

TechTarget

EXCHANGE 2010 SERVER ROLES EXPLAINED

→ The Flexibility of Roles

- → The Mailbox Server Role
- → The Hub Transport Server Role
- → The Unified Messaging Server Role

xchange Server has gone

through such major architectural and functional changes in its last two revisions that it's barely the same product it was back in 2003. Among the major changes along the way has been the addition of server roles, which significantly changed the way Exchange servers are deployed and administered.

Server roles were introduced in Exchange Server 2007 as a variation of the same feature found in Windows Server 2003 and later versions. There, roles allowed you to configure a server by adding only the functionality you needed. The server's reduced attack surface made administration easier and reduced the overall number of moving parts in the system.

Exchange 2003 did have roles, but only two: front-end and backend. Exchange 2007 and 2010 brought the total number of roles to five: client access, mailbox, unified messaging, hub transport and edge transport. If you're migrating an Exchange 2003 setup—and especially if Exchange 2003 or earlier is what you have the most experience with—the way these roles work may seem a little alien at first. But mastering them is key to making Exchange Server 2010 work for you.

THE FLEXIBILITY OF ROLES

What makes role-based management in Exchange useful is that it makes an Exchange installation that much more flexible. You can have all your roles—with one exception—on one physical server, or you can spread them out across multiple redundant servers.

For instance, to balance the storage and processing load needed for your mailboxes, you can have five mailbox-role servers to every one hub transport server. A simple Exchange 2010 installation would involve the client access, hub transport and mailbox roles—the absolute minimum allotment of roles needed for an Exchange setup—all on one physical machine. You can also add or decommission role instances as needed.

There are a couple of exceptions and caveats. The one inflexible role in Exchange is the edge transport role, which always has to run on its own server, although it's an optional Exchange role. Also, some roles—the unified messaging role in particular—may not perform as well on a server with other roles, depending on the server load. Finally, Exchange Server 2010 should be installed on a member server rather than a domain controller to keep the load of each component from affecting the other.

To better understand each of the roles, how they're deployed, when they can share the same server and when they shouldn't, let's take a look at each role in detail.

THE CLIENT ACCESS SERVER ROLE

The client access server (CAS) role handles incoming connections to Exchange 2010.This sounds like a simple—even mindless—duty, but as Exchange installations grow larger and more complex, the CAS role was developed to allow better segregation of duties and a degree of automatic load balancing and failover between Exchange servers. To that end, the CAS role han-

dles five basic duties:

• Outlook Web App. This service, previously known as Outlook Web Access, allows you to access your Exchange account from a Web browser. Exchange 2010 adds, among other things, policies for Outlook Web App mailboxes that control behaviors such as how attachments are handled.

Exchange ActiveSync clients.

This service synchronizes data between Exchange and mobile

devices with ActiveSync. Note that Windows Phone 7 clients don't support all ActiveSync policies, so if you have any such devices in your organization, you'll need to think ahead.

POP3 and IMAP4 clients. These are the legacy Internet email protocols that all of us should be familiar with. Note that in Exchange, POP3 and IMAP4 are installed but disabled by default as a security measure. You must set up and configure them manually, along with Simple Mail Transfer Protocol after the CAS role has been added.

• Availability service. This publishes users' free/busy data to Outlook 2007 and 2010 clients and even has its own API to allow third-party applications to use it. If you want to use the availability service in conjunction with load balancing—especially if the performance counters for the service show it's being strained—you'll need to make some changes to the system.

• Autodiscover service. This innocuous-sounding service is quite important. It allows Outlook 2007 and 2010 clients, as well as supported mobile devices, to be automatically configured with the needed user profiles: the users' names, connectivity settings for their mailboxes and a slew of other things that used to be configured by hand.

→ The Flexibility of Roles

→ The Mailbox Server Role

→ The Hub Transport Server Role

→ The Unified Messaging Server Role It also automatically keeps clients cognizant of any network connection changes in Exchange, so you don't have to do this manually.

The CAS role is quite demanding, so in high-performance Exchange scenarios there should be more than one CAS. Note that having a CAS array, or multiple CASes, doesn't automatically work as a load-balancing solution for Exchange. A CAS array is a valuable component, but you'll still need to set up a <u>full-blown loadbalancing solution</u>, whether via a hardware appliance, a reverse proxy or Windows Network Load Balancing—in a roughly descending order of preference.

CLIENT ACCESS SERVER ROLE: MORE RESOURCES

Dig deeper into the functions of the Client Access Server role in Exchange Server 2010 with the following resources from Microsoft's TechNet:

- Outlook Web App Mailbox Policies
- Exchange ActiveSync
- POP3 and IMAP4
- Availability Service
- Autodiscover Service

THE MAILBOX SERVER ROLE

The mailbox server role is in many ways the "guts" of an Exchange 2010 installation—it's where the mailboxes, public folders, address lists and books, and message records are all stored and indexed. For that reason, the planning you put into servers that have this role is crucial.

