Requirements for Recoverable Systems

Abstract

Recoverable Systems incorporate capabilities to repair or restore operating systems, applications, and user data that have been corrupted by malware or misconfiguration.

A Recoverable System is composed of applications called Recovery Agents that perform the servicing operations running on a foundation called a Recovery Capable Platform. The Recovery-Capable Platform protects Recovery Agents from malware so they can reliably be launched when they are needed.

Recoverable Systems capabilities can be applied to any system that includes protected firmware.

Contents

1 Overview

Resilience is the ability to prepare for and adapt to changing conditions as well as withstand and recover from disruptions.[1] In computer systems, anticipated disruptions include cyber-intrusions, accidents, and naturally occurring threats or incidents. Cyber-resilient systems incorporate a wide range of technologies and practices to protect against malware intrusions that might impair the ability of an organization to conduct its mission.

Despite decades of efforts by industry and academia, successful intrusions of computer systems are still commonplace. Once a cyber-intrusion occurs, a resilient computer system must be able to repair or compensate for the damage. This publication specifies the requirements for *Recoverable Systems*: systems that will be able to address the problems that arise *after* a successful destructive cyber-intrusion. Recoverable Systems repair or re-provision computing platforms that have been compromised by malware or misconfiguration.

Today's compromised systems are typically repaired using external media: a DVD, a USB thumb drive, or – in the case of mobile devices – a second computing system. However, initiating recovery using offline media is complex, does not scale well in enterprise data centers, and can introduce performance problems, so some users will be unable to self-recover, and recovery typically requires physical access to the system. These manual recovery processes also do not scale well, particularly in enterprise environments with headless or unattended systems (e.g., servers, network systems, etc.). While network-based recovery mechanisms are available (e.g., PXE-boot based repair and re-provisioning tools), these present their own deployment and security challenges.

Recoverable Systems simplify this process by ensuring that recovery applications are easily accessible but protected from malware that might delete them or interfere with their operation. Recoverable Systems may also allow secure and centralized management of the recovery capabilities. As such, Recoverable Systems provide the reliability of external-media based recovery but at lower operating cost and shorter downtime.

The recovery capabilities described here are applicable to broad classes of computing systems. The essential requirement is that the system can maintain the integrity of the recovery mechanisms described in this publication.

Recoverable Systems will be most effective if they are deployed and managed as part of a comprehensive cyber-resiliency strategy.

[1] Presidential Policy Directive 21.

2 Audience and Scope

2.1 Audience

The audience for this publication is computer-system manufacturers, as well as software vendors and service providers that wish to build Recoverable Systems. This publication also provides non-normative guidance for information systems managers that deploy Recoverable Systems.

2.2 Applicability

Recoverable System capabilities are applicable to a wide range of computer platforms from mobile devices to PCs, servers, and network equipment. Detailed requirements vary by system class and capabilities.

2.3 Scope

This publication includes requirements for the computer platform (hardware or system-manufacturer-provided firmware) as well as requirements for software or firmware applications called *Recovery Agents* that perform system repair. A system that meets the platform requirements alone is termed a *Recovery-Capable Platform*. If the system is also equipped with a conforming Recovery Agent, then the platform is called a *Recoverable System*. See Figure 1 for a schematic representation.

Figure 1: Recoverable System Building Blocks.

2.4 Limitations

This publication assumes that the Recovery-Capable Platform hardware and firmware is robust to cyber-attack and does not provide for firmware recovery if it is vulnerable. Implementers should be aware that the capabilities described in this publication can significantly reduce the disruption of an OS-level cyber-attack, so attackers are likely to devote more attention to subverting firmware.

Some Recoverable Systems will rely upon networked services for their proper operation. For example, Recovery Agents may restore backups from file servers, and IT managers may use centralized management infrastructure to configure and administer the recovery capabilities of their assets. This publication does *not* place requirements on the behavior or management of such infrastructure, but vendors and IT systems managers should be aware that compromise of this infrastructure might impair the operation of the associated Recoverable Systems.

2.5 Recoverable Systems as part of a Comprehensive Cyber-Resiliency Strategy

System recovery and repair is just one element of computer security incident response, which is itself just one part of a comprehensive cyber-resiliency strategy.[2] It is beyond the scope of this publication to describe best practices for designing and managing resilient systems and networks, but two considerations for maximizing the benefit of Recoverable Systems as part of a cyber-resiliency strategy are highlighted.

2.5.1 Infrastructure Remediation

Many Recoverable Systems will rely on network services for management, data, and programs. If individual Recoverable Systems are compromised, then the remaining network services and infrastructure will be available to help the systems remediate. However, if disruption is widespread, then infrastructure will have to be remediated before the relying systems can be restored. Further, if Recoverable Systems management infrastructure is compromised, then recovery capabilities may themselves be disrupted.

In this publication, we refer to the services and infrastructure that support Recoverable Systems as *Critical Recovery Infrastructure*. Enterprises, vendors, and service providers should treat this infrastructure as a *high-impact information system* as defined by NIST Special Publication 200.[3]

2.5.2 Recovery State Management and Remediation

System repair and recovery restores a system to a configuration determined by local or network-hosted data. The data used to restore the system may be the system-configuration at an earlier point in time (*e.g.* a backup), or may be a collection of operating system and program files for fresh installation or repair. In either case, the resulting configuration may still contain malware, software vulnerabilities, or exposure due to compromised administrative accounts that were present in the data files used to restore the system. A comprehensive resiliency strategy should include threat-eradication in the recovery process. For example, recovery images should be patched and up to date, and recovery state data should be scanned for malware. Systems may need to be isolated from other devices on the network during the recovery process to prevent the proliferation of malware across the network.

[2] NIST SP 800-61rev2, *Computer Security Incident Handling Guide*. August 2012.
[3] NIST SP 200, *Minimum Security Requirements for Federal Information and Information Systems*. March 2006.

3 Architecture Overview

In this section, we describe the architecture of Recoverable Systems. This section is not normative: the behavior and terms are informally introduced here and are precisely defined in Sections 4 and 5.

3.1 Basic Operation

Figure 2 is a simplified view of the normal operation and recovery of a Recoverable System.

In normal operation, the operating system and applications start through the “Normal Boot” path - exactly as platforms boot today - and the recovery capabilities lie dormant.

If recovery is required, a *Recovery-Capable Platform* will locate and run *Recovery Agents* that can examine and restore or repair the primary OS and its applications and user data. New platform capabilities either provide for the *protection* of embedded Recovery Agents from corruption or the *reliable download* of Recovery Agents from a network service.

