Cisco Business Edition 6000 and Cisco Business Edition 7000 Co-residency Policy Requirements

Published: August 14, 2015

In addition to Cisco Unified Communications (UC) applications sold with Cisco Business Edition 6000 (BE6000M, BE6000H, and BE6000S) and Cisco Business Edition 7000 (BE7000M and BE7000H), Cisco also allows the installation of a broader range of Cisco and third-party virtualized applications, subject to the conditions that are detailed in this document. This policy applies to any Cisco UCS server using embedded virtualization software licenses (either Cisco UC Virtualization Hypervisor that is shipped with BE6000 and BE7000 servers, or the Cisco UC Virtualization Foundation license that is sold separately), including all new and previously supplied BE6000M, BE6000H, BE7000M, and BE7000H servers.

For the purposes of this document, the following definitions apply:

- Business Edition applications means the Collaboration applications that are explicitly included in the BE6000 and BE7000 Solutions. These applications are preloaded on the server and are usually integrated with BE6000 starter licensing. Business Edition applications that are preloaded on BE6000M, BE6000H, BE7000M, and BE7000H servers are as follows:

  - Cisco Unified Communications Manager
  - Cisco Unity Connection
  - Cisco Unified Provisioning Manager Business Edition (8.x and 9.x releases only)
  - Cisco Prime Collaboration Provisioning (10.0 and later releases only)
  - Cisco Unified Communications Manager IM and Presence Service
  - Cisco Unified Contact Center Express
  - Cisco TelePresence Video Communication Server
  - Cisco Expressway
  - Cisco TelePresence Conductor
  - Cisco TelePresence Server Virtual Machine
  - Cisco TelePresence Management Suite
  - Cisco Paging Server
  - Cisco Emergency Responder

- – Cisco Unified Attendant Console

- – Cisco TelePresence Content Server

- BE6000S only supports a subset of the above applications. Core Business Edition applications that are preloaded on a BE6000S server are as follows:

  - – Cisco Unified Communications Manager

  - – Cisco Unity Connection

  - – Cisco Unified Communications Manager IM and Presence Service

  - – Cisco Prime Collaboration Provisioning

  - – Cisco Paging Server

- Non-Business Edition applications means any of the following three application types:

  - – Other Cisco Collaboration applications listed at www.cisco.com/go/uc-virtualized (such as MediaSense) that are not explicitly listed above as part of the BE6000 or BE7000 solutions.

  - – Virtualized third-party applications that are included in the Solution Partner Program (SPP), formerly known as Cisco Developer Network (CDN) Marketplace Solutions Catalog for Collaboration. A list of all permitted third-party Collaboration applications can be found here. (Select **Technology =Collaboration**.)
    Note: You may only use only third-party applications from the Collaboration category with the Business Edition embedded hypervisor licenses.

  - – Virtualized third-party applications that are offered through the Cisco SolutionsPlus Program and complementary to Collaboration are described at the following URL:
    http://www.cisco.com/web/partners/pr46/solutions_plus/index.html

- Coresident means "running different Collaboration applications in dedicated virtual machines on the same virtualized Business Edition physical server or host."

Cisco only supports the Core Business Edition applications listed above for BE6000S servers. No other applications– whether Cisco or third-party – are supported with BE6000S at this time, even if other virtualization software licenses are substituted.

On BE6000M, BE6000H, BE7000M, BE7000H servers (or any non-Business Edition Cisco UCS server) with embedded virtualization software licenses:

- Cisco supports all Business Edition applications listed above.

- Non-Business Edition applications are allowed if all rules in this document are followed. Note that Cisco TAC support is only for products purchased from Cisco with a valid, active maintenance contract.

For more details on supported Non-Business Edition applications, see the application links in the "At a Glance - Cisco Collaboration Virtualization Support" table at the following URL:

www.cisco.com/go/uc-virtualized

For permitted third-party applications, there is a maximum number of virtual machines allowed in a BE6000, BE7000, and Cisco UCS server deployment using Cisco Unified Communications Virtualization Hypervisor or Cisco Unified Communications Virtualization Foundation licensing.

