Repeatability Test

Young-Kyoon Suh
Data-Intensive Systems Lab
School of Computer Science & Engineering
Kyungpook National University

September 19, 2018

1 Description

This document characterizes execution times measured on a simple program in pure-computation mode, called *INC*, with increasing task lengths (up to 16,384 seconds from 1 second). For the characterization, we present various histograms of INC throughout this document. The main goal of examining these histograms is to make sure if two (or more) histograms of INC with the same task length have the same shape or not. If that's the case, then we can say that such an INC run is so called *repeatable*; in other words, *repeatability* is satisfied in our experimental settings. Another goal is to uncover and explicate several interesting structures behind the histograms. The third goal is to build a statistical distribution (or model) fitting in the histograms, so that later we are capable of predicting a concrete execution time via the model on an arbitrary algorithm with a given input on a real execution environment.

In our experiments, we used EMPv5 [1]. (That said, the second step of EMPv5 was on purpose omitted, just to obtain better histograms by retaining more samples.) In the protocol, we use taskstats C struct to get measures of a captured process. The taskstat's data is delivered via a netlink socket from the kernel space. The receive buffer for the socket is not robust for many observed processes [2]. Fortunately, there is an average of 95 processes per iteration of a run, which turns out to be fine with the struct. For a much more number of processes, the use of `/proc/`[pid]`/stat` is preferred, as (i) there are equivalent measures available in the `/proc` filesystem, and (ii) there's little constraint on the use as opposed to taskstats.

Now we show histograms of elapsed time (ET) and process time (PT) of INC via the EMPv5 protocol.

2 Histograms on the First Run

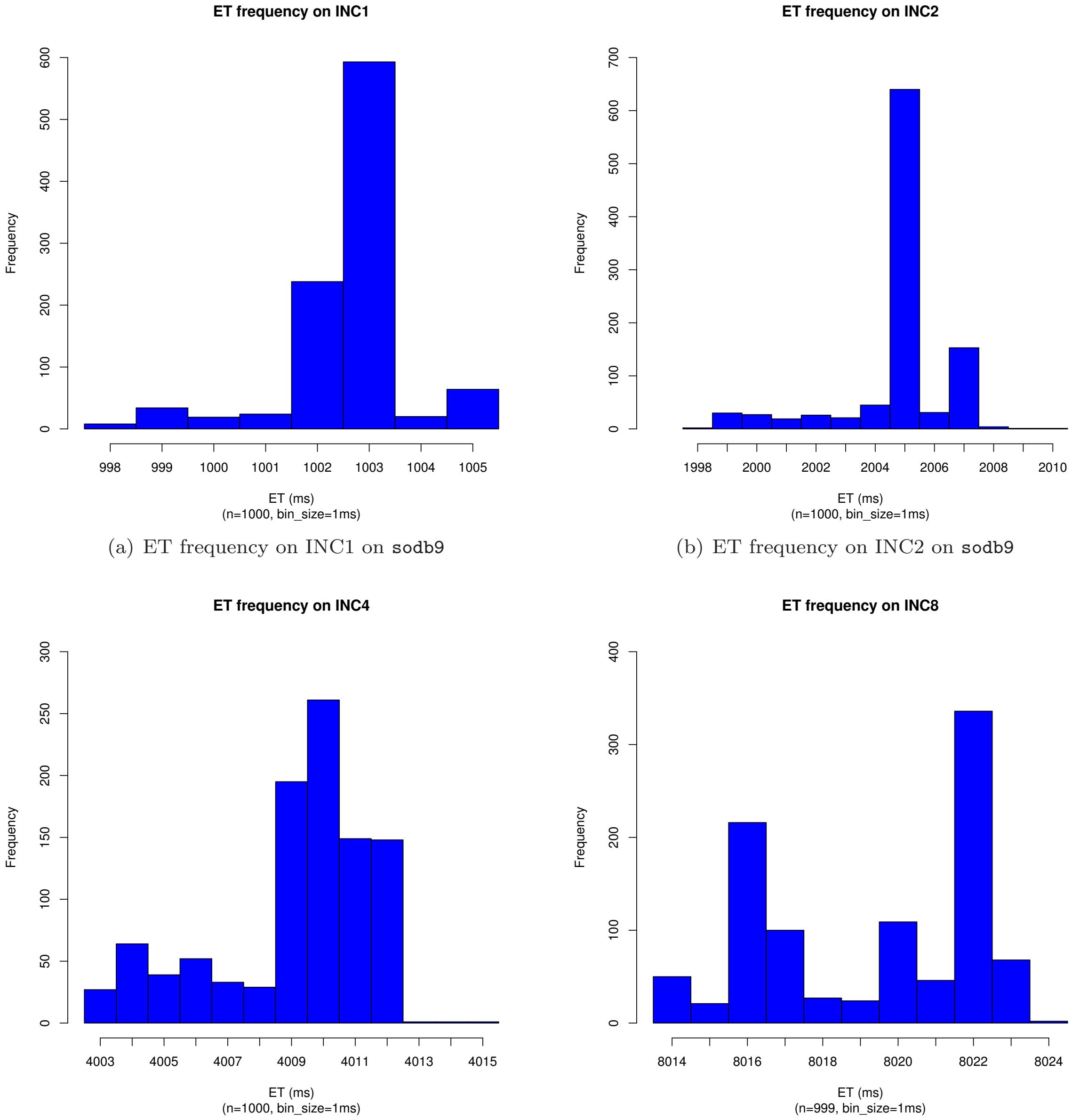
This section exhibits histograms on the first run of INC with its task length increasing from 1 second to 4,096 seconds. The detailed description of the base data is from Table 1.

Experiment Notes: Table 1 provides a short description of our experimental runs, on which the following histograms are based.

Machine	Task Length (sec)	Description	Experiment Period	Relevant Histograms
`sodb9` (plugged into *the upper left* power strip)	INC1∼INC64	1000 samples, each	2017-03-02 ∼ 2017-03-04	Figs. 1, 2, 5, and 6
`sodb9` (plugged into *the upper left* power strip)	INC128∼ INC1024	300 samples, each	2017-03-04 ∼ 2017-03-11	Figs. 3 and 7
`sodb10` (plugged into *the upper left* power strip)	INC2048	300 samples	2017-03-02 ∼ 2017-03-09	Figs. 4(a) and 8(a)
`sodb12` (plugged into *the upper right* power strip)	INC4096	300 samples	2017-02-13 ∼ 2017-02-27	Figs. 4(b) and 8(b)

