



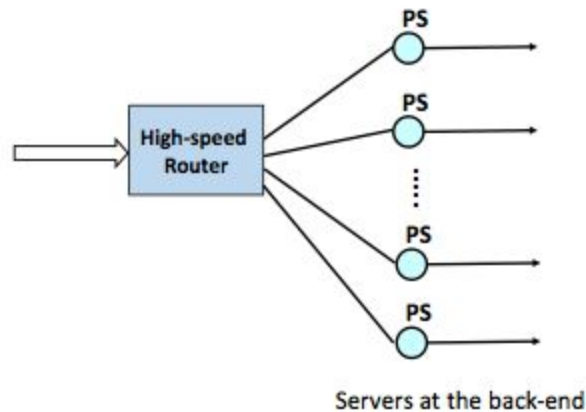
## **COMP9334 Capacity Planning of Computer System and Network**

### **Project 1 Server Farms**

**Z3492782**  
**Li Yu**

# 1.Object

- Simulate a Processor Sharing (PS) server farm and decide how many server should be opened in order to achieve the minimal response time.
- Use statistically sound method to compare the performance of different systems and find out which solution is better.



## 2. Project condition

- Server number:  $s \leq 10$
- Power budget:  $p = 2000$
- Clock frequency:  $f = 1.25 + 0.31 * (p/200 - 1)$  (measured in GHz)
- Under Round robin job assignment

## 3. Parameter

- The inter-arrival time:  $\{a_1, a_2, \dots, a_k, \dots, \dots\}$ 
  - 1) Each  $a_k$  is the product of two random numbers  $a_{1k}$  and  $a_{2k}$ ,  $a_k = a_{1k} a_{2k}$
  - 2) The sequence  $a_{1k}$  is exponentially distributed with a mean arrival rate 7.2 requests/s.
  - 3) The sequence  $a_{2k}$  is uniformly distributed in the interval  $[0.75, 1.17]$
- The service time:
  - 1) If the server is operating at 1 GHz, then the probability density function  $g(t)$  of the service time  $t$  of the requests is:

$$g(t) = \begin{cases} 0 & \text{for } t \leq \alpha_1 \\ \frac{\gamma}{t^\beta} & \text{for } \alpha_1 \leq t \leq \alpha_2 \\ 0 & \text{for } t \geq \alpha_2 \end{cases} \quad (2)$$

where  $\alpha_1 = 0.6$  0.43,  $\alpha_2 = 8$  0.98,  $\beta = 0.86$ , and

$$\gamma = \frac{1 - \beta}{\alpha_2^{1-\beta} - \alpha_1^{1-\beta}}$$

2) Compute the indefinite integral of  $g(t)$  to get  $G(t)$  which is cumulative distribution function (CDF). Let  $G(\alpha_1) = 0$  or  $G(\alpha_2) = 1$  to get the constant  $C$ . The result is shown in the formula below.

$$G(t) = \frac{\gamma}{1-\beta} * (t^{1-\beta} - \alpha_1^{1-\beta})$$

$$\text{where } r = 1.2889, \beta = 0.86, \alpha_1 = 0.43$$

3)  $G(t)$  is the generated uniformly distributed in the interval  $[0,1]$ , the function of service time  $t$ :

$$t = \sqrt[1-\beta]{\frac{1-\beta}{\gamma} * G(t) + \alpha_1^{1-\beta}}$$

$$\text{Where } r = 1.2889, \beta = 0.86, \alpha_1 = 0.43, G(t) \text{ in } [0,1]$$

- Number of servers opened :  $0 < s \leq 10$
- Simulation time :  $T_{end} = 4000$
- Replication times :  $r = 5, 10, 15, 20, 25, 30$
- Removed jobs number:  $w = 1, 500, 1000$

## 4. Simulation process

- Use Matlab to simulate and analysis the data of systems.
  - 1)  $s$  = the number of servers which are opened,  $s = 3$  to  $10$
  - 2)  $T_{end}$  = simulation time,  $T_{end} = 4000$
  - 3) replication\_id = replication number, from  $5, 10, 15, 20, 25$ , to  $30$
  - 4)  $w$  = how many jobs are non-steady,  $w = 1000$
  - 5)  $T_{ave}$  = the mean response time
- The functions and output data (files) are shown in the table below.

main functions		seed file or load file	input	output	Note
simulate	sim_PS	p1_rand_setting_arrival	s Tend	Tave file1_running_mean_7_4000	sub functions: find_serverNUM.m add_job_func.m delete_departure_func.m update_job_list_func.m update_next_departure_list_func
analysis steady values	transient	file1_running_mean_7_4000	running_mean jobs_num	plot a graph for analysing w	vary the value of w: from 1,100,500 to 1000
independent replications (Common random numbers method for s=3 to 10)	replication_random	-	Tend repli_times w	data1_replications_15	call sim_PS_remove_transient
	sim_PS_remove_transient	15 independent seeds	s Tend replication_id w	Tave file2_running_mean_transient_7_4000 running_mean_transient	Compare when System s = 3 to 10. this function will be called by "replication_random" for making replications
	interval_plot	data1_replications_15	from server num to server num eg.,(3,10), (5,8)	plot the graph of response time	-
independent replications (Common random numbers method comparing s=6 with s=7)	replication_common	-	Tend repli_times w	data2_replications_15_common data2_replications_30_common	call sim_PS_remove_transient_common
	sim_PS_remove_transient_common	common seed	s=6,s=7	Tave file3_running_mean_transient_7_4000_common	compare System s= 6 and System s= 7 this function will be called by "replication_random" for making replications
	interval_plot_common	data2_replications_15_common data2_replications_30_common	server 6 server 7 eg.,(6,7)	plot the graph of response time	-
analysis	table_Confidence_interval	data1_replications_15 data2_replications_15_common data2_replications_30_common		list = [lower; upper; T ; S ; ] Repli_list = [i,lower_diff, upper_diff, ]	for Common random numbers method ANALYSIS

- Code and file are in attachments

## 5. Analysis of the simulation results

### 5.1 Transient removal process

- Steady state value

1) When s servers opened, each job's response time continuously varies. Response time does not settle to a constant value. But, mean response time does settle.

