SIX KEY TRENDS MANIFESTING THE PLATFORM FOR INNOVATION

TECHNOLOGY TRENDS 2019

BY: ERIK EKUDDEN, CTO

Affordable and efficient connectivity is a fundamental component of digitalization and has become as important as clean water and electricity in creating a sustainable society of the future. Recognition of this fact is of critical importance as we enter a new era that is defined by the combinatorial effects of a multitude of transformative technologies in areas such as mobility, the Internet of Things (IoT), distributed computing and artificial intelligence (AI).

The universal connectivity network that we use today is built on voice and mobile broadband services that currently serve 9 billion connected devices globally. This technology is recognized and acknowledged for its availability, reliability, integrity and affordability, and it is trusted to handle sensitive and important information. Today's network provides pervasive global coverage on a scale with which no other technology can compete. It has quickly become a multipurpose network, ready and able to onboard all types of users, as well as supporting a large number of new use cases and a plethora of new technologies to meet any consumer or enterprise need. As such, it is ideally suited to serve as the foundation for future innovation in any application.

APPROPRIATE AND UNIVERSAL CONNECTIVITY

The multipurpose network is significantly more cost-efficient than specialized or dedicated network solutions, making it the most affordable solution to address society's needs across the spectrum from human-to-human to human-to-thing and thing-to-thing communication. It supports everything from traditional voice calls to immersive human-to-human communication experiences. In terms of human-to-thing communication, it enables everything from digital payments to voice-controlled digital assistants, as well as real-time sensitive drone control and high-quality media streaming.

With regard to IoT communication, the ubiquitous connectivity provided by the multipurpose network enables the creation of a physical world that is fully automated and programmable. Examples of this include massive sensor monitoring, fully autonomous physical processes such as self-driving cars and manufacturing robots, as well as digitally-embedded processes such as autonomous decision-making in tax returns.

KEY TECHNOLOGY TRENDS

In my view, the ongoing evolution toward the future network continues to rely heavily on the five key technology trends that I outlined in last year's trends article. Therefore, in this year's technology trends article, I have chosen to build on last year's conclusions and share my view of the future network platform in relation to those five trends, with one addition: distributed compute and storage.

CTO TECHNOLOGY **TRENDS 2019** ✱

Demanding use cases exemplified by trends 1 and 2

Today's networks are transforming into a platform where applications, processes and other technologies are developed, deployed and enhanced. For me, it is fundamental that the platform ensures affordable, reliable and trusted operation. Two use cases that I expect the network platform will need to support are trends 1 and 2: the Internet of Skills and cyber-physical systems (CPSs).

TREND #1: INTERNET OF SKILLS

The Internet of Skills has the potential to bridge the geographical distance between humans as well as between humans and things. A high quality of experience (QoE) is essential to create immersive interactions that allow humans to attend meetings remotely with the same ability to participate as if they were physically present. Humans have to trust the network to enable critical remote operations and interaction with things.

Self-driving vehicles will require a remote person to take over the driving or support in the decision-making if the autonomous system fails. Hence, tele-operation of robots and vehicles is needed at sea, on land and underground, as well as in the air. Remote human assistance is also required for tasks such as maintenance, troubleshooting and repairing across industrial, enterprise, health care and consumer domains. The Internet of Skills also applies to the ability to experience physical items remotely in applications such as online shopping and gaming.

High-quality and efficient capturing, transmission and rendering of visual, audio and haptic information is essential to the Internet of Skills. This information will be captured by multiple devices and it must be fused together to be reproduced remotely. A distributed environment for access, compute and storage of this information is therefore highly advantageous.

Haptic communications require latencies below 10ms in the most demanding scenarios. Large volumes of 3D visual data and high-frequency haptic data impose high network bandwidth and latency demands, both in the uplink and downlink.

A network platform with low-latency characteristics allows for large amounts of data to be quickly transmitted between devices. This means that more time can be spent on processing and performing analytics on the available information to enhance the experience.

Security and privacy are very important since the devices may capture sensitive visual, audio and haptic information. This information can relate to the user of the device or other users that share the same environment, including detailed characteristics of the user's physical environment such as their home or office, as well as insights into the user's daily activities.

