This paper was originally published in the Proceedings of the Computer Measurement Group's 1987 International Conference.

# THE USE OF NATURAL FORECASTING UNITS

John M. Reyland
Community Mutual Blue Cross & Blue Shield
Cincinnati, Ohio

## ABSTRACT

The use of Natural Forecasting Units (NFU's) within capacity planning programs has proven to be an effective procedure. This tutorial considers the identification and measurements of NFU's and their use to project future computer system requirements based on planned or statistically projected NFU changes. It continues with a general report of a real experience utilizing NFU's, how their use not only projected future computer system resources, but also reported unexpected capacity and performance problems for a converting IMS system. The tutorial is structured around two different approaches in the use of NFU's and the capacity planning program.

## INTRODUCTION

The NFU is a business element measurement from the business side of an organization which can be related to the use of computer systems resources by workloads. Examples of potential NFU's are purchase orders, production output counts, insurance contracts, accounting data, etc.. The purpose of NFU's can be twofold. The first deals with measuring and evaluating current system workloads as to their capacities and performance relative to selected business elements. The second deals with the impact of planned NFU changes as to their future impact on system configurations, workload capacities and performance, capital investments and budgets.

T. L. Lo [1] succinctly defines the NFU concept as attempts to link business activities/plans to DP activities/plans. Such linkage enables the data processing function to develop DP plans based on corporate plans, which is very important because of large capital investments needed to maintain adequate computer support.

The number of manufactured widgets might be directly related to a particular system workload and its CPU time. This would be expressed as a CPU time unit per manufactured widget. If this is a valid functional relationship between these two measurements, then one would observe during a given time period that a change in the number of manufactured widgets would cause a corresponding change in the total CPU time for the particular workload. Such a coorelation makes it possible to predict future system resources by workloads.

It is essential that one understands the functional relationship between NFU's and computer system resources by workloads. This is the basis of the NFU concept and its use to predict future computer system requirements. The above example of the academic widgets is simplistic, but it illustrates this functional relationship. The system resource side of this relationship is sometimes referred to as Data Processing Units (DPU's). Thus the relationship is between NFU's and DPU's.

## THE NFU COMPONENT

Depending upon the business environment, there are many possible approaches to measuring and evaluating NFU's, which will functionally relate to computer system resources by workloads. There are some general considerations when selecting potential NFU's.

The first consideration is measurability, which is the ability to identify and count the occurrences of NFU's. This must be done with ease of identification and with consistency from reliable sources.

For purposes of this tutorial, NFU's can be categorized as contractual services or product lines. An example of a business industry that provides contractual services would be the insurance industry. Product line business would be the manufacturing of cars or television sets.

Identification deals with the terminology about contractual services and/or product lines. This can prove to be difficult, especially in the areas of contractual services. Services like medical insurance plans, which have evolved over time, have an ever changing realm of terminology. Terminology difficulties can also be found in product line businesses. The history of product diversification, model numbers, product specifications can all contribute to the problem of identification. Such instability makes it difficult to identify and track NFU's over long periods of time.

Consistency deals with the measurability of NFU's across business segments of the same organization. This is especially true within organizations that provide essentially the same service by its business segments, but the service has been tailored for particular client groups. This is probably a lesser problem in product line businesses.

The second consideration deals with the quality of the functional relationship between a potential NFU and its DPU. This quality is statistically referred to as correlation. Examples of the quality of correlation are illustrated by the following environments:

* The best measure of correlation can be found in a dedicated online system, which has individually identified transaction codes that support specific services.
* Correlation would diminish if the above services were supported by an online system that simultaneously processed multiple functions with a limited number of transaction codes.
* Batch processing would probably provide the least amount of correlation. To identify and count types of input records by type process is a less dependable procedure.
* If contractual services and/or product lines do not have direct data processing support, then correlation must be determined by some indirect means. This might involve sales records, accounting data or the use of manufacturing resources.

The ability to identify and use NFU's from the business side of an organization is greatly dependent upon one's understanding of the business organization. One should study the formal, and if possible, the informal structure of the business organization. Within this study, one must be careful to identify line and staff functions. It is within the line function that most NFU's will be identified, even if measured by a staff function such as accounting. The ability to identify and measure NFU's will also depend to great extent on the dynamics of an organization. Business and non-business organizations must continuously change to be competitive.

It is risky to select NFU's by a trial-and-error approach. The risk involves the reliability of selected NFU's, and the credibility of the capacity program. NFU's only become reliable when they have a history of being useful and dependent upon to make valid projections of required computer system resources.

