

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

A VIRTUALIZATION APPROACH TO AUTO-SCALING PROBLEM



E071849

**DOUANGHATAI THIPPAYAWONG
THITHAT THEPPARAT
AMNART HARNPRASARNKIT**

เลขหมู่.....
เลขทะเบียน.....**71849**
วันเดือนปี.....**27 ส.ย. 2554**

b.....
i.....

**A SPECIAL PROJECT SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENT FOR DEGREE OF BACHELOR OF SCIENCE
IN COMPUTER SCIENCE
FACULTY OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
ACADEMIC YEAR 2009**

Title A virtualization approach to auto-scaling problem

Students Miss. Douanghatai Thipayawong
Mr. Thithat Thepparat
Mr. Amnart Harnprasarnkit

Degree Bachelor's Degree of Science

Program Computer Science

Academic Year 2009

Advisor Associate Professor Dr. Veera Boonjing

Abstract

The purpose of this project is to simulate feasibility of using virtualization technology to increase MTTF and CPU utilization. In our research, we use ARENA which is simulation software to build two different models. There are auto-scaling without server virtualization and auto-scaling with server virtualization. The result of our experiment shows that server virtualization can be increased the life time of server and also reduce the waiting time in queue.

Keywords: Mean Time To Failure (MTTF), Server Virtualization, Simulation.

Acknowledgement

The author would like to take this opportunity to express honest thanks to our teachers and people who gave constructive advice and full support in this research.

This project would not have been possible without the support, include suggestion from Assoc.Prof.Dr.Veera Boonjing, Our adviser for valuable guidance, previous assistance and support throughout this research.

We are greatly appreciate the entire professors who have invaluable knowledge while studying in the department of computer science, faculty of science, King Mongkut's Institute of Technology Ladkrabang. In the addition, thanks also to Mr.Sivadon Chaisiri for good and advise in cloud computing information

Especially, I would like to give my special thanks to my parent whose patient love enabled me to complete this work.

Dounghathai	Thippayawong
Thithat	Thepparat
Amnart	Harnprasarnkit

Table of contents

	Page
Abstract.....	I
Acknowledgement	II
Table of contents.....	III
List of tables.....	V
List of figures.....	VI
Chapter 1 Introduction	
1.1 Rationale	1
1.2 Objectives	1
1.3 Scope of study.....	2
1.4 Organization.....	2
Chapter 2 Related literature	
2.1 Cloud computing.....	3
2.2 Auto-scaling.....	4
2.3 Virtualization technology.....	5
2.4 Server consolidation.....	7
Chapter 3 A Simulation of server virtualization	
3.1 Problem description.....	9
3.2 Arrival process.....	10
3.3 Choose service.....	12
3.4 Assign to the right server.....	12
3.5 Create necessary resources.....	14
3.6 Decide the server.....	15
3.7 The process.....	17

3.8 The end of process.....	21
3.9 Model running set up.....	22
Chapter 4 Experimental result	
4.1 Experimental design.....	27
4.2 Experimental result.....	30
4.2.1 Mean time to failure.....	30
4.2.1.1 Mean time to failure of auto-scaling without server virtualization model.....	31
4.2.1.2 Mean time to failure of auto-scaling with server virtualization model.....	37
4.2.2 Key performance indicators	
4.2.2.1 Average workload waits time.....	39
4.2.2.2 Average wait time queue.....	39
4.2.2.3 Average number of waiting time.....	40
4.2.2.4 Average resource usage.....	41
4.3 Comparison.....	43
Chapter 5 Conclusion and recommendation	
5.1 Conclusion.....	44
5.2 Recommendation.....	44
Bibliography.....	45
Authors biography.....	47
Appendix.....	48

List of tables

	Page
Table 3-1: Workloads arrival schedule.....	10
Table 3-2: Detail of assignment for modules.....	13
Table 4-1: Failure rate of servers in auto –scaling without server virtualization model.....	29
Table 4-2: Failure rate of servers in auto –scaling with server virtualization model.....	29
Table 4-3: Running setup.....	30
Table 4-4: Show the number of failure of auto-scaling without server virtualization model.....	31
Table 4-5: Shown how to calculate the time before failure.....	33
Table 4-6: Show the number of failure and time before failure of auto-scaling with server virtualization model.....	37
Table 4-7: Show the increasing resource usage of each server type.....	42
Table 4-8: Comparison between auto-scaling without virtualization model and auto-scaling with virtualization model	43

List of figures

	Page
Figure2.1: Cloud computing characteristics.....	3
Figure2.2: Description of an auto-scaling feature.....	4
Figure3-1: Create schedule for workloads.....	10
Figure3-2: Schedule dialog.....	11
Figure3-3: Create module and create dialog.....	11
Figure3-4: Decide module and decide dialog.....	12
Figure3-5: Assign module and assign dialog	13
Figure3-6: Create resources for first model	14
Figure3-7: Create resources for second model	14
Figure3-8: Decide dialog with correct and complete values.....	16
Figure3-9: Hold module dialog	17
Figure3-10: Process module dialog of first model	17
Figure3-11: Process module dialog of second model	18
Figure3-12: Decide module for virtual server	19
Figure3-13: Pick station dialog	20
Figure3-14: Process module dialog of second model	20
Figure3-15: Station dialog	21
Figure3-16: Dispose dialog	21
Figure3-17: Project parameters setup	22
Figure3-18: Replication parameters setup.....	23
Figure3-19: Arena toolbar.....	24
Figure3-20: First model: Auto-scaling without virtualization.....	25

Figure3-21: Second model: Auto-scaling with virtualization.....26

Figure4-1: Auto - scaling without server virtualization model.....27

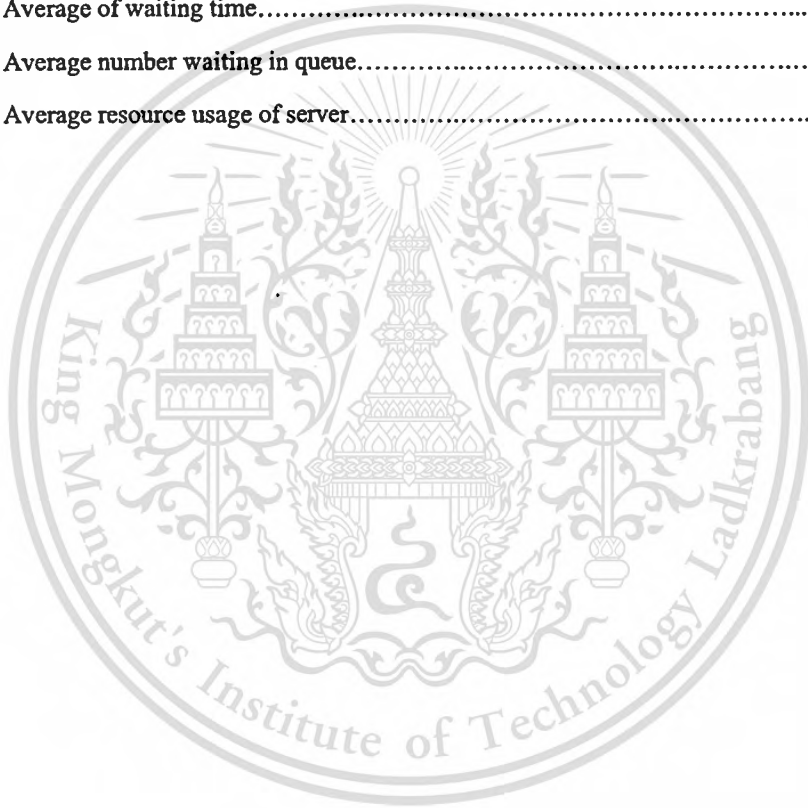
Figure4-2: Auto - scaling with server virtualization model.....28

Figure4-3: Average waiting time per entity.....39

Figure4-4: Average of waiting time.....40

Figure4-5: Average number waiting in queue.....41

Figure4-6: Average resource usage of server.....42



Chapter 1

Introduction

1.1 Rationale

Cloud computing is a style of computing where massively scalable IT-related capabilities are provided “as a service” across the internet to multiple external customers. However, there is one significant point that a cloud provider especially private cloud provider should realize is how the servers can support the heavy coming workloads at the peak time. Some public cloud providers provide the auto-scaling feature in order to support the heavy workloads. They use physical servers to scale up the system when workloads increase. It takes cost too much. However, it cannot guarantee that the servers will fail or not.

We offer using virtualization technology in server in order to reduce the failure of the server. Server in term of virtualization technology means creating virtual server in order to support heavy workloads at the peak time.

In our experiment, we study feasibility of using virtualization technology especially the server virtualization in order to reduce the server failure. We use simulation to do the experiment. We choose simulation technique because we can estimate the performance of server with using virtualization technology before real implementation. We build two different models. First model is auto-scaling without server virtualization and second model is auto-scaling with server virtualization model.

1.2 Objectives

The objective of this project is to simulate feasibility of virtualization technology to auto-scaling problem. The purpose of simulation is to prove that using virtualization technology in servers can increase the mean time to failure and also increase the CPU utilization.

1.3 Scope of study

In our research, we measure mean time to failure (MTTF) and CPU utilization by simulate two different models. There are Auto-scaling without virtualization technology and Auto-scaling with virtualization technology. We define ten rounds of replications. Each replication is 1 year and the system performs 24 hours per day. To support the heavy load, we have to use horizontal scale up technique. If we do not have enough servers to support, we have to increase server's capacity or add more servers. Our research concentrate on the private cloud computing and its environment only.

1.4 Organization

The rest of this special project consists of 4 chapters as follows:

Chapter 2 provides related literature of visualization approach to auto-scaling problem, cloud computing, auto-scaling, virtualization technology, server consolidation and mean time to failure.

Chapter 3 introduces the reader more insights into a virtualization approaching to auto-scaling problem.

Chapter 4 gives details of the experiment result.

Chapter 5 gives conclusion and recommendation.

Chapter 2

Related literature

2.1 Cloud computing

Cloud computing is a general term for anything that involves delivering hosted services over the internet. These services are broadly divided into three categories: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS). The name “Cloud Computing” was inspired by the cloud symbol that is often used to represent the internet in flowcharts and diagrams.

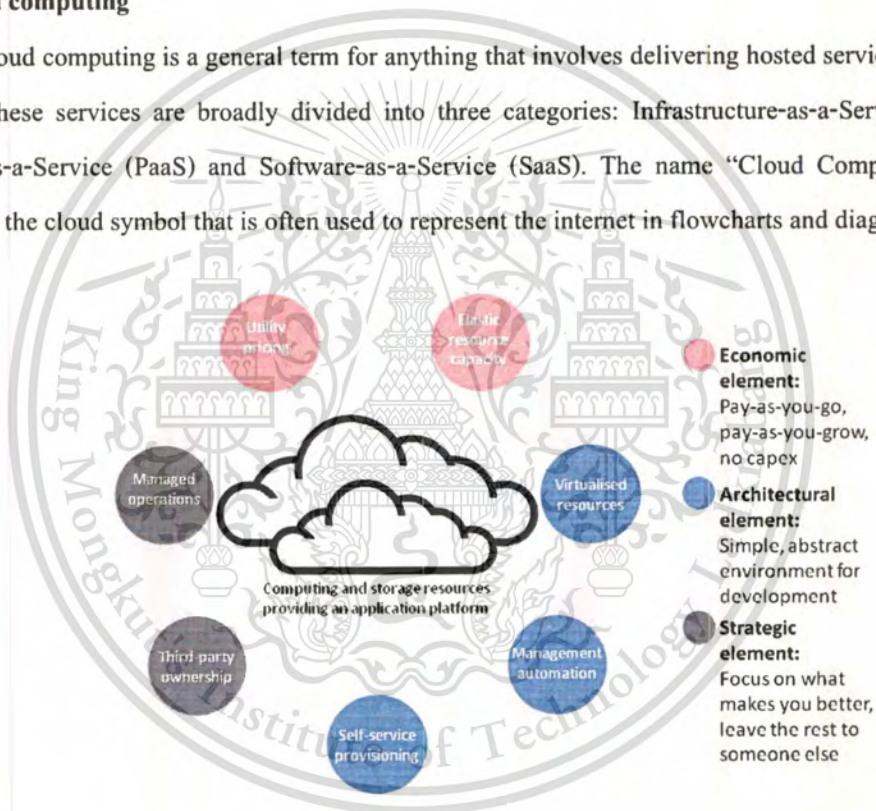


Figure 2.1: Cloud computing characteristics

A cloud service has three distinct characteristics that differentiate it from traditional hosting. It is sold on demand, typically by the minute or the hour; it is elastic – a user can have as much or as little of a service as they want at any given time; and the service is fully managed by the provider (the consumer needs nothing but a personal computer and Internet access). Significant innovations in

virtualization and distributed computing, as well as improved access to high-speed Internet and a weak economy, have accelerated interest in cloud computing.

There are three kinds of cloud computing. A public cloud sells services to anyone on the internet. A private cloud is a proprietary network or a data center that supplies hosted services to a limited number of people. When a service provider uses public cloud resources to create their private cloud, the result is called a virtual private cloud. Private or public, the goal of cloud computing is provide easy and scalable access to computing resources and IT services.

2.2 Auto-scaling

The objective of this paper is to simulate feasibility of virtualization technology to auto scaling. First you should understand what auto-scaling is. In this paper we present auto-scaling feature of Amazon EC2 as example.

What is auto-scaling?

Auto-scaling is features that scale-up the system when workload increases and scale-down the system when workload decreases. Amazon EC2 is a cloud provider that provides this feature by allows users to create triggers that automatically start and terminate instances based on pre-determined rules. Figure2.2 is describes auto-scaling feature.

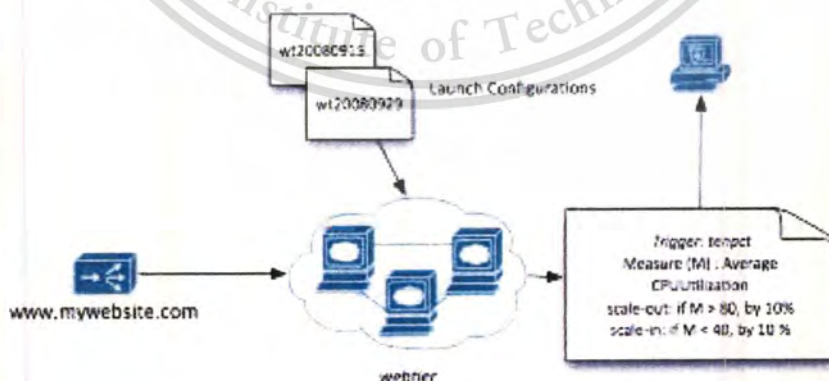


Figure 2.2: Description of an auto-scaling feature

Auto - scaling web services on Amazon EC2

This paper presents auto-scaling web services on Amazon EC2. Their approach to scale up web service is dynamic load balancing where the load balancer can itself made part of the group it is distributing the load across. So, membership changes in these application nodes can be detected by the load balancer. Traditionally, the list of nodes across which workload is distributed had to be provided to the load balancer. Such a load balancer is known as a static load balancer. Hence, new members could not join the application group at runtime. However, such a setup will not work economically on EC2 since the only way of achieving this is to reserve a set of elastic IP addresses, assign them to the application instances and provide those IP addresses to the load balancer. This defeat the idea of auto-scaling since new members cannot join the cluster and IP addresses of members cannot change to anything other than those assigned from the set of elastic IP addresses. Their approach enable membership discovery by the load balancer, hence making it a dynamic load balancer [5].

The paper above, they do not study the mean time to failure of servers on Amazon EC2. However, how we make sure that the servers can support the heavy loads at peak time. It does not take time to wait for recovering.

2.3 Virtualization technology

Our approach is implementing virtualization technology on servers in order to scale up the resource at the peak time. In term of virtualization refers to technologies designed to provide a layer of abstraction between computer hardware systems and the software running on them. By providing a logical view of computing resources, rather than a physical view, virtualization solutions make it possible to do a couple of very useful things Virtualization can allow you to trick your operating systems into thinking that a group of servers is a single pool of computing resources and they can allow you to run multiple operating systems simultaneously on a single machine.

Virtualization has its roots in partitioning, which divides a single physical server into multiple logical servers. Once the physical server is divided, each logical server can run an operating system and

applications independently. In the 1990s, virtualization was used primarily to re-create end-user environments on a single piece of mainframe hardware. If you were an IT administrator and you wanted to roll out new software, but you want to see how it would work on a Windows NT or a Linux machine, you should use virtualization technologies to create the various user environments,[9].

Advantages

The benefits of virtualization technology have been widely documented.

