INVENTION DISCLOSURE

1. **Invention Title.**
   Method for distributing device metadata across the Internet

2. **Invention Summary.**
   Using this invention equipment metadata such as the physical dimensions, power consumed and compatible accessories can be made available in a defined format for use in OSS tools such as Inventory and Discovery.

3. **Invention Description.**

   a. **Describe the invention in detail and/or attach a description, drawing(s) and/or diagram(s), if available.**

Communications network providers have a need to understand the physical and configuration characteristics of the devices that make up their network. For example they need this information to use in network planning and design tools, capacity management and to make decisions about space, power and cooling needs within their data centers. Manufacturers of network equipment typically make this information available in human readable format using datasheets and printed and electronic manuals and network providers enter this information into databases and tools manually so that they can then accomplish the above tasks.

This process is inefficient and error-prone and does not allow for the manufacturer to easily make changes to the documentation that would then be immediately reflected in the tools of the network provider. Using this invention a manufacturer could make this information available across the Internet on their website and the tools of the network provider that needed this information could pull it as needed. This would allow the manufacturer to keep the information up to date at all times with each network provider and it would allow the network provider to reduce errors in data entry and reduce the effort to create this device metadata. The invention is to create a formal data schema that all manufacturers could follow and make this schema then a part of the tool set that the network providers use to manage and plan their networks.

   b. **Why was the invention developed? What problem(s) does the invention solve? How is it better?**

This invention was developed to make the operations of our members more efficient and less error-prone as it relates to network planning, discovery and inventory. The invention solves a problem of manual creation of device metadata in proprietary formats of OSS tools and moves it into a standardized format. This is better than the alternatives because the network provider currently has to create device models based upon the tool vendor's schema and maintain them according to that schema. These formats act as a way to lock a provider into a particular tool and are expensive to maintain. Creation of a new piece of equipment or compatible accessory takes time and effort and is error-prone.

c. Briefly outline the potential commercial value and customers of the invention.

The customers of this invention would be the equipment manufacturers including the primary manufacturer as well as the peripheral manufacturers. A company that makes compatible parts for a variety of equipment for example could use this invention as a way to advertise that compatibility across the Internet for discovery by automated tools. The primary manufacturers would also benefit by giving them another channel for marketing information about their equipment.

Network providers are another customer because they will benefit from having standard data that is entered by the manufacturer. This will reduce effort on their part to create and maintain this information in proprietary formats in various tools and it will reduce errors made in these tools as well since the data will be kept up to date by the manufacturer. Having a standard format will also allow a marketplace of equipment information to become available for equipment whose manufacturer does not participate or have ceased operation since network providers could create the data themselves in the standard format and share among themselves, or through a company like CableLabs, the information and thus spread effort.

OSS tool vendors would also be customers since they would need to create interfaces in their software that could consume this standard format and incorporate the definitions into their tools. This would open a market for more competition in the OSS tools because the lack of proprietary formats would lead to lower costs of replacement.

4. **HOW is your invention different from existing products, processes, systems?**
The closest thing I could find to this is the standard format that Microsoft has created for making Windows compatible peripherals (printers, mice, etc.) visible to the Windows operating system. The differences that I see are that this is related to a single operating system contained on a single computer and what I am discussing here is related to an entire network across the Internet. See http://msdn.microsoft.com/en-us/windows/hardware/gg463145 for details on the Windows data format.

CableLabs®

B/OSS

USING SEMANTIC WEB TECHNOLOGY TO DISTRIBUTE DEVICE METADATA

prepared by Shane Furlong
Senior B/OSS Software Architect
s.furlong@cablelabs.com

June, 2011

[redacted]

DISCLAIMER

Table of Contents

List of Figures

List of Tables

This page intentionally left blank.

EXECUTIVE SUMMARY

Creating and maintaining metadata about network equipment for use in OSS tools is a manual process for service providers that is fraught with errors and incomplete data. There is no reason that it needs to be this way. Each manufacturer of equipment supplies the service providers with the information needed to populate this data either through user manuals or freely on the Internet through data sheets but the information is not available electronically and the OSS providers use proprietary schemas and formats for device templates and configuration rules. This paper introduces a plan to use Semantic Web technology to create an industry-standard device metadata format so that OSS tools can pull device models into their schemas automatically from the Internet.

