

Hardware Sizing Method for Information System in Korea

Jonghei Ra¹, Haeyong Jung² and Kwangdon Choi³

¹Department of e-Business, Gwangju University
e-mail : jhra@gwangju.ac.kr

²Department of MIS, Nazarene University
e-mail : hyjung@kornu.ac.kr

³Department of e-Business, Hansei University
e-mail : kdchoi@hansei.ac.kr

Abstract

Accurate system sizing are essential for higher efficiency of investment. Accurate system sizing benefits are generally viewed in terms of the avoidance of excess equipment and lost opportunity costs by not being able to support business needs. We are proposed calculating method for hardware components that is CPU, memory and system, data disk according to the application system types

Keyword: Hardware Sizing, System Performance, Information System

1. Introduction

The introduction of information system in public sector and private sector regardless of its organization characteristic is recognized as necessary infrastructure. The investment on information system is continuously expanding in the public sector for the improvement of productivity and the level of service in Korea. In addition, rapid transitions are taking place in the public sector from mainframe to client/server, Internet, Intranet, which resulted in putting more importance on the management of system performance and capacity management due to the complexity and increasing usage of the system. In other words, it is very important to calculate the required hardware resources accurately prior to the introduction of the system, because failure of managing the system performance and capacity sizing can result in high cost and resources waste and lower the productivity and cause distrust and credibility breakdown due to the poor service [1,2, 6,13].

According to report of Korean government, the level of resource utilization is very low, especially utilization of CPU is 46% on average, which stems from the lack of accurate resource capacity sizing before introducing information system. However, it is not easy to calculate the appropriate resource capacity sizing, because the resource capacity of information system should be determined based on business characteristic, estimated workload increasing rate, level of system usage of end users, technical architecture [5, 9, 10].

In general, resource capacity sizing of information system has been conducted by system suppliers or internal resources based on the non-standardized methodology, even though the cost of hardware generally takes up 30%~50% of total project cost. Particularly there is no standardized methodology on resource capacity sizing and CPU

performance evaluation, which is controversial when the public organizations consider building their own information system. Also the price of server varies depending on the CPU spec. Therefore, it is necessary to establish a methodology for required resources capacity for each project of the public organization, just like there is standard for software quotation for the public organization project [7, 8].

We proposed sizing methodology based on the empirical analysis. The empirical analysis has been conducted based on the survey by working level experts in the public organizations and debate with groups of experts using existing capacity sizing framework and hardware capacity sizing guidelines.

This paper is organized as follows. Section 2 introduces related works, Section 3 presents the research design, Section 4 provides hardware sizing method, and Section 5 concludes the paper with some future works.

2. Related works

We normally use different terminology for the identical implication of hardware resource capacity calculation, such as Capacity Planning, Capacity Sizing, and System Sizing, and there are few differences among them.

System capacity sizing determines system requirements such as CPU type or number, Disk type or volume and memory volume based on the concepts defined by organizations such as TPC(Transaction Processing Performance Council)[12], SPEC(Standard Performance Evaluation Corporation)[11] and IDEAs.

In other words, the system capacity sizing uses mathematical methodology based on business process and applications, which is different from the one which decides the required capacity sizing decided by system architecture and application.

Hardware sizing is studied when software vendor notify the most suitable hardware size for their package [9]. But there is no hardware sizing method to use for organization ordering the constructing IS, constituted with hardware and software.

Many research for software sizing, such as LOC(Line Of Code), FP(Function Point), etc. are exist, but research for hardware sizing is scarcely. The existing structured studies for capacity sizing include 『developing tool for capacity sizing(1994)』, 『study for hardware capacity sizing(2002)』, 『capacity sizing technique and framework study for information system(2003)』, 『The capacity sizing study by information system size(2004)』 which studies were done by working level researchers in the National Computerization Agency of Korea. However, this study has the limitation as a general guideline due to the lack of credibility and objectivity [7].

New performance metric is used with tpmC (transactions per minute) and OPS(Operations Per Second)[3,4], But for example of CPU, the most important element of hardware sizing, tpmC, most used metric have three problems. First, there is the discrepancy of CPU capacity sizing criteria. Second, some hardware suppliers does not apply TPC performance standard. Third, it is not reflect the web based software characteristics.

As a result of reviewing all the case studies and former studies, it is necessary to uphold objectivity and develop a concrete capacity sizing, and to differentiate the calculating methods based on business type and size.

3. Research design

This study had empirical analysis with group of experts in the public sector and suppliers in order to obtain and generalize the appropriateness of capacity sizing method. In addition, this study established the criteria for analysis and interpretation through debate with group of experts and validated objectivity of method and items, which should be included in the capacity sizing formula.

The detail hardware elements in this study are made up based on the reference of information from NCA, public organizations, System Integrator's as there was no capacity calculation formula in former research and standardized capacity formula. The detailed hardware capacity formula and elements for it are like Table 1.

Hardware was major items to be validated for appropriateness of method, including CPU, memory, disk, and the objectivity validations has been done through survey by experts from operational level people and capacity sizing experts from the public organizations.

