2030 VISION SERIES

THE EVOLUTION OF PRODUCTION SECURITY

Securing the 10-Year Vision for the Future of Media Production, Post and Creative Technologies

movie labs®

Motion Picture Laboratories, Inc. (MovieLabs) is a nonprofit technology research lab jointly run by Paramount Pictures, Sony Pictures Entertainment, Universal Studios, Walt Disney Pictures and Television and Warner Bros. Entertainment.

MovieLabs enables member studios to work together to evaluate new technologies and helps the industry develop next-generation content experiences for consumers, reduce costs, enhance security and improve workflows through advanced technologies.

CONTENTS

EXECUTIVE SUMMARY

SECURING PRODUCTION IN THE CLOUD

In a white paper published earlier this year titled "The Evolution of Media Creation" (referred to here as the "2030 Vision Paper"), MovieLabs and its member studios have laid out a bold vision for the future of filmmaking some 10 years out, with a call to action for the industry to collaborate appropriately to achieve our shared goals. The goals envisioned in the 2030 Vision Paper focus on requirements for implementing true cloud-native production workflows.

All workflows transcend organizations, and simply stated, protecting cloud production requires a new approach to security.

Our definition of security has two parts. The first is protection from malicious and unauthorized activity, which is the primary goal of information security. The second part is protection of the integrity of data, workflows, applications, and processes. For example, protecting the integrity of a workflow includes preventing the introduction of unauthorized applications or an unauthorized change to a workflow.

This document presents six primary security principles, which are the foundation of the security architecture required for the 2030 Vision Paper:

1. Security is intrinsic to every component of every workflow and does not inhibit the creative process.
2. The security architecture addresses requirements specific to cloud-based workflows.
3. Production workflows, processes, and assets are kept secure, even on untrusted infrastructure.
4. Content owners control security and workflow integrity.

5. Security can be scaled to appropriate levels and can integrate with existing security policy and management systems.
6. The security architecture limits the spread of any breach and adapts to the evolving threat and response landscape.

The first principle forms the cornerstone, encompassing the cybersecurity industry's dictum of *security by design* as well as the importance that secure cloud workflows enhance the creative process. Usability is as important in the security model as the security itself.

This white paper explains the principles outlined above and articulates practical ways to achieve each of them. As noted in the 2030 Vision Paper, MovieLabs recognizes that defining the security foundations for the 2030 Vision will require all stakeholders to work together constructively to design a solution. The common goal will be a core security architecture that builds on the principles expressed above and meets the requirements of the industry. For the new security architecture to be effective and integrated by design, this initial work must be completed and ready for implementation ahead of significant adoption of the principles in the 2030 Vision Paper.

SECTION 1
INTRODUCTION

The 2030 Vision Paper foresees that within 10 years, and likely a lot sooner, all assets will be stored in the cloud, and all processing of those assets will run in the cloud. Here, processing includes both services such as encoding and applications such as craft edit tools, with the latter running on virtual workstations. The cloud enables this processing to be combined into software-defined workflows.

The security architecture has two purposes, and they constitute the definition of security:

- Protection from malicious and unauthorized activity, such as the exfiltration of assets.
- Protection of the integrity of data, workflows, applications, and processes.

The objects that need to be protected fall into three categories:

1. Assets: Data and metadata that are created, processed, and output.
2. Processes: Software services and user-interacting applications that process assets (including automated tasks such as AI/ML processing).
3. Workflows: Orchestrated sets of processes acting on a set of assets.

The security threat to production is not solely the theft of assets. Simply protecting assets is not enough. Processes must be protected to ensure that their function is not subverted or their output redirected. Workflows must be protected so that processes are orchestrated as intended.

These cause us to look for a security architecture that protects assets, processes, and workflows.

Production in the cloud means that production workflows are abstracted from any facility's infrastructure and happen in the cloud. In Section 1: Challenges, we show that, from the security point of view, there is a distinction between *production in the cloud* and a facility using cloud resources *as part of its infrastructure*.

This document presents the application of Zero Trust architectures to production workflows. While this can leverage some existing tools and practices, it requires significant new ones. This paper introduces two concepts: encryption key management services (that manage the creation, secure storage and distribution of encryption keys) and authorized applications. It describes the role each fulfills and how they integrate with identity management.

The policies for these services can be configured with as much granularity as is required; they do not need to be global or system-wide. Each service is under the control of one entity, usually the content owner, and can be used for workflow security as required.

The document outlines the full extent of the security architecture, but the design is modular and can be applied as needed. The goal is for this security architecture to be both scalable and renewable. Modularity improves both. For scalability, the content owner has the freedom to decide where and when in the production workflow this security architecture is applied, what level of security is used, and at what granularity. The propagation of assets to the next stage of processing or review is controlled by the security system, and the decision to "publish" for review or further processing can be made by those working on those assets.

What we present in this document establishes the underlying principles of the architecture to secure production in the cloud. The architecture is designed to be durable and tailored to support cloud production. Furthermore, the principles of the 2030 Vision Paper and the security principles presented complement each other.

It is important to know that there is nothing in this document that requires a leap in security technology. The main components of the architecture can be drawn from emerging enterprise cybersecurity systems such as Zero Trust architectures and from entertainment industry technologies where the asset protection mechanism is part of the mature technology of digital rights management (DRM) systems used in consumer distribution.

While discussions of Zero Trust architecture often focus on its application to networks, its full application goes far beyond that. A Zero Trust architecture moves the security effort from securing the perimeter with firewalls, VPNs and web gateways to verification

and authorization, endpoint security and traffic inspection, and logging and constant risk assessment. The security is embedded throughout the system, in every piece of hardware and software and many of security best practices such as patching vulnerable systems are every bit as critical.

Securing the hardware, software and displays at the endpoints can leverage and build upon existing security practices for facilities and physical access. The security controls and assessment processes defined by entities such as the MPA and the Trusted Partner Network will continue to play an essential role in ensuring that best practices are followed for securing facilities and, as those practices are extended, for securing these new cloud-based workflows.

TERMINOLOGY

The following terms used throughout this document warrant definition of their usage:

Applications and Processes	An application is a software program that can perform one or many operations on an asset, including programs with a user interface. A process is one or more applications that work in combination to perform a distinct part of a workflow.
Asset	An asset is digital file containing data that is part of the process of producing content. It might be a single frame from a camera, an audio file, a script in digital form, the metadata describing some part of a piece of content, etc.
Authentication	The process of confirming the identity of a person or system.
Authorization	The process of determining whether an action is permitted.
Cloud	For the purposes of this document, cloud refers to internet addressable computation and storage resources. Those resources can be in a public or a private cloud, including on-premises, or a combination thereof.
Content Owner	The entity that has overall responsibility for the content being produced. It does not imply any rights ownership.
Enterprise	Any commercial organization, not necessarily a media company, with its own IT infrastructure.
Facility	An organization with its own IT infrastructure contributing to a production. This includes the studio, a postproduction house, a VFX company, etc.

Hybrid Cloud	For the purposes of this document, a hybrid cloud is a facility infrastructure that uses both on-premises and cloud services, where the cloud services are managed by and used exclusively by the facility. The cloud services are thus an extension of the on-premises infrastructure.
Identity and Access Management (IAM)	The system and its policies that identify, authenticate, and authorize individuals requiring access to the system.
(IT) Infrastructure:	A collection of workstations, servers, network attached storage, local and wide area data communications, network routers, firewalls, and so on. It may be on-premises or in a private data center.
Production in the Cloud	This term is used to describe the use of cloud platforms in production as envisaged in the 2030 Vision Paper. It is distinct from production using a *Hybrid Cloud*. Section 2 explains the distinction in detail.
Hyperscale Cloud Provider	The cloud services providers often referred to as 'public' cloud providers; since there is nothing inherently 'public' in their offering, that word may get in the way of understanding the role they fulfill.
Zero Trust	Zero Trust is a security model based on the principle of "never trust, always verify." Users, computers and software are not trusted until they have been verified (authenticated and authorized). This contrasts with the traditional approach of "trust, then verify."
Zero Trust Architecture	The Zero Trust security model applied to a system.
Zero Trust Network	The Zero Trust security model applied to the security of a network.

SECTION 2

CHALLENGES

PRODUCTION IN THE CLOUD IS A NEW PROBLEM

Traditional production is built on discrete workflows that happen within a facility (the studio, the postproduction house, a VFX company, etc.) on an infrastructure controlled by the facility and secured by its information security group. Assets are moved from one facility to another primarily by using some form of encrypted file transfer protocol or over private fiber networks. The infrastructure is certified as secure with a 'point-in-time' certification from an external vendor.

The cloud production workflows envisioned in the 2030 Vision Paper create new and innovative ways to create content, and they are not contained within facility boundaries. The workflows run in the cloud, with native cloud storage and with content and computation services that may be extended across cloud providers and may dynamically change before or during production ("software-defined workflows").

As will be explained in more detail in Section 3, securing a cloud workflow is not the same thing as an enterprise securing its infrastructure, whether that be a data center or a hybrid cloud.

THE 10-YEAR HORIZON

The 2030 Vision Paper looks at how media production will be transformed over the next 10 years. Much of that change will be enabled by cloud technology. Unfortunately, any prediction as to how cybersecurity will evolve over the next 10 years will be inadequate, as smart attackers will exploit vulnerabilities that defenders have overlooked or don't know about.

For this reason, the new security architecture needs to be modular, and each function can be replaced, if necessary, by a different or updated technology that provides the same utility.

While the 2030 Vision Paper looks out over the next 10 years, the security architecture needs be defined now before there is significant adoption of the technologies and processes it is designed to protect.

CYBERSECURITY THREAT ESCALATION

It should surprise no one that we are in a cybersecurity arms race.

On top of all the cybersecurity threats that any enterprise faces, the content industry has unique threats. The content industry is no longer threatened just by hackers out to prove something by stealing content ahead of its release and distributing it freely on peer-to-peer networks. Today's attackers are business-focused organized criminals, harvesting content to supply illegal distribution or stealing IP for other purposes.

Threat levels increase daily. For example:

- Cyber criminals have access to SaaS tools that raise their capability to a level once only the province of organized criminal gangs and nation-states.
- Software vulnerabilities in operating systems on computers and smartphones, open-source software, and many commonly used applications are found at a current rate of about 25 per day.[1]
- Even when vulnerabilities can be patched, organizations are not always able to effectively apply the patches throughout their infrastructure. The disastrous consequences of the WannaCry ransomware on the British National Health Service and the Equifax breach are examples of unpatched systems being exploited.
- The unique nature of our business means that productions are inherently multi-party, comprising dozens of individual companies and contributors, most of whom have their own IT systems, where content owners have limited visibility of configuration, patching, etc.
- Even the most secure environments are not immune to misconfiguration that can lead to unauthorized or even public access to sensitive IP. Common misconfigurations such as an open bucket policy or default settings can have disastrous consequences.

[1] Calculated from the number of vulnerabilities added to the Common Vulnerabilities and Exposures database in the first half of 2019. See https://cve.mitre.org/ for more information.

ARTIFICIAL INTELLIGENCE AND CYBERSECURITY

Cybersecurity threats will become more difficult to detect as AI becomes a widely available tool for attackers. For example, one type of AI system is the generative adversarial network (GAN). A GAN consists of two AI networks: the generative network and the discriminative network. In cybersecurity, the generative network attempts undetected intrusion, while the discriminative network attempts to detect intrusion.

