ZTE Intelligent Data Mid-End Platform for Carrier

White Paper

Contents

Telecom Industry Development Opportunities and Challenges

Digital Economy Rise

With the development of communication and digital technologies, networking and digitalization bring more wonderful and infinite possibilities to human beings, and push the world into the fully-connected informatization. Especially from 2020, the 5G infrastructure industry is empowering the trend. The "digital economy" extends from consumption to education, work, and production, such as mobile social service and diversified entertainment, short video platforms and live broadcast products, home health management, online diagnosis and treatment, distance learning, and telecommuting. Many analysts predicted that the industry would complete digital transformation in two years or more, but it has entered the stage of digital transformation within 10 weeks in 2020, endowing with intelligence. IDC predicts that with the maturity of 5G application, the digital industry will continue to grow rapidly. In 2022, 65% of global GDP will come from the digital field.

Opportunities of Digital Economy for Telecom Operators

5G is fueling digitalization and empowering thousands of industries from below aspects:

- From single-point and single-service to shared and diversified service operation, establish an intelligent and data-driven operation system.
- From extensive marketing to precise marketing, break down data silos to meet customer requirements in a quick response.
- From isolated channel to collaborative channel, develop multi-channel convergence.
- From passive response to active prevention, cover full-lifecycle customer services.
- From single product to cross-product synergy.

Challenges of Digital Economy for Telecom Operators

Telecom operators are facing the challenges of the digital economy, including:

- Accelerated channelization trend. A large number of Internet applications have a great impact on traditional telecom services, and telecom operators cannot obtain reasonable revenues that match their investment.

- IT architectures to be improved. The legacy isolated systems are hampering and slowing down the introduction of new services, and increasing complexity to data access and scalable deployment.

Digital Transformation is Becoming a Reasonable Choice

Facing the opportunities and challenges of the digital economy, digital transformation is an reasonable choice for telecom operators. Digital transformation is using the core technologies of new-generation big data and AI to effectively solving the problems of network management complexity and benefits in the future. The AI technology provides predictive capabilities, granular and multi-dimensional analysis, and process automation capabilities to optimize network operations and expansion. Big data and AI technologies are key elements to the development of operator network value.

Telecom operators are ongoing digital transformation, and their communications networks will serve the intelligent society in an all-round manner. Network O&M will evolve towards precise value-based operation, providing digital, intelligent, and integrated innovative new services.

- Empower operation capabilities to all industry and ecosystem partners.
- Work together to innovate in ecosystem development and business models.

VMAX Intelligent Data Mid-End Platform

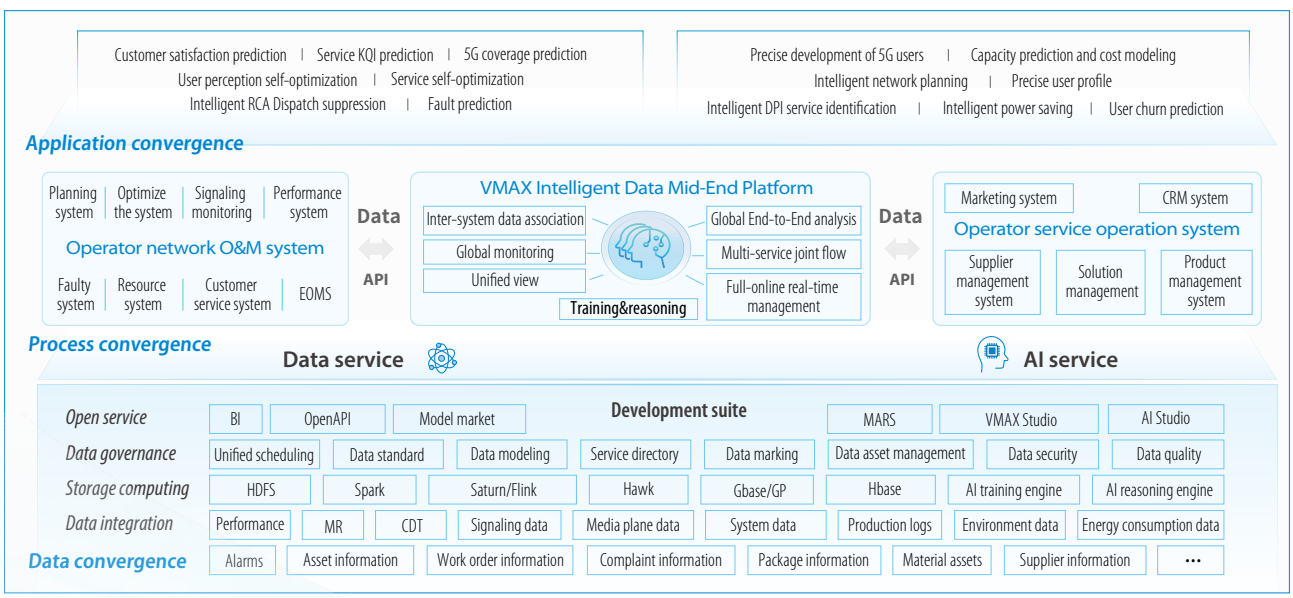
Intelligent Data Mid-End Platform is the Core Engine of Digital Transformation for Telecom Operators

Digital transformation is driven by the development of big data, artificial intelligence (AI). With data driving and AI empowering as the two engines, Intelligent Data Mid-End Platform breaks down data silos, continuously improves data quality, makes full use of data service value with data-AI convergence, empowering the evolution of IT architecture.

VMAX Platform Architecture

Digital transformation of telecom operators will be a long-term evolution, requiring gradual advancement based on the process of 5G network construction and cloud-network synergy with AI, big data, and other technologies.

To promote the digital transformation, ZTE proposed the AIVO (AI Insight and Value Operation) digital operation solution. Based on AI and big data technology, the solution integrates three dimensions of data resources, service process and application ecosystem to support the digital operation of the operator's network among OSS, BSS, and other sources.



The core concepts of the AIVO solution are presented in three aspects: data convergence, process convergence, and application convergence. In terms of data convergence, this solution can break down the data silos to support accurate, all-round data sources through VMAX(Value Multi-Analysis eXpert) Intelligent Data Mid-End Platform. As for process convergence, by building an integrated tool system, the operator's O&M and operation systems can be connected to realize the automation of the service process and the improvement of the O&M efficiency. The last one is application convergence, where AI is enabled to achieve pipeline visualization and application innovation and increase revenue.

VMAX Intelligent Data Mid-end Platform supports underlying data collection and apply data for applications, through connecting with the operators' system, to achieve internal asset operations, open external capabilities.

Function Description

VMAX Intelligent Data Mid-End Platform builds a fast, simple, AI-based integrated data analysis system for the telecom operator. VMAX consists of the following eight aspects:

01 Data integration

Provides high-efficiency collection and loading of heterogeneous data sources, and the unified data lakes.

02 Storage & Computing

Provides systematic and distributed offline storage computing, stream real-time computing, and machine learning capabilities.

03 Data governance

Provides data specification modeling, data audit, metadata management, asset management and big data lifecycle management capabilities to build a unified data scheme.

04 Development suite

Provides self-service analysis, self-service mining and self-service report capabilities, and supports the capability of visualized code development, so that the data analysis engineer can quickly complete BI and AI analysis by drag-and-drop.

05 Full-process management and control

Unified task scheduling is provided to meet the requirements of batch data and real-time stream services. Tenant management and data security mechanisms are provided to ensure data execution in a secure environment.

06 Intelligent modeling

Supports one-stop AI development, visualized drag-and-drop modeling, and professional code modeling. The training service engine provides automatic concurrent training capabilities, while the inference service engine provides AI model optimization and cloud-side collaborative deployment capabilities.

07 Data service

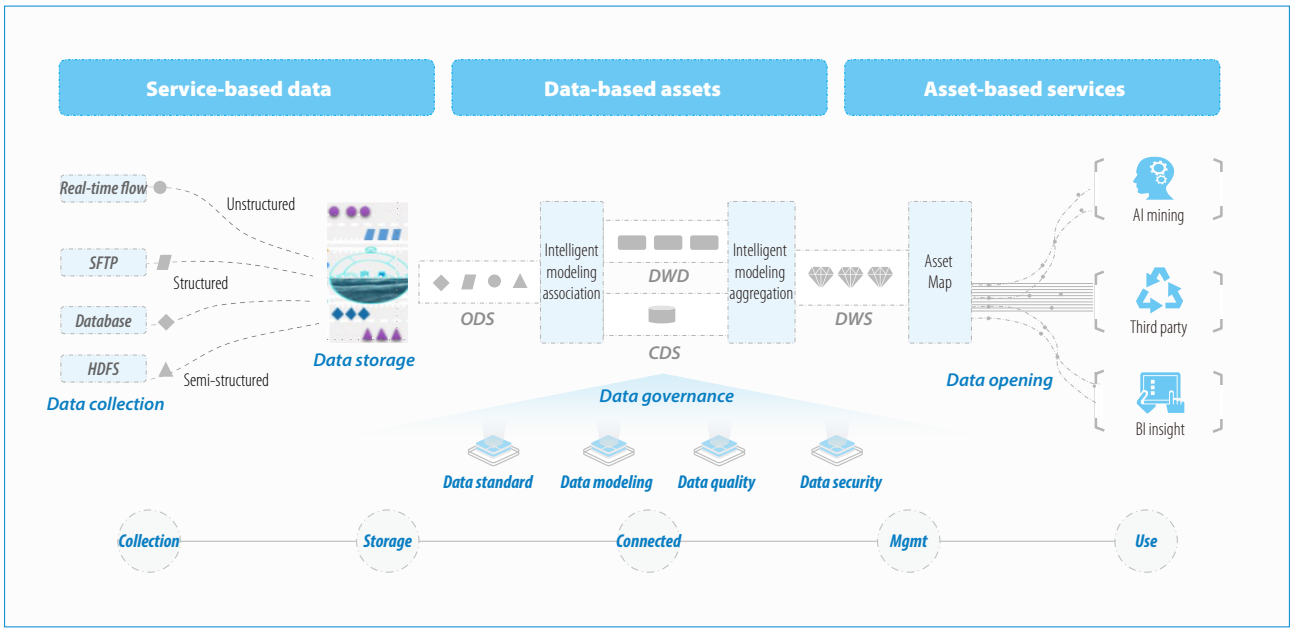
Includes two parts of data interface and tools. It provides an open environment and the visualized analysis platform, which solves the shortage of data sharing and analysis tools.

