Onome Imoniana, Joshua

Philosophical and methodological reflection on database capacity planning: a controllership standpoint

PHILOSOPHICAL AND METHODOLOGICAL REFLECTION ON DATABASE CAPACITY PLANNING: A CONTROLLERSHIP STANDPOINT

Joshua Onome Imoniana
Methodist University of São Paulo, Brazil

ABSTRACT

This paper analyzes the capacity-planning process of database management systems focusing the theoretical decision models that considers the variables constraints that could be taken slightly and other variable constraints that are indispensable during upgrade. In order to proceed with the aforementioned, we assume that capacity planning decision, which considers the general corporate, technical, users and executive views aiming at prioritizing a measurable relative importance of the planning constraints is capable of giving more economic added-value decision elements to the management. Thus, the optimization multicriteria analysis leads us to infer upon the reasonableness of the decision taken in regard to implementing budgetary planning in moment of business uncertainties. Results of this study indicate that corporate strategic views in decision-taking processes some times override some technical, users and executive views as relating to capacity planning in information technology management environment.

Keywords: database, capacity, planning, management control, systems.

1 BACKGROUND

As it is incumbent on the controllership department to consolidate operating budgets of an organization and also enhance the monitoring procedures, it is also a concern that the controllers department optimizes the available resources in order to implement the Management Control Systems and the Decision Support Systems (DSS). When we mention management control systems in this study we mean the following three groups of systems:

- Value Systems – Mostly used by the to-management to circulate the mission statements and credos to every *Dick and Harry* in the organization, for instance: Strategic Planning System.
- Administrative Systems – They are systems mostly accounted for by the middle-level managers, committed to attainment of operational, economic & finance and accounting transactions, for instance: Invoicing & Receivable, Purchases & Payable, Inventory, Payroll, Fixed Assets and General Ledger Systems.

- Performance Evaluation Systems – They are systems used by the top-management to monitor performance in key strategic areas of an organization, for instance: Balanced Score Card, Budgetary Systems.

According to Kimball (1997:58) decision support systems can be used for operational control, management control and strategic planning. In another point of view, one would suggests the application of Executive Information Systems (EIS) which is mentioned in Furlan (1998:98) as a tool to browse database for the presentation of information in a simple and friendly format, by addressing the needs of the higher executives thereby eliminating the need for intermediation.

The capacity planning therefore, being issues raised in a management control process of a database, involves the conjugation of the technical knowledge of the information technology and the financial know-how of the accounting or financial department to prudently stipulate how much it is worth spending, considering current and future operations already budgeted for.

As in Robert (2002:14-51) which analyzes a capacity–planning and pricing problem of a monopolist facing uncertain demand. His study incorporates "soft" and "hard" capacity constraints. In this model, soft constraints could be relaxed at a cost while hard constraints cannot be relaxed.

We do no have a specificity of environment being studied. As in Ahmed & Lochovsky (1998:69) the goal of the study was to explore a macro-level and not to compare specific facilities and functionality of object-oriented and relational 4GL approaches or database systems *per se*.

1.1 Research questions

How does the Controllers and the Chief Information Officers (CIOs) go by the decision processes to upgrade the capacity of the database of their organization?

How does one measure the effectiveness of the existing capacity?

1.2 Research Hypothesis

To assist the development of this study the following assumptions were taken:

H_1 The capacity planning decision, which considers the general corporate, technical, users and executive views aiming at prioritizing a measurable relative importance of the planning constraints is capable of giving a more economic added-value decision elements to the management.

H_2 Capacities planning of a database system should primarily reflect the volumes of transactions processed by the organization, any distance from this definitely would result in over projection and consequentially arriving at a redundant capacity that signifies a tied down capital.

H_3 Even though database dimensions are planned to match the operational activities of any organization, uncertainties that result in cash flow deficiencies may also result in a set back in investment projects that rescinds decision taken before hand.

1.3 Research Context

A database is composed of a collection of interrelated data files and joint operating procedures that enhances its maintenance with little of no difficulties. The tasks to retrieve and manipulate such data are being simplified day in day out, and this has resulted to the proposition and commercialization of data warehouse which conjugates methods and technology in a pragmatic, simple and user-friendly load-and-access environment. Data warehouse is a group of data that is theme-oriented,

integrated, however changes with time but not volatile, to support a decision-taking process Inmon (1997:75).

