THE BROADCAST BRIDGE
Connecting IT to Broadcast

Broadcast And The Metaverse

Essential Guide

EG
ESSENTIAL GUIDES

Introduction

By Tony Orme, Editor at The Broadcast Bridge

Maintaining backwards compatibility has been at the center of television technology since the first days of broadcasting. However, the new technology that is providing the foundations of the Metaverse is driving broadcasters to think about how they can make the most of these new opportunities.

The need to deliver programming to existing viewers while at the same time driving technology forwards has been televisions greatest strength and its Achilles' Heel. When color was developed engineers had to think of imaginative ways of simultaneously delivering monochrome signals, which led to color sub carrier and all its peculiarities, and when we moved to widescreen, we had to find methods of simultaneously delivering 4:3 and 16:9.

IP is not only an enabling technology, but also has the potential to allow broadcasters to break away from some historic working practices and deliver new and better immersive experiences for the viewer. One of the technologies that IP is providing access to is the Metaverse which has the potential to take broadcasting to new levels.

From its outset, advances in television have been motivated by improving the immersive experience. With each new generation of technology, either the sound or vision has been improved which in turn has delivered better experiences for the viewer. But there is now a whole new technology family that looks to take television to even greater heights, and that is the GPU based technology that is underpinning the progression of the Metaverse.

Viewers have become accustomed to passively watching television, but the 3D layer-on-top of the internet that the Metaverse is delivering, allows viewers to enter a new virtualized world. This might be a sports stadium or concert venue where the viewer can explore and walk around the event as if they were there. They might be able to choose their favorite seat or even meet virtually with their friends.

Although this may seem futuristic at the moment, vendors are ploughing massive resource into the development of the Metaverse. And not only does this have potential for broadcasters using the technology today, but there are many opportunities as we look to the future.

One of the challenges for the future involves deciding if television and the internet are still two separate entities. Traditionally this has been the case, however, the progression to IP distribution is starting to blur these lines. Some may think that television is working with the internet as television is being broadcast using OTT. But this is merely using the internet as a delivery mechanism and isn't embracing its full capability.

The progression of the Metaverse may well remove the blurring of the television and internet demarcation as the broadcaster's quest for delivering greater immersive experiences may have already made this decision. Layering a 3D virtualized world on top of the 2D internet has the potential to deliver exciting opportunities and is something broadcasters should be fully aware.

The Metaverse is starting to make in-roads into our everyday lives and it's only a matter of time before broadcasters will capitalize on the benefits for immersive viewing that it is sure to offer.

Tony Orme
Editor, The Broadcast Bridge

Tony Orme.

Broadcast And The Metaverse

By Tony Orme, Editor at The Broadcast Bridge

From the earliest public radio transmissions in the 1920s to 4K television of today, broadcasters have been working to continually deliver and improve the immersive experience. Television broadcasting has gone from black and white, to color, then HD and 4K, with massive improvements in audio all building on previous technologies to encourage viewer engagement and get closer to the event. The last hundred years, with all the technological advances in television, has taught broadcasters that viewers always want more.

Now, a new generation of ultra-high-power parallel compute and storage resource is forming the infrastructure to create the Metaverse that will deliver the ultimate in immersive and viewer experience. The Metaverse is rapidly driving adoption by bringing into existence a 3D overlay of the Internet to create a natural human interface that will provide an experience near indistinguishable from the real world.

The Metaverse end game may seem like a futuristic utopia, however, large parts of the enabling technology are actively being used today. And by

	VECTOR	RASTER
COMPRISED OF	MULTIPLE REFERENCE POINTS AND CURVES	PIXELS
SCALABILITY	SCALABLE WITHOUT QUALITY LOSS	LOSES QUALITY WHEN SCALED
CONVERTIBILITY	CONVERTIBLE TO RASTER	NOT CONVERTIBLE TO VECTOR

Figure 1 – Vector representation abstracts the image data from the display so that the scene is not dependent on the display (as is the case with raster images). The process of rendering the data will provide an image that can be displayed from the correct view-point and optimized for the target display.

adopting this technology, broadcasters will be empowered to develop the immersive experience that their viewers are continually demanding, thus driving greater viewing figures and potential growth.