2) The running mean is used for analysing the systems :

- Let us assume that the first  $m$  jobs constitute the transient part and there are  $N$  jobs altogether, we should revise the formula to compute the mean to

$$\frac{X(m+1) + X(m+2) + \dots + X(N)}{N - m}$$

3) Steady state value is the key to analysis different systems. By excluding the transient part of the data, the mean response time is more reliable to use.

- Choosing the number of w (the first jobs constitute the transient part)

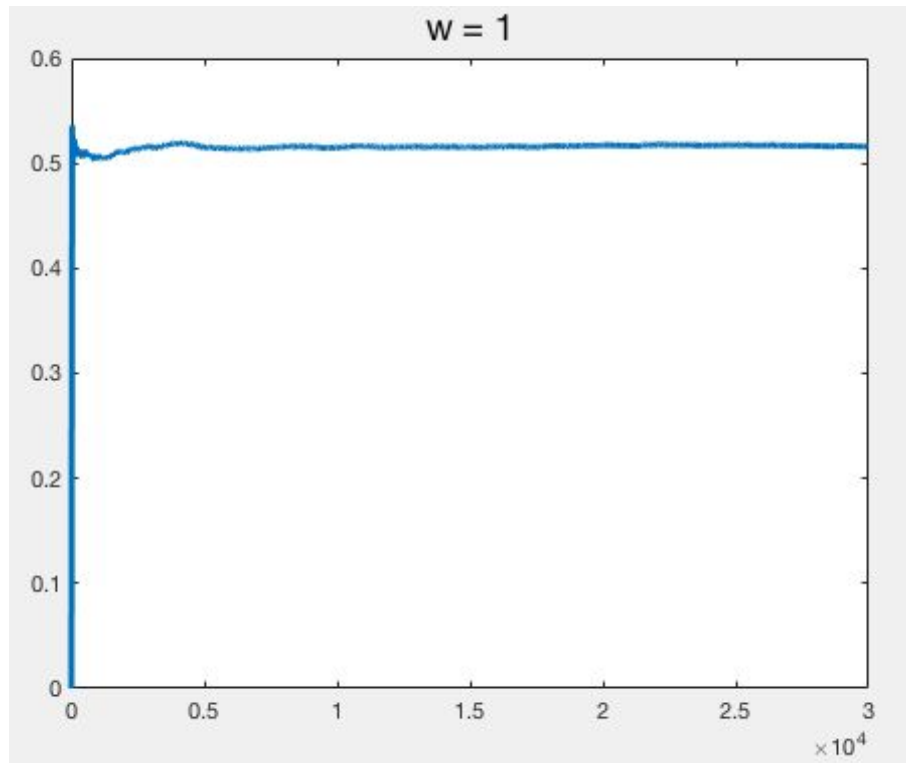
1) Condition :

s = 7 (s servers opened)

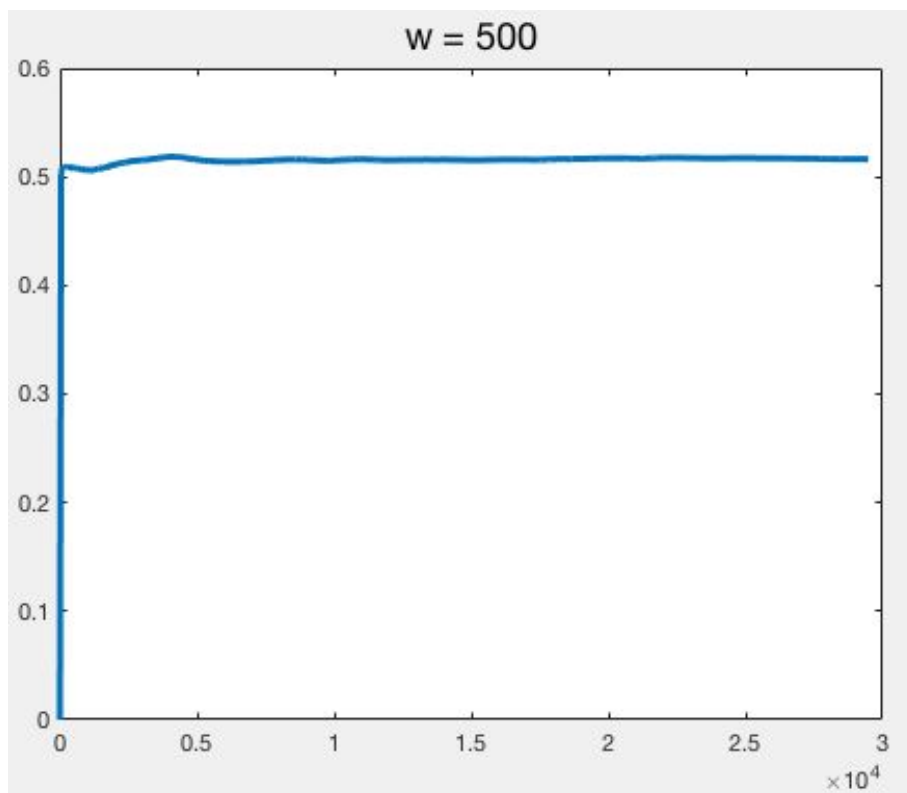
Tend = 4000 (the simulation period)