1. Cost - A number of companies have found they have had 'server sprawl' in recent years. As the price of hardware has fallen rapidly, new servers have tended to be assigned to single applications, which led to a profusion of servers. In fact it is estimated that the total cost of ownership of x86 servers has exceeded the purchase costs (Gillen, Humphreys & Waldman, 2006; Goth, 2007). These studies estimate that the TCO for x86 servers is between \$10,000 and \$15,000 per year, yet the most servers today are running at under 10 percent utilization. This indicates that the cost of managing physical servers will exceed their purchase costs. Virtualizing servers can aid in reducing these costs because it offers the possibility to consolidate a number of servers onto one physical host and also achieve higher utilization rates. VMware claim that their customers have seen consolidation ratios of 10-15:1 for production servers which allowed customers to significantly reduce their total cost of ownership [6].

2. Consolidating servers can also reduce energy consumption and lower overall spend as energy costs are continuing to rise. A study by Koomey (2007) shows that "aggregate electricity use for servers doubled over the period 2000 to 2005 both in the U.S. and worldwide". He also found that this was primarily due to an increase in the volume of servers rather than an increase in the power usage per unit. IDC expects server power and cooling costs to increase by an average of 11.2% per year to reach \$44.5 billion by 2010; when it will equal 70% of the overall new server spending for that year [9].

3. Virtualization offers effective and efficient system recovery. Virtual machine virtual disks are literally files that can be backed up and restored quickly for disaster recovery purposes. The virtual machine is stored on disk as a file that can be moved to any other disk area therefore system recovery is much easier than for traditional physical servers (Greiner, 2007)[6].

Drawbacks

1. While virtualization allows more servers to run on physical hardware, there is an initial investment required for virtualization, which should not be overlooked. These costs include:

- Licensing costs for enterprise features.
- Installation costs.
- Training costs.
- Storage costs.

The trend for virtualization hypervisors is to provide them with the hardware or OS. However there is a need for tools to manage the virtual servers and provide the enterprise functionality discussed, such as clustering and high availability. If you want to implement virtualization on a large scale then management tools will be required. It may also require new backup solution and server management tools [7].

2. There is an overhead associated with running the virtual machines, particularly for busy transaction based server. Therefore it may not be suitable for all servers. There is research to suggest that the new generation of x86 processors do not actually improve performance because they have no support for MMU virtualization and they do not co-exist with existing software techniques for MMU virtualization [9].

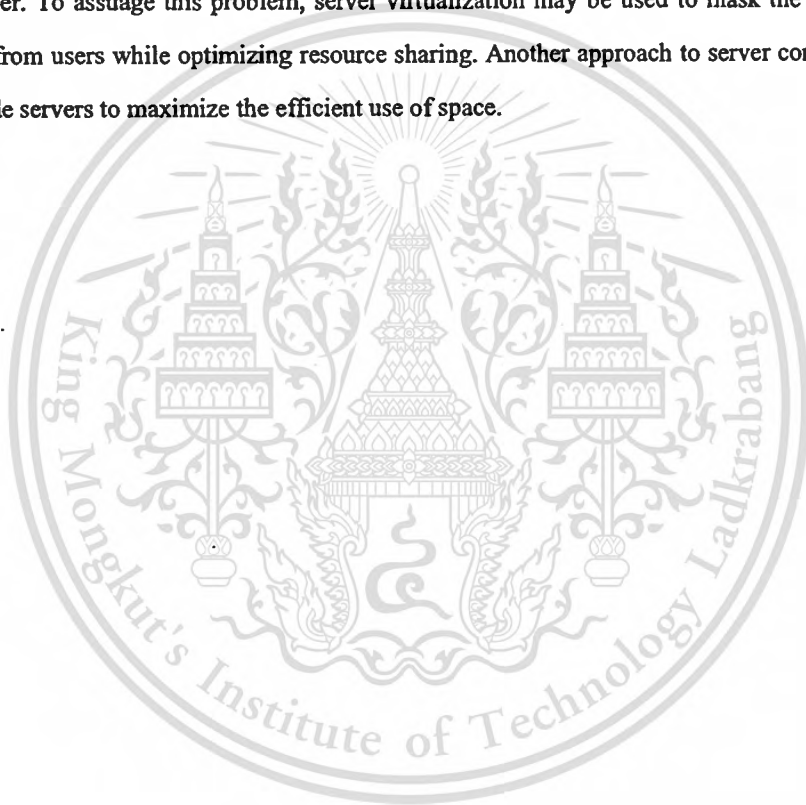
3. Management - The technology is relatively immature on x86 servers, and there is a learning curve with the new technology. It takes experience to learn how to best allocate resources. The specification and configuration of servers has changed significantly, it is necessary to consider physical hosts and virtual machines. Virtualization adds management complexity; innovation and standardization in management of virtual entities are required to realize full potential of virtualization [9].

2.4 Server consolidation

Server Consolidation is an approach to the efficient use of computer server resources in order to reduce the total number of servers or server locations that an organization requires. The practice

developed in response to the problem of server sprawl, a situation in which multiple, under-utilized servers take up more space and consume more resources than can be justified by their workload.

Although consolidation can substantially increase the efficient use of server resources, it may also result in complex configurations of data, applications, and servers that can be confusing for the average user. To assuage this problem, server virtualization may be used to mask the details of server resources from users while optimizing resource sharing. Another approach to server consolidation is the use of blade servers to maximize the efficient use of space.



Chapter 3

A simulation of server virtualization

This chapter gives the reader more insight into a simulation process using ARENA to evaluate the utilization of simple server model and virtualization server model in the same environment such as failure or the entity arrival.

3.1 Problem description

Our experiment has four kinds of servers. All servers use auto-scaling. Auto-scaling will start working when workloads in the system increase and need more capacity in order to avoid the failure of the system at a peak time. Each peak time of each service is not the same. Auto-scaling need the compatible server that have same properties such as OS. This mean every services must have secondary server that ready to scale up. So, we use virtual server instead of the real physical server. Virtual server will generate from the necessary server but less utilization at that time. A server able to has more than ten virtual servers. Each virtual server can operate different OS and provide different service. Auto-scaling combine with virtual server will use less number of servers. The problem is auto-scaling with virtual server can ensure the better number of failure and still work well than usual or not.

3.2 Arrival Process

The workload will arrive and enter to the system by the schedule below.

Table 3-1: Workloads arrival schedule

Time	Number of workloads arrival	Time	Number of workloads arrival
00:00 – 01:00	48	12:00 – 13:00	338
01:00 – 02:00	56	13:00 – 14:00	431
02:00 – 03:00	74	14:00 – 15:00	426
03:00 – 04:00	86	15:00 – 16:00	444
04:00 – 05:00	99	16:00 – 17:00	383
05:00 – 06:00	129	17:00 – 18:00	332
06:00 – 07:00	160	18:00 – 19:00	389
07:00 – 08:00	275	19:00 – 20:00	371
08:00 – 09:00	292	20:00 – 21:00	259
09:00 – 10:00	363	21:00 – 22:00	165
10:00 – 11:00	424	22:00 – 23:00	136
11:00 – 12:00	405	23:00 – 00:00	93

Table 3-1 is the workloads arrival per hour. There are many workloads enter to the system all the time. From this schedule, peak time of the system begin from 10:00 am – 12:00 pm and start again at 13:00 pm – 16:00 pm. After that incoming workloads and total number of workloads in system will decrease.

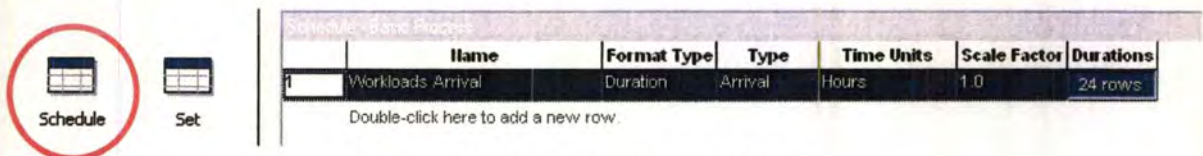


Figure 3-1: Create schedule for workloads

Select schedule spreadsheet icon and double click the message “Double-click here to add a new row” to create the new schedule for this system. After that, right click in the new row and select “Edit via dialog” to add the arrival data from Table 3-1.

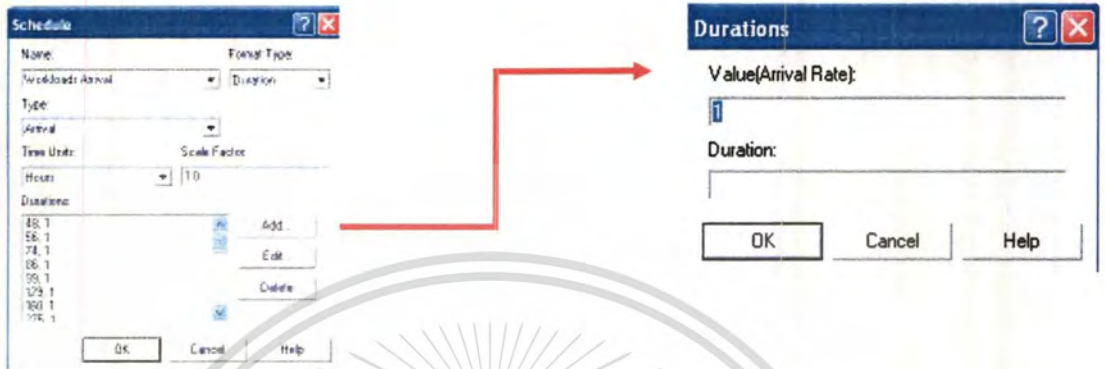


Figure 3-2: Schedule dialog

Figure 3-2 is the schedule dialog. It used to enter the data from the Table 3-1. To add the data you have to click Add button. Then the new dialog names “Durations” will show and require you to enter the number of arrival and the duration.

We use create module in Arena and named it “Workload Arrival”. This module used to create the workload for the system. In time between arrivals part, select “Schedule” as the type of arrival and choose the name of the schedule. This step will let this create module create the workload refer to the schedule. The capture of create workload step will be show in the Figure 3-3.

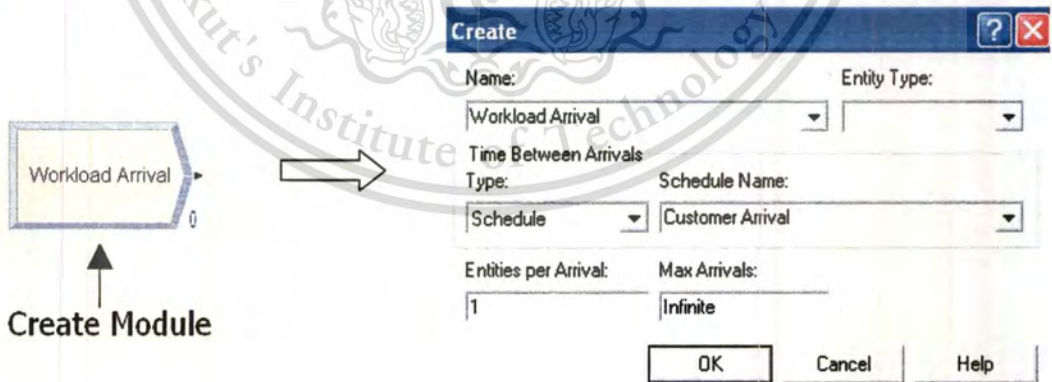


Figure 3-3: Create module and create dialog

3.3 Choose service

When workload was created, they have to choose one service to do. So, we use decide module for each workload. Decide module used to create the condition or the chance for the entered workload to choose one way to go to the next process. In our experiment, each workload has to choose one out four services by chance. We assume the chance for the less utilization server is 12% for minimum and 20% for maximum. We assume the chance for the normal utilization server is 33% and hard utilization server is 35%.

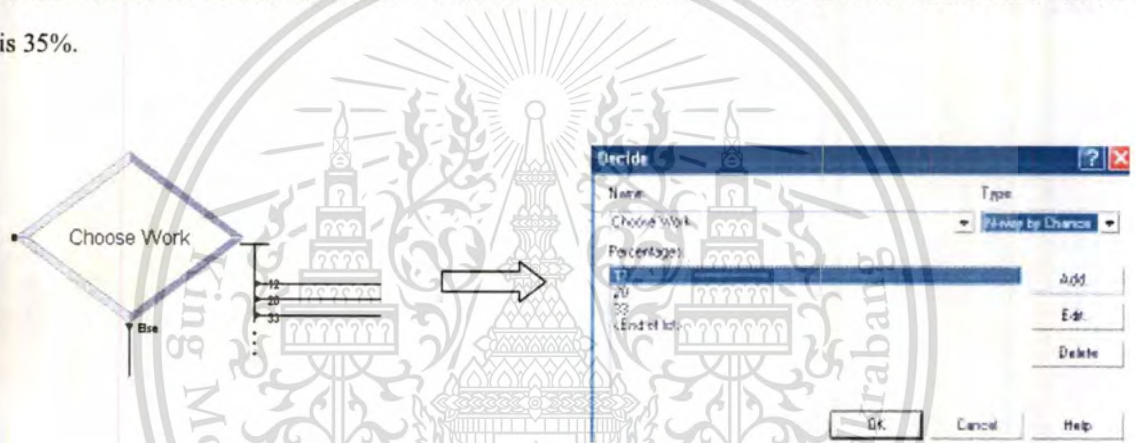


Figure 3-4: Decide module and decide dialog

We use decide module and named it as “Choose Work” and select “N-way by chance” because we have more than 2 choice of selection for workload. After that we add percentages to this module and then click OK button. We enter 12%, 20%, 33% only because the last value of chance will be the remaining of total number of entered number minus 100%. The remaining percentage’s way will show and named as “else”. This step will let the workload decide the work and go to the next step.

3.4 Assign to the right server

This step will begin after incoming workload has decided the work. Each workload will be assign the necessary attribute such as image, type of server. Our experiment decide to have four kinds of servers so, we have to create four assign modules in order to add the four assignment. We named four

modules as “File Serv”, “Mail Serv”, “DB Serv” and “Web Serv”. The assignment that we assign is entity picture and entity type. Table 3-2 below is the detail of assignment for each module.

Table 3-2: Detail of assignment for modules

Module's name	Entity picture	Entity type
File Serv	Blue page	Work1
Mail Serv	Green page	Work2
DB Serv	Red page	Work3
Web Serv	Yellow page	Work4

We enter the data in Table 3-2 by double click on the module body then the assign dialog will appear as show in Figure 3-5. Then, click add button and select entity picture. Click add button again to select entity type then click OK to finish this step.

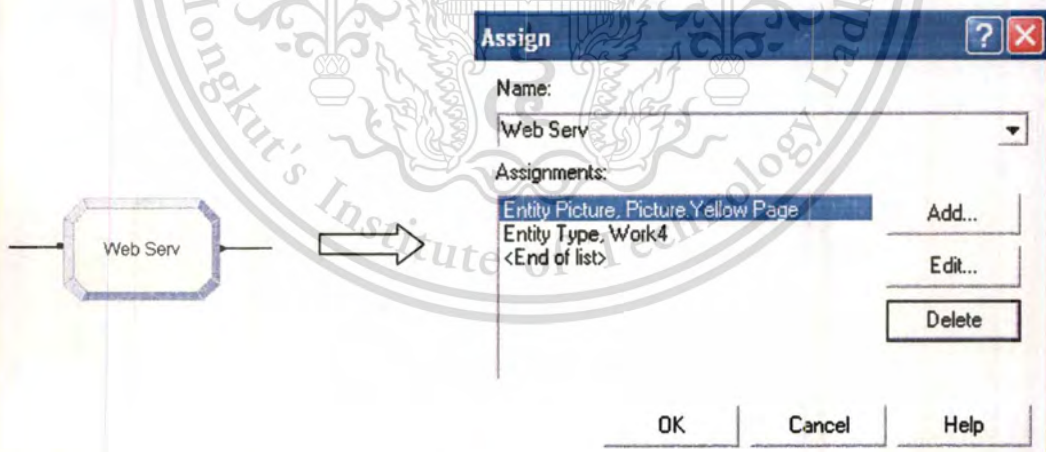


Figure 3-5: Assign module and assign dialog

Figure 3-5 is just one sample from four assign modules. The full version of assign layer will be show in the last page of this chapter.

3.5 Create necessary resources

Every process needs some resource to operate. Both first and second model need the different number of resources. The resources of first model compose of four primary servers and four secondary servers. Both primary and secondary servers have ten of capacity as show in Figure 3-6.

