EUROPEAN WEATHER CLOUD

CLOUD COMPUTING-BASED INFRASTRUCTURE, FOCUSED
ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

# The European Weather Cloud

*Collaborative Delivery and Operation of a Production Grade On-Premises OpenStack Cloud Environment*

# Invitation to Pre-Qualify (ItPQ)

| |
|---|
| Ref: ECMWF/ ITPQ/2020/307 |
| ISSUED BY:<br>ECMWF<br>Administration Department<br>Procurement Section |
| Date: 01/12/2020 |
| Version: Final |

# Table of Contents

1 Instructions for Responders ....5

*1.1 Introduction ....5*

*1.2 Introduction to ECMWF ....5*

*1.3 Introduction to European Weather Cloud ....5*

*1.4 Timetable ....6*

*1.5 Submission of Responses to this ItPQ ....7*

1.5.1 General ....7

1.5.2 ECMWF eprocurement portal and submission of responses ....7

*1.6 Evaluation method and selection criteria ....9*

*1.7 Required Information ....9*

1.7.1 Company and Contact Details ....9

1.7.2 Company Track Record ....10

1.7.3 Approach to Delivery ....10

1.7.4 Requirements ....10

1.7.5 Standards & Procedures ....10

1.7.6 Additional matters ....10

*1.8 Conditions for Responses ....11*

*1.9 Language for Response ....11*

*1.10 Expenses ....11*

*1.11 Confidentiality ....11*

*1.12 Warnings/disclaimers ....11*

2 Background to this Invitation to Pre-Qualify ....13

3 Current cloud capabilities within ECMWF ....13

*3.1 The Climate/Atmosphere Data Store (CDS/ADS) ....14*

*3.2 The WEkEO Data & Information Access Service (DIAS) ....15*

*3.3 The pilot European Weather Cloud (pEWC) ....15*

*3.4 Summary of Capabilities ....17*

*3.5 The Capability Goal ....17*

*3.6 Preparing Responses – General Guidance on Achieving Compliance ....17*

4 Mandatory High-Level Requirements - General Remarks ....19

*4.1 The Mandatory High-Level Requirements – Common Terminology ....19*

*4.2 The Mandatory High-Level Requirements - Summary List ....19*

*4.3 The Mandatory High-Level Requirements – Detailed Requirements ....22*

4.3.1 OpenStack APIs to Ussuri Release Standard or Later (**MHLR#1**) ....22

4.3.2 Three-Nines Unplanned Downtime (**MHLR#2**) ....22

4.3.3 Three-Nines Planned Downtime (**MHLR#3**) ....23

4.3.4 Two-Nines Disaster Recovery (**MHLR#4**) ....23

4.3.5 Repeatably Deployable On-Premises Using Full Automation (**MHLR#5**) ....23

4.3.6 Repeatably Deployable Without Internet Access (**MHLR#6**) ....24

4.3.7 Highly Scalable Without Performance Degradation (**MHLR#7**) ....25

4.3.8 Modern Networking Compliant (**MHLR#8**) ....25

4.3.9 Differing Hardware Flavours (**MHLR#9**) ....26

4.3.10 Differing Storage Flavours (**MHLR#10**) ....27

4.3.11 Software Defined Networking (**MHLR#11**) ....28

4.3.12 Data Centre Components, Sub-Systems, Practices & Introspection (**MHLR#12**) ....28

4.3.13 Server Farms & Soft Clusters – Metal Management (**MHLR#13**) ....29
4.3.14 Logistics & Equipment Supply Chains (**MHLR#14**) ....29
4.3.15 1st, 2nd and 3rd Line Support - Personnel (**MHLR#15**) ....30
4.3.16 Regulatory & Compliance (**MHLR#16**) ....30
4.3.17 Segmentable & Ring-Fenced (**MHLR#17**) ....31
4.3.18 High-Integrity Data Storage (**MHLR#18**) ....32
4.3.19 Resilient Operation Under non-OpenStack Failure Conditions (**MHLR#19**) ....32
4.3.20 Secure By Default –Secure Audited Operation & Misuse Detection (**MHLR#20**) ....32
4.3.21 Monitoring and Alerting (**MHLR#21**) ....33

5 Technical Requirements - General Remarks ....34

*5.1 Choice of OpenStack Distribution & Deployment Method (**TR#1**) ....34*
5.1.1 OpenStack Deployment using Triple-O / RedHat OpenStack Platform (RHOSP) (**TR#1a**) ....35
5.1.2 OpenStack Deployment using Kolla-Ansible & Kolla (**TR#1b**) ....35
5.1.3 OpenStack Deployment using OpenStack-Ansible (**TR#1c**) ....36
5.1.4 Other Strong OpenStack Deployment Candidates (**TR#1d**) ....36

*5.2 Choice of Ceph Distribution & Deployment Method (**TR#2**) ....37*

*5.3 Choice of Kubernetes Distribution & Deployment Method (**TR#3**) ....37*

*5.4 Containerisation using OpenShift / OKD (**TR#4**) (**OPTIONAL**) ....39*

*5.5 Operations & Maintenance (**TR#5**) ....39*

*5.6 Training (**TR#6**) ....40*

*5.7 Documentation (**TR#7**) ....40*

*5.8 Testing, Failures, Fault Analysis, Fixing & Patching (**TR#8**) ....40*

*5.9 Resource Accounting & Charging (**TR#9**) ....41*

*5.10 Fully Automated Cloud & Supporting DevOps Environments (**TR#10) (OPTIONAL)** ....41*

