

Deploying EMC Smarts Application Discovery Manager

Abstract: Customers are curious, rightfully, on how to optimally deploy EMC Smarts Application Discovery Manager in a “typical” enterprise application environment. This whitepaper addresses some key considerations for customers as part of their deployment strategy when deploying Application Discovery Manager. Since every customer environment is unique, the whitepaper does not attempt to make any specific recommendations but provides some alternatives customers should consider.

Copyright © 2007 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS IS.” EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

Sample Trademark page below.

EMC², EMC, ApplicationXtender, Celerra, CentraStar, CLARAlert, CLARiiON, Connectrix, Dantz, Direct Matrix Architecture, DiskXtender, Documentum, EmailXtender, EmailXtract, HighRoad, Legato, Navisphere, PowerPath, RepliStor, ResourcePak, Retrospect, Smarts, SnapView/IP, SRDF, Symmetrix, TimeFinder, VisualSAN, and where information lives are registered trademarks and EMC ControlCenter, EMC Developers Program, EMC OnCourse, EMC Proven, EMC Snap, EMC Storage Administrator, Access Logix, ArchiveXtender, Automated Resource Manager, AutoSwap, AVALONidm, C-Clip, Celerra Replicator, Centera, CLARevent, CopyCross, CopyPoint, DatabaseXtender, Direct Matrix, DiskXtender 2000, EDM, E-Lab, EmailXaminer, Engenuity, eRoom, FarPoint, FLARE, FullTime, InfoMover, MirrorView, NetWin, NetWorker, OnAlert, OpenScale, Powerlink, RepliCare, SafeLine, SAN Advisor, SAN Copy, SAN Manager, SDMS, SnapSure, SnapView, StorageScope, SupportMate, SymmAPI, SymmEnabler, Symmetrix DMX, and VisualSRM are trademarks of EMC Corporation.

All other brand names are trademarks or registered trademarks of their respective owners.

Part Number: S0078

Table of Contents

- Overview4**
- Application Discovery Manager Discovery Technologies4**
 - Passive Discovery 4
- Requirements4**
- Where Should I Connect To?5**
 - Connections to Core Switches 5
 - Connections to Distribution Routers and Switches 5
 - Connections to Access Switches 6
- Planning the Installation6**
- Validating Your Deployment Plan7**
 - Common Deployment Pitfalls 7
 - Active Discovery 7
 - Requirements 7
 - The Challenge and Risk 8
 - Planning Active Discovery 8
- Conclusion.....8**
- About EMC Smarts.....8**

Overview

Customers are curious, rightfully, on how to optimally deploy EMC Smarts Application Discovery Manager in a “typical” enterprise application environment. This whitepaper addresses some key considerations for customers as part of their deployment strategy when deploying Application Discovery Manager. Since every customer environment is unique, the whitepaper does not attempt to make any specific recommendations but provides some alternatives customers should consider.

The typical three-layered hierarchical network model includes:

- **Core layer**—The core layer is a high-speed switching backbone and should be designed to switch packets as fast as possible. This layer of the network should not perform any packet manipulation, such as access lists and filtering, that would slow down the switching of packets
- **Distribution layer**—This layer includes LAN-based routers and Layer 3 switches. This layer ensures that packets are properly routed between subnets and VLANs in your enterprise. This layer is also called the workgroup layer.
- **Access layer**—This layer includes hubs and switches. This layer is also called the desktop layer because it focuses on connecting client nodes, such as workstations to the network. This layer ensures that packets are delivered to end user computers.

See <http://www.cisco.com/univercd/cc/td/doc/cisintwk/idg4/nd2002.htm>

See <http://www.semsim.com/ccna/ccna-study-guide.asp?ain=57>

Application Discovery Manager Discovery Technologies

Application Discovery Manager is the de facto standard in application discovery and dependency mapping. Application Discovery Manager uses a hybrid passive and active approach to perform a complete, end-to-end discovery of the application environment including network resources, servers, services, and business applications. As part of the discovery process, also builds out a complete relationship and dependency model, as well as usage and demand models for the application environment.

Passive Discovery

Passive Discovery (PD) works by performing deep packet analysis of network packets that are collected through monitoring ports of switches. Since PD works without spidering, crawling or probing the network in any way, no load is placed on the network when performing PD. In fact, PD is completely passive in nature, as the name implies.

Requirements

Application Discovery Manager must see traffic of ALL desired application components to provide a complete picture of the application environment.

Modern networks are designed in a hierarchical topology. No single switch device processes all traffic passing through the network. As such, Application Discovery Manager PD requires connections to multiple switches for complete application mapping. The connections, however, must be made to the right switches in order to ensure optimal coverage and efficient deployment.

Conversely, missing a key switch may result in incomplete dependency and services mapping. Similarly, misconfiguring monitoring ports may also have adverse effects on the switch and potentially the network as well.