## COMPUTER SYSTEM RESOURCE COMPONENT (DPU)

The identification of a computer system resources unit (DPU) is found by examining workloads that directly support an NFU. This identification can become difficult because workloads may contain composites of work. A batch workload may contain both test and production work. An IMS system may have multiple IMS systems, or a transaction may perform multiple functions and call other transactions. To avoid this kind of problem, it is essential to identify specific workloads by assigning them to unique performance groups within the IPS of the System Resource Manager, a component of the IBM's MVS operating system. Non-IBM computer systems need to have similar features to isolate workloads.

Performance groups can be used to isolate types of workloads, which also isolates system resource utilization by such workloads. Depending upon the type of workload, not all system resources are reported by SMF/RMF at a great enough detail to be useful.

SMF/RMF will report by performance groups system resources utilization for batch workloads. Typically this includes CPU times, execution times, elapsed times, allocation times and for devices their I/O counts and service times, etc.. There still remains hundreds of other batch workload measurements.

TSO workloads are very much like batch. In addition to typical system resources being reported, TSO also reports Puts and Gets as terminal activity, and in some environments, TSO commands.

Beyond batch and TSO, data about the utilization of system resources must come from other dedicated sources. IMS and CICS have their own programs to reduce log data to meaningful reports. IMS PARS provides extensive data about individual transaction codes, elapsed times, processor times, data base calls, etc.. IMS also has DB/DC real time monitors that provide even greater detail than data reduction programs for log files. CICS systems have similar capabilities to obtain detail data about transactions and system resources.

There are communications packages that sample, and in some cases, report actual system resource utilization which can be used for communication workloads.

Most IBM compatible software vendors do provide some SMF recording about system resource utilization and transaction counts. This data, along with standard SMF and RMF data, can provide enough data to include such workloads within a capacity program.

## NFU PROJECTIONS

There are several approaches in the use of NFU's to project the need for future system resources.

The first approach is to relate NFU's to input counts such as IMS or CICS transactions. These counts are used to drive a queueing or simulation model to change workload demands within a model. Projected IMS or CICS transaction counts are determined by an NFU conversion factor, which is calculated by relating NFU's and actual transaction counts. Future system resource requirements are determined from the changes within the model.

Such models usually require a preprocessor program to reduce SMF, RMF or other system data to workload characteristics. These programs will also generate input files for models, which contain required system resource measurements by workload. Morino's model generator and BGS's Capture program are examples of such software.

The second approach is to statistically project NFU's for purposes of determining changes in DPU's. Such changes are then translated into new hardware and software requirements.

## NFU CONVERSION FOR MODELING

The needed functional relationship for the NFU is between the NFU counts and the model's input transactions. Such a relationship could be between policy holders and actual IMS transactions. There is also a need to mark this relationship at a point in time, possibly month end. All counts must also be scaled down as required by modeling software, which usually operate at hourly intervals. The selected hourly interval can represent a peak processing period or can average value.

More importantly the relationship must be functional. If there is an increase in policy holders and such policies are supported by a dedicated IMS system, then an increase in policy holders will cause a corresponding change in the number of IMS transactions. Such IMS transactions can be categorized as those that update policy information, inquire against the policy holder's databases and process policy holder's claims. The opposite is true, a decrease in the number of policy holders would result in a decrease in the number of supporting IMS transactions by transaction category.

These relationships must correlate either directly or indirectly. They may be simple, involving one NFU component or a string of individually weighted components. The more complicated a function becomes, the more difficult it is to measure and validate the relationship. The complexity of such relationships would be determined by the business environment. The mathematics usually involves the division of the transaction counts by the NFU component. This produces a conversion factor of IMS transactions per policy holder during the same time interval. This relationship can be expressed by categories of supporting IMS transactions. Conversion factors could be developed for IMS transactions per inquiry, IMS transactions per claim and IMS transactions per miscellaneous task. For each conversion factor there must have been an identifiable system workload with dedicated performance groups and IMS transaction codes.

Conversion factors from one time period to another should have approximate values. Historically they should have statistical distributions, which can be evaluated by standard statistical methods.

An example of a conversion process will serve to bring together the concepts about NFU's and their use for modeling. The example relates policy holders to IMS transactions. The process counts the NFU's, policy holders at month end and then scales selected IMS transactions to hourly averages. The NFU's are categorized as claims processing and non-claims processing. Dividing the hourly IMS transactions by the respective NFU will result in two conversion factors. The first would be average hourly IMS transaction for claims. The second would be the average hourly IMS transactions for non-claims processing. This assumes that the IMS system has been set up so that IMS transactions used to process claims have been uniquely defined and assigned to message processing regions (MPR's) and performance groups. The same would be true for the non-claims processing.

