Performance Benchmarks for Blackbaud CRM

Executive Summary

Blackbaud CRM is built to scale out to store millions of records and handle thousands of application users. To get the most out of the application, customers must allocate sufficient physical resources to their systems to ensure that ***Blackbaud CRM*** can achieve optimal performance. To lay the groundwork for successful implementations, Blackbaud conducted performance tests to evaluate the system requirements for different types of customers. The tests help to determine how to size their systems to achieve optimal performance results. This paper sets a framework for the factors to consider when planning for new ***Blackbaud CRM*** deployments and upgrades.

The performance guidelines are based on performance tests against various configurations of ***Blackbaud CRM*** with realistic workloads modeled after usage patterns at client sites. The tests adjusted system resources such as CPU cores, memory, disks, and network bandwidth to measure how variations in the physical configuration impact the responsiveness of the system. This paper provides detailed information about those tests and illustrates how variations in the physical resources affect the performance of ***Blackbaud CRM***. At a high level, the findings from the performance testing can be summarized with the performance guidelines and best practices in this section.

A ***Blackbaud CRM*** deployment contains multiple servers: The database server, web servers, reporting server, and data warehouse server are common components. Although each part of the system requires adequate configuration, experience shows that performance limitations are almost always caused by the database server. This paper focuses exclusively on the sizing of the database server.

Following the recommendations in this paper, and understanding how they are derived from the system testing, will make it very likely that your system provides satisfactory performance. Still, every deployment is unique, and therefore, no guarantees are given that a system thus configured will meet your expectations.

Recommendation for Memory

Sufficient memory is probably the single most important consideration for system performance.

To calculate the minimum amount of memory that you need, Blackbaud recommends a simple rule of thumb: Add up the space used by seven key tables (ADDRESS, CONSTITUENT, EMAILADDRESS, FINANCIALTRANSACTION, FINANCIALTRANSACTIONLINEITEM, JOURNALENTRY, and PHONE) that the system generally keeps in memory, including both the table data and the indexes on the table. Then adjust the estimate based on expectations for future growth. The result is a rough estimate of your minimum memory requirement. To calculate this rough estimate of the amount of memory your system needs, fill in your row counts in the following worksheet, multiply by the given bytes per row, and add. Then, because memory is so vital, Blackbaud recommends that you double your computed estimate. For more information, see the Varying Memory section.

Table Name	BytesPerRow		Number Rows		Space in Bytes
ADDRESS	1200	X		=	
CONSTITUENT	1100	X		=	
EMAILADDRESS	650	x		=	
FINANCIALTRANSACTION	1600	x		=	
FINANCIALTRANSACTIONLINEITEM	900	x		=	
JOURNALENTRY	1300	x		=	
PHONE	800	x		=	
Minimum memory					
Recommended memory	Min. Memory	x	2	=	

Note that these tables were chosen because the row counts are likely to be values that can be estimated prior to deployment. If you know your organization and its business processes and consult the existing data to migrate to ***Blackbaud CRM***, then you should be able to estimate these values. Be sure to allow for forecasted growth when you make the estimates.

Recommendation for Logical Processors

Performance testing shows that the need for CPU cores grows linearly with the number of users who concurrently access the system and that the need for CPU cores during nighttime workloads is limited by the serialization of the SQL statements that are executed.

For the work presented in this paper, hyperthreading was not used, so the terms "core" and "logical processor" are used interchangeably.

The optimal number of users per CPU core can vary for different organizations based on how active users are, but Blackbaud recommends the following rule of thumb to compute how many CPU cores are necessary for your organization: Take your highest anticipated number of concurrent users (remember to allow for growth), divide by 30, add 2 to the result to allow for a buffer as the upper limit approaches, and then round up to a multiple of 4. Remember that concurrent users are the users who access ***Blackbaud CRM*** at the same time, not all application users with accounts in the system.

$$\text{NumberCores} = ceiling(\frac{MaxConcurrentUsers}{30} + 2, 4)$$

The result is the number of cores that Blackbaud recommends. The following chart provides examples of the minimum requirement for CPU cores for different numbers of concurrent users.

Number of Users	Minimum Cores
1 to 60	4
61 to 180	8
181 to 300	12

301 to 420	16
421 to 540	20
541+	Use Formula

For more information, see the Varying CPU Cores section.

Recommendation for Disk Input/Output

Performance testing shows that average disk input/output is not a reliable way to anticipate system needs because the peak I/O bandwidth can be many times the average. Instead of planning for I/O based on the average, Blackbaud recommends that you plan for 50 times the average.

Since you don't necessarily know the average I/O in advance, Blackbaud recommends a rule of thumb calculation that is based on the overall size of your database. The peak data rate for disk I/O is about 0.0004 times the data space. (Not index space or reserved space, just data space.) To calculate the peak data rate, fill in the following worksheet with the database size and multiply by the conversion value.

DB Data size MB		Conversion		Max data rate, MB/sec
	X	0.0004	=	

The database size should be space-allocated to ***SQL Server***, which may be different from the size of the physical drives that support the database. For an existing installation, your database administrator will know the size of the database data space. For a new installation, work with your Blackbaud representatives to determine the size to plan for. Use the resulting value to talk to your system administrator or SAN administrator to be sure that sufficient I/O bandwidth is available.

For more information, see the Disk Input/Output section.

Recommendation for SSD Usage

Solid State Disks can have important benefits in a modern server system, but you need to understand the benefits to make a good cost/benefit tradeoff. For ***Blackbaud CRM***:

- System responsiveness generally sees a small benefit from SSD storage.
- When memory is constrained, the benefits of SSDs are greater. SSDs can delay the point when performance becomes unacceptable. Sufficient memory is preferred, but SSDs may be less expensive than memory upgrades. Or they may be an option when memory upgrades are not possible.
- The optimal configuration may be to place the Data drive on SSD storage and the Log drive and TempDB on spinning disks.

For more information, see the Spinning Disks vs. Solid State Disks section.

Recommendation for Network Bandwidth and Latency

1 GB/sec network interfaces are commonplace and are sufficient for ***Blackbaud CRM*** usage. There isn't really a reason to use anything else.

Consider the possible effects resulting from the physical distance between your users and the ***Blackbaud CRM*** servers that support them. Pages that require many round-trips to display can be slowed by network latency, no matter how fast the server is. If your users are not located near the servers, see the Network Bandwidth and Latency section.

Observed Response Times Based on Performance Recommendations

To observe the impact of the performance recommendations in this paper on end user response times, a test environment implemented the recommendations. In general, sub-second response times are desired for most user actions in ***Blackbaud CRM***, and the test environment achieved this benchmark. For details about the end user response times, see the End User Response Time Observations section.

Recommendation for System Monitoring

Following the recommendations in this paper should lead to a system configuration that meets your initial needs and provides good performance. But as time passes, your needs may change. Blackbaud recommends routinely collecting, and regularly checking, a small number of performance counters to monitor changing load on the system and be proactive in responding. Many excellent articles are available about performance counters and what to monitor. At minimum, Blackbaud suggests that you monitor items in the following list. (Shown as Object / Instance / Counter.)

- Processor / _Total / % Processor Time

  If the average often rises over 50 percent, then it's time for a CPU upgrade.

- SQLServer:Buffer Manager / / Page life expectancy

  If this falls below 300, then it's time for a memory upgrade.

- Network Interface / <your active interface> / Bytes Received / sec
- Network Interface / <your active interface> / Bytes Received / sec

  If the average for these often rises over 50 percent of capacity, then it's time for a network upgrade.

- PhysicalDisk / <your Data, Log and Temp drives> / Avg. Disk Bytes/sec
- PhysicalDisk / <your Data, Log and Temp drives> / Avg. Disk sec/Read
- PhysicalDisk / <your Data, Log and Temp drives> / Avg. Disk sec/Write

  If the average I/O rate often increases to more than 50 percent of capacity or if the latency starts to exceed a few milliseconds, then it's time for a storage upgrade.

Table of Contents

Overview

This paper draws its performance guidance from the results of performance tests against ***Blackbaud CRM*** workloads that mirror the average daily usage patterns that were observed from actual clients during their busy season. The first set of workloads is modeled after activity at a large cause organization, and the second set of workloads is modeled after activity at a midsize higher education foundation. Baselines were established for these workloads, and then the physical configuration of the system was altered to test the impact of system resources such as CPU cores and memory on the responsiveness of the system. For more information about the testing methodology, the system configuration of the test environment, and the workloads themselves, see the appendices at the end of the paper.

The sections in the main body of the paper focus on the results from the performance tests. First, the paper looks at the impact on user response times when the load on the baseline workloads is increased. Then it examines the results from running tests against ***Blackbaud CRM*** environments with varying amounts of memory available to ***SQL Server***. Next, it looks at the impact on performance when tests run with varying numbers of logical processors in the system. Then it examines the impact of using spinning disks vs. solid state disks. The paper also includes sections about network bandwidth latency, interactive workloads and business processes, and end user response times.

Varying Loads

This section provides an overview of results from running tests against ***Blackbaud CRM*** environments with varying loads placed on the system. For daytime workloads, the tests started with baseline workloads and then increased the number of users linearly, which increased the activity in the system. Nighttime workloads are not dependent on the number of users; those run through a set list of tasks.

For tests on the large cause organization workloads, the load scaled up to 3 times the normal load. For tests on the midsize higher education foundation workloads, the load scaled up to 6.5 times the normal load. The difference between the test results for the baseline workloads and the workloads with increased loads illustrates the impact of heavier loads on ***Blackbaud CRM*** environments. The tests measured the impact on user response times, CPU usage, and disk input/output.

Key Takeaways

The key takeaways from this series of performance tests include:

- CPU usage grows linearly as the number of active users increases. So the optimal number of logical processors for a ***Blackbaud CRM*** environment depends on how many users concurrently access the system. For information about how to estimate the optimal number of users per CPU core, see Varying CPU Cores.
- Response times follow Little's Law, which indicates that you should limit the load on critical resources to 50-60 percent.

- For nighttime workloads that primarily handle business processes, serialization limits CPU usage. This means that increasing the number of cores won't speed up the workloads.
- Disk input/output experiences significant bursts above and below the average I/O. Planning for I/O based on the average is not a reliable way to allocate resources. Instead, plan for the peaks.
- For the Midsize Higher Education Foundation Daytime Workload, TempDB usage increased as user load increased. And for both night workloads, significant TempDB usage occurred because the business processes commonly required temporary data storage. It is important to allocate sufficient resources to make sure TempDB can handle these scenarios.

For more information about the test results that led to these conclusions, see Appendix D.

Varying Memory

This section provides an overview of results from running tests against ***Blackbaud CRM*** environments with varying amounts of memory available to ***SQL Server***. The tests started with a baseline of 95 GB for the database and then reduced the amount of memory available to the ***SQL Server*** instance by adjusting the maximum memory setting. This setting has the same effect as if the server had less memory available.

The workloads were run with multiple memory settings, and the number of users was increased linearly for each memory setting. For tests on the large cause organization workloads, the memory settings included 60 GB, 40 GB, and 20 GB. The tests started with the baseline workload of 1.0 times the normal number of users, and the load increased linearly by a factor of 0.25 until it reached 3.0 times the normal number of users. For tests on the midsize higher education foundation workloads, the memory settings included 20 GB, 15 GB, and 10 GB. These tests started with the baseline workload of 1.0 times the normal number of users, and the load increased linearly by a factor of 0.5 until it reached 6.5 times the normal number of users.

The test results indicate the vital importance of sufficient memory in the system. When you do not have enough memory for the database to hold the needed data, the system must repeatedly fetch the data from disk storage, which is substantially slower. Thus it's vital to ensure that sufficient memory is available to support the desired operations on the system. How much is enough? For the large cause organization, 40 GB was sufficient for the daytime workloads, while 60 GB might be preferable for the nighttime workload. Meanwhile, 15 GB supported the daytime workloads for the midsize higher education foundation, while 20 GB might be better for the nighttime workload. Appendix E describes how those numbers were arrived at.

How can you know in advance how much memory to place in a new system, without the opportunity to take measurements? Based on the performance test results, Blackbaud can suggest a rule of thumb to help determine the minimum memory requirement for a ***Blackbaud CRM*** system. It works on the assumption that the system generally wants to keep certain key tables and their indexes in memory. Based on the size of those tables, you can compute a minimum memory estimate. While additional data is needed at times and not all indexes on all the key tables are needed, the calculation provides a useful starting point. For the customers that were the basis for the workloads in this paper, it came within about 30 percent of the minimum observed in the runs. Is two sites sufficient to conclude that this is a good rule of thumb? That is simply all the available evidence at this time. Considering how vital it is to have sufficient memory, Blackbaud recommends that you double the computed estimate anyway.

The rule of thumb works like this: Add up the space that seven key tables use both for data and for indexes on the tables. This provides *an estimate* of the minimum memory you need to run a ***Blackbaud CRM*** system. The tables are ADDRESS, CONSTITUENT, EMAILADDRESS, FINANCIALTRANSACTION, FINANCIALTRANSACTIONLINEITEM, JOURNALENTRY, and PHONE.

Of course that leads to the question: How big will each table be? You can estimate the size of each table based on estimates of certain basic entities, such as the number of constituents and the number of transactions. To estimate the size of the tables, multiply the number of those basic entities by the number of bytes per row of data.

In making these estimates, keep in mind that the system doesn't just track, for example, the number of constituent addresses stored. It also maintains a record of prior addresses in the table. And some constituents have multiple addresses, so the number of addresses might be 1.5 times the number of constituents – or some other number for your site. If you are moving from a previous system, many of the values should be easy to obtain from that system. And of course, remember that beyond the current row counts, your estimate should account for likely growth over time.

Table Name	BytesPerRow		Number Rows		Space in Bytes
ADDRESS	1200	X		=	
CONSTITUENT	1100	X		=	
EMAILADDRESS	650	x		=	
FINANCIALTRANSACTION	1600	x		=	
FINANCIALTRANSACTIONLINEITEM	900	x		=	
JOURNALENTRY	1300	x		=	
PHONE	800	x		=	
Minimum memory					
Recommended memory	Min. Memory	x	2	=	

So that's it! Fill in your row counts, multiply by the given bytes per row, and add. Now you have a rough estimate of the amount of memory the system needs. As previously noted, feel free to inflate the number – it's even encouraged because memory is so incredibly vital to good system performance. Blackbaud recommends that you double your computed estimate. Remember that sufficient memory is probably the single most important consideration from a system performance point of view.

After your system is in operation, you should routinely monitor memory on the database server. Use your tool of preference to capture system performance counters, including the ones for memory usage. The Page Life Expectancy counter for ***SQL Server*** is a useful tool to detect when memory demand is too high. A number of articles about ***SQL Server*** suggest that a Page Life Expectancy value below 300 seconds indicates too little memory, and this seems reasonable for ***Blackbaud CRM***.

Key Takeaways

The key takeaways from this series of performance tests include:

- While the exact memory requirement varies by organization, insufficient memory degrades system performance, particularly for user response times and disk input/output.
- With insufficient memory for both daytime and nighttime workloads, the system experiences a significant degradation of performance as the load increases.
- With insufficient memory, CPU usage increases somewhat from additional I/O overhead.
- With insufficient memory, disk I/O increases dramatically as the load increases, particularly the I/O for the Data drive.
- For nighttime workloads, the increased disk I/O necessitated by insufficient memory causes business processes to run longer.

For more information about the test results that led to these conclusions, see Appendix E.

Varying CPU Cores

This section provides an overview of results from running tests against ***Blackbaud CRM*** environments with a varying number of logical processors in the system. The tests started with a baseline of 12 logical processors, and then changed the number of available processors to 16, 8, and 4. The tests did not include hyperthreading, so one logical processor is the same as one core.

- In the Large Cause Organization Daytime Workload, the need for CPU cores grew with the number of users concurrently accessing the system at a rate of roughly 60 users per fully saturated core. After adjusting that figure by 50 percent based on Little's Law, a reasonable CPU loading level is roughly 30 users per core.
- In the Midsize Higher Education Foundation Daytime Workload, the need for CPU cores grew with the number of users concurrently accessing the system at a rate of roughly 100 users per fully saturated core. After adjusting that figure by 50 percent based on Little's Law, a reasonable CPU loading level is roughly 50 users per core.
- In the Large Cause Organization Nighttime Workload, the serialization of SQL statements limited the need for CPU cores. The largest number of cores required was 4. While other operations such as backups and index rebuilds may occur during the night, nighttime CPU usage from business processes was not significant.
- In the Midsize Higher Education Foundation Nighttime Workload, the serialization of SQL statements limited the need for CPU cores. The largest number of cores required was 2. While other operations such as backups and index rebuilds may occur during the night, nighttime CPU usage from business processes was not significant.

