December 2002 **File: SMS 16**

Secure DMZ Infrastructure Management

Service Management Strategies

Glenn O'Donnell

FOCAL POINT

Demilitarized zone (DMZ) infrastructures are common for secure hosting and other connections between the enterprise and the Internet. The security mechanisms necessary for these infrastructures pose special challenges to management efforts. Although firewall policies usually disallow passage of standard management traffic, alternatives can achieve both manageability and security. Consolidation with general infrastructure and application management (IAM) systems and processes is essential for operational efficiency and adequate end-to-end service-level management. Several methods are available to offer such consolidation while preserving the integrity of security measures.

CONTEXT

Infrastructure management normally deals with infrastructure contained within the enterprise boundaries. Communications between the management station and the managed infrastructure are simple in this scenario, using Simple Network Management Protocol (SNMP) as the protocol of choice for collecting management data, especially from network infrastructure. As business systems migrate beyond this safety zone, IAM systems must adapt beyond such heavy reliance on SNMP.

SNMP Version 1 (SNMP V. 1) suffers from well-known security vulnerabilities like clear-text data encapsulation, subject to unauthorized viewing, and weak authentication using easily intercepted “community strings” (i.e., passwords) to govern access. SNMP Version 3 (SNMP V. 3) solves these issues, but technologies that augment or replace SNMP are proving desirable options to pure SNMP. These options will grow across all infrastructure management, but their application is especially suited to DMZ management. Also, despite SNMP V. 3’s superior design, nearly 100% of production IAM systems exclusively use SNMP V. 1.

Currently, only 20% of Global 2000 (G2000) enterprises have implemented secure means to manage DMZ infrastructure from within the safe environment of the corporate network. Twenty percent are managing with security vulnerabilities exposed, and another 10% are using completely isolated management systems. The remaining 50% have little or no insight into objective performance or availability of DMZ infrastructure, with administration requiring close physical proximity or performed though vulnerable or inefficient channels.

Technology developments and operational maturity will compel stronger systems for properly managing DMZ infrastructure. By 2005/06, 50%-60% of G2000 enterprises will successfully integrate secure DMZ management into overall IAM efforts. Technology developments include new IAM tool packaging that enables inexpensive, broad distribution of data collection and processing.

Protocols are evolving, moving to XML and other object-oriented transmission methods to traverse security boundaries. SNMP will coexist indefinitely, but XML-based cross-boundary management communications will be ubiquitous by 2008.

SNMP will continue as a basic data collection protocol for network devices. Centralized management consolidation products (sometimes called managers of managers [MoMs]) are migrating to XML interfaces to support communication to distributed management components and third-party products.

Internet Control Message Protocol (ICMP), used to poll infrastructure components (i.e., “ping”) for availability, will also remain popular, though it has limitations that are being

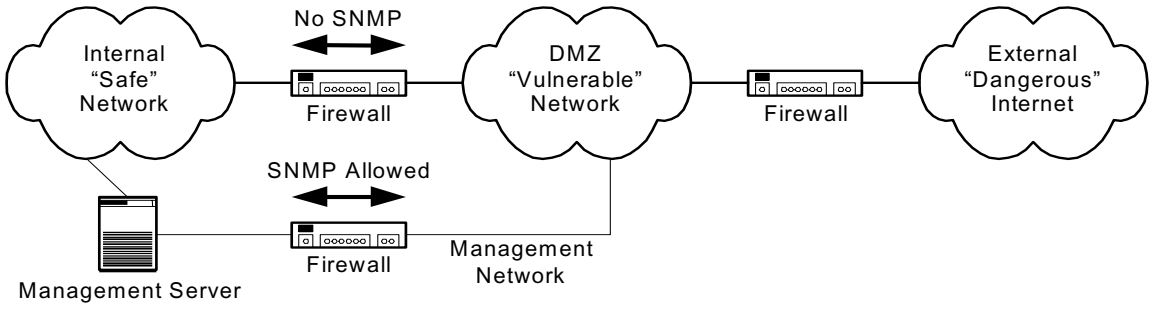
META Trend: During 2002-04, new infrastructure and application architectures (e.g., Web services, utility computing) will drive meta-management approaches (e.g., status aggregation, coordinated policy control). Through 2005, continued cross-boundary management demands (e.g., organizational, informational, technical) will drive process, sourcing, and instrumentation changes. By 2006, business perspective management will be pervasive.

overcome by more complex agents. Current agents (e.g., Tivoli, HP, BMC, NetIQ) enable “super pings” that indicate true resource availability. Simple ICMP echo requests can yield false-positive availability for servers.

