Solaris Trusted Extensions

Architectural Overview

Glenn Faden
April 2006

Table of Contents

Introduction

Solaris™ Trusted Extensions is an optionally-enabled layer of secure labeling technology that allows data security policies to be separated from data ownership. While it has its roots in the multilevel **Trusted Solaris 8** OS, it has been integrated into the standard Solaris 10 Operating System. This new approach allows the Solaris operating system to support both traditional Discretionary Access Control (*DAC*) policies based on ownership, as well as label-based Mandatory Access Control (*MAC*) policies. The label-based policies for file systems and networks are light-weight and have been implemented within the standard Solaris 10 kernel, services and utilities. Unless the Trusted Extensions layer is enabled, all labels are equal so the kernel is not configured to enforce the MAC policies. For efficiency, a boolean value is maintained in the kernel to indicate whether labeling comparisons should be used in policy enforcement.

When the label-based MAC policies are enabled, *all* data flows are restricted based on a comparison of the labels associated with the subjects requesting access and the objects containing the data. Like other multilevel operating systems, Trusted Extensions meets the requirements of the **Common Criteria Labeled Security Protection Profile** (**LSPP**) and the **Role-Based Access Protection Profile (RBAC)**. However, the Trusted Extensions implementation is unique in its ability to provide high assurance, while maximizing compatibility and minimizing overhead.

Compatibility

Maximum compatibility with thousands of Solaris applications is achieved by building on existing Solaris features and using existing industry standards. No new protocols are required nor new file system attributes. Applications do not need to be modified nor profiled to bring them into conformance with the MAC policy. Instead, the entire application environment is virtualized for each label through the use of Solaris Containers (*zones*). This approach is referred to as *polyinstantiation* because there can be an instance of each resource and service available at each label. However, there is also a unification principle known as a *single system image* which is applied to the entire operating environment. All the zones are centrally administered from a special, protected global zone which manages the *Trusted Computing Base* (TCB) known as the *Trusted Path*. The zones share a single LDAP directory in which network-wide policy is defined, as well as a single name service cache daemon for synchronizing local databases. All labeling policy and account management is done from within the Trusted Path. MAC policy enforcement is automatic in labeled zones and applies to all their processes, even those running as `root`. Access to the Global Zone (and hence Trusted Path applications) is restricted to administrative roles.

Efficiency

Minimal overhead is achieved by moving access control checks to a higher level of enforcement. For example, instead of maintaining labels on fine-grained objects like files and directories, Trusted Extensions associates labels with Solaris Containers (zones), and network endpoints. Each zone is assigned a unique sensitivity label and can be customized with its own set of file systems and network resources. Each mounted file system is automatically labeled by the kernel when it is mounted. The file system label is derived from the label of the zone or host which is sharing it. All files and directories within the mounted file system have the same label as their mount point. Because no explicit extensions to the file or file system structure are required, essentially any file system that works on Solaris 10 will work when Trusted Extensions label enforcement is enabled. This

includes Sun issued file systems such as UFS, ZFS, SAM-FS and QFS as well as third-party file systems.

Processes are uniquely labeled according to the zone in which they are executing. All processes within a zone (and their descendants) must have the same label, and are completely isolated from processes in other zones. Unlike other virtualization technologies, there is no performance penalty for executing within a zoned environment as there is no emulation required for a Container. Labeled zones can be instantiated quickly by cloning a copy of a default zone. Disk usage is minimized by sharing immutable instances of most system files and by utilizing copy-on-write technology for the rest.

File System MAC Policy

A zone's local file systems are only writable at the zone's label, but can be shared with labeled zones via loopback or NFS mounts. Loopback mounts are used between zones running on the same host, and multi-level NFS is used for access between hosts. File systems that are shared by all zones on a system are always mounted read-only. This policy provides both confidentiality and integrity protection. Such file systems are assigned the lowest administrative label, `ADMIN_LOW`. Similarly, file systems imported from lower-level zones are assigned the label of the zone with which they are shared.

File sharing between sets of Trusted Extensions systems using NFS can be symmetric. Corresponding zones on each system (with matching labels) can have read-write access to each other's shared file systems. Zones which dominate (have higher labels) than the owning zone can be granted read-only access (depending on per-zone policy settings).

Writing up to higher-level regular files is not possible because such files are never visible within a labeled zone. However, writing up is possible using named pipes which are loopback mounted into higher-level zones. This unidirectional conduit is useful for implementing one-way guards and for tamper-proof logging.