08 System management

Provides unified O&M management capabilities, and easily implements routine O&M management tasks such as system installation, upgrade, capacity expansion, monitoring, preventive maintenance, and troubleshooting.

Data Processing Architecture

VMAX supports one-stop big data lifecycle processing to ensure data collection, storage, connection and management that would use the data platform to provide service-based data, data-based assets, and asset-based services.



Features of VMAX

The VMAX Intelligenct Data Mid-End Platform has the following features:

Data integration , The VMAX adopts the mainstream big data technology, which provides good support for all types of data. The unified collection/ingestion, unified storage, unified scheduling, and separated storage, not only solve the data silo issue among all kinds of operator's systems, but also greatly reduce system integration costs.

Professional governance , One of the main principle of VMAX Intelligent Data Mid-End Platform is "unified governance and unified development," VMAX provides industry-leading telecom data governance solutions based on years of professional network service experience, From the beginning of system design phase, the data governance solution focuses on empowering service, formulates unified standard specifications and hierarchical modeling, sorts out service domains and entities from a global perspective, and outputs data from multiple dimensions, and outputs data models and assets with service insight.

Intelligent modeling, The VMAX Intelligent Data Mid-End Platform provides support for the whole process of AI application modeling, including data labeling, model training, model management and online reasoning service. In addition, it provides statistics and monitoring from resources to application layer to support production applications. Fully support machine learning, deep learning and intensive learning, built-in 150 + machine learning operators and mainstream deep learning framework, support drag and drop AI intelligent modeling.

Universal AI, The VMAX Intelligent Data Mid-End Platform Integrates the AI engines of both the big ring VSE (Very Huge Smart Engine) and the small ring LSE (Lite Smart Engine) . By combining the "large-scale" and "lightweight" deployment capabilities, it implements the flexible deployment from a cluster to cloud edge, and intelligent closed-loop of service applications. By way of automatic optimal model exploration and Adlik inference acceleration engine, the intelligent closed-loop optimal decisions can be found and timely executed.

Zero-sensing low-loss data masking, The VMAX Intelligent Data Mid-End Platform provides a trusted, manageable and controllable security system, and solves security problems of collection, storage, application and sharing in one stop. The first company through the DCA certification of big data desensitization testing with performance loss of only 3%.

Unified data service, VMAX Intelligent Data Mid-End Platform provides secure and easy-to-use data sharing and open services through OpenAPI. The subject data set is constructed based on the service logic of a metadata center to provide a unified data egress for external applications. The data difference between multiple data sources is filtered for users, so that data can be opened and upper-layer applications can be deployed easily on the platform.

Zero-coding visualized development, Based on big data, artificial intelligence and low-code technologies, the VMAX Intelligent Data Mid-End Platform creates an "one-stop visualized" data lifecycle operation platform, provides industrial-leading drag-and-drop development suite as well as real-time, easy-to-use, immersive and intelligent data services to achieve zero-coding data development.

Key Technologies

VMAX Intelligent Data Mid-End Platform has seven key technologies: Unified data integration standard, unified storage and calculation module, professional development and management process, full-process data management and control, AI intelligent modeling of easy-to-develop data service and all-round system management. It has unique advantages in data integration, stream batch processing, inventory and accounting separation, unified scheduling, data governance and self-development, and improves the capabilities of big data and AI service.

Unified Data Integration

Data integration is intended to provide multiple data sources with channels for entering and leaving full data or incremental data in complex network environments. According to different data sources, configuration and scheduling policies are visually formulated, and data is collected and loaded under the drive of the distributed workflow engine. It has the following technical advantages: Simple and easy to use, stable and efficient, secure and reliable, and elastic and scalable.

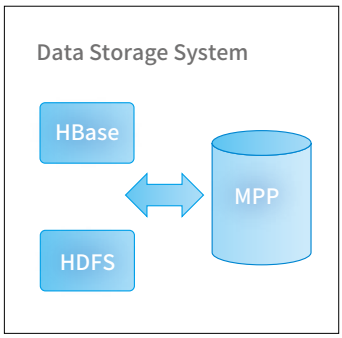
- Supports visualized online configuration access, data source collection management, and data source creation, modification, query, and deletion. The data source supports multiple structural and unstructured types such as FTP, SFTP, HDFS, relational database, Kafka, and real-time signaling. It also provides data access componentized plugging/unplugging capability to facilitate service expansion.
- Provides scheduling management and real-time monitoring of data collection tasks, including creating, modifying, copying, starting, stopping, and deleting collection tasks.
- Supports the scheduling management of task sets, enables easy management of collection of multiple data source instances of the same type, and supports the creation, modification, query, and deletion of task sets.
- Monitoring capabilities in each phase of the collection process, data source status monitoring, and file and field-based quality check to ensure data quality during the collection process.
- Based on a distributed computing framework, supports synchronization of task load balancing and all-memory in parallel, thus achieving higher efficiency. It has perfect traffic control and resource control measures.

  The data collection service is easy to use. Users only need to configure data sources and collection tasks through the visual interface, and it provides full-link management and maintenance capabilities for data sources and collection tasks, greatly reducing development and maintenance costs.

Unified Storage Computing

The processing layer adopts the "MPP+Hadoop" distributed cluster architecture. It integrates big data related components such as Zookeeper, Hadoop, Spark, HBase, YARN, Saturn and Gbase to implement distributed storage and parallel computing of mass data.

Data Storage



To solve the problem of storing, cleaning and value mining of mass heterogeneous data, the traditional centralized architecture should be transformed into a distributed architecture. It deploys industrial well-developed Hadoop+MPP hybrid architecture solution, and adopts different storage strategies according to different data types to provide structured, semi-structured and unstructured data storage capabilities.

In the hybrid architecture, the Hadoop platform provides a highly reliable and distributed storage and processing framework for mass data. The Hadoop cluster can be expanded to thousands of servers. The storage structure in the distributed framework of Hadoop is HDFS. HDFS adopts the multi-copy policy to ensure data security and reliability.

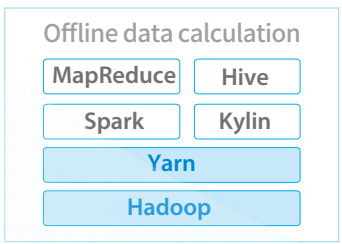
Based on HDFS, the distributed real-time database Hbase provides platform support for high concurrent retrieval and analysis and transaction support. Through multiple indexes, the Hbase supports multi-dimensional second-level query of massive data, including global index, full document index and combined index. In addition, the Hbase supports topology data storage and provides the graph retrieval function.

MPP database adopts multi-copy and Share Nothing architecture, and evolves into both OLAP and OLTP features. It is responsible for in-depth analysis, complicated query, KPI calculation, data mining and variable self-help analysis applications, and supports PB-level data storage and calculation.

Based on the above storage calculation model, the service flow is designed according to the features and advantages of Hadoop and MPP.

- Long term preservation on the Hadoop platform of cold data such as raw data and unstructured data. The analysis and processing of unstructured data, KPI and statistics analysis are completed on the Hadoop platform.
- The MPP database stores structured hot data such as summary data, and provides the ad hoc query capability.

Offline Data Calculation



Based on the Hadoop HDFS storage service, as a computing cluster, Hadoop Yarn is used to manage all the computing resources in the cluster in a unified manner. The resource queues using the tenant mechanism are used to effectively isolate resources and avoid the mutual impact of job contention for resources. With the data localization feature, the computing performance is effectively improved. The offline computing engine supports mainstream MapReduce, Hive, Spark and Kylin.

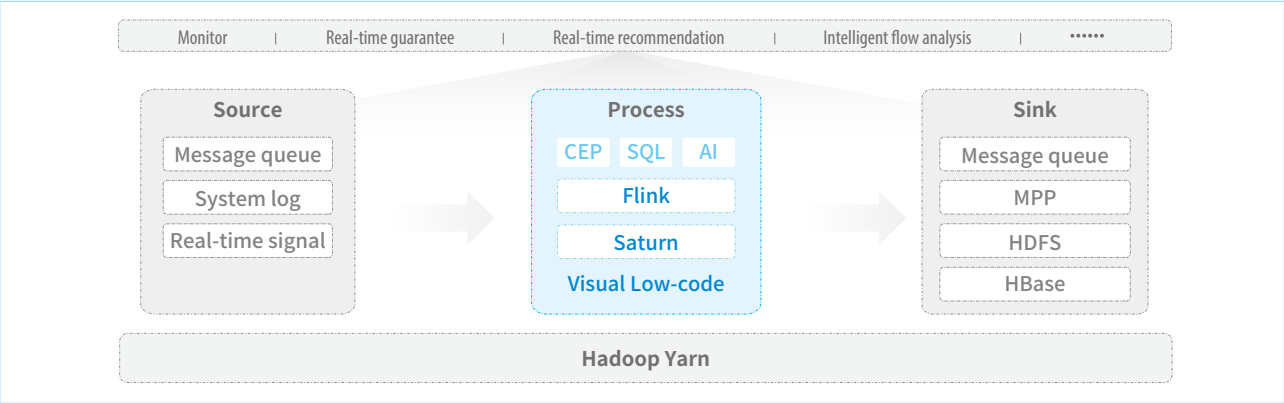
The Hadoop provides a distributed MapReduce calculation framework for parallel calculation of large-scale data sets (TB level, even PB level). It is designed to run batch jobs that last several hours. It encapsulates such technologies as parallel processing, fault tolerance processing, data localization optimization, and load balancing. It features low memory consumption and easy to use.