Various events called *Recovery Triggers* can cause a Recovery-Capable Platform to launch the Recovery Agent rather than the Primary OS. Recovery may be automatically initiated (such as when platform firmware or the OS detects corrupt or missing data), manually initiated (for example, when a local user presses a key at boot time), or remotely initiated by an authenticated system administrator using infrastructure management tools.

Once invoked, the Recovery Agent can repair, replace, or restore the primary OS and/or user data. Some Recovery Agents will perform all repair operations using local data. Other Recovery Agents will retrieve programs, settings, and user data from a public or private network service. Once a Recovery Agent’s work is done, it will reboot the system to restore normal system operation.

In the remainder of the Architecture Overview section we describe the functions and management of Recoverable Systems in more detail.

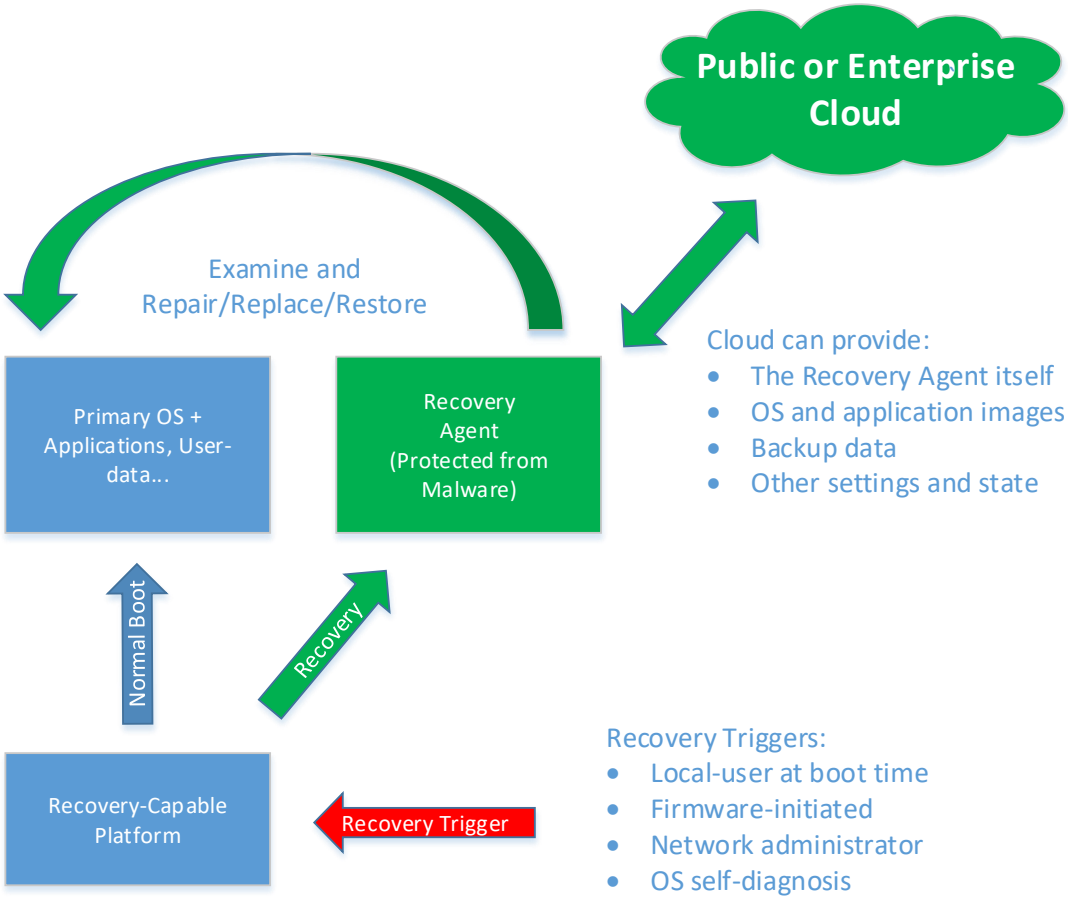

Figure 2: Recoverable Systems Architecture.

3.2 Recovery-Capable Platforms

Recovery-Capable Platforms provide foundational hardware and firmware mechanisms that allow recovery and other maintenance operations to be reliable, discoverable, and manageable.

Recoverable Systems must be robust to malware executing with operating system privileges (for example, rootkits) as well as being robust in the face of compromised operating-system administrative accounts. This robustness can be achieved by embedding all recovery-related behavior in system firmware and allowing little or no management by the operating system or other administrators. Systems that are more flexible may provide more end-user control through protected capabilities that require special privileges to administer (i.e. privileges beyond those normally held by an operating system.) This control is described in more detail in the Recovery Administration section below.

Recovery-Capable Platforms also provide interfaces called Recovery Triggers that can be used to signal that a recovery process should start. Depending on the type of system, interfaces may be provided to allow local users, operating systems, and properly authenticated remote administrators to initiate recovery. Some Recovery-Capable Platforms may allow Recovery Triggers to be enabled or disabled.

Full System Recovery typically involves re-association of recovered machines with enterprise or public network services. Current practice is to either require users or administrators enter passwords, re-associate machines under conditions of physical security, or have deployment servers send credentials to unauthenticated endpoints. It is recommended that Recovery-Capable Platforms contain hardware-based cryptographic system identifiers such as an 802.1AR token[1] or a Trusted Platform Module[2] (TPM) to both simplify and improve the security of system/service re-enrollment.

Finally, Recovery-Capable Platforms that use network-hosted Recovery Agents must have networking firmware capabilities to access the Recovery Agents when they are required.

3.3 Recovery Agents and Actions

Recovery Agents are the programs that perform recovery actions. Recovery Agents can perform a wide range of actions depending on their sophistication. Some of the actions that *must* be provided to meet the requirements in this publication are repair operations (in which system files and settings are examined and repaired if needed), operating system/application re-installations, and system recovery from backups.

Recovery Agents may be part of the platform's firmware or can be software utilities that are installed as part of system manufacture or operating system configuration. Recovery Agents can run on the host processor as long as malware, OS-corruption, or other incidents cannot interfere with them during storage, delivery, or execution. Recovery Agents can also execute on supplementary Service Processors such as Baseboard Management Controllers (BMCs) if contained within the trust domain.

Recovery Agents must be able to restore full operation of computing systems, which means that Recovery Agents need to be able to restore operating systems and applications as well as restoring enterprise or user data.

Recovery Agents may be monolithic applications or may be a collection of utilities that work together or in turn to restore operation of the system. If one Recovery Agent passes control to another Recovery Agent to continue the recovery process, then the prior Recovery Agent must ensure that (a) the latter Recovery Agent is authentic, (b) its execution environment is protected, and (c) any parameters that the latter Agent requires are protected from malware.

