5 December 2020

SPAN-AI-HARD™*

Rhett Sampson†
GT Systems pty ltd
Sydney, NSW, Australia
rhett@gtsys.co

Jaime Llorca
GT Systems pty ltd/New York University
Brooklyn, NY, USA
jllorca@gmail.com

ABSTRACT

GT Systems has been researching, designing, developing, and building physical (CD/DVD Manufacture on Demand) and Internet online content distribution systems for two decades. In 2012, we set out to solve a simple problem for a specific retail customer: building an online store for movies, music, and games; and solving the Internet distribution problems of Adaptive Bit Rate and the "spinning wheel of death" (buffering) by utilizing unused peer resources[1,2]. That was the start of a journey that has led us to solve much more than that. We have conducted our own, privately funded and Australian government subsidized R&D. We have brought Hollywood along on the "P2P" journey and they have been engaged, interested, and supportive.

We combined our initial, patented, Secure Peer Assist[3] concepts of distributed storage and peer-assisted streamload™ content based routing at the video application level with "state of the art" network architectures. We've integrated local AI for decentralized, self-optimizing, in-network storage and processing swarms; with global, optimizing ML models for joint storage, communication and computation resource allocation. We've proven that's the most efficient model, and the only one that scales to next gen applications[4]. We further developed and patented a Universal Content Distribution Network (UCDN) system, founded on our AI-driven, Secure Peer Assist Networking technology, SPAN-AI™, and our intelligent Hybrid Adaptive Routing Design, AI-HARD™[5]. We're building a Proof of Concept utilizing Nokia's new, open, datacenter switch fabric platform.

CCS CONCEPTS

• Networks• Architectures• Algorithms• Protocols• Performance evaluation • Services

KEYWORDS

Intelligent hybrid adaptive networks, Universal Content Distribution Network, UCDN, Information Centric Networking, ICN, Name Domain networking, NDN, distributed storage networks, Interplanetary Files System, IPFS, peer to peer, P2P, blockchain, cloud native, edge networks, Network Function Virtualization, NFV, universal naming and directory system.

The problem being addressed

The limitations of the current Internet and underlying/overlaying telco and CDN networks are well known[6]. The recent black swan events of COVID-19, that caused sudden and catastrophic changes in network load and distribution[7,8], have exposed just how fragile these heavily centralized and connection-oriented networks are. A consensus is emerging that now is the time to fundamentally re-evaluate Internet models, architectures, and protocols.

While there has been a plethora of global initiatives to address different aspects of these problems, none of them, on their own, will be capable of solving the whole problem. Our thesis is that we need to: put content at the

*SPAN, SPAN-AI, AI-HARD, SPAN-AI-HARD, "The network is the cloud", streamload, are trademarks of GT Systems. All other trademarks are the property of their respective owners.

†The Author's work has been supported by private funding, Australian state JobsforNSW and federal R&D grants

center of the architecture; exploit all available resources (core, edge, peer); enable elasticity (NFV/SDN); and embed ML/AI for self-optimized, adaptive use of next gen networking technologies.

What makes this problem interesting RIGHT NOW

The central thesis of the ICN movement has been proven beyond argument. All the telco, CDN, and Internet networks in the world couldn't deal with the "overnight" changes wrought by COVID-19. Global sovereign governments were forced, cap in hand, to ask the big global content distributors (Netflix, YouTube, etc) to reduce their streaming rates so that national networks could survive in the short term[9].

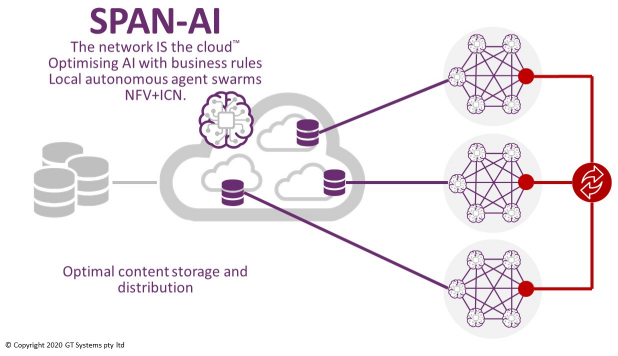
Telcos, ISPs and CDNs may argue "too fast/not enough time to reconfigure networks." Our response: it's about much more than that; and that is precisely what a modern network should be capable of: robust self-optimization, among other things. Add to that the fact that CDN and telco investment models are individually "running out of steam" and are having difficulty justifying the massive investment required to meet the coming scale of streaming, VR/AR, IoT, and 5G. The moment is ripe for a new approach.

