

PDC Project

Parallel Algorithms for Butterfly Computations

PERFORMANCE ANALYSIS

Tool Used : TAU

Hamdan Sajid	Muhammad Mujtaba	Maria Naeem
22i-0872	22i-2123	22i-0812

Section K

Github: <https://github.com/mmujtaba0085/parallel-butterfly-counter>

Execution Time & Speedup (MPI + OpenMP):

Dataset #1: Bipartite_Graph_50k

Serial Execution Time: 7.09174 seconds

Processes	Parallel (s)	Speedup
2	4.02249	1.76
4	3.49078	2.03
8	6.57926	1.07

- Best speedup at 4 processes (speedup 2.03).
- When increasing from 4 to 8 processes, speedup drops to 1.07 — meaning performance worsens with more processes.
- This suggests that for 50k data size, the overhead (like communication and synchronization between processes) start dominating when using too many processes (8 in this case).

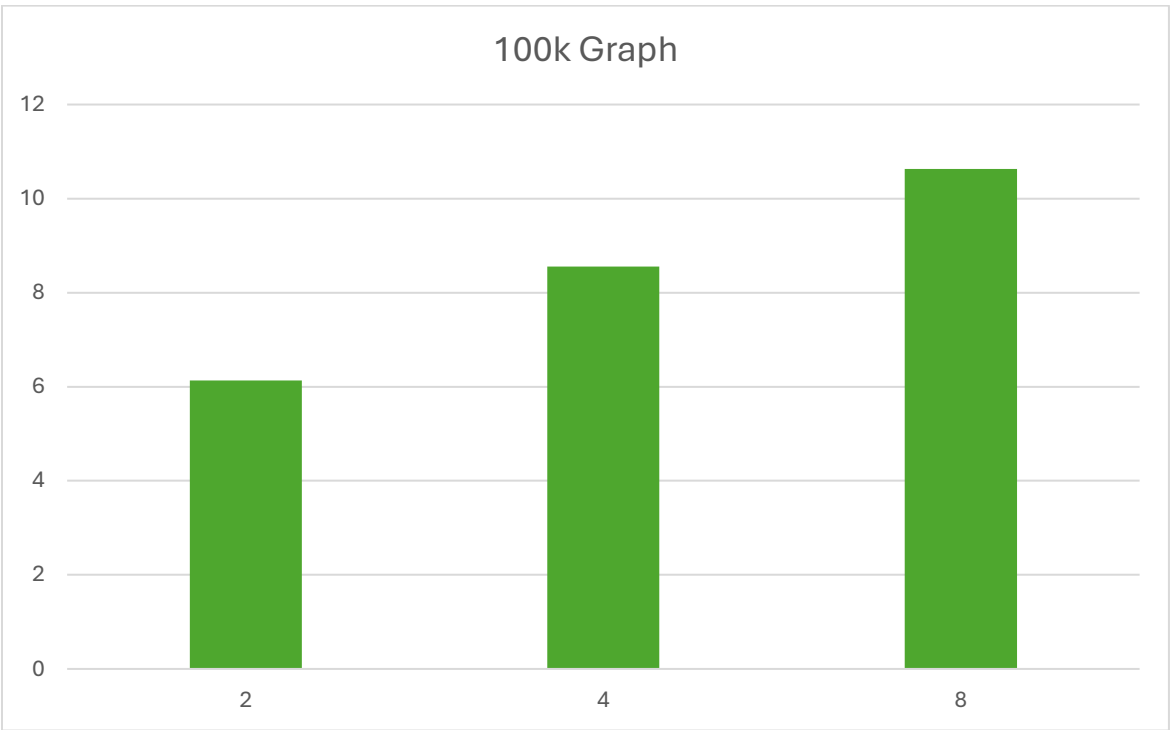
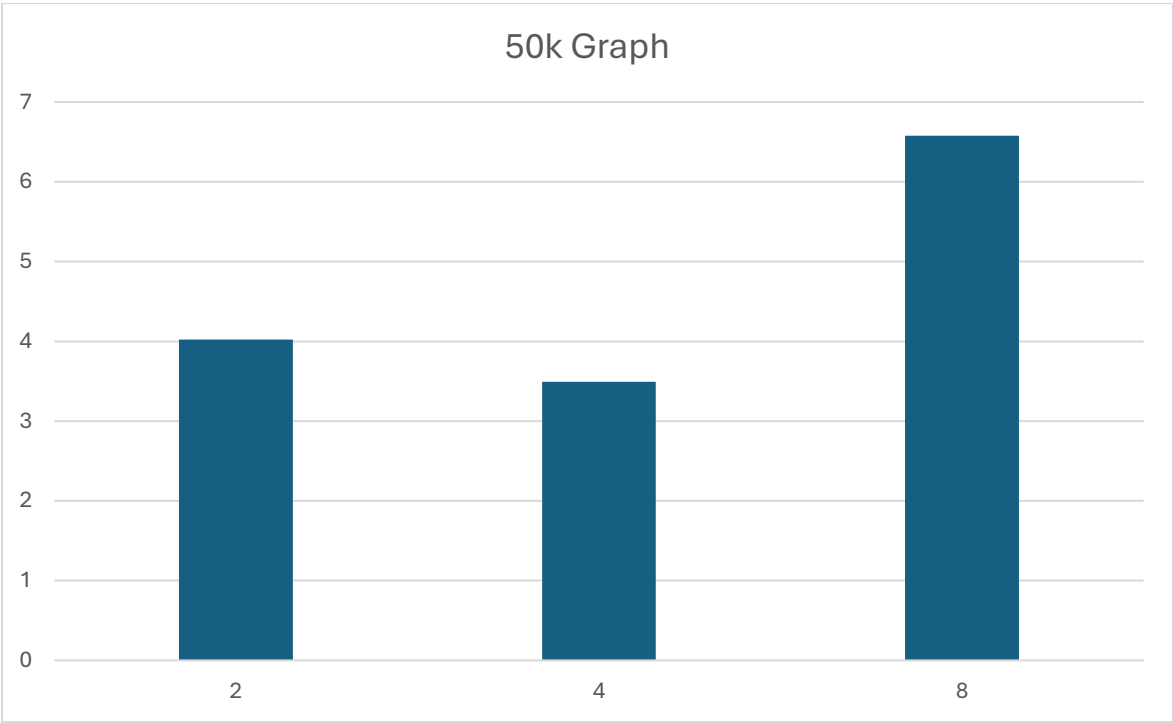
Dataset #2: Bipartite_Graph_100k