2) The graph of the running mean response time when choosing different value for  $w$  (the first jobs constitute the transient part) show below:

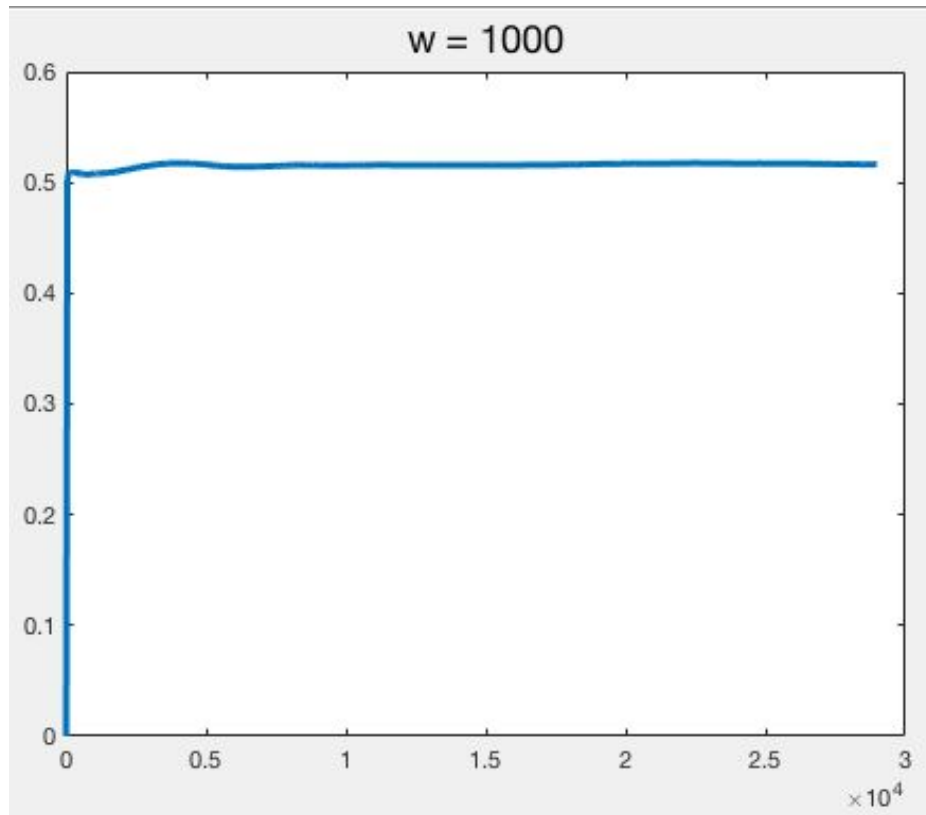
When  $w = 1$ , the graph shows there are non-steady state behaviour:



When  $w = 500$ , the line become more smooth but still not ideal.

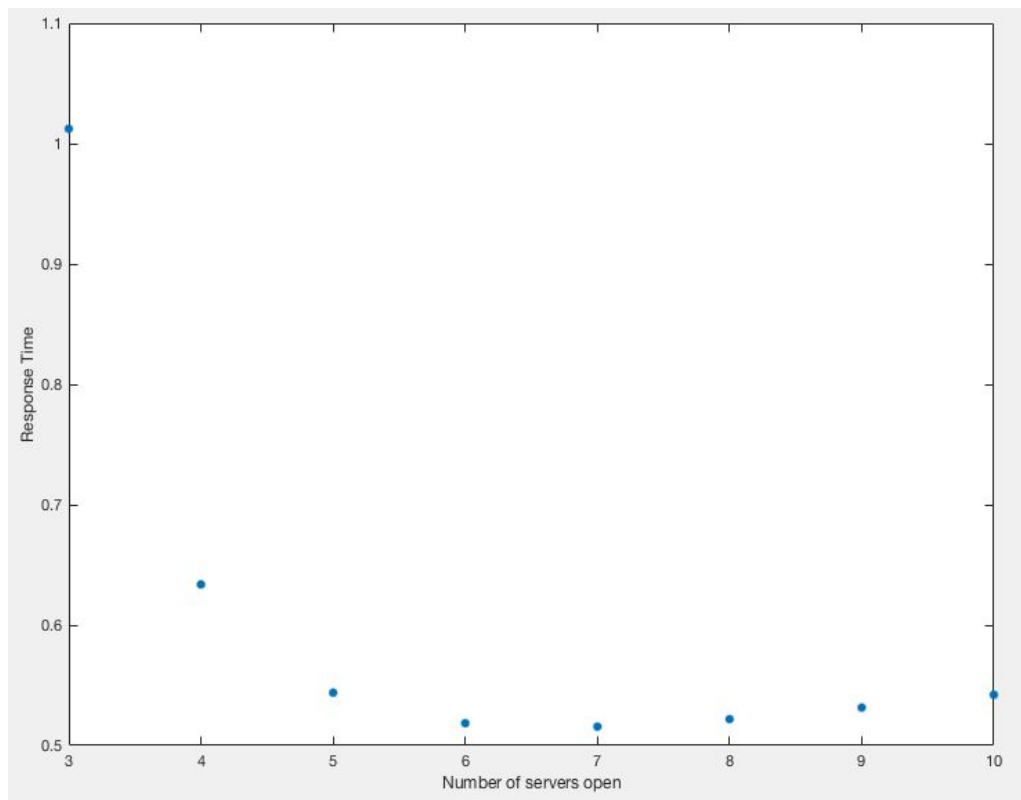


When  $w = 1000$ , the line is steady. Without the early part of the data is steady enough. Removing the transient part (exclude first 1000 jobs) and the left part which has around 29000 jobs (when  $s = 7$ ,  $T_{end} = 4000$ ) is steady to analysis.



