

ControlTower Appliance

**SunSM Managed Operations
Services Brief
September 2006**

Overview

As a key component to Sun's Managed Operations Services solution, the ControlTower Appliance (CTA) enables Sun to deliver remotely monitored and managed IT operations services 24 hours a day, seven days a week, to any location worldwide.

In order to effectively deliver its flexible suite of remote management services, Sun has developed the CTA as an important segment to its management framework. Using patented technology, the CTA represents Sun's third generation management platform and is the ideal system to deploy in both Internet and corporate data centers.

All successfully monitored devices and applications share three key monitoring metrics:

- **Reactive:** The device/application initiates the sending of an event to signify a change in state
- **Proactive:** An external entity connects to the device/application and queries its health
- **Predictive:** An external entity collects trended information over time to look for resource constraints or chronic issues occurring on the device/application.

The CTA platform provides the tools to effectively monitor these metrics.

CTA Architecture

Sun's CTA runs on a 1U server footprint Sun Server using the Solaris™ 10 Operating System. Sun developed the CTA to be extensible via plug-in software modules. Because the CTA uses a Java™-based architecture it can easily be extended to support new monitoring and management protocols as new industry standards develop, with minimal impact to the existing architecture. The CTA software has been developed through years of management experience, and is designed to handle a number of unique characteristics of remote telemetry monitoring. Some of the features in the platform that facilitate this include: stateful persistence of events, event filtering, and automatic load based reconfiguration.

Predictive Collectors

Predictive Collectors can query device and application agents to retrieve Key Performance Indicators, also known as KPIs, such as CPU usage, memory usage, network performance, and storage utilization. Over time these KPIs show the relative health of the system and can also help predict future scaling and growth requirements. Predictive monitoring provides us with early detection of system issues that may be caused by runaway processes, memory leaks, or other situations that can introduce a large strain on system resources (e.g. Denial of Service (DoS) attacks). Predictive data is available in graphical form within the Sun ControlPointSM management portal, where customers can view and analyze the performance of their managed operations assets.

Receptors

Receptors provide a reactive interface for devices and applications to send event data to the CTA. Sun developed these receptors to maximize performance and throughput, which are capable of handling hundreds of events per second. This design is necessitated by the fact that many reactive telemetry events are delivered over unreliable protocols such as UDP (ex: SNMP and SYSLOG) and therefore, there is only one opportunity to catch the event. The CTA software can intelligently re-allocate its resources as needed to accommodate dramatic increases in telemetry flow, in order to make sure that all events are received from the customer's environment.

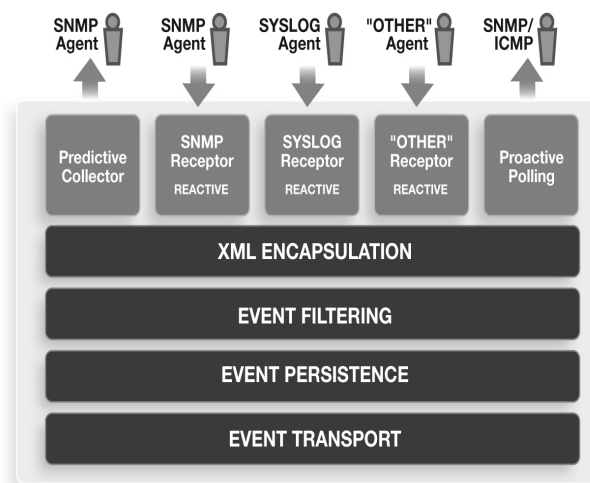


Figure 1: CTA Architecture, Designed for Flexibility

Application Polling

Application polling focuses on indicators of impaired performance from applications. The CTA has the ability to measure transaction performance via HTTP, HTTPS, JDBC, SNMP, ICMP and many other protocols. Should the performance of these application-level transactions fall below pre-defined thresholds, the CTA begins to exercise retry logic associated with each protocol. Should the retry also fail, an alarm is generated by the CTA.

In addition to monitoring for transaction performance times that are falling below an acceptable level, application polling can also monitor responses for the absence or presence of certain type of information – and generate an alert if the expected condition is not met. For instance, making sure that certain fields are always displayed on a web page or generating an error if a certain type of error message is found.

XML Encapsulation

The CTA normalizes events after they are polled or received by passing them through an XML encapsulation layer. Each event type is described using a unique Document Type Definition (DTD) for data source (e.g., SNMP, syslog, etc.). Through this mechanism of converting all message types to XML, future protocols and applications are easily and rapidly integrated into the CTA platform.

Event Filtering

Event filtering is another key characteristic of the CTA platform. In many cases, particularly with network and security devices, granular control of reactive events is significant, but not possible. For instance, when monitoring VPN devices many of the events received only signify the positive status of a VPN.