For example, a Recoverable System may provide a Recovery Agent launched by firmware that restores an operating system and then starts it. The Recoverable System may then also provide a second Recovery Agent that runs in the freshly repaired operating system to finish recovery by restoring the applications and user data. In this case, the first Recovery Agent must ensure that malware cannot interfere with the execution of the repaired operating system or the second Recovery Agent.

3.4 Recoverable System Management

Recoverable Systems must provide dependable services even when the main operating system is inoperable or has been subverted by malware. It follows that Recovery-Capable Platforms and Recovery Agents *may not* be directly managed by the (potentially untrustworthy) local operating system.

The *Recovery Administrator* is the entity that controls the configuration of the recovery capabilities of a Recoverable Platform. In the very simplest case, the Recoverable System will be pre-configured by the system vendor, and Recovery Administration by customers is neither supported nor required.

Other systems may allow customers to install and configure recovery capabilities. Simple systems may define a single administrative privilege for Recoverable System administration. Systems that are more complex may allow the Recovery Administrator to delegate responsibilities for some actions to other trustworthy parties. For example, an enterprise Recovery Administrator might delegate management of recovery to the equipment vendor.

Recoverable Systems may support more than one Recovery Agent, and each Agent may provide more than one type of recovery service. In such cases, the Recovery Administrator (or delegate) sets the *Recovery Policy* describing the recovery actions that should be performed. The actions may be set in advance, or the Recovery Administrator (or delegate) may be able to choose the action at the time of recovery.

Note that initiating (triggering) recovery usually does not involve a Recovery Administrator. Recovery Administrators are involved when systems are initially configured, and if changes to settings are required.

Figure 3 illustrates administration of the recovery capabilities of a platform with two Recovery Agents. The equipment vendor has pre-configured a platform with an embedded Recovery Agent in local protected storage. The IT department has added a configuration (a *Recovery Profile*) to use an enterprise-specific Recovery Agent on an enterprise file server. The IT department has also configured the Recovery Policy to indicate that the Recoverable Platform should first try the enterprise Recovery Agent. If this fails (for instance if the system is not on the corporate intranet), then the platform is configured to attempt the vendor-provided solution.

Figure 3: Recoverable System Management

Figure 3 also illustrates the use of public keys or certificates to authenticate authorized Recovery Administrators. Cryptography allows key-holders to authenticate themselves or their actions to the Recovery-Capable Platform without needing to trust the intervening software and networks. This allows

a remote key-holder to perform administrative actions even when the (potentially untrustworthy) operating system is running.

Recovery-Capable Platforms may use other techniques to authenticate Recovery Administrators. For example, some systems may use an out-of-band communication channel or allow administration from a management console presented by platform firmware. The essential requirement is that operating-system-level malware of arbitrary sophistication cannot disable or misconfigure the recovery capabilities.

This publication does not place requirements on the methods and security characteristics of the initial installation and configuration of Recovery Agents or the Recovery Administrator, because different systems and scenarios will require different tradeoffs between ease-of-deployment and security. However, platforms that spend an extended period in a state where malware can install a rogue Recovery Agent or Recovery Administrator will be at risk of a malware infection that is hard to remove.

3.5 Support for Backup and Restore in Recoverable Systems

Backups are part of most cyber-resiliency strategies, and the backup process is often automated and transparent to users.

Unfortunately, restoring from a backup after a destructive malware event (or other event that renders a platform unbootable) is complex and time-consuming because the system-restore application must be located and then run on the system. Recoverable Systems simplify this process: Recovery Agents are required to include functions to restore systems from backups. This means that system-restore can be automated or easily user-initiated.

This publication demands Recovery Agents that support reliable, discoverable, and manageable *restore* capability, but does *not* place requirements on how backups are performed and how the resulting backup state is managed. Vendors and managers of end-to-end backup and recovery services are advised to respect the following guidelines to improve cyber-resiliency in the face of malware or enterprise network compromise.

3.5.1 Creating Backups

Backup systems make local or network copies of the system and user state. The copying action may be disrupted if a system becomes compromised by malware, resulting in missing or subverted backup data.

Recoverable Systems do not themselves offer features to improve the reliability or integrity of backups: the benefits of a Recoverable System pertain exclusively to the process of system restoration. Solution and service providers should use standard best practices to mitigate disruption of the backup process: frequent backups, both full-system and differential backups, event-logging, and proper management of backup state.

3.5.2 Management and Protection of Backup State

Backup agents and services need to be able to write backup data to local storage or network services. Malware seeking maximum damage will attempt to use the same channels to delete or disrupt previously created backup data. If these efforts are successful, then system-restoration will fail or the resulting recovered systems may still contain malware.

It is outside the scope of this publication to dictate how these threats should be mitigated, but storage systems and services that allow Recoverable Systems to *add* new data but *do not* allow existing data to be modified or deleted is one suitable approach.

System and service providers should also attempt to mitigate threats arising from backups containing stealthy malware or backups that contain the same software vulnerabilities that were originally exploited.

Finally, the infrastructure and management of backup systems will themselves be the targets of cyber-attack and need to be carefully protected and managed.

4 Definitions

Platform

The hardware and Platform Firmware combination that constitutes the computing system.

Recovery-Capable Platform

A Platform that supports the protection, execution, configuration, and management of Recovery Agents and their actions.

Recovery-Capable Platforms must meet the requirements of Section 5.1.

Recovery Agent

Computer programs that perform recovery-related maintenance functions on the System Software, applications, and data.

Recovery Agents on Recovery-Capable Platforms can either be stored on the Platform (in a location where they are protected from System Software) or in a network location where they can be accessed when needed. Recovery Agents can execute on the Host Processor or on supplemental management processors.

Recovery Agents must meet the requirements of Section 5.2.

Recoverable System

The combination of a Recovery-Capable Platform and a least one Recovery Agent.

Recovery Firmware

Code provided by the manufacturer of a system or Platform that implements the functions of a Recovery-Capable Platform. Recovery Firmware may execute on the Platform's host processor and/or on non-host processors. If Recovery Firmware is updatable, then updates must be authenticated as originating from the same manufacturer.

This publication assumes that the Recovery Firmware is protected and remains in a state of integrity.

System Software

The privileged software supporting normal system operation. System Software may be a full operating system, a simple library operating system, a hypervisor, or an application running directly on the Platform's host processor. Recoverable Systems rely on Recovery Firmware to recover System Software if System Software is compromised by malware or misconfiguration.