Table 1. Items of Hardware Capacity Calculation and Survey

Items		Definition	A	B	C
CPU for WEB/ WAS server	Number of concurrent users	Number of concurrent users	o	o	o
	Application interface load correction	Load rate on interface communicating with another server	o	x	o
	Peak time load correction	Load rate considering peak time load upon significantly increasing abnormal accesses		x	x
	System redundancy	Redundancy for stable operation based on importance and urgency of business	o	o	o
	Number of operations per user	Number of operations per minute per person	o	o	o
CPU for OLTP	Number of concurrent users	Number of concurrent users	o	o	o
	Number of transactions	Number of transactions per minute per person	o	o	o
	Default tpmC correction	Correction based on system size to run system under optimized status	o	x	o
	Peak time load correction	Correction considering peak time of load	o	o	o
	Database size correction	Correction for the data size of transaction to be processed	o	x	o
	Application complexity correction	Correction for the complexity of program	o	o	x
	Application architecture correction	Correction for application architecture	x	o	x
	Application load correction	Actual user operation environmental correction	x	o	x
	Network correction	Correction for network bandwidth	o	o	x
	Cluster correction	Correction for outage under cluster environment	o	x	o
Redundancy correction	Redundancy for unexpected situation and expansion	o	o	o	
Memory	System area	Required memory space for OS, DBMS, Middleware, other utilities	o	o	o

	System managed area	Required memory space for system operation	x	x	o
	Number of users	Number of users using system	o	o	o
	Required memory per user	Required memory space for application, middleware, DBMS	o	o	o
	Buffer cash	Buffer cash size to reduce Disk I/O	o	x	x
	Cluster Correction	Redundancy considering increasing load for one system when the other system is down which is clustered together	o	x	x
	Redundancy correction	Redundancy for unexpected situation and expansion	o	o	o
System Disk	System O/S area	Disk space for O/S, system software	o	o	o
	Application S/W area	Disk space for middleware, DBMS, application S/W, various utilities	o	o	o
	SWAP area	Disk space for Swapping	o	o	o
	System Disk redundancy	Disk space for safe data management	o	o	o
Data Disk	Data area	Disk space for data store	o	o	o
	Backup area	Disk space for copied file and database in case of hardware outage	o	x	o
	RAID area	Parity space for RAID disk	o	o	o
	Data disk redundancy	Disk space for safe data management	o	o	o

The measurement for capacity sizing objectivity had been conducted based on a scale of one to five by questionnaire. The pilot testing for survey was conducted by 4 experts in public organizations, 2 experts from system suppliers. The capacity calculation formula and importance of formula elements have been finalized based on the assessment criteria after debating among groups of experts.

A population is composed of 2 groups including working level experts who have implemented and were operating the IS, and capacity calculation experts from suppliers. The working level people is limited to the public organizations in which they had rolled out and were running the IS, including government organizations founded by government organization law, local government, nonprofit government invested organizations and a government enterprises.

As for suppliers experts survey we limited to the major 9 suppliers, which supplied servers to the public organizations.

The distributed survey has returned across 5 areas from 61 participants from 29 government organizations and the 60 survey feedbacks had been used for survey analysis except one participant who has not yet had the IS in operation. As for suppliers experts survey, 44 survey feedbacks from 45 participants had been used for survey analysis. Those 45 participants consist of 5 from each supplier and SI enterprise lest we should overemphasize a specific enterprise.

As for working experience, experts from the public organizations had 12.8 years of working experience on average, and been working for 4.38 years on average in their current job position. Experts from the supplier organizations had 9.8 years of working experience on average, and been working for 5.15 years on average in their current job position.

The information systems for which 60 participants from government organization had implemented and were operating included internal use application which were 35

system taking up 58.3% and included the public service system which were 24 systems taking up 40%.

With classification by system area, unit applications are 33 taking up 55%, and enterprise application were 27 taking up 45%. Table. 2 show the classification based on purpose of information system and information system building area. In addition, the average age of server currently being used by the participants was 2.3 years and 9 various servers being used models are shown in Table 3.

Table 2. Classification by objective of information system and scope of system building

Classification		System Area		Total
		Unit App.	IT Infra.	
Objective of system	Improve efficiency of business process	23 (38.3%)	12 (20%)	35 (58.3%)
	Original service	9 (15%)	15 (25%)	24 (40%)
	Both	1 (1.7%)	-	1 (1.7%)
Total		33 (55%)	27 (45%)	60

Table 3. Distribution of Server model from the survey participants of the public organization

H/W vendor	Frequency	Rate
SUN Microsystems	20	33%
IBM	13	22%
HP	12	20%
Fjuzitu	6	10%
etc	9	15%
Total	60	100%

4. Hardware sizing method

4.1 Analysis on appropriateness and credibility

The appropriateness validation is to review whether measuring tool measures the intended concept appropriately. The validity can be classified abused on the assessment method as content validity, criterion-related validity, constructs validity. This study rather generates the formula which is used to compute objectively the most optimized system capacity for CPU, Memory, Disk for information system building than is research which defines and measures the concept. Thus, conceptual validity is no longer valid as this study is intended to validate the validity of formula and formula elements shown in capacity sizing guide line written in 2003. This study assume that contents validity and sampling validity of this research has been confirmed as 2003's research created the capacity calculation formula based on experience of capacity sizing for the public organization information system building and based on the debate among groups of experts. The credibility analysis results performed by this study are over 0.6 of Cronbach's α value shown in Table 4. From the personal analysis point of view this score is not high enough, but it does not pose a problem, as it is not a conceptual

research. In addition, original analysis item is maintained as there is no big difference on value variation of Cronbach's α when comparing to without individual element.

