总结

1. 发现nvvp和nvprof不支持在算力比较高（8.9）的显卡上使用，于是使用了nsight systems
2. 将测量的GPU运行时间调整为测量**数据拷贝时间和kernel运行时间**
3. 经过多次测试，原版运行时间大约为**675ms**；使用pinned memory的运行时间大约为**173ms**；使用breadth-first 2 streams和pinned memory的运行时间大约为**168ms**；使用breadth-first 3 streams和pinned memory的运行时间大约为**178ms**
4. 在使用2 streams的情况下，使用breadth-first和depth-first两种方式的运行时间没有测量出明显区别，通过nsight systems可以看出stream的运行顺序变了，时间上没有明显区别应该是和显卡有关
5. 对于原版和pinned memory版本来说，[6/8]+[7/8]的时间近似就是测出来的时间；而当stream的数量超过1时，这个约等式不再成立，根据stream的作用可以得出测量的时间理论上不再是kernel的时间和数据拷贝的时间线性和，根据nsight systems给出的timeline可知，此时的kernel执行和数据拷贝存在overlap，提高了运行效率。
6. 使用2streams比单纯使用pinned memory大概快了5ms，使用3streams又比单纯使用pinned memory大概慢了5ms。通过下一章节的数据可以看出stream增加会导致创建成本同步成本等成本增加，结果上慢了一点应该是因为3streams的排布不如2streams的效率高
7. 多次测量时发现3streams的用时不如2streams的稳定，应该和3不是2的幂有关

数据

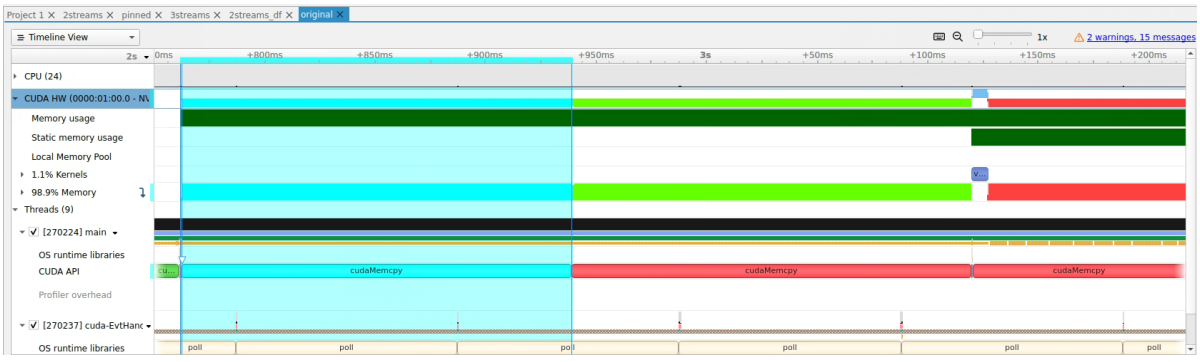
> 包含了命令行的分析总结和GUI的timeline

原版

```
$ nsys profile --stats=true ./main
[Vector addition of 104857600 elements]
A ok!B ok!C ok!Copy input data from the host memory to the CUDA device
CUDA kernel launch with 409600 blocks of 256 threads
Copy output data from the CUDA device to the host memory
GPU Time used:  674.0 ms
CPU Time used: 902.5 ms
Test PASSED
Done
Generating '/tmp/nsys-report-9fb9.qdstrm'
[1/8] [========================100%] report1.nsys-rep
[2/8] [========================100%] report1.sqlite
[3/8] Executing 'nvtxsum' stats report
SKIPPED: /home/admini/code/generalgpu/stream/build/report1.sqlite does not
contain NV Tools Extension (NVTX) data.