As a data warehouse framework based on Hadoop, Hive provides a convenient SQL-like analysis mechanism based on the MapReduce calculation engine to operate the large-scale data in the Hadoop file and lower the data analysis threshold.

The Spark distributed computing framework is more efficient, it is compatible with the HDFS distributed storage, and can be integrated into the Hadoop ecosystem. Unlike MapReduce, the intermediate outputs and results of the Job can be saved in the memory to avoid frequent reading from/writing into the HDFS. Therefore, the Spark is more suitable for iterative algorithm research such as data mining and machine learning.

Kylin is a multidimensional analysis engine based on Hadoop. It performs multidimensional modeling and analysis through scheduled Cube operation, and caches the analysis results to improve the query efficiency. Kylin supports multi-dimensional modeling and data query of class SQL, which is simple and easy to use.

Real-Time Stream Calculation



Real-time flow calculation provides low delay (millisecond level), high throughput and highly reliable distributed real-time calculation service. The platform uses two real-time processing engines, Flink and Saturn, to meet the requirements of various real-time flow calculation scenarios.

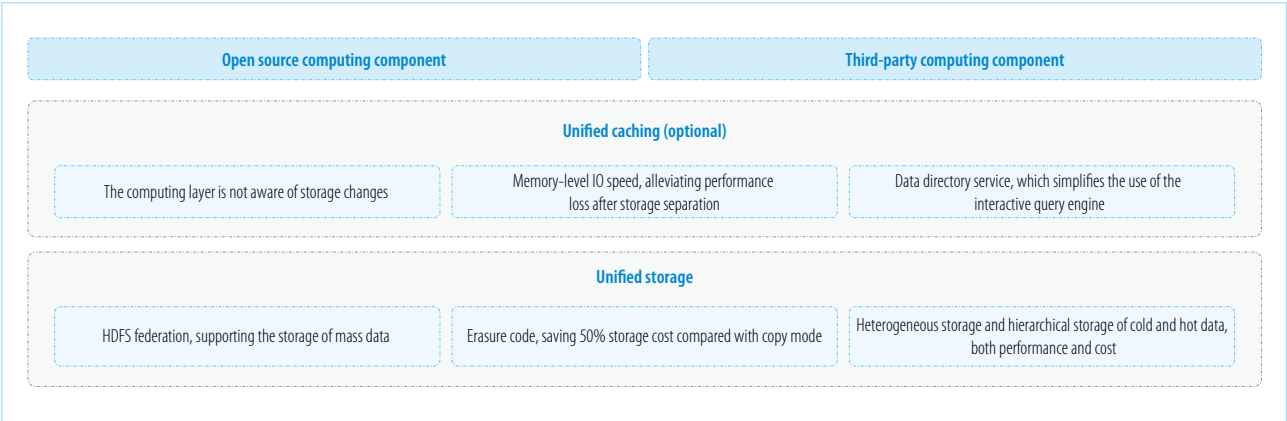
Based on Flink, the real-time processing engine is developed to support real-time processing of the general message queue system and some system logs.

Saturn is a high-performance real-time stream processing engine developed by ZTE to process signalling data.

Real-time stream calculation has the following advantages:

- Real-time stream processing includes complex transaction processing such as data conversion, processing, backfilling, and correlation..
- High throughput and low latency: single-node CAPS can support 300,000 to 500,000 processing capacity per second, achieving millisecond-level latency.
- Support in-memory computing and the evaluation of window aggregates.
- Supports AI real-time reasoning.
- Simple and easy to use, supports the development of real-time visual SQL tasks , and supports the visualized scheduling and management of real-time tasks.
- Security isolation and tenant security mechanisms ensure that all real-time operations are performed safely.

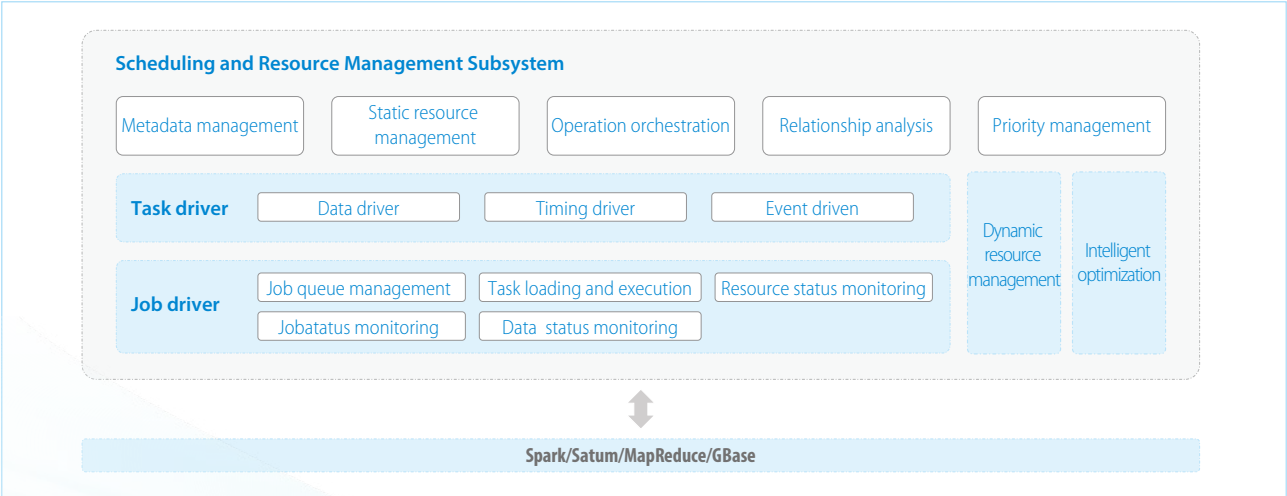
Disaggregated Storage and Compute Architecture



Decoupling storage and compute architecture uses dedicated servers for calculation and storage respectively. Compared with coupling storage and compute architecture, as the business grows, computing servers or storage servers can be independently expanded to achieve the purpose of reducing costs and increasing efficiency.

- The Alluxio distributed memory file system is introduced as the storage cache to provide data localization for calculation, share hot data, avoid intermediate data from being stored , and shorten IO path. This effectively solves the problem of performance degradation after storage separation.
- Achieve SSD-based HDFS write cache, the HDFS write performance is improved by 110%.
- The HDFS erasure code technology is introduced to reduce the storage size by 50%+ compared with the HDFS three-copy mode.
- The unified namespace technology is used to present objects such as HDFS file storage and Ozone/S3 to the outside in the form of unified file directory. Multiple bottom heterogeneous storage systems are filtered, so that application computing is not perceived.

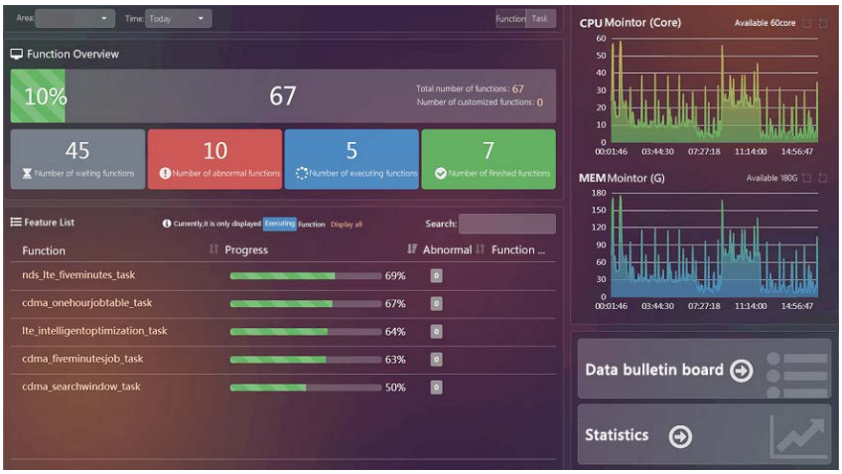
Unified Task Scheduling



The platform provides a unified scheduling engine to model and manage the full lifecycle of data, such as data integration, cleaning, analysis, and sharing. Various operations are abstract into jobs, and operate through job orchestration, driving, and execution. provides data lineage analysis and monitoring capabilities for various tasks.

Task scheduling has the following features:

- Provide unified task scheduling and management capabilities for various tasks (Shell, MR, Spark, SQL and Python).
- The system has the task orchestration capability, which allows multiple tasks to be orchestrated into one combined task, which can be driven flexibly in multiple ways. The tasks are restricted to each other. The task scheduling management can automatically start the follow-up tasks according to the execution status and results of each task.
- Supports offline big data scenario scheduling and real-time stream scenario scheduling and filter the difference between bottom-layer data storage engines.
- It supports dynamic resource allocation. Based on the history of task execution, it can automatically optimize task resource parameters through intelligent model for Top tasks with low running efficiency. The overall resource ratio can be optimized without manual intervention.
- Provides a monitoring view of task scheduling, and displays the operation and monitoring of various data processing tasks on the platform, including task status statistics, flow status statistics, flow definition statistics, operation status presentation, and task logs.



