GTSystems™ WHITEPAPER

SPAN-AI UCDN

AUTHORS

Rhett Sampson, CTO

Jaime Llorca, head of R&D

March 2024

1. INTRODUCTION

The Internet, and particularly its core protocol stacks, TCP/IP and DNS, are reaching their limits of growth. The Internet has evolved over 60 years as a network of dumb, application-agnostic networks; with rudimentary protocols and interconnect capabilities; and no inherent global management, adaptability, or optimisation. Attempts to scale it by "patching" it with add-ons and workarounds, including the World Wide Web, the cloud, and content distribution networks (CDNs), have resulted in increasing complexity, redundancy, centralisation, and fragility. This is evident in the increasing occurrence of "the spinning wheel of death" (buffering) and dropped frames, unbearable interaction delays, and global catastrophic failures of major telcos, ISPs, CDN providers and hyper scale web applications. To quote Beck and Moore[1] "*The information super-highway to hell was paved with good intentions*"!

A new approach is needed to deliver the scalability demanded by Internet growth and to deliver new applications such as industrial automation (aka IoT); network AI; extended reality (XR); 6G; blockchain; distributed web3 gaming; and the Metaverse. These next gen applications have latency, caching and processing requirements that are beyond the capability of most legacy, centralised, hierarchical, networks. Many have their own protocols that inter-communicate via IP which does not understand them. These next gen apps require a new approach to networking; one where the network is the cloud™ and therefore the computer; that provides a global, unified, platform for coming generations; on earth; and as we return to explore and colonise space. A Network of the Future.

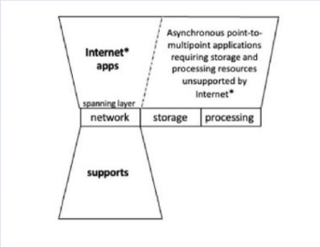
GT Systems has been researching and solving these problems for over a decade, working with leaders in the field at the Australian Commonwealth Scientific and Industrial Research Organisation (CSIRO); Protocol Labs; Bell Labs; UCLA; Hong Kong University at Guangzhou; University of Technology Sydney; the NDN project; and the Internet Research Task Force, Information Centric Networking Research Group; to design and build next generation content distribution network systems (ngCDN) that are content native, hyper-scalable, and intelligently interoperable. We call that the Universal Content Distribution Network (UCDN).

This whitepaper describes how intelligent, distributed, elastic, content native, autonomous, and self-optimising networks can be implemented alongside, and interoperable with, traditional networks, delivering substantial savings in network costs and improvements in efficiency and performance, even as the networks hyper-scale across a wide range of applications.

GT Systems is rolling out PoC and beta applications of the next generation UCDN based on our SPAN-AI technology. Our intelligent, fully distributed, and elastic network is autonomous and self-optimising. It is content native, works alongside, and is interoperable with, existing networks. It is the foundation for next generation industrial automation and augmented consumer experience applications. It will enable a new organisation of value on common infrastructure, illustrated in the diagram below:

[1] "How We Ruined The Internet" https://arxiv.org/abs/2306.01101

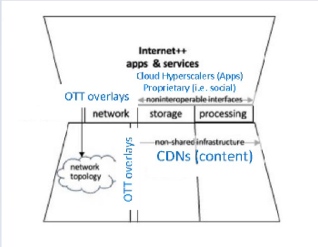
SPAN The network is the cloud™ – a new organization of value on common infrastructure



Original Internet

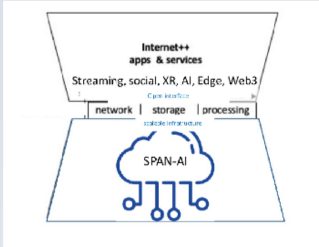
- IP Protocol stack
- "A network of networks"
- Reliable one-to-one Packet delivery
- Storage + Compute left to "upper layers" (App)

ORIGINAL GRAPHIC from "How We Ruined the Internet", Micah D. Beck et al Dept. of Electrical Engineering and Computer Science, University of Tennessee, Knoxville



Internet Overlays 90s – 00s – 10s

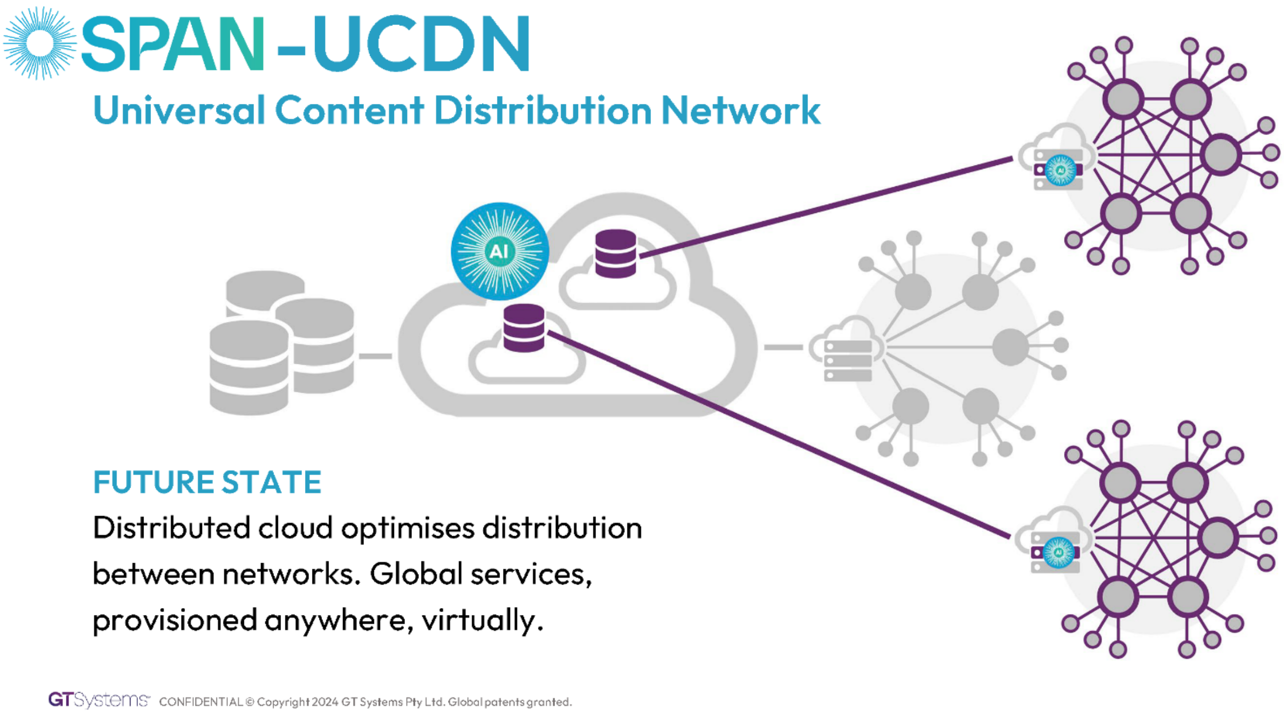
- Web to Social to Streaming
- CDN Overlays to scale Web then Streaming
- Apps require Storage + Compute + Overlays
- Rise of Cloud Hyperscalers
- Proprietary architectures
- Value flows to Infra (CDNs) and Cloud providers
- Cost structure posing challenges to Content Producers (studios)
- Service providers left with unending cycle of CAPEX investment vs stagnant / declining ARPU

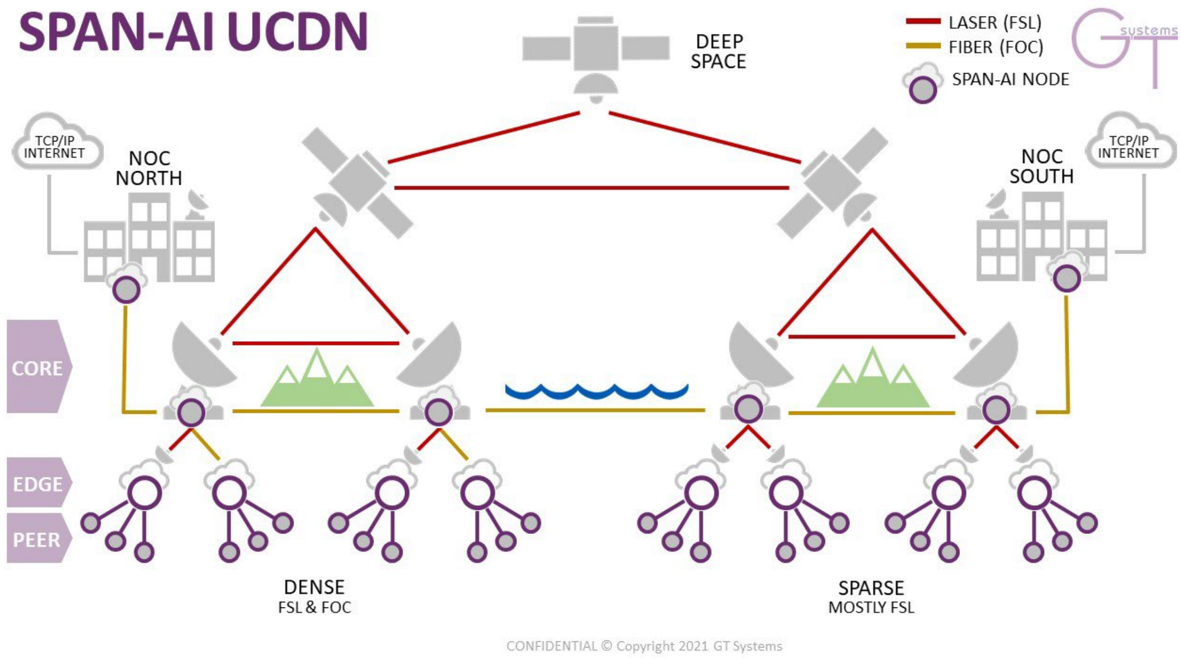


Internet Federated (20s & beyond)