[4/8] Executing 'osrtsum' stats report

 Time (%)  Total Time (ns)  Num Calls    Avg (ns)       Med (ns)      Min (ns)
   Max (ns)       StdDev (ns)             Name
 --------  ---------------  ---------  -------------  -------------  -----------
 -------------  ---------------  ----------------------
     36.4    1,871,878,762          2  935,939,381.0  935,939,381.0    1,726,358
 1,870,152,404  1,321,176,727.3  sem_wait
     33.2    1,703,924,428         27   63,108,312.1  100,202,545.0        1,620
   100,267,667     46,615,563.2  poll
```

```
     29.2    1,500,403,355          3  500,134,451.7  500,152,039.0  500,083,832
  500,167,484         44,512.9  pthread_cond_timedwait
      1.1       56,953,982        446      127,699.5        5,043.5        1,021
   17,756,478      1,168,657.0  ioctl
      0.0          903,761         27       33,472.6        7,885.0        3,223
      535,724        101,050.5  mmap64
      0.0          790,354          9       87,817.1       97,007.0        4,728
      270,348         82,779.0  sem_timedwait
      0.0          350,439         44        7,964.5        7,336.5        1,922
       25,047          3,941.2  open64
      0.0          329,887          5       65,977.4       56,914.0       34,540
      120,088         32,663.1  pthread_create
      0.0          162,678         20        8,133.9        3,184.0        1,090
       64,653         13,940.1  mmap
      0.0           91,623         26        3,524.0        2,845.0        1,136
        9,218          2,553.5  fopen
      0.0           65,996         11        5,999.6        3,947.0        1,318
       19,637          5,250.1  munmap
      0.0           42,311          4       10,577.8        8,134.5        3,029
       23,013          9,263.4  fgets
      0.0           17,251          5        3,450.2        2,703.0        1,105
        5,652          1,945.3  open
      0.0           12,062          2        6,031.0        6,031.0        3,641
        8,421          3,380.0  socket
      0.0           11,591          5        2,318.2        2,314.0        1,614
        3,527            751.9  fread
      0.0            7,056          1        7,056.0        7,056.0        7,056
        7,056              0.0  connect
      0.0            6,929          5        1,385.8        1,324.0        1,145
        1,844            267.2  fclose
      0.0            5,918          1        5,918.0        5,918.0        5,918
        5,918              0.0  pipe2
      0.0            5,409          3        1,803.0        1,916.0        1,269
        2,224            487.4  read
      0.0            3,138          2        1,569.0        1,569.0        1,353
        1,785            305.5  write
      0.0            2,141          1        2,141.0        2,141.0        2,141
        2,141              0.0  fopen64
      0.0            2,014          1        2,014.0        2,014.0        2,014
        2,014              0.0  fcntl

[5/8] Executing 'cudaapisum' stats report

 Time (%)  Total Time (ns)  Num Calls    Avg (ns)       Med (ns)      Min (ns)
  Max (ns)    StdDev (ns)            Name
 --------  ---------------  ---------  -------------  -------------  -----------
 -----------  ------------  ----------------------
     88.6      673,424,071          3  224,474,690.3  182,700,874.0  174,260,069
 316,463,128  79,776,038.4  cudaMemcpy
     10.7       81,183,968          2   40,591,984.0   40,591,984.0          696
  81,183,272  57,404,750.0  cudaEventCreate
      0.6        4,564,216          3    1,521,405.3    2,105,675.0      263,817
   2,194,724   1,090,013.2  cudaFree
      0.0          310,990          1      310,990.0      310,990.0      310,990
     310,990            0.0  cudaLaunchKernel
```

```
      0.0          205,717          3       68,572.3       57,842.0       49,911
      97,964       25,761.0  cudaMalloc
