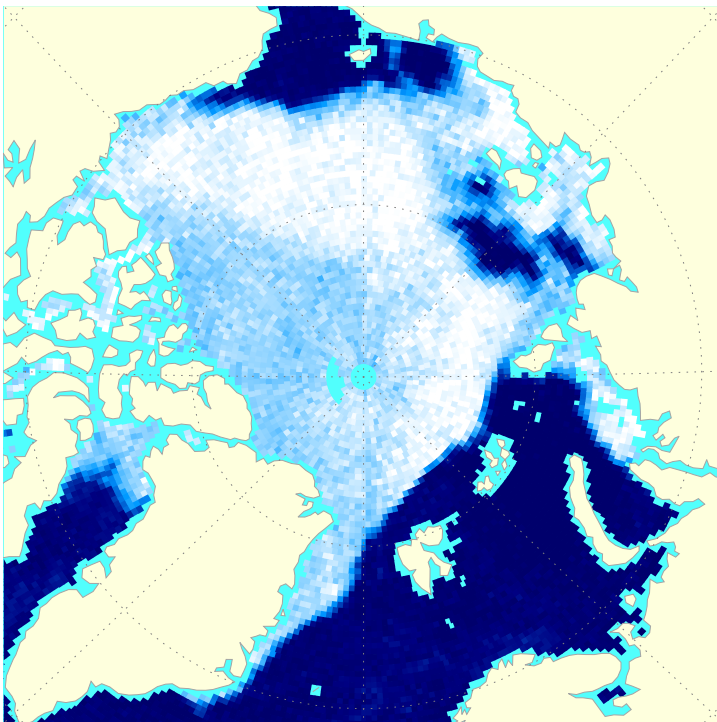


ECMWF Feature article

from Newsletter Number 177 – Autumn 2023

COMPUTING

Running a Global Broker as
part of the new WMO data
sharing solution



www.ecmwf.int/en/about/media-centre/media-resources

doi: 10.21957/rgm8d751xk

This article appeared in the Computing section of ECMWF Newsletter No. 177 – Autumn 2023, pp. 32–36.

Running a Global Broker as part of the new WMO data sharing solution

Rémy Giraud, David Podeur, Thierry Lacoste (all Météo-France)

The World Meteorological Organization (WMO) is upgrading its Global Telecommunication System (GTS) and the WMO Information System (WIS), which are used to transmit WMO data globally, to an Internet-based service called WIS 2.0 (WIS2). As described in a previous article in this Newsletter (<https://www.ecmwf.int/en/newsletter/176/computing/wis-20-wmo-data-sharing-21st-century>), in order to provide a reliable, efficient service for all WIS Users, the following Global Services have been defined:

- **Global Broker:** Meteorological centres will be responsible to make sure that all messages announcing the availability of new data and metadata can be easily obtained by all users. The Global Broker will provide a subscription service using the MQTT (Message Queuing Telemetry Transport) standard and Free and Open Source Software solution with an additional companion software (specific to WIS 2.0) to ensure uniqueness of messages as well as verifying the correct format of those messages.
- **Global Cache:** In order to provide quick and reliable access to core data as defined by the WMO Unified Data Policy, a copy of this data will be made available by a Global Cache. Storing data from originating WIS2 Nodes, the Global Cache will then make available the core data to all WIS Users.
- **Global Discovery Catalogue:** Each dataset available on WIS 2.0 must be described by a metadata record, using the OGC API - Records standard (soon to be ratified). The Global Discovery Catalogue will provide a discovery and metadata service using Free and Open Source Software, as well as provide quality assessment capabilities in support of continuous improvement of WIS 2.0 metadata.
- **Global Monitoring:** WIS 2.0 being an operational solution, it must be monitored. Each WIS2 Node and Global Service will provide metrics relevant to their operations. Global Monitoring Centres will collect the metrics and make available a visual dashboard presenting those metrics and alert the Centres when an unexpected event occurs in support of corrective action.

This article provides more detailed information on the architecture of WIS2 with a focus on one of the Global Services: the Global Broker and the instance of a Global Broker operated by Météo-France using the European Weather Cloud.