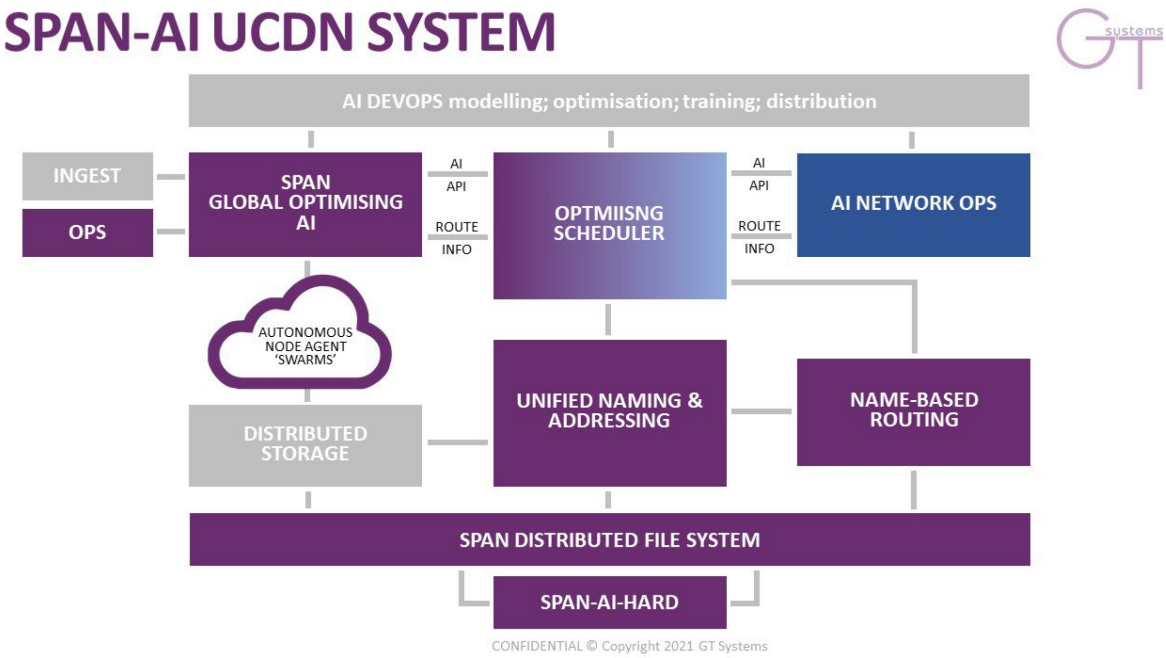
- SPAN-AI protocol stack
- "A federation of UCDN"
- Every node adds capacity to the system
- Positive economics for all industry participants
- Scalable CDNs, distributed Cloud
- Apps faster time-to-adoption: Experiences vs Infra
- Universal (global) common infrastructure
- Edge optimized: XR, AI, Web3, Autonomy, IoT
- Natively address Streaming industry (studios) concerns about distribution,

2. USE CASE DIAGRAMS





SPAN-AI Global Universal Content Distribution Network



SPAN-AI UCDN SYSTEM

3. BACKGROUND

The first wakeup call, unveiling the inefficiencies of current connection-oriented TCP/IP based network architectures, came from content distribution. The 2000's saw a rapid change in Internet usage, from point-to-point communication services or "conversations" between pairs of hosts, to shared content access.

Content distribution networks (CDNs) were developed as initial attempts to accommodate content delivery within the current Internet architecture. By deploying cache servers at geographically distributed network locations and storing content closer to end users, CDNs alleviated the congestion in the core network caused by multiple redundant point-to-point connections delivering the same content from origin servers to multiple end users. Route aggregation further helped reduce core congestion.

While CDNs have "sufficed" to provide acceptable quality of experience (QoE) for the distribution of relatively static web content, the increasing dominance of professional and user-generated video on demand (VoD) services - such as Netflix, Hulu, Disney+, YouTube - and more recently live video streaming - such as live sports/events, Meta(Facebook)/Instagram live, TikTok, Twitch - have exposed the inherent limitations of the current Internet and its overlay solutions. In essence, CDNs have moved the origin a few layers closer to the user, but what is known as the CDN "edge" is, at best, sitting at Internet exchange points (IXPs) at the telco core or in regional data centres. Redundant, content-unaware, point-to-point connections are still required to stream content from CDN "edge" caches, resulting in unbearable metro and aggregation network congestion and bandwidth costs, as well as server overload.

The exponentially increasing complexity of the Internet, CDN, and Domain Name Services is giving rise to more and more frequent failures due to configuration and deployment errors, both in CDNs and hyper scale web application networks. These failures usually have significant commercial consequences. It is only a matter of time before one has catastrophic consequences[2]. The Internet is becoming too complex for humans to manage alone.

This is only getting worse with the advent of next-generation system automation and augmented experience services, which not only increase the demand for high-quality video, but also add a new dimension of real-time interactivity, associated low latency requirements and increased complexity. The future of next generation networks is about real-time interaction with content, both for industrial and consumer applications. That future needs to be significantly simpler and more efficient.

4. SPAN HISTORY

In the early 2000's, the Australian Commonwealth Scientific Industrial Research Organisation (CSIRO's) head of R&D, Dr Terry Percival, was looking ahead to the next generation of networking solutions for the Internet. This "after-WiFi" research agenda included "peer assist" that combined peer to peer (p2p) techniques with an http "super server" and obtained very encouraging results for live video distribution.

Through his involvement in the program, and his own work in CD/DVD retail manufacture-on-demand (MoD) of movies and games, Rhett Sampson recognised that the peer-assist research had far-reaching potential benefits and applications. He founded GT Systems to further research, develop and patent Secure Peer Assist (SPA™), intelligent Secure Peer Assist Network (SPAN-AI™), intelligent Hybrid Adaptive Routing Design (AI-HARD™) and their ultimate expression, the Universal Content Distribution Network (UCDN).

At the same time other, complementary, global R&D projects were evolving:

- p2p and WebRTC (based on SCTP), combined with multi-CDNs, for live sport distribution (e.g. Peer5);
- distributed file systems such as the Interplanetary File System (IPFS) from Protocol Labs;
- distributed trust, blockchain protocols such as Bitcoin, Ethereum and Filecoin;
- structured hashed data models such as the Merkle Trees of IPFS and IPLD;
- content based networking such as Information Centric Networking (ICN), Content Centric Networking (CCN) and Named Domain Networking (NDN);
- multi-resource, graph-based network modelling by Jaime Llorca (now head of R&D at GT Systems) at the University of Maryland;
- and Jaime's work over a decade at Bell Labs, developing, modelling and optimising Software Defined Virtual Content Distribution Networks (SDvCDN).

5. A WORD ABOUT PATENTS

All the complementary, global R&D projects above fitted perfectly with the SPA and later SPAN-AI architectures described in our patents. Their integration with our patented innovations resulted in the UCDN architecture that is the subject of this white paper. Our intent with patents is not to stifle innovation. It is simply to protect our IP and its

[2] https://www.abc.net.au/news/2024-01-23/optus-network-outage-triple-0-emergency-calls-wider-than-thought/103380928

implementation, while receiving a reasonable return on some very significant investment, both cash and personal. We are very keen to find new models for R&D that combine some of the community incentive models above; recognising the shift from public to private R&D; and the role of voluntary R&D such as the IETF and IRTF. We will use the considerable revenue generated to provide stellar returns for investors and transform the Internet into a sovereign, self-funding body to protect, preserve and grow the most valuable man-made asset on the face of the planet.

6. ANOTHER WORD ON SOFTWARE DEFINED NETWORKS

Every network these days is a "Software Defined Network" (SDN). There are a myriad of versions of what that means. The real question becomes: "What is the optimum SDN"? The answer isn't simple. This white paper proposes what we think is the answer to that question.

7. DESIGN OBJECTIVES

A number of requirements were identified as necessary to deliver the increases in performance and efficiencies required for next-generation content delivery. These were incorporated into the design objectives so that SPA and SPAN-AI would:

1. be fully distributed; distributing and optimising storage (caching) and compute in the network
2. be content aware and switch/route based on that content awareness
3. be autonomous and self-optimising
4. be inherently secure
5. enable real time streaming of live and on-demand video with no spinning wheel of death (SWoD) or dropped frames and with minimal delay
6. enable next generation network applications such as real time gaming, web3 gaming, industrial automation and augmented consumer experiences such as extended reality and the metaverse
7. move from stateless, connection oriented, "dumb pipes" to stateful, connectionless, content-oriented, intelligent networking
8. remove costly and fragile DNS lookups and network bottlenecks
9. be able to inter-operate with a wide range of distributed storage technologies and platforms
10. be dynamic and reconfigurable (virtual/elastic)
11. make storage and distribution simple and transparent to content owners and publishers
12. be compatible and interoperable with existing TCP/IP networks
13. be open and intelligently inter-operable with any network using the SPAN-AI architecture and protocol stack, forming a Universal Content Distribution Network (UCDN)

8. EFFICIENCY

Our first generation, Secure Peer Assist (SPA), was initially focussed on movies. Video is still 80% of the problem and growing, but next-generation applications are increasingly imposing their own demands. Our modelling has proven that our current generation, Secure Peer Assist Network (SPAN), will continue to have much wider and more general applicability. Our modelling has also shown that the SPAN-AI design criteria deliver performance and efficiency increases resulting in substantial savings in network implementation and operational costs, even as the network hyper-scales across a wide range of network applications. Our early modelling has shown savings of 10-100X amounting to hundreds of millions of dollars for a single network; and latency improvements of 5-10X. This is only with the first stage of SPAN-AI implementation. There are two more stages that are projected to deliver still further significant additional cost savings and scalability.

9. FUNDAMENTAL PRINCIPLES

GT Systems' granted foundation patents for Secure Peer Assist (SPA and SPAN) form its core operating principles. Subsequent patents develop implementation and integration of these concepts into SPAN-AI and UCDN. These fundamental principles are the subjects of this paper:

1 ***Fully*** distributed. ***Every*** node is a "peer" incorporating the SPAN application (agent)

2 with storage and caching fully distributed in the network

3 with intelligent, distributed routing, caching, and resource allocation decisions supported by centralized global AI guidance

4 using hash tables or other efficient database mechanisms

5 switching/routing at the video packet level i.e. content-based routing (of video packets and other content)

6 combining distributed and centralised routing information and ***intelligence*** down to the video packet level

7 using global content addressing unifying a hash and a name. We call this a hashname.

8 each network appliance forming an ***intelligent*** node in a mesh network

9 that may be implemented as a virtual machine

10 described as distributed storage and distributed cloud computing

These are the claims of our granted global patents. In the words of the US patent examiner: "*The claimed invention is directed towards a comprehensive and multi-faceted content delivery system that is fully artificial intelligence (AI) driven, where AI components handle every facet of the content distribution means. These include naming, discovery, routing, publication, subscription, and security of content and network resources… no prior art found appears to teach or reasonably suggest the claimed invention, particularly the reliance upon a scalable hierarchical AI network consisting of numerous subsystems with additional embedded AI agents at every level of the distribution network.*" The patent is granted.

10. OUR INNOVATIONS

SPAN-AI is based on three simple innovations:

1. Tagging of content packets (e.g. video slice) with a hashname
2. Switching/routing of tagged content packets ***at the network level***
3. Application of Machine Reinforcement Learning (MRL) and Stochastic Optimisation (SO) ***inside*** the network to caching, switching/routing, and operations

11. DISTRIBUTED SYSTEMS NATURE OF SPAN-AI

SPAN draws on the best aspects of modern distributed systems while avoiding their limitations. SPAN autonomous agents operate on every node in the network. SPAN agents intelligently manage caching, operations, and switching/routing in the network. Operators download the SPAN-AI agent onto ***any open platform*** (e.g. X86; Tofino; Nokia SR Linux 7200; Broadcom switch; etc.). The operator stakes SPAN tokens to join the network. Proof of caching, processing, and routing earns SPAN tokens. Global optimising AI incurs fees. Proofs are very low cost and real time. Larger customers e.g. telcos pay an up-front fee.

