ViMaCs
Virtual Master Cooperation
Data Science

Big Data Acquisition: SCADA systems approach

Anatoliy Sachenko, Michael Dombrowski

West Ukrainian National University, Department for Information Computer Systems and Control, Research Institute for Intelligent Computer Systems

SCADA Systems

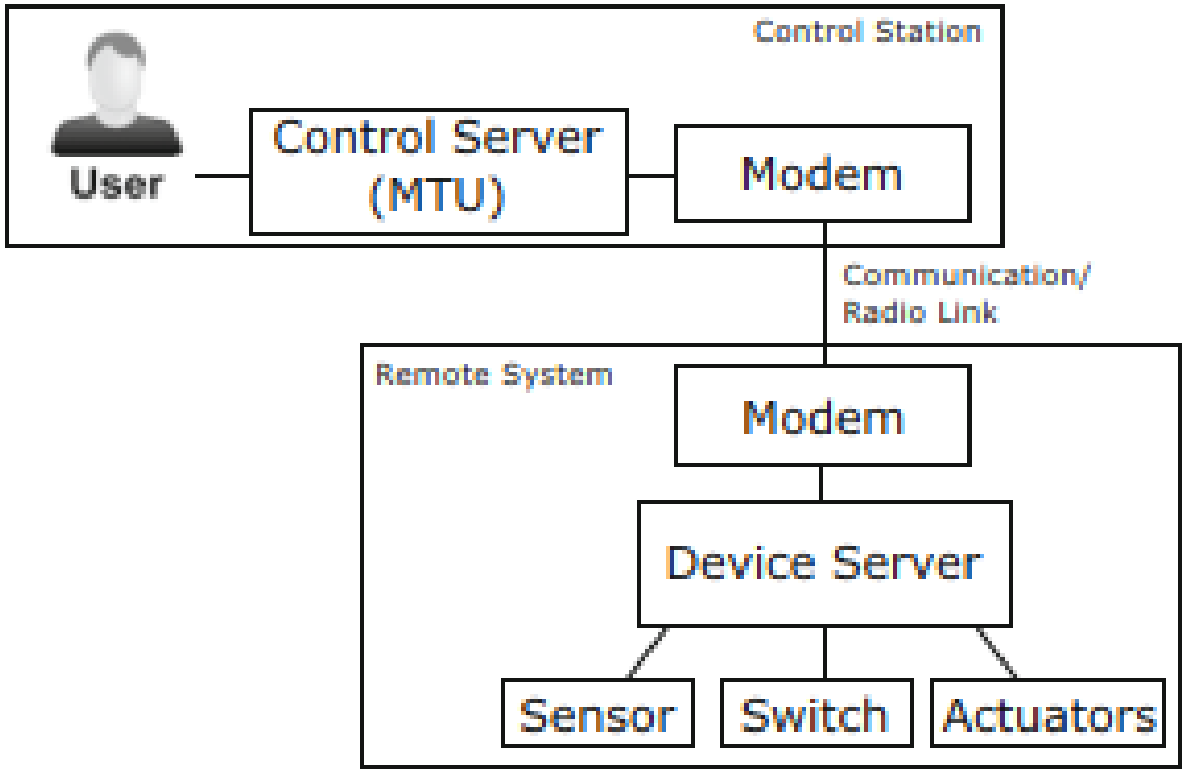
- **Supervisory Control and Data Acquisition** (SCADA) systems play an important role in monitoring of many industrial processes
- Applications include but are not limited to oil refineries, water treatment and distribution systems, transport systems, manufacturing plants, and electric power generations.
- A SCADA system gathers measurements—collects data from the deployed field devices—and sends them to a central site (normally a control center) for further processing and analysis.
- The information and status of the supervised and monitored processes are displayed at the base station or at the utility center.
- As such industrial systems are large and complex, a central master unit continuously and remotely monitors and controls different sections of the plant to guarantee their proper functioning.

SCADA Systems

SCADA is widely used in an industrial system that continuously monitors and controls many different sections of industrial infrastructure. Applications include but are not limited to oil refineries, water treatment and distribution systems, transport systems, manufacturing plants, and electric power generations, like areas of major applications of Big Data.

A SCADA system gathers measurements - collects data from the deployed field devices- and sends them to a central site (normally a control center) for further processing and analysis. The information and status of the supervised and monitored processes are displayed at the base station or at the utility center. As such industrial systems are large and complex, a central master unit continuously and remotely monitors and controls different sections of the plant to guarantee their proper functioning.

SCADA Systems



Simple SCADA system

A SCADA network provides the communication infrastructure for different field devices (e.g., PLCs and RTU) of a plant. These field devices are remotely monitored and controlled throughout the SCADA network. To make the network communication more efficient and secure, many modern computing technologies have evolved from a monolithic system to a distributed system and to the current inter-networked systems: IoT and Data Science based solutions.

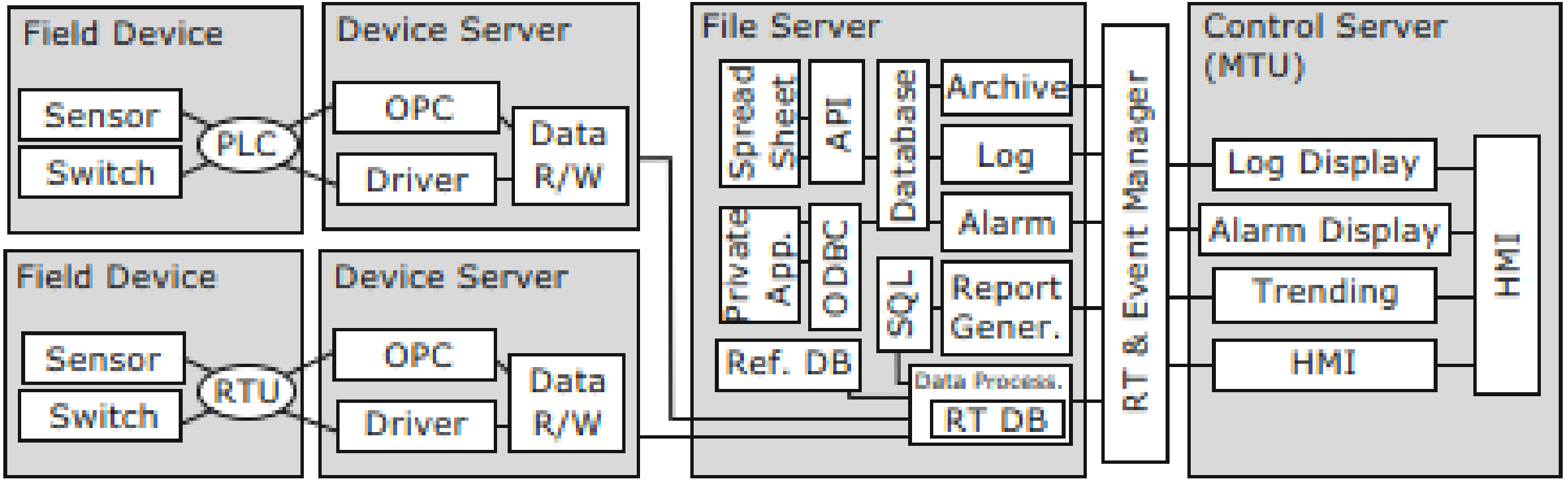
SCADA Components

As defined by the IEEE standard, the system is divided into a remote site and master station. The remote site consists of field devices connected to a device server. Communication between the field device and device server makes use of a SCADA communication standard (used by the driver). Collected information is stored in real-time and historical databases on the master station. Communication between components of the master station uses an internal communication protocol.

