Using Generative AI to drive person centric networking

1st Steve Jones
Data Driven Business
Capgemini
Phoenix, Arizona
steve.g.jones@capgemini.com

***Abstract*— The idea of person centric networking has been around for a while, but it has faced a number of key challenges in gaining widespread adoption. As the number of connected devices, and particularly personally connected, devices continues to increase exponentially the complexity and security challenges of traditional network approaches will require a shift from provider centric to user centric networking. The challenges to adoption of person centric networking have primarily been around ease-of-use, ease-of-configuration and ease of service portability. Generative AI is showing the potential to address these issues by providing mechanisms to dynamically engage and create configurations and solutions that previously required detailed engineering effort. The experience of users today is of a mountainous network of different networks, services and providers. This paper covers how Generative AI developments are already impacting certain areas and what is required for the shift from physical to person centric to be complete and present people with a "personal flat earth" from a networking perspective.**

Keywords—Generative AI, Telecoms, Networking, IPv6

I. The User Experience in a Personal Flat Earth

Lana is at home watching England play France in the Euros, all of her devices are already connected to her personal network, which is using a 6G fixed wireless solution. Her streaming service is authenticated against the network. A friend calls her and suggests she joins at watch party at their house, she gets in her connected car, and the game is automatically transferred to the radio. She reaches Louis' house and disaster has struck, his fibre connection is down. Lana asks her Smart Watch to show the game on the TV. Her personal connectivity provider sends a generative package to the TV which is able to provision a 6G eSIM and securely pull the TV into her network, authenticating the action via biometrics and her confirmation, meaning her streaming service is now available. After the game the eSIM and application are deprovisioned automatically as she leaves the house.

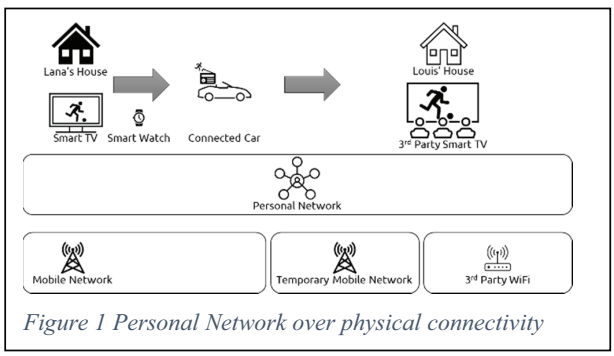

Figure 1 Personal Network over physical connectivity

In this personal flat earth, the first act of using a device is it authenticating and connecting to your network.

II. Today's Network is for Networks not People

Traditional approaches to networking have been based primarily based on location being a physical construct. Technologies such as Virtual Private Networks (VPN) [1] [2] enable people to remotely connect "into" a network, but the construct always required the routing to be linked to the physical location to provide security. This centers the network around its physical structures, and primarily binds between a, normally limited, number physical locations and into a single destination location.

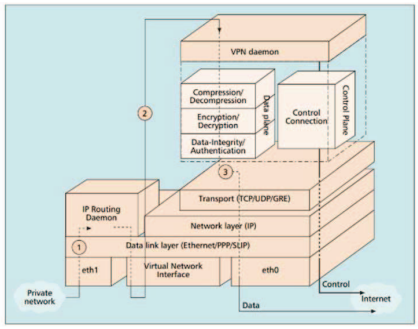

Figure 2 Software Architecture OSLV router [34]

Connectivity for people today is about a route onto the internet, where things are connected to a local network, and access to "personal" devices is mostly managed via intermediaries. Connecting devices together is normally done via an information sharing intermediary. The result of this is not just the trusting of third parties for sensitive information, but also the security of all of those connections into your personal networks. As a person moves from one network to another they are trusting that network to provide security and then leveraging point solutions, such as VPNs or "push-pull" architectures [3], to connect between disparate networks, none of these are trivial for the average user. This mess of devices and networks is set to increase to almost 30bn by 2028 [4]. Technologies such as 5G (and in future 6G) increasing the number of networks any individual will have to **physically** interact with.

The infrastructure complexity of services today is already far beyond the ability of normal individuals to understand, let alone manage, the question is how we can simplify the challenge **for the average user** and provide greater security but with simpler governance.