	Name	Type	Capacity	Busy / Hour	Idle / Hour	Per Use	StateSet Name	Failures	Report Statistics
1	Main Serv1	Fixed Capacity	10	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
2	Main Serv2	Fixed Capacity	10	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
3	Main Serv3	Fixed Capacity	10	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
4	Main Serv4	Fixed Capacity	10	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
5	Secondary Serv1	Fixed Capacity	10	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
6	Secondary Serv2	Fixed Capacity	10	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
7	Secondary Serv3	Fixed Capacity	10	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
8	Secondary Serv4	Fixed Capacity	10	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>

Figure 3-6: Create resources for first model

The resources of second model compose of four main servers which has ten capacity each and forty Virtual servers which has one capacity each as show in Figure 3-7

	Name	Type	Capacity	Busy / Hour	Idle / Hour	Per Use	StateSet Name	Failures	Report Statistics
1	Main Serv1	Fixed Capacity	10	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
2	Main Serv2	Fixed Capacity	10	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
3	Main Serv3	Fixed Capacity	10	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
4	Main Serv4	Fixed Capacity	10	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
5	VM1_1	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
6	VM1_2	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
7	VM1_3	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
8	VM1_5	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
9	VM2_1	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
10	VM2_2	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
11	VM2_3	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
12	VM2_4	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
13	VM3_1	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
14	VM3_2	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
15	VM3_3	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
16	VM3_4	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
17	VM4_1	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
18	VM4_2	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
19	VM4_3	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
20	VM4_8	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
21	VM1_4	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
22	VM1_6	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
23	VM1_7	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
24	VM1_9	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
25	VM1_8	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
26	VM1_10	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
27	VM2_5	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
28	VM2_6	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
29	VM2_7	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>
30	VM2_8	Fixed Capacity	1	0.0	0.0	0.0		1 rows	<input checked="" type="checkbox"/>

Figure 3-7: Create resources for second model

3.6 Decide the server

This step used to check the available capacity and status of the server. We use decide module again and build the expression in this module. The expression that we use to check the server status and available capacity is:

$$\text{NQ}(\text{Server } n.\text{Queue}) < 15 \ \&\& \ \text{STATE}(\text{Main Serv } n) \diamond -4$$

Where: n	Number of server that we want to check.
NQ	Current number in queue of server.
Server n .Queue	Name of queue.
&&	The logical AND operator indicates whether both operands are true.
\diamond	Not equal operator.
STATE	Current status of the server.
Main Serv n	Name of server that we want to check.

The expression “NQ(Server n .Queue) < 15” means the system not allow to have more than 15 workloads in the queue of the current server. The condition will return true if the current number of workload in the queue not exceed 15 at a time. We use First In First Out type of queue.

The expression “STATE(Main Serv n) \diamond -4” means the current status of server must not return value -4. The value -4 denote as FAILED status of resource in Arena. If the current state of the server which we want to check return -1(IDLE), -2(BUSY) and -3(INACTIVE) means the condition is true.

The “&&” denote the AND operator. This means both left and right expression must be true to satisfy the condition. If the condition is true, incoming workloads will assign to the main server and process their work, finish it and leave the system. If the condition is false, first and second model will provide different solution. For the first model when the condition is false, the system will scale up the capacity to the secondary server and incoming workloads will assign to the secondary server instead. For the second model when the condition is false, the system will scale up the capacity to another virtual

server in another main server. The selection of virtual server based on current number in queue of the main server.

In this step, we named the decide module as “Chk Serv *n* Status” and set the module type as “2-way by condition”. In the decide module’s dialog, select expression mode and enter the expression above in the value’s box and press OK button. The correct and complete value in this module will show in Figure 3-8.

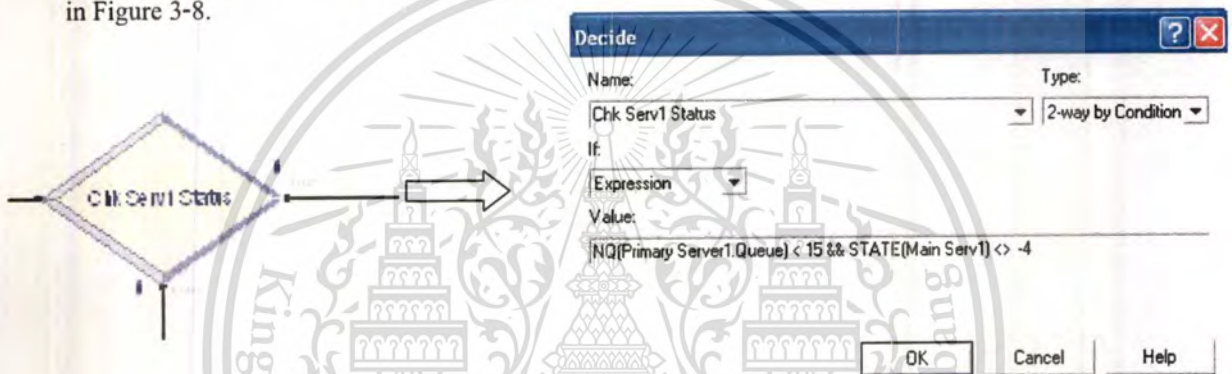


Figure 3-8: Decide dialog with correct and complete values

In case of all servers unavailable at the same time the result and solution of first and second model will be different. First model, server status will change to FAILED and cannot process anything for workloads. Current workloads and incoming workloads have to wait until the decided server back to the IDLE status and ready to perform again. Second model is different when the server status change to FAILED processing workloads and incoming workloads will be assign to another virtual server in another main server. This is the property of virtualization which used to operate virtual server. Every virtual server are in the same layer and able to work together. If some virtual servers are failed another virtual server which is not failed can support the remaining process without any problem. This property called “Live migration”.

So, we use Hold module as show in the Figure 3-9. In this module, we named it “Wait for available” and select “Scan for condition” mode and use the expression in the next page as a condition:

```
STATE(Main Serv1) <> -4 || STATE(Main Serv2) <> -4 || STATE(Main Serv3) <> -4 || STATE(Main Serv4) <> -4
```

The || denote the OR operator. This operator indicates whether either operand is true. Another description of parameter and operator was described in the previous expression.

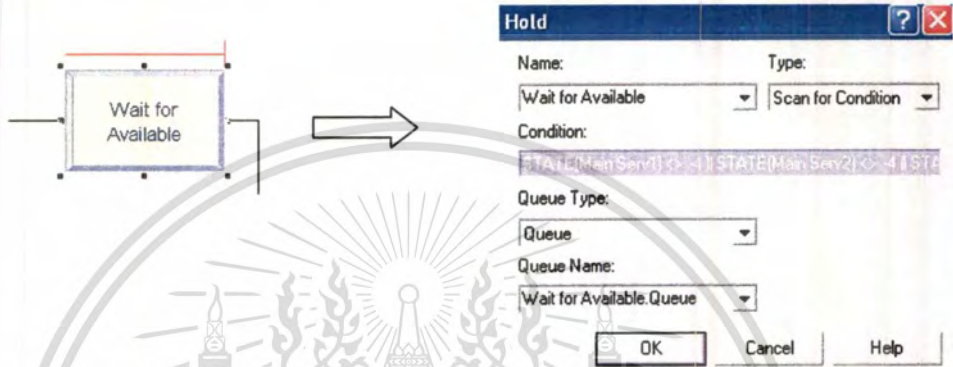


Figure 3-9: Hold module dialog

3.7 The process

The process step is just let the workload perform their work in the available server. This step of the first model and the second model is not the same.



Figure 3-10: Process module dialog of first model

Figure 3-10 show the detail about the parameter and required values for the first model. In the logic part, we select “Seize Delay Release” for the resources names “Main Serv1” and allow only one quantity of resource to use each time. Seize delay release means we want this module to take care of the workload’s seizing some number of units of a resource (after a possible wait in queue), then delay for a time representing the service time, and then release unit(s) of resource so that other workloads can seize it. We select triangular distribution as delay type. The minimum process time is 10 seconds, the most likely process time is 15 seconds and maximum process time is 20 seconds. That is all for the process step of first model.

The triangular distribution is commonly used in situations in which the exact form of the distribution is not known, but estimates for the minimum, maximum and most likely values are available. Our model has the situation that the processing service time for each workload is unknown but can estimate it.

The process’s detail of second model is not the same as Figure 3-10. The different is delay type, we select expression mode and enter the EXPO function and its mean is 3 seconds. Another different is second model does not has secondary server as show in Figure 3-11.

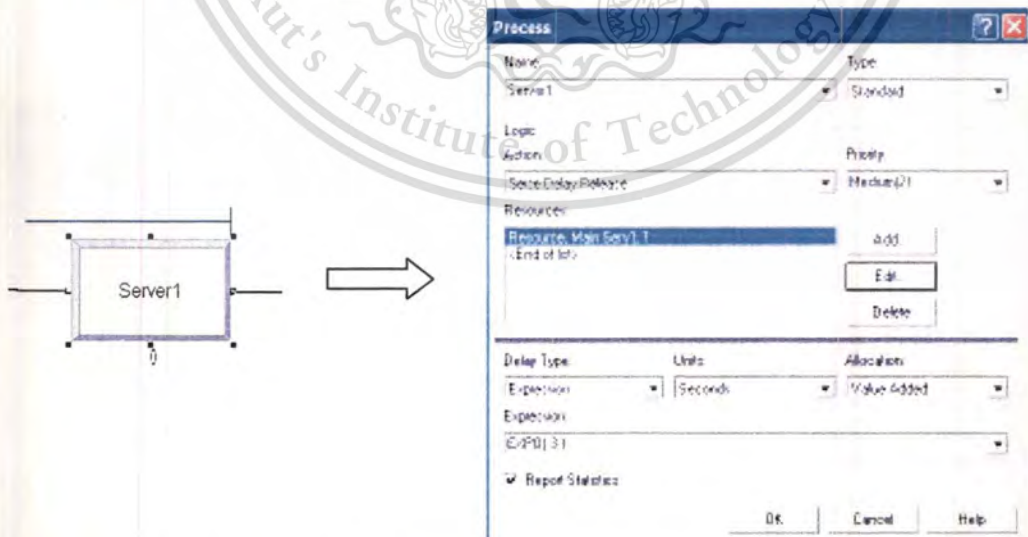


Figure 3-11: Process module dialog of second model

The full name of EXPO is EXPONENTIAL. This distribution is often used to model interevent times in random arrival and break-down processes. We use this distribution because each workload in second model has to assign to the virtual server which run on some physical server. This process may need the time to make the decision in order to assign workload to the suitable virtual server. The suitable servers which run on the virtual layer may operate in another physical server.

After physical main server decide the suitable server for the workload. The system will check the current number in virtual server's queue. So, we use "Decide" module again with 2-way by condition mode as show in Figure 3-12. The condition that we use to check the availability of virtual server is:

$$NQ(\text{Virtual Server } n.\text{Queue}) \leq 8$$

This expression mean if current number in queue of Virtual Server number n is less than or equal to 8, the condition will return true. If the condition is true, incoming workloads will go to use the virtual server number n . If the condition is false, incoming workloads will go to check another virtual server which can perform the service for them.

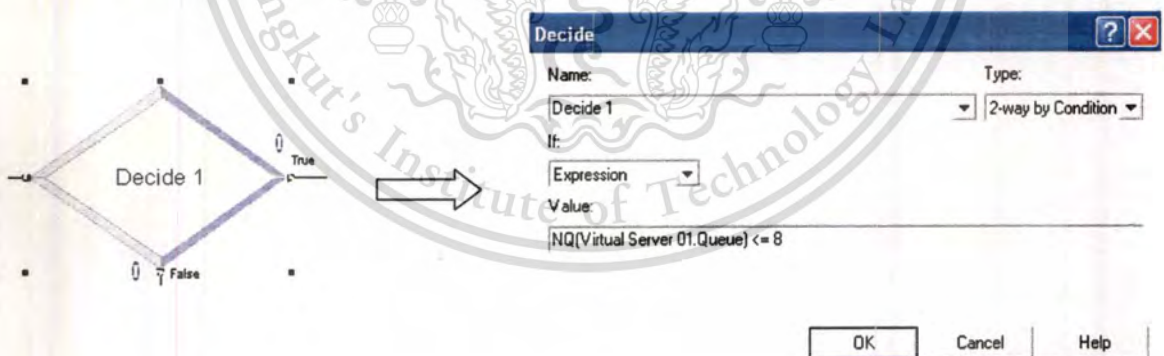


Figure 3-12: Decide module for virtual server

In case of every virtual server have number of waiting in queue more than 8. The system will assign the incoming workload to the less number of waiting in queue. To do so, we have to use "Pick Station" module as show in Figure 3-13.

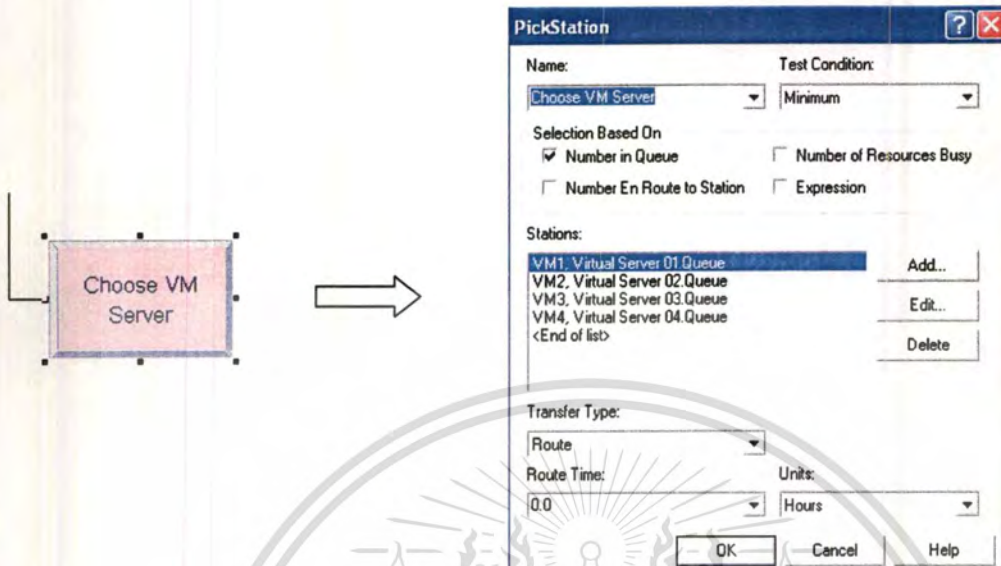


Figure 3-13: Pick station dialog

We named it as “Choose VM Server” and select minimum test condition. The selection of server based on number in queue of the station in the list and transfer by route. The previous statement means this module will select virtual server based on which virtual server (station) has less number in queue and assign the workload to that virtual server.

After the system decide the suitable server for incoming workload. The virtual server will perform their work under the same setting as Process module of first model as show in Figure 3-14

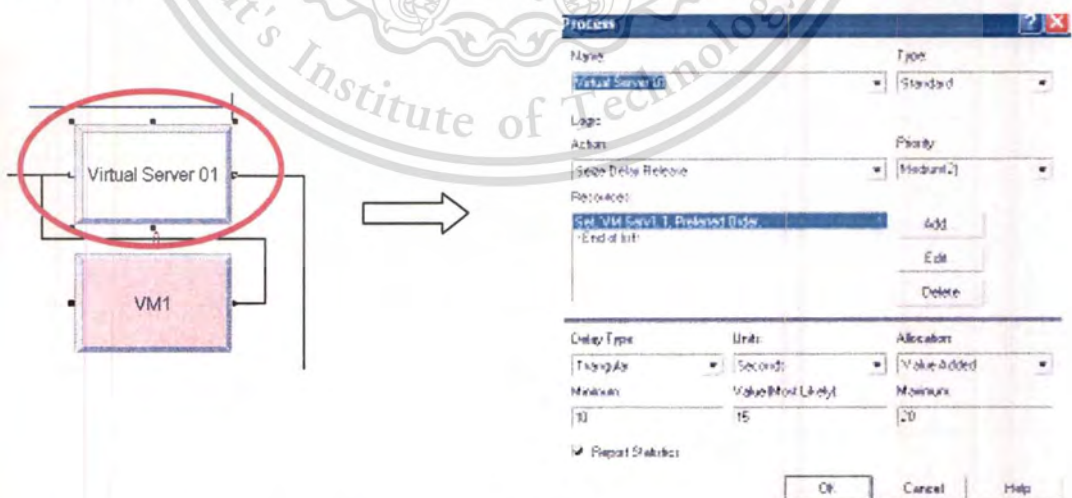


Figure 3-14: Process module dialog of second model

The pink module in Figure 3-14 is the station of this virtual server. This module use as the entrance of this virtual server in case of workload was assign by the “Pick Station” module. After workload enters to this station, it has to wait in the queue until its turn to use the virtual server in the process module. The setting of station module is very simple as show in Figure 3-15.