In a BE6000 or BE7000 deployment with one physical server, up to three third-party virtual machines may run on the server. For larger deployments, a maximum of three times the number of physical servers is permitted. The allowed quantity of third-party virtual machines can be deployed across physical servers in any combination. For example, with two physical servers, the six virtual machines can be distributed evenly across both, all, or one physical server.

Not all UC applications support co-residency, or they may have limited support of co-residency. See the "Application Co-residency Support Policy" for details:

http://docwiki.cisco.com/wiki/Unified_Communications_Virtualization_Sizing_Guidelines#Application_Co-residency_Support_Policy.

All other general UC virtualization rules apply—for example, VMware feature support and supported ESXi versions. See the following URL:

www.cisco.com/go/uc-virtualized.

All non-Business Edition applications must be qualified to run virtualized on VMware and must align with the virtualization software requirements for Cisco Collaboration that are outlined at the following URL:

http://docwiki.cisco.com/wiki/Unified_Communications_VMware_Requirements.

All applications must support ESXi 5.5 as a minimum and align with supported versions of Business Edition and Non-Business Edition Collaboration applications.

Customers that run a coresident deployment that includes third-party non-Business Edition applications must agree to temporarily reduce the number of virtual machines that are running on a host if Cisco deems it necessary for debugging purposes.

Customers must permanently reduce the number of virtual machines that are running on a host if Cisco determines that the host is overloaded.

If a customer is unwilling to agree to these requirements, Cisco TAC will not support the coresident deployment.

Support for third-party applications is provided by the vendor of the individual application.

Planning for Co-Residency

When planning a coresident deployment, you must consider four areas: CPU, RAM, storage and network.

For details on virtual to physical sizing rules in a co-residency context, see the following URL:

http://docwiki.cisco.com/wiki/Unified_Communications_Virtualization_Sizing_Guidelines.

The sizing rules refer to the "Tested Reference Configuration" hardware support approach, described at the following URL:

http://docwiki.cisco.com/wiki/UC_Virtualization_Supported_Hardware.

RAM

To ensure there is no over subscription of memory, you must set up all coresident virtual machines with a physical memory reservation equivalent to the vRAM setting. For example, if a virtual machine is configured with 4GB of vRAM, you must assign a physical memory reservation of 4GB.

To virtualize a BE6000 or BE7000 server, the hypervisor requires physical memory to host and run the virtual machines. To ensure that virtual machines have sufficient resources, this memory overhead must be taken into account to avoid resource oversubscription. ESXi 5.0 and 5.1 hosts must reserve 2 GB RAM for this overhead. ESXi 5.5 hosts must reserve 4 GB RAM.

Note The overhead reservation by ESXi hosts is not applicable to BE6000S. As BE6000S is a special configuration with deployment model restrictions, it does not ship with or require extra memory for ESXi as described above for other Business Edition models.

For example, if the BE6000 host running ESXi 5.1 has 32GB of physical RAM, only 30GB of this is available to virtual machines. For more information, see "Understanding Memory Overhead" at the following URL:

http://pubs.vmware.com/vsphere-51/index.jsp?topic=%2Fcom.vmware.vsphere.resmgmt.doc%2FGUID-98BD5A8A-260A-494F-BAAE-74781F5C4B87.html

CPU

A BE6000 deployment must have a one-to-one allocation of vCPUs to physical cores. For example, if you have a host with 16 physical cores, you can deploy any combination of virtual machines with a combined requirement of no more than 16 vCPUs. There is a special case for ESXi 5.0 and 5.1 if one or more of the virtual machines are running Cisco Unity Connection. In this case, you must reserve one vCPU for ESXi, leaving the remaining 15 vCPUs for the installed virtual machines. For servers using ESXi5.5, you may configure Unity Connection virtual machines to use High Latency Sensitivity. In this case and provided that at least one other virtual machine is configured with Normal Latency Sensitivity, you need not reserve a vCPU for ESXi.

