White Paper

Higher Profitability and Productivity in DI Workflows
SGI® and DataFrameworks™ Provide the Data Management Foundation for Digital Intermediate Workflow

Table of Contents

List of Figures

1.0 Introduction

The process of shooting on film and finishing in the digital realm has been available to the TV production community for many years, but the problems with moving massive amounts of film-resolution data to finish a feature film has meant that film studios have tended to utilize the process only for special effects sequences.

That has all changed - breakthroughs in data management, shared file systems, the price and performance of storage and advancements in film scanning and recording technology have created new workflow opportunities to finish features entirely in the digital realm.

A Digital Intermediate (DI) is the digital form of a timed intermediate created by a traditional chemical film lab. Unlike the film intermediate, various output formats can now be generated faster from the DI master with less quality loss. The DI environment enables a Producer, Director, or Director of Cinematography to sit with a colorist within the DI environment and see changes, interactively, as they are made.

Digitally assembling a complete motion picture presents some challenging data management problems. The digital intermediate process is aimed at transforming all digital image manipulation into an integrated data path in which images, metadata, and settings travel seamlessly through a ubiquitous pipeline from one workflow operation to another. The digital pipeline starts in the earliest days of pre-production and continues to the final moments of the finishing process. The DI process often involves collaboration methodology, both within a facility, and among facilities.

2.0 Overview of the DI Workflow

The DI workflow consists of three main components:

- **Ingest:** The DI process begins with the ingest or capture of film or electronic cinema camera data. An increasing percentage of content is captured and delivered digitally from HD, 2K, and 4K digital cameras. Content is typically delivered to the DI facility as 20 minute “reels” of film, which are then scanned into the system.
- **Image Processing:** Image processing encompasses the creative aspect of the DI workflow. Commercial and facility-proprietary software tools perform a variety of image processing functions:
  - Dust and scratch removal
  - Color correction
  - Editing
  - Compositing
  - Visual effects
  - Grain management
  - Titles and credits
  - Image restoration
  - Incorporating audio
- **Output:** With an edited and graded digital master, it is easy to make the highest quality deliverables for cinema and for other purposes like home video, DVD, airline versions, foreign language distribution, television broadcast, and compressed media such as MPEG and JPEG2000. The majority of feature films are still delivered to theaters on film. However, 2K and 4K digital projectors suitable for theatrical presentation are available today. Digital cinema delivery is rapidly approaching technical and business viability.

A schematic diagram of the DI workflow is presented in Figure 1.

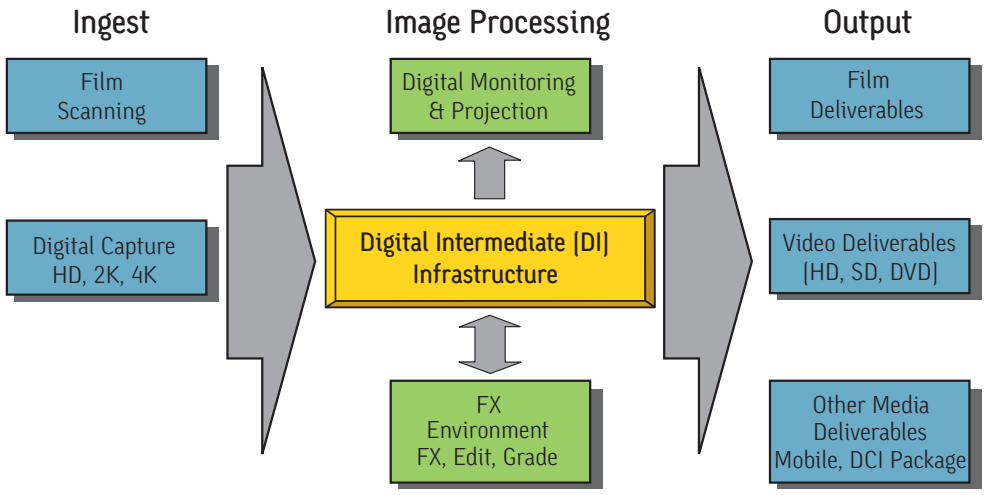

Figure 1. The Digital Intermediate Workflow

Figure 2. SGI® InfiniteStorage 6700 & 10000 and SGI® Altix® Family

2.1 Components of a DI System

A complete DI system consists of:

- Image capture devices (cameras and scanners)
- Computer Hardware Infrastructure (Servers, Storage, Networking, and Desktop systems). As shown in figure 2
- Software Infrastructure (interfaces for data and asset management)
- Creative Applications (software tools for color correction, restoration, editing and special effects)
- Monitoring Tools (computer monitors, digital projectors, audio monitors)
- Output Tools (Film recorders, videotape recorders, DVD burner drive bays, tape duplication, tape backup and archive systems)

This white paper focuses on how significant efficiencies in the DI workflow can be obtained through the hardware and software infrastructure provided by SGI and DataFrameworks.

2.2 Data Management in the DI Environment

The following describes a typical data management process within a DI facility:

- Data Managers assign storage resources to job. Individual artist workstations may have internal or direct attached storage (DAS). Departments or workgroups often have dedicated network attached storage (NAS) device and/or a storage area network (SAN)
- Data Managers are informed by the production team when and what content files to put on and take off the various storage systems based on the work shift requirements
- Data movements and copies are performed via system commands. Software scripts are often developed internally to facilitate the process
- Coordinators inform Data Managers when a project is completed and the data is archived
- Special effects artists usually manage renders and are responsible for associated file clean up activities
- Data Managers use system monitoring tools to scan for "abusive" users

2.3 Data Management Challenges within a DI Environment

Improving Workflow

Workflows that required data copying from system to system between steps are impractical. Content must be equally accessible to a wide variety of system platforms so that workflow is never impeded. For instance, in the course of a project a company might need to share the same media assets

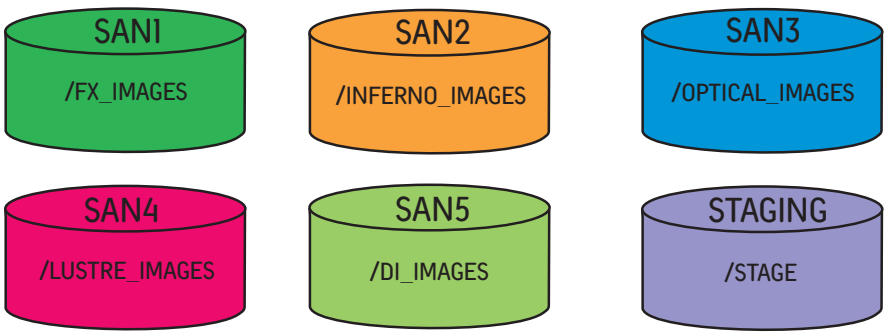

Figure 3. Storage Area Networks (SAN) Allocated by Workflow or Department

between a variety of systems running IRIX®, Linux, Mac® OSX, Window®, etc., and those systems may be in different geographic locations.

Important content must remain online and accessible at a moment's notice to ensure that critical deadlines are met. As far as the DI operator is concerned, content should be exactly where it needs to be, exactly when it is needed. This requires careful coordination of shared resources among multiple projects and creative staff. Traditional visual effects facilities often employed individuals (referred to as Data Wranglers) to manage location, movement, and tracking of data to insure the right images/audio is in the right place at the right time in support of production schedules.

Resource Management

As noted above, many DI Facilities dedicate resources on a physical basis, e.g., one storage system is dedicated to artists. A workflow strategy based on physically dedicating resources ultimately will:

- Reduce overall utilization rates
- Cause unnecessary data movement
- Limit production line flexibility / adaptability
- Result in unnecessary capital investment

Furthermore, a strategy of dividing storage into volumes often involves a lot of human interaction to coordinate what goes where and why. E-mails and phone calls to achieve this coordination can slow down production and introduce error.

