VAMC DATABASE ANALYSIS USING SAGG UTILITIES

John C. Kupecki, Senior Computer Specialist
Department of Veterans Affairs
Albany Information Systems Center
Albany, New York

Kornel A. Krechoweckyj, Senior Computer Specialist
Department of Veterans Affairs
Albany Information Systems Center
Newington, Connecticut

ABSTRACT

Various options and utilities have been developed as adjunct support tools to aid support personnel at the Information Systems Centers (ISCs) in utilizing the Statistical Analysis of Global Growth (SAGG) Project database. Through the use of these tools, extensive statistical analyses can be performed for the medical centers. For this paper, SAGG data previously collected from a randomly selected Veterans Administration Medical Center (VAMC) test site, was summarized and compared against averaged SAGG data from sites of similar size and complexity. Areas of comparison focused on total database size, database growth and current disk capacity. Trending studies were used in predicting potential problems and in charting the local growth fluctuations of key Decentralized Hospital Computer Program (DHCP) globals at the site. Emphasis was placed on ascertaining which globals consistently accounted for the major database growth at the VAMC test site, predicting future growth trends, and forecasting potential adverse impacts on Information Resource Management (IRM) resources.

SAGG OVERVIEW

There has been a continued expansion in the size and scope of the Decentralized Hospital Computer Program (DHCP) at the medical centers within the Department of Veterans Affairs (VA). Several years ago, the VA realized that reliable quantitative information had to be obtained in order to accurately track global database growth patterns at these medical centers. In response to this need, the Veterans Health Administration (VHA) developed the Statistical Analysis of Global Growth (SAGG) Project in early 1992.

Developed primarily as a statistical tool that examines global database sizes and efficiencies, SAGG incorporated other key features into the project, such as which files are in use, their current package versions, and the number of file entries. This fully automated MUMPS package regularly monitors DHCP global activity at each medical center with only minimal impact on the computer center's resources. The software is compatible with the current M database systems that are running at the Department of Veterans Administration (VA) medical centers and is easy to manage. Pertinent summary information relating to the captured data is immediately transmitted back to the participating site. The captured information is merged into a centralized SAGG database that utilizes the VA developed FileMan database management system. Subsequently, a variety of statistical analyses are performed and formulated into different reports.

A series of menu options and utilities have been developed as statistical tools for utilizing the SAGG database information captured from the medical centers. It is through the use of these adjunct tools that support personnel at the Information Systems Center (ISC) are better enabled to lend assistance to the medical centers regarding database analysis.

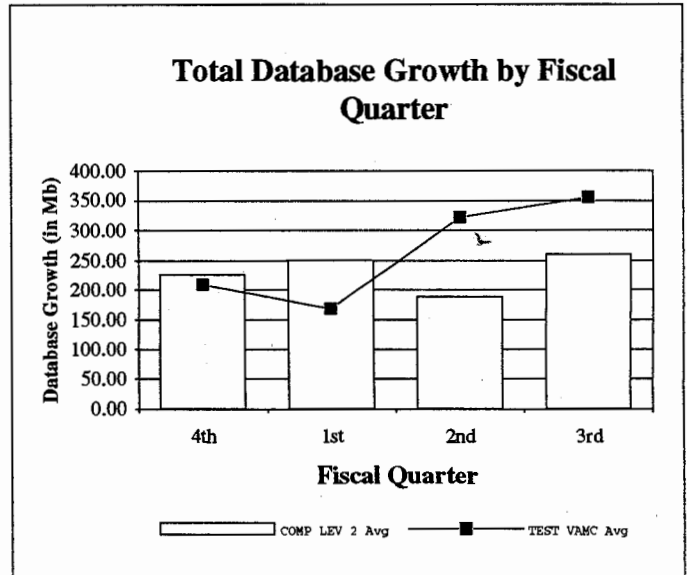**Table 1. Test VAMC vs. Complexity Level 2 Sites Database Size and Growth Summary**