## 5.2 Mean response time analysis

- Condition:  
 $s = 3$  to  $10$  ( $s$  is the number of servers opened, from 3 to 10)  
 Seed = one same seed `p1_rand_setting_arrival`  
 $w = 1000$  (removal jobs num)
- This graph below shows the mean response time of each system (opened servers from 3 to 10). Generally, it can be seen that when  $s = 6$ ,  $s = 7$  and  $s = 8$ , systems show a better performance than others ( $s = 3$ ,  $s = 4$ ,  $s = 5$ ,  $s = 9$ ,  $s = 10$ ).



### 5.3 Analysis independent replications to find the confidence interval

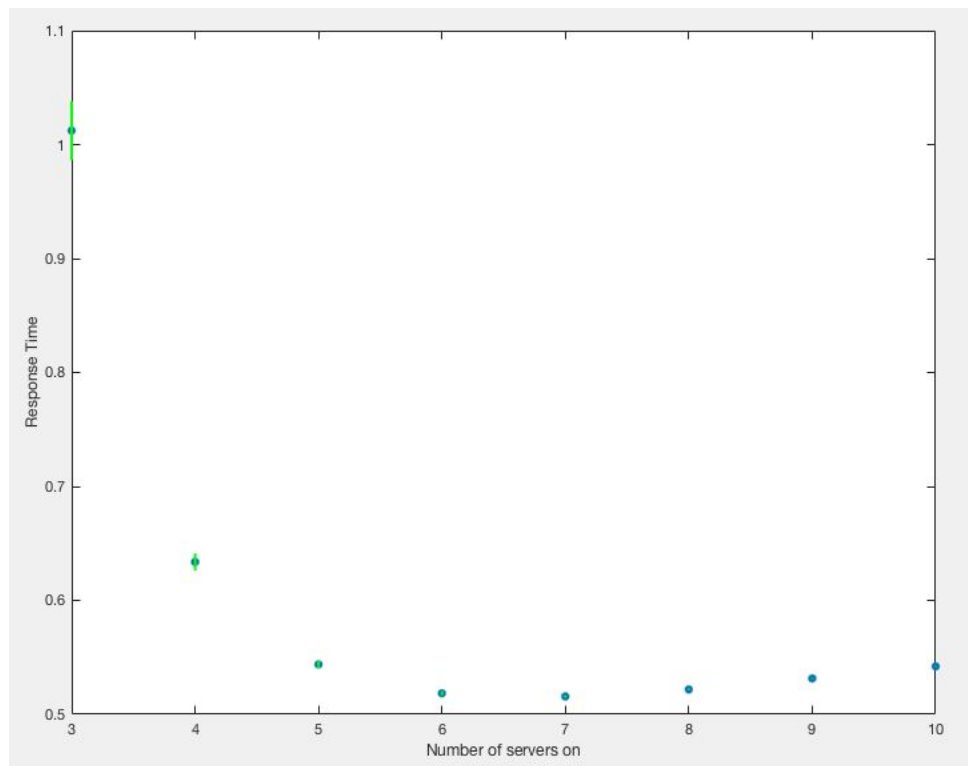
- Condition:  
 Independent replications times = 15 (start 5 times, then increase to 15)  
 Seed = 15 independent seeds  
 $s = 3$  to 10 ( $s$  is the number of servers opened, from 3 to 10)  
 $w = 1000$  (removal jobs num)  
 Seed = 15 independent seeds
- The table below shows the data of systems when different number of servers opened and different replication seed chosen. (The data of 15 sample replications is stored in [data1\\_replications\\_15.](#))



replication	s server opened							
	s=3	s=4	s=5	s=6	s=7	s=8	s=9	s=10
r=1	1.032	0.6387	0.5459	0.5191	0.5159	0.5214	0.531	0.5414
r=2	0.9496	0.6207	0.538	0.5146	0.5136	0.5204	0.5306	0.5418
r=3	0.9735	0.6265	0.5406	0.5156	0.5138	0.5204	0.5302	0.5411
r=4	1.0066	0.6276	0.5427	0.5178	0.5155	0.5219	0.5316	0.5424
r=5	1.0598	0.6485	0.5502	0.5225	0.5184	0.5239	0.5329	0.5433
r=6	1.0213	0.6356	0.5405	0.5175	0.5147	0.521	0.5306	0.5413
r=7	1.0072	0.6292	0.5438	0.5177	0.5151	0.5214	0.531	0.5415
r=8	0.9916	0.6244	0.5411	0.5177	0.5153	0.5217	0.5312	0.542
r=9	1.0302	0.6321	0.5431	0.5185	0.516	0.5221	0.5315	0.5421
r=10	1.0826	0.6563	0.5539	0.524	0.5187	0.5237	0.5326	0.543
r=11	0.9022	0.6046	0.533	0.5126	0.5125	0.52	0.5304	0.5413
r=12	1.0317	0.6399	0.547	0.5218	0.5182	0.5239	0.5331	0.5434
r=13	1.0071	0.6298	0.5407	0.517	0.5142	0.5209	0.5312	0.5421
r=14	1.0479	0.6419	0.546	0.5198	0.5163	0.5226	0.5323	0.5429
r=15	1.0407	0.6507	0.5508	0.5216	0.5173	0.5231	0.5318	0.5421
sample mean response time	1.0123	0.6338	0.5438	0.5185	0.5157	0.5219	0.5315	0.5421
lower 95% confidence interval	0.9864	0.6263	0.5408	0.5168	0.5146	0.5211	0.5309	0.5417
upper 95% confidence interval	1.0381	0.6412	0.5469	0.5203	0.5168	0.5226	0.532	0.5425
sample standard deviation	0.0451	0.013	0.0053	0.0031	0.0018	0.0013	0.0009	0.0008