Table 1: Notes on experiment runs used for histograms

2.1 ET



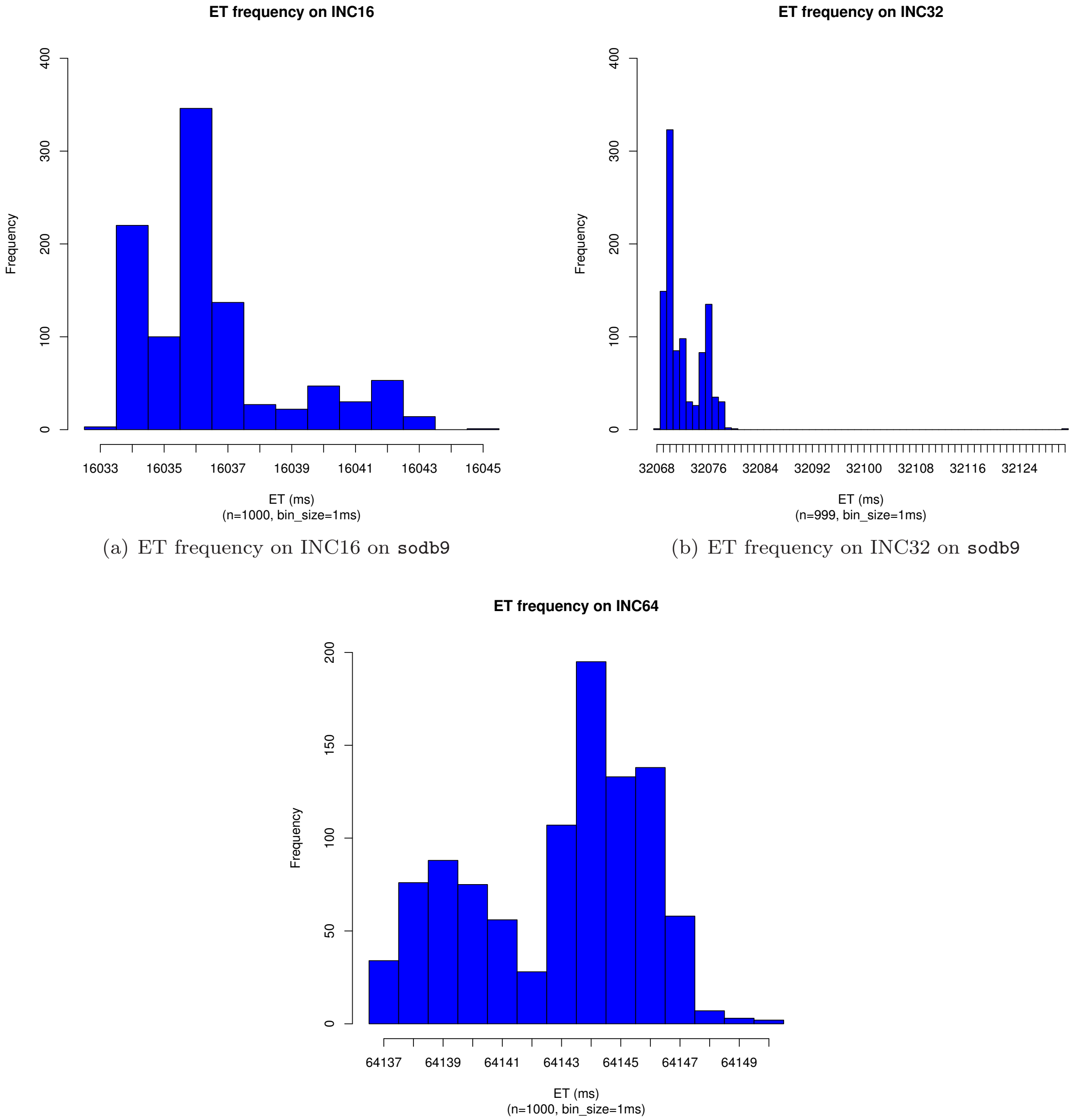
(a) ET frequency on INC1 on `sodb9`

(b) ET frequency on INC2 on `sodb9`

(c) ET frequency on INC4 on `sodb9`

(d) ET frequency on INC8 on `sodb9`

Figure 1: ET Histograms of INC1 ... INC8

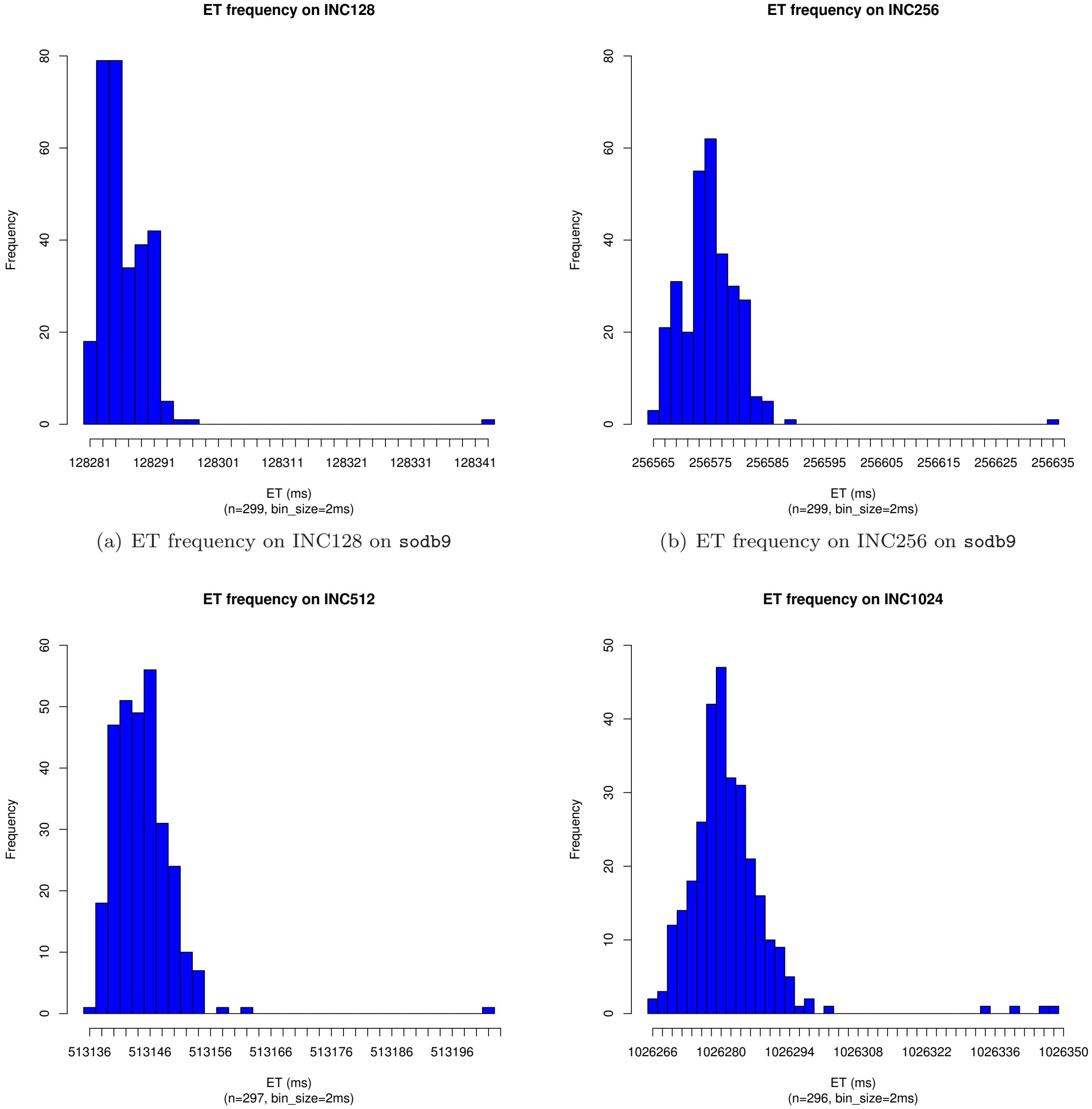


(a) ET frequency on INC16 on `sodb9`

(b) ET frequency on INC32 on `sodb9`

(c) ET frequency on INC64 on `sodb9`

Figure 2: ET Histograms of INC16 ... INC64



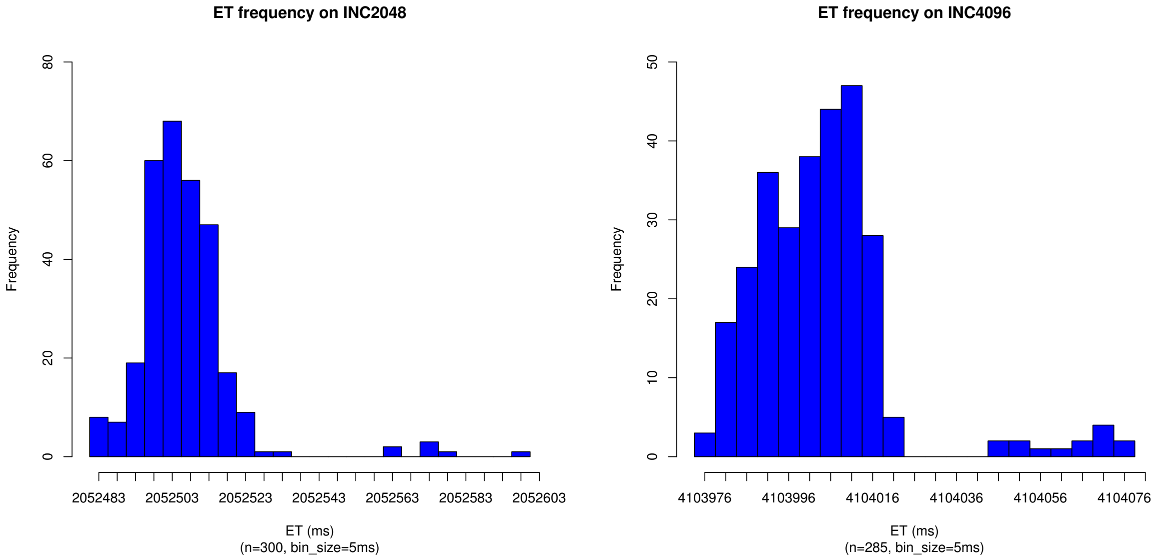
(a) ET frequency on INC128 on `sodb9`

(b) ET frequency on INC256 on `sodb9`

(c) ET frequency on INC512 on `sodb9`

(d) ET frequency on INC1024 on `sodb9`

Figure 3: ET Histograms of INC128 ... INC1024

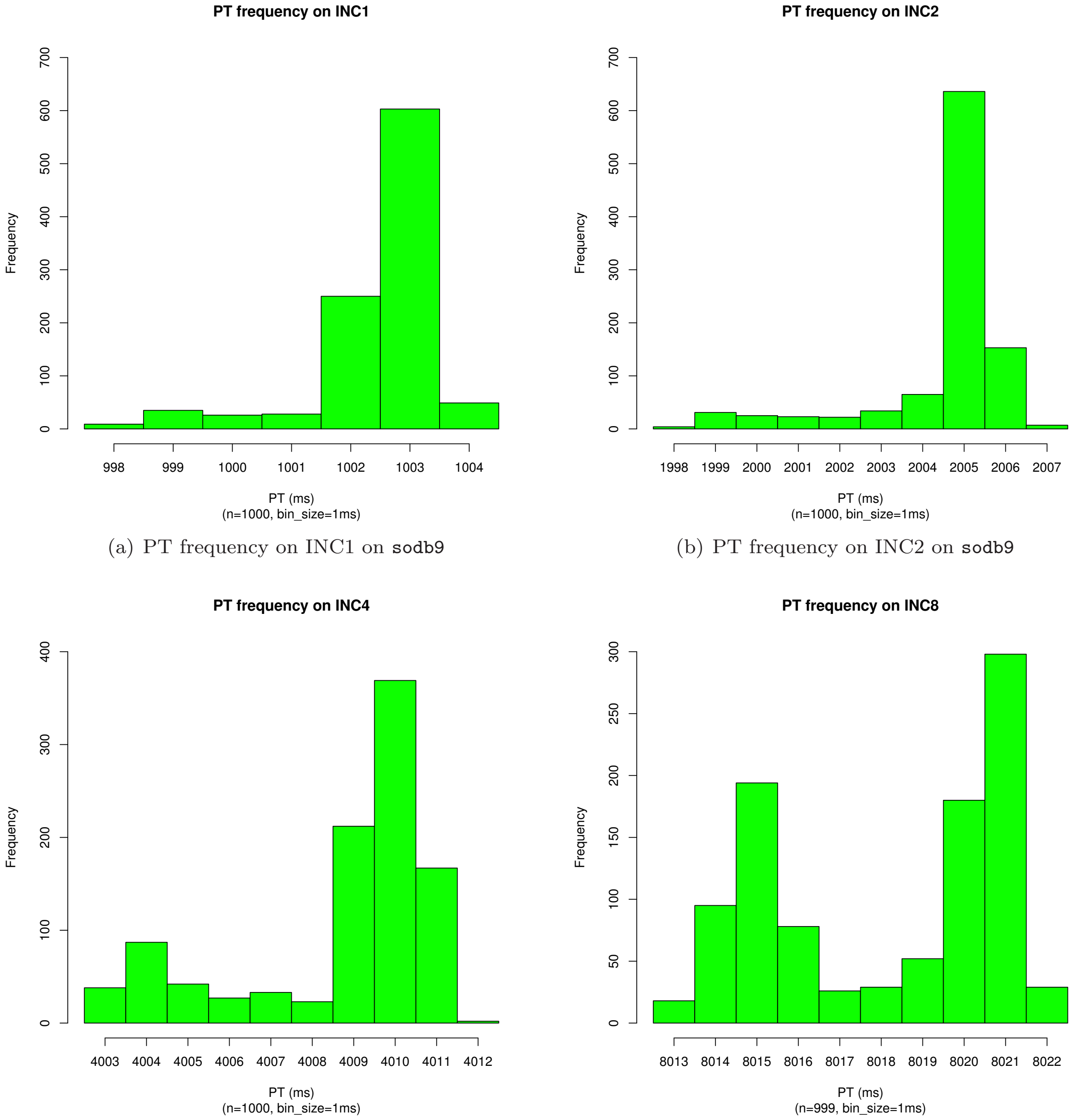


(a) ET frequency on INC2048 on `sodb10`

(b) ET frequency on INC4096 on `sodb12`

Figure 4: ET Histograms of INC2048 and INC4096

2.2 PT



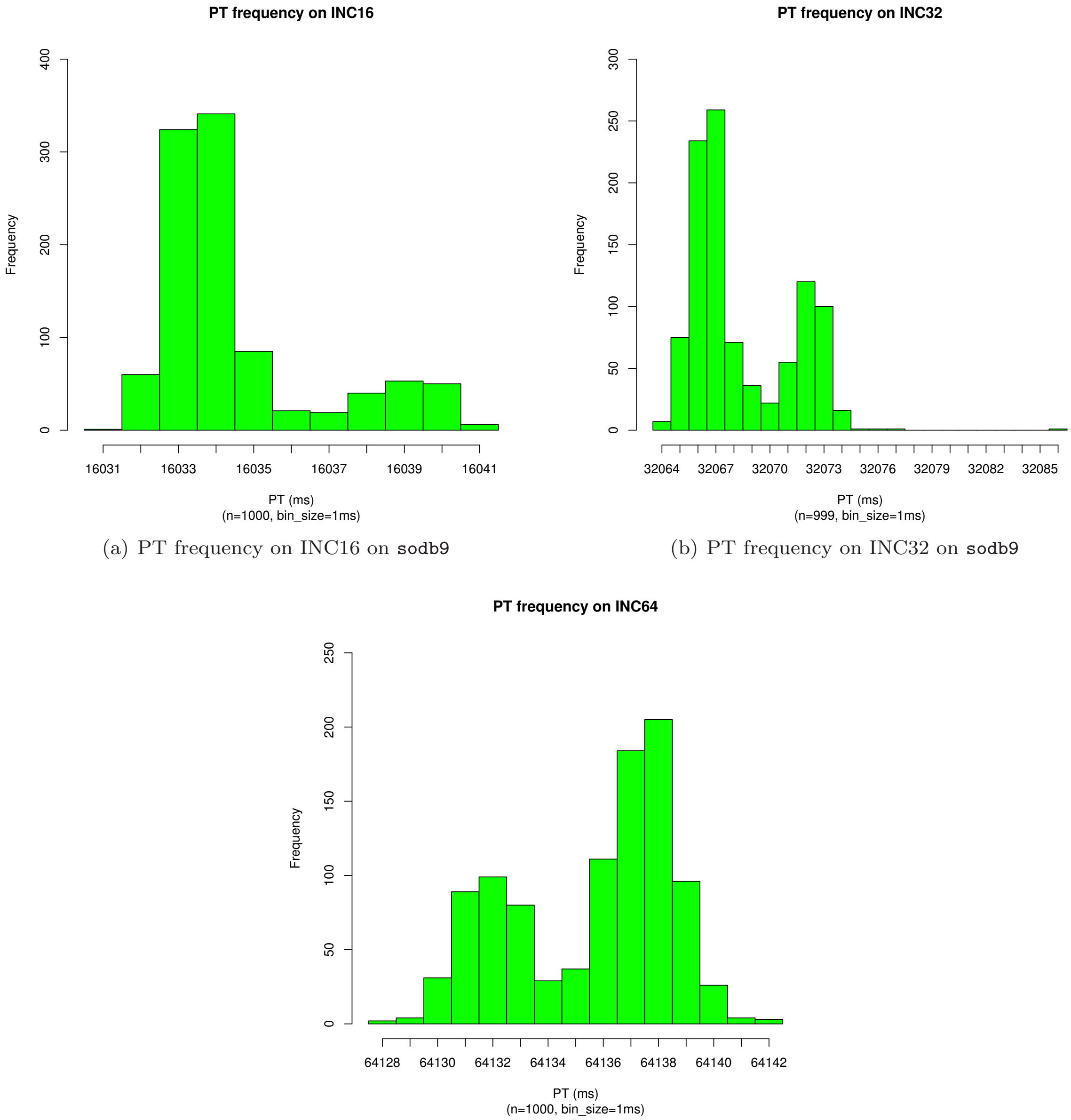
(a) PT frequency on INC1 on `sodb9`

(b) PT frequency on INC2 on `sodb9`

(c) PT frequency on INC4 on `sodb9`

(d) PT frequency on INC8 on `sodb9`

Figure 5: PT Histograms of INC1 ... INC8

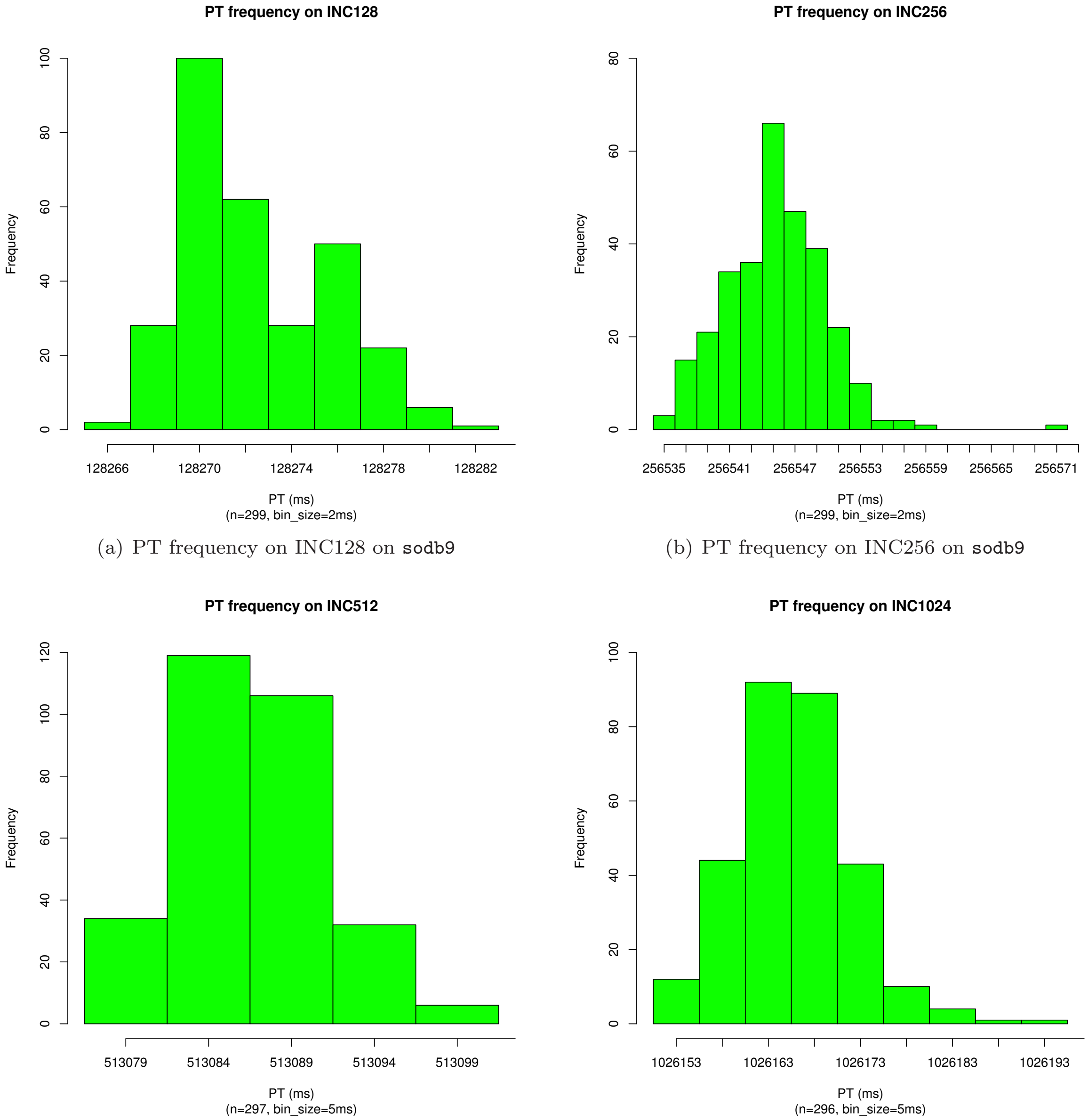


(a) PT frequency on INC16 on `sodb9`

(b) PT frequency on INC32 on `sodb9`

(c) PT frequency on INC64 on `sodb9`

Figure 6: PT Histograms of INC16 ... INC64



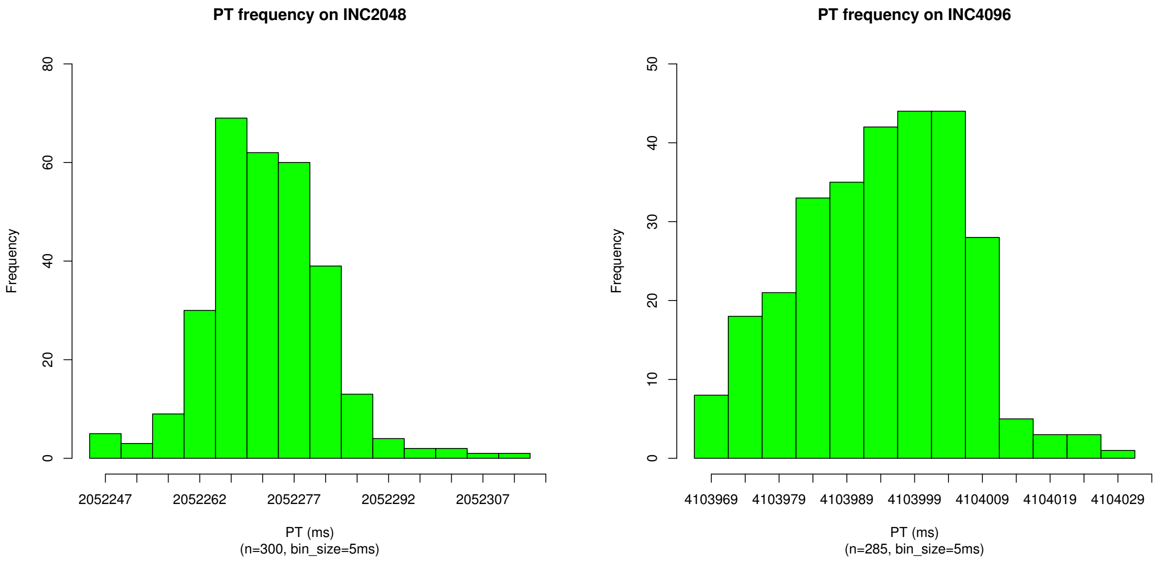
(a) PT frequency on INC128 on `sodb9`

(b) PT frequency on INC256 on `sodb9`

(c) PT frequency on INC512 on `sodb9`

(d) PT frequency on INC1024 on `sodb9`

Figure 7: PT Histograms of INC256 ... INC1024



(a) PT frequency on INC2048 on `sodb10`

(b) PT frequency on INC4096 on `sodb12`

Figure 8: PT Histograms of INC2048 and INC4096

2.3 Additional Histograms

This section exhibits the histograms of INC with some intermediate task lengths under 256 secs. These histograms are intended to investigate where the crossing and merge of two peaks that are consistently observed up to INC128 happened. Table 2 provides a description of the intermediate runs.

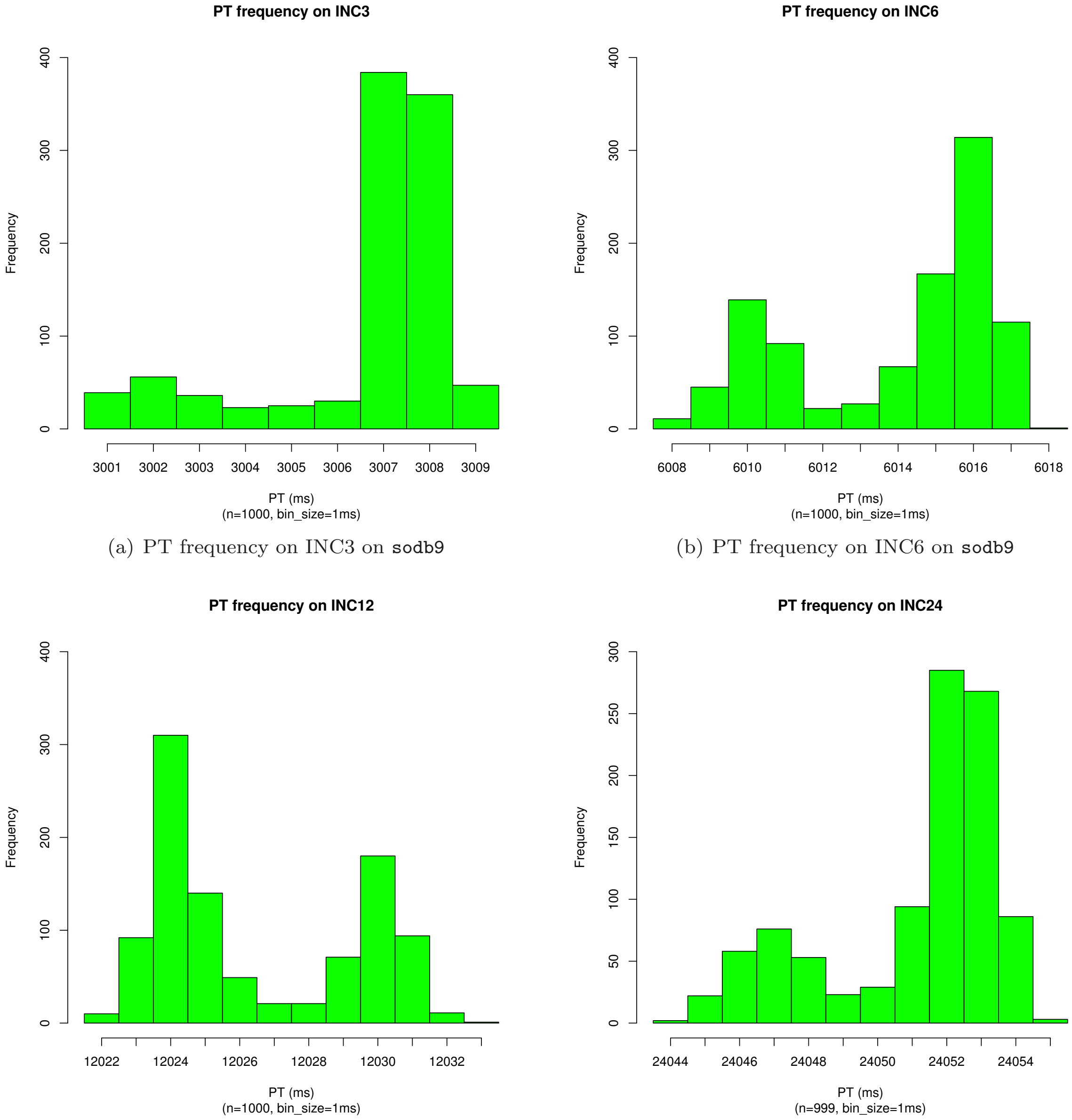
Machine	Task Length (sec)	Description	Experiment Period	Relevant Histograms
sodb9 (plugged into *the upper left* power strip)	INC96∼INC256	1000 samples, each	2017-05-24 ∼ 2017-06-06	Figs. 19, 22(d), and 23(a)
	INC3, 6, 12, 24, 48, 72, 80, 88, 104, 112, and 120	1000 samples, each	2017-06-07 ∼ 2017-06-16	Figs. 17(c), 18(a), 18(c), 19(a), 19(c), 20(a), 20(b), 20(c), 21(a), 21(b), and 21(d)

Table 2: Notes on experiment runs used for histograms

2.3.1 ET

Not available at this point, due to the labshelf server's unavailability for the time being.

