BUILDING A MEDIA SUPPLY CHAIN IN THE CLOUD:
A CASE STUDY

J C Derby

Discovery Communications LLC, USA

ABSTRACT

The rapid technical, social, and financial changes in the media industry are forcing all of us to re-examine our content supply chains. A decade ago our supply chains used tape decks and baseband video to create programs for linear television.   Today's broadcast supply chain uses a growing number of incoming media assets to create hundreds of file-based products for dozens of outlets. This case study examines Discovery Networks' transition from a traditional linear supply chain to an automated media factory built in "the cloud." It provides an overview of the tools Discovery uses to receive content from its suppliers using cloud-backed file transfer. This case study also details the architecture and capabilities of the cloud-based media factory that Discovery has built together with our partner, SDVI. It discusses the key benefits of scalability and cost control within Discovery's media factory and some of the media industry's obstacles that prevent it from fully embracing cloud workflows. This case study shares many lessons we've learned since launching the platform in the third quarter of 2016.

INTRODUCTION – THE NEED FOR A NEW GLOBAL SUPPLY CHAIN

In 2015, Discovery's supply chains needed an overhaul. Many of our facilities still relied on tape-based workflows and physical media delivery.   Some facilities had built their own file delivery workflows using file portals. Our US facility had standardized around file delivery on LTFS-formatted LTO data cartridges (Linear Tape File System / Linear Tape-Open), but other regions were hesitant to adopt that standard. LTO's complexities taxed our smaller suppliers. These smaller suppliers often struggled to make deliverables that met our standards.

Our different regional facilities each maintained their own standards for file delivery, and some accepted a wide variety of incoming file types. Some wanted suppliers to deliver files that were broadcast-ready, while others wanted master-quality files in each program's original frame rate. Facilities often changed their requirements to accommodate different suppliers, evaluating media files on a case by case basis. We wanted to build a global supply chain and a unified front door for content. These regional differences in standards and workflow were a major obstacle.

Discovery's supply chain is large, varied, and global. We take delivery of over 12,000 hours of programming and elements from over 600 different suppliers each year. These suppliers are a mix of large production conglomerates, independent producers, and program distributors. Suppliers might be major television studios with a staff dedicated to mastering and fulfilment. They might also be small producers with minimal experience of delivering files to broadcasters. Our business models depend on this supplier diversity, but the differences in technical capabilities and experience among suppliers meant that many deliveries arrived with technical errors. Suppliers often struggled to understand how their deliveries failed to meet our specifications and often couldn't understand how to fix them.

DESIGNING FOR THE FUTURE

Discovery needed a supply chain platform built for the future. This new platform needed to support four primary requirements.

- The platform had to support our global community of suppliers and their diverse levels of experience;

- The platform needed to leverage our global scale, consolidating resources and using a follow-the-sun model to maximize resource utilization;

- The platform's software and hardware resources had to scale to meet the demands of workflow peaks without requiring large capital investments;

- The platform had to meet the disparate needs and requirements of our global facilities, providing each facility with the types of files they needed and tailoring the file acceptance process to every facility's unique requirements.

BUILDING THE SOLUTION – DISCOVERY'S PROJECT ON RAMP

We began building the new delivery platform in mid-2015. Our design team quickly realized that the supply chain's size, reach, and variety could only be managed using a cloud-based solution. The design team chose SDVI's Rally supply chain product to manage the back end of the workflow, while our internal development team handled the front-end client interface.

The Discovery Producer's Portal – The UI for Cloud Delivery

We chose to build the user interface for the new cloud supply chain on top of the Discovery Producer's Portal, an existing application that our suppliers already used to deliver scripts, waivers, and budgets. Our internal development teams wanted to add the new media file delivery capabilities to this existing platform, creating a unified delivery experience for our suppliers.

We overcame several challenges as we worked to build these new features. First, we had to find a way to embed a secure transfer engine into the existing application. Our security policies require an obfuscated endpoint, encryption during transmission, and encryption at rest within the cloud storage. We chose Signiant Flight and the Signiant Transfer API to handle the transfer from the supplier to the cloud. The Transfer API allowed us to build custom file renaming schemes, custom upload queuing behavior, and a robust retry system.

Figure 1 - The Discovery Producer's Portal

We also built features to address issues that plagued previous workflows: metadata validation, multiple types of file deliverables, and inconsistency in file naming. The Producer's Portal pulled data from Discovery's content lifecycle database, populating the list of required deliverables for each programming deal based on the deal's contract. This made it easy for each supplier to quickly see what types of deliverable files (programs, textless material, bonus material, etc.) they owed us. Suppliers could choose their deliverable from the list, assigning files to each deliverable and entering metadata. The developers built form-based metadata entry screens that tailored the required metadata for each type of deliverable, ensuring that suppliers never had to enter metadata that was irrelevant or redundant. The metadata entry screens also validated supplier data, checking the integrity of segment times and time code markers. Finally, the Producer's portal eliminated all the problems caused by file name confusion. Suppliers could upload files with any name they pleased, with the Producer's Portal automatically changing the name during the upload to the "correct" name for each delivery. The Portal referred to the file by its original name when communicating with suppliers but provided its correct name to Discovery systems and employees.

