Boot Manager Technical Documentation

Aaron Klingaman <alk@cs.princeton.edu>

Table of Contents

1. Components

The entire Boot Manager system consists of several components that are designed to work together to provide the functionality outline in the Boot Manager PDN [1]. These consist of:

- A set of API calls available at PlanetLab Central
- A package to be run in the boot cd environment on nodes
- An appropriate user interface allowing administrators to create node configuration files

The previous implementation of the software responsible for installing and booting nodes consisted of a set of boot scripts that the boot cd would run, depending on the node's current boot state. The logic behind which script the node was sent to the node existed on the boot server in the form of PHP scripts. However, the intention with the new Boot Manager system is to send the same boot manager back for all nodes, in all boot states, each time the node starts. Then, the boot manager will run and detiremine which operations to perform on the node, based on the current boot state. There is no longer any boot state specific logic at PLC.

2. API Calls

Most of the API calls available as part of the PlanetLab Central API are intended to be run by users, and thus authentication for these calls is done with the user's email address and password. However, the API calls described below will be run by the nodes themselves, so a new authentication mechanism is re-

quired.

2.1. Authentication

As is done with other PLC API calls, the first parameter to all Boot Manager related calls will be an authentication structure, consisting of these named fields:

- method

  The authentication method, only 'hmac' is currently supported

- node_id

  The node id, contained on the configuration file.

- node_ip

  The node's primary IP address. This will be checked with the node_id against PLC records.

- value

  The authentication string, depending on method. For the 'hmac' method, a hash for the call, made from the parameters of the call the key contained on the configuration file.

Authentication is succesful if PLC is able to create the same hash from the values usings its own copy of the node key. If the hash values to not match, then either the keys do not match or the values of the call were modified in transmision and the node cannot be authenticated.

TODO: add specifics on how the hash value is produced from the parameters in the API call.

2.2. PLC API Calls

Full technical documentation of these functions can be found in the PlanetLab API documentation.

- BootUpdateNode(authentication, update_values)

  Update a node record, currenly only allowing the boot state to change.

- BootCheckAuthentication(authentication)

  Simply check to see if the node is recognized by the system and is authorized

- BootGetNodeDetails(authentication)

  Return details about a node, including its state, what networks the PLC database has configured for the node.

- BootNotifyOwners(authentication, message, include_pi, include_tech, include_support)

  Notify someone about an event that happened on the machine, and optionally include the site PIs, technical contacts, and PlanetLab Support

- BootUpdateNodeHardware(authentication, pci_entries)

  Send the set of hardware this node has and update the record at PLC.

3. Core Package

The Boot Manager core package, which is run on the nodes and contacts the Boot API as necessary, is responsible for the following major functional units:

- Installing nodes with alpina, the PlanetLab installer
- Putting a node into a debug state so administrators can track down problems
- Reconfiguring an already installed node to reflect new hardware, or changed network settings
- Booting an already installed node

3.1. Boot States

Each node always has one of four possible boot states.

1. 'new'

   The boot state cooresponds to a new node that has not yet been installed, but record of it does exist. When the boot manager starts, and the node is in this state, the user is prompted to continue with the installation. The intention here is to prevent a non-PlanetLab machine (like a user's desktop machine) from becoming inadvertantly wiped and installed with the PlanetLab node software.

2. 'reinstall'

   In this state, a node will reinstall the node software, erasing anything that might have been on the disk before.

3. 'boot'

   This state cooresponds with nodes that have sucessfully installed, and can be chain booted to the runtime node kernel.

