Abstract— Climate Change is an extraordinary challenge for the development of our socioeconomic environment. The compilation of comprehensive knowledge about our environment is a key importance for implementing of mitigation strategies. Long-term terrestrial observatories are supporting the systematic monitoring of environmental parameters. They are responsible for data collection, data analysis and subsequently for decision making. Not only the complex structure and the large volume of data streams but also the necessary integration of existing monitoring infrastructures for such observatories imply special technological challenges for today’s scientific data and information management. Recent developments of Information and Communication Technology provide important conceptual and technological input for the proper design and implementation of underlying monitoring and data infrastructures. To avoid constantly recurring system developments for such infrastructures, a general and integrated approach for a reference architecture concept is needed.

Keywords— monitoring, infrastructure, information system, reference architecture, environmental observatory

I. INTRODUCTION

The multiple processes of global change, and in this case climate change, are among the biggest challenges for mankind today. Various scenarios of the impacts of these changes were described in the last IPCC report, for example [5]. It is widely acknowledged in the scientific community that only comprehensive knowledge of all environmental systems and their processes constitutes a significant progress to counter these challenges. The focus hereby should not only consider short-term applications, but long-term environmental processes must also be taken into account. Only this broadening of the focus can lead to a significant positive impact on human society in the course of the next decades [2]. Therefore in the last decade many interdisciplinary long-term monitoring networks were established globally in terrestrial environmental sciences.

Such long-term environmental monitoring networks consist of single sensors or sensor networks to record a variety of environmental parameters from hydrology, pedology, climate, lower atmosphere, or others. The underlying monitoring and data infrastructures have to deal with various kinds of data management tasks from data acquisition, through data preprocessing to data visualization and data post processing. This has only been possible – in its whole characteristics - over the last ten years by the rapid development of Information and Communication Technology (ICT) [18]. ICT has revolutionized the entire process of data and information management, in particular the acquisition, processing, analysis and visualization of data as well as the integration of computer systems into such infrastructures [18]. It had immediate impact on the availability of sensor information, now almost in real time, on the processing power and speed. This technological progress opens new potentials, but also the need to an enhanced development of system architectures for monitoring and data infrastructures.

As one example the Terrestrial Environmental Observatories (TERENO) uses this potential for an interdisciplinary and long-term research project over four test sites within Germany supported by six research centers of the Helmholtz Association. Each contribution by a participating research center is called an observatory in organizing the monitoring activities. The single observatories are responsible to develop their solutions for monitoring and data infrastructure independently, while being interoperable with the TERENO data portal TEODOOR hosted by the Research Centre Jülich [7]. The challenge for TERENO and its observatories is to integrate the whole aspects of data management, data workflows, data modeling and visualizations in a proper design of a monitoring infrastructure. For TERENO Northeast, one of the sub-observatories within TERENO, we will develop existing infrastructure concepts further towards a general and integrated architecture concept. This concept can serve as a reference architecture model for TERENO Northeast and for future monitoring and data infrastructures.

In our contribution we analyse first the architectures of existing environmental monitoring systems and sensor networks, and further architecture approaches of a state-of-the-art Early Warning System (EWS). In section 3, we discuss the specific situation in the subproject TERENO Northeast. This discussion will be followed by highlighting the used and standardized interfaces and services with a particular reference to the sensor integration platform called Sensor Service Bus (SSB). Section 3 ends with explaining the initial steps to a general architecture model for an environmental monitoring...
and data infrastructure in TERENO Northeast under considering the existing concepts of TERENOs TEODOOR system architecture. Section 4 summarizes this contribution with a perspective to future developments.

II. MONITORING SYSTEMS AND ITS ARCHITECTURES

A. Recent Architectures of Environmental Sensing and Sensor Networks

Recent developments and implementations of adapted monitoring and data infrastructures were investigated by [1, 2, 3, 6, 10, and 12]. Two recent projects with participation of the German Research Centre for Geosciences (GFZ) in relation to its software architectures will be shortly explained here:

1.) CAWa

The project Central Asian Water (CAWa) is a regional monitoring network in Central Asia (Kazakhstan, Kyrgyzstan, Tajikistan, Uzbekistan, and Afghanistan) collecting observations of meteorological and hydrological parameters and delivering these to its regional users. Within CAWa a long-term monitoring of water resources and climate parameters is established at the scale of river basins in a network of continuously operated in-situ stations. The German Research Centre for Geosciences (GFZ) and Central Asian Institute for Applied Geosciences (CAIAG) coordinate the activities in cooperation with National Hydrometeorological Services (NHMS) of the Central Asian states. CAWa’s core component in their data infrastructure is the system operation, processing and archiving facility (SOPAF) which bases on the sensor web enabled (SWE) observation and measurement specification in combination with the SOS standard data model [15]. Several Java-based retrievers work in parallel and send the station data to the dispatcher which translates and writes the information into the SOPAF data model (Fig. 1)[13]. The ROMPS (Remotely Operated Multi-Parameter Stations) is an open station concept which provides a platform to add additional sensors, when necessary [13].