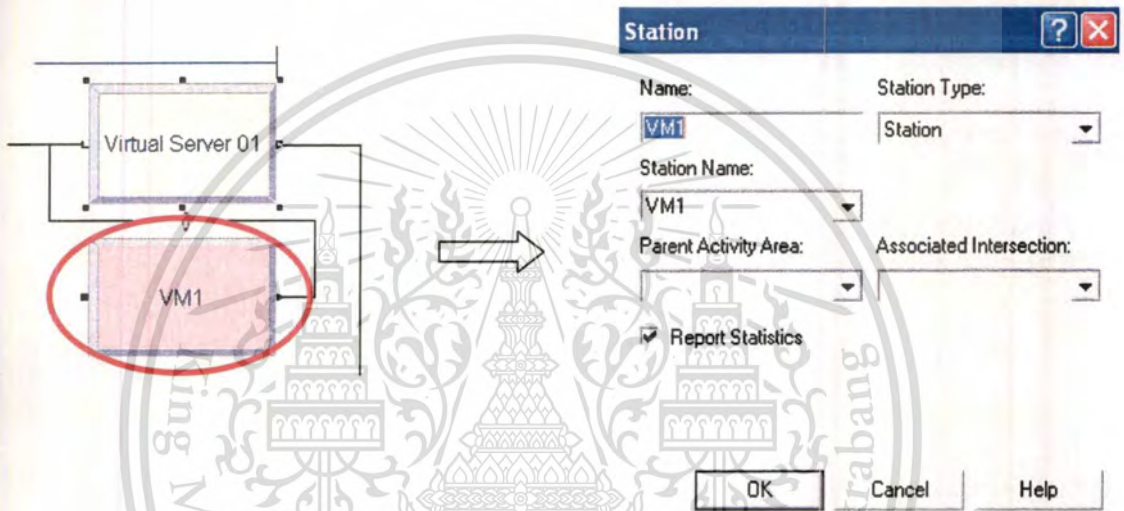


Figure 3-15: Station dialog

3.8 The end of process

After the server completely perform their services for workload. Every workload will leave the system. In the end of this process, we use “Dispose” module bring the workloads out of system. Dispose module does not need any setting. The only one thing that you can do with this module is named it. In our experiment we have named this dispose module as “End Process” as show in Figure 3-16.

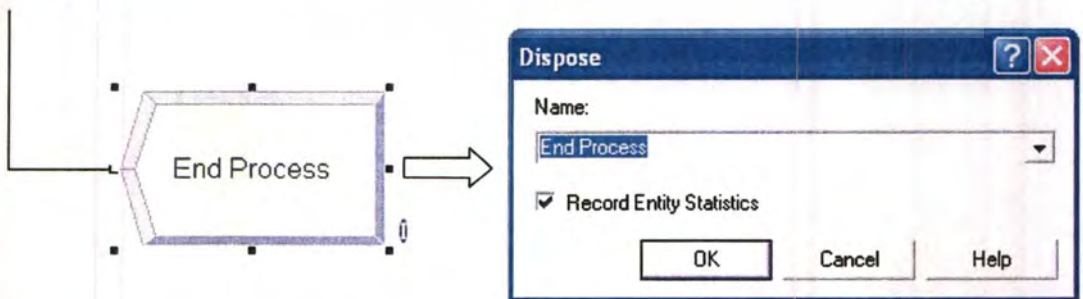


Figure 3-16: Dispose dialog

3.9 Model running set up

When everything is ready to simulate, we have to setup the necessary parameter for our model. This step is very important because this is part of system's environment. Project parameter and project statistics is setup here too.

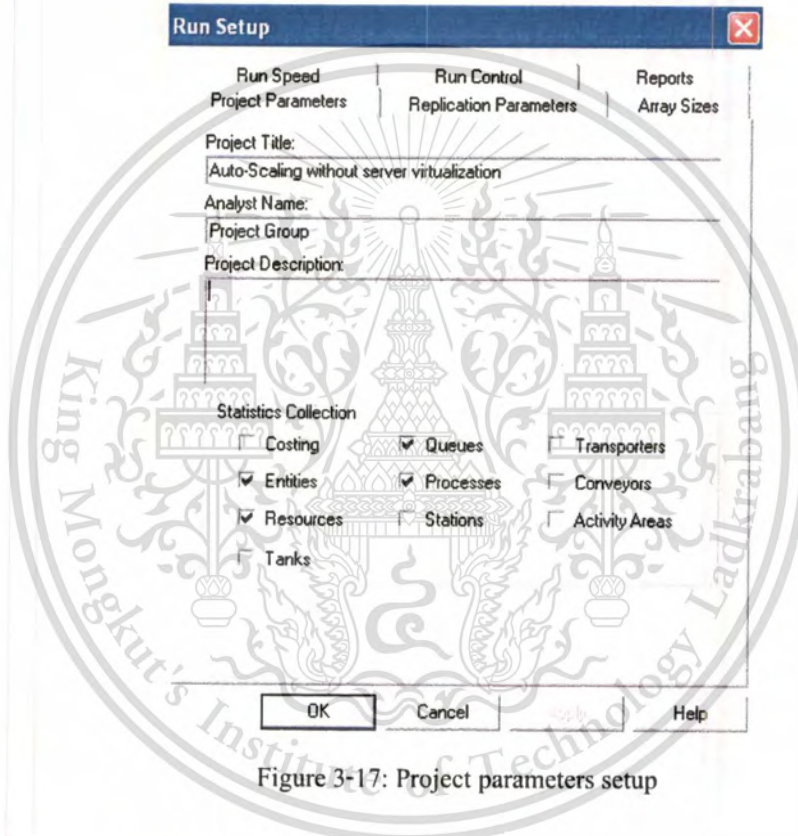


Figure 3-17: Project parameters setup

To setup the model running, click menu “Run” and select “Setup...” then run setup dialog will appear. In project parameters tab, you can enter the project title, analyst name and project description as you want. In statistics collection part, we select queue, entities, process and resources. This means we want to collect the statistics data in four areas of them.

- **Entities:** this is the statistics about workload; provide the information about total number of all type of work and number of workload using each server and virtual machine server.
- **Resources:** provide all the information about exist resources such as instantaneous utilization, number of busy, schedule utilization, etc.
- **Queues:** provide the information about waiting time in the queue and number of waiting in queue.
- **Processes:** providing the information about the process in the system.

Run Setup

Run Speed | Run Control | Reports
Project Parameters | Replication Parameters | Array Sizes

Number of Replications: Initialize Between Replications
 Statistics System

Start Date and Time:

Warm-up Period: Time Units:

Replication Length: Time Units:

Hours Per Day: Base Time Units:

Terminating Condition:

Figure 3-18: Replication parameters setup

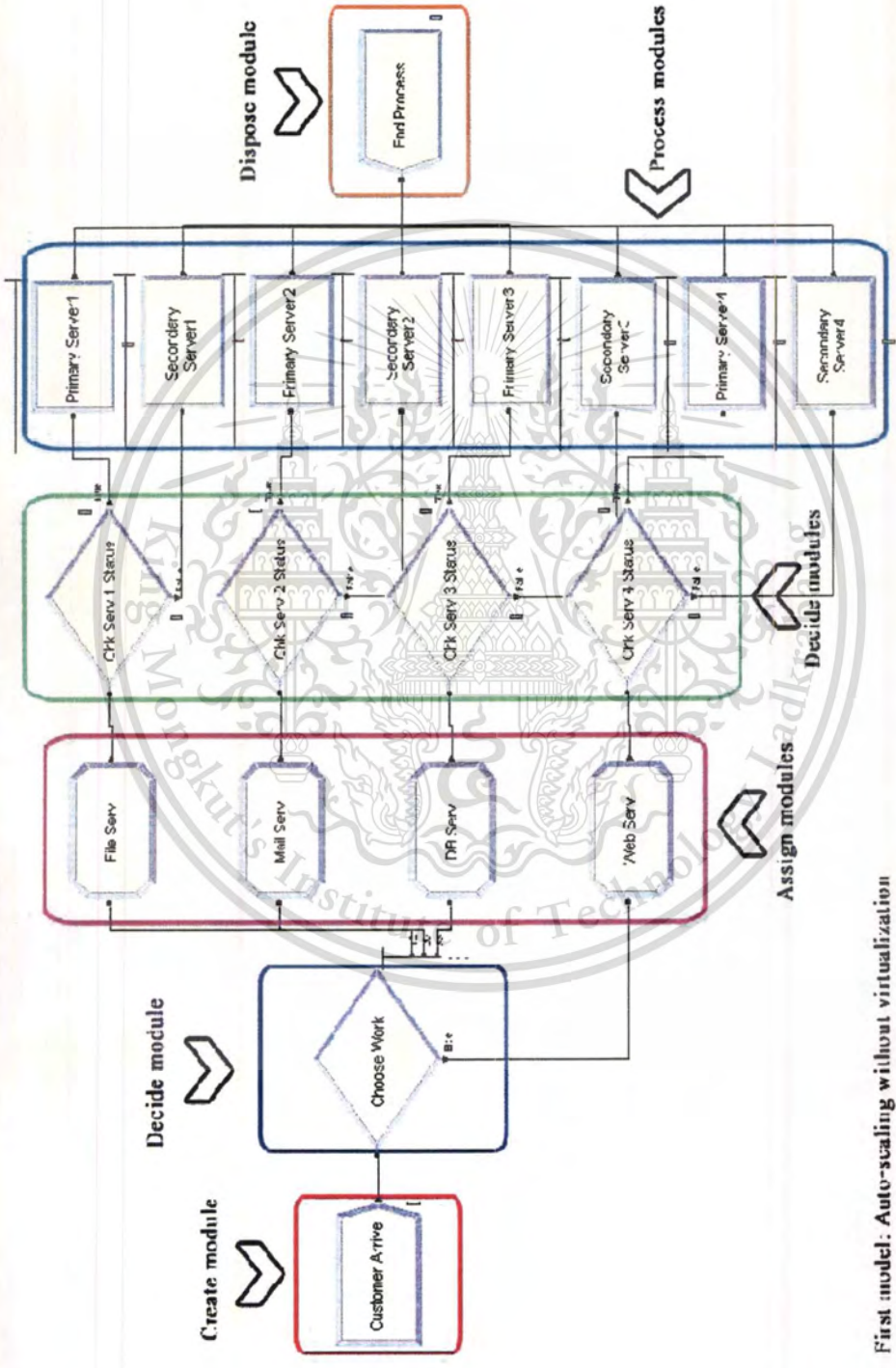
Before testing our model, we have to set up some necessary parameter for replication in order to make the testing more accurate as much as possible. The following is the list of necessary parameters for our model testing.

- Number of replication: enters the number of round that the model will run. The more replication numbers the more accuracy data you will get.
- Replication length: indicates the length per round of each replication of this model.
- Hours per day: indicates the detail of hours per day that our model will operate.
- Terminating condition: indicates the condition that will stop the running of the model.

After everything is ready press OK button and then press the ► button from the toolbar that show in Figure 3-19. Finally, let Arena simulate your model.

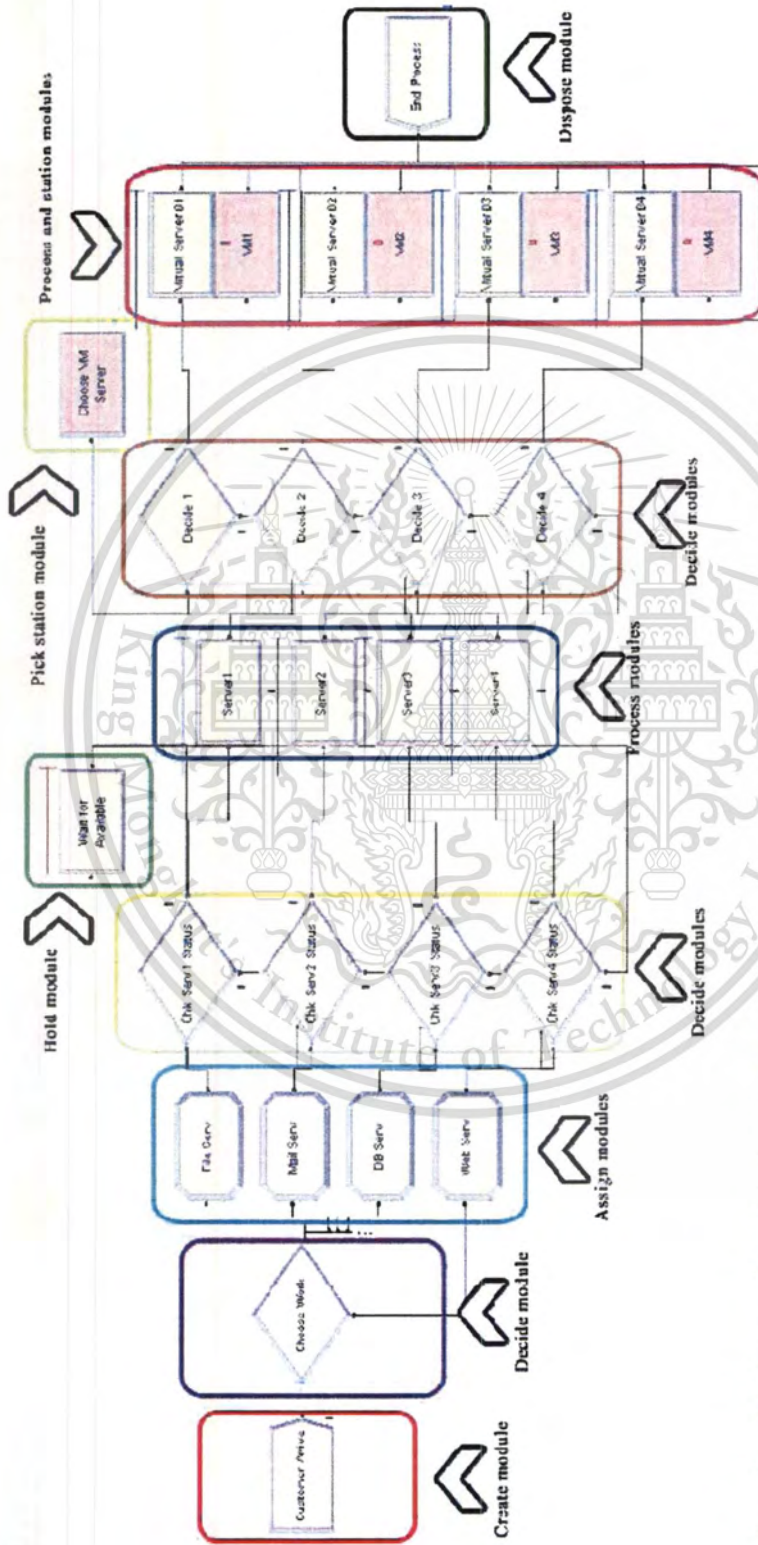


Figure 3-19: Arena toolbar



First model: Auto-scaling without virtualization

Figure 3-20: First model: Auto-scaling without virtualization



Second model: Auto-scaling with virtualization

Figure 3-21: Second model: Auto-scaling with virtualization

Chapter 4

Experimental result

4.1 Experimental design

In our experiment, we set up the experiment as the following step.

1. Install Arena Rockwell version 11.
2. Create 2 models: auto – scaling with server virtualization model and auto – scaling without server virtualization model.

The illustration of models is shown below. Figure 4 – 1 is illustrated of auto – scaling without server virtualization model. Figure 4 – 2 is illustrated of auto - scaling with server virtualization model.