Modern servers are multi-core systems that generally support at least four cores per CPU, and possibly six or eight cores. For example, the server in these experiments is a dual-processor system with eight cores per processor. To run experiments with fewer cores, only some of the cores were enabled when booting the system. Barring an extraordinary need, it's easy to configure a server with plenty of power for ***Blackbaud CRM***. However, it is usually not easy or cheap to add processors to an under-configured system, so it's probably best to start generously in this area.

Here's a rule of thumb to compute how many CPU cores you need:

- Take your highest anticipated number of concurrent users of the system, and divide by 30. Remember to allow for growth over time in your estimate.
- Add 2 to the result to allow for a buffer as the upper limit approaches.
- Round up to a multiple of 4.
- That's the number of cores to use.

Is two sites sufficient to conclude that this is a good rule of thumb? That is simply all the available evidence at this time.

Key Takeaways

The key takeaways from this series of performance tests include:

- While the exact CPU core requirement varies by organization, insufficient CPU degrades system performance, particularly for user response times.
- As predicted by Little's Law, it's best to keep average CPU usage below about 50 percent.
- With insufficient CPU cores for daytime workloads, the system experiences a significant degradation of user response times as the load increases.
- I/O was not a bottleneck with insufficient CPU cores, and it didn't change for the CPU variations.
- Nighttime workloads are not limited by the number of cores. The serialization of SQL statements during these workloads limits the need for CPU cores, and adjusting the number of available cores did not affect performance.

For more information about the test results that led to these conclusions, see Appendix F.

Disk Input/Output

In the performance tests described in the previous sections, disk I/O was analyzed alongside the impact of varying load, memory, and CPU cores on system performance. Based on the test results from the four workloads in the performance tests, Blackbaud can make the following observations about disk I/O:

- The maximum I/O bandwidth is many times the average. It is not wise to plan for the average. Instead, Blackbaud recommends that you plan for 50 times the average.

- Business processes, as exemplified by the nighttime workloads, create a higher I/O demand than interactive users. This is particularly notable for the peak measurements. You should plan for business process levels of activity.
- There is some similarity in the numbers if you compare the peak I/O rate of the night workloads with the size of the data in the database. Not index space or reserved space, just data space. For the nighttime workloads, the peak data rate is about 0.0004 times the data space.

Is two workloads enough to conclude that this is a good rule of thumb? That is simply all the available evidence at this time. Here is an example that uses the large cause organization's database.

	DB Data size MB		Conversion		Max data rate, MB/sec
Data	1000000	x	0.0004	=	400

For modern SAN systems or RAID arrays, these are not unreasonable data rates. Talk to your system administrator or SAN administrator to clarify that your storage architecture will meet the need. Since the Data drive I/O requirement is higher than the Temp drive requirement, it's reasonable to use that as the overall system requirement. Peak needs for the two drives do not seem to occur at the same time; see Figure 6 below for an example.

Spinning Disks vs. Solid State Disks

SSDs are a marvelous technology, and response times from SSD stay quite fast up to very high loads. As Little's Law predicts, the response time of traditional disks becomes longer as the load increases. However, workloads continue to see fast performance from SSDs even after the point where increases to the work queue cause slower responses from traditional disks. So when I/O bottlenecks occur, SSDs have great potential to address the problem.

To force I/O to become the bottleneck in performance tests, the available memory was decreased just like in the memory tests. This drives I/O up to see the impact of SSDs vs. traditional disks. In these experiments, the SSD holds data, log, and temp space for the database all on one drive.

This section provides an overview of test results from comparing spinning disks vs. solid state discs for ***Blackbaud CRM*** environments. The performance recommendations drawn from these tests include:

- System responsiveness sees a small benefit from SSD.
- When memory is constrained, the benefits of SSDs are greater. SSDs can delay the point when performance becomes unacceptable. Sufficient memory is preferred, but SSDs may be less expensive than memory upgrades. Or they may be an option when memory upgrades are not possible.
- The optimal configuration may be to place the Data drive on SSD storage and place the Log drive and TempDB on spinning disks.

Key Takeaways

The key takeaways from this series of performance tests include:

- SSDs provide better read performance than spinning disks.
- SSDs can mitigate the performance impact from insufficient memory.
- For nighttime workloads, SSD storage provides a modest to insignificant improvement in the performance of the workload. The workloads are primarily limited by the fact that the system is serialized on a single execution thread, so the lack of difference between spinning disks and SSDs is expected.

For more information about the test results that led to these conclusions, see Appendix G.

Network Bandwidth and Latency

When it comes to networking, most people think about network bandwidth – how many bytes/sec can the network interface send and receive? While that is important, it is also important to consider network latency – how long does it take data to get from the source to the destination? This can be crucial when operating over the Internet, where it takes physical time for data to reach the server from the client's workstation, and vice versa. This section makes a few observations about network bandwidth for ***Blackbaud CRM***, and then focuses on latency because that can be crucially important for users who are far from the data centers that serve them.

Key takeaways from this topic:

- 1 GB/sec network interfaces are commonplace and are sufficient for ***Blackbaud CRM*** usage.
- Consider the possible effects resulting from the physical distance between your users and the ***Blackbaud CRM*** servers that support them. Pages that require many round-trips to display can be slowed by network latency, no matter how fast the server is.

In terms of network bandwidth, the busiest 30-second interval seen in any baseline test was 6 MB/sec. These are measurements taken at the database server. Of course, there are busier moments – after all, the only possible instantaneous values are "in use" and "not in use." But as intervals go, this is not really very busy. These tests were done with very common 1 GB/sec network interfaces in the systems; 6 MB/sec is 48 MB/sec (converting bytes to bits) which is 0.05 GB/sec. In other words, the network interfaces only run at 5 percent utilization in the busiest 30-second interval in any baseline tests.

Recommendation: 1 GB/sec network interfaces are very common and are sufficient for ***Blackbaud CRM*** usage. Use them.

This paper is almost entirely about the configuration of the database server for ***Blackbaud CRM*** because it is the most crucial component in terms of achieving good performance, but areas outside of the database server can be crucial as well. When you deal with websites and web applications, higher network latency can affect the perceived performance of a given system. For example, if data takes 2 seconds to transit from a web server to a workstation, then no matter how fast the web server is, the end user won't be able to view the data for at least two seconds. When a client is far away from the server, this network latency can have a significant impact on end user response time. To determine how much network latency impacts end user response time, a series of interactive tests were conducted at varying levels of network latency with the workload for the large cause organization. The tests measured end user response times for functions such as page load, tab load, and record saving. The tests used an isolated lab environment with additional network latency added to the connection between the test agents (from which requests originate) and the web servers.

Several tests were conducted at various network latencies based on Verizon latency SLAs. For each latency, two tests were run, and the results were compared to the baseline.

- Baseline of 0ms (no network latency added) – This establishes a working baseline to show response time when the server and the clients are collocated.
- North America, close (15ms) – This models going up and down the Eastern coast of North America.
- Europe (30ms) – This models going across Europe (both the client and server are in Europe).
- North America, far (45ms) – This models going across North America.
- Transatlantic (90ms) – This models going across the Atlantic Ocean to England.
- Transpacific (110ms) – This models going across the Pacific Ocean to Australia from Los Angeles.
- High Latency (220ms) – This models the 110ms latency and doubles it for an overly high network latency test to establish a worst case.

The chart that follows illustrates how higher network latency causes end user response times to degrade. The chart displays the results of each test based on the Performance Index, which shows how well a test run compares to the baseline. A performance index of 1 is the baseline performance, and lower numbers indicate worse performance. For example, a Performance Index of 0.5 means the average user response time was twice as long.

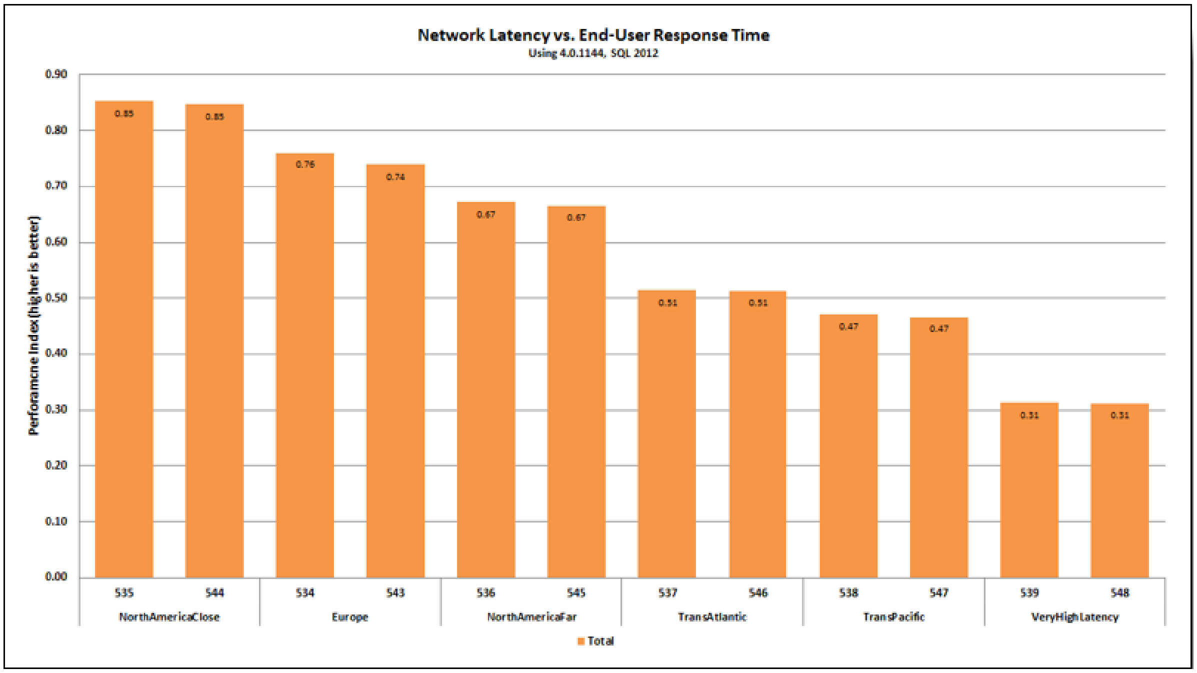

Figure 1

The test results confirm that end user response times get slower as network latency increases. They even indicate that higher network latencies can cause page loads to regress to a point where users might lose focus on tasks and browse away from pages while waiting for them to load. Most transactions got slower as network latency increased, but the ones that regressed the most were the actions that construct complicated pages via multiple HTTP requests, and hence multiple round-trips to the server, to gather information to construct specific data forms. The delays caused by higher network latency come from two sources: Round-trips that occur because the system makes multiple calls in sequence (to populate various parts of the screen), and round-trips that occur at the TCP layer to accommodate large amounts of data.

Interactive Workloads and Business Processes

This seems obvious in some ways, but it must be said: To the degree it can be arranged, you should run large business processes, smart field recalculations, data extracts, ETL jobs, and system maintenance functions such as index rebuilds during times when users are not also trying to perform interactive operations on the system. On a number of occasions when customers report of poor system performance, it turns out to be the result of large operations like these using too many system resources and slowing the system down for live users. This is one reason that Blackbaud recommends over-configuring the system – so that you have resources to spare if an unanticipated large job is executed. Still, it remains a best practice to run such workloads off-hours.

Both of the model sites used in this work do exactly that, and the naming of the “day” and “night” workloads reflects that.

Key takeaway from this topic:

- Plan to run large non-interactive workloads during off-hours.

End User Response Time Observations

If a site exactly follows the guidance in this paper to configure a ***Blackbaud CRM*** system, what responsiveness can they expect? The simple answer is that “it depends” because each site is unique. In particular, performance depends on the size of the database and whether business process activity occurs during the same time when users interact with the system.

But assuming that little or no business process activity occurs during the day, Blackbaud can set some general expectations by focusing on particular transactions in the data from the tests discussed in this paper. Up to this point, this paper has looked at interactive user transactions in the aggregate and simply used averages of all transactions during the course of a run. But now it isolates some transactions of interest to examine the responsiveness of those transactions.

Transaction Name	Description	Occurrences in a baseline run
ConstitSearch-Search	Search for a specific constituent, given a last name and post code.	Large cause organization: 971 Midsize higher education foundation: 1,725
AddPayment-ConfirmFormSession	Save a one-off Add a Payment form.	Large cause organization: 198 Midsize higher education foundation: n/a
EditBio-ConfirmFormSession	Save changes to an individual constituent’s first name and last name.	Large cause organization: 32 Midsize higher education foundation: n/a
ProspectSearch-Search	Search for a specific prospect, given the lookup ID.	Large cause organization: 23 Midsize higher education foundation: 171
RevenueHistoryTab-GetTab	Show the Revenue History tab for a constituent with revenue. This requires fetching the historical revenue information for the constituent.	Large cause organization: 62 Midsize higher education foundation: n/a
EditInteraction-ConfirmFormSession	Save changes to an interaction: Add a date and a short comment, and mark the interaction as completed.	Large cause organization: 60 Midsize higher education foundation: 41
EditAddress-ConfirmFormSession	Save changes to an address and its start date.	Large cause organization: 125 Midsize higher education foundation: 120
EditPostedPayment-ConfirmFormSession	Save changes to the amount of a posted payment. This is actually a complex operation that reflects a change to a payment that was already posted to the GL. The value of a stock donation is being revised.	Large cause organization: 65 Midsize higher education foundation: n/a

Validate-BatchEntryAction	Validate a batch after editing.	Large cause organization: n/a Midsize higher education foundation: 5
SaveBatch	Save a batch after editing.	Large cause organization: 9 Midsize higher education foundation: 5
RevenueSummaryTab-GetTab	Load the Revenue Summary tab that shows a donation summary for a constituent.	Large cause organization: 61 Midsize higher education foundation: n/a

Note that certain pages in ***Blackbaud CRM*** are configurable and that their complexity depends on the user-determined content. The Constituent page is an example. In ***Blackbaud CRM*** 4.0, the user experience when viewing a constituent was updated to use summary tiles in place of a multi-use summary section. These tiles load asynchronously, and completion of loading the Constituent page depends on the load times of the individual tiles in use. Hence the load time varies from site to site and even from user to user.

For the selected actions, here are average and 90th percentile response times (in seconds) from baseline runs. These measurements are from the baseline configuration, with no added network latency.

	Large Cause Organization Daytime Workload		Midize Higher Education Foundation Daytime Workload	
	ResponseMean	Response90th	ResponseMean	Response90th
AddPayment-ConfirmFormSession	0.40	0.56		
ConstitSearch-Search	0.12	0.10	0.04	0.04
EditAddress-ConfirmFormSession	0.23	0.29	0.09	0.12
EditBio-ConfirmFormSession	0.11	0.15		
EditInteraction-ConfirmFormSession	0.07	0.10	0.06	0.09
EditPostedPayment-ConfirmFormSession	6.56	9.36		
ProspectSearch-Search	0.12	0.34	0.39	0.41
RevenueHistoryTab-GetTab	1.99	2.88		
RevenueSummaryTab-GetTab	0.19	0.32		
SaveBatch*	0.08	0.14	4.35	7.87
Validate-BatchEntryAction			6.27	12.74

* Even though the users for both the large cause organization and the midsize higher education foundation edited and saved batches with 50 rows, the contents of the batches were very different and required vastly different processing for validation. This is why the execution times were so different between the two sites. The large cause organization batches include simple revenue that requires very little validation. For the higher education foundation, the revenue is applied to commitments that must be checked to validate each row of the batch, and each revenue row is associated with a note.

Appendices

The following appendices include additional information about the performance testing and guidelines described in this paper.

Appendix A – Performance Test Overview

This paper reflects conclusions drawn from performance testing. The performance tests run workloads that mirror the average daily use patterns of actual clients. To establish baselines for the workloads, multiple tests ran under consistent conditions and established baselines with average response times for all actions in the workloads. With the baselines in place, the available system resources such as CPU cores and memory were altered and new performance tests were run to determine the impact of the variations by comparing new test results to the baseline results. This technique determined how the physical configuration of ***Blackbaud CRM*** environments impacted the responsiveness of the systems.

The performance tests in this paper focus on the database server because database servers are usually the bottleneck in CRM systems. Tests on other areas are less likely to yield significant results.

When evaluating the results of performance tests, it is important to rely on Little's Law to help determine the optimal physical configuration for ***Blackbaud CRM*** environments. Little's Law indicates that you should never run critical resources at 100 percent utilization. As the load on critical resources increases, the queue for those resources also increases, and eventually this degrades performance. Little's Law suggests that you should limit the load on critical resources to 50 to 60 percent. In these performance tests, load curves were created by isolating various resources and increasing the load on them until the resources were in bottleneck conditions. The shape of the load curves were then analyzed to determine where the bottlenecks occurred and identify critical regions for the various resources. With that information in hand, rules of thumb could be devised for the resources to allow you to plan appropriately for ***Blackbaud CRM*** systems.

