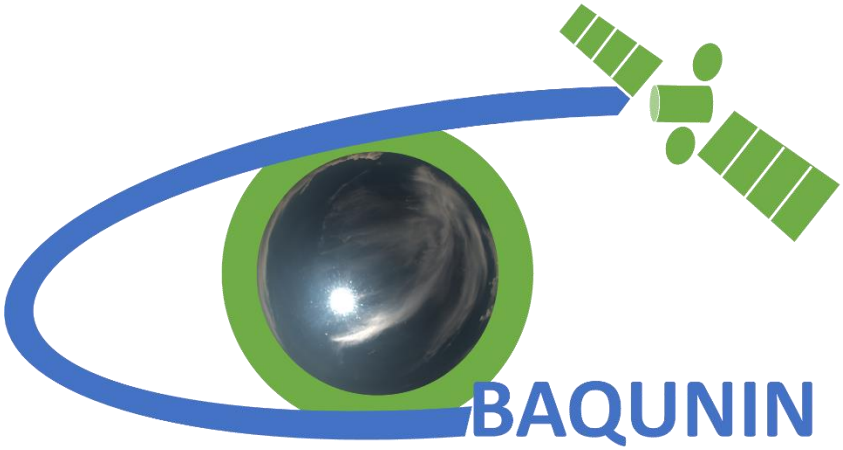
Customer	:	**ESRIN**	**Document Ref**	:	**BAQ-TEC-TEN-SER-023**
Contract No	:	**4000126749/19/I-NS**	**Issue Date**	:	**14 February 2020**
WP No	:	**All**	**Issue**	:	**1.0**



Title : **New Baqunin IT Architecture**

Author : Erminia De Grandis

Approval : Massimo Cardaci

Distribution : BAQUNIN project team

Serco Italia SpA
Sede Operativa: Via Sciadonna, 24-26 - Frascati (Roma)
Tel: +39 06 98354400 Fax:
www.serco.com

TABLE OF CONTENTS

Change History

This document shall be amended by releasing a new edition of the document in its entirety. The Amendment Record Sheet below records the history and issue status of this document.

ISSUE	DATE	REASON
1.0	14 Feb 2020	First version

1. INTRODUCTION

1.1 Purpose of the document

The purpose of this document is to define the system-level architectural design of the BAQUNIN infrastructure. An overview of the operational chain will be given with a detailed description of all involved parties, system interfaces, data flow, and protocols.

1.2 Background

The ESA Sensor Performance, Products and Algorithms (SPPA) section invested resources in the development and operation of ground-based instruments devoted to satellite atmospheric Cal/Val activities.

Thanks to the coordination of a team composed by SERCO, “Sapienza” University, CNR-ISAC and CNR-IIA researchers and remote sensing specialists, a joint instrumental super-site has been set up in the area of Rome called BAQUNIN (Boundary-layer Air Quality-analysis Using Network of INstruments)

The BAQUNIN Super-Site includes ground-based active and passive remote sensing instruments operating in synergy, offering quantitative and qualitative information on a wide range of atmospheric parameters. Therefore, a significant contribution is given to the atmospheric chemistry validation activities and the Planetary Boundary Layer (PBL) studies.

The instruments are located in three different sites: an urban component (Physics Department of ”Sapienza” University, Rome), a semi-rural component (CNR-ISAC, Tor Vergata) and a rural component (CNR-IIA, Montelibretti), each hosting at least one BAQUNIN Pandora instrument and, as in the case of Sapienza, a large number of other atmospheric remote sensing devices. Some of which are included in international networks such as AERONET, EUROSKYRAD, EUBREWNET, and PGN.

ACRONYMS

Acronym	Definition
APL	Atmospheric Physics Laboratory (at Sapienza)
BAQUNIN	Boundary-layer Air Quality-analysis Using Network of INstruments
AERONET	Aerosol Robotic Network
EUROSKYRAD	European Skynet Radiometers Network
EUBREWNET	European Brewer Network
PGN	Pandonia Global Network
CNR-ISAC	Istituto di Scienze dell'Atmosfera
CNR-IIA	Istituto sull'Inquinamento Atmosferico
WRF	Weather Research and Forecasting Model
LIDAR	Light Detection And Ranging
SODAR	Sonic Detection And Ranging
MFRSR	Multifilter Rotating Shadowband Radiometer
EVDC	ESA Atmospheric Validation Data Centre
DIAS	Data and Information Access Service
GEOMS	Generic Earth Observation Metadata Standard

2. CURRENT BAQUNIN INFRASTRUCTURE

In this paragraph, a high-level architecture of current BAQUNIN infrastructure is introduced and the main system components, their interfaces and protocols are described.