Publish and subscribe (pub/sub) in WIS2

In the weather, climate, hydrology and ocean community, where new data are almost constantly produced, it is key to make this new data available to customers as quickly as possible. Historically, the solution has been based on pushing the data from producers to consumers. This method has many merits, including in particular its simplicity. However, it also has a lot of drawbacks. For example, consumers cannot change easily what they receive, and the successful product distribution by producers depends on the availability of the IT systems of consumers.

When designing WIS2, its architects decided to replace the 'push by default' by 'pull by choice'. The new challenge was therefore to find a way to inform the user that new data are available.

Many messaging systems, such as X/Twitter and WhatsApp, are using a publish/subscribe approach. Consumers subscribe to channels, topics or feeds of interest, and producers publish their information. Consumers are immediately informed about any new information and can access it. WIS2 is following the same design principle. A data producer publishes a message in a topic, and users, by subscribing to the topic of interest, will be informed that new data are available. The message contains an HTTPS link to the resource. Global Services, in particular Global Caches and Global Brokers, are providing the required scalability and redundancy needed for WIS2 operations.

The WIS2 Notification Message

Each WIS2 Node will announce the availability of data using a WIS2 Notification Message. This message, which in the case shown below is from the Swedish Meteorological and Hydrological Institute (SMHI), is published on a local broker managed by SMHI:

```
{
  "id": "019918dc-0419-4bb1-aa70-5de539621087",
  "version": "v04",
  "geometry": {
    "type": "Point",
    "coordinates": [
      23.389945,
      67.20363
    ]
  },
  "properties": {
    "datetime": "2023-07-31T17:14:00Z",
    "hierarchy": "origin/a/wis2/swe/smhi/data/core/weather/surface-based-observations/synop",
    "data_id": "wis2/swe/smhi/data/core/weather/surface-based-observations/synop/WIGOS_0-20000-0-02095_2023-07-31T17:14:00",
    "integrity": {
      "method": "sha512",
      "value": "9cb67d6a6e5735f0166d93e07a19309cfe0824ca7f0b380d23eb74ebb6b59cbd26bec604e3feldebee7c4b34f839c19dfe09e24b021f93d69ddb3ad8eacf08a2"
    },
    "pubtime": "2023-07-31T17:28:27.204915Z"
  },
  "type": "Feature",
  "links": [
    {
      "href": "https://wis-tst.smhi.se/WIGOS_0-20000-0-02095_2023-07-31T17:14:00.bufr",
      "rel": "canonical",
      "type": "application/x-bufr"
    }
  ]
}
```

The Global Brokers are subscribing to this broker and, after checking its uniqueness, this message is published on the broker part of the Global Broker. WIS Users (including Global Caches) will be informed of the availability of the dataset if they subscribe to the Global Broker. They can download the data if they are of interest to them. Figure 1 summarises the flow of messages in WIS2.

