



# Truth

On machine epistemology and the importance  
of dependencies

Leon Towns-von Stauber  
Cascadia IT, March 2013  
<http://www.occam.com/sa/>



## Contents

Questions	4
The Repository of Truth	6
Dependencies	18
Automation	22
Review	28



## Legal Notices

- This presentation Copyright © 2013 Leon Towns-von Stauber. All rights reserved.
- All trademarks are the property of their respective owners.



## Introduction

- If you're involved in managing more than about 40 systems (physical and/or virtual), and you had access to a computer on your network, could you quickly:
  - Answer *precisely* how many systems you manage?
  - For each of those systems, provide (for example):
    - OS version
    - Hardware model
    - Amount of installed/allocated RAM
    - Physical location, to rack and U
    - Serial number
  - Determine the number of active login sessions on each system?
  - Determine whether configuration management is operating correctly on each system?



## Introduction

- I've rarely worked in shops where anyone could claim to answer such questions with 100% accuracy
- In the analogy of IT to manufacturing (see: Lean IT, theory of constraints, etc.), our systems are like the machines on the factory floor
  - Do you think that manufacturing companies don't have reliable knowledge about every significant piece of equipment in a factory: where it is, what it's doing, whether it's working as it should?



The Repository of Truth



## The Repository of Truth

- You need a single *repository* of truth about machines
- Sources of truth are multiple
  - Machine (RAM, CPU, MAC address, serial number, etc.)
  - Software (OS and version, installed software, configurations, etc.)
  - Physical environment (location, asset tag, network/power connections, temperature, etc.)
  - Organization (responsible persons, role of system, etc.)
  - The repository may (should) also be the source of some truths
- There should be exactly one source of truth for any datum
  - All other instances of that datum should be determined (programmatically if possible) from the designated source of truth
  - This is an ideal, which may not be completely possible in practice

7



## The Repository of Truth

- Most organizations have multiple independent repositories of data about machines, few of which are authoritative
  - For example, my current shop has over 20 that I've counted, including:
    - RackTables (2 different instances)
    - The hosts themselves
    - Virtual machine management systems
    - DNS zone files, Active Directory
    - Network device configurations (switches, load balancers, firewalls)
    - Software configurations (CFEngine, Nagios, Cacti, backup software)
    - Lists used by `dsh`, database account management tools, cache flush systems, and other automation

8



## The Repository of Truth

- Implementation

- The repository is most likely a database, with a web front end
- Features it should have (incomplete list)
  - CRUD functionality (Create/Read/Update/Delete)
  - Authentication
  - Authorization (differentiated access to functions and content)
  - Accounting: what was done and who did it (not necessary, but very useful)
- A plain database can do all of these, but SQL isn't the best user interface

9



## The Repository of Truth

- A repository of truth may be referred to as, or be part of, a:
  - Machine database
  - Configuration management database (CMDB) in ITIL-speak
  - Data center infrastructure management (DCIM) tool
- Examples
  - RackTables
  - dcTrack® (from Raritan)
  - DCATS (in-house application at previous shop)

10



# The Repository of Truth

Device #93 (Server) : www1.tuk.intelius.com



links: [N](#) [O](#) [N](#) [Z](#)

Device Details	
Details	Value
Device Name	www1.tuk.intelius.com (Online)
Location	Tukwila - Sabey
Class	server
Model	HP DL360 G5
Rack-U	2-3:12
S/N	
Asset Tag	
OS	
Group(s)	
Purpose	web
Pool	trans
PXE Boot	-no-
Root SSH From MCP1	YES
Device Comments	
Server Comments	
Children	None

Network Connections			
Hide Details			
Port Name	Patch Panel	Port	Rack:U
eth0	----> 2-8 PP7	12	
hostname www1.tuk.intelius.com			
IP			
MAC			
VLAN 452 : TUK-Servers			
DHCP YES			
ilo	----> 2-8 PP7	36	
hostname www1-ilo.tuk.intelius.com			
IP			
MAC			
VLAN 451 : TUK-ILO			
DHCP YES			

