GPU Bench

GPU Performance Details: Tesla M2075

Contents:

System Configuration

⚠ **Note that this is previously stored data and does not reflect your system configuration.**

MATLAB Release: R2016a

Host

Name	Intel(R) Xeon(R) CPU E5-2660 0 @ 2.20GHz
Clock	2201 MHz
Cache	2048 KB
NumProcessors	16
OSType	Windows
OSVersion	Microsoft Windows 7 Enterprise

GPU

Name	Tesla M2075
Clock	1147 MHz
NumProcessors	14
ComputeCapability	2.0
TotalMemory	5.25 GB
CUDAVersion	7.5
DriverVersion	8.17.13.5390 (353.90)

Results for MTimes (double)

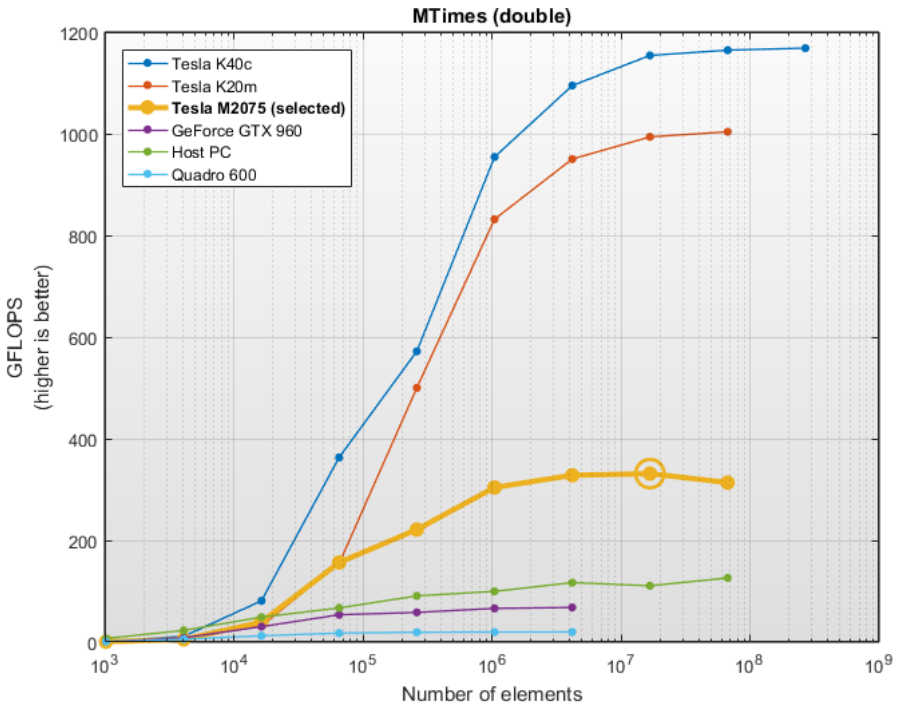
These results show the performance of the GPU or host PC when calculating a matrix multiplication of two NxN real matrices. The number of operation assumed to be `2*N^3 - N^2`.

This calculation is usually compute-bound, i.e. the performance depends mainly on how fast the GPU or host PC can perform floating-point operations.

Raw data for Tesla M2075 - MTimes (double)

Array size (elements)	Num Operations	Time (ms)	GigaFLOPS
1,024	64,512	0.11	0.57
4,096	520,192	0.10	5.13
16,384	4,177,920	0.11	39.37
65,536	33,488,896	0.21	156.41
262,144	268,173,312	1.21	221.35
1,048,576	2,146,435,072	7.06	304.08
4,194,304	17,175,674,880	52.34	328.12
16,777,216	137,422,176,256	414.74	331.34
67,108,864	1,099,444,518,912	3505.54	313.63

(N gigaflops = $Nx10^9$ operations per second)



Results for Backslash (double)