The network platform will also be very beneficial for enabling the positioning of devices, both outdoors and indoors. The network radio positioning information can be fused with information from the device's onboard sensors such as the camera and inertial sensors.

TREND #2: CYBER-PHYSICAL SYSTEMS

CPS results from the integration of different systems to control a physical process and uses feedback to adapt to new conditions in real time. This is achieved by integrating physical processes, networking and computation. A CPS generates and acquires data, so that the relevant elements involved have access to the appropriate information at the right time. Therefore, the CPS can autonomously determine its current operating status, and corrective actions are realized by the actuators. Information comes from sensors and from other related CPSs. The role of humans is to supervise the operation of the automated and self-organizing processes.

Communication is vital in CPSs to allow different and heterogeneous objects to exchange information with each other and with humans, at any time and in any conditions. Deterministic communication (in terms of latency, bandwidth and reliability) largely impacts the dynamic interactions between subsystems in CPSs. Minimizing the time it takes to perform control tasks is critical to ensuring that a system functions correctly.

The future network platform should provide the specific connectivity performance to guarantee CPS-critical requirements. As an example, latency criticality is an issue for all cases where a controller or complex AI must take decisions and actions in real time.

Each CPS has a specific architecture that requires an adaptive network platform. Hence, a specific ad-hoc design of indoor and/or outdoor coverage is required. In addition, network slicing will enable satisfying heterogeneous connectivity requirements on the same network, for any indoor or outdoor scenarios.

EXAMPLES OF CYBER-PHYSICAL SYSTEMS

PORTS OF THE FUTURE

Terminal port operations will increasingly consist of a mixture of physical machinery, robotics systems, automated vehicles, human-operated digital platforms and AI-based software systems. These elements will transform future ports into CPSs, creating a digital ecosystem comprised of various intelligent agents highly specialized in specific aspects of cargo loading/unloading and of the logistic chains.

AUTOMOTIVE

All new features in modern cars, such as advanced driver assistance systems and connected vehicle services, are based on electronics and software rather than on mechanical engineering innovations. Safety-critical functions, driver-assistance software and infotainment applications will run in specific and highly compartmentalized onboard modules that interact with a plethora of sensors and actuators. In this context, the future vehicle will increasingly take the form of a CPS for which the prevention of accidents is the main goal.

SMART MANUFACTURING

The factory of the future will be a set of interacting CPSs, where highly skilled workers will have direct insight into the operations of coordinated intelligent machines from a central control entity. Every functional aspect of a production chain will be affected – from design, to manufacturing, through to supply chains, and later extending to customer service and support. The smart factory will be hyper-connected, data-intensive and highly secure.

My vision of the future network platform

As I see it, the future network platform is characterized by its capability to instantaneously meet any application needs. It can handle huge amounts of data, scarce amounts of data, and everything in between. It will meet requirements for both open data and sensitive data, as well as all manner of needs related to uplink and downlink transmission. From real-time critical to non-critical, predefined to flexible air interface, preset to adaptive routing – the future network platform has it covered. Anyone and anything that can benefit from a connection should be able to access and use the network.

MAIN CHARACTERISTICS

The interconnect between different kinds of networks, from local to wide-area coverage, builds a global network that provides a platform for pervasive global services. The inherent mobility within and between the networks creates unprecedented coverage both indoors and outdoors. Utilizing all these network assets enables a distributed environment for access, compute and storage. These assets are virtualized, distributed across the network, and are made available where they are needed and are most efficient. Applications and processes are dynamically deployed throughout the network. Network slicing enables streamlined connections for different applications, enhancing the efficiency of the total usage of the network.

Autonomous deployment, operation and orchestration is an essential capability of the network platform to enable cost-efficiency. Just as important are the reliability and resilience to fulfill expectations from industry and society. Built-in, automated security functions protect the network and the integrity of its users from external threats.

