Ideas International Evaluation of IBM's Commercial Processing Workload (CPW) for Rating Relative Performance of IBM eServer iSeries Systems

June 2005

PREPARED FOR

IBM

TABLE OF CONTENTS

Ideas International (IDEAS) has spent more than 12 years evaluating the performance of computer systems and comparing the performance of one system with another.[1] IDEAS is also the only analyst firm that is an associate member of the Transaction Processing Performance Council (TPC), the organization that owns the TPC-C benchmark. Our long experience in tracking system performance has made us well aware of the strengths and weaknesses of the various benchmarking methodologies that are available in the industry.

IDEAS has developed its own metric for measuring systems performance called the Relative Performance Estimate or RPE. RPE covers a broad range of systems including Windows and UNIX. The RPE methodology also takes publicly available benchmarks into account. Specifically, to create the RPE values for IBM's iSeries, we start with IBM's own CPW numbers and then adjust them based on other benchmarks as well as our own analysis. We believe that our adjustments take into account the fact that IBM is conservative in its CPW data. We then use RPE in our products to help our customers assess the performance of competing systems.

This paper provides CPW background information and discusses the steps IBM has taken to make the measurements more reflective of today's customer production environments. In it, we clarify the differences between CPW and the various public benchmarks. Finally, we explain why we have confidence in the CPW data and describe how that data compares with our RPE data.

What Is CPW?

For the IBM eServer iSeries, one of the principal performance metrics is IBM's Commercial Performance Workload (CPW) rating system or benchmark. CPW measures the relative computing power and associated software capabilities of these systems in a commercial environment. The CPW rating of a system is generated using the measurements of a specific workload that is maintained internally within the iSeries Systems Performance group. This workload is rigidly defined for functionality, performance metrics, and price/performance metrics.

As experts in benchmarking and measurement, IDEAS believes that IBM's use of CPW provides a greater opportunity to accurately reflect relative performance than can be found in public benchmark results. A public benchmark's goal is to demonstrate the absolute highest system performance possible, whereas the primary goal with CPW is to show the relative performance between systems.

HISTORY OF RELATIVE PERFORMANCE MEASUREMENT ON iSERIES

In order to better understand why CPW provides such a viable measure on IBM eServer iSeries systems and software performance, it is important to discuss its heritage. The measurement of relative performance on the IBM eServer iSeries family of systems is directly related its successful predecessor midrange systems, the IBM AS/400, S/36, and S/38. These systems were originally measured and compared using the IBM internal benchmark RAMP-C (Requirements Approach to Measuring Performance – in COBOL).

RAMP-C measurements were used to compute relative performance ratings by comparing measured results of new systems with a baseline system. RAMP-C was tightly regulated within IBM to ensure that measurements from release to release, system to system, and even architecture to architecture were comparable.

Because of its original OLTP-focused design, RAMP-C proved to be an excellent measure of applications up to the early 1990s. As online transaction processing applications increased in complexity and sophistication, however, it became clear that a new, more complex workload design would need to replace RAMP-C.

After TPC-C was released, IBM determined that it would continue to use an internally controlled measure of relative commercial performance. IBM decided that the TPC-C benchmark would not be suitable for capacity planning (for the reasons discussed in the section "Advantages of CPW over Public Benchmarks"). However, IBM combined many of the attractive characteristics of the new TPC-C benchmark together with the consistency rules from RAMP-C to produce a new measurement, the Commercial Processing Workload or CPW, which debuted in 1996. Since that time, CPW has been used to provide relative ratings for the AS/400 and iSeries systems.

CPW measures a range of database applications, including simple and moderately complex updates, simple and moderately complex inquiries, realistic user interfaces, and a combination of interactive and batch activities. CPW tests the ability of the system to provide concurrent access to large numbers of users running a single group of programs.

CPW also provides a measurement of commitment control. Commitment control is an extension to the journaling function that enables users to ensure that all changes to a transaction are either complete or, if not complete, can be easily backed out. This is an important test as the use of commitment control adds additional journal entries to committed transactions resulting in additional CPU and I/O overhead. In addition, the amount of time that record level locks are held increases with the use of commitment control. Because of this additional overhead and possible additional record lock contention, adding commitment control will in many cases result in a noticeable degradation in performance for an application that is currently doing journaling.

