Creating a Fully-Automated, Fully-Orchestrated, Complex Media Transformation Solution

PixelStrings™

By Cinnafilm Inc. and Discovery

---

paradigm shift /'perə,dīm shĭft/ *noun*

a fundamental change in approach or underlying assumptions. (Oxford Languages)

This case study evaluates the tangible, positive business impacts experienced by Discovery over the past three years, as a result of a conscious investment into shifting the paradigm of how content is managed, delivered, and financially tracked.

This study is only a small snapshot of the entire process that Discovery has adopted. A string of new technologies, modified application techniques, intelligent machine cooperation, and orchestration, all built to serve the ability to scale both mechanically and fiscally. Cinnafilm has provided components of a very large operational "puzzle," built and commanded by the talented team within Discovery.

Note: *On April 8, 2022, Discovery and AT&T closed their WarnerMedia transaction to become Warner Bros. Discovery. The process, technology, and workflows described in this case study only reflect the legacy Discovery side of the business.*

Some Backstory

Like many media companies, Discovery initially built media workflows around legacy MAM systems and on-prem hardware and software. As the business of media & entertainment evolved to include a huge array of new products and services, Discovery transformed its workflows. Discovery embraced a new media supply chain philosophy focusing on building global workflows hosted on public cloud resources. The flexibility, scalability, and dynamic nature of these new supply chains helped Discovery transform the workflows for its over 500 linear networks and then launch Discovery+.

Discovery, alongside other large media companies with similar philosophies, pushed the industry away from on-prem hardware and perpetual software licenses to new "by the content minute" purchasing of software hosted in the cloud. Discovery's supply chain includes over 50 different services from dozens of vendors, chained together in a flexible orchestration fabric that creates workflows dynamically based on business needs.

Building a Better Mousetrap

"Cloud processing" represents an on-demand resource that networks and content distributors cannot ignore – so long as they have the right orchestration and processing tools to take advantage of it. Otherwise, fully automated cloud processing is just the pipe dream of the CTO or an empty directive from the board to "do something." So while Company A celebrated a handful of initial workflows that are processing in AWS or Azure, Discovery finalized the orchestration details on some of the most difficult delivery workflows a media company could possibly undertake. Simply put, Discovery became a global leader in the concept of driving 100% automated, cloud-based file acceptance, processing, and delivery.

This bold statement is backed up by a fully functional, automated file transformation system that scaled quickly and operated (mostly) as expected. A great example of the power of this new paradigm investment presented itself when Discovery acquired Scripps Networks. The system built by the Discovery Content Systems team, headed by Group Vice President Josh Derby, was going to have the additional stress of six networks' worth of daily video processing added to the queue that outsiders would certainly consider to already be full. Knowing they built a system that could easily scale, they flipped the proverbial switches and the system worked as expected, turning out video to meet the needs of a widely varied distribution model for both the legacy Discovery content and all of the added Scripps content.

While many companies take weeks/months to make decisions on processing, Discovery was able to close data centers and turn off countless racks of technology as their atomic automation simply scoffed at the massive increase in workload and turned out mind-numbing volumes of content that went through complex processing. All while simultaneously raising the quality bar for the new addition to the Discovery family.

Another example of the automation prowess created by this change in approach was when Discovery needed to convert its massive archive with more than 60,000 assets to bring the industry's largest, on-demand, OTT library online for Discovery+. While many companies would think about outsourcing such a daunting task or trying to find downtime in their schedule for processing the library, the automated system simply commandeered more cloud processing resources to achieve the task while the combined Discovery and Scripps daily volumes were delivered as scheduled and on time.

