Expediting AI for Operations in the Networks of Future

Sponsored by ZTE

In the past decade, Artificial Intelligence has morphed beyond academic researches and taken off in the industry. It has already garnered a track record of solving hard problems that were not possible before in various sectors including healthcare, ecommerce, financial services, etc. From businesses' point of view, the value proposition of this technology is also becoming glaringly clear – that AI based automation shall vastly outperform manual labor as well as rule based automation both in terms of productivity and precision, wherever data is abundant for model training.

IDC predicted that by 2020, organizations able to analyze all relevant data and deliver actionable information – effectively meaning those able to operationalize big data, will achieve an extra $430 billion in productivity benefits over their less analytically oriented peers. With now AI taking up the center stage of big data management and utilization, more organizations are in search of effective means to operationalize AI.

For telecom service providers, such need is more pressing as operational data and network events are already of humongous volume. The growing complexity of network topology and vendor ecosystems has further increased the demand for automated configuration, management and maintenance. With greater bandwidth to come in the much anticipated 5G era, more data traffic will be pumped through the networks of future within microseconds, exposing the operational vulnerabilities of the Telecom operators.

In this paper, we dive deeper into the role of AI automation can play in the telecom sector, and why AI for Operations represents a great opportunity for telcos to establish leader position in the future market. As an example of the evolving efforts towards AIOps for Networks, the solution architecture and highlights of the AI platform from ZTE: uSmartInsight are provided, together with concrete use cases where the platform has accelerated AI operationalization for telco operators with proven results.

Technology Perspective - Verticalized AI Process Automation to Prevail

The availability and processability of big data have accounted for recent growth spurt of cognitive and AI technologies. The time is now to focus on its adoption among businesses either to create value-add products, or to streamline internal business operations and uplift productivity.

As of now, all known AI applications are in the category of narrow AI, whereby the machine intelligence at work can only perform one, or a few specific tasks. Before the narrow AIs can combine forces and grow their capacities to cover more tasks, a key prospect of the technology adoption is the verticalization of narrow AI – to transform lower level AI solution components into fully integrated solution that works for end users facing very specific task scenarios.

Lower Level AI Solution Components are Not an Industry Focus

IDC sees the family of AI technologies to wrap around lower level solution components such as Speech Recognition, Computer Vision and Text Analytics, etc. enabled either by Machine Learning or Deep Learning. The open sourcing of many ML and DL libraries and frameworks like Scikit-Learn, TensorFlow, and Caffe have effectively reduced the technological barriers to reuse, adapt or package such solution components. At the same time, open sourcing renders commercial competition on these lower level capabilities pointless. Unlike academic research, industry organizations need a solution recipe that cannot be easily replicated or improved upon.

Competition Converges on Computation Power, Verticalized Data and Task Scenario

With AI algorithms and lower level solution components being largely taken out of the success formula for industry players, the three remaining differentiating factors for AI technologies to deliver real business value are data, task scenario and computation power. Among the three factors, data and task scenario are both specific to different verticals and LoBs. As a result, the next phase of AI development in the business world has to go vertical. It shall require technology buyers to invest in task level software solutions that consume from scenario specific data. IDC foresees

that by 2020, 10% of enterprise applications spending will be for new task level applications that incorporate software, data, and algorithms.

AI Based Business Process Automation to Prevail

With the right data, task scenario and computation power in place, AI technologies can be truly operationalized. In a recent series of IDC Cognitive/AI surveys conducted in the United States, Western Europe and Asia Pacific, 60% of industry respondents selected Process Automation as the primary expected benefits for the adoption of AI solutions, which is about 20 percentage points higher than other options including uncovering of new insights, improving of interactions and decision consistencies. Based on the aggregated survey results, it is estimated that 40% of worldwide organizations will be using AI based process automation solutions within the next 5 years' time. And the business processes that will be most impacted are customer support & service, management of security, sales and marketing.

40% of worldwide organizations will use AI based process automation solutions within the next 5 years' time
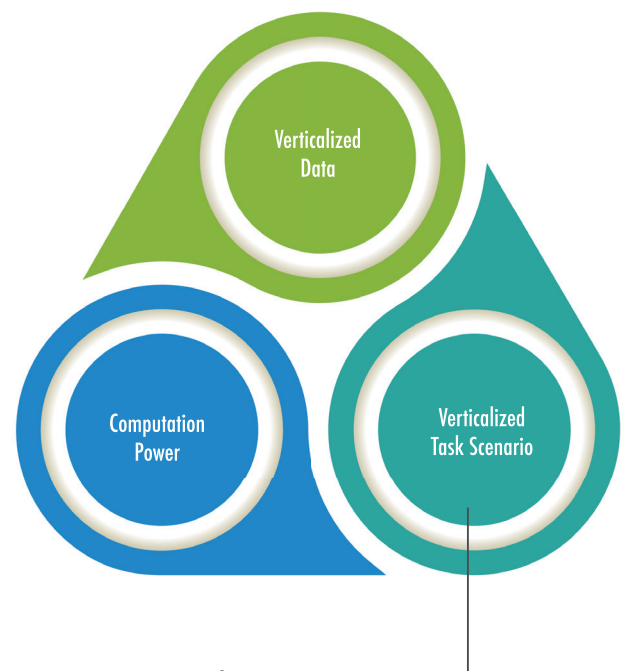
Verticalized task scenario and data will be increasingly important for AI to deliver business value.