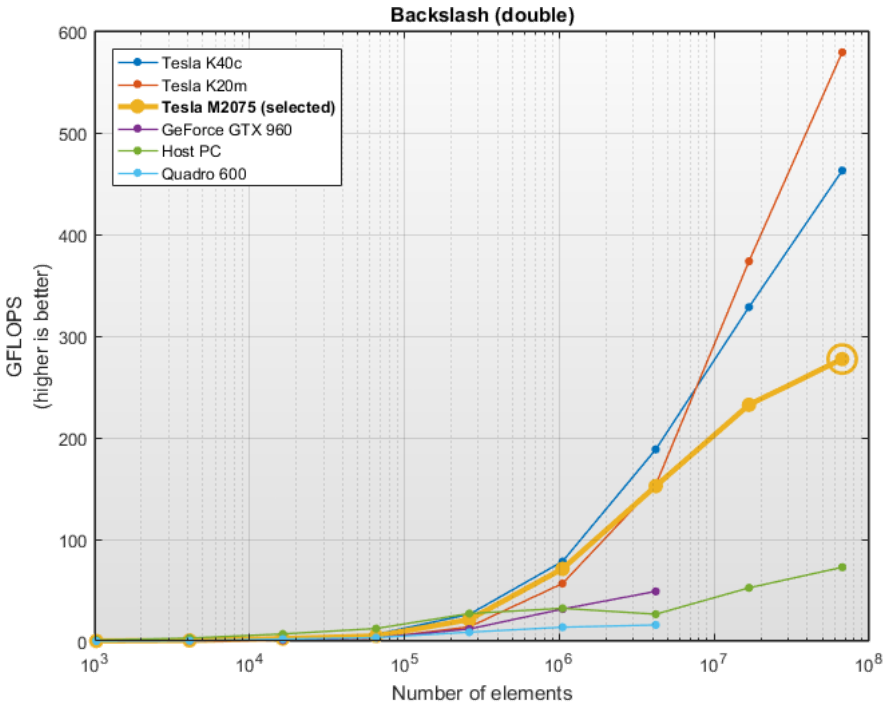
These results show the performance of the GPU or host PC when calculating the matrix left division of an `NxN` matrix with an `Nx1` vector. The number o operations is assumed to be `2/3*N^3 + 3/2*N^2`.

This calculation is usually compute-bound, i.e. the performance depends mainly on how fast the GPU or host PC can perform floating-point operations.

Raw data for Tesla M2075 - Backslash (double)

Array size (elements)	Num Operations	Time (ms)	GigaFLOPS
1,024	23,381	0.23	0.10
4,096	180,907	0.31	0.58
16,384	1,422,677	0.60	2.38
65,536	11,283,115	2.24	5.04
262,144	89,871,701	4.25	21.13
1,048,576	717,400,747	10.10	71.02
4,194,304	5,732,914,517	37.61	152.43
16,777,216	45,838,150,315	197.26	232.38
67,108,864	366,604,539,221	1321.71	277.37

(N gigaflops = Nx10^9 operations per second)



Results for FFT (double)

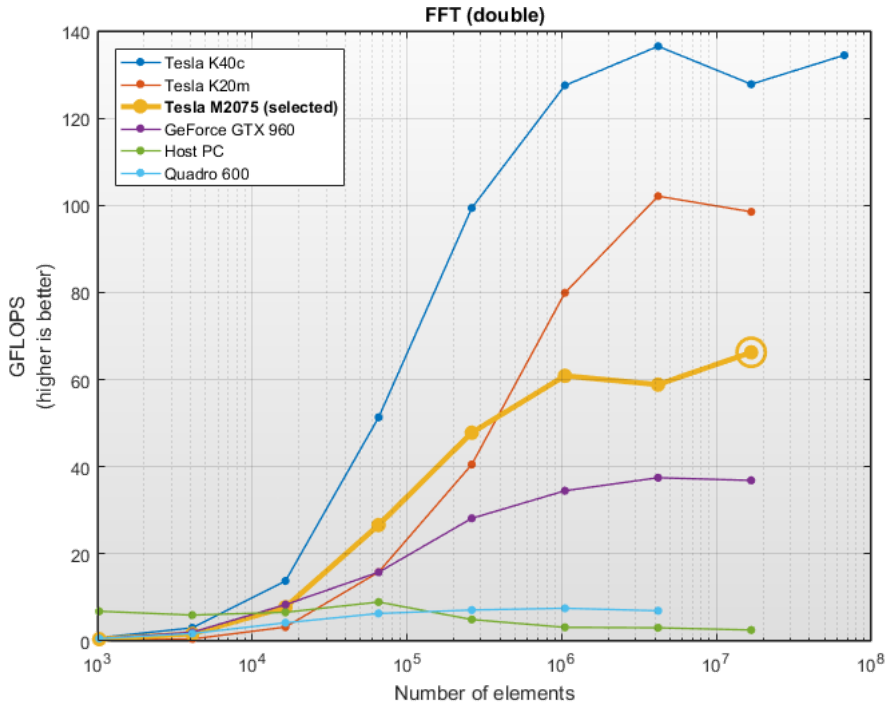
These results show the performance of the GPU or host PC when calculating the Fast-Fourier-Transform of a vector of complex numbers. The number operations for a vector of length `N` is assumed to be `5*N*log2(N)`.