Table 4. Analysis Result of Credibility

Classification	Number of survey questionnaire	Cronbach's α value
CPU for WEB/WAS	5 items	0.7267
CPU for OLTP server	12 items	0.7987
Memory	7 items	0.6116
System Disk	4 items	0.6992
Data Disk	4 items	0.7498

4.2 Calculating formula

We use all items in Table 1 as initial elements of calculating formula, and then validate it by questionnaire. As criteria for importance of formula elements, first we excluded the elements from the formula which are below 3.5 on average from both the public organizations and suppliers, and second, even though average is 3.5 and above we reviewed to see if there is big difference between the public organizations' and suppliers'. If big difference exists then the elements were removed from the formula and if not, the elements were included in the formula. Sizing elements of hardware are chosen by the importance analysis described above. For example, WEB/WAS CPU has 5, OLTP CPU has 10, Memory has 4 and Data disk has 4 sizing elements. We conducted difference analysis between two groups, experts from public organizations and experts from suppliers, to ensure whether there are differences in recognizing the validity of calculating formula and the importance of each element between them.

Table 5. Comparison between two groups of experts from the public organization and system supplier for importance of each element of capacity calculation formula of CPU for WEB/WAS server

Items	Classification	# of data	Total avg.	Avg. by group	Standard deviation	t-value	Variance validation																																																				
Appropriateness of formula	Public Organization experts	45	3.70	3.82	.58	1.909	.060*																																																				
	Suppliers' experts	37		3.54	.73			Importance of element	Number of simultaneous users	Public Organization experts	4.73	4.64	.53	1.884	.063*	Suppliers' experts	37	4.84	.44	Operations per user	Public Organization experts	4.40	3.81	.85	2.295	.024**	Suppliers' experts	37	4.30	1.10	Application load correction	Public Organization experts	4.19	3.55	.90	1.077	.285	Suppliers' experts	37	3.78	1.06	Peak time load correction	Public Organization experts	4.02	4.36	.74	-1.520	.605	Suppliers' experts	37	4.46	.99	System reserve rate	Public Organization experts	4.7	3.65	.70	-3.312	.756
Importance of element	Number of simultaneous users	Public Organization experts	4.73	4.64	.53	1.884	.063*																																																				
		Suppliers' experts		37	4.84				.44	Operations per user	Public Organization experts	4.40	3.81	.85	2.295	.024**	Suppliers' experts	37	4.30	1.10	Application load correction	Public Organization experts	4.19	3.55	.90	1.077	.285	Suppliers' experts	37	3.78	1.06	Peak time load correction	Public Organization experts	4.02	4.36	.74	-1.520	.605	Suppliers' experts	37	4.46	.99	System reserve rate	Public Organization experts	4.7	3.65	.70	-3.312	.756	Suppliers' experts	37	4.22	.63						
	Operations per user	Public Organization experts	4.40	3.81	.85	2.295	.024**																																																				
		Suppliers' experts		37	4.30				1.10	Application load correction	Public Organization experts	4.19	3.55	.90	1.077	.285	Suppliers' experts	37	3.78	1.06	Peak time load correction	Public Organization experts	4.02	4.36	.74	-1.520	.605	Suppliers' experts	37	4.46	.99	System reserve rate	Public Organization experts	4.7	3.65	.70	-3.312	.756	Suppliers' experts	37	4.22	.63																	
	Application load correction	Public Organization experts	4.19	3.55	.90	1.077	.285																																																				
		Suppliers' experts		37	3.78				1.06	Peak time load correction	Public Organization experts	4.02	4.36	.74	-1.520	.605	Suppliers' experts	37	4.46	.99	System reserve rate	Public Organization experts	4.7	3.65	.70	-3.312	.756	Suppliers' experts	37	4.22	.63																												
	Peak time load correction	Public Organization experts	4.02	4.36	.74	-1.520	.605																																																				
		Suppliers' experts		37	4.46				.99	System reserve rate	Public Organization experts	4.7	3.65	.70	-3.312	.756	Suppliers' experts	37	4.22	.63																																							
	System reserve rate	Public Organization experts	4.7	3.65	.70	-3.312	.756																																																				
		Suppliers' experts		37	4.22			.63																																																			

Note) * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$

For example, CPU capacity sizing for WEB/WAS is Table 3. According to the survey analysis, appropriateness of formula is 3.70 on average, which exceeds standard of 3.5 established by expert groups and the formula turned out to be appropriate. On the

other hand, importance of 5 elements of formula for CPU capacity calculation is above 3.5 on average. According to the results of analysis reviewing any difference on the recognition for formula elements, it turned out that there is a slight difference on simultaneous number of users, number of operations per users between these two groups. However, as the elements importance is above 3.5 from these two those 5 elements are included in the formula. Below is the CPU capacity calculation formula for WEB/WAS server based on the survey result.

$$\text{CPU Sizing} = \text{Number of simultaneous users} * (\text{application load correction} + \text{peak time load correction}) * \text{system reserve rate} * \text{Number of operations per user}$$

According to the CPU sizing for OLTP survey result analysis, the average of survey is 3.79, which exceeds standard of 3.5 established by group of experts, and it turned out to be appropriate. Importance of individual elements being used for CPU capacity calculation formula for OLTP is 3.5 and above on average for all 12 elements, such as number of processed transactions, number of simultaneous users, peak time correction, and reserve rate correction.