Because the number of vCPUs must not exceed the number of physical cores, you do not have to configure CPU reservations or limits. Under no circumstances may a physical core be oversubscribed.

Note: Some processors support Hyperthreading, which allows a hypervisor to see a physical core as two logical processors. Logical processors must not be used for co-residency planning.

For more details, see the "No Hardware Oversubscription" section at the following URL:

http://docwiki.cisco.com/wiki/Unified_Communications_Virtualization_Sizing_Guidelines.

Disk Space and Storage Performance

The server Direct Attached Storage (DAS) must supply the combined disk space and IOPS (Input Output Operations per Second) capacity for the virtual machines that will run on the host, while maintaining a minimum performance level.

It is unlikely that the latency requirements of the Business Edition applications will limit third-party applications. However, you must understand the latency and load requirements of the non-Business Edition applications before installation.

The DAS and RAID are configured during manufacturing, and no field changes are allowed to the BE6000 or BE7000 Unified Computing System Tested Reference Configuration (TRC). The BE6000 TRC is designed to meet the storage requirements of all BE6000 collaboration applications. The BE7000 TRC is designed to meet the storage requirements of higher capacity points of these applications as described at the following URL:

www.cisco.com/go/uc-virtualized

To ensure the reliable operation of Business Edition applications, disk latency must not exceed 20 ms.

Cisco recommends that deployments that include non-Business Edition applications be verified such that the kernel command latency does not exceed 3 ms and the physical device command latency does not exceed 20 ms under any conditions. For more details, see "Sizing Shared Storage" at the following URL:

http://docwiki.cisco.com/wiki/UC_Virtualization_Supported_Hardware#Storage

For example, when you test a TRC with DAS, all UC applications that are designed to run on the TRC are loaded with full traffic and software upgrades are run simultaneously on all virtual machines. This generates the highest IOPS load possible on the host, and if this test passes, it is very likely the DAS array can handle the I/O load of the specific set of virtual machines.

To summarize the requirements, the disk subsystem must supply the disk space that is required for all the applications and support the aggregate IOPS load that the virtual machines generate, while not impacting the latency requirements of the UC applications. Otherwise, some virtual machines must be removed.

Network

The aggregate networking load of the coresident virtual machines must not exceed the capacity of the physical networking interfaces. Generally, the I/O capacity of the physical network resources on any modern server is more than adequate to meet the needs of the virtual machines that are being hosted. For UC applications, see the "QOS Design Considerations" at the following URL:

http://docwiki.cisco.com/wiki/QoS_Design_Considerations_for_Virtual_UC_with_UCS

You must understand the networking requirements of the virtual machines that are deployed on the host and how to set up the host networking hardware to meet those needs. If Cisco determines that application performance problems are due to networking congestion within the host, then some VMs must be moved off the host.

Summary

Under certain conditions, Cisco allows coresident deployments with a mix of Business Edition and non-Business Edition applications. The requirements that are listed in this document provide the basis for a successful coresident deployment, but due to the unpredictable behavior of non-Business Edition applications and the impracticality of testing every possible combination of applications that could be deployed, Cisco cannot guarantee that following these guidelines alone will be sufficient.

A key principal of virtualization is based on sharing hardware resources among multiple virtual machines. When deploying only Business Edition applications on a host server, Cisco can guarantee levels of performance, because the applications are thoroughly tested and their behavior understood. Deploying coresident third-party applications introduces a degree of unpredictability that can be mitigated by following general principals of virtualization and the specific requirements in this document.

Ultimately, Cisco cannot guarantee that by following these guidelines alone, the virtual machines on a host will never be starved of resources. However, if this does occur, the only recourse will be to remove some of the virtual machines from the host such that the load is reduced. This can be done by moving some of the virtual machines to another host, or by moving all the virtual machines to a more powerful host.