Appendix B – System Configuration

To conduct the performance tests, a performance test lab was set up to mirror the hosting environment for ***Blackbaud CRM*** clients. The test lab is well-controlled for updates, virus scans, and other administrative activities in order to ensure repeatable and reliable tests.

The following table describes the servers and specifications for the test lab environment for the performance tests described in this paper. Note that the baseline uses 12 logical processors, not 16, and that it uses 95 GB of memory even though the system has more available. The baseline configuration also uses spinning discs instead of SSD.

SQL Server	• 2x Intel Xeon E5-2690 @ 2.9 GHz (16 cores) • 512 GB RAM (configured to use 95 GB) • Windows Server 2008 R2 SP1 Enterprise • ***SQL Server*** 2014 SP1 Developer • Visual Studio Team Test Agent 2013 • Storage configuration o Data drive: 16x 600 GB, 10K RPM, RAID 6, 4.4 TB usable

	○ Temp drive: 4x 600 GB, 10K RPM, RAID 1+0, 1.1 TB usable ○ Log drive: 4x 600 GB, 10K RPM, RAID 1+0, 1.1 TB usable ○ SSD: 4x 800 GB SSD, RAID 0, 3.0 TB usable
Web Servers (2)	• Virtual Servers • 2x Intel Xeon E5-2697 v2 @ 2.7 GHz (4 cores) • 8 GB RAM • Windows Server 2008 R2 SP1 Standard • Load balanced using Application Request Routing with weighted round robin even distribution • Visual Studio Team Test Agent 2013

Appendix C – Workloads

The workloads for the performance tests reflect the typical actions of ***Blackbaud CRM*** users. They include both interactive actions and business processes. The workloads are modeled after user activity at two customer sites where IIS logs were obtained from production web servers. Data from a month-long period was analyzed to determine the actions that users perform and the order that they perform them in. Copies of the production databases were masked to protect personally identifiable information while retaining the realistic data relationships in the databases. For each organization, the user activity plus masked databases were used to create daytime and nighttime workloads reflecting the activity that occurs during peak hours and off hours. Average activity levels for both of the workload model organizations were measured during their busy seasons.

The first set of workloads is modeled after activity at a large cause organization, and the second set of workloads is modeled after activity at a midsize higher education foundation. These two ***Blackbaud CRM*** customers represent two very different usage patterns, and the differences in the test results help to shed light on the physical resources necessary for different organizations.

Large Cause Organization Workloads

The workloads modeled after the large cause organization represent a large customer with a low-touch constituent contact model. The database is 2.3 TB with about 4.6 million constituents, and the deployment is hosted by Blackbaud.

- The daytime workload represents typical user actions and the typical order that they occur in. It includes interactive actions and business processes. The most common activities were selected and grouped based on the tasks that users were working on. Then activities were organized into scenarios, web performance tests, and transactions (individual user actions). This workload models 149 distinct transactions, and each runs between 1 and 500 times during a load test run. The workload contains the following web performance tests:
  - Add an Individual
  - Edit an Individual

- Add a Household
    - Edit a Household
    - Add an Organization
    - Add a Constituent Solicit Code
    - Add a Constituent Interaction
    - Edit a Constituent Organization
    - Edit a Constituent Address
    - Edit a Constituent Name Format
    - Edit Constituent Individual Biographical
    - Add a Payment
    - Edit a Posted Payment
    - Manual Batch Entry without Commit
    - Edit Reservation Comments
    - Edit Reservation Due Dates
    - Edit a Track Load
    - Add a Contact Report
    - Add a Deposit
    - Link Payments to a New Deposit
    - View a Constituent History Report
    - View an Itinerary Report
    - Assign Constituent Security Group Process
    - Que Process – All Imports
    - Smart Field Process
    - Global Change Process
- The nighttime workload represents typical activity overnight in the client system. It mainly includes business processes. Activities were organized into scenarios, web performance tests, and transactions just like in the daytime workload. The workload models 16 distinct transactions. The workload contains the following web performance tests:
    - Incremental Constituent Duplicate Search Process
    - Constituent Merge Process
    - Queue Process – Production Weekday
    - Queue Process – Production Nightly Was Scheduled Weekly
    - Queue Process – Scheduled Daily Processing – DM Queues Then Smart Fields

Midsize Higher Education Foundation Workloads

The workloads modeled after the midsize higher education foundation represent a mid-sized customer with a high-touch constituent contact model. The database is 250 GB with about 1 million constituents, and the deployment is hosted on premise by the foundation.

- The daytime workload represents typical user actions and the typical order that they occur in. It includes interactive actions and business processes. The most common activities were selected and grouped based on the tasks that users were working on. Then activities were organized into scenarios, web performance tests, and transactions (individual user actions). This workload models 119 distinct transactions, and each runs between 1 and 296 times during a load test run. The workload contains the following web performance tests:
    - Add an Individual
    - Add a Constituent Address
    - Edit a Constituent Address
    - Add a Constituent Phone
    - Edit a Constituent Phone
    - Add a Constituent Email Address
    - Edit a Constituent Email Address
    - Add a Constituent Interaction
    - Edit a Constituent Interaction
    - Add a Constituent Alias
    - Edit a Constituent Alias
    - Add a Constituent Solicit Code
    - Add a Constituent Name Format
    - Add a Constituent Attribute
    - Add a Constituent Educational Involvement
    - Add a Constituent Alternate Lookup ID
    - Add a Constituent Constituency
    - Add a Constituent Prospect Constituency
    - Add an Individual Relationship
    - Edit Constituent Individual Biographical
    - Edit Prospect Details
    - Edit an Opportunity
    - Edit a Step
    - Add Steward Recipient
    - Add a Deposit
    - Add a Registrant
    - Edit Revenue Recognition Credit
    - Manual Batch Entry with Commit
    - Batch Commit Process
- The nighttime workload represents typical activity overnight in a client system. It mainly includes business processes. Activities were organized into scenarios, web performance tests, and transactions just like in the daytime workload. The workload models 1083 distinct transactions that run 1 to 2 times during a load test run. The workload contains the following web performance tests:
    - Refresh Queue Process (to refresh smart fields and static selections)

- Drop Codes Refresh Queue Process (to refresh static selections)
- Exists in Queue Process (to refresh a large number of static selections)
- Geographic Areas Queue Process (to refresh a large number of static selections)

A critical component of the performance tests is the ability to increase the load and then analyze the impact of the increased load on various physical resources. The tables below illustrate how the number of users scaled up linearly to increase the impact on the system. For example, the baseline for the large cause organization daytime workload includes 160 users, and when that load increases by a factor of 1.5 times the normal load, the workload includes 240 users.

Large Cause Organization Daytime Workload

Ratio	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
Users	160	200	240	280	320	360	400	440	480
Users+BPs	164	204	244	284	324	364	404	444	484

Midsize Higher Education Foundation Daytime Workload

Ratio	1.00	1.50	2.00	2.50	3.00	3.50	4.00	4.50	5.00	5.50	6.00	6.50
Users	142	213	284	355	426	497	568	639	710	781	852	923

The performance tests for daytime workloads use consistent run times and warmup periods. The consistent run times ensure repeatable and reliable test results, and the warmup periods ensure a stable state when measurements are taken. Here's how the process works:

1. Restore the database from a backup before each test run.
2. Start the workload and then run it for a 15-minute warmup period to reach a stable state before taking any measurements.
3. Run the tests for another hour to record response times.
4. Shut down, restore the database, and start another test run.

Nighttime workloads run longer to allow the business processes to complete and do not use a warmup periods.

Appendix D – Varying Loads Test Results

This appendix describes the detailed results from testing ***Blackbaud CRM*** environments with varying loads on the system. The tests start with baseline workloads and then increase the number of users linearly to determine the impact that the increased load and its resultant activity have on the system.

Large Cause Organization Daytime Workloads

Key takeaways for this workload:

- CPU usage grows linearly with the number of active users.
- Response times follow Little's Law, so plan to stay below 50 percent CPU usage.
- Disk input/output experiences significant bursts above the average, so plan for the peaks.

To measure the impact of increased load on user response times for the large cause organization, the baseline workload starts with a reflection of the actual number of users at the organization. From this starting point of 1.0 times the normal number of users, the load increases linearly by a factor of 0.25 until it reaches 3.0 times the normal number of users.

As the following chart indicates, the average response time for the baseline is about 0.6 second per action, although this varies for different types of user actions. As the load increases, average response time gradually increases to more than 1 second per action when the load reaches 3 times the usual number of users.

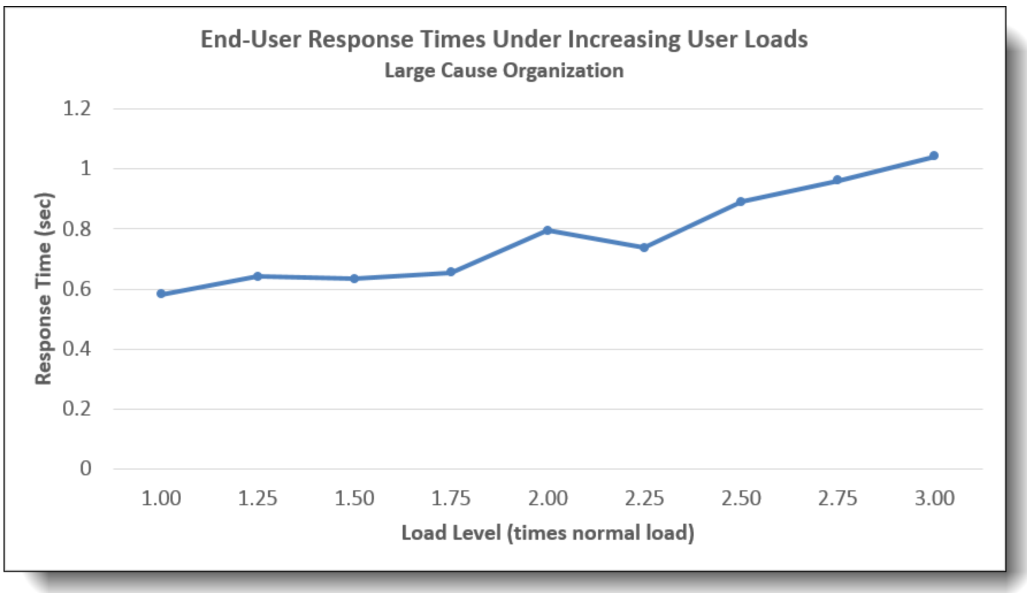

Figure 2

The tests measured CPU usage during each run, which is the percentage of available CPU time. As the next chart indicates, average CPU usage increases linearly as the number of users increases. The rate of increase is very roughly one fully loaded CPU core per 60 users. Little's Law indicates that you never want to plan for fully loaded CPUs, so this organization should plan for 30 users per CPU core.

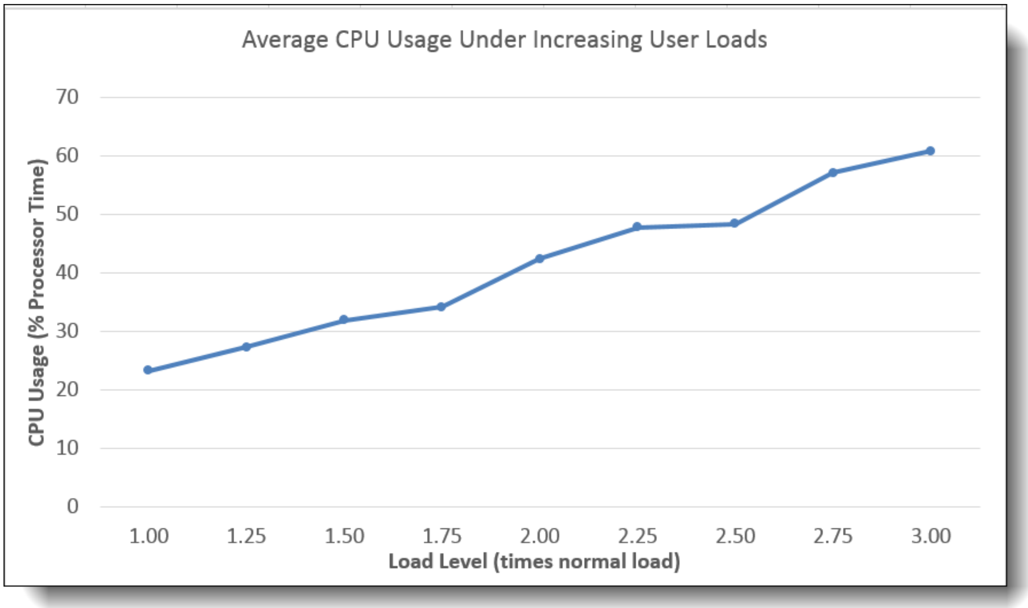

Figure 3

The following chart maps the CPU usage in 30-second sampling intervals during a single run. The chart shows that the CPU usage varies considerably over time. It is important to plan for CPU usage during the busiest times, not just for the average CPU usage. These workloads apply randomized-but-average loads throughout the run. In real-world applications, you must also consider higher-level patterns that push loads higher and lower at different times. For example, you likely will see a lull during lunch hours and a spike when users return from lunch at about the same time. Also, nonprofit organizations generally have a busy season around the end of a calendar year. Average activity levels for both of the workload model organizations were measured during their busy seasons.

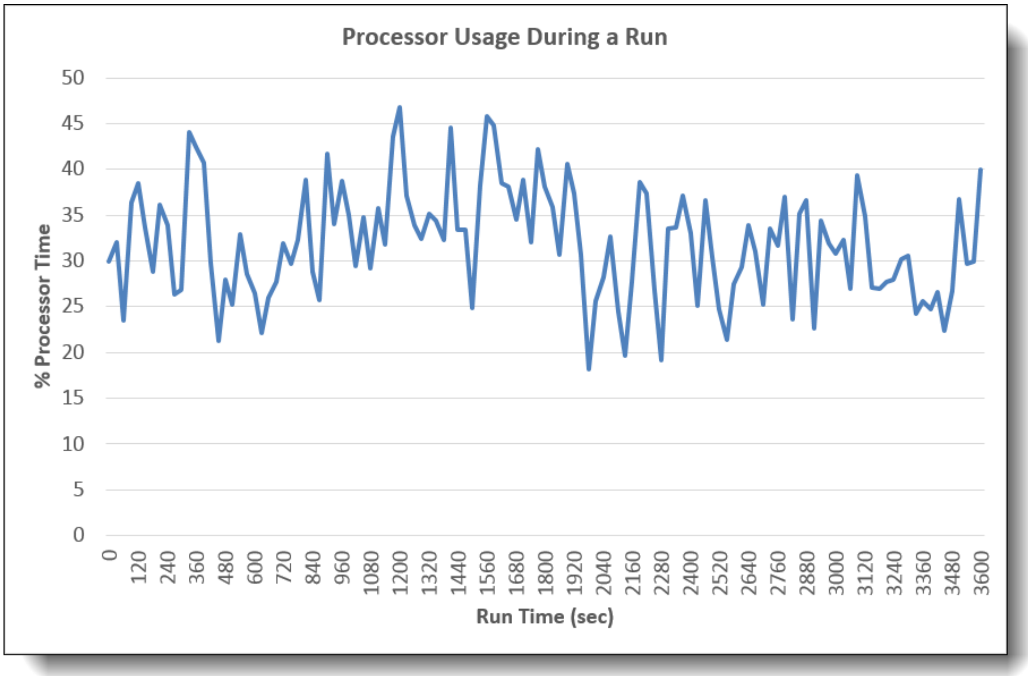

Figure 4

The tests also measure disk I/O bytes per second for the Data drive with ***Blackbaud CRM*** database data files, the Log drive with database log files, and the Temp drive with the TempDB data and log files.

As the chart below indicates, the average activity on the Data drive increases the most under the heavier load, but it remains well with the capabilities of a RAID5 disk system like this.

