MediaTek 6G Technology White Paper

Enabling Ambient Intelligence

Release date: 2023-06-22

Use of this document and any information contained therein is subject to the terms and conditions. This document is subject to change without notice.

Keywords

MediaTek Inc.

Postal address

No. 1, Dusing 1st Rd. , Hsinchu Science Park,

Hsinchu City, Taiwan 30078

MTK support office address

No. 1, Dusing 1st Rd. , Hsinchu Science Park,

Hsinchu City, Taiwan 30078

Internet

http://www.mediatek.com/

Table of Contents

List of Figures

1. Introduction to Ambient Intelligence

Technology is a primary characteristic of human beings, which at its core is intended to improve our lives. Whether incremental or disruptive, technological advancements form a virtuous cycle of constant innovation for ever greater expected benefits. But with technology usually comes a learning curve that stands in the way of humans adapting to and adopting technology. Where mass adoption of a technology is targeted, user-friendliness and/or engineered abstraction of the technology itself from the end user have become necessary conditions to fully enjoy the benefits the technology provides.

Computing and communication technologies have been the most transformative technologies of the last decades, permeating every aspect of our lives and society. These technologies offer unprecedented benefits that were once in the realm of science fiction. Their omnipresence has been made possible through making their inherent complexity gradually invisible to end users and their benefits and related gratification thus easily available. Advances in these technologies are reinforcing this trend, enabling further transformation of the human/technology interaction; instead of a fragmented environment with separate interfaces and a learning cycle for each device and system, all of the devices and systems shall work together as one. Where once the user had to adapt to a technology prior to adopting it, now technology is being defined that is engineered to adapt to us, and, in a way, eventually adopt us. Already today, smartphones and apps can learn and predict users' behavioral patterns to automate certain prompts and other configurations; cars can automatically adjust the driving mode to the driver, the traffic, weather and road conditions for the safest and most comfortable journey; homes can welcome us as we arrive from work, switching lights and music on; and voice assistants can react to our voice commands. These applications represent the beginnings of a pervasive technological change.

Ambient Intelligence (AmI) is the ultimate realization of the above trend, human-centric technology purpose-built to seamlessly blend into our lives, anticipating our needs rather than reacting to them, with no or minimal user input. It is built to learn about us, our needs and desires, in order to deliver the best response to satisfy them, for any intended application. To this end, AmI relies on the collection of data from a potentially wide variety of sources (including sensors, devices, applications, etc.) and the processing, analysis of and learning from these and other data in order to determine and deliver the best response in the most intuitive form possible to the end user.

Since inception, almost all the generations of wireless systems have been heavily focussed on an end-to-end and over-the-top model, in which the system moves bits between a client device and a distant server (often in an external cloud). AmI assumes greater flexibility in the location of intelligence and computational resources: The ability to adapt to and meet a user's needs will be embodied in devices and nodes scattered throughout the system, resulting in user services developing performance demands incompatible with the end-to-end model. 6G systems[1,2] should accordingly be designed from the beginning with the requirements of the AmI vision in mind.

These requirements include technology advancements in multiple areas, covering both communication and computing, both devices and networks. Existing mechanisms for integrating computation and intelligence into user services will not be equal to the performance and scaling requirements of true ambient intelligence. Large numbers of devices and nodes will need to cooperate to create strength in numbers, with intelligent functionality hosted at potentially any node of the system. A 6G system will be expected to provide users with seamless access to such intelligent functionality from anywhere in the system.

Our original 6G Vision[3] already touched upon the convergence between communication and computing in the 6G era. The present White Paper further dives into how we, at MediaTek, envision 6G will make Ambient Intelligence a reality. A number of use cases and the societal impact we anticipate of AmI are described first, along with the KPIs AmI requires. Then we discuss enabling technologies for AmI, including mechanisms for flexible communication, computation, and caching; device collaboration; implications for the protocol stack; physical layer communication technologies; and integrated sensing and communication.

[1] Next G Alliance Report: 6G Technologies
[2] Next G Alliance Report: 6G Distributed Cloud and Communications Systems
[3] MediaTek: 6G Vision Whitepaper

2. Use Cases and KPI for Ambient Intelligence

Ambient Intelligence will bring a fundamental shift in how humans and technology interact – going from our interacting with technology to technology interacting with us naturally. Ambient Intelligence is a profound and subtle concept where technology will simply “fade into the background”[4] of our everyday lives. In this section we describe how, through example use cases, we envision Ambient Intelligence will enter our lives and the main associated KPIs the 6G System will need to fulfill to realize this vision.

2.1 Use Cases of Ambient Intelligence

Ambient Intelligence lends itself to a large family of use cases, some of which are highlighted here.

Intelligent Homes: Homes contain a wide and growing array of devices and appliances that are increasingly smart and connected, designed to make our daily lives simply more comfortable. These often operate in isolation from one another, optimized for some given task and each of them relying on some specific, more or less advanced (and sometimes complicated) user interface. Ambient Intelligence will enable the creation of a smart ecosystem of interconnected devices and appliances that, rather than imposing their buttons and other knobs on us, will interact naturally with us through our senses, will sense our presence, will identify who we are, and accordingly will be able to seamlessly serve our needs. Not all these devices will need to be “smart” or to have the most advanced connectivity, so long as the ecosystem itself is smart and fully connected. Collectively and collaboratively, all these devices will help each other, learning how to best help us on a day-to-day basis. For instance, some may not be able to perform speech recognition to hear our commands, but they will be able to rely on the speech recognition facility of the ecosystem itself. This will allow intelligent homes that fully adapt to their occupants, their daily lives and their preferences: homes that automatically adjust ambient lighting to each person’s preference, brew each person their favorite coffee, push emergency notifications from loved ones on the TV screen during a family movie when no phone is in the TV room, contact first responders when detecting an emergency, disconnect an electrical plug when identifying a risk of electrocution, etc. Intelligent homes, imbued with AmI, will also assist elderly communities to improve the quality of their everyday life. Sensors equipped with AmI will automatically assess people’s physical conditions, detect any functional decline or frailty, schedule online therapy programs, monitor psychological conditions and report any physical or mental irregularities of the elderly community. This will help countries to prepare for societal sustainability in terms of population aging and prevent unstable situations in the foreseeable future.