Sequence of steps to complete the conversions and utilize the NFU for projections:

* Assign IMS transaction codes to MPR's as claims and non-claims processing. Assign MPR's to dedicated performance groups.
* Count IMS transactions by selected IMS transaction codes needed for claims and non-claims processing.
* Scale IMS transactions to hourly rates. This can be done by having available hourly counts that correspond to the desired hour selected to be used by the model. If hourly IMS transaction counts are unavailable, then daily IMS transaction counts can be scaled by use of measured factors. Experience in our environment indicates that 12.78% of the daily IMS transactions will provide a peak hourly rate of IMS transactions.
* Perform the respective divisions to obtain the conversion factors for the hourly number of IMS transactions per claim and non-claim processing.
* Define the performance groups to the enqueueing or simulation model and execute the data reduction programs needed to process SMF and RMF data for the model.
* Distribute IMS system resources to IMS workloads within the model. This is necessary because SMF and RMF recording of IMS I/O's and CPU times are aggregate values for a processing period. By means of the DC monitor, the percentage distribution of these system resources can be determined. Such IMS resources usually involve the CPU times of the control region to MPR's and the distribution of I/O's by DDNAME to storage devices.
* Validate the model from actual system performance values and system measurements.
* Project the IMS workloads by using the conversion factors with planned changes of NFU's. This involves the multiplication of the monthly conversion factor by the planned monthly changes in NFU's. This is done on a monthly increment, since NFU's and IMS transactions are counted at month end.
* It is assumed that the model is being executed with all its other workloads, and that they are changing or remaining the same during the projections. It is also assumed that with the current workloads, the system is operating at over 85% CPU utilization and that there is a shortage of DASD storage.
* As the NFU's are converted to greater number of IMS transactions, the system would saturate with infinite response times because of the lack of system resources. It would be necessary to add system resources and to reconfigure the total system to process the increasing number of IMS transactions, while maintaining desired performance levels.
* The model's report of changed system resources are the future system requirements to process the planned NFU changes.

## A REAL EXPERIENCE WITH NFU's AND MODELING

The above example is representative of our NFU program during the past 18 months. The purpose of our NFU program is to track, analyze and make recommendations for future system requirements, as part of a multi-year conversion effort. The conversion involved the replacement of an older IMS health insurance system, which does not have online claims processing, to a newer system with online claims processing and other desirable features.

There have been three stages which can be reported on in general terms, as means to present some realism of an on-going NFU program.

The first stage was early in the conversion schedule. Actual sample sizes of converted contracts were non-existent. In place of these samples, there was another health insurance system, similar to the new system, which was used to obtain system utilization and IMS data. There was not an official consolidated schedule for the conversion of contracts by the five business segments of the organization. By means of a telephone we were able to consolidate the individual conversion schedules.

The contract counts functionally related to IMS data. The conversion factors were stable and useful to compute IMS transaction levels as input to the model. The results of the model indicated an inordinate amount of system resources needed to process online claims in the outer time periods. This was disconcerning, but our measurements lack sufficient sample sizes and we used a mirror image of the new system for system utilization and IMS data.

The second stage proved to be very much like the first, because very little of the scheduled conversion had occurred.

The third stage was the start of a new game. A group was established to officially set a conversion schedule and track its progress. In doing so, a new term appeared that was different than ours. The term contract was replaced with the term policy holder. Technically they are different, which changed our initial NFU measurements. At this point we had to align our NFU measurements with the new terminology, and recalculate conversion factors.

With better sample sizes, an official schedule of NFU's, good functional relationships and conversion factors, we were able to do more reliable modeling. The output of the model still indicated an inordinate amount of system resources needed to process online claims.

This became the center of concern for the business segments, data operations and systems and programming. Capacity planning and performance management suggested a limited amount of online claims processing with the remainder in batch. This was unacceptable to the business segments because they had been promised full online claims processing.