Using todays Metaverse technology will bring viewers closer together for sports, concerts, and public events, and will not only create a greatly improved immersive experience, but also empower community engagement. This technology is significantly more than just enhanced VFX as ML, 3D animation, selective viewports, and avatars, to name but a few, have come together to allow broadcasters to massively improve their live programs and deliver the experience their viewers are demanding.

Virtual Worlds

One of the benefits of digitization of video tape into on-prem and cloud storage is that it has become much easier to process for AI applications. The data is parsed using convolution neural networks that learn aspects of the image to allow them to be classified, such as trees, cars, and highways. Which in turn provides the opportunity for image synthesis through AI technologies including generative adversarial networks (GANs).

In traditional broadcast workflows using green screen or LED wall technology, the designer has had to build the background images using recorded video or digitally generated images. This is incredibly expensive and time consuming as either a crew has to go out and record the GVs, which will need to be edited, or a graphic designer has to painstakingly create a background sequence frame by frame.

Contrast this to an AI method using virtual world technology consisting of high-end GPUs, servers and storage. Instead of creating a complete background, the graphic designer might create a viewport layered mask. For example, a scene with a car driving down the road might have one mask that outlines the road, another mask for the cars, another mask for the trees, and another mask for the building. The graphic designer would then parse the mask sequence through an AI image synthesizer that would replace the masks with the images it has created. This is far more than just replacing the masks with pictures of cars or trees as the images it synthesizes will be based on a library of vector mapped images that exist in 3D. Consequently, the images are represented as vector maps and only become visible at the point of image rendering, that is, when it is turned into a 2D image.

Due to the vector mapped nature of this technology, real-time viewport orientation can be applied. This allows the designer to move the visible rendered image so that it can be anywhere within the scene. This is further enhanced as the vector representation includes depth so that the viewport can be moved into and out of the image as well as around it. All in real time. Not only does this provide incredible 2D rendered images for the home viewer, but also has the potential to allow the viewer to enter into the scene, especially when using headsets, thus greatly improving the viewer immersive experience.

Broadcasting Future Virtual Worlds

Although the technology may be available now to help broadcasters create virtualized LED walls, the future offers many more interesting opportunities when considering how the metaverse is predicted to progress, especially when considering improving the immersive experience.

The metaverse can be thought of as adding a 3D layer to the 2D internet which in turn will allow viewers to enter the virtualized world that has been created. In broadcasting terms this could be a sports stadium where viewers swap seats and move from their homes to being in the stadium. Viewers then cease to be passive observers and instead become immersed in the game itself, allowing them to explore and experience the event almost at firsthand.

One of the key challenges for broadcasters is how do they take advantage of the future potential of the immersion that the metaverse promises?

In the same way that Google Maps have vehicles traversing roads around the world to provide street views, will broadcasters do the same with stadiums and venues? Thus, allowing them to build a massive library of images that can be used to form the virtualized worlds. Viewers could even choose their own seat in the stadium, go "backstage" to see how the game is being planned and the team strategies being adopted, and even join their friends to increase their enjoyment.

Once the historical limitations of thinking in a single viewport linear-2D environment have been removed, the true potential of moving and exploring within an interactive 3D virtual world becomes clear. And with this comes massive revenue opportunities for broadcasters.

Telestration and AI

Television is about telling stories and sports events are full of them. The slo-mo replay has provided viewers with analysis since engineers realized they could provide still frames using VT machines. Each event within a game can be replayed frame by frame to allow the commentators to analyze the players and provide a deeper insight into the game. And this immersive insight is further enhanced when Telestration is added to the mix.

Telestrators first appeared in the 1950s when physicist Leonard Reiffel used one to draw on a series of science shows for WTTW. Using analog storage CRTs and X-Y grid arrays, Reiffel was able to draw as he spoke to explain and educate. It didn't take long for the sports commentators to see the benefits of this technology and started to use it for major sports events. As technology improved, commentators were able to draw directly onto touch sensitive screens and even roll the video sequences backwards and forwards.

Pose estimation techniques used in ML are being used to predict the trajectory of the ball and the direction the players are moving in. This is providing commentators with an added dimension for improving storytelling and providing viewers with a deeper insight to how the game is progressing.

Telestrators are a rich source of information as the data they provide can be fed back into the real-time AI engines to enhance the richness of the data available, thus providing even deeper insights into the game. Highly enthusiastic sports viewers are constantly looking for new information about their favorite teams and players, and AI can analyze and find patterns much faster than humans. For example, "player A" may have run along the left wing five times during the three goals scored, although they may not be seen on the main program output, the AI engine would have recognized this as a pattern and would report it to the commentators and viewers.