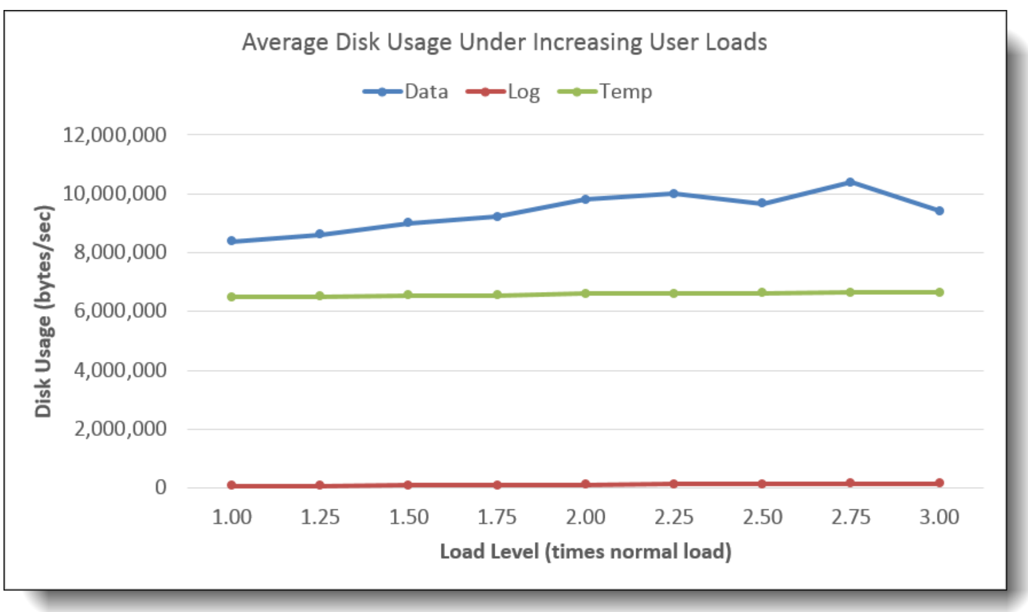

Figure 5

Keep in mind that the previous chart illustrates ***average*** disk activity during the duration of a run. As the next chart indicates, activity at any given time during a run can vary quite a bit. That chart shows disk bytes per second during 30-second sampling periods, and it is consistent with all of the Large Cause Organization Daytime Workload runs. During one interval, nearly 200 MB per second of data was transferred from the Data drive, but the average amount of data transferred from the Data drive was much lower. Similarly, the Temp drive hit more than 65 MB per second during one 30-second interval even though its average disk output per second was much lower. The chart also shows an increase in TempDB usage starting at about 950 seconds into the run and lasting until about 1900 seconds into the run. This increase occurs when smart fields are processed. Although the CPU load is slight from this business process, the usage of TempDB I/O is significant. High TempDB I/O seems to be a common feature among many business process that handle large amounts of data. It is very pronounced in the nighttime workload.

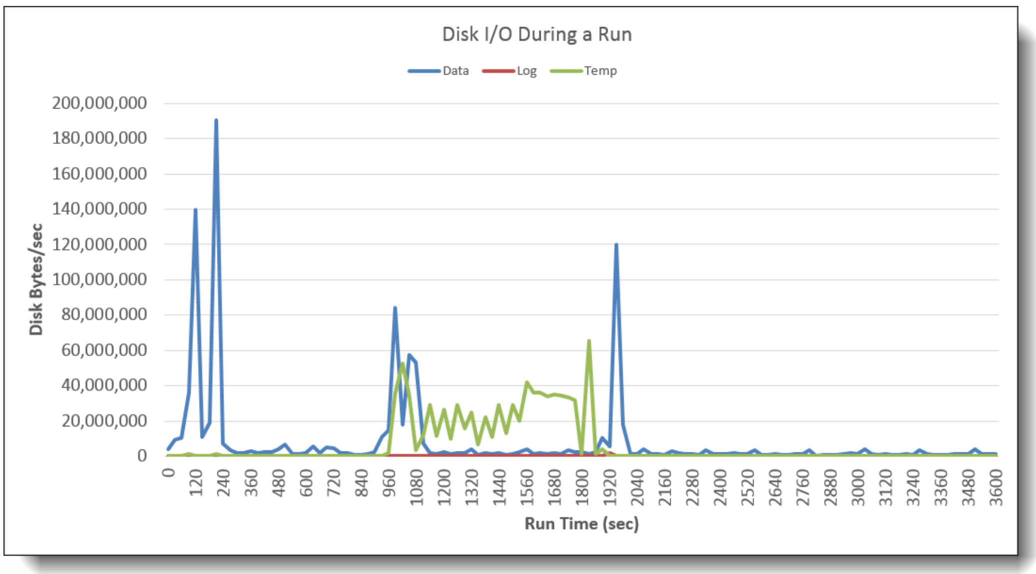

Figure 6

The final chart in this section illustrates the extreme relationship between the average I/O for each drive and the maximum I/O observed during testing. When it comes to I/O bandwidth, planning for the average is not nearly sufficient.

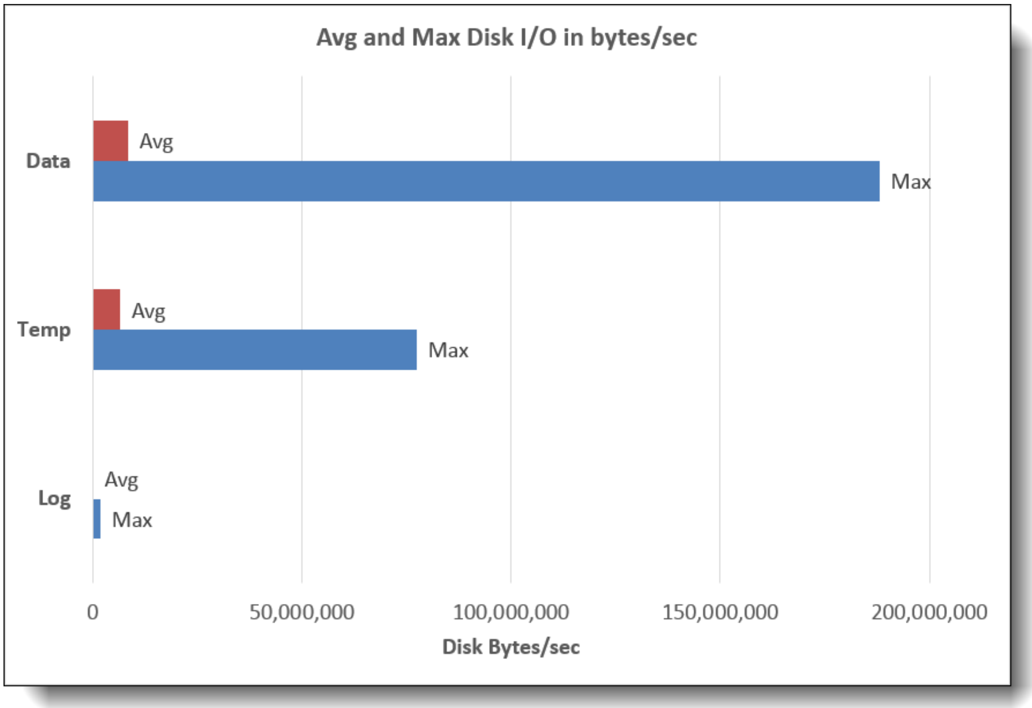

Figure 7

Large Cause Organization Nighttime Workloads

Key takeaways for this workload:

- CPU usage is limited by serialization, so more cores won't speed up this workload.
- Significant usage of TempDB occurs during this workload, and it is in more demand than the Data drive.

Each ***Blackbaud CRM*** customer has a unique set of business processes that handles various tasks such as data imports and exports, data-cleaning operations, re-computing smart fields, and ETL operations to maintain the data warehouse. Comparing the performance of these operations between sites has limited value because the work varies widely from organization to organization. However, some commonalities are worth noting in the large operations that commonly occur after hours. First, they tend to be serialized in terms of CPU usage. And second, the TempDB usage is relatively high.

For the Large Cause Organization Nighttime Workload, the average CPU usage is about 15 percent of available CPU time. However, this varies throughout the duration of the run as the following chart illustrates. The CPU usage often "plateaus" at certain levels, which is almost certainly because work on each long-running task is serialized on a single execution thread and cannot be parallelized to take advantage of multiple CPUs. With 12 logical processors in the system, one LP is 8.3 percent of the total capacity, and plateaus occur at about 8 percent, 17 percent, and 25 percent just like expected based on serialization when serialized tasks run concurrently. Increasing the number of cores won't improve performance because this nighttime workload includes a small number of very large processes that don't scale across multiple cores.

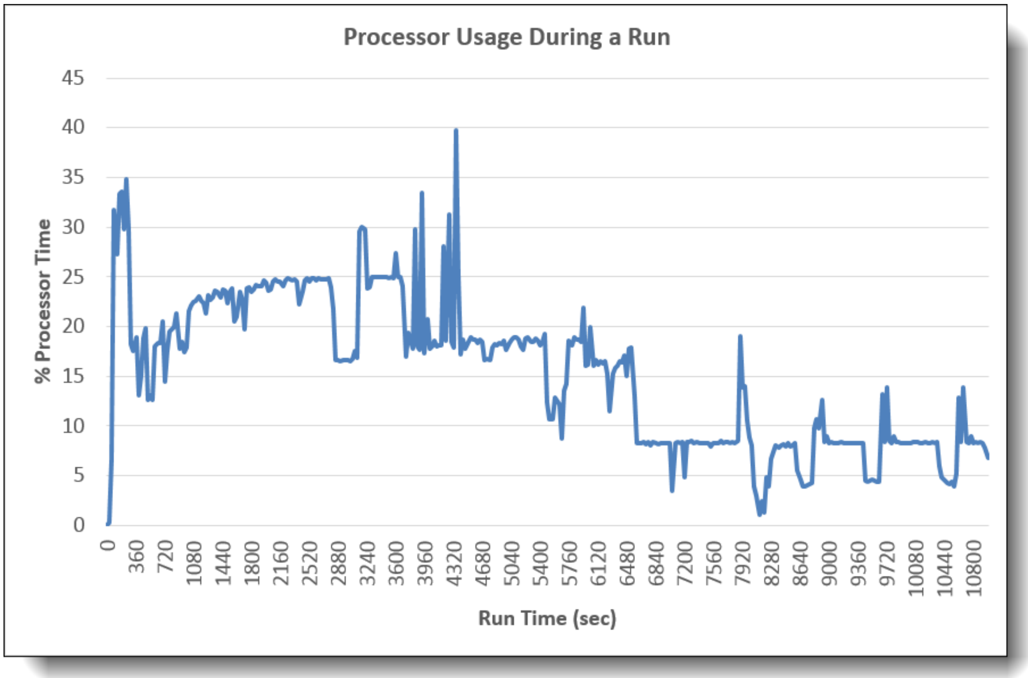

Figure 8

The next chart shows the average disk bytes per second for the Data drive with ***Blackbaud CRM*** database data files, the Log drive with database log files, and the Temp drive with the TempDB data and log files. Far more TempDB activity occurs in the business process-oriented nighttime workload than in the user activity-oriented daytime workload because business processes commonly run complex queries that require temporary data storage. You should not overlook the importance of TempDB in ***Blackbaud CRM*** systems, especially for complex business processes, even though additional memory in the server can reduce the need for TempDB in some cases.

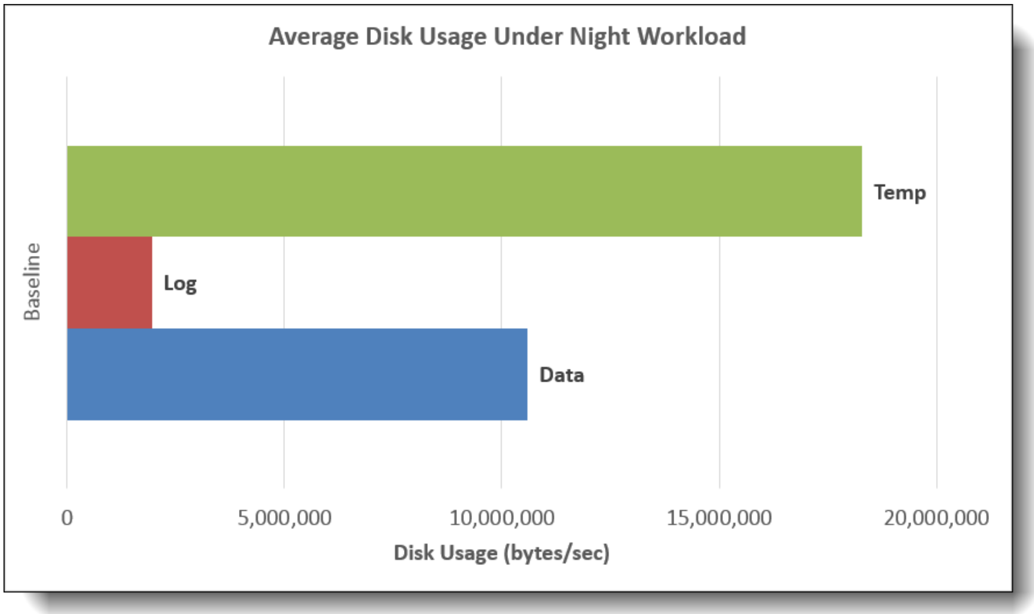

Figure 9

The next chart illustrates that the maximum disk throughput during the nighttime workload is many times higher than the average, just like in the daytime workload.

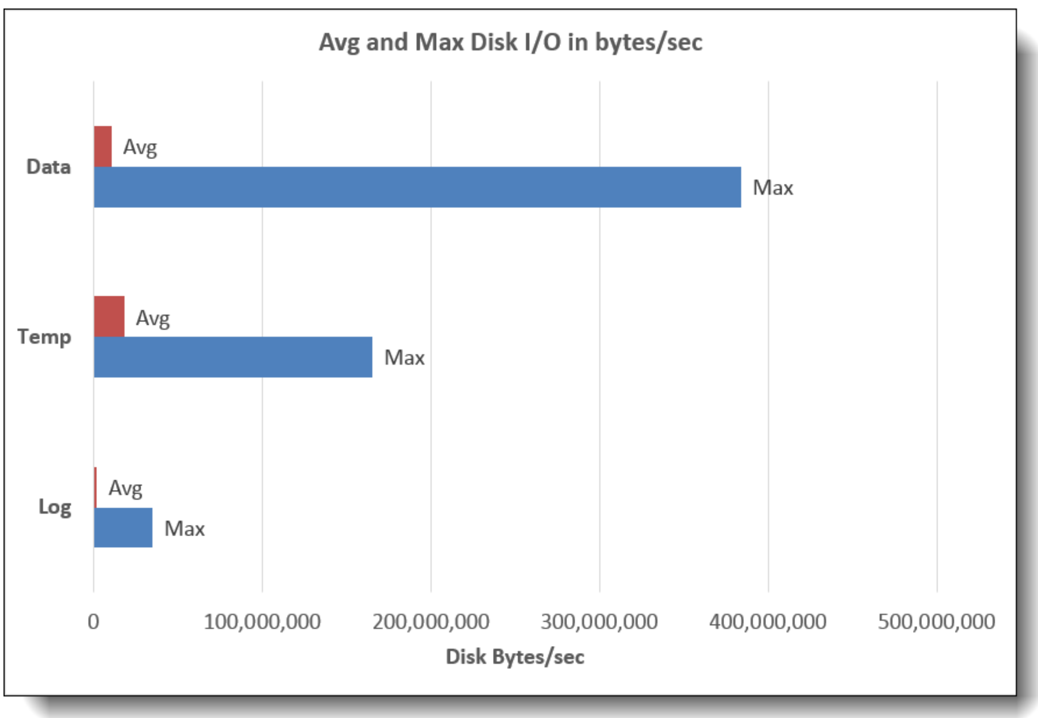

Figure 10

Midsize Higher Education Foundation Daytime Workload

Key takeaways for this workload:

- CPU usage grows linearly with the number of active users.
- Disk input/output experiences significant bursts above the average, so plan for the peaks.
- TempDB usage increases as user load increases, which is different from the large cause organization.

To measure the impact of increased load on user response times for the midsize higher education foundation, the tests start with the baseline workload that reflects the actual number of users at the organization. From this starting point of 1.0 times the normal number of users, the load increases linearly by a factor of 0.5 until it reached 6.5 times the normal number of users.

As the following chart indicates, the average response time for the baseline number of users was less than 0.2 seconds per action, although this varies for different types of user actions. (For this analysis, batch commits are not considered interactive operations, even though they occur during the day.) As the load increases, the average response time gradually increases to nearly 0.35 seconds per action when the load reaches 6.5 times the usual number of users.

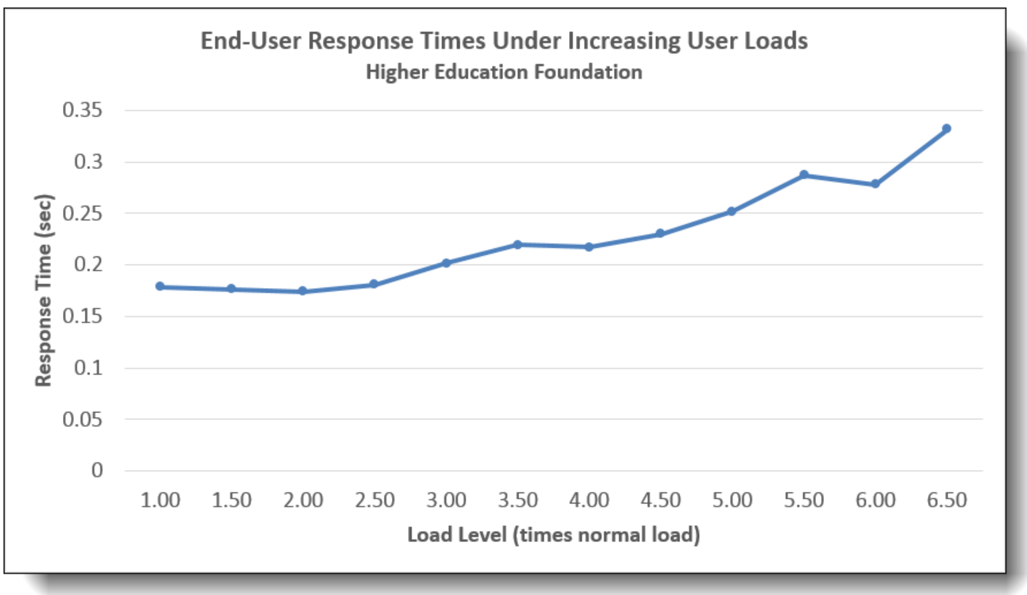

Figure 11

The tests measured CPU usage during each run, which is the percentage of available CPU time.