Smart Transportation: The automotive industry is gradually shifting its focus from performance to safety, connectivity, entertainment and passengers’ comfort. Ambient Intelligence is creating the platform for smart cars, equipped with autonomous (self) driving technology. Drivers of such intelligent cars will receive assistance and support to prevent accidents and traffic congestion. Fully autonomous vehicles will use efficient AmI, context sensing and context sharing to respond within the precise time to provide fast decisions and avoid the need for human intervention in control and driving. Future smart cars are expected to use Ambient Intelligence in a collaborative manner to form a group or platoon of connected cars and provide smart decisions on vehicular safety and roadside conditions based on combined understanding of their joint contexts.

Mission Critical: AmI is also expected to play a major role in emergency and mission-critical services in our society. Recently, the demand for support systems associated with mission-critical services has increased significantly. For example, there is a strong demand for supporting first responders. First responders are often required to make dynamic decisions in a mission-critical environment. AmI can help in improving this decision-making process by providing a wide variety of knowledge from different sources, according to the particular emergency situation. Similarly, AmI is expected to improve security systems, by automatically identifying suspicious persons or activities and alerting the nearest monitoring station. In future, AmI is also expected to have a vital role in the fast construction of virtual response teams using intelligent robots, as well as assistance for communication and coordination of individual team members, even if they span across different cities or even different countries.

---

[4] Walt Mossberg, “The Disappearing Computer”, 2017

Digital Twins: Digital twins are software models that closely resemble their corresponding real-life counterparts. Digital twins are gradually penetrating across a broad range of applications, which includes precision medicine, 3D immersive environments, digital agriculture and complex engineering design. AmI is expected to be at the heart of these digital twins, enabling them in continuous learning of the latest updates about their physical entities, by using information from sensors, monitoring equipment and software programs. AmI will assist in fusing the ideas of Internet of Things (IoT), metaverse and augmented and virtual reality (AR/VR) and create the up-to-date digital twins. In future, these digital twins will be used to study the effects on the corresponding physical objects at a much reduced cost and time.

As we can see above, bringing AmI to life will require a vast array of interconnected sensors and devices with varying degrees of communication, computing and caching capabilities that will cooperate with one another to serve our needs. It is a distributed and decentralized cooperative platform able to exploit every one of its components to deliver its mission. While somewhat comparable to a circuit board that relies on each of its interconnected parts to offer its functionality but that can fail if a single one of them fails, Ambient Intelligence will instead be malleable and resilient. We list hereafter the primary Key Performance Indicators (KPI) to make Ambient Intelligence a reality.

2.2 Major KPIs related to Ambient Intelligence

In order to realize the vision of AmI, the above-mentioned applications and use cases need to maintain some specific Quality of Service (QoS) and Quality of Experience (QoE). Maintaining QoS and QoE guarantees calls for some major KPI parameters. In the next subsection we investigate the list of major KPI parameters relevant to AmI.

Latency and jitter: In traditional networks latency represents the delay a data packet transiting through the network is subject to, while jitter is the variation of this delay. For a given service, latency is typically bounded by an upper limit beyond which the service can no longer be acceptably delivered. Jitter, even if the actual maximum delay is acceptable, can cause a number of problems like packet loss and out-of-order delivery, which can significantly disrupt the service being delivered. Ambient Intelligence is built on a heterogeneous network topology, where unlike in a traditional network, data sources and data destinations can vary greatly depending on the cooperation between network components; e.g., a pair of AR glasses can on one hand cooperate with a nearby smartphone for content rendering or for exploiting the smartphone's greater radio capabilities, and on the other hand rely on a software update from a remote server. Hence, depending on the cooperation model, the source and destination points are critical, for which latency and jitter requirements can greatly vary.

Practical data rate: Ambient Intelligence relies on the exchange of a vast amount of sensor data and other control information between varying source and destination points depending on the cooperation model (as explained above). It is therefore necessary to ensure that all the links involved in a given cooperative activity are able to sustain the load of data and information exchange in the underlying cooperation model. More than a peak data rate, it is therefore the practical data rate between the source and destination points that matters. In an Ambient Intelligence network, it is vital to establish least-effort data paths between source and destination such that the traffic load in the network can be reduced and data congestion thus averted. This will also contribute to minimizing latency and jitter.

Reliability: Depending on the use case addressed by a given Ambient Intelligence solution, the reliability with which sensor data and control information will be exchanged will vary. When combined with a tight latency requirement, a high reliability is generally particularly challenging if not impossible to achieve end-to-end in traditional network deployments. However, the topology of an Ambient Intelligence solution, where information can flow between source and destination with the least amount of intermediate nodes (e.g., proximal device-to-device communication rather than via network infrastructure), inherently provides a favorable environment for achieving more stringent combinations of latency and reliability for data that require them. This will open up a wider range of applications than achievable today.

Availability: The adjective "Ambient" suggests a notion of spatial presence in the environment where Ambient Intelligence is operating, but also a notion of temporal presence. In other words, Ambient Intelligence is intrinsically present in space and time, available to serve humans. A high degree of availability (hence resilience) is a defining criterion for Ambient Intelligence to exist.