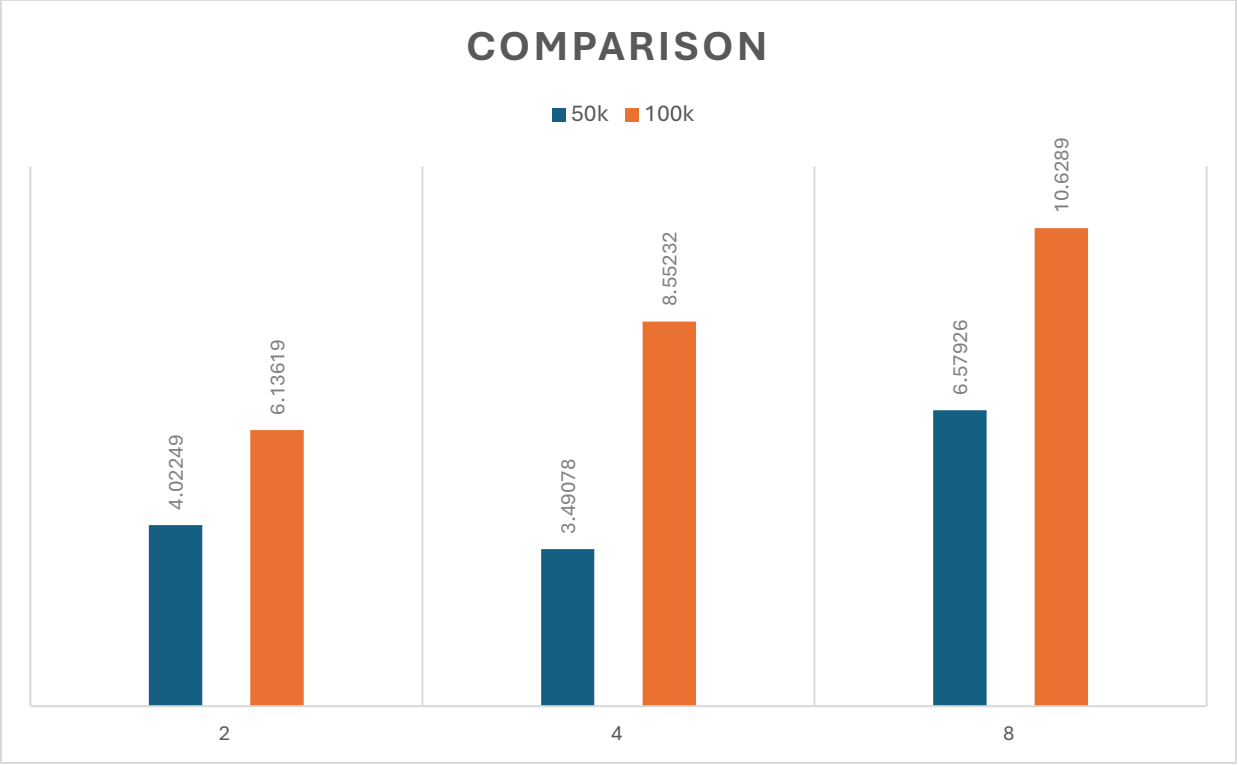
Serial Execution Time: 19.4348 seconds

Processes	Parallel (s)	Speedup
2	6.13619	3.16
4	8.55232	2.27
8	10.6289	1.82

- Best speedup at 2 processes (speedup 3.16) — excellent gain here.
- As you increase to 4 and then 8 processes, speedup **drops** (2.27 and 1.82 respectively).
- So even with a larger dataset (100k), adding more than 2 processes leads to diminishing returns — likely due to overhead again.

Visualized Performance:





Screenshots (With Profiling):

Serial

```
hamdan@main: ~/Unl/PDC/project/parallel-butterfly-counter/Serial
hamdan@main:~/Unl/PDC/project/parallel-butterfly-counter/Serial$ ./serial ../datasets/bipartite_graph_50k.txt
running
Loading graph from file: ../datasets/bipartite_graph_50k.txt
Bipartite Graph Statistics:
Left vertices: 25000
Right vertices: 25000
Weighted: Yes
Total edges: 249324
Butterfly count: 21920
Execution time: 7.01974 seconds
hamdan@main:~/Unl/PDC/project/parallel-butterfly-counter/Serial$ ./serial ../datasets/bipartite_graph_100k.txt
running
Loading graph from file: ../datasets/bipartite_graph_100k.txt
Bipartite Graph Statistics:
Left vertices: 50000
Right vertices: 50000
Weighted: Yes
Total edges: 498937
Butterfly count: 43163
Execution time: 19.4348 seconds
hamdan@main:~/Unl/PDC/project/parallel-butterfly-counter/Serial$
```

Parallel:

50k, 2 Processes

```
Loading graph from file: parallel-butterfly-counter/datasets/bipartite_graph_50k.txt
Bipartite Graph Statistics:
Left vertices: 25000
Right vertices: 25000
Weighted: Yes
Total edges: 249324
MPI distribution: 12500 vertices per process (+0 remainder)
Each process finished counting local wedges with OpenMP parallelism
Wedge counting time: 0.607069 seconds
Wedge pairs that could form butterflies locally: 1485/320146 (0.463851%)
Communication time: 2.35065 seconds