III. Defining a Flat Network

A flat network is one in which an individual's devices and services are authenticated against a single logical network and services are available to that individual through all appropriate

979-8-3503-5409-6/24/$31.00 ©2024 IEEE
DOI 10.1109/CAI59869.2024.00268

devices on that network. The user is unaware of the physical network configuration under their logical network and devices are able to transition across physical networks transparently.

A. The reality of compute and network

What this makes this approach now possible is the level of compute available both directly in devices and through the cloud services that back them. As the number of intelligent devices continues to exponentially increase, forecast to reach nearly 30bn by 2028 [4], so the power of those devices continues to track Moore's Law [5], mobile phones have moved in 20 years from simple airline information systems [6] to high powered AI devices in their own right. Even resource intensive actions such as computer gaming can now be portable between multiple devices using things such as Unity [7]m and when backed by cloud computing platforms the optimization of these services for any device is now common [8]. Meaning even relatively "dumb" devices have access to effectively infinite compute if required.

Consumer network connectivity has increased alongside compute [9] [10] with 4G and 5G providing high-speed wireless connectivity [11] and home broadband being widely available in many countries it is creating markets in which connection is ubiquitous. As we look to, Fiber delivering 10 Gb/s into buildings [12] and the adoption of 6G [13] it is reasonable to start designing consumer systems for an always connected and always computation available world.

IV. "I WANT TO RUN THIS ON THAT"

While technologies to address these challenges have been available for a while [14] the issue has been the ability of consumers to easily adopt such approaches. This can be summarized with the phrase "I want to run this on that", this requires us to address 4 core challenges:

1. Who – understanding the meaning of "I" in a given context
2. Here – the location of the "who" to understand device affinity
3. What – the service being referred to
4. Where – the target device

This is where Generative AI helps provide the non-technical interface onto deeply technical elements, this ease of use is why ChatGPT, a Generative AI Large Language Model, became the fastest ever adopted consumer technology [15]. They do not need to "understand" in a cognitive sense, but simply be able to mimic within proscribed domains [16], this makes it an ideal technology for the heterogenous challenge we are addressing.

A. Receiving the command and building context

The simplest part of the challenge is processing the request. In our Use Case Lana uses her Smart Watch to make the request. This then need to engage with Lana's active context to determine the meaning of the request. This is something that the current generation of LLMs are particularly good at [17] [18], meaning they can be used to track and manage a person's context, including location, and then have that leveraged by the LLM and other approaches to determine the right action for the request. With LLMs such as GPT4 enabling plugins for purpose specific tasks, such as mathematics [19], consumer facing LLMs can be seen as the engagement point for the overall system.

In our context of "Show the game on that TV" the LLM is able to parse the command, use its stored context to identify that "the game" is England v France, and recognize that the next challenge is to identify the TV. This is where short range technologies and inference approaches can help.

B. Short-range negotiation from here to where

The next challenge is to identify the device that is to be bonded to the network, there are a number of techniques that can help with this, from simple approaches leveraging short range approaches such as Bluetooth Low Energy [20] or more complex modern approaches using network inference and AI [21] [22]. For these challenges a generative AI will need to understand the capabilities available, the type of device being searched for, and the sort of security that will be required to complete the attachment. Once the Generative AI is aware of these capabilities it can create a custom solution to deploy to a local device, the Smart Watch in our use case, and use that to discover the TV. A standard bonding process could then be established.

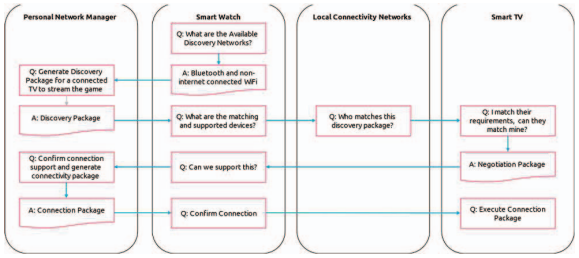

Figure 3 Simplified workflow on negotiated discovery through GenAI

The end result of this is that the device is now available, within security policies, on Lana's personal network. The underlying technologies for all of that are already available, the role of Generative AI is to simplify the connectivity to a level where it becomes trivial for consumers. This is an area where there has already been significant development with companies already announcing support for the Generative AI configuration and monitoring of Telco grade networks [23], the adaptation of technologies such as EnterpriseWeb to the consumer grade challenge would represent a simplification of a model designed to support the deployment of 5G networks. This final step in the process would be hidden from users, with the prompt required being created by other elements. Normal security approaches like the user confirming from their known device that the target device is correct "do you want bond with the TV showing the code 'Allez Anglais'?" would of course be required, but the actual negotiation and provision would be transparent.