This calculation is usually memory-bound, i.e. the performance depends mainly on how fast the GPU or host PC can read and write data.

Raw data for Tesla M2075 - FFT (double)

Array size (elements)	Num Operations	Time (ms)	GigaFLOPS
1,024	51,200	0.12	0.43
4,096	245,760	0.15	1.67
16,384	1,146,880	0.15	7.77
65,536	5,242,880	0.20	26.63
262,144	23,592,960	0.49	47.82
1,048,576	104,857,600	1.72	60.87
4,194,304	461,373,440	7.84	58.86
16,777,216	2,013,265,920	30.38	66.27

(N gigaflops = Nx10^9 operations per second)



Results for MTimes (single)

These results show the performance of the GPU or host PC when calculating a matrix multiplication of two NxN real matrices. The number of operation assumed to be `2*N^3 - N^2`.

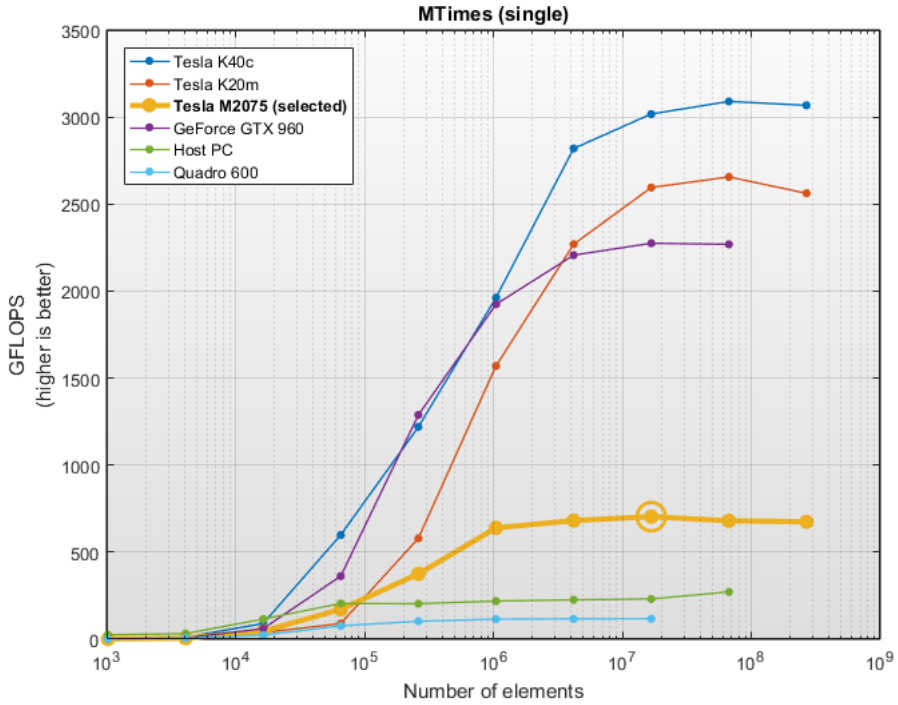
This calculation is usually compute-bound, i.e. the performance depends mainly on how fast the GPU or host PC can perform floating-point operations.