The players themselves are also generating data and tracking technology is used to record the movements. This is much more powerful than a video recording as the players movements are decomposed into motion vectors that can be easier to record and process. This opens the door to more ML processing as the data is fed into the ML engine and used to generate statistics for the commentators and viewers at home. All these concepts are already possible and being improved upon by the Metaverse technology.

Adding to this additional layer of information is speech recognition AI where a viewer may ask a question, such as "who is player 17 and what is their stats?". The speech recognition AI will respond with the necessary information either on screen or as part of a voice synthesis response. It's even possible to provide avitars to provide the requested stats.

Broadcast and Avatars

Although mind-controlled avatars of the Metaverse are still in the realm of science fiction, AI driven facial and pose avatars are very much a reality. The first question must be, why would a broadcaster want to use avatars? Apart from providing new and interesting light entertainment shows where performers can adopt alternate egos with facial recognition cameras and motion detecting body suits, there are some interesting applications in news and current affairs.

24/7 news is here to stay but one of the challenges that this presents is that a team of presenters often must be in the studio all day and all night. While this is possible, it's also very inefficient. Instead, think of a Metaverse type digital representation of a presenter that looks, behaves, and expresses themselves like the original, but they're not the real presenter, just their avatar.

The Metaverse technology to achieve this exists now as facial recognition and motion positioning can be learned by AI engines. A script can be written that is input into the AI engine so it can then create a real-life avatar simulating the presenter, and the current Metaverse technology is so good that it's becoming increasingly difficult to tell the difference.

For the avatar to be truly convincing it must move smoothly and the facial imaging must be in sympathy with the context and substance of the script. AI can already determine the attention of written words and it's only a short jump to add this to the avatar generator to provide emotional content to the face and body motion. Admittedly in news, the upper body is only usually seen but most of the facial gestures are replicated so that the presenter looks real. The intonation of their sentence constructions can be simulated to provide a further level of immersion.

This opens the possibility of using avatars for signing for the hearing impaired. Many governments throughout the world now mandate signing for a percentage of the broadcaster's output, and this percentage is only going to increase. The synthesized facial and hand gestures of the avatar can be created to show very convincing signing avatars in real time.

Universal Scene Description

USD (Universal Scene Description) is a 3D framework for describing, composing, simulating, and collaborating within 3D worlds as it unifies workflows and file formats to provide a programming language that is at the heart of the metaverse.

There are three main components of USD, the user API, rendering engine, and the scene specification. The designer creates a scene using the USD format through the API which is rendered, often using ray tracing technology, to provide the final sequenced images.

