

Estimate performance and capacity requirements for Excel Services in Microsoft SharePoint Server 2010

This document is provided “as-is”. Information and views expressed in this document, including URL and other Internet Web site references, may change without notice. You bear the risk of using it.

Some examples depicted herein are provided for illustration only and are fictitious.  No real association or connection is intended or should be inferred.

This document does not provide you with any legal rights to any intellectual property in any Microsoft product. You may copy and use this document for your internal, reference purposes.

© 2010 Microsoft Corporation. All rights reserved.

# Contents

Contents

[Contents 2](#_Toc260403161)

[Test Farm Characteristic 3](#_Toc260403162)

[Dataset 3](#_Toc260403163)

[Workload 4](#_Toc260403164)

[Recommended and Max Definitions 4](#_Toc260403165)

[Hardware Settings and Topology 5](#_Toc260403166)

[Lab Hardware 5](#_Toc260403167)

[Topology 6](#_Toc260403168)

[Scale Out Test results 6](#_Toc260403169)

[Scale Up Test Results 11](#_Toc260403170)

[Recommendations 12](#_Toc260403171)

[Hardware recommendations 13](#_Toc260403172)

[Performance-Related Excel Services Settings 13](#_Toc260403173)

[Common bottlenecks and their causes 15](#_Toc260403174)

[See Also 17](#_Toc260403175)

|  |
| --- |
|  |
| Estimate performance and capacity requirements for Excel Services |
|  |

This performance and capacity planning document provides guidance on the footprint that usage of Excel Services has on topologies running Microsoft® SharePoint® Server 2010.

In this article:

* [Test farm characteristic](#_Test_Farm_Characteristic)
* [Test results](#_Test_results)
* [Recommendations](#section4)
* [Troubleshooting](#_Troubleshooting_performance_and)

For general information about how to plan and run your capacity planning for Microsoft® SharePoint® Server 2010, see [Capacity management and sizing for SharePoint Server 2010](http://technet.microsoft.com/en-us/library/cc261700(office.14).aspx).

## Test farm characteristic

### Dataset

Excel Services capacity and performance is highly dependent on the makeup of the workbooks that are hosted on the service. The size of the workbook and the complexity of calculations have the most impact. Our testing used representative sizes and complexities, but every workbook will be different, and your capacity and performance will be dependent on the actual workbooks in use, and their specific size and complexity.

To evaluate our capacity profile, we simulated Excel Services workbooks on a farm dedicated to Excel Services (no other SharePoint tests running). Within this farm, we used three buckets of workbooks, Small, Large, and Very Large, based on workbook size and complexity:

|  |  |  |  |
| --- | --- | --- | --- |
| Workbook Characteristics | Small | Large | Very Large |
| Sheets | 1-3 | 1-5 | 1-20 |
| Columns | 10-20 | 10-500 | 10-1,000 |
| Rows | 10-40 | 10-10,000 | 100-30,000 |
| Calculated Cells | 0-20% | 0-70% | 0-70% |
| Number of Formats | 1-10 | 1-15 | 1-20 |
| Tables | 0-1 | 0-2 | 0-5 |
| Charts | 0-1 | 0-4 | 0-4 |
| Workbook Uses External Data | 0% | 20% | 50% |
| Workbook Uses a Pivot Table | 0% | 3% | 3% |
| Workbook Uses Conditional Formats | 0% | 10% | 20% |

In terms of storage, 2,000 SharePoint sites were created on our test farm, each containing one small, one large, and one very large workbook.

In addition, our dataset included SharePoint pages that included from 1 to 5 Excel Web Parts. The distribution of workbooks included on the SharePoint pages: 10% Small Workbooks, 90% Large Workbooks.

### Workload

To simulate application usage, workloads were created which would perform one or more of the following operations:

|  |  |  |
| --- | --- | --- |
| Action Mix | Small Workbook | Large and Very Large Workbook |
| View | 50% | 70% |
| Edit | 35% | 15% |
| Collaborative Viewing | 10% | 10% |
| Collaborative Editing | 5% | 5% |

In addition, 17% of all the workbooks have external data. For large and very large workbooks that have external data, we do a refresh 80% of the time; small workbooks do not contain external data.