Figure 1. Architectural schema of CAWa infrastructure from [13]

2.) SuMaRiO

The SuMaRiO (Sustainable Management of River Oases along the Tarim River) project is supporting oasis management along the Tarim River (West-China, Provence Xinjiang) under conditions of climatic and societal changes. Within the project the ecosystem services (ESS) of land and water management in the Tarim River Basin will be investigated by monitoring of various environmental parameters and subsequent model derivations. The infrastructure will combine monitoring data with a spatial decision support system (SDSS) to a scientific spatial data infrastructure (SSDI) based on service oriented architectures (SOA) principles. The schema of Fig. 2 gives a rough overview of this recently in planning infrastructure [14].

B. Similar Approaches in other Systems

Early warning systems (EWS) are specific applications of monitoring infrastructures. In recent years, a number of such systems, in particular for natural disasters like floods, storms, and tsunamis, have been established. Each of these systems consists also of a monitoring infrastructure. While many of these systems have always been rather particular and individually customized, the architecture approach of such EWS systems and its underlying infrastructures in [10] is more systematic. EWS can be subdivided into three main components, where the monitoring infrastructure is represented as a separate subsystem. The main components can be described using the structural architecture description concept of FMC (Fundamental Modeling Concept) [8] and subsequently combined to an overall reference architecture. According to [10] the term reference architecture can be seen as a sum of system design decisions that is also applicable in other related systems. Reference architectures include the system structure, the functional behavior, the interactions as well as nonfunctional requirements. See figure 3 for a FMC model of the monitoring infrastructure in a EWS after [10].

Figure 2. SSDI architecture based on SOA principles (according to [19])

Figure 3. Subsystem monitoring system in an EWS reference architecture (according to [10])
III. GENERAL ARCHITECTURE CONCEPT FOR TERENO NORTHEAST

A. Test Sites Northeast Lowland

The German Northeast Lowland region has been shaped by recurring glacial and periglacial processes for at least half a million years. Three major glaciations covered the entire region, the last glaciation occurring approximately 25 – 15 ka ago (Weichselian glaciation). Many lakes and river systems which are connected to the shallow ground water table mark this region. Land use types are arable land, planted pine forests, deciduous forests, and wetlands of high ecological value. Low annual amounts of precipitation in connection with major reconfigurations of the hydrologic system (d.amming and drainage) in the past make this region highly sensitive to the effects of climate change. The Northeastern Lowland Observatory provides unique natural laboratories to investigate relatively young country surfaces starting at the “zero-point” of landscape evolution. Within TERENO the sensitivity of the landscapes can be compared with that of older landscapes that have not been affected by glaciers [17].

The four test areas of TERENO Northeast Lowland differ significantly in intensity and history of their land use practices.

1) Müritz National Park represents a protected quasi-natural site;
2) Schorfheide-Chorin Biosphere Reserve is a well-controlled cultivated landscape;
3) The catchment of the river Uecker; and

In the four representative regions, areas are furnished to measure parameters such as soil moisture, ground water levels, precipitation and tension. Eddy-Covariance stations measure a number of climate parameters and gas concentrations. SoilNet observation networks collect ground temperature, soil moisture and other climate parameters. Data from tree cuts and core drilling also feed into the project TERENO.

B. Standardized Interfaces and Sensor Service Bus

An essential aspect in TERENO is the use of web services, which were standardized by the Open Geospatial Consortium (OGC)[12]. Web services supply standards and to query and access spatial data. With the use standards data can be transferred and integrated by other data holders without significant investments into the design of interfaces and data harmonization.

A uniform treatment of sensor data can be realized by the OGC Sensor Web Enablement (SWE) [16], which makes a number of standards and interface definitions available: Observation & Measurement (O&M) model for the description of observations and measurements, Sensor Model Language (SensorML) for the description of sensor systems, Sensor Observation Service (SOS) [15] for obtaining sensor observations, Sensor Planning Service (SPS) for tasking sensors, Web Notification Service (WNS) for asynchronous dialogues and Sensor Alert Service (SAS) for sending alerts.

In order to integrate all these services to our infrastructure we used the enterprise service bus (ESB) technology as integration platform. We call it Sensor Service Bus (SSB) which was initially developed as Tsunami Service Bus (TSB) in the German Indonesian Tsunami Early Warning System (GITEWS) project, where it has been deployed with great success [4]. TSB is the integration platform in GITEWS. The functional integration relies both on a Service Oriented Architecture (SOA) as well as on an integration concept, built on standardized encodings and protocols provided by Sensor Web Enablement (SWE) [4]. The following functional requirements were implemented [4]:

- Standardized Interface for fast accessing sensor data and for sensor tasking.
- Storage of all incoming sensor data, including post processing and quality check.
- Flexible integration mechanism of new or no foreseen sensor types.
- Operating time 24 h a day, 7 days a week.

