General Information

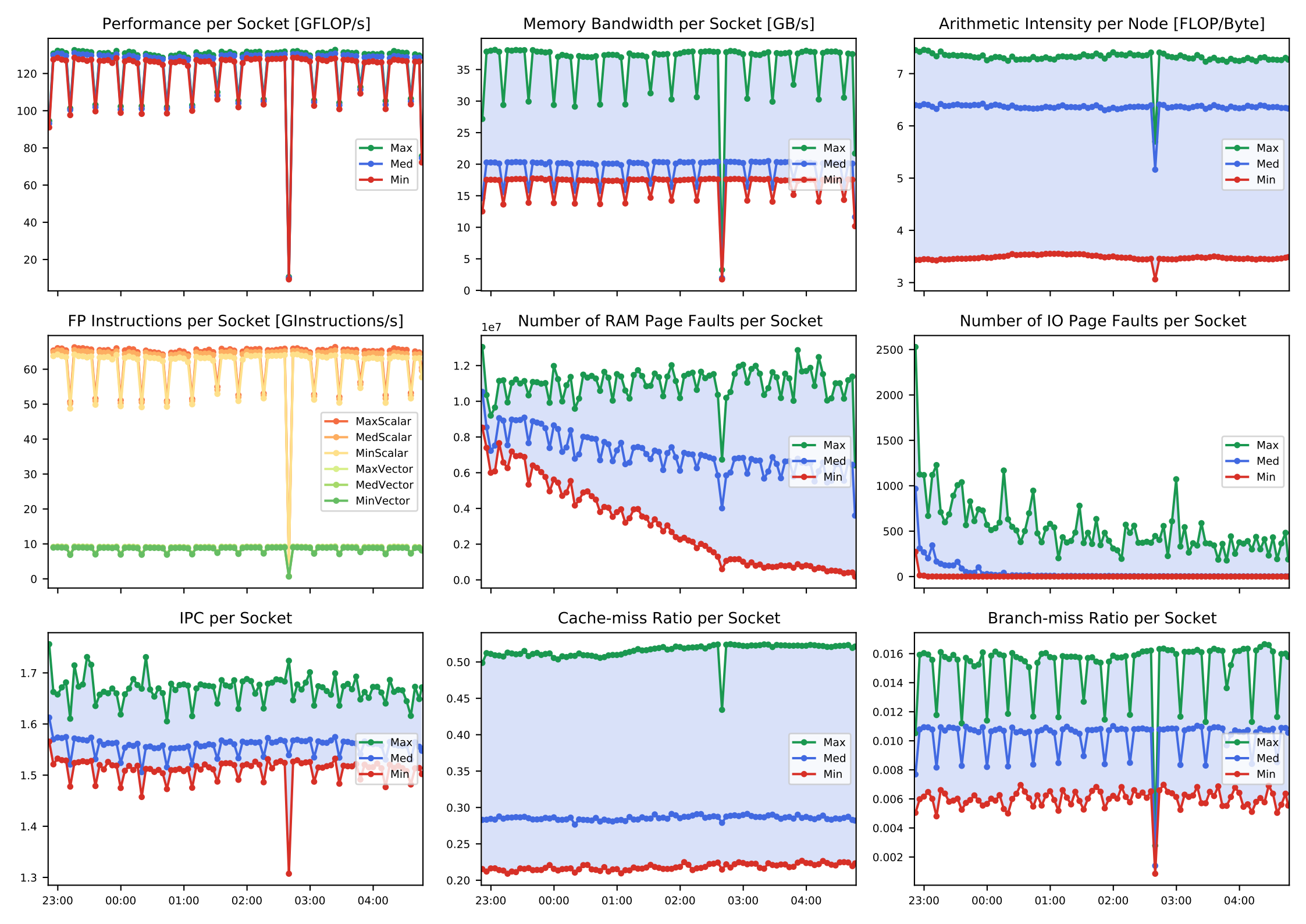
jobid	userid	groupid	partition	exit_code	jobname	exe
9777446	nkumar	mkp(10400)	general	143:0	MA30mi100-2D	smilei