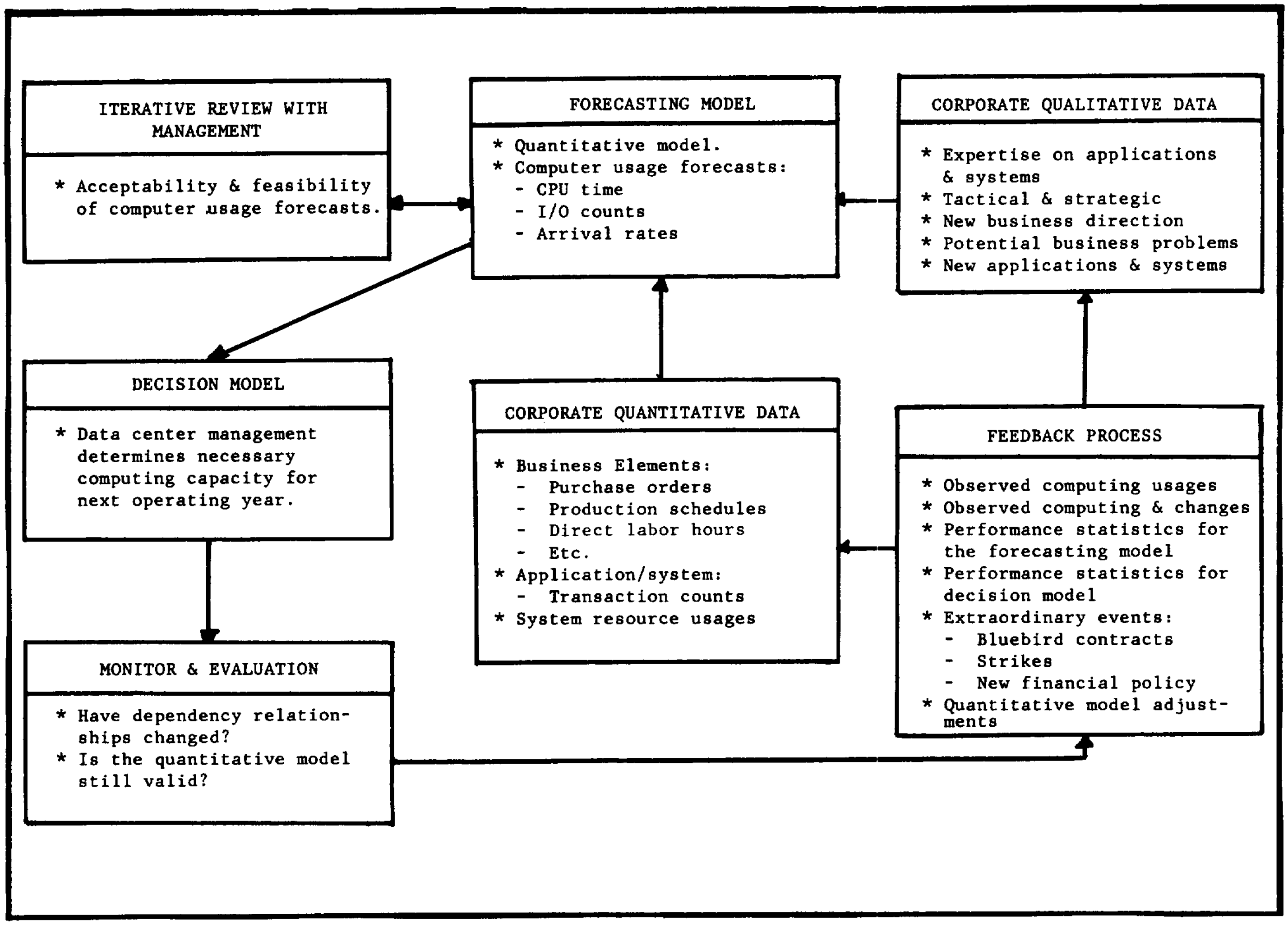
The system designers, data base people and programmers have studied and corrected this performance by a code change for claims processing. Today our projection models are reporting a totally convert system on one side of a 3090/400E with good response times.

## NFU FOR DPU STATISTICAL PROJECTIONS

T. L. Lo [1] made a presentation at the CMG86 conference which gave details about their capacity program and how they statistically projected NFU's to determine DPU's. In summary, their capacity program has the following steps:

* Identify business elements at NFU's and data collection.
* Determine DPU's and their data collection.
* NFU/DPU dependency analysis.
* Forecasting process.
* Determine capacity requirements.
* Iterative review and revision.

The following chart, Lo [1] is a high level flow of their capacity program.




The center piece of their program is the forecasting model and how it interfaces with other managerial functions of the flow. The forecasting model is designed around the dependency of the NFU and DPU, which is usually a linear relationship. In general, the forecasting process involves the following steps:

* Collect the NFU data from a historical source and use smoothing techniques to reduce randomness of the input data while determining projected values.
* Pass the smoothed NFU projected values to a Boole & Babbage program, the Workload Planner. This program produces projected DPU values (CPU times and I/O counts) from a historical source of system data. Such values are generated by a stepwise regression that uses the most affecting NFU's.
* Projected DPU values are used to determine future hardware and software system requirements.
* This process is iterated through the management flow for monitoring and evaluation, and for feedback.

## CONTRASTING THE NFU APPROACHES

The first approach, preparing NFU's for modeling, is practical in smaller business organizations, which are stable and have a limited number of potential NFU's. This approach is particularly good in those system environments where the performance and capacity planners have control over the SRM parameters.

The second approach, projecting NFU's with statistical processes, is used within large business organizations having many segments that dynamically change their organizational structures and strategies. Such organizations potentially have a large number of NFU's that probably do not directly relate to unique workloads.

## REFERENCES

[1] T. L. Lo and J. P. Alias, "Workload Forecasting Using NFU: A Capacity Planner's Perspective". 1986 CMG Conference, Dec. 1986.