

Notes from an Introductory Meeting on formats for digital video preservation related to the Preserving Digital Public Television project funded by NDIIPP

Held as an adjunct session to the PBS Technical Meetings that precede the National Association of Broadcasters annual meeting in Las Vegas, Nevada

April 17, 2005

Attending, with a few notes from the round of self introductions:

- **Jerry Butler**, from PBS, working on the Next Generation Interconnect Service (NGIS), for file based handling of programs
- **Glenn Clatworthy**, PBS archivist (NDIIPP project participant): notes the successful handoff to LC collections, in the past, of about 23,000 16mm representations of programs and about 20,000 2-inch videotapes from the 1970s, wants to move fwd with handoff of more recent content, and to see planning for file-based handoff in future
- **David Fallon** (?? not sure of spelling), director of operations at Milwaukee Public TV: we have a near-tapeless environment at our medium-sized public TV station. We want to figure out what we should do, we have been shooting everything in HD, and this gives our recent archive higher value. Second: I am the tech mgr for IEC's development of TC-100, multimedia activities, we want to see if developments from this NDIIPP can be internationalized
- **Carl Fleischhauer**, the NDIIPP program officer for the project at the Library of Congress, and the scribe for these notes
- **Bruce Jacobs**, chief technologist at Twin Cities Pub TV, interested in namespaces and other things in the digital environment
- **Jim Kutzner**, chief engr PBS (NDIIPP project participant)
- **Dave MacCarn**, WGBH (NDIIPP project participant): I have spent many years working on the issue of extracting essences
- **Jerome McDonough**, NYU (NDIIPP project participant): I work on digital library infrastructure in various ways
- **Bea Morse**, PBS senior dir of bdcst operations, (NDIIPP project participant)
- **Judy Reynolds**, producer of religion and ethics programs for WNET, based in Washington
- **Nan Rubin**, WNET (NDIIPP project participant), my job includes the WNET archive, exclusively focused on videotapes until very recently
- **Robin Schatzenbach**, NYU, director of the university's media services activity
- **Merrill Weiss**, consultant to CPB (Corp for Public Broadcasting), developer of TV technology for 30 yrs. I chaired an SMPTE and EBU task force on the exchange of programs as bitstreams and the move to file based production, and we published a 1998 report that has been influential (mentions of the task force report sprinkled in the notes that follow). See http://www.ebu.ch/trev_277-contents.html. I have also done other things thru SMPTE, and have worked on International Standard Audiovisual Number and the ISAN international agency, which is relevant to this effort. See <http://www.isan.org/>.

Introductory matters. There was a mixup about the room reservation for the meeting, so a few us were sitting around chatting before it began. One topic in this side conversation concerned the extent of data that public broadcasting produces or manages each year. Weiss said that PBS distributes 10,000 hours per year, but as much as 50 percent is recycled from prior years, so there may be 5,000 hrs/year of new program content. Dave MacCarn reported that at WGBH, they archive 10,000 hours per year, but that this is a mix of finished programs and the raw materials/source footage that is used to produce the programs.

The discussion in the meeting proper was led by Jim Kutzner and Jerry McDonough, with many contributions from others, including the scribe. A number of avenues or facets to the formats and wrappers topic were explored, in a natural but not always in a linear way, which gave the discussion

somewhat meandering quality. It was a meeting of the type that participants will replay in their memories, finding elements to emphasize and organize. Although there were some “next step” items identified at the end, I would also expect that the leaders will winnow their memories and these notes to define sub-themes or sub-topics as they organize future discussions.

McDonough led off with a recap of relevant factors for consideration regarding format selection. He made use of the categories that Caroline Arms and I had put on our LC Web site, and passed around a handout on this topic. The handout also compared a couple of encodings and a couple of wrappers. I am not sure that I would have spun the handout in the same way (he presented Motion JPEG 2000 as an encoding when I think of it was a wrapper) but the handout and its discussion served admirably as an introduction to the problems and issues in preservation. McDonough reinforced this message by talking about what libraries and archives mean by long term: we think in terms of centuries, but we understand that business organizations have a different frame of reference, looking at the shorter term with more intense worries about costs. McDonough talked about the idea of putting content away for the long term and the need to consider transcoding. He also talked about NYU’s examination of moving still images from RGB color space into one of the CIE color spaces (L*A*B or XYZ).

Bruce Jacobs responded with very useful challenging questions, as he did throughout the discussion. “Are we trying for one format?,” he asked. “Maybe for the wrapper. But surely not for the essence: 50 yrs from now there may be a better way to transcode, why transcode at this time?” McDonough replied with remarks on the archivist’s respect for the original artifact, indicating that an archive might well keep both the pre-transcoded as well as the migrated versions of a program.