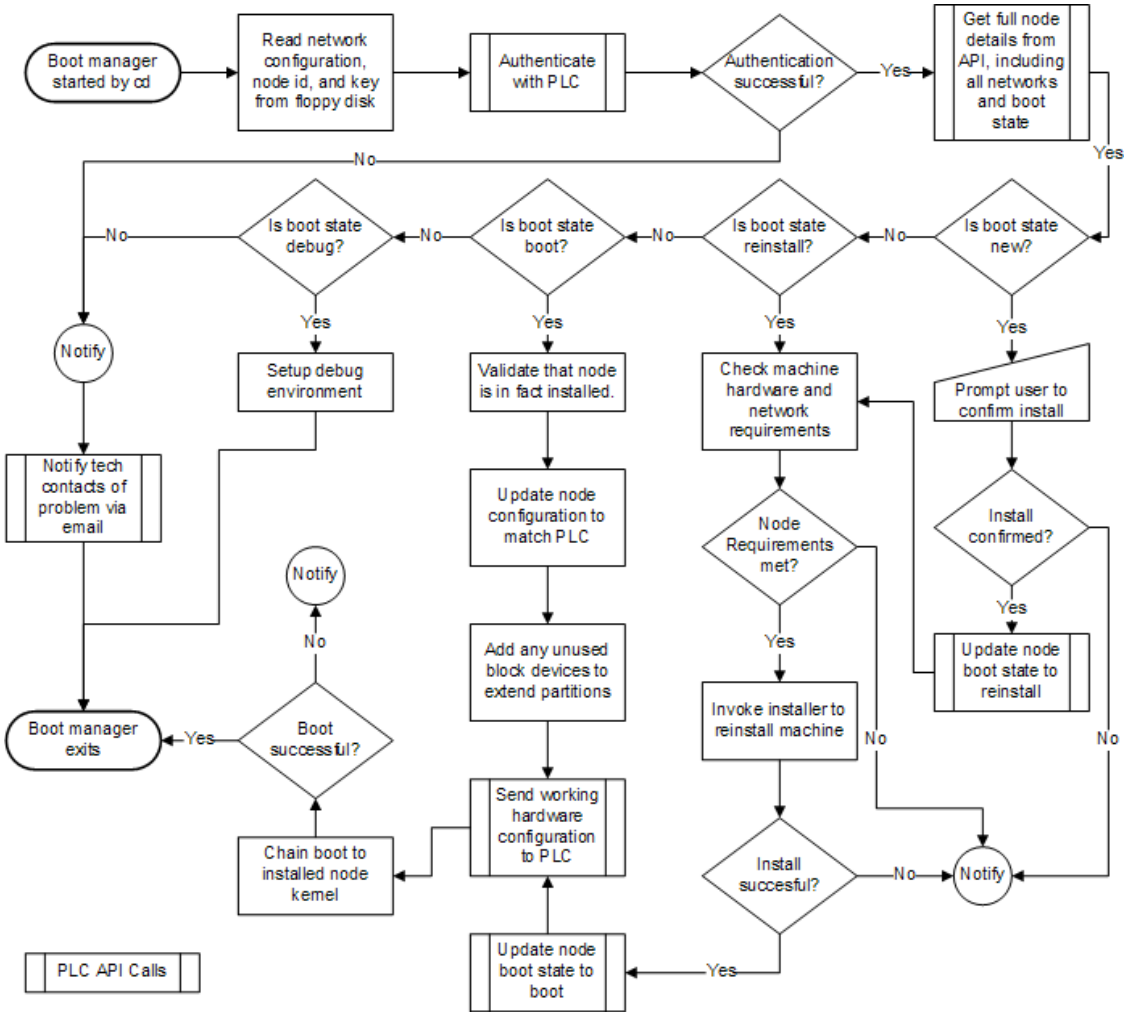
4. 'debug'

   Regardless of whether or not a machine has been installed, this state sets up a node to be debugged by administrators.

3.2. Flow Chart

Below is a high level flow chart of the boot manager, from the time it is executed to when it exits.

Figure 1. Boot Manager Flow Chart



3.3. Boot CD Environment

The boot manager needs to be able to operate under all currently supported boot cds. The new 3.0 cd contains software the current 2.x cds do not contain, including the Logical Volume Manager (LVM) client tools, RPM, and YUM, among other packages. Given this requirement, the boot cd will need to download as necessary the extra support files it needs to run. Depending on the size of these files, they may only be downloaded by specific steps in the flow chart in figure 1, and thus are not mentioned.

3.4. Node Configuration Files

To remain compatible with 2.x boot cds, the format and existing contents of the configuration files for the nodes will not change. There will be, however, the addition of three fields:

1. NET_DEVICE

   If present, use the device with the specified mac address to contact PLC. The network on this device will be setup. If not present, the device represented by 'eth0' will be used.