The Challenge of Managing the DMZ

The DMZ emerged with the growth of Internet access, as an architectural security buffer between the internal, safe enterprise network and the external, hazardous Internet (see Figure 1). Externally facing resources (e.g., Web servers, e-commerce servers) reside here in an environment that has an intermediate risk level for security breaches. The networks and servers within the DMZ are among the infrastructure most critical to the business, yet this infrastructure is often neglected in management efforts. The conflicting interests of management and security are seen as too tenuous. Actually, the two interests can be served, but new IAM approaches are required.

Figure 1 — Basic DMZ Architecture



Source: META Group

The three network domains are segmented by firewalls to prevent undesired traffic flows (e.g., SNMP, ICMP, RPC) while offering controlled passage of necessary traffic (e.g., HTTP, FTP, Telnet). User Datagram Protocol (UDP) is often disabled on firewalls, because common cross-boundary application traffic is based on the more structured Transmission Control Protocol (TCP). SNMP normally uses UDP. This fact, along with the aforementioned SNMP V. 1 security holes, causes security staff members to block SNMP from passing between the internal network and the DMZ network. This limits the ability to remotely manage the DMZ infrastructure. The ICMP is also often blocked for inbound connections to prevent certain attacks (e.g., denial of service).

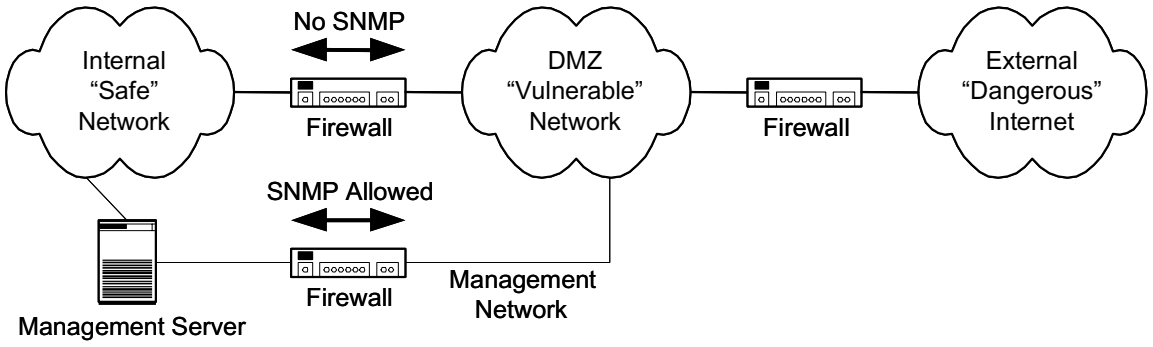
One solution to this conundrum is to open the firewall to pass SNMP, but this is unacceptable in most security policies. Alternatives must be implemented that conserve investments in SNMP for localized data collection, but enable better distribution of processing intelligence and safer cross-boundary communications.

Further complicating the situation, many DMZ architectures employ more than one security zone to provide additional security isolation. Regardless of the number of zones, each firewall represents a demarcation point where one side has a higher trust level than the other. Each zone of insecurity is treated as a mistrusted zone to the next zone up in security. All zones can employ similar management techniques, as long as these cascading security relationships are enforced. For the sake of simplicity, only single-zone architectures will be used for our examples.

Parallel Management Network

One common, effective solution uses a parallel network (see Figure 2) to pass only management traffic. This is also an expensive and complex proposition, because duplicate network equipment is necessary and meticulous control of configurations is critical to ensure true security.