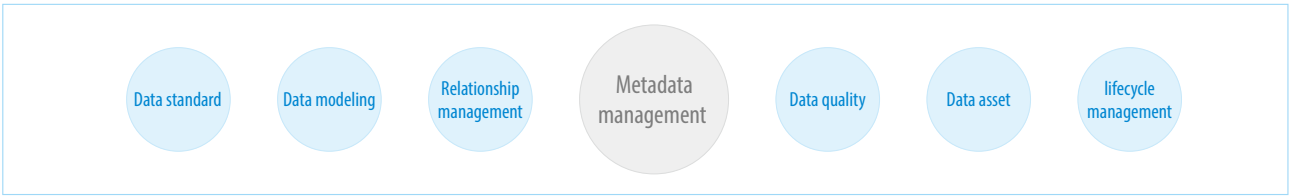
Batch Integration (XC)

VMAX Intelligent Data Mid-End Platform provides the capability of building an integrated batch capability with limited resources, and solves the problems of real-time and batch data coexistence, thus achieving unified batch storage, simplified system architecture, unified batch model and unified code.

- Yarn-based unified Spark and Flink computing resource scheduling; unified task scheduling to implement coordination of batch resources.
- ZSQL which is ZTE's self-developed UML unifies the programming languages of Spark and Flink. From the perspective of language specifications, ZSQL provides unified batch processing, so that one set of language can simultaneously complete the data development of both the stream and batch scenarios, greatly reducing the development complexity.

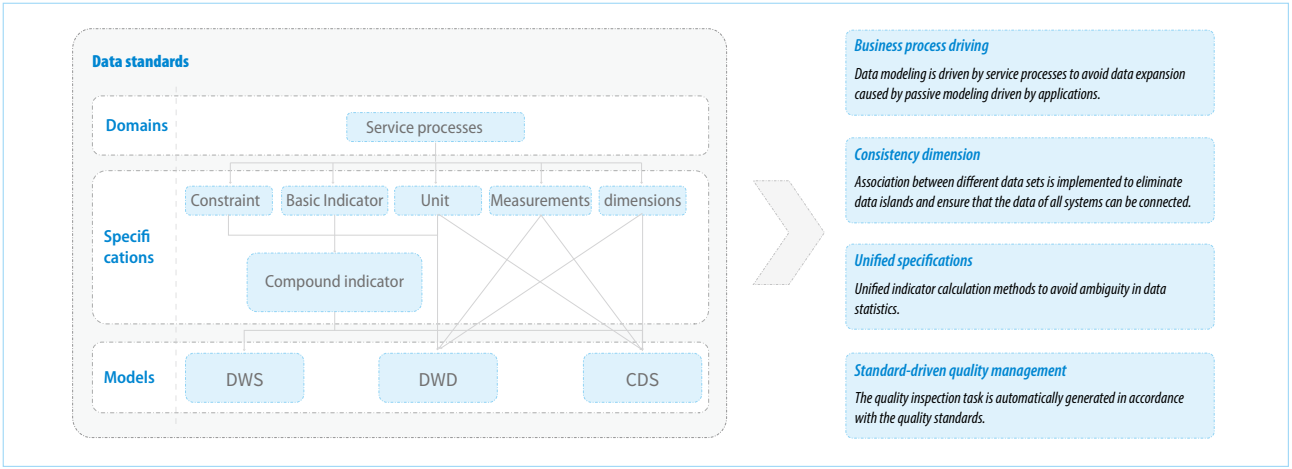
Professional Development and Governance

Development governance includes data modeling, data governance and data asset management. It is the core engine for data intelligence middle platform construction. It covers functions as standard defining, data modeling, data quality management, metadata management, data labeling, data asset management and lifecycle management, and create subject domain and label assets with unified standard, clear and stable model and manageable quality.



Data Standard

Data standard is the cornerstone of data development. Based on the experience in telecommunication business and the data characteristics , ZTE VMAX platform designs a complete set of data standard system and formulate data standard of O&M management and process control.



Data standard is built according to the following steps:

- Define the business scenario
- Define unified data specifications and data dimensions
- Set up data quality verification tasks

Establish the data modeling process driven by the business process, establish unified data specifications and consistency dimensions, set audit rules in accordance with quality standards, and automatically generate quality inspection tasks.

The VMAX Intelligent Data Mid-End platform unifies the real-time and offline data processing and data modeling. Both real-time and offline models can adopt the same set of data standards.

It eliminates data isolation,establishes the value of data extraction associated with multiple data sources through standard data.

Data Modeling

Data model is an extracted model which can reflects the association of service information (objects) in accordance with service requirements. That is, a subject fact table associated with other fact tables and dimension tables through the main/external key association. The VMAX platform standardizes the design of the data model on the dimension of consistency, and meets the following principles:

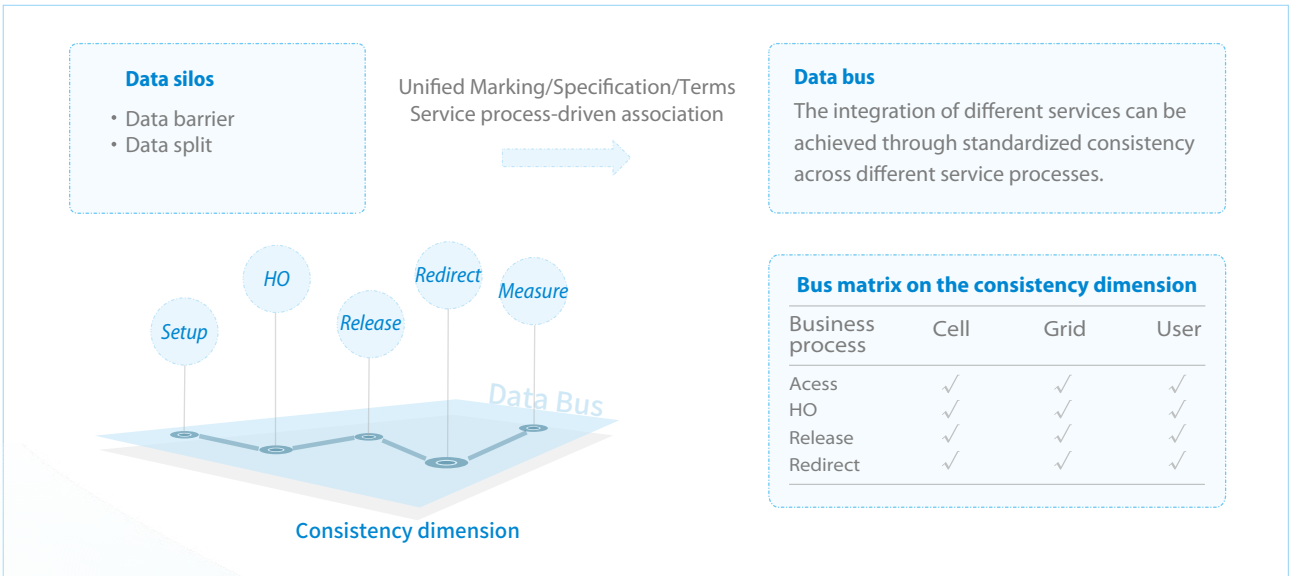
Dimension model should be built based on service requirements: The dimension model serves for service analysis. When creating a dimension model, you must create actual entities and dimension entities in accordance with the data used in service analysis.

Ensure that multiple indicators in each actual entity have the same granularity or level: Multiple indicators in the same actual entity must have the same granularity,otherwise ambiguity may occur during analysis, presentation, and summarization.

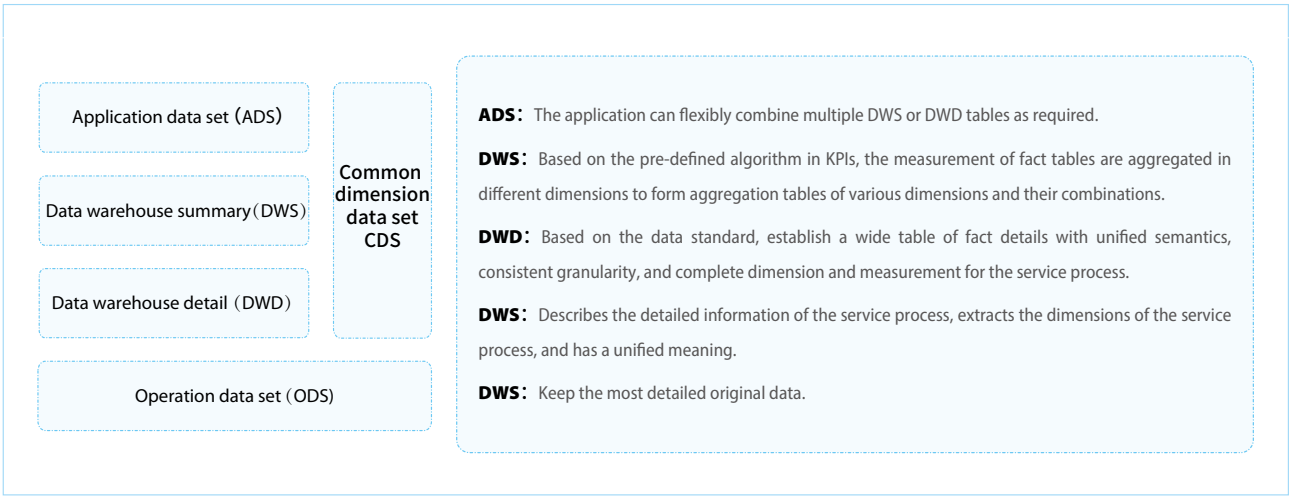
A consistent dimension shall be created: The entire platform uses a consistent dimension (also called general dimension, standard dimension, or reference dimension), and can be reused in all actual entities. In this way, data can be integrated from multiple service processes, eliminating redundancy design, and shortening the development process.