According to the analysis seeing if there were any differences in 6 elements between the public organizations and system suppliers, including number of simultaneous users, number of processed transactions, user complexity correction, application architecture correction, it turned out that there is quite a amount of difference. Below is the CPU capacity calculation formula for OLTP server based on the survey result.

$$\text{CPU Sizing} = \{(\text{Number of simultaneous users} * \text{Number of transactions processed}) * (\text{default tpmC correction} + \text{Peak Time correction} + \text{database size correction} + \text{application complexity correction} + \text{application load correction} + \text{network correction} + \text{cluster correction})\} * \text{reserve rate}$$

According to the analysis of memory sizing survey results, the average of survey is 3.70, which exceed the standard of 3.5 established by group of experts, and it turned out to be appropriate.

Importance of 7 elements, such as required memory per user, OS managed area, reserve rate, buffer cash, cluster correction which is used for memory calculation formula is above 3.5 in 5 elements, but importance of O/S managed area is 3.44 which is below standard of 3.5. Thus, the element of O/S managed area is excluded from the formula.

Table 6. Comparison between two groups of experts from the public organization and system supplier for importance of each element of capacity calculation formula of CPU for OLTP server

Items	Classification	# of data	Total avg.	Avg. by group	Standard deviation	t-value	Variance validation																																																																																																																																													
Formula appropriateness	A	44	3.79	3.73	.62	-.801	.425																																																																																																																																													
	B	41		3.85	.82			Importance of element	Number of simultaneous users	Public Organizations	46	4.47	4.28	1.07	-2.175	.039**	Supplies	41	4.68	.61	Number of transactions	Public Organizations	46	4.48	4.26	.93	-2.984	.004***	Supplies	41	4.73	.50	Default tpmC correction	Public Organizations	43	3.90	3.86	1.01	-.453	.652	Supplies	41	3.95	.80	User complexity correction	Public Organizations	44	3.69	3.43	1.00	-2.732	.008***	Supplies	41	3.98	.82	Peak-time correction	Public Organizations	44	4.35	4.27	.76	-.867	.389	Supplies	41	4.44	1.00	Database size correction	Public Organizations	44	3.88	3.82	.87	-.707	.482	Supplies	41	3.95	.86	Application complexity correction	Public Organizations	44	3.94	3.82	.79	-1.494	.139	Supplies	41	4.07	.79	Application structure correction	Public Organizations	44	3.55	3.30	1.02	-2.816	.006***	Supplies	41	3.83	.70	Application load correction	Public Organizations	44	3.86	3.70	.98	-1.820	.073*	Supplies	41	4.02	.61	Network correction	Public Organizations	44	3.64	3.55	1.02	-.982	.329	Supplies	41	3.73	.71	Cluster correction	Public Organizations	44	3.80	3.61	1.02	-1.831	.071*	Supplies	41	4.00	.92	Redundancy	Public Organizations	44	4.12	3.93	.93	-2.167	.033**
Importance of element	Number of simultaneous users	Public Organizations	46	4.47	4.28	1.07	-2.175			.039**																																																																																																																																										
		Supplies	41	4.68	.61	Number of transactions			Public Organizations		46	4.48	4.26	.93	-2.984	.004***	Supplies	41	4.73	.50	Default tpmC correction	Public Organizations	43	3.90	3.86	1.01	-.453	.652	Supplies	41	3.95	.80	User complexity correction	Public Organizations	44	3.69	3.43	1.00	-2.732	.008***	Supplies	41	3.98	.82	Peak-time correction	Public Organizations	44	4.35	4.27	.76	-.867	.389	Supplies	41	4.44	1.00	Database size correction	Public Organizations	44	3.88	3.82	.87	-.707	.482	Supplies	41	3.95	.86	Application complexity correction	Public Organizations	44	3.94	3.82	.79	-1.494	.139	Supplies	41	4.07	.79	Application structure correction	Public Organizations	44	3.55	3.30	1.02	-2.816	.006***	Supplies	41	3.83	.70	Application load correction	Public Organizations	44	3.86	3.70	.98	-1.820	.073*	Supplies	41	4.02	.61	Network correction	Public Organizations	44	3.64	3.55	1.02	-.982	.329	Supplies	41	3.73	.71	Cluster correction	Public Organizations	44	3.80	3.61	1.02	-1.831	.071*	Supplies	41	4.00	.92	Redundancy	Public Organizations	44	4.12	3.93	.93	-2.167	.033**	Supplies	41	4.32	.69								
	Number of transactions	Public Organizations	46	4.48	4.26		.93		-2.984	.004***																																																																																																																																										
		Supplies	41	4.73	.50	Default tpmC correction	Public Organizations				43	3.90	3.86	1.01	-.453	.652	Supplies	41	3.95	.80	User complexity correction	Public Organizations	44	3.69	3.43	1.00	-2.732	.008***	Supplies	41	3.98	.82	Peak-time correction	Public Organizations	44	4.35	4.27	.76	-.867	.389	Supplies	41	4.44	1.00	Database size correction	Public Organizations	44	3.88	3.82	.87	-.707	.482	Supplies	41	3.95	.86	Application complexity correction	Public Organizations	44	3.94	3.82	.79	-1.494	.139	Supplies	41	4.07	.79	Application structure correction	Public Organizations	44	3.55	3.30	1.02	-2.816	.006***	Supplies	41	3.83	.70	Application load correction	Public Organizations	44	3.86	3.70	.98	-1.820	.073*	Supplies	41	4.02	.61	Network correction	Public Organizations	44	3.64	3.55	1.02	-.982	.329	Supplies	41	3.73	.71	Cluster correction	Public Organizations	44	3.80	3.61	1.02	-1.831	.071*	Supplies	41	4.00	.92	Redundancy	Public Organizations	44	4.12	3.93	.93	-2.167	.033**	Supplies	41	4.32	.69																				
	Default tpmC correction	Public Organizations	43	3.90	3.86		1.01		-.453	.652																																																																																																																																										
		Supplies	41	3.95	.80	User complexity correction	Public Organizations				44	3.69	3.43	1.00	-2.732	.008***	Supplies	41	3.98	.82	Peak-time correction	Public Organizations	44	4.35	4.27	.76	-.867	.389	Supplies	41	4.44	1.00	Database size correction	Public Organizations	44	3.88	3.82	.87	-.707	.482	Supplies	41	3.95	.86	Application complexity correction	Public Organizations	44	3.94	3.82	.79	-1.494	.139	Supplies	41	4.07	.79	Application structure correction	Public Organizations	44	3.55	3.30	1.02	-2.816	.006***	Supplies	41	3.83	.70	Application load correction	Public Organizations	44	3.86	3.70	.98	-1.820	.073*	Supplies	41	4.02	.61	Network correction	Public Organizations	44	3.64	3.55	1.02	-.982	.329	Supplies	41	3.73	.71	Cluster correction	Public Organizations	44	3.80	3.61	1.02	-1.831	.071*	Supplies	41	4.00	.92	Redundancy	Public Organizations	44	4.12	3.93	.93	-2.167	.033**	Supplies	41	4.32	.69																																