There are three specific technology contributions from Cinnafilm that will be overviewed in this case study. These three components are critical for file delivery to the international audience Discovery reaches on a daily basis and for the maximization of revenue associated with higher-rated assets and series. Those technologies are:

- PixelStrings (Video transformation platform)
- Tachyon (Standards conversion)
- Wormhole (File run-length retiming)

TRANSFORMATION TOOLS

PixelStrings – The Platform

"When we discussed upgrading our RadiantGrid platform to PixelStrings in 2020, we were concerned that Cinnafilm's relative newness to the transcoding market would be of concern. While they had provided solid support on the RadiantGrid technology they had acquired, developing a new transcoding platform was another story, and our delivery chain had grown to depend on the Tachyon standards conversion and Wormhole file run-length retiming options in RadiantGrid. Simply put, we could not afford our implementations of these key features on SDVI's Rally system to take a step backward," stated Josh Derby. *"But not only did the functionality we need transfer to PixelStrings, everything took significant steps forward in audio and image quality. And with PixelStrings, we have been able to achieve higher levels of automation than before."*

PixelStrings is a 64-bit media transformation platform that uses enterprise-grade demuxers, decoders, codecs, and specialty processing image and audio plug-ins to recreate optimized video for new delivery requirements. By providing functionality that would require 2, 3, and sometimes 4 separate renders on other equipment, users can generate mezzanine, archive, or distribution level files in a single render to meet all of their complex delivery requirements. This toolset is a fundamental piece of the total architecture because it empowers Discovery with the ability to create any number of high-quality deliverables on demand, but with the assurance that the image and audio quality will be best in class on output once converted.

Since Discovery's deployment of PixelStrings is 100% cloud-based, it is driven solely through API commands on an AWS virtual machine, orchestrated through SDVI Rally. There are a handful of workflows that are called by Rally, and the intelligence of PixelStrings automatically identifies the source file characteristics through inspection and metadata and then delivers a new file to the new specifications.

On the surface, the statement "handful" might imply limited usage, but it is just the opposite.

> *"PixelStrings only needs to be told what the file needs to become, not what the source is composed of,"* said Adam Weyl, Senior Director of Discovery's New Product Development. *"Our technical specifications for our production partners are wide-ranging, and we know that PixelStrings has been honed to accept just about everything our partners throw at it. We know nothing is fool-proof, and a file or combination is going to be presented that errors out from time to time. But those situations are rare, and Cinnafilm has proven ultra-responsive in both squashing bugs and making adjustments to PixelStrings internal processing to accommodate user submission errors that we weren't aware of."*

Taking a closer look at the comment about just having a "handful of workflows," let's inspect a target of 1080i 25 XDCAM50. For this target, only one workflow is needed for every possible source file combination of resolution, frame rate, scan type, and codec/container. The source can be anything: 1080p 23.976 ProRes, 2160p 29.97 XAVC, or 720p 59.94 DVCPro, but only one workflow is needed for all 1080i 25 XDCAM50 deliveries from those sources. (Full disclosure, it's not just PixelStrings that makes automation at this level possible; the Tachyon plug-in is a big part of why a single workflow is possible for this target. We'll get into that shortly.)
Adam concluded, "Simplifying workflows to just a handful requires intelligent processing, and PixelStrings has expanded on the intelligence that began on the RadiantGrid platform years ago."

On-Premises or In the Cloud

PixelStrings API-driven structure provides flexibility to operate within many different environments effectively, either on-premises or in a completely virtual cloud environment. All functions that can be invoked through the UI are available via API, allowing it to easily be strung together with management or orchestration platforms that also connect to MAM and storage systems. SDVI's Rally is a perfect example of this and why systems like Rally are so important to the entire automation puzzle. Something has to manage the conveyor belts into and out of the PixelStrings platform, and SDVI Rally was paramount for the successful application of PixelStrings within Discovery's defined workflows. It is worth noting to not underestimate the complexity of managing licenses and pricing models from the variety of vendors that contribute to systems like Rally. The team at SDVI has effectively "cracked the licensing nut," and by taming the complexities of license management, they have added tremendously to the overall value proposition of Rally's orchestration capabilities.

Tachyon – The Frame Rate Conversion Power App

Tachyon is Cinnafilm's award-winning standards conversion plug-in for enterprise-grade transcoding solutions that provides intelligent analysis of content to accurately recreate the assets at a different resolution, scan type, or frame rate without affecting runtime. In the use cases for Discovery, it is used daily for the following:

- NTSC/PAL conversions
- 23.976/29.97i conversions
- 23.976/25 conversions
- "Progressify" interlaced content for OTT delivery on Discovery+

(It is important to discuss the basic differences of Tachyon versus typical standards conversion solutions – knowing what separates it will explain why Discovery has relied so heavily upon it for this solution set. We will also discuss the types of settings typically used for these larger automation workflows).

