An Introduction to DMTF Cloud Auditing using the CADF Event Model and Taxonomies

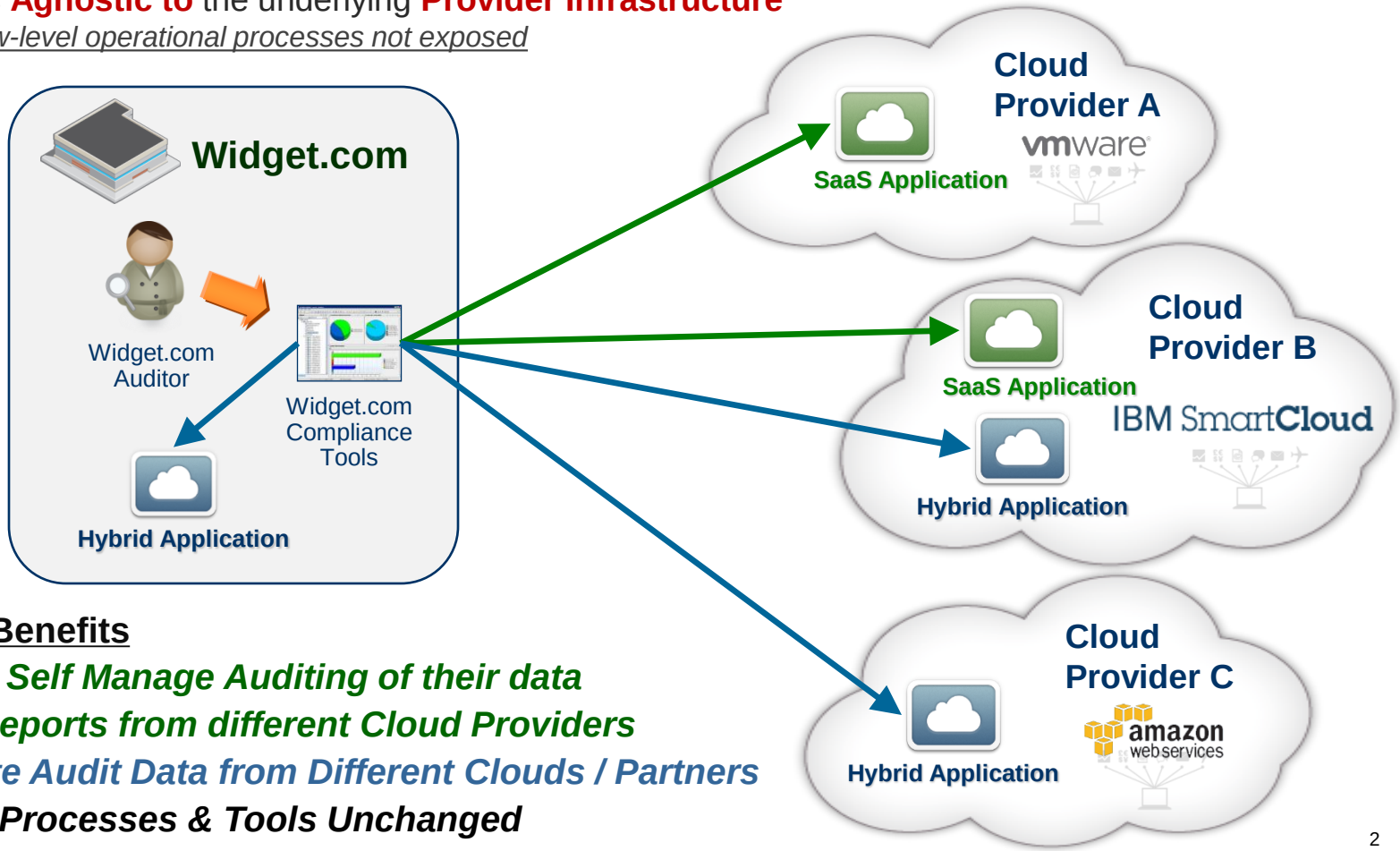
Matt Rutkowski,
IBM SWG Emerging Standards and Strategy
Co-chair DMTF CADF Working Group

Cloud Auditing: Customer Importance: Self-Managing Auditing Data on Clouds

Customers will not trust clouds to host their workloads / data without the ability to self-audit

Auditing using a standard such as CADF has many benefits:

- *Create and request **Customized Views** for Audit & Compliance Data*
  - *Track regional, industry and corporate policy compliance using standardized APIS / Reports*
- Key event data is **Normalized** and **Categorized** to support auditing of Hybrid Cloud Applications
  - *CADF assures consistent mappings across cloud components and cloud providers*
- Format is **Agnostic to** the underlying **Provider Infrastructure**
  - *Low-level operational processes not exposed*



Customer Benefits

- ✓ ***Ability to Self Manage Auditing of their data***
- ✓ ***Similar Reports from different Cloud Providers***
- ✓ ***Aggregate Audit Data from Different Clouds / Partners***
- ✓ ***Auditing Processes & Tools Unchanged***

Overview: Cloud Auditing Data Federation (CADF) Standard

Develop a normative audit event data model and compatible set of interfaces for federating events, logs and reports between cloud providers and customers

- Use **extensible classification systems** for real or virtual cloud-based IT resources and their interactions to **support cross-cloud / hybrid cloud analysis** allowing universal query of event data

Key Participants: IBM, NetIQ, Microsoft, CA, Huawei, VMware, Fujitsu, Citrix, EMC

Work Products

- ✓ **Published**: "*Use Cases Whitepaper v1.0*", Public Draft, July, 2012
- ✓ **Released**: "*Data Model and Interface Specification, v1.0b*", Public Draft 2, June 2013
  - ✓ *Added RESTful **Query Interface** finalized (Customer Self Management),*
  - ✓ Added support for "**Control Event Type**" to support linkage to customer policies
  - ✓ Added support for Event "**Tags**" to support customer & domain-specific views on data
  - ✓ Added support for "**Complex Targets**" (multiple heterogenous/homogenous)

Current Work

- **Target October 2013**: v1.0c Specification Draft for review.
  - Last scheduled v1.0 draft before "final" v1.0 status is sought from DMTF companies.
  - *Incorporates updates based upon integration work with OpenStack*

"As audit and compliance concerns grow for cloud-based applications or cloud-inclusive workflows, the importance of interoperability becomes more evident."

Key Features - of CADF Standard over Existing / Legacy Formats

Standardized Classification of Event Data using Extensible Taxonomies

- **Resources** – By the role played in the event (Initiator, Target, Observer, etc.)
- **Actions** – No confusion between events with similar sounding activities.
- **Outcomes** – Well-defined and unambiguous results for all activitie types.

Federation of Events Data from Hybrid Deployments

- **Tagging –** Customers can create **Orthogonal Views** via ***Path-based Tagging***
  - Ability to identify and track multiple **Domains of Interest** from same event data
    - *e.g. PCI, SoX, Local Corporate Policy, Regional Policy, Departmental Policy, etc.*
- **Resource Identity** – Resources **uniquely tracked via UUIDs**, not dependent on relative IP addresses.
- **Timezone-Aware** - Specifies how to create events from **different Timezones** and track any record changes
- **Geolocation Aware** – Track geolocation of resources using International Standards
  - Proving enforcement of **Regional Policies** for data and application hosting
- **Event Merge –** Instantly merge CADF event data from any hybrid deployment into consistent end-to-end logs

Self-Service, REST-based APIs for Log Management and Audit

✓ **Standardized Query using Standardized Path-based Expressions**

- Construct reports from any attribute within CADF data schema
- More than tracking Source, Target (IP Addresses) and timestamps (geolocs, resource types, policies, etc.)