      0.0           13,111          2        6,555.5        6,555.5        6,226
       6,885          466.0  cudaEventRecord
      0.0            3,080          1        3,080.0        3,080.0        3,080
       3,080            0.0  cudaEventSynchronize
      0.0            1,156          2          578.0          578.0          287
          869        411.5  cudaEventDestroy
      0.0              905          1          905.0          905.0          905
          905            0.0  cuModuleGetLoadingMode

[6/8] Executing 'gpukernsum' stats report

 Time (%)  Total Time (ns)  Instances   Avg (ns)     Med (ns)    Min (ns)   Max
(ns)   StdDev (ns)      GridXYZ          BlockXYZ
Name
 --------  ---------------  ---------  -----------  -----------  ---------  -----
----  -----------  ----------------  ---------------  ------------------------------
----------------------------
    100.0        7,083,616          1  7,083,616.0  7,083,616.0  7,083,616
 7,083,616          0.0  409600    1    1   256    1    1  vectorAdd(const double
*, const double *, double *, int)

[7/8] Executing 'gpumemtimesum' stats report

 Time (%)  Total Time (ns)  Count    Avg (ns)       Med (ns)      Min (ns)
Max (ns)    StdDev (ns)      Operation
 --------  ---------------  -----  -------------  -------------  -----------  ---
--------  -----------  ------------------
     53.6      356,676,468      2  178,338,234.0  178,338,234.0  174,126,935
 182,549,533  5,955,676.2  [CUDA memcpy HtoD]
     46.4      308,821,854      1  308,821,854.0  308,821,854.0  308,821,854
 308,821,854          0.0  [CUDA memcpy DtoH]

[8/8] Executing 'gpumemsizesum' stats report

 Total (MB)  Count  Avg (MB)  Med (MB)  Min (MB)  Max (MB)  StdDev (MB)
Operation
 ----------  -----  --------  --------  --------  --------  -----------  --------
----------
  1,677.722      2   838.861   838.861   838.861   838.861        0.000  [CUDA
memcpy HtoD]
    838.861      1   838.861   838.861   838.861   838.861        0.000  [CUDA
memcpy DtoH]
```

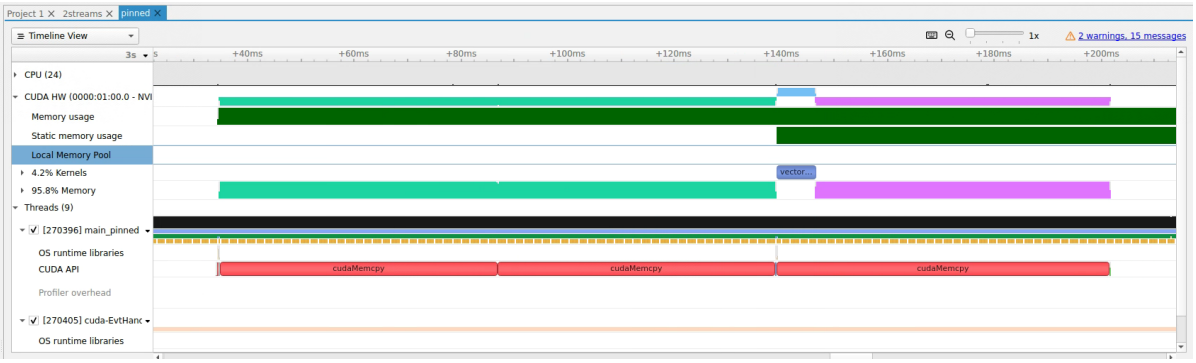


使用pinned memory

```
$ nsys profile --stats=true ./main_pinned
[Vector addition of 104857600 elements]
Copy input data from the host memory to the CUDA device
CUDA kernel launch with 409600 blocks of 256 threads
Copy output data from the CUDA device to the host memory
GPU Time used:  172.5 ms
CPU Time used: 913.8 ms
Test PASSED
Done
Generating '/tmp/nsys-report-3171.qdstrm'
[1/8] [========================100%] report10.nsys-rep
[2/8] [========================100%] report10.sqlite
[3/8] Executing 'nvtxsum' stats report
SKIPPED: /home/admini/code/generalgpu/stream/build/report10.sqlite does not
contain NV Tools Extension (NVTX) data.
[4/8] Executing 'osrtsum' stats report

 Time (%)  Total Time (ns)  Num Calls     Avg (ns)         Med (ns)       Min
(ns)      Max (ns)       StdDev (ns)             Name
 --------  ---------------  ---------  ---------------  ---------------  --------
---  -------------  ---------------  ----------------------
     33.7    4,218,830,129          2  2,109,415,064.5  2,109,415,064.5
 2,212,856  4,216,617,273  2,980,033,941.9  sem_wait
     32.0    4,008,072,448         50     80,161,449.0    100,203,911.0
 2,159    100,252,674     38,729,981.0  poll
     28.0    3,500,954,069          7    500,136,295.6    500,151,025.0
 500,064,024    500,172,517         40,436.2  pthread_cond_timedwait
      5.4      673,308,632        471      1,429,530.0          4,892.0
 1,008    191,342,389     15,095,320.8  ioctl
      0.9      115,111,404         26      4,427,361.7          8,188.5
 1,384     39,305,693     12,486,171.8  mmap
      0.0        1,072,952          9        119,216.9        106,741.0
71,567        237,809         51,550.8  sem_timedwait
      0.0          837,156         27         31,005.8          6,891.0
 3,690        505,067         95,274.6  mmap64