2. NODE_KEY

The unique, per-node key to be used during authentication and identity verification. This is a fixed length, random value that is only known to the node and PLC.

3. NODE_ID

   The PLC assigned node identifier.

Existing 2.x boot cds will look for the configuration files only on a floppy disk, and the file must be named 'planet.cnf'. The new 3.x boot cds, however, will initially look for a file named 'plnode.txt' on either a floppy disk, or burned onto the cd itself. Alternatively, it will fall back to looking for the original file name, 'planet.cnf'.

An example of a configuration file for a dhcp networked machine:

```
IP_METHOD="dhcp"
HOST_NAME="planetlab-1"
DOMAIN_NAME="cs.princeton.edu"
NET_DEVICE="00:06:5B:EC:33:BB"
NODE_KEY="79efbe871722771675de604a227db8386bc6ef482a4b74"
NODE_ID="121"
```

An example of a configuration file for the same machine, only with a statically assigned network address:

```
IP_METHOD="static"
IP_ADDRESS="128.112.139.71"
IP_GATEWAY="128.112.139.65"
IP_NETMASK="255.255.255.192"
IP_NETADDR="128.112.139.127"
IP_BROADCASTADDR="128.112.139.127"
IP_DNS1="128.112.136.10"
IP_DNS2="128.112.136.12"
HOST_NAME="planetlab-1"
DOMAIN_NAME="cs.princeton.edu"
NET_DEVICE="00:06:5B:EC:33:BB"
NODE_KEY="79efbe871722771675de604a227db8386bc6ef482a4b74"
NODE_ID="121"
```

4. User Interface for Node Management

4.1. Adding Nodes

New nodes are added to the system explicitly by either a PI or a tech contact, either directly through the API calls, or by using the appropriate interfaces on the website. As nodes are added, only their hostname and ip address are required to be entered. When the node is brought online, the records at PLC will be updated with the remaining information.

After a node is added, the user has the option of creating a configuration file for that node. This is done automatically, and the user is prompted to download and save the file. This file contains only the primary network interface information (necessary to contact PLC), and the per-node key.

The default boot state of a new node is 'new', which requires the user to confirm the installation at the node, by typing yes on the console. If this is not desired, as is the case with nodes in a co-location site, or for a large number of nodes being setup at the same time, the administrator can change the node state, after the entry is in the PLC records, from 'new' to 'reinstall'. This will bypass the confirmation screen,

and proceed directly to reinstall the machine (even if it already had a node installation on it).

4.2. Updating Node Network Settings

If the primary node network address must be updated, if the node is moved to a new network for example, then two steps must be performed to successfully complete the move:

1. The node network will need to be updated at PLC, either through the API directly or via the website.
2. Either the floppy file regenerated and put into the machine, or, update the existing floppy to match the new settings.

If the node ip address on the floppy does not Match the record at PLC, then the node will not boot until they do match. The intention here is to prevent a malicious user from taking the floppy disk, altering the network settings, and trying to bring up a new machine with the new settings.

On the other hand, if a non-primary network address needs to be updated, then simply updating the records at PLC will suffice. The boot manager, at next restart, will reconfigure the machine to match the PLC records.

4.3. Removing Nodes

Nodes are removed from the system by:

1. Deleting the record of the node at PLC
2. Shutting down the machine.

Once this is done, even if the machine attempts to come back online, it cannot be authorized with PLC and will not boot.

5. Common Scenarios

Below are common scenarios that the boot manager might encounter that would exist outside of the documented procedures for handling nodes. A full description of how they will be handled follows each.

- A configuration file from previously installed and functioning node is copied or moved to another machine, and the networks settings are updated on it (but the key is left the same).

  Since the authentication for a node consists of matching not only the node id, but the primary node ip, this step will fail, and the node will not allow the boot manager to be run. Instead, the new node must be created at PLC first, and a network configuration file for it must be generated, with its own node key.

- After a node is installed and running, the administrators mistakenly remove the cd and disk.

  The node installer clears all boot records from the disk, so the node will not boot. Typically, the bios will report no operating system.

Bibliography

[1] *The PlanetLab Boot Manager*. January 14, 2005. Aaron Klingaman.