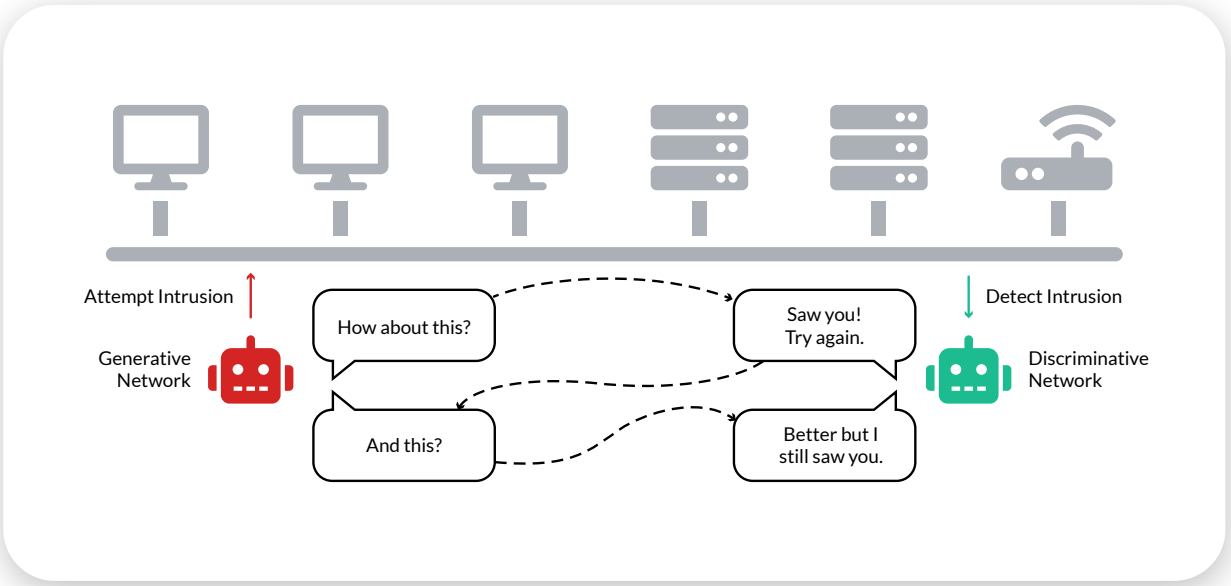

Figure 1: *A generative adversarial network learning cybersecurity*

The most well-known application of GANs is in the creation of synthetic faces in an Nvidia research project. The generative network creates a face that looks human; the discriminative network finds defects. As the process iterates, the faces become increasingly difficult to tell from real people.[2]

They learn collaboratively. The discriminative network informs the generative network how it detected the penetration, and the generative network adjusts accordingly, informing the discriminative network of how it changed.

GANs help both the defender and the attacker, and they will mean that intrusions are far more difficult to detect by conventional means.

[2] Nvidia, Generating Photorealistic Images of Fake Celebrities with Artificial Intelligence, October 30, 2017, https://news.developer.nvidia.com/generating-photorealistic-fake-celebrities-with-artificial-intelligence/

QUANTUM COMPUTING

Quantum computing is a computing platform that is radically different from the binary computing that has been at the core of computers since the first electronic digital ones of the 1940s. A quantum computer can perform certain kinds of parallel operations at a scale not achievable in conventional computers. For example, in theory, a quantum computer of sufficient size can defeat the mathematical properties that make certain types of encryption secure.

There are two primary classes of encryption algorithms: asymmetric and symmetric. Asymmetric encryption enables authentication, which is the foundation for the security of Internet communications between browsers and websites, between mail clients and servers, and between organizations. These asymmetric algorithms rely on the difficulty of factoring certain large numbers. Some security experts have expressed concern that within the next 10 years, quantum computers could be able to rapidly perform this factorization and crack current asymmetric encryption algorithms.[3] But the US National Academies of Sciences, Engineering, and Medicine state that it is "highly unexpected" that a quantum computer able to compromise current asymmetric cryptography will be built in the next decade.[4]

While asymmetric encryption is used for authentication, the actual encryption of data for transfer and storage primarily uses a symmetric algorithm called AES.

Symmetric cryptography algorithms, of which AES is the most widely used, make use of entirely different mathematical functions than asymmetric algorithms and do not rely on the difficulty of factoring large numbers.

Although quantum computing may threaten asymmetric encryption, current research suggests that quantum computing is not a scalable threat to symmetric encryption such as AES.[5]

3. ZDNet, IBM Warns of Instant Breaking of Encryption by Quantum Computers, May 18, 2018, https://www.zdnet.com/article/ibm-warns-of-instant-breaking-of-encryption-by-quantum-computers-move-your-data-today/

4. NASEM, "Quantum Computing: Progress and Prospects," 2019, p. 9, https://www.nap.edu/catalog/25196/quantum-computing-progress-and-prospects

5. Scott Fluhrer, "Reassessing Grover's Algorithm," IACR Cryptology ePrint, August 27, 2017, https://eprint.iacr.org/2017/811

SECTION 3

SECURITY AND PRODUCTION IN THE CLOUD

From here on, we use the term "security perimeter" to mean the conceptual line around the infrastructure within which the facility is responsible for maintaining security, and outside of which the facility is not responsible for security.

Earlier, we noted that there is a difference between a facility using cloud resources to augment or replace its infrastructure (which we defined as a "hybrid cloud") and *production in the cloud* as described in the 2030 Vision Paper. This section explains what that means from a security perspective.

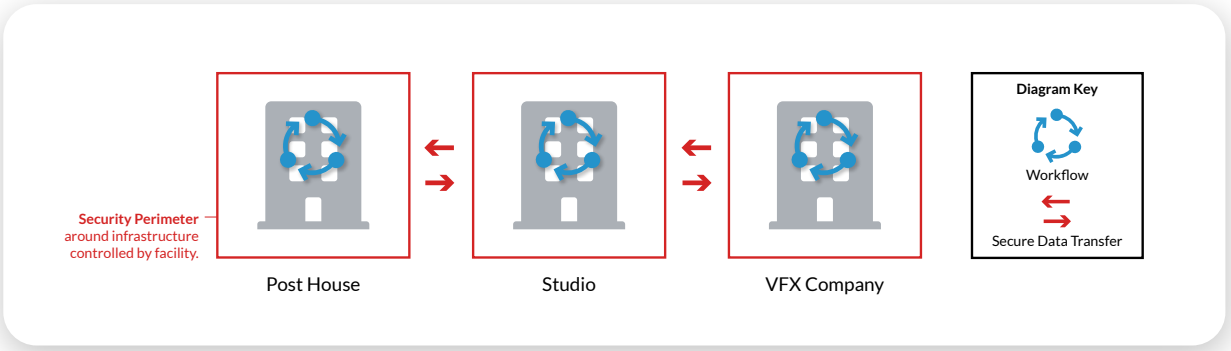

Figure 2: *Facilities operate within their security perimeters*

We first look at how security is implemented today in most facilities.

The security perimeter is the basis of most enterprise cybersecurity today and is the outermost wall of a defense-in-depth strategy.

Traditional production workflows are carried out on a facility's infrastructure: A process happens, and data is passed to the next step. If the next step is in a different facility, the data is transferred using a secure file transfer protocol. In terms of today's security model, it can look like this:

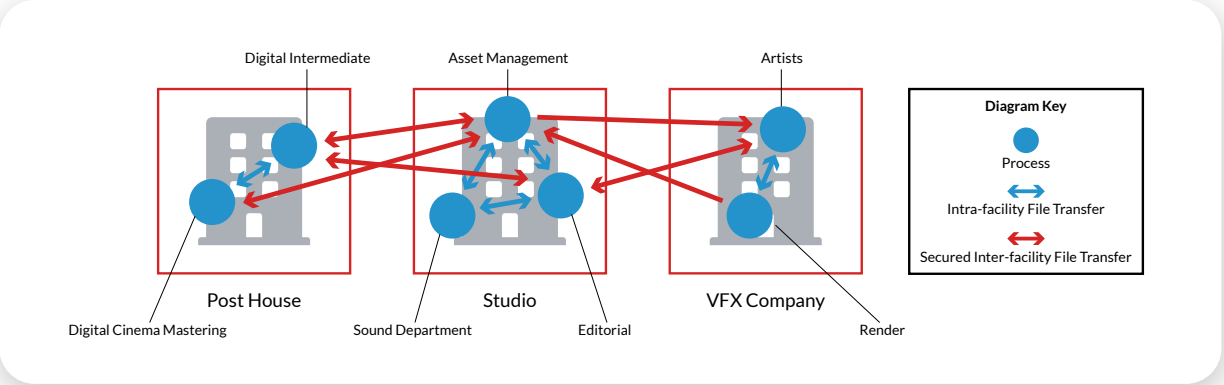

Figure 3: *Today's production workflow is protected by perimeter security*

Other than data transfer, nothing happens outside of a security perimeter. Thus, the workflow can be protected by the security perimeter model as long as the transfers are secure and human error to manage those transfers is minimized.

HYBRID CLOUD PRODUCTION

Today, we are seeing cloud resources being used in production to provide processing and storage resources to either supplement or replace the infrastructures within facilities. When this happens, the cloud resources are "leased" by the facility and used exclusively by it. While this makes a difference as to where computation and storage resources are located and how they are paid for, there is much less distinction between a facility's on-premises infrastructure and the cloud resources they use at the level of control and ownership. Let's add cloud resources to Figure 2.

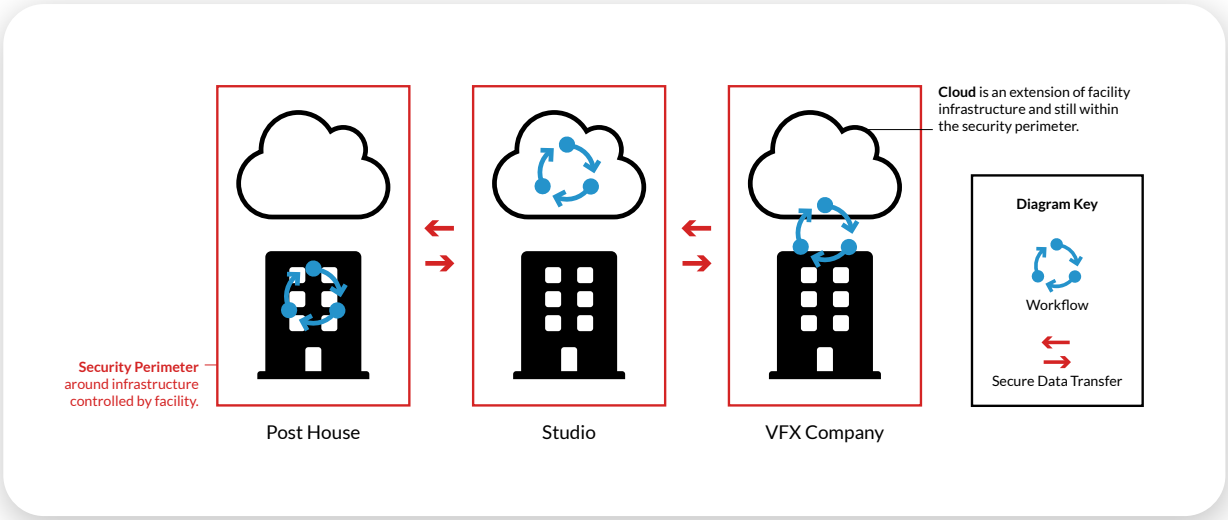

Figure 4: *Facility security perimeters expand with hybrid cloud to encompass their cloud environment*

The processing and storage are still inside the facility's security perimeter, since the cloud resources are completely controlled by the facility and for its exclusive use, even though the security of the cloud portion of a hybrid cloud is the responsibility of both the facility and the cloud provider.

We are looking beyond this.

PRODUCTION IN THE CLOUD

The 2030 Vision Paper envisages all work being done in the cloud and all assets being stored there. *t*d workflows and the resources they use are abstracted from any facility's infrastructure.