      0.0          350,722          5         70,144.4         61,090.0
30,247        133,131         40,148.9  pthread_create
      0.0          329,321         44          7,484.6          7,321.0
 1,813         16,303          3,092.1  open64
      0.0          109,739         17          6,455.2          4,434.0
 1,783         19,686          4,972.1  munmap
      0.0          107,127         30          3,570.9          2,567.5
 1,081         19,068          3,558.4  fopen
      0.0           44,352          4         11,088.0          9,451.0
 3,330         22,120          8,930.3  fgets
      0.0           23,051          5          4,610.2          4,044.0
 2,459          7,160          1,865.1  fread
      0.0           22,552          5          4,510.4          2,208.0
 1,236         11,171          4,143.1  open
      0.0           10,382          2          5,191.0          5,191.0
 2,820          7,562          3,353.1  socket
```

```
     0.0           7,352          3          2,450.7          1,236.0
 1,133          4,983          2,193.7  fclose
     0.0           7,180          1          7,180.0          7,180.0
 7,180          7,180              0.0  connect
     0.0           5,352          1          5,352.0          5,352.0
 5,352          5,352              0.0  pipe2
     0.0           4,978          3          1,659.3          1,854.0
 1,194          1,930            404.8  read
     0.0           4,410          2          2,205.0          2,205.0
 1,665          2,745            763.7  write
     0.0           3,491          1          3,491.0          3,491.0
 3,491          3,491              0.0  fopen64
     0.0           3,373          1          3,373.0          3,373.0
 3,373          3,373              0.0  fcntl
     0.0           1,040          1          1,040.0          1,040.0
 1,040          1,040              0.0  fflush

[5/8] Executing 'cudaapisum' stats report

 Time (%)  Total Time (ns)  Num Calls    Avg (ns)       Med (ns)      Min (ns)
  Max (ns)    StdDev (ns)            Name
 --------  ---------------  ---------  -------------  -------------  -----------
 -----------  ------------  ----------------------
     66.8      661,479,223          3  220,493,074.3  193,939,888.0  188,565,645
 278,973,690  50,716,934.3  cudaHostAlloc
     17.4      172,042,881          3   57,347,627.0   57,250,147.0   52,753,572
  62,039,162   4,643,562.4  cudaMemcpy
     15.3      151,670,130          3   50,556,710.0   49,837,796.0   49,608,299
  52,224,035   1,448,498.1  cudaFreeHost
      0.5        4,568,332          3    1,522,777.3    2,108,546.0      250,479
   2,209,307   1,102,993.9  cudaFree
      0.0          456,910          1      456,910.0      456,910.0      456,910
     456,910           0.0  cudaLaunchKernel
      0.0          285,434          3       95,144.7       51,202.0       49,726
     184,506      77,392.7  cudaMalloc
      0.0           16,719          2        8,359.5        8,359.5        3,944
      12,775       6,244.5  cudaEventRecord
      0.0           11,906          2        5,953.0        5,953.0          413
      11,493       7,834.7  cudaEventCreate
      0.0            2,783          1        2,783.0        2,783.0        2,783
       2,783           0.0  cudaEventSynchronize
      0.0            1,134          2          567.0          567.0          212
         922         502.0  cudaEventDestroy
      0.0            1,061          1        1,061.0        1,061.0        1,061
       1,061           0.0  cuModuleGetLoadingMode

[6/8] Executing 'gpukernsum' stats report

 Time (%)  Total Time (ns)  Instances   Avg (ns)     Med (ns)    Min (ns)   Max
(ns)   StdDev (ns)      GridXYZ          BlockXYZ
Name
 --------  ---------------  ---------  -----------  -----------  ---------  -----
----  -----------  ----------------  --------------  ----------------------------