Industry Perspective – Telco to Establish a New Identity as the Vanguard of Technology Adoption

The telecom companies provide connectivity which is the backbone of most digital disruptive innovations. Itself however is a sector where products and services have become largely commoditized. Even though many of their customers, individual consumers and enterprise accounts alike, remain to be locked in for the mid to long term, telecom service providers need to search and secure new growth areas and business models to make up for the lost ground to the Internet OTT (over the top) content providers, and more importantly to stay prepared for future changes.

The Need to Revamp & Reestablish Identity

Despite being able to offer a myriad of services including data center, OTT, communication services, cloud services, managed business and IT services, telecom service providers often have an obscured identity among the digital natives, where not many of their value-add capabilities can stand out in competition. For example, less people will relate telecom companies as the go to vendor for cloud services, or managed security services, or OTT content. Even for their traditional business strongholds, customer centricity could be lagging to secure continued loyalty of the Millennials and the Centennials.

It is time for telecom companies to revamp and re-establish themselves as a leader in the technological world. And the means to achieve that is continued digitalization to modernize legacy network operations and upgrade service offerings, so that the services provided will stay relevant and distinctive for the targeted customers. IDC sees the telecom sector to be an important aspect of digital transformation and the transition to 3rd platform accelerators in 2018. According IDC Futurescapes: Worldwide Telecommunications 2018 Predictions, many digital technologies such as AI, BDA (big data and analytics),
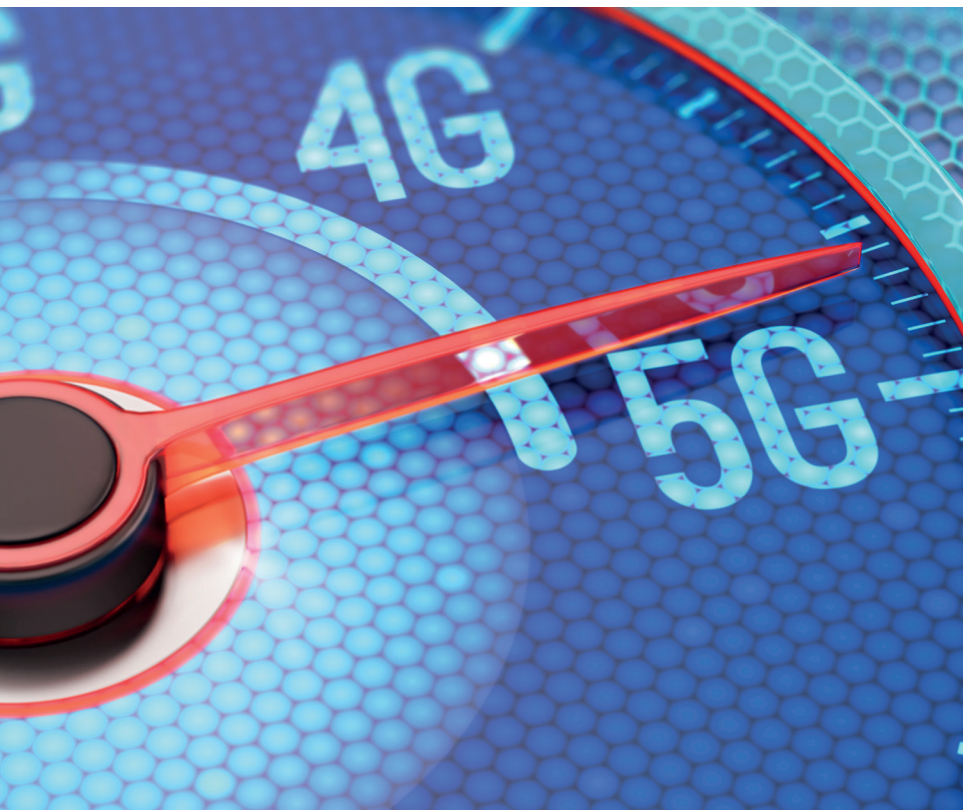
SDN (software defined networking), NFV (network function virtualization), and API are high on the list for telecom companies to leverage.