With thousands of files requiring conversion weekly, it is important for Discovery to know that a workflow specifying a 25 interlaced frame rate target is going to achieve that target with every file that is processed with it. There can be no concern if the video essence varies from file to file or even within a single file – it has to be handled correctly every time. This is where the intelligence in Tachyon is most impactful, as it allows Discovery to trust their workflows without hesitation for very complex frame mechanics that are being performed under the hood.

"Tachyon's intelligent analysis of every scene in the file allows us to create fewer workflows that will preserve the original, artistic intent of our production partners while satisfying our varied technical delivery requirements," said Josh Derby. *"When we set the 'Allow' functions in conjunction with the automated motion compensation in the workflow, we know Tachyon is going to use a specific algorithm or method only if it is the best solution for the essence of the video we are converting."* Adam Weyl continued, *"This intelligence gives us confidence that the Tachyon converted content is going to look and play back just like the original every time, no matter the target frame rate. Also, it's fast – Tachyon lives up to its name, providing the performance we require to keep up with our high volumes."*
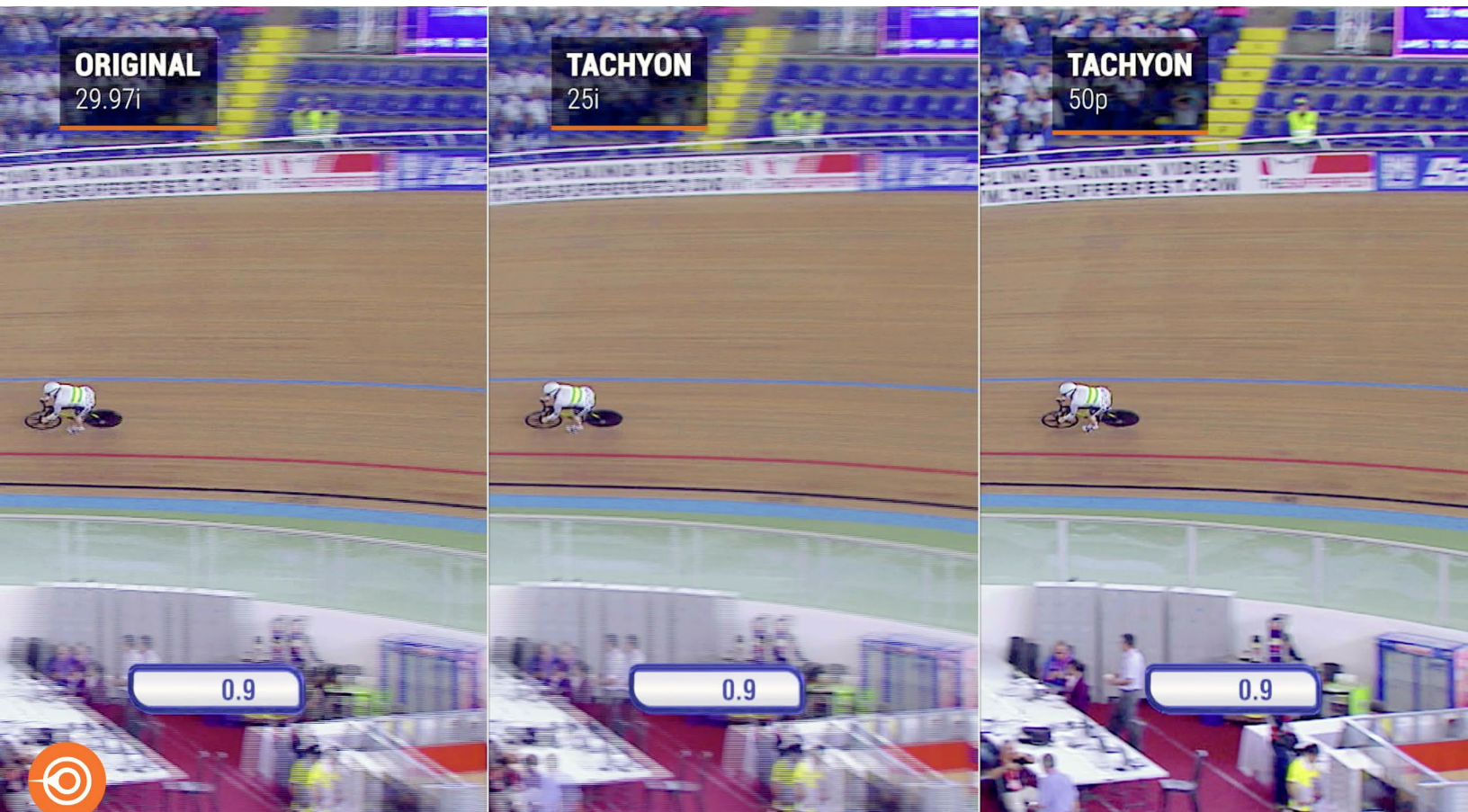
While there are frame drop/repeat or speed-based solutions for performing frame rate conversions, both of those result in an output that is visibly different from the source. Frame drop/repeat causes a stutter in motion, and speed conversion causes motion and audio that is