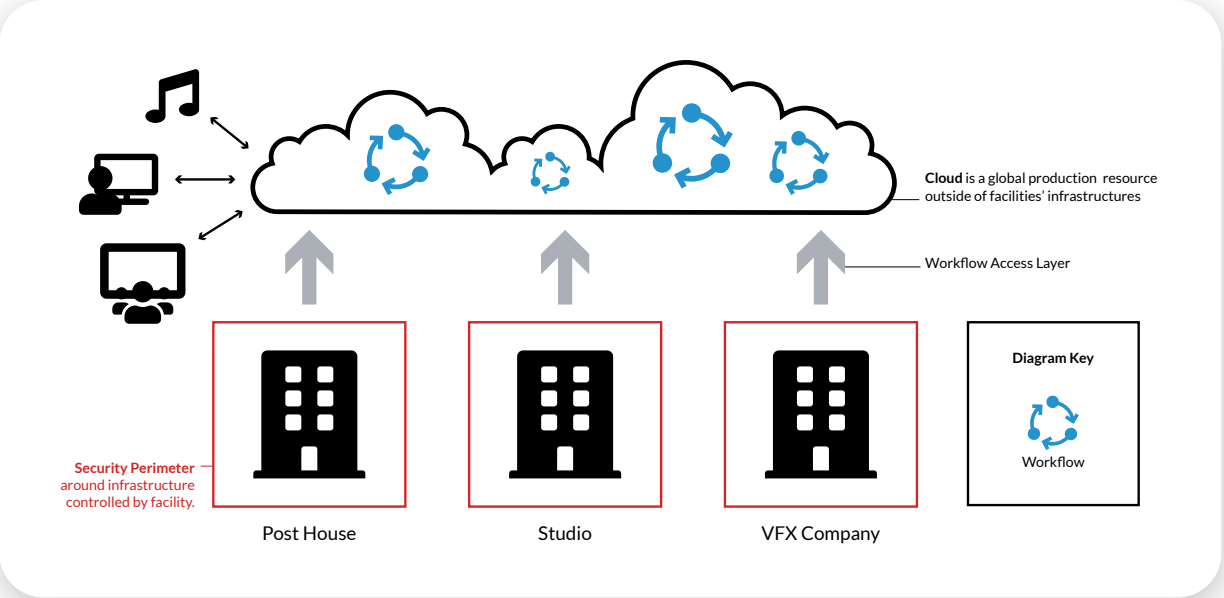

Figure 5: *Production in the cloud abstracts workflow control and access from facilities*

The use of a workflow access layer not only relocates the workflow from the facilities—it moves it beyond facilities' security perimeters. This is the challenge that a new security architecture needs to address, but crucially, acknowledging that this will be a phased migration, until we get to truly cloud-native workflows, we need to continue to protect the legacy systems in Fig 3 and Fig 4.

THE NEW WORKFLOW REQUIRES A NEW SECURITY MODEL

We are at a revolutionary moment. Technology is enabling content to be produced in new and exciting ways. At the same time, cybersecurity technology is changing rapidly, and Zero Trust platforms that satisfy the principles in this paper are available from many vendors. Forrester publishes a report of over a dozen vendors offering multiple components.[6]

Workflow innovation is only just beginning, but we cannot implement new workflows unless and until we can secure production in the cloud. This requires a new, unified security model. We have a point-in-time opportunity to rethink security and to be very deliberate to get it right.

6. Forrester Research, The Zero Trust eXtended (ZTX) Ecosystem, December 2018.

SECTION 4

THE SECURITY PRINCIPLES

In this section, we define six security principles that are the basis of the security architecture that will secure the 2030 Vision. Securing facility infrastructure is a well-understood discipline with existing recommended practices and certification programs, and there is no need to address it here. It is the goal of the security architecture to protect the production workflows envisioned in the 2030 Vision Paper, and as we have seen, these transcend the facility infrastructures used by traditional production.

SECURITY PRINCIPLE 1: SECURITY IS INTRINSIC TO EVERY COMPONENT OF EVERY WORKFLOW AND DOES NOT INHIBIT CREATIVE PROCESSES

The 2030 Vision Paper foresees new and better ways of producing content being invented and running on cloud platforms. When a new process or application emerges today, one of two things usually happens: either a way is found to secure it easily within the existing security model or sometimes, for expediency, it is unfortunately left less than fully protected.

The alternative is to design security natively into the processes and applications. This is called *security by design* and is a well-defined security discipline. Security as an add-on rarely yields the same level of security as security by design.

Figure 6: *Security as an add-on vs. security by design (those are gold bars)*

Security by design is an approach to designing systems where security is a foundational component of system design. The approach takes malicious practices for granted and makes no assumption as to the trustworthiness of users or what an attacker may or may not do.

> **"Applications without security architecture are as bridges constructed without finite element analysis and wind tunnel testing. Sure, they look like bridges, but they will fall down at the first flutter of a butterfly's wings. The need for application security in the form of security architecture is every bit as great as in building or bridge construction."**
>
> **Introduction to *Security by Design Principles*, Open Web Application Security Project (OWASP).**

If followed, security by design reduces security vulnerabilities and minimizes the effects of any that are discovered. Security by design applies to every part of the system, including application software, system configuration, and access controls.

There are several accepted methodologies available, and some are listed in the Suggested Reading section at the end of this document.

An outcome of security by design is that security measures do not get in the way of the user and therefore the creative process of content production.

However, we have deliberately not stated that the security is transparent to the user. The architecture needs to support transparency, but it is up to the content owner to determine how aware of security the user needs to be.

Table 1 below lists some relevant security by design principles drawn from those defined by the Open Web Application Security Project.[7]

[7] Open Web Application Security Project, Security by Design Principles, https://www.owasp.org/index.php/Security_by_Design_Principles

Security by Design Dictum	Explanation
Minimize attack surface area	Every feature that is added to an application adds a certain amount of risk to the overall application. The aim for secure development is to reduce the overall risk by reducing the attack surface area.
Establish secure defaults	By default, the experience should be secure, and it should be up to the user to reduce their security – if they are allowed.
Run with least privilege	Accounts should have the least amount of privilege required to perform their business processes.
Apply defense in depth	Even when a single control might be reasonable, more controls that approach risks in different fashions are better.
Fail securely	Applications regularly fail to process transactions for many reasons. How they fail can determine if an application is secure or not.
Don't trust services or infrastructure	Many organizations utilize the processing capabilities of third-party partners, who more than likely have different security policies and postures than you. Implicit trust of externally run systems is not warranted.
Separate duties	Certain roles have different levels of trust than normal users. Administrators are different from normal users. In general, administrators should not be users of the application.
Avoid security by obscurity	The security of key systems should not rely upon keeping details other than encryption keys hidden.
Keep security simple	Attack surface area and simplicity go hand in hand. Developers should avoid the use of double negatives and complex architectures when a simpler approach would be faster and simpler.
Fix security issues correctly	Once a security issue has been identified, it is important to develop a test for it and to understand the root cause of the issue. When design patterns are used, it is likely that the security issue is widespread among all code bases, so developing the right fix without introducing regressions is essential.

Table 1: *Foundational elements of security by design for production workflows*

In addition to elements above, which optimize the security of a design, aspects of implementation and deployment are essential to the overall success of a security system.

The ones below are drawn from the NSF-funded SecureCore project.[8] In particular for production workflows, security functions should not interfere with the efficiency or creativity of users. Ideally, users should not even be aware of the security controls unless they try to do something that is not authorized.

Make security usable	User interfaces for security functions and supporting services should be intuitive and user friendly. Security measures should not significantly impede efficient use of systems.
Scale costs to value	The financial investment in security should always be in line with the value of the assets under protection. The strength of mechanisms must be sufficient to satisfy the system requirement. Using mechanisms of greater strength than necessary may unnecessarily incur extra overhead.

Table 2: *Guidelines for implementation and deployment of security for production workflows*

SECURITY PRINCIPLE 2: THE SECURITY ARCHITECTURE ADDRESSES CHALLENGES SPECIFIC TO CLOUD-BASED WORKFLOWS

With this principle, we start to define our new approach to security. As we will see, this principle means that the security architecture can be implemented in a way that does not require perimeter security.

Security principle 2 has three parts to it:

a. Security is centered on workflows, rather than the infrastructure they run on.
b. Security is centered on assets, rather than their storage and transport.
c. The integrity of assets, processes, and workflows is protected.

PRINCIPLE 2(a): SECURITY IS CENTERED ON WORKFLOWS, RATHER THAN THE INFRASTRUCTURE THEY RUN ON

[8] Benzel et al., "Design Principles for Security," 2005, https://pdfs.semanticscholar.org/7830/cafce7da73aa5b137e2a5654f75877e306cd.pdf

A significant part of the 2030 Vision Paper is the software-defined workflow. A software-defined workflow, as the name implies, is configured out of software processes independent of specific hardware. That is, the workflow is not tied to specific pieces of hardware in specific places but can happen wherever the necessary processing, storage, and communications infrastructure exists.

These workflows are well suited for implementation on cloud platforms abstracted from any facility's infrastructure or any particular cloud provider. The workflows are well defined and consist of a set of processes each utilizing one or more applications. As the workflow is independent of hardware, so must be the security model. We want to minimize the attack surface, and stating the goal of the security correctly helps with that. For security principle 2(a), it is as follows:

> The security goal is to protect the workflow and ensure the integrity of the workflow whether it is running on a hyperscale cloud platform, on facility infrastructure, or on a hybrid of the two.

By protecting the workflow wherever it runs, the security is abstracted from the infrastructure, as is the case with the workflow itself. The security wraps around each workflow like cling wrap, reducing attack surfaces to the minimum. Let us look at this example of a workflow:

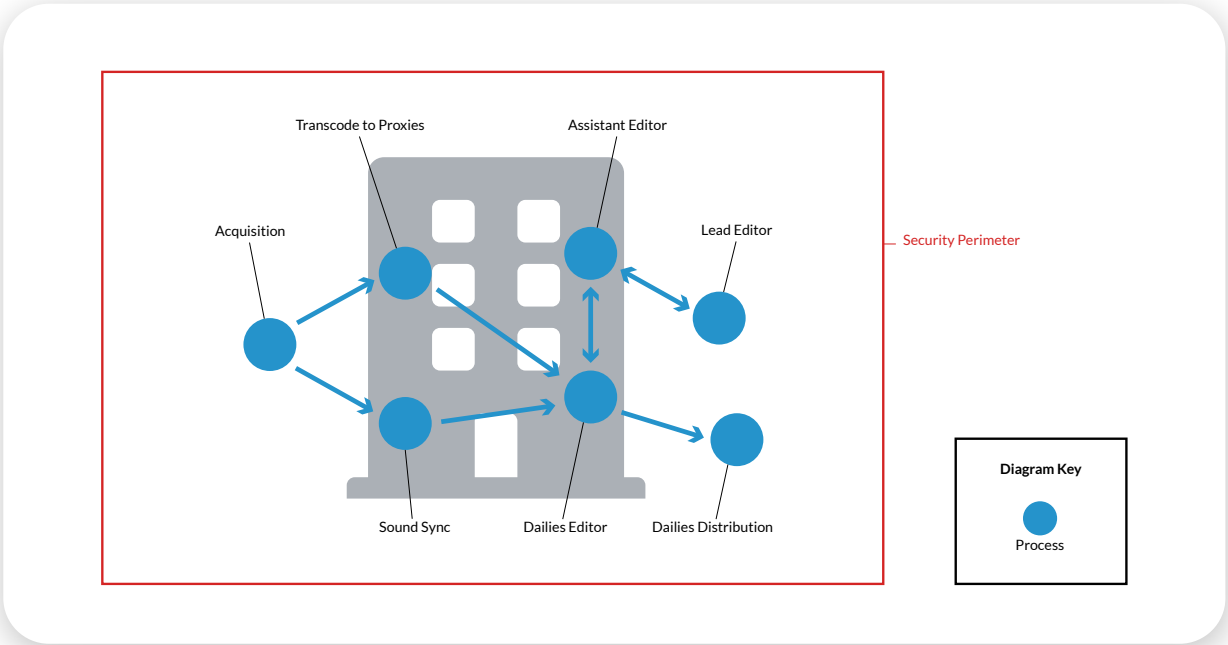

Figure 7: *Dailies workflow in a facility*

Clearly, since the workflow is running entirely on the facility's infrastructure (whether on-premises or in a hybrid cloud), it is within the security perimeter.

When the workflow is running in the cloud, we need a way to protect it, and we choose to do that the most secure and efficient way, which is to protect only what needs to be protected. That is shown in this next diagram.