Each workload includes think time between user actions of 10 seconds. This is different from other SharePoint Server 2010 capacity planning documents. Excel Services is stateful – we maintain the workbook in memory between user interactions – making it important to simulate a full user session and not just individual requests. For a single user workload, there are on average .2 requests per second.

For each workload, one of the 2,000 sites were randomly selected to run the test. Within that site, to select which application and which size application to choose between, the following percentages were used:

|  |  |
| --- | --- |
| Workbook Selection | % Usage |
| Small Workbook | 30% |
| Large Workbook | 55% |
| Dashboard | 10% |
| Very Large Workbook | 5% |

### Recommended and Max Definitions

For each configuration we ran two tests to determine a “green zone” or Recommended throughput that can be sustained and a “red zone” or Max throughput that can be tolerated for a short period of time but should be avoided.

To determine our red and green zone user loads, we first conducted a step test and stopped when the following conditions were met:

* For the green zone, we stopped at the point when any of the machines in our farm (WFE, ECS, or SQL Server) exceeded 50% CPU or the response time for the overall system exceeded 1 second.
* For the red zone, we stopped at the point where the successful RPS for the ECS machines in the farm were at a maximum. Past this point, the overall throughput for the farm started to go down and/or we would start to see failures from one of the tiers. Often the ECS’ max private bytes would be exceeded when we hit the red zone.

After conducting the step tests, we backed off these maximum values to run a longer constant load test of 1 hour. For the Green Zone, we used 75% of the load when we stopped the green zone test, and for the Red Zone, we used 65% of the load when we peaked in the step test. If the green zone test was limited by memory, and % CPU never exceeded 50%, we instead used 75% of the load number calculated for the red zone.

In for both Green and Red zones, and for both Scale Out and Scale Up tests, average response time was below .25 seconds.

## Hardware Settings and Topology

### Lab Hardware

To provide a high level of test-result detail, several farm configurations were used for testing. Farm configurations ranged from one to three front end Web servers, one to three application servers (in the case of Excel Services, ECS, or Excel Calculation Services), and a single database server computer that is running Microsoft SQL Server® 2008. In addition, testing was performed with 4 client computers. All server computers were 64-bit, and the client computers were 32-bit.

The following table lists the specific hardware that was used for our testing.

|  |  |  |  |
| --- | --- | --- | --- |
| Machine Role | CPU | Memory | Network |
| Front-end Web server | 2 proc/4 core 2.33 GHz Intel Xeon® | 8GB | 1 gig |
| ECS | 2 proc/4 core 2.33 GHz Intel Xeon® | 8GB | 1 gig |
| SQL Server | 4 proc/4 core 2.6GHz Intel Xeon® | 16GB | 1 gig |

### Topology

From our experience (and of course yours may vary), memory on the ECS tier and CPU on the WFE tier are the most important limiting factors for throughput. So we varied the number of machines in both the ECS and WFE tiers for scale out testing.

For scale up testing, a topology of 1 WFE to 1 ECS was used, and the number of processors and available memory in the ECS machine were varied.

ECS is not particularly hard on the SQL Server instance running SharePoint Server 2010, as the workbook is read a binary blob from SharePoint Server 2010 and placed in memory on the ECS tier (and further disk cached). At no time did SQL Server become a bottleneck.

### Scale Out Test results

The following tables show the test results of Excel Services in SharePoint Server 2010. For each group of tests, only certain specific variables are changed to show the progressive impact on farm performance.

Note that all the tests reported on in this article were conducted with think or wait time (see above for more details). This differs from the capacity planning results for other parts of SharePoint Server 2010.

For information about bottlenecks of Excel Services in SharePoint Server 2010, see the [Common bottlenecks and their causes](#bottlenecks) section later in this article.

#### Overall scale

The table below summarizes the impact of adding additional WFE and dedicated ECS machines to the farm. These throughput numbers are specifically for the ECS machines, and do not reflect the impact on the overall farm.

|  |  |  |
| --- | --- | --- |
| Topology | Baseline Max (RPS) | Baseline Recommended (RPS) |
| 1x1 | 38 | 31 |
| 1x2 | 35 | 26 |
| 1x3 | 28 | 21 |
| 2x1 | 57 | 35 |
| 2x2 | 62 | 46 |
| 2x3 | 52 | 39 |
| 3x1 | 51 | 32 |
| 3x2 | 81 | 69 |
| 3x3 | 83 | 64 |

#### Recommended Results

The following shows our results for recommended sustainable throughput.