1 INTRODUCTION

One of the problems associated with maintaining an accurate and useful inventory of the network elements of a service provider is the creation and maintenance of the metadata about that equipment. Attributes of a piece of network equipment that are useful for a service provider to have available include things such as:

- **Physical dimensions**
- **Slot and shelf configuration**
- **Model number(s)**
- **Amount of power consumed and heat generated**
- **Available line and control cards with slot compatibility rules**
- **SNMP OID**

This information is contained on the manufacturer's websites but it is locked into human readable formats only. In order for this information to be made available within an inventory or design tool inside of a service provider a person has to manually enter this information into the tool using whatever proprietary format or schema that the tool enforces for this purpose. It would be very useful if this inventory tool could pull the information from the manufacturer's website as needed and populate the information in the tool automatically using a defined format that all manufacturers could agree upon.

This would benefit the service provider as well as the manufacturer. For the service provider this would reduce the effort to maintain their inventory systems by making the creation of metadata an automated task. The service provider could also benefit in having more accurate information, less manual entry means less mistakes, and updates would be fixed at the manufacturer for all providers to pull simultaneously. The service provider would also have accurate design rules built into the metadata which would enforce for them the manufacturer's rules for configuration of a device such as a certain card type can only go into certain slots.

For a manufacturer a system such as this would allow them to publish information accurately about their products in one spot and make it available for all their customers. For example if a manufacturer is setting the end of life date for a certain model of router they could update the metadata for that router and know that all subscribed service providers would have received that notification. Also a manufacturer would know that a new piece of equipment or card was added to the available metadata and the information about it would become a part of the metadata that a service provider pulls from their website which, depending upon the consuming application at the service provider, could serve as a marketing tool for the manufacturer.

In order for this idea to become a reality the manufacturers and the service providers need to develop a common format for the information first. This format should be easily consumed by a computer program and be available on the Internet. The data format would need to be able to hold the version of the information both that of the underlying metadata as well as specific versions of the information from the manufacturer as a whole or for pieces of equipment. A company like CableLabs is in a position to help create this data format and facilitate the adoption of the format using the relationships we have with our members and vendors that supply equipment to them.

This paper describes this idea and a start to the data format in detail and explains how semantic web technology could be used to accomplish this purpose. The semantic web is described by the W3C as a "Web of Data" that is envisioned as a way to enable computers to do more useful work and to develop systems that can support trusted interactions over the network.[1] The standards behind the semantic web have been created by the W3C for just this purpose and this paper proposes to develop this metadata using these standards, specifically to create an ontology using the OWL format to maintain this equipment taxonomy system.

2 Semantic Web Technology

The term "semantic web" was coined by Tim Berners-Lee, the director of the World Wide Web Consortium (W3C) and the standards are maintained by that organization. The intent of these standards is to create a web of data that is just as useful to machines as the current web is to humans. The Semantic Web also applies to the set of technologies, data formats and rules, developed to realize this vision.

At its heart the semantic web uses a simple subject-predicate-object relationship to define the body of knowledge about a subject. These relationships, or triples, are defined in the metadata for a particular knowledge base. So for example a knowledge base about a school might include the following elements as shown in Table 1 below.

Table 1. Example Triples for a School Knowledge Base

SUBJECT	PREDICATE	OBJECT
School	Is Led By	Principal
School	Has Location	Street Address
Principal	Has Name	Name
School	Has Grade	Grade
Grade	Has Homeroom	Homeroom
Homeroom	Has Teacher	Teacher
Teacher	Has Name	Name
Homeroom	Has Student	Student
Student	Has Schedule	Schedule
Schedule	Has Class	Class

Note: This list is incomplete and meant only to show examples.

As can be seen even a simple concept like a school can have many triples and the uses of the data could be varied. Systems as diverse as the school's personnel management system, the class assignment system and a student grading system could all make use of a populated knowledge base for a school.

An ontology is a concept from philosophy used to describe what exists and how these things that exist relate to each other. In a similar manner the Semantic Web uses this term to describe a vocabulary that defines the concepts and relationships within a specific knowledge area. In this paper for example an ontology is described that defines the concepts and relationships related to the physical aspects of network equipment and related equipment.

2.1 METADATA VERSUS DATA

In the above example about a school the information about the school, the instances of the teachers, the principal, the students, etc. are data. The structured format for the storing that information in a knowledge base, the triples from the table above, is metadata and in the semantic web both metadata and data can be stored in the same files and available across the Internet.

For this paper the proposal is to create a metadata format for the maintenance of the knowledge base related to a particular manufacturer's line of network equipment. This format would then be used to create data on the specific models of equipment and accessories that the manufacturer carries. This data along with the metadata that defined the format would both be available across the Internet and could be pulled into an application in use at a service provider to become metadata that allows for the creation of instances of the specific equipment within the network inventory of that provider.