In the context of embedded systems, network equipment, etc., the operating system running on the device is sometimes termed firmware because it is installed by the device vendor. In this document, the essential difference between System Software and Recovery Firmware is that they are protected using different mechanisms or policies. Specifically, System Software may be compromised while Recovery Firmware remains in a state of integrity, allowing Recovery Firmware to repair (or bootstrap the repair

of) System Software. Embedded systems with firmware that can be partitioned into "System Software" and "Recovery Firmware" portions that meet the relevant requirements in this publication are Recoverable Systems.

Recoverable Systems must provide reliable recovery services even if System Software is controlled by an adversary via Malware or a compromised administrative account. Hence, System Software must not be able to directly or independently change Recovery Agents or associated Critical Recovery Settings.

Advanced Persistent Threat (APT)

An adversary possessing substantial resources, motivation, patience, and skill who intends to exfiltrate confidential information, or corrupt, subvert, or disable the platform, its software (OS or applications), or its firmware.

Platform Cryptographic Identities

An identity that uniquely identifies a given platform and allows it to cryptographically authenticate itself to another entity. These identities typically bind a unique system name to a public key.

Malware

Software designed to disrupt computer operation or gather sensitive information. The capabilities described here address Malware that has infiltrated System Software.

Service Processor

A microcontroller that system administrators can use to manage a Platform even when there is no operating system running.

Platform Identity

A permanent cryptographic identifier for the platform that can be used to establish secure communications or reliably identify a system even after a destructive Malware event.

Host Processor

The primary processing unit in a Platform. Host Processors are also called CPUs and APUs. System Software, as well as user applications runs on this processor.

Recovery Action

The operations that Recovery Agents perform on System Software, applications, and data. Recovery Actions include repairing and reinstalling operating systems and appropriate patches as well as restoring the user state.

Recovery State

Data used by the Recovery Agent to reconstitute a functioning computer platform. Recovery State may be stored in a local protected location, may be network-hosted, or some combination of the two. This publication uses the following informal categorizations of Recovery State.

Common State

Data needed to restore normal operation of the operating system and applications, but *not* restore enterprise or user configuration. Common state is typically composed of an initial software distribution (System Software and applications), ideally modified by of a sequence of patches and updates that repair defects.

Common State is generally not confidential but must be integrity-protected and be readily available when needed.

The user of a Recoverable Platform that has only applied Common State will need to perform additional manual or automated configuration steps to restore full operation.

Imprinting State

Data that must be applied in addition to Common State to customize the computing system for a specific user or enterprise. Imprinting State may include the URIs of important network services such as enterprise email servers, as well as the configuration of local user accounts, local administrator accounts, time zones, language-locales, etc. Some systems may require or allow local users to provide Imprinting State (e.g. a user name and password) as part of recovery.

Imprinting State may contain confidential data.

User State

All user and application data on the platform. For client systems, this may include documents, cached email, and other user configuration data and settings. For server systems, this may include server or application data – for example database files. For network equipment, examples include firewall rules or router configuration data.

Once Common, Imprinting, and User State have been applied, the system is fully functional.

User State will usually contain confidential data.

Combined State

The combination of Common, Imprinting, and User State. Used, for example, when describing the data produced during a full-system backup when state is intermixed.

Combined State will usually contain confidential data.

Recovery Agent Protected Storage

An optional Platform-protected capability for storing a local Recovery Agent. (Recovery Agent Protected Storage will not be needed if the Recovery Agent is network-hosted.)

Critical Recovery Settings

Security-critical configuration parameters that affect the recovery process. Although this publication does not mandate implementers use any particular settings, examples of Critical Recovery Settings include the network location and TLS/SSL-certificate of an enterprise backup image server.

Critical Recovery Storage

A facility for securely storing Critical Recovery Settings.

Recovery Profile

The combination of a Recovery Agent and its associated Critical Recovery Settings. See Figure 3.

Recovery Administrator Privilege

Privileges required to manage the recovery-related behavior of Recoverable Systems. In particular, this privilege is needed to change Critical Recovery Settings, Recovery Policy, Recovery State, and the Recovery Agent itself.

Implementers of Recoverable Platforms may define a single Recovery Administrator Privilege, or may allow delegation of a set of reduced privileges that can be individually assigned to autonomous parties such as end-users, software, and equipment vendors.

Adversaries that obtain Recover Administrator Privileges can reduce the dependability of recovery mechanisms or install malware that is hard to remove. As such, the privilege must be carefully controlled and should never be directly available to System Software once provisioning is complete.

Recovery Administrator

An entity with Recovery Administrator Privileges.

Recovery Policy

The policy that controls the recovery-related behavior of a Recovery-Capable Platform. Recovery Policy may describe the actions that the Platform will take if recovery is triggered, which Recovery Triggers are enabled, whether users demonstrating Unambiguous Physical Presence can perform recovery-related administration, and the set of authorized Recovery Agents or software publishers.

If the Recovery Policy is configurable, it is under the control of the Recovery Administrator, and *not* under the control of System Software.

Unambiguous Physical Presence

Interaction by a local user that cannot be impersonated by OS-present Malware.

Unambiguous Physical Presence requires a channel from the physically present user to the recovery capabilities that cannot be disturbed by (potentially untrustworthy) System Software. Two examples are a BIOS/Firmware-configuration screen and a GUI that is presented to the user by the Recovery Agent itself.

Some systems may allow Firmware configuration screens to be accessed remotely. Local System Software should not be able to use this channel to subvert the recovery capabilities, although the channel may provide other points of attack.

Recovery Trigger

A signal to a Recovery-Capable Platform that recovery should be initiated. A Recovery-Capable Platform may provide multiple Recovery Triggers. For example, triggers a boot-time key-press, an interface accessible to System Software, or a trigger that can be invoked remotely via a Service Processor.

Recovery Agent Loader

The component in a Recovery-Capable Platform that responds to a Recovery Trigger and locates, authenticates, and executes a Recovery Agent. If the Recovery Agent is stored remotely, the Recovery Agent Loader is also responsible for downloading the Recovery Agent from the network.

The Recovery Agent Loader can execute (suitably protected) on the main processor or it may execute on a supplemental management processor. The Recovery Agent Loader may consult a Recovery Policy to determine which Recovery Agent to start. The Recovery Agent Loader may also allow interactive users with Unambiguous Physical Presence to choose between installed Recovery Agents or Recovery Profiles.

Secure Boot

A computer initialization sequence in which software components are checked against a security policy before they are allowed to execute.

Critical Recovery Infrastructure

The people, systems, and services that manage the Recovery Agents, Recovery State, and Critical Recovery Settings of Recoverable Systems.