Site Type	Part A. Size Summary (megabytes)				
	3rd Qtr 92 Avg Size	4th Qtr 92 Avg Size	1st Qtr 93 Avg Size	2nd Qtr 93 Avg Size	3rd Qtr 93 Avg Size
Test VAMC	2,808.47	3,017.33	3,185.32	3,506.61	3,861.55
COMP LEV 2	2,737.41	2,963.37	3,214.54	3,402.98	3,663.75

Site Type	Part B. Growth Summary (megabytes)						
	3rd Qtr 92 Avg Size	4th Qtr 92 Growth	1st Qtr 93 Growth	2nd Qtr 93 Growth	3rd Qrt 93 Growth	Total Growth 3rd Qtr 92 3rd Qtr 93	Avg Quarterly Growth
Test VAMC	2,808.47	208.86	167.99	321.29	354.94	1,053.08	263.27
COMP LEV 2	2,737.41	225.96	251.17	188.44	260.77	926.34	231.59

Note: TEST VAMC refers to the test site's database averages

COMP LEV 2 refers to all Complexity Level 2 sites' database averages
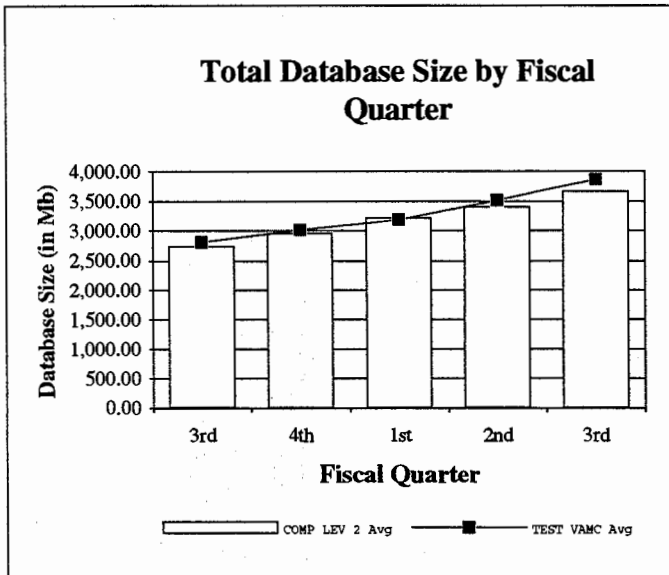
Figure 1. Test VAMC vs. Complexity Level 2 Sites Database Size

Figure 2. Test VAMC vs. Complexity Level 2 Sites Database Growth

ANALYSIS METHODS

The scope of the SAGG statistical utilities is broad and facilitates the examination of the databases at the medical centers. One set of options reviews several areas of an individual medical center and includes global, database, site-specific demographics, file, package, and current disk capacity information.

Another group of options focuses on all sites which have been intentionally grouped into four subsets called Complexity Levels. These Complexity Levels are based on factors which are dependent on the medical center's clinical makeup. By grouping the sites by Complexity Level, these options derive similar information that can be obtained for the individual sites within the subset.

Trending utilities have also been developed which track many aspects of global, package, and database growth over a monthly or yearly time span. Projections on future database and global sizes can calculated. Determinations on the fastest growing globals, and on database and package growth rate fluctuations, can easily be made. Likewise, the effects of purging, archiving, and data compression, can be accessed.

By using these options, a comprehensive analysis was conducted for a randomly selected VAMC in which their SAGG data was examined. For purposes of comparison, similar steps were taken to obtain corresponding values for the appropriate Complexity Level subset. In this case, the VAMC test site belonged to Complexity Level 2. Many aspects of this site's data were reviewed in order to isolate potential problems and to reveal trends. Appropriate steps were examined to reduce adverse impacts.

FINDINGS

Below is a brief summary of these results obtained for this site within its Complexity Level category. This illustrates some of the types of analyses that can be performed.