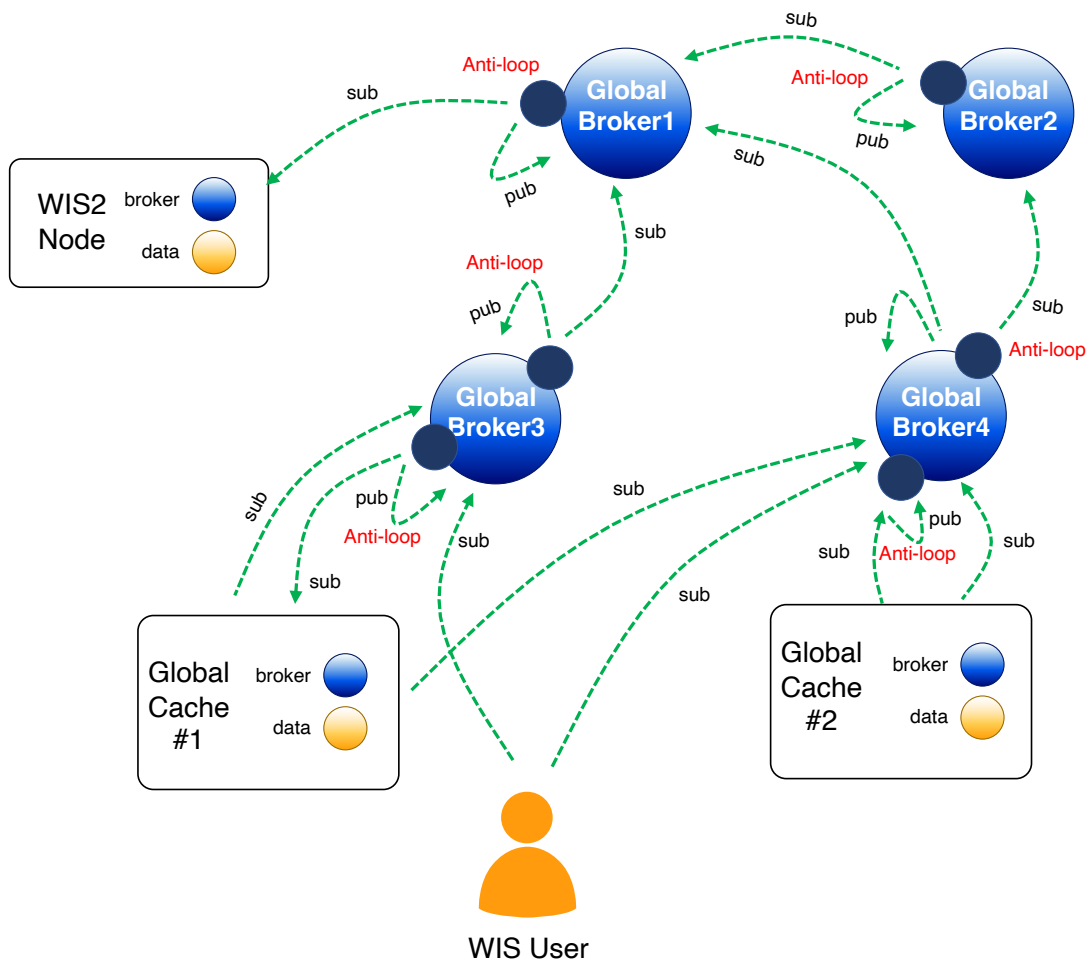


Figure 1 The flow of messages in WIS2 to alert a WIS User to the availability of data. The ‘anti-loop’ functionality is described below.

What constitutes a Global Broker?

A Global Broker is made up of two components:

- A highly available, scalable publish/subscribe broker
- Software ensuring deduplication of published messages by the WIS2 Node, the Global Cache and the other Global Brokers.

The broker part

Having decided to rely on a publish/subscribe approach, the next decision while designing WIS2 was to choose the underlying protocols. Considering that WIS2 has decided to leverage open standards whenever possible, the potential protocols considered were:

- AMQP (Advanced Message Queuing Protocol) version 1.0
- Message Queuing Telemetry Transport (MQTT) version 3.1.1 and version 5.0.

After thorough analysis and based on hands-on experience by various WMO Members, WMO experts decided to rely on MQTT 3.1.1 and 5.0 for the publish/subscribe protocol in WIS2. Originally, MQTT 3.1.1 was designed to support the Internet of Things (IoT). It is widely used by many industries to collect data from sensors located in cars or machinery. Nowadays, the two versions of the MQTT protocol coexist, and most of the tools available as Free and Open Source Software support both versions. For example Mosquitto, the broker used in the WIS2-in-a-box reference implementation of a WIS2 Node, is compliant with both versions. A notable exception to this is RabbitMQ, whose currently available releases do not support version 5.0. However, VMWare, the owner of RabbitMQ, has confirmed that an upcoming release of RabbitMQ will support MQTT 5.0. It should be released by the end of 2023.

Some of the other MQTT broker implementations are EMQX (free and enterprise version), HiveMQ (only commercial, licence-based) and VerneMQ (free for personal use and commercial licence otherwise). In order to benefit from support from a commercial company, Météo-France decided to rely on VerneMQ to run its Global Broker. Support has been tremendous and has helped us to have a reliable solution to run the broker. The VerneMQ broker has been deployed on a cluster of virtual machines to provide a redundant and scalable service.