2.3.2 PT



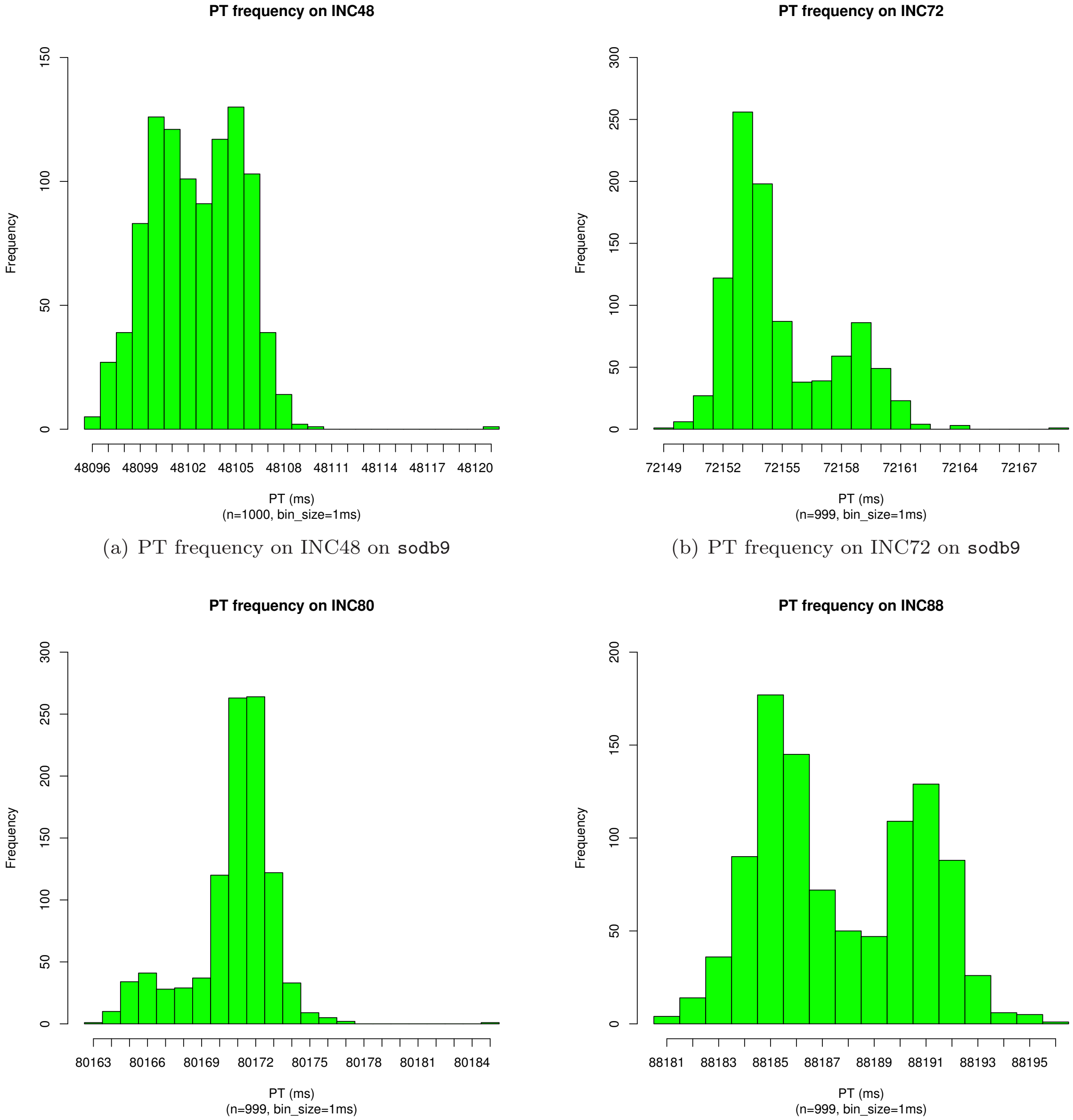
(a) PT frequency on INC3 on `sodb9`

(b) PT frequency on INC6 on `sodb9`

(c) PT frequency on INC12 on `sodb9`

(d) PT frequency on INC24 on `sodb9`

Figure 9: PT Histograms of INC3 ... INC24



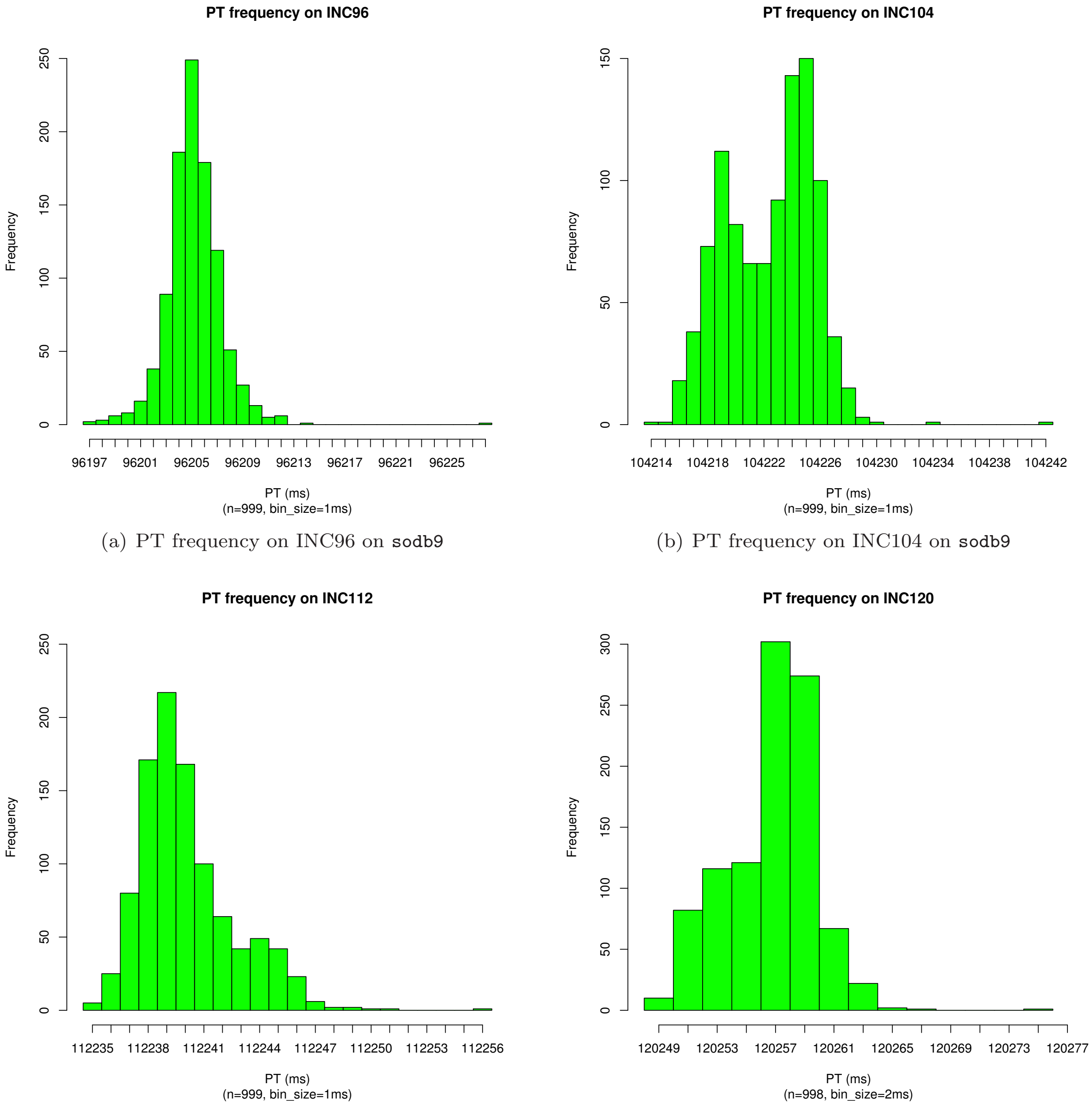
(a) PT frequency on INC48 on `sodb9`

(b) PT frequency on INC72 on `sodb9`

(c) PT frequency on INC80 on `sodb9`

(d) PT frequency on INC88 on `sodb9`

Figure 10: PT Histograms of INC48 ... INC88



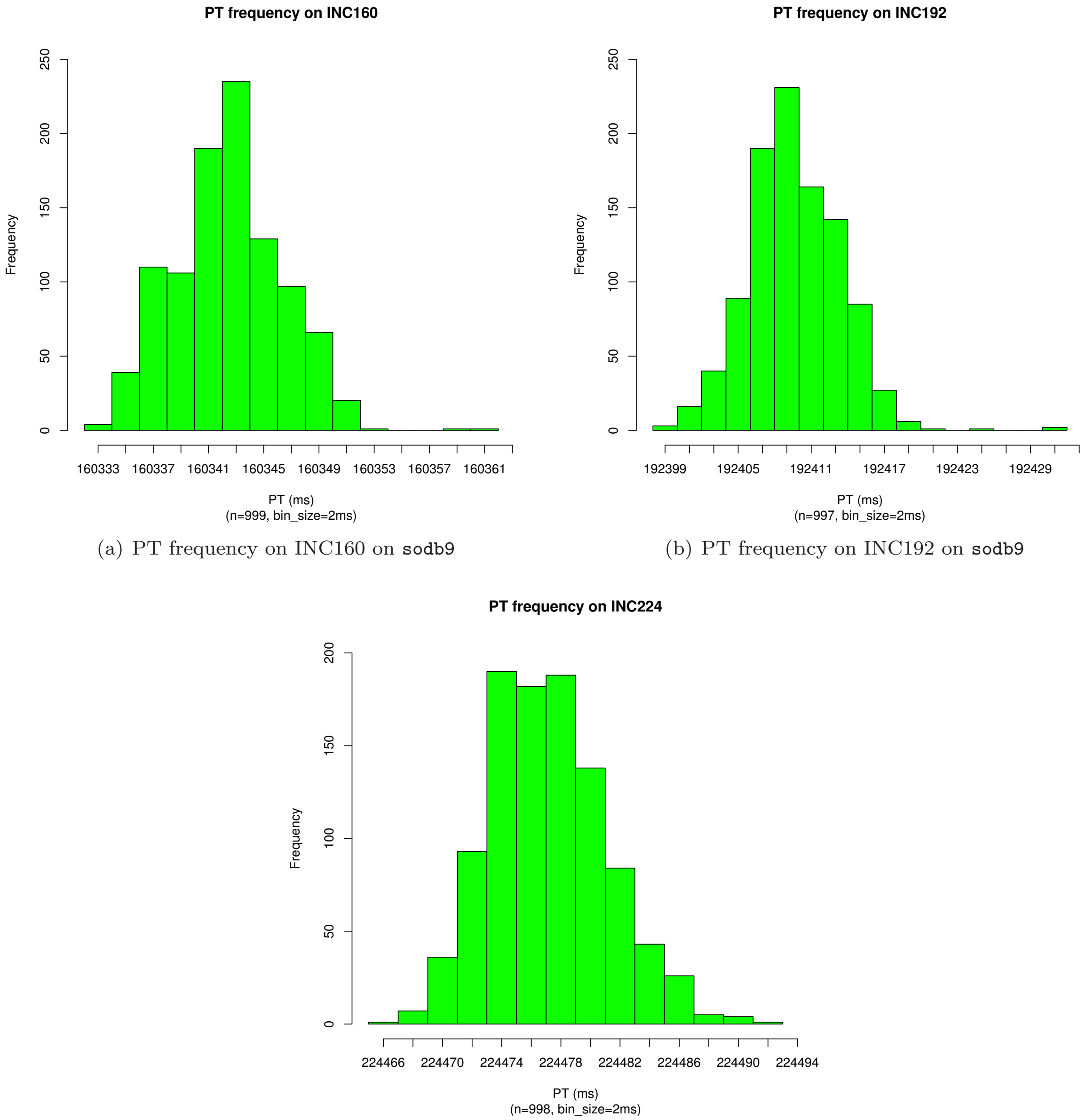
(a) PT frequency on INC96 on `sodb9`

(b) PT frequency on INC104 on `sodb9`

(c) PT frequency on INC112 on `sodb9`

(d) PT frequency on INC120 on `sodb9`

Figure 11: PT Histograms of INC96 ... INC120



(a) PT frequency on INC160 on `sodb9`

(b) PT frequency on INC192 on `sodb9`

(c) PT frequency on INC224 on `sodb9`

Figure 12: PT Histograms of INC169 ... INC224

2.4 Summary

Stacked Histograms in a 3D Plot: Figure 13 represents a 3D plot of collecting the histograms of PT in Section 2.2. Note that we additionally include the histograms from the task lengths of 8192 and 16384 seconds, which will be shown in Figures 38(c) and 38(d) (on page 47). See the legend in the right hand side of the figure, for more details about specific task length information. (To obtain the better shape, we could remove from those histograms the outliers that are identified via the second step of EMPv5.) The x-axis corresponds to the normalized PT relative to the minimum, ranging 0 to 100, the y-axis to task length in log scale, and the z-axis to the normalized frequency relative to the highest bar for each histogram.

One thing to notice is that there are some runs having a few empty bins in their respective histogram. For instance, INC32 has such an empty bin as shown in Figure 2(b). Same with INC256 in Figure 3(b). For such an INC run, the z-axis associated with its histogram has the zero value on the x-axis in Figure 13.

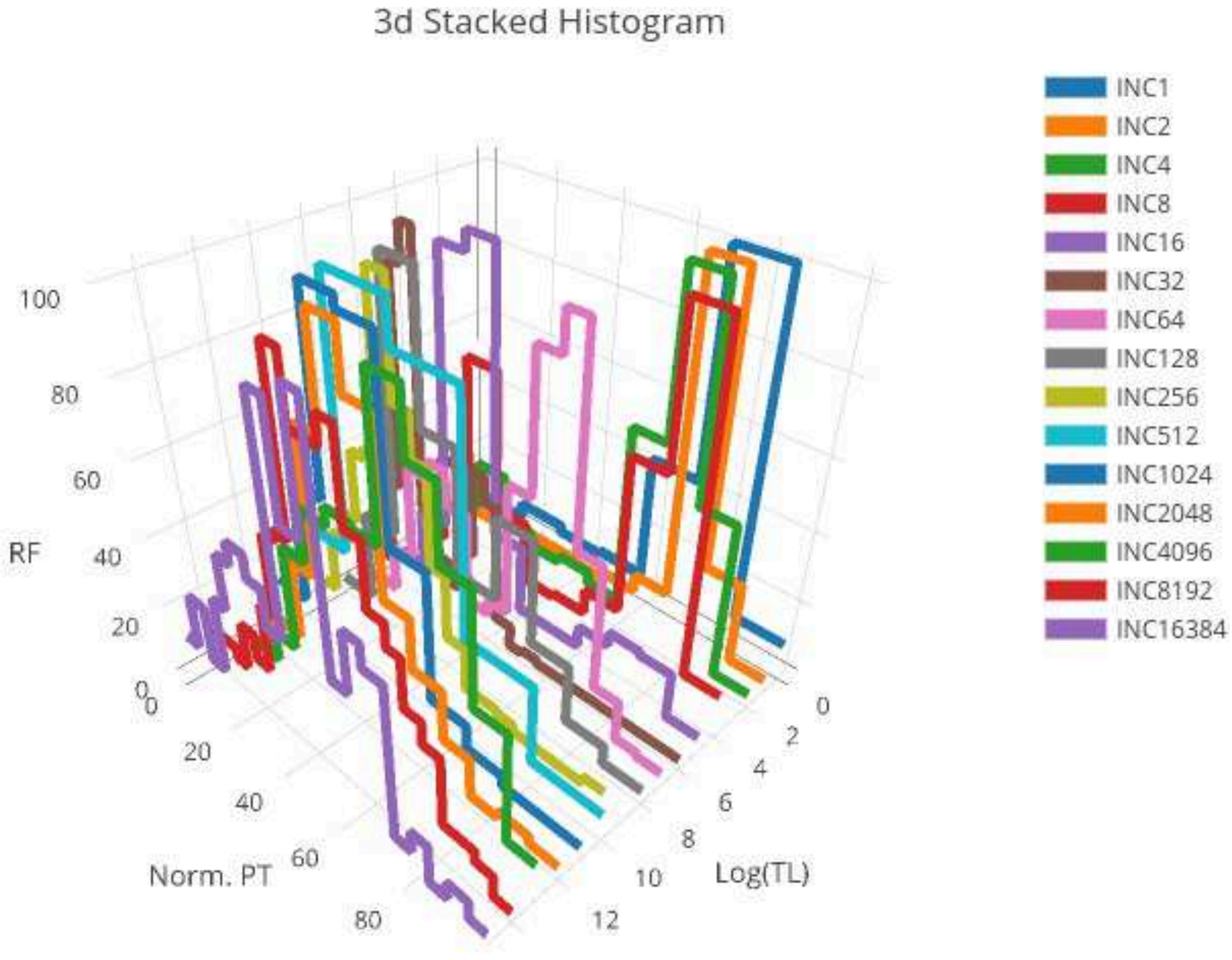


Figure 13: 3D histograms of PT on doubly-increasing INC on `sodb9` for up to INC1024, on `sodb10` for INC2048, on `sodb12` for INC4096

Figure 14 represents a 3D plot of collecting the histograms of user time only exhibited in Section 2.

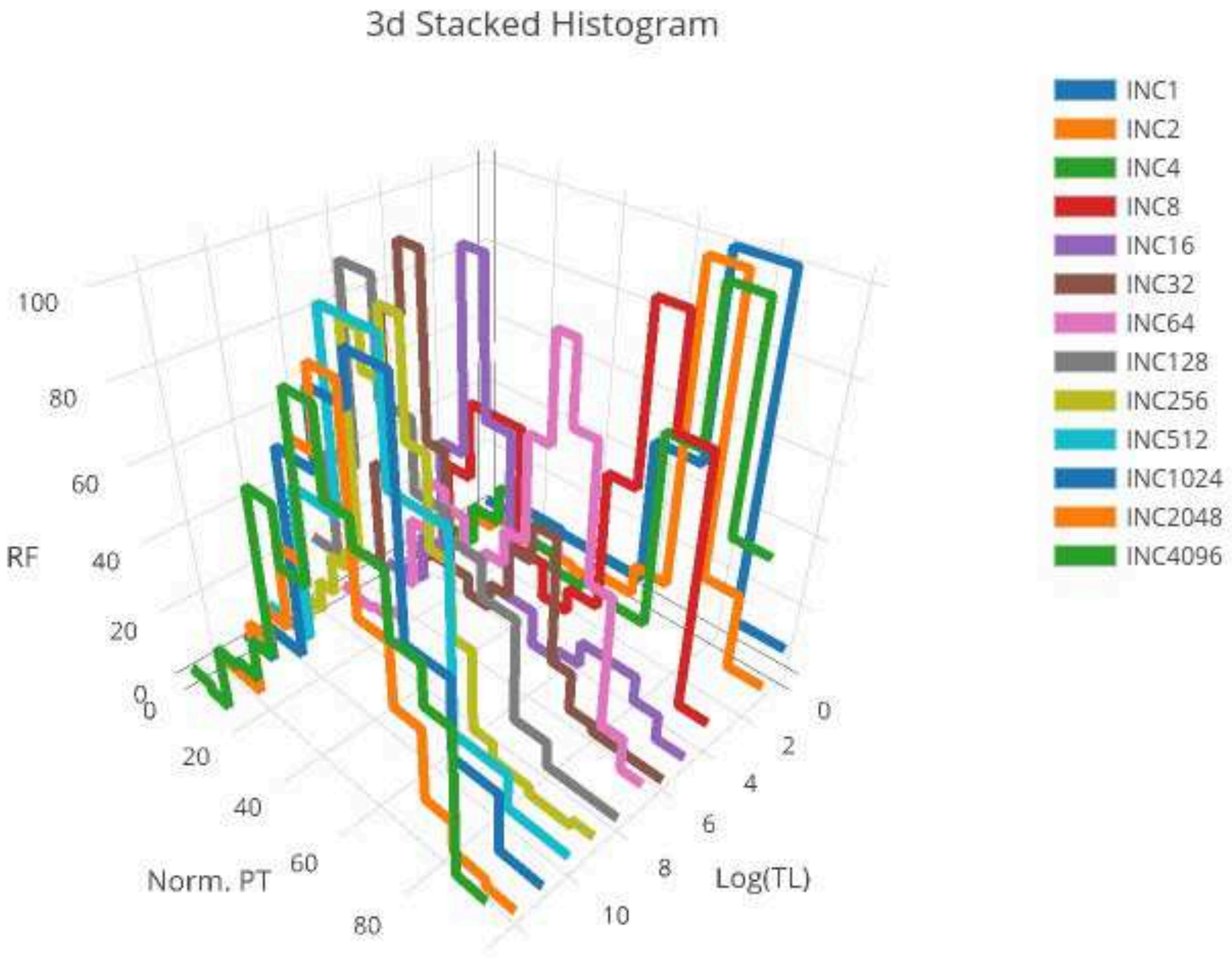
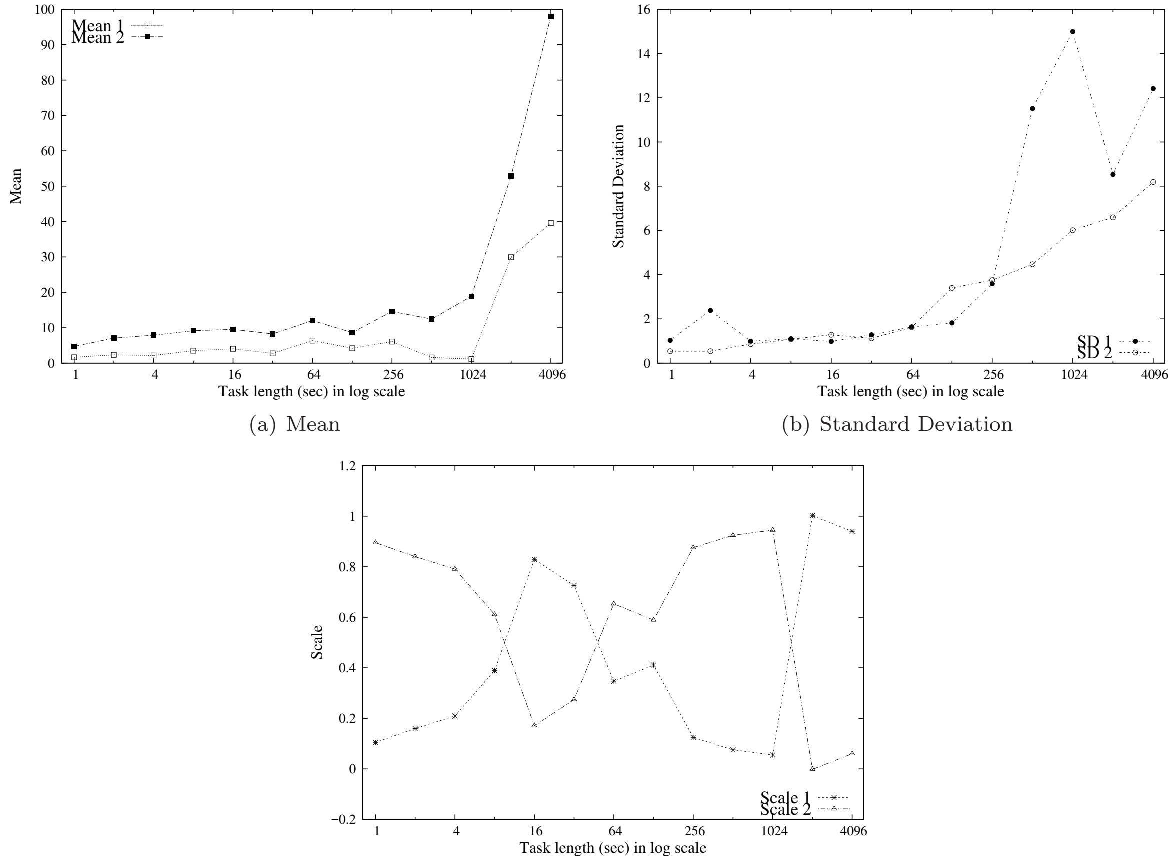


Figure 14: 3D histograms on INC on `sodb9` for up to INC1024, on `sodb10` for INC2048, on `sodb12` for INC4096: user time only

Fitting PT Histograms in Bi-Normal Distributions: We ran Rob's code to obtain a bi-normal fit for each INC's PT histogram in Section 2. Table 3 exhibits the total range of bins, the two means (modes), standard deviations, scales (weights) for the PT histogram, and figure numbers associated with each INC.

INC	Total Range	Mean 1	Mean 2	Std 1	Std 2	Scale 1	Scale 2	Figures
INC1	6	1.66	4.73	1.03	0.54	0.10	0.90	Fig. 5(a)
INC2	9	2.33	7.07	2.38	0.54	0.16	0.84	Fig. 5(b)
INC4	9	2.17	7.89	0.99	0.86	0.21	0.79	Fig. 5(c)
INC8	9	3.51	9.15	1.10	1.08	0.39	0.61	Fig. 5(d)
INC16	10	4.04	9.53	0.98	1.28	0.83	0.17	Fig. 6(a)
INC32	22	2.76	8.26	1.28	1.12	0.73	0.27	Fig. 6(b)
INC64	14	6.36	12.02	1.64	1.62	0.35	0.65	Fig. 6(c)
INC128	16	4.21	8.64	1.82	3.40	0.41	0.59	Fig. 7(a)
INC256	38	6.08	14.61	3.59	3.75	0.12	0.88	Fig. 7(b)
INC512	20	1.60	12.43	11.51	4.47	0.08	0.92	Fig. 7(c)
INC1024	40	1.13	18.88	14.99	6.01	0.06	0.94	Fig. 7(d)
INC2048	65	29.93	52.92	8.53	6.59	1.00	-0.00	Fig. 8(a)
INC4096	65	39.57	98.00	12.41	8.19	0.94	0.06	Fig. 8(b)

Table 3: Two means in mode computed for each INC



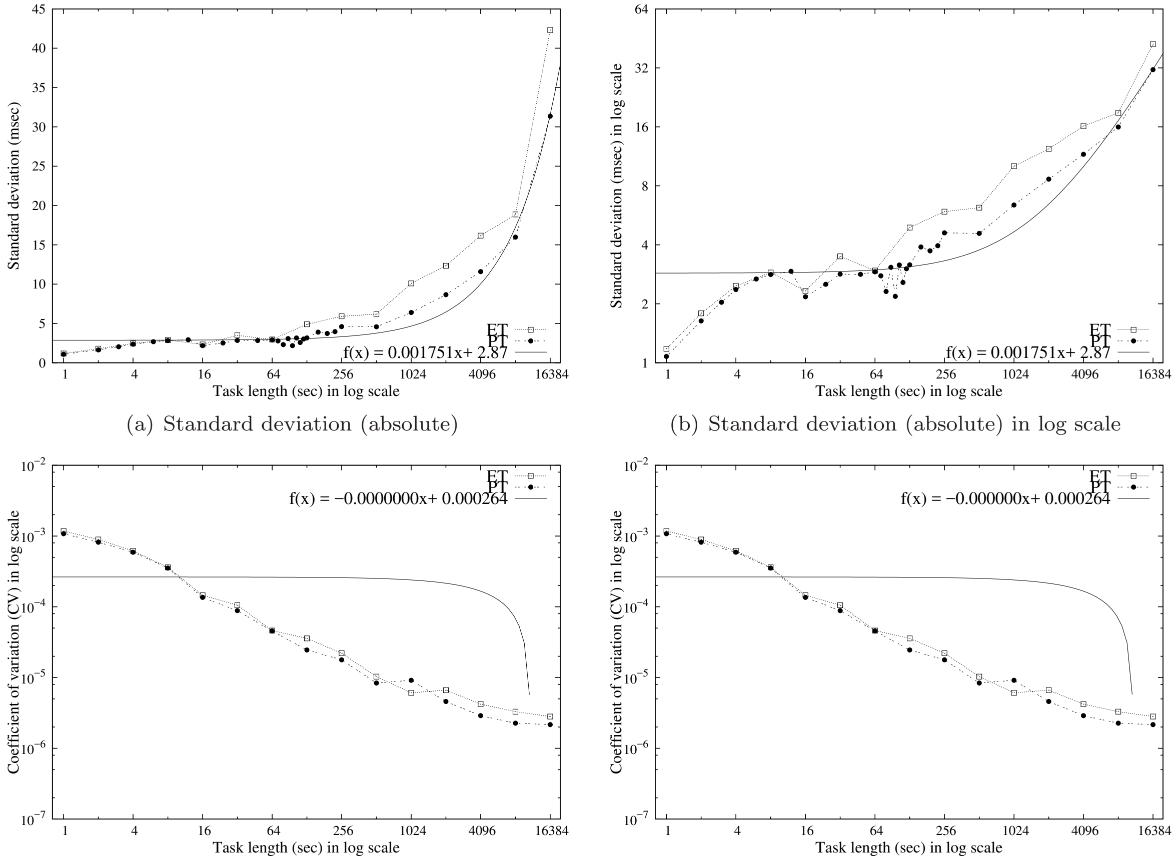
(a) Mean

(b) Standard Deviation

(c) Scale

Figure 15: Plots of two means, standard deviations, and scales for each INC over increasing task length: based on Table 3

Absolute and Relative Variance over Increasing Task Lengths: Figure 16 exhibits absolute and relative variances over increasing task lengths. More specifically, Figures 16(a) and 16(b) concern the PT standard deviations of all the runs including those two longest runs of INC8192 and INC16384 described in Table 12. The x-axis is task length, and the y-axis standard deviation; Figure 16(b) is taken in log scale. Figures 16(c) and 16(d) shows the relative variance of the same data set—*coefficient of variation* (= standard deviation / task length (or mean)). Both of the x and y axes in these two figures are taken in log scale. We also overlap a linear-square-fit of each case on the same figure.