Figure 2 — Parallel Management Network



Source: META Group

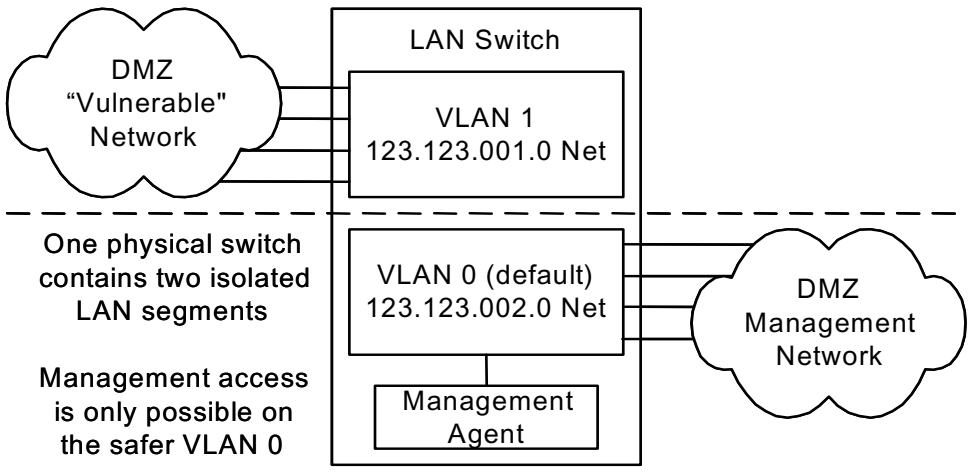
The management network must connect to all managed components in the DMZ through separate network interfaces from the production network. The networks must be functionally isolated, so no traffic can cross over from one network to the other. For optimum security, an additional firewall, with a different configuration from the production firewall, should be placed between the management network and the management server attached to the internal network. A similar supporting measure is to use a separate network interface card (NIC) on the management server itself. Servers and other managed computers inside the DMZ require an additional NIC. For routers and switches, these connections can be segmented from the production DMZ network, but more advanced devices are required. Most low-end network components (e.g., single-LAN switches) are unable to support this segmentation.

VLAN Segmentation for LAN Components

Although the parallel management network is separate from the production DMZ network, it still needs contact with the actual DMZ network components to enable assessment performance and availability of those components. Some type of separation is required to give this visibility while preserving a secure blockade against breaches.

Port-based virtual local-area networks (VLANs) offer a convenient mechanism to segment the management network from the production network (see Figure 3) while retaining secure manageability of the network devices themselves.

Figure 3 — VLAN-Based Traffic Segmentation



High-end switches, using the same physical device, support splitting the switch fabric into two or more logically segmented LANs. Many will support only management traffic on the default VLAN, so the production network will need to be moved to one of the secondary VLANs. Once this is done, the switch cannot provide a path for traffic to cross from one VLAN to the other, thus enabling a secured network at that point. Any traffic to and from the switch itself, including SNMP, will appear only on the management VLAN. Indeed, the switch's IP address will appear only on this VLAN. The switch is "invisible" to the production network, yet it continues to switch packets in a normal fashion.

Port-based VLANs apply static configurations to the switch. Some interpret this to be inherently more secure than newer, dynamic VLAN technologies (e.g., MPLS, 802.11q), where protocol spoofing can theoretically alter the VLAN configuration. With port-based VLANs, configuration changes require access to the IP or SNMP agent.

Strict Traffic Route Management

On servers, network routes must be carefully defined to direct traffic across the appropriate NIC. This can get tricky, but a good system administrator should easily be able to manage this configuration. Failure to properly configure routing could compromise security and possibly even disrupt the entire DMZ's performance.

Figure 4 shows an example of a hypothetical route table for a server with two NICs, one with the address 123.123.1.100 and the other with 123.123.2.100. Figure 5 graphically illustrates the parallel network structure for this hypothetical example.