either faster or slower than the original. If either of these conditions is not of concern, then they are certainly valid solutions. However, if preserving the original intent of the source is required, neither of those two low-technology solutions should be used. In contrast, Tachyon achieves artistic parity between source and target through a fully automated, multi-stage analysis engine when creating a new file. Those stages are as follows:

- Field dominance is analyzed on interlaced content
- Each unique scene in a file is identified, and each unique scene is slated for analysis
- Pulldown/Telecine patterns are detected and removed in each scene, exposing the original video essence
    - Broken patterns are identified
    - Each scene is matched against multiple pattern types to ensure maximum duplicated field/frame removal
    - Options are provided for difficult, composited images where pattern cadence was not considered when the composite was made
- Interlaced scenes are de-interlaced with zero resolution loss
    - Field Rate becomes Frame Rate (59.94 fields becomes 59.94p)
    - SD content is anti-aliased to help ensure horizontal lines remain horizontal
- Progressive scenes are identified
- Content that is already progressive needs no preparation prior to motion compensation

Cinnafilm calls the above stages the "progressification" step, where the video essence is identified scene-by-scene and made progressive through the process of pattern removal or de-interlacing. Once the frames and scenes are wholly progressive, they can be motion-compensated to the target frame rate specified in the workflow.

Motion compensation is at the heart of Tachyon's frame rate conversion. Now in its 10th release, Tachyon's phase correlation-based motion compensation engine has been tuned for varying source video essence rates and automatically adjusts based on source/target rates and the amount of motion present in each scene. Frame generation is performed with the correct amount of reliance on motion data and implementation of source frame fallback. This results in frames that have fluid motion while suppressing artifact generation.

While having an accurate motion compensation engine is critical to generating new frames for the specified frame rate, its value quickly diminishes unless it can keep track of the original, **artistic intent** of every scene within the asset. For example, a 29.97i file can contain a 23.976 video essence that had a Telecine pattern applied to it. If that 23.976 video essence Tachyon extracts is upconverted to 50 frames so it can be interlaced into a 25i container, the original filmic characteristics of the 23.976 would be lost, and the conversion will take on a decidedly "video" look. **The content producers would not be pleased with this result.** This is where Tachyon's Filmic Preservation Options are used.