As can be seen above, there is overhead associated with adding WFE machines to the farm, which is made up for as ECS machines are added. 1 WFE became the bottleneck when moving to 2 ECS machines, nullifying the additional capacity when adding a second and third ECS. Also note that 3 WFE did not add any more throughput, as ECS became the limiting factor.

As WFE machines are added, CPU load on each machine is reduced significantly. Note that 2 WFE x 3 ECS CPU load is reaching the maximum seen for 1 WFE, implying that the addition of another ECS would make the WFE tier the limiting factor. Remember that these results are for the “recommended” load, which is why CPU is maxing out at around 35% instead of at a higher level.

#### Maximum Results

The following shows our results for maximum peak throughput.

Similar to our recommended results, we see that 1 WFE is the limiting factor as we add a second and third ECS machine. Also note that, as with the recommended results, adding a third WFE does not add to our throughput as ECS is the limiting factor after the second WFE has been added.

In these results, we see that multiple WFEs are not becoming as heavily loaded as a single WFE configuration, implying that the ECS machines are the bottleneck after the second WFE has been added.

### Detailed Results

##### Recommended Results

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Overall** | **1x1** | **1x2** | **1x3** | **2x1** | **2x2** | **2x3** | **3x1** | **3x2** | **3x3** |
| Client Successful RPS | 30.56 | 34.55 | 31.67 | 26.03 | 45.94 | 68.37 | 20.71 | 38.82 | 63.70 |
| Client Response Time (sec) | 0.22 | 0.18 | 0.19 | 0.16 | 0.19 | 0.20 | 0.15 | 0.15 | 0.17 |
| TPS | 1.58 | 1.77 | 1.61 | 1.40 | 2.38 | 3.54 | 1.08 | 2.03 | 3.25 |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| **WFE Tier** | **1x1** | **1x2** | **1x3** | **2x1** | **2x2** | **2x3** | **3x1** | **3x2** | **3x3** |
| % CPU (average over all WFE machines) | 33.73 | 37.64 | 33.84 | 14.61 | 23.95 | 36.90 | 7.54 | 13.12 | 21.75 |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| **ECS Tier** | **1x1** | **1x2** | **1x3** | **2x1** | **2x2** | **2x3** | **3x1** | **3x2** | **3x3** |
| % CPU (average over all ECS machines) | 21.05 | 11.04 | 7.08 | 17.31 | 14.88 | 17.42 | 13.65 | 13.03 | 13.69 |
| Peak Private Bytes (max over all ECS machines) | 5.94E+09 | 5.82E+09 | 5.79E+09 | 5.87E+09 | 6.09E+09 | 5.92E+09 | 5.79E+09 | 5.91E+09 | 5.85E+09 |

##### Maximum Results

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Overall | 1x1 | 1x2 | 1x3 | 2x1 | 2x2 | 2x3 | 3x1 | 3x2 | 3x3 |
| Client Successful RPS | 37.85 | 56.70 | 51.17 | 35.19 | 62.04 | 81.31 | 27.79 | 51.62 | 82.58 |
| Client Response Time (sec) | 0.19 | 0.28 | 0.23 | 0.16 | 0.20 | 0.25 | 0.16 | 0.16 | 0.22 |
| TPS | 1.92 | 2.96 | 2.59 | 1.81 | 3.21 | 4.60 | 1.41 | 2.72 | 4.30 |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| WFE Tier | **1x1** | **1x2** | **1x3** | **2x1** | **2x2** | **2x3** | **3x1** | **3x2** | **3x3** |
| % CPU (average over all WFE machines) | 41.08 | 67.78 | 58.59 | 19.44 | 34.11 | 45.97 | 10.19 | 17.79 | 28.69 |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| ECS Tier | **1x1** | **1x2** | **1x3** | **2x1** | **2x2** | **2x3** | **3x1** | **3x2** | **3x3** |
| % CPU (average over all ECS machines) | 24.99 | 18.44 | 10.96 | 23.57 | 20.56 | 17.77 | 18.97 | 17.04 | 18.10 |
| Peak Private Bytes (max over all ECS machines) | 5.91E+09 | 5.85E+09 | 5.91E+09 | 5.88E+09 | 5.99E+09 | 6.50E+09 | 5.94E+09 | 5.94E+09 | 6.04E+09 |

### 

### Scale Up Test Results

We also measured the impact of adding CPUs and memory to the ECS tier. For these tests, a 1x1 topology was used.

