The Minimalist Reference Architecture

Contributed by Charles Meyer, Chief Technology Officer, Production, Grass Valley, A Belden Brand
Published 16 July, 2015

Note: This article is being made available as part of the activities of the Joint Task Force on Networked Media. The author is solely responsible for its content. While the information contained herein may be of interest to those following the work of the JT-NM, this document is published by its author, and is not being published by the JT-NM itself.

Reference Architectures, or RAs, are abstract. In some ways, the more abstract the better. Ideally, they are *machine to make machines.* In another line of thought, they should represent the simplest essence of the family of systems you are trying to represent with one common model.

First principles, Clarice. Simplicity. Read Marcus Aurelius. Of each particular thing ask: what is it in itself? What is its nature?

[1]

The OSI (Open Systems Interconnection) protocol stack is perhaps one key example of a minimalist reference architecture. The interfaces between layers are well defined. The behavior of each layer is easily modeled. Each layer has a number of standards and choosing

---

[1] Image from "Becoming Hannibal Lector," *Time*, http://content.time.com/time/arts/article/0,8599,1568581-2,00.html, accessed March 14, 2015.

a set of standards, one per layer, is similar in concept to a profile. Sometimes a profile itself might be a standard, such as SMPTE 2022-6.
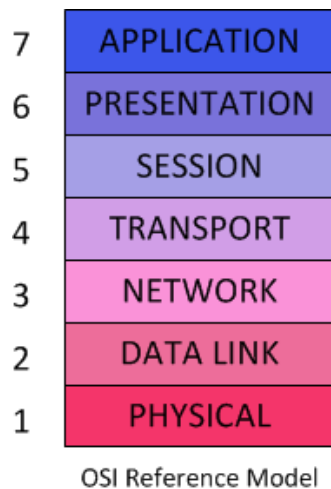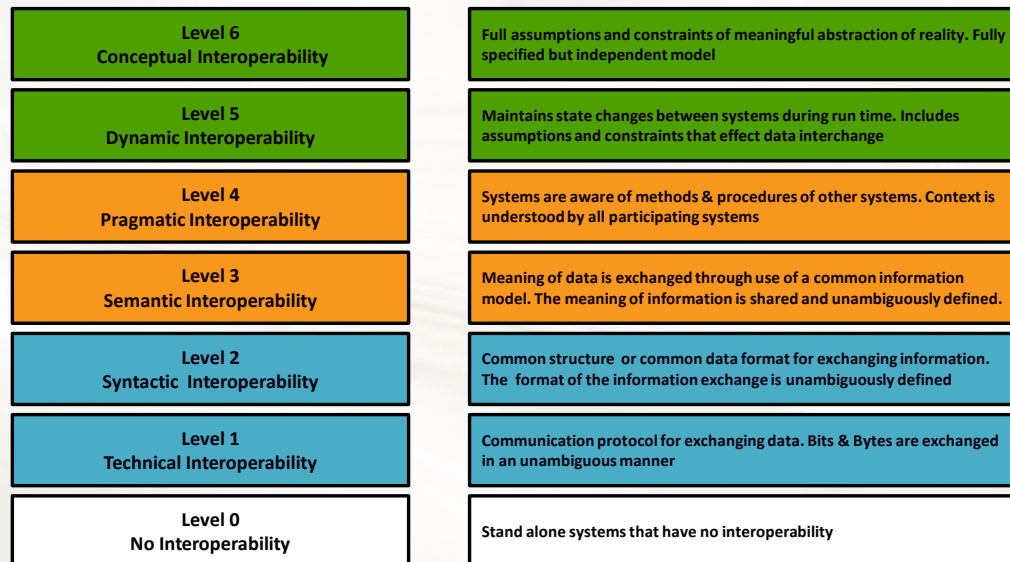

OSI Reference Model

Figure 1. The OSI Reference Model

In a good Reference Architecture, layers do not intermix. Clearly identified boundaries between the layers and functional capabilities contained in each layer are defined. This is not always easy, but the value of being able to substitute copper for fiber at the physical layer, as an example, is significant.

Ethernet, FDDI (Fiber Distributed Data Interface) and ARCNET (Attached Resources Computer NETwork) are but 3 of many link layer protocols. IP, ATM (Asynchronous Transfer Mode) and XNS (Xerox Network Systems) are but 3 of the many protocols in Layer 3, the network layer. Ethernet and IP are dominant. But, other combinations provide business value and should not be excluded. In fact, because the OSI stack is respected, it is possible to exchange layers, one with the other, as is the case of Ethernet over OTN (Optical Transport Network).   As IT technology is applied to more media workflows, there will be more variations to layers 4 through 7. It is not unusual for standards covering a new OSI Stack to reference other adjacent standards for interoperability between the layers. SMPTE 2022-6 is one such new standard.