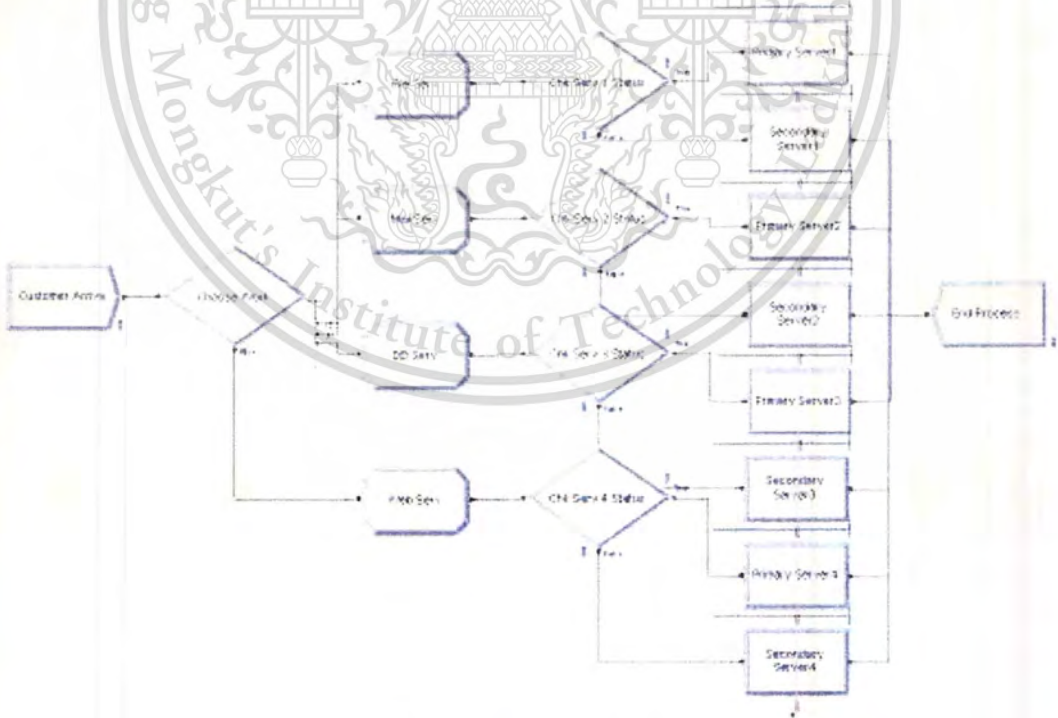


Figure 4 – 1: Auto – scaling without server virtualization model.

Table 4-1: Failure rate of servers in auto –scaling without server virtualization model

Server name	Type	Uptime	Units	Downtime	Units
File server	Time	EXPO(365)	Days	$-0.5 + \text{ERLA}(0.0204, 25)$	Hours
Secondary file server	Time	EXPO(8500)	Hours	$-0.5 + \text{ERLA}(0.0204, 25)$	Hours
Mail server	Time	EXPO(365)	Days	$-0.001 + \text{EXPO}(0.00374)$	Hours
Secondary mail server	Time	EXPO(8500)	Hours	$-0.001 + \text{EXPO}(0.00374)$	Hours
Database server	Time	EXPO(365)	Days	$\text{TRIA}(0.5 , 0.75 , 1)$	Hours
Secondary database server	Time	EXPO(8500)	Hours	$\text{TRIA}(0.5 , 0.75 , 1)$	Hours
Web server	Time	EXPO(365)	Days	$-0.5 + \text{LOGN}(0.58, 0.229)$	Hours
Secondary web server	Time	EXPO(8500)	Hours	$-0.5 + \text{LOGN}(0.58, 0.229)$	Hours

We also define the failure rate of servers in auto –scaling with server virtualization model. See on Table 4 – 2.

Table 4-2: Failure rate of servers in auto –scaling with server virtualization model

Server name	Type	Uptime	Units	Downtime	Units
File server	Time	EXPO(365)	Days	$-0.5 + \text{ERLA}(0.0204, 25)$	Hours
VM file server fail	Time	EXPO(8500)	Hours	$\text{TRIA}(0.05 , 0.25 , 0.5)$	Hours
Mail server	Time	EXPO(365)	Days	$-0.001 + \text{EXPO}(0.00374)$	Hours
VM mail server fail	Time	EXPO(8500)	Hours	$\text{TRIA}(0.05 , 0.25 , 0.5)$	Hours
Database server	Time	EXPO(365)	Days	$\text{TRIA}(0.5 , 0.75 , 1)$	Hours
VM DB server fail	Time	EXPO(8500)	Hours	$\text{TRIA}(0.05 , 0.25 , 0.5)$	Hours
Web server	Time	EXPO(365)	Days	$-0.5 + \text{LOGN}(0.58, 0.229)$	Hours
VM web server fail	Time	EXPO(8500)	Hours	$\text{TRIA}(0.05 , 0.25 , 0.5)$	Hours

Remark: VM stands for Virtual Machine. DB stands for Database.

5. We collect statistics in four areas. There are queues, entities, resources, processes.
6. We run both models in 10 replications. Each replication length has 365 days in time unit. Each day in simulation environment has 24 hours. So you can see the running setup on Table 4 -3 below.

Table 4-3: Running setup

Replication parameters name	Value
Number of replications	10
Replication length	365
Time units	Days
Hours per day	24
Base time units	Hours

7. Run models
8. Analyze and summarize the result.

4.2 Experimental result

After running the models in 10 replications then we measure them in two significant categories. There is mean time to failure and CPU utilization. However, we also measure the models in term of Average workload waits time, Average wait time queue and Average number of waiting time.

4.2.1 Mean time to failure

This is the first measurement. We calculate the mean time to failure (MTTF) from the formula below.

$$MTTF = \frac{\sum_{i=1}^N t_i}{N}$$

Where t is time before the server down.

N is the number of failure.

We collect the number of failure and failure time by observation in each model. You will see the topic later.

4.2.1.1 Mean time to failure of auto-scaling without server virtualization model

Table4-4 is shown the number of failure and failure time after finish running model in 10 replications. We collect the statistics data by observation. See on Table 4 – 4.

Table4-4: Show the number of failure of auto-scaling without server virtualization model

Replication no.	Error no.	Failure time(Hour)	Number of failure
1	1	513	7
	2	2,308	
	3	2,947	
	4	3,788	
	5	4,238	
	6	5,311	
	7	6,419	
2	1	824	5
	2	920	
	3	4,473	
	4	6,000	
	5	8,480	
3	1	59	10
	2	1,170	
	3	3,803	
	4	5,147	
	5	5,627	
	6	6,061	

Table4-4: Show the number of failure of auto-scaling without server virtualization model (Cont.)

Replication no.	Error no.	Failure time(Hour)	Number of failure
	7	6,274	
	8	6,927	
	9	8,221	
	10	8,503	
4	-	-	0
5	1	586	3
	2	1,880	
	3	7,037	
6	1	2,793	3
	2	6,998	
	3	8,005	
7	1	275	3
	2	3,275	
	3	5,718	
8	1	736	5
	2	2,099	
	3	2,393	
	4	5,920	
	5	8,055	
9	1	33	4

Table4-4: Show the number of failure of auto-scaling without server virtualization model (Cont.)

Replication no.	Error no.	Failure time(Hour)	Number of failure
	2	4,548	
	3	7,363	
	4	8,010	
10	1	4,445	4
	2	4,833	
	3	5,528	
	4	5,720	
Total			44

Next step we find the time before failure. Table4-5 is show how to calculate the time before failure.

Table4-5: Show how to calculate the time before failure

Replication no.	Failure time(Hour)	Time before failure
1	514	514
	2,308	1,794
	2,947	639
	3,788	841
	4,238	450
	5,311	1,073
	6,419	1,108
	Total	6,419

Table4-5: Show how to calculate the time before failure (Cont.)

Replication no.	Failure time(Hour)	Time before failure
2	824	824
	920	96
	4,473	3,553
	6,000	1,527
	8,480	2,480
	Total	8,480
3	59	59
	1,170	1,111
	3,803	2,633
	5,147	1,344
	5,627	480
	6,061	434
	6,274	213
	6,927	653
	8,221	1,294
	8,503	282
	Total	8,503
4	0	0
	Total	0

Table4-5: Show how to calculate the time before failure (Cont.)

Replication no.	Failure time(Hour)	Time before failure
5	586	586
	1,880	1,294
	7,037	5,157
	Total	7,037
6	2,793	2,793
	6,998	4,205
	8,005	1,007
	Total	8,005
7	275	275
	3,275	3,000
	5,718	2,443
	Total	5,718
8	736	736
	2,099	1,363
	2,393	294
	5,920	3,527
	8,055	2,135
	Total	8,055
9	33	33
	4,548	4,515
	7,363	2,815

Table4-5: Show how to calculate the time before failure (Cont.)

Replication no.	Failure time(Hour)	Time before failure
	8,010	647
	Total	8,010
10	4,445	4,445
	4,833	388
	5,528	695
	5,720	192
	Total	5,720
	Total in 10 replications	65,947

Therefore the mean time to failure of auto-scaling without server virtualization model is $65,947/44$ or equal to 1,498.80 hours per year in 10 replications.

We can say that in one year one replication, this system has 149.88 uptime hours before the system will down.

4.2.1.2 Mean time to failure of auto-scaling with server virtualization

In this model, we find the mean time to failure (MTTF) as same as previous topic. You can see the data which is shown on Table 4-6 below.

Table4-6: Show the number of failure and time before failure of auto-scaling with server virtualization model

Replication no.	Failure time(Hour)	Time before failure	Number of failure
1	880	880	
	1,560	680	
	3,367	1,807	
	Total	3,367	3
2	1,809	1,809	
	1,407	402	
	Total	2,211	2
3	5,074	5,074	
	6,059	985	
	8,602	2,543	
	Total	8,602	3
4	1,114	1,114	
	4,222	3,108	
	4,283	61	
	5,863	1,580	
	Total	5,863	4
5	1,220	1,220	
	Total	1,220	1

Table4-6: Show the number of failure and time before failure of auto-scaling with server virtualization model (cont.)

Replication no.	Failure time(Hour)	Time before failure	Number of failure
6	3,543	3,543	
	7,470	3,927	
	7,880	410	
	Total	7,880	3
7	0	0	
	Total	0	0
8	2,660	2,660	
	5,600	2,940	
	8,566	2,966	
	Total	8,566	3
9	1,800	1,800	
	3,469	1,669	
	Total	3,469	2
10	7,353	7,353	
	Total	7,353	1
	Total in 10 replications	48,531	22

We can calculate the MTTF of this model by using the data from Table 4-6.

$$\begin{aligned}
 \text{So MTTF} &= 48,531/22 \\
 &= 2,205.95455 \text{ hours/ 10 replications} \\
 \text{1 replication} &= 2,205.95/10 \\
 &= 220.60 \text{ hours}
 \end{aligned}$$

4.2.2 Key performance indicators

There are three additional measurements after running model. See the next topic later.

4.2.2.1 Average waiting time per entity

Figure4-3 depicts the average waiting time per entity which is a measurement of the model. We found that the average waiting time per entity of every server in auto-scaling without server virtualization model is higher than the average waiting time per entity of servers in auto-scaling with server virtualization model. This bar chart shows clearly the primary web server (physical server) using server virtualization (red bar), has the average waiting time per entity (workload) less than servers that do not use server virtualization (blue bar).

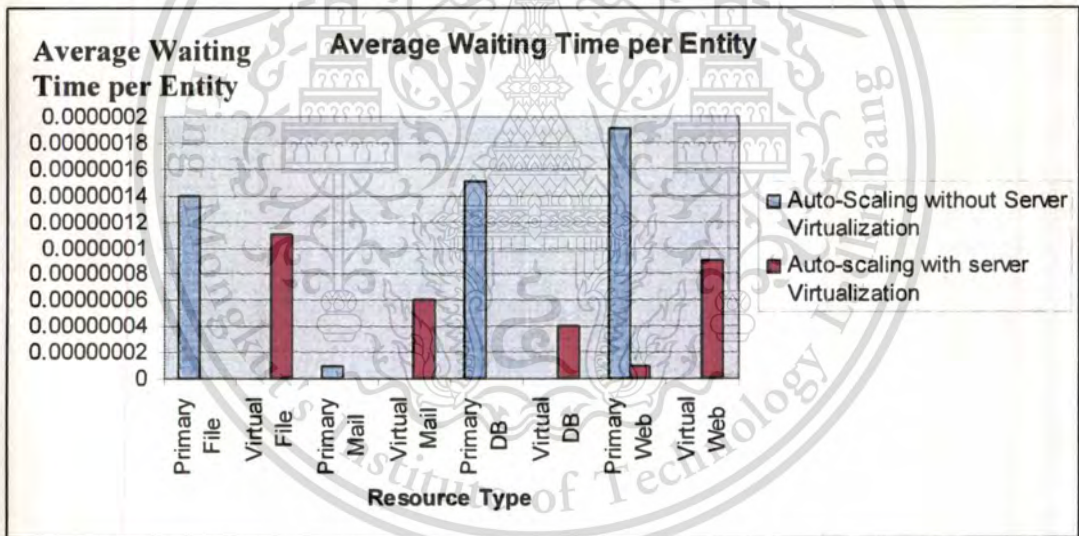


Figure4-3: Average waiting time per entity

4.2.2.2 Average wait time queue

This figure illustrates the average waiting time in queue which is mean time of workloads wait in queue for processing. We found that most of physical server in auto-scaling without server virtualization model has the average waiting time higher than the physical server in auto-scaling with server virtualization model. However, if we compare the other virtual servers in auto-scaling with server

virtualization model, we found that virtual database server and virtual web server still have to wait in queue because most of users often request service from them.

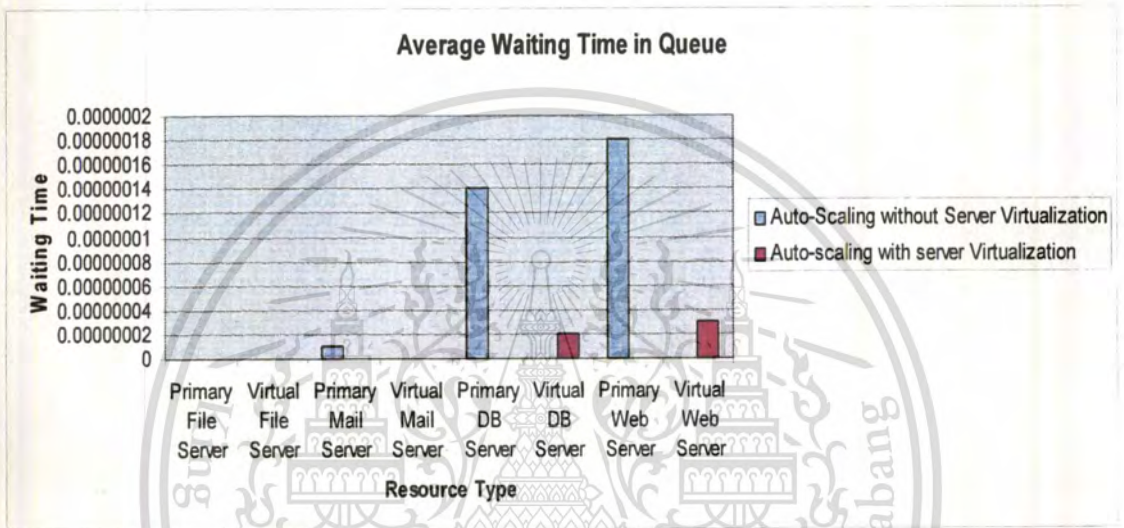


Figure 4-4: Average waiting time

4.2.2.3 Average number of waiting queue

In Figure 4-5 depicts the average number waiting queue of most servers in auto-scaling with server virtualization model which is less than the average number waiting queue of most servers in auto-scaling without server virtualization model. It can be seen that using server virtualization in order to support the heavy workloads at peak time reducing the number of waiting queue.

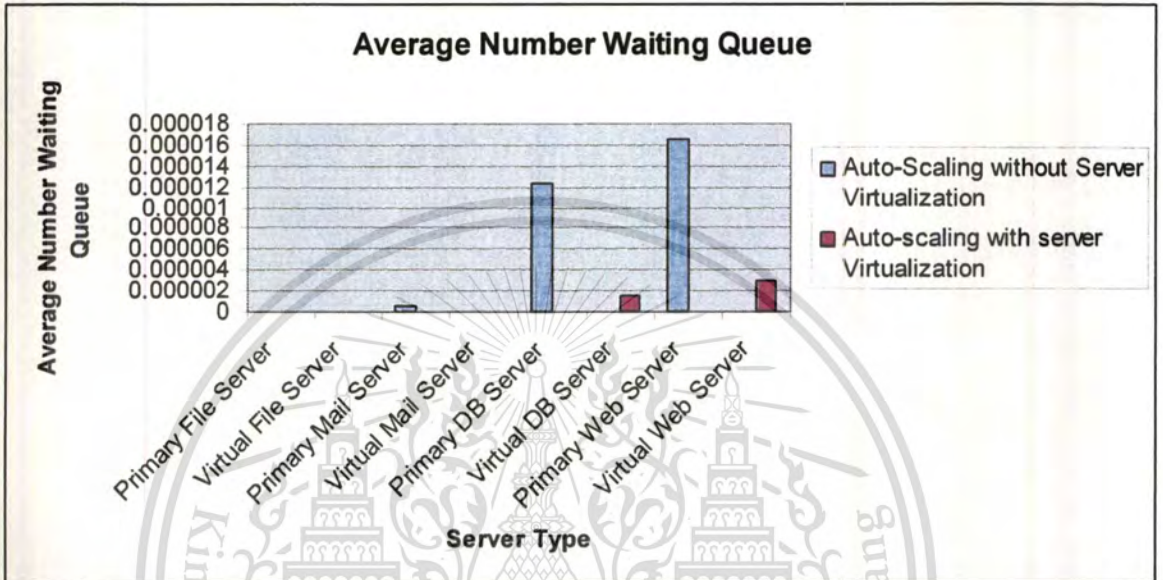


Figure4-5: Average number waiting in queue

4.2.2.4 Average resource usage

This measurement represents the CPU utilization. Figure4-6 depicts clearly that using server virtualization to support the heavy workloads increase the average resource usage of each server. The gap can be seen between the blue line (servers plus virtual servers, in auto-scaling with server virtualization model) and pink line (only physical server, in auto-scaling without server virtualization model). See on Figure 4 -6 below.