The need to distribute content and applications closer to the edge, in proximity to end users, is already pushing global networks towards a uniform, flatter pattern: hyperscale (centralized and edge) datacenter LANs; switched or shared layer 2/2.5 local loops (NBNs); "bridged" in the wide area; with 5G for mobile. We will show in our modelling and the PoC that this pattern is particularly suited to ICN and, more particularly, to NDN, and that it can be made suitable for P2P distributed storage networks.

The content distribution business has gone through several major cycles: broadcast (disaggregated) => cable (aggregated) => OTT (disaggregated). Given the current proliferation of direct to market, disaggregated, OTT streaming services, the market is ripe for new aggregator(s).

If you take a global, content centric, view and extend all these scenarios to their logical conclusions, a very interesting potential new model emerges: the telco networks, ISPs, CDNs (public and private), and the Internet TOGETHER could be the content aggregator; providing storage, on-demand and live streaming as a ***unified*** SERVICE to TVs, SMPs, PCs, game consoles, etc. (themselves forming part of the network). The publishers and content originators (Hollywood, Netflix, indies, game publishers, players, etc.) publish ONCE to the NETWORK. The network UI is a simple window with brands: Disney, 'Netflix original', Fortnite, Gods Unchained, indies, etc. Consumers search and choose the brand that can supply the **content at the QoS they want**. The network delivers it; and settles accounts with all parties: publisher, telco, ISP, CDN, game streamer, esports player, etc. Just like a phone call. **Everyone** contributes. And everyone benefits.

The critical question is: how do you do that? It requires a fairly radical rethink and a new approach to network construction and information dissemination, as well as some adjustments to business models. But it's time for that. And that's what we're building.



Our approach

Sound like science fiction? Well, we're about to do modelling and simulation to test and demonstrate the benefits and then build a working Proof of Concept. It's time to go from a 50 year old paradigm of isolated, dumb, point to point conversations where computers, memory

and storage were the scarce resource (TCP/IP) to what we care about today: a global, CONTENT CENTRIC, ***network*** storage and distribution model in a world where your phone has more power and storage than every computer on the original Internet. Once you do that, all the problems go away (including the spinning wheel of death!) and the ideal scenario becomes not only possible, but implementable.

Our key contributions

We have taken an open minded, global, content based networking view and developed SPAN-AI, an intelligent, hybrid, adaptive, Universal Content Distribution Network design, combining the best of: ICN/NDN; P2P distributed storage networks[10, 11]; blockchain; cloud native edge architectures; network function virtualization; and hyper scale web architectures. SPAN-AI proposes a new, universal naming system that unifies: name resolution (content hash and url to IP) routing for persistent data storage and access; and name based routing for fast content delivery; optimized according to application (storage or distribution) ***and required QoS***. With SPAN-AI, content lives in, gets stored and processed by, and is delivered over ***the network***. The network ***is*** the cloud™.

We have "glued" the most promising of the "island" solutions together into a single, cohesive solution: SPAN-AI; and we are proving that it is: Robust; Efficient; Flexible; Self-optimizing; Secure; and Hyper-scale.

To paraphrase Van Jacobson: SPAN-AI works over TCP/IP, and TCP/IP works over SPAN-AI. As it should, and must. We have drawn on ICN/NDN methods and distributed storage methods, particularly IPFS, so that they interwork and sit alongside each other and are compatible with TCP/IP.

We are building a PoC utilizing Nokia's new, open, datacenter fabric platform and switches. We will initially model and derive benchmarks for performance improvement on metrics such as: resource efficiency; utilization; network load; congestion; application throughput; latency; and response time.

We may run SPAN-AI as a TCP/IP overlay initially. Pretty quickly we will build production, native, SPAN-AI router/switch protocol stacks (including IPFS stacks) and run them alongside/on mature and robust TCP/IP and layer 2 stacks on the Nokia hyper scale network platform and switches[14]. The Nokia NetOps development kit (NDK) runs SR Linux under a telco-grade, cloud native, Kubernetes micro-services devops system and incorporates a network simulator to run initial production tests.

We will expand out to 3rd party global cloud infrastructure to test node/agent scale up to hyper scale. We will test on a production, next gen network and datacenter infrastructure. The intent is to build a new, global, Universal Content Distribution Network (UCDN). We already have a highly innovative, early adopter, "customer zero" for it. We believe there will be many more.