5 Requirements

The following two subsections describe the requirements for a Recoverable System. Section 5.1 contains requirements for the behavior of a *Recovery-Capable Platform*, and section 5.2 contains requirements for a conforming *Recovery Agent*.

In these sections, the normative terms "MUST", "MUST NOT", "SHOULD", "SHOULD NOT", and "MAY" are used in accordance with IETF RFC 2119.[3]

5.1 Recovery-Capable Platform Requirements

A Recovery-Capable Platform must meet the requirements in this section.

5.1.1 Cryptographic Algorithms

All cryptographic algorithms and key sizes used MUST be currently approved by NIST.

5.1.2 Platform Identity Requirements

Recovery-Capable Platforms MAY incorporate one or more unique platform cryptographic identities (e.g. using an 802.1AR Secure System Identity, or a Trusted Platform Module). For example, these identities may be used to re-associate recovered systems with enterprise networks, or to authenticate systems to servers containing Imprinting and User State.

If implemented:

- Firmware System Software and Recovery Agents MUST be able to use Platform Identities to cryptographically identify the system and establish secure communications.
- Recovery-Capable Platforms MAY be furnished with separate, dedicated Platform Identities for Firmware, System Software and Recovery Agents.
- Platform Identities MAY be used to report (attest) the identity and configuration of Firmware and/or software running on the Platform, as well as the Critical Recovery Settings.
- The Platform MUST protect Platform Identity private or secret keys from unauthorized access or modification to prevent the cloning or misuse of Platform Identities.
    - Platform Identity private or secret keys MUST be protected from modification by System software.
- If a Platform Identity is an asymmetric key, then the corresponding public key and certificate MUST be protected from modification by System Software.

Note that Cryptographic Platform Identities have privacy implications and vendors should enforce or enable appropriate control of their use.

5.1.3 Strong Authentication of Users and Administrators

Recovery-Capable Platforms need to provide dependable service recovery capabilities even when System Software has been compromised by Malware. If Recovery-Capable Platforms and Recovery Agents are configurable, then recovery configuration parameters and policies can only be modified or deleted by identified, authenticated, trusted, and appropriate privileged entities. These changes need to be communicated over trusted channels, including the channel provided by Unambiguous Physical Presence.

Recoverable Systems may allow local administration, remote administration, or both. An example of local administration is a user interacting with a Firmware (e.g. BIOS) setup screen. An example of remote administration is an authorized entity interacting with a Service Processor. Recoverable Systems may also include physical management interconnects that can be linked to external management infrastructure. For example, a Recoverable System server blade may be managed by a service processor in a rack backplane via a bus interconnect.

Remote administration may be performed using an out-of-band communication network or may be performed using System Software as an untrusted conduit. In either case, administrative actions or commands MUST use a cryptographically protected channel that ensures the confidentiality, integrity and authenticity of these communications.

- Recovery-Capable Platforms that support ***local administration***:
  - MUST ensure that only local users and administrators demonstrating Unambiguous Physical Presence can administer the recovery-related behavior of the platform.
  - SHOULD be able to authenticate local users and administrators using at least single-factor authentication.
- Recovery-Capable Platforms and Recovery Agents that support ***remote administration:***
  - MUST authenticate users or entities authorized to administer a system using cryptographic means.
    - Remote administration MAY employ (potentially untrustworthy) System Software as a communication conduit as long as the cryptographic protocols in use prevent System Software from intercepting or replaying authentication credentials, or modifying the administrative commands.
- Recovery-Capable Platforms that include a **bus interconnect for administration**:
  - MUST ensure that local System Software cannot use this channel for recoverable system administration.

NIST Special Publication 800-63-2, *Electronic Authentication Guideline*, provides technical guidelines on authentication tokens, processes and protocols for remote authentication to IT systems at four electronic authentication assurance levels. Some of the methods described at electronic authentication levels 3 and 4 would support implementation of the cryptographic authentication methods required for remote administration.

5.1.4 Recovery Administrator Privilege Requirements

The recovery-related behavior of Recoverable Systems is configured and managed by entities called *Recovery Administrators,* which are entities that hold Recovery Administrator Privileges. Since Recoverable Systems must provide dependable service in the face of Malware, System Software should *not* hold Recovery Administrator Privileges post-provisioning.

Simple systems may have fixed recovery-related behavior and not support customization. However, if customer-customization *is* supported, then Recovery-Capable Platforms MUST define a Recovery Administrator Privilege or group of Privileges that meet the following requirements:

- Recovery-Capable Platforms MUST authenticate Recovery Administrators and their actions using schemes that meet the requirements of section 5.1.3 (Strong Authentication of Users and Administrators)

- Recovery-Capable Platforms MAY define a single Recovery Administration Privilege that enables management of all recovery-related state, or MAY define a group of delegated privileges that can be individually assigned.

The specific actions requiring Recovery Administrator Privileges are detailed elsewhere in this publication. When an action demands Recovery Administrator Privileges, this SHALL be understood to mean that System Software is *not* itself authorized to perform the action.

Note that triggering recovery does *not* require Recovery Administrator Privileges.

5.1.5 Recovery of Compromised Recovery Capabilities

Recovery Capable Platforms provide dependable service in the face of OS-level Malware, but this means that compromise or loss of a Recovery Administrator privilege can result in misconfiguration that cannot be repaired using normal OS-present tools.

Vendors that allow customer-management of Recovery Capabilities SHOULD provide facilities to repair platforms with hijacked recovery capabilities. If implemented, these facilities MUST NOT be accessible to System Software (to prevent access by OS-level Malware), but MAY be accessible to users demonstrating Unambiguous Physical Presence or using other channels appropriate to the type of system and management model.

5.1.6 Provisioning

This publication does not impose requirements on how and when the Recovery Administrator Privilege is initially assigned and configured. In addition, the behavior of a Recoverable Platform *before* a Recovery Administrator is assigned is not defined by this publication. This flexibility is provided to support different models of manufacturing and customer provisioning.

Once recovery capabilities have been configured and Recovery Administrators have been assigned, Recovery Administrators are the only entities that can change the recovery-related configuration; this is to ensure that local Malware cannot disable or misconfigure the recovery capabilities. Note that some platforms may allow a Recovery Administrator (or users demonstrating Unambiguous Physical Presence) to “factory reset” the system to its initial configuration - e.g. one of the example configurations listed below.

This section contains no normative requirements, but describes three illustrative approaches to provisioning that can be compliant with this publication.