CPW Application Description

The CPW application simulates the database server of an online transaction processing (OLTP) environment. Requests for transactions are received from an outside source and processed by application service jobs on the database server. CPW is based, in part, on the business model from benchmarks owned and managed by the Transaction Processing Performance Council. However, there are substantive differences between this workload and public benchmarks that preclude drawing any correlation between them. The net result is an application that we believe provides an excellent indicator of transaction processing performance capacity in IBM's eServer iSeries product family.

Five business functions of varying complexity are simulated. These transactions are all executed by batch server jobs, although they could easily represent the type of transactions that might be performed interactively in a customer environment. Each of the transactions interacts with three to eight of the nine database files that are defined for the workload. Database functions and file sizes vary. Functions exercised are single and multiple row retrieval, single and multiple row insert, single row update, single row delete, journal, and commitment control. These operations are executed against files that vary from hundreds of rows to hundreds of millions of rows. Some files have multiple indexes, some only one. Some accesses are to the actual data and some take advantage of advanced functions such as index-only access.

Advantages of CPW over Public Benchmarks for Relative Ratings

We mentioned earlier that there were substantive differences between the CPW workload and the workloads of public benchmarks. Some of the key differences between CPW and public benchmarks are:

» Public benchmarks allow optimization as long as the external requirements are met. This means that one cannot separate the improvements that are attributable to the platform and those that result from tuning. IBM has tried to eliminate this effect in CPW in order to better represent relative improvements from release to release and system to system.

» Public benchmarks typically do not require full security, but many IBM customers tend to run on secure systems; consequently, a higher level of security is specified for the CPW workload, which translates to more realistic workloads being measured.

» Public benchmarks are super-tuned to obtain the best possible results for that specific benchmark. CPW tends to use more of the system defaults to better represent the way the system will actually be used by customers.

» Public benchmarks often use different applications for different sized systems and take advantage of all of the resources available on a particular system. CPW has been designed to run as the same application at all levels with approximately the same disk and memory resources per simulated user on all systems, enabling a true "apples to apples" comparison to be performed.

» Public benchmarks tend to stress extreme levels of scaling at very high CPU utilizations for very limited applications. To avoid misrepresenting the capacity of larger systems, CPW is conservatively measured at approximately 70% CPU utilization.

» Public benchmarks require extensive, sophisticated driver and middle tier configurations. In order to simplify the environment and add a small computational component to the workload, CPW is driven by a batch driver that is included as a part of the overall workload.

» A corollary to this discussion is the ability for any benchmark to relate to the customer environment, over time. During the 13 years TPC-C has been active, the analysts and designers running the benchmark have shortened the overall path of the benchmark to a fraction of what the first implementations experienced in 1992. At the same time, computing capacity has mushroomed and consumers have taken advantage of that capacity by building evermore-complex applications. The result is a divergence that has caused a benchmark that began as an accurate reflection of reality to stray well away from that path. So, while TPC-C continues to be the most robust, well controlled, transaction processing benchmark in the industry, its use and purpose is not always in line with the way clients plan to actually use the systems being measured.

Again, utilizing CPW, the net result is a measurement that IDEAS believes provides an excellent indicator of transaction processing performance capacity on eServer iSeries systems.

Why IDEAS Is Confident CPW Provides a Viable Measure of Performance

Of course, IBM has also improved CPW over the years. However, instead of focusing on reducing code paths to improve vendor "bragging rights" over published benchmarking results, IBM has enhanced CPW with new functions that are likely to be found in today's typical customer environments. IBM has documented the following changes to CPW:

» Robust security enforcement

» System Managed Access Path Protection (SMAPP)

» Redundant disk protection (RAIDx) on all database disks

» Use of default tuning settings instead of benchmark "hyper tuning"

» Use of enhanced journal/logging features

» Adjustments to the ratio of processor use and disk I/O to more accurately reflect customer reality

» Use of a consistent amount of memory per task, regardless of the amount configurable on the system

» Additional processing of information within transactions

» Additional background processing

» Execution at a processing level that is more representative of the customer environment

» Inclusion of two-phase commitment control for a fraction of transactions

IDEAS believes that all of these changes make the CPW measurement more closely resemble real-world production.