SDVI Rally – The Supply Chain Platform for The Cloud

The Producer's Portal provided our suppliers with a front door for cloud delivery, but SDVI's Rally platform formed the core of the new cloud supply chain. The Rally platform provided us with a flexible, scalable, and reliable solution for processing media files. At the beginning of the workflow the platform inspected the wrapper of each file, evaluating it against our requirements. If a file made it past this initial "file sniff," it entered the main

acceptance workflow. In the main workflow one or more QC software suites analyzed each file. The platform created multiple proxies for every file and transcoded files to the format needed by each Discovery facility. At the end of the workflow the platform delivered each file to delivery buckets for local Discovery facilities and sent out messages that checked the files into local facility MAMs (Media Asset Manager). Those MAMs could continue to drive workflow in the cloud or download the files for on-premises editing. As we expanded our cloud deployment to include playout and digital fulfillment we expected more and more media files to complete their entire workflow in the cloud, never reaching our on-premises storage.

Figure 2 - SDVI Rally Workflow Monitoring View

Rally proved to be an ideal platform for this complex acceptance workflow. It allowed each step in a workflow to contribute more information to the file's metadata, allowing later steps in the workflow to tailor themselves automatically to each file's properties. The platform used the powerful SDVI Evaluate language to allow Discovery to design this "smart" workflow. Evaluate is based on the Python general purpose programming language, but adds dozens of custom functions purpose-built for media workflows. These custom functions allowed relatively inexperienced Python coders to quickly build workflow steps that evaluated each media file's properties and the results of previous workflows.

Discovery used Evaluate to build up several types of workflow handlers. The workflow began with file classifiers and evaluators. The classifiers used the metadata in each file's wrapper to assign that file a "file type." The file type (DnX HD MXF, XDCAM HD MXF, etc.) became a critical piece of metadata that steered the file through the remainder of the workflow. Once a file was classified into a type it was evaluated against the requirements and standards for that type and checked for problematic file conditions.

Figure 3 - An SDVI Evaluate Preset in the Discovery workflow

We also used Evaluate to build workflow control scripts called Dynamic Launchers. Prior to each "heavy lifter" in the workflow (QC software, transcoder, or delivery engine) the Dynamic Launchers evaluated the file's properties and the results of previous steps in the workflow. The launchers used the file's class, deliverable type, ultimate delivery location, and the results of previous QC analysis to determine whether that file needed to go through the workflow step controlled by the launcher. If the file needed to use that workflow step the launcher used that same information to feed the correct preset to the cloud computing instance running the "heavy lifter" software. In many cases the Dynamic Launchers chose presets from among dozens of possible choices. Using Dynamic Launchers allowed us to simplify a workflow with hundreds of possible branches into a workflow that appeared almost linear.

Evaluate and SDVI's "bundle" features also allowed many parts of the workflow to handle problem files without the need for human intervention. We wrote Evaluate functions that inspected the results of QC software analysis and the files output from transcoders. These "Post-Task Evaluators" checked files and previous workflow steps for signs of trouble. They looked for key values from QC reports, rejecting some files based on those results and routing others to alternative workflows for correction. The Post-Task Evaluators also inspected transcoded outputs for key properties, ensuring that output files matched input files. If transcoded outputs didn't pass inspection the Post-Task Evaluator routed the file to an alternate transcoding product to see if that product would generate a better result.

The platform's integration of open messaging busses such as Amazon SNS (Barr (1)) allowed it to communicate easily with other systems. We used SNS to send messages back to the Producer's Portal, updating suppliers with status information in real time. The system also sent messages to our other content and MAM systems, allowing for

automated check-in and real-time visibility into the supply chain for Discovery staff and for the company's suppliers.

The platform's inherent scalability allowed us to quickly expand and contract software and computing resources on demand. We chose to deploy a pool of permanent software licenses for transcoders and QC tools within the platform. These permanent licenses were supplemented by "on demand" licenses. The platform's job management system assigned permanent licenses (with their lower cost) first, then expanded the resource pool to meet high demand based on our preferences and business rules. Because deliveries for multiple regions entered a common cloud environment, we could use licenses around the clock as suppliers from the Americas, Europe, and Asia delivered their content over the course of each day.

All these tools allowed Discovery to build an automated, flexible, and self-healing workflow. This helped us reduce time and human effort spent on content acceptance. The messaging components also allowed much of the communication with suppliers to move to a self-service model, with Rally passing along messages about issues with files directly to the Producer's Portal.