PDU Connections		
Name	PDU Port	
(This Device) ----> pdu2-3-2.tuk.intelius.com	4	
(This Device) ----> pdu2-3-2.tuk.intelius.com	12	

Change Requests		
Title	Change Date	Status
		Succeeded
		Succeeded
		Succeeded
		Succeeded
		Succeeded
		Succeeded

DCATS Device Details



# The Repository of Truth

Home Nagios Alerts Network Diagrams Recent Changes

Circuits Models Racks Rack Conflicts Sites VLANs PXE Servers Server Groups Server Properties

DCatsLive Home View Device

Edit Server Device #93 : www1.tuk.intelius.com

View device Edit Device Move Device History

-Device setting-

Device name (Use hostname as fqdn when applicable) -- this ought to be an identical match to the 'Hostname + Domain' on one of the NICs listed below:  
www1.tuk.intelius.com

Model: HP DL360 G5

Serial:

Asset Tag:

Rack: 2-3


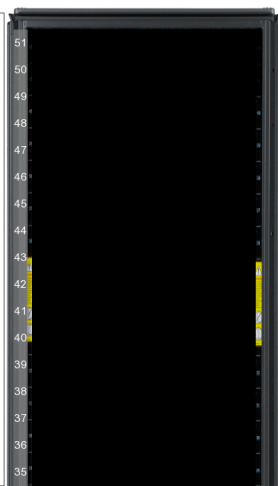
u: 12

In back of rack?  
If this device shares the same 'u' as another device and is located in the back of the rack, check this to prevent conflict warnings.

Status:  Online  Temporarily offline  Decommissioned

Device Comments:

Freeform comments are permitted here.

DCATS Device Editing



# The Repository of Truth

Server Info

OS:   
 Red Hat Enterprise Linux 4.4

Purpose:   
 web

This field is deprecated(will disappear) and is being replaced by a set of Properties affiliated with the Group below

Group(s):   
 addr   
 trans

Group(s) should correspond to the new cplengine classes

Pool:   
 trans

This field is deprecated(will disappear) and is being replaced by the Group above

Needs PXE Boot

PXE Server:   
 --none--

You can either specify one PXE Server for the host, or separate ones below per interface(which will be preferred if set)

root SSH from mcp1 works

Server Comments:   
 This field is to become free of restrictions - the Groups Properties will replace the tags that were recorded here

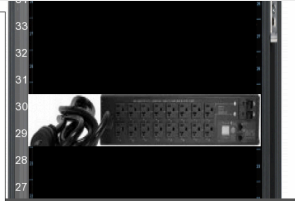
---

Network Interfaces

ADD MORE NICs

Any interfaces with the name left empty will be deleted.

Name*	Interface*	IP Address*	MAC*	Patch Panel	Port	VLAN	DHCP	PXE Boot	PXE Server
eth0	www1	uk.intelias.com		2-8 PP7	12	none	<input checked="" type="checkbox"/>	<input type="checkbox"/>	none
ilo	www1-ilo	uk.intelias.com		2-8 PP7	36	none	<input checked="" type="checkbox"/>	<input type="checkbox"/>	none
		uk.intelias.com		none		none	<input type="checkbox"/>	<input type="checkbox"/>	none



## DCATS Device Editing



# The Repository of Truth

View Device DcatsLive  
Home - View Device

View device Edit Device Move Device History

	devices rack_id	devices u	interfaces cxn_id	interfaces cxn_id	interfaces cxn_port	interfaces cxn_port	interfaces mac_addr	pduports pdudev_id	pduports pdudev_id	pduports port	pduports port
now	11	12	397	397	12	36		43	43	4	12
2010-06-24 11:14:50								41		2	
2010-04-30 16:11:10	9	16	7	7	16	40		39	39	16	8
2010-03-01 11:49:39											

This table shows what the values were prior to a change that "User" made on "Date"