2.1 System Level Architecture

Each BAQUNIN acquisition station, hosting a group of instruments, works independently following different steps before to disseminate the products. For this reason, the acquisition frame was divided, as schematised in Fig. 1, into different groups according to the different scenarios.

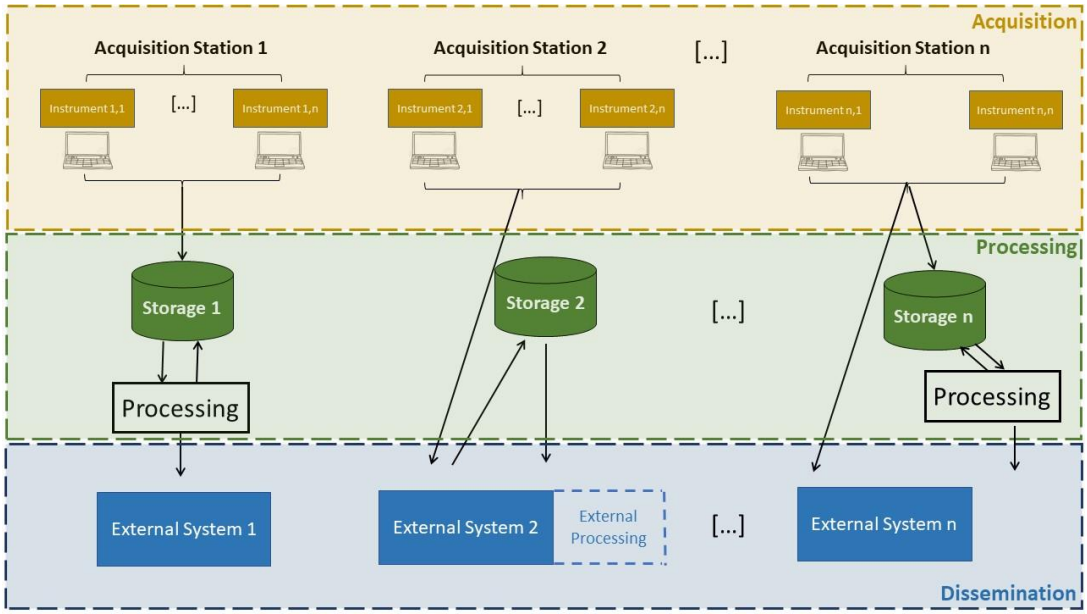


Figure 1: Current BAQUNIN architecture

In detail three main data flows are identified:

1. the acquired data are archived and sent to the processing facility where will be elaborated according to the requirements of external system where will be disseminated (format change, Level1 and Level2)
2. the acquired data are sent directly to the external system where will be processed, then sent to the acquisition station to be archived and disseminated.
3. the acquired data are disseminated and archived at the same time. An additional processing step can also take place before the dissemination.

Similarly, the block diagram in Fig. 2 represents the same scenarios in terms of consecutive steps within the BAQUNIN processes