Streamlining Management

Data management within the DI system involves data migration: from scanner to DI system, DI system to film recorder, or between multiple dedicated visual effects, editing and color correction systems within the DI environment. Data is typically acquired from a variety of sources and workflows, which all must be assimilated into the DI system. Ideally the visual effects will maintain consistent naming conventions and number of frames identical between the QuickTime/folders with low-resolution sequential frames and the high-resolution versions. Otherwise, reels may not auto conform, confusion may result regarding shot numbers, revisions, etc. and unnecessary hours will be lost or incorrectly billed.

Protecting Valuable Digital Assets

For DI facilities, digitally-stored content is the lifeblood of the company. Ensuring that critical digital assets are protected from user error, theft and disaster is now an essential element of business planning. Customers of DI services typically require stringent internal controls over file system access. Digital Intermediate providers often implement a network topology that physically separates the "production systems" from the "office system" where:

- Production systems typically contain the customers' actual image files
- Office systems typically support the business processes of the DI system

However, individuals not directly involved in the creative aspect often need to communicate with the client and must know information about where specific shots are located, the stage of the project, what resources are available or are in use – without having to disturb the engineers or the artists in the facility. Adding to the challenge, these scheduling personnel are prohibited actual physical access to project image files on these systems.

Accommodating Exploding Data Set Sizes

Across the media industry, organizations are faced not only with a huge store of digital assets, but also tremendous increases in the rate of content creation. Accommodating these trends requires scalable storage systems with the I/O capability necessary to move data to and from storage at guaranteed data rates that can accommodate high-resolution digital formats.

Due to the shear volume and dynamic turnover of data within the DI system, the manual data management task quickly becomes unmanageable and an automated level of data management is required to provide data access and storage utilization information to users.

Eventually the size of the data grows and the manual tasks performed by the team become increasingly unmanageable due to shear size and complexity.

Controlling Spiraling Costs

No matter how important the project, storage costs remains a critical factor. Cost-effective storage systems that meet the needs for a particular part of the workflow should be considered. Storage for playback of 4K files in real-time has very different characteristics than near-line storage repositories. All storage needs to scale to accommodate growing storage requirements without hampering access to digital assets in the process. These critical challenges are driving a widespread need for storage systems with:

- High-speed data sharing
- Wide area data sharing
- Support for heterogeneous filesystems
- Seamless hierarchical storage management
- Improved scalability and performance
- Simplified management of digital assets

3.0 The SGI and DataFrameworks Solution

SGI and DataFrameworks have combined their expertise to deliver the best-of-breed solution for optimizing the DI workflow.

SGI InfiniteStorage solutions deliver comprehensive data management designed to meet critical needs in all media market segment shares, providing cost-effective storage options with shared, fast access to content. SGI provides a unique data-centric storage architecture that addresses the challenges faced by the media industry. Critical digital assets are centralized to improve storage utilization, security and economy all while providing the scalability and performance to meet the needs of the most time-critical projects.

The DataFrameworks software offers a new paradigm for provisioning, monitoring, and reporting on data. DataFrameworks automation and virtualization capabilities reduce the IT workload required to manage file systems and storage. It allows a business-oriented, 'top down' view of corporate data, thereby improving the user's ability to find data and maintain a disciplined, standardized file structure. DataFrameworks provides the vehicle to introduce, systematic data organization, consistency and discipline into the production workflow.

3.1 High-speed, Shared Access

A key element of the solution is the SGI® InfiniteStorage Shared Filesystem CXFS™. CXFS provides high-speed, shared access to eliminate the bottlenecks that hamper data-intensive digital operations. CXFS supports many of the most widely used system platforms, so it works seamlessly in heterogeneous environments.

CXFS avoids bottlenecks associated with manual copying and dramatically streamlines workflow. The ability to concurrently access data without copying delays, significantly improves productivity. No need to move the uncompressed 2K or 4K data around the facility. Multiple workstations, using a wide variety of heterogeneous operating systems, can work sequentially and collaboratively on a single project without the need to copy the data --eliminating a traditional workflow bottleneck. This means that more productive work can be completed in the same amount of time – or larger more complex projects can be undertaken.