To provide a flexible architecture in TERENO Northeast for integration sensor services, new sensors and sensor data the TSB was adapted to the SSB. The SSB will be a main interface in TERENO Northeast. The SSB is divided into five components (according to TSB):

- “The Processing component receives incoming data as messages. The data is analyzed, processed, and stored into the database. Registered applications are then informed about the new available data or about system alerts (by WNS resp. SAS).
- The Provisioning component provides access to all sensor data for client applications in terms of the SOS interface.
- The Tasking component enables the uniform control of sensors. The Tasking component forwards requests to the sensor specific command adapter (via Sensor Manager Interface) and forwards the (asynchronous) results back to the client application.

![Figure 4. Overview of test sites of TERENO Northeast](image-url)
The Registry is the central provider for all sensor metadata. It also provides functions for the management of sensor metadata (create, modify, delete). The Registry stores its metadata into the Database.

The Database acts as a general storage for all sensor data and metadata.\textsuperscript{[4]}

The success of long-term projects like TERENO depends on a well-organized data management, data exchange between the involved part projects, and on the availability of the captured data. To allow this, a spatial data infrastructure TEOODOR was built as part of TERENO\textsuperscript{[9]}. TEOODOR bundles the data, provided by the different web services of the single observatories, providing tools for querying, visualizing and downloading data to make these available to the scientific community\textsuperscript{[9]}. Within the TERENO initiative the Northeastern Lowland of Germany is one part of four long-term terrestrial observatories, which collect huge amounts of environmental relevant data. To order, describe, exchange and publish the data for internal and external use TERENO Northeast will be create a Spatial Data Infrastructure. All observatories in TERENO using standardized interfaces to support data transfer and interoperability. Based on the experiences within the projects at the GFZ described above (2.1 & 2.2), in the following the initial step towards a reference architecture for TERENO Northeast will be developed. With this general concept similar projects can be developed as modular system in a much faster way.

C. General and Integrated Architecture Model

As outlined above, the architecture concepts for monitoring infrastructures has approaches quite similar to EWS. The basic components connect to sensor agent software (integration platform) together with different kind of web-services, even though the application context might be quite different. It may therefore be a good approach to foster the development of an integrated architecture model which can be also used as general reference model, analogous to EWS\textsuperscript{[10]}. To communicate these concepts, tools like FMC have proven to be useful, as they show sufficient detail while at the same time allowing a high level of abstraction. Additionally, few import design and engineering principles for system architectures should be considered. As published in [18], a step-wise refinement of design artifacts distinguished into analysis, abstract design, concrete design and engineering step, is essential. Keeping the afore mentioned in mind, we developed a general and integrated architecture model for TERENO Northeast as general reference model. On this abstract level of system design it is now possible to integrate additional sensors or sensor systems and external resources to our system, when necessary. In addition, the reference model is helpful to communicate in detail about system requirements with different stakeholders of the project.

An initial design may look like this: The measurement information is provided by physical sensors, sensor systems or by other monitoring platforms. Sensors trace data for phenomena of processes connected to an object of interest. In TERENO Northeast context this can be soil moisture, ground water levels, precipitation, or other parameters.

The Sensor Management component is responsible for sensor communication, sensor control and dynamically adding or removing of sensors. From raw data to the data storage requires processing steps, for example quality checks to remove spikes that are not physically possible for a given parameter, or to remove values that might be physically possible but are nonsensical for the specific monitoring location. The SSB will be used for processing of incoming data, data analysis and storage in the database. The monitoring data in the database will then be published via OGC Services\textsuperscript{[12]} like Web Feature Services (WFS), Web Map Services (WMS), Web Catalogue Services (CSW), and Sensor Observation Services (SOS). These interfaces are standard services for spatial data search and access which facilitate the internal and external exchange with other observatories, data infrastructures, web portals or information systems.

![Monitoring infrastructure design as reference architecture model](image-url)

Figure 5. Monitoring infrastructure design as reference architecture model
IV. PERSPECTIVES

Within TERENO large amounts of environmental relevant parameters will be collected by four terrestrial environmental observatories. The observatories are the data originators, their respective operating institutes will be independently responsible for the construction of their data infrastructures and monitoring. The data of all observatories are published through OGC web services in the main data portal TEODOOR [9].

TERENO Northeast data infrastructure has the objective to build an integrated and general reference architecture as reference model for its monitoring and data infrastructure. This will allow the transfer of the reference architecture to future monitoring projects within the German Research Centre for Geosciences (GFZ), and other research centers. An initial step in developing such reference architecture could be shown here. Next steps will be the refinement of single design steps, a clear definition of an underlying information model, and finally the implementation within TERENO Northeast.

REFERENCES


