Exchange Remote Management and Monitoring

Remote Management

If your company awards Coyote Creek the responsibility of remotely managing your Exchange 2010 network, here's a brief summary of the native tools we'll use.

- First, to access your network, we use Windows Remote Management (WinRM) 2.0. WinRM allows our engineers to connect securely to your network using an Internet Information Service (IIS) virtual directory from wherever they may be.
- The core tool for managing Exchange—whether locally or remotely—is the *Exchange Management Shell* (EMS). The EMS is a command-line interface, much like the interface for an OS shell, that lets us as administrators use both simple commands and scripts to manage any aspect of your Exchange network. The EMS utilizes Windows PowerShell 2.0, allowing remote administrators to run commands directly on any server in your Exchange network without having to have EMS binaries installed on the client they're using. The EMS defines any configurable entity on the network—a user mailbox, a server, a connector—as an *object*. This allows pipelining, another OS: making the output of one command or script become the input of another. Together with pre-scripted cmdlets (pronounced "command-lets") the object-oriented structure of the EMS makes configuring any aspect of your ES 2010 network far easier to do than earlier versions. The EMS can also run any *.exe* command in the Windows shell.

- Built in top of the EMS is the Exchange Management Console (EMC). The EMC is a GUI that enables most though not all EMS commands to be run from a menu. We use it for day-to-day administrative tasks.

Monitoring

At Coyote Creek, we prepare a carefully thought out and comprehensive monitoring solution for Exchange 2010 right at implementation, not as an afterthought. Properly designed monitoring allows administrators to identify, troubleshoot, and repair issues before end users are impacted. We believe comprehensive monitoring is worth the time it takes by reducing service outages that can translate into financial costs down the line. A Coyote Creek solution will do the following:

- *Identify performance issues*. The faster the solution helps us determine the cause, the less time to resolution. We don't think our clients should rely on their end users to alert them to issues; it causes user dissatisfaction and drives up help desk calls.
- *Identify growth trends*. Over time, usage patterns and business needs change, and hardware may need to be modified to address these changes. Trending helps forecast usage changes and allows administrators to fix an environment before the changes cause problems.
- *Track performance against established service level agreements* (SLAs). When we monitor Exchange, we report on how the network is meeting SLA-defined metrics. This may help with upgrades and other areas that need cost justifications.

- *Track configuration changes*. Your company's Exchange environment may have dozens or even hundreds of servers, and they all need ongoing maintenance. Even with a fraction of those numbers, ensuring that settings remain consistent across servers is a constant challenge. To keep service running smoothly, our engineers also review changes against known best practices.

Monitoring Exchange 2010 requires a solution that can gather information from dependent services, such as Active Directory or DNS. If your company doesn't already have a comprehensive solution, we recommend System Center Operations Manager 2007 R2 (SCOM) from Microsoft. SCOM provides end-to-end service management that includes monitoring, troubleshooting, and reporting tools. SCOM uses *management packs* to extend the base framework for specific applications like Exchange 2010. A management pack includes all of the rules, knowledge, reports, and tasks needed to monitor a product. The Exchange 2010 management pack has a complete health model, diagnostic alerts, and service-level reports that we use to smoothly operate and monitor Exchange.

Exchange Performance Monitor, Server by Server

When monitoring Exchange Server 2010, we pay close attention to the aspects of server performance that are the most important. Performance counters and threshold values can identify potential issues as well as isolate their root causes for troubleshooting. We prefer to create a baseline measurement of server performance during normal operations. After the baseline has been established, we set thresholds to we'll know—and you'll know—when performance metrics are not met.

Performance Data for the Mailbox Server

Many performance counters are available for the Mailbox server. The key storage metrics we look at are those that give us average time for reading data from the active database file or from a passive database file, and the rate of page faults that can’t be serviced because no pages are available for allocation from the cache. Otherwise, our focus is primarily on storage response times. If the disk subsystem is not meeting demand, fixing the problem may require additional disks, faster disks, or modifying the disk configuration.

If Remote Procedure Call (RPC) counters indicate a problem, there are several possible causes. This is what we do in each case:

- *Storage subsystem*: We ensure that I/O read/write latencies are not excessive and correlate the storage-based counters with RPC counters to see whether they align.
- *Network components*: We check network card settings for errors, dropped packets, network speed, and duplex settings.
- *CPU*: We assess whether the CPU is running near capacity; if it’s overtaxed, it won’t be able to process RPC calls.
- *Applications*: We identify applications that generate lots of RPC calls. Then we use Client Throttling Policies to prevent the application from consuming excessive server resources.

Performance Data for the Hub Transport and Edge Transport Servers

The key counters for transport servers are for queues; monitoring transport queues helps our engineers ensure timely message delivery. We look at how many messages from all queues are waiting for delivery: how many are in the active remote delivery queues, how many in the submission queue, how many are in a retry state in the remote delivery queues, and how many are in the poison queue.

Performance Data for the Client Access Server

The key counters for the Client Access Server role center on client services such as Outlook Web App and Exchange Web Services. We monitor metrics like how many times the client-service application has been restarted during the Web server's lifetime, how long in milliseconds the most recent request was waiting in the queue, the number of requests in the application request queue, the average time for a search to complete, and the average time (in milliseconds) ECP took to respond to a request during the sampling period.

Performance Data for the Unified Messaging Server

The key counters for the Unified Messaging role are about UM availability. We check the percentage of mailbox connection attempts that failed in the last hour, the percentage of messages successfully processed by UM in the last hour, and the number of failed attempts to access Active Directory.

Monitoring Exchange in the Enterprise IT Environment

Overall, in addition to these details, we look at Exchange 2010 like any other Windows server and check the "big four" first: CPU, memory, disk, and network. For example, if the network is saturated or the server is memory-starved, any Exchange-specific counters, such as RPC latency or message queues, are very likely only symptoms of the underlying problem.

When troubleshooting Exchange, we often also need to troubleshoot related services. Active Directory and the operating system are both intimately linked to Exchange. And the whole Exchange network , the AD forest, and the OS all depend on hardware fundamentals—server CPUs, server memory, and disks. For this reason, we recommend that your IT staff creates operational-level agreements (OLAs). OLAs are similar to SLAs—except that OLAs are between internal groups working to support an SLA, and the SLA is a performance agreement with a business group. OLAs are created to ensure that core activities performed by different support teams are aligned to meet the relevant business SLAs. At Coyote Creek, we have the knowledge to assist other support groups and know where the interdependencies are.