Information about how Open-AudIT processes and stores data

How do we process and store data?

NOTE - Updated for 1.12.8 with new fields and logic.

Each system (computer, network device, printer, et al) has an entry in the "system" table. Each system (from the "system" table) has an "id" column. This value is unique - it's an auto incrementing id. A system is determined to be unique by a the table below.

A system is audited and the result submitted to the server. The first table processed is the "system" table. The "id" is determined and passed (along with the other details) to each other section (table). Every table has two timestamp columns, "first_seen" and "last_seen". The "first_seen" value is populated whenever an insert occurs - hence this value reflects the first time an item was reported in the audit script. The "last_seen" value is inserted when an item is first seen, or updated when an item is seen in subsequent audit script(s). There is an "audit_log" table that contains details of each time an audit is submitted (including timestamp). Each sub-table also contains a 'current' column which is an enum with possible values of 'y' and 'n'.

So, for an example - "hard_drive".

- The system.id is retrieved, along with the timestamp of the previous audit submission and the "status" column.
- For each entry in the hard_drive audit result, the database is queried.
- It checks for hard drive model, serial, index and size.
- These values vary according to the item being processed - see the PHP page at /code_igniter/application/models/m_devices_components.php
- If it gets a match on the above values, combined with component.status = 'y' and the system.id and a system.status of "production", then an existing entry exists for this piece of equipment.
- In the case of hard drives, it simply updates the current flag to 'y' to reflect the component is still current.
- If it does not get a match, it does an insert of the relevant details.

So, we can determine if something is currently installed - the current column is 'y'.

We can determine when something was detected - the "first_seen".

We can determine if something was installed after the initial audit - first seen will be different.

We can determine if something is not currently installed, but previously was - current = 'n'.

We can determine the last time we detected an item - last_seen.

At any given point, we can determine what was on a system - by using the audit_log table and selecting the relevant components based on first_seen and last_seen.

So, that's how we determine what's on or has been on a system.

How do we determine device uniqueness?

When we receive data about a device we check the following columns for matches. If we get a match and the existing entry has a status of 'production', we update this device.

The code for this currently resides in code_igniter/application/models/m_system.php.

Devices are considered the same if they have the following attributes in common: UUID & hostname, dbus_identifier*, FQDN, serial & device type, MAC address and config item**, ip address and config item**, hostname and config item**.

* In 1.12.8 we use the dbus_uuid in Linux to determine uniqueness. This is being reverted in 1.12.8.1 because ESX does not recreate this identifier upon cloning a mchine, hence possibly causing false positive matching.

** In the configuration of Open-AudIT you can select discovery_hostname_match (and mac, ip) to enable this matching.

What do we use for a name?

Where possible, the first option will be chosen and where possible on subsequent audits, will be changed to the first option: hostname, dns_hostname, sysName.