The Joint Taskforce on Networked (Professional) Media (JT-NM) is making significant progress in developing a Reference Architecture for media facilities integrating IT technology and equipment. This is more complex than an OSI protocol stack. But perhaps, if the fundamental basics of media production can be modeled, then a minimum model can be developed.

Levels of Conceptual Interoperability Model (LCIM) for Systems Engineering

Level 6 **Conceptual Interoperability**	Full assumptions and constraints of meaningful abstraction of reality. Fully specified but independent model
Level 5 **Dynamic Interoperability**	Maintains state changes between systems during run time. Includes assumptions and constraints that effect data interchange
Level 4 **Pragmatic Interoperability**	Systems are aware of methods & procedures of other systems. Context is understood by all participating systems
Level 3 **Semantic Interoperability**	Meaning of data is exchanged through use of a common information model. The meaning of information is shared and unambiguously defined.
Level 2 **Syntactic Interoperability**	Common structure or common data format for exchanging information. The format of the information exchange is unambiguously defined
Level 1 **Technical Interoperability**	Communication protocol for exchanging data. Bits & Bytes are exchanged in an unambiguous manner
Level 0 **No Interoperability**	Stand alone systems that have no interoperability

VMASC (Virginia Modeling, analysis and simulation centre)

rti

© 2012 Real-Time Innovations, Inc.

1

Figure 2. Levels of Conceptual Interoperability Model (LCIM) for Systems Engineering (courtesy of Real-Time Innovations, Inc.).

At the highest level, this 7-Layer stack serves as a general approach for a Reference Architecture.   It is a Reference Architecture for reference architectures.   This is not double-speak. In business, the goal is to generate as much revenue as possible from goods and services offered. If there is a pattern, or repetition in the business model, then make a process and repeat (aka "rinse and repeat").

In our professional networked media industry today, the super-ordinate goals may be expressed as "Monetize Content."   But, before we jump into live video production factories, here is how a different factory exploited these techniques.

In my experience with NVISION, we built machines called routers, master control engines and modular products. NVISION was also vertically integrated with in-house manufacturing. Essentially, NVISION was a machine for building machines. Layers of the machine, our business systems, could be upgraded without changing layers above or below. For example, during the RoHS[2] transition, NVISION used big data methodology to programmatically replicate part numbers with a simple change to indicate lead free. This information was embedded in program files automatically by existence of the new part numbers, which were generated programmatically. Existing documentation was cloned, and modified programmatically with the new, or alias part number. The assembly line robots knew which solder paste and which oven profiles to use based on these changes. The schematic process did not change. The design process did not change. The purchasing process did not change, except to buy only lead free materials.   Many, many things did not change. Only about 50,000 components changed. Then released BOMs (Bill of Materials) and machine programming changed algorithmically and programmatically. Once the 90% point was reached, the toggle switch between the originals and the clones was turned off. 10% of the designs were sold as end-of-life. It was simply an A to B switch on the preselect bus for master control.

Based on understanding the fundamentals of PCB (Printed Circuit Board) manufacturing and exploiting an open data culture, NVISION pulled data into smart documentation and in so doing, created a self-documenting machine for building machines.

The change to lead-free manufacturing took planning, but the system provided the ability to create change programmatically, and far more efficiently than had we manually replicated every schematic, schematic symbol, part number, and the like. These manual double entry systems might have been the alternative were it not for a well-designed system and big data methods so often talked about today.

---

[2] Restriction of Hazardous Substances, http://www.rohsguide.com, pronounced "Ro-Haas".