✓ **Beyond Access or Activity Reports**

- Defines **Metric** (e.g., usage and performance data) and **Control** events that map to external policies (security, operational, business, or other).
- *Metrics and Measurement data are **compliant with NIST metric standards** (in development)*

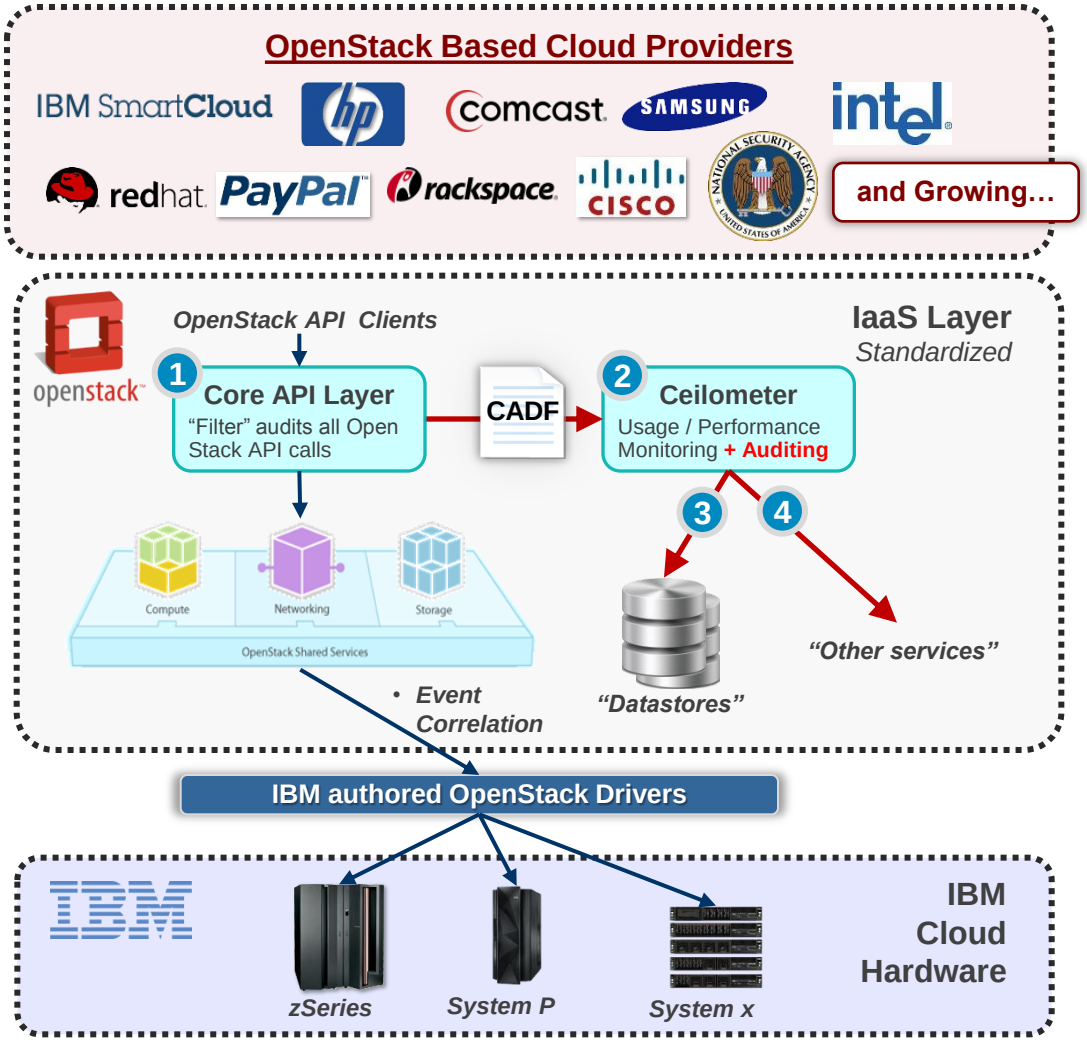
OpenStack Integration – API Auditing using the CADF Model, JSON Format

Project Ceilometer

- OpenStack's Aggregator of Performance and Usage Metrics
- **Delivered API Audit "Plug-in"**:
  - in "Havana" Release 4Q 2013
  - Fully tested for **Nova** (compute)
  - Works for any component, including: *Network (Neutron), Storage (Cinder), etc.*

OpenStack Integration (Completed)

1. **"Audit filter" in Core Components APIs generate audit data in CADF format**
2. **Ceilometer receives data from agents and filters listening to core components.**
3. **Ceilometer dispatches event data using CADF format to one or more "datastores".**
4. **Dispatchers can be added to send CADF audit data to other services for analysis.**



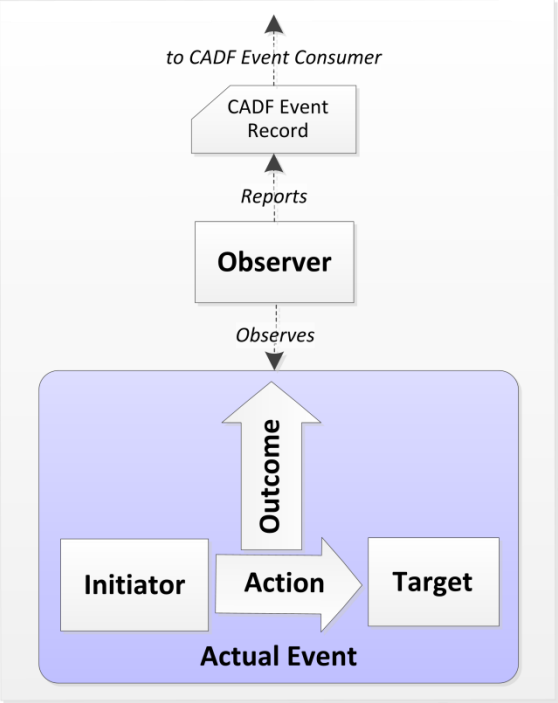
OpenStack is IBM's Strategic IaaS Platform for SmartCloud

OpenStack Core API Specs: http://docs.openstack.org/api/api-specs.html
Project Ceilometer: https://wiki.openstack.org/wiki/Ceilometer

CADF Event Model Components

Event model is common to all CADF Event Types (i.e. “activity”, “monitor” and “control”)

Conceptual Model



Model Components

CADF Event Model	
Component	**Definition**
OBSERVER	The RESOURCE that generates the CADF Event Record based on its observation (directly or indirectly) of the Actual Event.
INITIATOR	The RESOURCE that initiated, originated, or instigated the event's ACTION, according to the OBSERVER.
ACTION	The operation or activity the INITIATOR has performed, attempted to perform or has pending against the event's TARGET, according to the OBSERVER
TARGET	The RESOURCE against which the ACTION of a CADF Event Record was performed, was attempted, or is pending, according to the OBSERVER.
OUTCOME	The result or status of the ACTION against the TARGET, according to the OBSERVER.

CADF specification and Event Model are extensible

- New “event types” can be defined for other domains that extend this model
- Profiles of the base spec. can be published to describe proper use in other domains

"CSI for Clouds" - *How CADF standard expresses the 7 "W"s of audit and compliance*

"W" Component	CADF Mandatory Component	CADF Optional Components (where applicable)	Comment
What	event.action event.outcome	event.reason (e.g. severity, reason code, policy id)	"what" activity occurred; "what" was the result
When	event.eventTime	reporter.timestamp (for each reporter that modifies the record)	"when" did it happen
Who	initiator.id initiator.type	Initiator.id (id, name): (basic) initiator.credential (token): (detailed) initiator.credential.assertions (precise)	"who" (person or service) initiated the action
OnWhat	target.id target.type		"onWhat" resource did the activity target
Where	reporter.id reporter.type		"where" did the activity get (observed) logged
FromWhere		initiator.addresses (basic) initiator.host (agents, platforms, …) (detailed) Initiator.geolocation (precise)	
ToWhere		target.addresses (basic) target.host (agents, platforms, …) (detailed) target.geolocation (precise)	

CADF provides methods to "Extend" the event data (format) to carry domain-specific information

CADF Extension Type	CADF Optional Component	Purpose
Attachment	event.attachments	For adding domain specific structured or unstructured data. If structured, the type can be supplied and referenced by the CADF Query API.
Tags	event.tags	For adding domain-specific identifiers and classifications that enable domain specific identification and can be used with the CADF Query API to construct custom reports.