THE NETWORK PLATFORM OFFERING

The network platform offers a wide range of capabilities to all its users. It provides a seamless universal connectivity fabric with almost unlimited, scalable and affordable distributed compute and storage. Sensors and actuators can be attached anywhere throughout the network. Latency can be optimized by interacting with the control of access, compute and storage. Embedded into the platform is a distributed intelligence that supports users with insights and reasoning.

The addressability and reachability capabilities make it possible to connect anyone or anything regardless of location and time. Together with the inherent security and availability, the network platform can also meet communication needs relating to secure identification of users and networks. It also provides the scalability to automatically adapt to the exact needs of individual users and applications. As an example, adaptive power consumption is enabled by a flexible air interface. Another example is automated life-cycle management of devices, users and applications. This guarantees the most cost-efficient solution for users, in both the long and short term.

The network platform offering is consumed through an automated digital marketplace. Network services and data are available through consistent and open business interfaces for the applications (APIs). Data, such as location, connectivity conditions and user behavior, can be made available from the network platform.

With all these capabilities, the network platform offers the most accessible and valuable foundation for future innovation.

Four technologies evolving the network platform: Trends 3-6

In my view, four technology areas are crucial to the evolution of the future network platform, represented by trends 3 to 6: distributed compute and storage, ubiquitous radio access, security assurance and zero-touch networks.

TREND #3: DISTRIBUTED COMPUTE AND STORAGE

Future applications will require new processing capabilities from the network in order to reduce the amount of data that needs to be communicated, provide low latency, and increase robustness and security.

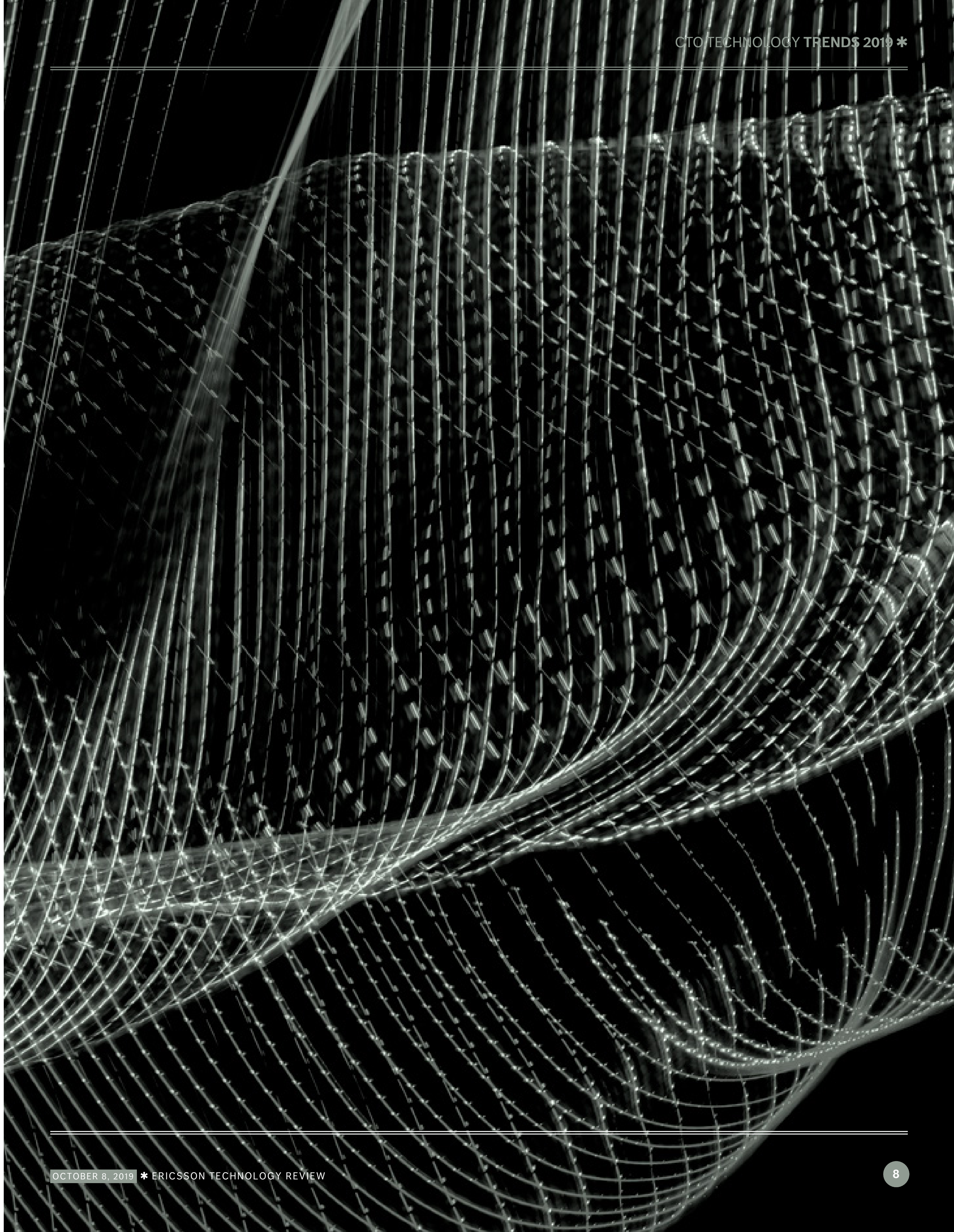
Today's processors and accelerators will eventually experience the end of Moore's Law, and new heterogeneous computing solutions will emerge. Commodity hardware has been joined by a highly heterogeneous set of specialized chipsets – often referred to as accelerators – that are optimized for a certain class of applications. For example, data-intensive applications such as machine learning (ML)/AI or augmented reality/virtual reality can take advantage of the massive parallelization offered by Graphical Processing Units or Tensor Processing Units. Latency-sensitive applications can utilize computation pattern reuse offered by either custom-designed integrated circuits or field-programmable integrated circuits.