Furthermore, DataFrameworks provides the ability to efficiently manage limited production infrastructure resources, for multiple projects where everyone shares the same resources. The software provides the tool set to assist the data management process of managing location, movement, and tracking of data to ensure the right images/audio files are in the right place at the right time. A file folder can be created for each reel of film and subsequent creative processes on that content can be monitored, as shown in Figure 4.

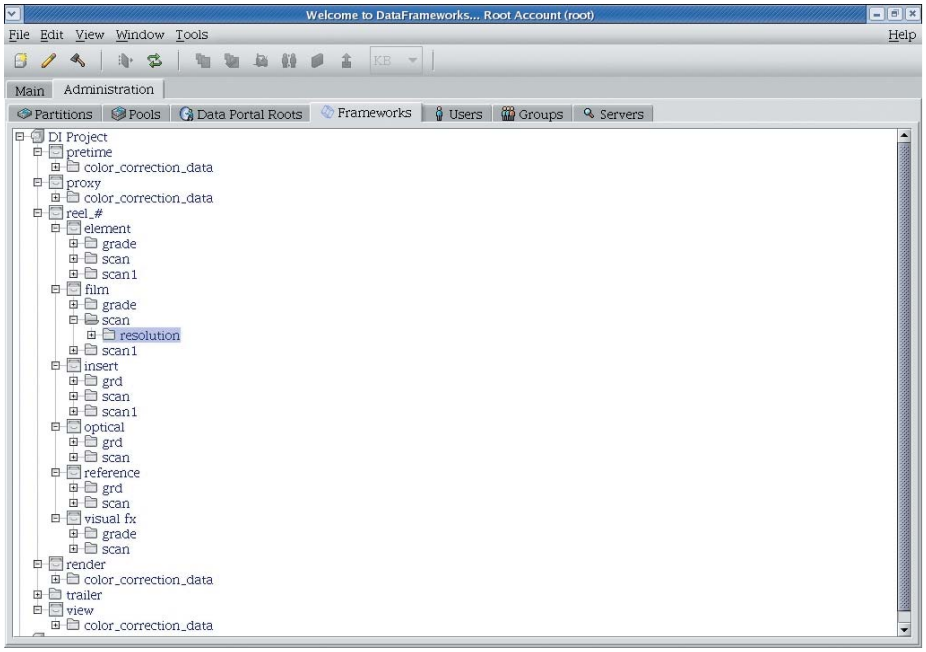

Figure 4. Screen shot example of DataFrameworks Project Management Tool

3.2 Intelligent Consolidation

SGI is in the business of delivering optimal efficiency for storing, organizing, accessing and managing digital assets. The storage architecture is tailored to reduce storage complexity, streamline management, increase performance, improve availability, and reduce TCO. By eliminating the bottlenecks that limit other storage solutions, SGI delivers unparalleled data performance for the creative talents of the DI organization to ensure success.

SGI's unique approach enables all digital assets to be managed from a central, consolidated storage architecture for:

- Better disk utilization
- Improved load balancing
- Zero data replication
- Reduced storage capacity requirements
- Reduced management costs

Intelligent Consolidation give the DI facility the ability to scale its digital infrastructure independently in different dimensions over time, including storage capacity, bandwidth, performance, connectivity, and supported operating systems – providing virtually unlimited growth paths to meet current and future storage needs.

3.3 Resource Management

Traditional NAS and SAN systems do not easily interoperate. As noted above, many organizations maintain separate storage pools for each function. SGI brings NAS and SAN together with its architecture, unifying storage systems to enable efficient workflow, enhance collaboration, and improve data management. Content files can quickly migrate from NAS to SAN storage as bandwidth requirements change along the DI production pipeline.