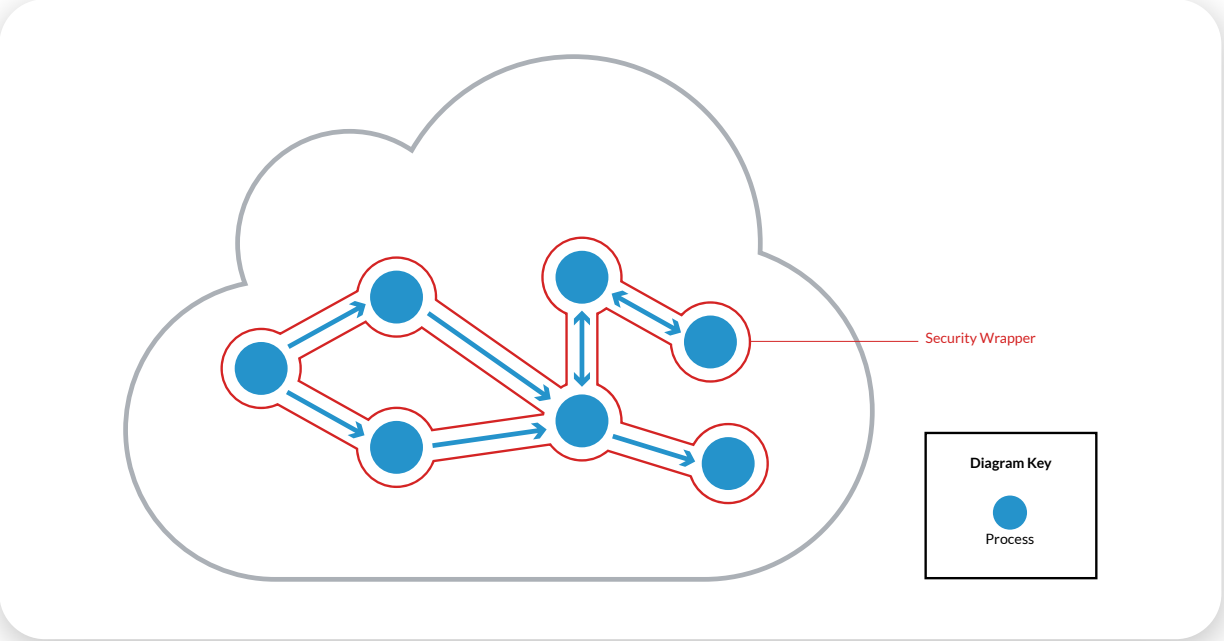

Figure 8: *Securing the workflow above by protecting each process*

PRINCIPLE 2(b): SECURITY IS CENTERED ON ASSETS, RATHER THAN THEIR STORAGE AND TRANSPORT

Above, we drew a distinction between protecting a workflow and protecting an infrastructure. A second distinction is between protecting assets (usually files) and protecting the storage (buckets, volumes, hard drives, etc.) where the assets reside.

The preferred method of protecting assets stored in the cloud today is encryption-at-rest, which is a storage-centric function; data is encrypted in a cloud bucket, an object store, or a storage volume. When that data is moved, it is decrypted, transferred, and re-encrypted if the new storage is encrypted. Encryption is part of the write function, and decryption is part of the read function. The robustness of encryption-at-rest may not be consistent across all the types of storage used, but that is generally not a significant risk if best practices are followed.

Similarly, encryption in transit is a data communications–centric function; what is encrypted is the data stream, regardless of how many files are being sent. Encryption is part of the transmit function, and decryption is part of the receive function. A bad choice of transfer protocol or insecure key management practices are major risks but can be managed by following best practices. From the viewpoint of data encryption, the right combination of encryption at rest and encryption in transit only leaves the data unencrypted for a short period of time between reading from storage to encryption for transmission and vice versa.

However, there are more serious security risks. Encryption-at-rest and encryption in transmission protect against external attackers. But it is file access controls that control file access by system users. This means that if someone or a process can access the file, for example to copy it, they can read the data.

Access controls tend to be built around the security models of the platform. In this model, it is not possible to allow someone to read a file so that they can copy it, a utility function, without giving them the ability to read the data the file contains, a higher-level function. To put that another way, even following the principle of least privilege, there is no differentiation between being able to manage a file and being able to read its contents.

A common attack is privilege escalation, where an intruder who has penetrated the security perimeter has exploited a security weakness to elevate their access rights on the system from user to administrator. The weakness could be at any one of many places in the system.

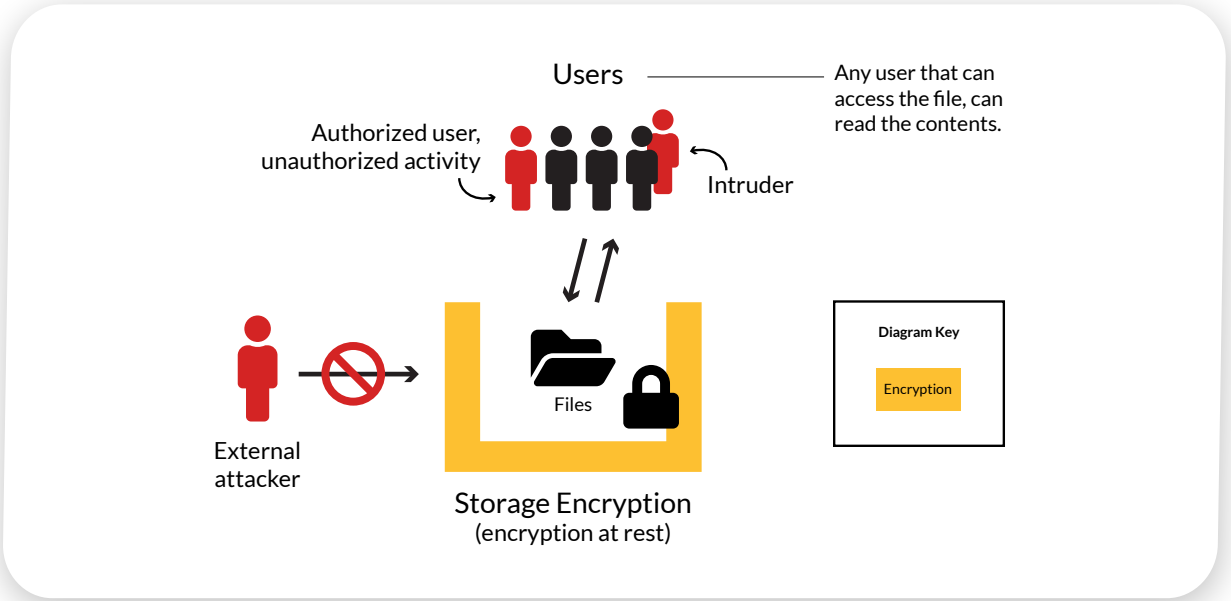

Figure 9: *Encryption at rest protects from external attack but is vulnerable to inside attack*

In Figure 9, we can see that encryption at rest protects the file from an external attacker, but it does not protect the file from an intruder who has gained the same privileges as an authorized user. Furthermore, it does not protect the file from unauthorized use by an authorized user.

Most successful attacks on DRMs exploit weaknesses in the implementation of the player. Consumer devices are highly cost-sensitive, leading to a level of robustness in DRM implementation that is often, at best, the minimum required.

However, the renewable security of DRMs supports rapid remediation if any breach does occur in a particular implementation.

Asset-centric security means having a security mechanism that can enforce different access control on individual assets. One of the better mechanisms for this is to encrypt each asset at the file level and to only decrypt it when it goes into a process (i.e., application) and to encrypt new or modified assets when they come out. The right to access the contents of a file is separated from system access controls. Decryption rights are granted only to users, applications, and processes authorized to read the file contents at the time access is requested.

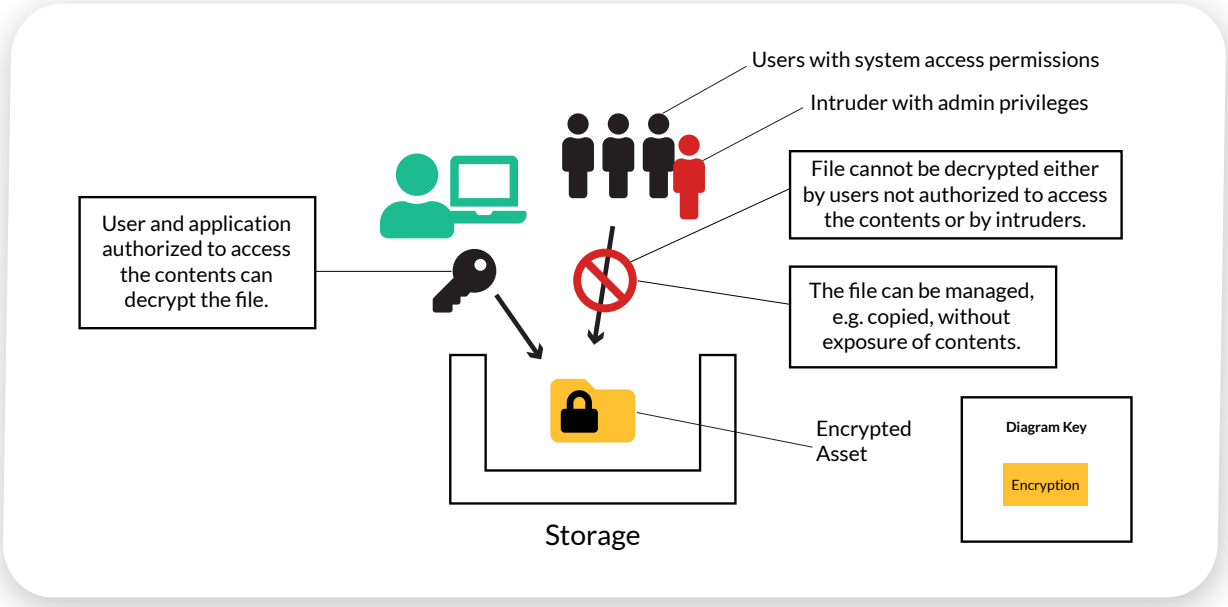

Figure 10: *Asset encryption separates file access controls from controls on access to file contents.*

In Figure 10, we see that while the intruder with escalated privileges can copy the file, as can any user with the appropriate access, they cannot decrypt the contents. Thus, asset management is unaffected, but permission to decrypt the file is a separate function with separate authentication. Decryption rights are "need-to-know." The permission can be granular to the level of each asset.

This is the way the DRM (digital rights management) systems used to protect content delivered to the consumer operate. DRM technologies protect the distribution of the key used to encrypt the content files and do not process encrypted files. An encrypted file needs no protection (since it is encrypted) and can be streamed, downloaded, or distributed on physical media because, without the encryption keys, the file is a useless collection of bits. The risks to an encrypted file (assuming a robust encryption algorithm like AES is used) are to be found in key management and the security of the processes where the content is encrypted and decrypted.

The security of the asset is managed by the encryption key management service, separating asset security from system access controls. This reduces complexity (security by design) and means that asset access can be delegated to, for example, productions without jeopardizing the content owner's enterprise security.

Furthermore, asset encryption enables use cases that are not possible with file access controls alone. For example, different parts of a file can be left unencrypted, allowing wider access to metadata, or encrypted with different keys, allowing separately controlled access to metadata and video assets.

The security architecture should not mandate specific policies. Content owners can choose the granularity of access control (e.g., how many different keys are used) and how broadly access is granted. It's also important that access control actions that "publish" out assets, such as dailies, can be under the control of those producing them. These are independent of the access control mechanism(s) themselves.

We've used the example here of asset encryption, but a scalable security architecture could include the other access control mechanisms (in a later section, we discuss a mechanism called microsegmentation). When asset encryption is used, it is up to the content owner to decide which assets to encrypt and whether they operate the key management service or delegate it, perhaps to the process owner. These are not global services.

Lastly, while the security architecture could mandate a single standard encryption algorithm (e.g., AES-256[9]), the key management protocol does not need to be standardized, and multiple key management systems could coexist for the same asset.[10] This is already the case for DRM systems when there is not a single DRM system that works on all the player platforms supported by a service. In such cases, there is one encryption key for each piece of content but more than one mechanism for securely delivering that key to the player.

PRINCIPLE 2(c): THE INTEGRITY OF ASSETS, PROCESSES, AND WORKFLOWS IS PROTECTED

Securing the integrity of the workflow means that the workflow does the desired actions on the desired assets using the authorized processes.

The primary threat to production is the unauthorized access to and copying of assets. The first two parts of this security principle are designed to address that.