Ensure that each entity has an associated time dimension: Each business event has a time stamp, and the platform stores a large amount of historical data. Each fact is related to a time point. Therefore, each actual entity must have a time dimension associated with it. The granularity can be yearly, quarterly, monthly, daily, or hourly.

The VMAX platform integrates the isolated service data through the data bus to form an information model with the service entity as the core and the basic attributes and extended attributes as the auxiliary through the dimension of unified data standard and consistency. In this way, we can perform multi-dimensional value extraction for the isolated and discrete data, so as to create a set of subject-oriented data model.

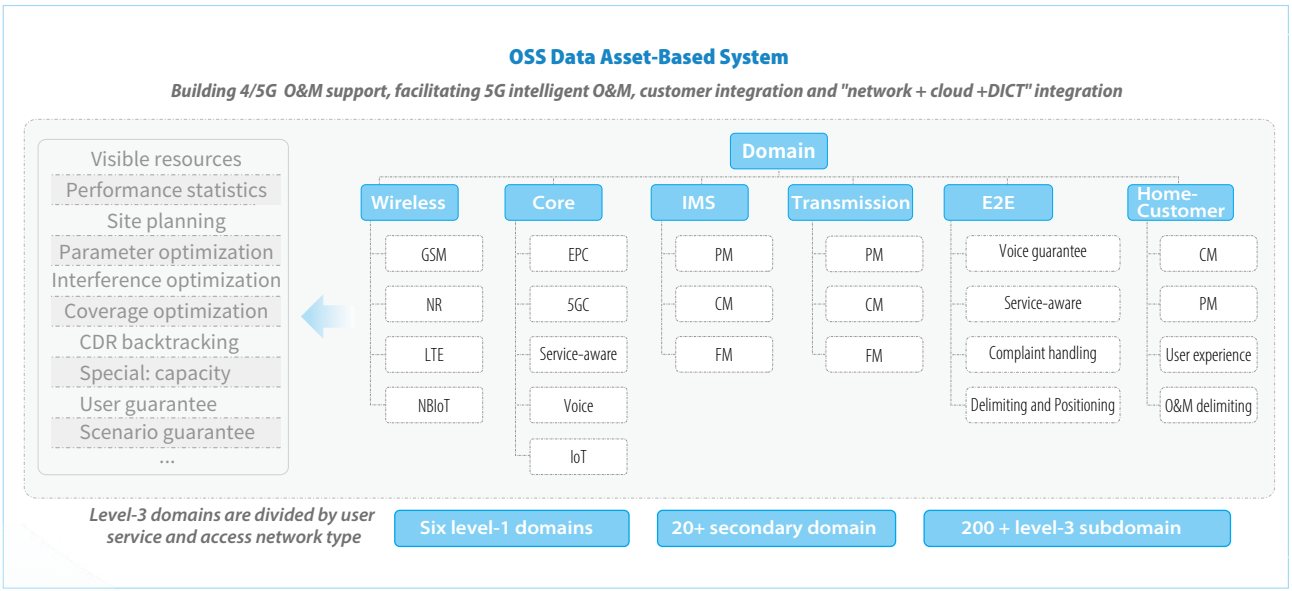


The key to successful data modeling is to establish a reasonable data architecture. By building a hierarchical architecture, redundant storage can be reduced, computing resources can be reduced, and computing time can be shortened. A unified data base can be used to better manage and share data.

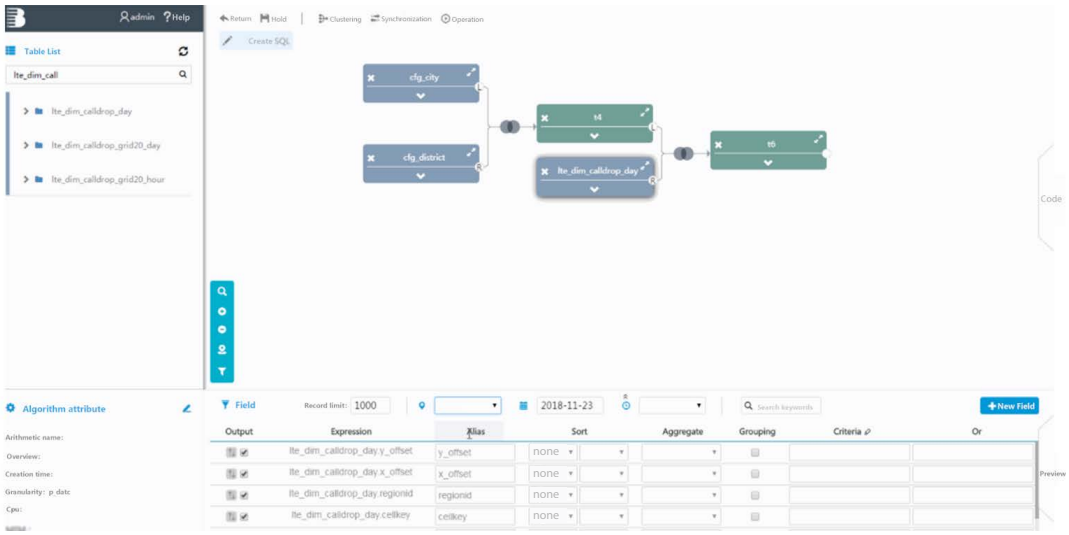


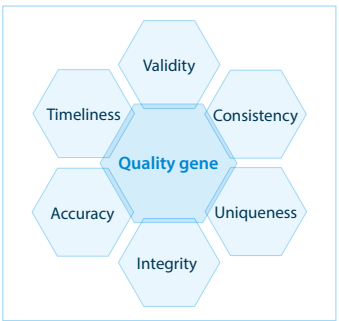
Data Model

The ZTE VMAX platform builds a data model based on dimension modeling and its professional understanding, provides standardized and structured data assets, supports user perception evaluation, poor quality analysis and end-to-end delimiting and positioning in the 4G, 5G, home broadband and enterprise customer fields, and achieves integrated intelligent O&M of the network.



The VMAX platform provides the visual data modeling capability, which can develop the data model quickly and agilely by dragging in graphical mode.





Data Quality Management

Data quality management provides an end-to-end data quality audit system, and data audit rule management, data audit task management, data quality monitoring, and quality root cause analysis function, which is the core to data governance.

The platform measures data quality from six indicators: Integrity, validity, consistency , accuracy, uniqueness, and timeliness . Through data quality management, clean data and clear structure can be achieved and it lays the foundation for data value development.

Metadata Management

Metadata describes data, including data label descriptions, database tables and field definitions, service definitions, and relationship between them. The unified metadata can create a data map portal that can find data, so that the data lineage and influence analysis map of the data and a set of data quality scoring system can be used conveniently.

Metadata management is the basis of enterprise data governance. It manages the description information of the full lifecycle of data collection, storage, development and governance, and open sharing in a unified manner. It provides metadata classification, kinship and impact analysis capabilities, helping users clarify data relationships, reducing the difficulty of using metadata, and improving user experience. Metadata management functions include: Metadata collection, automatic kinship association resolution, version management flow, metadata view management, metadata data analysis and metadata external service.

Metadata collection:

Collect full metadata from RDBMS and big data platform, including data entities (system, database, table, and field) in the process and logical relationship in the processing procedure of data entities.

Automatic relationship analysis:

For collected metadata, the analysis process is started automatically, and a network of relationship between entity object tables and fields is created to support impact analysis and full chain analysis.

Metadata version management:

Metadata change contents can be analyzed periodically and automatically. Metadata change contents and their impact scopes can be output to support metadata change monitoring applications. Historical versions can be viewed and compared, so that users can monitor the changes at any time and eliminate potential problems.

Metadata quality management:

An effective metadata quality check mechanism is established to discover, report, and handle metadata quality problems in a timely manner, ensuring the quality of metadata. Metadata quality audit reports can be provided, including summary reports and detailed reports.

Metadata analysis:

It provides abundant metadata analysis functions, including kinship analysis, impact analysis, whole chain analysis, correlation analysis and attribute value difference analysis. It analyzes the origin and impact of the metadata, so as to quickly identify the value and help the user to use, evaluate and clean up the data assets efficiently and accurately.

Metadata query service:

Provides the metadata query and usage analysis functions, besides it provides the metadata keyword retrieval and kinship analysis functions in a standardized manner to support invoking from internal and external systems.

Data Assets

Data asset management presents the data model systematically in the data asset module through collection, integration, and processing, provides asset map and asset retrieval functions.It constructs clear, complete, and high-quality data asset views of enterprise-level, implements classification management of service data, and creates yellow pages that provide convenient retrieval.

Builds model structure to provide a clear data structure for data developers.

Build asset **panorama to help CXO level executives manage the overall distribution of assets.**

Build value stream **control to present visual data streams to data managers.**

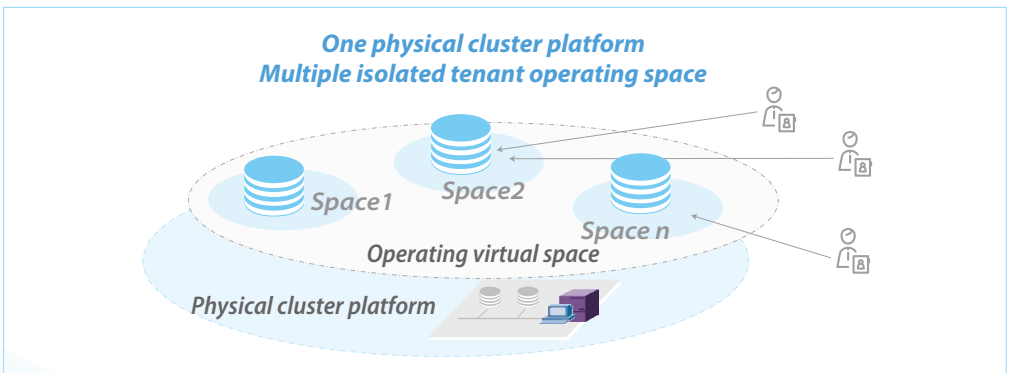
Life Cycle Management

Through periodic collection and quantitative evaluation, life cycle management builds an asset operation center to complete an optimal close loop of data collection, storage, utilization to value release, and achieves systematic and visualized data asset management.The machine learning technology is used to comprehensively evaluate the quotation popularity of data entities and realize intelligent data O&M policies. During the "offline" data clearing process ,based on the relationship of metadata and kinship link , the knowledge map should be combined to implement the reasonable recycling of related models and applications, thus effectively releasing computing and storage resources.