----------------------------
```

```
    100.0        7,085,974          1  7,085,974.0  7,085,974.0  7,085,974
 7,085,974          0.0  409600    1    1   256    1    1  vectorAdd(const double
*, const double *, double *, int)

[7/8] Executing 'gpumemtimesum' stats report

 Time (%)  Total Time (ns)  Count    Avg (ns)      Med (ns)     Min (ns)    Max
(ns)   StdDev (ns)      Operation
 --------  ---------------  -----  ------------  ------------  ----------  ------
----  -----------  ------------------
     66.7      109,913,699      2  54,956,849.5  54,956,849.5  52,674,590
 57,239,109  3,227,602.3  [CUDA memcpy HtoD]
     33.3       54,949,107      1  54,949,107.0  54,949,107.0  54,949,107
 54,949,107          0.0  [CUDA memcpy DtoH]

[8/8] Executing 'gpumemsizesum' stats report

 Total (MB)  Count  Avg (MB)  Med (MB)  Min (MB)  Max (MB)  StdDev (MB)
Operation
 ----------  -----  --------  --------  --------  --------  -----------  --------
----------
  1,677.722      2   838.861   838.861   838.861   838.861        0.000  [CUDA
memcpy HtoD]
    838.861      1   838.861   838.861   838.861   838.861        0.000  [CUDA
memcpy DtoH]
```



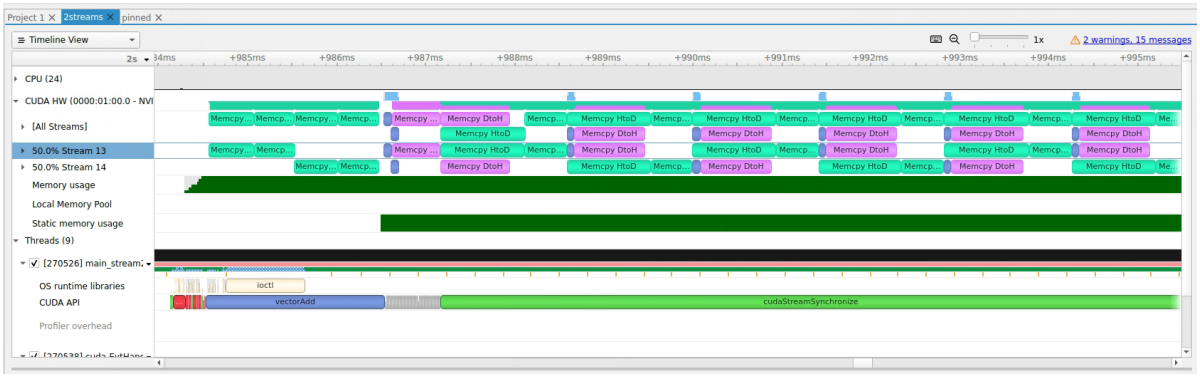
使用breadth-first 2 streams with pinned memory

```
$ nsys profile --stats=true ./main_stream2
[Using 2 streams]
[Vector addition of 104857600 elements]
Copy input data from the host memory to the CUDA device
GPU Time used:  168.2 ms
CPU Time used: 905.8 ms
Test PASSED
Done
Generating '/tmp/nsys-report-b538.qdstrm'
[1/8] [========================100%] report7.nsys-rep
[2/8] [========================100%] report7.sqlite
[3/8] Executing 'nvtxsum' stats report
SKIPPED: /home/admini/code/generalgpu/stream/build/report7.sqlite does not
contain NV Tools Extension (NVTX) data.
[4/8] Executing 'osrtsum' stats report
```

```
 Time (%)  Total Time (ns)  Num Calls     Avg (ns)         Med (ns)       Min
(ns)      Max (ns)       StdDev (ns)             Name
 --------  ---------------  ---------  ---------------  ---------------  --------
---  -------------  ---------------  ----------------------
     33.7    4,216,700,547          2  2,108,350,273.5  2,108,350,273.5
 1,792,022  4,214,908,525  2,979,123,249.2  sem_wait
     32.1    4,007,147,145         50     80,142,942.9    100,157,232.0
 1,964    100,265,535     38,740,124.4  poll
     28.0    3,501,004,838          7    500,143,548.3    500,154,704.0
 500,094,356    500,185,178         34,309.3  pthread_cond_timedwait
      5.3      659,203,703        486      1,356,386.2          5,123.0
 1,019    190,723,960     14,746,022.2  ioctl
      0.9      113,875,176         27      4,217,599.1          8,649.0
 1,068     38,445,928     12,130,762.8  mmap
      0.0          971,456          9        107,939.6        101,601.0
52,169        240,729         54,751.3  sem_timedwait
      0.0          858,116         27         31,782.1          6,438.0
 3,813        504,523         95,179.2  mmap64
      0.0          357,565         44          8,126.5          7,419.0
 1,878         19,493          3,914.3  open64
      0.0          316,729          5         63,345.8         54,978.0
41,915        102,779         24,852.5  pthread_create
      0.0          132,251         26          5,086.6          3,083.0
 1,164         45,746          8,667.5  fopen
      0.0          113,997          3         37,999.0          2,656.0
 1,433        109,908         62,278.0  read
      0.0           85,139         12          7,094.9          5,697.5
 3,732         15,034          3,687.5  munmap
      0.0           56,817          5         11,363.4         10,827.0
 2,442         22,150          8,509.9  fgets
      0.0           34,889          5          6,977.8          7,271.0
 2,761         11,161          3,522.7  fread
      0.0           21,559          5          4,311.8          2,161.0