The initial findings focused on comparing the total database size and global growth patterns for the VAMC test site against Complexity Level 2 averages. The database size and growth summary data shown in Table 1, and graphically depicted in Figure 1, strongly illustrates a steady average increase in overall database size for both the VAMC test site and for all Complexity Level 2 sites. When the VAMC test site is compared against all Complexity Level 2 sites, the test site shows a much greater database growth for the last two fiscal quarters. This is shown in Table 1, Part B, and is further illustrated in Figure 2.

Another study was conducted to compare database growth values for the test site against total disk storage capacity. The same steps were then performed for all Complexity Level 2 medical centers. In this review, the average quarterly database growth values were calculated over the past fiscal year and extrapolated to project data growth over the next three fiscal years. These projected totals were then compared against the global disk space capacity for each grouping. The results from these calculations are summarized in Table 2. This analysis indicates that by the end of the 3rd quarter of FY96, the global disk storage capacity for the VAMC test site, as compared with other Complexity Level 2 sites, will be utilized as follows:

VAMC Test Site: 65%
Complexity Level 2 Sites: 68%

It should be noted that the above projections are based on the average quarterly data growth values derived from collected SAGG Project data from all participating Complexity Level 2 medical centers and current total disk space capacity information. It also assumes that the same global growth rates will be generated from the current DHCP packages. The average quarterly growth rate can vary between quarters because of a variety of situations such as package usage, newly released and/or mandated packages, purging and archiving, and the total number of sites participating in the SAGG Project for this Complexity Level.

Table 2. Projected Database Growth (in gigabytes)

COMPLEXITY LEVEL	CURRENT DATABASE SIZE June 1993 *	AVERAGE QUARTERLY GROWTH	PROJECTED DATABASE SIZE 3rd Q FY94	PROJECTED DATABASE SIZE 3rd Q FY95	PROJECTED DATABASE SIZE 3rd Q FY96	GLOBAL DISK CAPACITY June 1993 **
TEST VAMC	3.86	0.26	4.90	5.94	6.98	10.7
COMP LEV 2	3.66	0.23	4.59	5.52	6.45	9.43

* Value for TEST VAMC was the current database size

Value for Complexity Level 2 sites was the averaged current database size

** Value for TEST VAMC was the current global disk capacity

Value for Complexity Level 2 sites was the averaged global disk capacity

Another aspect of the analysis centered on ascertaining which globals may have been pivotal for the sudden growth rate increase at the test site in the last two fiscal quarters. Results revealed nine rapidly growing globals that span several key DHCP packages as the major contributors. The examination was restricted to identifying globals with growth rates of approximately 5,000 DSM (Digital Standard MUMPS) blocks or more for all months within the last two fiscal quarters of this study. Table 3 summarizes these findings. Not surprising, many of these globals belong to the fastest growing DHCP packages. The top nine fastest growing globals for the VAMC test site are presented here in package alphabetical order. They are: Accounts Receivable (^PRCA), IFCAP-Fiscal (^PRC and ^PRCF), Integrating Billing (^IB), Laboratory (^LR and ^LRO), MailMan (^XMB), and Pharmacy (^PS and ^PSRX).

Of the VAMC test site's nine fastest growing globals over the last two fiscal quarters, it appears that purging and/or archiving was limited to only several occasions for the ^LRO, ^PRCF, and ^PS globals. As illustrated in Figure 3, the ^LRO, and ^PRCF globals displayed negative growth during January 93, April 93, and June 93, indicating a size reduction in these globals. The ^PS global only displays a size reduction in January 93. For

the remaining six globals, purging and/or archiving was either minimally accomplished or absent. Figure 3 clearly shows the growth rate of these nine globals over a six month period and indicates the beneficial effects of purging/archiving during the January, April, and June time frames. These global growth rate fluctuations can be clearly seen in Figure 4 which illustrates the changes in a 3D scheme.

In Figure 5, the sizes of the nine globals in June 93 were combined to arrive at the percentage that each contributes to the total DHCP database size. The results show that these nine globals comprise over one half of the VAMC test site's entire database. This illustrates how critical these nine globals are in comprising the overall database at the VAMC test site.