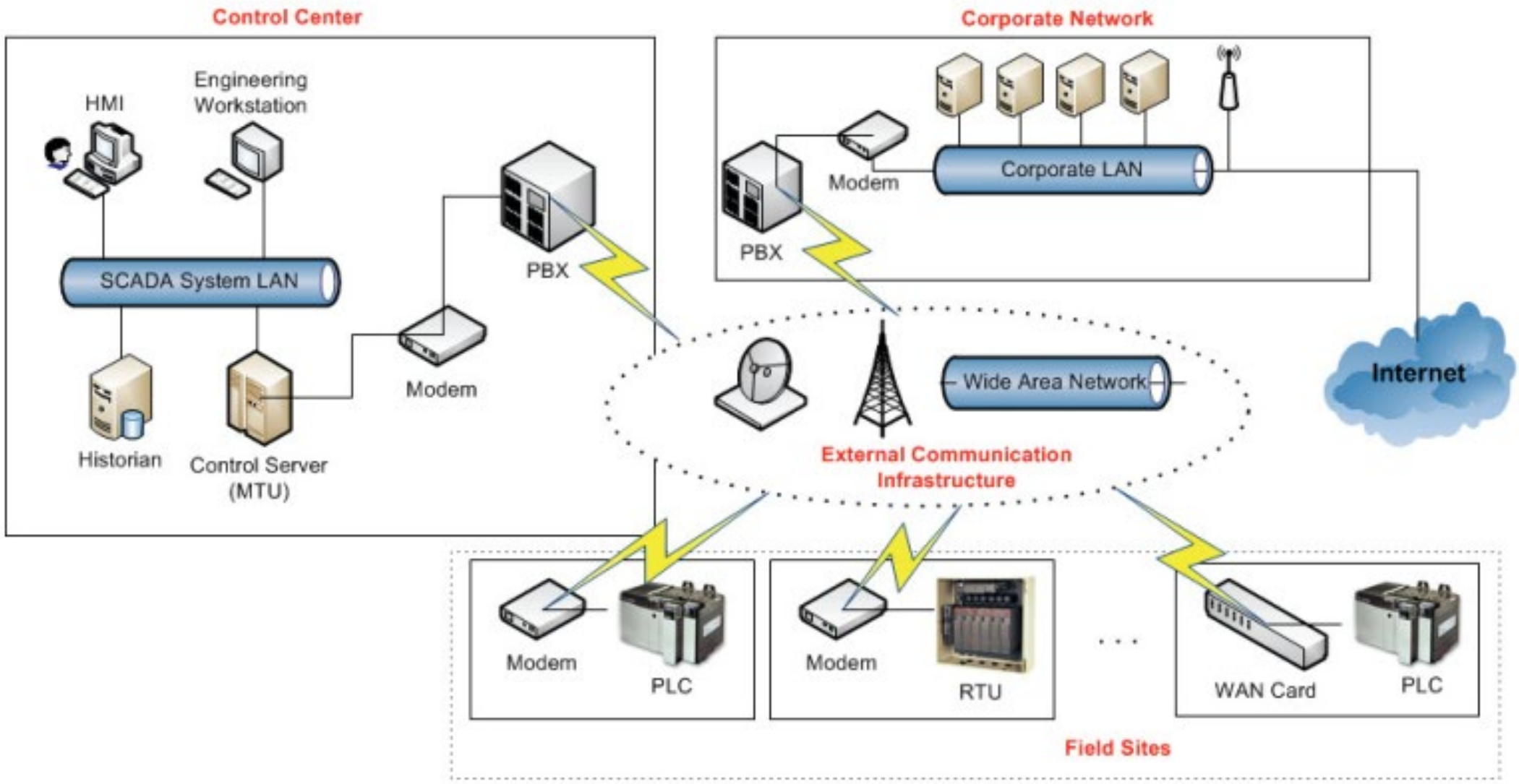
The main components of a typical SCADA system include the following: Master Terminal Unit (MTU), Programmable Logic Controller (PLCs), Remote Terminal Unit (RTU),Intelligent Electronic Device(IED),Human Machine Interface (HMI),and Communication Media.

Every SCADA system uses a large number of field devices. A field device must be able to understand and collect data from sensors, switches and actuators.