CONCLUSION

We have been working with CSIRO/NICTA[1], Protocol Labs[12,13] and others and would love to work with the rest of the Internet, telco, ISP, CDN and cloud communities; including the ICN/NDN and blockchain communities. We're building the PoC on an open Nokia platform[14] where our protocol stacks and routing applications will sit alongside Nokia's rock solid TCP/IP and Internet stacks and benefit from telco grade infrastructure. We're privately funded and are raising venture capital and Australian government grants for the PoC. **We have "stood on the shoulders of giants" to get to here, and now we need community help to refine and build the PoC. There's plenty more challenges to solve!**

ACKNOWLEDGMENTS

We thank CSIRO/NICTA for their early pioneering work envisaging and implementing peer assist[1]. We acknowledge and thank Protocol Labs for their pioneering work on distributed storage systems[11] and crypto powered network markets[16]. We also thank Van Jacobson, the ICN movement and the NDN project for providing many of the missing pieces.

REFERENCES

[1] Borghol Y., Ardon S. (CSIRO/NICTA), Carlsson N., Mahanti A. (2010) Toward Efficient On-Demand Streaming with BitTorrent. In: Crovella M., Feeney L.M., Rubenstein D., Raghavan S.V. (eds) NETWORKING 2010. NETWORKING 2010. Lecture Notes in Computer Science, vol 6091. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-12963-6_5

[2] Stefan Lederer, Christopher Mueller, Christian Timmerer, and Hermann Hellwagner, 2014. Adaptive Multimedia Streaming in Information-Centric Networks. IEEE Network, (November/December 2014), 91-96, 0890-8044/14/$25.00

[3] Rhett Sampson, Australian patent application number 2015342719, filed 2015-11-04

[4] J. Llorca, K. Guan, G. Atkinson and D. C. Kilper, "Energy benefit of distributed in-network processing for personalized media service delivery," 2012 IEEE International Conference on Communications (ICC), Ottawa, ON, 2012, pp. 2901-2906, doi: 10.1109/ICC.2012.6364502.

[5] Rhett Sampson, [more authors to be added including Jaime Llorca] Australian provisional patent application number 2020901494, filed 2020-05-09

[6] Cisco, Visual Networking Index (VNI) 2017-2022 retrieved from https://www.cisco.com/c/dam/m/en_us/network-intelligence/service-provider/digital-transformation/knowledge-network-webinars/pdfs/1213-business-services-ckn.pdf

[7] Gartner, Gartner CFO Survey Reveals 74% Intend to Shift Some Employees to Remote Work Permanently, from https://www.gartner.com/en/newsroom/press-releases/2020-04-03-gartner-cfo-surey-reveals-74-percent-of-organizations-to-shift-some-employees-to-remote-work-permanently2

[8] Todd Spangler, Post-Pandemic, Americans Say They'll Keep Streaming More TV — and Go to Movies, Concerts Less, retrieved Apr 24, 2020 from https://variety.com/2020/digital/news/post-pandemic-streaming-tv-increase-1234589195/

[9] Josh Taylor, Australian government asks Netflix and Stan to reduce data to avoid broadband overload, Friday 20 March 2020 from https://www.theguardian.com/media/2020/mar/20/australian-government-asks-netflix-and-stan-to-reduce-data-to-avoid-broadband-overload

[10] Onur Ascigil, Sergi Reñé, Michał Król, George Pavlou, Lixia Zhang, Toru Hasegawa, Yuki Koizumi, Kentaro Kita, 2019, Towards Peer-to-Peer Content Retrieval Markets: Enhancing IPFS with ICN. ICN '19, 6th ACM Conference on Information-Centric Networking, September 24–26, 2019, Macao, China.. ACM ISBN 978-1-4503-6970-1/19/09

[11] IPFS a peer-to-peer hypermedia protocol designed to make the web faster, safer, and more open, from https://ipfs.io/ 08/10/2020.

[12] Protocol Labs RFP [particularly the problem statement] from https://github.com/protocol/research-RFPs/blob/master/RFPs/rfp-7-MLDHT.md

[13] Protocol Labs RFP [particularly the problem statement] from https://github.com/protocol/research-RFPs/blob/master/RFPs/rfp-8-pubsub.md

[14] Nokia, Nokia Data Centre Fabric, from https://www.nokia.com/networks/solutions/data-center-fabric/#overview

[15] GT Systems [early work on a CE device and service using Secure Peer Assist] www.blust.tv

[16] Filecoin: a decentralized storage network designed to store humanity's most important information, from https://filecoin.io/