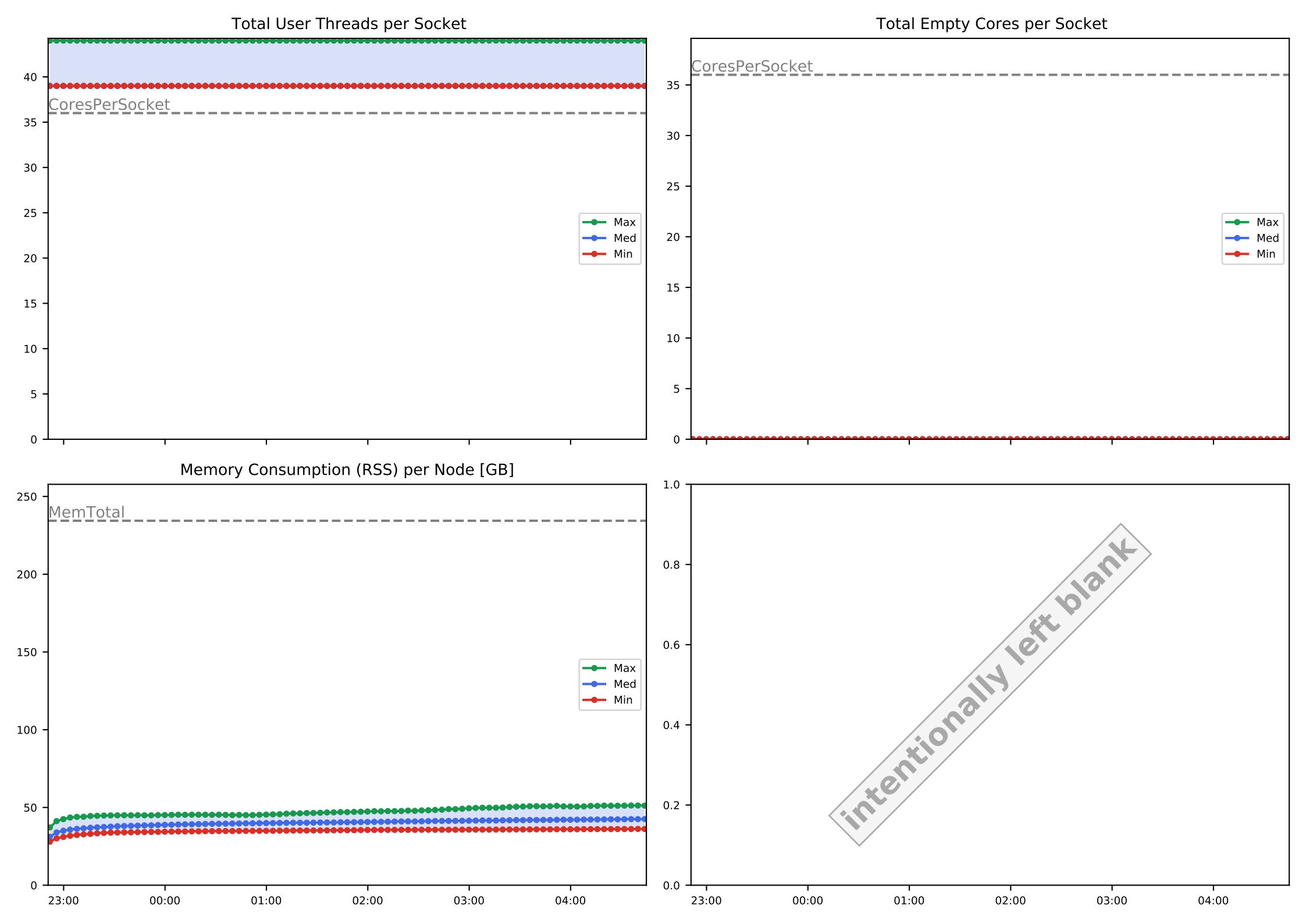
Timings

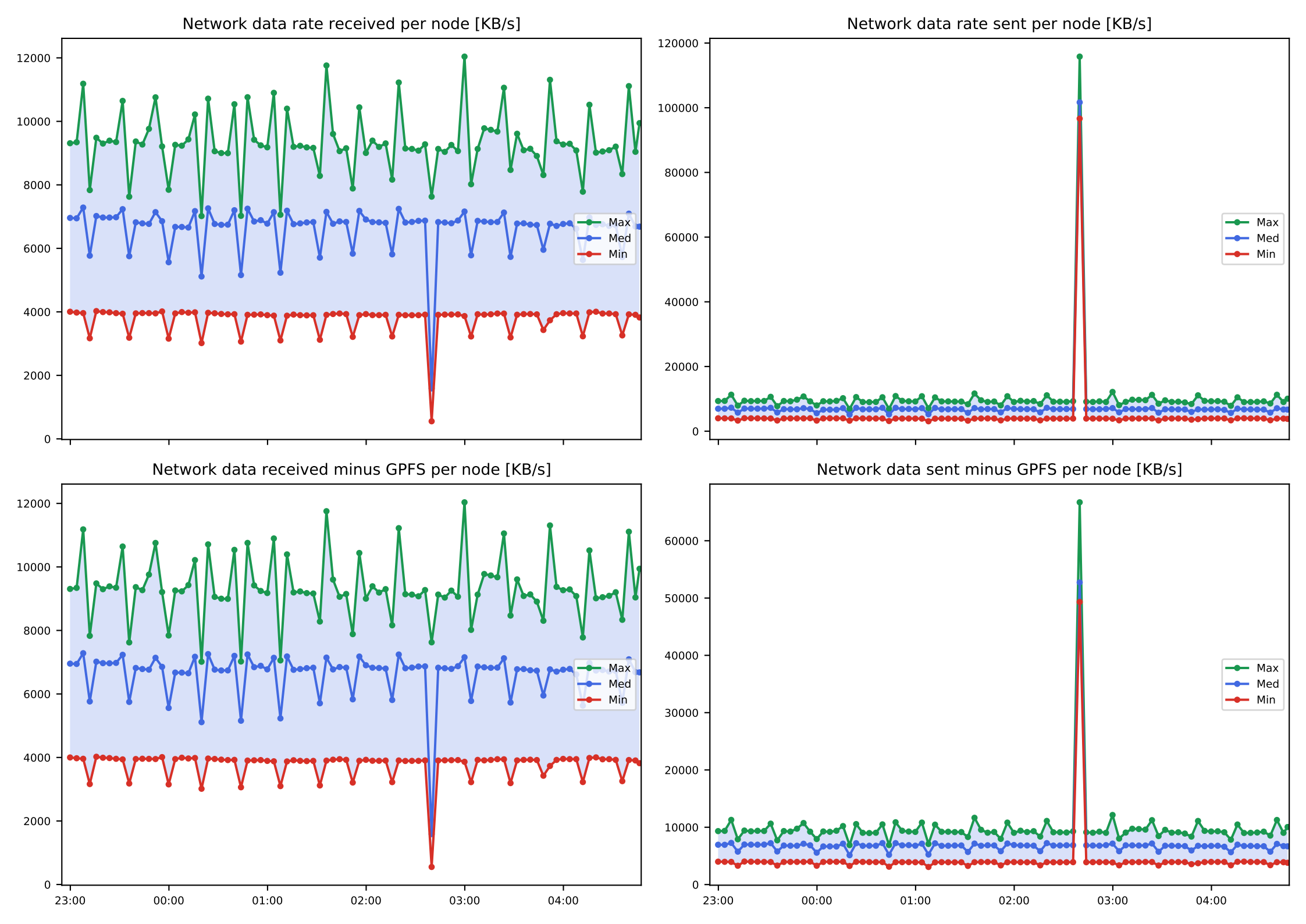
first_sample	last_sample	jobstart	jobend	wallclock	timelimit	core_hours
2024-03-19 22:48:00	2024-03-20 04:46:27	2024-03-19 22:44:47	2024-03-20 04:46:17	06:01:29	1-00:00:00	111048.21