The Ticking Clock of 5G on the Horizon

Full deployment of 5G network is not yet a reality. However, the clock is already ticking for network operators and everyone in the value chain to prepare for the many changes it will bring along. With greater bandwidth, lower latency and wider connectivity, more data will be pumped through the networks in a faster speed. UHD (ultra-high definition) streaming, Autonomous driving, AR/VR (augmented reality and virtual reality) for workplaces, large scale industrial automation – what are not possible today will become possible tomorrow.

New business models will surface, new applications will flourish, Telcos must not stay content as the connectivity provider, as what they had reacted when faced by the boom of OTT mobile applications. An anchoring role in the forming ecosystem to get revenue share from solution and application consumptions will be critical. Nonetheless, such a role will not come by easily, operation digitalization and transformation have to be a key investment area, to pave the way for full platform capabilities that gel enterprises and businesses together.

Even for telcos not venturing into new business models, they need to be aware, if not already, that 5G will expose their vulnerability in handling the flood of data, or the surge of workload. Increased bandwidth is nothing if the network is prone to congestion or down-time. The telco systems in days of 5G must have the ability to intelligently figure out how to move data traffic from one place to another, and pinpoint issues that can arise from anywhere of the many-layered technology stacks.

AI Based Automation to Probe Through Multiple Technology Stacks

Among the many digital acceleration technologies, AI based Automation is the key enabler to modernize network operations, as it is able to accommodate the many complications introduced by different waves of digitalization:

Customer Centricity as a key driver of Digital Transformation brings along increased complexity in vendor ecosystems, that their services have to interact with all kinds of hardware and mediums possible. This complex ecosystem often proves to be problematic for telco's support team, where multivendor support requires higher level capacity to understand the interactions between technology layers.

....

IoT and the proliferation of edge devices continue to create variations in network topology, making it increasingly hard to conduct fault diagnosis and implement rule based policies.

....

SDN and NFV initiatives decouple network services from physical units. As a result, multiple instances will be running on a single system, it will increase both the cost of failure and difficulty to recover all services affected.

....

With the massive spike of data traffic to come in the much anticipated 5G era, more events will be traversing different "edges" of the networks at the microsecond level. Customer expectation on service level at the same time shall only become higher, requiring more detection and correction of issues to be automated.

....

The combination of the above factors pose aggregated challenges for service providers to manage network operation and prevent service degradation. The scope of problem has become too complex for rule based means to render effective help. AI capabilities have to be incorporated to provide intelligent detection automation. It shall serve as the single lens that probes through multiple technology stacks and vendor systems, at the same time managing the growth in scale of future networks.

Operating the Future Networks

IDC believes the ability to harness AI will become imperative for operators of future 5G networks. The application prospects are discussed below covering three functional areas:

Network Operation, Maintenance and Security will be the immediate areas for AI based Automation to create value. By ingesting customer and network event data that well represents the fast-evolving network status, AI based automation has the highest efficiency (e.g. time required) to scrutinize logs, extract patterns, learn and adapt continuously through model retraining and refinement, and to respond promptly to events in different types of networks – traditional and software-defined.

....

Network Configuration, Orchestration and Optimization will see the most benefit from AI in the 5G context. Theoretically AI can take over service planning, deployment and management including Global Service and Resource Orchestration, Virtual Network Slicing, Cross Layer Optimization and Customized Mobility Load Balancing, to optimize resource allocation and utilization. The capability of AI shall grow along with the continued virtualization of networks to achieve new level of orchestration - autonomous, predictive, and end-to-end.