*5.11 Segmentation and Zoning of Ring-Fenced Resources (**TR#11)** ....42*

*5.12 Architectures Types – Cloud Efficiency (**TR#12) (OPTIONAL)** ....42*

6 Indicative Pricing ....44

*6.1 General Approach to Pricing ....44*

*6.2 Manpower Types, Amounts and Pricing ....44*

*6.3 Equipment Pricing ....44*

*6.4 Software Pricing ....46*
6.4.1 Open Source ....46
6.4.2 Community Editions ....46
6.4.3 Enterprise Editions ....46

Appendix A Cross-Reference to Responders Documentation ....47

Appendix B OpenStack Components Diagram ....54

# Index of Tables

Table 1: CDS/ADS System Scale ........ 14

Table 2: Pilot European Weather Cloud System Scale ........ 16

Table 3: Scale-Out System Scale ........ 21

Table 4: Types of Artefacts Required to Build Production Clouds ........ 24

# 1 Instructions for Responders

## 1.1 Introduction

This Invitation to Pre-Qualify (ItPQ) has been prepared by the European Centre for Medium-Range Weather Forecasts, (governed by its Convention and associated Protocol on Privileges and Immunities which came into force on 1 November 1975, as amended on 6 June 2010) ("ECMWF").

The purpose of this ItPQ is:

- To state ECMWF's requirements for a Collaboratively Delivered and Operated Production Grade On-Premises OpenStack Cloud Environment
- To solicit responses from the widest possible range of potential suppliers who are able to meet those requirements
- To use those responses to draw up a shortlist of potential suppliers who will be invited to tender for the planned follow-up Invitation to Tender (ItT)

**In recognition of the scale of this requirement, of the work involved in responding to it, and the fact that such work will not be compensated, ECMWF has simplified the compliance process for this stage only. Any subsequent ItT will be conducted under the normal ECMWF rules for such procedures. The simplified process to be used at this stage will be described further later in the document.**

ECMWF is not looking for **detailed** architectures or designs at this stage, though in some cases **high-level** architectures and designs are called for. There will be no commitment on the part of responders. Responses will, however, assist ECMWF in understanding the market and available solutions and in supporting its decision to tender and in what form for certain aspects of the European Weather Cloud.

Responders will, however, be required to provide indicative pricing as called for under various sections of this document. They will also be required to give an overall indicative price. They should be aware that price will carry a significant weight during any subsequent ItT evaluation

As a part of the assessment of the responses to this ItPQ, ECMWF may seek to meet with responders[1] to clarify and expand on their responses. Finally, and based on all the inputs received, ECMWF intends to issue a tender in the second quarter of 2021 for the procurement of certain aspects of the European Weather Cloud.

## 1.2 Introduction to ECMWF

ECMWF is an independent inter-governmental organisation supported by 34 Member States. Information on ECMWF's activities can be found at https://www.ecmwf.int/en/about. Information about the European Weather Cloud can be found at https://europeanweather.cloud.

## 1.3 Introduction to European Weather Cloud

The European Weather Cloud will deliver data access and cloud-based processing capabilities for the European Meteorological Infrastructure (EMI) and its users. The EMI comprises

[1] At the time of writing, expected to be conducted by tele-conference due to pandemic restrictions.

ECMWF, EUMETSAT, and the National Meteorological Services of their respective Member States (ECMWF[2]) (EUMETSAT[3]).

As the EMI collects more detailed and frequent weather and climate observations and develops enhanced prediction capabilities and services, it is increasingly facing challenges to provide sufficient infrastructure to store, manage and process very large datasets.

At the same time, technological progress offers new possibilities to enable harmonised online access to data across large data centres that have been joined together by high-speed wide-area networks. Working on data in the cloud enables new types of capabilities, including running software close to the data, rather than downloading vast amounts of data locally and needing a local infrastructure to process it.

ECMWF and EUMETSAT hold an ever-growing data store for meteorological applications and so they have started the pilot phase of the "European Weather Cloud" to make it easier to work on weather and climate big-data in a cloud-based infrastructure. The pilot phase also endeavours to develop the set of terms and conditions for operation and use of such an infrastructure, and for joining a federation of meteorological data centres.

The aim is to maximise the value generated by Member State investments, rationalise the usage of data access infrastructure and related developments, and foster new forms of data-driven collaboration across the EMI and users of meteorological data.

### 1.4 Timetable

ECMWF envisages the following timetable for this ItPQ and for any subsequent tender that may follow as a result:

| **Closing date/time for submission of responses to the ItPQ** | **26th February 2020, 15:00 (UK local time)** |
|---|---|
| Evaluation of responses and drafting of Invitation to Tender documentation | April/May 2021 |
| Publication of Invitation to Tender | Quarter 2 2021 |
| Negotiation of service details with preferred bidder(s) | Quarter 3 2021 |
| Award of contract by | Quarter 4 2021 |

ECMWF may, at its own absolute discretion, amend the dates mentioned above and in such an event ECMWF will notify all responders who have provided ECMWF with an email address for communication of additional information.

[2] Austria, Belgium, Croatia, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, The Netherlands, Norway, Portugal, Serbia, Slovenia, Spain, Sweden, Switzerland, Turkey and United Kingdom.

[3] Austria, Belgium, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, The Netherlands, Norway, Poland, Portugal, Romania, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey and United Kingdom.