Date	User	Table	Key Name	Key Value	Previous Values	Revert
2010-06-24 11:14:50		interfaces	id	104	• mac_addr => 00:1C:C4:79:FD:FE	Revert
2010-06-24 11:14:50		pduports	id	69	• pdudev_id => 41 • port => 2	Revert
2010-04-30 16:11:10		devices	id	93	• rack_id => 9 • u => 16	Revert
2010-04-30 16:11:10		interfaces	id	104	• cxn_id => 7 • cxn_port => 16	Revert
2010-04-30 16:11:10		interfaces	id	105	• cxn_id => 7 • cxn_port => 40	Revert
2010-04-30 16:11:10		pduports	id	69	• pdudev_id => 39 • port => 16	Revert
2010-04-30 16:11:10		pduports	id	70		Revert

## DCATS Device History



# The Repository of Truth

DCATS admin DcatsLive  
Home - DCATS admin

Manage circuits | Manage models | Manage rack models | Manage racks | Manage sites | Manage pxe\_servers | Manage Groups | Manage Properties

View models | Add models

Class	Model & Make	U Size	Amps Idle	Amps Full	Watts Idle	Watts Full	BTUs Out	On	Off	Decomm
Usage blade chassis	Cisco UCS 5108	6	-	15.50	-	2500.00	999.99	-	-	-
Usage cable management	Extended Cables 1U	1	-	-	-	-	-	-	-	-
Usage cable management	Extended Cables 2U	2	-	-	-	-	-	-	-	-
Usage cable management	Ortronics 1RU	1	-	-	-	-	-	24	-	-
Usage cable management	Ortronics 2RU	2	-	-	-	-	-	28	-	-
Usage console server	APC AP5015 KVM Console	1	-	-	-	-	-	-	1	-
Usage console server	Aten MasterviewPlus KVM	1	-	-	-	-	-	-	4	-
Usage console server	Avocent Cyclades CS4016	1	-	-	-	-	-	-	-	-
Usage console server	Cyclades CS4016	1	-	-	-	-	-	-	2	-
Usage console server	Raritan Dominion KX 2 IP KVM	1	-	0.60	-	72.00	-	-	1	-
Usage console server	Raritan Dominion SX-8	1	-	-	-	6.00	-	-	-	1
Usage console server	Raritan Dominion SXA-16	1	-	-	-	8.00	-	-	2	-
Usage console server	Raritan Dominion SXA-32-DLM	1	-	0.60	-	9.38	-	-	1	-
Usage console server	Raritan Dominion SXA-48	1	-	-	-	12.50	-	-	1	-
Usage emu	APC AP9319	1	-	-	-	-	-	-	1	-
Usage emu	APC NetBotz 355	0	-	-	-	-	-	-	1	-
Usage emu	NTI Enviromux	1	-	-	-	-	-	-	1	-

## DCATS Administration



# The Repository of Truth

2 results for search 'www1.tuk.intelius.com'

Displaying previous search. Clear results -

SEARCH | ADVANCED

Keyword: www1.tuk.intelius.com SEARCH  
 Searches the class, hostname, devicename, IP, IP/CIDR, MAC, or pool. Leave blank to show all devices.

Status: Operational (Online & Temp Offline) | Location: All Locations |  Group by device

Search Criteria  
 If you want to add a column just for viewing, add the Field and leave the Value blank.

Field	Operator	Value	Method	Sort	Show column?
Device	contains		AND		<input checked="" type="checkbox"/>
Device	contains		AND		<input checked="" type="checkbox"/>

Saved searches  
 Private? view saved search

Hostname (Device name)	IP addr/MAC	Rack	Site	Pool
<a href="#">view server</a> www1.tuk.intelius.com		2-3-12	Tukwila - Sabey	trans
www1-ilo.tuk.intelius.com				

## DCATS Advanced Search





## The Repository of Truth

- Quote from 2009 Forrester Research report, “Knocking the NOC: Enter the New Operations Center” (p. 4)

In all cases, however, it is best to make sure that all tools, regardless of their domains of control, cooperate with other related tools. A good case in point is the unifying role of the configuration management database (CMDB) or configuration management systems (CMS). In the design phases of a new or modifying service, SMEs and design tools must consult a common CMS from the very beginning. This presents a snapshot of reality on which one can build. This snapshot is modified in the tools to reflect the newly proposed situation. The new design then passes through the necessary change and release processes to move into production. When the changes are actually implemented, the CMS must be synchronized and reconciled in real time to maintain an accurate perspective of “the truth” within the CMS. If “the truth” is not true, then success is false.