PixelStrings™

The Filmic Preservation Options in Tachyon are designed to do as their name states and are noted as "Allow" options in the user interface. As mentioned by Josh Derby, when these options are invoked, Tachyon only uses them if the video essence requires them to preserve the original intent of the source file. Using the example in the previous paragraph, if a scene is 29.97i but was originally captured at 23.976 and had Telecine applied when the 23.976 is extracted, that essence needs to maintain the filmic look at whatever the new target frame rate is. For this purpose, Cinnafilm has created Allow settings for all frame rates:

- 48p and 50p fps targets would use "Allow Frame Double"
  - Filmic sources are motion comped to 24 or 25 fps, respectively, and then would be doubled to fit into the 48 or 50 fps container

- 59.94p fps targets would use "Allow 4:6 Insertion"
  - Filmic sources are motion comped to 23.976/24, and then a 4:6 ratio is applied to the frames to achieve 59.94/60p

- 25i and 29.97i fps targets would use "Allow 2:2 Insertion (PsF)"
  - Filmic sources are motion comped to 25p or 29.97p, are then doubled, and the doubled images are placed into the upper and lower field of the interlaced frame – creating a progressive image in an interlaced container (aka Progressive Segmented Frames)

- 29.97i fps targets would use "Allow 2:3 Insertion
  - Filmic sources are motion comped to 23.976p, and they have the standard 5:4 ratio pattern applied to achieve the "3-2" look (3 progressive frames followed by 2 interlaced frames)

Allow options are what make high degrees of automation possible. In fact, all of the Allow options listed above can be turned on for every workflow, and they will only be used when it is duly appropriate. Simply put, there is zero chance the wrong Allow will ever be used. Tachyon's ability to know when to apply filmic preservation algorithms is what permits one workflow to accept such a broad range of input, and it's part of what makes Tachyon the stand-out standards conversion solution for enterprise transcoders.

Wormhole – The Surgical Retiming App

With motion compensation at the heart of Tachyon's conversions that allow new video streams to be created without affecting the run length, Wormhole is the opposite: Allowing frame rate to remain static while allowing a variance in run time. PixelStrings incorporation of Wormhole's retiming capabilities for Audio, Video, and Captions (embedded or sidecar) in a single render **enables content owners/distributors to adjust the run time while not altering the storyline or having to return to edit**.

Broadcasters like Discovery understand the need to adjust run times to meet a certain market's timing requirements. One market may need to shorten an asset by 30 seconds every 30 minutes, while another may need to lengthen a project to meet minimum regulatory run-time

requirements. In either case, Wormhole is the broadcast standard for making those run-time adjustments without having to go back to editorial and without having to wait for the production team to approve the changes.

Wormhole Audio

While Cinnafilm's ability to generate frames from motion data can be used to create a new video stream that runs any % faster or slower than the original, audio is an area for Cinnafilm partner Skywalker Sound, one of the world's most renowned names in audio mastering and processing. Adam Weyl commented:

"In moving to PixelStrings, we were able to get an instant upgrade on Wormhole's audio processing. The Skywalker Sound Tools audio processing in PixelStrings has more sonic accuracy than any other audio pitch correction and retiming solution we have ever used. We don't have to worry about LKFS and True Peak targets like we've had to with other solutions. There aren't any dropped samples and there is never a tunneling effect, phasing, or cross-talk between any of the channels. If we view the new audio program that is generated during the Wormhole process on a scope, the dynamics of the audio correlate directly with the original. Viewers have no idea there has been a retime on the asset because visually and audibly, it looks and sounds like the original asset."

PixelStrings audio capabilities have taken significant strides forward from the limited functionality found in RadiantGrid. Unlimited channel mixing/routing, audio loudness processing, audio pitch correction/retiming, and audio upmixing are provided in PixelStrings by Skywalker Sound Tools, Skywalker Sound's first foray into commercial software.

Scott Levine, Manager of Software Engineering at Skywalker Sound, said:

"When we partnered with Cinnafilm, we wanted to partner with a company that had the same passion for video as we did for audio. When they asked us to provide a retiming solution in addition to our original slate of features, we knew quality audio retiming was not a simple task. In order for it to be believable, Skywalker Sound has identified nearly every aspect of retiming that could cause the result to have a materially different sound than the original source material. We think you'll notice the quality difference in our solution."

When Wormhole jobs are processed in PixelStrings, Skywalker Sound Tools is automatically invoked, and the audio wizardry begins. First, there are a few checks concerning the audio channel mapping/routing:

- If the channel routing is going to be the same as the source file, Skywalker is able to use the loudness characteristics of each channel in the source as a reference. So when the new audio program is created at a different runtime, it understands how to create the audio streams so that they mirror the original loudness characteristics.