(a) Standard deviation (absolute)

(b) Standard deviation (absolute) in log scale

(c) Coefficient of variation in log scale (relative): divisor=mean

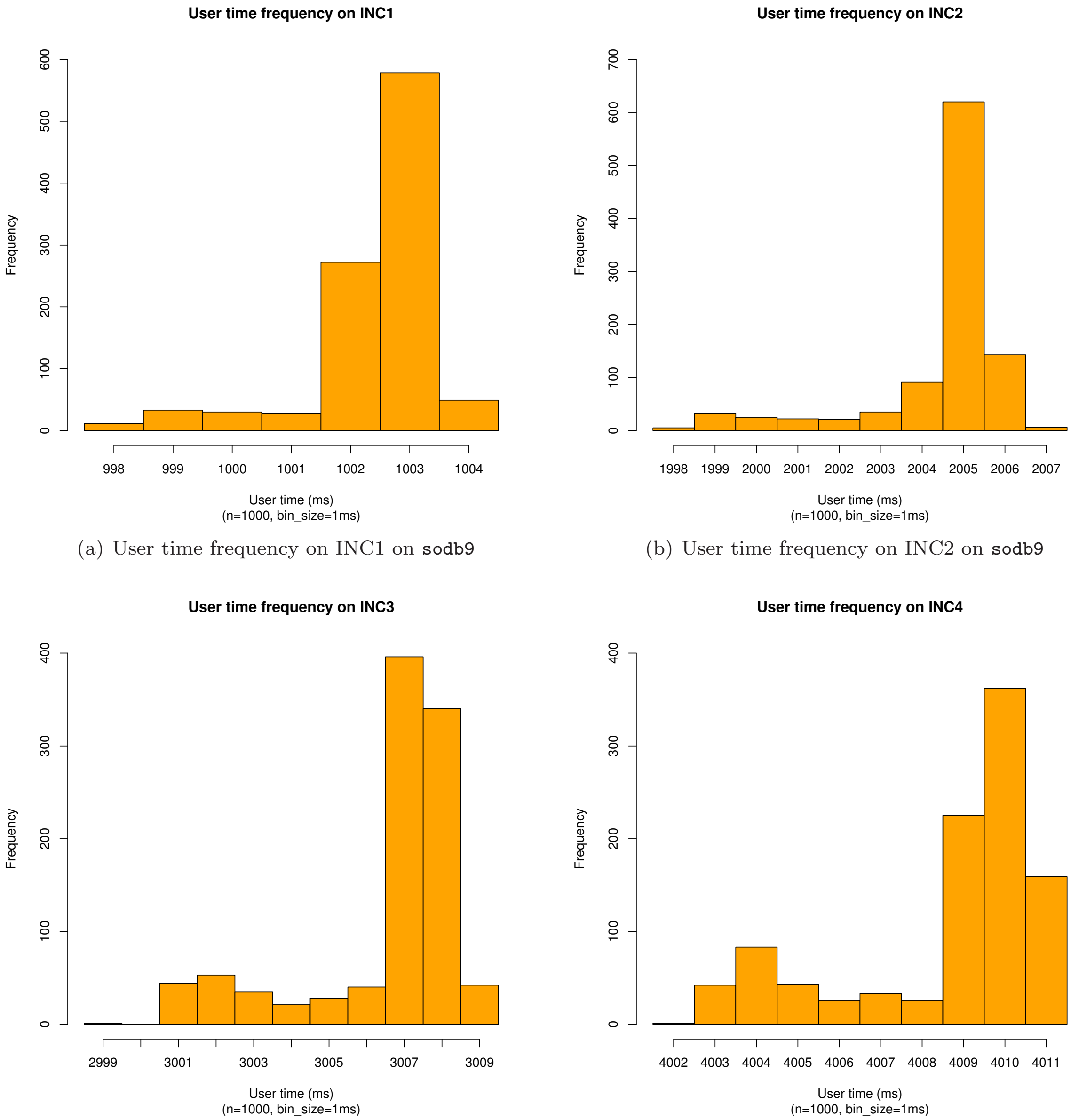
(d) Coefficient of variation in log scale (relative): divisor=task length

Figure 16: Absolute and relative variances

2.5 Histograms of User and System Time

This section exhibits histograms on user and system time of the INC runs described in Table 1.

2.5.1 User Time



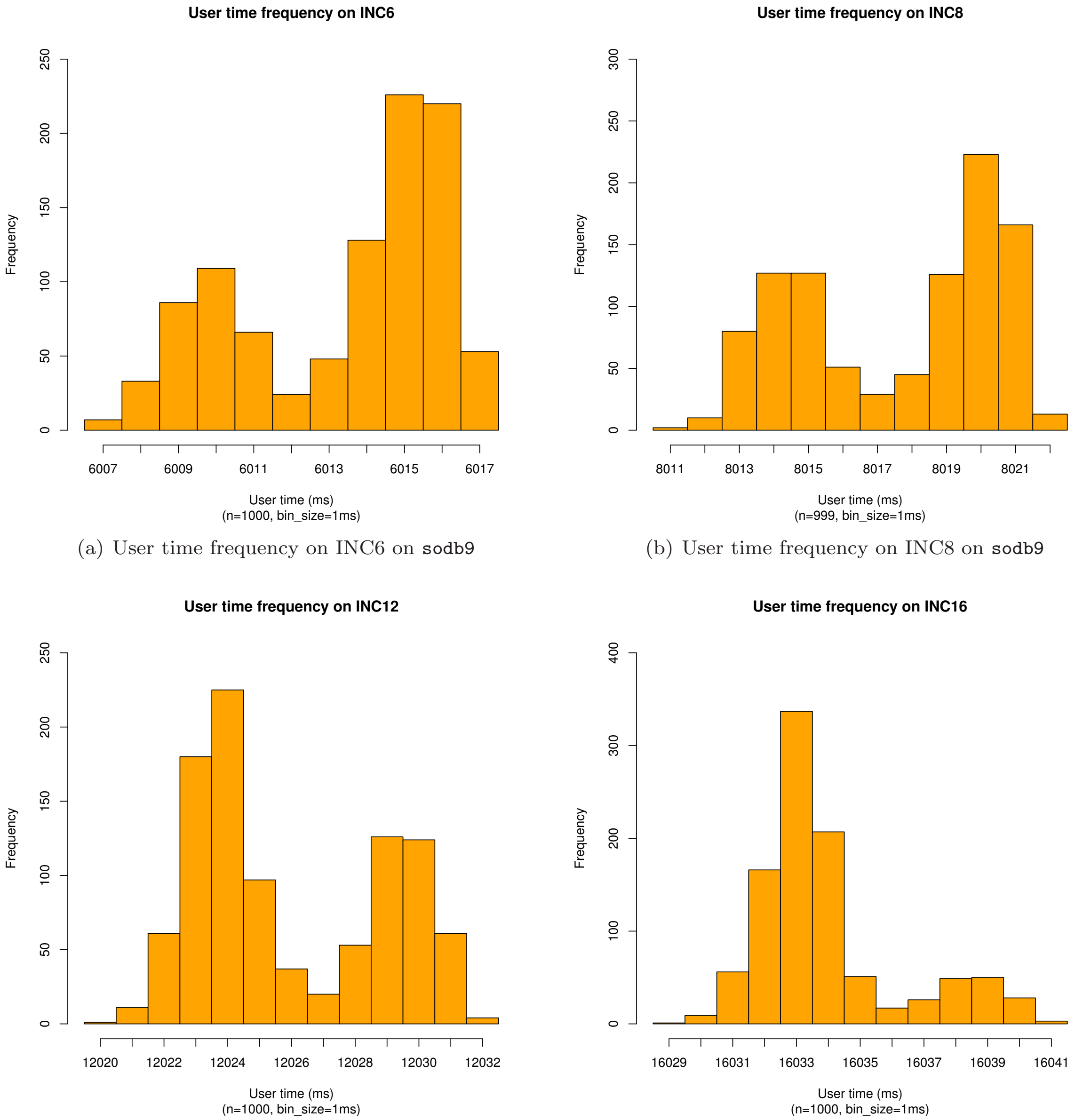
(a) User time frequency on INC1 on `sodb9`

(b) User time frequency on INC2 on `sodb9`

(c) User time frequency on INC3 on `sodb9`

(d) User time frequency on INC4 on `sodb9`

Figure 17: User Time Histograms of INC1 ... INC4



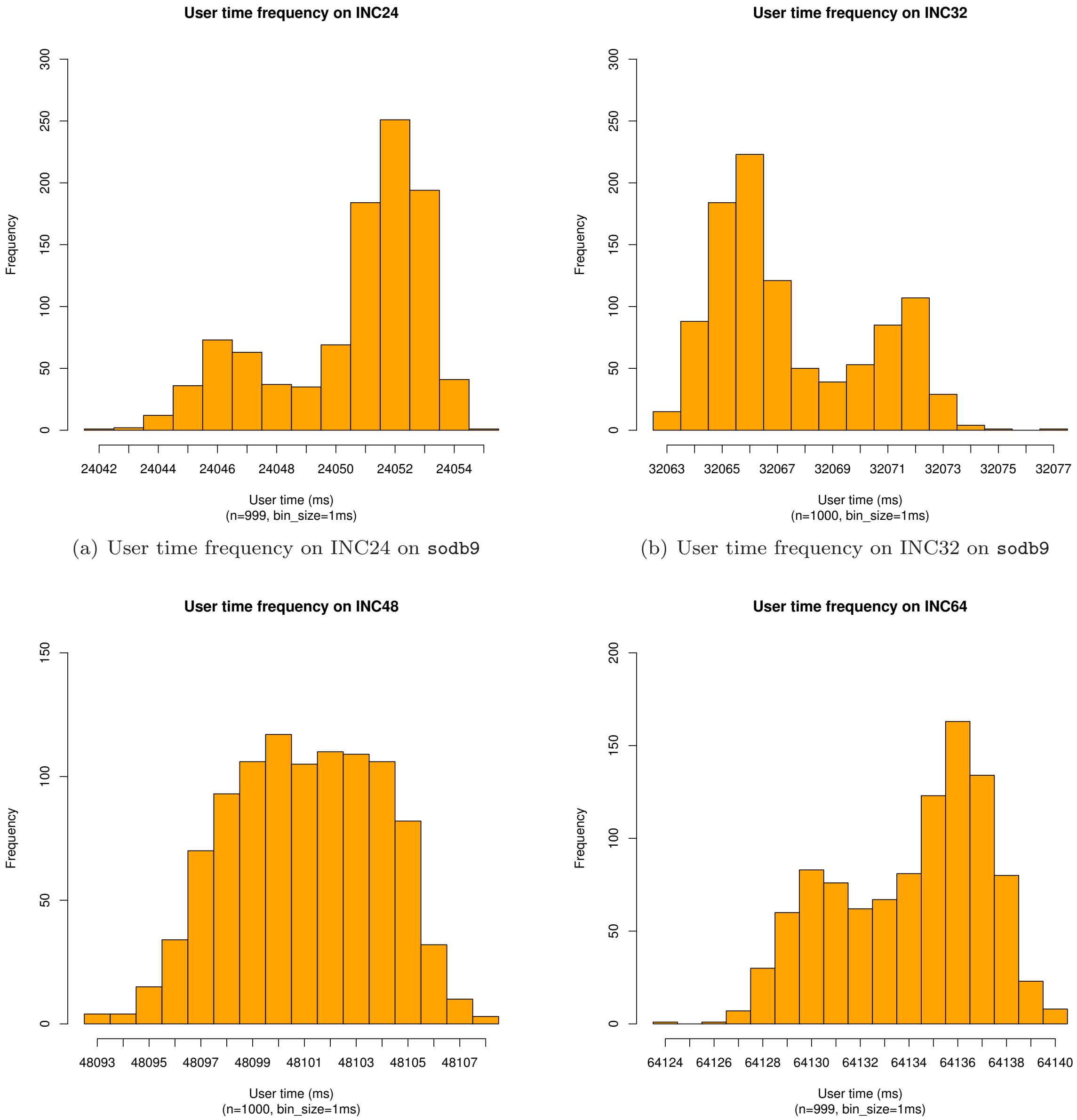
(a) User time frequency on INC6 on `sodb9`

(b) User time frequency on INC8 on `sodb9`

(c) User time frequency on INC12 on `sodb9`

(d) User time frequency on INC16 on `sodb9`

Figure 18: User Time Histograms of INC6 ... INC16



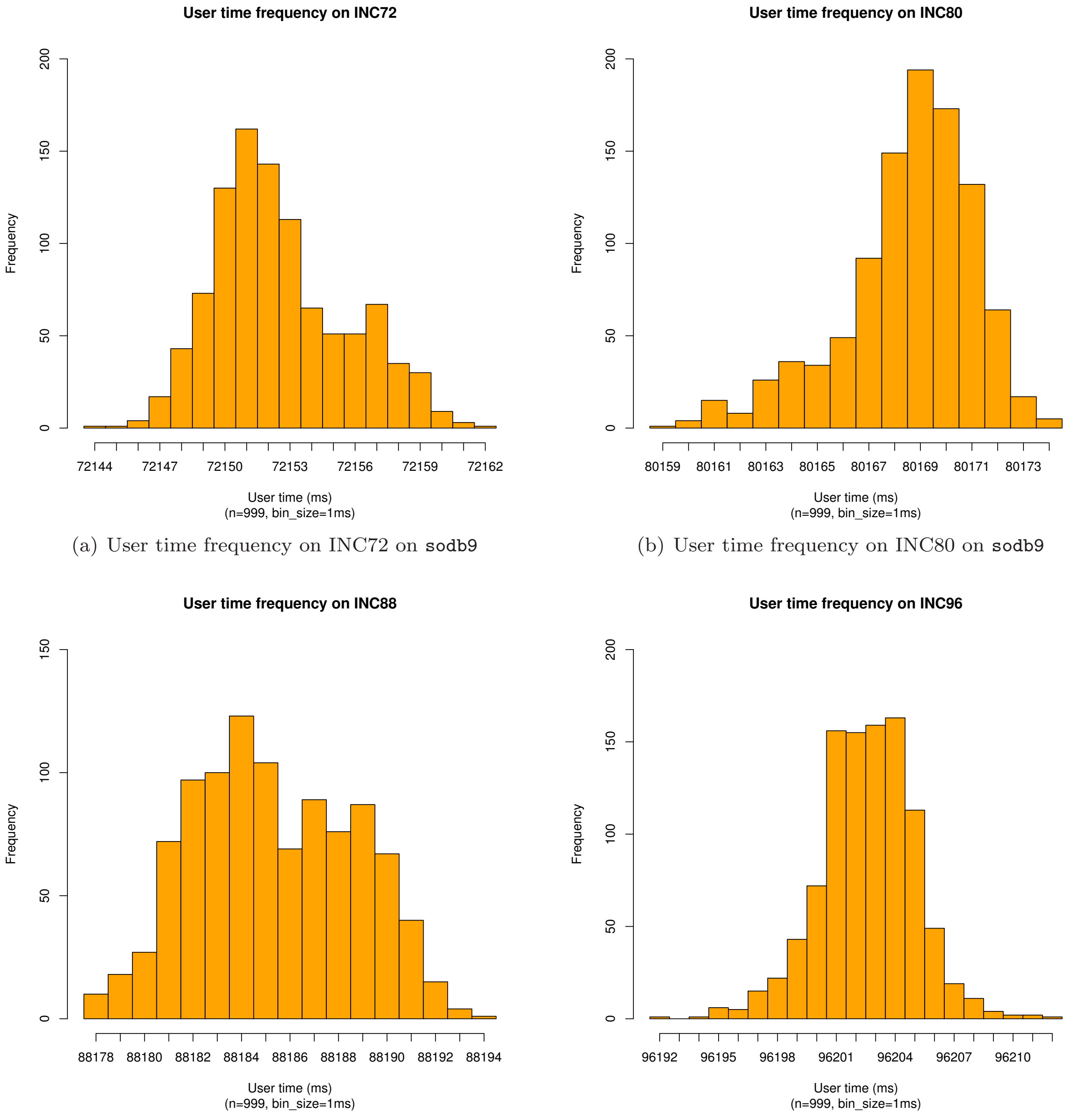
(a) User time frequency on INC24 on `sodb9`

(b) User time frequency on INC32 on `sodb9`

(c) User time frequency on INC48 on `sodb9`

(d) User time frequency on INC64 on `sodb9`

Figure 19: User Time Histograms of INC24 ... INC64



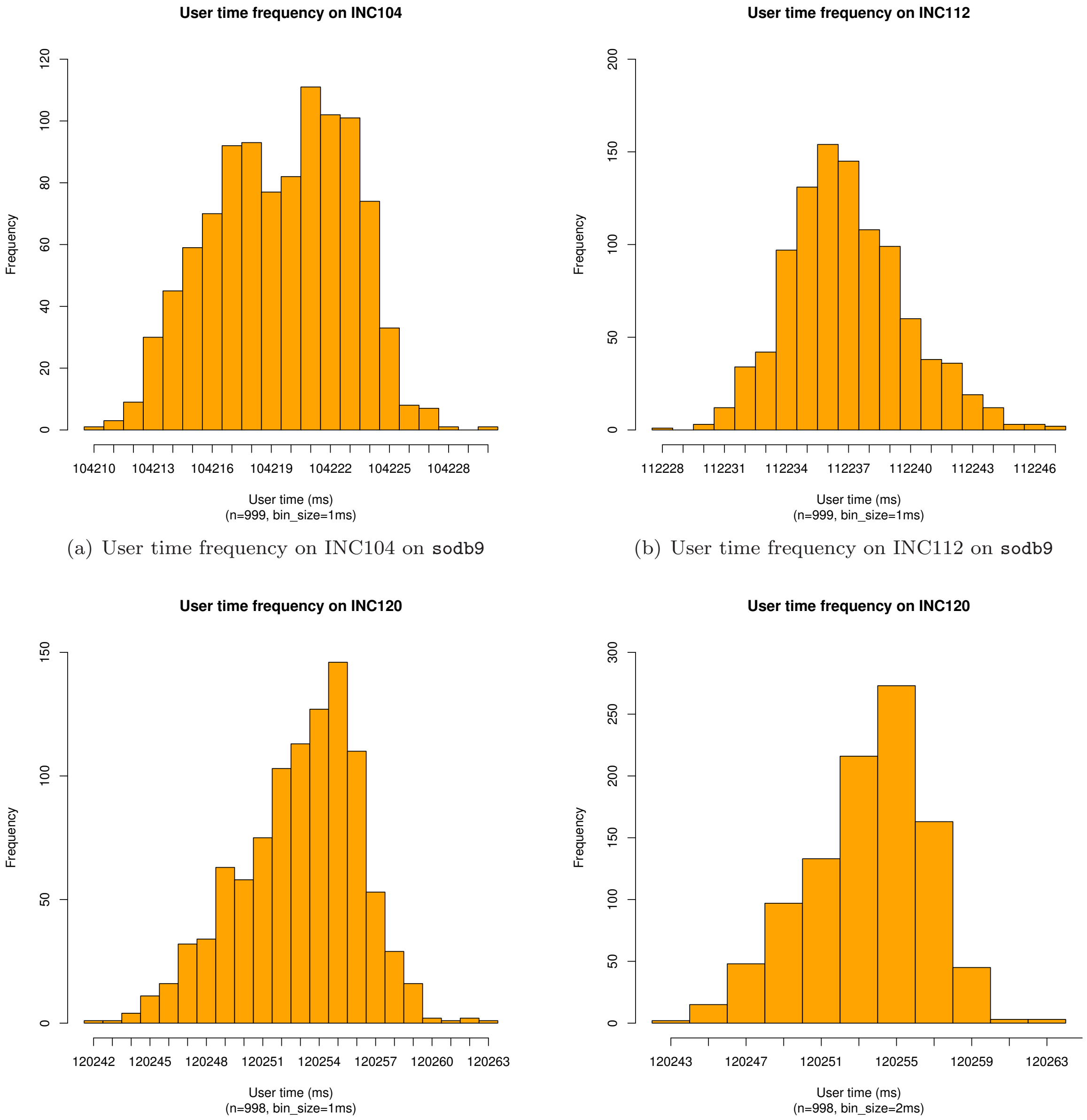
(a) User time frequency on INC72 on `sodb9`

(b) User time frequency on INC80 on `sodb9`

(c) User time frequency on INC88 on `sodb9`

(d) User time frequency on INC96 on `sodb9`

Figure 20: User Time Histograms of INC72 ... INC96



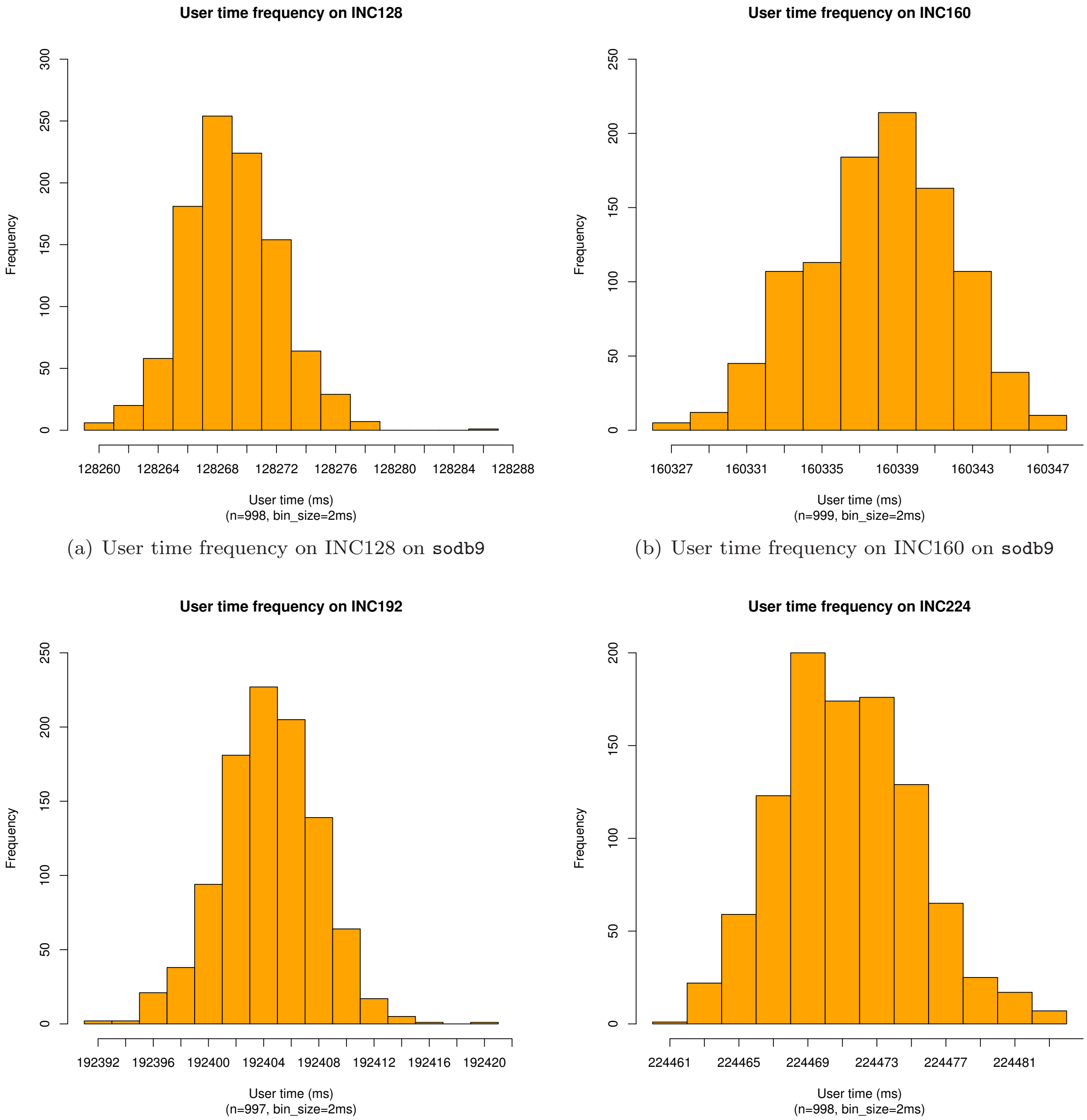
(a) User time frequency on INC104 on `sodb9`

(b) User time frequency on INC112 on `sodb9`

(c) User time frequency on INC120 on `sodb9` with bin size = 1 ms

(d) User time frequency on INC120 on `sodb9` with bin size = 2 ms

Figure 21: User Time Histograms of INC104 and INC120: for the same 120-sec task length, we used *two different* bin sizes to see the influence of the choice of the bin size on the shape of histogram.