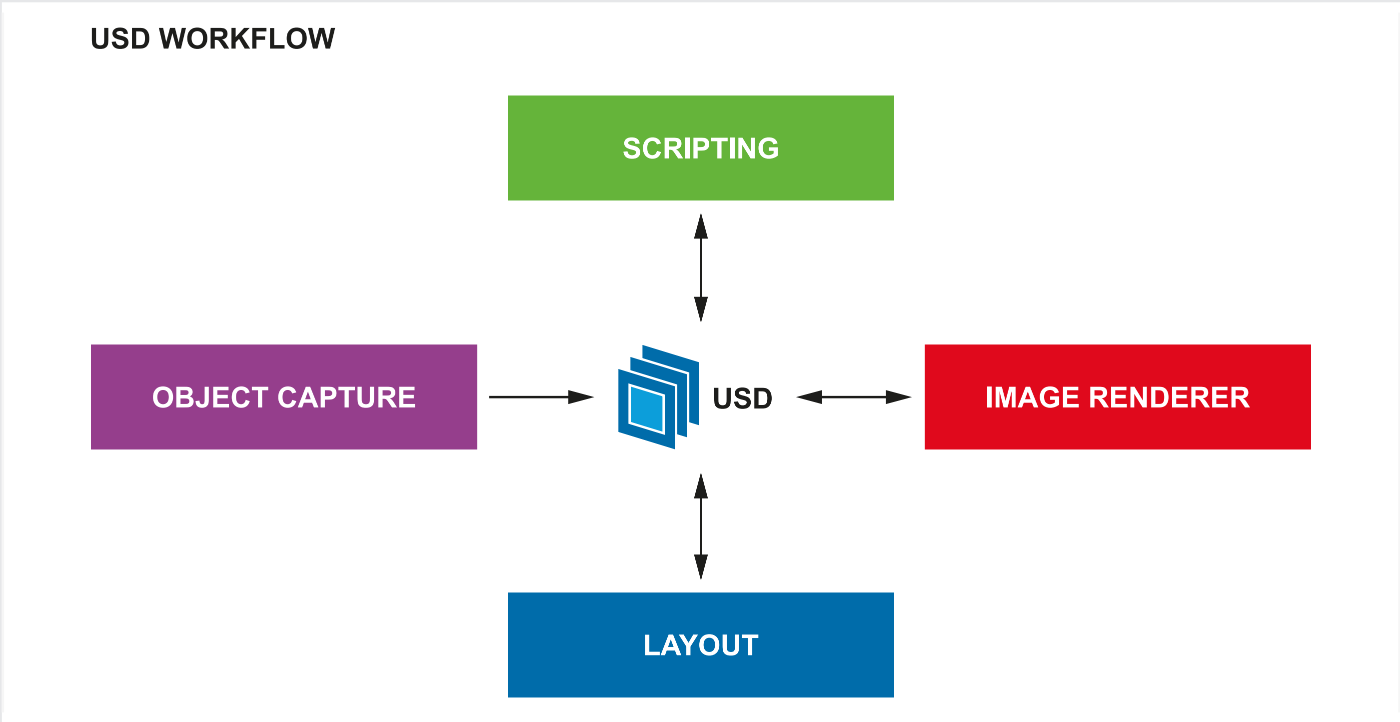


Figure 2 – USD provides a method of collaboration through open standards to create scene data for virtualized images. USD is used extensively in Metaverse design and is highly portable to broadcast television applications.

The scene can be compared to a theatre which includes the stage, props and lights where viewers observe the performance through the viewport of the stage. The scene is a database consisting of defined objects that can be layered to provide a hierarchical tree structure.

The USD creates a hierarchical structure of files that specify the scene so designers who are collaborating effectively share these files. For example, if one designer is creating the higher-level image of a street, another may be creating the representation of a car. When the car is finished, the USD description file is added to the hierarchical structure thus including it in the final render.

It's important to note that the USD files are not image files but text-style representation of the objects and layers that make up the scene. This allows the ray tracing rendering engine to adopt different viewports based on the designers' parameters.

The designer can also call upon vast libraries containing billions of USD represented images, and add their own attributes to provide an impressive array of potential image renders. This makes distribution of the images incredibly efficient as it's mainly the USD file that is being distributed and the localized rendering engine turns these relatively small files into the final images.

An interesting application for broadcasters using this technology is in LED wall virtual productions. Not only can the scene be created by designers collaborating from all over the world, but the scene has depth associated with it which can be rendered in real-time. This will allow the cameras to track into the scene to give a convincing depth of field.

Advanced Resource

The Metaverse requires a huge amount of storage for both creating the 3D virtualized environment and all the user data created by the user and this is delivering a major benefit for broadcasters. Data falls into a two-stage process, there is the creation of the virtualized environment and then the acquisition of the user data such as head and eye movements and hand gestures.

Creation of the virtualized environment requires extensive storage. As an example, one of the very earliest virtualized environments was Microsoft Flight Simulator that fitted onto a single floppy disk, the images of the instrumentation and views through the aircraft window were very primitive and blocky. Fast forward twenty odd years and the same product provides a highly immersive experience with Microsoft modelling 1.5 billion houses and over 2 trillion trees, this highly complex virtual world consumes around 2.5 petabytes of storage. Flight Simulator is highly constrained which only demonstrates the massive amount of storage that is being designed for the Metaverse world.

This expansion is not only seeing a massive amount of R&D investment in storage technology but is also seeing a significant increase in the processing power needed to make the virtualized world. We must remember that the user will also be generating massive amounts of data that has to be processed in near real-time as well as stored for fast access.

The GPUs needed to process data for the Metaverse creators are orders of magnitude more powerful than standard graphics computers. The servers hosting the GPUs are intrinsically aligned to them to not only provide visual rendering but also provide the AI processing capabilities that ML demands, and this is a core requirement of any Metaverse technology.