In another contribution to JT-NM, the following architecture was shared by SRG (the Swiss Broadcasting Corp. in English).
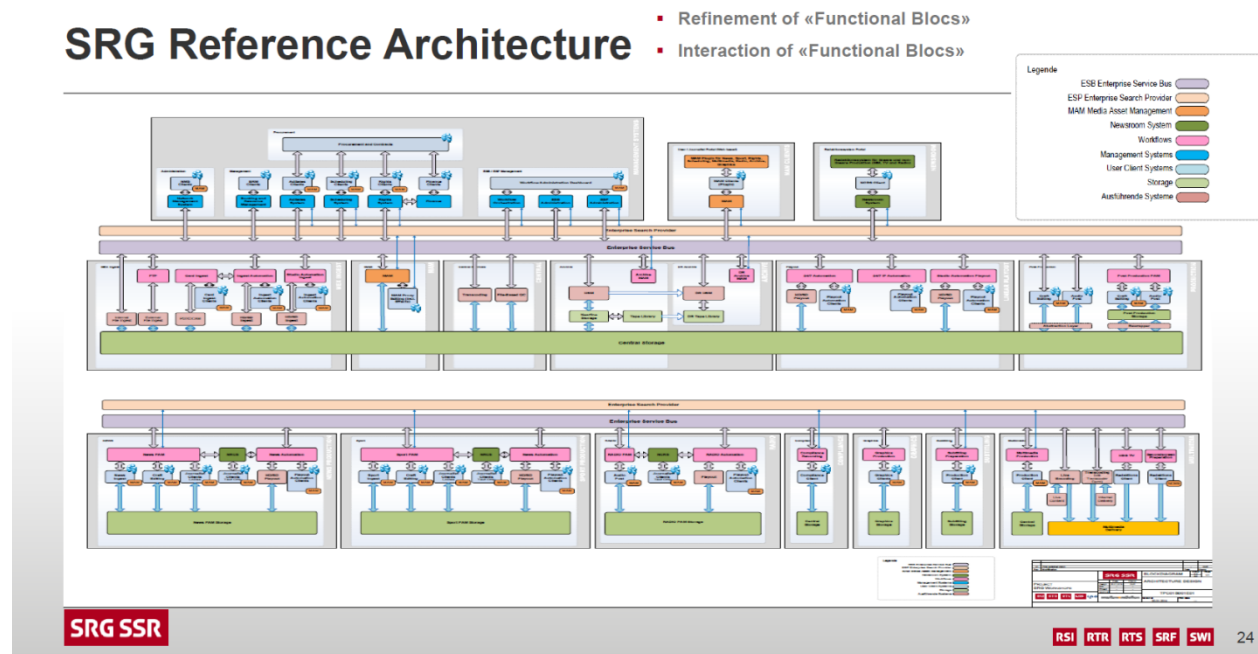

Figure 3. SRG Reference Model (courtesy of SRG SSR).

SRG discussed how, as a company, they spent significant time and effort to completely understand their business needs and their processes. They identified how material and information moved through their facility. They identified common patterns, processes and information transfers. This is indeed an abstract model, and it is minimal for SRG and it is likely SRG specific. But it is an inspiring example of how our industry can apply these modeling tools, which in many ways are exactly the same ones used for MRP and ERP systems deployed in industrial manufacturing facilities as described for NVISION.