As the chart below shows, average CPU usage increases linearly as the number of users increases. The rate of increase is very roughly one fully loaded CPU core per 100 users. As Little's Law indicates, you never want to plan for fully loaded CPUs, so this organization should plan for 50 users per CPU core.

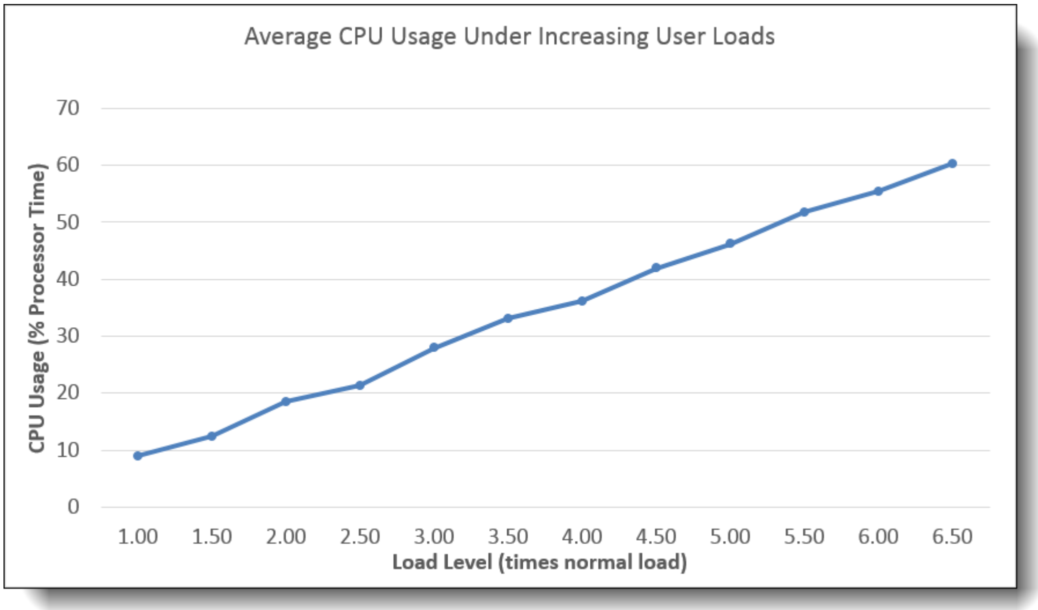

Figure 12

As with the Large Cause Organization Daytime Workload, CPU usage varies considerably over time. For brevity, a chart that illustrates this point again was excluded, but keep in mind that it is important to plan for CPU usage during the busiest times, not just for the average CPU usage.

The tests also measured the disk bytes per second for the Data drive with ***Blackbaud CRM*** database data files, the Log drive with database log files, and the Temp drive with the TempDB data and log files.

As the chart below indicates, the activity on the Data drive increases under the heavier load, but it remains well within the capabilities of a RAID5 disk system like this. However, the usage of TempDB increases much faster. This highlights how important it is that the drive holding TempDB not only has enough capacity, but also has enough I/O bandwidth.

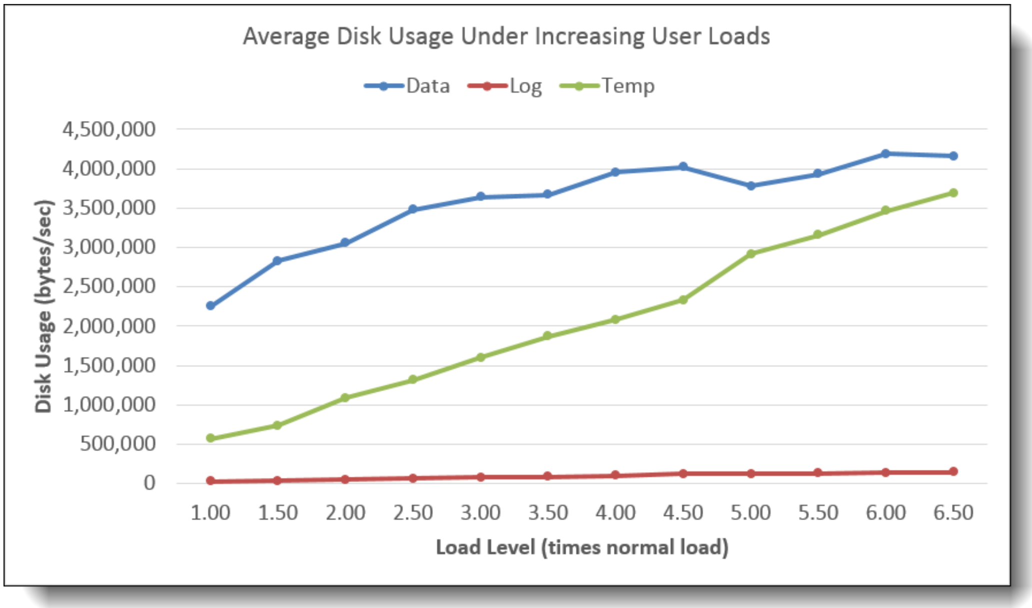

Figure 13

The chart illustrates a different pattern from the large cause organization and shows that the two sites use the system quite differently. It is not easy to look at one site and make predictions for another. However, both organizations saw the same extreme relationship between the average I/O and the maximum I/O observed.

The chart below shows that the maximum disk throughput is once again many times higher than the average, which means that planning for the average is not sufficient.

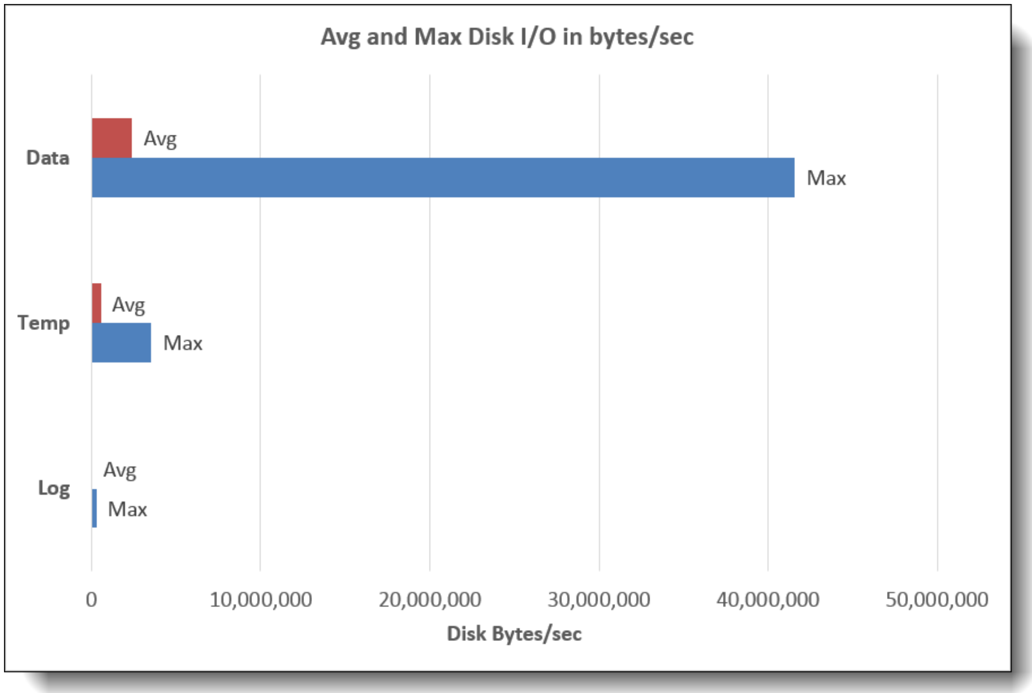

Figure 14

Midsize Higher Education Foundation Nighttime Workload

Key takeaways for this workload:

- CPU usage is limited by serialization, so more cores won't speed up this workload.
- Disk input/output experiences significant bursts above the average, so plan for the peaks.

As the Large Cause Organization Nighttime Workload section explained, ***Blackbaud CRM*** organizations have unique sets of business processes that vary widely, and comparing their performance between sites has limited value. The main similarities to note about the large operations that typically occur after hours is that they are serialized in terms of CPU usage and that the TempDB usage is relatively high.

For the Midsize Higher Education Foundation Nighttime Workload, CPU usage averages about 8.5 percent of available CPU time, although this varies throughout the duration of the run. Since this is a 12-core system configuration, this usage is almost exactly the equivalent of one core. The chart below illustrates CPU usage during a single run, and there is almost always the equivalent of one core (LP) being used, with small variations up or down from that level. The CPU usage "plateaus" at about 8.5 percent of total capacity, which is almost certainly because the work on each long-running task is serialized on a single execution thread and cannot be parallelized to take advantage of multiple CPUs. Increasing the number of cores won't improve performance because the nighttime workload contains a small number of large processes that don't scale across multiple cores.

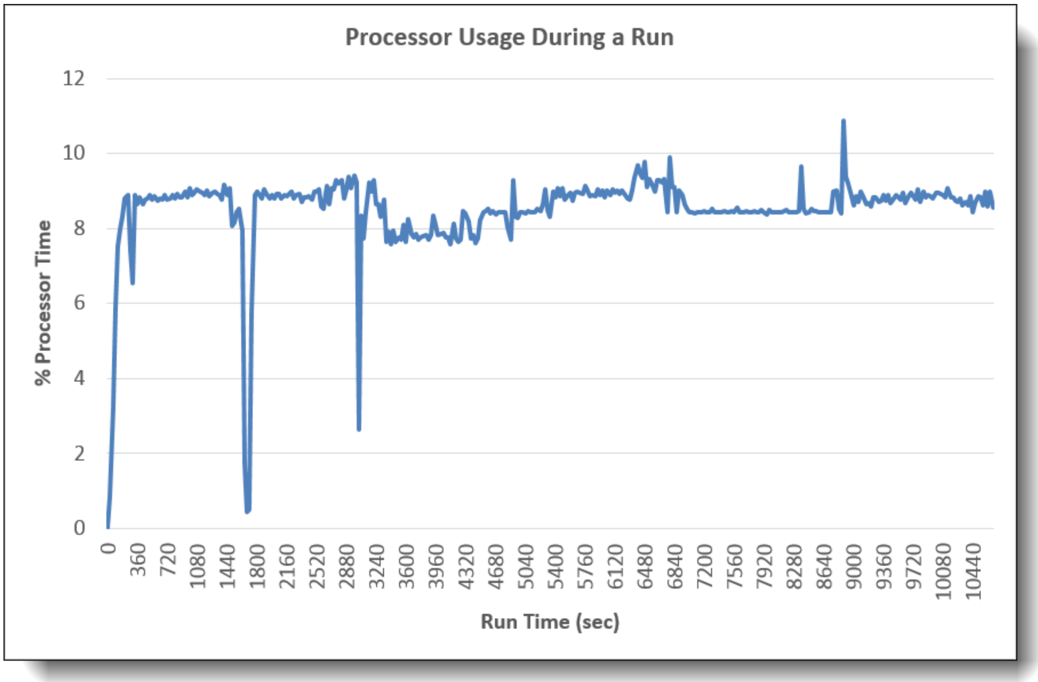

Figure 15

The next chart shows the average disk bytes per second for the Data drive with ***Blackbaud CRM*** database data files, the Log drive with database log files, and the Temp drive with the TempDB data and log files. Note that TempDB activity here is less significant than in the Large Cause Organization Nighttime Workload and that the disk activity is much lower. This is probably because the workload is serialized on one core!

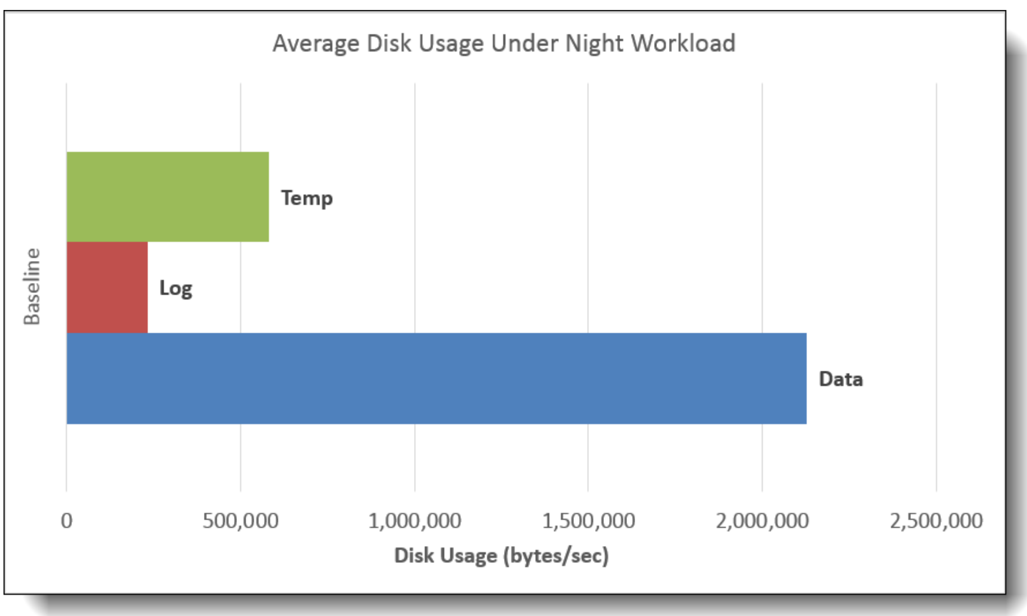

Figure 16

Finally, the last chart in this section illustrates that the maximum disk throughput is once again many times higher than the average disk throughput.

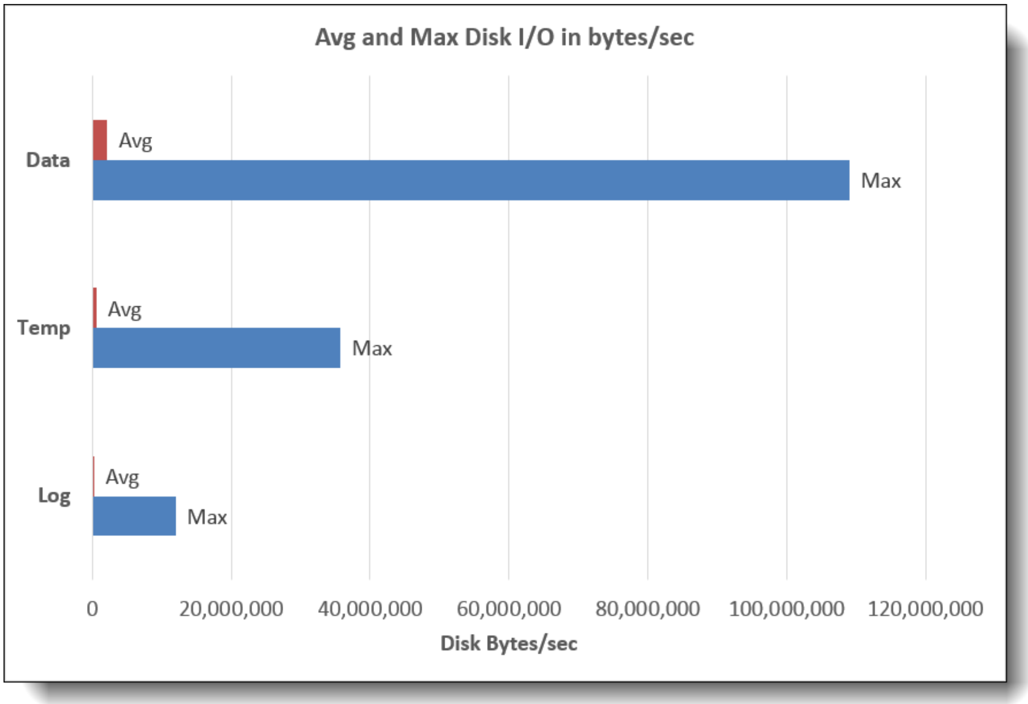

Figure 17

Appendix E – Varying Memory Test Results

This appendix describes the detailed results from testing ***Blackbaud CRM*** environments with varying amounts of memory available to ***SQL Server***. The tests start with a baseline of 95 GB for the database and then reduce the amount of memory available to the ***SQL Server*** instance by adjusting the maximum memory setting. This setting has the same effect as if the server had less memory available. For example, executing *`sp_configure 'max server memory', 40960`* limits ***SQL Server*** to use no more than 40 GB of memory.