Figure 4 — Example Parallel Network Structure

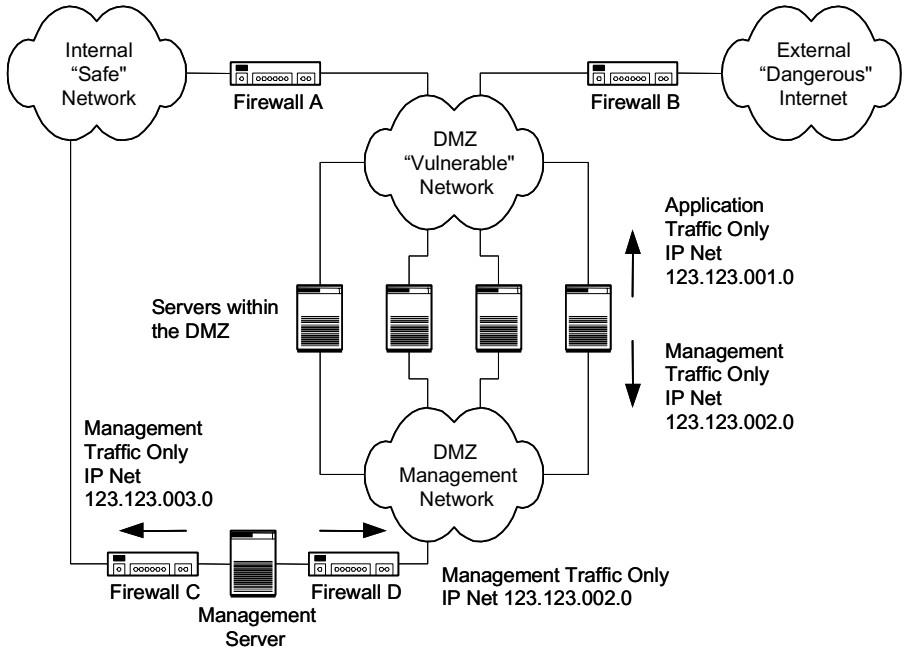
Active Routes:

Network Destination	Netmask	Gateway	Interface	Metric
0.0.0.0	0.0.0.0	123.123.1.1	123.123.1.100	1
127.0.0.0	255.0.0.0	127.0.0.1	127.0.0.1	1
123.123.1.0	255.255.255.0	123.123.1.100	123.123.1.100	1
123.123.1.100	255.255.255.255	127.0.0.1	127.0.0.1	1
123.123.2.0	255.255.255.0	123.123.2.100	123.123.2.100	1
123.123.2.100	255.255.255.255	127.0.0.1	127.0.0.1	1
123.123.1.255	255.255.255.255	123.123.1.100	123.123.1.100	1
224.0.0.0	224.0.0.0	123.123.1.100	123.123.1.100	1
255.255.255.255	255.255.255.255	123.123.1.100	123.123.1.100	1

Default Gateway: 123.123.1.1

Source: META Group

Figure 5 — Management Network Routing Example



Source: META Group

With Class C subnetting, 123.123.1.0 and 123.123.2.0 are separate networks. The default network is 123.123.1.0, so all traffic from this server, with one notable exception, will transmit to this network. That notable exception concerns the second network. Traffic destined for the 123.123.2.0 network, and only that traffic, will take this alternate route. Therefore, the 123.123.1.0 network is appropriate for the production DMZ network and the 123.123.2.0 network is the management network.

WARNING: Servers with multiple NICs sometimes attempt to behave as routers, receiving traffic on one NIC and relaying it on the other. Such behavior will wreak havoc on the network by interfering with the routing behavior of legitimate routers. Users must ensure that routing services on the server are explicitly disabled (e.g., the routed process must be disabled on Unix or Linux servers).
The management agents on these servers will be configured to communicate their data back to the centralized management server. If the management server's address is on the 123.123.2.0 network, the traffic will automatically traverse the management network.

Localizing Management Intelligence Within the DMZ
The ideal solution gives the management system a presence directly in the DMZ. Simplicity, performance, and costs improve in such a scenario. Localized intelligence collects and processes management data before the processed information and event notifications are forwarded to the centralized management station inside the safe network. This forwarding is carried over a secure tunnel (e.g., VPN, secure socket) that is allowed passage by the firewall. No separate management network is required, and complex routing issues do not exist.

The major drawback to this solution is the immaturity of the enabling technology. Therefore, full realization of such a scenario is currently unattainable. Some management products (e.g., SMARTS' InCharge, Peregrine's Xanadu, HP's OpenView Operations) can partially fulfill this solution. Further functional consolidation (e.g., network and systems), industrywide standards compliance (e.g., the DMTF's CIM and WBEM initiatives), and product development are needed, however, to complete this evolutionary stage. Proprietary communications between management servers must yield to standards-based communications to enable simpler firewall pass-through and integration with other management products.