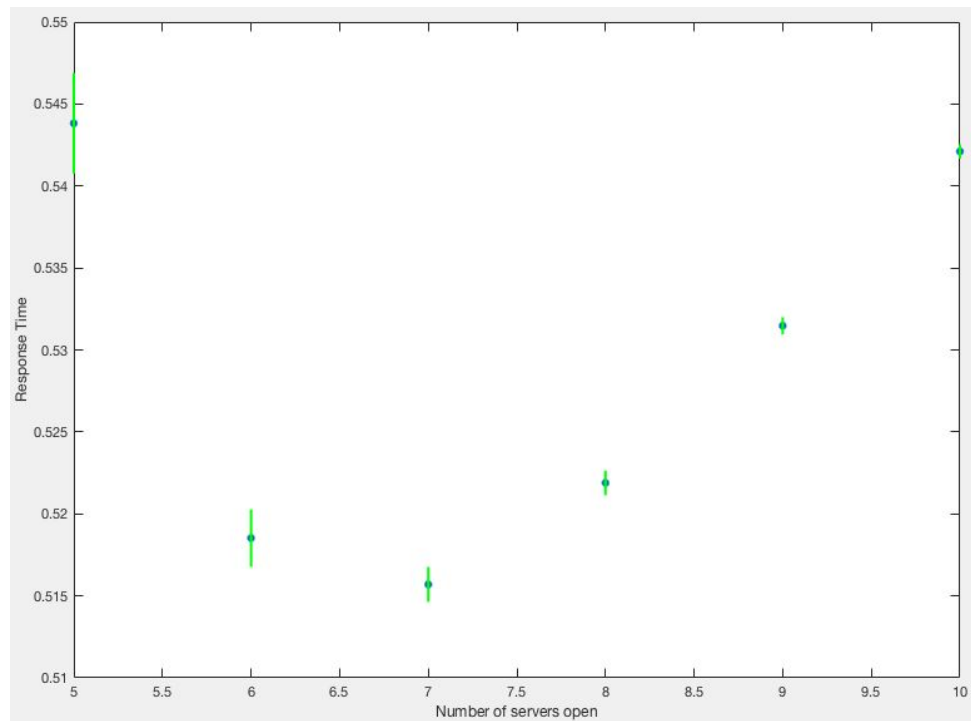
- Graphs analysis: ( The blue dot shows the mean response time. The green line shows the 95% confidence interval which is computed by 15 independent experiments).

1) The graph below illustrates that  $s = 3$  and  $s = 4$  has larger mean response time than  $s = 5$  to  $10$ . The 95% confidence intervals for  $s = 3$  and  $s = 4$  are not overlap. System of  $s = 4$  has the less response time than  $s = 3$ .