Butterfly counting time: 0.271449 seconds
Butterfly count: 21920
Total hybrid parallel execution time: 4.02249 seconds
Showing performance analysis with pprof...
Reading Profile files in profile.*

NODE 0;CONTEXT 0;THREAD 0:
-----
%Time   Exclusive   Inclusive   #Call   #Subrs   Inclusive Name
      msec    total msec
-----
100.0    2,098      21,600        1        8   21600270 .TAU application
80.0    17,276     17,276        1        0   17276398 MPI_Init_thread()
 7.2     1,544     1,544        2        0   772278 MPI_Recv()
 3.0       16        655        2        2   327652 MPI_Bcast()
 3.0       639        639        2        0   319564 MPI_Collective Sync
 0.1        25        25        1        0   25237 MPI_Finalize()
 0.0       0.029    0.029        1        0   29 MPI_Comm_rank()
 0.0       0.028    0.028        1        0   28 MPI_Comm_size()
-----

USER EVENTS Profile :NODE 0, CONTEXT 0, THREAD 0
-----
NumSamples  MaxValue  MinValue  MeanValue  Std. Dev.  Event Name
-----
          2    1E+05         4    5E+04    5E+04  Message size for broadcast
-----

NODE 1;CONTEXT 0;THREAD 0:
-----
%Time   Exclusive   Inclusive   #Call   #Subrs   Inclusive Name
      msec    total msec
-----
100.0    1,743      2,883        1        8   2883982 .TAU application
21.5      620        620        1        0   620943 MPI_Init_thread()
17.3      497        497        2        0   248758 MPI_Send()
 0.6       16        16        2        2   8469 MPI_Bcast()
 0.2        4        4        1        0   4634 MPI_Finalize()
 0.0       0.506    0.506        2        0   253 MPI_Collective Sync
 0.0       0.001    0.001        1        0   1 MPI_Comm_rank()
 0.0       0.001    0.001        1        0   1 MPI_Comm_size()
-----

USER EVENTS Profile :NODE 1, CONTEXT 0, THREAD 0
-----
NumSamples  MaxValue  MinValue  MeanValue  Std. Dev.  Event Name
-----
```