## 1.5 Submission of Responses to this ItPQ

### 1.5.1 General

Before submitting their response to this ItPQ, responders should be aware of the following considerations:

- Responses to this ItPQ are voluntary.
- This ItPQ is being issued to enable ECMWF to:
  - better understand the market,
  - better understand available solutions, and
  - for the purpose of drawing up a shortlist of pre-qualified potential bidders who may subsequently be invited to respond to the follow-up ItT.

  **This ItPQ should not be construed as a solicitation or as an obligation on the part of ECMWF in any way.**
- ECMWF will use the information submitted in response to this ItPQ at its discretion within standard confidentiality rules.
- The information submitted will be analysed and may be shared both internally and with external partners or incorporated into future documentation, as appropriate and at the discretion of ECMWF (see also Section 1.7.3).
- Proprietary, confidential, or sensitive material/information should be marked as such and highlighted in the response. ECMWF prefers not receiving such material/information at this ItPQ stage if at all possible. **Furthermore, under no circumstances is classified material to be submitted at this ItPQ stage.**
- ECMWF confirms that Responders shall only be able to bid on the follow-up ItT if they have made it onto the shortlist of successful Responders resulting from this ItPQ exercise.

### 1.5.2 ECMWF eprocurement portal and submission of responses

Responses to this ItPQ must be submitted via ECMWF's eProcurement Portal no later than the closing date/time specified in Section 1.4 (**Closing date/time for submission of responses to the ItPQ).** Instructions on how to access the Portal for this ItPQ, and any subsequent procurement exercise, are included below:

**(a) Supplier registration on the eProcurement Portal**

To take part in any procurement exercise, including this ItPQ, and to be able to express an interest in an opportunity, first you must register on the eProcurement Portal (ProContract) with details of your organisation. Registration is free and the process is managed by the portal provider Due North. The eProcurement Portal can be accessed from one of the following links: https://procontract.due-north.com or https://procurement.ecmwf.int (the latter redirects to the ProContract Opportunities page).

When the registration is accepted then you will receive an email containing a reminder of your username and the link to access the opportunity portal. Note that once you are registered, you will be able to see all the opportunities available in ProContract for various buyer organisations beside ECMWF. You will be able to narrow your search results to opportunities issued by ECMWF by selecting the corresponding portal or organisation on the ProContract Opportunities page.

Further guidance for suppliers is available at https://supplierhelp.due-north.com

Registration does not commit you to respond to any procurement exercisess. ECMWF will not interpret your registration as an intention to respond to an opportunity. However, you can "Register intent" to notify ECMWF that you intend to respond if you so wish.

Once an opportunity is published you must log on to the portal and express interest to obtain access to the relevant documents and any subsequent clarifications.

Tenderers who have registered an interest in a particular procurement exercise using the eProcurement Portal will be kept informed of any developments about it, including any updates to the tender documentation and any clarifications that are issued. Tenderers must read all documents and comply with ECMWF's instructions with regard to the submission of their proposals. ECMWF reserves the right to reject a proposal that does not substantially comply with the conditions that are part of the procurement exercise.

ECMWF has also developed a document, providing step-by-step guidance to suppliers about how to navigate the eProcurement Portal for:

- finding an opportunity launched by ECMWF ;
- accessing the tender documents on the Portal;
- submitting a response to an opportunity on the Portal;
- accessing the messaging board of an opportunity.

The guidance document can be found at the following link: https://www.ecmwf.int/en/about/suppliers

**(b) Online questionnaire for preparation of responses on the portal**

Online questionnaire of the opportunity in the portal is where tenderers prepare their responses by answering questions about their organisations and proposals and uploading documents for their responses in accordance with the instructions therein. It must be answered online. You can download it to look at the questions but you cannot complete the questionnaire offline and upload it back to the portal. It does not have to be answered completely in one session. You can start to answer it, then save it and complete it later.

The portal will remember your responses to questions in the questionnaire so that if you respond to more than one opportunity and any questions are the same you will not have to respond to those questions again (though you are able to change your answer). Having completed the questionnaire online you can either save it for later submission or submit it straight away **(recommended, as it can be amended later).** Once you complete your response, **you must click the "Submit response" button and the status of your response should read "Submitted".** Until the closing date you can change answers to the questionnaire and submit a new version. ECMWF is able to see only the version of the answers that is current at the closing date.

**(c) Clarification questions**

All correspondence is conducted via the e-Procurement Portal. No other form of communication will be accepted.

Any questions concerning this opportunity ("Clarification Questions") must be submitted via the eProcurement Portal and must be received by ECMWF more than 10 working days before the closing date. ECMWF will respond via the portal within five working days and will send the question and answer to all suppliers who have expressed an interest in the opportunity unless the question is specific to a supplier's proprietary solution. The identity of the questioner will not be revealed.

Please note that ECMWF reserves the right to contact responders during the evaluation process in order to clarify points in their response, therefore the contact point should be monitored during this period.

#### (d) Timeliness of response

As the ItPQ will be used for shortlisting, ECMWF will not consider any late or partial responses to this opportunity (unless this is due to a technical issue caused by either ECMWF or their Portal) nor will it consider requests for extension of the time or date fixed for the submission of responses. It may, however, at its own absolute discretion extend the time or date fixed for submission and in such an event ECMWF will notify all interested parties via the e-Procurement Portal.

ECMWF encourages responders to submit the questionnaire soonest, even if it is in draft format, as it can still be amended, and added to, up to response close.