The anti-loop function

In a typical deployment of MQTT, the broker, the publishing part and the subscribing part are managed by the same entity. In particular, it is important to note that there is very little protection to prevent a publisher from flooding the broker with many messages.

First, to avoid potential flooding by publishers, it was decided that no WIS2 data provider would be allowed to publish on a broker that is not managed by the data provider itself.

Second, considering the extent of WIS2, an architecture with a single redundant broker ensuring all of the exchange of messages was discarded, and a design with multiple brokers providing a distributed, redundant, and reliable operation was adopted. Considering that there is no standard method to copy messages between brokers, so that all messages are available on all brokers, a specific WIS2 solution called the 'anti-loop' has been designed.

The anti-loop tool:

- subscribes to as many brokers as needed; the brokers might be part of a WIS2 Node, a Global Cache or another Global Broker
- publishes to its local broker after having checked that the message has not yet been published to the broker.

In the WIS2 Notification Message example above, the id of the message is used to implement this 'anti-loop' feature.

Météo-France implementation of the anti-loop function

As this part is specific to WIS2, there is no off-the-shelf software that implements this feature. In late 2022, the anti-loop function was developed to be able to deploy the Global Broker feature during the pilot phase of WIS2, starting in early 2023. A flow-based, low-code solution based on the open-source tool Node-RED was used. This has ensured a rapid take-off of the WIS2 pilot phase while providing the required features.

A flow is a succession of nodes providing high-level components, such as MQTT subscription, variable manipulation, MQTT publication, and OpenMetrics (Prometheus) monitoring. It is also possible to develop specialised functions in javascript for features that are not part of a pre-existing node. As part of the Node-RED flow, Redis (another open-source tool) is used to store and then detect the uniqueness of the id providing the core feature of the anti-loop function.

The flow shown in Figure 2 is the most basic implementation of the anti-loop function. It is very compact and extremely easy to understand while providing the advanced features required. After six months of use in the pilot phase, it has been extremely reliable. What started as a test implementation can now be considered as production-ready and mature enough to be used in production for the upcoming phases of WIS2. The code and a docker container providing the anti-loop feature is available on GitHub (<https://github.com/golfvert/WIS2-GlobalBroker-Redundancy>) and Docker Hub (<https://hub.docker.com/r/>)

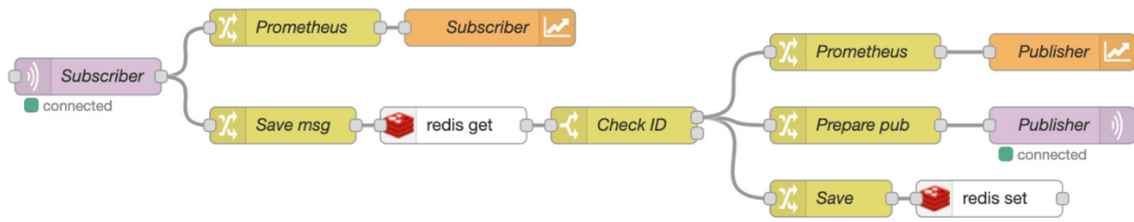


Figure 2 Node-RED flows are read from left to right. The message is received by the Subscriber (MQTT Subscribe) node, connected to the remote WIS2 Node. The message is sent to the downstream nodes to make available statistics (Prometheus and orange Subscriber nodes) and also to the other stream where the anti-loop function is performed. The Save msg node formats the message to query the redis database (redis get). The Check ID is checking whether the message has already been seen or not. Out of this node, the upper connection is used when the message was not received before. Prometheus and orange Publisher are providing statistics. Prepare pub and Publisher (MQTT Publish) are used to publish the received message to the broker of the Global Broker. Then Save and redis set are used to store the message on the redis database so that further messages with the same id will be discarded.

[golfvert/wis2gb](https://github.com/golfvert/wis2gb).