Large Cause Organization Daytime Workloads

Key takeaways for this workload:

- 40 GB of memory is sufficient for this workload; 20 GB is insufficient.
- With insufficient memory, response times increase.
- With insufficient memory, CPU usage increases somewhat from additional I/O overhead.
- With insufficient memory, disk I/O increases dramatically, particularly for the Data drive.

To measure the impact of increased load for different memory variations at the large cause organization, workloads ran for multiple memory settings and increased the number of users linearly. Memory settings included 60 GB, 40 GB, and 20 GB. The tests started with the baseline workload of 1.0

times the normal number of users and increased the load linearly by a factor of 0.25 until it reached 3.0 times the normal number of users.

The tests measured the impact of increased load on user response times for each memory variation. As the chart below indicates, the average response times seen by interactive users on the ***Blackbaud CRM*** system was not greatly affected when the number of users increased until memory was reduced to 20 GB. The system consistently achieved sub-second average response times for each memory variation until response times started to increase significantly for the 20 GB memory setting.

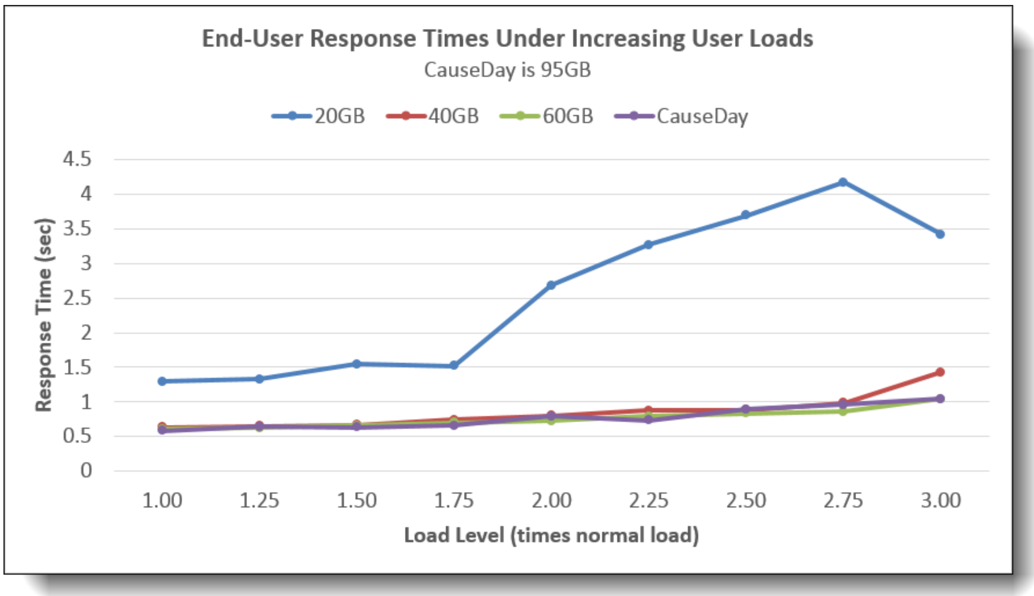

Figure 18

The tests measured CPU usage for each memory variation. The next chart shows the average CPU usage as the number of users increases for each memory variation. Even though the system performs the same amount of work for each memory variation, CPU usage is higher for the 20 GB memory setting. The system is 10 to 20 percent less efficient in CPU usage because of the additional I/O that it must perform.

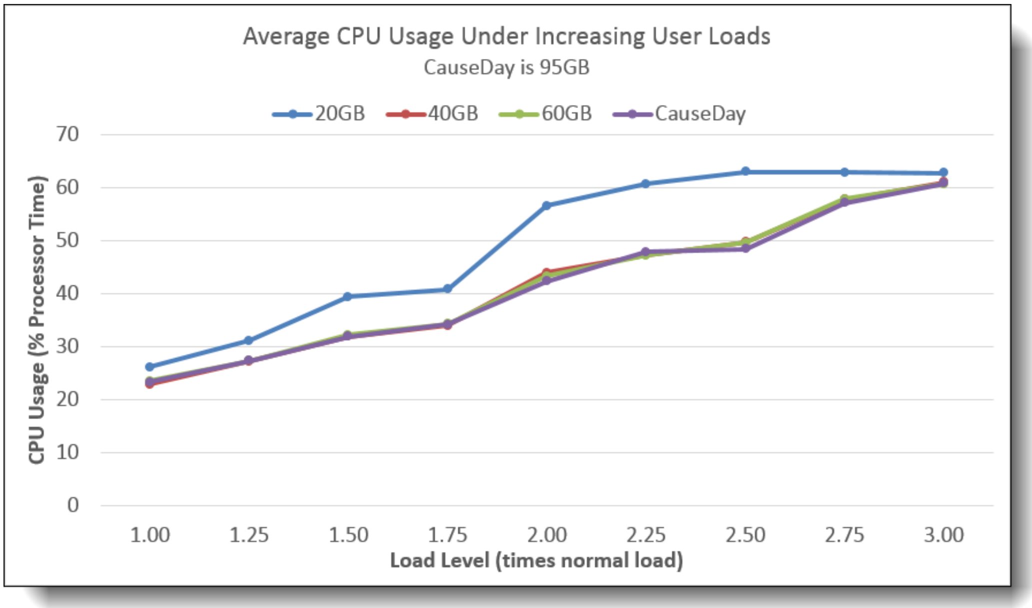

Figure 19

The tests measured average disk bytes per second for each memory variation. The following chart shows average disk bytes per second on the Data drive that stores ***Blackbaud CRM*** database data files for each memory setting. When the tests reduce memory to 40 GB, ***SQL Server*** can no longer keep essential data in memory and must read from disk more often. When memory is reduced further to 20 GB memory, this effect becomes catastrophic as the load increases and the I/O capacity of the drive is reached at an average of about 140 MB/s. As noted previously, the peak usage will be higher than the average.

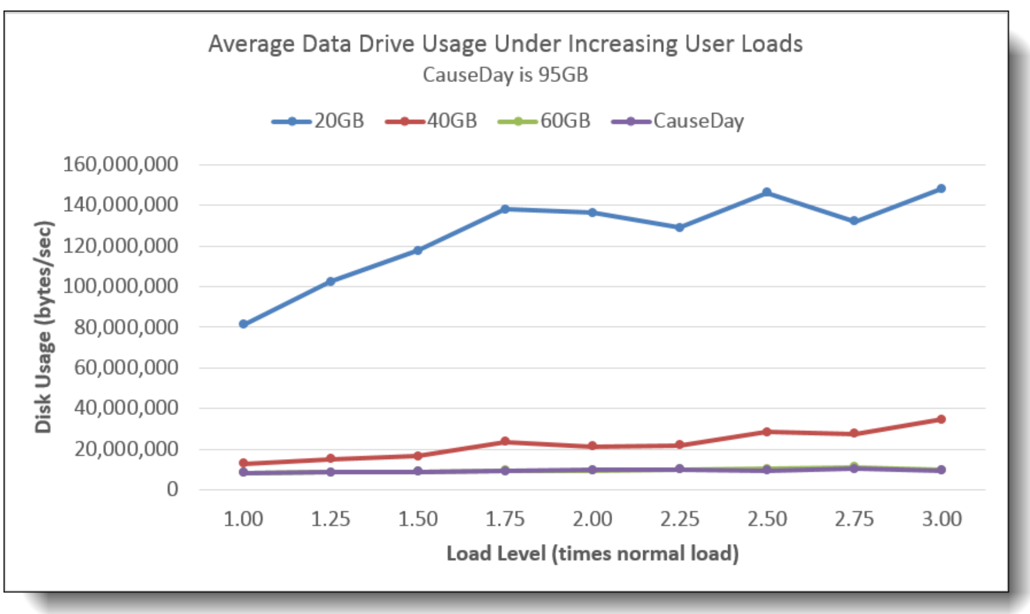

Figure 20

Large Cause Organization Nighttime Workloads

Key takeaways for this workload:

- 40 GB of memory is minimally sufficient for this workload, with a decrease in performance just beginning to occur.
- With insufficient memory, disk I/O increases dramatically, particularly for the Data drive.
- The increased disk I/O causes business processes to run longer.

For the Large Cause Organization Nighttime Workload, the tests measured the impact of varying amounts of memory available to ***SQL Server*** on the time that it takes to run after-hours business processes. The memory settings included 60 GB, 40 GB, and 20 GB.

The tests measured the impact of each memory variation on business process completion times. The chart below gives the sum of times to run the business processes in this workload at each memory variation. Reducing the memory from the baseline's 95 GB to 60 GB does not impact system performance. However, a slight performance degradation occurs when the memory is reduced to 40 GB. And at 20 GB, a considerable increase in run times occurs.

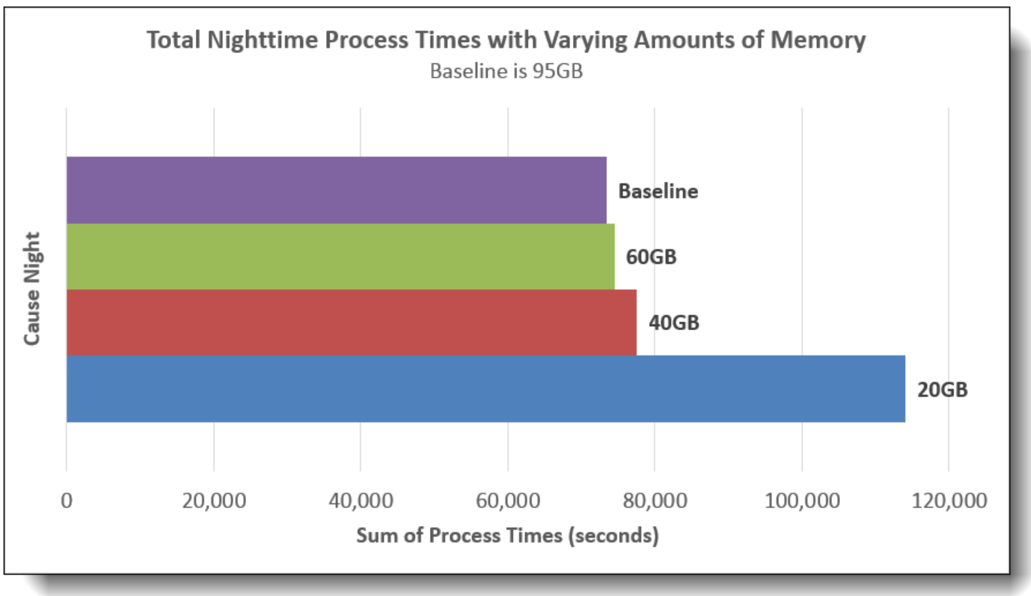

Figure 21

The tests measured the impact of each memory variation on CPU usage. The chart below illustrates the average CPU usage to run the workload at each memory variation. In a nutshell, CPU usage does not change as the memory is increased or decreased. The average CPU usage appears to be lower at 20 GB memory, but the reason is that the work was spread out over a longer time. The same amount of work was done, but the jobs took longer to complete as shown above.

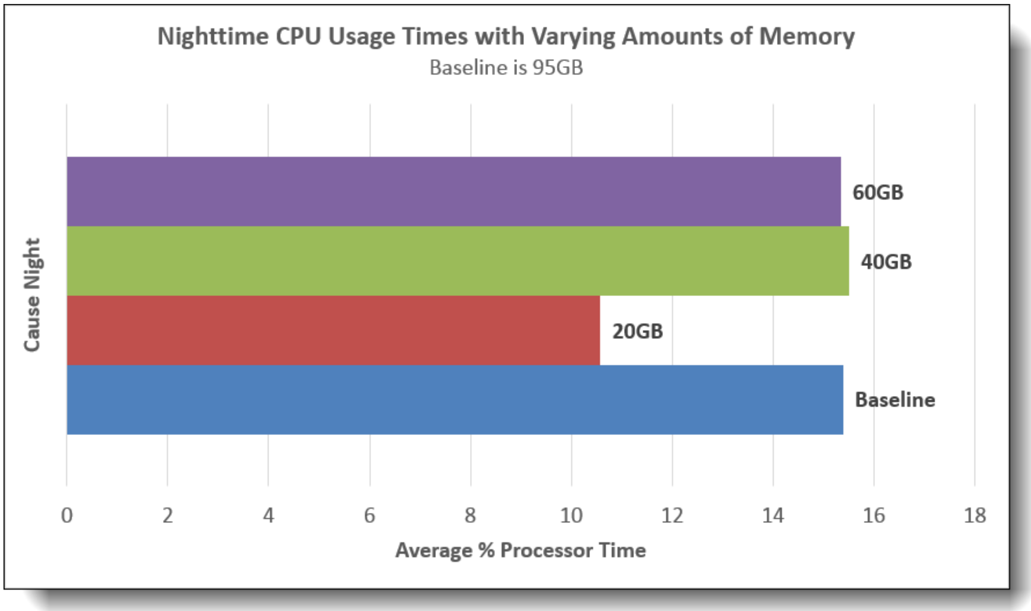

Figure 22

The tests measured the impact of each memory variation on average disk bytes per second. The chart below illustrates the average I/O at each memory variation for the Data drive with ***Blackbaud CRM*** database data files, the Log drive with database log files, and the Temp drive with the TempDB data and log files. As expected, disk usage increases as the memory is reduced, particularly on the Data drive. As the amount of I/O on the Data drive increases, the system slows down, which corresponds to the longer total run time as memory decreases. The apparent drop in Temp and Log drive activity at the 20 GB level reflects that the system took longer to do the work, which resulted in a lower average similar to the results for the CPU usage.

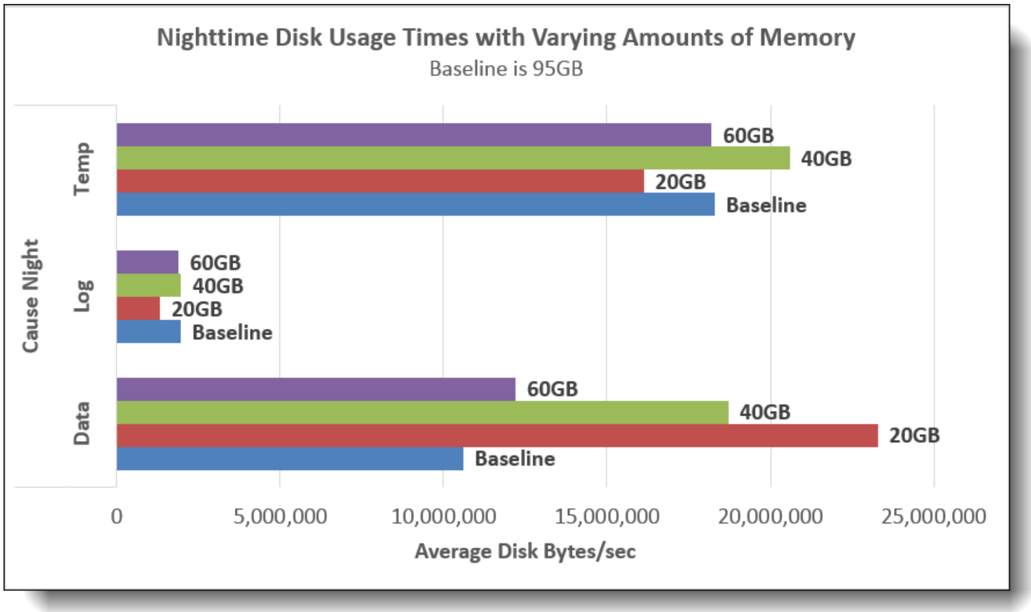

Figure 23

Midsize Higher Education Foundation Daytime Workload

Key takeaways for this workload:

- 15 GB of memory is sufficient for this workload; 10 GB is barely sufficient and at the lowest load levels only.
- With insufficient memory, response times increase dramatically.
- With insufficient memory, disk I/O increases dramatically, particularly for the Data drive.

To measure the impact of increased load for different memory variations at the midsize higher education foundation, workloads ran for multiple memory settings and increased the number of users linearly. Memory settings included 20 GB, 15 GB, and 10GB. The tests started with the baseline workload of 1.0 times the normal number of users and increased the load linearly by a factor of 0.5 until it reached 6.5 times the normal number of users.