Generally, mailbox databases and storage use are radically different in Exchange 2010 than in Exchange 2003. The new database availability group architecture provides better by-default protection for mailboxes, and Exchange 2010 adds support for 512e-format disks for enhanced performance. For those and other reasons, you need to think about the mailbox server role as the central building block for an Exchange 2010 installation, and be aware that it's free of the assumptions you might bring to it from previous versions of Exchange.

Microsoft has a slew of tools for mailbox server capacity planning. Most notable is the <u>mailbox server</u> <u>role requirements calculator</u>, a constantly updated Excel spreadsheet template that lets you plug in figures that reflect your ideal mailbox server configuration. In addition, it produces recommendations for CPU, memory and storage requirements.

Note that, even with this calculator, you still have to understand

→ The Flexibility of Roles

→ The Mailbox Server Role

→ The Hub Transport Server Role

→ The Unified Messaging Server Role The Flexibility of Roles

→ The Mailbox Server Role

→ The Hub Transport Server Role

→ The Unified Messaging Server Role the concepts involved in planning an Exchange Server deployment. If you're not familiar with Exchange 2010 concepts like high availability and site resilience, learn about those first (see "Enhancing Reliability With Database Availability Groups" below). They're central to how the mailbox server role works, and with them you'll have a far better idea of why certain things are needed.

One aspect of planning that's slightly overlooked with a mailbox server is network topology. Mailbox servers talk directly to servers that use every other Exchange role, so they should be at the heart of any topology planning you're putting together for Exchange. This includes, as you can imagine, your Active Directory servers.

Note that mailbox servers aren't accessed directly by Outlook for the sake of their mailboxes, since that task is handled via the CASrole computers. But Outlook *does* talk directly to mailbox server role machines for the sake of accessing public folder information, so bear this in mind when planning your network architecture.

ENHANCING RELIABILITY WITH DATABASE AVAILABILITY GROUPS

THE MOST CRUCIAL new high-availability feature in Exchange Server 2010 is the database availability group (DAG), which gives Exchange servers the reliability of clusters without the headaches associated with traditional clustering. DAGs gang together groups of Exchange mailbox servers to allow automatic database replication among the group's servers and automatic failover across the servers in the DAG. Microsoft bills DAG technology as the base component of the high-availability and site resilience framework built into Microsoft Exchange Server 2010.

Having DAGs simplifies many procedures when it comes to Exchange reliability. With DAGs, you can allow for failover without first having to build clustered servers and install Exchange on them. In Exchange 2010, you can add servers to a DAG at any time and allow them to replicate between each other. And because replication is more granular in Exchange Server 2010, you don't have to replicate every mailbox.

Additionally, with DAGs, databases are not tied to a single server; they can be replicated up to 16 times across different mailbox servers. Microsoft claims that a DAG with three servers better protects Exchange and recovers from disaster faster than a single, conventionally backed-up server. <u>—SEARCHEXCHANGE.COM</u>.

THE HUB TRANSPORT SERVER ROLE

The hub transport server role governs the flow of Exchange mail within an organization. In other words, it handles mail before it's actually delivered to an inbox within the organization or sent to another server across the Internet. To the uninitiated, this role might sound like pure busy work, but it's actually a way to dedicate a server to handle some tasks that are normally intermixed with other machines.

Regulatory compliance. Sarbanes-Oxley-Act-type regulations and internal organizational message policies—data retention, confidentiality, discovery compliance and compartmentalization—have become increasingly important in companies of just about every size. The hub transport role is, in part, a way of dedicating a server role to handling those issues and segregating how they're handled from the mailboxes themselves.

Mail flow and categorization.

This includes things like expanding address lists, forwarding addresses and applying other mail-handling policies. Placing these tasks in the hub transport means the actual mailbox servers don't have to be burdened with them.

Shadow redundancy for mes-

sages. This feature, new to Exchange 2010, adds new highavailability protections for messages while they're in transit. If one hub transport goes down, another one—assuming one is available can pick up where it left off without losing messages that were being delivered during the failure. This also allows servers with the hub transport role to be taken offline without worrying about losing mail.

• Antispam and antivirus. While these tasks are normally the province of the edge transport roles, hub transport role servers can also run antivirus agents—sold separately—and provide antispam protection.

If any of these tasks or behaviors are critical to your organization, think about having more than one hub transport in your Exchange setup.