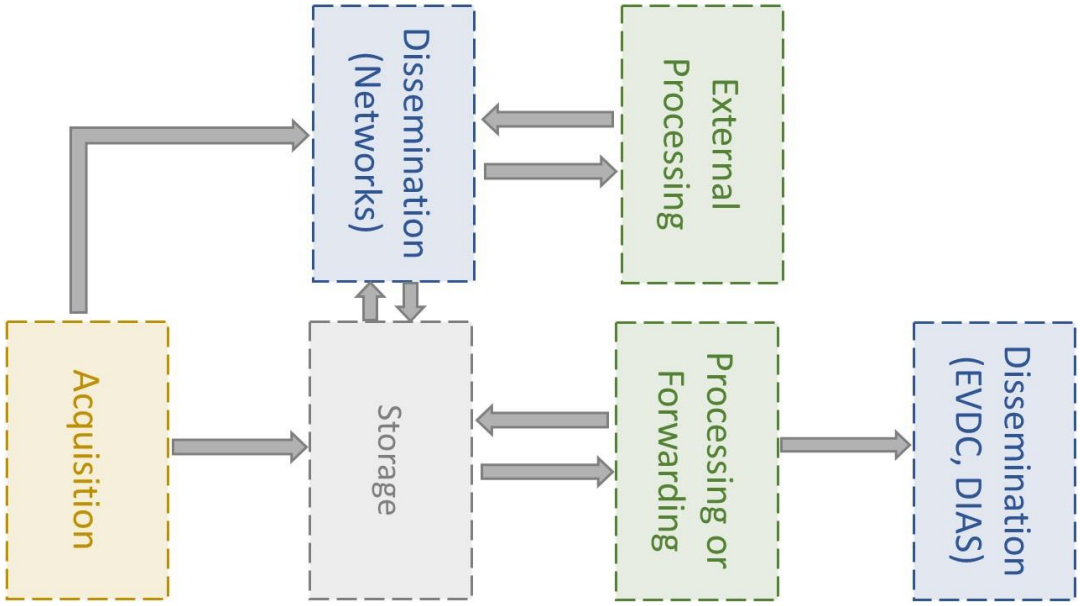


Figure 2: Block Diagram of BAQUNIN processes

The acquired data are archived or directly disseminated. Within the dissemination facility the data can be processed and then archived into the BAQUNIN storage. Finally, the archived data can be forwarded to the dissemination facility (data previously processed within the networks) or processed before to be disseminated.

2.2 System components

The acquisition stations included in the BAQUNIN operational chain are:

- *Physics Department of Sapienza University Rome*: urban component hosting most of the instruments of the BAQUNIN framework
- *CNR-ISAC Tor Vergata, Rome*: semi-rural component hosting the Pan155 pandora instrument
- *CNR-IIA, Montelibretti*: rural component hosting the Pan 138 pandora instrument
- *ESA/Esrin, Frascati* which provides real-time simulations based on actual or idealised atmospheric conditions produced by the WRF model, a mesoscale weather prediction system designed for atmospheric research and operational forecasting applications.

The instrumental set-up which composes the BAQUNIN framework is constituted by the following instruments:

- *Raman and elastic LIDAR:* a remote sensing device which uses the light pulses emitted by lasers to observe the atmosphere
- *SODAR:* an active remote sensing instrument based on the interaction of acoustic waves with the atmosphere, capable of retrieving three-dimensional wind filed and thermodynamic structure up to the first 1000m of atmosphere
- *MFRSR:* a radiometer passive instrument measuring global and diffuse components of solar irradiance
- *POM-Prede*: a sun-sky radiometer measuring direct solar and diffuse sky irradiances at several wavelengths and scattering angles
- *Brewer:* a spectrometer measuring the intensity of attenuated solar radiation in the UV and visible region
- *PANDORA:* a dual spectrometer system taking measurement and performing observations pointing anywhere in the sky as direct sun, zenith sky and principal plane
- *CIMEL:* a radiometer measuring direct solar irradiance and performing sky observations
- *All Sky Camera:* a specialized camera capturing a photograph of the entire sky used in meteorology to study cloud cover, current level of UV radiation, fractional cloud coverage, sky polarization, cloud base hight and wind speed at cloud heights
- *Pyranometer:* a radiometer measuring the hemispherical solar radiation
- *Meteorological sensors:* a weather station providing relative humidity, temperature and wind speed

Each instrument is monitored by the acquisition station where the data are collected, archived and sent to the processing and dissemination facilities. Within the acquisition process, a daily checking procedure guarantees consistency of archived data and ensures prompt handling of missing data.

The **processing facility** is a customized environment where a set of tools and libraries are available to analyse and process the data.

The acquired and processed products are disseminated towards the following **external systems**:

- *ESA Atmospheric Validation Data Centre (EVDC):* a central long-term repository, fully compatible with GEOMS, for archiving and exchange of data within validation activities of satellite atmospheric composition products. Powerful tools for extraction, conversion and archival of a large amount of EO data are provided
- *International Networks*: global ground-based remote sensing networks such as AERONET, EUROSKYRAD, EUBREWNET and PGN
- *ONDA Data and Information Access Service (DIAS):* cloud environment offering free solutions to the users for easily accessing, downloading and processing all Copernicus data.
  Currently the BAQUNIN products to be stored into ONDA DIAS are archived in a temporary server located in ESA/ESRIN but they are due to be migrated in the ONDA platform.