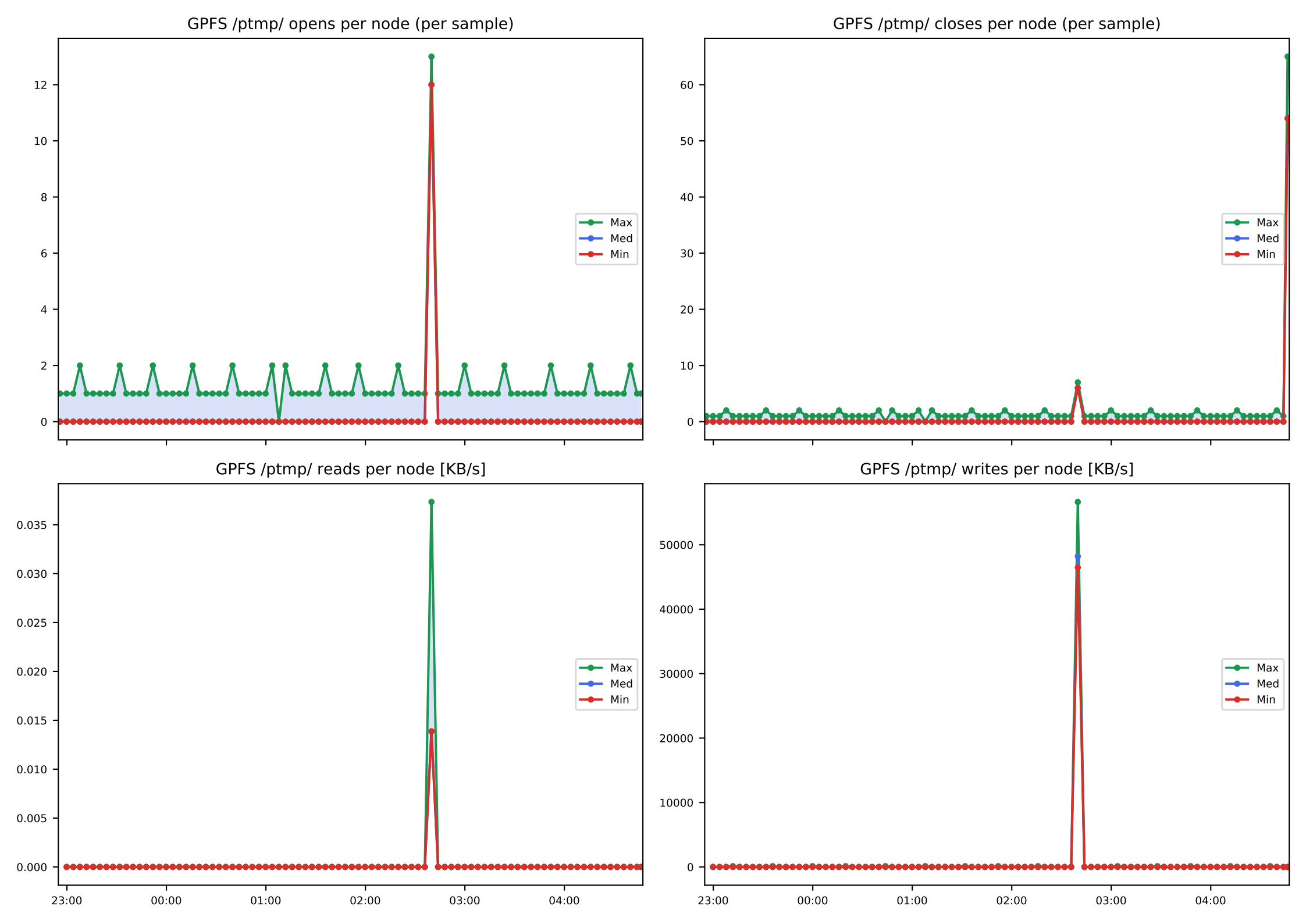
Node Information

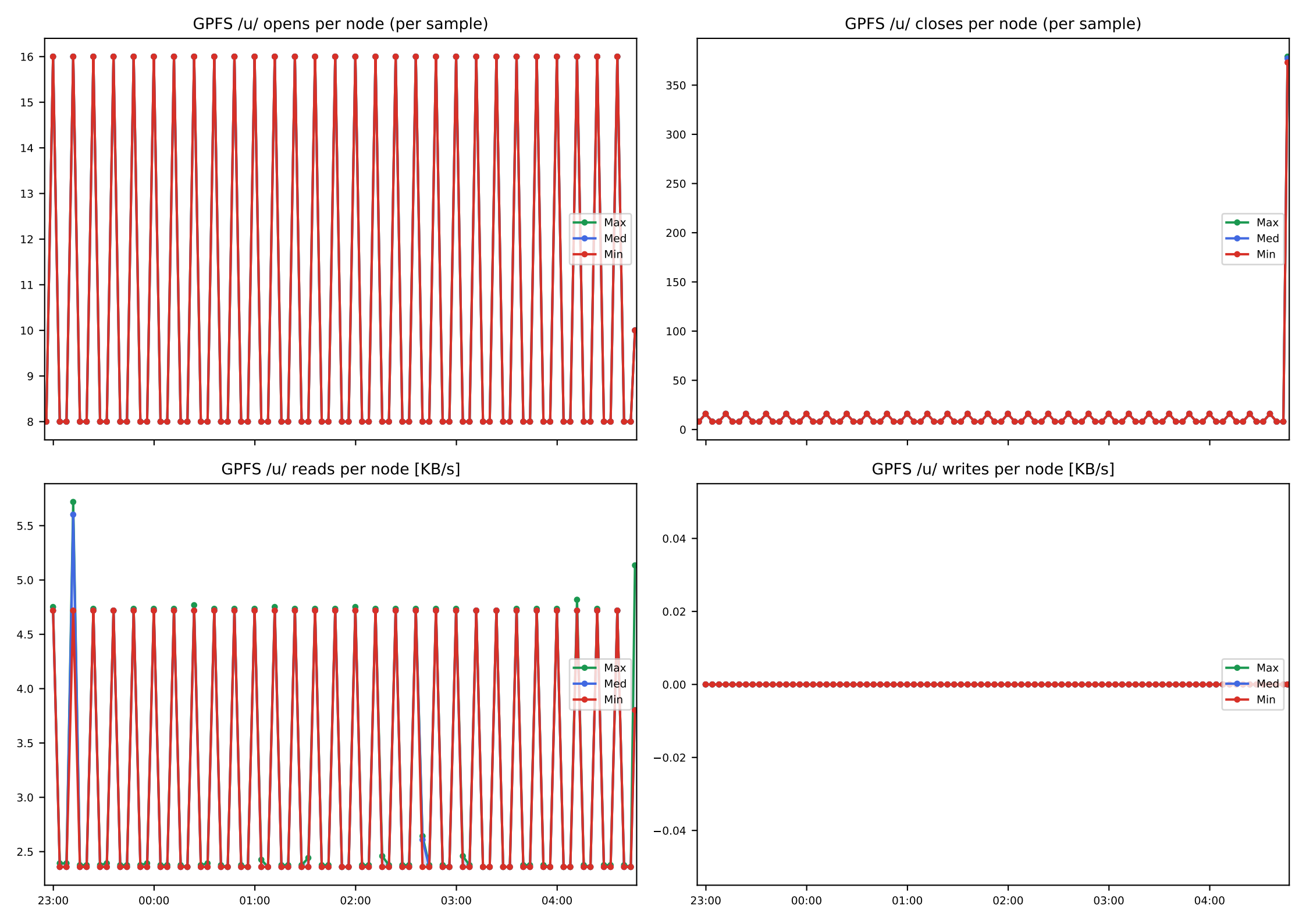
nnodes	sockets_per_node	cores_per_node	threads_per_core	ntasks_per_node	njobsteps	memory_per_node_GB
256.0	2.0	72.0	2.0	6.0	1.0	234.38