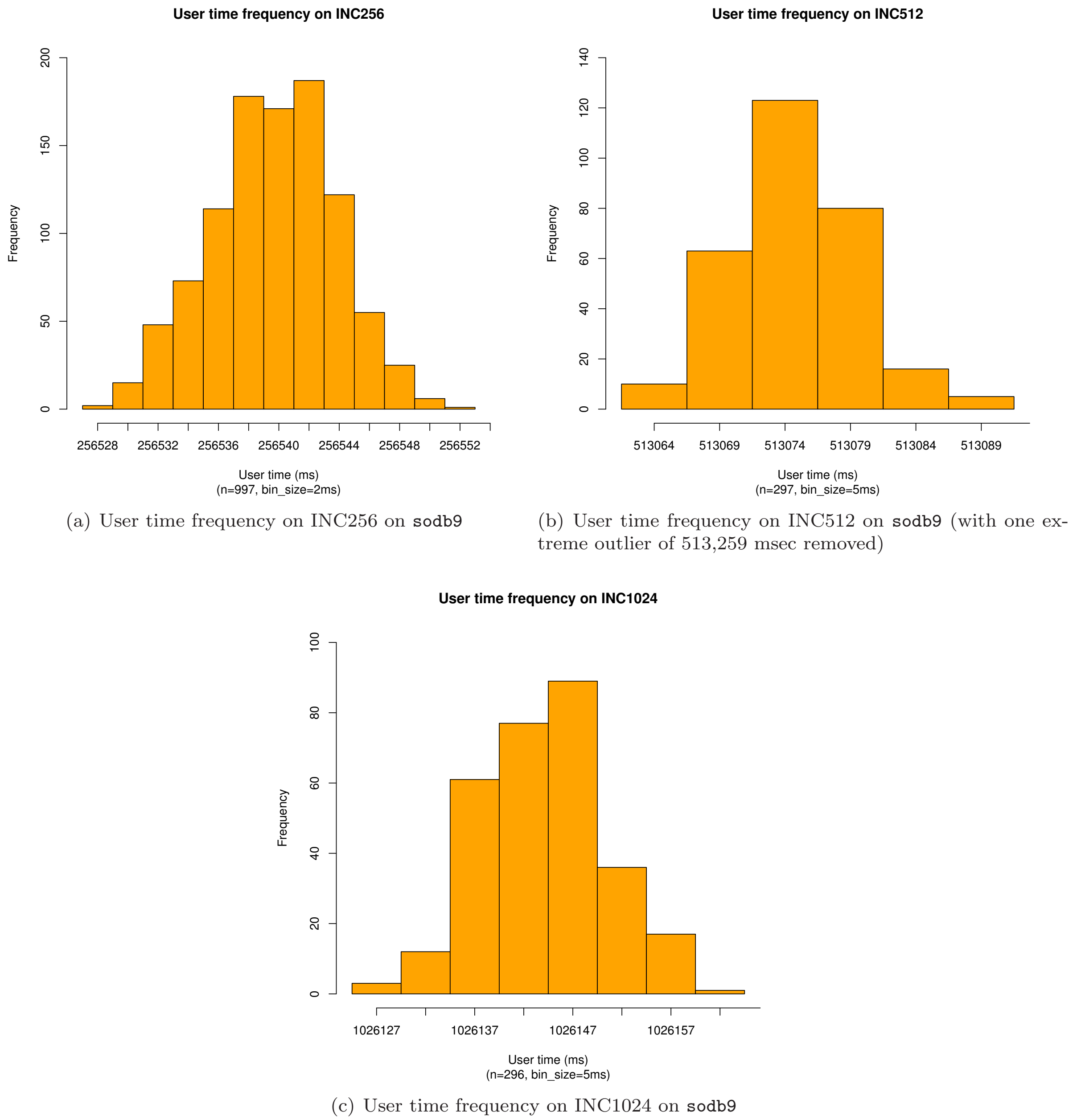
(a) User time frequency on INC128 on `sodb9`

(b) User time frequency on INC160 on `sodb9`

(c) User time frequency on INC192 on `sodb9`

(d) User time frequency on INC224 on `sodb9`

Figure 22: User Time Histograms of INC160 ... INC224

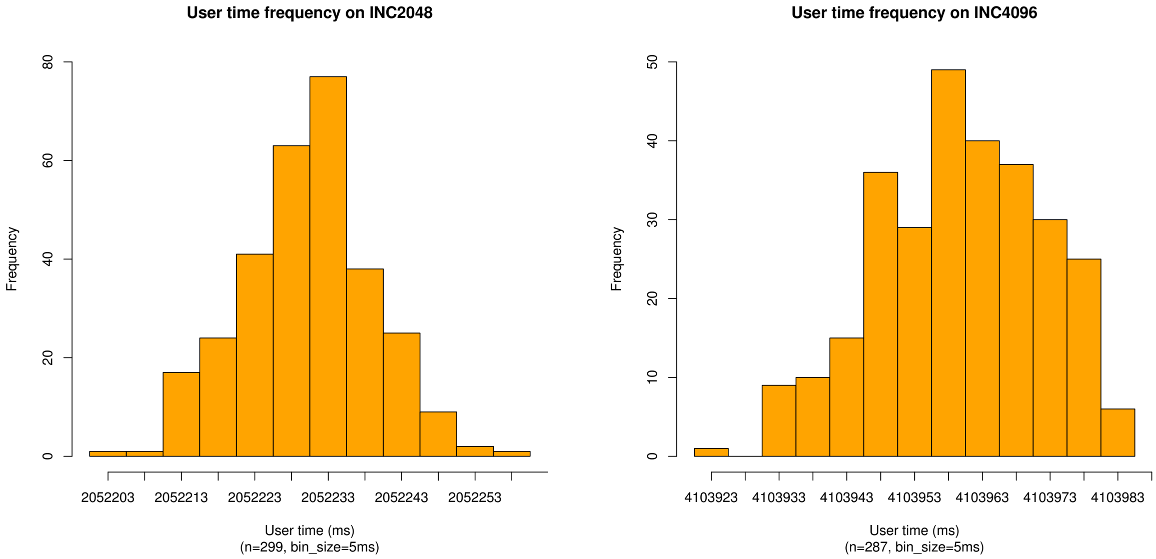


(a) User time frequency on INC256 on `sodb9`

(b) User time frequency on INC512 on `sodb9` (with one extreme outlier of 513,259 msec removed)

(c) User time frequency on INC1024 on `sodb9`

Figure 23: User Time Histograms of INC256 ... INC1024

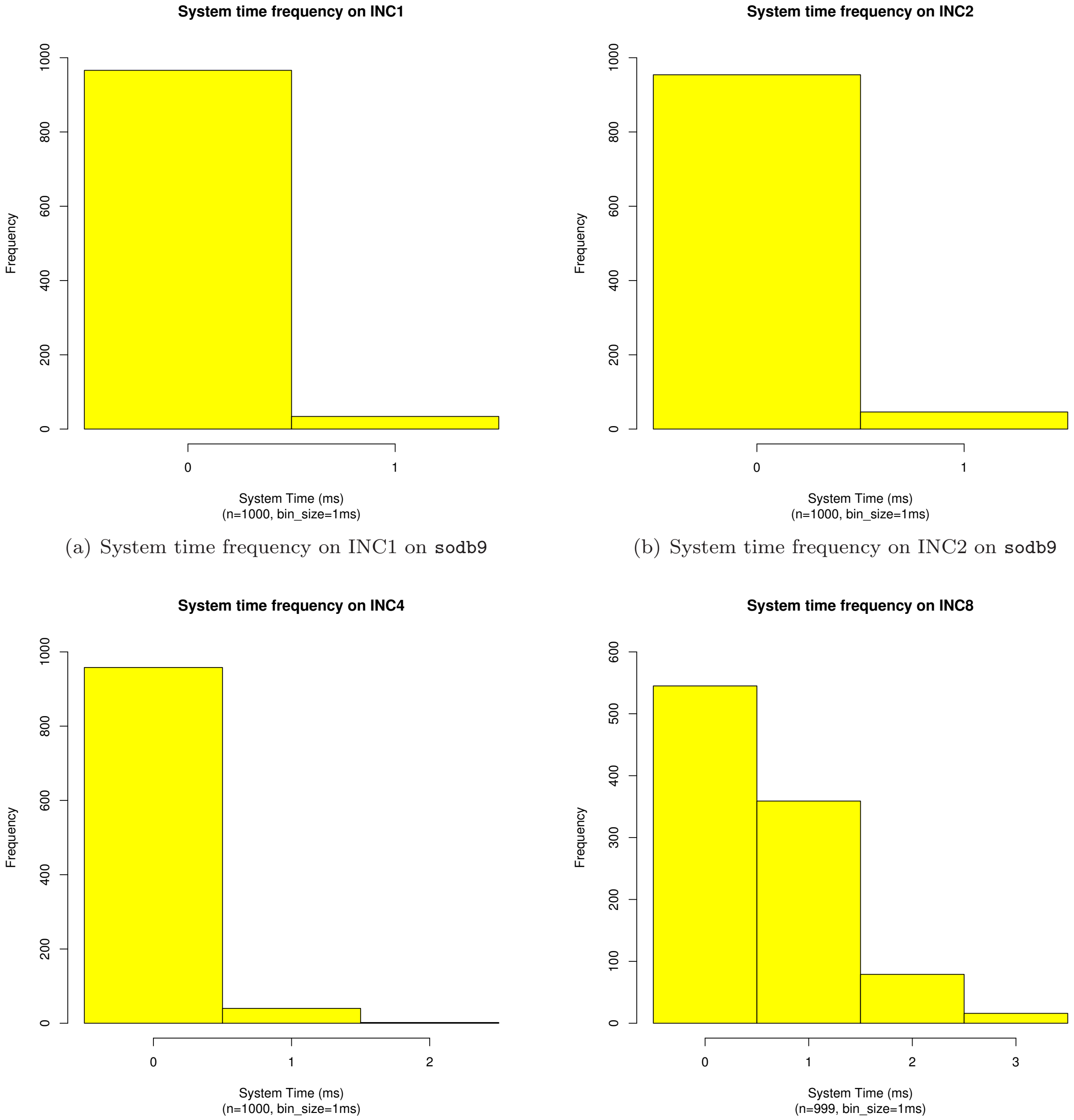


(a) User time frequency on INC2048 on `sodb10`

(b) User time frequency on INC4096 on `sodb12`

Figure 24: User Time Histograms of INC512 ... INC4096

2.5.2 System Time



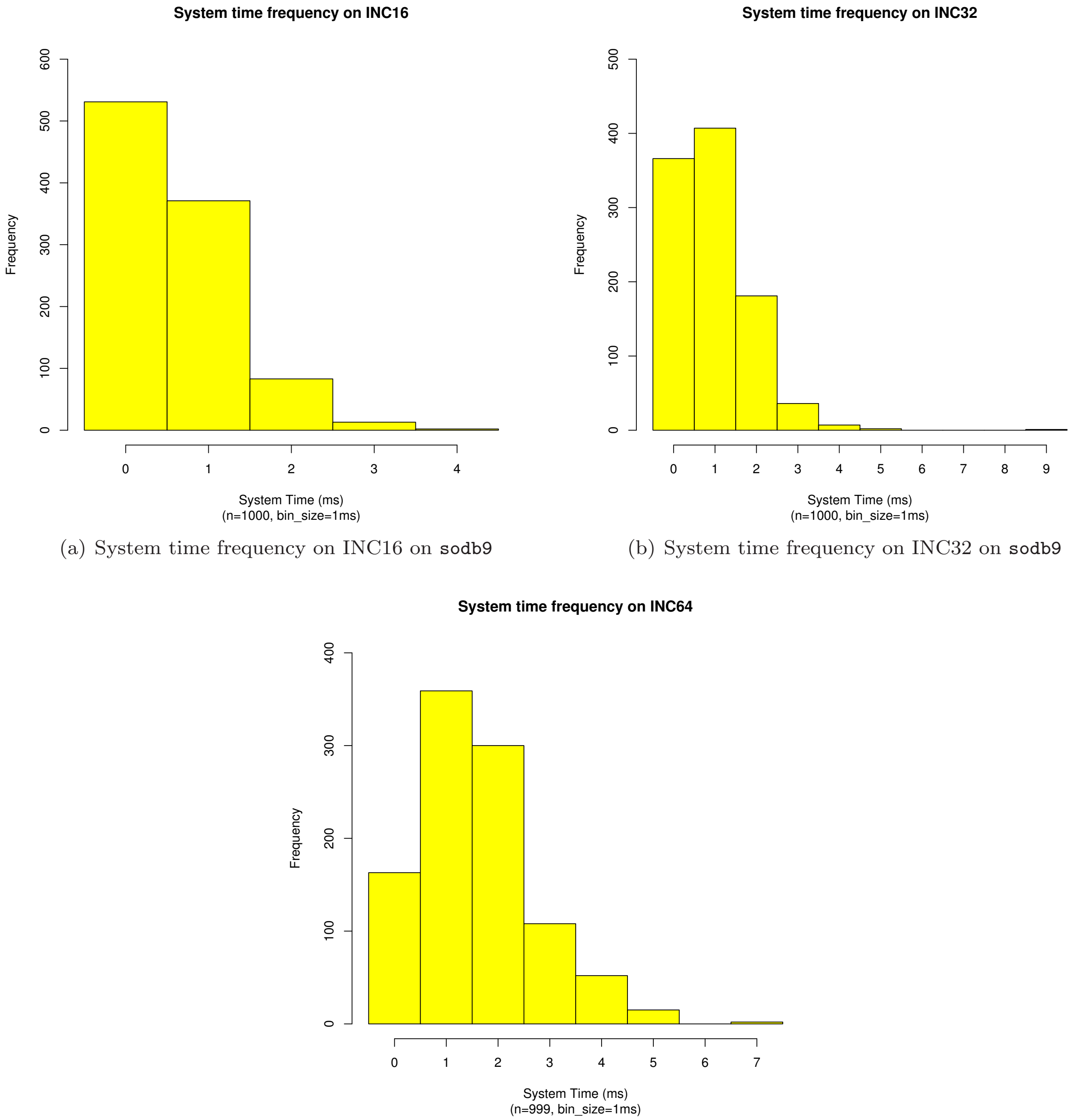
(a) System time frequency on INC1 on `sodb9`

(b) System time frequency on INC2 on `sodb9`

(c) System time frequency on INC4 on `sodb9`

(d) System time frequency on INC8 on `sodb9`

Figure 25: System Time Histograms of INC1 ... INC8

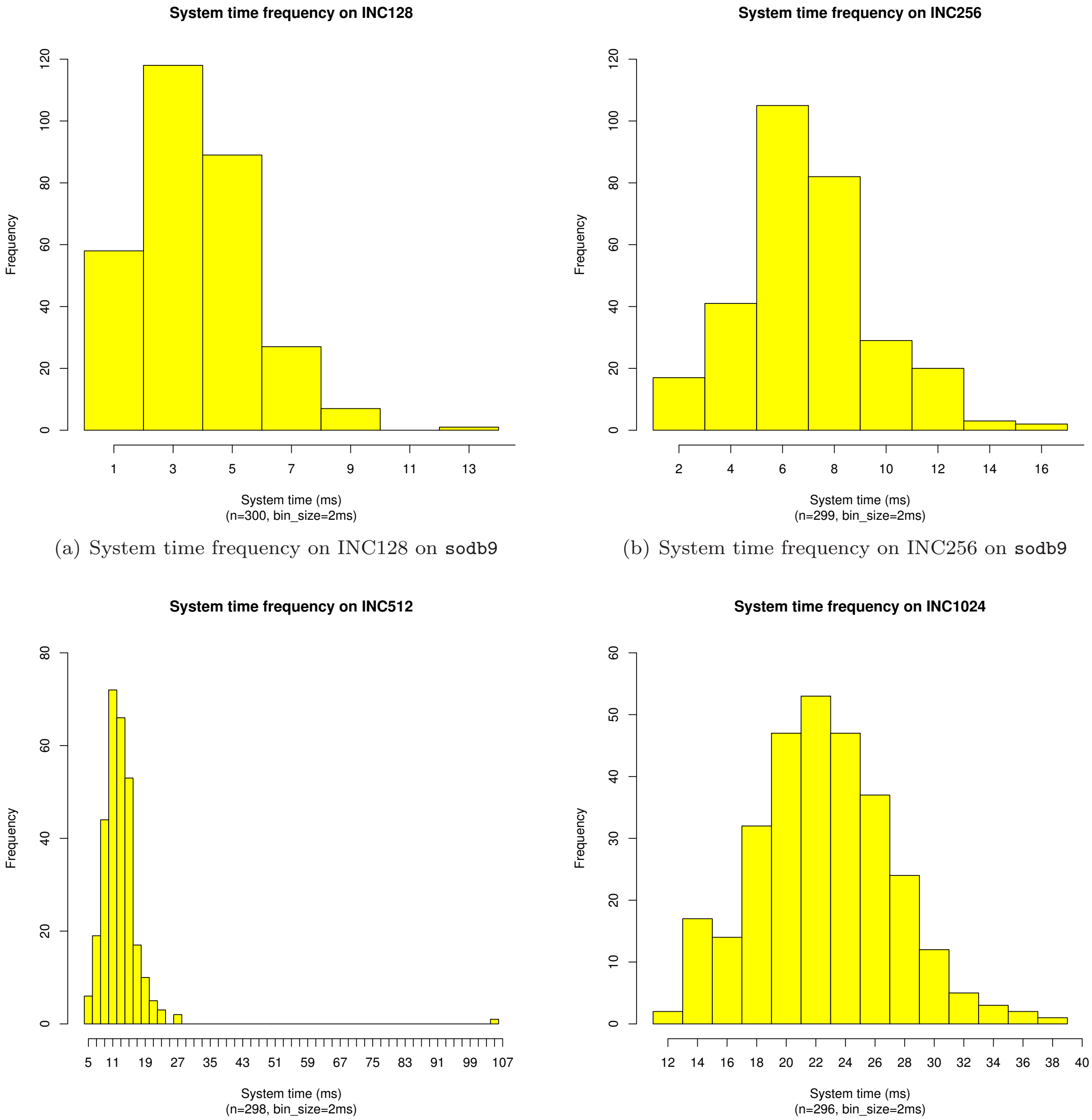


(a) System time frequency on INC16 on `sodb9`

(b) System time frequency on INC32 on `sodb9`

(c) System time frequency on INC64 on `sodb9`

Figure 26: System Time Histograms of INC16 ... INC64



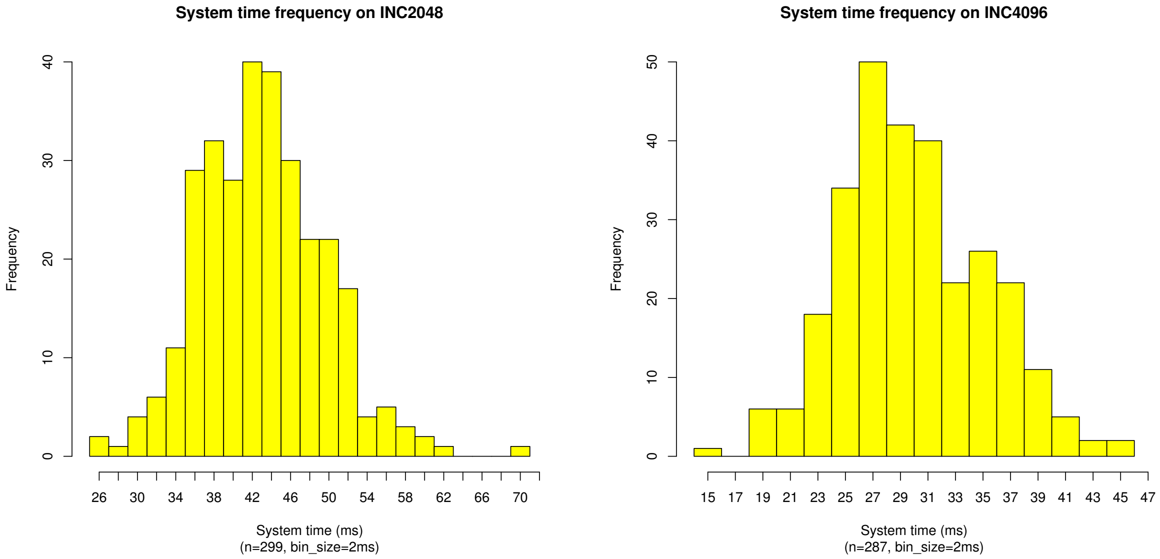
(a) System time frequency on INC128 on `sodb9`

(b) System time frequency on INC256 on `sodb9`

(c) System time frequency on INC512 on `sodb9`

(d) System time frequency on INC1024 on `sodb9`

Figure 27: System Time Histograms of INC256 ... INC1024



(a) System time frequency on INC2048 on `sodb9`

(b) System time frequency on INC4096 on `sodb9`

Figure 28: System Time Histograms of INC2048 and INC4096

2.5.3 Correlational Analysis on Selected Runs

In this section we perform correlational analysis between the user and system time of INC2048 and INC4096 runs and selected measures of the daemons captured on the two runs.