Lastly, the end-users access the final products browsing the catalogue in each dissemination system.

3. TARGET ARCHITECTURE

The main goal of the new architecture is to centralise the data storage in a star schema where the current BAQUNIN infrastructure could be harmonised.

In this paragraph the new BAQUNIN framework is described. Details on interfaces, protocols and processes, are also provided.

3.1 System Level Architecture

Each system component belonging to the BAQUNIN framework is reported in Fig. 3 as an element integrated into the new infrastructure.

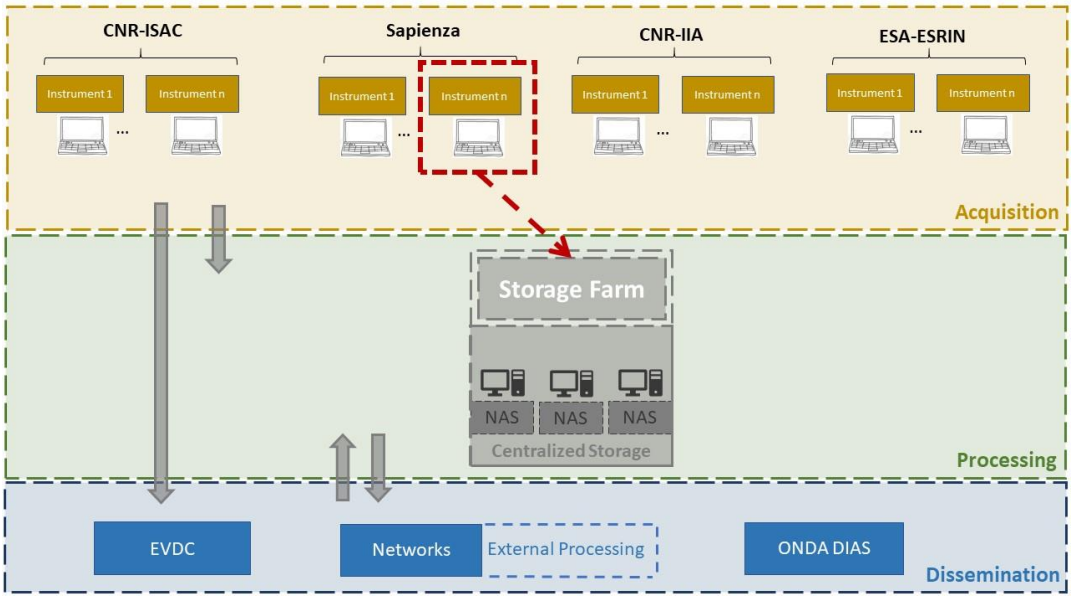


Figure 3: Target BAQUNIN Architecture

The data flow respects the same schema described in the previous paragraph but in the new architecture the data are collected in a centralised storage instead of locally in each acquisition facility. In addition, the data flow is managed by a single node (storage farm), in order to have a standard interface towards all integrated systems and a straightforward plug-in for all new instruments to be integrated (dash red line).

The storage will be located at "Sapienza" University. The hardware infrastructure, designed according to the pre-requisites of the Physics Department, will consist of three replicated HW systems configured to be compliant with all network security and data redundancy requirements. The data exchange between the internal and external systems respect to the firewall will occur via SFTP/SCP protocols (Fig. 4).

A third storage will be located at the Serco premises and configured in mirroring with the main system with a view to a future disaster and recovery plan.