Where Should I Connect To?

The Application Discovery Manager appliance includes three input connections that are used to connect to monitoring ports. The three connections can be easily extended to over 100 monitoring port connections by using a device called an Aggregation Switch (A-SWITCH). Aggregation switches are standard monitoring port enabled switches that are configured to perform multiple monitoring port aggregation and provide a single output stream. Each Application Discovery Manager monitoring port can be connected either directly to a switch's monitoring port or to an aggregation switch's output.

Since the aggregation process takes multiple high bandwidth feeds and has a single output, the majority of the packets are lost in the aggregation process (the original packets of the network are not lost of course). However, with Application Discovery Manager statistical sampling approach, the Passive Discovery engine is capable of performing an accurate discovery even with less than 1 percent of the traffic.

There is no silver bullet answer on where to connect the Application Discovery Manager appliance. This is primarily because the desired servers may be connected to Access, Distribution and core switches. But here are some suggestions and considerations.

Connections to Core Switches

Core switches provide visibility to traffic going between different Distribution level devices – in a typical environment, this includes the traffic between different data centers (or data centers regions), clients and geographically separated networks.

The servers of a modern distributed application are usually installed within the same data center boundaries. Packets going between the different components of the distributed application are therefore not likely to require processing by core switches. A connection to a core level switch is therefore not a good source to map these kinds of dependencies.

It is also very common for an application to interact with another application (or applications). If the other application resides within the same data center boundary, this interaction is typically not processed by the core switches. However, if the other application resides in a different data center or in a different geographical location, then the interaction between these two applications is typically passing through the core switches. Core switches are therefore a good source for mapping application-to-application dependencies for applications residing in different data centers.

Finally, clients usually reside outside of the data center and are likely to use the core level switches to access the applications that reside in the data center. Core level switches are therefore a very good source for mapping client to application dependencies.

One other consideration - core switches usually come in pairs to provide high availability and load balancing. In normal operation, the load is balanced between the switches and it is therefore recommended to connect to all core switches.

To summarize, in a typical environment, core level switches are a good source for client to application dependencies as well as application-to-application dependencies for applications residing in different data centers. Core level switches are typically not a good source for mapping dependencies amongst the different servers of a single application as well as mapping inter application dependencies for applications residing in the same data center.

Connections to Distribution Routers and Switches

Distribution level switches provide visibility to traffic going between access level switches and core switches, as well as traffic going between different distribution level switches. Distribution switches are usually the entry points for data center regions and subnets.

Since all clients accessing the data center must pass through distribution level switches, these switches are a good source for mapping client to application dependencies.

In addition, applications connecting to other applications in different data centers will also go through distribution level switches making them a good source for mapping this kind of dependencies.

However, as with core switches, different components of a distributed application are typically installed within the same subnet and do not require intervention of distribution level switches. Therefore, distribution level switches are also not a good source for mapping dependencies between the different application components.

As with core switches, distribution level devices typically come in pairs as well. If you decide to connect to distribution level devices, we highly recommend connecting to both devices of each pair.

Connections to Access Switches

Access level switches provide complete visibility into the servers or workstations that connect to them. Since every packet that going in or out of these servers must pass through the access switch they are directly connected to, they are the ideal source for mapping these server dependencies.

However, an organization may have many access level switches, making connecting to each one of them an impractical implementation. Luckily, the majority of the access switches are used to connect workstations (clients) to the network and these do not require monitoring.

We recommend connecting to all access level switches in each data center (note – the access level switches that are part of the data center, but not those that provide access to standard clients). Connecting to all Access level switches in the data center would provide complete visibility to all dependencies among the different components of the application. In addition, they provide visibility into clients using the application as well as dependencies in other applications in different data centers.

It is possible, though not recommended, to skip connections to a few access level switches. This is only recommended if you are absolutely sure that the switch to which you do not connect does not serve any component of the applications you would like to map.

Planning the Installation

While EMC provides a detailed installation checklist prior to installation, the checklist alone is NOT enough to ensure a successful deployment. A good understanding of the application environment and the network environment is required:

- Identify the list of key applications being discovered
- Application owners should also provide IPs of main servers that are part of the application
- Identify with network administrator which switches are used by the IPs provided earlier
- Identify switches that are used by the customer including available monitoring ports
- Identify the switches that carry client->application traffic if that's also something customers would like to see

In cases where there are more than three switches to connect to, it may be necessary to acquire an aggregation switch that will act as an aggregation point for the multiple monitoring ports. This might be as simple as a basic hub, a typical switch, or a more specialized aggregation switch to aggregate multiple monitoring ports.

To provide complete visibility into a specific data center, it is recommended to connect to all switches supporting the data center. You may exclude switches providing access to workstations only, as client-to-client interdependency is usually of no interest.

