.45	2.45	0.05	2.58	18.45
8192	1000	3.05	5.52	3.12	3.12	0.09	3.27	18.45
16384	1000	4.23	6.77	4.30	4.30	0.09	4.42	18.45

32768	1000	6.17	10.61	6.25	6.26	0.15	6.44	18.45
65536	1000	9.31	11.90	9.38	9.39	0.09	9.51	18.45
131072	1000	14.60	19.28	14.75	14.76	0.18	14.92	18.45
262144	1000	25.19	28.16	25.35	25.36	0.12	25.56	18.45
524288	1000	46.34	49.18	46.49	46.50	0.12	46.67	18.45
1048576	1000	88.65	91.48	88.80	88.81	0.11	88.97	18.45
2097152	1000	173.22	176.11	173.41	173.41	0.12	173.60	18.45
4194304	1000	342.46	345.37	342.62	342.63	0.12	342.82	18.45
8388608	1000	680.88	685.48	681.05	681.06	0.17	681.24	18.45

---192.168.50.5---

Send Latency Test

Dual-port : OFF Device : mlx5_0
 Number of qps : 1 Transport type : IB
 Connection type : RC Using SRQ : OFF
 TX depth : 1
 Mtu : 4096[B]
 Link type : IB
 Max inline data : 236[B]
 rdma_cm QPs : OFF
 Data ex. method : Ethernet

local address: LID 0x03 QPN 0x0340 PSN 0x1bbf6c

remote address: LID 0x06 QPN 0x00a6 PSN 0x38199f

#bytes	#iterations	t_min[usec]	t_max[usec]	t_typical[usec]	t_avg[usec]	t_stdev[usec]	99% percentile[usec]	99.9% percentile[usec]
2	1000	1.30	4.86	1.42	1.43	0.13	1.64	18.45

4	1000	1.07	2.71	1.15	1.16	0.10	1.56	18.45
8	1000	1.05	3.76	1.11	1.12	0.09	1.20	18.45
16	1000	1.04	3.05	1.10	1.10	0.07	1.18	18.45
32	1000	1.04	3.35	1.10	1.11	0.09	1.25	18.45
64	1000	1.11	2.94	1.17	1.18	0.07	1.29	18.45
128	1000	1.14	2.19	1.20	1.20	0.05	1.30	18.45
256	1000	1.46	2.99	1.52	1.53	0.06	1.64	18.45
512	1000	1.55	3.42	1.61	1.62	0.07	1.72	18.45
1024	1000	1.67	2.64	1.73	1.74	0.05	1.85	18.45
2048	1000	1.91	3.48	1.97	1.98	0.07	2.10	18.45
4096	1000	2.37	3.32	2.43	2.43	0.05	2.54	18.45
8192	1000	3.04	5.50	3.11	3.12	0.09	3.22	18.45
16384	1000	4.22	6.81	4.29	4.30	0.09	4.43	18.45
32768	1000	6.16	10.17	6.23	6.24	0.15	6.41	18.45
65536	1000	9.19	11.81	9.39	9.39	0.09	9.54	18.45
131072	1000	14.62	18.94	14.75	14.76	0.16	14.90	18.45
262144	1000	25.15	32.70	25.33	25.35	0.26	25.50	18.45
524288	1000	46.30	49.15	46.48	46.49	0.10	46.65	18.45
1048576	1000	88.62	93.97	88.79	88.80	0.19	88.98	18.45
2097152	1000	173.26	176.16	173.39	173.40	0.10	173.55	18.45
4194304	1000	342.47	345.31	342.62	342.63	0.10	342.77	18.45
8388608	1000	680.86	686.72	681.04	681.05	0.23	681.22	18.45