Operational Scenario: Live Studio Production (Current) – Platform Zoom
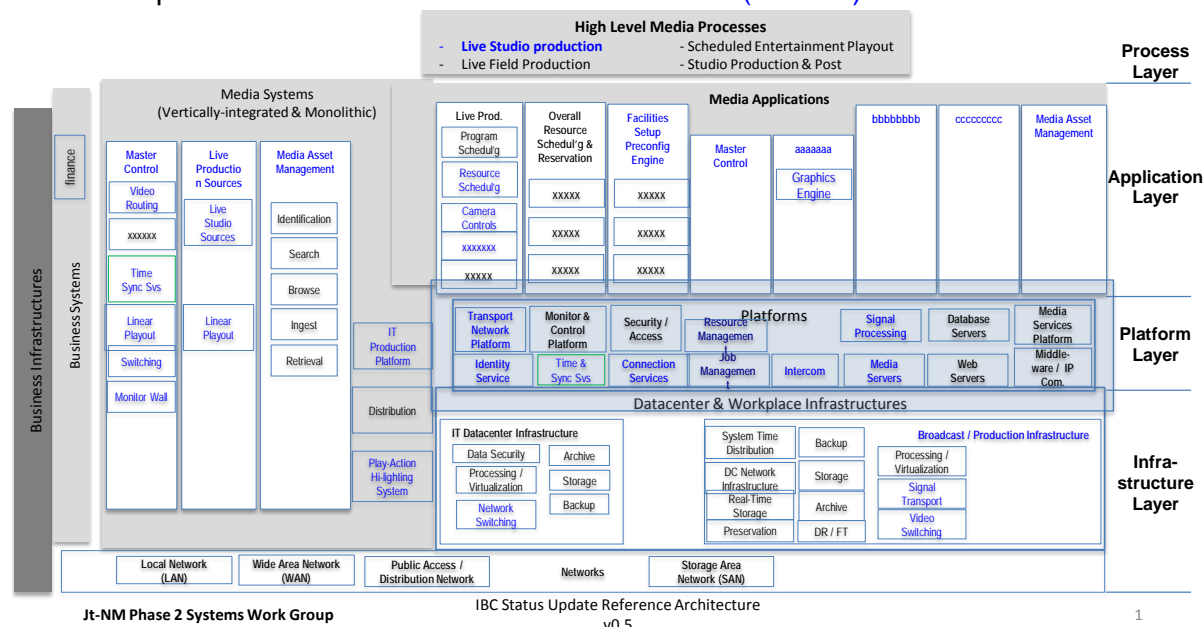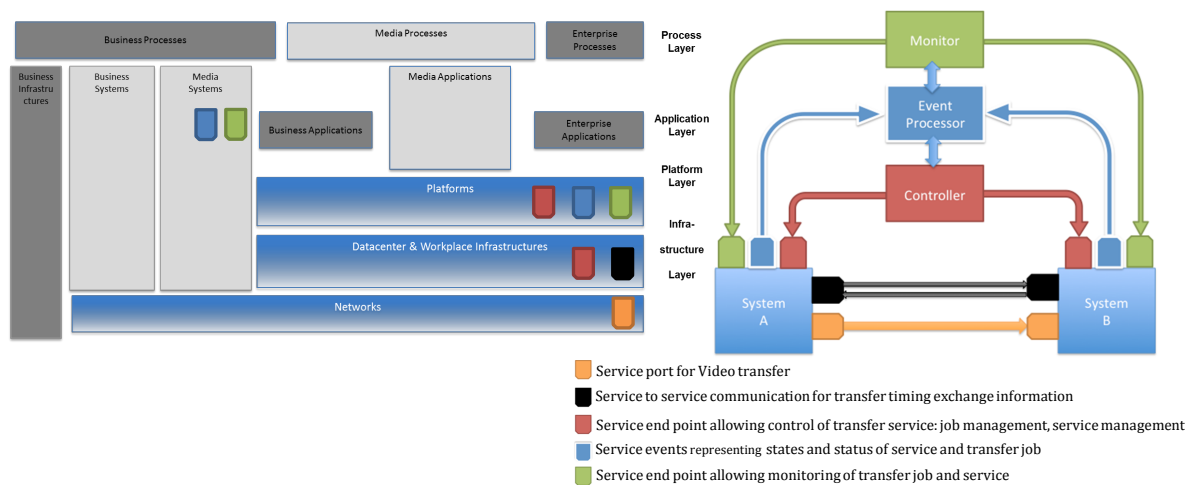

Figure 4. Operational Scenario: Live Studio Production (Current) – Platform Zoom Work-in-Progress (Courtesy of the JT-NM Phase 2 Systems Work Group).

In the above model, the basic layers of an information system are represented. Using this example, it is possible to map the processes used to various layers, identify the information flow, describe the capabilities of the process, and defined the required interoperability. This example shows one approach for representing a model for interoperability. The boxes inside each layer represent the elements required for a general case of video production. By examining their relationships within and between the layers, the model is refined and validated.

This simple model sets the foundation for the Minimum Viable Architecture (MVA). To test this model, we need the minimum behavior of creating and monetizing content. Finally, if we model this core requirement correctly, then common programming models can be employed, enabling multiple vendors, to more easily ensure interoperability between their products and the equipment purchased from them. Even public data models, such as EBU Core[3], and FIMS[4] can find a home in such models.

---

[3] EBU Core (EBU Core Metadata Set, https://tech.ebu.ch/MetadataEbuCore) "Combined with the EBU Class Conceptual Data Model (CCDM) of simple business objects, EBUCore provides the appropriate framework for descriptive and technical metadata for use in Service Oriented Architectures and also in audiovisual ontologies for semantic web and linked data developments."

Where the **Structure Layered Model** and **SOA** meet

Figure 5. Where the Structure Layered Model and SOA Meet (courtesy of the JT-NM Phase 2 Systems Work Group).

The model above (courtesy of Loic Barbou & Triskel and Dr. Karl Schubert & TechNova Consulting) is used to offer more detail on how the interfaces between layers operate, and what information may be transferred. Underneath all of this exists a bare minimum behavioral model. The last step is to create just this.

So, Is it possible to simplify content monetization? What is its essence?

It starts with human factors. Sight and hearing, even smell, taste and feeling might be considered. Anthropological considerations could lead one to conclude that pictures are so important to us because they were the first method of recorded information interchanged. Music next, it provided patterns to sound, making it easy to remember and music crosses language barriers. Then we move to the written word, and so on.

---

[4] FIMS (Framework for Interoperable Media Services, http://www.amwa.tv/projects/FIMS_projects.shtml) is a joint task force by the Advanced Media Workflow Association and the European Broadcast Union (EBU).