GPUs are not only used for their rendering capabilities but form an intrinsic part of any ML system. ML lays heavily on a branch of mathematics called linear algebra and GPUs with their ray tracing, shading, and parallel hardware accelerators use these functions extensively. GPU vendors have even started to include tensor manipulation hardware accelerators to further speed up ML training and inference. All this leads to hugely complex workstations and servers that need to be tuned to deliver real-time processing that the Metaverse technology is delivering.

Combined with the storage needs, the Metaverse technology is taking high-end resources to another level, which is something broadcasters are benefiting from both now and in the future.

Broadcast Reliability

Broadcasters work in 24/7 environments where the loss of program output is not something that can be tolerated. With so much high-value content and paying Ads being broadcast every day, the infrastructure must be resilient and reliable.

System reliability is another area where broadcasters can take advantage of the continuing development of Metaverse technology. To achieve the immersive effect, systems must be reliable both in terms of up time, and processing speeds. Although some latency is inevitable, research has demonstrated that humans can easily adapt to small amounts of latency if it is predictable and determinate. And Metaverse development is focused on delivering this as part of the immersive experience.

Conclusion

The Metaverse, as an absolute concept may still be work in progress, but we are starting to see layered 3D immersive experiences provided on top of a 2D internet. By design, the technology is leading the delivery of virtualized worlds and avatars as the infrastructure must be there to allow the creatives to generate the content. But broadcasters can benefit from the technology today and start trialing some of the enhanced immersive experiences that the Metaverse is delivering now and will continue to deliver and improve upon in the future.

The Sponsors Perspective

Metaverse: The Next Chapter Of Broadcast

By Cindy Olivo - Global Media and Entertainment Marketing Manager - Dell Technologies
and Sepi Motamedi - Global Broadcast Industry Marketing and Strategy - NVIDIA

The broadcast industry has been evolving towards greater immersion since its inception.

As technology evolves, so do the capabilities of broadcasters and content creators to bring audiences inside of their stories.

Today's immersive technologies and methodologies span the entire content pipeline from creation to delivery to consumption. They include 360-video capture, volumetric video capture, and spatial audio in content creation; augmented reality (AR) and virtual reality (VR) headsets for extended reality (XR) content distribution; and speech AI and data-driven visualization for interactivity during content consumption.

These technologies combine to deliver another layer to the viewing experience. This evolution of immersion in broadcast is taking us closer to the metaverse - the 3D Internet. The next step in this evolution is a shared virtual world. This is where audiences are headed, examples of which we're seeing in other parts of the media and entertainment industry, from gaming to advertising to film.

Technologies That Enable The Metaverse

Delivering experiences to audiences in a 3D Internet requires using several technologies, many of which are already in use by media companies today.

Accelerated Computing

At its core is accelerated computing to process vast quantities of data generated in creating, delivering and consuming the virtual experience.

The Dell Precision 7865 workstation combined with the NVIDIA RTXTM 6000 Ada Generation GPU delivers impressive computing capabilities ideal for content creation and dissemination in the metaverse. This computational powerhouse boasts up to 64 cores of processing power via its AMD Threadripper CPU and comes equipped with up to two RTX 6000 professional graphics cards. This provides the computational resources to render high-quality, detailed 3D graphics, to train and deploy AI models, and to encode video to multiple platforms, including head-mounted devices. The Precision 7865 can easily exceed all the stringent power requirements for these workflows.

Content creators in this space need workstations that can handle vast amounts of data and provide extreme processing power, to handle the complex tasks required to create immersive virtual environments, avatars, and interactive experiences. The Precision 7865 workstation equipped with the RTX 6000 is more than up to the task. Its quick processing and rendering capabilities enable content delivery that is both fast and efficient, ensuring that users have a seamless and immersive experience.