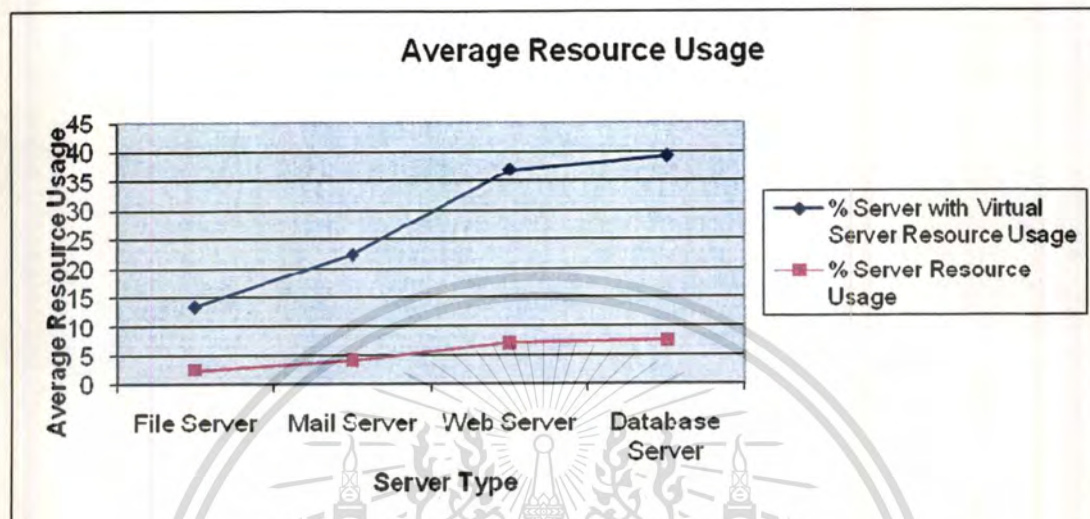


Figure 4-6: Average resource usage of server

Moreover, we show the percentage of the increasing resource usage in each different server type. The reason of showing this percentage is to give your understanding better. See on Table 4 -6.

Table4-7: Show the increasing resource usage of each server type

Server's name	% Server with virtual server resource usage	% Server resource usage	% Increasing resource usage
File server	13.438071	2.572906	10.86517
Mail server	22.37105	4.286299	18.08475
Web server	36.917985	7.079374	29.83861
Database server	39.185817	7.504679	31.68114

From the table above, we measure the overall CPU utilization in different models. We found the CPU utilization of auto – scaling with server virtualization model has increased 90.5% from the CPU utilization of auto – scaling without server virtualization model.

4.3 Comparison

We summarize that auto-scaling with virtualization model provides the number of MTTF and many types of waiting time are more effectively and efficiency. The result of the work in process is auto-scaling without virtualization model is more flexible because auto-scaling with virtualization model has to find the server that available from the existing physical server. On the other hand, auto-scaling without virtualization has not to do because each type of work has its own primary and secondary server. So, that is all the reasons why WIP of auto – scaling with virtualization model is less flexible than auto – scaling with virtualization. The comparison is show in Table 4-8.

Table4-8: Comparison between auto-scaling without virtualization model and auto-scaling with virtualization model

Comparison	Auto – scaling without virtualization model	Auto – scaling with virtualization model
Mean time to failure (Average time before failure / Number of failure)	149.88 hours	220.60 hours
Average waiting time per entity	Incoming workload spent much waiting time.	Incoming workload spent shorter waiting time.
Average waiting time in queue	Long time waiting in current queue of workload.	Shorter waiting time in current queue of workload.
Average number of waiting in queue	Longer queue.	Shorter queue.
Average resource usage of server	Less effective usage. Less utilization.	More effective usage. More utilization.
Work in process	Less number of works in process. More system flexibility.	More number of works in process. Less system flexibility.
Average wait time of each type of work	0.0000001225 hours	0.0000000775 hours

Chapter 5

Conclusion and recommendation

5.1 Conclusion

In this paper, we simulate the feasibility of virtualization technology especially server virtualization to auto –scaling problem. The purpose of simulation is to prove that using virtualization technology in servers can increase the mean time to failure and also increase the CPU utilization. We use ARENA software as a tool to do experiment. We create two different models. There are auto –scaling without server virtualization and auto – scaling with server virtualization. Auto –scaling without server virtualization is a model that does not implement virtualization technology. Otherwise auto – scaling with server virtualization is a model that implements virtualization technology and share the cross server's resource type. We run the models in 10 replications. Each replication length has 365 days. Each simulated day has 24 hours.

The result of our experiment shows that implementing virtualization technology on servers can be increased the mean time to failure (MTTF) (220.60 hours per year). Otherwise the mean time to failure of none implementing virtualization technology on servers is 149.88 hours per year. In addition, the CPU utilization of servers in auto-scaling with virtual server model is increased 90.5%.

Besides, we found that implementing virtualization in server can reduce the waiting time in queue, a number of waiting in queue and waiting time per entity. However, our experiment concentrates on server virtualization only.

5.2 Recommendation

In this paper, we recommend to use real input data for accurate output. The replication length should more than 1 year to study. However, this paper concentrates on server virtualization only. In reality virtualization technology is not only server virtualization but also a storage virtualization. It is an abstract storage, separating out logical data access from physical access. So this paper does not include storage virtualization.

Bibliography

- [1] **A different more critical look at auto-scaling now on Amazon EC2.** [Online] Available at: <http://tek-tips.nethawk.net/blog/a-different-more-critical-look-at-auto-scaling-now-on-amazon-ec2> (Accessed at: 4 October 2009).
- [2] **Amazon EC2.** [Online] Available at: <http://www.amazon.com/gp/browse.html?node=201590011> (Accessed: 20 September 2009).
- [3] **Amazon EC2 New Features: Elastic Load Balancing, Auto-Scaling, and Amazon CloudWatch.** [Online] Available at: www.azurejournal.com/2009/05/amazon-ec2-new-features-elastic-autoscaling.html (Accessed at: 9 December 2009).
- [4] **Auto Scaling Developer Guide API Version 2009-05-15.** [Online] Available at: <http://docs.amazonwebservices.com> (Accessed at: 4th September 2009).
- [5] **Autoscaling Web service Amazon EC2.** [Online] Available at: people.apache.org/~azeez/autoscaling-web-services-azeez.pdf (Accessed at: 4 October 2009).
- [6] **HPC: High Performance Computing.** [Online] Available at: <http://www.ddj.com/hpc-high-performance-computing/208800138> (Accessed at: 5 September 2009).
- [7] **IT Labour.** [Online] Available at: <http://www.oknation.net/blog/itlabour> (Accessed 2 October 2009).
- [8] **Nati Shalom's Blog: Its time for auto scaling – avoid peak load provisioning for web applications.** [Online] Available at: http://natishalom.typepod.com/nati_shalom_blog/2009/03/its-time-for-auto-scaling-avoid-peak-load-provision.html (Accessed at: 7 October 2009).
- [9] **Server Virtualization Cloud Computing.** [Online] Available at: <http://malaysiavm.com/blog/server-virtualization-cloud-computing/>, (Accessed at 4 September 2009).
- [10] **VMware.** [Online] Available at: <http://vmware.com/solutions/consolidation/> (Accessed at: 3rd October 2009).

[11] What is cluster HPC & Virtualization. [Online] Available at: <http://tn-home.blogspot.com/2009/08/what-is-cluster-hpc-ha-virtualization.html> (Accessed at: 5th September 2009).

[12] Kelton, W. D., Sadowski, R. P. and Sturrock, D. T., 2007. **Simulation with Arena**, fourth edition, McGraw-Hill Higher Education.



Authors Biography

Name: Douanghatai Thippayawong

Date of Birth: June 15th, 1987

Place of Birth: Uttaradit, Thailand

Education: 2000-2006, High School and Junior High School
Uttaraditdaruni School
Uttaradit, Thailand

Name: Thithat Thepparat

Date of Birth: February 24th, 1988

Place of Birth: Sakaeo, Thailand

Education: 2000-2006, High School and Junior High School
Kanchanaphisekwittayalai Chachoengsao School
Chachoengsao, Thailand

Name: Amnart Harnprasamkit

Date of Birth: December 29th, 1987

Place of Birth: Phuket, Thailand

Education: 2000-2005, High School and Junior High School
Phuketwittayalai School
Phuket, Thailand



This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Auto – Scaling without Server Virtualization Model Report



This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Auto

Replications: 10

Time Units: Hours

Key Performance Indicators

System

Number Out

Average

2,254,151



This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Values Across All Replications

Auto

Replications: 10 Time Units: Hours

Entity

Time

VA Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Work1	0.00416668	0.00	0.00416536	0.00416774	0.00277842	0.00555455
Work2	0.00416665	0.00	0.00416572	0.00416929	0.00277802	0.00555480
Work3	0.00416662	0.00	0.00416532	0.00416788	0.00277871	0.00555465
Work4	0.00416656	0.00	0.00416509	0.00416744	0.00277877	0.00555459

NVA Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Work1	0.00	0.00	0.00	0.00	0.00	0.00
Work2	0.00	0.00	0.00	0.00	0.00	0.00
Work3	0.00	0.00	0.00	0.00	0.00	0.00
Work4	0.00	0.00	0.00	0.00	0.00	0.00

Wait Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Work1	0.00000014	0.00	0.00	0.00000084	0.00	0.2269
Work2	0.00000001	0.00	0.00000001	0.00000002	0.00	0.00212423
Work3	0.00000015	0.00	0.00000012	0.00000019	0.00	0.00595484
Work4	0.00000019	0.00	0.00000016	0.00000027	0.00	0.04953847

Transfer Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Work1	0.00	0.00	0.00	0.00	0.00	0.00
Work2	0.00	0.00	0.00	0.00	0.00	0.00
Work3	0.00	0.00	0.00	0.00	0.00	0.00
Work4	0.00	0.00	0.00	0.00	0.00	0.00

Other Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Work1	0.00	0.00	0.00	0.00	0.00	0.00
Work2	0.00	0.00	0.00	0.00	0.00	0.00
Work3	0.00	0.00	0.00	0.00	0.00	0.00
Work4	0.00	0.00	0.00	0.00	0.00	0.00

Total Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Work1	0.00416682	0.00	0.00416584	0.00416823	0.00277842	0.2312
Work2	0.00416666	0.00	0.00416575	0.00416929	0.00277802	0.00681385
Work3	0.00416677	0.00	0.00416549	0.00416800	0.00277871	0.00994693
Work4	0.00416675	0.00	0.00416527	0.00416763	0.00277877	0.05326460

Other

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Values Across All Replications

Auto

Replications: 10 Time Units: Hours

Entity

Other

Number In	Average	Half Width	Minimum Average	Maximum Average
customer	2254151.50	991.00	2252684.00	2256608.00
Work1	270466.30	444.16	269815.00	271738.00
Work2	450577.60	435.44	449497.00	451478.00
Work3	744192.40	867.25	742498.00	746686.00
Work4	788915.20	574.03	787722.00	790413.00



Number Out	Average	Half Width	Minimum Average	Maximum Average
customer	2254151.50	991.00	2252684.00	2256608.00
Work1	270466.10	444.29	269815.00	271738.00
Work2	450577.50	435.49	449497.00	451478.00
Work3	744192.30	867.25	742498.00	746686.00
Work4	788915.00	574.12	787722.00	790413.00

WIP	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
customer	0.00	0.00	0.00	0.00	0.00	1.0000
Work1	0.1287	0.00	0.1283	0.1292	0.00	6.0000
Work2	0.2143	0.00	0.2138	0.2147	0.00	7.0000
Work3	0.3540	0.00	0.3532	0.3551	0.00	8.0000
Work4	0.3753	0.00	0.3747	0.3758	0.00	8.0000

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Values Across All Replications

Auto

Replications: 10 Time Units: Hours

Process

Time per Entity

VA Time Per Entity	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Primary Server1	0.00416667	0.00	0.00416536	0.00416774	0.00277842	0.00555455
Primary Server2	0.00416665	0.00	0.00416572	0.00416929	0.00277802	0.00555480
Primary Server3	0.00416662	0.00	0.00416532	0.00416788	0.00277871	0.00555465
Primary Server4	0.00416656	0.00	0.00416509	0.00416744	0.00277877	0.00555459
Secondary Server1	0.00172670	0.00	0.00	0.00455313	0.00	0.00531071
Secondary Server2	0.00037935	0.00	0.00	0.00379349	0.00	0.00379349
Secondary Server3	0.00076066	0.00	0.00	0.00478673	0.00	0.00478673
Secondary Server4	0.00153221	0.00	0.00	0.00401158	0.00	0.00540363
Wait Time Per Entity	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Primary Server1	0.00000014	0.00	0.00	0.00000084	0.00	0.2269
Primary Server2	0.00000001	0.00	0.00000001	0.00000002	0.00	0.00212423
Primary Server3	0.00000015	0.00	0.00000012	0.00000019	0.00	0.00595484
Primary Server4	0.00000019	0.00	0.00000016	0.00000027	0.00	0.04953847
Secondary Server1	0.00	0.00	0.00	0.00	0.00	0.00
Secondary Server2	0.00	0.00	0.00	0.00	0.00	0.00
Secondary Server3	0.00	0.00	0.00	0.00	0.00	0.00
Secondary Server4	0.00	0.00	0.00	0.00	0.00	0.00
Total Time Per Entity	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Primary Server1	0.00416682	0.00	0.00416583	0.00416823	0.00277842	0.2312
Primary Server2	0.00416666	0.00	0.00416575	0.00416929	0.00277802	0.00681385
Primary Server3	0.00416677	0.00	0.00416549	0.00416800	0.00277871	0.00994693
Primary Server4	0.00416675	0.00	0.00416527	0.00416763	0.00277877	0.05326460
Secondary Server1	0.00172670	0.00	0.00	0.00455313	0.00	0.00531071
Secondary Server2	0.00037935	0.00	0.00	0.00379349	0.00	0.00379349
Secondary Server3	0.00076066	0.00	0.00	0.00478673	0.00	0.00478673
Secondary Server4	0.00153221	0.00	0.00	0.00401158	0.00	0.00540363

Accumulated Time

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Values Across All Replications

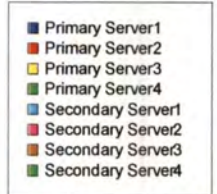
Auto

Replications: 10 Time Units: Hours

Process

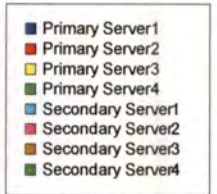
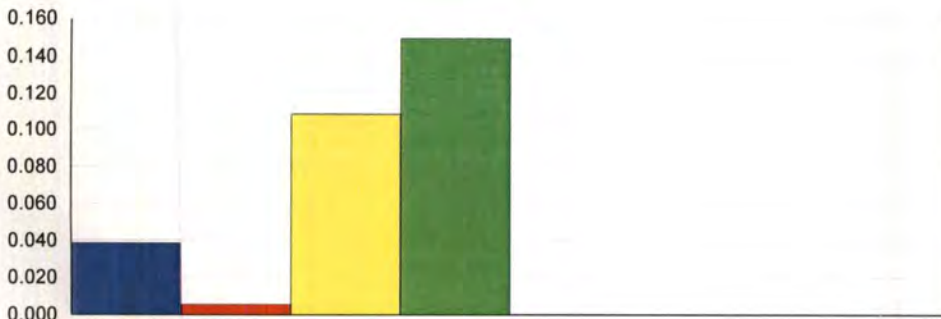
Accumulated Time