50k, 4 Processes

```

Loading graph from file: parallel-butterfly-counter/datasets/bipartite_graph_50k.txt
Bipartite Graph Statistics:
Left vertices: 25000
Right vertices: 25000
Weighted: Yes
Total edges: 249324
MPI distribution: 6250 vertices per process (+0 remainder)
Each process finished counting local wedges with OpenMP parallelism
Wedge counting time: 0.254914 seconds
Wedge pairs that could form butterflies locally: 148/103842 (0.142524%)
Communication time: 2.46318 seconds
Butterfly counting time: 0.229541 seconds
Butterfly count: 21920
Total hybrid parallel execution time: 3.49078 seconds
Showing performance analysis with pprof...
Reading Profile files in profile.*

```

NODE 0;CONTEXT 0;THREAD 0:

%Time	Exclusive msec	Inclusive total msec	#Call	#Subrs	Inclusive Name usec/call
100.0	3,164	12,011	1	12	12011300 .TAU application
67.5	8,108	8,108	1	0	8108150 MPI_Init_thread()
4.3	513	513	6	0	85641 MPI_Recv()
1.8	102	217	2	2	108768 MPI_Bcast()
1.0	115	115	2	0	57678 MPI_Collective Sync
0.1	6	6	1	0	6932 MPI_Finalize()
0.0	0.039	0.039	1	0	39 MPI_Comm_rank()
0.0	0.033	0.033	1	0	33 MPI_Comm_size()

USER EVENTS Profile :NODE 0, CONTEXT 0, THREAD 0

NumSamples	MaxValue	MinValue	MeanValue	Std. Dev.	Event Name
2	1E+05	4	5E+04	5E+04	Message size for broadcast

NODE 1;CONTEXT 0;THREAD 0:

%Time	Exclusive msec	Inclusive total msec	#Call	#Subrs	Inclusive Name usec/call
100.0	888	2,995	1	8	2995504 .TAU application
39.1	1,171	1,171	1	0	1171060 MPI_Finalize()
16.5	493	493	2	0	246670 MPI_Send()
10.2	305	305	1	0	305767 MPI_Init_thread()
4.6	20	136	2	2	68350 MPI_Bcast()
3.9	116	116	2	0	58137 MPI_Collective Sync
0.0	0.003	0.003	1	0	3 MPI_Comm_rank()
0.0	0.002	0.002	1	0	2 MPI_Comm_size()

USER EVENTS Profile :NODE 1, CONTEXT 0, THREAD 0

50k, 8 Processes

```
Loading graph from file: parallel-butterfly-counter/datasets/bipartite_graph_50k.txt
Bipartite Graph Statistics:
Left vertices: 25000
Right vertices: 25000
Weighted: Yes
Total edges: 249324
MPI distribution: 3125 vertices per process (+0 remainder)
Each process finished counting local wedges with OpenMP parallelism
Wedge counting time: 0.209784 seconds
Wedge pairs that could form butterflies locally: 20/34146 (0.058572%)
Communication time: 3.92859 seconds
Butterfly counting time: 0.406966 seconds
Butterfly count: 21920
Total hybrid parallel execution time: 6.57926 seconds
Showing performance analysis with pprof...
Reading Profile files in profile.*
```

NODE 0;CONTEXT 0;THREAD 0:

%Time	Exclusive msec	Inclusive total msec	#Call	#Subrs	Inclusive Name usec/call
100.0	4,487	1:17.234	1	20	77234816 .TAU application
91.1	1:10.342	1:10.342	1	0	70342032 MPI_Init_thread()
1.5	221	1,125	2	2	562580 MPI_Bcast()
1.3	981	981	14	0	70133 MPI_Recv()
1.2	904	904	2	0	452072 MPI Collective Sync
0.4	298	298	1	0	298667 MPI_Finalize()
0.0	0.019	0.019	1	0	19 MPI_Comm_rank()
0.0	0.017	0.017	1	0	17 MPI_Comm_size()