CADF Event Model – The "Reporter Chain"

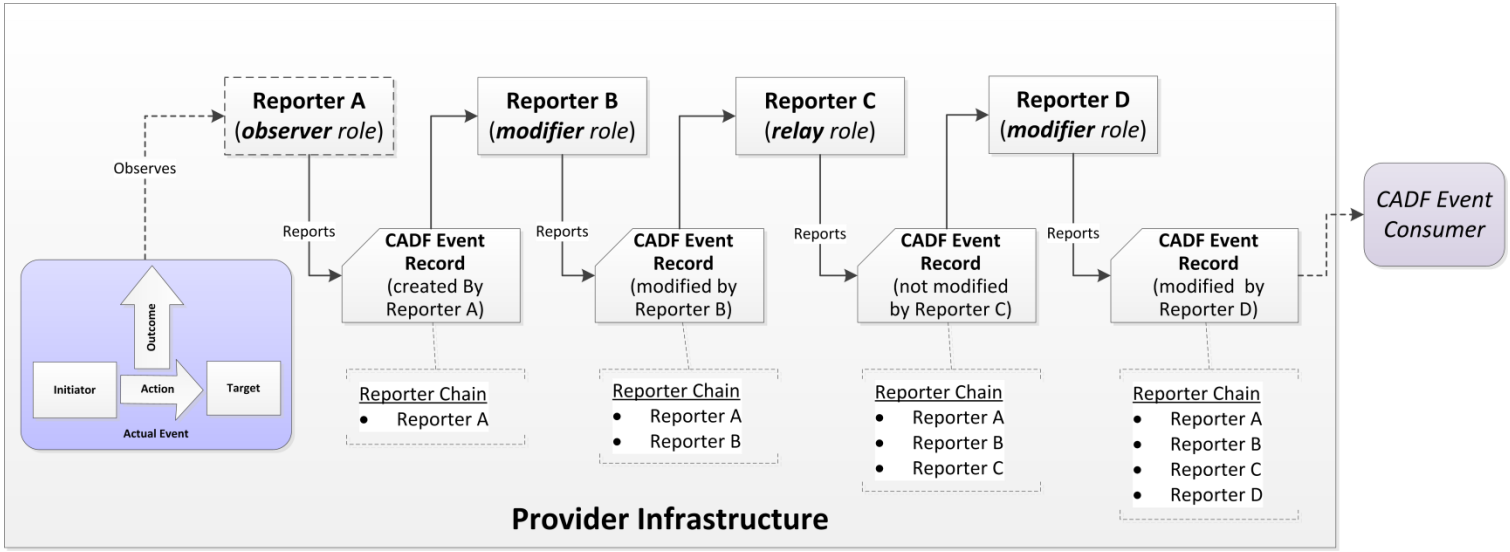
Cloud provider architectures are generally layered in a way such that many Actual Events may occur at the lower layers, which are close to the infrastructure components and services. Additionally, operational systems and processes may span many layers of the architecture, each with critical information that would be valuable to associate with audit events.

The CADF Event Model recognizes that many components may assist in constructing and surfacing the CADF Event Record before it is presented to the end consumer. These components can each be viewed as CADF Event Record REPORTERS each serving a specified role in raising the CADF Event Record as part of a sequential chain of REPORTER components.

Valid Roles for a REPORTER resource

Reporter Role	CADF Definition
observer	A REPORTER that fulfills the role of OBSERVER.
modifier	A REPORTER that adds, modifies or augments information in the CADF Event Record for the purposes of normalization or federation.
relay	A REPORTER that passes the CADF Event Record to another REPORTER or to end record consumer without modifying the information in the CADF Event Record.

Conceptual "chain" of reporters that may "handle" the event record before reaching the consumer rce



Additional Event Model Components: *Measurements and Metrics*

Measurements are an optional component of the CADF Event Type, but are essential (required) for any CADF Event Record that is classified as a "**monitor**" type event.

Event Component	CADF Definition
MEASUREMENT	An entity that contains statistical or measurement information for the TARGET resource(s) that are being monitored. The measurement should be based upon a defined metric (a method of measurement).

Metric data type

The Metric data type describes the rules and processes for measuring some activity or resource, resulting in the generation of some values (captured by the Measurement type). A set of metric instances may be associated with an Event Log, and referred to by individual events.

Property	Type	Required	Description
metricId	cadf:Identifier	Yes	The identifier for the metric (allows reuse) Metric data is designed so that it can be described once, for example in the context of a CADF Log, and referenced by the multiple CADF Event (records) the log contains..
unit	xs:string	Yes	The metrics unit (e.g., "msec.", "Hz", "GB", etc.)
name	xs:string	No	A descriptive name for metric (e.g., “Response Time in Milliseconds", "Storage Capacity in Gigabytes", etc.)
annotations	cadf:Map	No	User-defined metric information. The same “key” SHALL NOT be used more than once within a "annotation" property.

Measurement data type

The Measurement type is intended to hold the values generated by the application of a metric in a particular context (e.g., for a resource or during an activity). The CADF Event Record includes a property that is capable of holding measurements represented by this type.

Property	Type	Required	Description
result	xs:any	Yes	The quantitative or qualitative result of a measurement from applying the associated metric. The measure value could be boolean, integer, double, a scalar value, etc.
metric	cadf:Metric	Optional	The property describes the metric used in generating the measurement result (if no “metricId” property is provided)
metricId	cadf:Identifier	Optional	This property identifies a CADF Metric by reference and whose definition exists outside the event record itself (e.g., within the same CADF Log or Report).
calculatedBy	cadf:Resource	No	An optional description of the resource that calculated the measurement

Additional Event Model Components: *Geolocation*

Geolocation information reveals a resource's physical location and can be obtained through various technologies. It is widely used in context-sensitive content delivery, enforcing location-based access restrictions on services, and fraud detection and prevention along with addressing concerns regarding security and privacy especially in countries/regions with compliance legislation and regulation. It is crucial to report geolocation information unambiguously in an audit trail.

Geolocation data type

Property	Type	Required	Description
id	xs:anyURI	No	Optional identifier for a geolocation.
latitude	xs:string	No	Indicates the latitude of a geolocation. Geolocation MAY be provided in a pair of latitude and longitude. Latitude values adhere to the format based on ISO 6709:2008
longitude	xs:string	No	Indicates the longitude of a geolocation. Geolocation MAY be provided in a pair of latitude and longitude. Longitude values adhere to the format based on ISO 6709:2008
elevation	xs:double	No	Indicates the elevation of a geolocation in meters. Elevation at or above the sea level shall be designated using a plus sign (+), or no sign. Elevation below the sea level shall be designated using a minus sign (−).
accuracy	xs:double	No	Indicates the accuracy of a geolocation in meters. Geolocation expresses the resource location to a reasonable degree of accuracy.
city	xs:string	No	Indicates the city of a geolocation.
state	xs:string	No	Indicates the state/province of a geolocation
regionICANN	xs:string	No	Indicates a region (e.g., a country, a sovereign state, a dependent territory or a special area of geographical interest) of a geolocation. The value used to indicate the region SHOULD match the ICANN country code top level domain (ccTLD) naming convention [IANA-ccTLD].
annotations	cadf:Map	No	Indicates user-defined geolocation information (e.g., building name, room number). The same "key" SHALL NOT be used more than once within a "annotation" property.

JSON Example:

```
"geolocation": {
  "latitude": "+37.37",
  "longitude": "-122.04",
  "elevation": "10"
}
```

XML Example:

```
<geolocation
  city="Sunnyvale"
  state="CA"
  regionICANN="us"
  <annotation key="building" value="B2"/>
  <annotation key="room" value="201"/>
</geolocation>
```

CADF Query Interface and Syntax (basics)

Query Interface : implementations only require a “filter” parameter:

```
?$filter=expression
```

Query Syntax: “filter” parameter defines an “XPath like” expression

```
Filter       ::= AndExpr ( 'or' Filter )* ;
AndExpr      ::= Comp ( 'and' AndExpr )*
Comp         ::= Attribute Op Value | Value Op Attribute | '(' Filter ')'
Op           ::= '<' | '<=' | '=' | '>=' | '>' | '!='
Attribute    ::= ? property name ? | PropertyPath
PropertyPath ::= ? property name ? | ? property name ? “[” Index “]” | ? property name ? “/” PropertyPath |
                 ? property name ? “[” Index “]” “/” PropertyPath
Index        ::= '*'  |  IntValue
Value        ::= IntValue | DateValue | StringValue | BoolValue | PathValue

PathValue    ::= “ PValue “ | ‘ PValue ‘
PValue       ::= StrValue | StrValue “/” PValue | StrValue “//” PValue | “//” PValue | “*”

IntValue     ::= /[0-9]+/
DateValue    ::= ? as defined by XML Schema ?
StringValue  ::= "StrValue" | 'StrValue'
StrValue     ::= ? character string without “ nor ‘ ?
BoolValue    ::= 'true' | 'false'
```

Query Syntax: Operators

```
'or', 'and'                        : Boolean value/attribute
'<', '<=', '=', '>=', ">', '!='    : Integer and date value/attribute
'=', '!='                          : String value/attribute
```

Example: “Time query window”

To search for events that occurred on or after 2012-07-22:

```
/events/Event?$filter=eventTime>=”2012-07-22T00:00:00-02:00”
```

CADF Path-Based, Extensible Taxonomies

CADF defines three taxonomies designed to provide the basis for a domain extensible, path-based mechanism to name resources, actions and that appear in audit events in order to enable normative classification and query of events data.