DataFrameworks controls and monitors disk space consumption, as well as setting up and enforcing directory structures, so growing the storage becomes much less of a risk. DataFrameworks provides historical trending reports to enable companies to make better judgments on when growing storage is necessary in a proactive versus reactive manner. An example is shown in Figure 5.

The DataFrameworks reporting feature shows current disk space consumption and historical disk space consumption trends. While it is easy to understand space consumption on an individual data volume, it is often a burden to figure out what a specific project or department is consuming across multiple data volumes. With DataFrameworks, it is a very simple to track and monitor space consumption from a business perspective, not just a storage device perspective, allowing the DI facility to evaluate space consumption from any level of a business' organization.

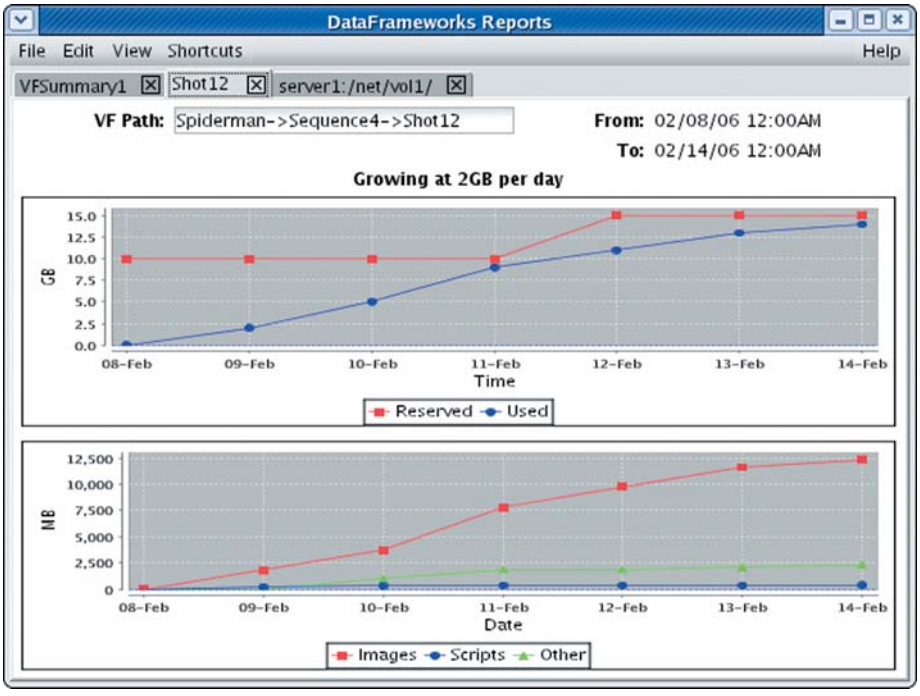

Figure 5. Screen Shot of DataFrameworks Storage Utilization Report

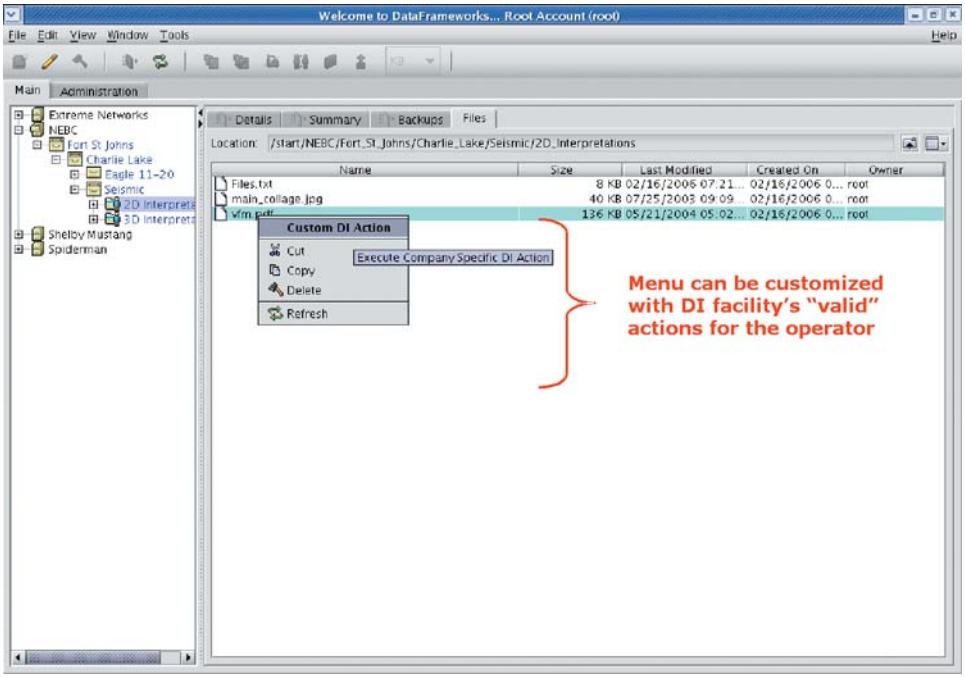

Figure 6: Screen Shot Example of DataFrameworks Customizable Menu

3.4 Automation

DataFrameworks provides the foundation to introduce automation to improve efficiency and minimize human error. Data organization schemes are consistent, predictable, and repeatable. The interface is intuitive, so even non-technical users easily navigate through the production systems. The interface provides features to facilitate troubleshooting operations, and allows a facility to customize workflow to further automate and reduce error. As shown in Figure 6, the DataFrameworks GUI provides many places where a company may fully customize actions to execute on a selected set of data, presenting only the "valid" operations to any given user. These operations typically spawn automated workflow activities.

3.5 Security