Encoding and transcoding. Merrill Weiss emerged as a very knowledgeable and wise discussant throughout the meeting. In this early moment, he expressed puzzlement at what he understood McDonough to be proposing. “What if you have video encoded as MPEG-2, at 50 MB/s, I-frame only? I can see how there is sense that JPEG 2000 could do something at a higher quality, but why encode up? That will add artifacts. If you want that higher quality, ought you not go to the original source and encode as JPEG 2000 from the start?” Jacobs added, “Well, you’d have to go the camera.”

McDonough introduced the idea of full-sampled data, what is called 4:4:4 encoding, a term that refers to the sampling applied to luminance and chrominance data and in this discussion it was generally associated with color space. “If your data set is other than 4:4:4,” McDonough said, “you will lose data each time you migrate.” Jacobs said, “Well you lose data from the camera forward, at many steps in the process.” Kutzner said, “We could have the idea of 4:4:4 ‘on the project road map’ but it may not be practical to do anything about this. There are degrees of loss that vary according to what you do. Jacobs noted that even the Cinealta 4:4:4 (is this a camera? editing system?) throws away information, there is always loss.

Weiss said, “I can see how you might want to preserve, not MPEG-2, but baseband video.” Jacobs interjected, “What’s the advantage?” One of the two men noted that you could get back to 4:2:2 from 4:4:4, but you’ll never get any more data there. “Do you want to store 4:4:4 in a your data space, one quarter of this is not used, empty data, do you really want to store it?” McDonough talked about chromaticities and white point information and at this point (or a little earlier) used terms like “putting away” files in a repository (I cannot recall the exact phrase). Jacobs was unconvinced: “Why bother?” He was joined by MacCarn who said that 4:4:4 was not practical. “In our organizations we can’t do this,” MacCarn added, “and we can’t store away where no one touches the content. In our work, we need a practical activity, and we need to recycle our content.” Weiss referred to the idea of a deep archive (which in fact McDonough probably didn’t intend to convey) using electrical circuit terminology: “It is a diode doorway, the content only goes one way.”

Regarding the timing of a transcoding action, MacCarn said, "If I have the source code, I can decode in the future, I can get at the encoded encoded for twenty years or so, and I may have better tools to use in the future." Weiss repeated his earlier question regarding transcoding from 4:2:2 to 4:4:4: "Why manipulate the color space? Why not just preserve the information about the white points?" McDonough talked about the importance of getting content out of lossy compression, and said that as an archivist, he wants to get into a single format, for ease of management.

The discussion of color space, digital sampling, and transformations went on. McDonough talked about an archive's migration of bitstreams over time. "You cannot do bit for bit copying successfully," he said, "unless you have the widest possible colorspace, e.g., 4:4:4." Weiss said "Well, if you have an 8-bit input, you will have to store as 16-bit in order to protect you wider colorspace. By doing a color space transform you are creating artifacting, unless you have enough bit depth."

McDonough responded by showing images of some test charts he had (the ones that were shown at the 2003 DLF Forum in Albuquerque), that showed the deterioration after a number of migrations. Weiss asked follow-up questions, saying he was not sure he understood what he was being shown. "Did you start with an SDI pixel matrix?," he asked. McDonough explained that these were movements of content to and from RGB color space. "Oh," Weiss said, "then what you are seeing is filter effects . . . I guess that is what you would expect here." McDonough said, "We went from 4:4:4 to 4:2:2 to 4:4:4 again." Weiss said, "You are seeing the effects of low pass filters and decimation. You will see effects like this. They are not phase coherent, and you would have that effect." McDonough said, "That is why I don't want to rely on 4:2:2." Weiss said, "Well, you would convert once to 4:2:2, and leave it in 4:2:2." McDonough argued that these color space changes would be needed over time. MacCarn asked, "Why change the encoding when you migrate?"

This led to an exchange of ideas on migration and migration triggers. McDonough noted that physical media are not important in this kind of digital environment. Weiss noted that MPEG-2 has limited lifetime, and so will its follow-on formats. That is why he was drawn to the idea of a picture-based, baseband structure, like CCIR601. "Save the picture matrix," he said, "that makes sense." McDonough called attention to the low cost of storage. David Falond (spelling) said something about DigiBeta and making copies without loss. Regarding Weiss's comments about CCIR601 streams, I asked about whether this was what happened with the uncompressed video stream in MXF (SMPTE 384M), and Weiss agreed. Kutzner noted that MPEG-2 is growing obsolete, and that we are caught in a tension between the near and far term. We are soon going to see MPEG-4 come into play.