Process Management and Control

Tenant Management

The unified tenant management function creates different tenants such as marketing team, CS team and tenant resources for the platform, meeting the requirements of resource density management and the security isolation of each user for data processing. It implements dynamic configuration and management, resource isolation, and resource usage statistics for physical and logical resources between tenants in a cluster, and provides users with secure and reliable multi-tenant services.



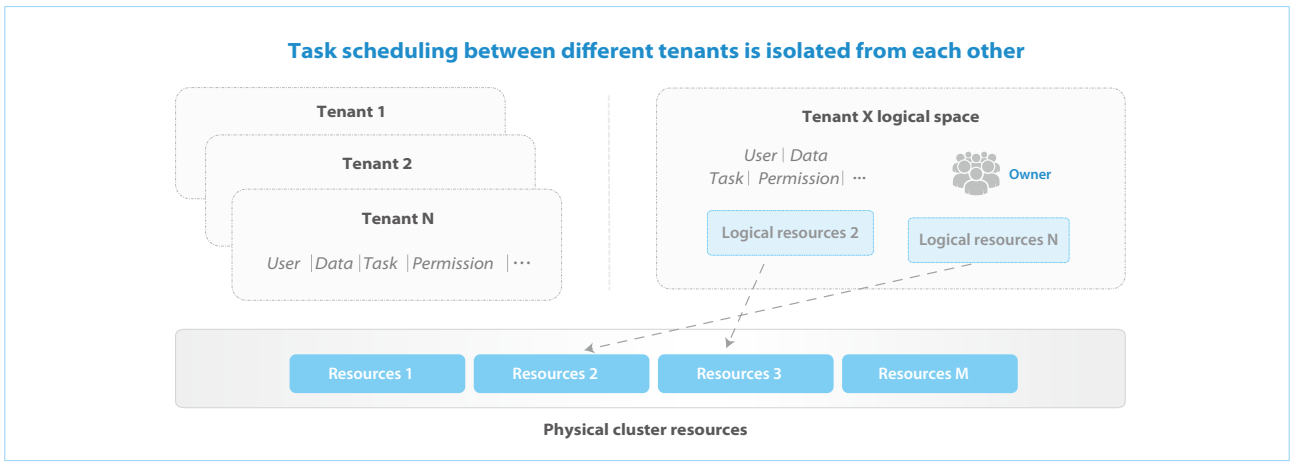
The tenant mechanism based on VMAX platform has the following features:

Tenant permission model: For computing and storage resources, tenants can apply for resource quotas as required and manage their own resources independently. For data, tenants can manage their own data independently to separate them from other tenants and ensure data security.

Tenant management capability: Includes tenant information management and tenant environment management.

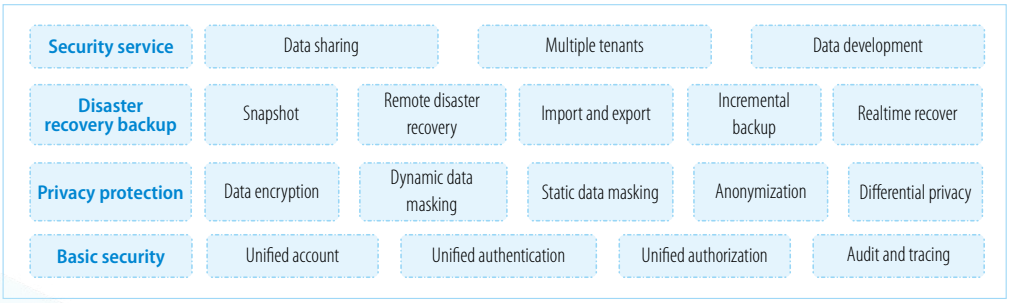
Tenant isolation capability: Supports performance SLA isolation, fault SLA isolation, data isolation, and security isolation. Isolated tenants are deployed in accordance with the requirements of virtual resource pools in terms of the use of physical resources. In this way, the sharing and scalability of physical resource pools can be implemented, and resource utilization can be optimized on the basis of tenant isolation.

Tenant operation monitoring capability: It provides the resource configuration and usage of tenants, performance monitoring, task operation monitoring and presents the overview of the tenants managed by the cluster. It supports graphical display of the allocation of resources such as CPU, memory and HDFS used by the tenants in the cluster.



Data Security Management

Through an all-round security system developed, this system sets security levels, and provides data security services that can be managed and controlled beforehand, and can be queried afterwards, to avoid data leakage risks. The ZTE VMAX security system consists of four parts: Basic security, privacy protection, disaster recovery backup and security service.



The VMAX platform is the first to create an integrated security solution for big data masking, providing an end-to-end security guarantee system to protect sensitive information of operators.

- Fully transparent to users and applications
- One-stop solution to collection, storage, application, and sharing security problems
- The anonymous performance loss of 3% ranks first in the industry
- 13 core invention patents
- 100% compatible TPCDS-99 test

Data Service

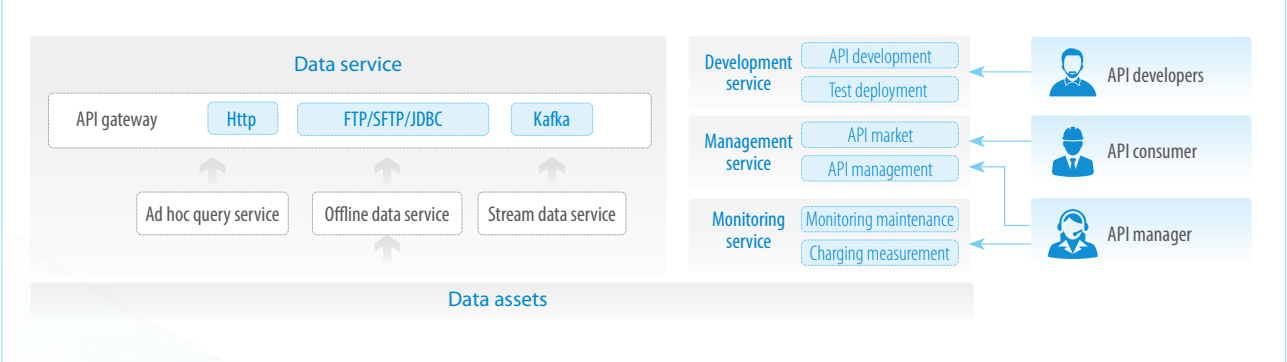
The data service provides two open capabilities:

Data openness: meets the requirements of multi-manufacturer and multi-application access, and to provide the functions of service development, deployment, authorized access, monitoring and other aspects required for data openness. It can provide unified data opening capabilities through OpenAPI, and achieve unified management and control of the entire process covering service design, service realization, service publishing, and consumption subscription.

Tool openness: provides basic tool capability support for application developers. It can provide an immersive data open tool suite, including self-service query analysis and interactive data mining, and support BI and AI analysis capabilities based on service requirements.

Center of Capability Openness

The center of capability openness provides data service interfaces through OpenAPI, to realize the data assets-based service, and supports the technical platform of enterprise internal management and value-added business cooperation with external partners, to realize internal asset operation and external capability opening.



OpenAPI provides the full lifecycle solution of API, including four functional components: Data service, development service, management service and monitoring service.

Data service supports the opening of different types of data, including HDFS, Spark, HBase, MPP database and Flink. It supports multiple open interfaces and modes to meet different development scenarios.

Development service provides developers with the capability of rapid API development and encapsulation through a visual developer platform.

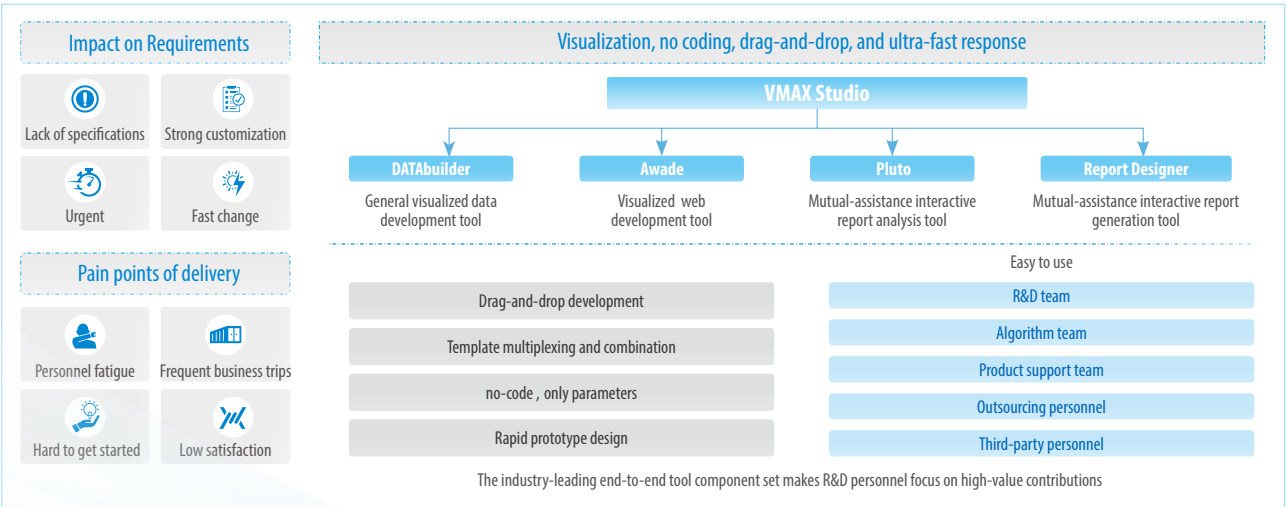
Management service provides API browsing, application, management, and other full-process function sets for API administrators and consumers.