Database management systems have evolved from a simple and specific computer application to a modern and complex technological data processing environment, now it is dynamic, a part and parcel of convergent economic resources of any competitive organization. For that reason, interest in the dispensation of planning techniques to measure its capacity has also grown along side.

A database is considered as a system, which facilitates the storage of a greater volume of information about organizations or business unit having programs, in alliance with its application systems, enables its users to manipulate files, including the following facilities:

- Avoid redundant data;
- Facilitate the retrieval of data and if possible restrict on need to know basis;
- Exclusion of obsolete items;
- Reorganization of data files;
- Concatenation of similar files;
- Execution of simple functions to generate reports for the monitoring of the database
- Possible simultaneous sharing of various resources existing in the data base by other users in different locations
- Restriction of access privileges to modify database

With the outcome of data warehouse therefore, data manipulation is enhanced. According to Cavalcanti *et al* (1998:4-45) it is fundamental to create mini-dimensions to ensure performance; the authors mentioned that there are most commonly used attributes for searching, for example, those attributes known as demographics (Marital status, age, and gender, per capita) of a client (dimension).

1.4   Capacity of database

Capacity of a database is the maximum amount of the work a controlled database management system can withhold in a manageable unit.

Apart from the capacity of the database, performance indexes that are monitored are:

- Turnaround time – time interval between the instant a program is read into the system for processing and the instant the program execution is completed;
- Response time – is the time interval from the instant a command inputs to the system (strikes enter) to the instant the corresponding reply appears on the terminal. See graph 1. The X axis normally shows the loading capacity, and the Y with the corresponding response time;

Graph 1 - System capacity and Response time
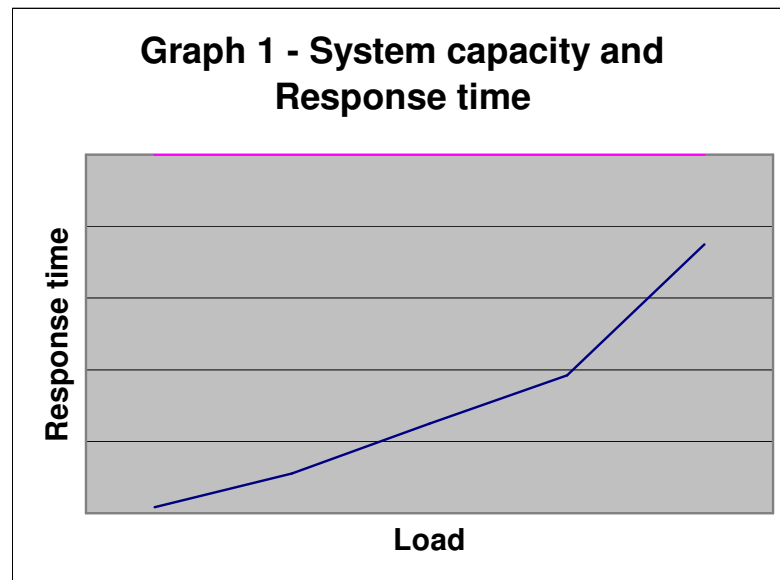
- Availability – the percentage of the total time during which the user has access to the system.
- Throughput – is the measure of the systems productivity or efficiency as related to the existing backlogs. See graph 2. The X axis normally shows the loading capacity, and the Y with the corresponding throughput;