Accum VA Time	Average	Half Width	Minimum Average	Maximum Average
Primary Server1	1126.93	1.87	1123.86	1132.10
Primary Server2	1877.40	1.85	1872.70	1880.87
Primary Server3	3100.77	3.51	3093.79	3110.86
Primary Server4	3287.05	2.30	3282.01	3292.14
Secondary Server1	0.01213838	0.01	0.00	0.04343222
Secondary Server2	0.00037935	0.00	0.00	0.00379349
Secondary Server3	0.00076066	0.00	0.00	0.00478673
Secondary Server4	0.01267573	0.02	0.00	0.1003



Accum Wait Time

Accum Wait Time	Average	Half Width	Minimum Average	Maximum Average
Primary Server1	0.03840818	0.06	0.00	0.2269
Primary Server2	0.00544136	0.00	0.00269695	0.01118502
Primary Server3	0.1083	0.01	0.08789861	0.1376
Primary Server4	0.1496	0.02	0.1239	0.2105
Secondary Server1	0.00	0.00	0.00	0.00
Secondary Server2	0.00	0.00	0.00	0.00
Secondary Server3	0.00	0.00	0.00	0.00
Secondary Server4	0.00	0.00	0.00	0.00



Other

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Values Across All Replications

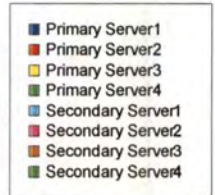
Auto

Replications: 10 Time Units: Hours

Process

Other

Number In	Average	Half Width	Minimum Average	Maximum Average
Primary Server1	270463.50	445.40	269810.00	271738.00
Primary Server2	450577.50	435.58	449496.00	451478.00
Primary Server3	744192.20	867.23	742498.00	746686.00
Primary Server4	788912.00	573.99	787719.00	790412.00
Secondary Server1	2.8000	2.74	0.00	10.0000
Secondary Server2	0.1000	0.23	0.00	1.0000
Secondary Server3	0.2000	0.30	0.00	1.0000
Secondary Server4	3.2000	5.55	0.00	25.0000



Number Out	Average	Half Width	Minimum Average	Maximum Average
Primary Server1	270463.30	445.53	269810.00	271738.00
Primary Server2	450577.40	435.63	449496.00	451478.00
Primary Server3	744192.10	867.23	742498.00	746686.00
Primary Server4	788911.80	574.08	787719.00	790412.00
Secondary Server1	2.8000	2.74	0.00	10.0000
Secondary Server2	0.1000	0.23	0.00	1.0000
Secondary Server3	0.2000	0.30	0.00	1.0000
Secondary Server4	3.2000	5.55	0.00	25.0000

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Values Across All Replications

Auto

Replications: 10 Time Units: Hours

Queue

Time

Waiting Time

	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Primary Server1.Queue	0.00000000	0.00	0.00	0.00000000	0.00	0.00047273
Primary Server2.Queue	0.00000001	0.00	0.00000001	0.00000002	0.00	0.00212423
Primary Server3.Queue	0.00000014	0.00	0.00000012	0.00000019	0.00	0.00349908
Primary Server4.Queue	0.00000018	0.00	0.00000016	0.00000022	0.00	0.00329301
Secondary Server1.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Secondary Server2.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Secondary Server3.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Secondary Server4.Queue	0.00	0.00	0.00	0.00	0.00	0.00

Other

Number Waiting

	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Primary Server1.Queue	0.00000001	0.00	0.00	0.00000005	0.00	1.0000
Primary Server2.Queue	0.00000062	0.00	0.00000031	0.00000128	0.00	2.0000
Primary Server3.Queue	0.00001229	0.00	0.00001003	0.00001571	0.00	3.0000
Primary Server4.Queue	0.00001652	0.00	0.00001414	0.00001942	0.00	3.0000
Secondary Server1.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Secondary Server2.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Secondary Server3.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Secondary Server4.Queue	0.00	0.00	0.00	0.00	0.00	0.00

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Auto

Replications: 10 Time Units: Hours

Resource

Usage

Instantaneous Utilization

	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Main Serv1	0.02572906	0.00	0.02565881	0.02584706	0.00	1.0000
Main Serv2	0.04286299	0.00	0.04275561	0.04294233	0.00	1.0000
Main Serv3	0.07079374	0.00	0.07063457	0.07102418	0.00	1.0000
Main Serv4	0.07504679	0.00	0.07493176	0.07516290	0.00	1.0000
Secondary Serv1	0.00000028	0.00	0.00	0.00000099	0.00	0.6000
Secondary Serv2	0.00000001	0.00	0.00	0.00000009	0.00	0.2000
Secondary Serv3	0.00000002	0.00	0.00	0.00000011	0.00	0.2000
Secondary Serv4	0.00000029	0.00	0.00	0.00000229	0.00	0.4000

Number Busy

	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Main Serv1	0.1286	0.00	0.1283	0.1292	0.00	5.0000
Main Serv2	0.2143	0.00	0.2138	0.2147	0.00	5.0000
Main Serv3	0.3540	0.00	0.3532	0.3551	0.00	5.0000
Main Serv4	0.3752	0.00	0.3747	0.3758	0.00	5.0000
Secondary Serv1	0.00000139	0.00	0.00	0.00000496	0.00	3.0000
Secondary Serv2	0.00000004	0.00	0.00	0.00000043	0.00	1.0000
Secondary Serv3	0.00000009	0.00	0.00	0.00000055	0.00	1.0000
Secondary Serv4	0.00000145	0.00	0.00	0.00001145	0.00	2.0000

Number Scheduled

	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Main Serv1	5.0000	0.00	5.0000	5.0000	5.0000	5.0000
Main Serv2	5.0000	0.00	5.0000	5.0000	5.0000	5.0000
Main Serv3	5.0000	0.00	5.0000	5.0000	5.0000	5.0000
Main Serv4	5.0000	0.00	5.0000	5.0000	5.0000	5.0000
Secondary Serv1	5.0000	0.00	5.0000	5.0000	5.0000	5.0000
Secondary Serv2	5.0000	0.00	5.0000	5.0000	5.0000	5.0000
Secondary Serv3	5.0000	0.00	5.0000	5.0000	5.0000	5.0000
Secondary Serv4	5.0000	0.00	5.0000	5.0000	5.0000	5.0000

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Values Across All Replications

Auto

Replications: 10 Time Units: Hours

Resource

Usage

Scheduled Utilization

	Average	Half Width	Minimum Average	Maximum Average
Main Serv1	0.02572906	0.00	0.02565881	0.02584706
Main Serv2	0.04286299	0.00	0.04275561	0.04294233
Main Serv3	0.07079374	0.00	0.07063457	0.07102418
Main Serv4	0.07504679	0.00	0.07493176	0.07516290
Secondary Serv1	0.00000028	0.00	0.00	0.00000099
Secondary Serv2	0.00000001	0.00	0.00	0.00000009
Secondary Serv3	0.00000002	0.00	0.00	0.00000011
Secondary Serv4	0.00000029	0.00	0.00	0.00000229



Total Number Seized

	Average	Half Width	Minimum Average	Maximum Average
Main Serv1	270463.50	445.40	269810.00	271738.00
Main Serv2	450577.50	435.58	449496.00	451478.00
Main Serv3	744192.20	867.23	742498.00	746686.00
Main Serv4	788912.00	573.99	787719.00	790412.00
Secondary Serv1	2.8000	2.74	0.00	10.0000
Secondary Serv2	0.1000	0.23	0.00	1.0000
Secondary Serv3	0.2000	0.30	0.00	1.0000
Secondary Serv4	3.2000	5.55	0.00	25.0000



This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use

The logo of King Mongkut's Institute of Technology Ladkrabang is a circular emblem. It features a central five-tiered stupa with a sunburst above it. The stupa is flanked by two smaller stupa-like structures, each with a flame-like top. The entire emblem is surrounded by ornate, symmetrical floral and scrollwork patterns. The text "King Mongkut's Institute of Technology Ladkrabang" is written in a circular path around the inner edge of the emblem.

Auto – Scaling with Server Virtualization Model Report

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Auto

Replications: 10

Time Units: Hours

Key Performance Indicators

System

Number Out

Average

2,262,349



This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use

Values Across All Replications

Auto

Replications: 10 Time Units: Hours

Entity

Time

VA Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Work1	0.00472181	0.00	0.00472006	0.00472342	0.00279335	0.01349617
Work2	0.00472221	0.00	0.00471942	0.00472433	0.00278450	0.01241133
Work3	0.00472173	0.00	0.00472038	0.00472354	0.00278356	0.01312927
Work4	0.00472266	0.00	0.00472109	0.00472416	0.00278167	0.01308116
NVA Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Work1	0.00	0.00	0.00	0.00	0.00	0.00
Work2	0.00	0.00	0.00	0.00	0.00	0.00
Work3	0.00	0.00	0.00	0.00	0.00	0.00
Work4	0.00	0.00	0.00	0.00	0.00	0.00
Wait Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Work1	0.00000011	0.00	0.00	0.00000093	0.00	0.1354
Work2	0.00000006	0.00	0.00	0.00000030	0.00	0.06755639
Work3	0.00000004	0.00	0.00000001	0.00000014	0.00	0.0967
Work4	0.00000010	0.00	0.00000002	0.00000026	0.00	0.1280
Transfer Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Work1	0.00	0.00	0.00	0.00	0.00	0.00
Work2	0.00	0.00	0.00	0.00	0.00	0.00
Work3	0.00	0.00	0.00	0.00	0.00	0.00
Work4	0.00	0.00	0.00	0.00	0.00	0.00
Other Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Work1	0.00	0.00	0.00	0.00	0.00	0.00
Work2	0.00	0.00	0.00	0.00	0.00	0.00
Work3	0.00	0.00	0.00	0.00	0.00	0.00
Work4	0.00	0.00	0.00	0.00	0.00	0.00
Total Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Work1	0.00472192	0.00	0.00472006	0.00472342	0.00279335	0.1406
Work2	0.00472227	0.00	0.00471942	0.00472433	0.00278450	0.07199176
Work3	0.00472177	0.00	0.00472038	0.00472368	0.00278356	0.1018
Work4	0.00472275	0.00	0.00472111	0.00472419	0.00278167	0.1318

Other

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Values Across All Replications

Auto

Replications: 10 Time Units: Hours

Entity

Other

Number In	Average	Half Width	Minimum Average	Maximum Average
customer	2262349.40	943.54	2260634.00	2265241.00
Work1	339419.30	404.88	338481.00	340307.00
Work2	226120.10	352.24	225351.00	226867.00
Work3	791691.00	518.98	790243.00	793007.00
Work4	905119.00	641.00	904002.00	906673.00



Number Out	Average	Half Width	Minimum Average	Maximum Average
customer	2262349.40	943.54	2260634.00	2265241.00
Work1	339419.20	404.86	338481.00	340307.00
Work2	226120.10	352.24	225351.00	226867.00
Work3	791690.90	519.14	790242.00	793007.00
Work4	905118.80	640.88	904002.00	906673.00

WIP	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
customer	0.00	0.00	0.00	0.00	0.00	1.0000
Work1	0.1830	0.00	0.1824	0.1834	0.00	6.0000
Work2	0.1219	0.00	0.1215	0.1223	0.00	6.0000
Work3	0.4267	0.00	0.4260	0.4274	0.00	10.0000
Work4	0.4880	0.00	0.4872	0.4889	0.00	9.0000

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Values Across All Replications

Auto

Replications: 10 Time Units: Hours

Process

Time per Entity

VA Time Per Entity	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Server1	0.00055500	0.00	0.00055402	0.00055625	0.00000000	0.00853069
Server2	0.00055575	0.00	0.00055266	0.00055780	0.00000000	0.00787371
Server3	0.00055530	0.00	0.00055450	0.00055620	0.00000000	0.00890090
Server4	0.00055553	0.00	0.00055453	0.00055608	0.00000000	0.00861593
Virtual Server 01	0.00416681	0.00	0.00416474	0.00416824	0.00277832	0.00555496
Virtual Server 02	0.00416646	0.00	0.00416547	0.00416779	0.00277829	0.00555464
Virtual Server 03	0.00416643	0.00	0.00416469	0.00416733	0.00277842	0.00555459
Virtual Server 04	0.00416713	0.00	0.00416590	0.00416828	0.00277827	0.00555462

Wait Time Per Entity	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Server1	0.00	0.00	0.00	0.00	0.00	0.00
Server2	0.00000000	0.00	0.00	0.00000003	0.00	0.00635799
Server3	0.00	0.00	0.00	0.00	0.00	0.00
Server4	0.00000001	0.00	0.00	0.00000010	0.00	0.08962156
Virtual Server 01	0.00000011	0.00	0.00	0.00000093	0.00	0.1354
Virtual Server 02	0.00000006	0.00	0.00	0.00000030	0.00	0.06755639
Virtual Server 03	0.00000004	0.00	0.00000001	0.00000014	0.00	0.0967
Virtual Server 04	0.00000009	0.00	0.00000002	0.00000024	0.00	0.1280

Total Time Per Entity	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Server1	0.00055500	0.00	0.00055402	0.00055625	0.00000000	0.00853069
Server2	0.00055575	0.00	0.00055266	0.00055780	0.00000000	0.00787371
Server3	0.00055530	0.00	0.00055450	0.00055620	0.00000000	0.00890090
Server4	0.00055554	0.00	0.00055463	0.00055608	0.00000000	0.0911
Virtual Server 01	0.00416692	0.00	0.00416474	0.00416824	0.00277832	0.1406
Virtual Server 02	0.00416652	0.00	0.00416554	0.00416779	0.00277829	0.07182236
Virtual Server 03	0.00416647	0.00	0.00416470	0.00416747	0.00277842	0.1011
Virtual Server 04	0.00416721	0.00	0.00416593	0.00416831	0.00277827	0.1314

Accumulated Time

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Auto

Replications: 10 Time Units: Hours

Process

Accumulated Time

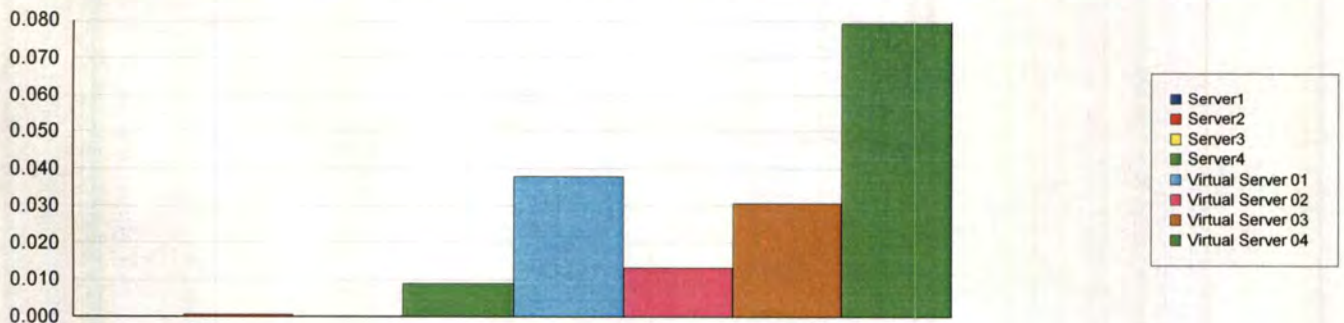
Accum VA Time

	Average	Half Width	Minimum Average	Maximum Average
Server1	188.39	0.26	187.88	188.96
Server2	125.67	0.31	124.93	126.21
Server3	439.61	0.37	438.92	440.51
Server4	502.83	0.57	501.76	504.05
Virtual Server 01	1414.34	1.92	1409.71	1418.41
Virtual Server 02	942.12	1.38	939.16	945.02
Virtual Server 03	3298.40	2.07	3292.74	3304.06
Virtual Server 04	3771.82	2.74	3766.41	3778.53



Accum Wait Time

	Average	Half Width	Minimum Average	Maximum Average
Server1	0.00	0.00	0.00	0.00
Server2	0.00063580	0.00	0.00	0.00635799
Server3	0.00	0.00	0.00	0.00
Server4	0.00896216	0.02	0.00	0.08962156
Virtual Server 01	0.03780313	0.07	0.00	0.3143
Virtual Server 02	0.01316702	0.02	0.00	0.06755639
Virtual Server 03	0.03066838	0.03	0.00424635	0.1113
Virtual Server 04	0.07924731	0.05	0.01909771	0.2204



Other

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Auto

Replications: 10 Time Units: Hours

Process

Other

Number In	Average	Half Width	Minimum Average	Maximum Average
Server1	339430.70	415.79	338487.00	340382.00
Server2	226120.80	352.06	225351.00	226867.00
Server3	791660.30	515.81	790243.00	792991.00
Server4	905137.60	624.65	904104.00	906660.00
Virtual Server 01	339430.70	415.79	338487.00	340382.00
Virtual Server 02	226120.80	352.06	225351.00	226867.00
Virtual Server 03	791660.30	515.81	790243.00	792991.00
Virtual Server 04	905137.60	624.65	904104.00	906660.00



Number Out	Average	Half Width	Minimum Average	Maximum Average
Server1	339430.70	415.79	338487.00	340382.00
Server2	226120.80	352.06	225351.00	226867.00
Server3	791660.30	515.81	790243.00	792991.00
Server4	905137.60	624.65	904104.00	906660.00
Virtual Server 01	339430.60	415.76	338487.00	340382.00
Virtual Server 02	226120.80	352.06	225351.00	226867.00
Virtual Server 03	791660.20	515.96	790242.00	792991.00
Virtual Server 04	905137.40	624.54	904104.00	906660.00

This material is reserved for educational use only, not allowed for commercial use.