Monitoring service provides API managers with a unified monitoring capability for the run-time API.

The platform manages external open data interfaces through OpenAPI to achieve unified management of users, data permissions and data interfaces, and provides market-based management methods for data monetization and value mining.

Data Analysis Service

The VMAX platform provides a set of simple and easy-to-use data application development kits, providing powerful self-service analysis and application development capabilities. The efficiency of data reports is greatly improved by virtue of visualization, non-coding, drag-and-drop, and high-speed response modes.

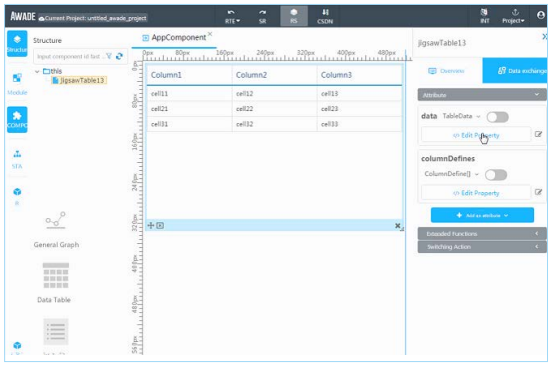


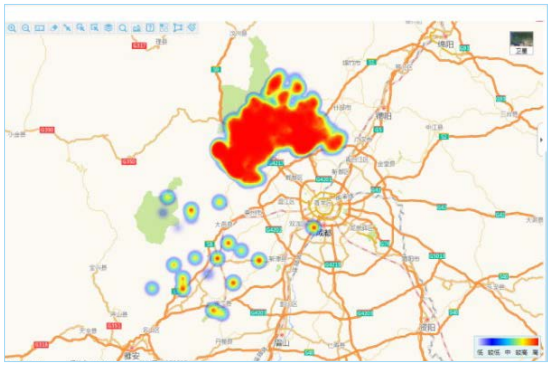
The tool suite has the following advantages:

Visualized interface. The algorithm exploration and data analysis can be completed quickly without coding through the drag-and-drop operation.

Visualized page layout supports dynamic page adjustment and lowers the threshold for professional and complicated page development. Functions can be easily reused and shared through templates.

Diverse interactive reports can be customized flexibly, and can be generated automatically or manually. Reports can be displayed in various graphs such as cells, grids, cell lines, and heat maps.





Data Mining Service

The VMAX provides a one-stop platform for AI algorithm exploration and intelligent application development, and provides a visualized online interactive machine learning development environment. It integrates multiple deep learning frameworks, and includes abundant operator libraries. Through simple drag-and-drop operations, the platform can explore data algorithm models, reduce data exploration thresholds, and help users quickly explore and develop algorithms for service applications online.

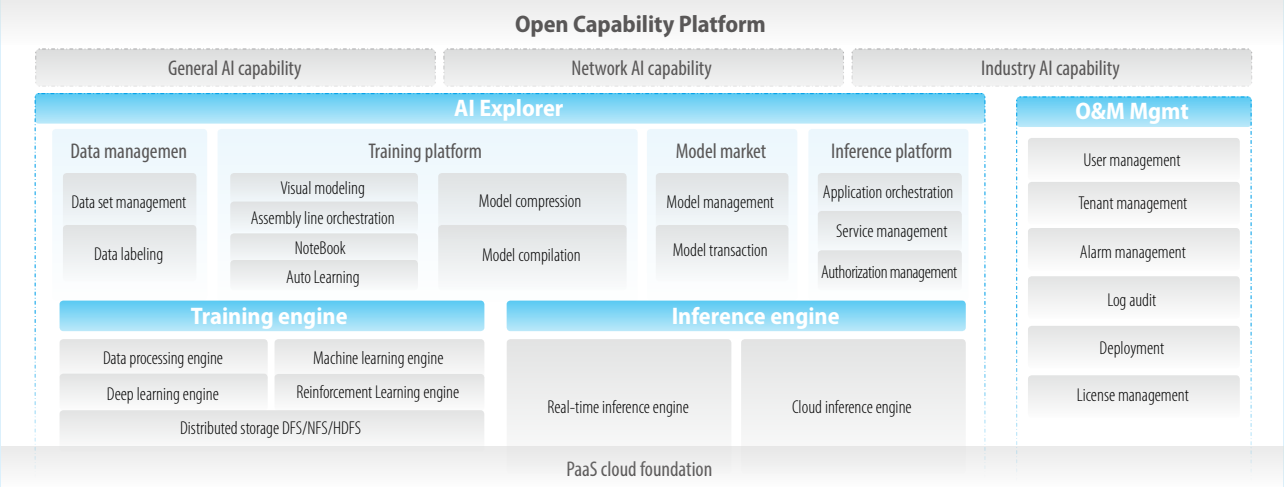
AI Intelligent Modeling

The AI technology has three key elements, computing power, algorithm and data. ZTE VMAX Intelligent Data Mid-End Platform focuses on such key attributes as high efficiency, scalability, smooth evolution and ease of use, to ensure long-term technological advance, continuous service empowerment, fast service capability development and convenient operation and maintenance.

AI intelligent modeling uses the micro-service architecture, and docker mirroring is used as the engine carrier. The training engine and the reasoning engine provide basic training and reasoning capabilities. Based on standard design, the training engine can be replaced by third-party engines, and supports new hardware, framework, algorithm, and application environment, and long-term expansion and evolution.

The built-in general AI capability is developed by the training platform, encapsulated as a standard AI capability, and exported as a Docker image. External applications can access the AI capability service through REST, GRPC and other protocols.

The main functional components of intelligent modeling include the data platform, training platform, model management platform, and reasoning platform. Decoupled components for smooth upgrade.



Data management platform: Collects, manages, and labels data. It provides access adaptation to multiple different types of data (FTP/RMDB/HDFS/csv etc.), and supports structural and unstructured data labeling. The data types supported include text/image/voice/video, and the labeling mode supports manual data labeling and automatic labeling.

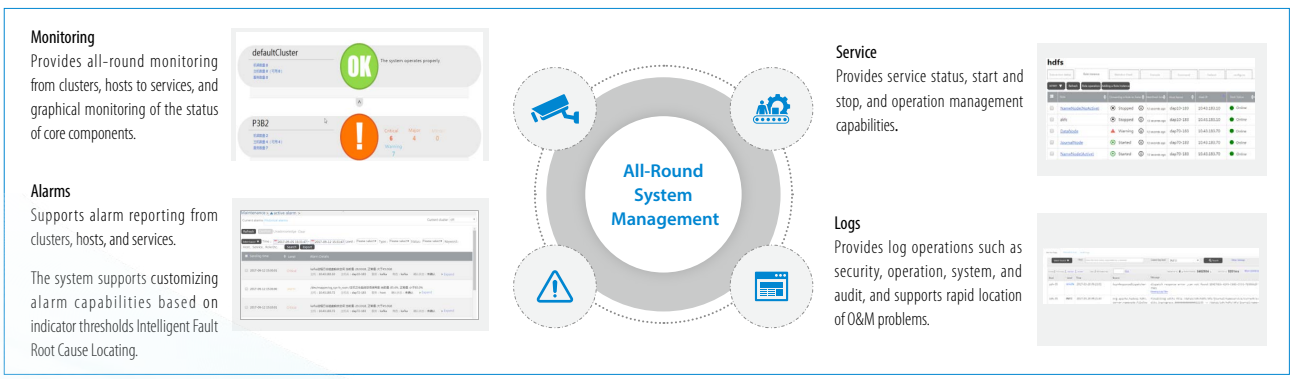
Training platform: Mainly used for model establishment, training, compilation and optimization. The platform supports model building both by using a visual modeling tool or direct programming, and then use the data set from data management platform to train the model to ensure that the accuracy meets the requirements. Finally, use the integrated compilation tool of the platform to compile and optimize the trained model, so that its operational performance (such as delay and throughput) can meet requirements. Finally, the model can be published to the model market for sharing by other users.

Model market: Used for model management and sharing. The model market is a trading platform for model suppliers and demanders. The model can come from the model trained by the training platform mentioned above, or come from the commercial model purchased from a third party. The requiring party can download the required model from the model market.

Inference platform: Used for model deployment and service management (such as startup, stop, expansion, load balancing, rolling upgrade, and fault migration). The inference platform can download the corresponding model from the model market and deploy the model as a service that external applications can access.

All-Round System Management

System management provides unified installation, deployment, management, and maintenance capabilities, including automatic installation, security management, alarm management, intelligent root cause analysis, platform monitoring, service management, host management, preventive maintenance, and resource scheduling policy management.