Mobility: Depending on the use case, cooperation mode and source/destination, diverse mobility scenarios can be identified, where the topology of the Ambient Intelligence network varies in time as its components move in relation to each other. It is thus necessary that communication paths can malleably and seamlessly shift as a result of such mobility whilst ensuring data rate, latency and jitter will not be prohibitively affected. As a result of mobility, some components could also dynamically leave or enter this network, which must therefore be able to reconfigure itself where and when needed in a seamless manner.

3. Enabling Technologies for Ambient Intelligence

A wireless network supporting the use cases and KPIs mentioned in Section 2 requires numerous architectural changes and innovations. Next-generation wireless networks are expected to span across multiple heterogenous and distributed platforms with varying capabilities. Designing such a wireless network calls for a flexible and programmable architecture for deployment in local personal-area networks (PANs) as well as in large-scale wide-area networks (WANs). To achieve this programmable network, we are exploring a flexible architecture for communication, computing and caching.

3.1 Flexible Communication, Computing and Caching

Building on the principle of separate user and control planes, introduced late in the 4G lifecycle, the 5G architecture introduced a total separation from the outset, namely separating the data session handler (SMF[5]) in the control plane from the data handler (UPF[6]) in the user plane. This enabled scaling the user plane independently from the control plane, allowing flexible deployment and/or selection of UPFs as necessary according to anticipated data demands and latency requirements, without in turn having to scale the SMF deployment accordingly. Fundamentally, such UPF distribution also enabled reducing the data latency and data traffic load through the network. However, this only reduced latency within the 5G system, while end-to-end latency remained unchanged, due to application servers being deployed in the cloud (internet). Another major change was necessary to reduce end-to-end latency, critical for some applications (e.g., XR): edge computing. Multi-access Edge Computing (MEC) in 5G enables deployment of application servers close to the end users, at or in the Radio Access Network (RAN) (i.e., "edge") by means of MEC server platforms that offer the computing and caching capabilities necessary for these application servers to operate.

In order to harness the latency benefits, MEC uses a local UPF, which enables direct transmission of data traffic between the device and a MEC server, without using a UPF up in the core network. A local UPF also helps in reducing the data traffic load through the network and in turn averting the risk of data congestion. Introduction of the local UPF is an important step to enable local execution of computing tasks, yielding significant latency and data traffic reduction, thereby not only improving the efficiency of existing services, but also enabling new latency-constrained services. However, MEC is defined based on an end-to-end data model, with computing being entirely performed in an application server, even if associated with a network edge. This can be well suited for compute-intensive, low-latency (e.g., max ~10ms e2e) tasks; but it may not be suited for even more time-critical activity, like rendering of AR content (<10ms refresh), and it may also be overkill for less time-critical applications. Computing and caching capabilities are also becoming more widespread and diverse. On one hand, they are increasingly abundant in mobile devices, e.g., smartphones acting as "local hubs" for tethered devices, like wearables, as well as in connected cars, voice assistants, smart TVs, network equipment, etc. On the other hand, these capabilities can also be strictly limited in devices restricted by form factor and/or battery size, such as IoT sensors and actuators, wearables like AR glasses, etc. These energy- and computationally-constrained devices therefore require assistance to execute tasks (e.g., exchanging data with a remote end, rendering content).

Ambient Intelligence, which requires data collection, processing and learning from a wide variety of sources and delivering a timely and intuitive response to the end user, will therefore be communication-, computing- and caching-intensive, and in turn require a fundamentally different approach from the existing 5G MEC. Our 6G vision lies in a fluid and flexible distribution of communication, computing or caching resources across any point in the system, such that a given task can be executed in a timely manner no matter the computing, communication and storage demand.

---

[5] Session Management Function – manages data sessions between the device and the network, where data sessions are anchored in the UPF

[6] User Plane Function – data gateway, anchor of data sessions for traffic exchange between the device and a data network (e.g., the internet), incl. all related tasks (e.g., IP address allocation, QoS handling and packet marking, etc.)

The benefits are multi-fold, for instance:

- Network operators and network vendors can expand their current communication-centric services to a multi-dimensional service that bundles communication, computing and caching resources.
- Further reduction of the traffic load across the network and of the associated risk of data congestion is possible.
- Support becomes possible for a new generation of applications that require ultra-high data rates, stringent latencies and intense computing and caching resources, without mandating massive processing, storage or power resources at the devices themselves.
- New device form factors and new device usage will be possible, e.g., thin devices (thin clients) that can leverage the communication, computing and caching capabilities of their environment.

In order to realize this vision, the 6G system architecture will accommodate flexible termination of user plane connectivity at points throughout the network, where the termination point(s) can vary depending on scheduled communication, computing and caching resources on different nodes, based on service requirements including computational intensity and latency sensitivity.