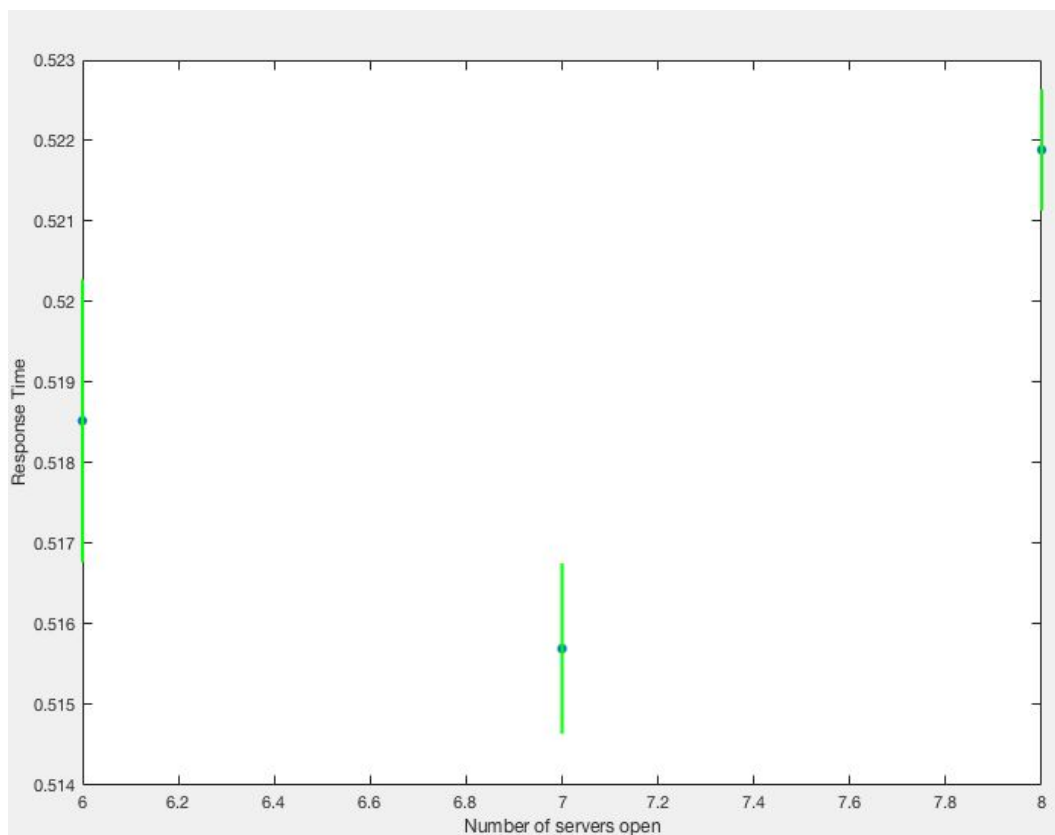




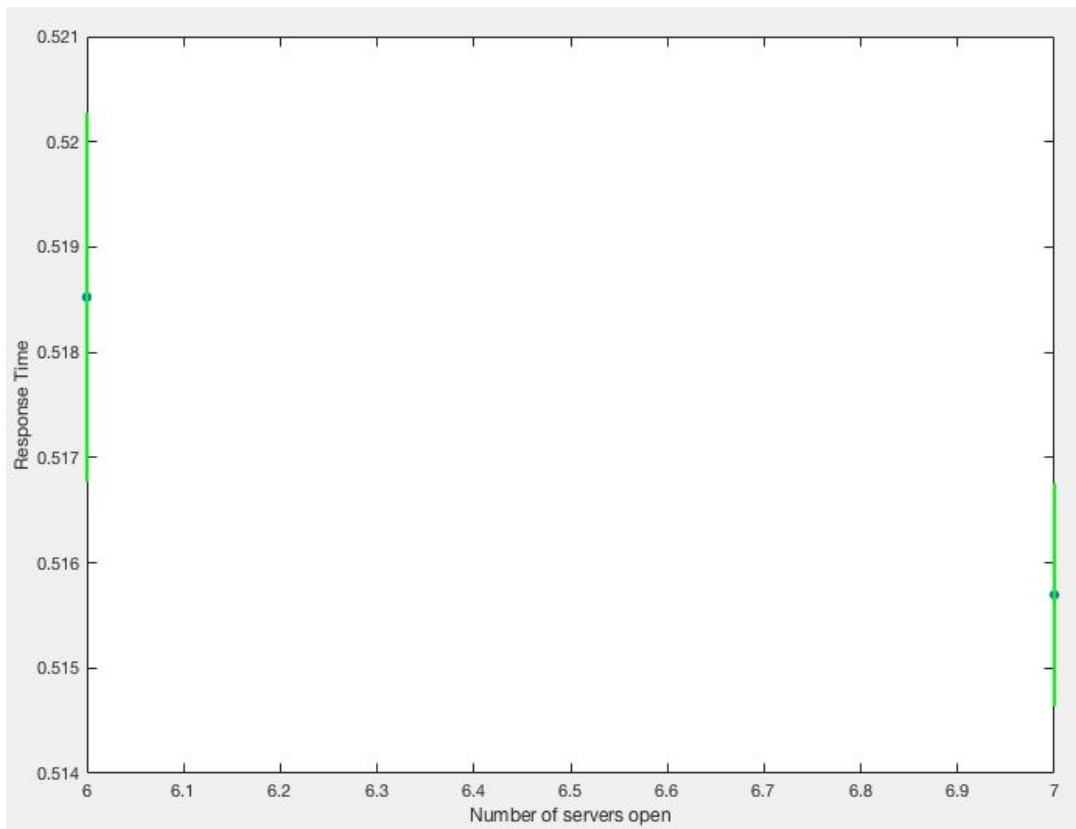
2) Next, the analysis focuses on systems  $s = 5$  to  $s = 10$ . System  $s = 5$ ,  $s = 9$  and  $s = 10$  have larger mean response time than systems  $s = 6$ ,  $s = 7$  and  $s = 8$ .



3) When we take a further look at system  $s = 6$ ,  $s = 7$  and  $s = 8$ , the graph shows that the performance if system  $s = 8$  is not as good as the other two ( $s = 6$  and  $s = 7$ ).



4) From the graph of system  $s = 6$  and  $s = 7$ , it is hard to tell which system has the better performance as it seems that the 95% confidence interval of each system is overlap.



5) After checking the lower bound of  $s = 6$  and the upper bound of  $s = 7$ , it can be confirmed that these two systems have overlapping areas but the mean of each one is not in the confidence interval of each other. This means that when taking 15 independent replications, system  $s = 6$  and  $s = 7$  have similar performance. The next step is to do a t-test.