The next step of heterogeneous computing will involve new computing paradigms such as neuromorphic processors that yield low power consumption, fast inference and event-driven information processing. Another emerging technology is photonic computing. Photons are used instead of electrons, thus avoiding the latency of the electron-switching times. Quantum processor-based acceleration of compute-intensive and latency-sensitive algorithms will eventually become a reality. By exploiting the quantum mechanics principles such as superposition and entanglement, quantum processors promise exponential growth of computing power for a certain class of problems.

The emergence of universal memories will offer the capacity and persistency features of storage, combined with byte-addressability and increased access speed of memory. Programs written for persistent memories can remove the distinction between runtime data structures and offline data storage structures, resulting in faster start-up times and recovery in case of failover. Advancements in non-volatile memory technologies will be crucial to meet strict latency requirements.

The increasing disparity of central processing unit speeds versus memory access speeds will lead to memory-centric compute architectures. Compute units will be embedded inside the memory or the storage fabrics. This will not only increase performance, but also lead to significant energy-efficiency gains by reducing the data movement of traditional compute-centric architectures.

Efficiently developing applications for a distributed compute environment will require new programming models. Programs will benefit from separating the intent of the application from the how



CTO TECHNOLOGY TRENDS 2019 ✱

and where of the physical network. Today, intent-based networking uses Service Level Agreements and policies to define the intent of network operations. The network configures, monitors and troubleshoots issues in the network to fulfill these intents. In the future, there will be more cloud services managed by intent-based operations to evolve toward more advanced automation.

The network platform will benefit from the seamless integration of specialized compute and storage hardware to boost performance for a wider range of emerging, complex applications. The advanced compute and storage capabilities will be moved to the edge of the network, closer to where the data is generated. Further, the network will be able to support developers with efficient and transparent programming models. Edge-native applications will be designed from the ground up to fully capitalize on compute and storage resources anywhere.

TREND #4: UBIQUITOUS RADIO ACCESS

Improved indoor coverage, maximal energy efficiency, fiber-like performance and support for both small cells and a wide range of new use cases are key features of the 5G networks that are currently being rolled out. These networks will be the baseline for future radio networks and the network platform itself.

Future wireless access networks will consist of a wide range of different types of nodes jointly providing wireless access coverage. Devices will in many cases have simultaneous connectivity to multiple network nodes, including different access technologies, for enhanced performance and reliability. Wireless technology will also be used for the connectivity between the network nodes, as a complement to fiber-based connectivity.