2.2 Semantic Web Stack

Figure 1 below shows the Semantic Web Stack. The specifications that underpin this stack are from the W3C.

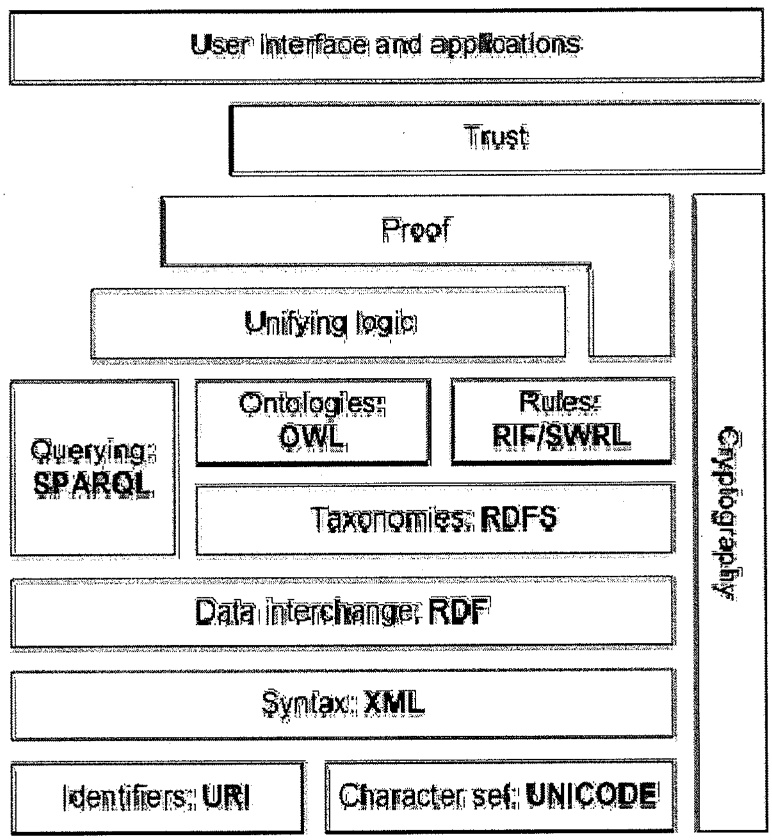

Figure 1. Semantic Web Stack

The stack builds upon the familiar elements of URIs for identification and XML and XML Schema definitions for syntax. RDF is a language for expressing data models which can be expressed in XML syntax (although there are other formats for expressing RDF). RDF Schema, RDFS, is a vocabulary for describing RDF-based resources.

OWL, or the Web Ontology Language, is an XML Schema based language for describing properties and classes as well as the relationships (the predicates in the triples) between these properties and classes (the subjects and objects in the triples). SPARQL is a query language with similar syntax to the familiar SQL and RIF/SWRL are rules based languages that infer relationships between classes and instances in an ontology.

2.3 Examples of Ontologies

There are many ontologies and use of Semantic Web technologies across the Internet to date and more coming every day. Most of the work has been done within the life sciences including work done in pharmaceuticals and the medical field but there have also been projects to create ontologies to be used within fields as diverse as law and human resources. The W3C maintains a good listing of ontology projects at http://www.w3.org/2001/sw/sweo/public/UseCases/.

3 Usage Scenarios

Creating a standard format for equipment metadata and making it readily available across the Internet is a useful idea for multiple scenarios for both manufacturers and service providers. The following are just some examples of how this data could be used to make operations for the service provider and manufacturer more efficient.

3.1 Data Center Planning

Planning for the heating, cooling and space needs of a data center can be a complex task. An accurate source of the requirements and dimensions of equipment to be placed in a data center is necessary to accomplish this task. Currently this information can be found on the datasheets and user manuals of the manufacturer of the equipment but has to be manually translated into systems that do the space planning and design.

If instead this information was available on the Internet in a defined OWL format then the applications being used by the service provider could pull that information as needed by the engineer. This can be seen in Figure 2 below.