	User complexity correction	Public Organizations	44	3.69	3.43		1.00		-2.732	.008***																																																																																																																																										
		Supplies	41	3.98	.82	Peak-time correction	Public Organizations				44	4.35	4.27	.76	-.867	.389	Supplies	41	4.44	1.00	Database size correction	Public Organizations	44	3.88	3.82	.87	-.707	.482	Supplies	41	3.95	.86	Application complexity correction	Public Organizations	44	3.94	3.82	.79	-1.494	.139	Supplies	41	4.07	.79	Application structure correction	Public Organizations	44	3.55	3.30	1.02	-2.816	.006***	Supplies	41	3.83	.70	Application load correction	Public Organizations	44	3.86	3.70	.98	-1.820	.073*	Supplies	41	4.02	.61	Network correction	Public Organizations	44	3.64	3.55	1.02	-.982	.329	Supplies	41	3.73	.71	Cluster correction	Public Organizations	44	3.80	3.61	1.02	-1.831	.071*	Supplies	41	4.00	.92	Redundancy	Public Organizations	44	4.12	3.93	.93	-2.167	.033**	Supplies	41	4.32	.69																																												
	Peak-time correction	Public Organizations	44	4.35	4.27		.76		-.867	.389																																																																																																																																										
		Supplies	41	4.44	1.00	Database size correction	Public Organizations				44	3.88	3.82	.87	-.707	.482	Supplies	41	3.95	.86	Application complexity correction	Public Organizations	44	3.94	3.82	.79	-1.494	.139	Supplies	41	4.07	.79	Application structure correction	Public Organizations	44	3.55	3.30	1.02	-2.816	.006***	Supplies	41	3.83	.70	Application load correction	Public Organizations	44	3.86	3.70	.98	-1.820	.073*	Supplies	41	4.02	.61	Network correction	Public Organizations	44	3.64	3.55	1.02	-.982	.329	Supplies	41	3.73	.71	Cluster correction	Public Organizations	44	3.80	3.61	1.02	-1.831	.071*	Supplies	41	4.00	.92	Redundancy	Public Organizations	44	4.12	3.93	.93	-2.167	.033**	Supplies	41	4.32	.69																																																								
	Database size correction	Public Organizations	44	3.88	3.82		.87		-.707	.482																																																																																																																																										
		Supplies	41	3.95	.86	Application complexity correction	Public Organizations				44	3.94	3.82	.79	-1.494	.139	Supplies	41	4.07	.79	Application structure correction	Public Organizations	44	3.55	3.30	1.02	-2.816	.006***	Supplies	41	3.83	.70	Application load correction	Public Organizations	44	3.86	3.70	.98	-1.820	.073*	Supplies	41	4.02	.61	Network correction	Public Organizations	44	3.64	3.55	1.02	-.982	.329	Supplies	41	3.73	.71	Cluster correction	Public Organizations	44	3.80	3.61	1.02	-1.831	.071*	Supplies	41	4.00	.92	Redundancy	Public Organizations	44	4.12	3.93	.93	-2.167	.033**	Supplies	41	4.32	.69																																																																				
	Application complexity correction	Public Organizations	44	3.94	3.82		.79		-1.494	.139																																																																																																																																										
		Supplies	41	4.07	.79	Application structure correction	Public Organizations				44	3.55	3.30	1.02	-2.816	.006***	Supplies	41	3.83	.70	Application load correction	Public Organizations	44	3.86	3.70	.98	-1.820	.073*	Supplies	41	4.02	.61	Network correction	Public Organizations	44	3.64	3.55	1.02	-.982	.329	Supplies	41	3.73	.71	Cluster correction	Public Organizations	44	3.80	3.61	1.02	-1.831	.071*	Supplies	41	4.00	.92	Redundancy	Public Organizations	44	4.12	3.93	.93	-2.167	.033**	Supplies	41	4.32	.69																																																																																
	Application structure correction	Public Organizations	44	3.55	3.30		1.02		-2.816	.006***																																																																																																																																										
		Supplies	41	3.83	.70	Application load correction	Public Organizations				44	3.86	3.70	.98	-1.820	.073*	Supplies	41	4.02	.61	Network correction	Public Organizations	44	3.64	3.55	1.02	-.982	.329	Supplies	41	3.73	.71	Cluster correction	Public Organizations	44	3.80	3.61	1.02	-1.831	.071*	Supplies	41	4.00	.92	Redundancy	Public Organizations	44	4.12	3.93	.93	-2.167	.033**	Supplies	41	4.32	.69																																																																																												
	Application load correction	Public Organizations	44	3.86	3.70		.98		-1.820	.073*																																																																																																																																										
		Supplies	41	4.02	.61	Network correction	Public Organizations				44	3.64	3.55	1.02	-.982	.329	Supplies	41	3.73	.71	Cluster correction	Public Organizations	44	3.80	3.61	1.02	-1.831	.071*	Supplies	41	4.00	.92	Redundancy	Public Organizations	44	4.12	3.93	.93	-2.167	.033**	Supplies	41	4.32	.69																																																																																																								
	Network correction	Public Organizations	44	3.64	3.55		1.02		-.982	.329																																																																																																																																										
		Supplies	41	3.73	.71	Cluster correction	Public Organizations	44			3.80	3.61	1.02	-1.831	.071*	Supplies	41	4.00	.92	Redundancy	Public Organizations	44	4.12	3.93	.93	-2.167	.033**	Supplies	41	4.32	.69																																																																																																																					
Cluster correction	Public Organizations	44	3.80	3.61	1.02		-1.831	.071*																																																																																																																																												
	Supplies	41	4.00	.92	Redundancy	Public Organizations			44	4.12	3.93	.93	-2.167	.033**	Supplies	41	4.32	.69																																																																																																																																		
Redundancy	Public Organizations	44	4.12	3.93		.93	-2.167	.033**																																																																																																																																												
	Supplies	41	4.32	.69																																																																																																																																																