....
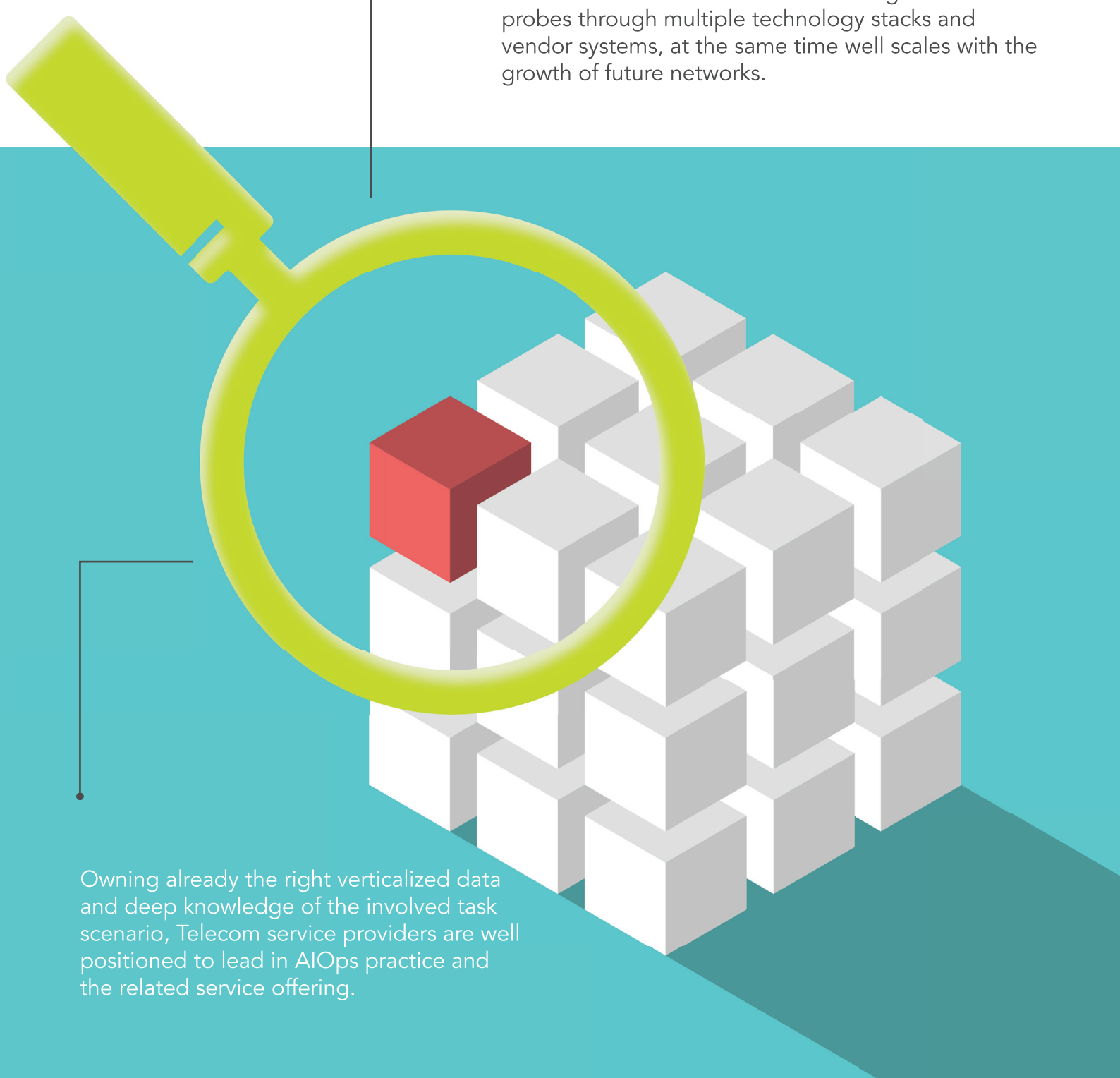
Business Operation and Support will be vastly upgraded with AI embedded in customer engagement interfaces, knowledge bases, job dispatch systems, customer churn predictors, and service life cycle analytics to ensure customer centricity and uplift productivity. AI will federate diverse data sources, create tailored customer experience and curated data services, enabling Telcos to enter adjacent verticals such as lifestyle services.

....

Telecom service providers have the opportunity to become the vanguard of AI for IT Operations, or AIOps. While DevOps (Developers for IT Operations) has recruited many adherents in the past few years, it will not remain as relevant in the networks of future, where data, events, and system changes will outpace even the experts. Increasingly, we have to admit the expanding big data world is not the best arena for human brains to excel. Owning already the right verticalized data and deep knowledge of the involved task scenario, Telecom service providers are well positioned to lead in AIOps practice and the related service offering. This presents a potential opportunity to establish leadership position in the future market.
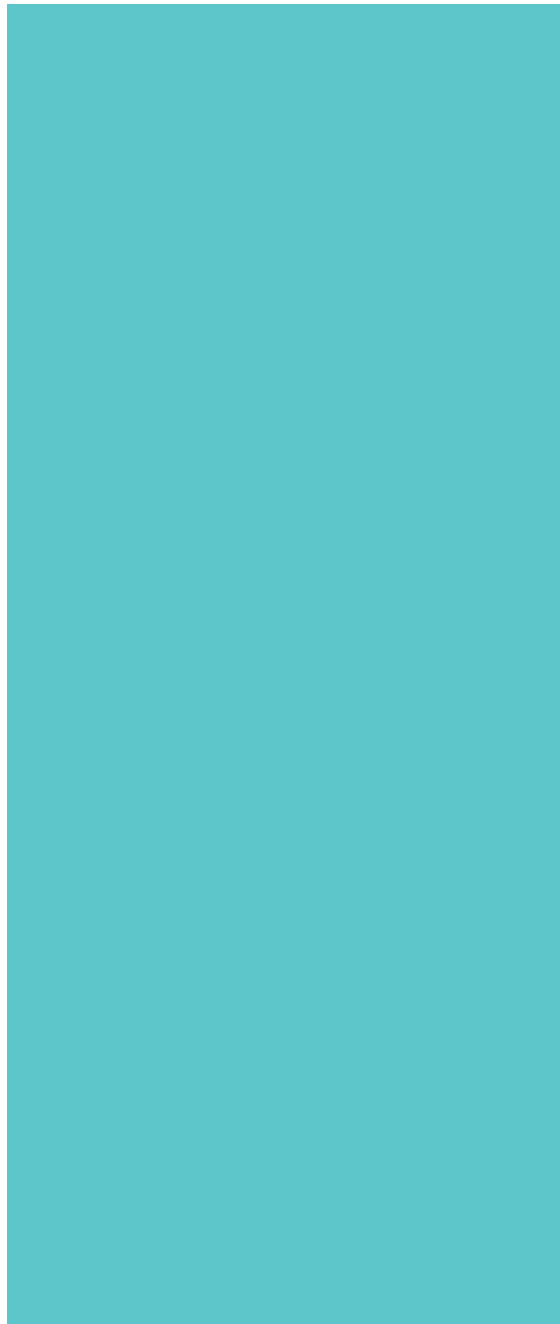
AI Based Automation serves as the single lens that probes through multiple technology stacks and vendor systems, at the same time well scales with the growth of future networks.