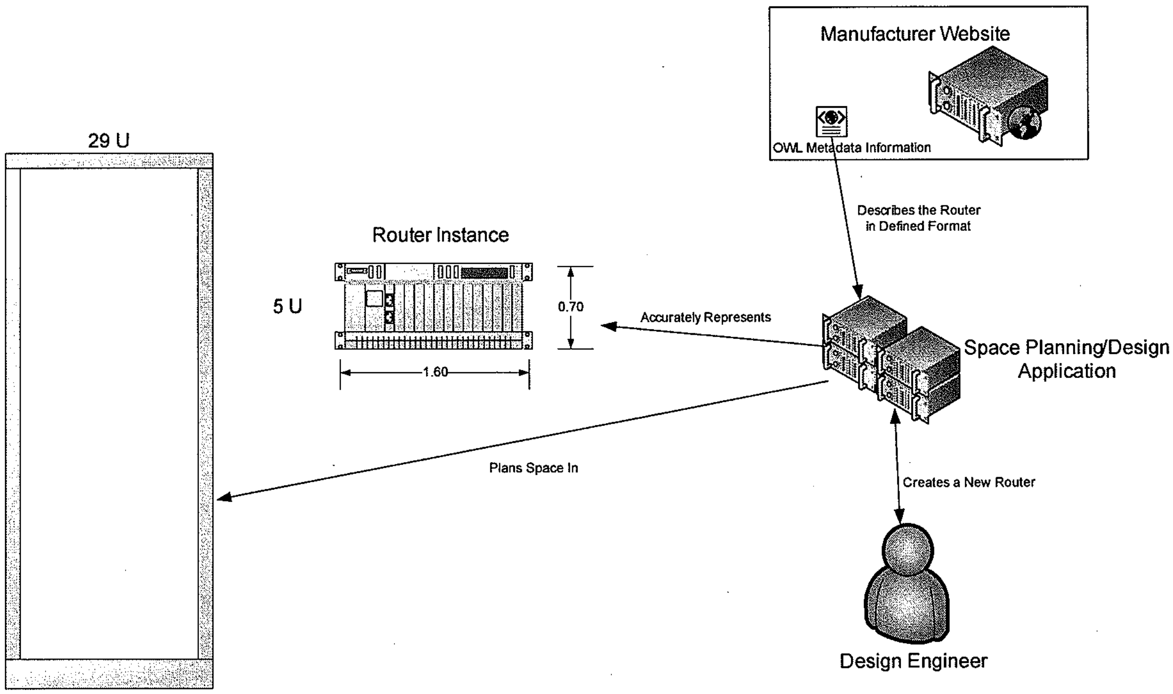

Figure 2. Data Center Planning Scenario

In this scenario the engineer would have an accurate set of dimensions for that model of router and would ensure that there was space for the router in the rack in the data center along with power and cooling for the router as well. If the router needed additional planning, line card types and compatibility rules, this is also possible using this approach.

3.2 Resource Design

Designing new resources to use in a service provider's network is a complex task. The design engineers have to take into account the current and future demand for services as well as the availability of capacity within the data centers and the current network. One of the feeds to this process is the physical capacity for a piece of equipment according to the manufacturer along with the line and control cards that go into the equipment.

With a standard format for this data that is available from the manufacturer's website the design engineer can use an application that pulls this data as needed and makes it impossible for the engineer to create a design that is not feasible. Semantic Web technologies can ensure that the engineer does not forget a part or violate a constraint on the equipment (planning for a card in a slot that is incompatible for example).

This scenario can be seen in Figure 3 below.