Pre-Provisioned

All recovery-related behavior of the Recoverable System is pre-configured by the manufacturer. If ongoing management of the recovery capabilities is required, then the manufacturer will serve as the Recovery Administrator.

Manufacturers may allow owners to replace or remove the manufacturer-supplied recovery solutions, for example by demonstrating Unambiguous Physical Presence.

Bare-Metal

A Recovery Capable Platform is sold without Recovery Agents, Recovery Profiles, or an installed Recovery Administrator. Tools and utilities can be used to install Recovery Profiles and assign an appropriate Recovery Administrator.

Bare-Metal Recovery-Capable systems are not resilient in their initial, default configuration; explicit configuration by the end user is required to provide recovery capabilities. End users who do not intend to use recovery capabilities may need to configure the platform to prevent a malicious entity from hijacking these capabilities.

Partially Provisioned

The Recoverable System is Pre-Provisioned, but is sold in a state where the customer can either accept the manufacturer default, add to it, or install their own recovery solution and Recovery Administrator. If a vendor uses this approach, it should be easy or automatic for users to accept the manufacturer default since none of the Malware-protection requirements of this publication apply until provisioning is complete.

This option assumes that most customers will choose the manufacturer default, but allows customers with specialized needs to easily override or adapt the default when the systems are initially deployed.

Partially Provisioned Systems are not resilient until provisioning is complete, so vendors may consider designs in which defaults are automatically accepted after a short period of use.

5.1.7 Critical Recovery Settings - Protection and Management

Critical Recovery Settings are configuration parameters held in Critical Recovery Storage that control the recovery process. For example, a Recovery-Capable Platform might define a Critical Recovery Setting that designates the network address of a service hosting a Recovery Agent. Recoverable Systems must provide robust services in the face of Malware, so modification of Critical Recovery Settings requires Recovery Administrator Privileges.

Not all systems will support customer configuration, but if implemented:

- Critical Recovery Settings MUST only be modifiable by entities holding Recovery Administrator Privileges (and specifically MUST NOT be independently modifiable by System Software.)
- Critical Recovery Settings that contain or may contain confidential data MUST NOT be readable by System Software.
- If a Recovery-Capable Platform defines a group of delegated Recovery Administrator Privileges that can be assigned to individual entities, then management of Critical Recovery Settings SHOULD be partitioned and management of the subsets SHOULD be assigned to appropriate entities.

(The latter requirement seeks to limit the harm that can be done by a compromised delegated-Recovery Administrator.)

Critical Recovery Storage may be implemented by any means that meets the requirements above. If cryptographic techniques are employed, then algorithms used MUST meet the requirements of section 5.1.1.

5.1.8 Installation and Management of Recovery Profiles

A Recovery Profile is the combination of a Recovery Agent and its associated Critical Recovery Settings. Recovery Administrators are the entities that are authorized to install, manage, and remove Recovery Profiles.

- A Recovery-Capable Platform MUST either:
  1) Be pre-configured with a Recovery Profile with capabilities that meet the requirements of this publication, or
  2) Allow customer installation of a Recovery Profile.
- A Recovery-Capable Platform MAY support the installation of additional Recovery Profiles.
- Once a System is provisioned, Recovery Profiles MUST only be modifiable by Recovery Administrators.
- Recoverable Platforms SHOULD support authoritative reporting of installed Recovery Profiles, for example using a TPM.

5.1.9 Recovery Policy

The Recovery Policy is the subset Critical Recovery Settings that dictate the overall behavior of Recoverable Platforms. The Recovery Policy is under the control of Recovery Administrators. Simple systems might not support customer-configuration, but if implemented, the Recovery Policy:

- Once a System is provisioned, MUST only be modifiable by Recovery Administrators.
- MAY allow Recovery Triggers to be selectively enabled and disabled.
- SHOULD specify the set of administrative actions that local users (i.e. users demonstrating Unambiguous Physical Presence) are authorized to perform.
- MUST cryptographically specify the authorized Recovery Agents or Recovery Agent publishers.
- If public-key infrastructure (i.e. certificates) are used to identify authorized Recovery Agent publishers, then the Recovery-Capable Platform MUST allow for selective revocation of specific Recovery Agents or publishers.
- If the Recovery Agent is stored remotely, then:
  - The network address of the Recovery Agent MUST be held in Critical Recovery Storage.
  - The Recovery Agent Loader MUST cryptographically verify the integrity and authenticity of the Recovery Agent.
  - The Recovery Agent Loader SHOULD download the Recovery Agent over a cryptographically protected channel that verifies the identity of the service and protects the integrity of communications.
  - The identity of the cryptographic authority for validating services or the authenticity of Recovery Agents themselves MUST be held in Critical Recovery Storage.
- If a Recovery-Capable Platform supports more than one Recovery Profile, then the Recovery Policy:
  - MUST indicate a default Recovery Profile,
  - SHOULD dictate an order of invocation if a Recovery Agent fails or is inaccessible,
  - MAY indicate that users demonstrating Unambiguous Physical Presence can choose amongst installed Recovery Profiles.

    - MAY indicate that users demonstrating Unambiguous Presence can override the specified behavior and indicate an alternative network address or local system as the source of the Recovery Agent (e.g. a USB-storage system.)
- The Recovery Profile invoked MAY depend on the specific Recovery Trigger than was invoked, and MAY allow the Recovery Administrator to override the default behavior at the time of recovery.

5.1.10 Recovery Agent Loader

The Recovery Agent Loader is responsible for locating and starting a Recovery Agent when signaled to do so by a Recovery Trigger. This loader will typically be part of the platform, either part of the platform boot firmware or the firmware on a Service Processor. As an essential part of the recovery process, the Recovery Agent Loader needs protection to ensure it will function reliably.

- The Recovery Agent Loader MUST be protected from System Software.
- The Recovery Agent Loader MUST start a Recovery Agent when recovery is signaled through a Recovery Trigger per Recovery Policy.
- The Recovery Agent Loader MUST interpret and enforce requirements in the Recovery Policy section that relate to the acquisition, authentication, and invocation of Recovery Agents.
- The Recovery Agent Loader MUST provide Recovery Agents with an execution environment that is protected from any Malware that was may have been present in System Software prior to recovery. This:
    - MAY be accomplished using a system reset that results in a protected environment for the Recovery Agents,
    - MAY be accomplished by executing the Recovery Agents on a Service Processor, or
    - MAY be accomplished by other means.
- Recovery Agent Loaders SHOULD securely log recovery Trigger events.
- If more than one type of Recovery Trigger is provided by the Platform, the Recovery Agent Loader SHOULD communicate to the Recovery Agent which triggering event occurred.