DataFrameworks provides reporting tools that can be configured separately from the authorization and access levels. This allows dustomer service representatives to have visibility of the status of the file system data without the ability (or privilege) to actually access the files within the production systems themselves.

3.6 Integration into Existing Infrastructure

Each facility will have different combinations of infrastructure, workflow, and client requirements, and therefore will need a solution that is customizable and open.

DataFrameworks understands the difficulty of providing a "shrink wrap" solution that meets every company's changing workflows and requirements right out of the box, and is designed to provide a proven, stable, core foundation to support rapid customization, development, and deployment.

Every company and project is slightly different and tools need to adapt. DataFrameworks integrates with the a DI facility's existing software and hardware infrastructure components, such as:

- Asset Management Systems
- Batch Processing Systems
- Scanners
- Production Tracking Systems
- Backup Systems

When considering commercial software, DI facilities need to evaluate whether the software is customizable enough to support their workflows. DataFrameworks provides a flexible architecture with the necessary interfaces to incorporate a company's unique business logic.

As illustrated in Figure 7, DataFrameworks provides the following interfaces that provides a stable, flexible foundation, allowing facilities to integrate with existing infrastructure:

- **XML-RPC Interface:** DataFrameworks provides an XML-RPC interface that gives other applications access to the services and functionality provided by DataFrameworks.
- **Command Line Interface:** The DataFrameworks command line interface provides all the common functions of the system in an efficient and easy-to-learn way across all supported hardware platforms.
- **Customizable Graphical User Interface:** The DataFrameworks GUI provides many places where a company may fully customize actions to execute on a selected set of data, depending on their staff and customers' unique requirements.