A realization of the generalized SCADA architecture is shown

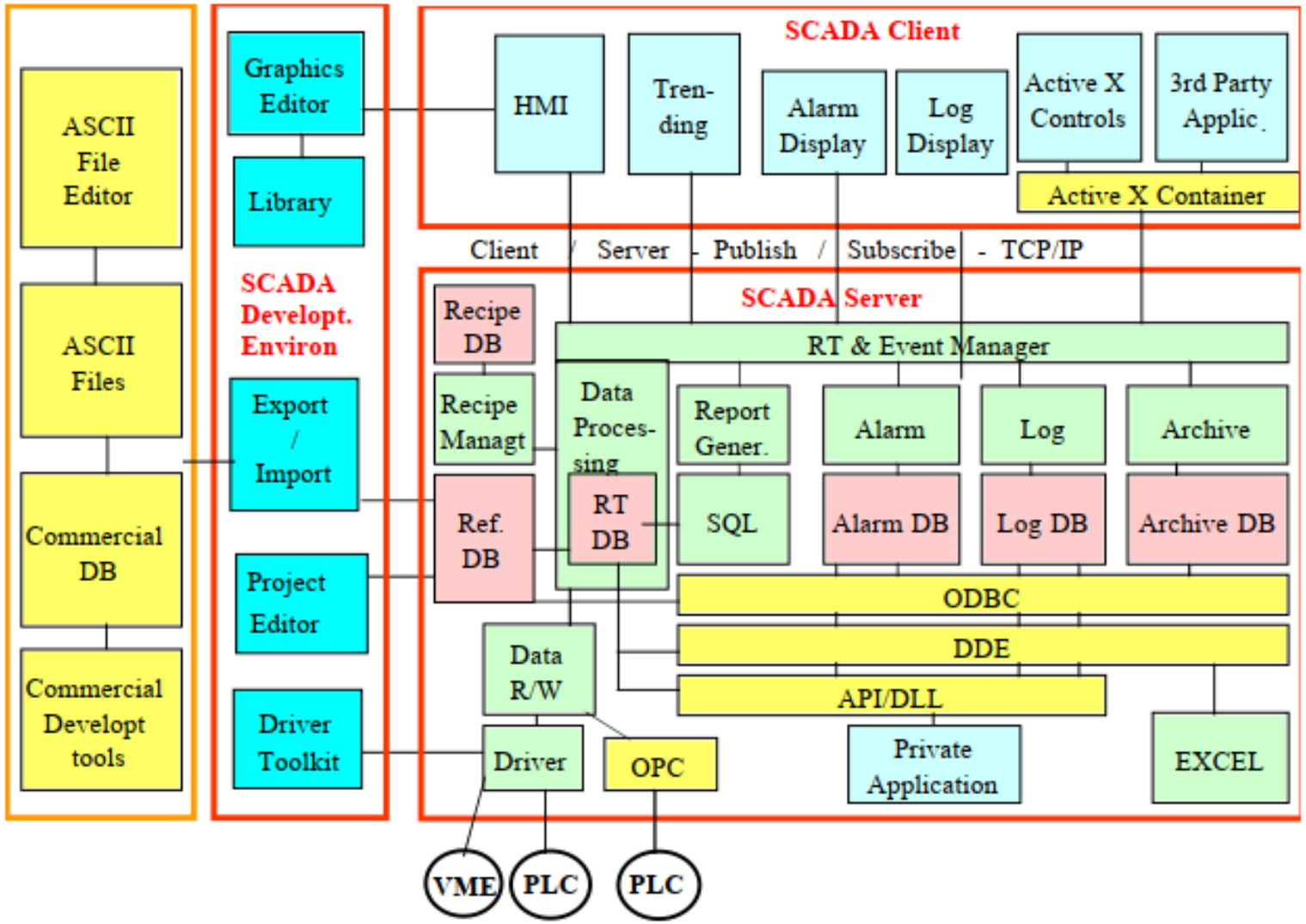


Typical architecture of a SCADA system

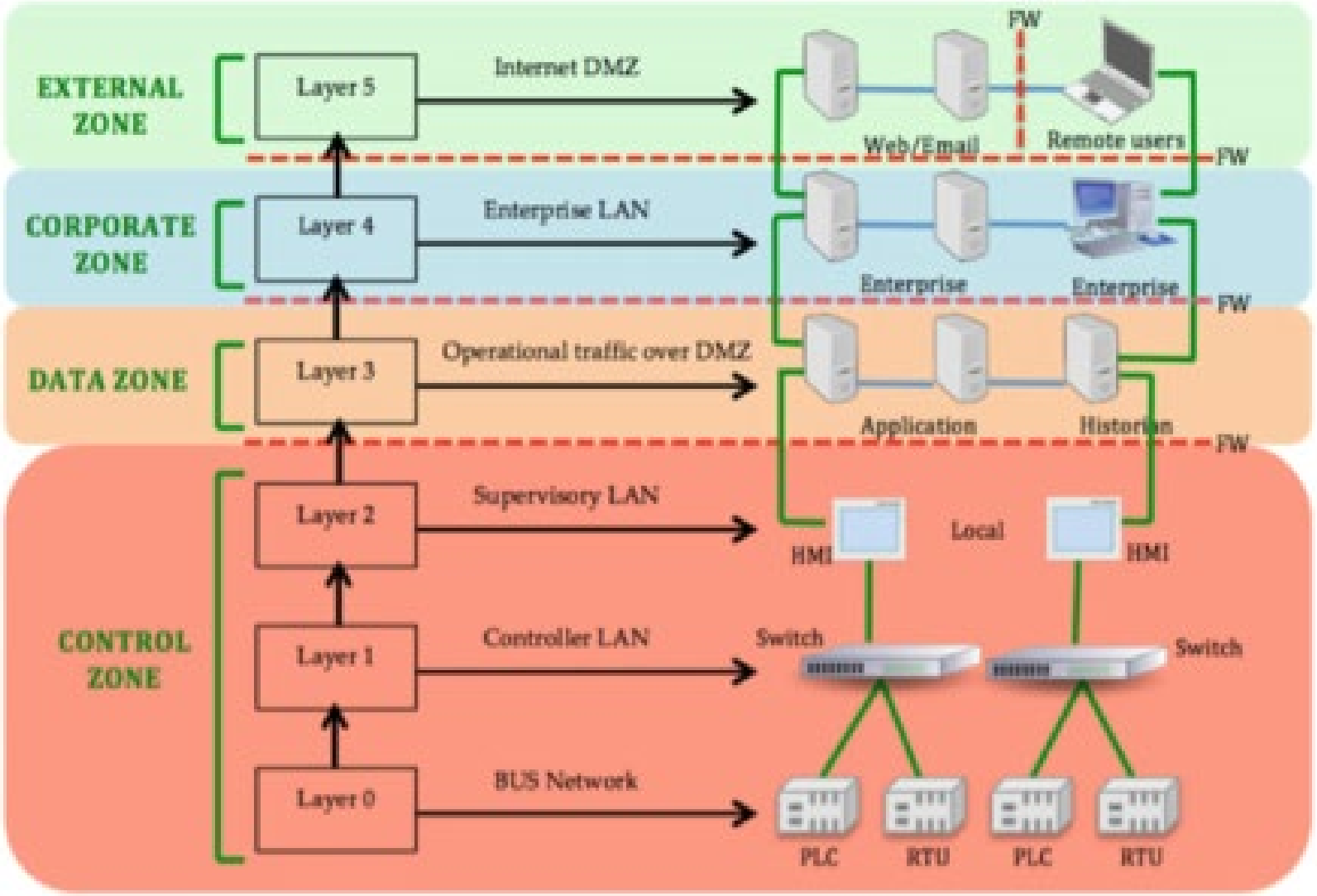


A typical SCADA Architecture in a simplified logical view.

SCADA Generic Software Architecture



SCADA communications zones



SCADA Characteristics

There are numerous SCADA systems available (both open-source and commercial), each with different features, architectures, hardware and standard support. Both open-source and commercial SCADA systems provide similar functionality and requirements.

Current SCADA systems provide users with the following functionalities:

- Connectivity to physical devices;
- Monitoring and Analytics;
- Data Storage.

SCADA systems must be designed and implemented to support the following non-functional requirements:

- Real-time;
- High availability;
- Reliability;
- High Security.