Our results above show that adding additional CPUs, although of course helping, is not having a significant impact on overall throughput.

However, adding memory does have a significant impact on throughput, especially at peak times as indicated by the red zone line. In this test, the same hardware was used throughout, but the Maximum Private Bytes for the Excel Services process was limited. Since workbooks are kept in memory, the size of the workbooks will have a significant impact on how many workbooks, and thus how many users, a given ECS machine can handle.

## Recommendations

This section provides general performance and capacity recommendations.

Note that Excel Services capacity and performance is highly dependent on the makeup of the workbooks that are hosted on the service. The size of the workbook and the complexity of calculations have the most impact. Our testing used representative sizes and complexities, but every workbook will be different, and your capacity and performance will be dependent on the actual workbooks in use, and their specific size and complexity.

### Hardware recommendations

Excel Services uses standard hardware for both Web Front Ends and Application Servers – no special requirements are necessary. General SharePoint Server 2010 guidelines on CPU number, speed, and memory are applicable for machines in the ECS tier. However be aware that one of the first bottlenecks an ECS machine is likely to run into is memory – this may be a place to add additional resources. But before doing so, we recommend testing with a representative set of workbooks from your organization, as the size and complexity of the workbook will have a large impact on how much more capacity the addition of memory is likely to have.

To increase the capacity and performance of one of the starting-point topologies, you can do one of two things. You can either scale up by increasing the capacity of your existing server computers or scale out by adding additional servers to the topology. This section describes the general performance characteristics of several scaled-out topologies.

The sample topologies represent the following common ways to scale out a topology for an Excel Services scenario:

* To provide for more user load, check the CPU and memory for the existing Excel Services application servers. Add additional memory if CPU is not a concern, or add additional CPU if memory is not a concern. If both memory and CPU are reaching their upper bounds, additional ECS machines may be warranted. This can be done to the point that the WFE’s become the bottleneck, and then add additional WFE machines as needed.
* In our tests, SQL Server was not a bottleneck. Excel Services does not make large demands on the database tier, as workbooks are read/written as whole documents, and workbooks are held in memory throughout the user’s session.

### Performance-Related Excel Services Settings

One of the ways to control the performance characteristics of Excel Services is to control how memory is utilized. Each of the following can be set through SharePoint Central Administration > Application Management: Manage Service Applications > Excel Services > Global Settings:

* Maximum Private Bytes

By default, ECS will use up to 50% of the memory on the machine. If the machine is shared with other services, it may make sense to lower this number. If the machine is not being shared, is dedicated to ECS, and is showing signs that memory may be a limiting factor, then increasing this number may make sense. In any case, experimenting by adjusting this number can guide the administrator to making the necessary changes in order to better scale.

* Memory Cache Threshold

ECS will cache unused objects (for example, read only workbooks for which all sessions have timed out) in memory. By default, ECS will use 90% of the “Maximum Private Bytes” for this purpose. Lowering this number can improve overall performance if the machine is hosting other services besides ECS. Increasing this number increases the chances that the workbook being requested will already be in memory and will not need to be to reloaded from the SharePoint content database.

* Maximum Unused Object Age

By default, ECS will keep objects in the memory cache as long as possible. To reduce the ECS footprint, in particular with other services running on the same machine, it may make more sense to impose a cap on how long objects are cached in memory.

There are also throttles available to control the maximum size of a workbook and the lifetime of a session, which in turn controls how long a workbook is held in memory. These settings are scoped to each trusted location and are not global. These settings can be set through SharePoint Central Administration > Application Management: Manage Service Applications > Excel Services > Trusted Locations, and then Edit the settings for each trusted locations.