1. **CADF Resource Taxonomy**
   - Normalized classification type "names" for the resource types that participate on an event (e.g. INITIATOR, TARGET, REPORTER)
   - *Enables Resource-based Query* by type of resource

2. **CADF Action Taxonomy**
   - Normalized names used to describe actions or activities performed on resources
   - *Enables Activity-based Query*

3. **CADF Outcome Taxonomy**
   - Normalize the names used to describe outcomes of activities
   - *Enables Outcome-based Query*

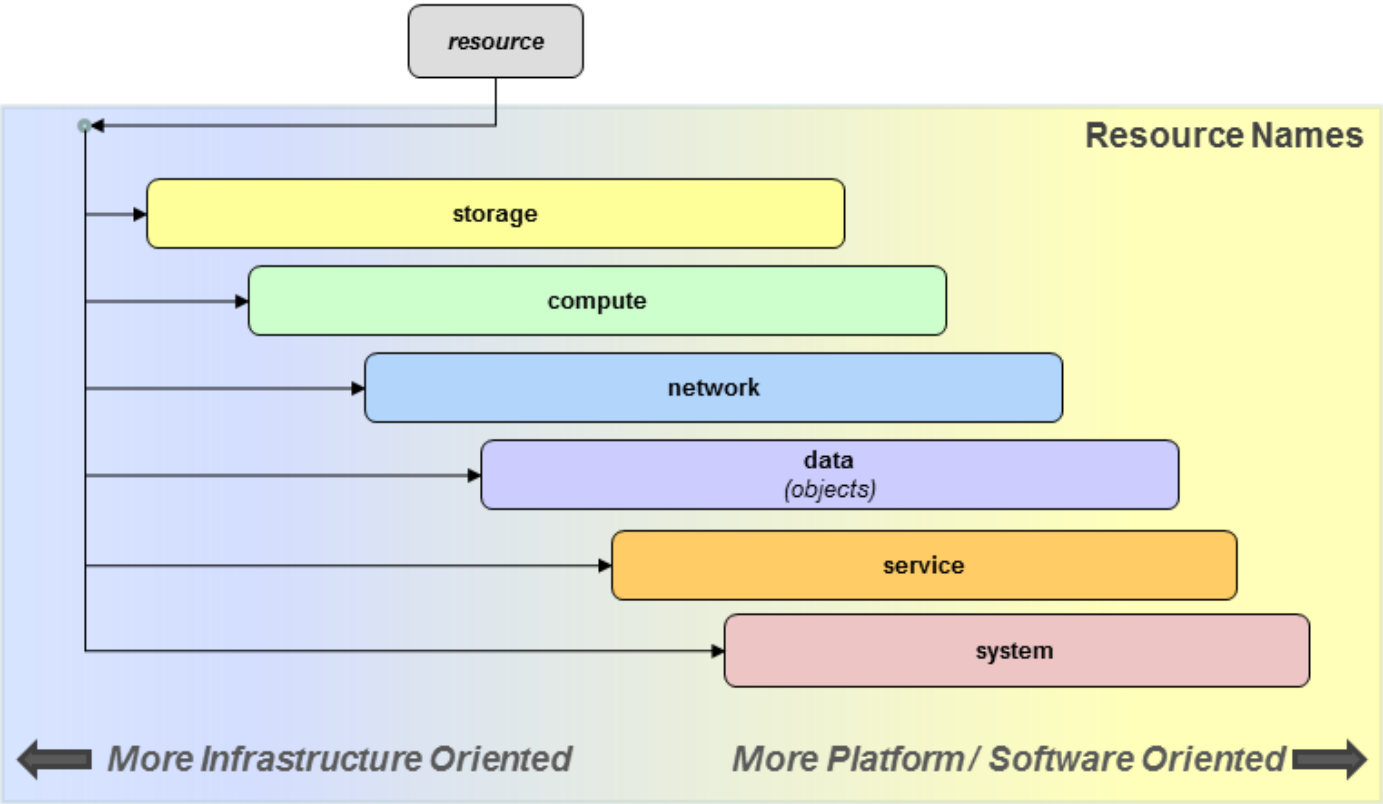
To assure every "name" in a taxonomy is unique when federated, each value is assumed to have the following implied absolute domain namespace:

Taxonomy Name	Taxonomy URI
resource	http://schemas.dmtf.org/cloud/audit/1.0/taxonomy/resource/
action	http://schemas.dmtf.org/cloud/audit/1.0/taxonomy/action/
outcome	http://schemas.dmtf.org/cloud/audit/1.0/taxonomy/outcome/

All taxonomies are *extensible* for any domain-specific compliance framework

CADF Logical Resource Taxonomy - Classification Methodology

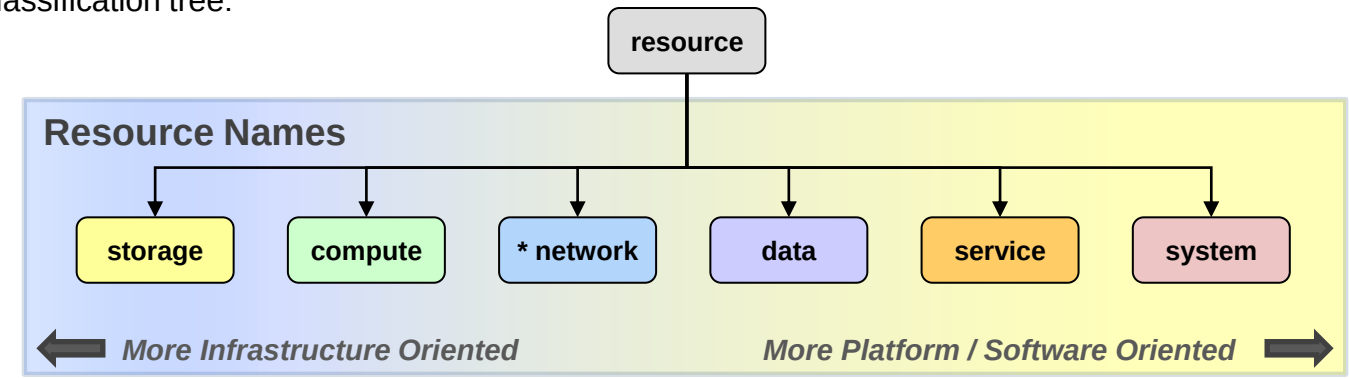
Taxonomy logical names can represent resources that may overlap cloud Infrastructure, platform & software service deployment models depending on provider architecture



The diagram attempts to convey that resources that may be named under these top-level nodes can represent resources some providers may consider more "infrastructure oriented" and offer as via an IaaS service model, whereas other providers may consider more "platform oriented" and offer them via PaaS or SaaS service models..

CADF Logical Resource Taxonomy - Top-Level Classifications Definitions

This diagram shows the top-level resource classifications as child nodes under the "resource" node of the CADF Resource Taxonomy's classification tree:

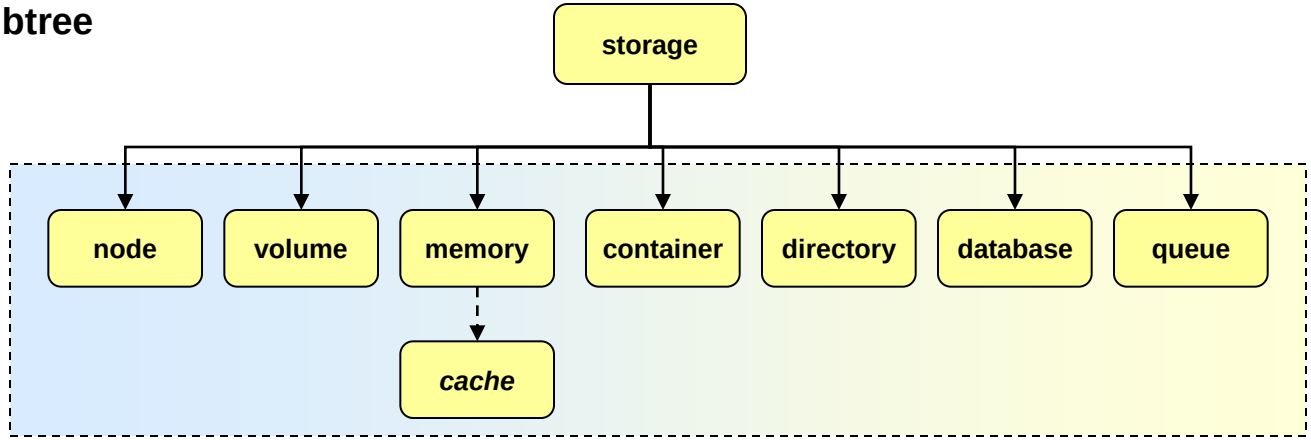


Name	Description
storage	Logical resources that represent storage containers
compute	Logical resources that are used to perform logical operations or calculations on data
network	Logical resources that interconnect computer systems, terminals, and other equipment allowing information to be exchanged.
data	Logical named sets of information (objectified data) that are referenced and managed by services.
service	Logical set of operations, packaged into a single entity, that provides access to and management of cloud resources (for a given domain).
system	Logical resources that are a combination of several other [cloud] resources that operate as a functional whole, this combination being manageable (created, operated, audited, etc.) as a unit i.e. offering some operations that could activate lower-level operations over each of the sub-resources.
unknown	Indicates that the OBSERVER of the event is not, to the best of its ability, able to classify a resource that contributed to the actual event it is reporting on using any other valid resource taxonomy value. Note: This value SHOULD only be used as a last resort, and when using another classification value from the CADF Resource Taxonomy is not possible.