replication	s server opened							
	s=3	s=4	s=5	s=6	s=7	s=8	s=9	s=10
r=1	1.032	0.6387	0.5459	0.5191	0.5159	0.5214	0.531	0.5414
r=2	0.9496	0.6207	0.538	0.5146	0.5136	0.5204	0.5306	0.5418
r=3	0.9735	0.6265	0.5406	0.5156	0.5138	0.5204	0.5302	0.5411
r=4	1.0066	0.6276	0.5427	0.5178	0.5155	0.5219	0.5316	0.5424
r=5	1.0598	0.6485	0.5502	0.5225	0.5184	0.5239	0.5329	0.5433
r=6	1.0213	0.6356	0.5405	0.5175	0.5147	0.521	0.5306	0.5413
r=7	1.0072	0.6292	0.5438	0.5177	0.5151	0.5214	0.531	0.5415
r=8	0.9916	0.6244	0.5411	0.5177	0.5153	0.5217	0.5312	0.542
r=9	1.0302	0.6321	0.5431	0.5185	0.516	0.5221	0.5315	0.5421
r=10	1.0826	0.6563	0.5539	0.524	0.5187	0.5237	0.5326	0.543
r=11	0.9022	0.6046	0.533	0.5126	0.5125	0.52	0.5304	0.5413
r=12	1.0317	0.6399	0.547	0.5218	0.5182	0.5239	0.5331	0.5434
r=13	1.0071	0.6298	0.5407	0.517	0.5142	0.5209	0.5312	0.5421
r=14	1.0479	0.6419	0.546	0.5198	0.5163	0.5226	0.5323	0.5429
r=15	1.0407	0.6507	0.5508	0.5216	0.5173	0.5231	0.5318	0.5421
sample mean response time	1.0123	0.6338	0.5438	0.5185	0.5157	0.5219	0.5315	0.5421
lower 95% confidence interval	0.9864	0.6263	0.5408	0.5168	0.5146	0.5211	0.5309	0.5417
upper 95% confidence interval	1.0381	0.6412	0.5469	0.5203	0.5168	0.5226	0.532	0.5425
sample standard deviation	0.0451	0.013	0.0053	0.0031	0.0018	0.0013	0.0009	0.0008

## 5.4 95% Confidence interval analysis

- Compute the 95% confidence interval for 15 replications in all cases (s = 3 to 10) by using different random numbers.
- The 95% confidence interval is computed by using the Sample Mean response time , Sample Standard Deviation which followed by the formula and the figure from Student's t-distribution table.

- Compute the sample mean

$$\hat{T} = \frac{\sum_{i=1}^n T(i)}{n}$$

- And the sample standard deviation

$$\hat{S} = \sqrt{\frac{\sum_{i=1}^n (\hat{T} - T(i))^2}{n - 1}}$$

**Note: for sample standard deviation, (n-1) is in the denominator, not n.**

- There is a probability (1-α) that the mean response time that you want to estimate lies in the interval

$$\left[ \hat{T} - t_{n-1, 1-\frac{\alpha}{2}} \frac{\hat{S}}{\sqrt{n}}, \hat{T} + t_{n-1, 1-\frac{\alpha}{2}} \frac{\hat{S}}{\sqrt{n}} \right]$$

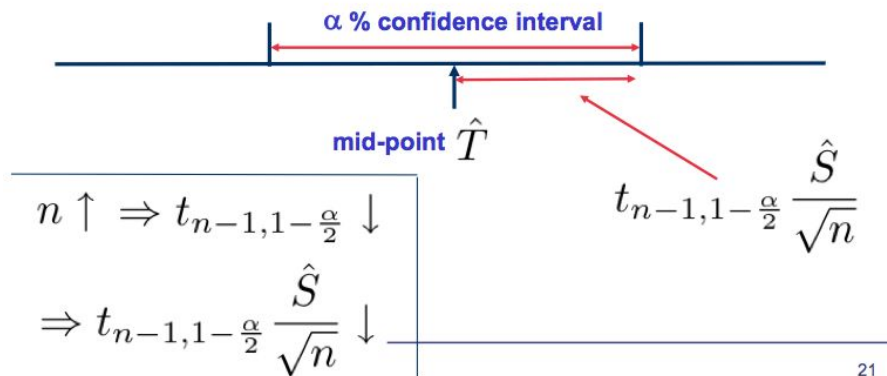
**t Table**

cum. prob one-tail two-tails	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
df	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

- The reason of the 95% confidence interval of system  $s = 6$  and  $s = 7$  are overlap is the interval is not narrow enough.

- Confidence interval

$$\left[ \hat{T} - t_{n-1, 1-\frac{\alpha}{2}} \frac{\hat{S}}{\sqrt{n}}, \hat{T} + t_{n-1, 1-\frac{\alpha}{2}} \frac{\hat{S}}{\sqrt{n}} \right]$$



21

- There are two ways to get a narrow interval.

- 1) Have more replications ( bigger  $n$ ).
- 2) Increase the simulation time (bigger  $t$  )