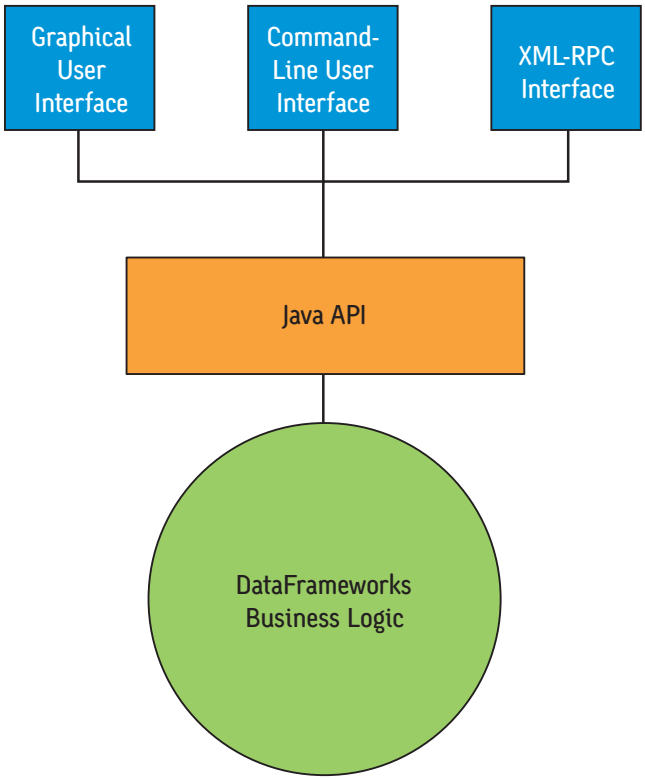

Figure 7: DataFrameworks Interface Schema and Underlying Technology

4.0 New Business Opportunities through Data Management: Digital Dailies

DataFrameworks provides a solid underlying architecture enabling a company to incorporate additional functionality. Because it is customizable, DataFrameworks can easily become the central hub for data management for almost any kind of workflow, such as Digital Dailies.

Daily reviews (or Dailies) of the camera footage are an important aspect of film production. Dailies provide the production team the ability to review the raw materials which provide the foundation the quality of the final product. Dailies provide the ability to check for errors and plan the next day's shoot. When images are captured via film cameras there are two methods to review dailies, 1) printing to film and projecting the dailies from film prints, or 2) generate digital dailies and reviewing from digital projector or monitor. Many DI facilities are pursuing a data-centric workflow from pre-production to post and are becoming involved in digital dailies and introducing other new services into their pipelines which create a data-centric workflow from the start of a film production.

Digital Dailies will introduce additional data management challenges. Within the DI process the data is often completed within one week to two months away from the movie's release, but for Digital Dailies process, the data generated on a daily basis is constantly evolving and growing. The entire Digital Dailies process might be a month or a year away from the final project's release, this extended life span of data, increases the difficulty and expenses associated with data management. Organized constructs and automation will be a requirement.

5.0 Conclusion

The SGI InfiniteStorage CXFS shared filesystem and storage hardware coupled with DataFrameworks file system and storage management software enables higher productivity for the creative and production staff and greater utilization of the physical assets within your facility. A leading DI facility reported 25% improved efficiency as a result of implementing the SGI and DataFrameworks solution.

"Our industry applies art to technology. We need something to take the technology out of technology, and provide our artist with a very simple view, instead of requiring them to learn the "/dev/dsk1" disk context up front." DataFrameworks allows us to put a project language interface, as opposed to a disk language interface, into the process of storing and finding data". – Leading Digital Intermediate Company

The combined solution also makes it easier for schedulers and customer service representatives to accurately track the project in-process and report to the customer without having to impact the engineering resources. Project producers are now able to better manage and track their data storage usage and the facility owner now has more accurate insight into their storage resources per project.

As film production continues to take advantage of digital tools and processes, the time between the completion of the DI project and the delivery of the content to the cinemas is only getting tighter. As more facilities deal with multiple secure 2K and 4K projects the demand on your resources becomes even greater. From managing hundreds of terabytes of content data, sharing files across multiple applications and systems, there is a demand for better management. Being able to consolidate your data, more accurately track your resources and increase efficiencies throughout the DI process brings a higher degree of knowledge, security and greater profitability to industry leading DI facilities. By combining the cutting edge technology of SGI InfiniteStorage and DataFrameworks' workflow solution, you can be assured that you are increasing productivity and lowering costs.