SCADA and Big Data

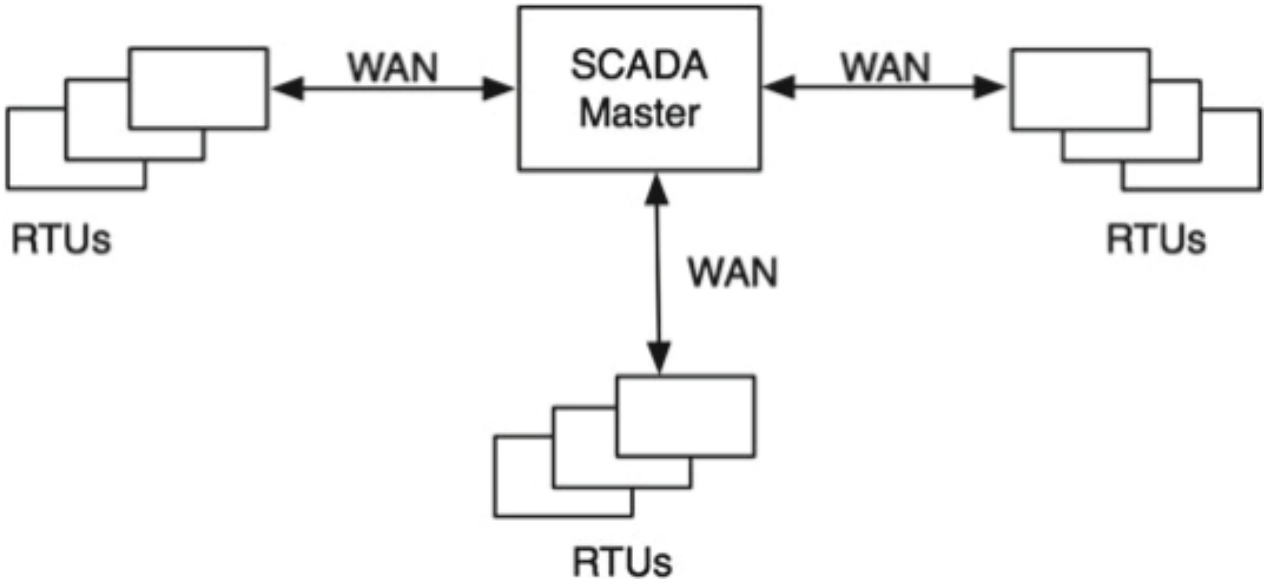
During operation, SCADA systems generates big data which need to be stored and analyzed. A typical SCADA system consists of 100,000 sensors polling at a rate of 1 s.

Assuming a data size of 8 bytes, over the course of an hour, this SCADA system will generate 2.8 GB of data. Some SCADA systems can generate even larger amounts of data, with large numbers of sensors polling at rates of 0.01 s, and even shorter periods.

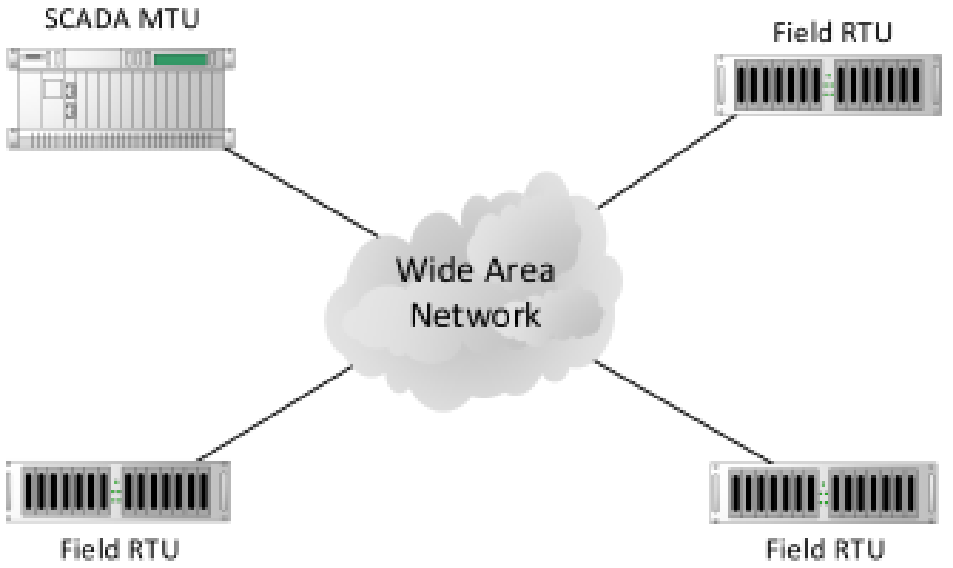
Big data is a paradigm applied to datasets of these large sizes. Therefore, there is a need for new software tools and infrastructure to collect, store, manage, and process these datasets in an acceptable time period. These datasets are of the number of V type, like

When using SCADA based big data there is a need for analytics to deal with: duplicate, erroneous, and inconsistent data, data trustworthiness (e.g., integrity, reputable source), and provide data validation. Clouds can be used to support big data storage and analytics.