Note: the name value “resource” (tree root) implies the absolute absolute name:

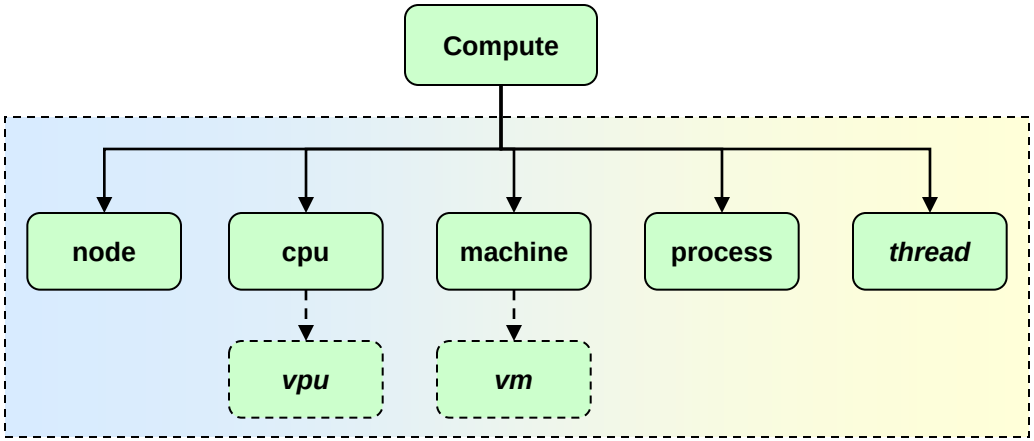
- “http://schemas.dmtf.org/cloud/audit/1.0/taxonomy/resource/”

Storage Subtree



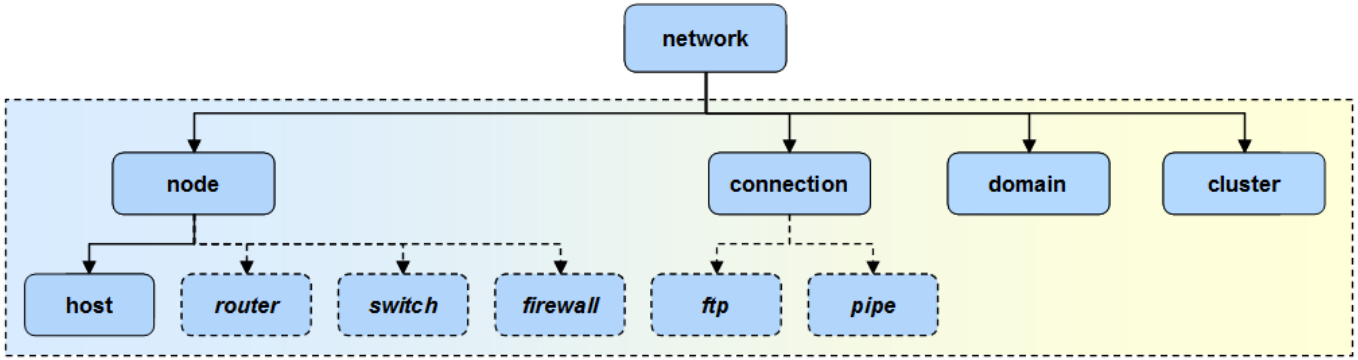
Name	Description
node	Logical resource that contains the necessary processing components to store data.
volume	Logical unit of persistent data storage that is may or may not be physically removable from the computer or storage system.
memory	Logical unit of data storage that is used for dynamically processing data.
container	Logical unit of storage where data objects are deposited and organized for persistent storage.
directory	Logical storage used to organize records about resources (e.g., files, subscribers, etc.) along with their locations and other metadata. Typically, these records are organized in a hierarchical structure.
database	Logical storage used to organize data to a model (schema) that reflects relevant aspects of a specific real-world application.
queue	Logical storage of a list of data awaiting processing.

Compute Subtree



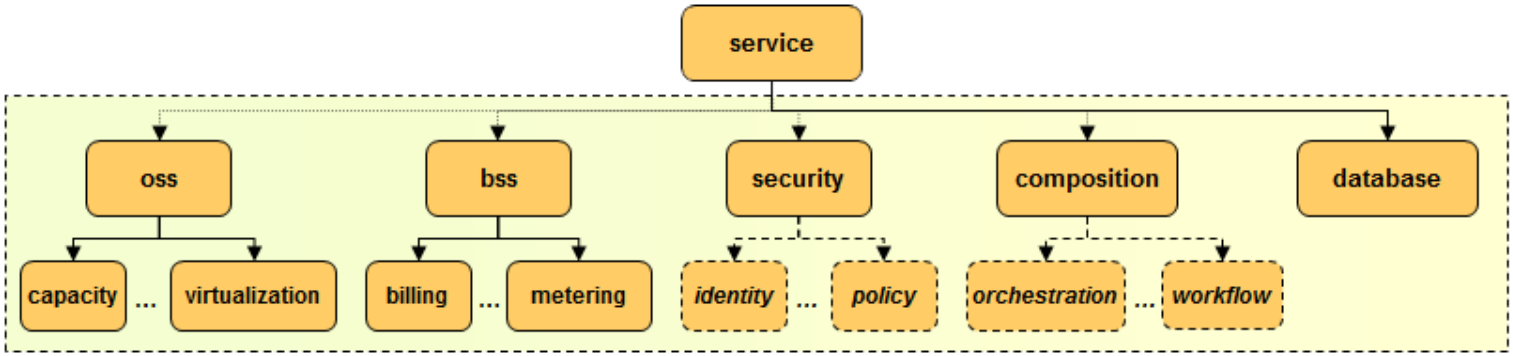
Name	Description
node	Logical resource that contains the necessary processing components to execute a workload.
cpu	Logical resource that represents a unit processing power that can consume a workload.
machine	Logical resource that encapsulates both CPU and Memory.
process	An instance of a granular workload, such as an application or service, that is being executed.
thread	A separable function of a running process that shares its virtual address space and system resources.

Network Subtree



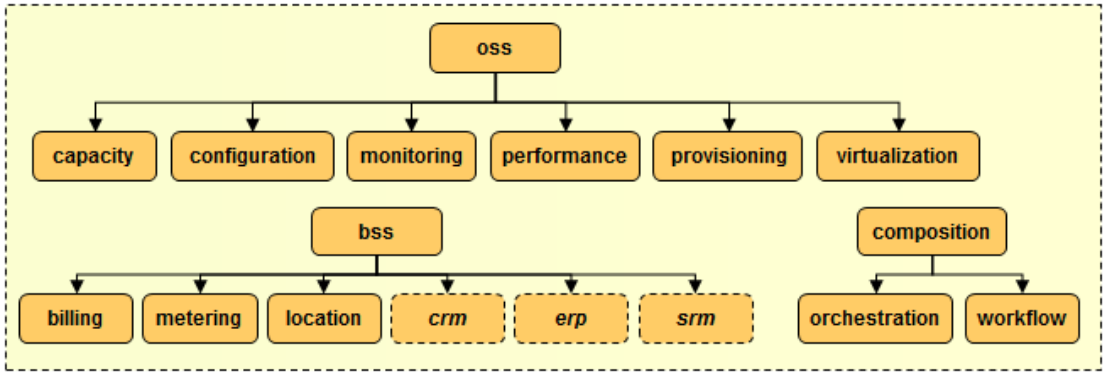
Name	Description
node	A logical resource that can be networked and provide services on data from network connections. A node may export zero or more endpoints (zero implies it is has not been provisioned).
host	A network node that can perform operations or calculations on data. **Note:** Network “nodes” should not attempt to describe details of compute or storage functions; specific compute and storage nodes exist that better suit this purpose).
connection	A single network interaction involving two or more endpoints (sources and destinations).
domain	Represents a logical grouping of networked resources
cluster	Represents a logical combination of tightly coupled, network resources.