C. A new service on a new destination

Once we have determined the target, and bonded it to the network, we then need to ensure that the required service, in this case a specific stream, can be displayed. To do this we need to understand the capabilities of the target device, and the requirements of the service to be deployed. This then enables a Generative AI to create an application specifically for this problem [24], and then confirmed within the target environment by creating a generative test [25]. With our

Smart TV example this might involve creating a full application to run on the platform, or if the TV does not have the compute required for that, a streaming shell that connects through a logical gateway deployed to Lana's network. This cannot include complex application creation, but can include the wrappers within which such applications can run, for instance if a device is able to support the Unity gaming engine then a generative wrapper can be created to create a platform specific deployment of the core code. For our streaming service the question will be whether an app needs to be installed, for instance via an app store, or just a direct streaming connection using one of the many available streaming platforms. The base functionality of this is already available, Generative AI will just simplify the transition between devices.

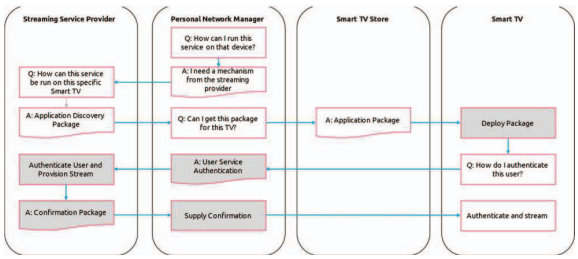

Figure 4 Generative negotiation to create a deployable package on the target device (grey = non generative)

Most of these facilities are already available, streaming to a connected TV already prompts to install applications for instance, the only shift would be the ability for a personal network to act as the authentication and security point for services, rather than requiring device specific authentication for an entire network. With the target device now on Lana's logical IPv6 personal network the services remain authenticated only to her. This enables greater degrees of privacy, and greater degrees of content control.

V. Generative AI gives scalable usage to semantic description languages

For all of this to succeed there will need to be a standardization of the language of both network and service description. While LLMs can handle some degree of language inference, future systems will be much more efficient and effective if the service, device and network capabilities are more akin to an API with additional contractual definitions than simply natural language. These service descriptions would include more than simply the API, but instead the functional and non-functional boundaries required for the service, as described previously in "Toward an acceptable definition of Service" [26]. This ability to extend service description to include requirements means that devices can advertise their capabilities, not just their existence, and therefore the right form of service can be developed.

Device Description Languages [27] [28] and Service Description Languages [29] [30] are not new, and efforts like the Semantic Web [31] have long tried to address the challenge of descriptive content associated with information assets. What Generative AI can do is give scale to these efforts, by enabling such approaches to be much more rapidly incorporated into new features and functions and leveraging the power of those technologies to provide the base semantic information for future Generative services.

Traditional approaches have required more binary association of source and destination models, which gave issues of unplanned obsolescence and competing standards resulting in a lack of market adoption. LLMs are already extremely successful in human language translation, so domain specific translations on limited ontologies should not prove to be a major barrier. With Generative AI leveraging semantic description languages, within proscribed domains, we can create specifically tailored technical LLM models that can continually adapt to new services, new approaches and still make them available on older devices. This could also prevent "walled gardens" being created by market dominant players.

VI. The impact of flat networks on consumer adoption

Security and Privacy [32] [33] are two of the primary challenges and concerns for consumers when looking at the explosion of connected devices. The shift to a flat-network approach would radically simplify this challenge, leveraging Generative AI to provide a simple and natural language interface to the consumer, with the technical details being managed and guaranteed by the provider of their personal flat network. This would alter the dynamic of the market from just brand centric ecosystems towards a greater degree of interoperability centered around connectivity providers and consumers. The ability to continually tie services back to specific consumers would also help some of the challenges faced by streamers and other services providers where traditional username/password approaches have resulted in service control problems. For consumers the advantage of a flat-network will be the ubiquity of service independent of the specific mechanism of connection and a guarantee of security being in-built into their personal network.