USER EVENTS Profile :NODE 0, CONTEXT 0, THREAD 0

NumSamples	MaxValue	MinValue	MeanValue	Std. Dev.	Event Name
2	1E+05	4	5E+04	5E+04	Message size for broadcast

NODE 1;CONTEXT 0;THREAD 0:

%Time	Exclusive msec	Inclusive total msec	#Call	#Subrs	Inclusive Name usec/call
100.0	615	14,085	1	8	14085296 .TAU application
78.4	11,048	11,048	1	0	11048432 MPI_Finalize()
8.1	1,134	1,134	1	0	1134298 MPI_Init_thread()
8.0	1,128	1,128	2	0	564112 MPI_Send()
1.1	69	159	2	2	79510 MPI_Bcast()
0.6	89	89	2	0	44932 MPI Collective Sync
0.0	0.002	0.002	1	0	2 MPI_Comm_rank()
0.0	0.001	0.001	1	0	1 MPI_Comm_size()

100k, 2 Processes

```
Loading graph from file: parallel-butterfly-counter/datasets/bipartite_graph_100k.txt
Bipartite Graph Statistics:
Left vertices: 50000
Right vertices: 50000
Weighted: Yes
Total edges: 497555
MPI distribution: 25000 vertices per process (+0 remainder)
Each process finished counting local wedges with OpenMP parallelism
Wedge counting time: 1.18353 seconds
Wedge pairs that could form butterflies locally: 2779/635771 (0.437107%)
Communication time: 3.64238 seconds
Butterfly counting time: 0.331469 seconds
Butterfly count: 43179
Total hybrid parallel execution time: 6.13619 seconds
Showing performance analysis with pprof...
Reading Profile files in profile.*
```

NODE 0;CONTEXT 0;THREAD 0:

%Time	Exclusive msec	Inclusive total msec	#Call	#Subrs	Inclusive Name usec/call
100.0	4,796	8,439	1	8	8439266 .TAU application
23.0	1,940	1,940	2	0	970360 MPI_Recv()
18.2	1,531	1,531	1	0	1531940 MPI_Init_thread()
1.4	120	120	1	0	120848 MPI_Finalize()
0.6	40	49	2	2	24520 MPI_Bcast()
0.1	8	8	2	0	4412 MPI Collective Sync
0.0	0.029	0.029	1	0	29 MPI_Comm_size()
0.0	0.026	0.026	1	0	26 MPI_Comm_rank()

USER EVENTS Profile :NODE 0, CONTEXT 0, THREAD 0

NumSamples	MaxValue	MinValue	MeanValue	Std. Dev.	Event Name
2	2E+05	4	1E+05	1E+05	Message size for broadcast

NODE 1;CONTEXT 0;THREAD 0:

%Time	Exclusive msec	Inclusive total msec	#Call	#Subrs	Inclusive Name usec/call
100.0	4,029	4,994	1	8	4994432 .TAU application
10.6	527	527	2	0	263840 MPI_Send()
4.5	29	225	2	2	112586 MPI_Bcast()
4.2	207	207	1	0	207727 MPI_Init_thread()
3.9	195	195	2	0	97796 MPI Collective Sync
0.1	4	4	1	0	4232 MPI_Finalize()
0.0	0.001	0.001	1	0	1 MPI_Comm_rank()
0.0	0.001	0.001	1	0	1 MPI_Comm_size()

USER EVENTS Profile :NODE 1, CONTEXT 0, THREAD 0

100k, 4 Processes

```

Each MPI process will use up to 4 OpenMP threads
Loading graph from file: parallel-butterfly-counter/datasets/bipartite_graph_100k.txt
Bipartite Graph Statistics:
Left vertices: 50000
Right vertices: 50000
Weighted: Yes
Total edges: 497555
MPI distribution: 12500 vertices per process (+0 remainder)
Each process finished counting local wedges with OpenMP parallelism
Wedge counting time: 0.39856 seconds
Wedge pairs that could form butterflies locally: 239/206716 (0.115618%)
Communication time: 6.24442 seconds
Butterfly counting time: 0.917805 seconds
Butterfly count: 43179
Total hybrid parallel execution time: 8.55232 seconds
Showing performance analysis with pprof...
Reading Profile files in profile.*