However, it is still necessary to ensure the integrity of the workflow. Workflow integrity ensures that what comes out of the work is what was intended, nothing goes astray, and there are no unauthorized changes to parameters and no extra processes added. A violation of workflow integrity does not have to mean that content is leaked. It might mean that content is corrupted (for example, the music is replaced, or the edit is modified), that content is sent to an unauthorized location, that incorrect content is sent to an authorized location, or that a workflow process is substituted without authorization.

One way of doing this is to extend access control policies from just users to combinations of users and applications. For access control enforced by encryption, this could mean that both the application and the user need to authenticate themselves to the key server before receiving the decryption key.

9. The Advanced Encryption Standard (AES) is defined both in FIPS PUB 197: Advanced Encryption Standard (AES) and ISO/IEC 18033-3: Block ciphers. 256 refers to the key length and is the key length of choice.

10. For example, the Common Encryption Scheme (CENC) is the MPEG-DASH streaming standard.

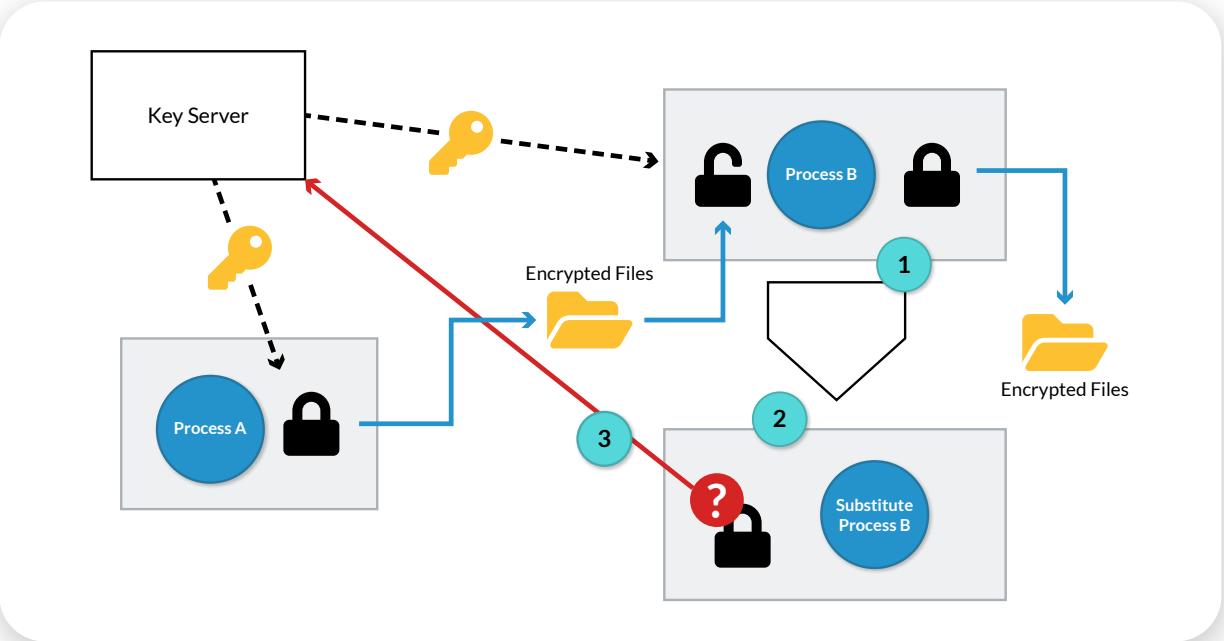

Figure 11: *Example of process substitution (the numbers are referenced in the text below)*

Figure 11 shows an example of this. The approved workflow is process A, followed by process B. Files created in process A are encrypted on output, then saved to cloud storage. Process B reads the encrypted files, decrypts them with keys obtained from an encryption key management service, performs its function, and then encrypts the outputs. If an unauthorized change is made to the workflow, this is what happens:

- At (1), the owner of process B decides to use a substitute for process B. The substitute might be a different process or outsourcing to a vendor.
- At (2), the substitute process B requires keys in order to decrypt the files it is going to process. It cannot use the keys supplied to process B because the keys are delivered in a protected manner that is specific to the instantiation of process B (this follows the DRM model).
- At (3), the substitute process B requests encryption keys from the encryption key management service. At this point, it is up to the content owner to grant or deny that request. If the request is denied (i.e., the change is not authorized), the substitute process B cannot be used. This enforces both asset security and workflow integrity.

SECURITY PRINCIPLE 3: PRODUCTION WORKFLOWS, PROCESSES, AND ASSETS ARE SECURE, EVEN ON UNTRUSTED INFRASTRUCTURE

> **"There's an old saying in information security: 'We want our network to be like an M&M, with a hard crunchy outside and a soft chewy center.' For a generation of information security professionals, this was the motto we grew up with. It was a motto based on trust and the assumption that malicious individuals wouldn't get past the 'hard crunchy outside.' In today's new threat landscape, this is no longer an effective way of enforcing security. Once an attacker gets past the shell, he has access to all the resources in our network. We've built strong perimeters, but well-organized cybercriminals have recruited insiders and developed new attack methods that easily pierce our current security protections."**
>
> **John Kindervag, Forrester Research, 2010.**

The security architecture must support securing workflows and assets regardless of whether the infrastructure of servers, workstations, storage, and data communications is secure. Securing activity on an untrusted infrastructure is part of our everyday lives. Financial transactions over the Internet do not rely on the Internet being secure, and it would be foolhardy if they did. The web browser is secure, the financial institution web server is secure, and the HTTPS connection between the two is secure. It does not matter than the Internet connection is not.

Some of the best concepts for security on untrusted infrastructure were first articulated in John Kindervag's 2010 paper, "No More Chewy Centers: Introducing the Zero Trust Model of Information Security." That paper opens with *"there is a simple philosophy at the core of Zero Trust: Security professionals must stop trusting packets as if they were people."*

The Zero Trust architecture is centered on the belief that nothing should be automatically trusted either inside or outside of any security perimeter. Instead, the rule is to verify anything and everything trying to connect to a system before granting access.

At the center is the philosophy that you don't trust anything until you know who that user or system is and whether they're authorized. In this security model, no access is granted until the network knows who is asking. In the case of our security architecture, that means no process can join a workflow until it is verified. In 2009, Google responded to a sophisticated cyberattack called Operation Aurora with an internal initiative to reimagine their security architecture with regards to how employees and devices access internal applications. The result was BeyondCorp (https://beyondcorp.com/), a Zero Trust security framework modeled by Google for its enterprise that shifts access controls from the perimeter to individual devices and users. The result allows employees to work securely from any location without the need for a traditional VPN.

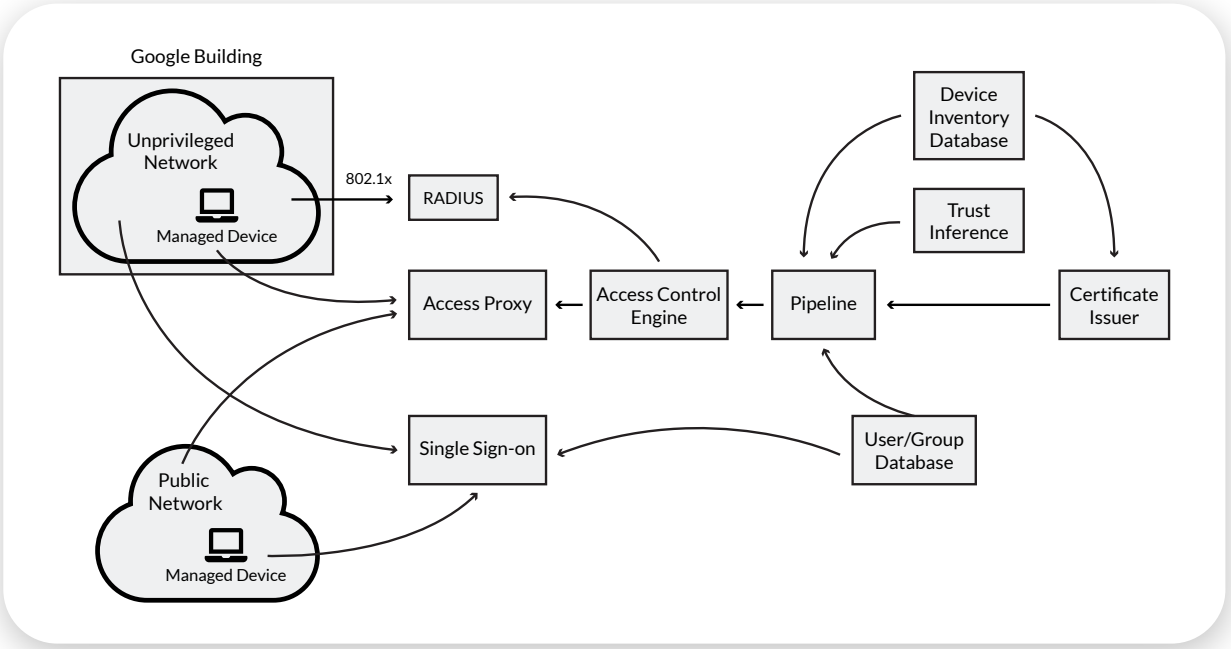

Figure 12: *BeyondCorp components and access flow (Source: BeyondCorp)*

For our purposes, Zero Trust security is an ideal approach. It places the control of security in the hands of the content owner, it is part of the security architecture, and it can be implemented through specification of the security of the applications and other components that are used in the workflow. Granularity is controlled by policy, not technical constraints.

Google is not the only major technology company to use a Zero Trust architecture. The model is becoming a core tool for enterprises securing processes and data in the cloud.

While Zero Trust models are straightforward to explain, many in IT see deploying them in the enterprise as a challenge. It is not necessarily easy to decide where, when, and how to implement a Zero Trust architecture in the enterprise infrastructure.

> **"Unlike the traditional perimeter security model, BeyondCorp dispels the notion of network segmentation as the primary mechanism for protecting sensitive resources. Instead, all applications are deployed to the public Internet, accessible through a user and device-centric authentication and authorization workflow."**
>
> **BeyondCorp website, https://beyondcorp.com/.**

Enterprise IT operations are built to support any application the enterprise may need. Applications are used by different departments in ways that are partially or completely unknown to the IT department. For example, the IT department provides email but does not know, or need to know, how a department makes use of email. That means many different interacting functions are running at once, and many of the functions and interactions are unmapped.

By comparison, the workflows of the 2030 Vision Paper can be mapped out. For the most part, they are assembled from processes that utilize specific, known applications. And while there are many nuances, the relatively small number of core workflow tasks makes the use of the Zero Trust architecture in our security model more straightforward.

The flip side of establishing trust in a Zero Trust architecture preventing access by everyone or everything that has not established trust. A range of security tools can be used for this. Trust in endpoints can be established using fully authenticated software boot chains, signed operating system and application software, and physical security that restricts access to hardware and displays. Continuous risk assessment, dynamic authorization and activity monitoring complete the picture.

SECURITY PRINCIPLE 4: THE CONTENT OWNER CONTROLS SECURITY AND WORKFLOW INTEGRITY

In a traditional workflow, security is the responsibility of the facility as a matter of course. Content owners exert control over security though contractual obligations, which may include a requirement for certification. The content owner has no insight into the operation of the security other than through an audit or a post-breach investigation.

Under this principle, authentication and authorization need to be at the center of the new security architecture. The security can then be built on services, including those that provide authentication, provide authorization according to policies, and manage asset encryption keys.