Note) * p : 0.1 ** p : 0.05 *** p : 0.01

Table 7. Comparison between two groups of experts from the public organization and system supplier for importance of each element of capacity calculation formula of memory

Items	Classification	# of data	Total avg.	Avg. by group	Standard deviation	t-value	Variance validation																																																																																	
Formula appropriateness	Public Organizations	59	3.70	3.69	.65	-.125	.901																																																																																	
	Supplies	45		3.71	.66			Importance of element	Number of users	Public Organizations	60	4.57	4.58	.56	.246	.806	Supplies	45	4.40	.72	Required memory per user	Public Organizations	60	4.41	4.40	.72	-.163	.871	Supplies	45	4.42	.66	System area	Public Organizations	60	4.45	4.50	.75	.829	.409	Supplies	45	4.38	.75	Redundancy	Public Organizations	60	4.02	3.85	.76	-2.763	.007***	Supplies	45	4.24	.68	Buffer cash	Public Organizations	60	3.84	3.85	.78	.191	.849	Supplies	45	3.82	.68	Cluster correction	Public Organizations	60	3.64	3.47	1.03	-2.184	.031**	Supplies	45	3.87	.84	System managed area	Public Organizations	60	3.44	3.47	1.05	.368	.714
Importance of element	Number of users	Public Organizations	60	4.57	4.58	.56	.246			.806																																																																														
		Supplies	45	4.40	.72	Required memory per user			Public Organizations		60	4.41	4.40	.72	-.163	.871	Supplies	45	4.42	.66	System area	Public Organizations	60	4.45	4.50	.75	.829	.409	Supplies	45	4.38	.75	Redundancy	Public Organizations	60	4.02	3.85	.76	-2.763	.007***	Supplies	45	4.24	.68	Buffer cash	Public Organizations	60	3.84	3.85	.78	.191	.849	Supplies	45	3.82	.68	Cluster correction	Public Organizations	60	3.64	3.47	1.03	-2.184	.031**	Supplies	45	3.87	.84	System managed area	Public Organizations	60	3.44	3.47	1.05	.368	.714	Supplies	45	3.40	.81								
	Required memory per user	Public Organizations	60	4.41	4.40		.72		-.163	.871																																																																														
		Supplies	45	4.42	.66	System area	Public Organizations				60	4.45	4.50	.75	.829	.409	Supplies	45	4.38	.75	Redundancy	Public Organizations	60	4.02	3.85	.76	-2.763	.007***	Supplies	45	4.24	.68	Buffer cash	Public Organizations	60	3.84	3.85	.78	.191	.849	Supplies	45	3.82	.68	Cluster correction	Public Organizations	60	3.64	3.47	1.03	-2.184	.031**	Supplies	45	3.87	.84	System managed area	Public Organizations	60	3.44	3.47	1.05	.368	.714	Supplies	45	3.40	.81																				
	System area	Public Organizations	60	4.45	4.50		.75		.829	.409																																																																														
		Supplies	45	4.38	.75	Redundancy	Public Organizations				60	4.02	3.85	.76	-2.763	.007***	Supplies	45	4.24	.68	Buffer cash	Public Organizations	60	3.84	3.85	.78	.191	.849	Supplies	45	3.82	.68	Cluster correction	Public Organizations	60	3.64	3.47	1.03	-2.184	.031**	Supplies	45	3.87	.84	System managed area	Public Organizations	60	3.44	3.47	1.05	.368	.714	Supplies	45	3.40	.81																																
	Redundancy	Public Organizations	60	4.02	3.85		.76		-2.763	.007***																																																																														
		Supplies	45	4.24	.68	Buffer cash	Public Organizations				60	3.84	3.85	.78	.191	.849	Supplies	45	3.82	.68	Cluster correction	Public Organizations	60	3.64	3.47	1.03	-2.184	.031**	Supplies	45	3.87	.84	System managed area	Public Organizations	60	3.44	3.47	1.05	.368	.714	Supplies	45	3.40	.81																																												
	Buffer cash	Public Organizations	60	3.84	3.85		.78		.191	.849																																																																														
		Supplies	45	3.82	.68	Cluster correction	Public Organizations	60			3.64	3.47	1.03	-2.184	.031**	Supplies	45	3.87	.84	System managed area	Public Organizations	60	3.44	3.47	1.05	.368	.714	Supplies	45	3.40	.81																																																									
Cluster correction	Public Organizations	60	3.64	3.47	1.03		-2.184	.031**																																																																																
	Supplies	45	3.87	.84	System managed area	Public Organizations			60	3.44	3.47	1.05	.368	.714	Supplies	45	3.40	.81																																																																						
System managed area	Public Organizations	60	3.44	3.47		1.05	.368	.714																																																																																
	Supplies	45	3.40	.81																																																																																				