Graph 2 - System Capacity and Throughput
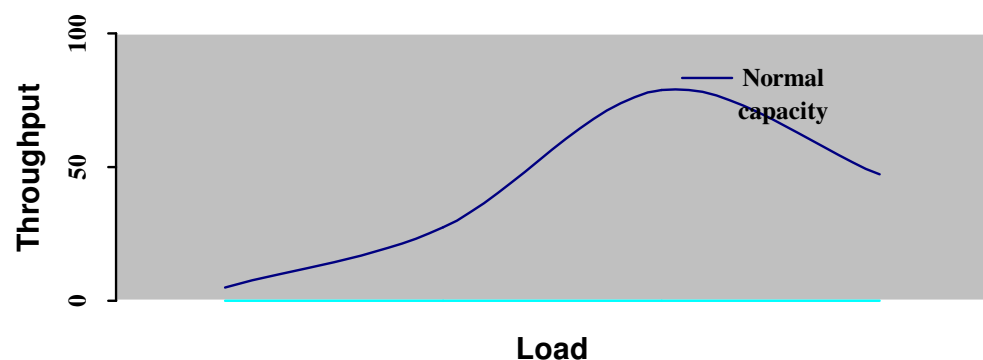
1.5   Capacity Limit

Capacity limit refers to the limit (or threshold) of utilizable capacity; is a point at which any additional workload is expected to result in excessive degradation of system performance. Lam & Chan (1990:51).

The question that triggers capacity decisions therefore is what level of utilization will prevent inordinate queuing delay of those resources that may impede the responsiveness of the application systems?

According to the just mentioned authors, a simple methodology based on the capacity limit rules of thumb is quite commonly used for computer performance evaluation and capacity upgrade decisions. They cited that:

system is monitored and the utilization tracked;

projected utilization levels resulted from an extrapolation of the utilization curve plus utilization estimation for applications are compared.

1.6   Capacity Planning

Capacity planning is a whole-brain process and utilizes both subjective and objective dimensions of consciousness, the feelings and the facts. Planning requires judgment and judgmental activities necessarily require both the abilities to analyze components Browning (1994). To illustrate his concept, he associated his words to the diagram shown in Figure 1 – Capacity Planning Modalities.

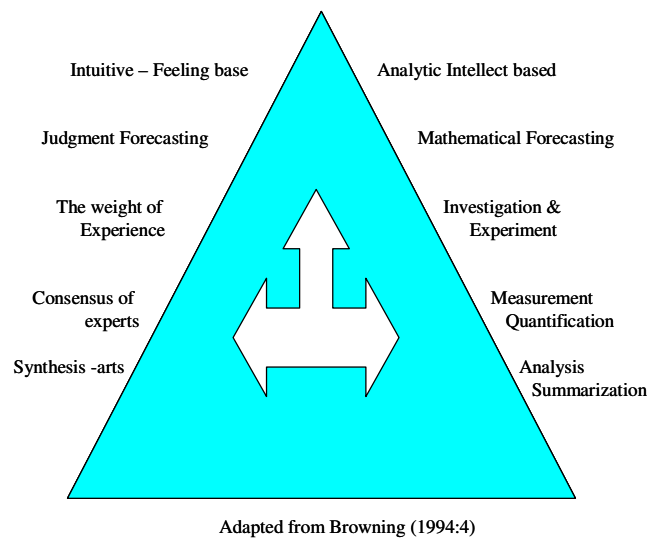Adapted from Browning (1994:4)