SCADA architecture evolution

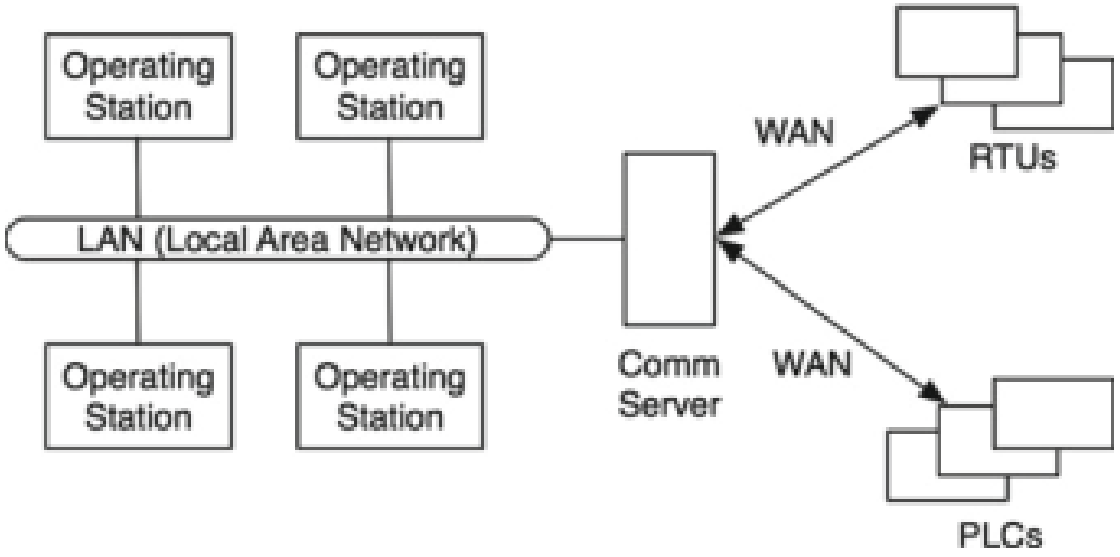


Monolithic SCADA system

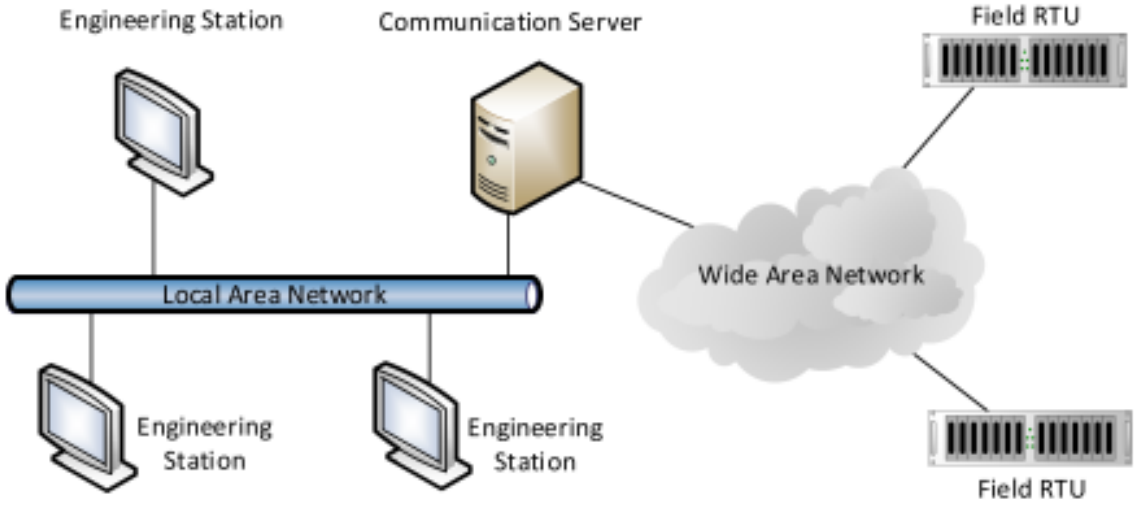


The first generation of SCADA architecture.

SCADA architecture evolution (2)

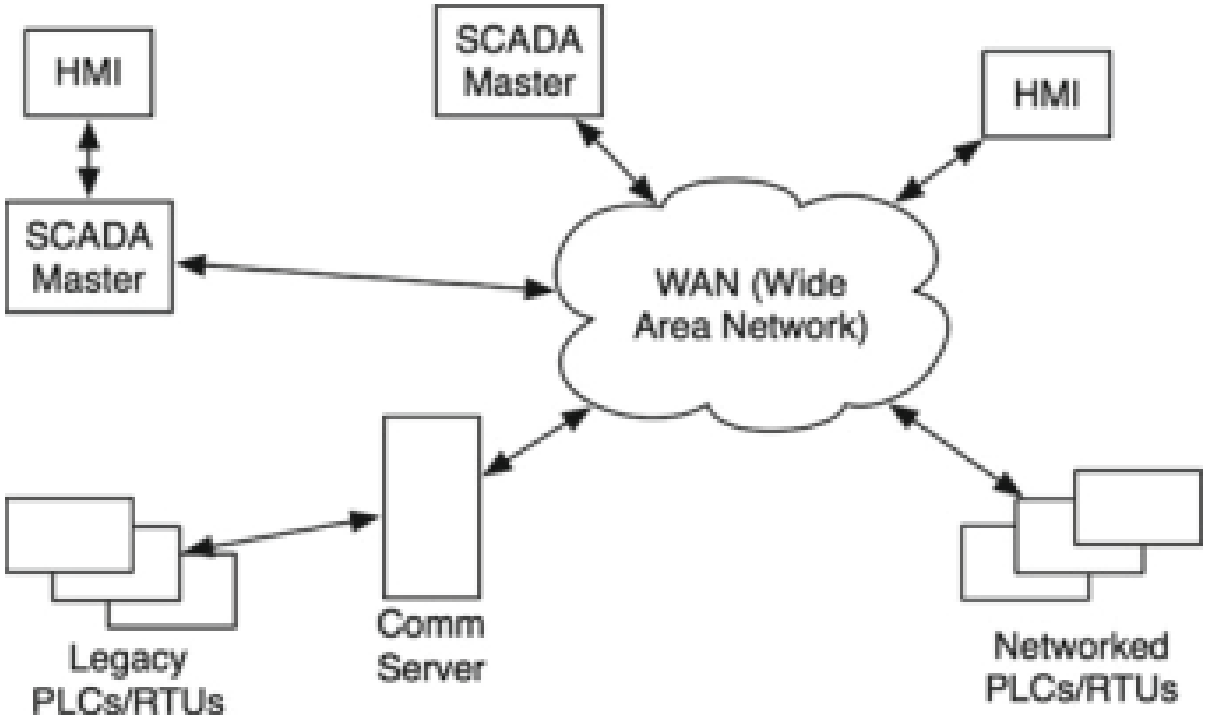


Distributed SCADA system

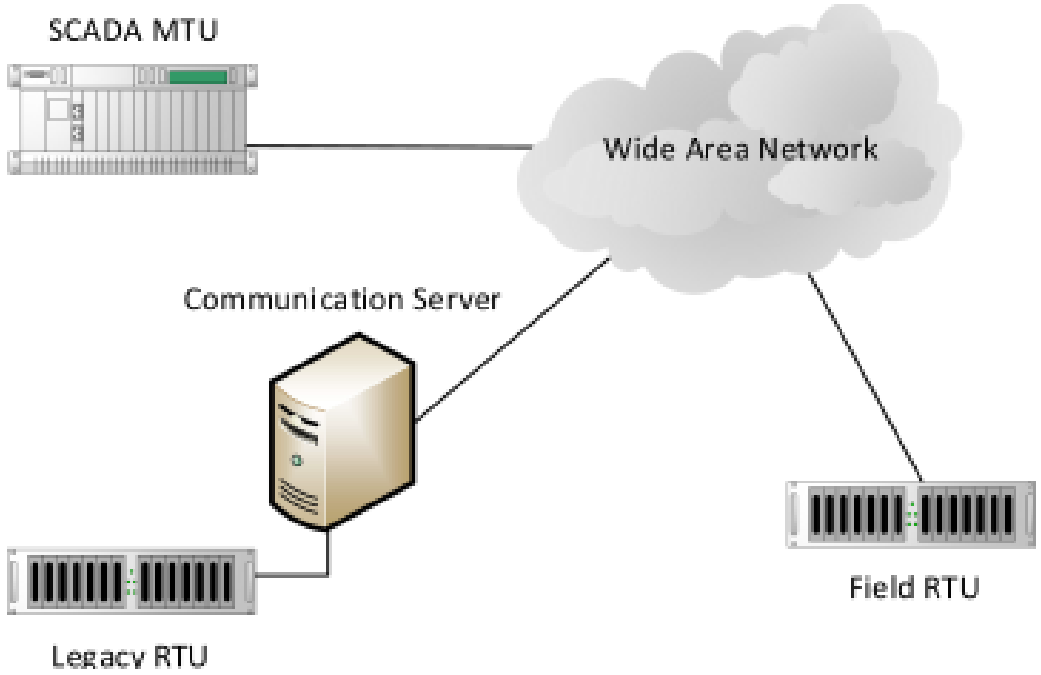


The second generation of SCADA architecture.