Whenever alterations were made to the internals of the workload, IBM conducted extensive bridging measurements to ensure that the relative ratings continued to be accurate across the entire family of AS/400 and iSeries systems.

To gage if CPW is accurately representing customer workloads, IBM has assembled one of the largest collections of consistent customer performance information of any platform in the industry. IBM has performance data from thousands of customers who participate in Performance Management (PM) eServer iSeries (http://www-1.ibm.com/servers/eserver/iseries/pm/). With this performance management service that submits data to IBM, customers have an easy means to track their resource usage and predict growth. In addition, IBM has an extensive archive of data to help ensure that internally measured workloads accurately reflect real-world workload characteristics.

Furthermore, CPW ratings also play a prominent role in IBM's own sizing tool for iSeries systems, the IBM eServer Workload Estimator (http://www.ibm.com/eserver/iseries/support/estimator). This tool is available to customers as well as IBM Business Partners, and IBM estimates that it is used for over 1000 sizing estimates each week.

Where to Find Additional CPW Information

Information on CPW can be obtained from a number of different places, but the most current source is IBM's iSeries Performance Capabilities Reference Manual, which can be found at the iSeries Performance Management website: http://www-1.ibm.com/servers/eserver/iseries/perfmgmt/resource.htm.

IDEAS Relative Performance Estimates

The IDEAS OLTP Relative Performance Estimate (RPE) is a cross-platform numerical ranking methodology used for expressing vendor intra-range server performance in a normalized manner.

RPE is derived by mapping vendor in-series performance estimates with known benchmark performance results to provide a more comprehensive view of relative positioning of competing products. This view allows a user to gain an understanding of the relative positioning of system configurations that may not have been tested in a public benchmark.

IDEAS has maintained the RPE index since the mid 1990s and also publishes a monthly updated spreadsheet tool that can be used to apply historic and current RPE data to server consolidation projects.

In Q305, we plan to replace the current RPE methodology with a more advanced multi-component metric called RPE2, which will still include the current RPE data as one of the components. We also plan to continue to use CPW as the basis for our ratings of the eServer i5 models.

The table below shows the current RPE values for a selection of i5 e-server models. IDEAS considers the IBM CPW index to be a very reliable representation of intra-model relationships and scale-up factors within individual models, and therefore does not need to adjust the data within the RPE metric. Hence, scaling relationships using RPE or CPW will be identical.

The sample data points in the table are intended to illustrate:

» Relationships between the reference iSeries 840-12 banking system and the proposed i570 server (1:2.23)

» The maximum growth path within the proposed i570 model (1:1.6)

» The current overall performance spectrum of i5 Series from the entry i520 to the high-end i595 (1:330)

	Processors	RPE	CPW
iSeries 840 600 MHz - 16 MB (12)	12	12674	12000
AS/400e 840 500 MHz - 8 MB (24)	24	17426	16500
eServer i5 0520 (4U) 1.5 GHz - 1.9 MB (1)	1	528	500
eServer i5 0570 (12U) 1.65 GHz - 36 MB (9)	9	28305	26800
eServer i5 0570 (16U) 1.65 GHz - 36 MB(16)	16	46471	44000
eServer i5 0595 1.65 GHz - 36 MB (64)	64	174268	165000

Summary

We have discussed IBM's CPW and given our assessment of its utility. IDEAS believes that CPW is a valid metric for performance estimates. As stated, we use CPW in our own products. Our reasons can be summarized as follows:

» There are thousands of CPW data points for AS/400 and iSeries systems in IBM's database generated by customers who have used CPW over the years for capacity planning. CPW has a long history.

» IBM has made continual efforts to make CPW reflective of real world requirements and to keep it up to date.

» CPW data is based on measurements. IBM measures the points that it documents.

» In our judgment, the CPW measurement methodology is conservative.

» Finally, IBM has documented the CPW methodology and it is very extensive and easy to understand.

---

1 See our website for a complete description of our offerings: www.ideasinternational.com.

Ideas International, Inc.
(Formerly D.H. Brown Associates, Inc.)
222 Grace Church Street
Port Chester, NY 10552
Tel + 1 914 937 4302
Fax +1 914 937 2485
www.ideasinternational.com