Values Across All Replications

Auto

Replications: 10 Time Units: Hours

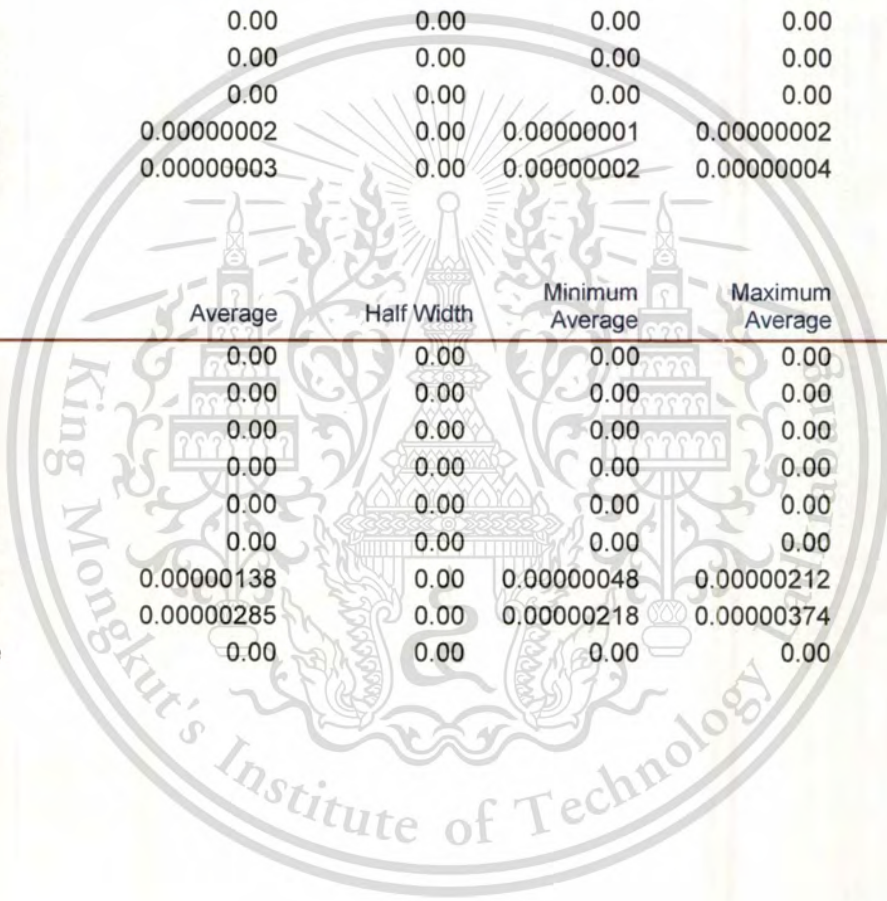
Queue

Time

Waiting Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Server1.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Server2.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Server3.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Server4.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Virtual Server 01.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Virtual Server 02.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Virtual Server 03.Queue	0.00000002	0.00	0.00000001	0.00000002	0.00	0.00255620
Virtual Server 04.Queue	0.00000003	0.00	0.00000002	0.00000004	0.00	0.00271035

Other

Number Waiting	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Server1.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Server2.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Server3.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Server4.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Virtual Server 01.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Virtual Server 02.Queue	0.00	0.00	0.00	0.00	0.00	0.00
Virtual Server 03.Queue	0.00000138	0.00	0.00000048	0.00000212	0.00	3.0000
Virtual Server 04.Queue	0.00000285	0.00	0.00000218	0.00000374	0.00	3.0000
Wait for Available.Queue	0.00	0.00	0.00	0.00	0.00	0.00



Values Across All Replications

Auto

Replications: 10 Time Units: Hours

Resource

Usage

Instantaneous Utilization	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Main Serv1	0.00215052	0.00	0.00214472	0.00215713	0.00	0.4000
Main Serv2	0.00143455	0.00	0.00142617	0.00144073	0.00	0.4000
Main Serv3	0.00501837	0.00	0.00501047	0.00502865	0.00	0.5000
Main Serv4	0.00574008	0.00	0.00572780	0.00575402	0.00	0.5000
VM1_1	0.1341	0.00	0.1338	0.1344	0.00	1.0000
VM1_2	0.02433952	0.00	0.02406433	0.02455147	0.00	1.0000
VM1_3	0.00281652	0.00	0.00275437	0.00285607	0.00	1.0000
VM1_4	0.00023108	0.00	0.00021664	0.00023913	0.00	1.0000
VM1_5	0.00001547	0.00	0.00001173	0.00001884	0.00	1.0000
VM1_6	0.00000034	0.00	0.00	0.00000103	0.00	1.0000
VM2_1	0.0946	0.00	0.0943	0.0948	0.00	1.0000
VM2_2	0.01197880	0.00	0.01189368	0.01206401	0.00	1.0000
VM2_3	0.00092120	0.00	0.00090396	0.00093768	0.00	1.0000
VM2_4	0.00004970	0.00	0.00004055	0.00006153	0.00	1.0000
VM2_5	0.00000208	0.00	0.00000037	0.00000429	0.00	1.0000
VM2_6	0.00000004	0.00	0.00	0.00000043	0.00	1.0000
VM3_1	0.2568	0.00	0.2565	0.2571	0.00	1.0000
VM3_2	0.0910	0.00	0.0908	0.0913	0.00	1.0000
VM3_3	0.02352042	0.00	0.02344425	0.02359367	0.00	1.0000
VM3_4	0.00448450	0.00	0.00441719	0.00454978	0.00	1.0000
VM3_5	0.00065782	0.00	0.00063101	0.00067805	0.00	1.0000
VM3_6	0.00007988	0.00	0.00007071	0.00008809	0.00	1.0000
VM4_1	0.2812	0.00	0.2810	0.2817	0.00	1.0000
VM4_2	0.1093	0.00	0.1090	0.1097	0.00	1.0000
VM4_3	0.03180630	0.00	0.03162753	0.03199612	0.00	1.0000
VM4_4	0.00693568	0.00	0.00686752	0.00700545	0.00	1.0000
VM4_5	0.00115345	0.00	0.00110728	0.00121603	0.00	1.0000
VM4_6	0.00015427	0.00	0.00014446	0.00017244	0.00	1.0000

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Values Across All Replications

Auto

Replications: 10 Time Units: Hours

Resource

Usage

Number Busy	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Main Serv1	0.02150520	0.00	0.02144720	0.02157132	0.00	4.0000
Main Serv2	0.01434555	0.00	0.01426174	0.01440732	0.00	4.0000
Main Serv3	0.05018375	0.00	0.05010465	0.05028653	0.00	5.0000
Main Serv4	0.05740083	0.00	0.05727803	0.05754016	0.00	5.0000
VM1_1	0.1341	0.00	0.1338	0.1344	0.00	1.0000
VM1_2	0.02433952	0.00	0.02406433	0.02455147	0.00	1.0000
VM1_3	0.00281652	0.00	0.00275437	0.00285607	0.00	1.0000
VM1_4	0.00023108	0.00	0.00021664	0.00023913	0.00	1.0000
VM1_5	0.00001547	0.00	0.00001173	0.00001884	0.00	1.0000
VM1_6	0.00000034	0.00	0.00	0.00000103	0.00	1.0000
VM2_1	0.0946	0.00	0.0943	0.0948	0.00	1.0000
VM2_2	0.01197880	0.00	0.01189368	0.01206401	0.00	1.0000
VM2_3	0.00092120	0.00	0.00090396	0.00093768	0.00	1.0000
VM2_4	0.00004970	0.00	0.00004055	0.00006153	0.00	1.0000
VM2_5	0.00000208	0.00	0.00000037	0.00000429	0.00	1.0000
VM2_6	0.00000004	0.00	0.00	0.00000043	0.00	1.0000
VM3_1	0.2568	0.00	0.2565	0.2571	0.00	1.0000
VM3_2	0.0910	0.00	0.0908	0.0913	0.00	1.0000
VM3_3	0.02352042	0.00	0.02344425	0.02359367	0.00	1.0000
VM3_4	0.00448450	0.00	0.00441719	0.00454978	0.00	1.0000
VM3_5	0.00065782	0.00	0.00063101	0.00067805	0.00	1.0000
VM3_6	0.00007988	0.00	0.00007071	0.00008809	0.00	1.0000
VM4_1	0.2812	0.00	0.2810	0.2817	0.00	1.0000
VM4_2	0.1093	0.00	0.1090	0.1097	0.00	1.0000
VM4_3	0.03180630	0.00	0.03162753	0.03199612	0.00	1.0000
VM4_4	0.00693568	0.00	0.00686752	0.00700545	0.00	1.0000
VM4_5	0.00115345	0.00	0.00110728	0.00121603	0.00	1.0000
VM4_6	0.00015427	0.00	0.00014446	0.00017244	0.00	1.0000

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Values Across All Replications

Auto

Replications: 10 Time Units: Hours

Resource

Usage

Number Scheduled	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Main Serv1	10.0000	0.00	10.0000	10.0000	10.0000	10.0000
Main Serv2	10.0000	0.00	10.0000	10.0000	10.0000	10.0000
Main Serv3	10.0000	0.00	10.0000	10.0000	10.0000	10.0000
Main Serv4	10.0000	0.00	10.0000	10.0000	10.0000	10.0000
VM1_1	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM1_2	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM1_3	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM1_4	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM1_5	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM1_6	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM2_1	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM2_2	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM2_3	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM2_4	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM2_5	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM2_6	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM3_1	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM3_2	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM3_3	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM3_4	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM3_5	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM3_6	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM4_1	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM4_2	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM4_3	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM4_4	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM4_5	1.0000	0.00	1.0000	1.0000	1.0000	1.0000
VM4_6	1.0000	0.00	1.0000	1.0000	1.0000	1.0000

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Values Across All Replications

Auto

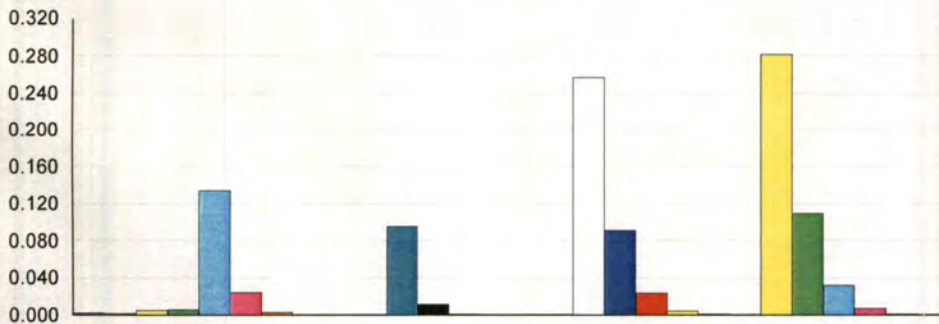
Replications: 10 Time Units: Hours

Resource

Usage

Scheduled Utilization

	Average	Half Width	Minimum Average	Maximum Average
Main Serv1	0.00215052	0.00	0.00214472	0.00215713
Main Serv2	0.00143455	0.00	0.00142617	0.00144073
Main Serv3	0.00501837	0.00	0.00501047	0.00502865
Main Serv4	0.00574008	0.00	0.00572780	0.00575402
VM1_1	0.1341	0.00	0.1338	0.1344
VM1_2	0.02433952	0.00	0.02406433	0.02455147
VM1_3	0.00281652	0.00	0.00275437	0.00285607
VM1_4	0.00023108	0.00	0.00021664	0.00023913
VM1_5	0.00001547	0.00	0.00001173	0.00001884
VM1_6	0.00000034	0.00	0.00	0.00000103
VM2_1	0.0946	0.00	0.0943	0.0948
VM2_2	0.01197880	0.00	0.01189368	0.01206401
VM2_3	0.00092120	0.00	0.00090396	0.00093768
VM2_4	0.00004970	0.00	0.00004055	0.00006153
VM2_5	0.00000208	0.00	0.00000037	0.00000429
VM2_6	0.00000004	0.00	0.00	0.00000043
VM3_1	0.2568	0.00	0.2565	0.2571
VM3_2	0.0910	0.00	0.0908	0.0913
VM3_3	0.02352042	0.00	0.02344425	0.02359367
VM3_4	0.00448450	0.00	0.00441719	0.00454978
VM3_5	0.00065782	0.00	0.00063101	0.00067805
VM3_6	0.00007988	0.00	0.00007071	0.00008809
VM4_1	0.2812	0.00	0.2810	0.2817
VM4_2	0.1093	0.00	0.1090	0.1097
VM4_3	0.03180630	0.00	0.03162753	0.03199612
VM4_4	0.00693568	0.00	0.00686752	0.00700545
VM4_5	0.00115345	0.00	0.00110728	0.00121603
VM4_6	0.00015427	0.00	0.00014446	0.00017244



This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

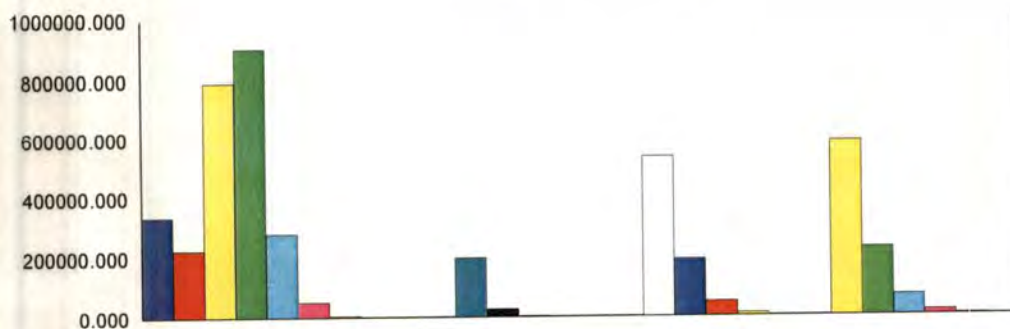
Auto

Replications: 10 Time Units: Hours

Resource

Usage

Total Number Seized	Average	Half Width	Minimum Average	Maximum Average
Main Serv1	339430.70	415.79	338487.00	340382.00
Main Serv2	226120.80	352.06	225351.00	226867.00
Main Serv3	791660.30	515.81	790243.00	792991.00
Main Serv4	905137.60	624.65	904104.00	906660.00
VM1_1	281814.70	264.39	281420.00	282426.00
VM1_2	51175.10	218.49	50639.00	51609.00
VM1_3	5922.10	49.81	5807.00	6007.00
VM1_4	485.70	11.65	454.00	505.00
VM1_5	32.4000	3.63	25.0000	39.0000
VM1_6	0.7000	0.59	0.00	2.0000
VM2_1	198883.40	268.02	198318.00	199452.00
VM2_2	25192.60	104.49	24970.00	25394.00
VM2_3	1936.20	21.05	1898.00	1972.00
VM2_4	104.00	10.17	86.0000	128.00
VM2_5	4.5000	1.89	1.0000	9.0000
VM2_6	0.1000	0.23	0.00	1.0000
VM3_1	539894.10	278.54	539140.00	540450.00
VM3_2	191326.50	222.52	190946.00	191961.00
VM3_3	49460.60	78.94	49306.00	49623.00
VM3_4	9430.00	67.50	9303.00	9561.00
VM3_5	1381.10	25.50	1321.00	1430.00
VM3_6	168.00	8.46	150.00	185.00
VM4_1	591172.80	273.31	590751.00	591994.00
VM4_2	229767.60	314.57	229262.00	230686.00
VM4_3	66856.80	175.19	66445.00	67303.00
VM4_4	14589.40	64.06	14470.00	14708.00
VM4_5	2426.70	48.71	2329.00	2559.00
VM4_6	324.30	11.28	302.00	360.00



This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.