Technical failure, including of a computer, browser, e-mail system or internet connection, is not a valid reason for late or failed submission of a response, unless as a result of a failure of the ECMWF's eProcurement Portal, and in the case that there was no reasonable course of action the responder could have taken to submit the response on time. **It is important that you do not leave the submission of your response to the last minute.**

## 1.6 Evaluation method and selection criteria

| **Evaluation criteria** | **Weights** |
|---|---|
| Track record | 20 |
| Quality of response<br>• *completeness of the response*<br>• *compliance with mandatory requirements*<br>• *added value in terms of information/solutions provided*<br>• *understanding of, and ability to deliver, ECMWF's requirements* | 80 |

## 1.7 Required Information

Parties wishing to respond to this ItPQ are requested to structure their response as described below. Responses should not be overly detailed (in the event that we progress to tender, full proposals will be sought at that stage), but should be sufficient to enable ECMWF to understand the options available. Please do not include any marketing or similar brochures or other materials in your response unless they are specific to your response and/or to one of the questions.

### 1.7.1 Company and Contact Details

Please confirm the full name and address of your company, together with details of the person at your organisation who can be contacted by ECMWF in relation to your response, and to whom any clarification questions that may arise may be addressed. Please give their name, position/title, address, telephone number and email address.

### 1.7.2 Company Track Record

Staff resources

Please give details of your staff numbers, skills and locations relevant to the services you envisage in your response to this ItPQ. Please include CVs or an outline of the particular qualifications and experience by key staff proposed.

Please set out any key skill or employee dependencies and the availability of replacement skills in those areas, particularly as regards the envisaged permanent point of contact.

Customers and references

Please describe in brief terms your company's history and your recent provision of the services you envisage providing in your response to this ItPQ. Please supply a list of three customers to whom similar services to those you envisage in your response to this ItPQ have been provided.

ECMWF reserves the right to seek references from one or all of these customers.

### 1.7.3 Approach to Delivery

Please describe in brief terms the types of hardware, software and services you could offer in line with the background and scope stated in this document. This should include an indication of an implementation budget with breakdown for each of the features presented and an annual budget for all commercial licences. **Please note that ECMWF has a strong preference for not using proprietary licensed software or software that otherwise carries a fee for use, whether one time or subscription based.**

### 1.7.4 Requirements

Responders' replies should take the form of a brief, but sufficiently detailed, response to each of the questions set out in Section 2, 3, 4, 5, 6 to enable ECMWF to fully understand what services are available, can be delivered and how.

It would greatly assist ECMWF's analysis of your responses if you completed the table in Appendix A which is a cross-reference from the requirements (**MHLR**s & **TR**s) in this document to the relevant section in your responses.

### 1.7.5 Standards & Procedures

Ordinarily, ECMWF would ask for Standards & Procedures information at the Invitation to Tender (ITT) stage of a project. However, on this occasion, Responders are encouraged to give any supporting Standards & Procedures information that they feel lends weight to the views expressed in their submissions.

### 1.7.6 Additional matters

Please set out any additional information or other relevant matters which you think have not been adequately addressed in the ItPQ and/or merit further consideration in your response.

## 1.8 Conditions for Responses

This exercise will be conducted in accordance with ECMWF's procedures as set out in these documents and no other procedures will apply.

Responders shall not:

- Consult, communicate or agree with any other Responder on any matter whatsoever related to this procurement exercise;
- Make any attempt to induce any other person or organisation to submit or not to submit a response.

## 1.9 Language for Response

All response documents, manuals and diagram labelling shall be written in English.

## 1.10 Expenses

ECMWF will not reimburse expenses incurred in connection with the preparation and submission of the response to this ItPQ. ECMWF accepts no liability whatsoever, whether in contract, tort or otherwise in relation to this ItPQ or in respect of any costs, damages or expenses incurred by responders or any third party.

## 1.11 Confidentiality

ECMWF reserves the right to retain all documents submitted by responders in response to this ItPQ. Any information in such documents that is proprietary and confidential to the responder will be handled confidentially by ECMWF provided it is clearly and specifically identified as such. Such obligation shall not apply if such information is or was obtained from other sources that do not bind ECMWF as to confidentiality or if the information is in the public domain. ECMWF may make responses available for evaluation purposes to authorised people, including its governing body, committees, and professional advisers in addition to ECMWF’s own personnel under the same conditions of confidentiality. Please also note that all personally identifiable information (PII) requested by ECMWF or provided by responders will be treated in accordance with the ECMWF Policy on Personally Identifiable Information Protection (PIIP). It is available on www.ecmwf.int . ECMWF shall process all PII submitted by your response for the sole purposes of assessing your response. In doing so, ECMWF may share such PII with consultants or external advisors.

## 1.12 Warnings/disclaimers

Nothing contained in this ItPQ or any other communication made between the responders and ECMWF or its representatives shall constitute an agreement, contract or representation between ECMWF and any other party (except for a formal award of a contract made in writing by ECMWF). Receipt by a responder of this ItPQ does not imply the existence of a contract or commitment by or with ECMWF for any purpose. ECMWF reserves the right to change any aspect of, or cease, the ItPQ at any time. While ECMWF has taken all reasonable steps to ensure, as at the date of this document, that the facts which are contained in this ItPQ are true and accurate in all material respects, ECMWF does not make any representation or warranty as to the accuracy and completeness or otherwise of this ItPQ or the reasonableness of any assumptions on which this document may be based. ECMWF accepts no liability to responders whatsoever and however arising and whether resulting from the use of this ItPQ, or any omissions from or deficiencies in this document. ECMWF may use the information included in responses for any reasonable purpose connected with this ItPQ or subsequent tender.