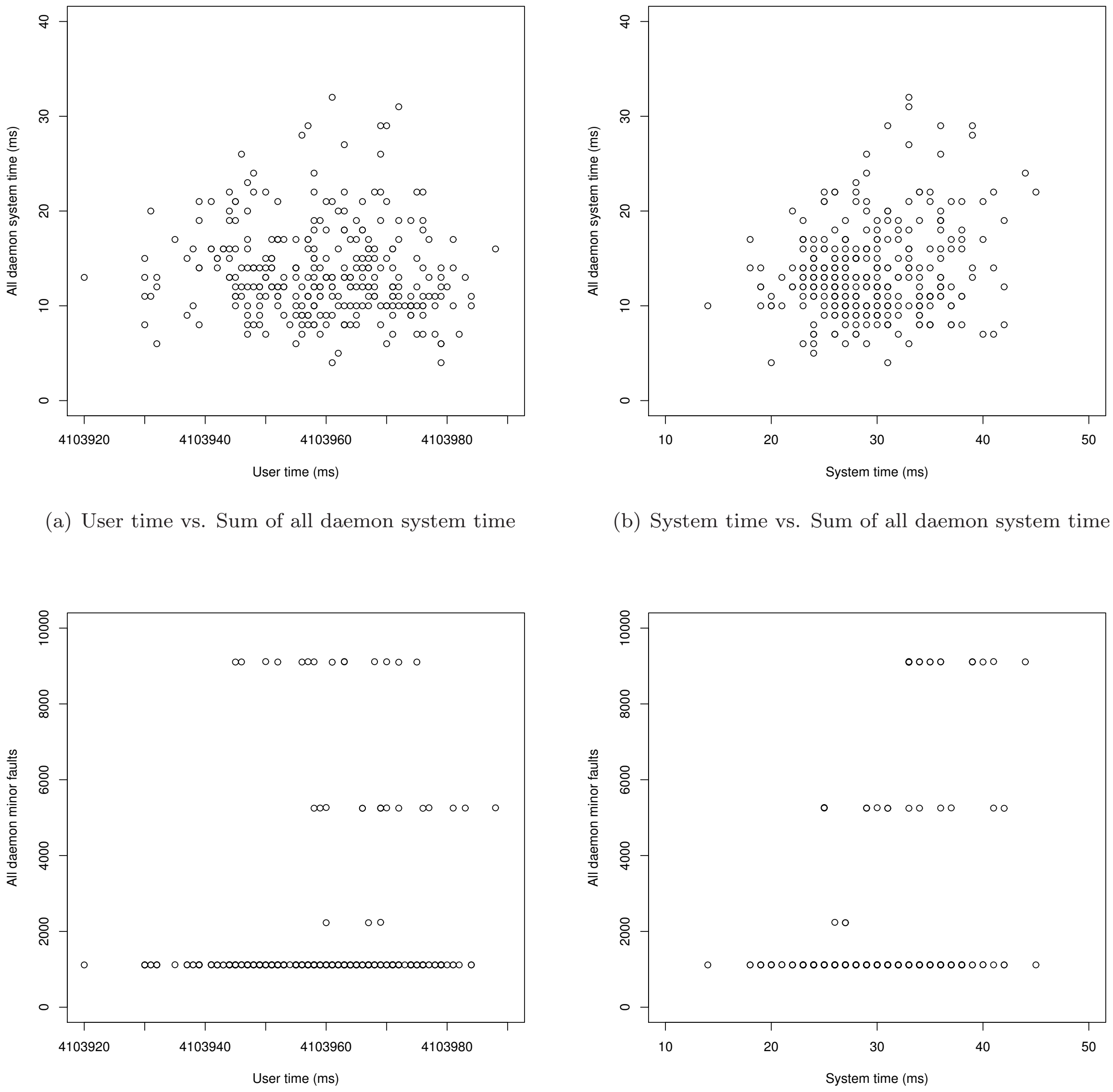
	INC2048's u time	INC2048's s time
INC2048's u time	-	0.05
daemon u time	0.56	0.61
daemon s time	0.48	0.57
daemon minor faults	0.53	0.62
daemon major faults	0.55	0.59
daemon read bytes	0.55	0.59
daemon read char	0.56	0.61
daemon read sys calls	0.57	0.63
daemon write bytes	0.57	0.64
daemon write char	0.53	0.62

Table 4: Correlation of user and system time of INC2048 with some daemon measures

	INC4096's u time	INC4096's s time
INC4096's u time	-	-0.30
daemon u time	0.1	**0.3**
daemon s time	-0.09	**0.19**
daemon minor faults	**0.11**	**0.32**
daemon read char	0.1	**0.32**
daemon read sys calls	0.11	**0.32**
daemon write bytes	0	**0.26**
daemon write char	0.11	**0.32**

Table 5: Correlation of user and system time of INC4096 with some daemon measures

Scatter Plots on Some Significant Correlations: The following scatter plots correspond to the correlations bold in Table 5.



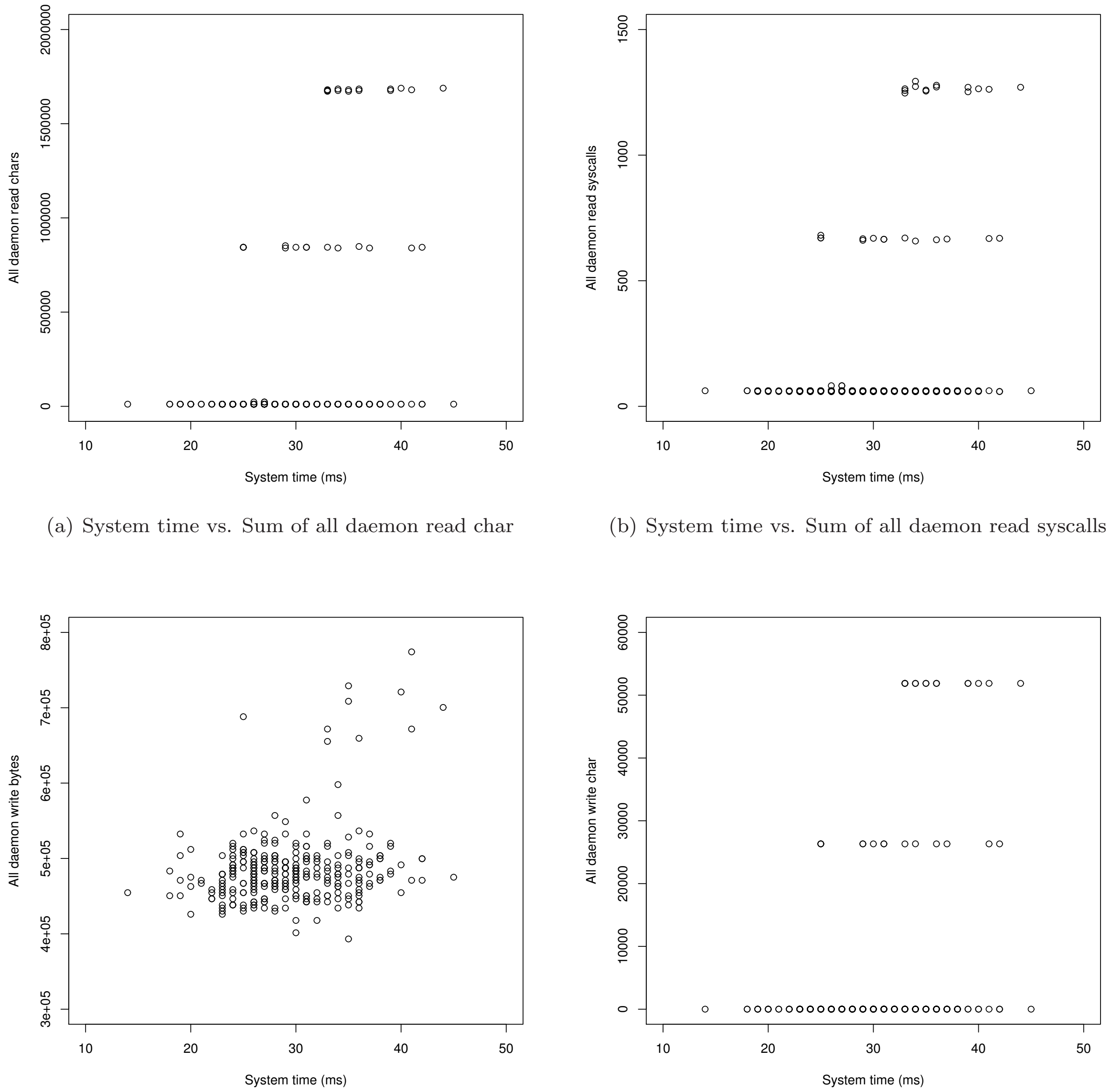
(a) User time vs. Sum of all daemon system time

(b) System time vs. Sum of all daemon system time

(c) User time vs. Sum of all daemon minor faults

(d) System time vs. Sum of all daemon minor faults

Figure 29: Scatter plots between measures on INC4096



(a) System time vs. Sum of all daemon read char

(b) System time vs. Sum of all daemon read syscalls

(c) System time vs. Sum of all daemon write bytes

(d) System time vs. Sum of all daemon write char

Figure 30: Scatter plots between measures on INC4096

Detailed Measures on Some Significant Samples: The following tables show the breakdown of measures of processes observed on some distinct samples (leftmost or rightmost) in the selected runs.

proc name (id)	u time	s time	min flt	maj flt	r bytes	r char	r sysc	w bytes	w char	w sysc
INC4096 (3559)	4103920	36	0	0	0	1	1	0	0	0
proc_monitor (25917)	194	6	2	0	0	0	0	0	0	0
md0_raid1 (484)	0	3	0	0	0	0	0	0	0	0
java (3549)	2	1	1093	0	0	11480	20	0	0	0
cifsd (1927)	0	1	0	0	0	0	0	0	0	0
kblockd/0 (16)	0	1	0	0	0	0	0	0	0	0
ntpd (28232)	0	1	1	0	0	0	42	4096	7	0

Table 6: Observed values of measures of processes on the leftmost sample in Fig. 29(c)

proc name (id)	u time	s time	min flt	maj flt	r bytes	r char	r sysc	w bytes	w char	w sysc
INC4096 (3559)	4103984	19	0	0	0	1	1	0	0	0
proc_monitor (25917)	190	6	2	0	0	0	0	0	0	0
java (3549)	2	1	1093	0	0	11480	20	0	0	0
md0_raid1 (484)	0	3	0	0	0	0	0	0	0	0
ntpd (28232)	0	0	1	0	0	0	39	4096	7	0
java (4108)	0	0	20	0	0	0	0	0	0	0

Table 7: Observed values of measures of processes on the second rightmost sample in Fig. 29(c)

proc name (id)	u time	s time	min flt	maj flt	r bytes	r char	r sysc	w bytes	w char	w sysc
INC4096 (3559)	4103988	25	0	0	0	1	1	0	0	0
proc_monitor (25917)	194	2	2	0	0	0	0	0	0	0
sshd (3609)	8	4	1382	0	0	512357	400	0	20881	0
bash (3611)	4	1	835	0	0	283911	155	0	136	0
md0_raid1 (484)	0	3	0	0	0	0	0	0	0	0
java (3549)	2	1	1093	0	0	11480	20	0	0	0
cifsd (1927)	0	2	0	0	0	0	0	0	0	0
java (3606)	0	1	22	0	0	0	0	0	0	0
jbd2/md0-8 (497)	0	1	0	0	0	0	0	450560	0	0
kblockd/0 (16)	0	1	0	0	0	0	0	0	0	0
grep (3617)	1	0	311	0	0	5417	11	0	0	0
bash (3612)	0	0	158	0	0	0	0	0	0	0
consoletype (3613)	0	0	127	0	0	1956	6	0	7	0
bash (3614)	0	0	174	0	0	0	0	0	0	0
uname (3615)	0	0	189	0	0	1956	6	0	7	0
sshd (3610)	0	0	425	0	0	22656	29	0	4630	0
sshd (2105)	0	0	14	0	0	0	0	0	594	0
bash (3618)	0	0	170	0	0	0	0	0	0	0
id (3619)	0	0	225	0	0	4352	12	0	2	0
ntpd (28232)	0	0	1	0	0	0	42	4096	7	0
bash (3616)	0	0	131	0	0	0	0	0	61	0

Table 8: Observed values of measures of processes on the rightmost sample in Fig. 29(c)

proc name (id)	u time	s time	min flt	maj flt	r bytes	r char	r sysc	w bytes	w char	w sysc
INC4096 (3559)	4103981	14	0	0	0	1	1	0	0	0
proc_monitor (25917)	194	2	2	0	0	0	0	0	0	0
md0_raid1 (484)	0	5	0	0	0	0	0	0	0	0
java (3549)	2	1	1093	0	0	11480	20	0	0	0
flush-9:0 (3548)	0	2	0	0	0	0	0	0	0	0
ntpd (28232)	0	0	1	0	0	0	42	4096	7	0
java (4585)	0	0	20	0	0	0	0	0	0	0

Table 9: Observed values of measures of processes on the leftmost sample in Fig. 30(c)

proc name (id)	u time	s time	min flt	maj flt	r bytes	r char	r sysc	w bytes	w char	w sysc
INC4096 (3559)	4103958	44	0	0	0	1	1	0	0	0
proc_monitor (25917)	194	2	2	0	0	0	0	0	0	0
sshd (4877)	10	4	1382	0	0	519568	392	0	20868	0
sshd (4886)	10	3	1383	0	0	515456	393	0	20868	0
grep (4888)	6	1	994	0	0	287034	154	0	136	0
grep (4879)	3	2	991	0	0	286801	159	0	136	0
java (3549)	2	1	1093	0	0	11480	20	0	0	0
md0_raid1 (484)	0	3	0	0	0	0	0	0	0	0
sshd (4878)	2	1	424	0	0	22403	26	0	4268	0
jbd2/md0-8 (497)	0	2	0	0	0	0	0	696320	0	0
flush-9:0 (3548)	0	2	0	0	0	0	0	0	0	0
cifsd (1927)	0	1	0	0	0	0	0	0	0	0
kblockd/0 (16)	0	1	0	0	0	0	0	0	0	0
grep (4892)	0	1	309	0	0	5417	11	0	0	0
grep (4883)	1	0	310	0	0	5417	11	0	0	0
id (4894)	0	0	226	0	0	4352	12	0	2	0
bash (4884)	0	0	164	0	0	0	0	0	0	0
id (4885)	0	0	226	0	0	0	12	0	2	0
sshd (2105)	0	0	28	0	0	0	0	0	1188	0
sshd (4887)	0	0	425	0	0	22403	26	0	4268	0
ntpd (20232)	0	0	1	0	0	0	42	4096	7	0
bash (4880)	0	0	165	0	0	0	0	0	0	0
uname (4890)	0	0	190	0	0	1956	6	0	7	0
bash (4891)	0	0	128	0	0	0	0	0	61	0
java (4874)	0	0	20	0	0	0	0	0	0	0
bash (4893)	0	0	164	0	0	0	0	0	0	0
uname (4881)	0	0	191	0	0	1956	6	0	7	0
bash (4889)	0	0	165	0	0	0	0	0	0	0

Table 10: Observed values of measures of processes on the second rightmost sample in Fig. 30(c)

proc name (id)	u time	s time	min flt	maj flt	r bytes	r char	r sysc	w bytes	w char	w sysc
INC4096 (3559)	4103948	45	0	0	0	1	1	0	0	0
proc_monitor (25917)	190	10	2	0	0	0	0	0	0	0
md0_raid1 (484)	0	5	0	0	0	0	0	0	0	0
jbd2/md0-8 (497)	0	3	0	0	0	0	0	471040	0	0
java (3549)	2	1	1093	0	0	11480	20	0	0	0
cifsd (1927)	0	2	0	0	0	0	0	0	0	0
flush-9:0 (3548)	0	1	0	0	0	0	0	0	0	0
java (5022)	1	0	21	0	0	0	0	0	0	0
ntpd (28232)	0	0	1	0	0	0	42	4096	7	0

Table 11: Observed values of measures of processes on the rightmost sample in Fig. 30(c)

3 Histograms on the Second Run

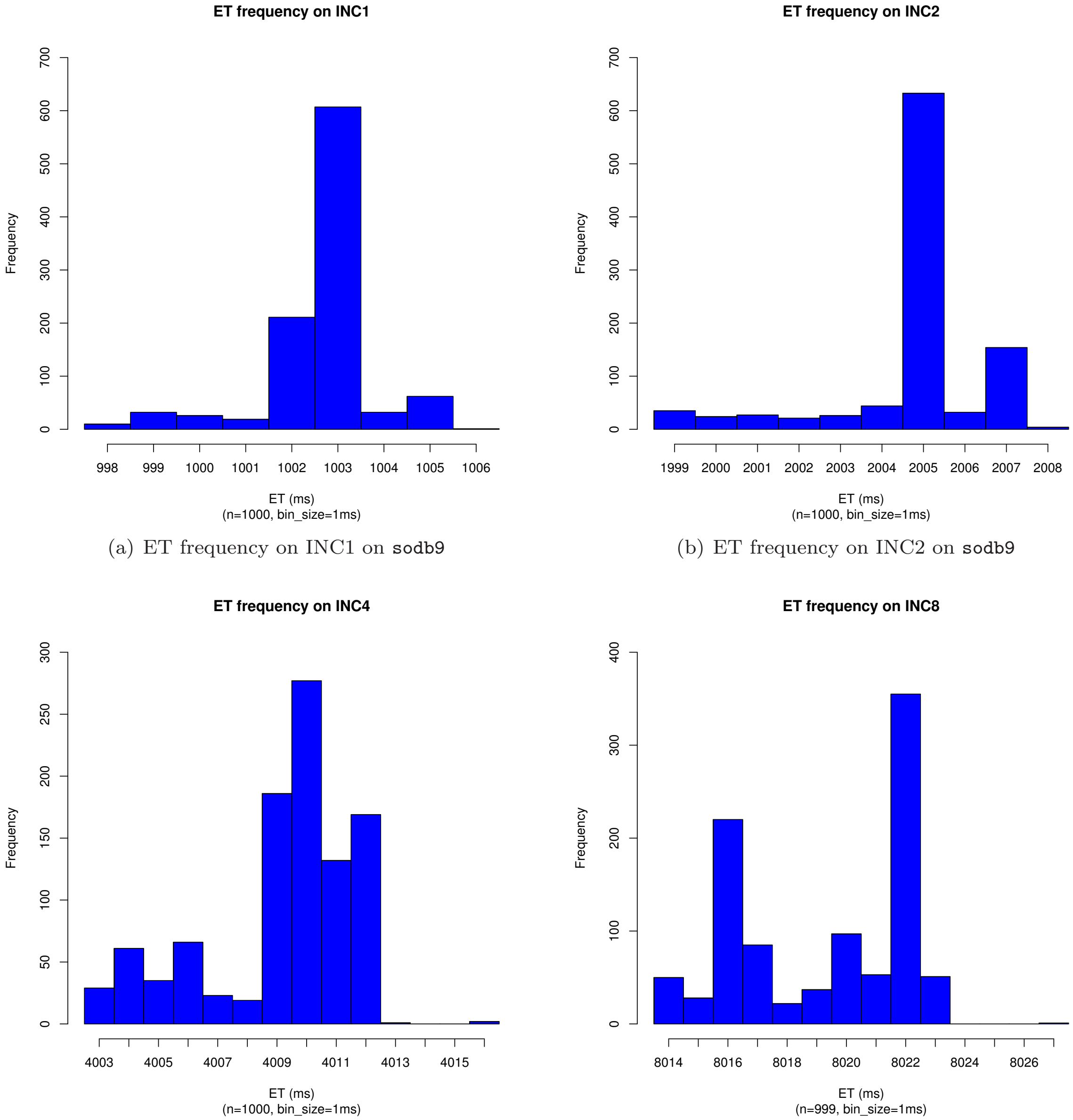
This section exhibits histograms on the second run of INC with its task length increasing from 1 second to 4096 seconds, via EMPv5 without Step 2. The detailed description of the base data is from Table 12.