Performance per Socket [GFLOP/s]
Memory Bandwidth per Socket [GB/s]
Arithmetic Intensity per Node [FLOP/Byte]
FP Instructions per Socket [GInstructions/s]
Number of RAM Page Faults per Socket
1e7
Number of IO Page Faults per Socket
IPC per Socket
Cache-miss Ratio per Socket
Branch-miss Ratio per Socket
Max
Med
Min
MaxScalar
MedScalar
MinScalar
MaxVector
MedVector
MinVector
23:00
00:00
01:00
02:00
03:00
04:00


Total User Threads per Socket
CoresPerSocket
Max
Med
Min
Total Empty Cores per Socket
CoresPerSocket
Max
Med
Min
Memory Consumption (RSS) per Node [GB]
MemTotal
Max
Med
Min
intentionally left blank
23:00
00:00
01:00
02:00
03:00
04:00


Network data rate received per node [KB/s]
Network data rate sent per node [KB/s]
Network data received minus GPFS per node [KB/s]
Network data sent minus GPFS per node [KB/s]
Max
Med
Min
0
2000
4000
6000
8000
10000
12000
20000
40000
60000
80000
100000
120000
30000
50000
23:00
00:00
01:00
02:00
03:00
04:00


GPFS /ptmp/ opens per node (per sample)
GPFS /ptmp/ closes per node (per sample)
GPFS /ptmp/ reads per node [KB/s]
GPFS /ptmp/ writes per node [KB/s]
Max
Med
Min
0
2
4
6
8
10
12
0
10
20
30
40
50
60
0.000
0.005
0.010
0.015
0.020
0.025
0.030
0.035
0
10000
20000
30000
40000
50000
23:00
00:00
01:00
02:00
03:00
04:00


GPFS /u/ opens per node (per sample)
GPFS /u/ closes per node (per sample)
GPFS /u/ reads per node [KB/s]
GPFS /u/ writes per node [KB/s]
Max
Med
Min
23:00
00:00
01:00
02:00
03:00
04:00

Explanation of the tables and plots

Timelines are max-median-min of all the samples for each timestamp for either each socket of a node (N#S#) or for each node (N#)

P1. General Information: table containing general information of the job, derived from jobid. 'exe' field represents the most frequent executable name during the job duration detected by 'ps' tool. Note that each sample reports a single value belonging to the executable which used most of the cputime.

P1. Timings: most important timings regarding the job. 'timelimit' field is the reservation time, while core hours are computed as ('wallclock' * 'cores' * 'nnodes') / 1_hour.

P1. (if GPU available) GPU Information: technical information about the GPU device 0 of the node.

P1. Node Information: information about the number of nodes, sockets, cores and threads, mostly retrieved from the SLURM submission script. 'njobsteps' field is how many SLURM steps were executed inside the job(how many 'srun' commands').

P1. (only available on Broadwell and newer CPUs, for jobs with smaller number of nodes) Performance of Each Socket: for each socket of a node, showing the maximum and median values of the GFLOPS, memory bandwidth(CPU - RAM bandwidth collected from uncore events), and arithmetic intensity(GFLOPS / Memory_Bandwidth). Arithmetic intensity is the same(avg) for all sockets of a node, as it is hard to know which memory access belongs to whom.