Figure 1 - Capacity Planning Modalities

In a broader perspective, capacity planning could be four-fold: a) Corporate view b) Technical view c) User view; and Executive view. See Figure 2 below.   As demonstrated in the mentioned figure, the corporate view which unites all the subsets of the individual views is concerned with the business optimisation of resources, value-added and sustainability of the activities; in other words, it is the going-concern view. The technical view is concerned with the usage of the computer resources (hardwares, softwares, material, etc.) with a hope to registering its effectiveness. The user view is related with the availability of the resources when needed to meet his day to day economic transactions, normally this view is concerned with the response time and the throughput.   The executive point of view is results oriented, monitors the cost-effectiveness of the operating functions and tries to draw a line between the budgeted and the realized values so as to decide about forecast adjustments and better measure returns on investments.
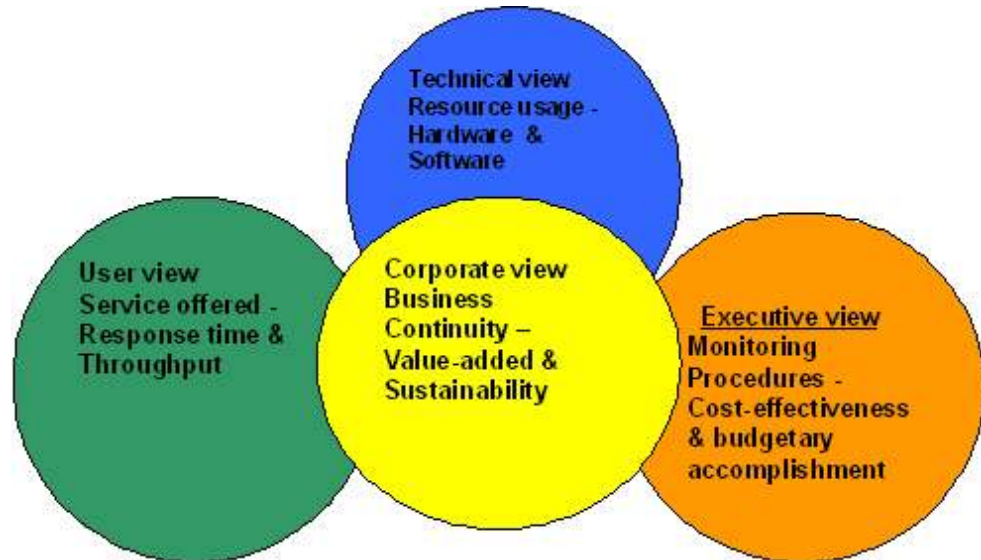
Figure 2 - The Four views of Capacity Planning

Workload forecasting to enhance decision-making

Generally, workload forecasting is a most strainous task in information technology management. Some arbitrarily define it to protect personal interest or Systems Analysts define without the clear consent of the user. Saviani (1992:62) user accuses the Systems Analysts for presenting political solutions previously defined by them which are incapable of prioritise their requirements for information, define individual objectives, forget their responsibilities and escape compromise.

For controllership point of view, it is important in order to present a more realistic scenario mainly to portray its financial consequences, since it is a problem worth resolving to avoid processing stagnation that could result in lost of earnings. Evidently, in day to day management control procedures means are being devised to ameliorate critical situations such as the installation of intelligent terminals that shifts the responsibilities from the CPUs. As in Lam & Chan (1990:34) workload forecasting is typically the weakest element in the computer capacity planning process. A common workload measure, CPU hours consumed is too abstract for the user community and cannot be related easily to other business measures (e.g. Invoicing) in the organization.

Normally, workload increases are as a result of:

Future applications based on historical data;

Latent workload, a portion of the load that increases as a result of malfunctioning or poor performance.

Methods used to forecast a workload ranges from personal judgements and analysis (feeling) or intuitive to the mathematical (scientific) such as the Box-Jenkins method and objective methods (operating figure analysis).

In order to assist decision-making, various methods such as the following are used to forecast workload, they are: Time series, Exponential method, Box-Jenkins, Structural method, Delphi to mention just a few.

Time series regression method – This is based on the premise that forecasted items are directly related to a workload data.

Thus:

$$\hat{W}_t = a + bt.$$
………………………………………………………………………….(1)

where $\hat{W}_t$ represents forecasted workload at a time *t*.

and "*a*" and "*b*" are constants estimated by the least squares method utilizing historical workload data.

Exponential smoothing models – This is recommended for the short and medium range forecasting.

It shows that:

$$d = \frac{2}{(N+1)}$$
……………………………………………………………………...(2)

where *N* is the number of periods in a cycle and "*d*" is the smoothing constraint. These constraints could be defined as:

Volatile (small) assign weights (light) according to feelings;

Stable (large) assign weight (heavy) according to feelings.

Box-Jenkins method – This philosophical method postulates a generalized model that explains a large number of time series data. The (B-J) tentative data results are normally analysed for their reasonableness and if found cohesive it is used as a base fore forecasting.

2 METHODOLOGY

The rational behind this study is the exploratory nature. It shows how applications systems development and its maintainability is increasingly becoming very important in highly competitive environments. What is alarming is the approximately 50% cost needed to put any hardware into operation, representing the cost of software implementation. Normally, we are conversant with project implementation groups working behind schedule and overrun of budget is not also strange.