SCADA architecture evolution (3)

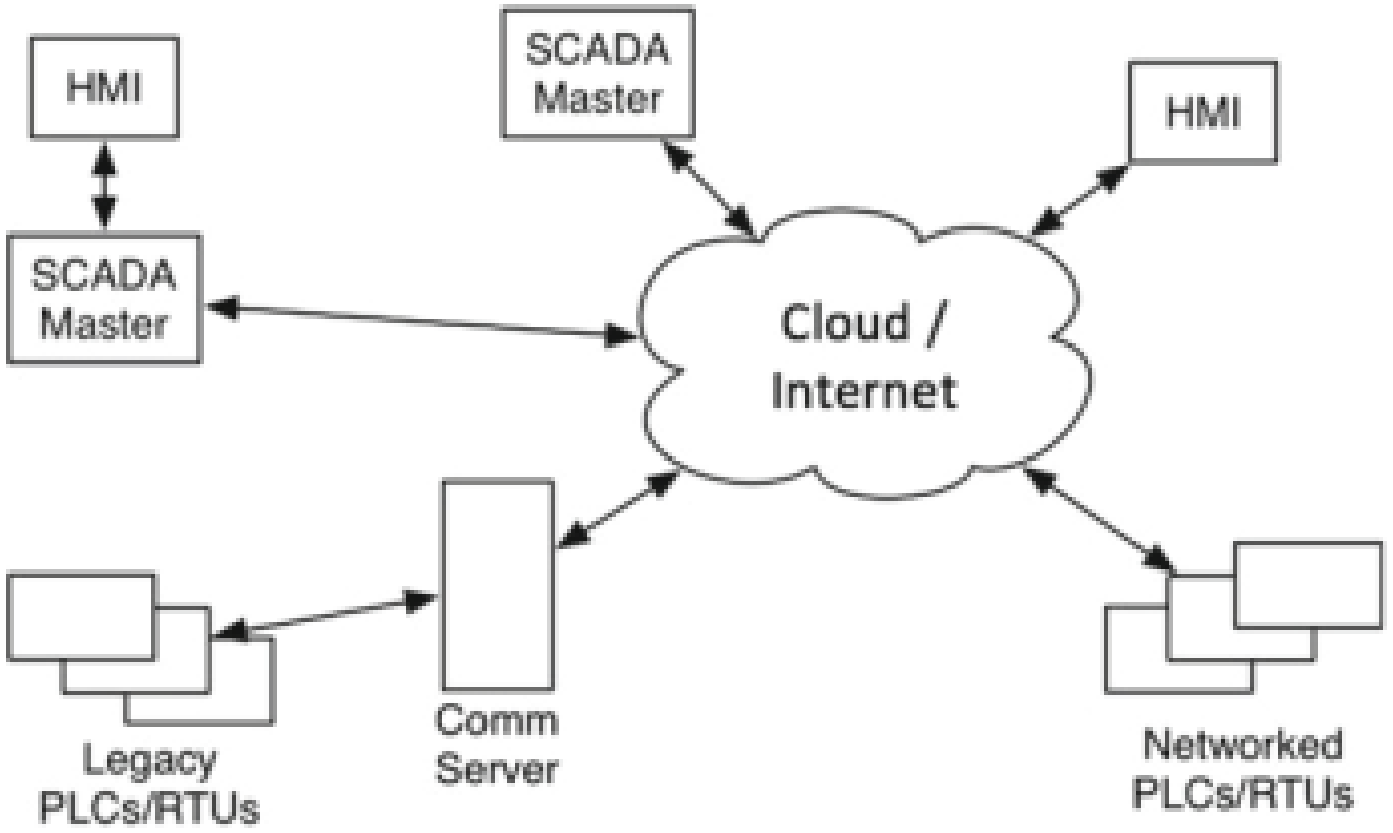


Networked SCADA system



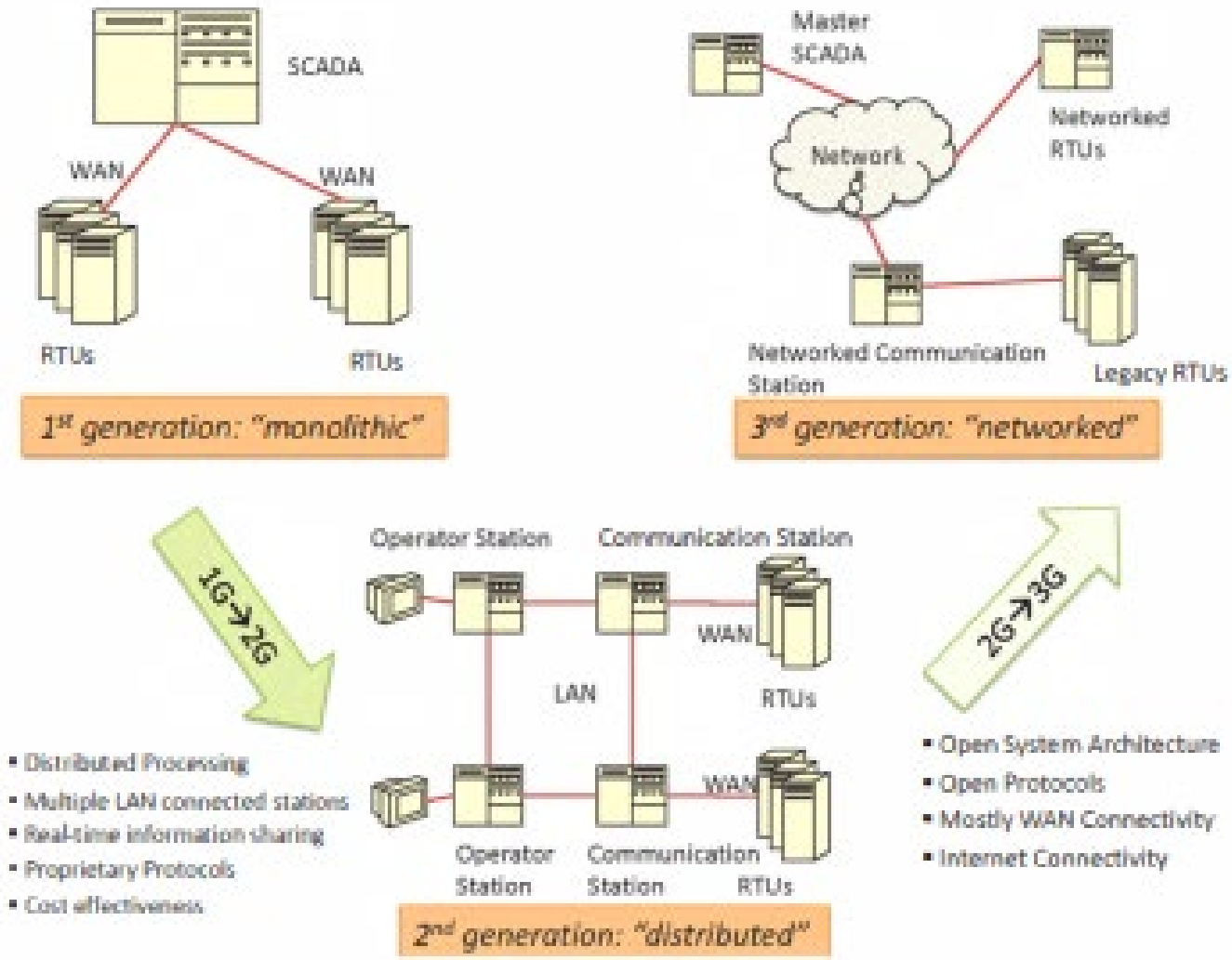
The third generation of SCADA architecture