```

NODE 0;CONTEXT 0;THREAD 0:

%Time	Exclusive msec	Inclusive total msec	#Call	#Subrs	Inclusive Name usec/call
100.0	8,892	11,915	1	12	11915599 .TAU application
18.0	2,145	2,145	1	0	2145060 MPI_Init_thread()
5.8	692	692	6	0	115365 MPI_Recv()
1.0	23	118	2	2	59220 MPI_Bcast()
0.8	94	94	2	0	47361 MPI_Collective Sync
0.6	67	67	1	0	67115 MPI_Finalize()
0.0	0.026	0.026	1	0	26 MPI_Comm_size()
0.0	0.024	0.024	1	0	24 MPI_Comm_rank()

USER EVENTS Profile :NODE 0, CONTEXT 0, THREAD 0

NumSamples	MaxValue	MinValue	MeanValue	Std. Dev.	Event Name
2	2E+05	4	1E+05	1E+05	Message size for broadcast

NODE 1;CONTEXT 0;THREAD 0:

%Time	Exclusive msec	Inclusive total msec	#Call	#Subrs	Inclusive Name usec/call
100.0	1,753	5,866	1	8	5866074 .TAU application
46.7	2,742	2,742	1	0	2742300 MPI_Finalize()
10.0	585	585	2	0	292826 MPI_Send()
7.4	25	431	2	2	215722 MPI_Bcast()
6.9	406	406	2	0	203102 MPI_Collective Sync
6.0	353	353	1	0	353666 MPI_Init_thread()
0.0	0.002	0.002	1	0	2 MPI_Comm_rank()
0.0	0	0	1	0	0 MPI_Comm_size()

USER EVENTS Profile :NODE 1, CONTEXT 0, THREAD 0

100k, 8 Processes

```
loading graph from file: parallel-butterfly-counter/datasets/bipartite_graph_100k.txt
Bipartite Graph Statistics:
Left vertices: 50000
Right vertices: 50000
Weighted: Yes
Total edges: 497555
MPI distribution: 6250 vertices per process (+0 remainder)
Each process finished counting local wedges with OpenMP parallelism
Wedge counting time: 0.153988 seconds
Wedge pairs that could form butterflies locally: 17/68782 (0.0247158%)
Communication time: 7.88711 seconds
Butterfly counting time: 0.472087 seconds
Butterfly count: 43179
Total hybrid parallel execution time: 10.6289 seconds
Showing performance analysis with pprof...
Loading Profile files in profile.*
```

NODE 0;CONTEXT 0;THREAD 0:

Time	Exclusive msec	Inclusive total msec	#Call	#Subrs	Inclusive usec/call	Name
100.0	9,372	14,376	1	20	14376703	.TAU application
14.8	2,131	2,131	1	0	2131235	MPI_Init_thread()
13.8	1,990	1,990	14	0	142209	MPI_Recv()
6.0	730	860	2	2	430429	MPI_Bcast()
0.9	129	129	2	0	64932	MPI_Collective Sync
0.1	20	20	1	0	20322	MPI_Finalize()
0.0	0.584	0.584	1	0	584	MPI_Comm_size()
0.0	0.03	0.03	1	0	30	MPI_Comm_rank()

USER EVENTS Profile :NODE 0, CONTEXT 0, THREAD 0

NumSamples	MaxValue	MinValue	MeanValue	Std. Dev.	Event Name
2	2E+05	4	1E+05	1E+05	Message size for broadcast

NODE 1;CONTEXT 0;THREAD 0:

Time	Exclusive msec	Inclusive total msec	#Call	#Subrs	Inclusive usec/call	Name
100.0	1,360	12,284	1	8	12284122	.TAU application
57.4	7,049	7,049	1	0	7049442	MPI_Finalize()
16.9	2,070	2,070	2	0	1035019	MPI_Send()
8.7	186	1,064	2	2	532152	MPI_Bcast()
7.1	877	877	2	0	438900	MPI_Collective Sync
6.0	740	740	1	0	740038	MPI_Init_thread()
0.0	0.002	0.002	1	0	2	MPI_Comm_rank()
0.0	0.002	0.002	1	0	2	MPI_Comm_size()

USER EVENTS Profile :NODE 1, CONTEXT 0, THREAD 0