## 2 Background to this Invitation to Pre-Qualify

ECMWF is issuing this Invitation to Pre-Qualify (ItPQ) at an important stage in its increasing adoption of cloud technology. Unlike many adopters of cloud, ECMWF, its Member States and its Partners face a particular challenge with respect to the amount of data to be processed and disseminated.

There is a substantial archive of historical data that is used on a regular basis – this accounts for more than 300 PB of near-line and on-line storage held in the ECMWF data centre.

Each day, raw data is acquired from many sensors worldwide, and production processed data is generated by the centres High-Performance Computers (HPCs), and this is disseminated to Member States and Partners as well as being added to the archive. Typically, this adds approximately ~35TB of new data each day.

The data is subsequently applied to many different challenging areas of meteorology and climatology research requiring substantial compute resources and generating still more data that needs to be archived for future reference.

Typically, public cloud offerings are best suited to application areas where the amount of data to be stored in the cloud is limited to terabytes and where the data can be cropped and/or taken off-line on an automated schedule to save money. Furthermore, they are best suited to situations where the data, once stored in the cloud, is processed in the cloud and the results that are shipped off-cloud are much smaller, saving on network egress costs.

ECMWF does not have such a usage profile. It requires High-Performance Computers (HPCs) to fulfil its core mission, the HPCs generate substantial amounts of data on a daily basis, and that has to be stored and disseminated to downstream users with minimum cost and maximum efficiency.

Accordingly, ECMWF has heavily favoured the use of **on-premises private OpenStack cloud technology** to date so as to maximise networking performance, minimise networking costs, and minimise long-term data storage costs.

But that, in turn, presents challenges including such considerations as skillset acquisition, equipment procurement, capacity planning, service resilience and service security.

Given the above context, and to recap from Section 1.1, the purpose of this ItPQ is:

- To state ECMWF's requirements for a Collaboratively Delivered and Operated Production Grade On-Premises OpenStack Cloud Environment
- To solicit responses from the widest possible range of potential suppliers who are able to meet those requirements
- To use those responses to draw up a shortlist of potential suppliers who will be invited to tender for the planned follow-up Invitation to Tender (ItT)

# 3 Current cloud capabilities within ECMWF

At the time of writing, ECMWF hosts in its Reading Data Centre the following on-premises OpenStack cloud systems:

- The Climate Data Store (CDS) / Atmospheric Data Store (ADS)
- The WEkEO Data & Information Access Service (DIAS)
- The pilot European Weather Cloud (pEWC)

Each is briefly described in the following sections.

## 3.1 The Climate/Atmosphere Data Store (CDS/ADS)

The Climate and Atmosphere Data Stores are European Union initiatives under the auspices of the Copernicus Programme. Development and operation of these services were awarded to ECMWF by the EU in 2015 and have been operational within the ECMWF Reading Data Centre since 2017. They are a highly-visible Internet-facing pair of services with some 30,000 registered users and growing. They have been especially heavily used during the COVID Pandemic by researchers investigating the interactions of virus propagation and climatic/atmosphere effects. Users have access to a large range of packaged tools and data sources providing them with high-performance, low-cost access to a wealth of data; however, they are not given unrestrained access to the cloud API and as such may not, for example, start their own dedicated virtual machines or containers.

The CDS/ADS is implemented using a collection of virtual machines, networks and large data volumes within an OpenStack Pike + Ceph Luminous cluster all orchestrated using Heat and automated with Puppet. Operation of the applications and their further development is under the control of small in-house teams at ECMWF augmented by a small number of contractors.

The total data stored within CDS/ADS' Ceph is in excess of 3 PB on a mix of hard disks (HDD) and solid-state disks (SSD). Data is distributed to maximise the benefit of the available SSD capacity, which represents 10% of the total storage capacity.

The service capacity is being expanded over the next six months by some 50%. In addition, it is being migrated to ECMWF's new data centre in Bologna. It will also be upgraded to OpenStack Stein during the migration process.

It will be joined by a small-scale development system that is otherwise functionally identical save for the addition of OpenStack's container (Zun), container networking (Kuryr) and container orchestration (Magnum) capabilities. It is expected that CDS/ADS will be re-engineered using this system into a hyper-scalable container-based application, and that in due course the production cluster will be upgraded to match the new developments.

The scale of the CDS/ADS system once expanded will be as summarised in Table 1:

| OpenStack - Type of Resource | Amount of Resource |
|---|---:|
| Servers | 104 |
| Cores | 5424 |
| Memory | 19,200 GB |
| NIC Speed | Multiples of bonded 25Gb |

| Ceph - Type of Resource | Amount of Resource |
|---|---:|
| Hard Disk (usable) | 7,007 TB |
| Solid-State Disk (usable) | 480 TB |
| NIC Speed | Multiples of bonded 25 Gb |

Table 1: CDS/ADS System Scale

The CDS/ADS service supporting infrastructure is supplied under a lease arrangement with a European provider of on-premises OpenStack cloud services. The provider provisions, maintains and manages the equipment and the OpenStack + Ceph software and provides a service to ECMWF for administrative level changes.