 1,272         11,048          4,026.4  open
      0.0           12,461          3          4,153.7          1,925.0
 1,053          9,483          4,635.9  fclose
      0.0           10,621          2          5,310.5          5,310.5
 2,739          7,882          3,636.7  socket
      0.0            6,495          2          3,247.5          3,247.5
 1,026          5,469          3,141.7  fcntl
      0.0            6,287          1          6,287.0          6,287.0
 6,287          6,287              0.0  connect
      0.0            4,964          1          4,964.0          4,964.0
 4,964          4,964              0.0  pipe2
      0.0            4,697          3          1,565.7          1,560.0
 1,065          2,072            503.5  write
      0.0            2,177          1          2,177.0          2,177.0
 2,177          2,177              0.0  fopen64

[5/8] Executing 'cudaapisum' stats report

 Time (%)  Total Time (ns)  Num Calls    Avg (ns)       Med (ns)      Min (ns)
  Max (ns)     StdDev (ns)            Name
 --------  ---------------  ---------  -------------  -------------  -----------
 -----------  -------------  ----------------------
```

```
     67.5      655,500,064          3  218,500,021.3  189,408,152.0  188,918,602
 277,173,310   50,813,148.1  cudaHostAlloc
     17.0      165,402,785          2   82,701,392.5   82,701,392.5    2,006,154
 163,396,631  114,120,300.7  cudaStreamSynchronize
     15.1      146,382,491          3   48,794,163.7   48,410,921.0   48,264,711
  49,706,859      793,790.9  cudaFreeHost
      0.2        2,265,903        100       22,659.0        1,432.0        1,303
   2,117,081      211,558.0  cudaLaunchKernel
      0.1          870,543          6      145,090.5      152,592.5      106,198
     155,931       19,206.8  cudaFree
      0.0          411,664        300        1,372.2        1,251.0        1,111
      15,376          901.6  cudaMemcpyAsync
      0.0          253,447          6       42,241.2       24,633.5       20,003
     116,901       37,705.3  cudaMalloc
      0.0           41,473          2       20,736.5       20,736.5        3,078
      38,395       24,972.9  cudaStreamCreate
      0.0           25,099          2       12,549.5       12,549.5       12,413
      12,686          193.0  cudaEventRecord
      0.0           18,547          2        9,273.5        9,273.5        2,236
      16,311        9,952.5  cudaStreamDestroy
      0.0            3,962          1        3,962.0        3,962.0        3,962
       3,962            0.0  cudaEventSynchronize
      0.0            3,849          2        1,924.5        1,924.5          291
       3,558        2,310.1  cudaEventCreate
      0.0            1,601          2          800.5          800.5          278
       1,323          738.9  cudaEventDestroy
      0.0            1,549          1        1,549.0        1,549.0        1,549
       1,549            0.0  cuModuleGetLoadingMode

[6/8] Executing 'gpukernsum' stats report

 Time (%)  Total Time (ns)  Instances  Avg (ns)  Med (ns)  Min (ns)  Max (ns)
StdDev (ns)     GridXYZ         BlockXYZ                                  Name

 --------  ---------------  ---------  --------  --------  --------  --------  --
--------  -------------  -------------  ----------------------------------------
------------------
    100.0        7,949,572        100  79,495.7  79,125.0    72,182    86,420
 1,842.8  4096    1    1   256    1    1  vectorAdd(const double *, const double
*, double *, int)

[7/8] Executing 'gpumemtimesum' stats report

 Time (%)  Total Time (ns)  Count  Avg (ns)   Med (ns)   Min (ns)  Max (ns)
StdDev (ns)      Operation
 --------  ---------------  -----  ---------  ---------  --------  ---------  ---
--------  ------------------
     64.6      165,767,568    200  828,837.8  930,461.5   445,698  1,700,914
 334,854.8  [CUDA memcpy HtoD]