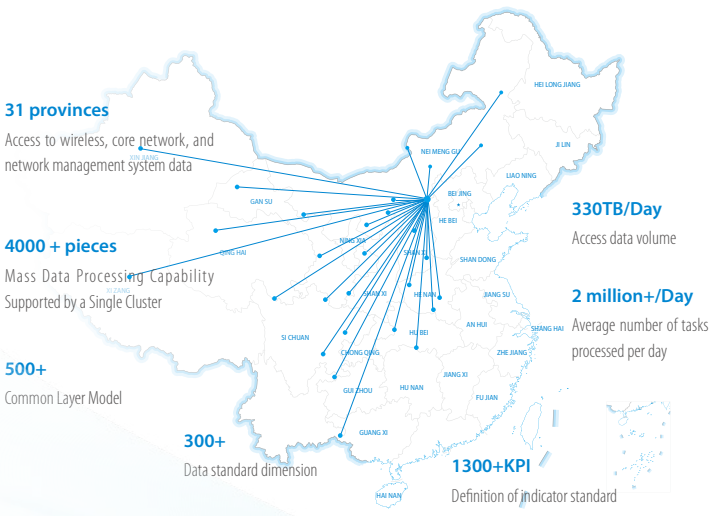
Use Case of VMAX Intelligence Data Mid-End Platform

Centrialized Data Analysis for 31 Provinces

The HQ of one of the major operator aggregates a large amount of data from 31 provinces, but the quality of the data is poor and cannot effectively support business agility.

With the support of VMAX platform the group has improved a lot in their digital and intelligent transformation:

- Access the wireless, DPI, NM and other system data of 31 provinces nationwide, with the data volume exceeding 330 TB/day.
- Support mass data processing capability. A single cluster server has more than 4000 servers, and the daily average task processing capacity is 2 million+/day.
- Unified data governance standard of O-domain , covering 500 + public layer models, 300 + dimensions and definition of 1300 standards, supporting rapid development of planning, construction, maintenance and optimization applications.
- All-round security capabilities guarantee the use and application development of multiple third-party data.
- Apart from the national O&M domain data governance, VMAX also provides the network coverage quality dashboard to present the network coverage quality status nationally. It also help the RF team to analyze the network coverage in multiple dimensions and support the TOPN finding of the worest cells with RF problems, such as weak coverage, over shooting, overlapping and MOD3 interference cells, and provide detail to help network optimization team to improve network performance.

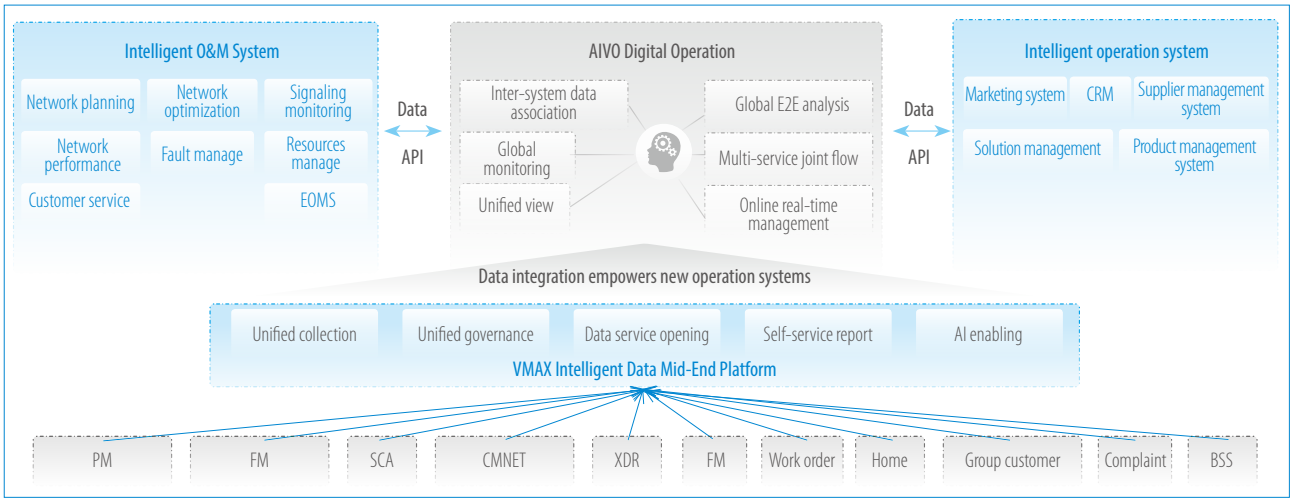


- Access the wireless, DPI, NM and other system data of 31 provinces in the country, unify the data governance standard O-domain data, and support the rapid development of planning, construction, maintenance and optimization applications.
- Guarantee the use and application development of multiple third-party data.
- We have analyzed the positioning and behaviors of users across the country, providing a basis for making decisions on government COVID-19 prevention and control.

Data Sharing Platform Construction in a Province of China Mobile

Data cross correlation and unified modeling are implemented through the access of DPI signaling, wireless, core network, NM configuration, alarm, performance, fault, and resource data to implement end-to-end problem delimitation, quality evaluation, and user identification.

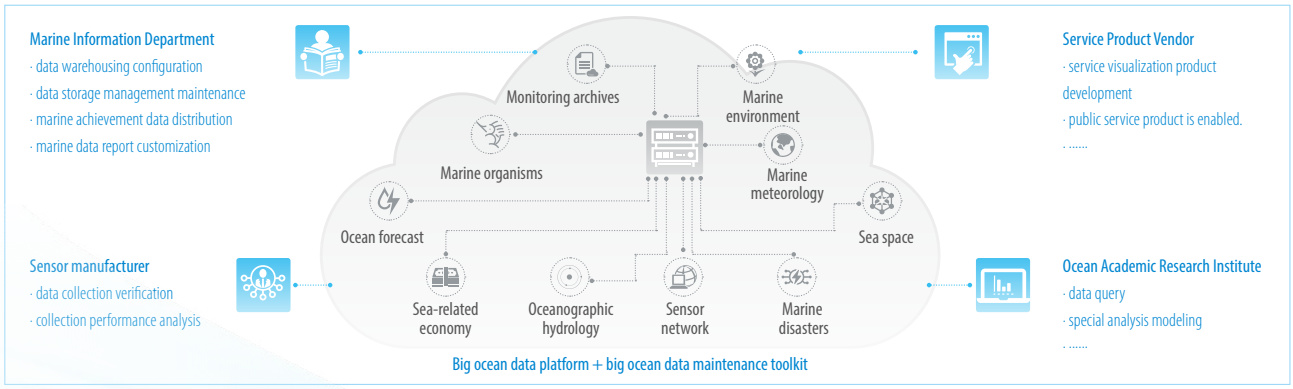
VMAX platform implements data governance from four aspects: Building foundation, capability enablement, operation support and innovation oriented. It conducts research on innovative topics related to data value realization, establishes a unified support platform, empowers intelligent network O&M and intelligent operation, and completes the overall digital transformation of the network.



VMAX Intelligence Data Mid-End Platform Construction for an Marine Bureau in China

VMAX Intelligent Data Mid-End Platform access to six categories of pan-ocean and four categories of related data field, which covering the access and management of a variety of structured and unstructured data. Establish the first data standard management system in the marine industry to connect sensor equipment manufacturers, marine information departments, business product manufacturers and scientific research units.

Through the data management and zero coding capabilities of VMAX Intelligence Data Mid-End Platform, it efficiently supports the development of marine big data applications.

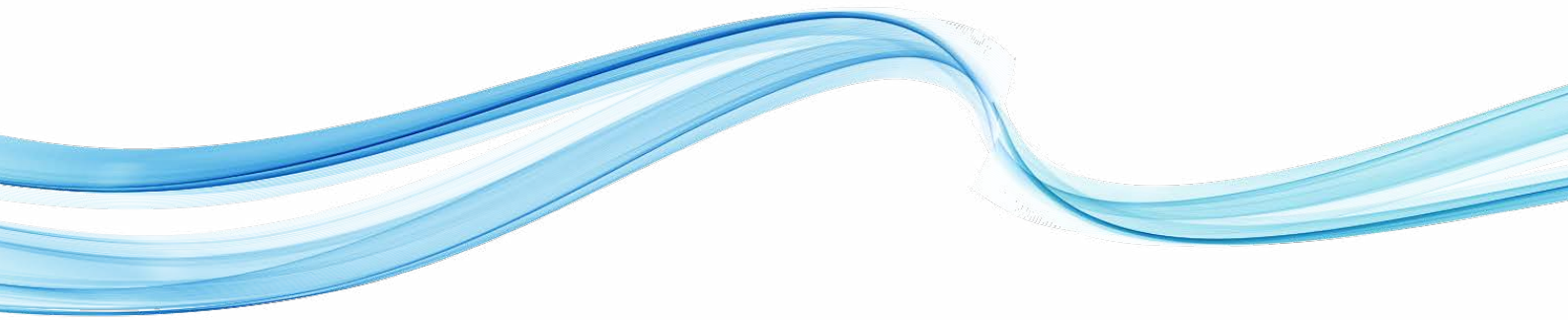


Future Prospects

As a long-term strategic partner of operators, ZTE has been committed to the R&D of new technologies, aiming to help global operators meet new challenges and opportunities. VMAX Intelligence Data Mid-End Platform is both capability building and ecological building. It will take openness, cooperation and win-win, work hand in hand with operators and partners, and actively participate in enterprise digital transformation.

With the aim of openness, cooperation and win-win, we will jointly work with operators and partners to actively participate in the process of enterprise digital and intelligence transformation.

Acronyms	Full Name
AI	Artificial Intelligence
API	Application Programming Interface
IDC	Internet Data Center
IT	Information Technology
MPP	Analytical Massively Parallel Processing Databases
O&M	operation and maintenance
VMAX	Value Multi-Analysis eXpert



ZTE ZTE CORPORATION

NO.55, Hi-tech Road South, ShenZhen,P.R. China Postcode:518057
Website:www.zte.com.cn Tel:+86-755-26770000 Fax:+86-755-26771999