The chart below illustrates the impact of increased load on user response times for each memory variation. Even 10 GB is minimally sufficient to run the ***Blackbaud CRM*** environment at normal load levels, but the low amount of memory has a dramatic effect on response times as the load increases. (Note that performance deteriorated to the point that tests were not run with 10 GB memory after the 5x load level.)

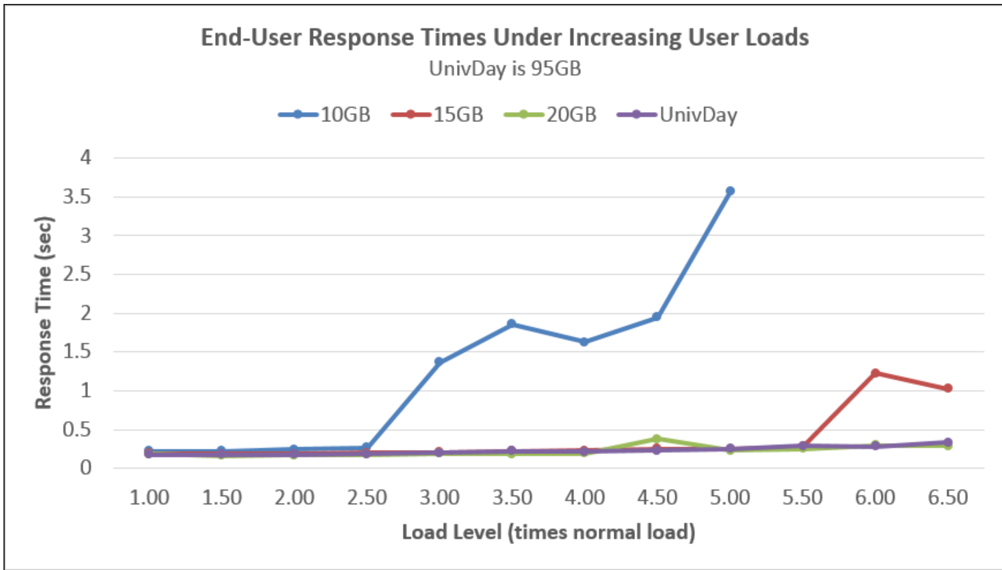

Figure 24

The amount of memory in the system did not impact CPU usage. In the interest of brevity, the chart was omitted.

The following chart shows average disk bytes per second on the Data drive as the load increases for each memory setting. As with the Large Cause Organization Daytime Workload, the increased load on the Midsize Higher Education Foundation Daytime Workload primarily stresses the Data drive that stores ***Blackbaud CRM*** database data files. When the system does not have enough space to hold data in memory, it must request it repeatedly from disk. Activity on the Log drive with database log files remains minimal and Temp drive activity is also significantly less than the Data drive, so these are left off the chart.

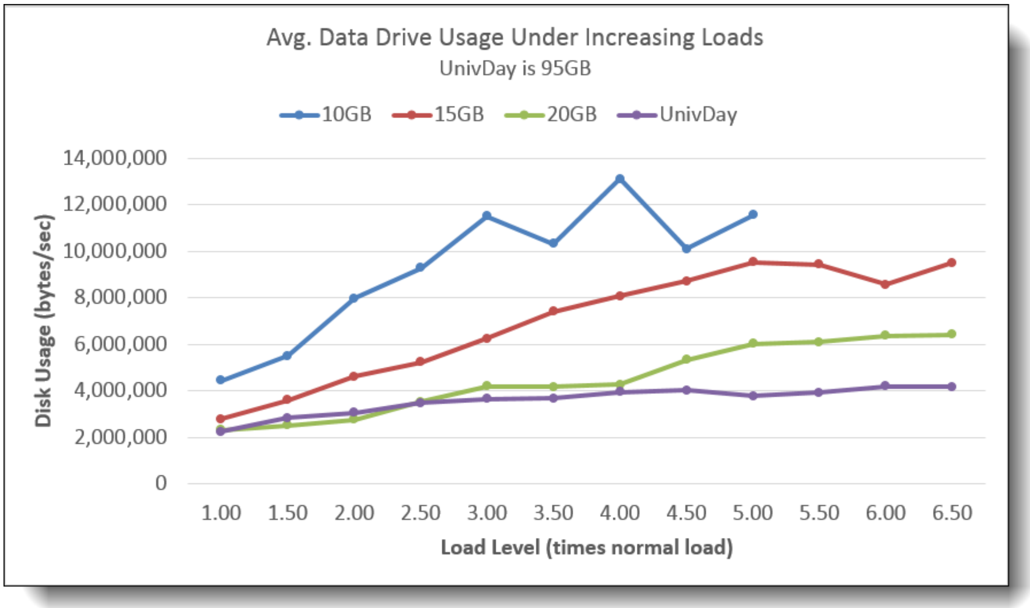

Figure 25

As noted previously, the peak usage will be higher than these averages.

Midsize Higher Education Foundation Nighttime Workload

Key takeaways for this workload:

- 20 GB of memory is sufficient for this workload; at 15 GB, a deterioration of performance is noticeable.
- With insufficient memory, disk I/O increases, particularly for the Data drive.

For the Midsize Higher Education Foundation Nighttime Workload, the tests measured the impact of varying amounts of memory available to ***SQL Server*** on the time that it takes to run after-hours business processes. Memory settings included 20 GB, 15 GB, and 10 GB.

The tests measured the impact of each memory variation on business process completion times. The chart below gives the sum of times to run the business processes in this workload at each memory variation. It shows that the total run time for the nighttime business processes increases with each decrease in the available memory. The slowdown is not dramatic, but it shows that extra effort is required to accomplish the work.

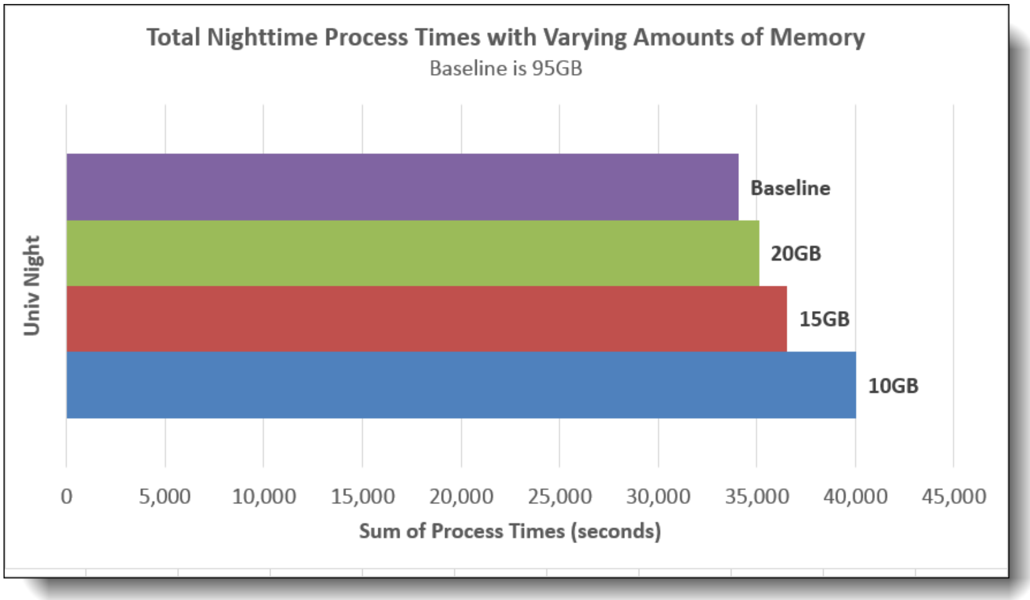

Figure 26

As noted previously, increasing the number of CPU cores won't improve the performance of nighttime workloads that include a small number of large processes that don't scale across multiple cores. The chart is omitted for brevity.

The tests measured the impact of each memory variation on average disk bytes per second. The chart below illustrates the average I/O at each memory variation for the Data drive with ***Blackbaud CRM*** database data files, the Log drive with database log files, and the Temp drive with the TempDB data and log files. As expected, disk usage once again increases as memory is reduced, with the increase in I/O being most significant for the Data drive.

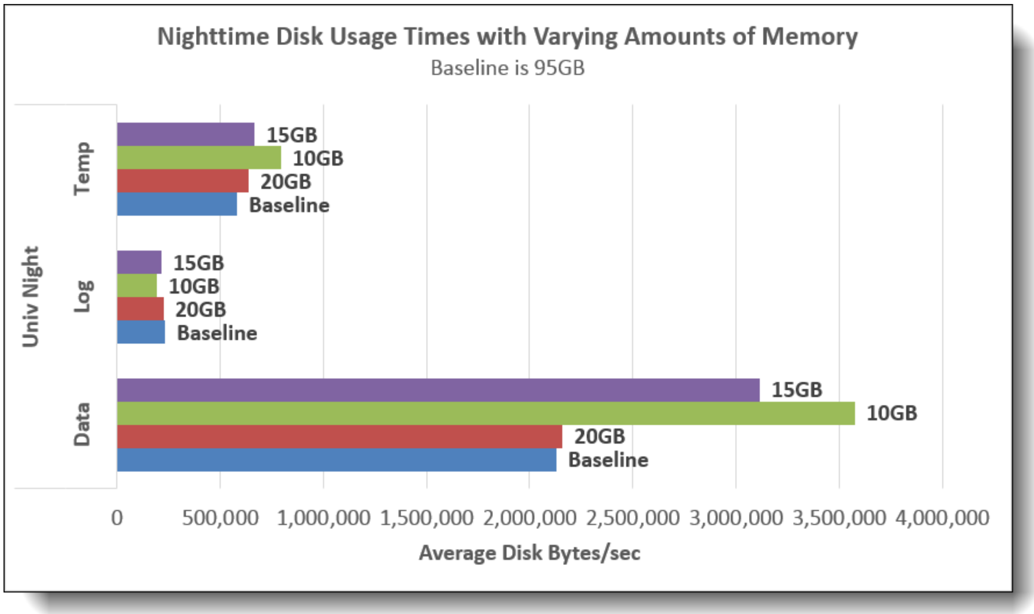

Figure 27

Appendix F – Varying CPU Cores Test Results

This appendix describes the detailed results from testing ***Blackbaud CRM*** environments with a varying number of logical processors in the system. The tests started with baseline of 12 logical processors, and then changed the number of available processors to 16, 8, and 4. The tests did not include hyperthreading, so one logical processor is the same as one core.

Large Cause Organization Daytime Workloads

Key takeaways for this workload:

- Eight cores is sufficient for this workload; 16 provides good performance at the highest load levels.
- As predicted by Little's Law, it's best to keep average CPU usage below about 50 percent.

To measure the impact of increased load for different CPU variations at the large cause organization, multiple workloads ran for each CPU setting and increased the number of users linearly. The tests started with the baseline workload of 1.0 times the normal number of users and increased the load linearly by a factor of 0.25 until it reached 3.0 times the normal number of users.

The tests measured the impact of increased load on user response times for each CPU variation. As the chart below indicates, 4 cores is insufficient for good response time with this workload at almost any load level.

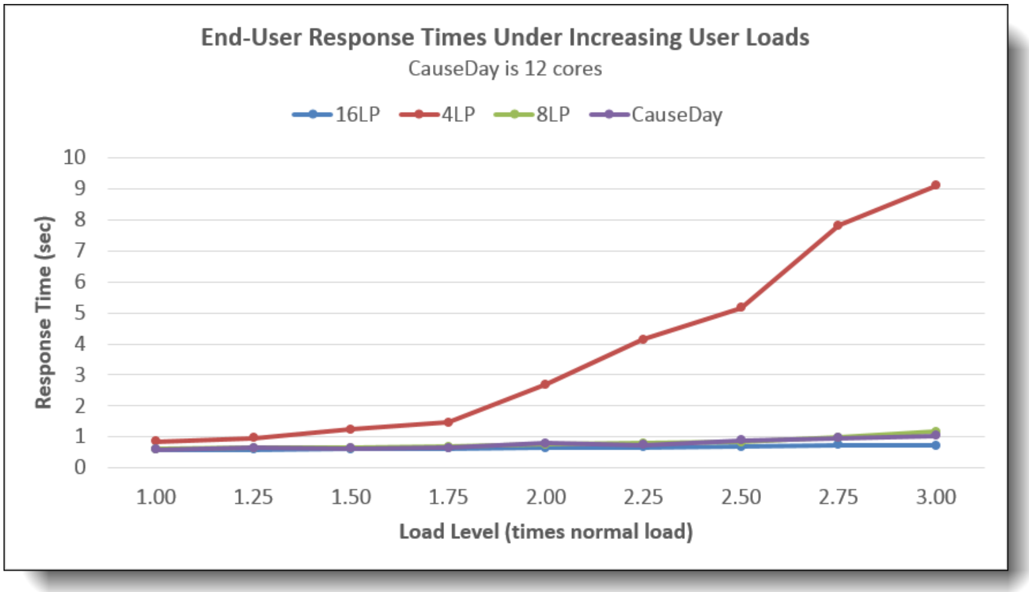

Figure 28

The next chart excludes the line for 4 cores to illustrate the other results in more detail. The chart shows that 8 or 12 cores work fine but that at the highest load levels, some advantage occurs with 16 cores.

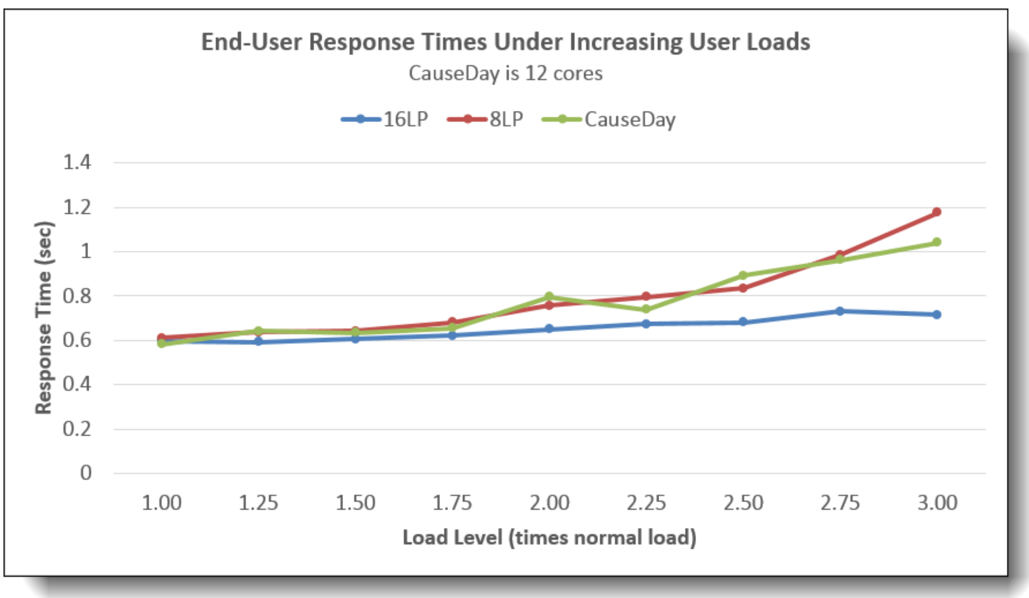

Figure 29

The charts on processor usage provide a beautiful example of Little's Law. With 4 cores, the system starts out in the response-time danger zone of more than 50 percent. With 8 and 12 cores, more capacity is available before the load reaches about 50 percent processor time. With 16 cores, the highest load level with this workload does not push the processor time into the danger zone.

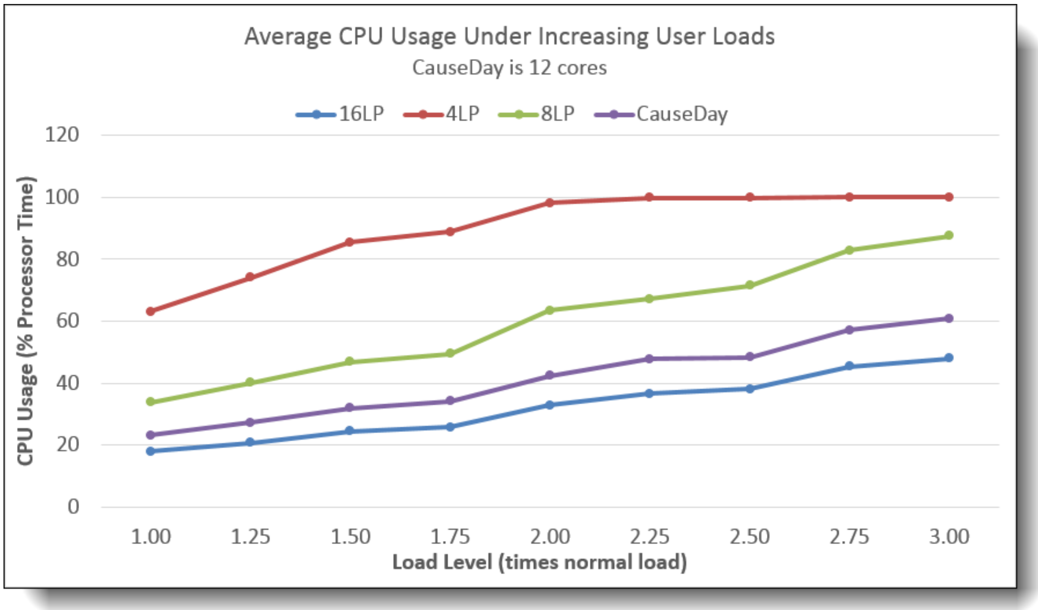

Figure 30

As for disk I/O, this workload doesn't show much. I/O was not the bottleneck and that didn't change for the CPU variations.