Extended Reality (XR)

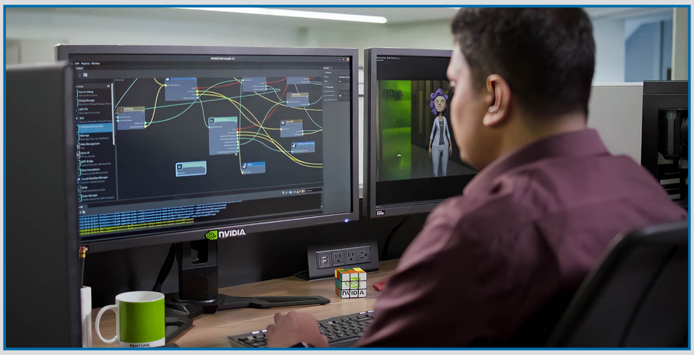
The Precision 7865 workstation supports a full range of professional NVIDIA graphics cards, making it capable of running high-resolution VR software and headsets. To provide customers with the best experience, Dell offers the Ready for VR program, ensuring customers choose a workstation and professional graphics card capable of driving superior VR experiences.

CloudXR, NVIDIA's streaming technology, delivers VR and AR across 5G and Wi-Fi networks. Built on NVIDIA RTX technology, CloudXR is fully scalable for data center and edge networks.

With the CloudXR SDK, extended reality content from OpenVR applications can be streamed to Android and Windows devices, dynamically adjusting to network conditions for maximum image quality and frame rates. This frees users from traditional VR and AR confines, streaming complex experiences from remote servers across 5G and Wi-Fi networks to any device, wirelessly.

Platform For Metaverse Application

NVIDIA Omniverse™ Enterprise is a scalable, end-to-end platform enabling enterprises to build and operate metaverse applications. Omniverse is a real-time, large-scale virtual world simulation engine. A computing platform that enables 3D designers and teams to better connect and build custom 3D content creation pipelines.



Omniverse unlocks the entire scope of today's 3D workflows – touching every single industry whether for building 3D assets and worlds or operating digital twins. Omniverse Enterprise is open, interoperable – built on Universal Scene Description – an open 3D framework and the foundation of 3D worlds, and MDL. It is easily extensible and customizable – customers can inspect, tweak, customize, build all our apps and extensions that we offer as source from scratch.

It is scalable from workstation including the Precision 7865 workstation, to data center, to cloud – and Omniverse performance scales just as much as the compute you throw at it. The platform can be deployed across hybrid infrastructure – and soon will be accessible from anywhere.

Avatars

Intelligent, lifelike avatars are a critical component of the broadcast metaverse. They heighten audience engagement in this new environment. Some examples of how interactive avatars can be used in broadcast include virtual news anchors, commentators and meteorologists that can be created to deliver the news in a virtual studio, virtual hosts who can guide participants through an event, and virtual celebrities and television personalities that can interact with fans about what they just watched.

NVIDIA Omniverse Avatar Cloud Engine (ACE) offers the fastest and most versatile solution for creating interactive avatars and digital human applications at-scale. Broadcasters can leverage

ACE to animate 2D or 3D characters, and to give them the ability to speak and interact with users. The ACE end-to-end avatar development suite enables seamless integration and deployment, allowing broadcasters to build, configure, and deploy avatar applications across any engine in any public or private cloud.

Artificial Intelligence

AI is one of the cornerstones of the broadcast metaverse. AI is woven into many of the applications mentioned above, but it deserves a specific call out. Speech AI, computer vision, recommendation engines, and more come together to accelerate content production, increase the accessibility of content as it gets distributed, and make content more personalized in its consumption. These technologies will allow audiences to navigate through virtual environments, interact with avatars, request the content they want to see or ask for recommendations, and more.

The Time To Build Towards The Metaverse Is Now

With its immersive and interactive nature, the metaverse represents the next chapter for content delivery and consumption in broadcast. With the increasing popularity of virtual and augmented reality and interactive visualizations, viewers are looking for immersive experiences beyond traditional TV and streaming. By creating virtual worlds and interactive avatars, broadcasters can further engage existing audiences, expand their reach to new demographics, and create new revenue streams.

As the metaverse becomes more accessible, it can revolutionize entertainment. Broadcasters must be at the forefront of this change to stay in step with audiences in the ever-evolving media landscape.

Cindy Olivo - Global Media and Entertainment Marketing Manager - Dell Technologies.

Sepi Motamedi - Global Broadcast Industry Marketing and Strategy - NVIDIA.

THE BROADCAST BRIDGE
Connecting IT to Broadcast

Themed Content Collection

EG
ESSENTIAL GUIDES

MEDIA

ARTICLES

For hundreds more high quality original articles and Essential Guides like this please visit:

thebroadcastbridge.com

03/2023

Supported by:

Dell Technologies NVIDIA