Raw data for Tesla M2075 - MTimes (single)

Array size (elements)	Num Operations	Time (ms)	GigaFLOPS
1,024	64,512	0.10	0.62
4,096	520,192	0.10	5.30
16,384	4,177,920	0.09	45.33

65,536	33,488,896	0.20	171.72
262,144	268,173,312	0.72	374.87
1,048,576	2,146,435,072	3.36	639.35
4,194,304	17,175,674,880	25.21	681.40
16,777,216	137,422,176,256	195.26	703.80
67,108,864	1,099,444,518,912	1616.43	680.17
268,435,456	8,795,824,586,752	13035.71	674.75

(N gigaflops = Nx10^9 operations per second)



Results for Backslash (single)

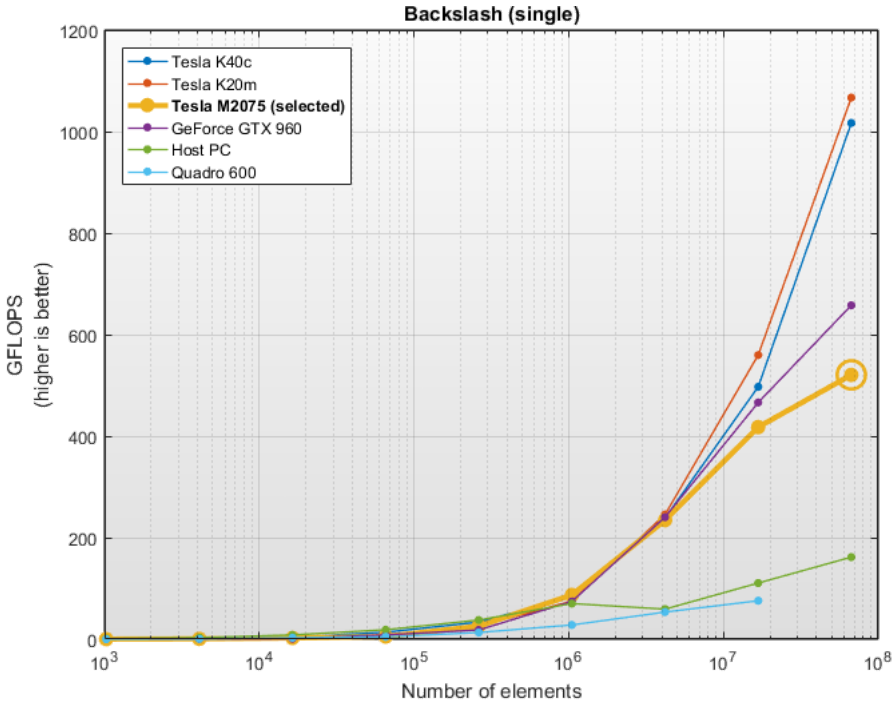
These results show the performance of the GPU or host PC when calculating the matrix left division of an NxN matrix with an Nx1 vector. The number o operations is assumed to be 2/3*N^3 + 3/2*N^2.

This calculation is usually compute-bound, i.e. the performance depends mainly on how fast the GPU or host PC can perform floating-point operations.

Raw data for Tesla M2075 - Backslash (single)

Array size (elements)	Num Operations	Time (ms)	GigaFLOPS
1,024	23,381	0.23	0.10
4,096	180,907	0.28	0.65
16,384	1,422,677	0.59	2.42
65,536	11,283,115	2.22	5.09
262,144	89,871,701	3.49	25.73
1,048,576	717,400,747	8.23	87.15
4,194,304	5,732,914,517	24.45	234.48
16,777,216	45,838,150,315	109.72	417.77
67,108,864	366,604,539,221	704.08	520.69

(N gigaflops = Nx10^9 operations per second)



Results for FFT (single)

These results show the performance of the GPU or host PC when calculating the Fast-Fourier-Transform of a vector of complex numbers. The number operations for a vector of length N is assumed to be 5*N*log2(N).

This calculation is usually memory-bound, i.e. the performance depends mainly on how fast the GPU or host PC can read and write data.

Raw data for Tesla M2075 - FFT (single)

Array size (elements)	Num Operations	Time (ms)	GigaFLOPS
1,024	51,200	0.10	0.53
4,096	245,760	0.10	2.55
16,384	1,146,880	0.15	7.90
65,536	5,242,880	0.15	35.09
262,144	23,592,960	0.33	70.70

1,048,576	104,857,600	0.84	124.33
4,194,304	461,373,440	3.20	144.18
16,777,216	2,013,265,920	12.30	163.71
67,108,864	8,724,152,320	87.85	99.31

(N gigaflops = $Nx10^9$ operations per second)