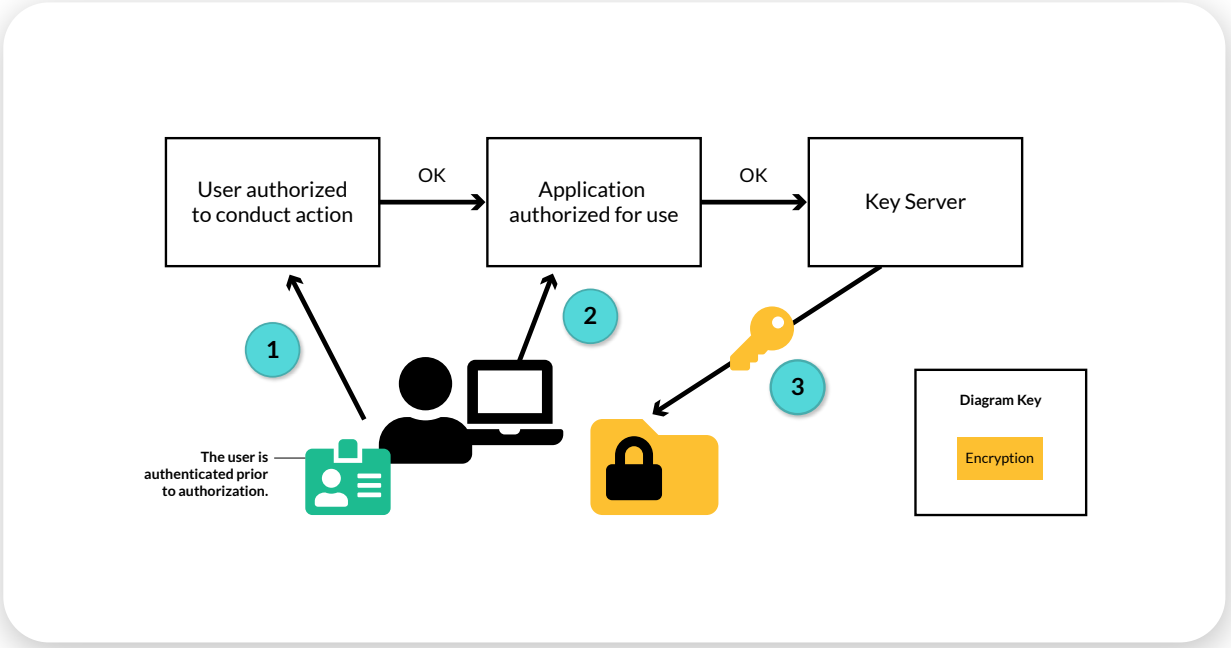

Figure 13: *The authorization process that grants access to file content*

Figure 13 shows the three steps that could allow access to a protected asset:

- At (1), an authenticated user is authorized to conduct the task they wish to do. The granularity of the resulting authorization is defined solely by the policies configured in the authentication server: this can vary from authenticating that they are permitted to work on the project to authenticating that they can access an asset.
- At (2), an authenticated application is authorized. This can happen asynchronously to (1). Authorization means that the application is approved to conduct the action the user is about to perform. The granularity of the authorization decision could be such that, for example, any version of the application is authorized, one version is authorized, or one version running on a specific machine is authorized. It may include verification that the application has not been changed.
- At (3), the key server supplies a key if both the user authorization and the application authorization approve access.

When you want to stream content from an OTT service, the same steps are followed. (1) The service's entitlement server determines whether you are permitted to stream the content you want to watch. Once the entitlement has been approved, (2) the DRM system checks that the version of the player is not one that has been compromised. If it is OK, then (3) the DRM playback license servers delivers decryption keys for the content to the player.

In the security architecture, the use of the two authorizations would be optional. And the security architecture would not mandate the use of either.

For example, user authentication could be tied to an identity management service that manages all employees, and if any version of any application is permitted, no application authentication is necessary.

Content owners can operate their own services, or they can outsource or delegate the services to others. The authentication may be tied into a content owner's identity management system.

As an example, some options for the operation of key management services are show in Figure 14.

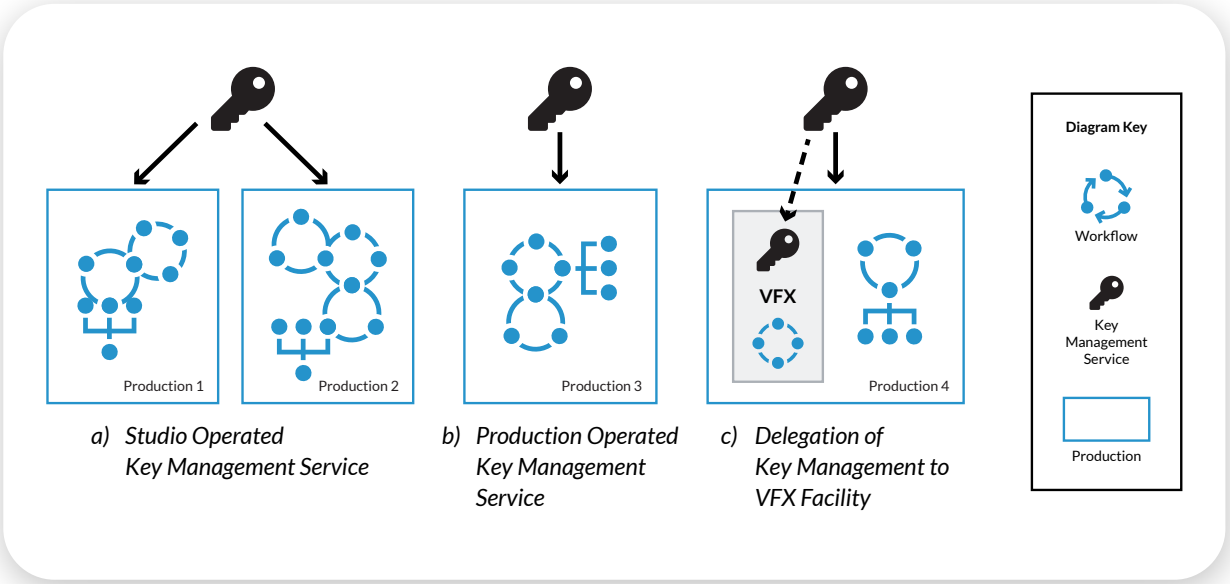

Figure 14: *Examples of different ways to manage key management services*

However, if the content owner chooses to stay in control of those services, they have control over the security and the workflow integrity.

These security controls could work together with the building blocks of the 2030 Vision Paper to yield a significant improvement in workflow management in terms of efficiency, security and, potentially, cost.

SECURITY PRINCIPLE 5: THE SECURITY CAN BE SCALED TO APPROPRIATE LEVELS AND CAN INTEGRATE WITH EXISTING SECURITY POLICY AND MANAGEMENT SYSTEMS

This principle defines the relationship between the security architecture and the content owner's requirements and security management systems.

The System for Cross-domain Identity Management (SCIM) can be used to share information about user attributes, attribute schemas, and group membership. Attributes could range from user contact information to group membership. Attribute values and group assignments can change, adding to the challenge of maintaining the relevant data across multiple identity domains.

The SCIM standard has grown in popularity and importance as organizations use more SaaS tools.

From Wikipedia (edited)

While quantifying the cost of security is outside the scope of the security architecture, it is quite reasonable to expect a correlation between the level of security and the cost of security. How security is to be scaled is a decision that comes from risk management: risk assessment, risk tolerance and the cost of security must be aligned. Risk assessment would typically determine that both the potential financial loss associated with a security breach and the probability of a security breach are larger for a major motion picture than they are for, say, a competitive cooking TV series. The motion picture has less risk tolerance than the TV series, and the lower the risk tolerance, the higher the level of security needed.

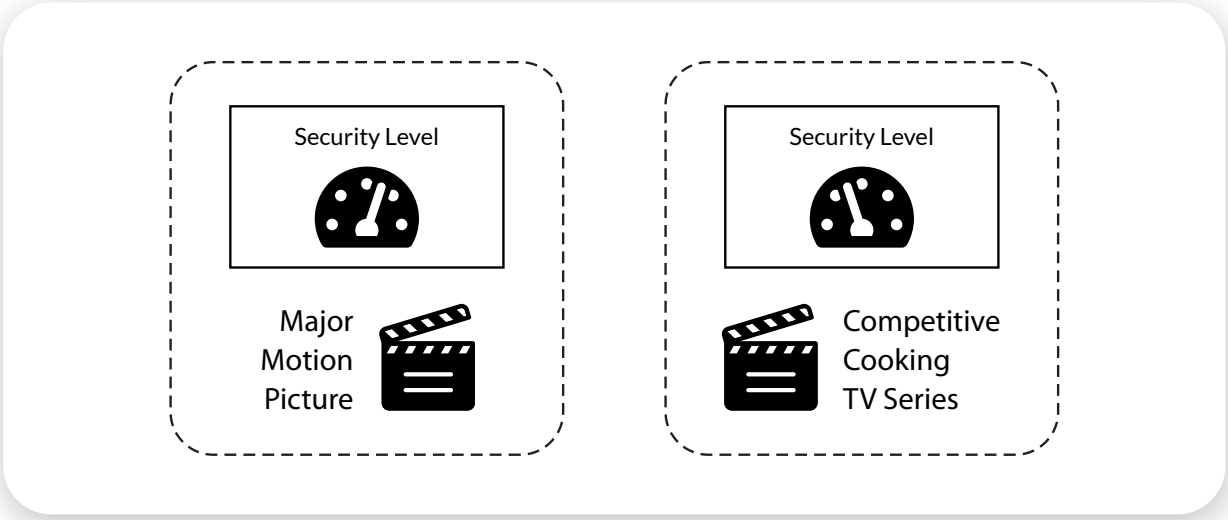

Figure 15: *Different productions have different security requirements*

Within a production, the attractiveness of an asset to an attacker varies by its type. For example, an individual camera frame would be less attractive than the dailies with sync sound, which in turn would be less attractive than a rough cut. And the likelihood of a leak may vary with the workflow or the way it is deployed. Scalability permits the security to be dialed in based upon assessments of the likelihood of a particular asset leaking and its impact.

The security architecture needs to be modular so the content owner can scale the security through the selection of security elements. Of course, there may be prerequisites: asset encryption requires a key management service, and the security model needs a way to authenticate users to implement the Zero Trust architecture.

The security architecture does not exist in isolation, and while it can, there is no inherent reason why it needs to supplant content owners' security management systems. Instead, it supports and integrates with existing security management. A content owner may have global security policies; they may set security policies by production, by process, or by the nature of the asset. The security policies for the editorial department may be different than those of the VFX department. And a policy framework needs capabilities to enable those working on assets to grant access to others for review (e.g., to publish dailies) or to propagate assets to the next stage of processing.

To enable this across multiple platforms, the security architecture will need common mechanisms to express and communicate security policies.

The security architecture is based on the principle of "verify, then trust," and so at the core of the security architecture is authentication. Establishing trust starts with identity and access management (IAM); however, the security architecture does not specify the IAM system. It is expected that the identity management will be maintained external to the security architecture, for example by the content owner's own IAM systems. The identity requirements apply not only to users, the way IAM is used in traditional security models, but also to other components, most notably applications.

Although no specific IAM system is required by the security architecture, cross-domain IAM is desirable. This has been of interest in cybersecurity and IT circles for at least 10 years. One standard is the System for Cross-domain Identity Management (SCIM). The standard is currently managed by an IETF working group. The most recent version was published in 2015 as IETF RFCs 7643 and 7644 along with the use case RFC 7642.

It is not clear if SCIM or any other standard is gaining widespread adoption, and the authors acknowledge that further research is needed.

SECURITY PRINCIPLE 6: THE SECURITY ARCHITECTURE LIMITS THE SPREAD OF ANY BREACH AND IS ADAPTABLE TO THE EVOLVING THREAT AND RESPONSE LANDSCAPE

This security principle is fundamental to the efficacy and longevity of any security system.

The reality is that there no way to make a computer system used by humans that is 100% secure, and it is a fallacy to claim otherwise. The only way to secure a computer with 100% effectiveness is to turn it off, lock it in a steel box, and destroy the key.

For any security system to continue to be secure, it must be able to adapt to the permanent compromise of one of its components. Unfortunately, we don't know what form that compromise will come in, what components it will affect, or when it will happen. This is the unknown unknown, a concept that Donald Rumsfeld brought to fame. It is taken from psychology[11] and is commonly used by intelligence professionals.

We prepare for the arrival of an unknown unknown through a security architecture that is modular and can be rapidly adapted to face the new threat.

It is demonstrable that just about every type of security system has suffered some type of breach. Any security system protecting something that is of value to someone will be attacked with ever-increasing sophistication. The odds that a breach will occur have proven to be high. While every effort must be devoted to preventing a breach, any breach that happens must be rapidly contained.