For example, a deployment in a datacenter that uses 20 access and distribution level switches and is connected to two core switches should be implemented by connecting to all 22 switches. Additional switches that connect clients (workstations) to the network are not required in this setup. Since this includes more than three switches, the deployment must also include an Aggregation Switch.

Validating Your Deployment Plan

Once deployed, you can start the discovery process. It is important to validate that all servers connected to the switches you are monitored are being discovered by the Passive Discovery component. If they do not show up, the two most common reasons are:

- The server is connected to a switch that is not monitored – ensure that the servers are connected to a switch that Application Discovery Manager is monitoring. Ensure that the switch is configured to perform port monitoring on the relevant ports.
- The server is not active – try creating load on the servers from remote clients

Common Deployment Pitfalls

“You are monitoring all VLAN traffic”

VLAN traffic may span multiple switches. You may see all VLAN 5 traffic from Switch A but not VLAN 5 traffic from Switch B

“You are monitoring all traffic from our core switches”

The majority of traffic among servers does not reach the core switches at all! As discussed above, it is more likely that traffic from Access switches is also needed to provide a complete picture.

“You should now see traffic of server X” (but it doesn’t show in Application Discovery Manager)

Validate with the network admin to ensure the correct switches have been configured for monitoring

Use tcpdump to test whether server X is monitored

“I’ve connected an aggregation switch, but now no traffic is passing thru my network”

When two switches are connected to each other in a network environment, they are identified by a number that determines which is responsible for maintaining the hierarchy of the network. If a switch configured to act as aggregation switch is connected to an existing DC switch, without SPAN ports configured on the existing DC switch, the aggregation switch can in essence take over the network and claim to be the “lead role.” In this case, some of the switches may assume a sub-ordinate role and stop responding to requests, potentially knocking services off line. It is therefore vital that SPAN ports are configured prior to connecting an aggregation switch to DC switches.

An aggregation switch can sit un-configured and can be connected to the network without causing any issues so long as the connections from the switches that make up the infrastructure of the DC are fully configured as SPANs.

Active Discovery

Active Discovery (AD) is an agentless, remote discovery technique that uses credentials to log on to remote servers to retrieve detailed configuration information. Application Discovery Manager supports SNMP, SSH, Telnet and WMI protocols in AD. In addition, Application Discovery Manager also provides IP Scanning as a supplementary discovery protocol to discover servers and open ports for a specified IP range.

Requirements

AD requires credentials or SNMP community strings to remotely connect to servers to retrieve detailed configuration information for hardware configurations of the servers and the applications running on the servers. AD will not work without the requisite community strings and credentials. In addition, if the credentials provided do not

have the required permissions to run commands needed to extract the configuration information, AD will not work either.

The Challenge and Risk

The biggest challenge with AD is mapping key servers and getting credentials for them. AD will *not* work without the correct credentials or if not all credentials are provided for the servers. In addition, firewalls/NAT and other connectivity issues may also prevent complete discovery.

Planning Active Discovery

Prior to enabling AD, it is important to consider the following:

- Identify and map the key servers against which you would like to run AD. You may use the topology discovered by the Passive Discovery to identify key servers of your applications.
- Decide on the required credentials and community strings:
 - SSH, WMI and Telnet are preferred over SNMP since they provide controlled access to the servers, and also provide more detailed configuration information
 - SSH and Telnet are required for detailed software configuration information
- Ensure that potential connectivity blockers (such as firewall and NAT) are located and considered prior to enabling AD. It may be necessary to configure the firewall(s) to allow this traffic.
- Test the credentials prior to enabling AD
- For Actively Discovered hosts, make sure all information is available

Conclusion

Although every customer application network environment is unique, and requires a certain amount of thought and planning, it is by no means difficult. The success of an Application Discovery Manager deployment is highly dependent on clear understanding of the network and application environment, careful planning and identification of the deployment locations, identification of the key applications, servers and credentials, and the methodical implementation according to plan.

About EMC Smarts

EMC Smarts plays a crucial role in managing your information infrastructure by automating the discovery, understanding, and mapping of the complex relationships that exist among business processes, applications, and the IT infrastructure.

With EMC Smarts solutions, organizations gain the visibility needed to accelerate and increase ROI on their highest-priority IT service and cost management initiatives. Offering the easiest, most-comprehensive solution in the industry, EMC Smarts technology allows organizations to:

- Accelerate ITIL and CMDB standardization
- Reduce costs—up to 25 percent in the first year
- Maximize resource utilization
- Mitigate risks and ensure business continuity
- Enhance business agility and IT service delivery by accelerating and simplifying initiatives that support business service management and data center automation

Getting Started

To learn more about how the EMC Smarts Solution for Data Center Audits—and other EMC Smarts solutions—can positively impact your business and IT operations, contact your local EMC or EMC Smarts sales representative, or visit our websites at www.EMC.com and www.smarts.com.