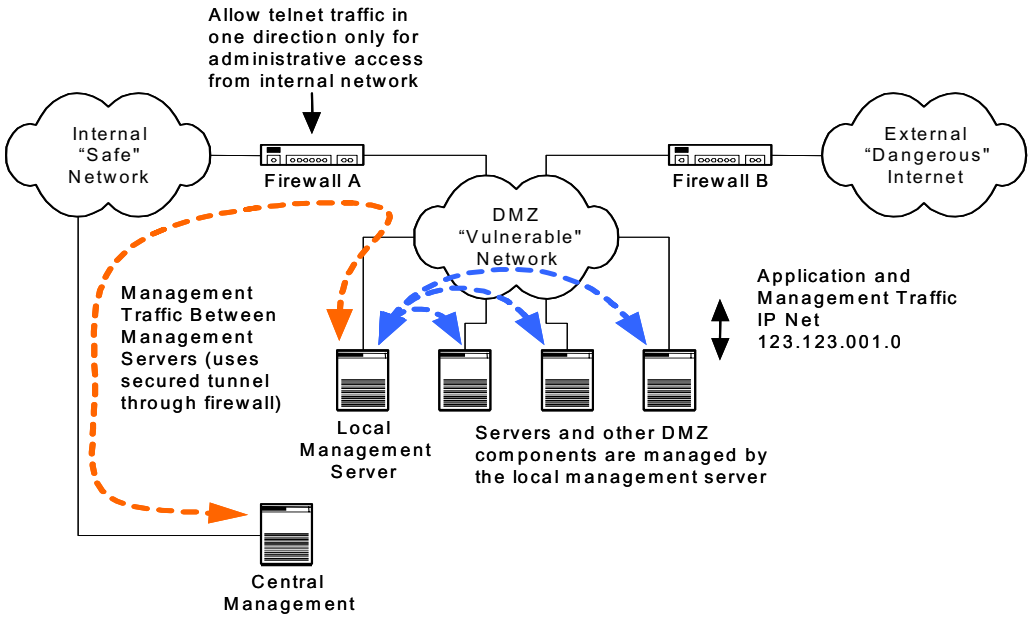
Secure protocols alone are insufficient. The tools themselves need tight authentication and auditing mechanisms as well as strong physical security. The local DMZ management system offers a potential alternative path into the safe network if compromised. Security measures within the tools will help prevent this scenario.

Although management traffic is normally a small fraction of total traffic, higher security levels, broader coverage, and more precise measurements will increase management traffic. All three factors are likely to be present in DMZ management scenarios. With a parallel management network, this is no issue, but if in-band management is desired, such traffic loads should be considered. Localized presence will reduce the impact of this management traffic because the majority of the traffic remains within the DMZ. Only summary information and correlated notifications are forwarded to the central manager.

Local DMZ Management Stations
The most easily achieved instantiation of localized intelligence uses a management server within the DMZ (see Figure 6). This server should act as a proxy for the centralized management server, not as a standalone manager. Management software architecture must also support this hierarchical distribution model.

The local server is a common Unix, Windows, or Linux server running the management application. A low-end server should suffice, because the managed domain is small. Still, the administrative cost and software license of the additional server can be high. With consolidated management across technology domains (e.g., server, network, database) remaining elusive, it is conceivable that multiple management servers may be required. The cost of such a solution must be considered. Continued evolution will drive these costs lower, especially as consolidation gains momentum.

Figure 6 — Local DMZ Management Server



Source: META Group

Proxy Agents

Localized management intelligence does not need to be a full management system. Lightweight agents (e.g., Tivoli's Firewall Proxy Toolkit) can act as local data collection points to funnel management data back to the central management server via a secure connection through the firewall. With lower costs and resource requirements than a full management station, this model is attractive for scaleable management distribution.

Management Appliances

Appliance-based management systems present an attractive option to localized management intelligence (see SMS Delta 1054). Only a few management solutions are currently packaged as appliances, but by 2008, appliances will perform 70% of management data collection and 60% of management processing. Time to value, simple operation, and lower initial cost will make the appliance model the chosen packaging for localized DMZ management intelligence by 2006.

Appliances need to cooperate with a centralized manager, possibly also packaged as an appliance. After simple initial configuration to initiate this dialog, all configuration of the appliance should be done through its Web interface or through the central manager. Security tends to be better than a general-purpose computer running management software because they present no access to the operating system.