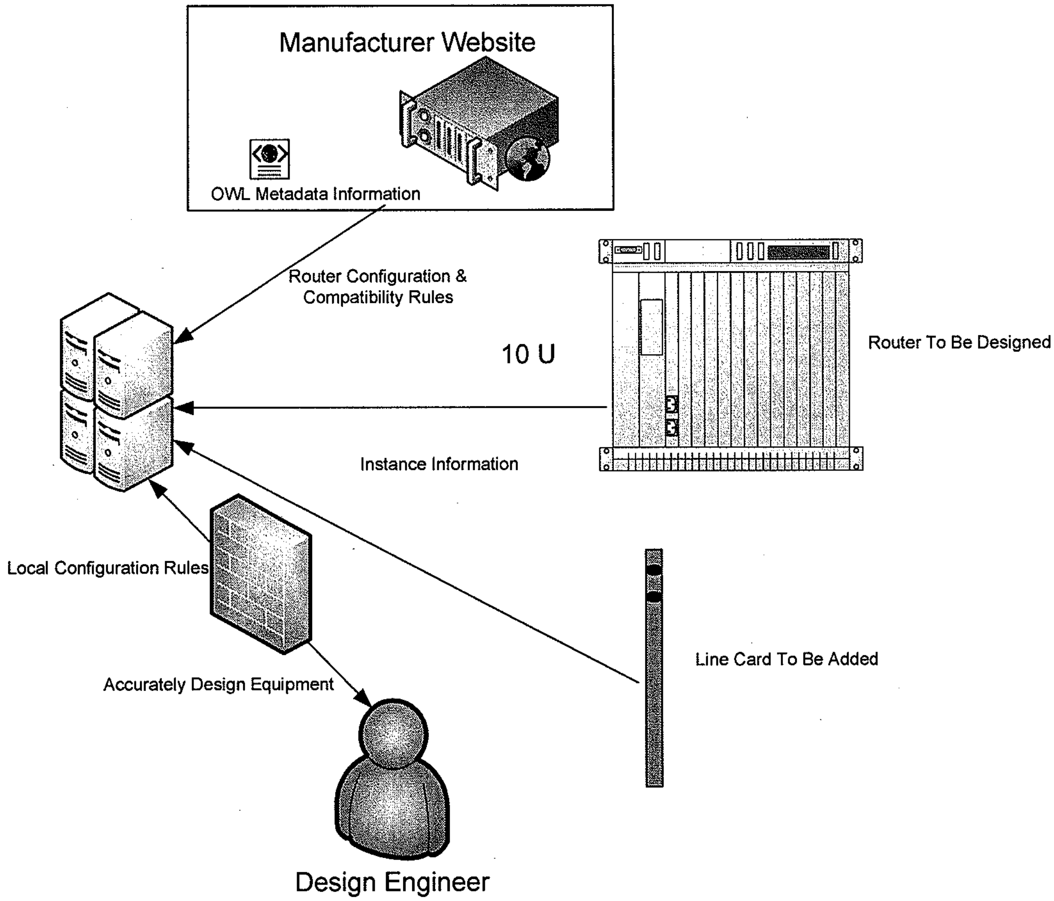

Figure 3. Resource Design Scenario

In this scenario the design application that the engineer is accessing could still be augmented with design rules that have been created by the service provider to ensure that the design not only is valid per the manufacturer but does not violate the individual design rules of the provider. For example if a service provider creates a design rule that says all 10GB line cards must have a fail-over card then this local design rule could be added to the metadata within the application of the service provider.

3.3 Manufacturer Obsolescence

Manufacturers set unsupported and unavailable dates on equipment all the time. Generally these notices come out on their website and go out in emails and notifications to customers. In many cases the manufacturer creates a replacement product at the same time as the notification of obsolescence. In this scenario the manufacturer updates their metadata on their website as well as sending the same notifications that they do today which has the effect of updating the design applications of the service provider when they pull the information as seen in Figure 4.