The version used today is more advanced and provides additional features. In particular, what started

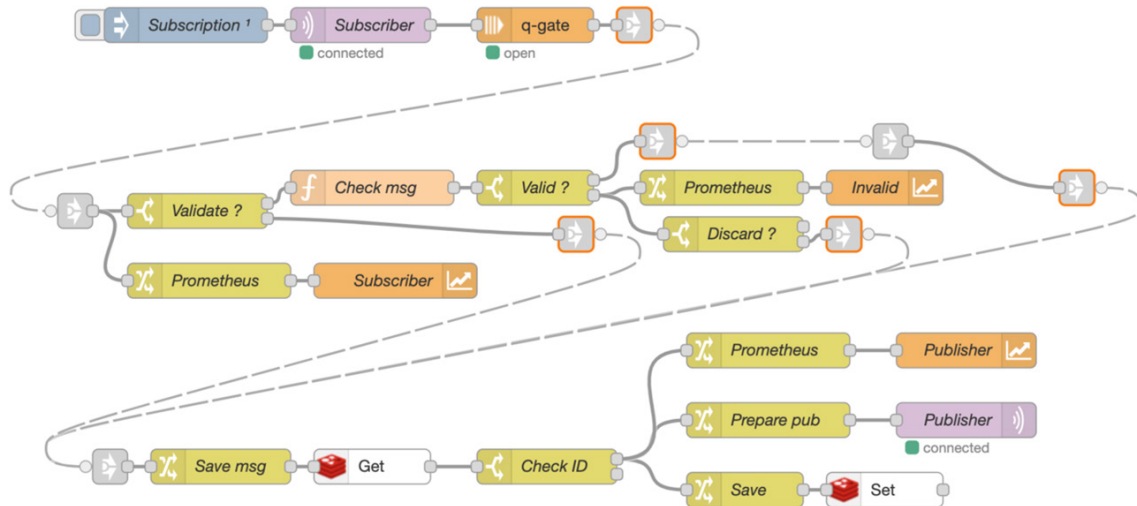
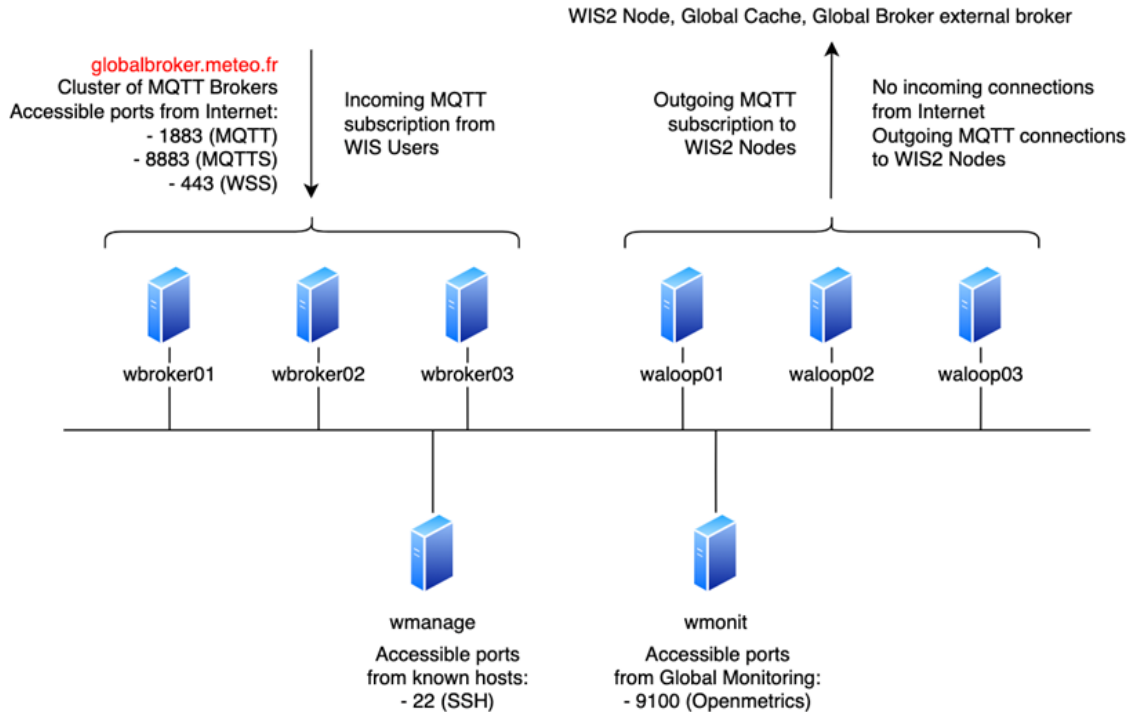