17



## Dependencies



## Dependencies

- Epistemology
  - How do we know what we think we know?
  - How do we know if it's true?
- How do you ensure that the data in the truth repository is complete and accurate?
  - I can guarantee that if you're depending on human beings keeping this information up-to-date, you'll be sorely disappointed
- The answer: Dependencies
  - Dependencies are good, and you should actively create them

19



## Dependencies

- Now I'm going to qualify that statement to make it less crazy
- "Hard" dependencies, that have real-time impact on functioning services, should be avoided
  - Loss of availability or correctness on the part of the truth repository shouldn't break a working service
- "Soft" dependencies are encouraged
  - Inaccurate or incomplete data in the truth repository *should* break processes
    - System builds, DNS additions, code releases, etc.

20



## Dependencies

- Dependencies are good as long as you can detect and remediate root causes of problems before they get serious
- However, the dependencies must cause some pain; that's their purpose
  - The pain alerts you to untruth in the repository, thus enforcing accuracy
    - How do we know that what we know is true? This is the key
  - Without pain, you quickly accumulate untruth in the repository, and then lose trust in it
    - Without pain, people skip steps, forget or delay updates, etc.
- Dependencies lead to management by *fact* rather than hope

21



## Automation

- Use automation to help enforce accuracy and completeness
  - Build automation over time
- Three categories of automation
  - Dependencies
  - Data gatherers
  - Consistency checks

- Dependencies
  - Automatically build configurations, or implement processes based on data in the truth repository
  - Examples:
    - Scripts to generate DNS zone files and DHCP configs
      - Make the repository the source of truth for IP addresses (host IP addresses, at least)
      - Include checks for duplicate IP assignments, malformed addresses, properly qualified domain names, valid subnets, etc.
        - » You may want additional checks for changes to existing records beyond those for creation of new ones
      - If IP address data in the repository is incorrect, the script will fail, or the host won't get on the network, or something else will break
        - » Make failures obvious
    - Monitoring applications
      - Generate host lists and groups from the truth repository for Nagios, Ganglia, etc.

- Dependencies (cont'd.)

- Examples:

- Host management tools

- Data used by tools should come from the repository

- `rshell` is a tool used to manage UNIX/Linux hosts in parallel

- »Build or modify an `rshell` extension to use the repository as its host data source

- If the data is incorrect, it'll probably be noticed in a frequently-used tool

- Code release tools

- Determine hosts that need code by checking the repository for roles and status

- Change management

- Tie CM systems into the truth repository

- Change-detection systems can reference lists of machines from the truth repository associated with CM requests

- CM requests can feed comments into the truth repository

25

- Data gatherers

- If the repository isn't the source of truth for particular data, gather that data from the source automatically whenever you can

- Examples:

- Scripts to gather hardware models, installed RAM and CPU, serial numbers, OS revisions, etc., and update the truth repository

- Optionally, alert you to updates of incorrect information in the repository

26

- Consistency checks

- If the repository isn't the source of truth, and the truth can't be gathered automatically, at least check whether the data in the repository is plausible

- Examples:

- Check racked device locations (rack and U)

- Check for valid formatting and values

- Two devices can't be in the same location

- » Include consideration of device height

- Check network switch or patch panel ports, or PDU sockets

- Check for valid formatting and values

- Two devices can't share the same port

- Single-port hosts can't have multiple network/power connections



Review

- You need a single *repository* of truth about machines
  - Sources of truth are multiple, but there should be a single source for each datum
  - Your truth repository may be called a machine database, a CMDB, or be part of a DCIM tool
- Dependencies help ensure truth
  - Actively create them, and make them painful (but not fatal!)
  - The goal is to manage by fact rather than hope
- Use automation to ensure accuracy and completeness
  - Consider automation of dependencies, data gathering, and consistency checks



## Truth

On machine epistemology and the importance  
of dependencies