12. PHYSICAL ARCHITECTURE

12.1. PRINCIPLES

Flattening

Networks globally are "flattening". That is, they are trending towards hyper-scale layer 2 Ethernet networks in data centres, running a mix of open and proprietary protocols on increasingly "white label" open switches and servers. Local, customer access networks are trending towards layer 2 Ethernet and layer 2.5 MPLS networks connecting regions via points of interconnect (PoIs). In between are legacy telco and CDN networks with centralised, "unified core TCP/IP" architectures and ***many*** intervening layers that essentially only add delay and cost, and reduce performance. Hyper scalers are "bridging" their layer 2 networks, mostly using VxLAN, over these intermediary layer 3 networks. OTT service providers and CDNs are still using TCP/IP at the "edge". This is all ***incredibly outmoded and inefficient***.

What is needed is a universal network and content protocol that works natively at layer 2 or 2.5 in the data centre or local access network and over layer 3 in the wide area, as we migrate to next gen networks and protocols. Ideally, that protocol would be part of a system that is "content aware" and intelligently distributes storage and caching into the network, such that the distinction between network, storage and processing disappears. ***The network becomes the cloud™*** and it is fully optimised.

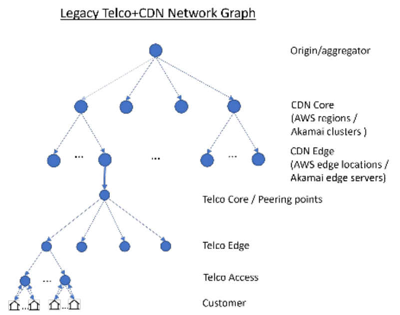
SPAN-AI is that protocol and system.

Inversion

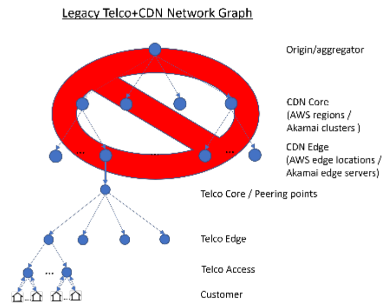
Most network and cloud service providers currently view large regional data centres or Internet Exchange Points (IXPs) as "the edge". We regard the regional DCs as the ***top*** of our network and the IXPs as gateways at best. In order to meet the QoS required by next gen applications, the network needs to be flattened (see above) and inverted, with at most two or three hops from the top of the network to the consumer. Relevant content needs to be optimally cached close to the consumer, according to demand, in real time.

Counter-intuitively, this next generation network architecture results in very significant capex and opex savings and reductions in latency. Fortunately, there are next generation service providers and equipment vendors building the infrastructure required for this and we are working with them. Legacy networks can also be transitioned in stages and transformed into next generation networks, and we are doing that too.

MOST NETWORKS LOOK LIKE THIS



WHEN THEY SHOULD LOOK LIKE THIS
to enable next gen applications: IoT; AR; Metaverse



Access

Layer 2

The ideal fixed local access network is regional layer 2 Ethernet via coax or fibre. High speed Ethernet provides an ideal SPAN network for applications such as video and gaming. SPAN is a native layer 2 broadcast protocol with mechanisms built in for load balancing, loop prevention, etc. Local points of interconnect form our access layer and the true edge of the SPAN network.

Layer 2.5 MPLS

SPAN can also work over layer 2.5 MPLS networks, interconnecting via SPAN multi-protocol service routers/switches to form a seamless SPAN network. Customers on different MPLS VPNs appear as a single, homogenous SPAN network.

Backbone

SPAN requires a high-speed backbone to meet next gen application SLAs. Terrestrial and undersea fibre cable services are ideal for this, as are terrestrial and space based (satellite) free space optics (lasers). Any optical switching or multiplexing is advantageous, as is AI based network optimisation e.g. to route around weather, etc. SPAN can interface to the backbone network AI to optimise whole of network virtual service placement and content caching, or it can simply operate over the layer 2 backbone as presented without deeper knowledge of the network. Ultimate benefit comes when we integrate both.

12.2. CORE AND DISTRIBUTED ORIGIN

The core of a SPAN network is the "edge" of most legacy networks and the location of many CDN "distributed" origin servers. SPAN fully distributes the origin servers to the ***real edge*** and adds storage to ***every node,*** optimally caching content throughout the network. The network ***is*** the cloud™. We are literally turning network architectures "on their head". Content owners publish ***once*** to the network and SPAN takes care of secure, distributed, storage and caching to meet ***specified SLAs*.** This starts with moving published content from ingest to optimum core, access, and edge nodes. Every SPAN node has processing, storage. and routing capabilities to varying degrees. Distribution will depend on contracted QoS, current and forecast demand, network load, and state in real time. Big publishing events can be flagged and content securely pre-cached in the network. At release time, all a consumer may be doing is buying an access code to a locally stored, secure, piece of content. Live events are "cascade streamed" via SPAN's inherent streaming architecture. Core nodes act as local distribution points and may be regional or edge data centres or other network infrastructure. The more participating nodes, the better SPAN works.

12.3. ACCESS EDGE

The access network is the new edge of next gen networks. More often than not, it will be one of the ideal locations to cache and process content, reducing backhaul, storage, and compute costs and significantly reducing latency/delay/lag. It will require dynamic optimisation and reconfiguration of service processing and caching in real time, according to demand, load, and network state. This will require very different network nodes and intelligence at the edge, communicating network status, load, and demand to a global optimising AIs from the intelligent, autonomous SPAN agents on every device. It will also be responsible for localised, "edge training" of the SPAN MRL agents.

12.4. PEERS / FAR EDGE

Any node in a SPAN network with a SPAN agent can be a "peer", but typically this term is used to describe consumer appliances (including industrial consumers) or consumer premise equipment (CPE) hosting SPAN agents at the "extreme edge" of the network. This could be as far out as space, Mars or beyond. On earth it will largely comprise phones, laptops, game consoles, headsets, and other UI devices at home, work, and on the go. It includes CPE such as home gateways, streaming appliances, wireless modems, home routers, and combinations thereof. In industrial settings, it will comprise the billions of automated devices controlled by the network. This includes autonomous vehicles, with their demanding enhanced vision and control requirements, and what we call robots today.

Peers run next gen applications enabled with SPAN-AI via SPAN SDKs distributed to application developers, e.g. Disney+. This enables the peer applications to communicate natively with the SPAN-AI network and the wider Universal Content Distribution Network (UCDN). Peers equipped with SPAN-AI can also spoof IP for applications that do not yet have the SPAN stack.

Peers essentially anchor and characterise demand at the extreme edge of the network. They provide the information via the SPAN agent that is required by the rest of the network to operate efficiently. They are the ultimate end-user for the services provided by the network and originators of data critical to control and operation of the network. They will increasingly also be originators of content for distribution by the network, e.g. User Generated Content

(UCG) and industrial automation sensing data[3]. Once again, they will be very different from today's consumer devices. Many of the network service functions characterised as "peer to peer" today will push back to edge access devices, some of which will be located in the home/factory/habitat and others in fixed and mobile edge access points.

12.5. INTERCONNECT TO OTHER NETWORKS

TCP/IP

SPAN is designed to be backward compatible and interoperable with TCP/IP. SPAN-AI-HARD™ is our AI-driven, hybrid, adaptive routing design that combines the benefits of name-resolution-based routing and storage (i.e. routing by resolving names and addresses by lookup, e.g. IP and DNS or hashes and IPFS); combined with name-based-routing and distribution, i.e. intelligent local and global routing by content names and interest without costly lookups. The unified naming and addressing system is what glues this all together.

At the simplest level, we can "tunnel" SPAN-AI protocol traffic over TCP/IP networks that do not have SPAN-AI routing capability using existing methods such as datagrams, segment routing, etc. However, SPAN-AI is designed to work as a dual protocol stack, alongside TCP/IP, on open routers and switches. Dual stack nodes can route SPAN-AI packets natively. Where a SPAN-AI route is not available, SPAN-AI content streams can be packaged as TCP/IP connections and tunnelled over the TCP/IP network, analogous to what is happening today. A major focus of our work will be to integrate SPAN with TCP/IP using existing mechanisms such as QoS fields in IPv4 and IPv6. This will allow inter-operation of legacy IP networks with next gen SPAN-AI networks.

NBN

Many countries have built national broadband networks to bring the benefits of high-speed broadband to most of their population. In Australia, this is being built by NBNCo. It is a fibre optic, coax, and satellite layer 2 and 2.5 MPLS network, currently connected to 80% of the population. GT Systems has developed detailed designs to inter-connect via SPAN service routers/switches to the Australian NBN to create a single, homogeneous, SPAN-AI UCDN. Service Providers can offer any SPAN virtual service without owning any infrastructure. We are partnering with network equipment vendors and NBN operators to build specific SPAN service routers/switches for each network. In Australia, we are working with Nokia, who built the NBN; and the Broadband Technology Research Unit (BTRU) at the University of Technology Sydney (UTS), led by Professor Ray Owen, a previous CTO of the NBN. Working with the BRTU we will build a model for every layer 2 "cable" operator to implement SPAN-AI.

National networks

In countries without national broadband networks, there are usually a small number of large carriers with connection to a large percentage of the population. There are only a very small number of network equipment

[3] https://urgentcomm.com/2023/08/05/the-edge-vs-cloud-debate-unleashing-the-potential-of-on-machine-computing/#:~:text=In%202025%2C%20Gartner%20expects%20that,be%20accelerated%20by%20the%20edge.

vendors to these carriers. We will very quickly develop interconnect design patterns for all of them, enabling rapid reach to large segments of the global population.

For countries and regions with sparse telecommunications infrastructure, we will use a mix of terrestrial and space (satellite) free space laser networks to provide broadband UCDN and Internet access.

Hyper scale DCs and the evolution to the Metaverse

Most hyper scale data centres (DCs) run high speed Ethernet networks and a mix of proprietary and open protocols on open switches and servers. At the simplest level, we can ingest content from the DC into the SPAN-AI UCDN for publication (storage) and distribution via the network. At deeper levels of integration, we can licence/build SPAN-AI technology into the DC to make it a SPAN-AI network node on the UCDN (see fig. 1c).

The Metaverse will exist initially on, and in, hyper scale data centres. But it must be experienced in real time, in augmented reality, by unlimited audiences, at the edge. The only way to achieve that is with a SPAN-AI UCDN with fully distributed processing. This will require co-operation with Metaverse developers and enablers to distribute core processing to the edge. SPAN-AI is designed for exactly that.