Owning already the right verticalized data and deep knowledge of the involved task scenario, Telecom service providers are well positioned to lead in AIOps practice and the related service offering.

Solution Overview – uSmartInsight Platform from ZTE

ZTE is a world leading solution provider for networks, terminals and telecommunications. The company has been investing heavily on 5G, pushing the trials of 5G networks in many locations around the world. Foresights in a future of hyper-connectivity have also led the company to orchestrate many years' research and development on Big Data and AI into one proprietary platform – uSmartInsight, to help expedite AI operationalization for its customers across different industries.
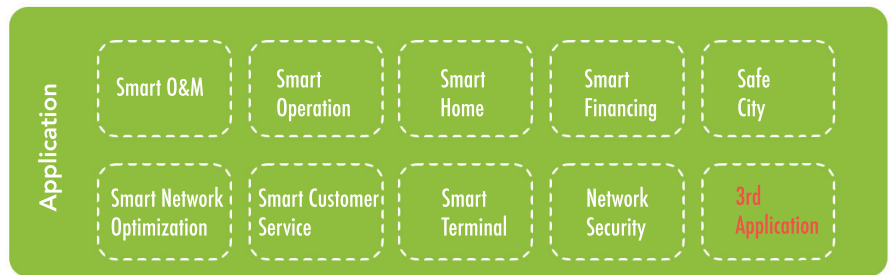
Architecture Brief

The uSmartInsight Platform of ZTE consists of three layers:

The Infrastructure Layer that supports multitype deployment including bare metal, cloud and container;

The Storage, Computing & Service Layer that encapsulates end-to-end data flow for AI model building, ranging from storage, preprocessing, tagging, feature engineering, to training, evaluation and deployment;

And lastly the Application Layer which caters to verticalized scenarios including AI for Operations, Smart Home, Smart Financing, Smart City, and AI for Cybersecurity.

Figure: uSmartInsight

Figure: uSmartInsight
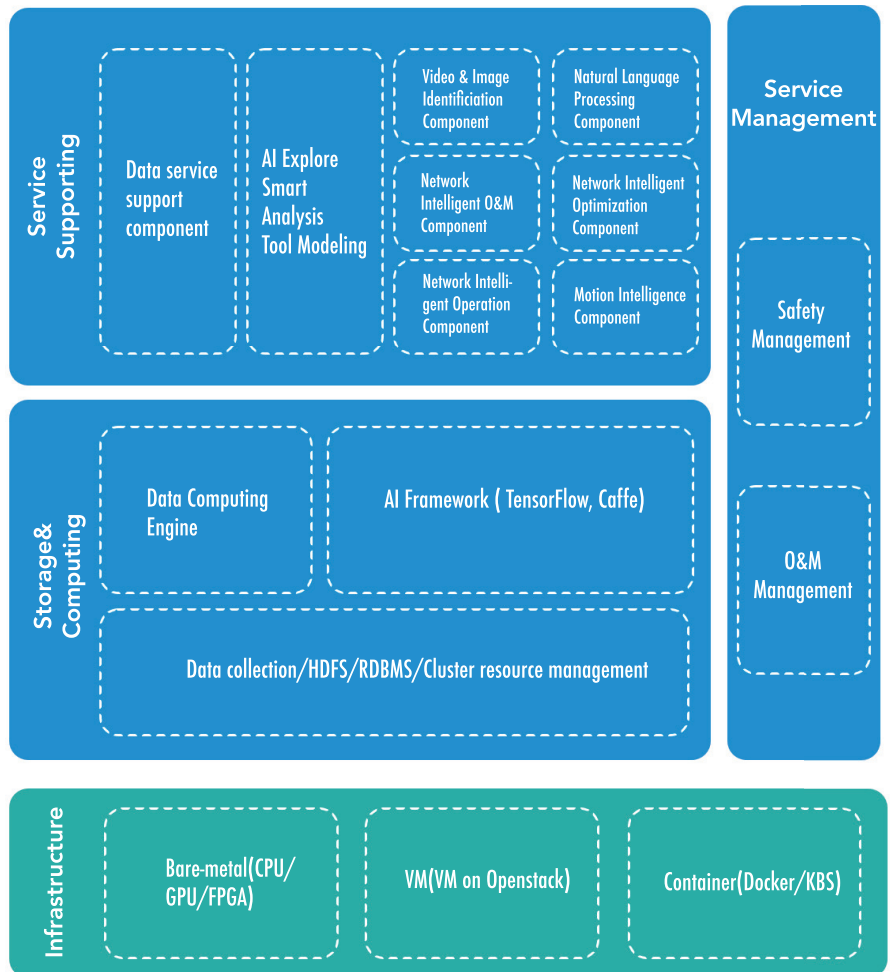
The platform aims to provide agile model building, fast deployment and flexibility to accommodate verticalized task scenarios, as well as third-party application development.