SCADA architecture evolution (4)



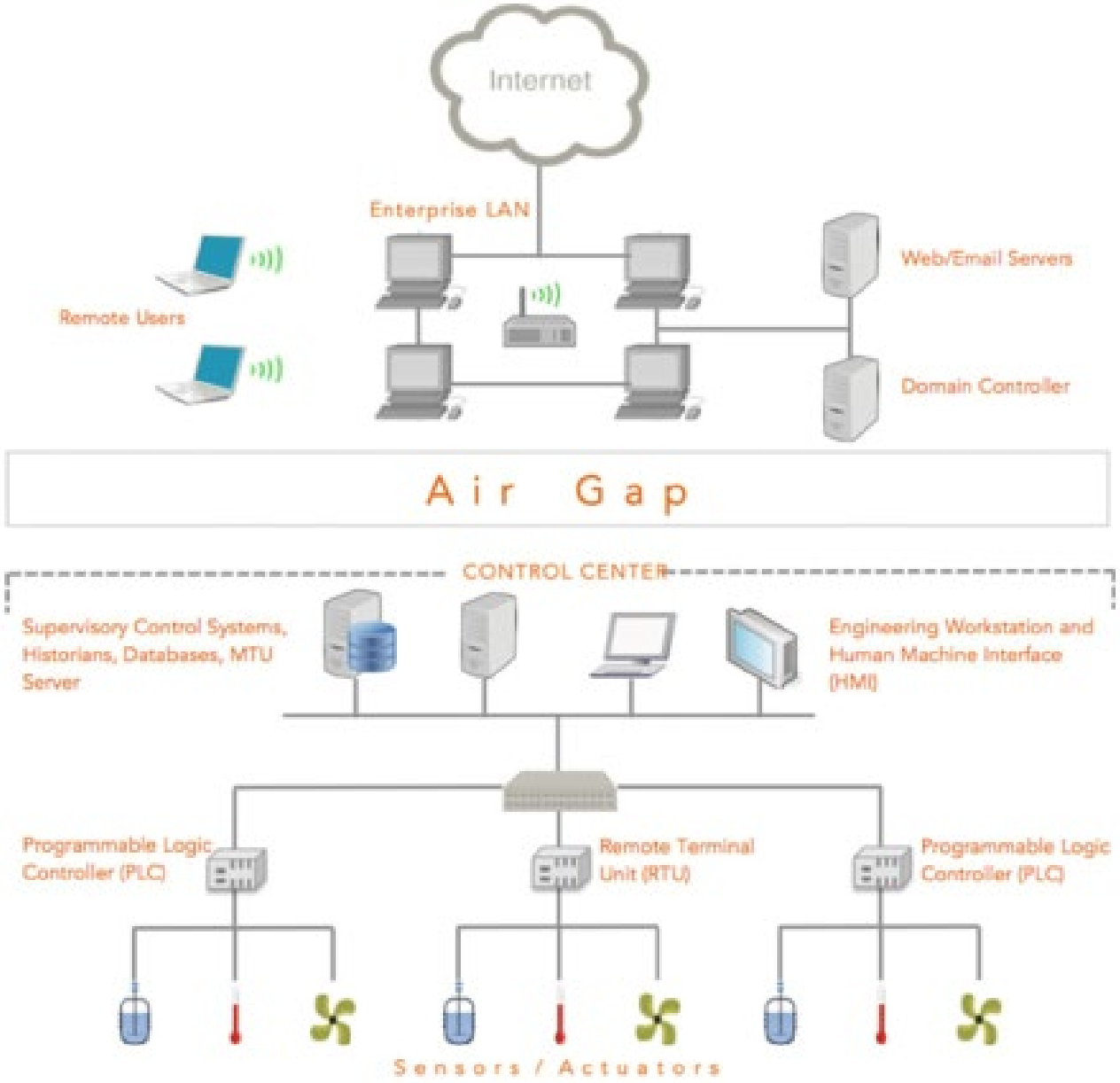
SCADA system operating over IoT

SCADA architecture evolution (5)