Emerging Platforms

Platforms are emerging to enable this. Pixar's Universal Scene Description (USD) [is]… the first publicly available software that addresses the need to robustly and scalably interchange and augment arbitrary 3D scenes that may be composed from many elemental assets." https://graphics.pixar.com/usd/release/intro.html

Nvidia's "Omniverse [is]… a scalable, multi-GPU real-time reference development platform for 3D simulation and design collaboration in industrial automation… based on Pixar's USD and NVIDIA RTX™ technology." The beauty of this is that it allows the exchange of industrial Metaverse assets, worlds, and changes in those worlds (e.g. a robot's movement in a factory) in a highly efficient "metalanguage and metadata" that significantly reduce bandwidth requirements and ensure interoperability. In short, Omniverse is the highly sophisticated and efficient ontology or "language" of the industrial metaverse. https://developer.nvidia.com/nvidia-omniverse-platform

We also need a "source of truth" for the Metaverse, a repository of everything that makes up the Metaverse and its sub-universes that enables publishers to "share and modify representations of virtual worlds". We could keep those sources of truth in their respective sub-universes and use USD and Omniverse to exchange information about them. Or we could do what Nvidia have done and have a central repository, Nucleus, https://docs.omniverse.nvidia.com/prod_nucleus/prod_nucleus/overview.html

Nucleus solves the N^2 interface problem (for 100 sub-universes to exchange data with each other we need 100^2 or 10,000 interface connections, speaking all their different languages). But it raises questions of scale. A central repository would have to be VERY BIG and VERY FAST to cater for the entire "Metaverse of possibilities". This could be beyond even the current capabilities of Nvidia super-computers, at least in a centralised architecture. The solution to this would be a ***network*** that can store the Metaverse assets and understands Omniverse commands. That network is SPAN-AI™. We "speak Metaverse" and we optimise content caching and routing ***in the network***. The network *is* the cloud™. We also speak native blockchain and Merkle Tree; in fact, any content-based storage and distribution "language".

Where SPAN-AI fits in the Metaverse

SPAN-AI™ is fully distributed, elastic, content native, and intelligent. ***EVERY*** node in a SPAN-AI™ Universal Content Distribution Network (UCDN) is autonomous and self-optimising. Every device on the network, whether it is a server, router, switch, gateway, PC, phone, VR/AR headset, etc. is a compute, storage, and routing element of the network. That makes the SPAN-AI™ network a giant, distributed Metaverse computer, far more powerful than any centralised super-computer, and far more efficient, because it is as close to consumers as possible. SPAN-AI™ fundamentally understands commands such as "publish" and "subscribe" (the core commands of Nucleus) and also understands Omniverse/Metaverse commands to create and share content. In short, SPAN-AI is the "glue" of the Metaverse that brings all the sub-universes together by speaking native Metaverse and caching Metaverse content and assets in the network. Nucleus would be a "super-node" in that network, as would every other world.

Virtual movie and world production

Hollywood has been grappling with these production issues for decades. Many modern movies are a mix of live actors and scenes with "green screened" computer generated characters and scenes. Motion Picture Laboratories (Hollywood's super techies) have published a vision for a future production platform that has many features in common with the Metaverse described above. It lays out a path for "The Evolution of Media Creation…. to enable seismic changes in media workflows with one objective in mind – to empower storytellers to tell more amazing stories while delivering at a speed and efficiency not possible today*"*. https://movielabs.com/production-technology/

A perfect example of this is Sony's recently released Ghostbusters machinima clip[4]. Produced using the Unity game engine and an unspecified AI, it produced in a day what would have taken weeks to shoot in real life in New York. The next step is to render this locally in real time. Web3 game producers are working on this already. It requires a fully distributed web3 game engine with local rendering and a distributed network capable of understanding and distributing complex abstract scene representations. That is SPAN-AI.

We can do all this now. We just need to "join up the pieces" and build it. The good news is that GT Systems is working with pioneers to build a global, next gen network to do exactly this. We have the foundations of a set of standards. We can extend them to cover this vision. More importantly, we have an architecture and high-level design to build it. This white paper is, in part, a call to the movie, gaming, blockchain and metaverse industries to do exactly that: come together to define the standards, platforms, and network on which to actually build the Metaverse; and then to build it.

[4] https://ghostbustersnews.com/2023/08/15/sony-shares-new-ghostbusters-short-film-made-using-real-time-game-engine-technology/

Toni Parisi from Unity has published a good set of "rules": There is ONE Metaverse; it is open; it is [enabled by] a network; [that network] is the Internet.
https://medium.com/meta-verses/the-seven-rules-of-the-metaverse-7d4e06fa864c

Satellite

Next generation satellites, employing free space laser optics instead of radio, solve many of the problems previously associated with satellites, and open up whole new possibilities. All-optical satellites enable high-capacity, multiplexed, secure, broadband backbones; cross-border and cross-geography. These are ideal as backbones for the UCDN and perform as an equivalent infrastructure to perfectly complement terrestrial fibre optic networks. Inter-connect will require very specialised, integrated ground stations and we are working with operators to build these. AI operational control and optimisation of the layer 2 satellite network adds another level of optimisation to the network.

Laser satellites also open up space communications as we return to the moon, Mars and beyond in coming decades. We will need to get high value content back to earth and ideally provide all the services we enjoy on earth, but in space. "Google space" on your 6G space mobile anyone? Next episode of Asimov's Foundation on your laptop in your Mars habitat? We will also need to carry out remote controlled robotic mining on the moon. The only way that can be done is by replicating the UCDN in moon orbit and on the surface of the moon. Networks of laser satellites running SPAN-AI will form the space backbone component of the UCDN. It will be capable of operating autonomously for the inevitable periods of disconnection, re-synching as connection to the "mother network" is re-established. New episodes of Foundation will be scheduled for delivery alongside high priority broadcasts of landing on new worlds. Mine vehicle operators will remotely control vehicles via the space component of the UCDN.

Combined with terrestrial, free space, point to point, laser optic networks (see below), satellites also help solve the problem of getting Internet access to the 40% of the world that currently does not have it.

Fibre optic cable

Fibre optic cable (FOC) is the perfect terrestrial backbone for the UCDN. Layer 2 SPAN enabled switches will interface directly to it at line speed via high-speed optical interfaces. Optical multiplexing and/or switching will add another level of optimisation to the network. We are currently working with Nokia to provide this capability and will work with all equipment vendors and hyper-scale platforms to build out the UCDN.

Point to point free space laser optics

Free space laser optics (FSO/FSL) have a multitude of applications, terrestrially and in space. As for fibre optic cable, Layer 2 SPAN enabled switches will interface directly to FSL at line speed via high-speed optical or Ethernet interfaces. Optical multiplexing and/or switching may add another level of optimisation to the network. We are currently working with terrestrial and satellite operators to provide this capability and will work with all equipment vendors and hyper-scale platforms.

Free space lasers present unique service opportunities, both at the optical transmission and network level. FSO/FSL are highly complementary when converged with fibre networks in both spectrum (1550nm), and service modes (WDM). Partner software solutions are also addressing the most pressing service challenge – atmospheric interference – and are ready for demonstration.

FSO/FSL applications include: horizontal lateral pointing between a ground station and a premise or data centre; from ground station to satellite; and inter-satellite links relaying high throughput data faster than terrestrial fibre over long distances for improved long-haul latency. FSO/FSL are a proven solution with significant benefits for high throughput, elastic, next generation data networks.

CDNs

Even the oldest legacy CDNs can be extended and easily transformed into next gen networks, becoming a significant part of the UCDN. Next gen CDN services can be provided virtually on the UCDN without owning infrastructure. Existing infrastructure can be simply transformed by the addition of SPAN-AI agents. Merging of multiple CDNs into a unified network (UCDN) allows the transition to pure throughput, distance independent, charging models. Newer CDNs are already implementing the first level of our innovations: tagging video packets with a hash and then "reverse proxying" them. We are working with them to implement the second and third levels of SPAN innovation.

The Network of the Future

The array of Internet, telco, CDN, web, and cloud services offered today to consumers, business, and government are invariably geographically and commercially fragmented. As a result, co-ordinated, global, provisioning, operation, maintenance, optimisation, and management of networks is a nightmare that falls to content owners and distributors to manage. The exponential increase in data flows and the shift to data distribution to the edge, underscored by COVID19, has stretched these services to breaking point. As we return to, and further explore, space, these problems will be exponentially exacerbated.

Next generation, distributed trust networks such as Bitcoin, Ethereum, Filecoin, etc, are raising fundamental questions around sovereign identity; ownership of data; governance of the Internet, the monetary system, technology businesses, and society itself. Commercialisation of space is leading to new challenges. It's time to take a universal view and produce the platforms for next generations to communicate simply and effectively wherever they are. The primary impediments to such a unified network are cost and architecture. This white paper addresses the latter. The network we are building addresses the former. It will provide a platform for the provision of network services for society for the foreseeable future.

13. SECURITY

The SPAN-AI UCDN is designed for security from the ground up. It supports centralised and distributed trust models, sovereign identity, and sovereign data ownership.

13.1. Encryption and DRM

Ingest

Any content can be encrypted at ingest and remain encrypted on request. High value content is encrypted at ingest by default. It remains encrypted in transit and at rest until unlocked by a key controlled by the publisher and tied to the user and platform. For example, movies are encrypted using Motion Picture labs (MPLABS) approved ingestion systems and DRM such as Microsoft PlayReady. GT Systems was the first in the world to achieve Hollywood studio approval for high value (4K UHD) studio master ingest for distribution on open Windows/Intel platforms using peer to peer technology. Encryption and decryption are anchored in hardware root of trust and trusted operating environments. Keys are managed by the publisher-controlled DRM system with additional security provided by UCDN sovereign identity and laser quantum key distribution.

13.2. Storage and distribution

Once encrypted, content remains encrypted as it is stored and distributed throughout the SPAN-AI UCDN network. The network was designed from the ground up to meet and exceed demanding standards of content owners and publishers for security and encryption.

13.3. Root of trust and trusted execution environments

SPAN DRM encryption and decryption is anchored by trusted platform hardware modules (TPM). A wide variety of chip architectures implement hardware TPM. TPMs enable the building of trusted execution environments in which DRM keys can be managed and encryption/decryption performed securely. Laser networks provide highly secure key distribution and quantum key management.

13.4. Trustless storage and distribution

Next gen distributed storage networks rely on “trustless” or “zero trust” security models. That is, public-key cryptography and a consensus mechanism such as a blockchain that distributes trust to reach consensus. SPAN-AI is designed to support this security model in both storage and distribution of content. This enables support of blockchain based systems. The UCDN is also designed to evolve to a distributed trust model to enable reward for contribution of resources to the network.

13.5. Quantum key distribution

An end-to-end, controlled, laser light network provides opportunities for secure quantum key distribution and, eventually, quantum information distribution. Once quantum keys can be distributed easily and efficiently, they enable the use of higher security symmetric key algorithms such as AES, rather than public key encryption algorithms.

13.6. Proofs

“Trustless” security relies on proofs. To date, these have been complex, slow, and expensive, both in cost of transactions and the power required to verify and maintain those trust networks. This limits scale, and the work required is also often not actually productive. SPAN-AI’s distributed technology improves the power efficiency of legacy, centralised networks. SPAN-AI proofs of caching and switching will be designed to be lightweight and efficient.

13.7. Space time storage

The Filecoin distributed storage network already employs a proof of storage over time (space time). If we use IPFS as-is, we may need to us the Filecoin proof. However, there are several possible improvements to IPFS in scaling and distribution that will require a new or modified proof. Ideally, we would develop these in collaboration with Protocol Labs.