THE EDGE TRANSPORT SERVER ROLE

An edge transport server (ETS) role is a must for large organizations that use Exchange as their mail gateway between their organization and the rest of the world. The ETS is roughly akin to a firewall or other network-protection appliance: It manages all incoming data from the Internet at large and works as a first line of defense against spam and viruses. It also

→ The Flexibility of Roles

→ The Mailbox Server Role

→ The Hub Transport Server Role

→ The Unified Messaging Server Role The Flexibility of Roles

- → The Mailbox Server Role
- → The Hub Transport Server Role
- → The Unified Messaging Server Role

handles outbound mail, performs preliminary message categorization and applies any mail-flow rules that have been set up.

If possible, the ETS role should always be on its own server. This is not just for practical reasons, although it's true that the ETS can be a demanding role. It's because the ETS sits outside of any network-protection system you might have—in the <u>DMZ</u>. For safety's sake, no other Exchange role should be placed outside your network-protection system.

Note that the ETS is an optional role, not a mandatory one. It's useful to have an ETS in a large organization, but if all your Exchange roles are housed on one server, you can live without it. You can instead enable the edge agent on whatever hub transport server role you have set up to accomplish the same functions, but you'll incur the cost of imposing that much more of a load on the hub transport server machine.

Another reason to set up an ETS is if you want to securely exchange mail with another Exchange setup. Conversations between ETS-role machines are encrypted by default, so if you have two Exchange Server setups in different physical locations, ETS roles for both are a must.

THE UNIFIED MESSAGING ROLE

Unified messaging in Exchange

allows your inbox and your voicemail to coexist more closely and lets you manage the latter with the same ease as the former using the same tools. Voice messages can be played through PC speakers, headsets or any phone; special rules can be established for voicemail; voicemail can be protected through Active Directory Rights Management; and much more.

A unified messaging server in Exchange speaks to your private branch exchange (PBX) network, typically through an IP gateway, and routes messages to your other Exchange role servers—client access, hub transport and mailbox. For that reason, the placement and connectivity of your unified messaging-role server is important. The PBX gateway and the unified messaging servers should always be as physically close as possible. Likewise, unified messaging servers should be as close as possible to the other Exchange role servers. Microsoft recommends that the total round-trip time between an IP gateway for your PBX and the unified messaging server is not more than 300 milliseconds.

Unified messaging roles tend to be pretty demanding with system resources, so machines with such roles should be dedicated to unified messaging only and should not share other Exchange roles. The general rule of thumb for provisioning unified messaging roles is by call volume: A single physical unified messaging server can handle about 100 concurrent voice calls. Microsoft also makes <u>recommendations</u> for call volume based on the codecs used for storing voice messages.

Finally, unified messaging servers are unique in that they depend on the connectivity of an entirely external system that has nothing to do with Exchange—that is, your PBX. The more capacity your PBX can support, the more unified messaging servers you may need to satisfy that demand.

The most important thing to understand about Exchange 2010 server roles is their diversity of function. While many of the roles can coexist on the same hardware, they're not freely interchangeable; vou wouldn't want to put a hightraffic unified messaging role on hardware formerly used for a lowtraffic mailbox server. Each role has its own specific demands, which need to be evaluated both individually—e.g., edge transport or not? and in light of the kind of Exchange installation you're putting together. Exchange isn't a monolith anymore.

Serdar Yegulalp has written about computers and IT for more than 15 years for a variety of publications, including SearchWinIT.com, SearchExchange.com, InformationWeek and Windows magazine. Check out his blog at <u>GenjiPress.com</u>.



Mastering Exchange Server Roles is a SearchExchange.com e-publication.

Margie Semilof Editorial Director

Lauren Horwitz Executive Editor

Christine Cignoli Senior Features Editor

Phil Sweeney Managing Editor, Special Products

Jeannette Beltran, Eugene Demaitre and Martha Moore Associate Managing Editors

> **Linda Koury** Director of Online Design

Marc Laplante Publisher <u>mlaplante@techtarget.com</u>

TechTarget USA 275 Grove Street, Newton, MA 02466 <u>www.techtarget.com</u>

© 2012 TechTarget Inc. No part of this publication may be transmitted or reproduced in any form or by any means without written permission from the publisher. TechTarget reprints are available through The YGS Group.

ABOUT TECHTARGET: TechTarget publishes media for information technology professionals. More than 100 focused websites enable quick access to a deep store of news, advice and analysis about the technologies, products and processes crucial to your job. Our live and virtual events give you direct access to independent expert commentary and advice. At IT Knowledge Exchange, our social community, you can get advice and share solutions with peers and experts.

→ The Flexibility of Roles

- → The Mailbox Server Role
- → The Hub Transport Server Role
- → The Unified Messaging Server Role