Machine	Task Length (sec)	Description	Experiment Period	Relevant Histograms
`sodb9` (plugged into *the upper left* power strip)	INC1∼INC64	1000 samples, each	2017-03-13 ∼ 2017-03-14	Figs. 31, 32, 35, and 36
`sodb9` (plugged into *the upper left* power strip)	INC512∼INC1024	300 samples, each	2017-03-17 ∼ 2017-03-21	Figs. 33 and 37
`sodb10` (plugged into *the upper left* power strip)	INC2048	300 samples	2017-03-13 ∼ 2017-03-20	Figs. 34(a) and 38(a)
`sodb12` (plugged into *the upper right* power strip)	INC4096	300 samples	2017-03-02 ∼ 2017-03-17	Figs. 34(b) and 38(b)
`sodb10` (plugged into *the upper left* power strip)	INC8192	261 samples	2017-04-27 ∼ 2017-05-21	Figs. 34(c) and 38(c)
`sodb12` (plugged into *the upper right* power strip)	INC16384	130 samples	2017-04-27 ∼ 2017-05-21	Figs. 34(d) and 38(d)

Table 12: Notes on experiment runs used for histograms

INC8192/INC16384 unfortunately stopped in the middle of their runs due to a frozen vnc problem. So couldn't finish 300 samples.