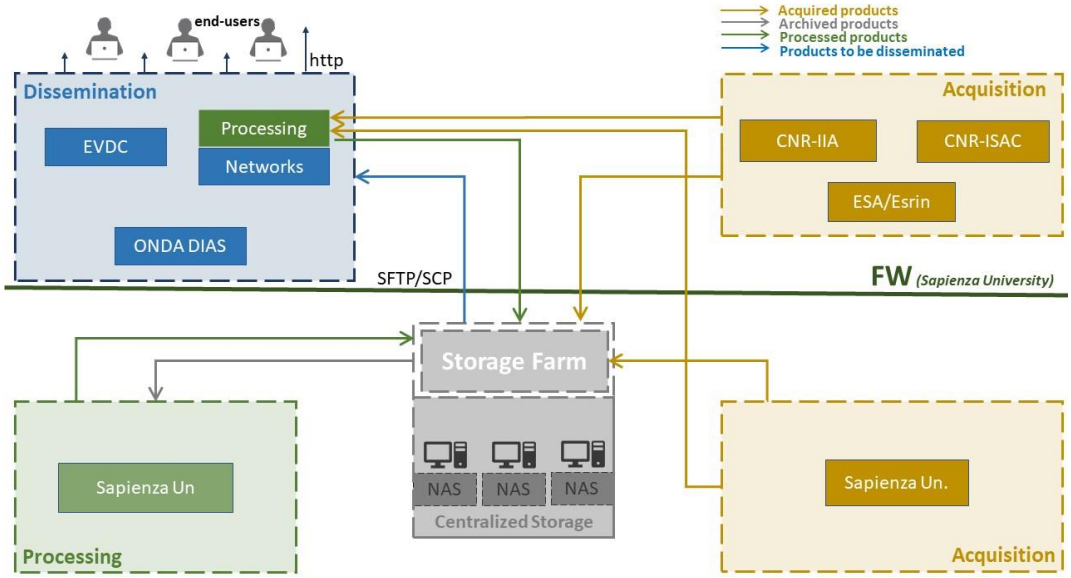


Figure 4: BAQUNIN Systems Architecture

Finally, the disseminated products will be retrieved by end-users which will access the catalogues of each external system via HTTP protocol.

The data dissemination process does not duplicate the products, i.e. each system only distributes products not provided by the others two. In Tab. 1 each product is therefore associated with a dissemination facility, the related instrument and acquisition facility.

Table 1 BAQUNIN products

Product	Instrument	Acquisition	Dissemination
O_3 surface, tropospheric and total column	Pandora	Sapienza CNR-ISAC CNR-IIA	EVDC
O_3 total column	Brewer	Sapienza	Networks – EUBREWNET
NO_2 surface, tropospheric and total column	Pandora	Sapienza CNR-ISAC CNR-IIA	EVDC
NO_2 total column	Brewer	Sapienza	ONDA DIAS
SO_2 surface, tropospheric and total column	Pandora	Sapienza CNR-ISAC CNR-IIA	Networks – PGN

HCOH surface, tropospheric and total column	Pandora	Sapienza CNR-ISAC CNR-IIA	Networks – PGN
H_2O total column, H_2O profile	CIMEL	Sapienza	Networks – AERONET
	LIDAR (profile)	Sapienza	EVDC (Low resolution data)
			ONDA DIAS (High resolution data)
Aerosol Optical Dept (AOD)	CIMEL	Sapienza	Networks – AERONET
	PREDE	Sapienza	Networks – EUROSKYRAD
	MFRSR	Sapienza	ONDA DIAS
	LIDAR	Sapienza	ONDA DIAS
Aerosol backscattering and extinction profiles	LIDAR	Sapienza	EVDC (Low resolution data)
			BAQUNIN DB (High resolution data)
Scattering and Absorption Angstrom Exponent	CIMEL	Sapienza	Networks – AERONET
Angstrom Exponent	CIMEL	Sapienza	Networks – AERONET
	PREDE	Sapienza	Networks – EUROSKYRAD
Single Scattering Albedo, Volume size distribution	CIMEL	Sapienza	Networks – AERONET
Real and imaginary part of Refractive Index, Phase Function	PREDE	Sapienza	Networks – EUROSKYRAD
Solar Irradiance	Pyranometer	Sapienza	ONDA DIAS
Spectral Radiance	Pandora	Sapienza	Networks – PGN
UV Dose, UV index	Brewer	Sapienza	BAQUNIN DB
Cloud top/bottom	Lidar	Sapienza	ONDA DIAS (High resolution data)
Cloud fraction	All Sky Camera	Sapienza	EVDC
Thermal Turbulence, Wind Speed and Direction	Sodar	Sapienza	EVDC (Low resolution data)
			ONDA DIAS (High resolution data)
Surface air temperature, humidity, pressure and wind	Metereological sensors	ESA/Esrin	ONDA DIAS

The implementation status for each product is updated and published on the BAQUNIN web site (http://www.baqunin.eu/products) where all information on the project are also available.

The BAQUNIN web site is currently implemented in a web hosting environment but will be migrated in a dedicated Virtual Machine belonging to the ONDA DIAS infrastructure: within this fully customizable development environment the end-users will be able to access directly the full Copernicus data archive and perform long-term time series analysis and cal/val activities using the available BAQUNIN archive.

End of Document