MacCarn asserted that he has to come from the practical side. "Indeed in the future we can store everything, but today it is the acquisition of material that we have to deal with. The creators and producers are focused on making programs. Today virtually all of them create in MPEG-2 and DV, this is what comes out of the cameras, that is the acquisition format, practically that is what we have to deal with." Kutzner asked, "What ought we ask for, from the manufacturers?" Someone said that they would not listen to us. Bea Morse remembered that the industry had listened to public tv on multicasting, an idea that had not yet struck the commercial broadcasters.

McDonough noted the importance of the Digital Cinema Initiative (DCI). "This will drive things, and I bet that the next generation for TV will follow that path, and that we will see people go to 4:4:4." MacCarn wasn't so sure, pointing out that "three to five years ago, we thought we would go for HD, and that we would be using 1.5 GB/s recorders. But now the cameras do HD compressed at from 150 to 600 MB/s, the outputs turn out to be far smaller."

Acquisition. This launched a bit of a sub-discussion on what the industry calls acquisition, what the Library of Congress life cycle analysis calls *get* and *produce*. It refers to getting the footage in the first

place, generally via a camera. It is roughly equivalent to what the motion picture industry calls *production*. The underlying question from MacCarn and others is “why archive at higher quality than your acquisition quality” which MacCarn also calls *acquisition bandwidth*.

Jerry Butler talked about the new SONY SR recorders, that I think he said had been developed with the [old] digital cinema standard in mind, and offering higher resolution. “The SR system uses MPEG-4 to accomplish this. It works at 440 MB/s, higher than HDCAM, but moved to MPEG-4 for acquisition. I don’t understand why you are talking so much about 50 MB/s I-frame, SONY jumped on this once but what is special about that?” Weiss reported, “The 50 MB/s idea came from the SMPTE-EBU task force, we looked at 25 and 50 MB/s, we looked at MPEG and Panasonic’s DV [DVCPRO, I think]. In both cases, if you worked at 50, you did not have bad quality decreases. You could go thru seven generations, which is taken to be what happens in a typical production process. We made quality judgments using a CCIR impairment scale. This was about EFP level production.” [Does that mean “electronic field production” or something, i.e., as for newsgathering?]

Butler said, “I work with acquisition bandwidth, I don’t want to lose anything. I choose a format for distribution, not for archiving.” McDonough asked, “What is the highest quality being produced today?” Weiss replied, “1.5 GB/s for HD, actually 1.485, just to be exact.” Butler added, “As a practical matter, most is lower, in typical work, the highest acquisitions I see are 135 MB/s for HDCAM, at PBS we might get 270 MB/s data sent to us but it was originally acquired at 100 MB/s.” Weiss asked, “Does the SONY SR machine do 600 MB/s?”

MacCarn said that at WGBH, “We record at the levels that the medium gives us, limits are partly from capacity of recording media. The capacity forces everyone to lower bit rates than they might otherwise choose. I wonder if we are at a dip right now, and that we will see data rates go up when the acquisition equipments offers better storage. But acquisition is what we deal with: we have only a few formats, really, MPEG-2 and DV.” I asked about normalizing, do you bring streams into a common format. Butler assented that one would, and MacCarn said, “We keep the source material, as well as doing some normalizing.”

Sound/audio. At this point, Kutzner suggest that the group move to next steps. But Butler wanted to ask about sound, noting that McDonough’s handout had really been limited to picture information.

“Should we take audio back to baseband,” Butler asked, “Or, say, keep it in Dolby?” I said that I had just been reading the SMPTE specs for what amounted to expansions of the payload using the AES3 interface. I asked MacCarn what he had at WGBH, what got archived there, and he said that their soundtrack and related audio was “all over the map, and was not normalized.” Weiss said that the majority of stuff produced is uncompressed [and is not in one of those special modified AES3 categories]. Butler pointed out that in the acquisition mode, most sound was limited to two channel (generally stereo) and not matrixed audio. Matrixed audio was created in post-production (editing). SMPTE saw that limits in recording devices forced the use of the AES3 interface for sound and this led to the SMPTE expansions that permit matrixed audio to be carried where normally stereo is the limit.

Wrappers. MacCarn then returned to the **next steps** question, and highlighted the possibility that wrappers (as distinct from *encodings*) might be the best starting point for the next discussions.