It is clear that more emphasis should be placed on reducing the size of these six globals as an aid in database management.

CONCLUSION

The test site's average database growth rate for the last two fiscal quarters was markedly increased when compared against their previous two fiscal quarters and

Table 3. VAMC Test Site's Fastest Growing Globals

Growth Summary (in DSM blocks)								
Global	Size	FY93 2nd Quarter			FY93 3rd Quarter			Size
	Jan-93	Jan-93	Feb-93	Mar-93	Apr-93	May-93	Jun-93	Jun-93
IB	99,354	6,322	6,888	6,779	7,219	7,423	7,344	141,329
LR	227,204	4,900	5,439	6,187	5,837	6,035	5,822	261,424
LRO	138,097	-19,673	17,296	19,409	-27,395	19,371	-7,940	139,165
PRC	166,351	3,897	4,441	4,491	4,855	5,418	4,472	193,925
PRCA	152,681	8,574	12,173	13,343	11,823	12,682	11,890	223,166
PRCF	30,056	226	929	6,028	-16,940	7,154	-4,911	22,542
PS	184,400	-9,580	8,931	10,834	10,409	10,734	10,388	226,116
PSRX	145,397	6,136	6,467	6,491	6,714	7,076	6,692	184,973
XMB	383,316	6,319	16,441	28,501	33,126	22,925	38,480	529,108
All Others	1,624,179	30,540	49,191	61,078	52,706	49,691	54,363	1,921,748
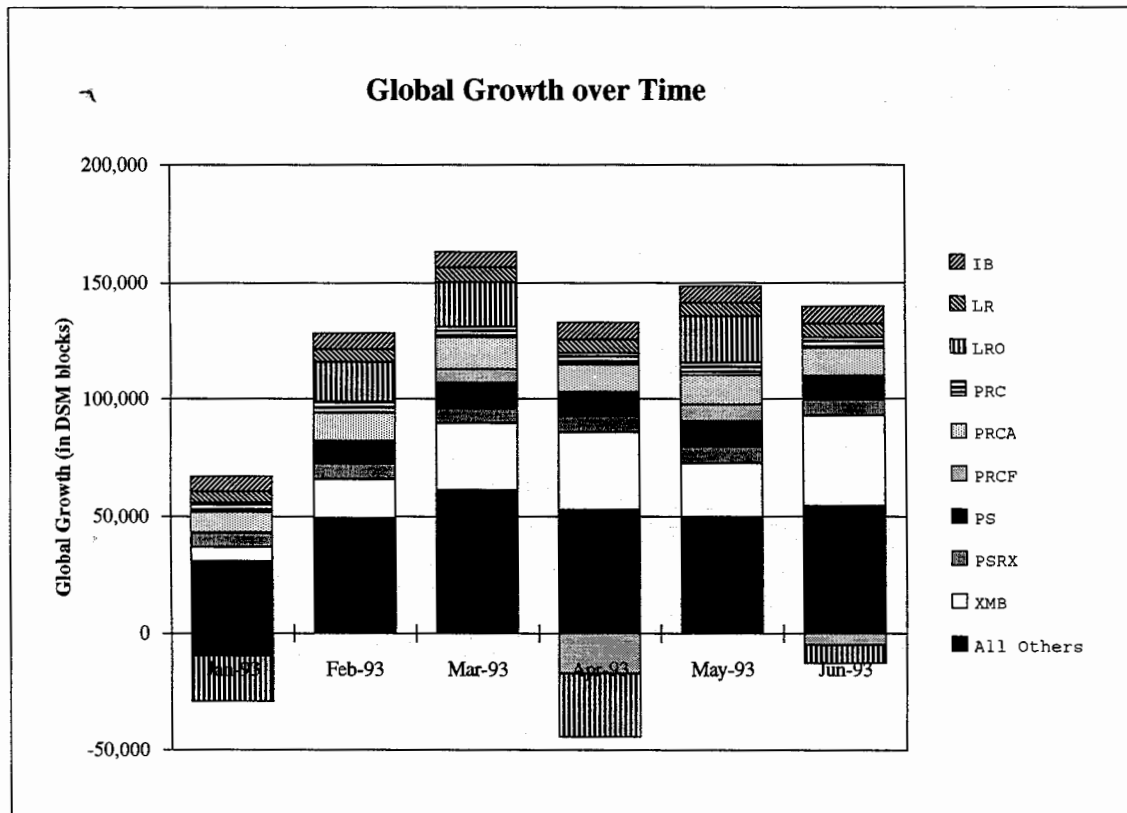
Figure 3. Growth of Selected Globals at VAMC Test Site