Large Cause Organization Nighttime Workloads

The nighttime workload is not limited by the number of cores. The serialization of SQL statements that run during nighttime workloads limits the need for CPU cores. Adjusting the number of available cores did not affect performance, and the tests found that even 4 cores was sufficient to achieve good response times during nighttime workloads.

Midsize Higher Education Foundation Daytime Workloads

Key takeaways for this workload:

- Four cores is sufficient for this workload at normal load levels; only the highest load levels require 12 or 16 cores.
- As predicted by Little's Law, it's best to keep average CPU usage below about 50 percent.

To measure the impact of increased load for different CPU variations at the midsize higher education foundation, multiple workloads ran for each CPU setting and increased the number of users linearly. The tests started with the baseline workload of 1.0 times the normal number of users and increased the load linearly by a factor of 0.5 until it reached 6.5 times the normal number of users.

The tests measured the impact of increased load on user response times for each CPU variation. As the chart below indicates, 4 cores is sufficient to get good response times with this workload until the load level reaches about 3 times the normal load.

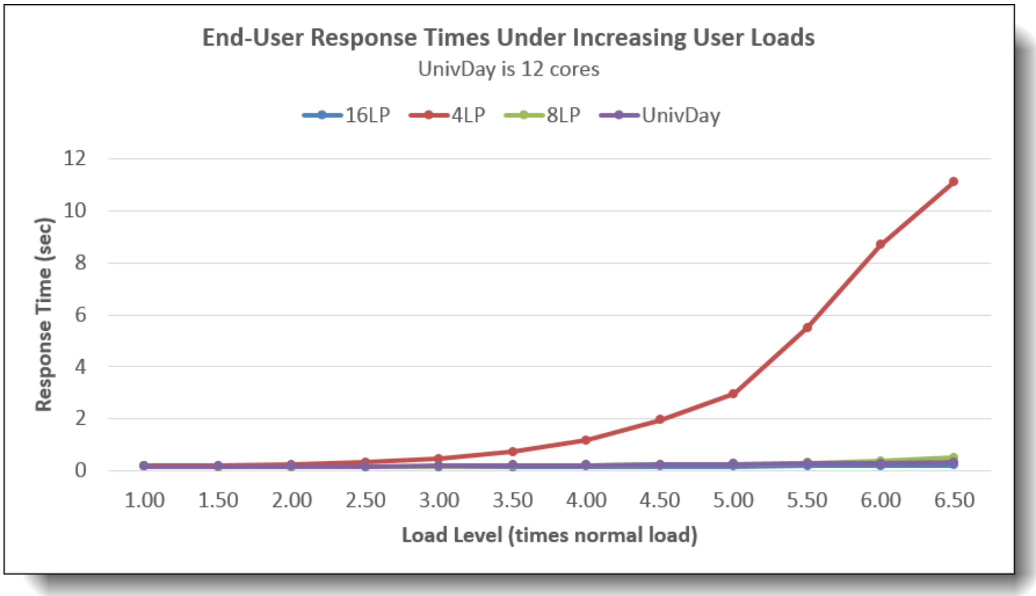

Figure 31

The next chart excludes the line for 4 cores to illustrate the other results in more detail. It shows that 8 or 12 cores work fine but that at the highest load levels, an advantage occurs with 16 cores.

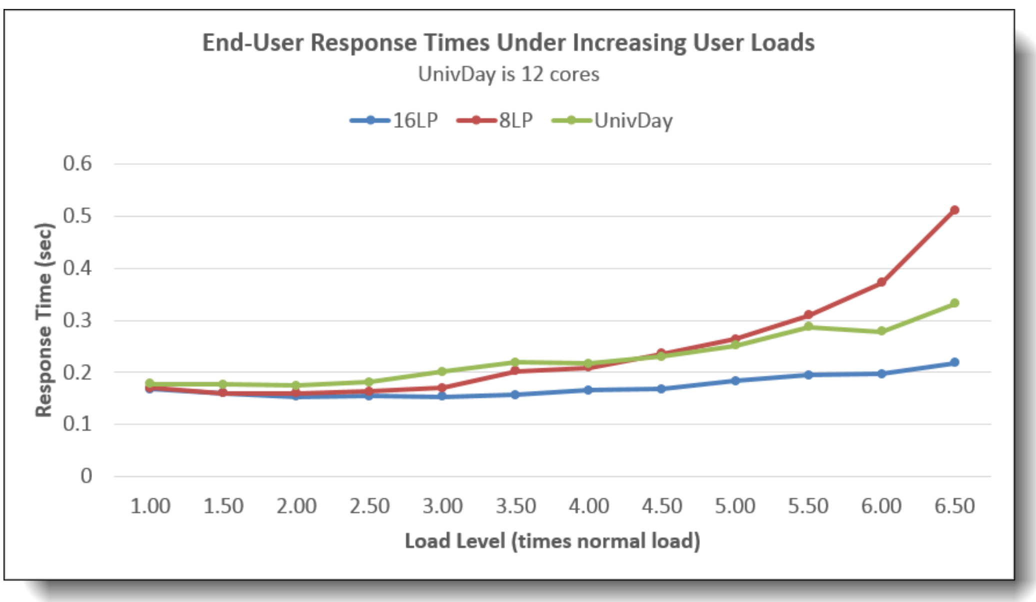

Figure 32

The charts on processor usage provide another example of Little's Law. With 4 cores, the system has sufficient resources to give good response until it reaches about 60 percent processor usage. After that, response times worsen. With more than 8 cores, even the highest load level with this workload does not push the processor time into the danger zone.

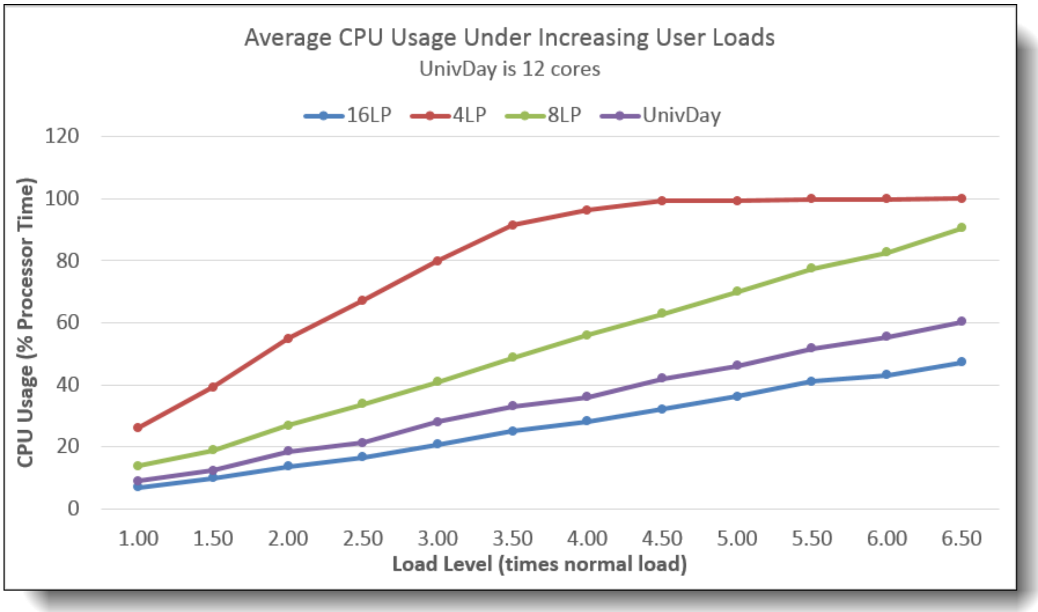

Figure 33

As for disk I/O, this workload doesn't show much. I/O was not the bottleneck and that didn't change for the CPU variations.

Midsize Higher Education Foundation Nighttime Workloads

The nighttime workload is not limited by the number of cores. The serialization of SQL statements that run during nighttime workloads limits the need for CPU cores. Adjusting the number of available cores did not affect performance, and the tests found that even 4 cores was sufficient to achieve good response times during nighttime workloads.

Appendix G – Spinning Disks vs. Solid State Disks Test Results

This appendix describes the detailed results from comparing spinning disks vs. solid state discs for ***Blackbaud CRM*** environments. The tests force I/O to become a bottleneck by decreasing the available memory. This drives I/O up to see the impact of SSDs vs. traditional disks. In these tests, the SSD holds data, log, and temp space for the database all on one drive. Separate SSD drives do not exist for data, log, and temp usage, so the reader must mentally combine the data, log, and temp measurements for spinning disks to compare to the SSD.

Large Cause Organization Daytime Workloads

Key takeaways for this workload:

- As the system requires more I/O, SSDs provide better read performance than spinning disks.
- SSDs for the Data drive can lessen the performance impact from insufficient memory.
- Because of the slower write speed of many SSDs, database Log files may be better placed on spinning disks, and it's likely that TempDB files should be as well.

When sufficient memory is available to run the workload, performance under this load is slightly better with SSDs than with spinning disks. However, the difference is not significant.

The chart below shows the average response times for user actions at the lowest memory condition tested (20 GB) and compares them to the baseline. Remember that 20 GB is a severe memory limit for this workload. As the load increases, the system slows down dramatically (as in the memory tests) because the lack of memory forces more disk reads to be done. However, when SSDs replace spinning disks, the workload runs better – not as fast as with no memory constraint, but decidedly better than with spinning disks.

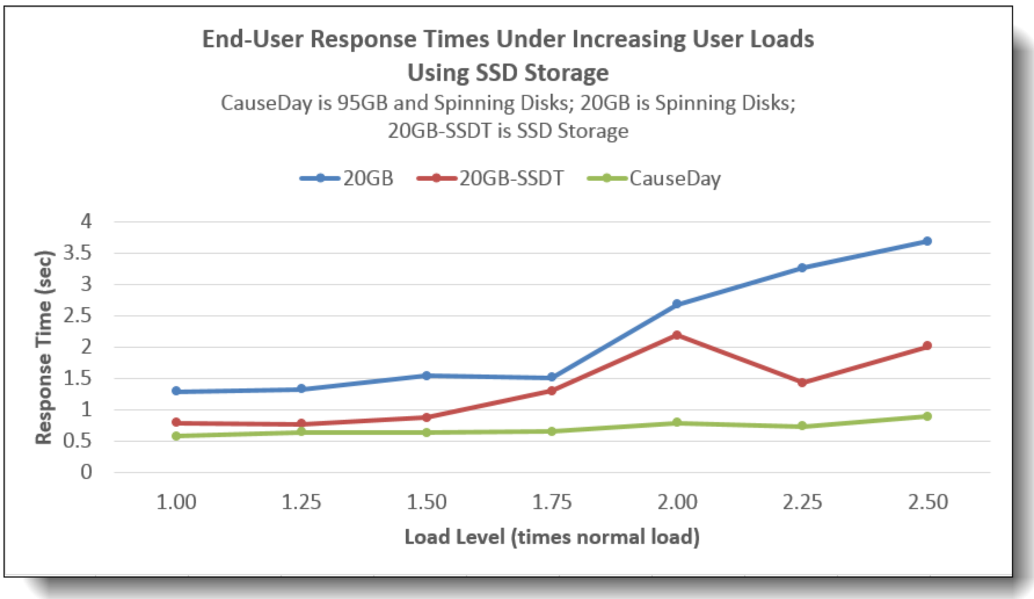

Figure 34

For processor usage, the processor time used is not changed by the use of SSD storage, even in limited memory cases. That chart is omitted for brevity.

As for I/O activity, the I/O to the SSD drive is the same as the I/O to the Data, Log, and Temp drives of a non-SSD configuration. This is the expected result, so the chart is omitted for brevity. One point to note though is that for ***SQL Server*** installations, the Log drive and Data drive should be different devices to ensure that data is not lost in the event of a device failure. The test configuration added more load to the SSD drive, but it did not follow recommended practices in that regard. In addition, SSD drives are commonly slower for write operations than for read operations. The SSD was slower to write than the spinning disks on this system, so for performance reasons, it might be better to put the log data on a spinning disk than an SSD. And depending on the read/write ratio, it might be better to have TempDB data on a spinning disk as well.

Large Cause Organization Nighttime Workloads

Key takeaway for this workload:

- SSD storage provides a modest improvement in the performance of this workload.

As observed earlier, the nighttime workload is primarily limited by the fact that it is serialized to a single execution thread. Speeding up the I/O subsystem shouldn't make much difference, and that is exactly what the results indicate. The much faster I/O subsystem made about a 15 percent difference in run times for the nighttime workload.

For processor usage, the processor time used is not changed by the use of SSD storage, even in limited memory cases. That chart is omitted for brevity.

As for I/O activity, the I/O to the SSD drive is the same as the I/O to the Data, Log, and Temp drives of a non-SSD configuration. This is the expected result, the chart is omitted for brevity. One point to note though is that because of the greater amount of TempDB activity in this workload, which is about evenly split between reads and writes, it might have performed better if TempDB was on a spinning disk rather than an SSD. This was not tested.

Midsize Higher Education Foundation Daytime Workloads

Key takeaway for this workload:

- SSD storage provides a modest improvement in the performance of this workload.

When sufficient memory is available to run the workload, performance under this load is somewhat better using SSDs than spinning disks, as the next chart shows. Because of the added expense of SSDs, this modest improvement may not be worthwhile.

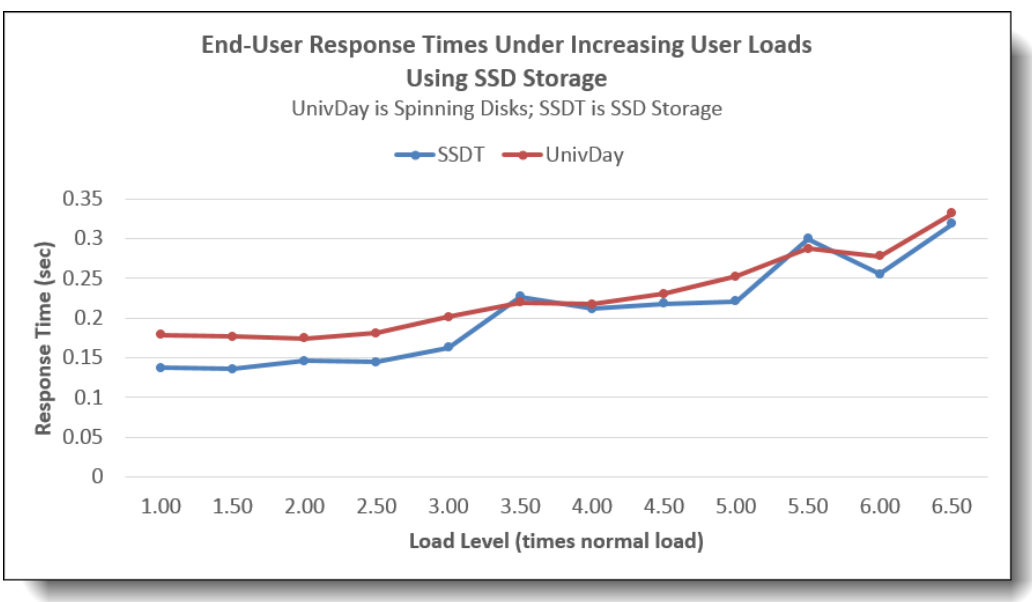

Figure 35

When the amount of memory was constrained, SSD storage kept response times acceptable for longer just like in the Large Cause Organization Daytime Workload. That chart is omitted for brevity.

Likewise, observations about CPU usage and I/O usage were the same as for the Large Cause Organization Daytime Workload, and the charts is omitted for brevity.

Midsize Higher Education Foundation Nighttime Workloads

Key takeaway for this workload:

- SSD storage provides an insignificant improvement in the performance of this workload.

As observed earlier, a nighttime workload is primarily limited by the fact that it is serialized to a single execution thread. Speeding up the I/O subsystem should not make much difference, and that is exactly what the results indicate. The run time difference here is even less significant than in the Large Cause Organization Nighttime Workload – in the area of 1 percent to 2 percent.

Likewise, observations about CPU usage and I/O usage were the same as for the Large Cause Organization Nighttime Workload, and the charts are omitted for brevity.