The following is an example of the labels assigned to the mount points in a zone called `needtoknow`, whose label is `CONFIDENTIAL : NEED TO KNOW`. It dominates two zones, `internal` and `public`.

Mount Point	Access	Sensitivity Label
/	Read/Write	CONFIDENTIAL : NEED TO KNOW
/kernel	Read Only	ADMIN_LOW
/lib	Read Only	ADMIN_LOW
/opt	Read Only	ADMIN_LOW
/platform	Read Only	ADMIN_LOW
/sbin	Read Only	ADMIN_LOW
/usr	Read Only	ADMIN_LOW
/var/tsol/doors	Read Only	ADMIN_LOW
/tmp	Read/Write	CONFIDENTIAL : NEED TO KNOW
/var/run	Read/Write	CONFIDENTIAL : NEED TO KNOW

Mount Point	Access	Sensitivity Label
/home/gfaden	Read/Write	CONFIDENTIAL : NEED TO KNOW
/zone/public/export/home/gfaden	Read Only	PUBLIC
/zone/internal/export/home/gfaden	Read Only	CONFIDENTIAL : INTERNAL USE ONLY

Table 1: Labeled mount attributes

To prevent configuration errors and to simplify system administration, there are no interfaces for specifying the labels of mount points. Instead, the kernel determines the labels of all mount points based on host and zone labels, and ensures that the MAC policy is correctly implemented.

Configurable Zone Privileges

By default each labeled zone is completely isolated from all other labeled zones because their labels are required to be unique. No process in a zone can view or signal processes running in other zones. There are no privileges available for any process in a labeled zone to write to lower-level files. However, such policies as reading from files in lower-level zones, exporting directories to higher-level zones, and moving files into higher-level zones can be enabled by specifying the privileges available to each zone when it is booted[1] .

Privileges available to a zone can, in turn, be assigned to processes in the zone. However, a zone's privilege limit is an upper bound that applies to all processes (even root-owned) that are run in the zone. All policies which affect multiple zones, such as sharing of directories, are administered via the Trusted Path.

Label Ranges and Sets

Labels consist of hierarchical components called classifications (or levels) and a non-hierarchical components called compartments (or categories). The mapping of names to classifications and compartments is specified in a database which is private to the Trusted Path. The internal structure of labels is deliberately opaque to users and applications and might change in a future release. At least 256 classifications and 256 compartment bits are supported.

When two labels are compared, the first label can be greater than, less than, equal to, or disjoint from the second label. Classifications are compared as integers, and compartments are compared as bit masks. Labels are disjoint when each contains at least one compartment bit which is not present in the other.

A label range can be specified by an upper bound (called a clearance), and a lower bound. Administrative roles can use the Trusted Path to assign label ranges to users, network attributes, workstations, and allocatable devices.

1 The policy for reading down is configurable because it is not always appropriate. For example, it could result in processes at a higher level executing lower-level applications which manipulate the higher-level data. One way to mitigate that risk is to configure the zone without the privilege to do read down mounts. An additional way is to specify the `noexec` mount option for lower-level mounts.

Network Labeling

As in previous versions of Trusted Solaris software, remote hosts can be single level or multilevel. Single level hosts have an implicit label assigned to them based on their network or IP address. Non-label aware systems, such as workstations running Microsoft Windows™, are assigned a specific label for communications purposes. Multilevel hosts are trusted to operate at a range of labels, and explicitly specify the label of every network packet when communicating with other trusted systems. Packet labels are specified using the Commercial IP Security Option (CIPSO) which encapsulates a sensitivity label as an IP option. CIPSO is specified in the FIPS 188 Standard and is supported by Trusted Solaris 8 and other labeled systems.

When specifying the labeling policy for network attributes, both label ranges and sets of disjoint labels can be enumerated. This ability to precisely define the labeling policy is required to support various multilevel configurations including guards, NFS servers, Sun Ray servers, name servers, print servers, workstations, and high-assurance grid computing. An administrator can also assign a label range to a router even if the router does not interpret labels. Although zones have unique labels, specific multilevel services can be configured for each zone.

The network attributes database is maintained in an `LDAP` directory and shared by all trusted systems comprising a network of multilevel systems. `IPsec` can be used to authenticate the source IP addresses associated with incoming network packets. IPsec enforces integrity protection, and is used to encrypt data on multilevel networks.