     35.4       90,989,003    100  909,890.0  787,554.0   596,013  1,467,986
 218,980.9  [CUDA memcpy DtoH]

[8/8] Executing 'gpumemsizesum' stats report

 Total (MB)  Count  Avg (MB)  Med (MB)  Min (MB)  Max (MB)  StdDev (MB)
Operation
```

```
 ----------  -----  --------  --------  --------  --------  -----------  --------
----------
  1,677.722    200     8.389     8.389     8.389     8.389        0.000  [CUDA
memcpy HtoD]
    838.861    100     8.389     8.389     8.389     8.389        0.000  [CUDA
memcpy DtoH]
```

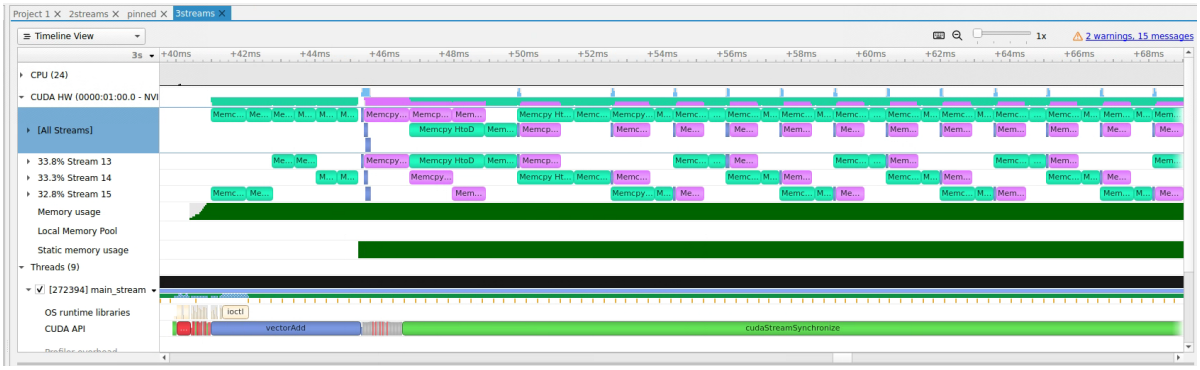


使用breadth-first 3 streams with pinned memory

```
$ nsys profile --stats=true ./main_stream
[Using 3 streams]
[Vector addition of 104857600 elements]
Copy input data from the host memory to the CUDA device
GPU Time used:  178.8 ms
CPU Time used: 896.9 ms
Test PASSED
Done
Generating '/tmp/nsys-report-a07b.qdstrm'
[1/8] [========================100%] report13.nsys-rep
[2/8] [========================100%] report13.sqlite
[3/8] Executing 'nvtxsum' stats report
SKIPPED: /home/admini/code/generalgpu/stream/build/report13.sqlite does not
contain NV Tools Extension (NVTX) data.
[4/8] Executing 'osrtsum' stats report

 Time (%)  Total Time (ns)  Num Calls     Avg (ns)         Med (ns)        Min
(ns)      Max (ns)       StdDev (ns)             Name
 --------  ---------------  ---------  ---------------  ---------------  --------
---  -------------  ---------------  ----------------------
     33.6    4,186,965,546          2  2,093,482,773.0  2,093,482,773.0
 1,742,187  4,185,223,359  2,958,167,905.7  sem_wait
     32.2    4,008,465,393         50     80,169,307.9    100,200,993.5
 1,752    100,258,175     38,841,113.4  poll
     28.1    3,500,932,635          7    500,133,233.6    500,146,474.0
 500,083,535    500,155,643         28,284.4  pthread_cond_timedwait
      5.2      650,852,838        502      1,296,519.6          5,386.0
 1,008    189,760,444     14,390,698.1  ioctl
      0.9      114,510,670         28      4,089,666.8          7,610.5
 1,249     40,036,719     12,003,406.1  mmap
      0.0          823,544         27         30,501.6          8,058.0
 3,773        476,307         89,647.8  mmap64
```

```
      0.0          710,264          9         78,918.2         78,169.0
 4,815        258,025         79,526.3  sem_timedwait
      0.0          355,792         44          8,086.2          7,275.0
 1,843         20,589          3,909.0  open64
      0.0          269,955          5         53,991.0         55,289.0
31,977         74,228         16,884.3  pthread_create
      0.0          134,231         28          4,794.0          1,887.0
 1,019         62,922         11,533.0  fopen
      0.0          123,533         14          8,823.8          6,649.5
 2,371         24,455          7,031.4  munmap
      0.0           58,341          5         11,668.2         10,388.0
 4,060         21,645          6,573.5  fread
      0.0           50,422          4         12,605.5         12,723.0
 3,010         21,966          8,324.8  fgets
      0.0           21,804          5          4,360.8          2,440.0
 1,505         10,193          3,659.8  open
      0.0           11,854          2          5,927.0          5,927.0
 3,733          8,121          3,102.8  socket
      0.0            9,904          2          4,952.0          4,952.0
 1,075          8,829          5,482.9  fclose
      0.0            9,719          2          4,859.5          4,859.5
 1,264          8,455          5,084.8  fcntl
      0.0            7,548          1          7,548.0          7,548.0
 7,548          7,548              0.0  connect
      0.0            6,291          3          2,097.0          1,918.0
 1,189          3,184          1,009.5  read
      0.0            4,977          1          4,977.0          4,977.0
 4,977          4,977              0.0  pipe2