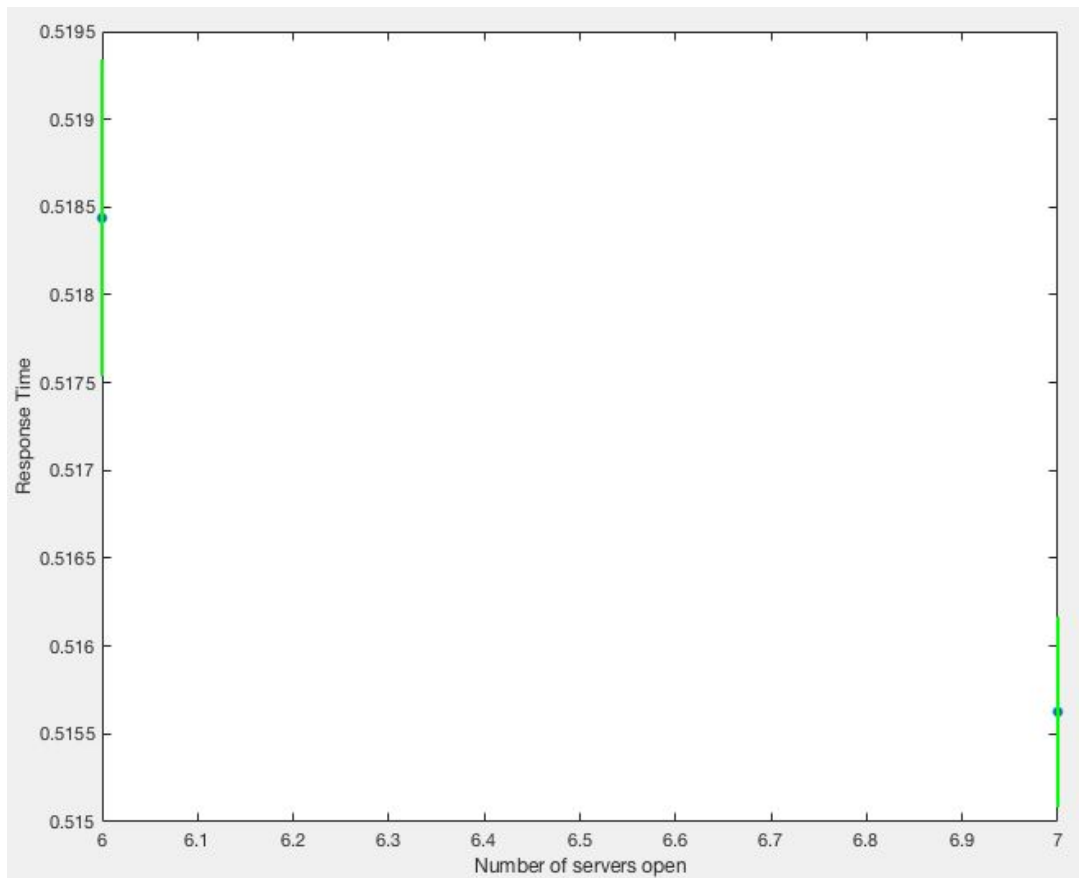
## 5.5 Do t-test and compare two systems by increasing the replication times

- Condition:  
Independent replications times = 30 (start 5 times, then increase to 15 then to 30)  
 $s = 6$  or  $7$  ( $s$  is the number of servers opened)  
 $w = 1000$  (removal jobs num)  
Seed = 30 independent replication seed (Common random numbers method: in order to reduce the variance between two systems)
- The table below shows the data of system  $s = 6$  and  $s = 7$  with the common seed. (The data of 30 sample replications for  $s = 6$  and  $s = 7$  is stored in [data2\\_replications\\_30\\_common](#).) The numbers of last column are negative. This means system  $s = 7$  has the less response time than  $s = 6$ .

replication $r$	EMRT $s = 6$	EMRT $s = 7$	(EMRT $s = 7$ ) - (EMRT $s = 6$ )
$r=1$	0.5191	0.5159	-0.0032
$r=2$	0.5146	0.5136	-0.001
$r=3$	0.5156	0.5138	-0.0018
$r=4$	0.5178	0.5155	-0.0023
$r=5$	0.5225	0.5184	-0.0041
$r=6$	0.5175	0.5147	-0.0028
$r=7$	0.5177	0.5151	-0.0026
$r=8$	0.5177	0.5153	-0.0024
$r=9$	0.5185	0.516	-0.0025
$r=10$	0.524	0.5187	-0.0053
$r=11$	0.5126	0.5125	-0.0001
$r=12$	0.5218	0.5182	-0.0036
$r=13$	0.517	0.5142	-0.0028
$r=14$	0.5198	0.5163	-0.0035
$r=15$	0.5172	0.5154	-0.0018
$r=16$	0.5203	0.5174	-0.0029
$r=17$	0.5199	0.5167	-0.0032
$r=18$	0.5183	0.5157	-0.0026
$r=19$	0.5155	0.5139	-0.0016
$r=20$	0.5175	0.5152	-0.0023
$r=21$	0.5186	0.5158	-0.0028
$r=22$	0.5181	0.5154	-0.0027
$r=23$	0.5171	0.5148	-0.0023
$r=24$	0.5168	0.5145	-0.0023
$r=25$	0.5187	0.5154	-0.0033
$r=26$	0.52	0.5161	-0.0039
$r=27$	0.5212	0.5167	-0.0045
$r=28$	0.5201	0.5163	-0.0038
$r=29$	0.5169	0.5145	-0.0024
$r=30$	0.5206	0.5168	-0.0038
sample mean response time	0.5184	0.5156	-0.0028
lower 95% confidence interval	0.5175	0.5151	-0.0024
upper 95% confidence interval	0.5193	0.5162	-0.0031
sample standard deviation	0.0024	0.0014	

- From the table above, it can be concluded that the estimated mean response time (EMRT) of system  $s = 7$  is always better than the system  $s = 6$  's in 30 common seed replications.

- The graph below illustrates that the confidence intervals for system  $s = 6$  and  $s = 7$  are not overlap anymore.



- The table below shows the data of the 95% confidence intervals when comparing EMRT  $s=6$  and EMRT  $s=7$  by choosing different replication numbers using Common random number. It can be seen that the intervals between the lower and upper are decreased with the grown of independent replication times. The minimal interval displays when  $r = 30$  (30 times replications).

# independent replications	95% Confidence interval of (EMRT $s=6$ ) - (EMRT $s=7$ ) using CNR	
	lower	upper
15	-0.0034	-0.0019
20	-0.0032	-0.0021
25	-0.0031	-0.0022
30	-0.0032	-0.0024

## 5. Conclusion

The server farm which consists of a high-speed router and a number of computing servers has a 2000 watt power budget. When open 7 servers ( $s = 7$ ) there is 95% confident that this server farm can maintain a good response time by given conditions in the specification of the project.