Multilevel Networking

Zones can be configured to share a single IP address, or they can be assigned unique IP addresses. Similarly, they can share the same physical network interface, or can be configured to use separate network interfaces. Both shared and per-zone IP addresses can be used concurrently, with different labeling policies for each IP address. Solaris Zones technology allows multiple zones to share a single network interface through the use of virtual interfaces.

Sharing of IP addresses is possible in Trusted Extensions because each packet is implicitly or explicitly labeled. When a packet is received, the kernel uses the label of the packet to determine the appropriate zone to which it should be delivered. Sharing a single IP address for all zones is convenient for workstations and laptops, especially when DHCP is used. This simplifies deployment into infrastructures with limited IP addresses.

Using per-zone IP addresses is required when separate networks are in use, and might be appropriate when multilevel services are being provided. To enable multilevel services, a database of multilevel ports is maintained via the Trusted Path. A multilevel port is a special kind of reserved port whose multilevel semantics are administratively controlled. For each IP address, a range of labels, as well as explicit labels outside of that range can be configured for use by multilevel services. A privileged server can bind to a multilevel port using any IP addresses that are assigned to the server's zone. The server can receive requests at these labels and reply to any request. For multilevel TCP services, the reply is automatically sent using the label of the request without requiring any special programming in the server. For multilevel UDP services, the server must set a socket option to indicate the label of the reply. In either case, the server can query the kernel to determine the label of

each request and then restrict the reply accordingly.

The following block diagram shows some typical network paths that can be administratively configured.