3.1 ET



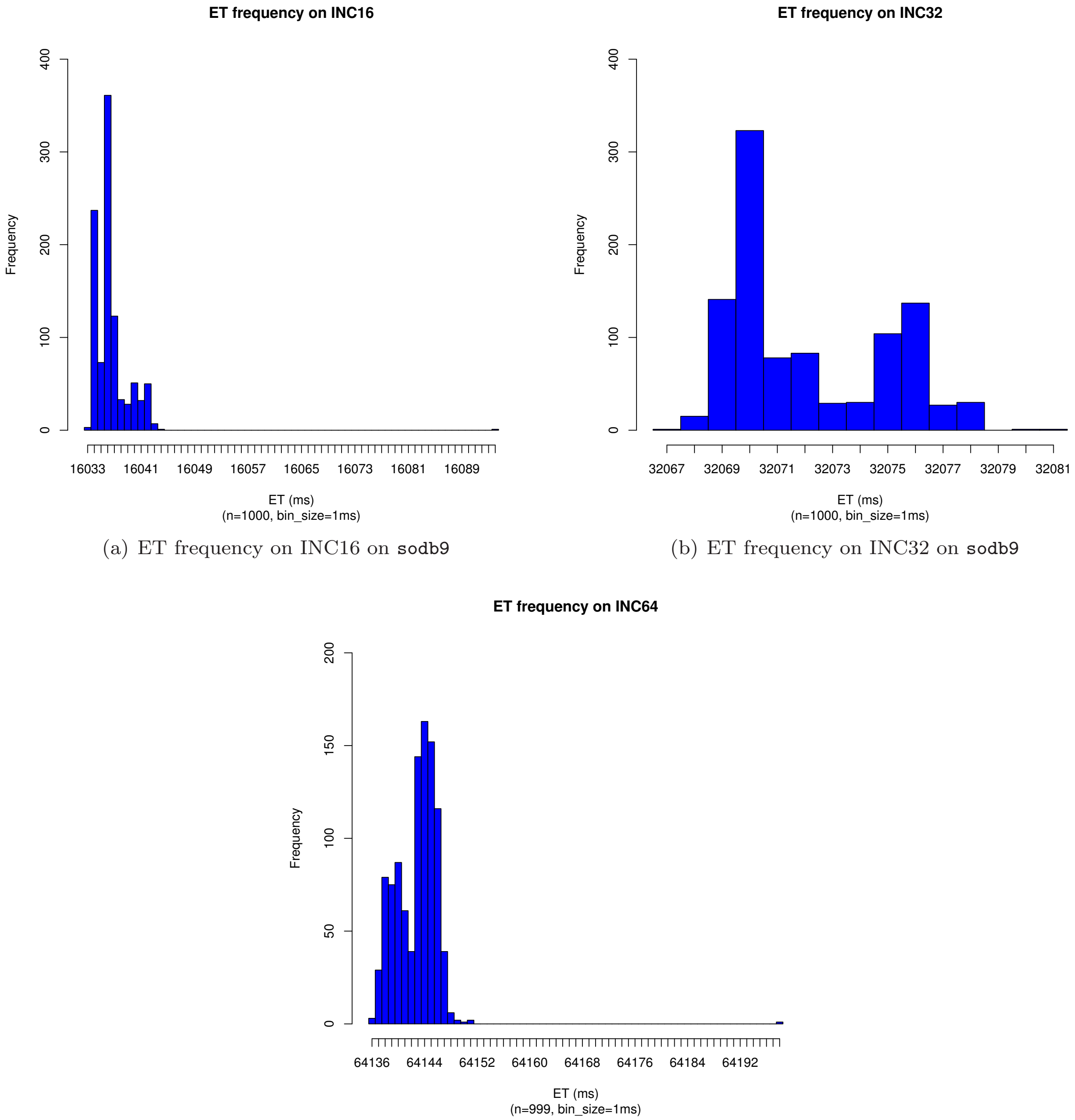
(a) ET frequency on INC1 on `sodb9`

(b) ET frequency on INC2 on `sodb9`

(c) ET frequency on INC4 on `sodb9`

(d) ET frequency on INC8 on `sodb9`

Figure 31: ET Histograms of INC1 ... INC8

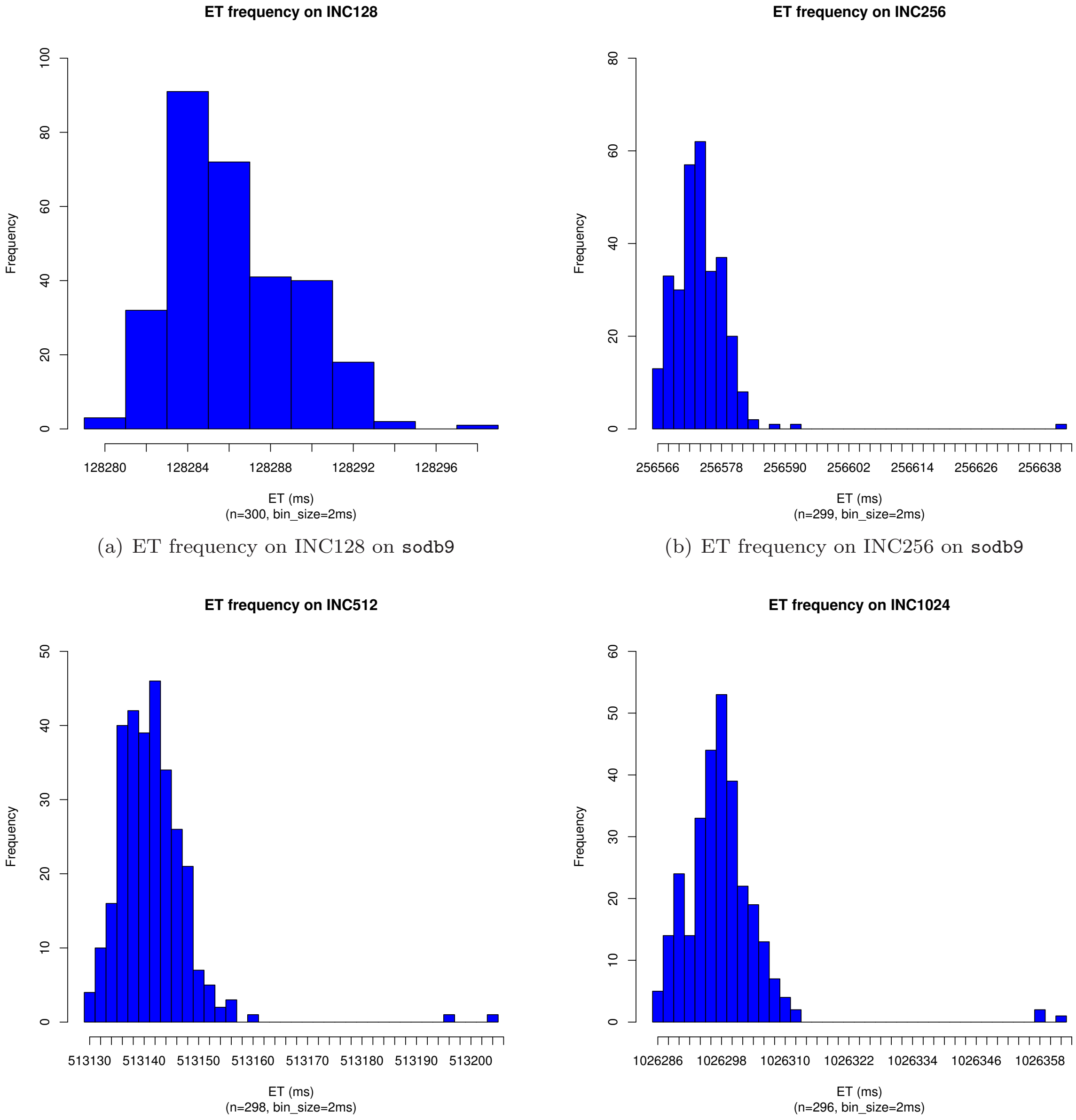


(a) ET frequency on INC16 on `sodb9`

(b) ET frequency on INC32 on `sodb9`

(c) ET frequency on INC64 on `sodb9`

Figure 32: ET Histograms of INC16 ... INC64



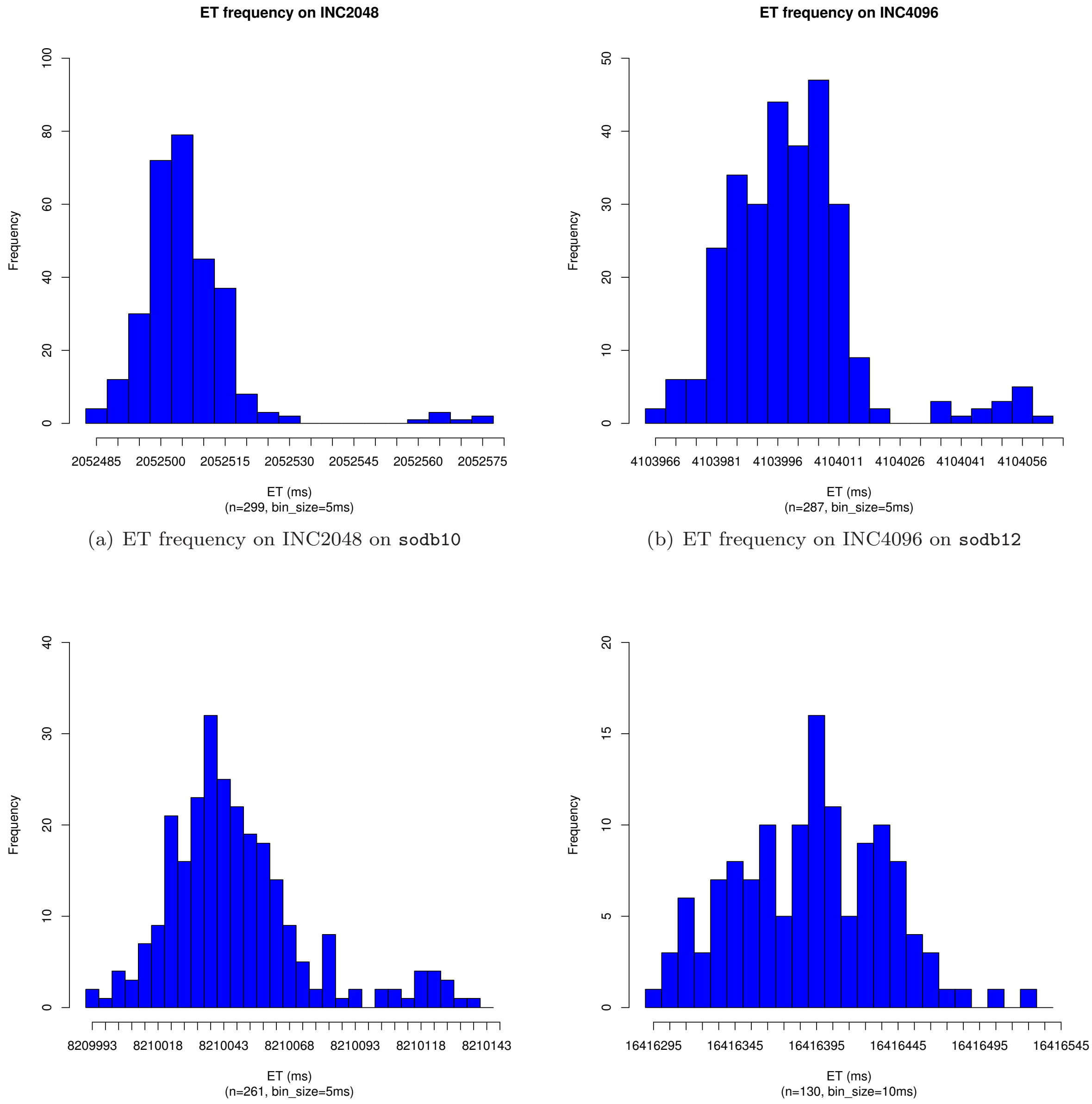
(a) ET frequency on INC128 on `sodb9`

(b) ET frequency on INC256 on `sodb9`

(c) ET frequency on INC512 on `sodb9`

(d) ET frequency on INC1024 on `sodb9`

Figure 33: ET Histograms of INC128 ... INC1024



(a) ET frequency on INC2048 on `sodb10`

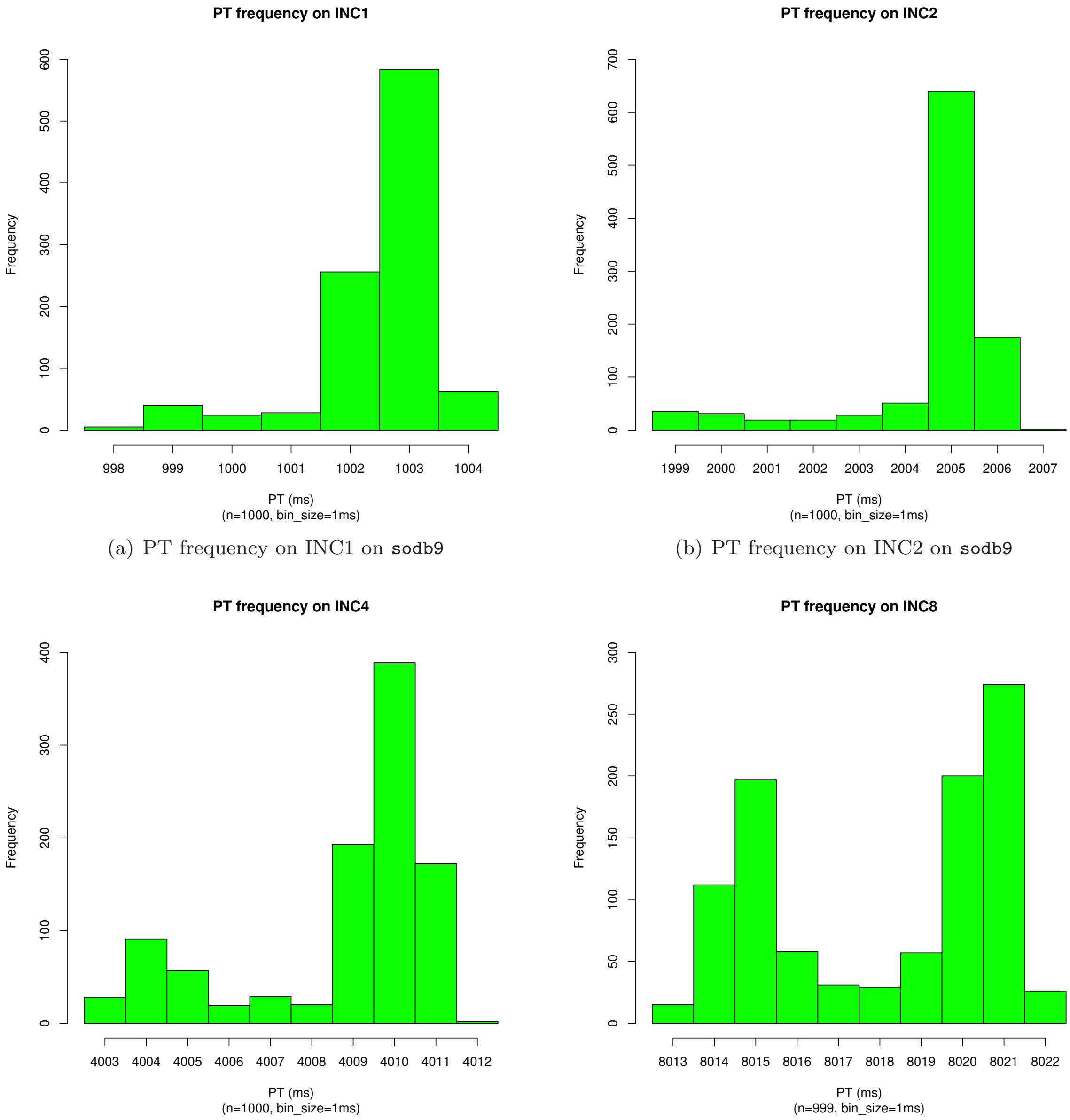
(b) ET frequency on INC4096 on `sodb12`

(c) ET frequency on INC8192 on `sodb10`

(d) ET frequency on INC16384 on `sodb12`

Figure 34: ET Histograms of INC2048 ... INC16384

3.2 PT



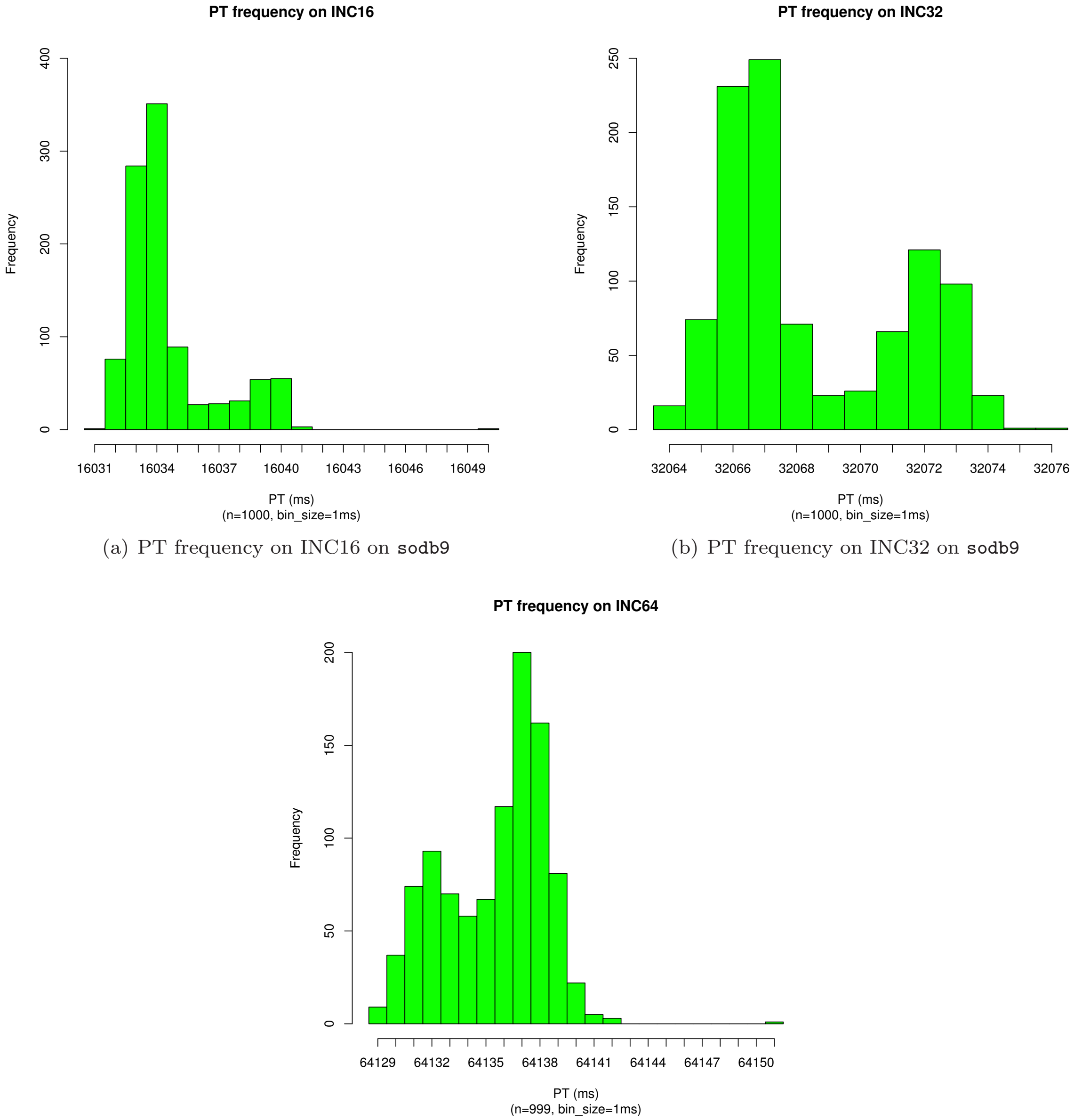
(a) PT frequency on INC1 on `sodb9`

(b) PT frequency on INC2 on `sodb9`

(c) PT frequency on INC4 on `sodb9`

(d) PT frequency on INC8 on `sodb9`

Figure 35: PT Histograms of INC1 ... INC8

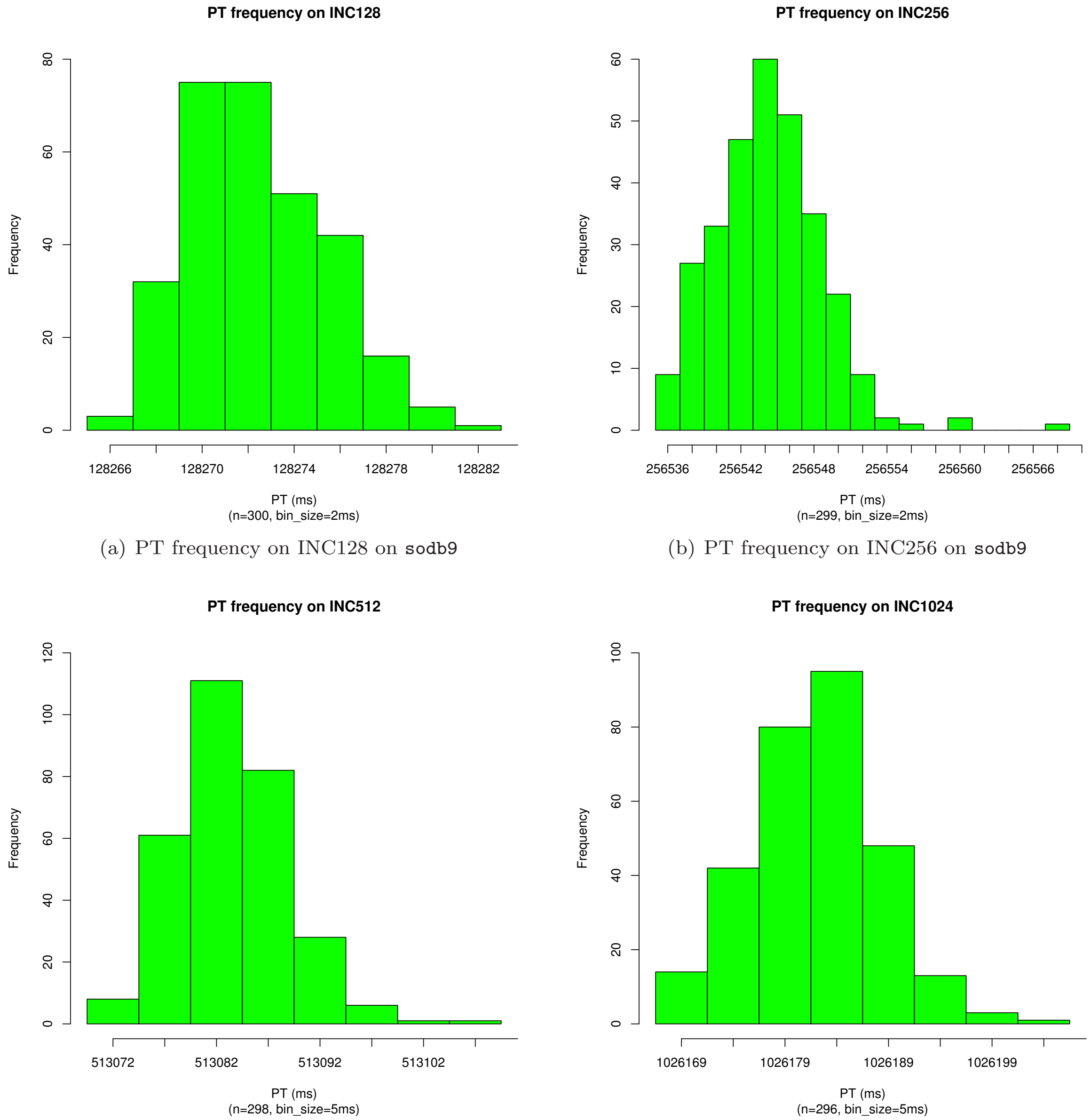


(a) PT frequency on INC16 on `sodb9`

(b) PT frequency on INC32 on `sodb9`

(c) PT frequency on INC64 on `sodb9`

Figure 36: PT Histograms of INC16 ... INC64



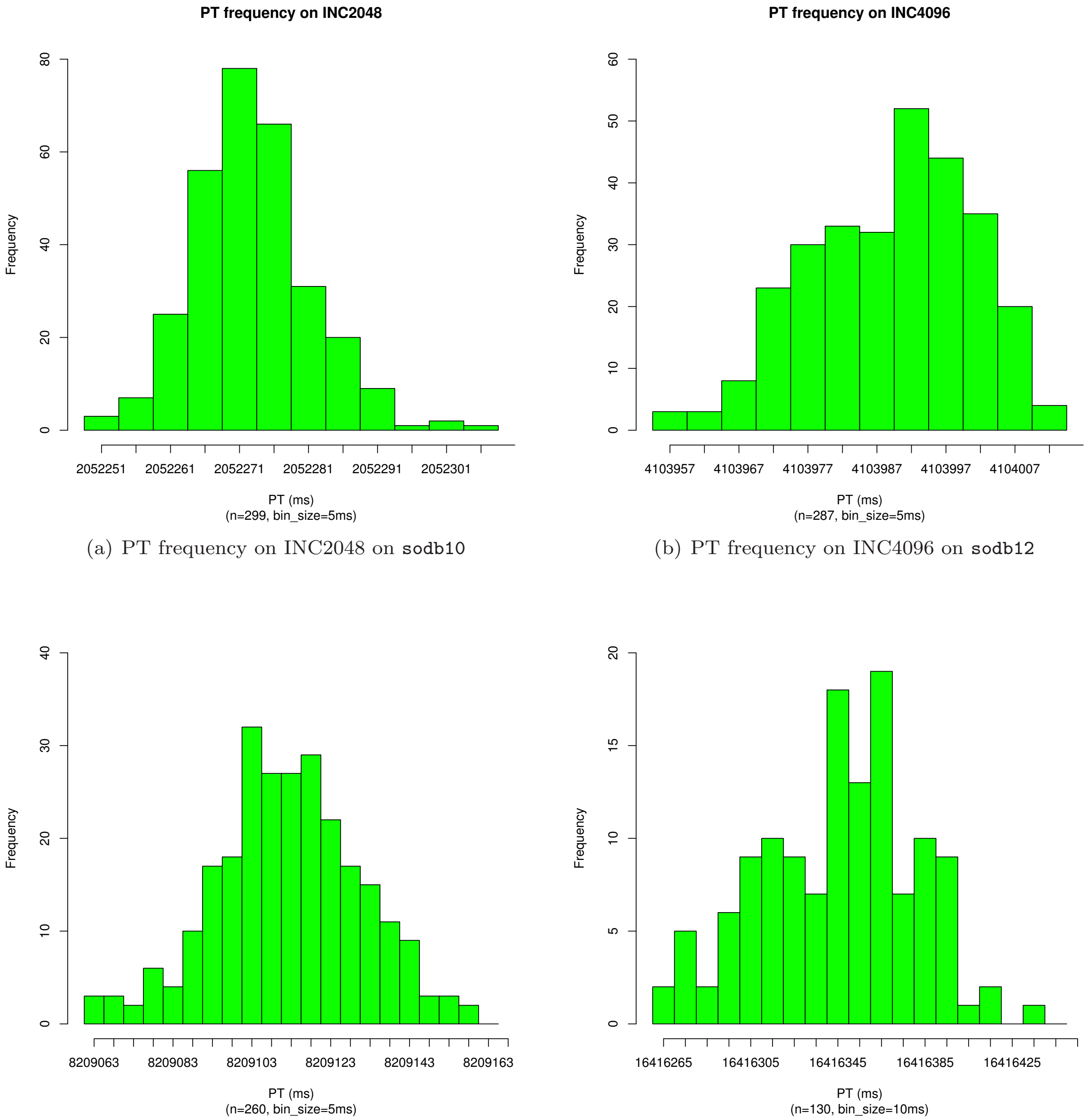
(a) PT frequency on INC128 on `sodb9`

(b) PT frequency on INC256 on `sodb9`

(c) PT frequency on INC512 on `sodb9`

(d) PT frequency on INC1024 on `sodb9`

Figure 37: PT Histograms of INC256 ... INC1024



(a) PT frequency on INC2048 on `sodb10`

(b) PT frequency on INC4096 on `sodb12`

(c) PT frequency on INC8192 on `sodb10`

(d) PT frequency on INC16384 on `sodb12`

Figure 38: PT Histograms of INC2048 ... INC16384

3.2.1 Analysis

In this section we look into what happened inside the peaks observed in a certain histogram. We consider Figure 35(d) for this study. In the figure, we see the peaks at 8015 msec, 8020 msec, and 8021 msec.

Table 13 shows captured daemons and their runtime statistics per bin of figure. Note that bin is at the unit of PT. It appears that the peaks are definitely correlated with (1) appearances of some daemons and (2) times that those daemons co-ran with INC8.

TASK_LEN	BIN (PT)	DAEMON	MIN_PT	MAX_PT	AVG_PT	STD_PT	Counts
INC8	8013	jbd2/md0-8	1	1	1	0	1
INC8	8013	kslowd000	1	1	1	0	1
INC8	8013	md0_raid1	1	1	1	0	17
INC8	8013	proc_monitor	196	200	197.72	1.07	18
INC8	8014	jbd2/md0-8	1	1	1	0	5
INC8	8014	kslowd000	1	1	1	0	35
INC8	8014	kslowd001	1	1	1	0	26
INC8	8014	md0_raid1	1	1	1	0	58
INC8	8014	proc_monitor	196	200	197.31	1.06	95
INC8	**8015**	java	2	7	4.5	3.54	2
INC8	**8015**	jbd2/md0-8	1	1	1	0	2
INC8	**8015**	kslowd000	1	1	1	0	86
INC8	**8015**	kslowd001	1	1	1	0	89
INC8	**8015**	md0_raid1	1	1	1	0	18
INC8	**8015**	proc_monitor	196	200	197.28	1.01	194
INC8	8016	kslowd000	1	1	1	0	36
INC8	8016	kslowd001	1	1	1	0	40
INC8	8016	md0_raid1	1	1	1	0	8
INC8	8016	proc_monitor	196	200	196.45	.95	78
INC8	8017	kslowd000	1	1	1	0	11
INC8	8017	kslowd001	1	1	1	0	10
INC8	8017	md0_raid1	1	1	1	0	3
INC8	8017	proc_monitor	196	200	197.15	1.16	26
INC8	8018	kslowd000	1	1	1	0	13
INC8	8018	kslowd001	1	1	1	0	9
INC8	8018	md0_raid1	1	1	1	0	6
INC8	8018	proc_monitor	196	200	197.24	1.27	29
INC8	8019	jbd2/md0-8	1	1	1	0	3
INC8	8019	kslowd000	1	1	1	0	9
INC8	8019	kslowd001	1	2	1.06	.24	18
INC8	8019	md0_raid1	1	1	1	0	27
INC8	8019	proc_monitor	196	200	197.1	1.18	52
INC8	**8020**	jbd2/md0-8	1	1	1	0	8
INC8	**8020**	kslowd000	1	1	1	0	52
INC8	**8020**	kslowd001	1	1	1	0	57
INC8	**8020**	md0_raid1	1	1	1	0	91
INC8	**8020**	proc_monitor	196	200	197.03	1.02	180
INC8	**8021**	cifsd	1	1	1	0	1
INC8	**8021**	java	2	37	19.5	24.75	2
INC8	**8021**	kslowd000	1	1	1	0	146
INC8	**8021**	kslowd001	1	1	1	0	143
INC8	**8021**	md0_raid1	1	1	1	0	11
INC8	**8021**	proc_monitor	196	198	197.15	.98	299
INC8	8022	kslowd000	1	1	1	0	20
INC8	8022	kslowd001	1	1	1	0	9
INC8	8022	proc_monitor	196	198	196.07	.37	29

Table 13: Daemons observed from the INC8 run

3.2.2 Analysis on Some Outliers Found from INC16 and INC64 on sodb9

In this section we investigate what happened about some outlying samples on INC16 and INC64 shown in Figures 36(a) and 36(c).

Regarding INC16, we looked into detailed information of processes that were captured at one execution corresponding to the bin of 16,050 (frequency = 1) msec in PT. For the purpose of comparison, we also explored processes captured at the second-highest bin of 16,041 msec in PT.

Table 14 shows the overall information of the captured processes at the two bins. As indicated by the table, there were some infrequent daemon processes such as sshd, grep, md0_raid1, cifsd, running for some time when the INC16 program was run.

Iteration number	Process ID	Process Name	PT (in msec)
77 (, 209, or 569)	10185	incr_work (=INC16)	16,041
	28525	proc_monitor	196
	167	kslowd001	1
	166	kslowd000	1
703	10185	incr_work (=INC16)	16,050
	28525	proc_monitor	196
	12322	sshd	14
	12313	sshd	13
	12324	grep	6
	472	md0_raid1	2
	12314	sshd	2
	12315	grep	2
	12319	grep	1
	167	kslowd001	1
	12323	sshd	1
	166	kslowd000	1
	12328	grep	1
	1925	cifsd	1

Table 14: Daemons observed at the second-rightmost and rightmost bins in Figure 36(a)

Regarding INC64, we looked into detailed information of processes that were captured at one execution corresponding to the bin of 64,151 (frequency = 1) msec in PT. For the purpose of comparison, we also explored processes captured at the second-highest bin of 16,041 msec in PT.

Table 15 shows the overall information of the captured processes at the two bins. As indicated by the table, there were also the same kind of infrequent daemon processes such as sshd, grep, md0_raid1 as well as java running for some time when the INC64 program was run. It is conjectured that for a longer INC program, it is more frequent to have java with some positive time in PT.

Iteration number	Process ID	Process Name	PT (in msec)
795 (, 125, or 236)	16256	incr_work (=INC64)	16,142
	28525	proc_monitor	198
	167	kslowd001	4
	166	kslowd000	3
(125)	(472)	(md0_raid1)	(2)
	16246	java	2
759	16256	incr_work (=INC64)	64,151
	28525	proc_monitor	198
	19320	sshd	13
	19329	sshd	12
	19331	grep	5
	19322	grep	5
	167	kslowd001	3
	166	kslowd000	3
	16246	java	2
	2067	sshd	2
	19335	grep	1
	19330	sshd	1
	472	md0_raid1	1
	19326	grep	1

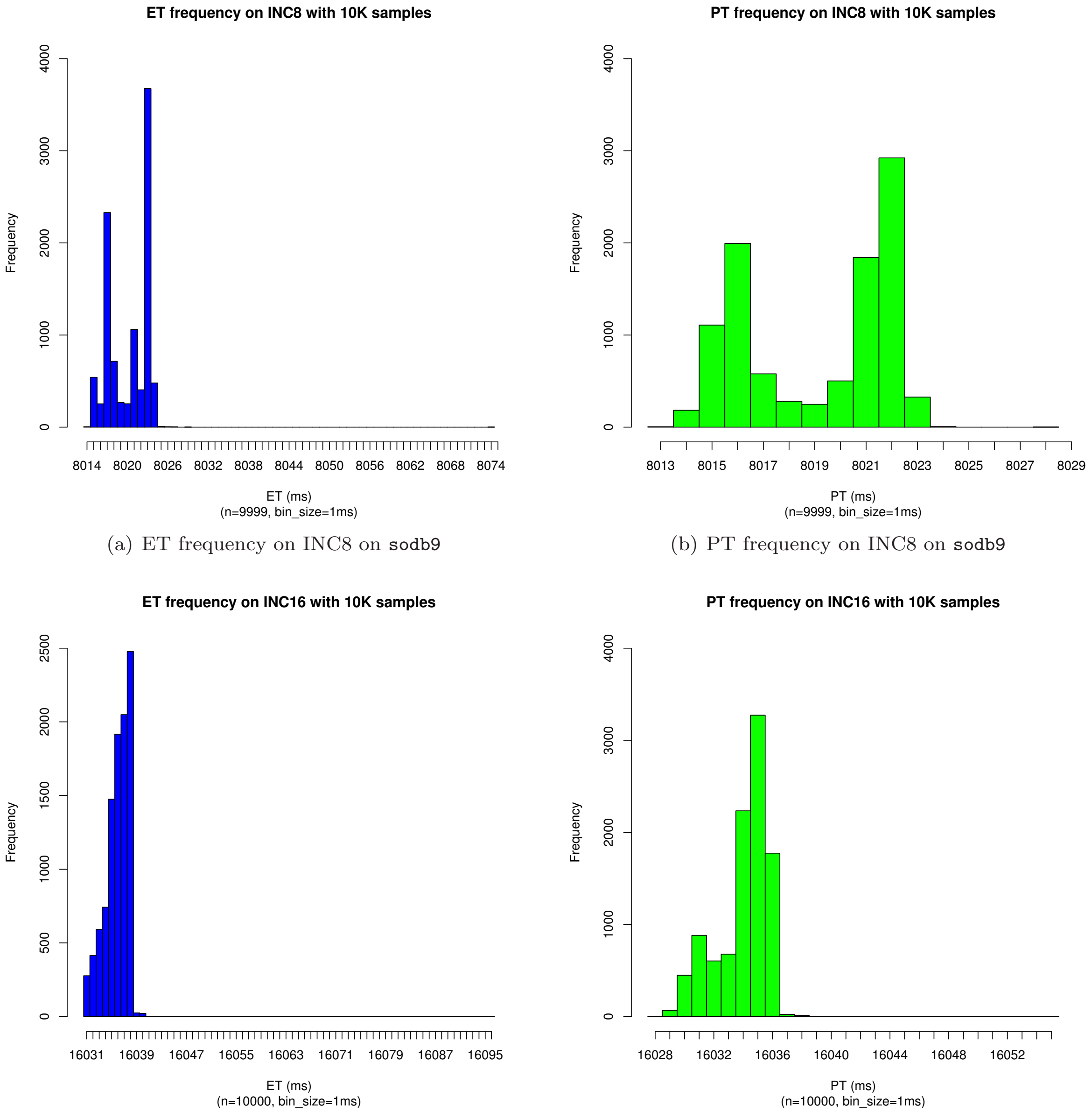
Table 15: Daemons observed at the second-rightmost and rightmost bins in Figure 36(c)

4 Histograms with 10,000 samples

This section exhibits histograms on two runs of INC, each with 8 and 16 seconds as its task length, having 10,000 repetitions. The detailed description of the base data is from Table 16.

Machine	Task Length (sec)	Description	Experiment Period	Relevant Histograms
`sodb9` (plugged into *the upper left* power strip)	INC8	10,000 samples	2017-03-29 ∼ 2017-03-30	Figs. 39(a) and 39(b)
`sodb10` (plugged into *the upper left* power strip)	INC16	10,000 samples	2017-03-29 ∼ 2017-03-31	Figs. 39(c) and 39(d)

Table 16: Notes on experiment runs used for histograms



(a) ET frequency on INC8 on `sodb9`

(b) PT frequency on INC8 on `sodb9`

(c) ET frequency on INC16 on `sodb10`

(d) PT frequency on INC16 on `sodb10`

Figure 39: Histograms of INC8 and INC16 with 10,000 samples

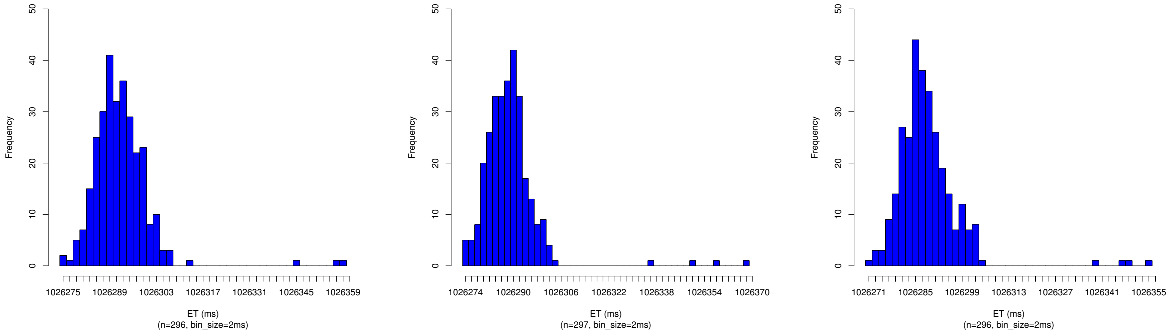
5 Histograms on Consecutive INC1024 Runs

This section exhibits histograms on (three) consecutive runs of INC1024. The detailed description of the base data is from Table 12.

Machine	Task Length (sec)	Description	Experiment Period	Relevant Histograms
`sodb9` (plugged into *the upper left* power strip)	INC1024	300 samples, each	2017-04-12 ∼ 2017-04-23	Figs. 40 and 41

Table 17: Notes on experiment runs used for histograms

5.1 ET



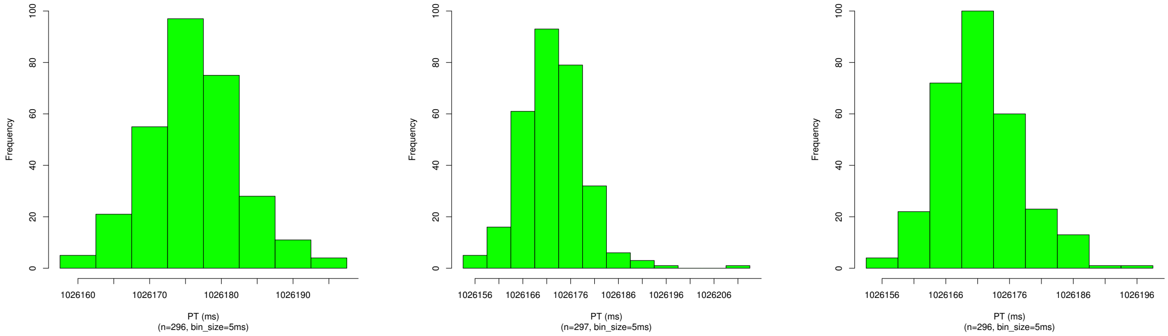
(a) ET frequency on INC1024-Run1 on `sodb9`

(b) ET frequency on INC1024-Run2 on `sodb9`

(c) ET frequency on INC1024-Run3 on `sodb9`

Figure 40: ET Histograms of Three Consecutive INC1024 Runs

5.2 PT



(a) PT frequency on INC1024-Run1 on `sodb9`

(b) PT frequency on INC1024-Run2 on `sodb9`

(c) PT frequency on INC1024-Run3 on `sodb9`

Figure 41: PT Histograms of Three Consecutive INC1024 Runs

References

[1] Young-Kyoon Suh, Richard T. Snodgrass, John Kececioglu, Peter J. Downey, Rob S. Maier, and Cheng Yi, “EMP: Execution Time Measurement Protocol for Compute-Bound Programs”, in *Software: Practice and Experience*, 47(4):559–597, 2017.

[2] Sabah Currim,Richard T. Snodgrass, Young-Kyoon Suh, and Rui Zhang, “DBMS Metrology: Measuring Query Time”, in *ACM Transactions on Database Systems*, 42(1):3:1–42(+8), 2017.