Global Growth (in DSM blocks)
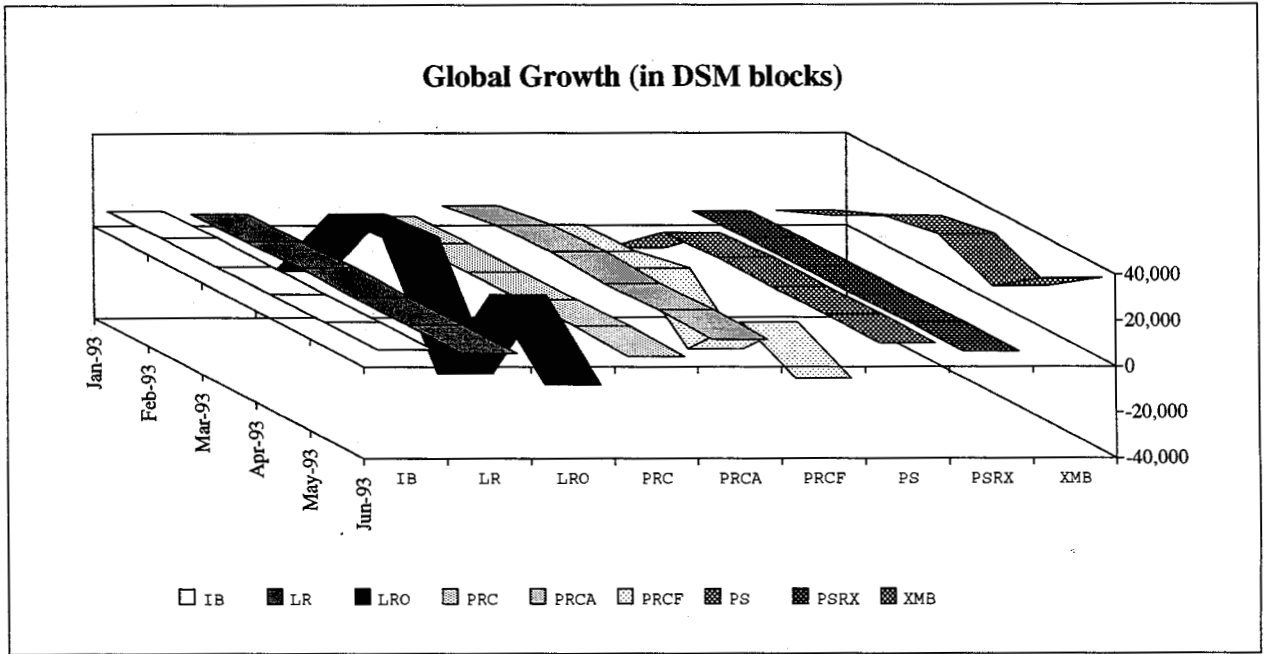
Figure 4. Growth of Selected Globals at VAMC Test Site