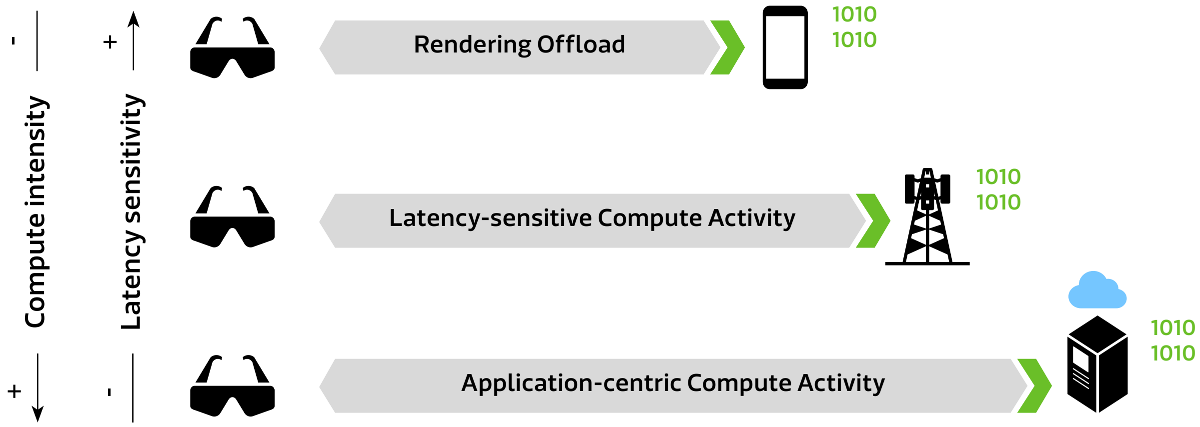

Figure 1: Flexible Architecture for Distributed Computing

As shown in Figure 1, a pair of AR glasses, which present an intuitive interface to the user, could flexibly offload compute activity across various nodes in the system, depending on the type of computation required (e.g., computational intensity vs. latency sensitivity and node capability), in a manner that is completely transparent to the end user (in this case the person wearing the AR glasses). For example, it could offload latency-critical rendering computation to a smartphone in proximity that is able to perform rendering, while offloading more compute-intensive but less latency-sensitive computational tasks to a RAN node and/or compute-intensive tasks to an application server.

With the evolution of MEC and edge intelligence, communication, computing, and caching (storage) capabilities will gradually proliferate at the edge of the network, enabling provisioning of brand new services and applications. Our 6G vision of flexible communication, computing and caching is illustrated in Figure 2, which displays a highly heterogeneous and distributed system, ranging from device/hyperlocal clouds, containing extremely local in-body devices, to wide area networks. Services in these hyperlocal clouds may be offered and consumed locally in real time and thus will not necessarily require connectivity with remote service platforms. This vision will require diverse forms of radio access, from device-to-device multi-hop communication to more traditional device-network communication. As motivated in our original 6G Vision[7], wireless access convergence will be instrumental in bringing this vision to fruition through hybrid nodes and unified radio access with flexible computing, communication and caching (FCCC).

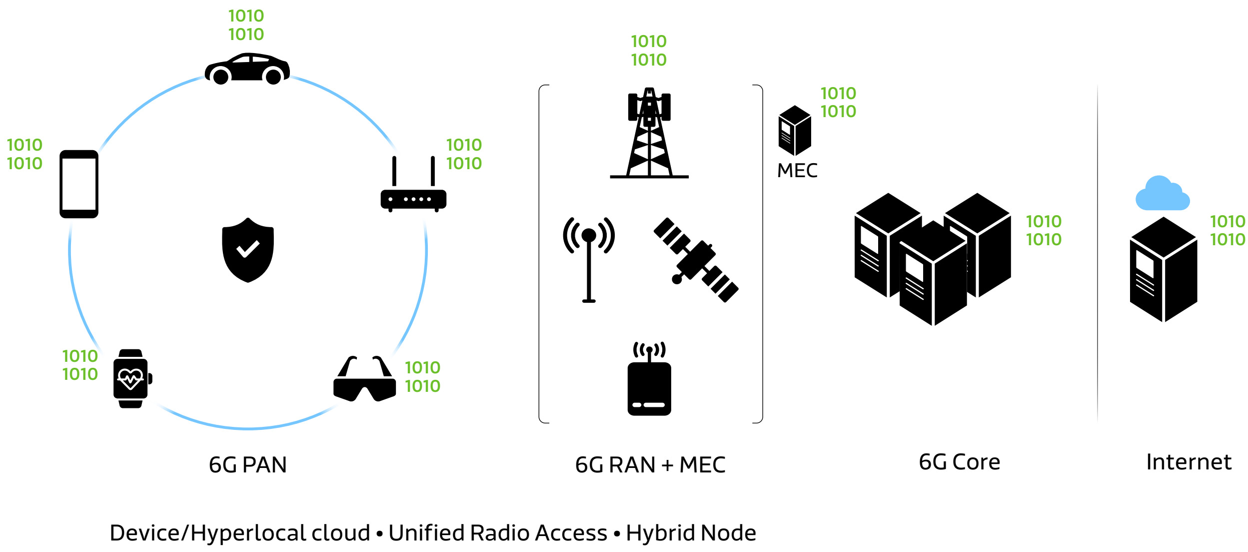

Figure 2: Flexible Communication, Computing and Caching in the 6G System

3.2 Device Collaboration for Flexible Computing

The vision of FCCC, as a prerequisite of AmI, calls for transforming the existing protocol stack into a scalable structure, having flexible termination points, which treats computational transactions as data and routes them throughout the 6G system homogeneously. To realize this vision, a single routing mechanism should govern data (or control signalling) routing throughout the system, covering both device-device and device-network packet routing. Such a unified adaptation layer will act as a multipath routing protocol, supporting a computation model in which a data consumer node divides the computational transactions across multiple termination nodes. Each node that is enabled with routing capabilities will instantiate the adaptation layer. Accordingly, a user plane data packet or a signalling message is addressed to a particular termination node, and other intermediate nodes route the packet/message until it reaches the termination node. Each node can also instantiate certain additional functions related to termination of the packet, e.g., security, header compression, etc. These functions may be embodied in the adaptation layer itself or in upper layers. An example of the use of such a universal adaptation layer is shown in Figure 3.