VII. The impact of flat networks on service and connectivity providers

The shift to a flat network model would create a new segment of connectivity provider, the Virtual Personal Network Provider (VPNP) a company that would not have to own any infrastructure but could provide ubiquitous connectivity to consumers above whatever network they are leveraging. This could disintermediate consumers from existing physical providers, as well as rendering VPN providers obsolete in the consumer space. When looking at the most likely providers of such facilitates then mobile platform providers such as Apple and Google are one option, but other hyperscalers and hi-tech networking providers, as well as Telcos, would have to strongly consider how to remain relevant and connected to consumers when such a transition occurs.

VIII. Conclusion and Further Study

Flat-Networking approaches have existed for a relatively long time, as have standardized approaches to service, device and semantic description. The challenge for these technologies for widespread adoption has always been the mechanism via which they have been made simple for consumers. This is where Generative AI offers immense opportunity to transform the connected consumer landscape.

We are already seeing efforts on Generative Networking, Product Development, Code, Testing and consumer avatars being deployed individually. The question and challenge will be how to construct these elements together into a unified data, service and AI ecosystem which enables a consumer, rather than product, centric approach to connectivity. Particular areas of focus will be on the security of virtual personal networks, and device description standardization to enable greater degrees of service portability.

IX. References

[1] V. S. L. C.-L. G. S. P. a. M. E. S. D. Wood, "Virtual Private Networks," in *1988 International Conference on Private Switching Systems and Networks*, London, 1988.

[2] R. Isaacs, "Lightweight, dynamic and programmable virtual private networks," in *2000 IEEE Third Conference on Open Architectures and Network Programming. Proceedings (Cat. No.00EX401)*, Tel Aviv, 2000.

[3] C. Whittington, E. Cady, D. Ratchen and Y. Dawhi, "Constructing universal secure and private data networks across consumer systems," in *2018 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, 2018.

[4] IOT Analytics, "State of IoT-Spring 2023," IoT Analytics, 5 2023. [Online]. Available: https://iot-analytics.com/number-connected-iot-devices/. [Accessed 6 2023].

[5] G. E. Moore, "Cramming more components onto integrated circuits," *Electronics,* 19 April 1965.

[6] S. Jones, "Next generation airline information using consumer devices to keep passengers informed and as an additional channel to market," in *IEEE International Conference on Mobile Data Management*, Berkley, CA, 2004.

[7] S. S. a. A. Kaur, "Game Development using Unity Game Engine," in *3rd International Conference on Computing, Analytics and Networks (ICAN)*, Rajpura, 2022.

[8] Y. Han, D. Guo, W. Cai, X. Wang and V. C. M. Leung, "Virtual Machine Placement Optimization in Mobile Cloud Gaming Through QoE-Oriented Resource Competition," *IEEE Transactions on Cloud Computing,* pp. 2204-2218, July-Sept 2022.

[9] P. Guturu, "Explosive Wireless Consumer Demand for Network Bandwidth-Fifth Generation and Beyond [Future Directions]," *IEEE Consumer Electronics Magazine,* vol. 6, no. 2, pp. 27-31, 20178.

[10] B. Rikic, D. Samardžija, O. Čadovski and T. Maruna, "Cellular network bandwidth prediction in consumer applications," in *2021 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, 2021.

[11] J. Janković, Ž. Ilić, A. Oračević, S. A. Kazmi and R. Hussain, "Effects of Differentiated 5G Services on Computational and Radio Resource Allocation Performance," *IEEE Transactions on Network and Service Managemen,* vol. 18, no. 2, pp. 2226-2241, 2021.

[12] C. He, Y. Liu, M. Zhi, Y. Zeng, Z. Ren, X. Wang and P. Li, "Q-band Millimeter Wave Communication: Enabling 10 Gb/s Home Network in Fiber-to- The-Room Scenario," in *2021 IEEE MTT-S International Wireless Symposium (IWS)*, Nanjing, 2021.

[13] W. Saad, M. Bennis and M. Chen, "A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems," *IEEE Network,* vol. 32, no. 3, pp. 134-142, 2020.

[14] S. B. RJ Atkinson, "Identifier-Locator Network Protocol (ILNP) Architectural Description (RFC6740)," November 2012. [Online]. Available: https://datatracker.ietf.org/doc/html/rfc6740.

[15] UBS, "ChatGPT sets record for fastest-growing user base - analyst note," Reuters, 2 2 2023. [Online]. Available: https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01. [Accessed 14 6 2023].