13.8. Routing/switching

As it evolves to a fully distributed, fully autonomous network with contribution of resources by anyone, the SPAN-AI UCDN will require a routing/switching proof. Just as SPAN-AI's distributed technology improves the power efficiency of legacy, centralised networks, SPAN-AI routing/switching proofs will be lightweight and efficient.

13.9. Green power

While SPAN-AI's distributed technology improves the power efficiency of legacy, centralised networks and enables co-location with renewable, green power, that is not a guarantee in itself that the network uses renewable power. Today, this is based in certification by traditional certifying authorities. Since SPAN-AI nodes will support both traditional and distributed security models and will likely be present at many renewable energy locations to enable next gen industrial automation and communications, they may play a useful role in developing green power proofs.

13.10. Data centric security

Like many aspects of TCP/IP, security has been "grafted on" after the fact. Content, or named data based networking, provides significant inherent improvements in security. The threat/attack surface is significantly reduced. Secure connection is managed at the edge, based on sovereign identity models. Data centric security is lighter weight, more robust and able to traverse networks without security loss. Distributed networks are much harder to attack. This provides the basis for many security innovations that will be the subject of a separate white paper.

13.11. No DNS

ANY node in the SPAN-AI UCDN is capable of resolving a content name to a publisher, either internally, by network interest request, or via the intelligent, hierarchical, unified name resolution systems. This integration of content name resolution and name-based systems means there is no single point of failure for name resolution. EVERY node performs name resolution, either autonomously or as part of an intelligent, mesh network. This completely avoids the fragility and inefficiency of name lookup systems such as the World Wide Web (www) Domain Name System (DNS).

13.12. Sovereign identity, sovereign data, and fairness

To date, networks and hyper scale web applications have largely relied on centralised, federated security models and identity management by the web app or by third parties. Data ownership is often ceded from the author/owner to the web application owner, who profits enormously from it. Networks and applications are owned by a small number of players and do not prioritise fairness.

The SPAN-AI UCDN is based in sovereign identity and sovereign data ownership. The SPAN-AI UCDN provides a simple and elegant method, based on IPID, to resolve identity, using IPFS and IPLD. The IPID method conforms to the WC3 Decentralised Identifier (DID) specification which states in part: “Decentralized identifiers (DIDs) are a new type of identifier that enables verifiable, decentralized digital identity. A DID refers to any subject (e.g., a person, organization, thing, data model, abstract entity, etc.) as determined by the controller of the DID. In contrast to typical, federated identifiers, DIDs have been designed so that they may be decoupled from centralized registries, identity providers, and certificate authorities. Specifically, while other parties might be used to help enable the discovery of information related to a DID [e.g. a blockchain], the design enables the controller of a DID to prove control over it without requiring permission from any other party. DIDs are URIs that associate a DID subject with a DID document, allowing trustable interactions associated with that subject.”

The SPAN-AI UCDN, like the content based networking and distributed systems technologies, platforms, and philosophies it draws on, is designed to promote fairness as far as possible.

13.13. Signing by publisher/owner

Every SPAN content packet is signed by the publisher/owner. This provides network-based authentication of data and source, vastly improving network, content owner, and user security.

13.14. Hashed node and content IDs

A fundamental part of all SPAN content tags and node IDs from day one has been a hash. This provides an immutable, traceable, and routable identity for every component of a SPAN-AI network. When combined with content name-based techniques and formed into interoperable linked data structures such as Merkle/Verkle trees and Merkle forests, based on a single name space such as IPLD, it provides the foundation of the UCDN and the possibility of spanning all next generation networks.

14. Network security

14.1. Real time status reporting

AmI-Rendezvous is the UCDN universal discovery system that provides smart discovery, configuration, and self-healing services to bootstrap nodes and resources; discover peers and services; and maintain DHT, naming and pub/sub in the SPAN-AI UCDN. It combines peer-level self-healing intelligence and edge-level smart discovery to determine where to host distributed AmI-Rendezvous services. AmI-Rendezvous is ideally co-hosted with edge-level naming and intelligence services. Peers register on initialisation, after performing host name and DHT discovery. Data partitioning and service placement may also be guided by naming.

14.2. AI threat detection

SPAN-AI agents provide enormous amounts of information about the network that is ideally suited to machine learning based threat detection. Content based networking further increases levels of awareness and security.

14.3. Network pruning

AmI-Rendezvous smart discovery and self-healing improves scalability and churn-resilience with little impact on the SPAN routing scheme, other than tuning overlay degree, fanout, and probability weights. Embedded plugins for self-healing and smart discovery strengthen peer discovery, activation, and lifecycle management of the overlay. The pubsub protocol is embedded with pluggable metric aggregators, actuators, and triage for smart discovery. A periodic heartbeat disseminates mesh-health metrics to AmI-Rendezvous. Aggregated metrics are used by Machine Learning models to classify, rank, and recommend peer connections and disconnections.

15. SYSTEM COMPONENTS

15.1. INGEST

All next generation networks process ***all*** content at ingest to enable optimum classification, manipulation, processing, storage, and distribution of that content. This model was originally inspired by early distributed storage networks such as BitTorrent and GitHub, the CSIRO's work on peer assist, by OTT video packaging using Adaptive Bit Rate (ABR) techniques, and by distributed storage networks such as the IPFS based Filecoin. However, in next gen networks these have been ***taken to a whole new level***.

The Filecoin distributed storage network, for example, processes files via its UnixFS distributed file system to segment them, attach hash tags (content IDs or CIDs) and store them in Merkle tree structures. Filecoin is a highly efficient and successful distributed storage system and token economy. It is optimised for storing valuable data safely. While Filecoin is excellent for secure distributed storage, it needs a few improvements for real time content distribution from that storage to consumers.

The SPAN-AI UCDN ingests and packages data to optimise it for ***both*** publishing to the network (distributed storage) and distribution from that storage to consumers e.g. video streaming (see UCDN OSI model diagram). At its simplest (our original foundation patent description), content is securely ingested, "sliced" (packaged), encrypted, tagged with a **hashname** (see Unified Naming and Addressing), and cached in the network by SPAN-AI according to requested SLAs, network state, current and forecast demand. This caching is continually optimised in real time by SPAN-AI, preparatory to distribution on request by a user or users.

15.2. SPAN-AI UCDN COMMANDS

Since the network IS the cloud™, (which is the computer) it needs an instruction (command) set. The two most fundamental are PUBLISH (store content in the network) and SUBSCRIBE (retrieve content from the network, either in real time or on-demand). (See the UCDN OSI model). The two most fundamental instructions are:

15.3. EXAMPLE COMMANDS

1. *PUBLISH <publisher> <content name; type; version> <master file; format; security> <QoS: encode; bitrate; reach;> <persistence> <permissions> <signature>*
2. *SUBSCRIBE <publisher if known> <content name; type; version> <QoS> <subscribe location> <signature/key>*
3. *etc etc etc*

15.4. SPAN-AI OPERATIONS

Modelling, Training, and Devops

Mathematical network models were developed as a flexible way of representing objects and their relationships that are not restricted to being a hierarchy. Network models are the foundation of SPAN-AI and the UCDN; both to verify optimum design and to optimise the operating network. Our models have been developed over decades of research and development of next gen networks in PhD work, real world modelling and optimisation, state of the art development at Nokia Bell Laboratories and now in the field. They are based on novel application of multi-resourced network graphs and directed acyclic service graphs to model and optimise both virtual service provisioning and content caching in software defined networks.

The models will be used initially to verify and then seed optimal virtual network designs into SPAN-AI. The models will be used for initial training of the machine learning (ML) algorithms and form the initial basis of the ML models. As operational data is collected, this will be used to inform, modify, and optimise the base models. SPAN agents will be trained by the models before distribution. Their correct functioning in the network will be verified by the models. This will form the basis of the SPAN-AI devops system.

Unified Naming And Addressing

Addressing scheme

The UCDN system uses a Unified Naming and Discovery (UND) System that i) maps mutable human readable names (e.g., domain names, content names) to immutable self-certifying content identifiers (CIDs), and ii) enables routing CIDs through both name-resolution and name-based routing subsystems, by iii) combining a name and a CID in such a way as to optimise routing and/or storage. We call this a "**hashname**".

Discovery

SPAN-AI employs an AI-driven, universal, discovery system, whose core component, Ambient Intelligence Rendezvous (AmI-Rendezvous), provides smart discovery, configuration, and self-organization services. AmI-Rendezvous services bootstrap nodes and resources, discover peers and services, maintain DHT and pub/sub overlays. AmI-rendezvous combines peer-level self-healing intelligence and edge-level smart discovery. SPAN-AI intelligence determines where to host distributed AmI-Rendezvous services.

Publish/subscribe

SPAN-AI uses an AI-driven Pub/Sub system for asynchronous multi-party communication services that support control plane dissemination: directory updates (names, discovery, configuration) and intelligence updates (optimization/control operations); as well as data plane dissemination: collaborative applications (messaging, video-conferencing, social networks, blogs, forums).

15.5. GLOBAL OPTIMISING AI

SPAN-AI orchestrates the adaptive operation of the routing and pub/sub systems via a family of pluggable, containerised, hierarchical (local/edge/global) autonomous and self-optimising AI agents that provide monitoring, prediction, optimization, and control services with varying degrees of awareness and optimization capabilities at peer, edge, and core network levels.

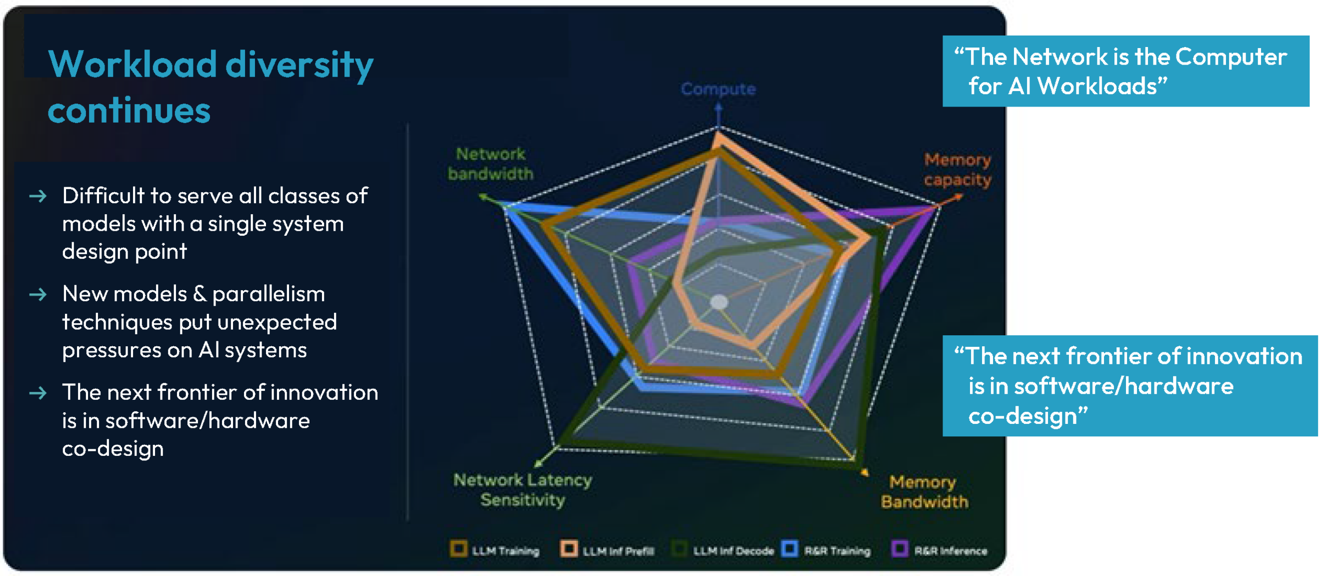
SPAN-AI incorporates a global, optimising intelligence with a simulation, training, and development pipeline that enables cloud-level replication of runtime environments, simulation, testing, and training of AI models, that can then be plugged into peer/edge/core network nodes for real-time optimization and control.