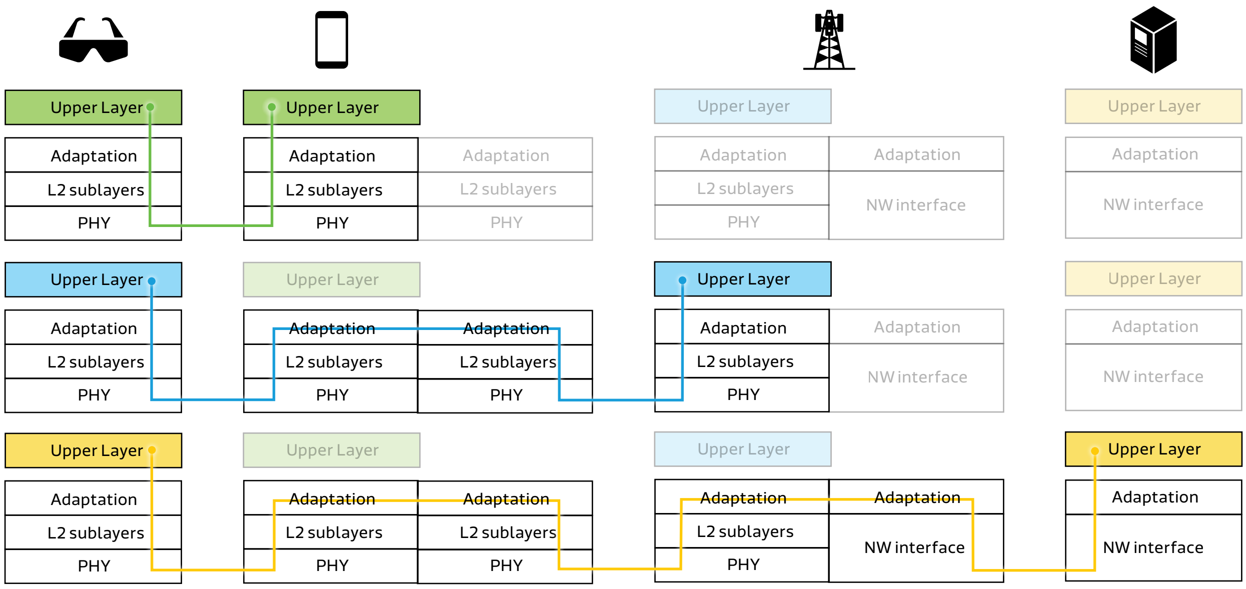

Figure 3: Adaptation Protocol for Distributed Computing in 6G

[7] MediaTek: 6G Vision Whitepaper, op. cit.

Expanding the protocol architecture throughout the entire system allows routing a diversity of protocols, in addition to user-plane data, using a common protocol architecture. This is a tangible change from today's "hard" node relationships, where communication (for computing or other purposes) is performed on specific reference points in the device and RAN domains. Instead, we envision computational transactions to be distributed over a multitude of termination nodes augmenting the capabilities of a data consumer node. As a result, the whole system interacts like a mesh, wherein termination functions (e.g., security) and associated computing are provided throughout by the adaptation layer and/or upper layers, rather than embedded in the lower-layer protocols that are used for point-to-point data delivery.

This eliminates the "nested layers" effect and develops a more homogeneous protocol architecture, with computing distributed and coordinated across multiple network nodes.

3.3 Protocol Stack for Device Collaboration

We next take a look at the layer 2 (L2) sublayer and the transport of data between nodes in a distributed system of collaborative devices. In NR, the protocol stack is composed of several layers, each associated with specific functionality. Data arrive into the L2 sublayer classified into flows based on their TCP/UDP/IP headers. These data flows are mapped to radio bearers which form the basis of differential handling of data in L2. All data flows that can be jointly handled are merged into a radio bearer, simplifying handling of flows within the stack. Further L2 functionality such as security, duplication, Automatic Repeat Request (ARQ) and data prioritization can be controlled for a radio bearer. The list of functionality present in the NR stack is quite significant and can be very flexibly configured.

As we consider a distributed architecture, there are new and interesting questions to consider. For example, do we handle data from different users differently? What data reliability should be targeted and how do we meet it? How would we schedule data at different nodes in the system? There are also new challenges to meet with a mesh-type architecture. Data rates and packet rates at intermediate nodes, especially as we get close to or at the radio interface, can quickly rise with additional remote nodes in the system, as will the number of radio bearers to handle. These questions and challenges warrant a closer look at the protocol stack in the L2 sublayer, which has largely remained architecturally unchanged for over 20 years.