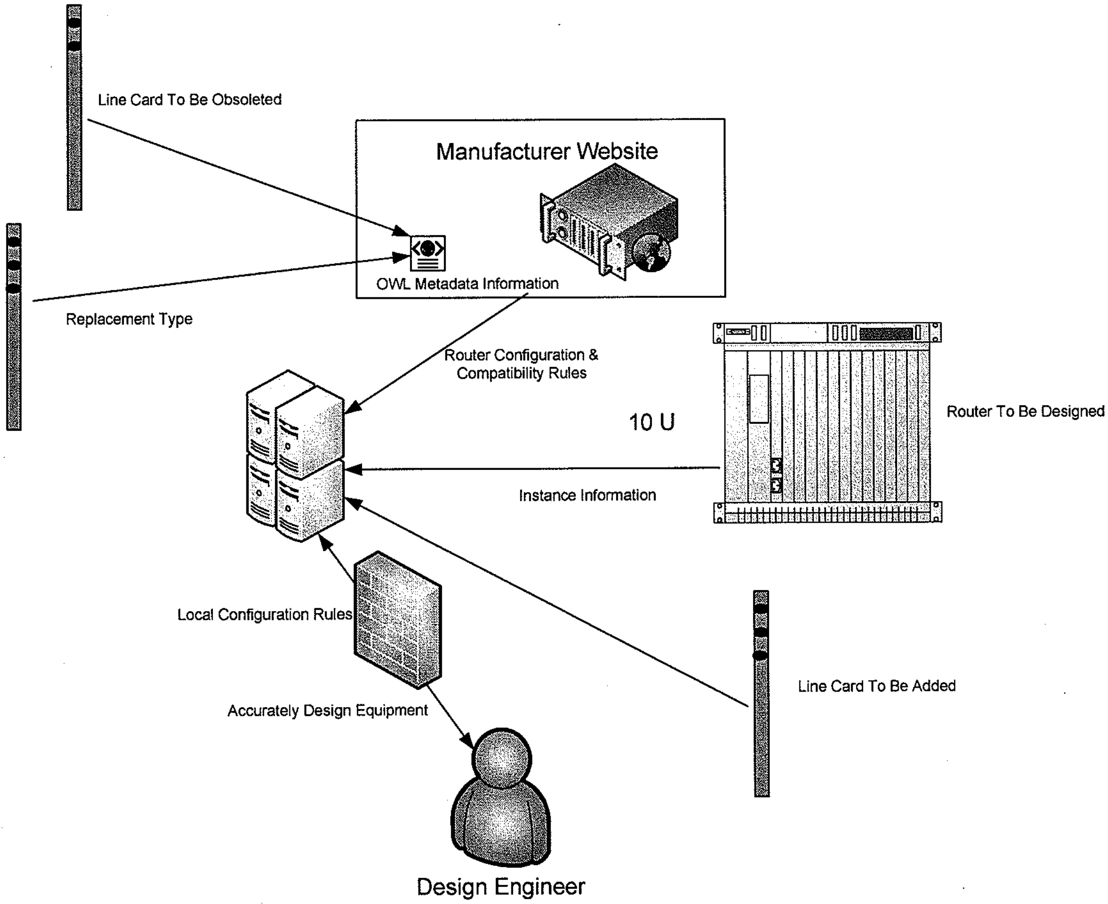

Figure 4. Manufacturer Obsolescence Scenario

In this scenario it should still be possible, through the use of local rules and extensions to the metadata, for the service provider to designate a replacement part as approved or unapproved for use within the service provider's network. The mere fact that the manufacturer has made the replacement available does not necessarily mean it should be used in a design by the engineer but the metadata would allow for the service provider to make their decisions with up to date information.

This same scenario could be used to notify the service provider of brand new equipment availability not just replacement and a service provider could build their design application in a way that shows them what is available and whether that equipment has been evaluated by their company and approved for use. By using this standard format and making the data available across the Internet the manufacturer can open up an entirely new sales channel through the design applications of the service providers.

3.4 CableLabs Maintains Metadata Repository

It may be some time before manufacturers support this standard format but it could be possible to begin using it even before it is widely adopted. If all of the member companies adopt the use of the data in their design and planning applications it will increase pressure on manufacturers to begin making the data available but until then CableLabs could maintain a catalog of created metadata for different manufacturers on its website that could be pulled by members.

This data could be created at different members and shared with each other through CableLabs website and CableLabs could maintain this repository until a manufacturer begins to create their own. This could also serve as a repository for orphaned equipment information for manufacturers that never agree to participate or those that have ceased operations but still have equipment in the networks of service providers and available through other channels.

This scenario is shown in Figure 5 below.