Using KVM Switches

Keyboard-video-mouse (KVM) switches are popular for providing console access to servers and network components from a central point. Initially hard-wired to a central console in the data center, KVM switches have evolved to support console access from remote locations. The hardwired connections to the managed devices remain, but the interface is client-side software instead of an actual keyboard, monitor, and mouse. The preferred communications method between the KVM switch and client is TCP/IP, with the client software emulating the console of the managed system. In-band communication is the default channel, though out-of-band PPP (via dial-up modem or another asynchronous channel) is usually required as a backup method.

In the case of DMZ management, in-band communication must be treated as any other management traffic. If the protocol is secure to the satisfaction of security staff, the firewall can allow a direct connection to the KVM switch. Otherwise, it must be treated as vulnerable, similar to SNMP.

Also, the KVM switch needs security itself. Strong authentication is required to ensure secure access to the switch and each managed device. Better KVM switches support role-based privileges for access to individual managed devices and secured protocols, so they represent a suitable option for certain configuration management tasks. However, they do not fulfill fault and performance management needs. These still depend on traditional IAM tools.

Firewall Policy Management

In any scenario, tight control over firewall policies is imperative. A high degree of security can be achieved by opening selective TCP ports (maybe even some UDP ports) on the firewall for specific addresses. The two-pronged restriction of address and port, when supported by management application security mechanisms, significantly minimizes risk.

A good example is NetIQ's AppManager product, which uses encrypted data over selected TCP ports between the management server and the agents on the managed servers. The firewall is configured to open those two ports for only the managed DMZ servers. The NetIQ server additionally checks the validity of the encrypted traffic. In the unlikely event of someone breaching the firewall restriction, the NetIQ application will generate an alert notification if the traffic is improperly encrypted. These alerts can be investigated for possible attacks.

Management Traffic Management

Cross-firewall management traffic should be initiated from within the safe confines of the internal network whenever possible. Traffic must obviously flow in both directions, but only responses, not requests, should flow inward. Traffic initiated from within the DMZ is potentially dangerous. Transactions from the internal network are more easily controlled, and they explicitly identify target nodes within the DMZ. Hijacking these transactions is extremely difficult, even if the DMZ has been severely compromised.

This unidirectional management flow is not always possible. If inbound transactions are unavoidable, protocols must be secure and well understood by both the management architects and security staff. Inbound SNMP is not an acceptable mechanism.

Most management systems rely heavily on periodic polling to determine availability. Poll requests from the internal safe network to DMZ components are acceptable. The timely identification of failures is determined by the polling interval. Common periods are 10 or 15 minutes. Failures can thus be missed for up to 10 or 15 minutes.

Asynchronous notifications (e.g., SNMP trap, XML event message) can solve this problem if the originating device has enough stability to generate the message (total catastrophic failures render the device incapable of generating the notification). The problem with this asynchronous method is that the notification is traversing the firewall in the undesired direction. A localized DMZ management presence mitigates this problem because it has a secure channel established with the central manager. Without this local presence, a suitable compromise is to reduce the polling interval for DMZ components. SNMP traps should not be attempted in the wrong direction. More secure protocols can be used only if strong authentication is used to the satisfaction of security staff.

Configuration Management

Configuration management requires strong authentication to ensure proper identities and privileges. Command-line interfaces (e.g., Unix log-in, Cisco IOS) use telnet, which passes freely through most firewalls in the outbound direction. Management protocols (e.g., SNMP, XML) can pass in this direction, but authentication is still paramount. Because SNMP V. 1 authentication is so weak, it is often avoided for configuration. This is a major reason for the persistence of command-line interfaces for configuration. More advanced protocols (e.g., XML, secure RPC) help. The better server management products lead network management products in the use of these more secure mechanisms.

Many network configuration management tools are basically just attractive graphical user interface (GUI) front ends to command-line interfaces like Cisco's IOS. Management functions are constructed as a script of command-line instructions to be sent along a telnet session to the managed device.

Unifying DMZ Management With Overall IAM

It is important to unify management efforts whenever possible. Operational efficiency and the need for end-to-end visibility are driving the requirement to incorporate DMZ management into overall enterprise management. The DMZ can neither be viewed nor managed as an island of the enterprise. It is integral to the entire IT infrastructure and e-business efforts.

Bottom Line

Integrating DMZ infrastructure management into other infrastructure and application management remains difficult and confusing, but the effort is essential to optimize operational performance while protecting security measures.

Business Impact: Business continuity and performance are strengthened when IT operations have better unified infrastructure visibility.