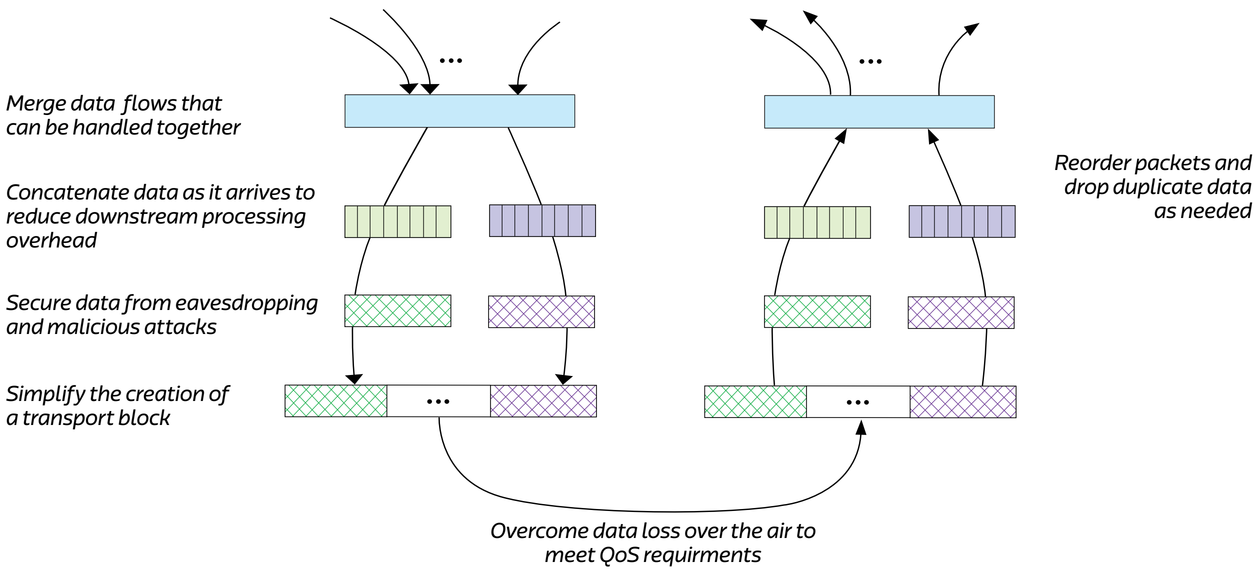

Figure 4: Functionality of the Protocol Stack

Functionality within the NR protocol stack runs at the granularity of a packet. Such an approach was seen as simpler and more hardware friendly for implementation. However, as packets from remote nodes accumulate in nodes closer to and over the radio interface, the processing overhead in the protocol stack can become quite significant, resulting in an unacceptably high power consumption penalty for such nodes in the system. A rethink of data handling in the stack is needed, such as concatenating packets as they arrive, to ensure that the design is scalable at relay nodes while still retaining the desirable property of being hardware friendly.

Another aspect to consider is the simplification of actual functionality that runs in the protocol stack. For example, we currently have two mechanisms for reliability in the protocol stack, i.e., Hybrid ARQ (HARQ) at MAC and ARQ at Radio Link Control (RLC). We need to consider whether we continue to require two mechanisms going forward, or whether these can be combined into one single mechanism for reliability handling. Similarly, we also need to take a closer look at scheduling of data, which is currently based on a scheduling request and buffer status reporting mechanism and can take several steps before scheduling of data transmission can take place. It is important to reduce scheduling latencies where possible to have a responsive system, as these latencies can add up across a distributed network.

In order to enable network operators to pivot from a traditional communication-centric service model to an FCCC-type service model, it is imperative that the protocol stack design and its operation for collaborative, distributed services is not different from that used for data transport in traditional service models. Therefore, collaborative services should be considered as a fundamental design requirement when discussing the protocol stack for 6G rather than attempting a clunky bolt-on to the baseline 6G protocol stack at a later point in time.

3.4 Physical Layer Communications for Device Collaboration

The vision of a distributed 6G system, imbued with AmI and flexible termination points, described in the previous subsections, assumes widespread communication between different points in the system, including coordinated communication between devices in close proximity. This coordinated set-up can also be exploited to augment capabilities of a device, if it has limitations in its physical layer capabilities.

In such a flexible network, devices need to be capable of discovering and communicating with other devices to accomplish end-to-end service. The coordination between nodes may be realized at the network side, the device side, or both. Multiple transmission/reception points (M-TRPs), dynamic point selection (DPS), non-coherent joint transmission (NCJT), and coherent joint transmission (CJT) are some key physical layer enablers to achieve collaborative communications on the network side. Similarly, collaborative communications at the device side can be enabled, and wireless-based techniquesfor connections among the collaborative devices are preferred for end-terminal user experience and deployments.

Currently, several wireless connection techniques are already available to connect collaborative devices, including sidelink, RF repeater (layer 1 or L1 relay), L2/L3 relays, WiFi, Bluetooth, etc. However, conventional L2/L3 relay and tethering from the Uu (device to network) interface to either WiFi or sidelink (device to device) interface will introduce additional latency, which may prevent the service from meeting demanding latency and data rate requirements for real-time services like AR/VR. On the other hand, an RF repeater leads to almost no additional delay, but it is difficult to realize on portable devices. In this section we would like to introduce L1 relay as a frequency-translation Amplify and Forward (AF) repeater to aggregate capability among collaborative devices, while latency introduced by the advanced relaying is minimized.

Figure 5 is a simple illustration of the principle of device-based radio collaboration between two devices (here shown in downlink direction), in which a set of AR glasses leverages the higher radio capabilities of the proximal device to which it is tethered to augment its own capabilities and resulting performance. In this illustration, the AR glasses are connected to the network and to a nearby smartphone such that the downlink radio signals between the network and the glasses are also received (dotted line) by the smartphone (under control of the glasses), which forwards them for combining by the AR glasses themselves. In uplink direction a similar principle can be exploited as well, where the base station exploits augmented signals resulting from the combination of the signals emanating directly from the AR glasses themselves and the corresponding ones from the proxy smartphone.

In other words, high- and low-capability devices can collaborate using L1 relay as a frequency-translation AF repeater to obtain path diversity gain, e.g., a high-capability device (such as a smartphone or CPE) can serve as a proxy to provide