Network coverage will be further extended by making use of intermediate devices to forward data to devices outside the coverage of the basic network. Device cooperation can be used to create virtual large antenna array by combining the antennas of multiple devices, which requires tight synchronization. As the network is becoming increasingly dense with a greater amount of small low-power network nodes, and with devices contributing to the overall connectivity, the border between devices and network nodes may be more diffuse.

Key to the management of this kind of massive heterogenous network, with a much more mesh-like connectivity, will be the development and utilization of advanced AI functionality. This will enable the network to evolve and adapt over time to new requirements and changes in the environment.

Operation above 100GHz will enable terabit-per-second data rates, although only for truly short-range connectivity. There are currently implementation challenges for this frequency range, such as how to generate substantial power and the heat dissipation, considering the inherently small dimensions of the components, including antennas. The extension to higher-frequency operation and use of beam-formed transmissions will enable enhancements in spectrum sharing.

In the higher layers of RANs and core networks, the evolution toward cloud-native implementation and automation continues. Network interfaces are moving away from traditional point-to-point interfaces toward more services-based application interfaces decoupled from underlying transport connections. Cloud-native implementation of stateless network functions use external context storage for redundancy and context management for different events, such as context relocation when mobile.

Beyond the primary task of providing wireless connectivity, the radio-access infrastructure will also be capable of delivering other services. This is already happening today, in part, with the introduction of location-based services as a complement to GPS. The combination of high-frequency band networks and dense deployments will make it possible to dramatically enhance the accuracy down to sub-meter level. Other service examples include time synchronization, time-sensitive networking, the collection of complementary information about local weather conditions and the creation of radar-like scans of the environment.

TREND #5: SECURITY ASSURANCE

The need for protection and assurance (or even compliance) is growing rapidly as business and society increasingly rely on universal connectivity and compute. Today, there is intense activity to explore the potential of AI and ML to protect systems and networks. There is large-scale adoption of these technologies in areas such as network threat detection and threat intelligence extraction, while other areas such as continuous authentication appear less mature. While AI technologies can provide a wide range of benefits, it is important to note that they can also be used by adversaries to find avenues of attack that specifically target ML systems.

In these autonomous networks, security assurance procedures play the important role of verifying security properties of the network platform. One challenge lies in the network architectures, based on cloud technologies and virtualization, which are introducing requirements for continuous compliance verification in a dynamic environment. At the same time, security assurance needs to be rooted in the evidence collected in the network slices supporting different industries. AI and ML technologies will bring automation of assurance and compliance verification to the network platform.

In the world of cloud computing, enclave and confidential computing hardware solutions that provide a root of trust are currently being packaged in pre-commercial cloud solutions. These technologies have the potential to become prevalent when addressing security concerns for processing in the cloud. Conceptually similar trusted computing technologies are also moving into IoT devices.

The trend toward encryption everywhere continues with reports of up to 90 percent usage of HTTPS. A substantially different protocol stack on the internet is expected in a few years, with QUIC and DoH as the dominant protocols, protected by newly standardized post-quantum algorithms. Since current security protocols are not suited for constrained IoT nodes and devices, the industry is working to standardize new lightweight application layer protocols.

At the same time, remotely managed eUICC (embedded Universal Integrated Circuit Card) based SIM identities in IoT devices are increasingly being deployed for network access. Modern SIMs based on the eUICC, and later the even more cost-effective iUICC (integrated Universal Integrated Circuit Card), will form the trust anchors for secure identities and network access in five to seven years.

Mission-critical use cases and regulatory demands, as well as cloud and edge computing, are the driving forces behind the trust and assurance technologies that are being developed and becoming integral parts of the network platform.

TREND #6: ZERO-TOUCH NETWORKS

A zero-touch network is capable of self-management and is controlled by business intents. Data-driven control logic makes it possible to design the system without the need for human configuration, as well as to provide a higher degree of information granularity. Applying AI technologies will enable zero-touch automation of network life-cycle management, including optimizing system performance, predicting upcoming faults and enabling preventive actions.