P2. (only available on Broadwell and newer CPUs) Performance per socket[GFLOP / s]: Sum of all floating point operations from perf events 'fp_*', divided by the measuring time.

P2. Memory Bandwidth per socket[GB / s]: CPU - RAM memory bandwidth collected from perf 'uncore' events. Be aware of the potential NUMA effects.

P2. (only available on Broadwell and newer CPUs) Arithmetic Intensity per node[FLOP / Byte]: GFLOPS / Memory_Bandwidth.

P2. (only available on Broadwell and newer CPUs) FP instructions for average socket[GInstructions / s]: The rate of the total amount of scalar('fp_s' and 'fp_d') and vectorized instructions.

P2. Number of RAM Page Faults per socket: Collected from perf event 'minor_faults'.

P2. Number of IO Page Faults per socket: Collected from perf event 'major_faults'.

P2. IPC per socket: Instructions Per Cycle, computed by dividing perf event 'instructions' with perf event 'cycles'.

P2. Cache - miss ratio per socket: Computed by dividing perf event 'cache-misses' with perf event 'cache-references'.

P2. Branch - miss ratio per socket: Computed by dividing perf event 'branch-misses' with perf event 'branches'.

P3. Total user threads per socket: All user threads reported by 'ps', so this number is typically slightly bigger than the number of application threads.

P3. Total empty cores per socket: Number of cores which do not run any of the user threads reported by 'ps'.

P3. Memory consumption (RSS) per node[GB]: Resident - Set - Size(RSS) occupancy of the user processes, reported by 'ps'. Be aware of what RSS actually takes into account and what it doesn't.

P4. Network data rate received per node[KB / s]: Received data in KB between two samples, divided by the sampling interval. Values are either read from InfiniBand or OmniPath counters.

P4. Network data rate sent per node[KB / s]: Sent data in KB between two samples, divided by the sampling interval. Values are either read from InfiniBand or OmniPath counters.

P4. Network data received minus GPFS per node[KB / s]: Received data in KB between two samples, subtracting GPFS traffic, divided by the sampling interval. Roughly MPI data received.

P4. Network data sent minus GPFS per node[KB / s]: Sent data in KB between two samples, subtracting GPFS traffic, divided by the sampling interval. Roughly MPI data sent.

P5 - P6. GPFS /ptmp/ | /u/ opens per node(per sample): Number of opens between two samples. Collected from mmpmon event 'opens'.

P5 - P6. GPFS /ptmp/ | /u/ closes per node(per sample): Number of closes between two samples. Collected from mmpmon event 'closes'.

P5 - P6. GPFS /ptmp/ | /u/ reads per node[KB / s]: Read data in KB between two samples, divided by the sampling interval. Collected from mmpmon event 'reads'.

P5 - P6. GPFS /ptmp/ | /u/ writes per node[KB / s]: Written data in KB between two samples, divided by the sampling interval. Collected from mmpmon event 'writes'.

P7. (if GPU available) GPU utilization per device: GPU utilization reported at the specified timestamp by 'nvidia-smi'. Value 1 represents 100 % utilization.

P7. (if GPU available) GPU memory per device[GB]: Memory occupancy of the GPU, reported at the specified timestamp by 'nvidia-smi'.

P7. (if GPU available) GPU sm-occupancy: The fraction of resident warps on a multiprocessor, relative to the maximum number of concurrent warps supported.

P7. (if GPU available) GPU sm-active: The fraction of time at least one warp was active on a multiprocessor, averaged over all multiprocessors.

P7. (if GPU available) GPU dram-active: The fraction of cycles where data was sent to or received from device memory.

P7. (if GPU available) GPU fp32-active: The fraction of cycles the FMA (FP32 (single precision), and integer) pipe was active.

P7. (if GPU available) GPU fp64-active: The fraction of cycles the FP64 (double precision) pipe was active.

P7. (if GPU available) GPU tensor-active: The fraction of cycles the tensor (HMMA / IMMA) pipe was active.