SCADA system evolution

Modern Conceptual Architecture of a SCADA System

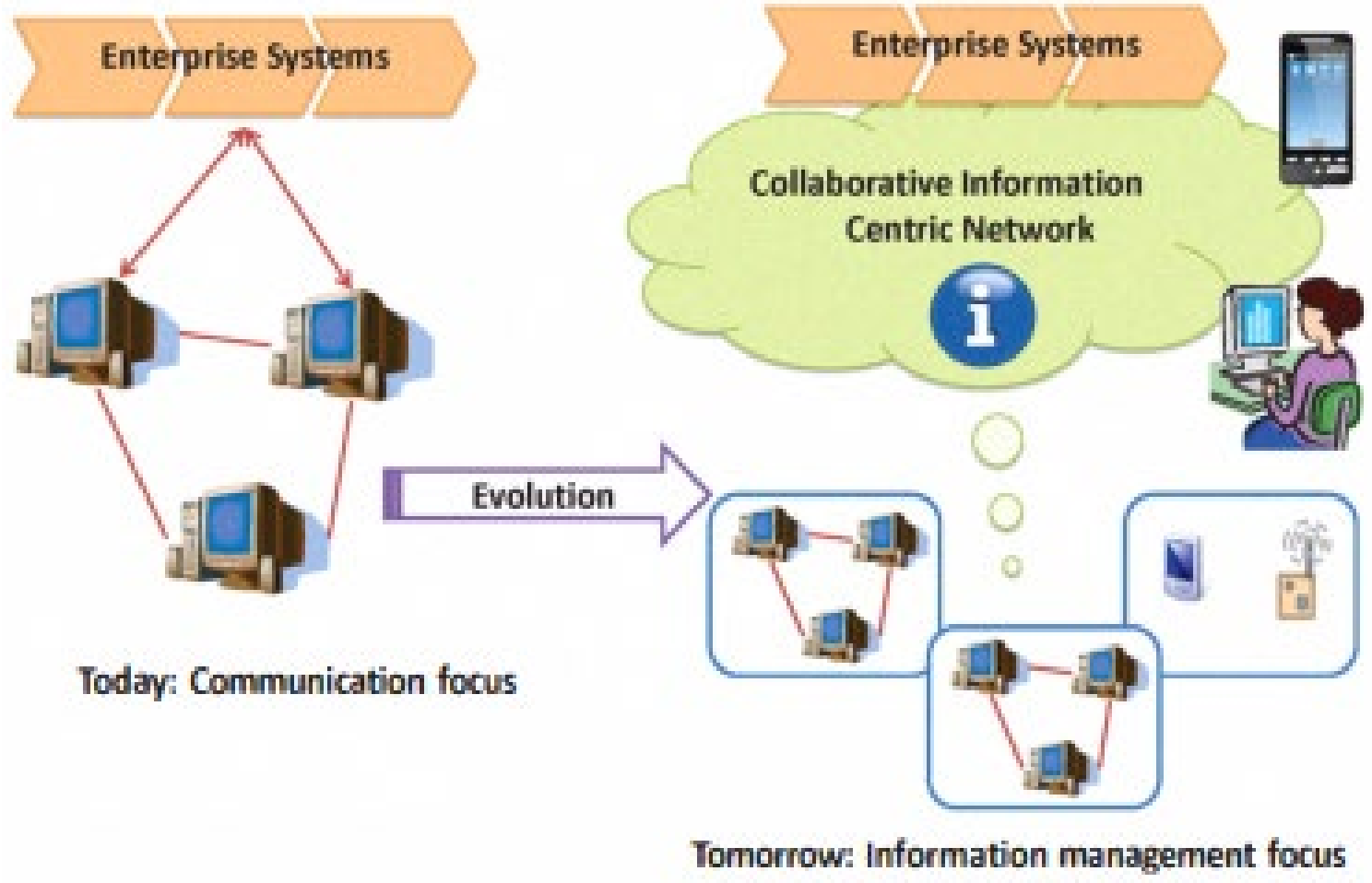


Conceptual architecture of a typical SCADA system

Modern SCADA systems comprise of a series of vital components, both hardware and software, that allow operations to be carried out successfully.

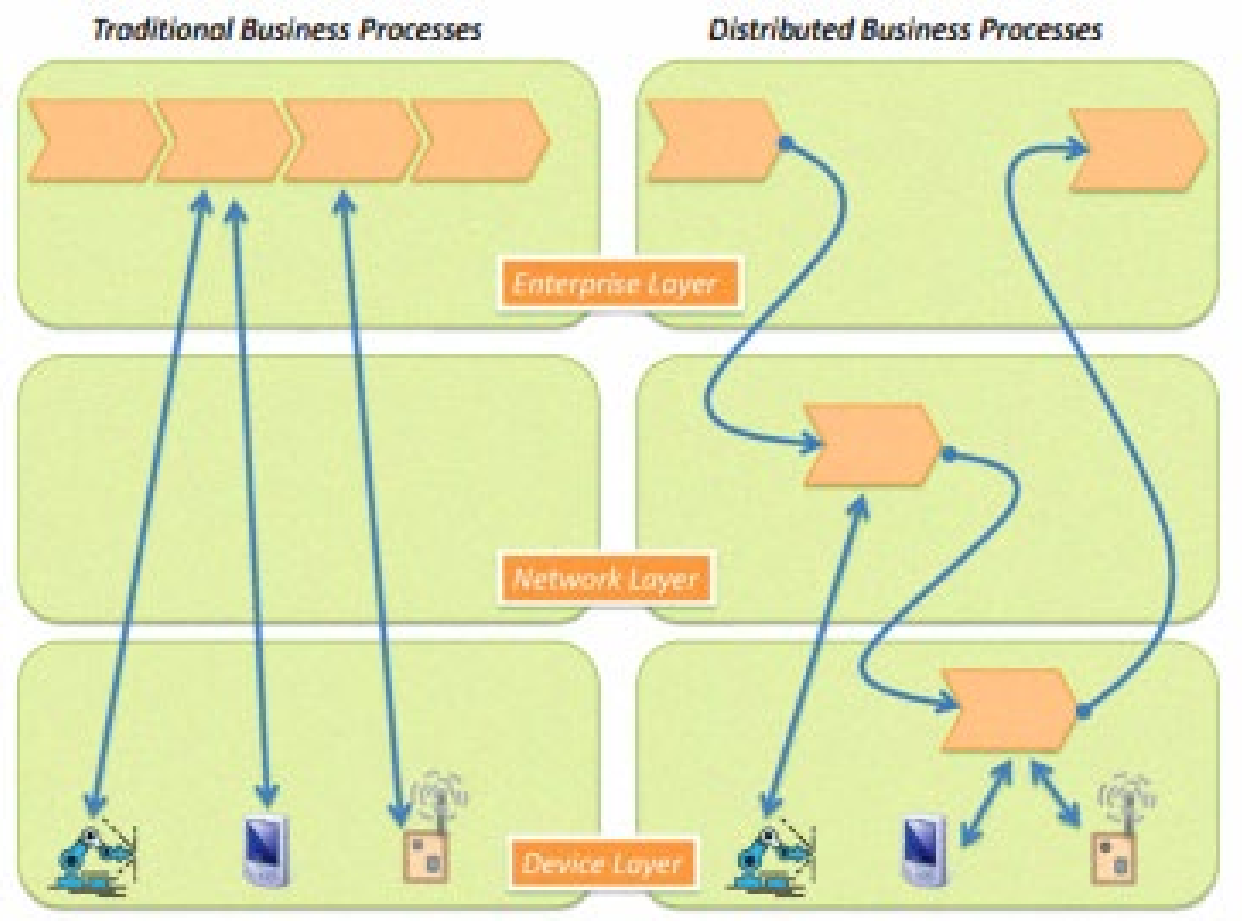
These components can be divided into two main sections within a SCADA system; the control Centre; and the field sites.

modern trend SCADA system view



IT trend: information driven interaction

SCADA architecture evolution



Distributed Business Processes on-device and in-network

Thank you for attention!