5.1.11 Recovery Trigger

Recovery Triggers signal the Platform to invoke the Recovery Agent Loader, which will then locate, authenticate, and execute a Recovery Agent. Recovery Triggers can be unauthenticated and untrusted. Malicious entities or code operating at the platform or OS level may be capable of invoking some of these triggers. As such, it is the responsibility of the Recovery Agent Loader and the Recovery Agent to determine what recovery actions are taken. Additionally, some Recoverable Systems may allow Recovery Triggers to be selectively enabled or disabled. This behavior is specified in section 5.1.9 – Recovery Policy.

Recovery Triggers that are remotely accessible, and might be used for large-scale denial of service, require strong authentication.

This subsection identifies requirements for Recovery Triggers that can be invoked by:

1) The local interactive user,
2) Platform firmware,
3) System Software, and
4) Remote administrators

At least one Recovery Trigger MUST be provided for a platform to be compliant, and Recovery-Capable Platforms MAY define additional triggers as appropriate. For each trigger that is implemented, this section defines relevant requirements.

5.1.11.1 Local Interactive User Initiated Recovery

- Recovery-Capable Platforms designed for local interactive use SHOULD provide a Recovery Trigger that is accessible to the local user (for example, a button.)
- If implemented, the Recovery-Capable Platform MUST ensure that this trigger can only be invoked by users demonstrating Unambiguous Physical Presence.

5.1.11.2 Platform-Initiated Recovery

- Recoverable System Firmware SHOULD provide a trigger that can be invoked by the Platform when normal boot fails.
- If implemented, the platform MUST be capable of invoking this trigger when it cannot find a suitable boot application, when the boot application fails to run, or when it detects a problem with the boot application.

5.1.11.3 Software-Initiated Recovery

- Recovery-Capable Platforms SHOULD define a Recovery Trigger that is accessible to operating-system loaders and other System Software.

5.1.11.4 Remotely-Initiated Recovery

If a Recovery-Capable Platform supports remotely initiated recovery, for example via a Service Processor:

- The Recovery-Capable Platform MUST authenticate that the requesting entity is authorized to invoke recovery according to the requirements in Section 5.1.3.

5.1.12 Recovery Agent Protected Storage

If a Recovery-Capable Platform supports embedded Recovery Agents, then Recovery Agent Protected Storage MUST be provided. Recovery Agent Protected Storage:

- MUST NOT be writable or modifiable by System Software.
- SHOULD be accessible to Recovery Administrators (to allow Recovery Agents to be installed and serviced.)

5.1.13 Recovery Agent Servicing

Recovery Agents may need to be serviced to correct problems or add functionality. Recovery Administrators (or their delegates) are responsible for authorizing these changes to Recoverable Systems.

If the Recovery Agent is hosted on a network location, servicing can be accomplished by updating the agent stored on that server.

If a Recovery Agent is embedded, then it SHOULD be serviceable by the Recovery Administrator. If the Recovery Agent is serviceable, then updates to the Recovery Agent MUST use an authenticated update process, absent physical intervention, that meets the following requirements:

- Updates to the Recovery Agent MUST be digitally signed by Recovery Agent Publishers or Recovery Administrators.
    - Recovery Agents MUST be signed in conformance with NIST SP 800-89, Recommendation for Obtaining Assurances for Digital Signature Applications [SP800-89], using an approved digital signature algorithm as specified in NIST FIPS 186-4, Digital Signature Standard, that provides at least 112 bits of security strength, in accordance with NIST SP 800-131A, Transitions: Recommendation for Transitioning the Use of Cryptographic Algorithms and Key Lengths [SP800-131A].
- The Platform or Recovery Agent MUST provide a Root of Trust for Update (RTU) capable of authenticating updates to the Recovery Agent.
    - The RTU MUST be protected from unauthorized modification. In particular, System Software MUST NOT be capable of bypassing protections on the RTU.
    - If the RTU is implemented in, or protected by, system Firmware, then the system Firmware MUST meet the applicable requirements in NIST SP 800-147 or NIST SP 800-147B.
    - If the RTU is implemented in, or protected by, other Platform Firmware, then updates to that Firmware MUST be by way of an authenticated update process, absent physical intervention.
- RTU MUST contain the digital signature verification algorithm and a key store, or the key store MUST be held in Critical Recovery Settings. The key store MUST include the public key needed to verify the signature of the Recovery Agent or an approved cryptographic hash of the key.
    - In the case of limited storage within an RTU, verification of a Recovery Agent may be established in two stages: (1) verification of the cryptographic hash of the supplied public key, and (2) verification of the signature of the Recovery Agent using the supplied public key.
- The Authenticated update process MUST ensure that updates to the Recovery Agent have been digitally signed and that the digital signature can be verified using a key in the RTU prior to writing the updated agent to Recovery Agent Protected Storage.
- Updates or other servicing actions SHOULD be logged securely.
- The authenticated update process SHALL be the exclusive mechanism for modifying Recovery Agents, absent physical intervention by authorized administrators.

5.1.14 Recovery State Protection

Recoverable Systems may store Recovery State locally, remotely, or a combination of the two. Additionally, some Recovery State may be provided by end users of interactive platforms (for example, user accounts or email names and passwords.)

Remote-Network-Hosted Recovery State

This publication does not place requirements on the management of Recovery State stored on network servers, but the following considerations will lead to reliable and effective recovery:

- Network servers and services hosting Recovery State will themselves be subject to cyber-attack and should be treated as Critical Recovery Infrastructure in their configuration, management and use.
- Common State should be updated to most recent (patched) versions whenever possible.

- State should be scanned for Malware.
- Imprinting and User-State containing confidential data should only be delivered to properly authenticated endpoints using known Platform Identities.
- Imprinting and User-State containing confidential data should only be delivered to endpoints that are running appropriately configured authorized System Software.

Embedded Recovery State

Recovery-Capable Platforms may provide local storage for Recovery State. If implemented, solution providers MUST protect embedded recovery state from destruction or subversion by Malware.

5.2 Recovery Agent Requirements

Recovery Agents are applications that perform servicing operations on system state. The range of possible servicing and architecture options for Recovery Agents is large; however, certain baseline functionality is required of conforming Recovery Agents in Recoverable Systems, as described in this section.

Recovery Agents depend upon the foundational security properties of a Recovery Capable Platform to provide reliable service, as described in section 5.1.

5.2.1 Cooperating Recovery Agents

A Recovery Agent can be a single monolithic application in Firmware and/or Software, or can be a collection of applications and workflows that accomplish a Recovery Action. For example, an initial Recovery Agent might restore the base operating system and then start the freshly restored OS so that an OS-resident Recovery Agent can restore the remaining Imprinting and User State.