The long-term goal of SPAN-AI is to provide a marketplace for pluggable AI agents to enable open, flexible innovation in the optimization and control of universal networks.

15.6. DISTRIBUTED AI

Training of AI models is currently highly centralised, with widely varying workload characteristics. Massive numbers of centralised accelerators (GPUs and DPUs) are required. Like video, these workloads must eventually be distributed if AI is to scale to the levels required. Local training will be essential for real time AI. SPAN-AI is ideally suited for this.

SPAN has massive application to enable DISTRIBUTED AI…



https://www.delloro.com/ocp-2023-key-takeaways-the-network-is-the-computer-for-ai-workloads/
Source: Meta at OCP 2023

15.7. SCHEDULER

The network is the cloud™ which is the computer. Scheduling is central to the efficient operation of a UCDN. This becomes even more critical as industrial automation, enhanced experience, and AI demand ever more significant processing capabilities at the edge in real time. We are working with leaders in the field of scheduling and optimisation to build a fully distributed cloud computing UCDN that works for all next gen applications.

15.8. SPAN-AI AGENTS

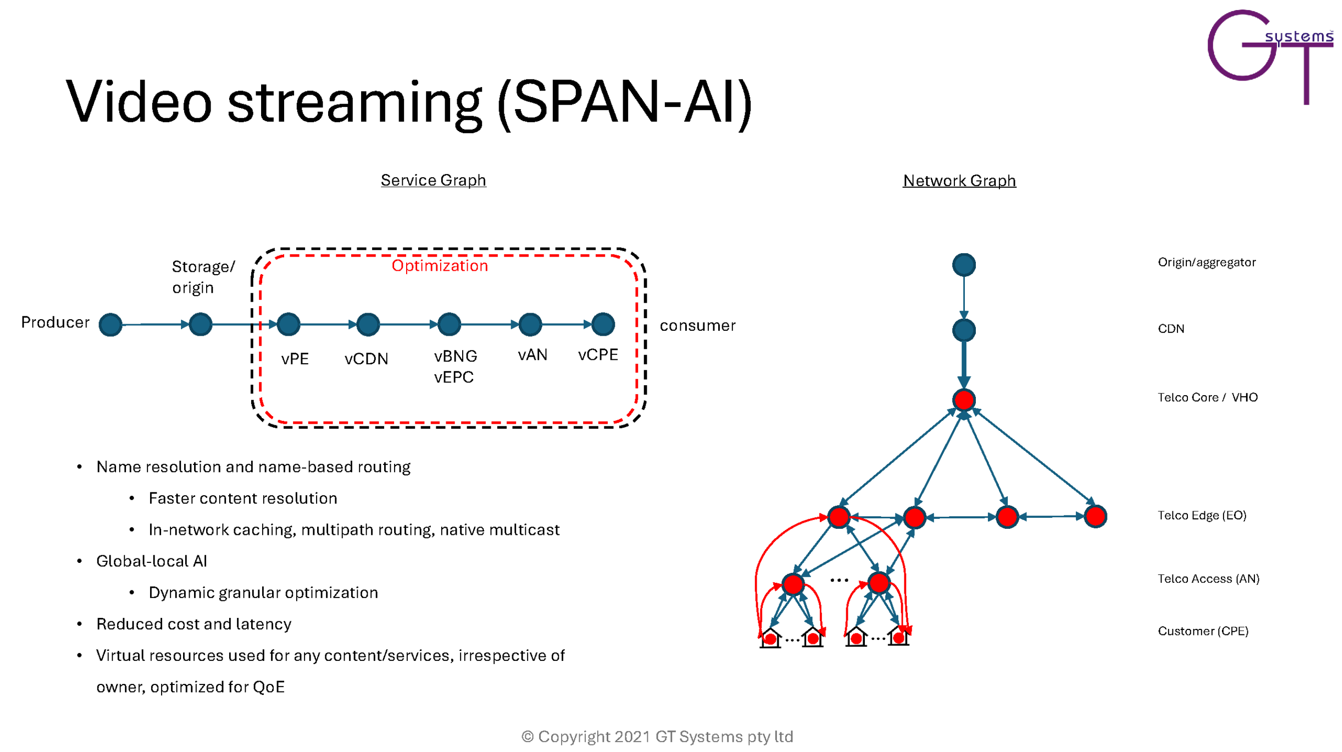
SPAN-AI agents provide local intelligence at edge and peer level to provide local monitoring and fast reactions for survivable operation, self-organization, and optimisation. They employ fast and simple rules on nodes with varying capabilities to enable ambient swarm intelligence and emergent behaviour. SPAN edge agents also provide higher level services supporting discovery, bootstrap, configuration, resource allocation, role assignment, storage decisions, routing hints, pub/sub memberships, naming, etc.

15.9. DISTRIBUTED STORAGE: SPAN-DFS

The core of SPAN has always been its distributed file system. This system combines the benefits of name-resolution-based storage and routing for scalable, available, accessible, distributed storage; with the advantages of name-based storage and routing for fast, reliable content delivery. It draws on the work of established distributed storage systems such as IPFS, with improvements in scaling and speed and the addition of parameters critical to QoS such as network distance and location. It is based on a Unified Naming and Discovery (UND) system that combines the benefits of name resolution based routing and storage with name based routing and distribution and scales to the levels required for trillions of users and devices and "zillions" of pieces of content.

15.10. DISTRIBUTED ORIGIN STORE

Video origin stores are currently only "distributed" as far as CDN or content publishers have been able to penetrate telco ISP networks. This has been against massive resistance from telcos and ISPs to hosting and carrying this infrastructure and traffic for free. This is usually only as far as regional data centres or telco Internet Exchange Points (IXPs). That is too far from consumers and there isn't room for everybody. GT Systems is developing the world's first fully distributed origin store. Coupled with SPAN-AIs distributed file system and MRL optimisation of caching, it will enable unprecedented levels of performance in on-demand and live video services hosted by telcos and ISPs. These new video services will form the core of the UCDN.



15.11. SPAN-AI-HARD ROUTING

SPAN-AI-HARD is our intelligent, hybrid, adaptive routing design that combines the benefits of name-resolution-based routing and storage (i.e. routing by resolving names and addresses by lookup e.g. IP and DNS or CID (hashes) and IPFS); with name based routing and distribution i.e. intelligent local and global routing by content names and interest without costly lookups. The unified naming and addressing system is what glues this all together.

SPAN agents will apply reinforcement learning to content-based routing, optimising routing and storage decisions based on actual network traffic and state. The global AI will also make inference decisions to forecast load. This forecast load data will influence global and local caching and routing decisions.

15.12. NETWORK OPERATING SYSTEM: UCDNOS

SPAN and UCDN assume a universal, homogenous network of intelligent networks implementing the SPAN-AI protocol and inter-working with legacy TCP/IP networks. See UCDN figures (1) and (1c). These sub-networks operate by means of intelligent SPAN “agents” implementing a dual, interoperable, SPAN-AI and TCP/IP stack on every device in the network: routers; switches; servers; gateways; CPE (PCs, phones); etc. Every device will have its own OS. There is a proliferation of operating systems at the network device level: versions of Linux; Nokia SR Linux; Ubuntu; Nvidia Cumulus; Arista EOS; specialised NOS; Windows (server and client); Apple iOS, Android, Sonic, etc.

SPAN-AI needs to operate and inter-operate across all these platforms, with APIs at the network and AI level to form the UCDN (fig 1). Ideally, there will be one universal, intelligent NOS (UCDNOS) that interoperates with the multitude of proprietary and open NOS. Initially this may require many versions. In the medium term it will migrate to containerised microservices and may include Wasm-like methods to enable running in legacy browsers. SPAN agents operating on every device under every NOS as containerised micro-services will report device capability and status to multiple inter-operating global SPAN-AIs and data lakes:

UCDNOS telemetry data

1) Location
2) NOS
3) File system(s)
4) Hardware platform
5) Processor(s)
6) Storage
7) Caching
8) Routing capability
9) Route info
10) Demand
11) Load
12) Operational status
13) Others TBA

15.13. SPAN-AI APIs AND INTERCONNECT

The UCDN is designed to be an intelligently inter-operating network of intelligent SPAN-AI networks (see fig 1). This inter-operation is achieved via APIs at two levels: the SPAN-AI-HARD routing level and the AI level. The AI API in turn breaks down into two levels: local agent and global optimising AI. Routing information must be exchanged to permit intelligent, stateful routing between networks. And the optimising AIs must talk to each other and share operational information. AI APIs are a relatively new concept and will be developed as part of the AI devops system.

15.14. UCDN MARKETS AND ECONOMY

The UCDN will initially be built and run by interested private parties and stakeholders. In order to hyper scale, it must become a fully open, distributed system, similar to existing distributed storage systems and markets such as Filecoin. This will require the development of trustless proofs such as Proof of Routing (PoR), in addition to versions of existing storage proofs such as Proof of Space-Time. These proofs will be light weight and use much reduced amounts of green power. Because of its efficiency and fully distributed architecture, SPAN-AI significantly reduces power consumption and enables co-location with renewable energy generation. This will require new proofs such as Proof of Green Energy. It will also require new token economies, management systems and governance. This will be the subject of a separate white paper.

16. POSSIBLE NETWORK SERVICE USE CASES - APPLICATIONS

16.1. AI

AI workloads are currently scaling exponentially. Training of AI models is currently highly centralised, with widely varying workload characteristics. Massive numbers of centralised accelerators (GPUs and DPUs) are required. Like video, these workloads must eventually be distributed if AI is to scale to the levels required. Local training will be essential for real time AI. Because it is AI aware and enables distributed processing of AI agents, SPAN-AI is ideally suited for this.

16.2. Web3

Every web3 distributed application is an Information Centric Network. Blockchains are chains of transactions which are small pieces of content: “A owes B $X”. Or bigger pieces e.g. Filecoin. Every web3 application we have seen may have its own internal protocol but communicates over IP. This is like a bank talking to another bank via a tin can and a piece of string. Layer 2 and layer zero applications are uniting disparate blockchains, but they need a universal protocol that speaks blockchain and IP and can talk to future next generation applications. That protocol is SPAN-AI.