Event filtering enables the CTA to look for specific patterns in incoming telemetry messages to prevent these events from entering the systems. By carefully filtering telemetry events, the CTA can reduce the amount of telemetry event processing required so that the key reactive information of failed conditions can always have priority. Multiple event filters can be applied to incoming telemetry streams, and are processed in a priority order with options to permit messages through, deny them completely, or log them to the local filesystem for later review.

Event Transport

Transport of events occurs between the CTA and our ControlTower Servers (CTS) located at Sun datacenters. All telemetry data from the CTA is pushed to the CTS, preventing the need for any inbound connections into the customer's environment in order to provide monitoring and reporting services.

Sun believes strongly in using open and auditable protocols between systems. To this end, all communication between CTA and CTS takes place over the HTTPS protocol. Using open standards and secured CTS servers allows customers to easily implement security controls to only allow outbound HTTPS traffic between the CTA and CTS.

Event Persistence

The CTA typically delivers event information back to Sun via an IP network connection. Due to the nature of IP networking, it is not always possible to guarantee the availability of an outbound connection. The persistence layer works closely with the transport layer to queue events for delivery. Events remain in the queue indefinitely until the CTA has both delivered the event to the CTS, and received a checksum based acknowledgement from the CTS that the message was received properly. This is known as the CTA/CTS handshake protocol and is the means by which we can guarantee to not lose event data from a customer's environment due to impediments at the network layer. The Sun ControlCenter systems will trigger an alarm should loss of connectivity occur to a CTA in the field, alerting our engineers to begin investigating - in conjunction with our customer - the nature of the communication issue between our customer's network and Sun.

Management Access

In addition to monitoring devices and applications, the CTA also provides a secure management channel called JumpGate for customers who engage our management services. JumpGate builds on the fundamental security principles defined by the C.I.A. model—Confidentiality, Integrity, Availability—to ensure that all action performed on the customer’s behalf is completed in a secure manner and is fully accountable.

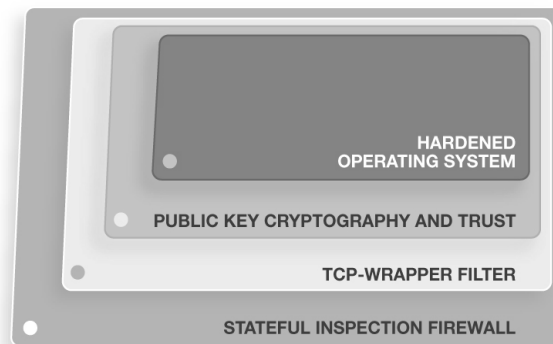


Figure 2: Maximum Security Design

Bastion Hosts

Access to JumpGate is restricted to only Sun hardened bastion hosts located within Sun datacenters by use of three independent security controls (firewall, tcp_wrappers, and public key trust). Bastion hosts require engineers to provide a strong authentication token before access is permitted. This authentication and all access are logged to a central logging server providing detailed audit trails of actions performed by Sun Managed Operations engineers. The customer is required to permit inbound SSH2 traffic from the JumpGate bastion hosts to the CTA.

Management traffic between the JumpGate bastion host and the CTA is encrypted using the 3DES algorithm often referred to as “military grade encryption”. An extra layer of security is added by the CTA as it only permits access from the JumpGate bastion hosts using both firewall and tcp_wrapper technologies and requiring a public key (PKI) based authentication for each engineer before access is permitted. Each login (both successful and failures) is treated as a significant event that the CTA sends back to the CTS for additional audit log creation.

Application and Infrastructure Management

Through the use of the management access architecture described in the previous two sections, the Sun engineer is able to launch a native application management capabilities in order to manage the device within your environment. The CTA provides management capabilities via Windows Terminal Services (RDP), SSH, Telnet, and HTTP. The actual remote management protocol traffic (e.g.: RDP, SSH, Telnet, HTTP) occurs between the CTA and the customers' device, while the rendering of the output (keyboard, video, and mouse) takes places at the JumpGate bastion host. At no point does the bastion host within Sun's network interact directly with the customer's device.

Each management protocol requires the Sun engineer to also authenticate again to the customer's device using native authentication provided by that device. Sun requires a dedicated account on each device for management access in order to provide full accountability of actions.

The combination of Sun and the customer authentication records generates a fully accountable trail of actions performed by Sun on behalf of the customer.

Summary

The CTA encapsulates Sun's operational experience in developing a management and monitoring platform capable of delivering service to today's technology with an eye on emerging advancements in both management and monitoring services and protocols. Its small footprint and support for unreliable transport makes it ideal for delivery in both an Internet data center and a corporate datacenter.