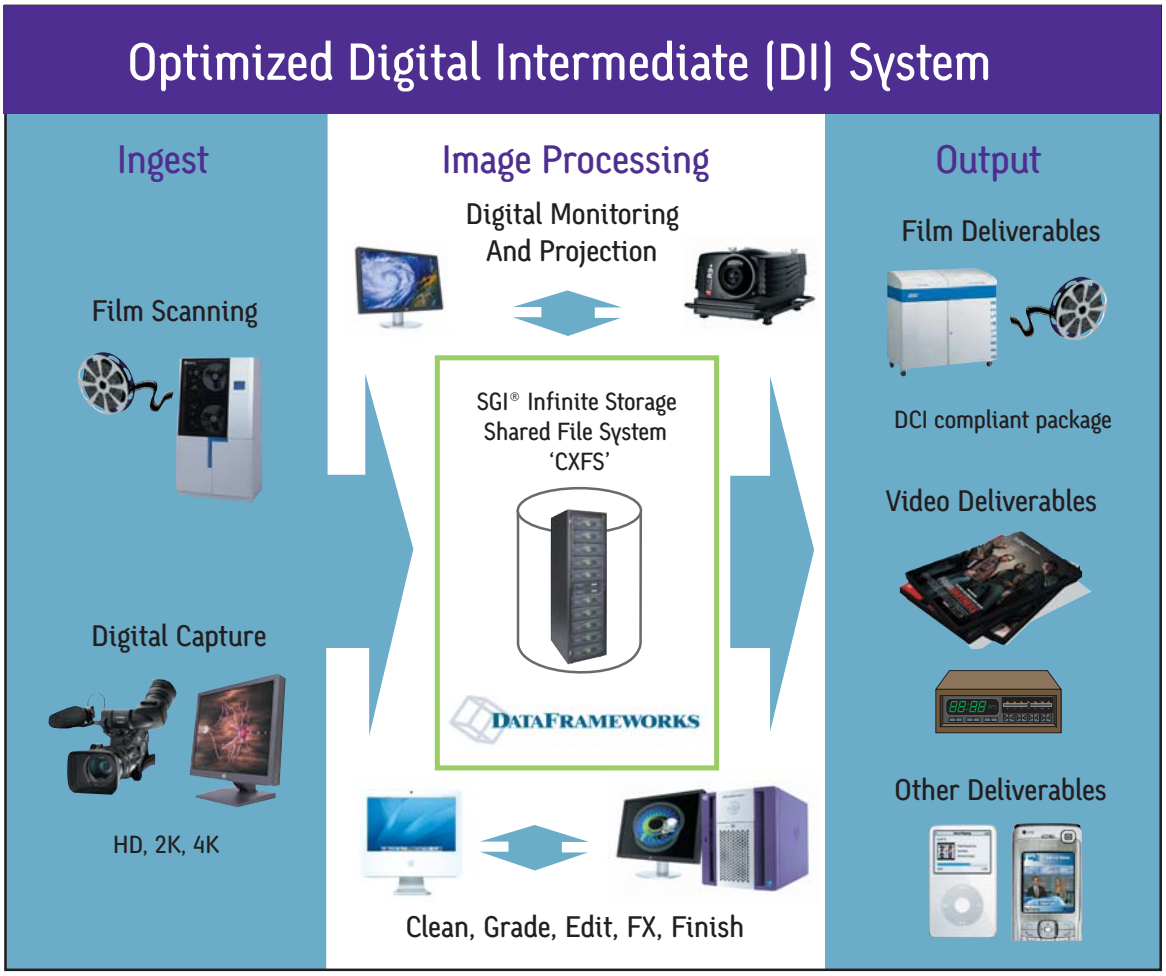

Figure 8. Optimized DI System

Corporate Office
1500 Crittenden Lane
Mountain View, CA 94043
(650) 960-1980
www.sgi.com

North America +1 800.800.7441
Latin America +55 11.5509.1455
Europe +44 118.912.7500
Japan +81 3.5488.1811
Asia Pacific +1 650.933.3000

3927 [04.07.2006]

J15144