[16] C. Veres, "Large Language Models are Not Models of Natural Language: They are Corpus Models," *IEEE Access,* vol. 10, pp. 61970-61979, 2022.

[17] S. Cai, S. Venugopalan, K. Tomanek, A. Narayanan, M. Morris and M. Brenner, "Context-Aware Abbreviation Expansion Using Large Language Models," in *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Seattle, 2022.

[18] L. Friedman, S. Ahuja, D. Allen, Z. Tan, H. Sidahmed, C. Long, J. Xie, G. Schubiner, A. Patel, L. Harsh, B. Chu, Z. Chen and M. Tiwari, "Leveraging Large Language Models in Conversational Recommender Systems," 6 2023. [Online]. Available: https://arxiv.org/abs/2305.07961. [Accessed 6 2023].

[19] S. Wolfram, "ChatGPT Gets Its "Wolfram Superpowers"!," Wolfram|Alpha, 23 3 2023. [Online]. Available: https://writings.stephenwolfram.com/2023/03/chatgpt-gets-its-wolfram-superpowers/. [Accessed 15 6 2023].

[20] J. Liu, C. Chen and Y. Ma, "Modeling Neighbor Discovery in Bluetooth Low Energy Networks," *IEEE Communications Letters,* vol. 16, no. 9, pp. 1439-1441, 2012.

[21] C. Tang, W. Sun, X. Zhang, J. Zheng, W. Wu and J. Sun, "A novel fingerprint positioning method applying vision-based definition for WIFI-based localization," *IEEE Sensors Journal (early access),* 2023.

[22] Y. Zhang, M. Fan and C. Xu, "Intelligent indoor localization based on CSI via radio images and deep learning," *Measurement Science and Technology,* vol. 34, no. 8, 2023.

[23] EnterpriseWeb, "erpriseWeb Premieres the First Telco-grade Demo of Generative AI for Network Service Orchestration," Enterprise Web, 16 5 2023. [Online]. Available: https://www.einpresswire.com/article/631034600/enterpriseweb-premieres-the-first-telco-grade-demo-of-generative-ai-for-network-service-orchestration. [Accessed 6 2023].

[24] V. Bilgram and F. Laarmann, "Accelerating Innovation with Generative AI: AI-augmented Digital Prototyping and Innovation Methods," *IEEE Engineering Management Review,* 2023.

[25] X. Guo, "Towards Automated Software Testing with Generative Adversarial Networks," in *2021 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks - Supplemental Volume (DSN-S)*, Taipei, 2021.

[26] S. Jones, "Toward an acceptable definition of service," *IEEE Software,* vol. 22, no. 3, pp. 87-93, 2005.

[27] P. Nye, "Device Description Language: Factorizing the Control of Arbitrary Networked Devices," in *2014 International Conference on Future Internet of Things and Cloud*, Barcelona, 2014.

[28] C. Chen and A. Helal, "Device Integration in SODA Using the Device Description Language," in *2009 Ninth Annual International Symposium on Applications and the Internet*, Bellevue, 2009.

[29] F. Hang and L. Zhao, "Supporting End-User Service Composition with Automatically Generated Composite Service Descriptions," in *2014 IEEE International Conference on Web Services*, Anchorage, 2014.

[30] L. Sun, H. Dong and J. Ashraf, "Survey of Service Description Languages and Their Issues in Cloud Computing," in *2012 Eighth International Conference on Semantics, Knowledge and Grids*, Beijing, 2012.

[31] D. Fensel, J. A. Hendler, H. Lieberman, W. Wahlster and T. Berners-Lee, Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential, MIT Press, 2005.

[32] P. M. Di Gangi, B. A. Wech, J. D. Hamrick, J. L. Worrell and S. H. Goh, "Risk perceptions about personal Internet-of-Things: Research directions from a multi-panel Delphi study directions from a multi-panel Delphi study," *Journal of Cybersecurity Education, Research and Practice,* vol. 2022, no. 2, p. Article 5, 2023.

[33] S. J. Philip, T. (. Luu and T. Carte, "There's No place like home: Understanding users' intentions toward securing internet-of-things (IoT) smart home networks," *Computers in Human Behavior,* vol. 139, p. 107551, 2023.

[34] S. Khanvilkar and A. Khokhar, "Virtual private networks: an overview with performance evaluation," *IEEE Communications Magazine,* vol. 42, no. 10, pp. 146-154, 2004.