* Maximum Workbook Size
* Maximum Chart or Image Size

By default, ECS is limited to 10 MB or smaller workbooks and 1 MB or smaller charts/images. Obviously allowing larger workbooks and larger charts/images puts more strain on the available memory of the ECS tier machines. However, there may be users in your organization that need these settings to be increased in order for ECS to work with their particular workbooks.

* Session Timeout

By decreasing the session timeout, memory is freed for either the unused object cache or other services faster.

* Volatile Function Cache Lifetime

Volatile functions are functions that can change their value with each successive recalculation of the workbook, for example date/time functions, random number generators, etc. Because of the load this could generate on the server, ECS does not recalculate these values for each recalculation, instead caching the last values for a short amount of time. Increasing this lifetime can reduce the load on the server, but this is dependent on having workbooks that use volatile functions.

* Allow External Data

ECS has the ability to draw on external data sources. However, the time required to draw upon the external source can be significant, with potentially a large amount of data returned. If external data is allowed, there are a number of additional settings that can help throttle the impact of this feature.

### Common bottlenecks and their causes

During performance testing, several different common *bottlenecks* were revealed. A bottleneck is a condition in which the capacity of a particular constituent of a farm is reached. This causes a plateau or decrease in farm throughput.

The following table lists some common bottlenecks and describes their causes and possible resolutions.

#### Troubleshooting performance and scalability

|  |  |  |
| --- | --- | --- |
| Bottleneck | Cause | Resolution |
| ECS Memory | Excel Services holds each workbook in memory throughout the user’s session. A large number of workbooks, or large workbooks, can cause ECS to consume all available memory causing the actually consumed “Private Bytes” to exceed “Maximum Private Bytes.” | Scale Up with more memory in the ECS tier machines, or Scale Out with the addition of more ECS boxes. The choice will partially depend on if CPU is also reaching a maximum. |
| ECS CPU | Excel Services can be dependent on a large amount of processing in the application tier, depending on the number and complexity of workbooks. | Increase the number of CPUs and/or cores in the existing ECS boxes, or add additional ECS boxes. |
| Web server CPU utilization | When a Web server is overloaded with user requests, average CPU utilization will approach 100 percent. This prevents the Web server from responding to requests quickly and can cause timeouts and error messages on client computers. | This issue can be resolved in one of two ways. You can add additional Web servers to the farm to distribute user load, or you can scale up the Web server or servers by adding higher-speed processors. |

#### Performance monitoring

To help you determine when you have to scale up or scale out your system, use performance counters to monitor the health of your system. Use the information in the following tables to determine which performance counters to monitor, and to which process the performance counters should be applied.

##### Front-end Web server

The following table shows performance counters and processes to monitor for front-end Web servers in your farm.

|  |  |  |
| --- | --- | --- |
| Performance counter | Apply to object | Notes |
| % Processor time | Processor(w3wp) | Shows the percentage of elapsed time that this thread used the processor to execute instructions. |
| % Processor Time | Processor(\_Total) | Shows the percentage of elapsed time that all threads on the machine used the processor to execute instructions. |
| Private Bytes | Process(w3wp) | This value should not approach the Max Private Bytes set for w3wp processes; if it does, further investigation is needed into what component is consuming the memory. |

##### Excel Calculation Services (ECS)

The following table shows performance counters and processes to monitor for Application Servers, or Excel Calculation Services (ECS) in this case, within your farm.

|  |  |  |
| --- | --- | --- |
| Performance counter | Apply to object | Notes |
| % Processor Time | Processor(\_Total) | Shows the percentage of elapsed time that all threads on the machine used the processor to execute instructions. |
| % Processor Time | Process(w3wp) | The ECS runs within its own w3wp process, and it will be obvious which w3wp process this is as it will be getting the bulk of the CPU time. |
| Avg. Disk Queue Length | PhysicalDisk(\_Total) | Watch for excessive disk writing due to logging. |
| Private Bytes | Process(w3wp) | Excel Services caches workbooks in memory, until the user’s session expires (the timeout for which is configurable). If a large amount of data is being processed through the ECS, then memory consumption for the ECS’ w3wp will rise. |

##### SQL Server

As we have previously described, Excel Services is relatively light on the SQL Server tier, as workbooks are read once into memory on the ECS tier for the duration of the user’s session. Follow general SharePoint Server 2010 guidelines for monitoring and troubleshooting of the SQL Server tier.