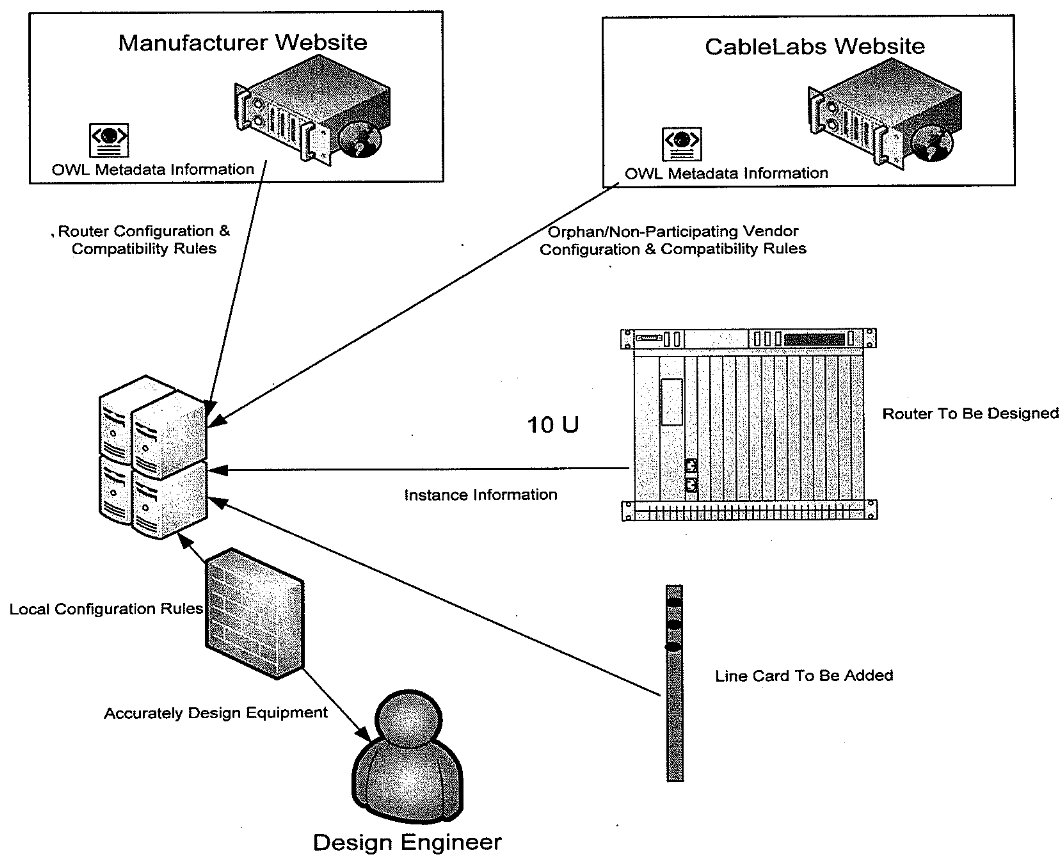

Figure 5. CableLabs Maintained Metadata Repository Scenario

3.5 Assisted Discovery

Another scenario that making this metadata available could help to solve is the problem with identifying equipment that is discovered on the network of a service provider. All service providers struggle with the problem of unknown and unmanaged equipment proliferating on their network and there are tools that they employ to try to help with that problem. In most cases these tools are some variation of an SNMP "crawler" that reaches across a network using various pointers and methods to find equipment that answers a query.

The problem with this approach is that the tool sometimes has to be loaded with the equipment's configuration and type before it can be discovered which is something of a Catch-22 since by definition the equipment is unknown which might also mean the type is unknown. By including the manufacturer's SNMP OID in the metadata for a device the discovery tool could have access to all available equipment types making the discovery of unknown equipment on the network much less likely.

This scenario is shown in Figure 6 below.

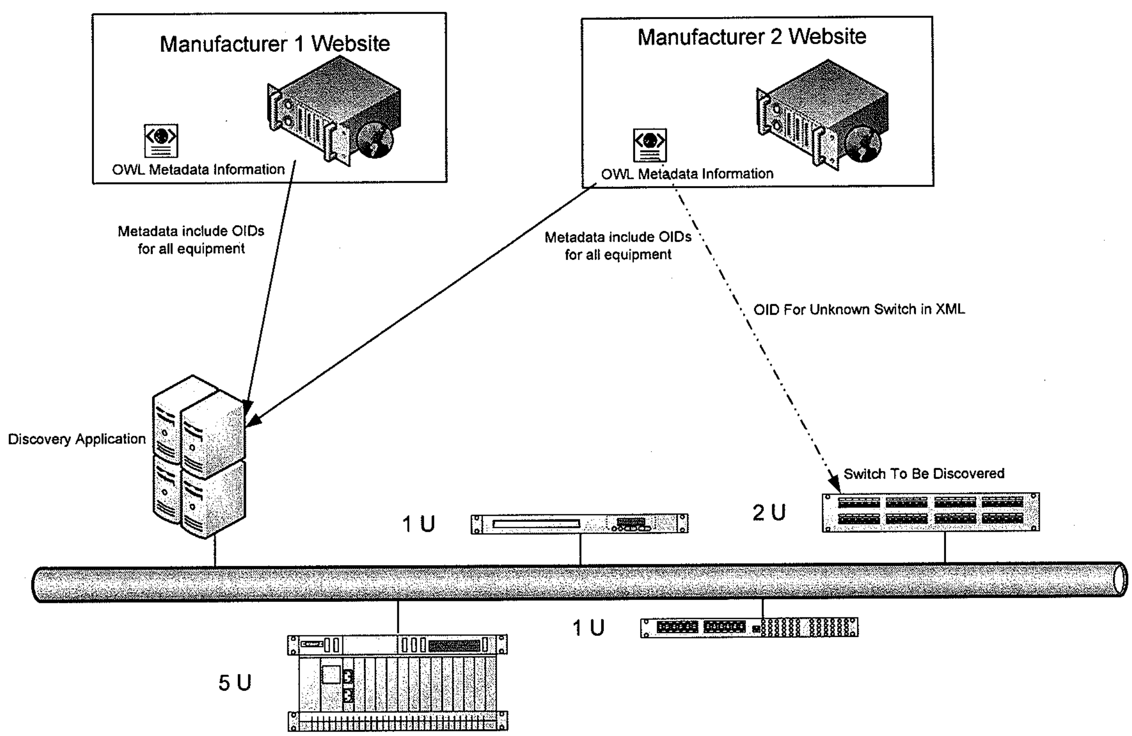

Figure 6. Assisted Discovery Scenario

Assisting in discovery is extremely valuable in helping service providers locate "orphaned" equipment and make sure that they have an accurate inventory for the purposes of support and maintenance contracts.