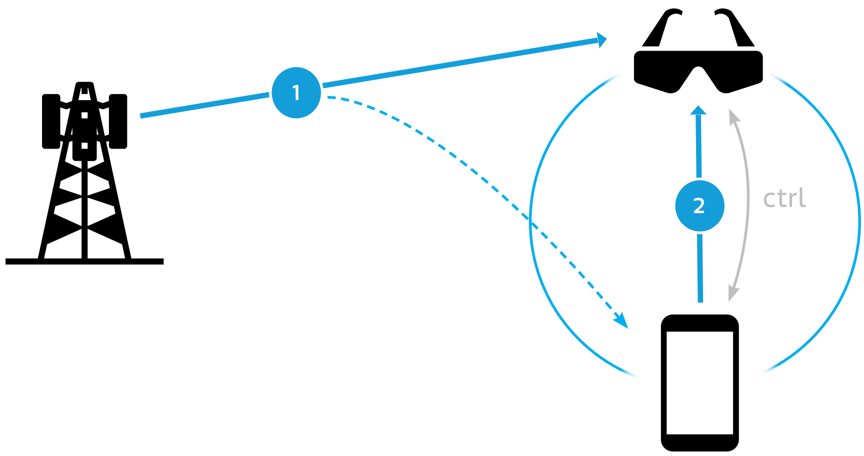

Figure 5: Device-based Radio Collaboration Principle (Downlink)

a good-quality data path for a lower-capability device (such as AR glasses or other wearables). In the above example, the smartphone acts as an advanced AF repeater to improve signal quality by rejecting interference from unexpected directions, and subsequently forwarding the processed signal to the lower-capability AR glasses.

It should be noted that designing a smart phone with full-duplex AF functionality and input/output signal in the same band is a challenge. Hence, instead of in-band AF (i.e., where (1) and (2) would operate on the same carrier inter-band frequency-translation AF can be a viable option where (1) and (2) operate on separate carriers. Similarly, a few collaborative devices in proximity can be grouped to form a single virtual end device, equipped with many antennas. This also requires exploring the functionality of frequency-translating AF to connect collaborative devices in proximity.

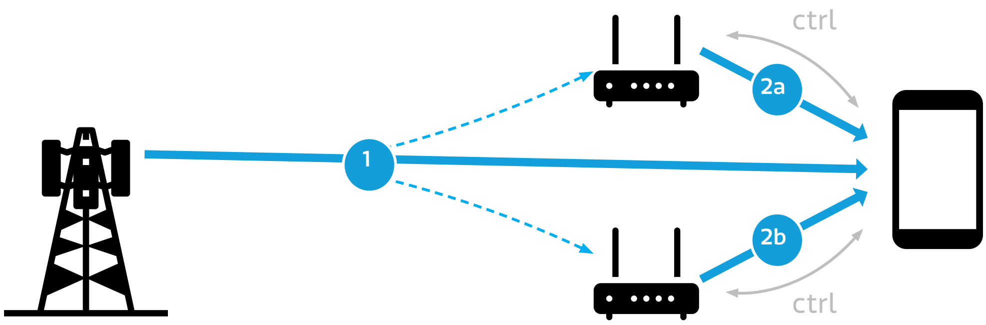

Figure 6: Device Collaboration using Multiple Rx Antennas (Downlink)

Figure 6 shows an example of antenna aggregation across multiple devices in an indoor scenario, where two indoor CPEs and one smartphone collaborate to receive, at the smartphone, a resulting high-rank data signal from the base station. The smartphone utilizes its carrier aggregation to receive signals (1), (2a), (2b) simultaneously in three bands and processes the received signals jointly, as if having three sets of Rx antennas. The base station transmits a signal (1) to an end device that appears to be equipped with three sets of Rx antennas in a single band. The concept of aggregating antenna capability can be extended to public networks, although device discovery and security issues can be more challenging.

Collaborative communications across multiple devices in close proximity can also exploit short range communication (SRC). Radio Frequency Identification (RFID), ultra-wide band (UWB), near field communications (NFC), IEEE 802.11 WLANs, Bluetooth and ZigBee are some of the promising solutions for the challenges of AmI in future networks. In the future SRC is expected to be evolved and built on a unified 6G WAN + SRC lower-PHY design to avoid additional efforts and cost on hardware design. In Sub-6GHz and mmWave spectrum, SRC can share spectrum with 5G and/or 6G systems (including refarmed 6G systems: SRC+5G+6G), or use dedicated spectrum such as the ITS band shared with sidelink (SL), with all cases requiring some coexistence measures.

Sub-THz is a new, emerging spectrum for 6G, with the potential for extreme speed, proximity communication and high-resolution sensing, all of which would be required for some services such as immersive XR. We expect that in 6G wireless, a unified low-PHY design and implementation of SRC and cellular communication will take place with shared hardware implementation (potentially different parameter profiles). SRC-enabled device collaboration will help in HARQ transmission/re-transmission from the hybrid node based on the unified air interface, SL tethering for XR and mesh/multi-path for coverage extension, throughput enhancement, and power saving.

3.5 Integrated Sensing and Communication for Ambient Intelligence

Most of the AmI use cases and applications mentioned in Section 2.1, like intelligent homes, smart city, connected cars, and assisted living and caregiving, demand fast and reliable communication, as well as high-accuracy sensing capabilities. While communication and sensing are traditionally viewed as two subjects, their underlying techniques are quite similar in terms of signal processing, hardware components and channel characteristics. Sensing and communication functionalities can thus be integrated by exploiting common hardware and radio resources. Integrated sensing and communication (ISAC) envisions more cost efficiency, compact size, less power consumption, spectrum sharing, improved performance from mutual assistance, and safety due to enhanced information exchange.