      0.0            4,076          1          4,076.0          4,076.0
 4,076          4,076              0.0  fopen64
      0.0            3,016          2          1,508.0          1,508.0
 1,283          1,733            318.2  write
      0.0            1,002          1          1,002.0          1,002.0
 1,002          1,002              0.0  bind

[5/8] Executing 'cudaapisum' stats report

 Time (%)  Total Time (ns)  Num Calls    Avg (ns)       Med (ns)      Min (ns)
  Max (ns)    StdDev (ns)            Name
 --------  ---------------  ---------  -------------  -------------  -----------
 -----------  ------------  ----------------------
     66.1      643,273,511          3  214,424,503.7  193,410,487.0  186,957,316
 262,905,708  42,109,752.2  cudaHostAlloc
     17.8      173,154,857          3   57,718,285.7          978.0          402
 173,153,477  99,969,808.2  cudaStreamSynchronize
     15.3      148,948,984          3   49,649,661.3   48,677,567.0   47,536,326
  52,735,091   2,732,309.4  cudaFreeHost
      0.5        5,011,884        100       50,118.8        1,696.5        1,334
   4,835,132     483,334.9  cudaLaunchKernel
      0.1        1,340,812          9      148,979.1      152,502.0      117,379
     164,148      13,399.9  cudaFree
      0.1          757,061          9       84,117.9       37,770.0       28,239
     388,126     117,024.6  cudaMalloc
      0.1          515,217        300        1,717.4        1,443.0        1,185
      40,155       2,402.7  cudaMemcpyAsync
```

```
      0.0          143,010          3       47,670.0        8,063.0        2,199
     132,748      73,738.0  cudaStreamCreate
      0.0           43,083          2       21,541.5       21,541.5        5,216
      37,867      23,087.7  cudaEventRecord
      0.0           28,084          3        9,361.3        2,505.0        1,342
      24,237      12,895.8  cudaStreamDestroy
      0.0           11,183          2        5,591.5        5,591.5          304
      10,879       7,477.7  cudaEventCreate
      0.0            2,973          1        2,973.0        2,973.0        2,973
       2,973           0.0  cudaEventSynchronize
      0.0            1,056          2          528.0          528.0          220
         836         435.6  cudaEventDestroy
      0.0              954          1          954.0          954.0          954
         954           0.0  cuModuleGetLoadingMode

[6/8] Executing 'gpukernsum' stats report

 Time (%)  Total Time (ns)  Instances  Avg (ns)  Med (ns)  Min (ns)  Max (ns)
StdDev (ns)     GridXYZ         BlockXYZ                                  Name

 --------  ---------------  ---------  --------  --------  --------  --------  --
---------  -------------  --------------  ----------------------------------------
------------------
    100.0        8,070,334        100  80,703.3  79,731.0    77,171    94,929
 2,560.5  4096    1    1   256    1    1  vectorAdd(const double *, const double
*, double *, int)

[7/8] Executing 'gpumemtimesum' stats report

 Time (%)  Total Time (ns)  Count  Avg (ns)   Med (ns)   Min (ns)  Max (ns)
StdDev (ns)      Operation
 --------  ---------------  -----  ---------  ---------  --------  ---------  ---
--------  ------------------
     66.3      176,471,298    200  882,356.5  978,240.0   448,119  2,578,555
 355,294.3  [CUDA memcpy HtoD]
     33.7       89,836,346    100  898,363.5  789,983.0   464,692  1,461,233
 200,773.8  [CUDA memcpy DtoH]

[8/8] Executing 'gpumemsizesum' stats report

 Total (MB)  Count  Avg (MB)  Med (MB)  Min (MB)  Max (MB)  StdDev (MB)
Operation
 ----------  -----  --------  --------  --------  --------  -----------  --------
----------
  1,677.722    200     8.389     8.389     8.389     8.389        0.000  [CUDA
memcpy HtoD]
    838.861    100     8.389     8.389     8.389     8.389        0.000  [CUDA
memcpy DtoH]
```



使用depth-first 2 streams with pinned memory