Capability Highlight

The uSmartInsight Platform is well developed to compete in the computation power, data and task scenario fronts.

On *computation power required for model training*, uSmartInsight Platform supports heterogeneous setup of CPU, GPU and FPGA, containerized HPC (high performance computing) clusters with 100G RDMA interconnectivity. The Platform is equipped with a Distributed Training Engine, which is the result of ZTE proprietary research incorporating TensorFlow and Caffe frameworks, to enable super scale model training. The parallel efficiency of the Distributed Training Engine for synthesized data has reached above 92% for 16 GPUs.

On *computation power required for model inferencing*, uSmartInsight provides a suite of solutions including GPU acceleration for central data centers; FGPA acceleration for edge data centers that runs on low energy; and embedded models for edge devices. In January 2017, its FGPA acceleration solution that runs on OpenCL framework and Intel Arria 10 achieved a new benchmark performance for CNN (convolutional neural networks) inferencing – 1,000 faces recognized per second with high accuracy. In less a year, ZTE pushed the bar further to recognize 3,500 faces per second.
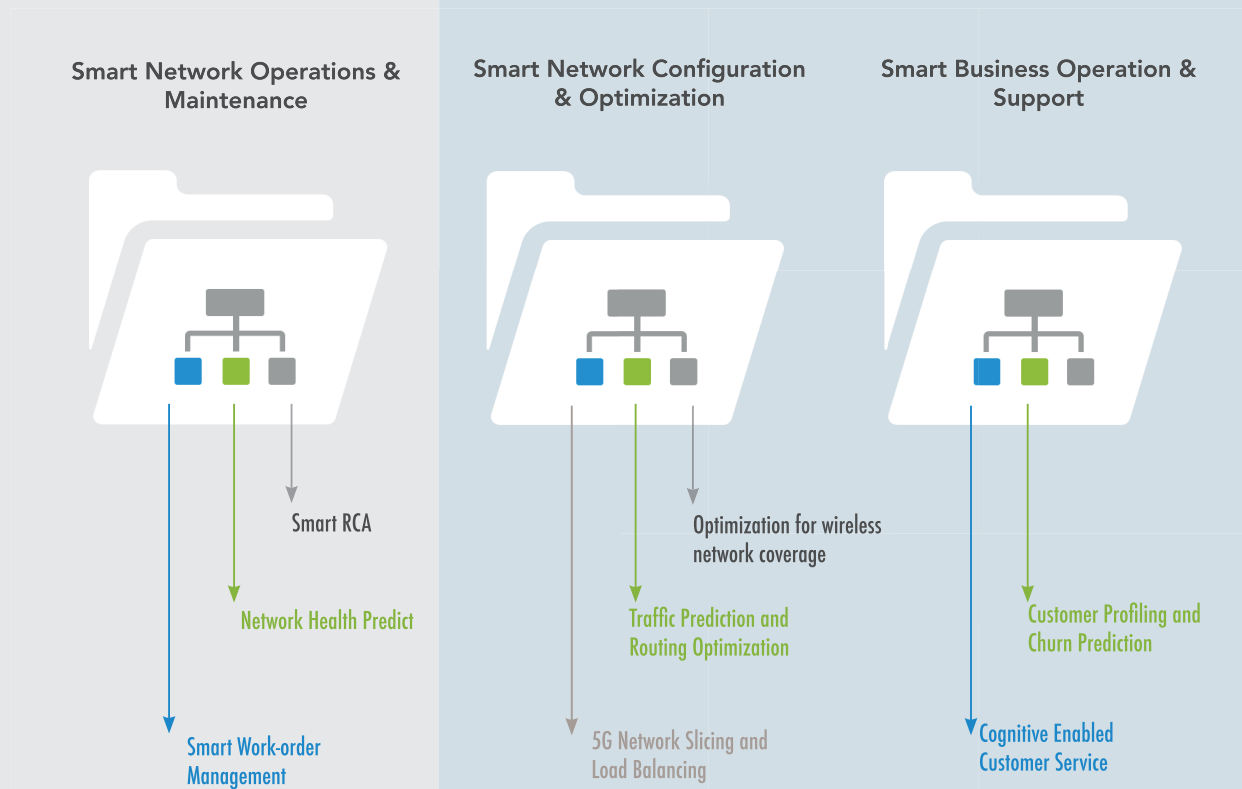
On *the data and task scenario front*, uSmartInsight has a clear focus on telecom vertical, including Smart Network Operations & Maintenance, Smart Network Configuration & Optimization, and Smart Business Operation & Support (Figure below). Smart RCA (root cause analysis) for example is part of Smart Network Operations & Maintenance, it mines network operation data, events, knowledge base and issue logs to uncover associations across multi-domain, multi-vendor, multi-service layer, and locate root issues with greater efficiency and precision. It is a breakthrough of traditional rule based system maintenance that has become overly slow and costly to deal with the intrinsic and increased complexity of communication networks.
....