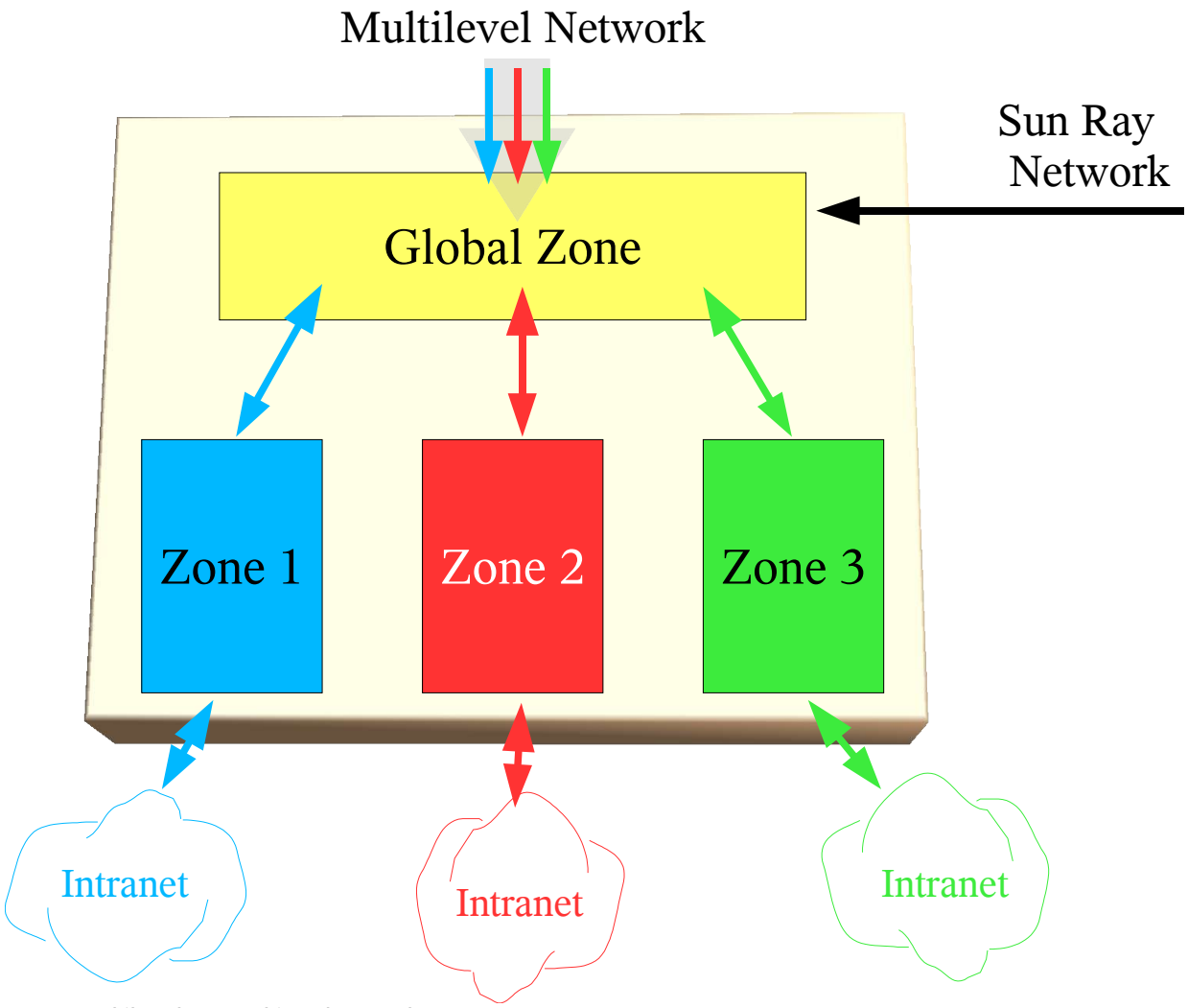

Figure 1: Multilevel networking data paths

Multilevel Services

By default Solaris 10 with Trusted Extensions enables the following multilevel services :

- X11 Window System
    - Common Desktop Environment
    - Java Desktop System
- Printing
    - Internet Protocol Printing

    - BSD Protocol Printing
- Network File System
- Sun Directory Server
- Label Translation Service
- Name Service Cache Daemon

All other services are polyinstantiated in each zone. However, additional multilevel services such as Web Servers and Secure Shell can be enabled administratively via the Trusted Path.

Multilevel Desktop Sessions

Users can log in via the Trusted Path and can be authorized to select their multilevel desktop preference (CDE or JDS). Once authenticated they are presented with an option to select an explicit label or a range of labels within their clearance and the label range of their workstation or desktop unit. The window system initiates a user session in the zone whose label corresponds to the user's default or minimum label.

The window system provides menus for interacting with the Trusted Path to change the label of the current workspace or to create additional labeled workspaces. For each selected label, the window system starts another user session in the corresponding zone. All of these user sessions run concurrently and are subjects of the user's identity that was established during the initial authentication. Each window is visibly labeled according to the zone or host with which it is associated. Although users can simultaneously interact with windows running in multiple zones, the applications themselves remain isolated.

Attempts to cut and paste data, or drag and drop files between clients running in different zone are mediated by the Trusted Path. Specific authorizations are required for upgrading or downgrading selections and files, and are prohibited by default. The following screen shot shows an authorized user interacting with the Trusted Path to upgrade a selection:

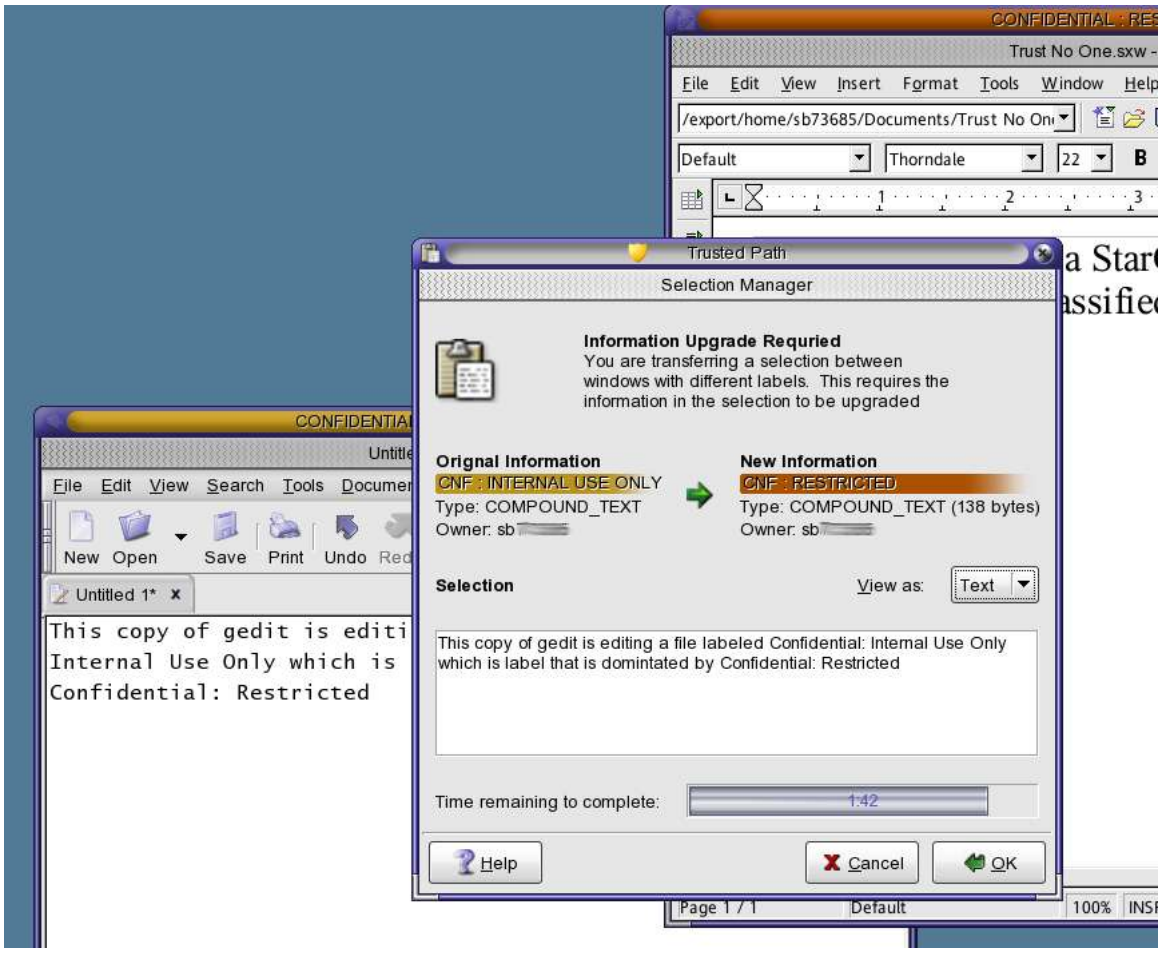

Figure 2: Multilevel Cut and Paste in Trusted JDS

Roles

By using the Trusted Path menu, authorized users can assume one or more roles which they have been assigned. For each assumed role a secondary authentication is required. Once authenticated, the window system creates a new administrative workspace for the role and starts another session. These administrative workspaces are protected from interference by untrusted X11 clients and non-role user logins. For roles who are cleared for all labels, their sessions are initiated as Trusted Path processes. Each role has a limited set of Role Based Access Control [RBAC] rights which restrict its access. Typically, two or more cooperating roles can be used to configure the system. For example, a system administrator role creates accounts and zones, while a security administrator assigns labels to them. Roles with sufficient rights can configure aspects of the MAC and DAC policies that apply to one or more zones

Devices

Devices represent a security threat because they can be used to import and export data from the system. In Trusted Extensions, removable media devices are administered through the Trusted Path menu. The window system provides a Trusted Path interface for device allocation which provides fine-grained access to specific

devices based on user authorizations and label ranges. For example, a user can be authorized to allocate the audio system (speaker and microphone) at a single level. Hot pluggable devices such as USB flash memory drives are also managed by the Trusted Path user interface. An authorized user can request to have such devices mounted into a zone whose label is within the user's label range and the device's label range. As an extra security measure, the raw device is not available within the labeled zone. This capability protects the integrity of the mounted file system and prevents unauthorized access.

Printing

Each printer is assigned a label range from which it will accept requests. Multilevel printers can accept jobs from labeled zones or remote hosts whose labels fall within their range. Each job can be encapsulated between reliably matching banner and trailer pages which indicate the label and handling caveats for the output. Each page can be automatically labeled with headers and footers corresponding to the sensitivity of the data.

Auditing

Trusted Extensions audit records are compatible with standard Solaris OS. They include the labels of subjects and objects, and additional label-related events. The auditing system is configured via the Trusted Path and is transparent to users and roles running in labeled zones. The auditing system is robust and cannot be tampered with by processes running in labeled zones. Even processes with all privileges cannot observe the audit trail nor tamper with any records.

Conclusion

Solaris Trusted Extensions builds on the security features in Solaris 10 in an upward compatible fashion. Users of Solaris 10 can enable Trusted Extensions by installing the Trusted Extensions administrative packages. Because it is built-into the Solaris OS, all of the latest Solaris functionality is supported by Trusted Extensions, and all hardware platforms are supported. The MAC policy is enforced automatically and transparently without requiring customized applications or complex application profiles that are difficult to deploy and maintain. This new approach makes it possible to deliver a cost-effective and interoperable system that simultaneously provides ease-of-use and a high level of information assurance.

Sun Microsystems, Inc. Network Circle, Santa Clara, CA 95054 USA Phone 1-650-960-1300 or 1-800-555-9SUN Web sun.com