“We need the wrapper discussion,” MacCarn said, “maybe this is a good place to start, maybe easier to reach consensus than the encoding thing.” I said that I agreed, noting that our discussion of encodings, part of which had to do with deferring transcoding migrations, highlighted the importance of the metadata that documented the encodings employed. This important metadata would most likely be carried in the wrapper.

Metadata (meaning tech-administrative metadata) **and metadata encoding.** I asked about the SMPTE metadata as represented in the organization's recommended practice RP210. This is the way metadata is referenced or incorporated in the MXF wrapper. Weiss said that the SMPTE: metadata dictionary work is ongoing, and is carried out by his committee. He highlighted some very interesting developments (!) now being worked on by the committee. These included a web services approach for the dictionary, I think this meant for getting from the code to definition, but it may be broader than that. He also talked about the need to update the dictionary more readily (was there a web services side to this?) and the concept of dynamic documents, intended to shorten the committee process for adding entries, trying to get it down to 90 days. As an example of the web service at work, Weiss said, "Suppose you get an edit decision list, has a new item, the EDL system doesn't recognize this item, it can automatically go to the data dictionary web service, and bring the needed information back to the system in a manner that is invisible to the operator." McDonough asked about the relationship of web services to the use of XML. Weiss said that the dictionary is (soon will be?) expressing its content both as KLV and XML.

MacCarn asked: "I'm a little confused about the dictionary, you can put fields of information, but is it references to external namespaces? How to I reference the dictionary, does it give me the namespace, is there a pointer?" Weiss explained that classes 13 or 14 in the dictionary are about information developed by outside organizations, i.e., not part of the SMPTE process. If I understood him correctly, the definitions are based on the SMPTE universal label (defined in an SMPTE standard), this label is used to identify the K in the KLV (key-length-value method for encoding metadata) construct. "If an organization wants to register a structure in class 13," Weiss said, "it has to publish everything about the term, but it is a part of the dictionary they have to manage. For example, the BBC is doing this." Class 14 definitions, Weiss said are ones that a non-SMPTE organization will manage but not publish, thus dark data is possible. He added something like this: "One level down they have to say what the node is but they do not have to give the leaves."

Someone asked about the new public broadcasting metadata structure, called PB Core (related to Dublin Core). Weiss said that he thought that PB Core ought to be integrated into another space within the SMPTE metadata set, because of the way PB Core is defined. Somehow the idea of synonyms came up, and Weiss said that the SMPTE committee tries to avoid them, but there is a column added to the matrix [for synonyms], perhaps we could do this with PB Core.

McDonough noted that in other communities, like information science and computing, XML is preferred as a means for encoding. Weiss said, "Well, have a party," meaning that interested persons should help themselves to work up alternate or translated encodings. [Not sure, but Weiss may have been thinking not only of the SMPTE data dictionary but also implementations like the MXF wrapper.] "The product [metadata dictionary], based on that old SMPTE-EBU task force," Weiss said, "is strongly oriented to KLV, and is not drawn to XML. AAF and metadata dictionary [did he also say MXF?]," Weiss added, "all come out of that task force. But Motion JPEG 2000 is not in that picture, it is not part of how we were moving to file-based structure."

I asked about the preference for KLV over, say, XML. Weiss pointed out that "in a limited bandwidth channel, also in storage but less so, there are limited resources, we have to be efficient, and KLV is efficient, while XML is verbose. In a tape format," he added, "you get two or three lines of active video space, you can only send so much data." "Nevertheless," Weiss said, "there is an effort to express the dictionary both ways, there is a lot of work on the XML side right now." I remarked that in the digital cinema DCI spec, I had been struck by how there was an XML structure for the whole package, kept separate from the essence files, which were in MXF wrappers. I said this reminded me of how we had used METS and also how MPEG-21 is employed.

The meeting was beginning to wind down and Judy Reynolds mused about the high-blown ideas expressed thus far as compared to the production of her WNET series on religion and ethics. She has no connection to a tapeless archiving system. Rather, her team is still shooting analog BetaCam SP tapes, putting the material into an AVID non-linear editing system, broadcasting via playout from the AVID, then dumping a copy of the final program back to videotape.

Next steps. Time was getting short and the following were mentioned:

- McDonough asked for a survey to document what metadata and what formats are being dealt with by producing stations.
- MacCarn promised a conference call on wrappers in 2-3 weeks.
- McDonough asked if he could send NYU's newly hired programmer (who will contribute to the NDIIPP project) to visit WGBH (MacCarn said, "sure.")

* * * * *