Figure: uSmartInsight-Suite of Applications for Telecom Operators

Smart Network Operations & Maintenance

Smart RCA

Network Health Predict

Smart Work-order Management

Smart Network Configuration & Optimization

Optimization for wireless network coverage

Traffic Prediction and Routing Optimization

5G Network Slicing and Load Balancing

Smart Business Operation & Support

Customer Profiling and Churn Prediction

Cognitive Enabled Customer Service

• SUCCESS STORIES •

Smart RCA *has been adopted by a Telecom Operator in Guangdong, China, for a number of projects covering IP bearer networks, wireless networks and NFV core networks. Millions of historical system warning records spanning a 6-month timeframe were used for model training, over 40 new association models for issue identification were uncovered and validated in operations. The training and validation were completed within seven man-days, a more than 300% improvement of productivity. The association models were able to locate issue triggers traversing virtual networks and physical networks, wireless networks and core networks, and across service and business layers.*

Network Health Prediction and Smart Work-order Management *were co-developed with and implemented at China Telecom Sichuan. Serving over 16million mobile users, the company faced increased pressure in its wireless network operation, configuration and support; largely due to the scattered nature of system data, climbing cost in staff training and retainment, and nebulous perception of network performance confounded by multiple vendors and problematic end devices. The application ingests system logs, alerts, measurement report, call detail trace, and customer complaint tickets to identify issue-prone service zones proactively and automatically. With the issue prone zones located based on customers' feedback, follow-up workflows can be standardized and automated including intensive network traffic monitoring, real time coverage assessment, parameter adjustment suggestion and job dispatch routing optimization. Since the solution implementation, the company had already seen an annual saving of operation cost about US$5million, a 50% reduction of customer complaints, with also 40% less operational staff to make all these happen.*

Ongoing Research

Understanding that the operation of upcoming 5G networks will be heavily dependent on AI capabilities, to cover infrastructure management, network and business function management, and operation support, ZTE has made AI in the context of 5G a strategic investment priority.

Global 5G Partners - ZTE has established 5G strategic partnerships with major telecom operators including China Mobile, China Telecom, China Unicom, Softbank (Japan), Telefonica (Spain), Telenet (Belgium), Orange (France), and SK Telecom (South Korea). In October 2017, ZTE announced its plan to deploy the very first 5G pre-commercial network of Europe, together with two leading telecom operators of Italy – Windtre and Open Fiber. An innovation center is to be set up in L'Aquila, Italy to accelerate the construction of an end-to-end 5G ecosystem.

High Value Scenarios - Targeting a fully autonomous network with coverage and capacity optimized, ZTE has identified following high value application scenarios for continuous research and development:

AI Aided 5G Massive MIMO Beamforming – MIMO setup of large scale antenna systems can amplify the capacity of wireless networks without requiring extra spectrum. Beamforming technology is used to achieve spectrum selectivity, and reduce network performance bottleneck in heavily populated areas. With AI agents trained on traffic distribution and inter-cell interference, massive MIMO systems can change the antenna patterns adaptively. As have been validated in pre-5G LTE systems, users of SmartInsight platform can make better beamforming inferences, and receive detailed suggestions of antenna alignment parameters for either manual or automated execution.

Smart Mobility Load Balancing – Radio Frequency Fingerprint can uniquely identify physical transmitters. It will stay largely as the basis to achieve better load balancing in the Future networks. uSmartInsight plans to address Smart MLB from two aspects:

First is to select the target users and load balance cells according to the granularity of RF fingerprint grid. This approach is more accurate, elaborate and efficient when compared with load balancing based on the granularity of cells. It will also come with faster speed of load balancing and enhanced traffic steering effect. It shall require a RF fingerprint database, which is to be constructed based on historical data of signal quality and other related KPIs.