Ironically, there is no web in SPAN-AI enabled web3. Everything is an application that talks to the ***network*** via an API. Maintenance of state and the storage of data where it is needed by distributed applications are removed from the application and taken care of ***by the network***. Data migrates from where it was generated to where it is needed; automatically; ***once***.

16.3. Web3 distributed games

This is a very interesting application. It requires a distributed rendering web3 game engine and a distributed network capable of understanding and distributing complex abstract scene representations. That is SPAN-AI. It has many applications beyond gaming. It’s also much harder than people thought. We have some ideas on how to solve this.

16.4. Real time studio quality broadcast video

For example, super high-quality video of the next man and woman on the moon broadcast from the moon to what will presumably (hopefully?) be an audience of billions in real time. Which of course extends out to live sports broadcast; the Super Bowl, the World Cup, and Tour de France being great examples with big global audiences. CDNs struggle to carry even 10% of these peak loads without significant degradation of QoS[5]. If we are going to transition from FTA and cable broadcast to digital[6], then radical retooling is required. SPAN-AI is that toolbox.

(HISTORICAL NOTE: the video of the first moon landing was broadcast in 1969 to a global audience of 650 million via the CSIRO Parkes radio telescope in Australia)

16.5. On demand video in broadcast quality any resolution no buffering

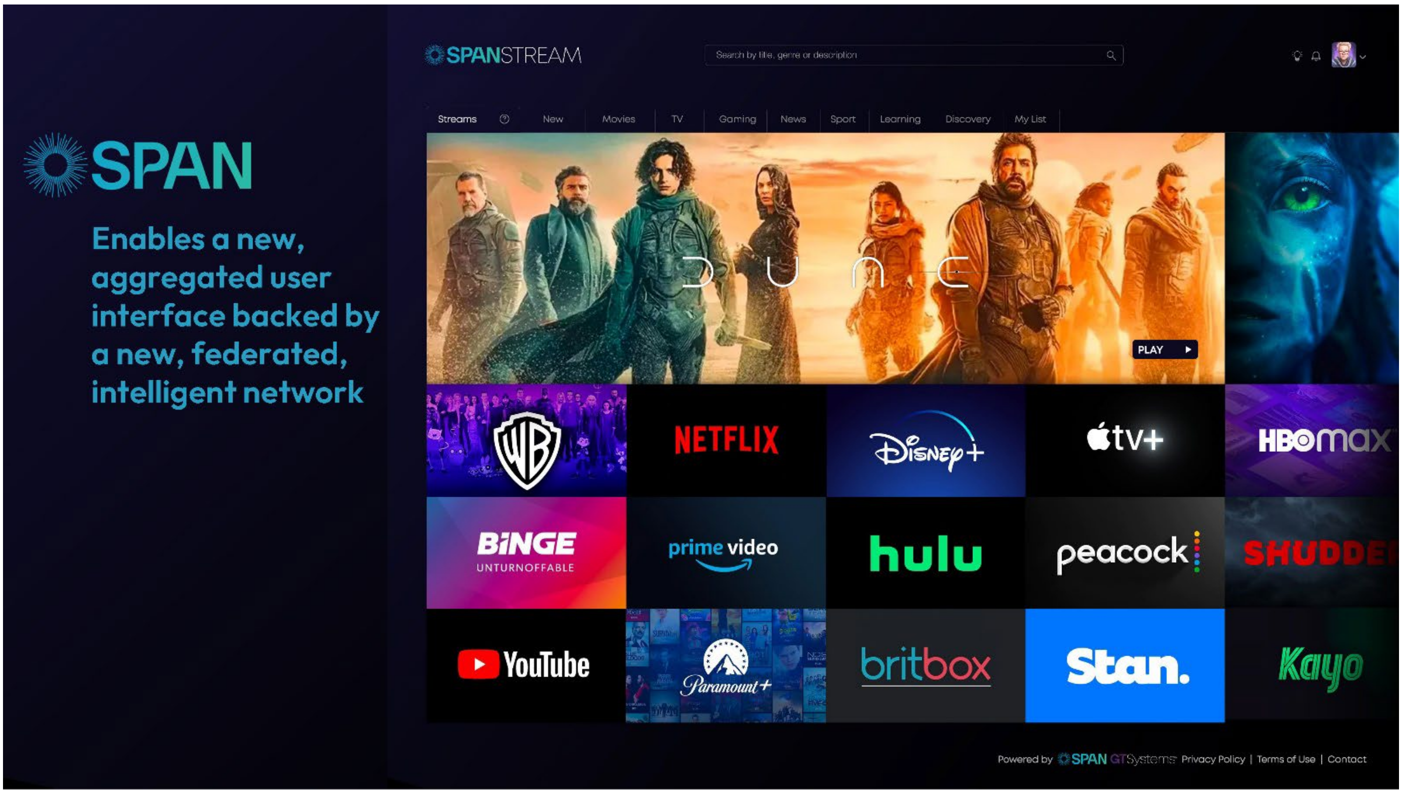
The example of super premium quality 8K HDR WCG 100fps non-interlaced movies streamed at 100mbps with no SWoD, dropped frames, or ABR fuzziness. That’s glossy, print magazine quality movies on your TV. If you haven’t already, go to your local mall and take a look at a new, high quality, 8K TV demo e.g. Samsung 8K QLED. They are stunning. No CDN can currently do this.

16.6. A new Universal Content Distribution Network fronted by a new User Interface

Hollywood is screaming for a Content Distribution Network that does what everyone has promised but no-one has yet delivered: studio quality; delivered anywhere; at a reasonable cost. David Zaslav (Warner Bros CEO) has called for

[5] https://www.businessinsider.com/paramount-super-bowl-streaming-error-issues-why-explained-netflix-2024-2#:~:text=Some%20Paramount%20Plus%20users%20complained,them%20do%20more%20of%20it.
[6] Telco engineers describe this as “threading an elephant through the eye of a needle”

it. Bob Iger desperately needs it for Disney+. So does Paramount+. SPAN-AI and the Universal Content Distribution Network enable it. It looks like this:



16.7. Industrial automation (aka IoT)

The problem with IoT comes when you go from low volume traditional “control systems” to next gen Industrial Automation which comprises real time modelling and video monitoring of a production line or factory, combined with real time AI analysis and action. This is the hardest next gen problem of all. This is almost the same problem as real time distributed gaming combined with autonomous car control. Interestingly, web3 distributed gaming promises to solve many of these problems, and we are working on it with web3 gaming partners.

16.8. The Metaverse

There is a lot of fuss about the metaverse at the moment. The good news is that the pieces we need to build it have finally arrived. So, the fuss is probably justified this time around. *Everyone* is asking: “What is the Metaverse? What are the pieces we need to build it? And how do we do that?”

It doesn’t take long to work out that, if we are going to join up a diverse set of existing sub-universes, such as Fortnight, Roblox, Minecraft, World of Warcraft, even Facebook (oops Meta), into the Metaverse, then they will all need to “talk” to each other. Talk how? They need to be able to exchange assets, characters (avatars), reputation, skills, experience, tools, weapons, clothing, potions, loot; and in the case of business worlds such as architecture, they need to be able to exchange drawings and models that interwork and follow *exactly* the same set of rules and dimensions. In real time. Most importantly, we need to be able to buy and pay for stuff in all the different stores in all the different sub-universes as we move between them. Blockchain and crypto currencies are an obvious candidate for that.

The solution to this would be a ***network*** that can store the Metaverse assets and understand and communicate complex scene representations. That network is SPAN-AI™. We speak Metaverse and we optimise content caching and routing ***in the network***. The network ***is*** the cloud™. We also speak native blockchain and Merkle Tree; in fact, any content-based storage and distribution "language".

SPAN-AI™ is fully distributed, elastic, content native, and intelligent. ***EVERY*** node in a SPAN-AI™ Universal Content Distribution Network (UCDN) is autonomous and self-optimising. Every device on the network, whether it is a server, router, switch, gateway, PC, phone, VR/AR headset, etc., is a compute, storage, and routing element of the network. That makes the SPAN-AI™ network a giant, distributed Metaverse computer, far more powerful than any centralised super-computer and far more efficient because it is as close to consumers as possible. SPAN-AI™ fundamentally understands commands such as "publish" and "subscribe", and also understands Metaverse commands to create and share 3D world content in real time. In short, SPAN-AI is the "glue" of the Metaverse that brings all the sub-universes together by speaking native Metaverse and caching Metaverse content and assets in the network.

17. SPAN UCDN NETWORK SERVICES

Because the SPAN-AI network service and control layer is abstracted from the physical layer, it enables the delivery of virtual network services wherever there is a SPAN agent. Because SPAN-AI is also distributed, intelligent, self-optimising and ***fast***, it enables the delivery of new network services that cannot be provisioned on legacy networks and still meet the demanding SLAs of these new services. Many of these services are simply not possible to provide on legacy networks.

Examples of this are:

17.1. Publishing services.

Content owners can now publish (store and distribute) their premium quality content via (to) a SPAN-AI network direct to their customers, bypassing aggregators, resellers, cloud service providers, etc. and becoming direct virtual service providers to their customers. These virtual publishing services are delivered via SPAN-AI, a better-than-telco grade, next gen CDN (ngCDN) with no degradation of picture quality from the original premium encode and no buffering. Publishers are charged by volume and QoS but there is no reseller/aggregator margin. The network is the cloud™. For example, a US publisher could publish a highly secure, glossy magazine print quality movie direct to customers in Asia, resulting in savings of 30-50% of RRP in reseller/aggregator margins.

Live sports broadcasters can broadcast direct to their global audience with milliseconds delay and premium picture quality. Audience size is unlimited. The more viewers, the better SPAN-AI works. Broadcasters are charged by volume and QoS for a better-than-telco grade service with no additional CDN charges.

Game publishers can increase their audience size, reach and performance with true edge gaming services. VR/AR performance is improved. Holograms become real. The metaverse becomes possible.

User generated content (UGC) originators (gamers, social video, corporates, etc.) can broadcast to a global audience in premium quality. There is no limit to stream quality or audience size. UGC originators are charged by volume and QoS so their costs scale with audience size and revenue while delivering premium quality services.

17.2. Virtual network services

Much like Mobile Virtual Network Operators (MVNOs) today, ***anyone*** can become a global virtual service provider (VSP) and deliver telco and cloud services via the SPAN-AI UCDN network to ***anywhere*** there is a SPAN agent. VSPs are charged wholesale rates by volume and QoS, enabling resale of any service that can be provisioned on SPAN-AI. Because SPAN-AI services are fully automated, abstracted and model driven, almost any service you can imagine can be provisioned almost immediately. Services can be defined and maintained via YANG models, including QoS specification. SPAN-AI is fully model driven, including pre-provisioning proof of concept (PoC) models.

17.3. Universal Content Distribution Network (UCDN) services

SPAN-AI networks are designed to join up via APIs, both at the network level and the optimising AI level. As the number of SPAN-AI networks grows, they form a fully homogeneous, global, intelligent, optimised Universal Content Distribution Network (UCDN). A global, optimised, virtual network allows for the provision of previously unimagined services. Virtual classrooms in Africa. Secure cross border and geographic obstacle hopping communications. Global or local gamer "LAN parties" and esports. Holographic concerts. New services will only be limited by imagination. The satellite and free space laser networks described here further enhance SPAN-AI's global service capabilities.