## 3.2 The WEkEO Data & Information Access Service (DIAS)

As part of the European Union’s Copernicus Programme, ECMWF, together with EUMETSAT and Mercator Ocean International, have joined forces to implement a Data and Information Access Services (DIAS) Platform called WEkEO.

WEkEO is a distributed cloud-computing infrastructure used to process and make the data generated by the Copernicus Environmental Services (CMEMS, CAMS, C3S and CLMS) accessible to users.

WEkEO also provides privileged access to derived products from these services and to all satellite data from the Copernicus Sentinel satellites operated by EUMETSAT and ESA.

Within the partnership, ECMWF is responsible for the procurement of the software to implement Data Access Services, Processing and Tools. ECMWF developed the requirements for the software and awarded and managed a contract for the implementation of those requirements to a European provider of software engineering services. The services have been integrated with the delivery platform by our partner EUMETSAT and their suppliers. The service is now live on the Internet.

The delivery platform uses OpenStack + Ceph. End-users are able to access a substantial portfolio of original Copernicus Programme and Sentinel satellite data using a harmonised data access mechanism and common data catalogue. They are able to harness compute resources to process that data without the networking and storage costs associated with public cloud offerings. They have access to a wide range of tools and technologies in the areas of DevOps, including data access, artificial intelligence, machine learning, workflow, compilation, build and more. The CDS/ADS data is also available through the WEkEO DIAS to allow bespoke processing of that data in unison with all other DIAS data using a common data access mechanism.

The underlying platform is delivered by EUMETSAT’s suppliers and is managed and operated on a daily basis by those suppliers. ECMWF hosts one of three peers of the system in its Reading Data Centre to provide high-speed networking access to the ECMWF hosted archived data. ECMWF does not manage or operate the OpenStack + Ceph cluster.

## 3.3 The pilot European Weather Cloud (pEWC)

The pilot European Weather Cloud is a joint initiative between ECMWF and EUMETSTAT sponsored by their respective Member States. The principal objectives of this initiative are to:

- deliver general-purpose compute resources (virtual machines, GPU-accelerated virtual machines, and orchestrated containers) as close *networking-wise* to the data archives as possible but accessible from anywhere on the Internet (for authorised evaluators only at this stage),
- act as a proof-of-concept on which use cases can be implemented and investigated and from which better decisions can be made about how to deliver a fit-for-purpose, value-for-money production-scale European Weather Cloud,
- act as a Morpheus test-bed for establishing ways of enhancing and simplifying the user experience when interacting with OpenStack, and

- inform decision-makers on the scale and complexity of the challenge of delivering such systems and thereby to assess how best to configure future manpower and other resources for operating such a production-scale system.

As an indicator of the scale of activity at this time, the following is a sample of the use case evaluations currently underway on the pEWC:

- Forecast and climatology of cloud cover for Energy & Spatial sectors
- Comparison of forecast and scatterometer derived surface wind
- Assessment of bias correction schemes for assimilation of radiance data
- Forecast data post-processing & Web Services using OGC standards
- Forecast data preparation for local RMI dispersion model
- Host climate explorer web application
- Mesoscale convective System tracking
- Nowcasting Collaborative Platform
- Meteosat Collaborative Platform
- Forecast data processing
- Storm surge model
- European Summer of Weather Code
- Radar data portal

The pEWC physical characteristics are shown in Table 2 (note the presence of GPUs in support of some use cases):

| **OpenStack - Type of Resource** | **Amount of Resource** |
|---|---:|
| Servers | 44 |
| Cores | 2920 |
| Memory | 21,000 GB |
| GPUs (NVIDIA TESLA V100 VOLTA architecture).[4] | On 5x2 compute servers |
| NIC Speed | Multiples of bonded 25Gb |

| **Ceph – Type of Resource** | **Amount of Resource** |
|---|---:|
| OSD Servers | 23 |
| Hard Disk (raw) | 668 TB |
| Solid-State Disk (raw) | 251 TB |
| NIC Speed | Multiples of bonded 25 Gb |

Table 2: Pilot European Weather Cloud System Scale

The OpenStack version deployed on the pilot is Ussuri as implemented by RedHat in their RHOSP product (note that the pilot uses the open-source variant Triple-O).

The Ceph version deployed is Nautilus. It is deployed separately from OpenStack and is accessible through public interfaces and isolated VLANs.

[4] Examples of acceptable specifications for the production system include: NVIDIA A100 TENSOR CORE GPU and AMD Instinct™ MI100 Accelerator

Per project build-on-demand Kubernetes clusters are supported and in extensive use by several of the abovementioned use cases.

The pilot is in successful operation serving the above use case evaluations and has been called into pre-production service to provide resources for Croatia during recovery from the earthquake that struck in March 2020.

The in-house team that built the pilot consists of two cloud engineers (~ 1.5 FTE), with significant other resources made available on an as-needed basis (~ 2 FTE). The cloud engineers are RedHat RHOSP trained up to advanced level.

Support to end-users is provided on a best-efforts basis, though this limitation is simply due to limited available manpower; use case evaluators are happy with the service they are receiving.

There is a smaller-scale test environment on which investigative work is performed, on which system upgrades are planned and rehearsed, and on which network integration tests are performed.

Planning is underway for migration of the service to Bologna when the new data centre becomes available. It is expected that the Bologna deployment will be one or more production-scale systems accompanied by suitable test and development environments and spread across the two service independent data halls.