- If there is any channel downmixing or replication/omissions, then the render is going to be an entirely new audio composition that does not have a reference. In this situation, the PixelStrings workflow validator will throw an error that a loudness program has to be selected.

While the workflows at Discovery do not call for significant channel mapping/alteration, the flexibility Skywalker Sound Tools presents will be able to handle any compounded audio workflow requirements Discovery may have.

| Assets | Workflows | **Modules** | Jobs | | Transactions | Settings | Support | Operations | Users | Finance |

Skywalker Sound™ Tools   Basic Audio Encoding   Tachyon Wormhole™   Data Burn-in   Cinecert™ IMF

Select Template

| | × ▾ | New | 📋 Duplicate | 🗑 Delete | 💾 Save |

Template Name

Template Description

Audio Templates are selected during job submission. An audio template can be used as many or few times as you would like, based on how specific the source/target layout is and the configuration of your source content.

| ➦ Export Template JSON | ⬇ Import Template JSON |

Define Input Soundfield Groups   ⇥

Linking input groups will disable re-ordering and removing input soundfield groups.

i.1   + add group

Define Output Soundfield Groups

Linking input groups will disable re-ordering, removing ouput soundfield groups and Force Discrete Tracks.

+ add group   **1**   add label (optional)

Wormhole Video and Interface – A Clear Picture of Complex Operations

The same scene-by-scene, analytic intelligence found in Tachyon is present in Wormhole – in fact, it's the same engine, just with different constraints around frame packaging. So, in addition to file retiming, a frame rate or scan type conversion can easily take place in the same render.

What makes Wormhole such an effective retiming solution is the combination of the motion compensation engine and the interface that allows for either global retiming settings that apply the same percentage of adjustment to an entire file or the granularity of user-defined, scene-by-scene runtime adjustments.

Name ⓘ	Source In ⓘ	Out ⓘ	Duration ⓘ	Adj. Method ⓘ	Non-prog. ⓘ	Retime Value	Output In ⓘ
Color bars	00:58:30:00	01:00:00:00	00:01:30:01	No Retime ▾	🟠	00:00:00:00	00:58:30
				-- Select One -- No Retime Auto Shorten Lengthen Target Slug Insertion Exclude from output			
Scene 1	01:00:00:01	01:03:30:00	00:03:30:00		⚪	00:00:10:00	01:00:00
Scene 2	01:03:30:01	01:08:00:00	00:04:30:00		⚪	00:00:14:00	01:03:20

The screenshot above provides a partial view of the Segmented Retiming Mode, which is used by Discovery for the scores of files that are retimed every week. The user interface shows timecodes are entered for each segment, and the retiming method for that segment is selected. Of course, in the fully automated Discovery environment, all of those entries and selections are performed through an API instruction set.

Here is a high-level description of the SDVI Rally automated process:

- Files are delivered to Discovery in components:
  - Muxed audio and video
  - Sidecar Scenarist Closed Caption file
  - Cut sheet annotating the timecodes of color bars/slates, segments, program breaks, and credits
- Assets are automatically logged in Rally and await instruction for processing
- Based on the popularity and distribution needs of the show, files are automatically categorized as needing retiming or not. If a file is categorized as needing retiming:

PixelStrings™

- Overall time adjustment is already known, based on if the file fits into a half-hour or one-hour time slot
- Rally generates a PixelStrings JSON using a pre-built workflow + Skysound audio template + a Wormhole Retime Decision List (RDL)
  - Wormhole RDL was created by incorporating the Cut Sheet provided by the Production Partner
- Rally launches a GPU-enabled AWS virtual machine that has PixelStrings running on it
- Program components and PixelStrings JSON are delivered to the virtual machine
- The virtual machine completes processing and delivers the file to QC for a final check

SUMMARY

Josh Derby closed with this:

"PixelStrings is a critical part of Discovery's content supply chain. It provides the quality, scalability, and flexibility that our company needs to send content to over 200 countries in more than 50 languages. Cinnafilm has taken a major step forward in embracing the benefits of a cloud-based supply chain, and we're happy to continue our partnership."