It is on this note that The Software Engineering Institute (SEI) in Pittsburgh was asked by the United States Air Force in 1986 to study the systematic way to evaluate software contractors. And as a result of this request, a Capacity Maturity Model (CMM) was invented, see Figure 3.

Level	Characteristic	Challenges
5 Optimizing	Feedback - the process includes back-up and returns	Continuous improvement
4 Controlled	Quantitative -the process is measured	-Technical innovation - Defect preventions - Process change mgmt
3 Defined	Qualitative - the process is described and institutionalized	- Process measurement - Process analysis - Quality management
2 Repeatable	Intuitive -the process is person -dependent but managed	- Peer review - Intergroup coordination - Product engineering - Integrate mgmt - Training program - Org. process Definition - Org. process focus
1 Initial	- Chaotic -Ad hoc	- Requirement mgmt - Quality assurance - Subcontractor mgmt - Project planning - Project tracking & oversight - Configuration mgmt

Adapted from Baskerville & Pries-Heje (1999:26)

Figure 3 – Capacity Maturity Model

According to Barkville & Pries-Heje (1999:26) maturity levels are composed of one or more key process areas, if the organizations develop these areas, the software process is known to improve. As a methodological approach to this study therefore, we crave the indulgence of the points of views presented by the mentioned authors thereby expatiating on levels 3, 4 and 5 of the capacity maturity model.

3 OPTIMIZATION OF THE DATABASE RESOURCES ENVISAGING BUDGETARY CONSTRAINTS

In as much as the operating environment works according to organization policies, files segments describes the space an application is occupying in a database and in terms of optimization it gives a greater concern as regards capacity planning. For instance, if the space extension is arbitrarily used to recharge operating departments, meaning the users, the accounting department that uses more of database to house accounting information systems and financial information that requires retention for several years, should bare the greater part of the cost of database systems. Thus, other criteria proposed for adoption should be unanimously accepted and should also suite the laid down policies of such establishment.

Monitoring of the optimization procedures of the database systems involves the follow-up of the performance through the tools of DMS – Database Management Systems, characterized by:

- Coordination of the maintenance of database architecture (definition, creation, exclusion and classification of data that would be stored in it);
- Definition of access policies, confidentiality, and data integrity (description of data elements, access keys, etc.);
- Maintenance (physical reorganization and routine backup) and documentation of databases;
- Coordination of data administration activities, users and systems programming;
- Development and maintenance of standards

Therefore, the mechanisms or better say major characteristics of capacity planning summarized in the views shown in figure 3, are eminently qualitative variables and in order to be controllable certain elements of measurability has to be established, thus, for decision-taking process, multicriteria methods could assist in this process.

For this reason, we revive the premise that capacity upgrade go-ahead signal depends on the sum of the relative importance and the associated value functions of the corporate view, technical, users and the executive views.

Thus, as already mentioned in the literatures the perception of capacity is subjective, and so there is a need to stipulate a level of importance which the planners in the assessment process could define by attributing relative importance to the general views:

$$W_{cus} = w^j cv^j + w^j tv^j + w^j uv^j + w^j ev^j$$

(3)

for $\sum_{j=1}^{2} w^j = 1$, and w^j is the relative importance (weight) of the general views.

The capacity upgrade go-ahead signal *cus* is a function of the criteria values of corporate views *cv*, technical view *tv*, users view *uv*, executive views *ev*, and relative weights w^j_i of each view, such as:

$$cus^j = w_i^j v_k (cv) + w_i^j v_k (tv) + w_i^j v_k (uv) + w_i^j v_k (ev)$$

(4)

that can assume the aggregate form:

$$cus^j = \sum_i^m w_i^j v_k (cpc_i^j)$$

(5)

for $\sum_{j=1}^{m} w_i^j = 1$ and $0 < w^j_i < 1$, where *cpc* is the capacity planning criteria for updating database of the planning general views j.