## 3.4 Summary of Capabilities

In summary, in the area of OpenStack cloud technology, ECMWF has:

- industry-leading experience of building and operating HPC-scale data centres and all their associated systems and services,
- significant experience of building Internet-facing OpenStack-based cloud services at scale and supporting large (30,000+) user communities,
- substantial experience of procuring and/or leasing large-scale data processing systems, and
- significant experience of building pilot OpenStack cloud systems and delivering use case evaluation capabilities to tight budgets and timescales and to medium-scale user communities.

## 3.5 The Capability Goal

**ECMWF now wishes to extend its capacity for delivering on-premises OpenStack cloud computing technology at high-scale, fully production-ready, and supportable 24x365. In addition, it wishes to ensure that all workflows starting from a cloud's specification (i.e. servers, storage and networking) and ending in a running cloud are as automated as possible.**

The remainder of this ItPQ will cover a number of areas that ECMWF considers key to achieving the above **Capability Goal**.

## 3.6 Preparing Responses – General Guidance on Achieving Compliance

Responders will be given explicit compliance guidance throughout the remainder of this document on how to respond to each requirement.

However, as a general rule, ECMWF has tried to minimise the work involved for responders by allowing a simple confirmation that a requirement will be met where the responder considers that to be a sufficient response.

Where a responder wishes to provide an improved alternative to the stated requirement and/or has sought clarification and is responding to a clarified requirement, a fuller response will be required.

If a requirement is thought to be unsound or otherwise misguided, responders are advised to bring this to the attention of ECMWF during the response preparation phase. Including such a challenge in the formal response risks being found non-compliant.

**Please note that all further compliance guidance in the remainder of this document is given in green.**

# 4 Mandatory High-Level Requirements - General Remarks

This section describes the **mandatory high-level requirements (MHLRs)** that any on-premises production OpenStack cloud computing environment must meet in order to be acceptable to ECMWF and its Member States and Partners.

These **mandatory high-level requirements**:

- will be elaborated on in the remainder of this ItPQ,
- should be interpreted as a set of **minimum** requirements that must be met by any production deployment, and
- should also be interpreted as a guide to the **maximum investment** ECMWF is willing to make in such deployments. (For example, ECMWF does not intend to try and achieve *not-stop-operation* of the kind found in the carrier-grade telecoms sector; the expense would not be justifiable at this stage of maturity. Accordingly, ECMWF specifies a less stringent availability requirement.)

## 4.1 The Mandatory High-Level Requirements – Common Terminology

Throughout the remainder of this ItPQ, the terms cluster, platform, and environment will be used extensively. They shall have the following broad meanings:

- Cluster – a group of interconnected systems
  - sharing a common OpenStack Control Plane
    or
  - participating in Ceph as an OSD/MGR/etc
    or
  - sharing a common Kubernetes Control Plane,
- Platform – a group of inter-connected systems possibly supporting multiple clusters of OpenStack, Ceph, Kubernetes or other types of clusters (e.g. Kafka, Spark),
- Environment – the totality of physical resources and network services and the platforms that they rely on.

More specific elaboration will be given where needed, for example, where a situation does not fit neatly into the above categories.

## 4.2 The Mandatory High-Level Requirements - Summary List

All ECMWF **on-premises OpenStack cloud computing environments** shall:

- deliver a fully functioning set of OpenStack APIs to a **minimum** of the Ussuri release standard and including all the core components highlighted in orange in Figure 1 (**MHLR#1**),
- perform to an availability standard of *three-nines* in respect of unplanned downtime (**MHLR#2**),
- perform to an availability standard of *two-nines* in respect of planned downtime. In other words, all such environments shall be manageable and maintainable (but not disaster recoverable) within an annual aggregate *two-nines* downtime window of approximately one day per quarter (**MHLR#3**),

- be disaster recoverable to an availability standard of *two-nines* downtime within one single event per year (**MHLR#4**),
- be repeatably deployable from on-premises artefact repositories on factory-reset machinery using only automation tools – there shall be no manual intervention save for selecting cluster hardware membership and for initiating the automation processes (**MHLR#5**) (note that manual steps are acceptable where the cost of avoidance would be prohibitive),
- be repeatably (re-)deployable without access to the Internet (**MHLR#6**) (so-called isolated network deployment),
- be scalable up to the sizes shown in Table 3 without aggregate performance degrading (**MHLR#7**) (so-called law of diminishing returns),
- be compliant with modern systems networking practices including software-defined networking (SDN), multi-pathing everywhere, automatic failover and full IPv6 readiness (**MHLR#8**) (**for example, OpenStack must be able to operate irrespective of system sessions or other downtimes on a single network path within the underlying infrastructure)**,
- be capable of delivering differing hardware flavours according to user-specified needs (**MHLR#9**) (for example GPUs, extra-large memory configurations),
- be capable of providing differing storage flavours according to user-specified needs (**MHLR#10**) (for example, local ephemeral versus high-performance SSD-based Ceph),
- be capable of delivering differing networking topologies according to user-specified needs (**MHLR#11**) (so-called software-defined networking – SDN - at the user level),
- be constructed using industry-standard components and sub-systems, be configured to standardised practices and be introspectable for configuration correctness checking (**MHLR#12**) (to simplify data-centre equipment and cabling management),
- be *server-farm* based allowing the construction of individual clusters exclusively using software control - sometimes known as *metal management* (**MHLR#13**) (to avoid physical data centre changes when a new cluster needs to be built),
- be backed-up by a robust equipment supply chain composed of warm-standby, on-site, near-off-site and supplier resources sufficient to ensure the availability requirements are *comfortably* met (**MHLR#14**),
- be supported by appropriately skilled first-line (operations), second-line (diagnostics and in-depth support) and third-line (development) staff whether on-payroll or on a direct contract or on a support contract (**MHLR#15**),
- be compliant with all laws in all applicable jurisdictions with regards to data protection, prevention of misuse, storage of illegal material and all other relevant matters (**MHLR#16**),
- be segmentable in such a way that the Member States and Partners can pay for and obtain a guaranteed level of ring-fenced resources on-premises at ECMWF but integrated within their own environments (**MHLR#17**),