For today, pictures and sounds make up our media experience. And, the archived versions of these must be included. As soon as we can adapt new MEMs[5], or chemistry advances, or as yet unpredicted technology to record and play back more human senses, these will become part of our media experience.

The key is that all we need for any media "sense" to be monetized is to record it, and play it back, and do so with reasonable time accuracy. It would of course be advantageous to audit the processing and distribution path it follows.

That is pretty simple.

What is the essence of live production?

In a simplistic model for real time production, the requirement might be one camera, one microphone and one recorder. The number, type and/or quality of each is additional information, but not necessary for a basic model.

This simple recording needs time. Advertising is sold based on time, both for duration and prime time viewing. Time is required in the future in order to recreate new programs from the original recordings.

Monetization can be simplified as well.

Total Revenue = Unique Media * (Number of views of that media) * Revenue/View.

This must be summed over the number of ways the media is viewed, which is a list, but not that big a list.

Live Play, traditional advertising

Replay (Re-runs)

M-Screen VOD

Internet Push

Internet Pull

Personal Archive (Store and publish, aka cloud)

Movie Theatres

There might be one or two more, but as a general, minimal list, most things can fit one of these categories.

---

[5] MEMS (Micro-Electro-Mechanical Systems) is a technology that allows miniaturization using techniques of microfabrication to create functional elements that are are miniaturized structures, sensors, actuators, and microelectronics.

In the Internet world the two extremes of Live Content Monetization could be as simple as this:

> <u>One Production, One Presentation, Global Distribution</u>. This is essentially the current model for classic paid advertising television. A major sporting championship is produced once, typically presented in only one way, and broadcast to the world.

Or,

> <u>Multiple Presentations, Targeted Distribution, Viewer Production</u>. Imagine a crowd-sourced event where each person generates his or her unique presentation. The audience self-selects by subscription to the program and the final production is carried out by the viewer on their home smart TV.

The cost elements associated with either case are the same: cost of access bandwidth, cost of computation/production resources and cost of media rights.

Processing media adds value. There is a simple and short list for this as well.

Time Shifting

Scaling

Mixing

Formatting

Encoding

Wrapping

Distributing

This is a finite set, each with simple-to-describe capabilities, and each with defined adjustments, settings and configurations.

Time is complicated. *But is it really?*

Time is considered to be very challenging and complicated; and, it certainly is if one considers atomic time, and Precision Time Protocol, (PTP) and GPS Time, and UTC time and Satellite ephemeris adjustments, and leap seconds and all the other science and methods of telling time. But operationally, the following set of operational requirements is sufficient:

> <u>Mutually align Identical Media Types</u>. Think of this as pixel time. This is the ability to time align signals for digital processing.

> <u>Mutually Align Related Media Elements within Human Perception</u>. This can be thought of as bundle, or presentation time, but is easily understood as lip-sync.

> <u>Align Media to Real Time of Day</u>.   Sometimes this is referred to as "appointment TV." In reality, this is the electronic programing guide that is used by a personal video recorder

to know when the broadcast plant automation system is going to air a program. It is the automated alignment of your recorder at home to the recorder used to play out content. This is "computers talking to computers," so that we can enjoy time shifted viewing.

It would seem that the abstract basis for a general Reference Architecture does exist. This means that a general behavioral model for systems can be developed to support the layered model shown earlier. This layered model may then be used to define the interoperable capabilities of, and the interfaces between, the layers.

As part of the JT-NM effort, each layer can then refer to a number of possible related standards and allow for future standards to fit within the model. Additionally, groups of standards supporting common use profiles can be identified. This makes it easier for users to adopt and deploy a model with confidence. These profiles do not need to be implemented as provided, but many businesses will adopt them because interoperability testing can be performed and these test results can be normalized. The best ones might even become standards of their own. A super user who wants to get the most of the Reference Architecture can still do so by defining their own profile or by using a trade secret method to replace a layer, or two, in the reference. By using a systems approach, and a minimalist modeling approach in combination, there is an opportunity to generate a long-term viable Reference Architecture.

In Summary

There is no one answer for how to monetize content. But the tools needed to create, produce, transport, and distribute audio and video are finite and, in fact, not really that great in number. How media is processed can be very unique. This is the artistic, or creative factor that differentiates between media types, styles and content. And the elements that guide how media is consumed are based on human factors and time. It seems very logical to conclude that a viable Reference Architecture can be developed which enables media manufacturers to provide interoperability for their customers. It is indeed amazing to consider how such a simple set of principles leads to the enormous diversity of media offered to today's consumer.