Service Subtree



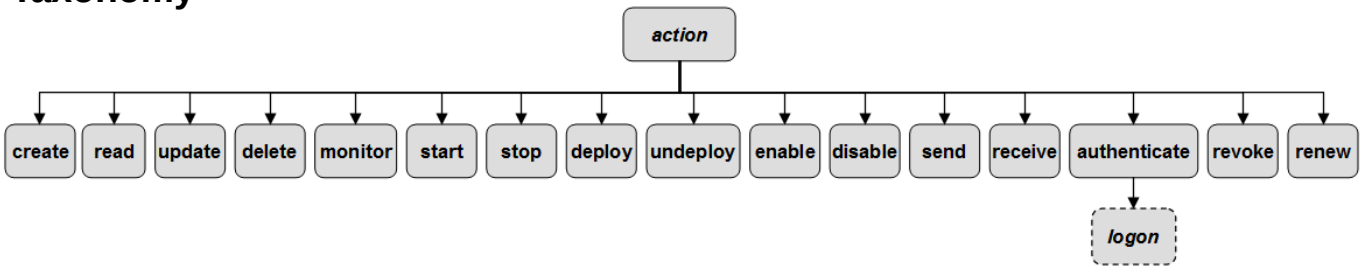
Name	Descriptive Name	Description
oss	**Operational Support Services (OSS)**	The logical classification grouping for services that are identified to support operations including communication, control, analysis, etc.
bss	**Business Support Services (BSS)**	The logical classification grouping for services that are identified to support business activities.
security	**Security Services** *(or Sec-as-a-Service)*	The logical classification grouping for security services including Identity Mgmt., Policy Mgmt., Authentication, Authorization, Access Mgmt., etc. (a.k.a. “Security-as-a-Service”)
composition	**Composition Services**	The logical classification grouping for services that supports the compositing of independent services into a new service offering
database	**Database Services** *(or DB-as-a-Service)*	Database services that permits substitutability to various provider implementations.

Service Subtree (Continued)



Name	Description
capacity	Operational services that ensure that the resource capacity allocated to an application (including compute, storage and networking resources) matches its current utilization.
configuration	Operational services that manage and monitor configuration changes on applications to avoid incompatibilities that can result in reduced performance or compliance failures.
logging	Operational services that capture or record information and identifying data about actions that occur in a system. This includes data that could be or contribute to auditable event records,
monitoring	Operational services that monitor for ensure the availability of services and that they are provided in accordance with terms of Service License Agreements (SLAs).
virtualization	Operational services that manage virtualization of compute, storage and network infrastructure.
location	Business services to manage the location, physical or virtual, of cloud based resources as well as clients (e.g., mobile devices).
billing	Business services to manage different types of charges for cloud based resources relevant to a given customer.
metering	Business Services to manage the measurement of cloud based resources (e.g., utilization, transactions, performance, etc.), often to determine how to bill for service usage.
orchestration	Composition services that automate the management of complex applications, services, platforms and/or infrastructures to align them to fulfill business and service agreements and operational policies.
workflow	Composition services that sequence connected steps that support management of a document (e.g., transaction, order, service template, etc.) through a complex system of applications, services, platforms and/or infrastructures.
crm	*Customer Relationship Mgmt. (CRM) Services (example extension of the "bss" classification)*
erp	*Enterprise Risk Mgmt. (ERM) Services (example extension of the "bss" classification)*
srm	*Service Request Mgmt. (SRM) Services (example extension of the "bss" classification)*

CADF Action Taxonomy



Value	Description
create	The target resource described in the event was created (or an attempt was made to do so) by the initiator resource.
read	Data was read from the target resource by the initiating resource (or an attempt was made to do so).
update	One or more of the target resource's properties were modified or changed by the initiator resource.
delete	The target resource described in the event was deleted (or an attempt was made to do so) by the initiator resource.
backup	The target resource described in the event is being persisted to storage without regard to environment, context or state at the time of storage.
capture	The target resource described in the event is being persisted to storage along with relevant environment and state information (e.g. program settings, network state, memory/cache, etc.). Conceptually, a “snapshot” of the resource is being captured at a moment in time.
configure	The target resource described in the event is being set-up to enable it to run on a particular environment or for a particular application or use.
deploy	The target resource is being positioned or made available for use by the initiator resource, but not yet started.
disable	The initiator resource is causing the target resource [that has been started] to disallow or block some set of functions.
enable	The target resource (that has been started) is being changed by the initiator resource to allow or permit some set of functions.
monitor	The target resource is the subject of a monitoring action from the initiating resource.
restore	The initiator is requesting the target resource (or some portion of it) be restored from persistent storage.
start	The target resource is being made functional by the initiator resource and able to perform or execute operations.
stop	The initiator resource is causing the target resource to no longer be functional or able to perform or execute operations.
undeploy	The initiator resource is causing the target resource to no longer be positioned or available for use.
receive	The initiator resource is receiving a message or data from the target resource.
send	The initiator resource is transmitting a message or data to the target resource.
authenticate	A security request used to establish an initiator’s identity and/or credentials to the target resource against a trusted authority.
renew	A security request from the initiator resource to renew a resource’s identity, credentials, or related attributes or privileges sent to the target resource (an authority).
revoke	A security request from the initiator resource to remove entitlements or privileges from a resource’s identity and/or credentials sent to the target resource.
allow	Indicates that the initiating resource has allowed access to the target resource.
deny	Indicates that the initiating resource has denied access to the target resource.
evaluate	The evaluation or application of a policy, rule, or algorithm to a set of inputs.
notify	Indicates that the initiating resource has sent a notification based on some policy or algorithm application – perhaps an alert to indicate a system problem.
unknown	Indicates that the OBSERVER of the event is not, to the best of its ability, able to classify the exact action for the actual event it is reporting.

Color Key

- General resource management (e.g. CRUD)
- Workload and data management
- Messaging
- Security - Identity
- Security - Policy

CADF Valid Actions

'backup',
'capture',
‘create,
'configure',
‘read’,
‘read/list’,
‘update’,
‘delete’,
'monitor',
'start',
‘stop',
'deploy',
'undeploy',
'enable',
'disable',
'send',
'receive',
'authenticate',
'authenticate/login',
'revoke',
'renew',
'restore',
'evaluate',
'allow',
'deny',
'notify'

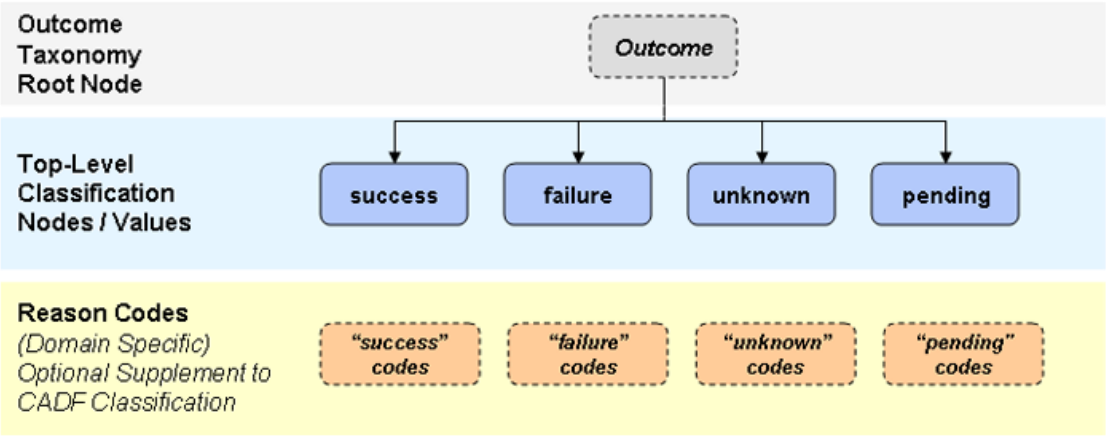
NOTE: CADF Action values are really “paths”. This set of “root” path action values **can be extended** if needed using path extension.