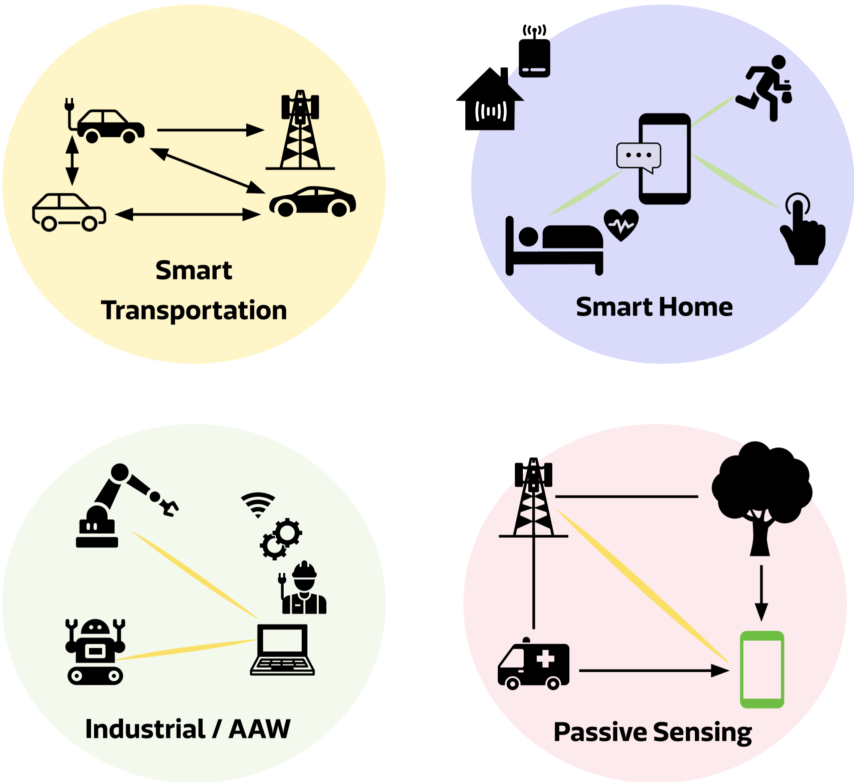

Figure 7: Integrated Sensing and Communications (ISAC) with AmI

Figure 7 illustrates how ISAC will be useful to improve the efficiency of some use cases mentioned in Section 2.1. ISAC is expected to enrich the smart transportation by providing information on obstacle detection and avoidance and efficient traffic monitoring. Similarly, in intelligent homes, smart industry (Ambient Assisted Working or AAW), ageing and caregiving scenarios, which require a smart ecosystem of interconnected devices and appliances, efficient sensing is essential, and the combination of efficient sensing with fast communication makes AmI more accessible. ISAC also helps in enhancing physical communications by obtaining knowledge of obstacles and scatters, and subsequently performing more accurate beam management and maximum permissible exposure control. Finally, automated high-resolution sensing, coupled with fast and efficient communications, will be useful in gesture and activity.

recognitions, required in XR and other applications involving fast human-computer interactions.

Enabling ISAC for AmI envisages research in several key technology directions. To start with, channel modeling for ISAC requires ray-tracing techniques, as statistical methods are not sufficient for sensing. Secondly, the 6G framework needs to consider holistically associated detection and estimation algorithms with unified waveform, reference signal and transceiver design, such that ISAC may attain the highest possible efficiency. To achieve active sensing, full-duplex ISAC transceiver design and its Tx-to-Rx performance characterization must be studied for implementation feasibility. Beyond far-field assumptions of antenna arrays, sensing could necessitate near-field signal processing under increased array size and operating frequency range, further enabling beam focusing techniques for multiple access. It is expected that research in new physical-layer, medium-access-control and cross-layer procedures to support and coordinate communication and sensing will grow substantially.

4. Conclusion

The needs and desires of users in the timeframe of 6G systems will advance to an extraordinary degree, and the technologies to meet these user requirements will involve significant complexity, which must be managed within the system in the AmI paradigm rather than devolving on users themselves. From the user perspective, 6G is expected to be responsive to the intentions of the user rather than presenting a set of complex interfaces as an obstacle to accomplishing those intentions. Elements within the system (e.g., devices, network nodes) will cooperate in a seamless manner to realize the promise of the evolving technology in a manner that transparently optimizes itself to meet user needs and desires.

AmI is expected to find a wide variety of use cases, especially in the domain of cyber-physical control applications that bring advanced capabilities of the system into unstructured environments in everyday life. Caregiving applications, connected cars and smart cities, and industrial and mission-critical applications all represent use cases where AmI can provide frictionless operation between the computing and human spheres.

AmI depends on the capabilities of many entities (devices and/or network components) cooperating, with the interfaces and coordination transparent to the user, so that the system as a whole responds to the user's intentions. Accordingly, nodes in the 6G system must be able to communicate flexibly, sharing information and computational tasks where needed, while still meeting the privacy requirements applicable to the user. This expectation of flexible communication, computing and caching will be reflected in the architecture of the system. Device collaboration in a hyperlocal cloud will require that devices be able to communicate with one another in the same manner as communicating with network nodes, and also that the protocol stack as a whole be designed for efficiency and scalability, with flexible routing in mind, in a way that does not unduly burden implementations.

The physical layer, too, will be affected by the need to support AmI. Devices will need to communicate with multiple peers having different communication capabilities and characteristics, requiring coordination between diverse nodes in a variety of scenarios. Short range communication, in an evolved form that shares characteristics with the 6G WAN physical layer, is expected to be used for device collaboration. The importance of sensing for AmI will be reflected in its integration with communication, allowing hardware and radio resources to be shared between the two operations for efficiency. This integration is expected to require new channel models, detection and estimation algorithms, and signal design.

The development of AmI will require progress in additional areas. For example, robust assurances of security and privacy, which are already critical in wireless systems, will be central to the usability of AmI, due to the diffusion of services through the system that may rely on sensitive user data. The role of artificial intelligence and machine learning will expand, as AI-enabled devices and network nodes offer their capabilities to user-facing services to enable AmI functionality.

Ambient Intelligence will appear in the 6G landscape as both a driver of technologies and an enabler of new system capabilities. It is imperative that the needs of a world of Ambient Intelligence be considered from the beginning of the 6G system design process, including support of the fundamental technological building blocks that enable the Ambient Intelligence concept.