---

[11] The Johari Window. Created by Joseph Luft and Harrington Ingham in 1955.

Limiting the proliferation of a breach places requirements on the security architecture:

1. It cannot assume that penetration can be prevented.
2. It must enable implementations resistant to proliferation.
3. It must have mechanisms that support the incident response plan.

Obviously, unauthorized access to a resource is a breach by any metric, but in traditional security, an intruder gaining access to the network is a breach. An intruder gaining access to a network secured with a Zero Trust model is not a breach because the network traffic is untrusted by default. Even if there is unauthorized access to a resource, a properly implemented Zero Trust architecture contains that breach because only the minimum necessary access rights are allocated.

> *Some of the most severe cases of network breaches could have been prevented using basic zero trust principles – for example, had [Zero Trust Architecture] access rules been applied to Edward Snowden, he would have been unable to obtain the broad range of documents that he released to the public. Instead, he was given "system administrator" privileges within the NSA network, which provided him blanket access to resources and files.*
>
> *The Road to Zero Trust (Security), Kurt DelBene, Milo Medin, Richard Murray, Defense Innovation Board (DIB),*[12] *July 9, 2019*

That is a powerful statement that comes from an organization charged with helping the US Department of Defense improve its effectiveness.

[12] "The DIB is a Federal Advisory Committee established to provide independent advice to the Secretary of Defense. Statements, opinions, conclusions, and recommendations in the report cited do not necessarily represent the official position of the Department of Defense." Quoted from the cited paper.

A breach might result in an intruder gaining encryption keys for some production assets, perhaps by compromising the security of an application that processed the content and therefore needed to decrypt it. However, proper key management dictates that only a certain set of assets are encrypted with the same key. We refer to this as key diversity. In the most secure case, each asset is encrypted with its own encryption key, but it may be expeditious to encrypt all the frames in a shot with the same key. The scope of the loss from a compromised key is determined by how many assets are encrypted with the same key.

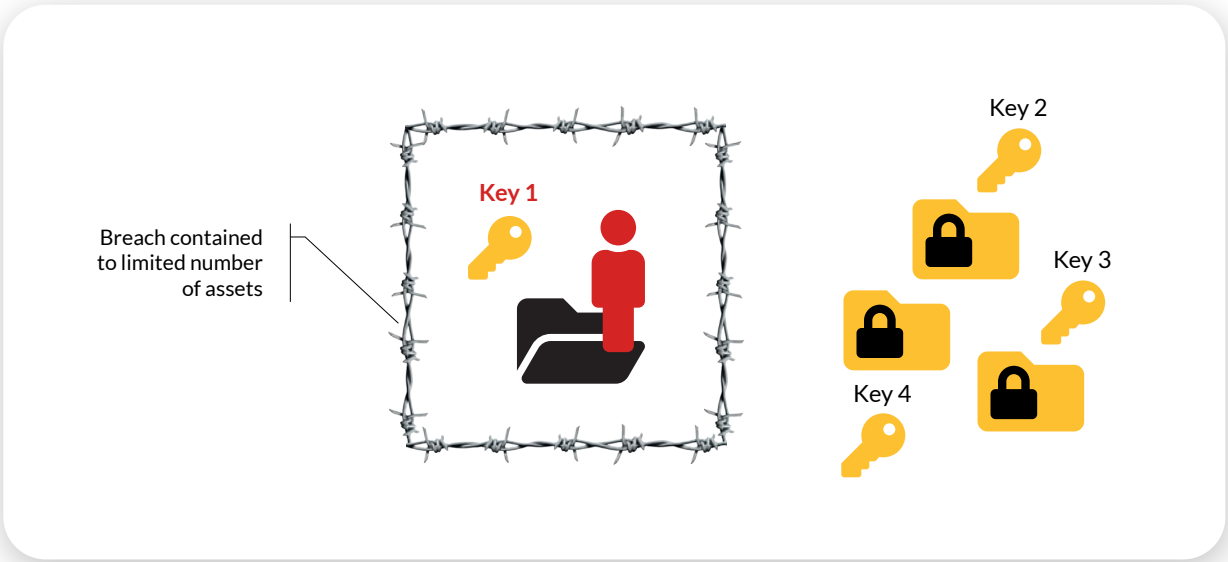

Figure 16: *Different productions have different security requirements*

In this diagram, the intruder has obtained Key 1 and can decrypt the contents of the assets encrypted with Key 1, but the intruder cannot decrypt the other assets, which are encrypted with different keys. To access other assets, the attack must be repeated to gain control of another encryption key.

Security can be increased further through key rotation, in which content is periodically re-encrypted with new keys. Key diversity and key rotation also permit the level of security to be scaled to the value of the assets and the risk to them. Obviously, this approach puts key management services very high on an attacker's target list, and therefore, those services must be extremely well protected.

It is always essential that good security practices are followed regardless of the security model and that security issues are remediated correctly and promptly. The logging and analysis that are part of architecture Zero Trust may help in root cause identification but other things, such as determining the extent of a security issue across all code base is outside of the scope of the security model.

If quantum computing defeats asymmetric encryption, as discussed in the Challenges section, it would represent a catastrophic failure of one part of the implementation of the security architecture. However, it is the implementation that is threatened, not the architecture itself. The architecture needs a mechanism for authentication and exchange of encryption material, and that will likely use asymmetric encryption, but the architecture is not bound to that technology (of course, the threat of quantum computing to asymmetric encryption is not an unknown unknown; it is a known unknown, and preparations are already under way).

The security architecture of the 2030 Vision Paper is modular and implemented using standard security building blocks. When a component is found to have irreparable security problems, as would be the case for a protocol using asymmetric encryption defeated by quantum computing, the protocol can be replaced with a new one.

This is the path that the web has taken when security vulnerabilities are discovered. For example, the underlying security protocol that enables HTTPS to protect web transactions is on its seventh revision, and older versions have been deprecated.

SECTION 5
SCALING SECURITY

Just as productions scale in commercial value, budget, or risk tolerance, security must too. It must scale to production size, budget, and acceptable risk.

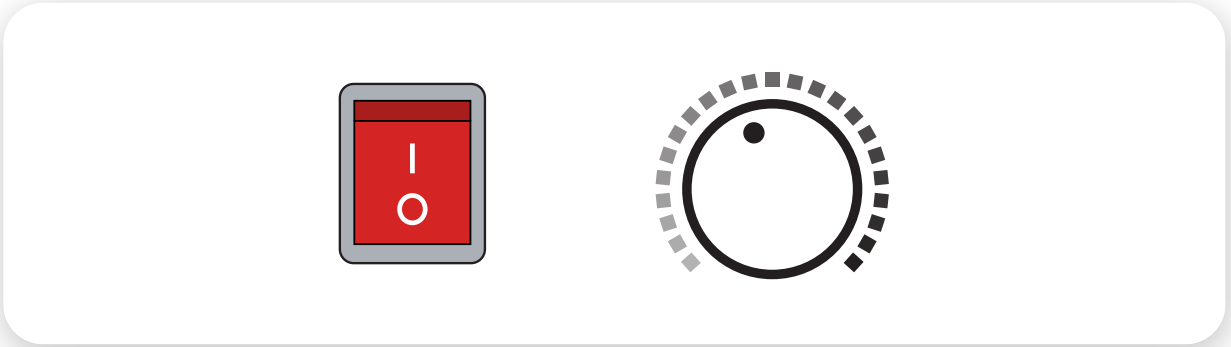

Figure 17: *Non-scalable vs. scalable*

In discussing Security Principle 5, we briefly compared the risk tolerance for a major motion picture and a competitive cooking TV series. The motion picture has less risk tolerance than the TV series, and the lower the risk tolerance, the higher the level of security needed. Security comes at a cost, and while the relationship is context-specific, higher security typically costs more one way or another.

The security model supports many different levels of security, allowing the content owner to select relevant measures, how those measured are used, and which implementation is deployed.

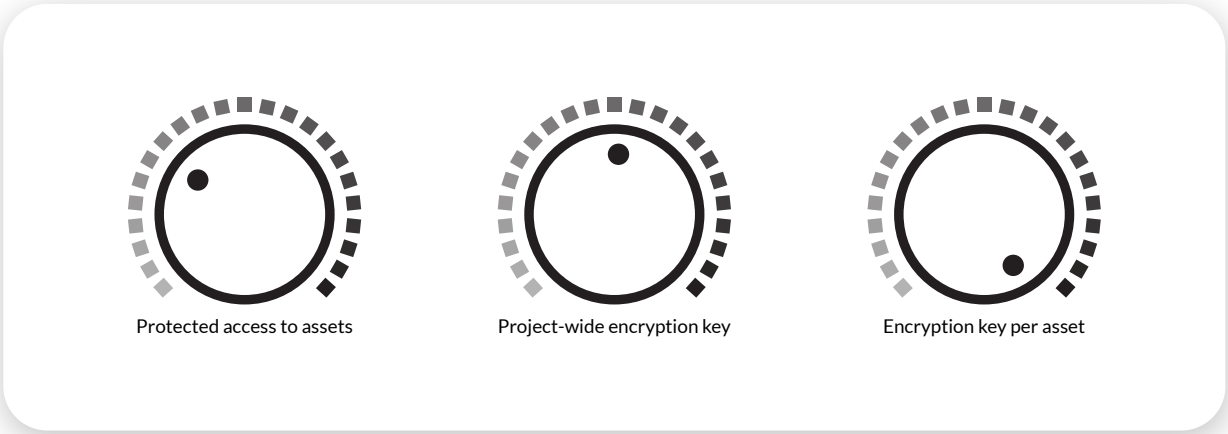

Figure 18: *Scaling security through different mechanisms for asset protection.*

Decisions about the required level of security will be an outcome of formal or informal risk analysis and do not have to remain constant across a production. For example, the value of the season opener and the season finale may be higher to an attacker than that of the episodes in between.

By way of example, here are some ways that the implementation of the principles can be scaled:

Security can be realized in a way that follows the principles without fully utilizing encryption, especially if a lesser degree of security is an acceptable risk. Encryption and key management featured prominently in the explanation of the principles, but encryption is not a principle. It is in the description because it is the most secure way of implementing the principles and the easiest way of explaining features of the principles. Encryption serves a dual function. The properties of an encrypted file mean that

1. The contents are not at risk from anyone that does not possess the encryption key.
2. Controlling access to the key controls access to the asset and does so without relying on system-level access rules.

One of the themes of this paper is that protecting an entire cloud production system is aspirational rather than realistic, and the security model as embodied in Security Principle 2 reduces attack surfaces. It is much easier to protect an asset than a system, and encrypting the asset is a reliable way of doing so.

Principle 2 describes the difference between traditional system-level security and our security model in regard to asset protection. A risk that traditional security does not prevent is unauthorized activity by an authorized user. If access controls are granular to the level of user, application, and time, then encryption is not the only way to implement a system that follows principle 2, especially if a higher level of risk is acceptable.

The Road to Zero Trust (Security) from the Defense Innovation Board, quoted in Principle 6, states that a Zero Trust architecture would have prevented Snowden from doing what he did. It does not say that encryption at the file level is required to achieve that.

However, it would be a mistake to believe that access controls are as secure as encryption. An authorized user conducting unauthorized activity with an unencrypted asset they have a legitimate need to access may not be prevented by access controls at any level. It can be prevented by encryption if decryption is in the application because when the asset is stored outside of the application in a location that the user has authorized access to, it is encrypted.

Without encryption, it may be impossible to prevent the unauthorized activity that encryption will prevent; however, the anomaly monitoring essential to a Zero Trust architecture may and should catch some of that activity.

There are reasons why encryption is better than access controls for both security and production management:

- The way encryption is used and managed in the security model is system-independent. For the purposes of the security model, system-level access controls have shortcomings. While access controls on different system follow similar models (user, group, and public), the way they operate is not universal, and the differences can be subtle.
- The key management can be a localized function decoupled from the system administration of the entire organization. It can be controlled by the content owner (i.e., the production) without jeopardizing organizational security. Conversely, organization-level changes to access permissions will not jeopardize production.
- Key management can be agile and automated. For example, if a task is assigned to a user who then needs immediate access to an asset, access can be automatically authorized by the production scheduling system if it is coupled to the key management system.
- The 2030 Vision Paper sees cloud resources as crossing organizational boundaries, and those resources may not be under the control of a single system administration group.