For example,

If a more granular “monitor” action is needed, all you need do is add a path segment to the base action.

e.g. “monitor/poll” the “poll” portion was added as a path to better qualify the type of monitor action.

CADF Outcome Taxonomy



Value	Description
success	The attempted action completed successfully with the expected results.
failure	The attempted action failed due to some form of operational system failure or because the action was denied, blocked or refused in some way.
unknown	The outcome of the attempted action is unknown and it is not expected that it will ever be known.
pending	The outcome of the attempted action is unknown, but it is expected that it will be known at some point in the future. A future event correlated with the current event may provide additional detail.

Backup Slides Follow

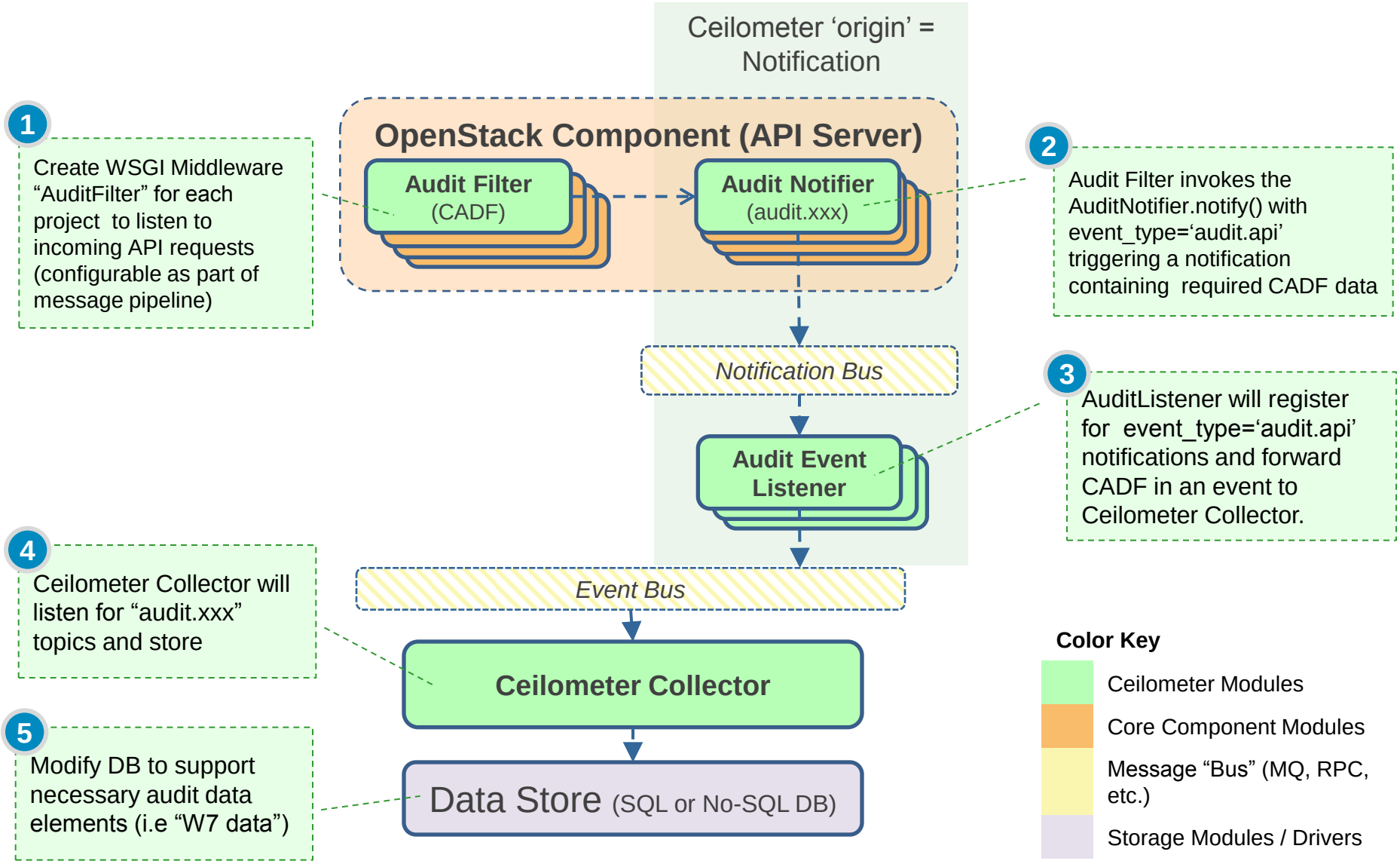
Cloud Audit Data Federation (CADF) Update

Significant Updates in WIP2 from WIP1

- **Event Model to addresses distinct "*complex targets*" use cases, including "actions" that affect:**
  - More than one homogenous or heterogenous resources
  - A target resource that needs a 2nd resource described to provide valuable context
    - Appendix with examples of treatment created.
- **"Tagging" support**
  - Customers can create Orthogonal Views via Path-based "Tag" elements
  - Ability to identify and track multiple Domains of Interest from same event data
    - e.g. PCI, SoX, Local Corporate Policy, Regional Policy, Departmental Policy, etc.
- **Updates to CADF Resource Classification Taxonomy to support DMTF CIMI Entities**
  - Additional supporting use cases and updates to data model
- **Appendix on how CIM Indications can be mapped to CADF**

How the Audit Filter Pushes Audit Events to Ceilometer

CADF Audit Data is "pushed" through Ceilometer's notification path (no delay)



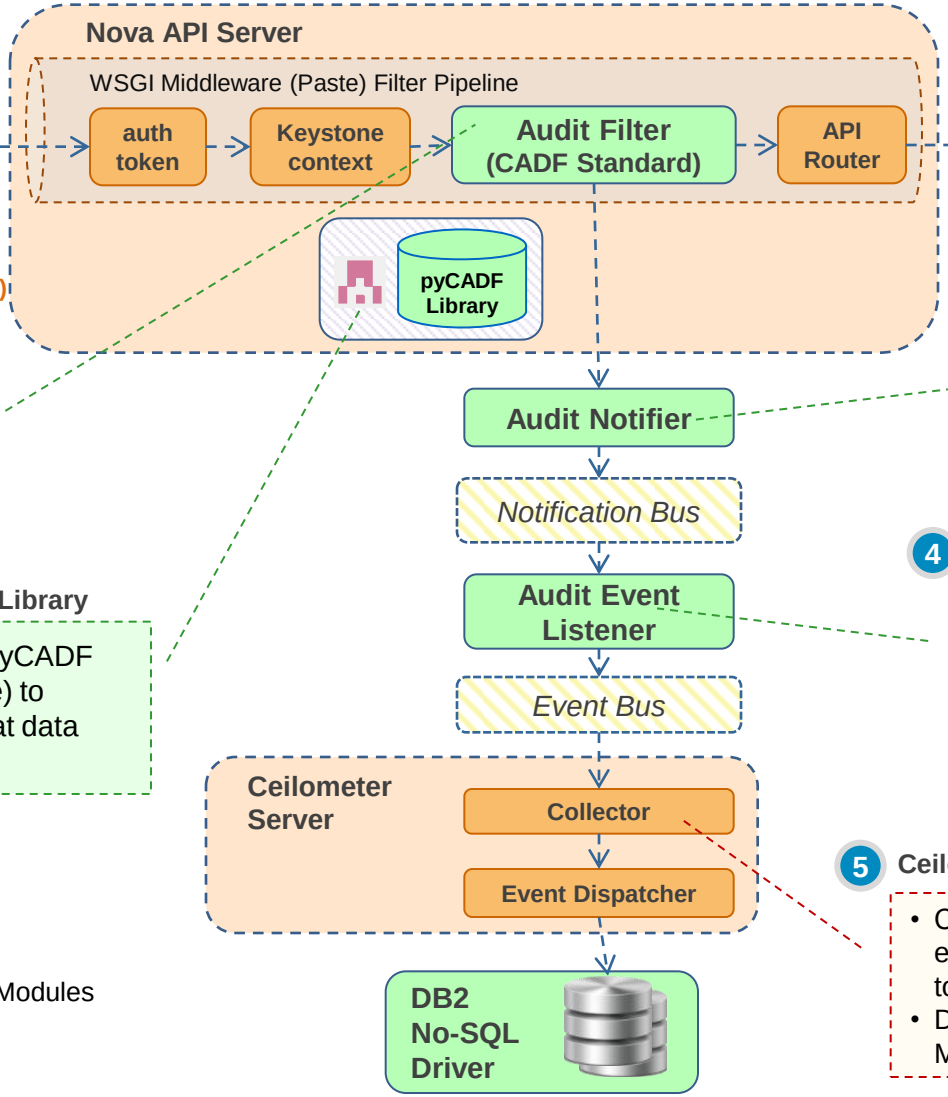
API Auditing with Ceilometer – How it Works…

Incoming Nova Client APIs

- Nova API requests enter "pipeline"
- Keystone authenticates and adds security data

"Backend" Routers

- Request are sent to the "real servers" that manage the actual hardware (e.g. Power)



1 Audit Filter (OSLO Common)

- Audit Filter Intercepts API request
- Maps API and Keystone data to CADF Event Record

2 pyCADF Stackforge Library

- Audit Filter uses pyCADF library (Stackforge) to validate and format data into JSON.