17.4. Space communication services

The NASA Johnson Space Centre Vision for Space Exploration is to "increase robotic missions in the solar system and return humans to the moon before sending them to Mars….. [including] robotic exploration missions across the solar system to support human exploration and further scientific discoveries….. to search for evidence of life, understand the history of the solar system **and search for resources.**"

Man's renewed journey into space over coming decades will require whole new communications infrastructure. Laser satellites and ground systems will play a critical role and will also bring new capabilities to terrestrial telecommunications. SPAN-AI is ideally suited to both space and terrestrial applications of free space lasers and the next generation services they enable. Our R&D team are innovators in free space optical communications, as well as fields such as sharded networks and protocol buffers that will be the building blocks of the new space age. Just as SPAN-AI enables the provision of virtual network services on earth, so it will in space.

17.5. Highly secure publishing and distribution services

SPAN-AI was designed to be secure "from the ground up". All the above services incorporate the SPAN-AI security model:

1. Security rooted in trusted hardware platforms and operating environments
2. Encryption of data at rest and in transmission to the most stringent commercial standards. We have full approval for high value content by major Hollywood studios.
3. Signing of all content by publishers

4. Built on platforms that enable quantum key distribution and encryption with high resistance to interception
5. Distributed systems that are much more resilient to denial-of-service (DoS) attack
6. Content based networking that reduces threat surface

18. FUTURE SERVICES

18.1. Sharded networks and databases

Protocol Labs and Laser Light Communications have championed the requirement for networks to work on the moon and Mars as well as they do on Earth. If we're going to the moon and Mars, we're not going to be in touch all the time. Both on earth and in space, the cloud/network needs to be capable of having large chunks "broken off" for extended periods of time, operate autonomously, and re-synch when they join back up. This also has application in scaling distributed networks.

18.2. Protocol buffers

A mechanism, invented by Google, to "serialise" (put down a comms line) structured data e.g. a database. Everyone is starting to use them. Same as we've done for movies. If we can invent a ***network*** standard for protocol buffers, this achieves what we are looking for: a way to ***intelligently*** move really big files around the UCDN, with caching and whatever else we need to do. Nvidia have done some work that sounds like it's heading in this direction.

18.3. The Metaverse

The ultimate manifestation of ***all of this*** is what has been described recently as "The Metaverse". Far from being some hokey evolution of Second Life or even Fortnight, it is, in fact, what the SPAN-AI UCDN enables: ***experience and control of virtual worlds, at the real edge, in augmented reality, for unlimited audiences, in real time.*** Those virtual worlds need to be open and inter-operable, with portability and tradability of avatars, skins, gear, swag, loot, weapons, tools, skills, virtual goods, services, models, and economies across worlds. SPAN-AI is the foundation for this to happen. A good set of rules for the Metaverse can be found here https://medium.com/meta-verses/the-seven-rules-of-the-metaverse-7d4e06fa864c We agree with the author on many points, especially that the Metaverse is a network and that it is the [next generation] Internet (rules 7 & 8)! Although he says there is no need to reinvent the Internet, he is closer to the truth when he says "more likely…. Is a series of evolutionary advancements here and there, ***punctuated by an occasional revolutionary leap via the introduction of an entirely new capability***". Replacement of the narrow waist IP protocol with SPAN seems to qualify! As we have said, evolution AND revolution.

19. DRAWINGS

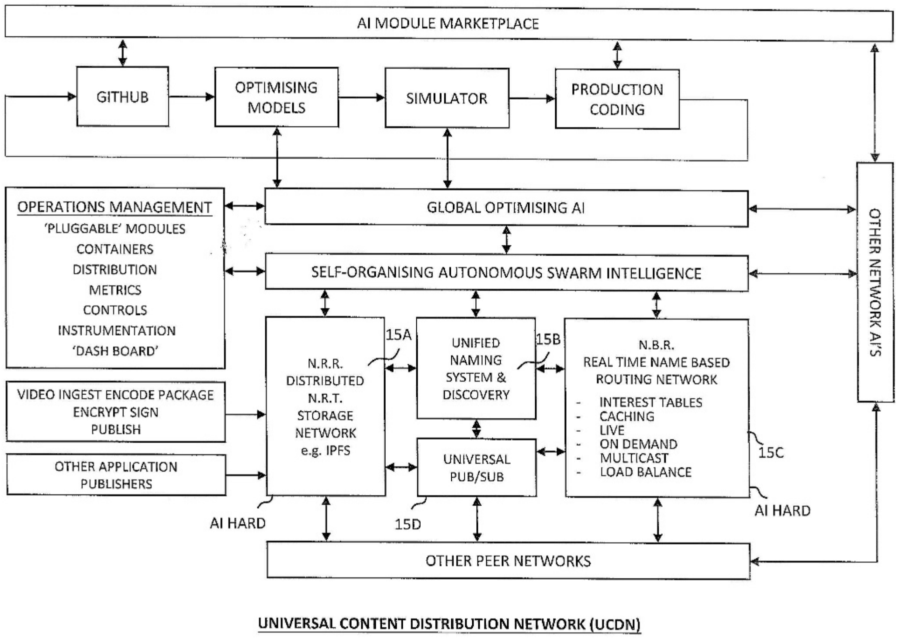


Fig. 1

1/28

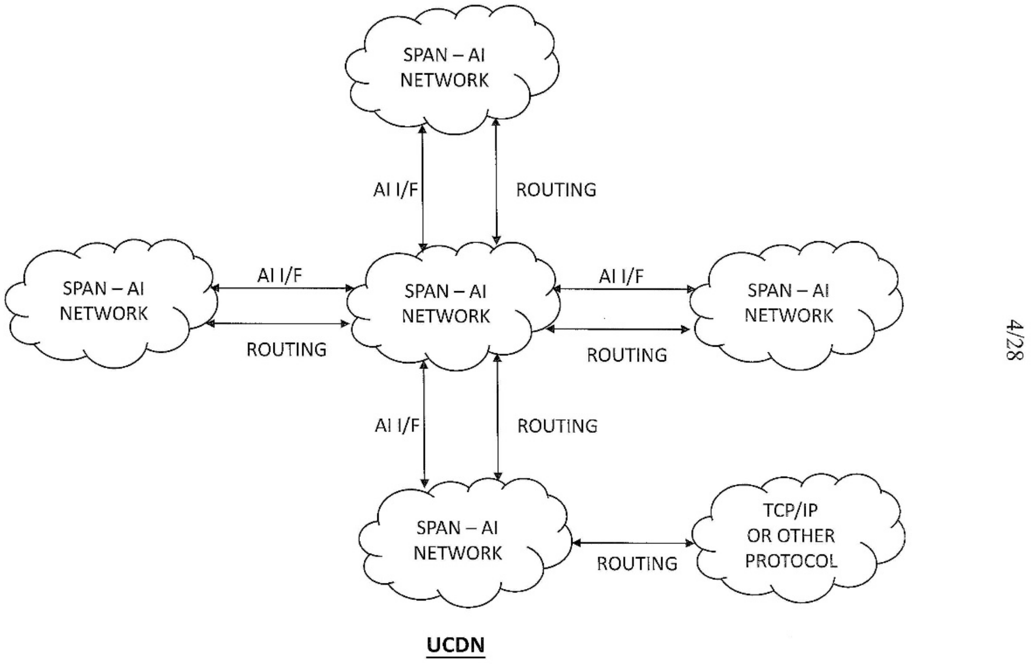
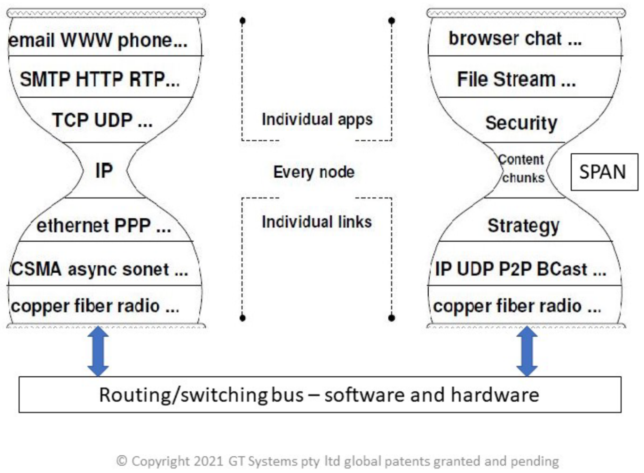


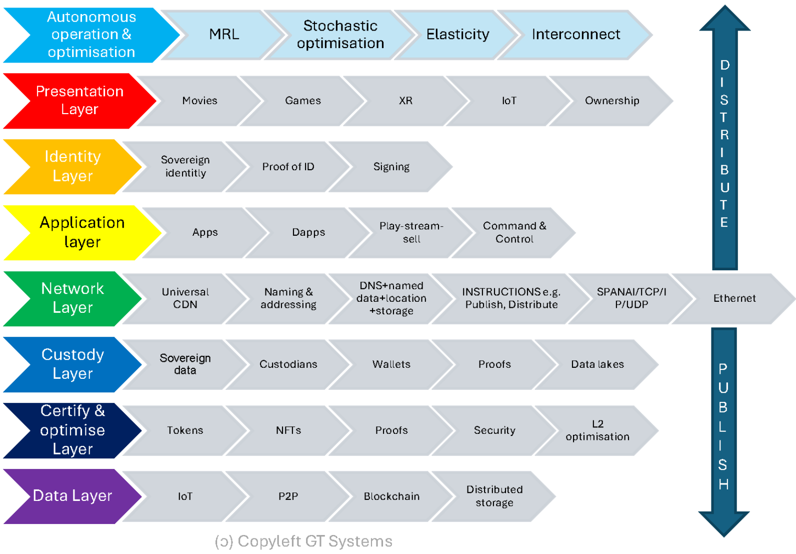
Fig. 1C

Dual stacks IP and SPAN on every device



SPAN-AI unites the current and next gen Internet in a new model – (W.I.P.)

- AI optimisation
- TCP/IP, UDP
- IPFS, IPLD
- Filecoin, Ethereum and other blockchains
- P2P
- IoT management
- Content based storage and distribution



SPAN-AI UCDN NEW OSI MODEL

20. GLOSSARY

AmI-Rendezvous Ambient Intelligence Rendezvous

CSIRO Australia's Commonwealth Scientific Industrial Research Organisation

DC data centre

DRM digital rights management

FOC Fibre optic cable

FSO/FSL Free space laser optics

Hashname a "tag" that is a combination of a hash and a content name to enable name resolution routing and name based routing of a piece of content (this may be the first use of the word?)

IPLD Inter-Planetary Linked Data https://ipld.io/

Metaverse the next generation of open, interoperable 3D worlds providing enhanced experiences

NOS network operating system

OTT Over the Top (video distribution via the Internet)

SPAN-AI Secure Peer Assist Network with Artificial Intelligence

UCDN Universal Content Distribution Network

VSP Virtual Service Provider