Alternatively, AI can be used to predict the future load of cells. When the load predictions are available for both the source cells and target cells, they can be used to prevent the ping-pong effect.

5G Intelligent Virtual Network Slicing – Considering the multitude of innovative services the future 5G network is going to support, it is critical to dynamically partition the physical network into different virtual networks to optimize resource utilization by different performance requirements. For example, video streaming is sensitive to delays while metering data from IoT devices has strict requirement on packet loss. To meet the diverse requirements of various services in the future network, ZTE is building up AI-assisted slicing capabilities so that it can be scheduled and deployed in a way that is agile, automatic, context-aware, and self-healing.

—— *Future Outlook – Network of Pervasive Intelligence*

Networks underpins close to all modern-day business operations, and their importance shall only elevate in the coming years. 5G technology will further release the power of computing on devices and in the cloud, bringing pervasive intelligence closer to reality.

For telecom service providers, the support and management of the future network shall face aggregated challenges. Reduction of manual operation is one of the top priorities for any telcos making ways for 5G, and most of the reduction has to leverage Artificial Intelligence. Not only that AI as a software agent works well with data of any size and any type, but it can also extract patterns from data without being explicitly programmed, which is a natural fit for the heterogenous and highly distributed environment of 5G networks.

While few vendors at current stage can provide a comprehensive AIOps solution, what ZTE has achieved with uSmartInsight is well on track to evolve into an AIOps platform for Networks, to enhance and gradually automate a broad range of network operation tasks including application performance monitoring, health checks in software-defined networks, root cause analysis for performance issues, capacity and coverage enhancement, service routing optimization, load balancing, and more.

IDC sees that AIOps for Networks will be the immediate next stage of the evolution of network management to ensure responsiveness, timeliness, proactiveness and quality of services in the context of 5G.

Essential Guidance

Network Operations are in for a major change over the next few years. AIOps will gradually come in to join force with DevOps to deal with the climbing need of automation. In the telecom sector, this is driven by the frustration with traditional network management techniques that have become too slow and costly to deal with the many complexities introduced by digitalization initiatives, the great momentum of AI technologies becoming verticalized, and the pressing need to deliver the ultra-high speed promised by 5G with quality of experience.

IDC sees AI for Operations represents an untapped opportunity for telecom service providers to establish leadership in the networks of future, and recommends the following to the aspiring companies:

Build up a foundation for Big Data - AIOps or AI for Operations in generation has to sit on top of a big data platform to ingest and aggregate operational data for historical analysis, pattern recognition and model building. It is important to have big data capabilities, including federation, preprocessing and governance, established or covered in the proposed AIOps solutions. For example, while graph analytics is nothing new, the ability to apply AI and machine learning on top of billions of edges in graphs allows for telcos to make faster, better and more accurate predictions in areas such as network operations visualization and fraud detection.

Clarify the core task scenarios – Verticalized task scenarios and relevant data are the two prerequisites for AI solutions to deliver business value. Businesses are to come up with a workable scope to include a few core operational tasks where training data is amenable, and performance improvement is measurable. For example, in subscriber care, AI can be applied to analyze subscriber experience, service quality, device, activity and usage behavior in order to allow telcos to deliver personalized experiences in customer care management to reduce churn and improve brand loyalty.

Upskill the operation staff – AIOps is not in a position to replace operation and maintenance engineers. Its success is largely dependent on the domain expertise and judgement of operations staff, especially experts who have thorough understanding of the system configurations, interactions and upgrades. The implementation of AIOps should aim to deal with the explosive growth and diversification of network events, and grant operational staff more time to become expert in multi-domains. For example, the citizen data scientist is a concept that telcos need to start embracing and adopt across business units; AI tools can be used to empower domain experts and users to make better and faster data-driven decisions.

Leverage platforms with proven AI capabilities and computation power – Businesses are advised against investment on lower level solution component of AI, instead to go for platforms where AI capabilities based on proven open source frameworks have been verticalized, with also the necessary computation power to support large scale model training and real-time inferencing. For example, pre-trained models optimized for specific use cases can be quickly deployed in production; and further enhanced over time as new data is received without having to start from scratch.

Collaborate to innovate for the longer run – Collaborative innovation, or co-innovation involving multiple entities, has become the new norm of innovation for enterprises. Comparing with incremental innovations within an organization, co-innovation has a greater success rate and shorter runway leading to productization. For telecom operators, co-innovation with the technology community shall expedite AI adoptions in the context of 5G. They should also bear in mind that machine intelligence needs time to mature. Expectations on AI for Operations from the lines of businesses are to be carefully managed, by for example communicating project milestones regularly such as in-scoped tasks, data readiness, model performance and retraining schedules.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.