THE RESULTS

Discovery first deployed the new supply chain platform in June of 2016, bringing a few key suppliers onboard to complete testing on the system. Over the remainder of 2016 we added more suppliers to the system. By November of 2016 we were taking delivery of over 80% of US content and over 50% of European content through the new system. As of April 2017, we have processed over 15,000 media deliverable files through this cloud-based platform. We've onboarded over 250 suppliers to the new platform.

Our suppliers have been generally pleased with the platform. They've appreciated the nearly instant feedback that the "file sniff" provides. They've also appreciated the reduced wait time for human QC evaluations, sped up by integrated QC metadata and automated MAM check-in. And of course, none of them have missed chasing the FedEx truck each day.

The direct connection from the suppliers to the public cloud also removes any concerns about network bottlenecks. Suppliers can upload at whatever speed their connection to Amazon's cloud allows them to maintain. While transfer rates for our supplier community average about 50 Mb/s, we have suppliers who have average transfer rates of over 500 Mb/s.

This new supply chain platform is driven directly by data from our business systems. The platform determines each file's workflow based on what it automatically pulls from those business systems and from analysis steps along the way. This has greatly reduced our reliance on human operators to start, monitor, and guide the workflow. Operators don't interpret metadata, as they would in a traditional MAM. The platform itself interprets the metadata and uses it to make autonomous decisions. This reduces the need for human labor and speeds processing and time to market.

LESSONS LEARNED

Discovery has learned many lessons from our year of operating experience with the cloud-based supply chain. These lessons have pushed the company to improve the product and influenced our future.

File Interchange and Compatibility

Our first year has taught us that there are many wrong ways to create both QuickTime and MXF files. In the early months of the project we continually ran into files from our supplier community that broke our workflows. Our ability to customize the workflow allowed us to improve the handling of these files. In some cases, we strengthened the inspection in our file sniff and automated QC tools, looking for problem conditions and rejecting files. In other cases, we used the Rally workflow scripting to automatically send files through correction workflows or routed problem files to tools that we knew were tolerant of a file's problems. We learned that all workflows needed to have a "failure route" that allowed the work to get rerouted or fail gracefully.

Cloud Licensing

Traditional broadcast software vendors are still struggling to adapt to cloud business models. While most vendors support installing and licensing software on virtual machines, many have difficulty with the concept of metered usage. In many cases those vendors must trust another company to report metered usage, and building that trust takes a good relationship. Some broadcast software vendors struggle to create effective and competitive per-minute pricing, creating pricing models that don't provide good value. If more broadcasters transition to flexible cloud workflow platforms, it may force traditional vendors to adapt more quickly. In a flexible workflow framework like SDVI Rally it is relatively easy to swap one vendor's product for another. This creates some competitive incentive for vendors to provide software that performs well and at competitive prices. When every vendor knows that they are three clicks and a strategically placed hash sign away from a breakup it creates a natural sense of competition. In the future Discovery will only work with software vendors who can provide competitive prices and flexible licenses for on-demand cloud licenses.

Understand Dependencies and Plan For Availability Issues

While public cloud resources generally have high availability, they do break down. Amazon's February 2017 S3 outage (Whitwam (2)) was challenging for many companies, including Discovery. While we didn't lose any data or disrupt our core operations, that event pushed us to ensure that our cloud platforms had sufficient regional diversity to continue working even if major portions of the Amazon ecosystem became unavailable.

It's also important to understand the dependencies of connected systems and vendors. Our upload capabilities were interrupted this year by the failure of PubNub, a common messaging service (PubNub (3)). While we didn't directly use PubNub, tools within our framework relied on its messaging.

THE FUTURE

Discovery continues to expand the scope, reach, and volume of its cloud supply chain. In the next few months we'll be adding automated standards conversion, time compression, and product transcoding to the workflow. Our goal is an automated media factory, where workflows build and modify themselves based on an understanding of the necessary output files and the available ingredients. As the broadcast industry evolves further and further away from its roots in linear television, this type of scalable, flexible, and automated media factory will become the core of any effective content workflow.

REFERENCES

1. Barr, J. (07 April 2010). Introducing the Amazon Simple Notification Service [Web log post]. Retrieved from https://aws.amazon.com/blogs/aws/introducing-the-amazon-simple-notification-service/

2. Whitwam, R. (28 February 2017). Amazon S3 Outage Has Broken A Large Chunk of the Internet. Retrieved from https://www.forbes.com/sites/ryanwhitwam/2017/02/28/amazon-s3-outage-has-broken-a-large-chunk-of-the-internet/#6fa29f7dc467

3. PubNub. (28 February 2017). Errors and Timeouts for History, Channel Groups, and Access Manager [Incident Report for PubNub]. Retrieved from http://status.pubnub.com/incidents/nd5nyf27j38l