All that having been said, let us be clear that while encryption is better than relying solely on access controls, employing sufficiently granular and properly managed access controls instead of encryption is one way that the security can be scaled. It is a question of meeting particular risk tolerance thresholds.

Another way that security can be scaled is in the definition of an asset. Up until now, this paper has used the term "asset" in the way that it is used in media production: a camera frame, a video clip, an audio stem, a metadata record. However, what is meant is that an asset is the smallest "thing," the atomic particle, that needs to be protected. Within the context of the security model, an entire episode of our archetypal cooking show could be viewed as a single asset. In that case, the asset, the atomic particle, encapsulates everything that was created in the course of production (formally defined, the security asset is a set of media assets).

Being unable to prevent unauthorized activity by an authorized user, as would be the case, may be an acceptable risk. However, please note that preventing such activity and detecting such activity are different functions, and in the security model, they are independent.

Our last example of scalability is in the implementation of the Zero Trust architecture. The premise of Zero Trust is to verify anything and everything trying to connect to a system before granting access. One way to grant access is by allowing the entity to join a microsegment.

Microsegments are network segments that are accessible only to trusted users and systems. They are typically implemented using virtual local area networks (VLANs), which behave as an isolated physical LAN but are implemented virtually within the network system's routers and switches. This is in contrast with traditional methods, where anything connected to a "trusted" physical port is trusted. In a VLAN, trust is a property of the device, not the physical connection. With properly configured VLANs, which can be tricky, a microsegment can operate securely on an untrusted network, which is Principle 3.

Since this paper does not address implementation, the use of microsegmentation across the Internet is for discussion elsewhere.

SECTION 6

AUTHENTICATION AND AUTHORIZATION

Authentication and authorization are foundational components of the security model.

- Authentication means I have confirmed you are who you say you are.
- Authorization means I will allow you to do something.

Authentication is performed by the entity being authenticated, demonstrating that it

1. Knows something: e.g., for a person, a shared secret like a password, or for a computer system, a private key.
2. Has something: e.g., a physical device such as a FIDO[13]-compliant key device or a cell phone to receive text messages.
3. Is something: e.g., for a person, a biometric method such as facial recognition, or for code or data, a signed, cryptographic hash.

Authorization frequently manifests itself in the form of access controls, for example, role-based systems, where membership in a group, such as administrators, grants specific permissions. The management of authentication and authorization both require a highly robust level of security protection.

Often, authentication and authorization are wrapped up together. A driver license is a form of ID (authentication) that authorizes the holder to drive a motor vehicle.

13. "The FIDO ('Fast IDentity Online') Alliance is an open industry association launched in February 2013 whose mission is to develop and promote authentication standards that help reduce the world's over-reliance on passwords. FIDO addresses the lack of interoperability among strong authentication devices and reduces the problems users face creating and remembering multiple usernames and passwords." Wikipedia entry.

Figure 19: *A driver license provides authentication and authorization*

Enterprise identity and access management systems can provide both authentication and authorization.

However, this is not necessarily the best option for the security model. There are several reasons for this:

Identity and access management (IAM) is the discipline that enables the right individuals to access the right resources at the right times for the right reasons.

Gartner Glossary

- Authentication of individuals might come from the identity management systems of multiple organizations, for example, by the production, the post house, and the studio, or from a federation of those systems.
- Authorization may be highly dynamic. For example, an artist may be authorized to access assets while they are working on a shot but not once they have completed the task.
- In the proposed security model, authentication and authorization requirements can apply to users, systems, and applications.

While it is not a requirement that authentication and authorizations be performed by separate systems, the security model describes them as separate functions to allow sufficient flexibility to address the full range of use cases.

SECTION 7
IMPLICATIONS

The security architecture needs to take a new approach to production security, but since it draws on existing security technologies, much of it can be designed and implemented without need of any significant invention. Early definition of the security architecture will enable the appropriate mechanisms to be developed.

Applications and systems that today rely on being inside a security perimeter will need to be adapted to operate as a part of this new security architecture. However, there are implementation options that can wrap an application to minimize the integration effort. Of course, to run successfully in the cloud, some applications, particularly those using a virtual workstation, need to be modified to run efficiently in the new environment, which is an opportunity to integrate security elements into or around them.

Effectively implementing security requires more than just the specification of technologies. The processes for deploying and managing them is an integral part of ensuring security. Therefore, standards and processes for measuring the effectiveness of implementation and deployment of the security model are needed. For enterprise and facility security, there are standards such as ISO 27001.

In the traditional production arena, the MPA has for many years defined security requirements that have then been employed to audit production vendors and facilities. Studios often have their own additional requirements and audit programs. More recently, the MPA and CDSA launched a worldwide program, the Trusted Partner Network, to accredit auditors in assessing facilities against a common set of security requirements. These programs and the security assessments they provide will continue to be vital for securing facilities. As they extend to encompass the additional requirements outlined in this paper, they will also be essential for securing cloud workflows.

Our security model and consumer DRM have the common goal of protecting the security of assets on an untrusted infrastructure and ensuring appropriate levels of security in applications. Looking at DRM, we can identify possible areas where additional practices and governance may be needed.

DRMs have developed technical and legal constructs to improve their security and renewability. These usually take the form of compliance and robustness requirements. Together, these determine how resistant a player has to be to attempts to circumvent the security of the DRM. Ensuring the robustness of Zero Trust approaches like asset-level encryption will likely require more technical guidelines for implementation and assessment than has been the case for facility and perimeter security approaches. Also, most systems for consumer DRM and for application authentication utilize cryptographic certificates. In the case of consumer DRM, these are usually managed by trust authorities, often operated by the DRM and device providers.

In any event, we believe that the security architecture will benefit from the engagement of the application vendors and a regime to ensure compliance.

The security architecture requires robust means of authenticating users, applications, and devices and validating participation in workflows. These are managed individually, directly or by proxy, for each production by the content owner and will benefit from work being undertaken to create interoperable identification and access control management. In addition, facility security controls will continue to play an important role in securing the endpoints.

SECTION 8
BENEFITS

The security architecture addresses the new challenge of securing cloud production workflows and does it in a manner that is scalable and can be implemented incrementally. Shifting security certification to that of applications sidesteps the challenge of applying perimeter security models to cloud workflows.

The greatest benefits are secure and sustainable workflows. The workflows of the 2030 Vision Paper show what is coming, and a new security architecture will make them more secure and easier to secure and place the content owner in control of the security and the workflow.

There are many synergies between the workflow and the security architecture: the workflow enables the new security architecture, and the security architecture protects the new workflow.

- The authentication services that support the security architecture tie into audit systems recording not just general activity but potentially activity down to file access.
- When an asset is accessed through a link as the result of the resolution of the asset's name to a location, the link has intrinsic security, and access can be authenticated at the point of name resolution.

The 2030 Vision Paper envisages a virtualized production environment where work is spread among many individuals and groups working remotely from the main production.

The security architecture can be a catalyst for marketplace solutions to sourcing talent. By working in the cloud in an environment protected by the security architecture, individual artists and small companies can participate in the ecosystem without compromising security.

We expect that the security architecture will have limited dependence on security features offered by cloud providers, thereby enabling multi-cloud operation.

SECTION 9
CONCLUSION

The industry is at a turning point on cloud production. This presents both an opportunity to move to new workflows and an urgent need for a security architecture designed for production in the cloud. To achieve this will take vision and leadership by the content owners.

This document lays out an approach to security that is complementary to the 2030 Vision Paper, but whereas that document deals with changes in workflows and how creative work may be done in the future, this security document is designed for a different approach in which any workflow should be supportive without creative users hitting a security barrier. That notion of a scalable, dynamic security system that can support future innovations in applications and workflows is a core foundation as we work to build this system over the next few years.

SUGGESTED READING

SECURITY DESIGN

The Open Web Application Security Project.
https://www.owasp.org/.

Design Principles for Security, Benzel et al., 2005.
https://pdfs.semanticscholar.org/7830/cafce7da73aa5b137e2a5654f75877e306cd.pdf.

Usable Security: How to Get It, Communications of the ACM | November 2009, Butler Lampson.

When Security Gets in the Way, essay on jnd.org. August 2009, Don Norman.
https://jnd.org/when_security_gets_in_the_way/.

Systems Security Engineering Vol 1 and Vol 2 (US National Institute of Standards and Technology (NIST)):

1. Systems Security Engineering: Considerations for a Multidisciplinary Approach in the Engineering of Trustworthy Secure Systems, NIST SP 800-160 Vol 1.
2. Systems Security Engineering: Cyber Resiliency Considerations for the Engineering of Trustworthy Secure Systems, NIST SP 800-160 Vol 2.

ZERO TRUST

No More Chewy Centers: Introducing the Zero Trust Model of Information Security, Forrester Research, Inc, 2010, John Kindervag.

The Road to Zero Trust (Security), Defense Innovation Board (DIB). 9 July 2019, Kurt DelBene, Milo Medin, and Richard Murray. *Zero Trust Networks: Building Secure Systems in Untrusted Networks* (O'Reilly Media 978-1491962190), Evan Gilman and Doug Barth.

BeyondCorp: A New Approach to Enterprise Security, Google research publication, 2014, Rory Ward and Betsy Beyer.

This and other research papers on BeyondCorp can be found at https://cloud.google.com/beyondcorp/.

Micro-segmentation for Dummies, (John Wiley & Sons, Inc. 978-1-119-44854-9), Matt De Vincentis.

Available as a free download from VMWare. http://learn.vmware.com/41021_REG. Registration required.

FUTURE TECHNOLOGIES

Quantum Computing and Cryptography (Schneier on Security, September 14, 2018), Bruce Schneier.

https://www.schneier.com/blog/archives/2018/09/quantum_computi_2.html

Generative Adversarial Networks, Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville. and Yoshua Bengio. Département d'informatique et de recherche opérationnelle, Université de Montréal.

Available for download at https://arxiv.org/pdf/1406.2661.pdf.

DRM

Microsoft PlayReady Product Documentation

1. Microsoft PlayReady Content Protection Technology. Download
2. Developing Microsoft PlayReady Clients. Download
3. Content Decryption Module Interface Specification. Download
4. Interoperability, Digital Rights Management and the Web. Download

Widevine DRM

https://www.widevine.com/solutions/widevine-drm.

PlayReady and Widevine DRM are two of several DRM technologies in regular use. These references are offered to the reader wishing to find out more about DRMs; no endorsement of either PlayReady or Widevine DRM can be implied.

CONTRIBUTORS

This paper would not have been possible without the invaluable support and contributions from the following executives:

Anthony Anderson, *Universal Studios*

Shadi Almassizadeh, *Disney*

Mark Arana, *Disney*

Bill Baggelaar, *Sony Pictures*

Richard Berger, *MovieLabs*

Bryan Blank, *Sony Pictures*

Annie Chang, *Universal Studios*

Susan Cheng, *Warner Bros*

Daniel De La Rosa, *Sony Pictures*

Daphne Dentz, *Warner Bros*

Eddie Drake, *Marvel Studios*

Anthony Guarino, *Paramount*

Aaron Kim, *Disney*

Jaclyn Knag, *Paramount*

Jim Helman, *MovieLabs*

Arjun Ramamurthy, *Disney*

Horst Sarubin, *Universal Studios*

Craig Seidel, *MovieLabs*

Jason Shea, *Disney*

Adam Slohn, *Warner Bros*

Spencer Stephens, *TechXMedia*

Angel Stone, *Warner Bros*

Mike Trainotti, *Paramount*

Mark Turner, *Entertainment Technologists*

MovieLabs is a jointly run industry lab that enables member studios to work together to understand new technologies and drive the right ones to adoption. We help set the bar for future technology and then define specifications, standards, and workflows that deliver the industry's vision. Our goal is always to empower storytellers with new technologies that help deliver the best of future media.

Email: info@movielabs.com

Twitter: @ MovieLabsNews

Website: movielabs.com