 [redacted]

4 Support From OSS Vendors

Beyond the support of the equipment manufacturers this idea needs the support of vendors that create OSS tools used in processes like Resource Design and Assign, Resource Inventory Management and Space Planning. Each of these vendors already has proprietary formats and schemas for creating and working with device metadata that will have to adjust to support the ontology. While this could take some time it should be possible for each vendor to create tools that pull the data from the Internet and convert it into their proprietary format, creating files or loading information in their database for example, in a batch fashion. Over time it should be possible for the vendor to refactor their applications to pull the data as needed from the Internet automatically.

5 Proposed Equipment Ontology

This paper proposes that a metadata format be created using OWL to show not only the data elements but the relationships between related equipment. Semantic Web technology is suited for this purpose because of its ability to encapsulate rules within its syntax better than pure XML. For example a device may have ports on it that are conditionally active based upon whether an SFP is in a slot or not. This sort of relationship can be easily modeled in OWL and not as easily expressed in pure XML.

The standardization of the metadata will take a full CableLabs project with input and advice from members and vendors but this paper does propose the following elements be created in the metadata.

5.1 Data Elements

The ontology developed to describe network equipment and related peripherals should include the following data elements:

- **Equipment Chassis**

  There should be a different top-level object for an element that is independently orderable from the manufacturer. The equipment configuration with the available ports that are on the chassis as well as any management and serial ports should also be represented. The number of slots on a chassis and their orientation (front and back, horizontal and vertical) should also be modeled.

- **Line and Supervisory Cards**

  Cards that are independently orderable should be represented as separate objects. If the cards must come as pairs (front and back for example) but are different part numbers they should be separate objects and the relationship should be managed in rules. The number and types of ports that are on a card should be included in the data.

- **Small Form Pluggables**

  All forms of SFPs such as GBICs, SFPs, XFPs, etc. should be modeled as independent items if they are orderable separately from the container element that holds them.

- **Fans**

  If a fan is a separately orderable or field replaceable unit it should be included in the model as a separate object.

- **Power Supplies**

  If a power supply is a separately orderable or field replaceable unit it should be included in the model as a separate object.

- **Install Kits**

  Installation kits that are ordered separately should be modeled.

- **Brackets**

  Mounting brackets that are ordered separately should be modeled.

- **Power Cords**

  Power cords that are ordered separately should be modeled.

5.2 EQUIPMENT ATTRIBUTES

The exact list of attributes with data types and ranges if applicable is to be determined within the context of a collaborative project between CableLabs, members and manufacturers but attributes such as the ones below would be a part:

- **Physical dimensions**
- **Number of slots and types of slots**
- **Model number(s)**
- **Number of ports and types of ports**
- **Amount of power consumed and heat generated**
- **SNMP OID**

5.3 DATA RELATIONSHIPS

Semantic Web technology is useful for categorizing not only a list of attributes but also codifying a set of rules within its syntax. For example a router may have a list of compatible cards that can be deployed in a set of available slots on its chassis. These rules may be complex based upon combinations and whether or not other attributes are set to different values such as a specific firmware version is needed. All of these complexities can be modeled in the ontology.

6 For Future Consideration

Beyond the creation of an equipment metadata ontology and working to gain acceptance in both the vendor and the OSS software community for the format this project could lead to the standardization of an asset lifecycle management system for network elements. Once the manufacturers have agreed to use a defined format for the metadata it would be a small move for them to extend that model to start creating "birth certificates" for this same equipment as it is sold to a service provider. These certificates would use an extended version of the metadata model to include elements that uniquely identify the specific instance of the equipment and could be used to track the equipment from order to installation and ultimate disposition.

A system such as this would give the service provider better asset control by keeping an up to date inventory between the manufacturer and the service provider of equipment installed. The common format for the data is only one aspect of this idea and the process whereby the certificate flows through the procurement process, including how this impacts resellers, would need to be considered. Additionally for this to work the service provider would need to change internal business processes to incorporate the maintenance of the certificate throughout the lifecycle of the equipment.

7 References

1. Semantic Web Standards - http://www.w3.org/standards/semanticweb/