The performance of a data-driven zero-touch function can increase by utilizing the wider network data from many local clients, but this needs to be balanced against the cost and time associated with transferring large volumes of data.

One approach is to design distributed ML solutions, such as federated learning, which makes it possible to generate a network-wide global ML model. Training is done on local clients, and the need to transfer data is limited to model updates, instead of raw data.

With reinforcement learning, it is possible to design a solution that responds to unforeseen environments, which can be used to automate or optimize a specific process. A reinforcement learning agent learns how to act optimally given the system state information and reward function, focusing on finding a balance between exploration of uncharted territory and exploitation of current knowledge. The requirements on reliability and safety will, however, set limits on the applicability.

Robots are used to interface with the network infrastructure, collaborate with humans and utilize AI to perform physical inspections, determine fault causes, predict future faults and plan maintenance work. Computer-vision techniques enable, for example, automated cell tower inspection, while machine reasoning is used to plan and execute drone flight. Techniques to generalize and transfer lessons learned can be used to increase performance from one tower inspection to another. These AI-based robot systems will collaborate with humans, thereby increasing their safety and efficiency.

An intent-based approach similar to the one referenced in trend 3 (distributed compute and storage) allows human users to interact with the AI system that is part of zero-touch applications. Domain modeling, knowledge representation and reasoning (together with ML) are used to create a cognitive layer for humans to interact with the system using high-level intents. The system is capable of evaluating and executing strategies in line with an intent, based on lower-level key performance indicator (KPI) predictions. By complementing ML with machine reasoning, the system can be designed to express why certain decisions were taken and is a way to implement explainable AI.

Trustworthy ML models that fulfill zero-touch aspects need to be built in line with the need for privacy and legislative rules for how data can be exposed or moved. New specialized hardware for accelerating ML training and inference will improve performance and reduce energy consumption in a well-designed zero-touch network platform. Recent progress in AI has shown new promising possibilities to design for zero touch. Many challenges need to be overcome, however, and the value and efficiency of traditionally designed control logic should not be underestimated.

CTO TECHNOLOGY TRENDS 2019 ✱

CONCLUSION

No other technology in the world today can provide pervasive global coverage on a scale comparable to that of the network platform, and it is my firm belief that it is ideally suited to serve as the innovation platform for both current and future applications. The technology evolution characterized by this year's trends points toward the future definition of 6G.

Much more cost-efficient than specialized or dedicated network solutions, the network platform is clearly the most affordable solution to address society's needs across the spectrum from human-to-human to human-to-thing and thing-to-thing communication. One of its major advantages is that it is available through an open marketplace that is accessible to anyone, anywhere, at any time.

The multipurpose network is rapidly emerging as a secure, robust and reliable platform where applications, processes and other technologies can be developed, deployed and managed. The Internet of Skills and cyber-physical systems – trends 1 and 2 – are important examples of use cases that it needs to support.

A key characteristic of the future network platform will be its ability to instantaneously meet any application need, anytime. Four technology areas – trends 3-6 – are playing critical roles in its ongoing evolution: distributed compute and storage, ubiquitous radio access, security assurance and zero-touch networks.

Self-driving vehicles, intelligent manufacturing robots and real-time drone control are just a few examples of the myriad of ways in which the multipurpose network is enabling the automation of the physical world and, ultimately, the creation of a sustainable society of the future.

◆ As Group CTO, Erik Ekudden is responsible for setting the direction of technology leadership for the Ericsson Group. His experience of working with technology leadership globally influences the strategic decisions and investments in, for example, mobility, distributed cloud, artificial intelligence and the Internet of Things. This builds on his decades-long career in technology strategies and industry activities. Ekudden joined Ericsson in 1993 and has held various management positions in the company, including Head of Technology Strategy, Chief Technology Officer Americas in Santa Clara (USA), and Head of Standardization and Industry. He is also a member of the Royal Swedish Academy of Engineering Sciences and the publisher of Ericsson Technology Review.

ERIK EKUDDEN

SENIOR VICE PRESIDENT, CHIEF TECHNOLOGY OFFICER AND HEAD OF GROUP FUNCTION TECHNOLOGY