3 Audit Notifier

- Audit Notifier receives CADF audit event (message) and puts it into "Audit Channel"

4 Audit Listener

- Listens for "Audit" notification messages and places them on Event Bus for Ceilometer Collector.

5 Ceilometer

- Collector retrieves "Audit" events and dispatches them to the configured datastore(s)
- DB2 No-SQL, Mongo, MySQL, etc.

Color Key

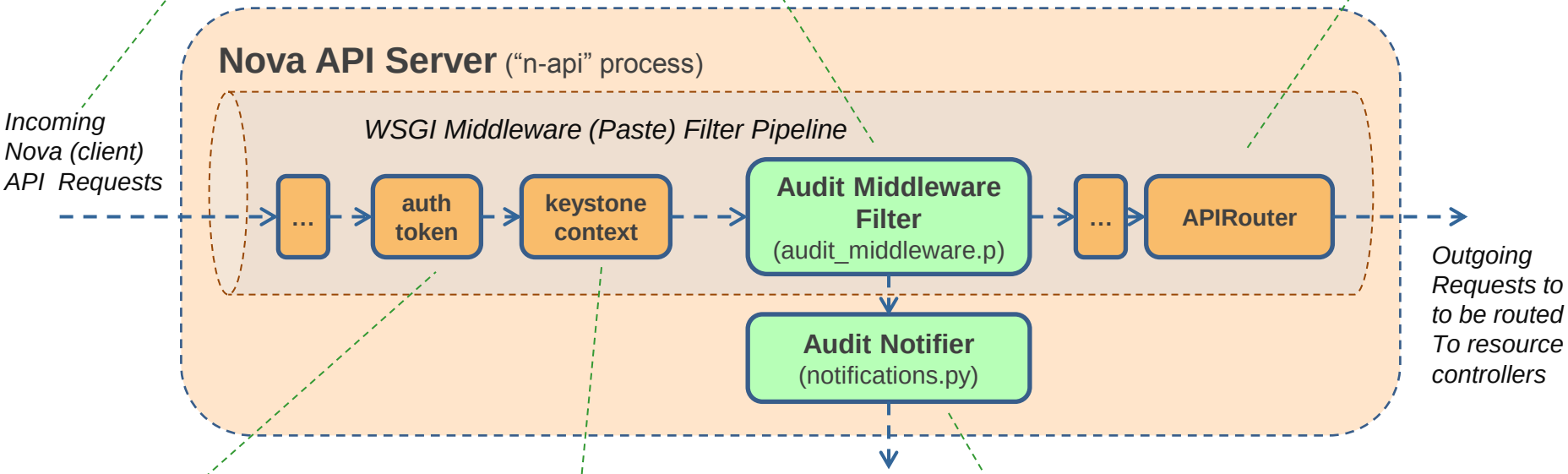
- Existing Nova / Ceilometer Modules
- *NEW* Cloud Audit Modules

How the CADF Audit Middleware Filter Works in the Nova API Pipeline

1 The WSGI "Paste" framework handles Nova API Client requests and passes them to the Nova API Server "pipeline".

4 Audit Filter is added to pipeline after standard "keystone" processing. It uses URI and request data to map to CADF data model including data inserted by "authtoken" and "keystonecontext" and creates a CADF Event Record

6 The API request continues down filter pipeline and eventually is routed to the correct resources (controllers)



2 The (shared) Keystone Filter validates "tokens" and adds account, tenant and user data to request (e.g. user name and ID).

3 Builds a Keystone request context using the data inserted by "authtoken" filter and places it in the request environment (e.g. the "nova.context").

5 The Audit Filter invokes the Audit Notifier with the CADF Event Record which adds it to the standard message payload with the "audit" topic which the Ceilometer Event Listener is registered to receive.

Color Key

- Nova Modules
- Ceilometer Modules

Sample Cloud Audit (CADF) Record produced by the OpenStack Audit Middleware Filter

Nova API Request (example) as "raw" CADF Event Record


```
"CADF_EVENT" :
{
"typeURI" : "http://schemas.dmtf.org/cloud/audit/1.0/event",
"id" : "9e929943-6903-50ad-af9e-90b68bf8ec59"
"eventType" : "activity",
"eventTime" : "2013-08-20T20:52:57.048554+0000",
"action" : "read",
"outcome" : "pending",
"observer": { "id": "target" },
"target" : {
        "typeURI" : "service/compute",
        "id" : "011438ffa2654c179bfef42d0aa150c8",
        "name" : "nova"
        "addresses" : [
                { "url" : "http://10.0.2.15:8774/v2/fac23fee740f45c88e3240d84f920dff", "name" : "admin" },
                { "url" : "http://10.0.2.15:8774/v2/fac23fee740f45c88e3240d84f920dff", "name" : "private" },
                { "url" : "http://10.0.2.15:8774/v2/fac23fee740f45c88e3240d84f920dff", "name" : "public" }  ],
},
"initiator" : {
        "typeURI" : "service/security/account/user",
        "id" : "11ba1e4cc6da4d4c953c158cbde6684a",
        "tenant" : "fac23fee740f45c88e3240d84f920dff"
        "name" : "ceilometer",
        "credential" : { "token" : "MIIQBgYJKoZIhvcNAQcCoIIP9z …… ", "identity_status" : "Confirmed" },
        "host" : { "agent" : "python-novaclient", "address" : "10.0.2.15" },
},
"tags" : [ "correlation_id?value=90681628-fd05-52da-938b-83ef458df26f"],
}
```


Color Code / Meaning
Direct Mapping to CADF
Mapping using CADF Extension
Mapping using CADF "Tag" Extension
CADF Specific value (needed)

How the CADF Event Record answered the 7 “W”s of Audit (for a Nova API request)

“W” Component	CADF Event Mandatory Property	value	
	typeURI	"http://schemas.dmtf.org/cloud/audit/1.0/event	Identifies specification and version for proper parsing and interpretation.
	id	"9e929943-6903-50ad-af9e-90b68bf8ec59“	Identifies this event uniquely (UUID) for federation & reference on queries.
	eventType	“activity”	
What	action	“read”	from the CADF Action Taxonomy
	outcome	“pending”	From the CADF Outcome Taxonomy
When	eventTime	"2013-08-20T20:52:57.048554+0000"	UTC timestamp generated by Audit Middleware filter (includes TimeZone offset)
Who	initiator.typeURI	"service/security/account/user"	From the CADF Resource Taxonomy
	Initiator.id	"11ba1e4cc6da4d4c953c158cbde6684a"	
	initiator.tentant	"fac23fee740f45c88e3240d84f920dff"	
	Initiator.name	"ceilometer"	A Ceilometer agent (“pollster”) called the Nova API as part of a
	Initiator.credential	{ "token" : "MIIQBgYJKoZIhvcNAQcCoIIP9z …… ”, “identity_status”: “confirmed”)	
OnWhat	target.typeURI	“service/compute"	
	target.id	"011438ffa2654c179bfef42d0aa150c8"	
	target.name	“nova”	
Where	observer.id	“target”	TARGET is also the observer (i.e. the case for most OpenStack API pipelines)
FromWhere	initiator.host	{ "agent" : "python-novaclient", "address" : "10.0.2.15“ }	
ToWhere	target.addresses	[{ "url" : “http://10.0.2.15:8774/v2/fac2 ...”, "name" : "admin“ }, { "url" : "http://10.0.2.15:8774/v2/fac2 ...", "name" : "private“ }, { "url" : "http://10.0.2.15:8774/v2/fac2... ", "name" : "public“ }],	
	tags[]	"**correlation_id**?value=90681628-fd05-52da-938b-83ef458df26f“,	Allows “correlation” of this event by other layers that log this ID.