Global Percentages for June 93
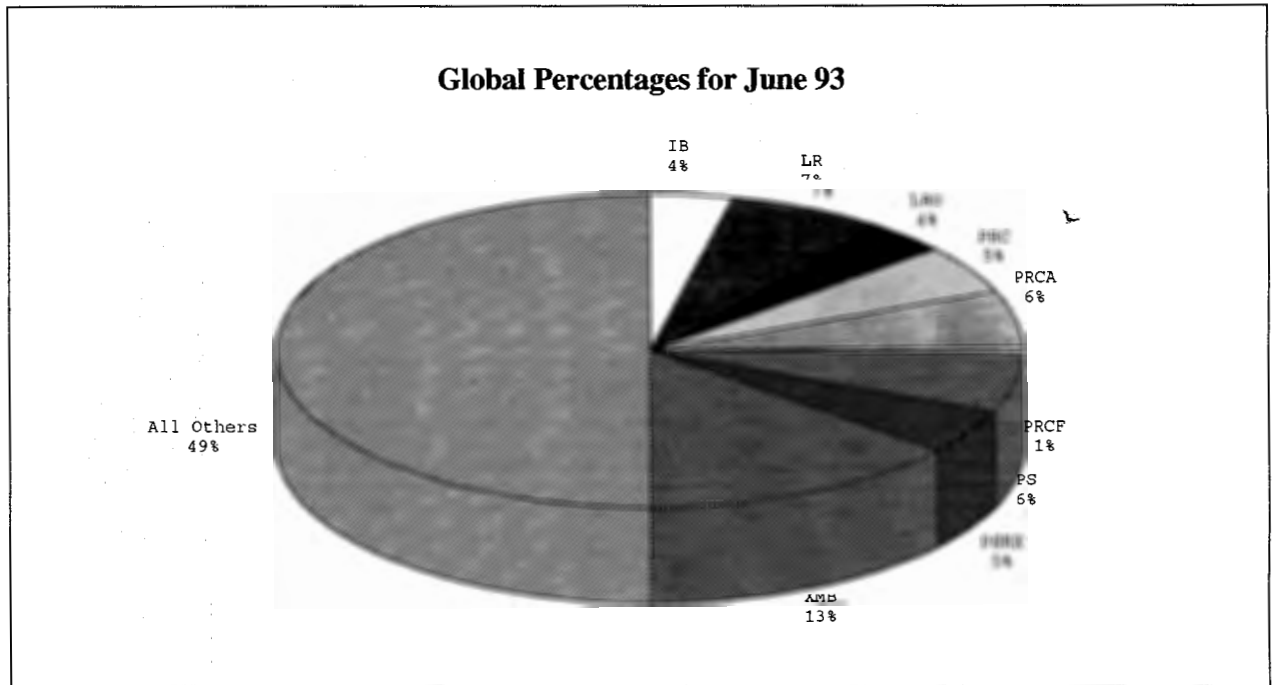
Figure 5. Global Size Percentage at VAMC Test Site

against the average for all Complexity Level 2 sites. A continued growth at this site could potentially lead to future problems. Many times, additional data on a computer system translates to additional overhead involving IRM resources. Assuming that the average database growth at the VAMC test site continues at the same rate, the test site will have exhausted approximately 65% of the available global disk storage by the 3rd quarter of FY96. Continued growth at the current rate of projection will exhaust all of the available global disk storage sometime in FY99. Factors such as a continued expansion of DHCP and increased hospital activities will necessitate the reevaluation of these findings for this VAMC test site.

The test site's average database growth rate for the last two fiscal quarters was markedly increased when compared against their previous two fiscal quarters and against the average for all Complexity Level 2 sites. Results revealed that nine globals were the major contributors to the database increase. These globals were found to comprise slightly more than half of the site's entire database size. Of the VAMC test site's nine fastest growing globals over the last two fiscal quarters, it appears that purging and/or archiving was limited to only several occasions for the ^LRO, ^PRCF, and ^PS globals. For the remaining six globals, purging and/or archiving was either minimally accomplished or absent. In the DHCP packages that contain these remaining globals, most do offer some form of purging/archiving capabilities. As a first step, it is clear that more emphasis should be placed on reducing the size of these six globals.

This aspect of database management will help decrease the size of the database and limit some of the potential effects of increased system overhead. Although the test site is not in danger of exhausting their available global disk storage in the near future, careful attention to the current situation will prevent future disk capacity storage problems.