Figure 3 This is version 2 of the Node-RED-based anti-loop used by the Météo-France Global Broker. It has three improvements. First, the new version is using redis in a cluster. The Get and Set (with the red cube) are reading and writing on the redis database. Second, it is also possible to run multiple instances of the same container connected to the same WIS2 Node in an active/active manner. All instances are connected (Subscriber node) to the WIS2 Node. The q-gate node is either in ‘open’ mode, and the messages are sent to the downstream nodes, or in ‘queue’ mode, and the messages are kept in a holding queue. A separate flow, not shown here, is checking whether this instance is the primary one (and therefore process the messages) or a secondary one (and therefore queue the messages). If a secondary node is promoted to primary (the previous primary not being available any more), then the q-gate is ‘flushed’ (all queued messages are sent) and then put in ‘open’ mode. Thanks to this queuing mechanism, no message will be lost even in case of failure of the primary instance. The third improvement of the flow is the capability to Validate the messages. If configured, the messages can be syntactically checked (Check msg) and then Discard if the message is not compliant with the standardized format. In this case, statistics will also be produced for Invalid messages.

as being a single point of failure is now working in an active/passive mode on a Docker set of virtual machines. The current version of the flow is shown in Figure 3.

Running Météo-France Global Broker on the European Weather Cloud

Taking advantage of the OpenStack-based European Weather Cloud provided by ECMWF, Météo-France



- Traefik is installed on all hosts (except wmanage)
- Redis (in docker) is installed on wbroker0x and waloop0x and is configured in cluster
- On waloop0x, the antiloop software is running in primary/secondary mode
- wmonit is gathering all metrics and stores in Prometheus, metrics can be collected by Global Monitoring

- MQTT (Message-Queueing Telemetry Transport)
- MQTTS (The Secure Socket Layer based version of MQTT)
- SSH (Secure SHell)
- WSS (WebSocket Secure)

Figure 4 The architecture of the Météo-France Global Broker.

has decided to run the Global Broker first in Reading (December 2022 – September 2023) and then in Bologna. ECMWF is providing the security layer (firewall) and the load-balancing layer (Octavia service in Open Stack). The overall architecture is shown in Figure 4.

The three ‘wbroker’ hosts are hosting the VerneMQ software and are clustered to provide the reliable, scalable and redundant MQTT 3.1.1 and MQTT 5.0 pub/sub protocol. The three ‘waloop’ servers are

hosting the anti-loop docker containers. There is one primary container per WIS2 Node, running on one of the three hosts. Traefik, another open-source tool, is used extensively to load-balance the traffic between the hosts. The two additional servers ('wmanage' and 'wmonit') are used to manage and monitor the entire environment.

Conclusion

Having decided to use open standards and, as a consequence, being able to use off-the-shelf software is one of the key choices made by the architects of WIS2. The Météo-France Global Broker is an excellent example of the benefits of those choices. Built around VerneMQ, Node-RED, Redis, and Traefik, developing a reliable and scalable Global Broker has been easy and quick.

Thanks to the support offered by ECMWF (European Weather Cloud team and networking team), Météo-France has provided to the WIS2 community the first example of a Global Broker. During the pilot phase, the level of service has been extremely high. Running at ECMWF on the European Weather Cloud in Bologna, a state-of-the-art facility, the Global Broker will provide Météo-France with the environment needed to run one of the Global Services critical for the success of WIS2 and the upcoming migration from the GTS and WIS.

© Copyright 2023

European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, RG2 9AX, UK

The content of this document, excluding images representing individuals, is available for use under a Creative Commons Attribution 4.0 International Public License. See the terms at <https://creativecommons.org/licenses/by/4.0/>. To request permission to use images representing individuals, please contact pressoffice@ecmwf.int.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error or omission or for loss or damage arising from its use.