4 DEFINITION OF RELATIVE IMPORTANCE OF PLANNING VIEW W^J AND CAPACITY PLANNING CRITERIA W^J_I

The estimation of weights or relative importance of planning view and capacity planning criteria is formulated through the *Analytic Hierarchy Process* (AHP), as in Saaty (1990, 1991). The process is initialized with parity comparison between planning

view w^j and capacity planning criteria w^j_i, according to the relative importance judgment scale of values shown in table 1.

Table 1 – Relative importance judgement scale.

Ratings of Absolute Importance	Explanation
a to $b = 1$	Two activities contribute equally to the objective and are of the highest importance.
a to $b = 3$	Experience and judgement moderately favours a over b.
a to $b = 5$	Experience and judgement strongly favours a over b.
a to $b = 7$	a is strongly favoured over b and its dominance is demonstrated in practice.
a to $b = 9$	The evidence favouring a over b is the highest possible.
a to $b = 2,4,6,8$	When compromise is needed.
Reciprocals	If activity a has one of the above numbers assigned to it when compared with activity b, then b has the reciprocal value when compared to a.

Source: Adapted from Saaty (1990, 1991)

Definition of value function of capacity planning criteria $v(cpc^j_i)$

The value function of capacity planning criteria is established through cardinal values attributed to standard view. In tables 2 below are shown standard situations of the optimization criteria, defined for the study with their respective value functions and the decision attributes.

Table 2 – Standard Circumstances of cpc^j_i.

Optimization criteria cpc^j_i	Standard Circumstances	Value Function $v(rc^j_i)$	Decision Attribute
Corporate	Business continuity and assurance	1,00	Very Significant
	Value-added business operation	0,75	Significant
	Sustainability of the operational level	0,50	Moderate
Technical	Resource usage and efficiency	1,00	Very Significant
	Software resources implementation	0,75	Significant
	Hardware resources implementation	0,50	Moderate
Users	Systems usage and recharge to users	1,00	Very Significant
	Applications throughput	0,75	Significant
	Applications response time	0,50	Moderate
Executive	Management Control Systems cost-effectiveness	1,00	Very Significant
	Budgetary accomplishments	0,75	Significant
	Performance Evaluation & monitoring procedures	0,50	Moderate

5 SIMULATION AND COMMENTS

In order to explain the exercises performed to test the assumptions of this research, the following additional methodological steps were reiterated:

Comparison of the relative importance of planning view w^j and capacity planning criteria w^j_i are essential to enhance a conclusion about the necessity to upgrade a data base.

Definition of value functions $v(rc^j_i)$ of the views raised concerning the four dimensions of capacity planning is important to give a score even though subjective, to a decision process that needs quantitative results.

An association of decision attributes and the reasonableness of the rules of thumb should be highlighted.

In the day to day activities of an organization note that steps a and b are standardized to enhance the assessment of capacity planning decision. Thus, standard information should be constantly reviewed because of the periodical mutation of business environment, internal or external to the organization and value references of the decision taker. And steps c should be developed for every specific controllership optimization processes.

Upon application of the approach therefore, see Table 3 a go-ahead or not decision was reached as we sum up the relative importance of the views raised as the supporting points for the database capacity upgrade.

Table 3 – Comparison of relative importance of planning view and capacity optimization criteria

Optimization Criteria cpc^j_i	Standard Circumstances	Value Function $v(rc^j_i)$	Decision Attribute
Corporate	1. Business continuity and assurance strategy	1,00	Very Significant
w^j_i 0,545 =	2. Value-added business operation	0,75	Significant
v^j_i= 1	3. Sustainability of the operational level	0,50	Moderate
Technical	1. Human resource usage and its efficiency	1,00	Very Significant
w^j_i 0,258 =	2. Software resources implementation	0,75	Significant
v^j_i= 1	3. Hardware resources implementation	0,50	Moderate
Users	1. Systems usage and recharge to users	1,00	Very Significant
w^j_i 0,105 =	2. Applications throughput	0,75	Significant
v^j_i= 1	3. Applications response time	0,50	Moderate
Executive	1. Management Control Systems cost-effectiveness	1,00	Very Significant
w^j_i 0,082. =	2. Budgetary accomplishments	0,75	Significant