If a Recovery Action is accomplished using a series of cooperating Recovery Agents then earlier Recovery Agents are responsible for ensuring the both the authenticity and integrity of the next Agent, as well as the integrity and security of the environment in which it runs.

Specifically, if recovery is structured as a series of cooperating Recovery Agents:

- Earlier Recovery Agents MUST provide an execution environment in which later Recovery Agents can execute reliably.
- Earlier Recovery Agents MUST ensure the authenticity and integrity of later Recovery Agents prior to their execution.
- Earlier Recovery Agents MAY pass parameters to later Recovery Agents.

5.2.2 Data at Rest Protection

If a Recoverable System protects Data using encryption (or by other means), then this section details how Recovery and Data at Rest Protections should interact. The following requirements apply to Recovery Agent actions that service existing state (either by modifying the local version or by restoring state).

- Protections for data and state (e.g. cryptographic protection or other access policies) MUST be re-applied or restored upon successful completion of Recovery Actions, OR MUST be maintained throughout Recovery

    - Note that if a system is restored from a backup, then this might result in access protections being reverted to those in effect at an earlier time.
- If cryptographic keys or credentials are needed for Recovery (for example, a disk encryption key), then the keys or credentials MUST be protected by the recovery capabilities of the System
    - However, if keys of credentials are provided by an external entity (for example, a console user entering a recovery key, or a recovery key supplied by a network service after authenticated that the system is authorized), then these keys or credentials do not need to be protected from the entity that supplied the information while recovery is in process.
- If cryptography is used for data or key protection, then the cryptographic strength of protection during and after recovery MUST be as high as that used by the system being recovered.

If the Recovery Agent installs new software and data (as opposed to repairing an existing configuration), then the Recovery Agent MUST either meet the servicing requirements above for Recovery Agents that service existing state, or erase, or cryptographically erase all prior state.

5.3 Recovery Actions

This section enumerates the recovery and servicing features of conforming Recovery Agents (or a set of cooperating Recovery Agents that meet the requirements of section 5.2.1).

Common State Recovery:

Recovery Agents:

- MUST be able to install System Software and applications based on Common Recovery State, discarding or erasing all previous state.
- SHOULD be able to install and/or repair System Software and applications based on Common Recovery State, while endeavoring to *preserve or restore* Imprinting and User state.
- MUST check the authenticity and integrity of Common Recovery State using cryptographic means (hashing, certificate validation, etc.) The governing security policy (e.g. root certificate, or image hash) MUST be embedded in the Recovery Agent, or MUST be a Critical Recovery Setting (consistent with the requirements in section 5.1.7).
- MUST be able to configure any Platform settings that are needed to properly execute the repaired or newly installed System Software.

For example, on a UEFI-system this would include the boot-related variables.

Imprinting State Recovery

In addition to the Common State Recovery Requirements:

- Recovery Agents MUST be able to apply available Imprinting State (for example, to establish accounts, re-associate Recoverable Systems with essential services, etc.)
- If Imprinting State is stored locally, the Recovery Agents MUST validate its integrity, for example by validating a digital signature.
- If Imprinting State is stored remotely, then Recovery Agents MUST either validate state integrity (for example, by validating a digital signature), or use a cryptographically protected channel to

download the Imprinting State from an authenticated and authorized source (for example, an IT-department server.)

- Imprinting State Recovery MAY be an automatic part of the recovery process, or MAY require inputs from local users or administrators to perform these actions (for example, an account name and password).

User State Recovery

In addition to the Imprinting and Common State Recovery Requirements:

- Recovery Agents MUST be able to restore available User State.
- If User State is stored remotely, then Recovery Agents MUST either validate state integrity (for example, by validating a digital signature), or use a cryptographically protected channel to download the Imprinting State from an authenticated and authorized source (for example, an IT-department server.)
- Recovery Agents MAY be an automatic part of the recovery process, or MAY require inputs from local users to perform these actions (for example, an account name and password.)

5.3.1 Considerations for Recovery State Management

The ultimate configuration of a recovered system will depend on the Recovery State data that the Recovery Agents use to reconstitute the system. The management of the Recovery State data is beyond the scope of this publication, but solutions should ensure that Recovery State repositories contain up-to-date and patched images, and is free from Malware.

5.3.2 Platform Identity and Attestation

Recovery State typically contains confidential information, so network-based recovery services need to ensure that it can only be accessed by authorized systems running known and authorized software. Recovery-Capable Platforms accommodate these requirements by (optionally) including a Cryptographic Platform Identity such as an 802.1AR security token, or Trusted Platform Module (TPM.)

If the Recovery-Capable Platform includes a Cryptographic Platform Identity or Identities, Recovery Agents:

- SHOULD support strong authentication of the Platform to network services.

If the Recovery-Capable Platform includes a Cryptographic Platform Identity or Identities that support attestation, Recovery Agents:

- SHOULD support attestation of the Recovery Agent identity and security configuration to network services.

5.3.3 Parameterized Recovery Agents

Recovery Agents may be fixed-function or their behavior may be controlled by Recovery Administrators through configuration. For example, a Recovery Agent that restores a backup from a network service may require a parameter that denotes the network address of the backup data.

If Recovery Agents employ security-critical parameters, then:

- The parameters MUST be treated as Critical Recovery Settings under the control of Recovery Administrators to ensure that they cannot be modified by Malware.

Selection of Recovery Actions

If a Recovery Agent offers more than one recovery service, then the Recovery Policy determines the actions that are to be taken.

The Recovery Policy MUST include configuration settings that determine the specific Recovery Action or actions to be taken. The actions MAY depend on how recovery was triggered, and MAY allow a user with Unambiguous Physical Presence to select the Recovery Action.

For interactive client systems, the Recovery Policy for triggers that might occur accidentally or are potentially accessible to Malware SHOULD require that a User with Unambiguous Physical Presence or a Recovery Administrator selects or confirms the Recovery Action.

Network-Initiated Recovery by an authenticated administrator SHOULD need no further confirmation.

5.3.4 Logging

Recovery Agents SHOULD securely log Recovery Actions and outcomes.

6 Conclusions

Systems that meet the requirements set forth in this publication will provide users and system administrators' discoverable and reliable machine recovery for many classes of cyber-attack.

---

[1] 802.1AR Secure System Identity. Ed Mike Borza and Max Pritikin.

[2] TPM Library Specification Trusted Computing Group.

[3] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", Internet Engineering Task Force RFC 2119, March 1997. http://www.ietf.org/rfc/rfc2119.txt.