Note) * p : 0.1 ** p : 0.05 *** p : 0.01

According to the analysis to see if there is any difference between the public organizations and system suppliers, it turned out that there was significant difference on reserve rate and cluster correction. Importance of reserve rate is 3.5, and importance of cluster correction is 3.4, which is below standard of 3.5, which is excluded from the

Memory Sizing = {system area + (number of users)} * buffer cash * reserve rate

formula. Therefore, it is determined to include 5 elements except O/S managed area and cluster correction. Below is the Memory capacity sizing formula based on the survey result.

According to the analysis of system disk sizing survey results, the appropriateness of capacity calculation formula on average from the survey is 3.77, which exceeds the standard of 3.5 established by group of experts, and this formula turned out to be appropriate.

Importance of 4 elements, such as application S/W area, reserve rate, system O/S area, SAWP is 3.5 and above on average. According to the analysis to see if there is any difference between the public organizations and system suppliers, it turned out that there is quite amount of difference on both application S/W area and reserve rate.

Table 8. Comparison between two groups of experts from the public organization and system supplier for importance of each element of capacity calculation formula of system disk

Items	Classification	# of data	Total avg.	Avg. by group	Standard deviation	t-value	Variance validation																																													
Formula appropriateness	Public Organizations	61	3.77	3.85	.65	1.494	.138																																													
	Supplies	45		3.67	.60			importance of element	Application S/W area	Public Organizations	61	4.44	4.56	.67	2.005	.048*	Supplies	45	4.29	.69	Redundancy	Public Organizations	60	4.29	4.40	.62	2.078	.040*	Supplies	45	4.13	.69	System O/S area	Public Organizations	61	4.22	4.30	.74	.198	.843	Supplies	45	4.27	.72	SWAP area	Public Organizations	61	4.25	4.20	.73	-.331	.741
importance of element	Application S/W area	Public Organizations	61	4.44	4.56	.67	2.005			.048*																																										
		Supplies	45		4.29	.69			Redundancy		Public Organizations	60	4.29	4.40	.62	2.078	.040*	Supplies	45	4.13	.69	System O/S area	Public Organizations	61	4.22	4.30	.74	.198	.843	Supplies	45	4.27	.72	SWAP area	Public Organizations	61	4.25	4.20	.73	-.331	.741	Supplies	45	4.24	.74							
	Redundancy	Public Organizations	60	4.29	4.40	.62	2.078			.040*																																										
		Supplies	45		4.13	.69			System O/S area		Public Organizations	61	4.22	4.30	.74	.198	.843	Supplies	45	4.27	.72	SWAP area	Public Organizations	61	4.25	4.20	.73	-.331	.741	Supplies	45	4.24	.74																			
	System O/S area	Public Organizations	61	4.22	4.30	.74	.198			.843																																										
		Supplies	45		4.27	.72			SWAP area		Public Organizations	61	4.25	4.20	.73	-.331	.741	Supplies	45	4.24	.74																															
	SWAP area	Public Organizations	61	4.25	4.20	.73	-.331			.741																																										
		Supplies	45		4.24	.74																																														

Note) * p : 0.1 ** p : 0.05 *** p : 0.01

However, importance of 2 elements is determined to be included in the formula as those two from the public organization's response is 3.5 and above. Below is the system Disk capacity calculation formula based on the above survey results.