$v^j_i = 1$	3. Performance evaluation & monitoring procedures	0,50	Moderate
	Capacity Upgrade Signal =	1,000	High Priority

Upgrade Signal	Numeric Interval	Value Function
Less Priority	$0,50 <= W_{cus} < 0,75$	0,50
Medium Priority	$0,75 <= W_{cus} < 1,00$	0,75
High Priority	$W_{cus} = 1,00$	1,00

Taking an organization *per se*, where its decision takers are trying to reach a consensus as regards a capacity upgrade of its database systems have to do a lot of brain storming exercises. In the aforementioned case, having raised the issues, and various views collected, a homogenous view for the need to upgrade the management control systems data base was pinpointed. This as a result of the relative importance corporate view (0,545) relating to business continuity and assurance strategy issue, technical view (0,258) relating to human resources usage and its efficiency, users view (0,105) relating to systems usage and recharge to users enhance, and the executive view (0,082) management control systems cost-effectiveness. These relative scores, as one could name them, when summed up (1,0) shows a high priority capacity upgrade signal. Thus, it looks very evident that the business continuity and its assurance strategies that affect the going-concern issues have been considered. If the scores had been less, it would have shown either a medium or less priority capacity upgrade decisions.

6 CONCLUSION

It is clear that this study suggests that a focus on capacity planning of database management systems could be of importance for organizations not just as a tool for resources optimization but avoids an arbitrary decision-taking processes concerning issues that have a more subjective measurement criteria.

Even though cautions should be taken with regards to interpretation of findings arising from exploratory study of this nature, results makes us to infer that some corporate view in decision taken processes some times override some technical, users and executive views as relating to capacity planning in information technology management environment.

Therefore, we confirm the assumption that capacity planning decision models, which considers the general corporate, technical, users and executive views aiming at prioritizing a measurable relative importance of the planning constraints is capable of giving a more economic added-value decision element to the management.

This study sparks interest into issues that pose problems to the controllership department and the information technology management during budgetary planning and the monitoring procedures; nevertheless, it abridges the existing distance between the day to day treatment of the issue in the organizations and the scientific proofs.

These findings suggests several avenues for future researches relating to capacity planning in data base environment and associated themes, especially those having to do with information technology resources as a whole and cost allocation.

7 REFERENCE

Ahmed, Iftkhar & Lochovsky, H. Frederick. A comparative Case Stady of Three Database Application Development Environments. DATA BASE for Advances in Information Systems – Fall, Vol 29, no.4 pg. 69, Atlanta: 1998.

Baskville, Richard & Pries Heje, Jan. Knowledge Capacity and maturity in software management, The Database for Advances in information Systems Spring, vol 30 no. 2 pg. 26. Atlanta: 1999.

Browning, Tim. Capacity Planning for Computer Systems. Academic Press, London: pg 4-50, 1994.

Cavalcanti, Maria Cláudia, Klein, Lawrence Zordan e Alenquer, Pablo Lopes. Princípios de Projetos para um Data Warehouse Dimensional: Rio de Janeiro, pg. 5-45, UFRJ,1998.

Robert, F. Gox. Capacity planning and pricing under uncertainty. Journal of Management Accounting Research. Sarasota: Vol. 14 pg. 59, 2002

Kimball, Ralph. The Data Warehouse Toolkit. New York: John Willey & Sons, pg. 58, 1997.

Lam, Shui F. & Chan Hung K. Computer Capacity Planning – Theory and Practice. Academic Press, Toronto: pg 34-51, 1990.

Furlan, José David. Sistemas de Informações Executivas. São Paulo: Makron Books, pg. 98, 1994.

Inmon, William H. Como Construir o Data Warehouse. Rio de Janeiro: Campus, pg. 75, 1997.

Saaty, Thomas L. Método de Análise Hierárquica. São Paulo: McGraw-Hill, pg. 99, 1991.

____, Thomas L. How to Make a Decision: The Analytic Hierarchy Process. European Journal of Operational Research, North Holland, V. 48, 1990, p. 9-26.

Saviani, J. Roberto. O Analista de Negócio e de Informação. Atlas, São Paulo: pg. 10-62, 1992.