System Disk sizing = (application S/W area + system O/S area + SWAP area) * reserve rate

According to the analysis of data disk sizing survey results, the appropriateness of capacity calculation formula on average from the survey is 3.78 which exceeds the standard of 3.5 established by a group of experts, and this formula turned out to be appropriate.

Importance of 4 elements, such as Data area, reserve rate, backup area, reserve RAID rate is 3.5 and above on average. According to the analysis to see if there is any difference between the public organizations and system suppliers on recognition view point, it turned out that there is no significant difference on 4 elements so these 4 elements are determined to be included in the formula. Below is the Data Disk capacity calculation formula based on the above analysis results.

Table 9. Comparison between two groups of experts from the public organization and system supplier for importance of each element of capacity calculation formula of system disk

Items	Classification	Number of data	Total average	Average by group	Standard deviation	t-value	variance validation	
Formula appropriateness	Public Organizations	61	3.78	3.75	.65	-4.59	.647	
	Supplies	38		3.82	.65			
Importance of element	Data area	Public Organizations	61	4.71	4.72	.52	.289	.773
		Supplies			45	4.69		
	Redundancy	Public Organizations	60	4.36	4.38	.67	.279	.781
		Supplies			44	4.32		
	Backup area	Public Organizations	61	4.22	4.28	.80	.905	.368
		Supplies			45	4.13		
	RAID redundancy	Public Organizations	60	4.21	4.18	.81	-.378	.706
		Supplies			45	4.24		

※) * p : 0.1, ** p : 0.05, *** p : 0.01

5. Conclusions

In this paper, we made efforts to increase appropriate budget execution and investment efficiency of public organizations conducting IS project.

For this end, we proposed hardware sizing method and conducted research for empirically extracting capacity calculation formula of elements such as CPU, memory and disk.

The significance of this study is that the validation was conducted through a survey and analysis of experts from both the public organizations and system suppliers to obtain appropriateness of existing capacity calculation formula.

We also plan to make integration guideline of IS sizing with software as well as hardware. For this purpose, methods for linking related criteria, hardware sizing, software cost and maintenance cost, and integrating them would be required.

6. References

- [1] Canturk I., Apler B., Margaert M., “Long-term Workload phases : Duration Prediction and Application to DVFS”, IEEE Micro, (2005) 39-51.
- [2] Carrington, L., Sanvely, A., Wolter, N: “A Performance Prediction Framework for Science Application”, Future Generation Computer System, Vol. 22, Issue 3, (2006) 336-345.
- [3] Compaq, “Sizing a thin client Server Computing Solution Deploying Compaq ProLiant DL series Servers”, (2001).
- [4] Daniel A. Menasc, Capacity Planning for Web Performance : metrics, models, and methods, Prentice Hall, (1998).
- [5] Doro G. Feitelson, “Metric and workload efforts on computer systems evaluation”, Computer, Vol.36, No.9 (2003) 18-25.
- [6] Doro G. Feitelson, Dan Tsafir, “Workload Sanitation for Performance Evaluation”, ISPASS (2006) 221-230.
- [7] J.H Ra, G.D choi, “An Exploratory Study on Sizing Method for Information System” Journal of Korea SI Society, Vol 3(2), pp. 9-20, 2004.
- [8] Missbach, M., Gibbels, P., Karnstädt J., Stelzel J., and Wagenblast, T., Adaptive Hardware Infrastructures for SAP, SAP Press (2005)

[9] National Computerization Agency, "A Study on Hardware Sizing Technique And Framework for Information Systems", (2003.)

[10] Rowe, L. and Jain, R.: "ACM SIGMM Retreat Report," ACM Trans. Multimedia Computing, Communications, and Applications, vol. 1, no. 1, (2005) 3-13.

[11] Standard Performance Evaluation Corporation, <http://www.spec.org>

[12] Transaction Processing Performance Council, <http://www.tpc.org>

U. Lublin and D. G. Fetelson, "The workload on parallel supercomputers : modeling the characteristics of rigid job", J. Parallel & Distributed Computing, Vol.63, No.11 (2003) 1105-1122.

Authors



Jonghei Ra

Received a B.S. degree in information engineering from Sungkyunkwan University, Korea, 1990, and M.S. degree in Information engineering from Sungkyunkwan University, Korea, 1992 and Ph D. degree in Information engineering from Sungkyunkwan University, Korea, 2001. In 2003 he joined the faculty of Gwangju University, Korea where he is currently a professor in Department of e-Business. His research interests include System performance,

Web computing, information system audit, Data Mining.



Haeyong Jung

Received a B.S. degree in Management Information System from Kwangwoon University, Korea, 1996, and M.S. degree in Management Information System Kwangwoon University, Korea. 1998 and Ph D. degree in Management Information System from Kwangwoon University, Korea, 2003. In 2003 he joined the faculty of Korea Nazarene University, Korea where he is currently a professor in Department of Management information system. His research interests include Information system evaluation, E-Learning, Management Innovation thru

Information Technology.



Kwangdon Choi

Received a BA degree in Business Administration from Kwangwoon University, Korea, 1985, and MBA degree in Information Management Systems from Hankuk University of Foreign Studies, Korea, 1987 and Ph D. degree in Business Administration from Kwangwoon University, Korea, 2001. In 2002 he joined the faculty of Hansei University, Korea where he currently a professor in Department of Electronic Business.