- be capable of delivering well-defined data integrity levels in line with industry standards (**MHLR#18**),
- be capable of uninterrupted operation throughout periods of essential maintenance on other (non-OpenStack) systems such as the enterprise network (**MHLR#19**),
- be capable of secure, audited operation at all times and of detecting and alerting most common forms of misuse (**MHLR#20**), and
- be capable of raising alerts and of being monitored in real-time (**MHLR#21**)

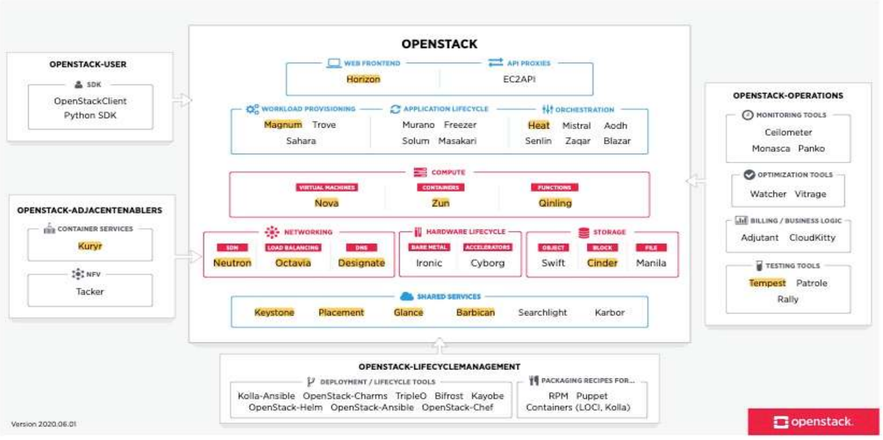


Figure 1: OpenStack Components – Mandatory Highlighted[5] [6]

| OpenStack - Type of Resource | Amount of Resource |
|---|---:|
| Servers | 500 |
| Cores | 25,000 |
| Memory | 100,000 GB |
| NIC Speed | Multiples of bonded 25 Gb |

| Storage[7] – Type of Resource | Amount of Resource |
|---|---:|
| Hard Disk (usable) | 40,000 TB |
| Solid-State Disk (usable) | 2,500 TB |
| NIC Speed | Multiples of bonded 25 Gb |

| GPU | Amount of Resource |
|---|---:|
| Examples of acceptable specifications include: NVIDIA A100 TENSOR CORE GPU and AMD Instinct™ MI100 Accelerator | 32 |

Table 3: Scale-Out System Scale

[5] DNS is a mandatory requirement; however, it may be implemented by means other than OpenStack Designate.

[6] Diagram is repeated in Appendix B for ease of reading

[7] Note that responders are encouraged to consider all storage options, but should include Ceph in their submissions

## 4.3 The Mandatory High-Level Requirements – Detailed Requirements

Each section below discusses in further detail each of the Mandatory High-Level Requirements mentioned above; Responders are requested to comment on the statements below and to add further details where they feel there are gaps in the analysis.

### 4.3.1 OpenStack APIs to Ussuri Release Standard or Later (MHLR#1)

ECMWF expects to commission its first **on-premises** production European Weather Cloud environment in Quarter 1 2022.

It wishes to ensure that the environment that it launches is up-to-date but not at the *bleeding-edge* of OpenStack developments.

Accordingly, it believes that the correct choice of OpenStack version is Ussuri (or Wallaby) since, at the time of writing, OpenStack Ussuri is described on openstack.org as Maintained now and entering Extended Maintenance in November 2021. Typically, that should imply a 3-year window of Extended Maintenance leading to an End-of-Life date of 2024.

There must be no deviations from the core functionality of the OpenStack Ussuri APIs. They must remain compatible with established standard tools in the industry, such as Terraform or Morpheus.

The base OS must be either CentOS 8 (CentOS 7 or earlier is not acceptable) or Ubuntu 20.04 LTS (18.04 LTS or earlier or any non-LTS release is not acceptable).

ECMWF wishes to ensure that it has in place a strategy for performing upgrades both to OpenStack and to Ceph – be it for functional reasons or support reasons. It also wishes to ensure that bug fixes and/or backports can be accommodated on a timely basis.

Responders are required at minimum to confirm that they "Will Comply" with the above.

Responders may wish to enhance their response by:

- verifying and/or clarifying ECMWF's understanding of OpenStack and supporting OS releases, and
- describing appropriate strategies for upgrading, bug fixing and/or backporting in OpenStack cloud environments.

### 4.3.2 Three-Nines Unplanned Downtime (MHLR#2)

ECMWF believes that three-nines availability free from **unplanned** downtime is appropriate at the infrastructure level for production cloud computing environments; this figure discounts failures of supporting systems such as networks, power sources and HVAC. ECMWF's experience to date with on-premises cloud computing suggests that this figure can be comfortably achieved. ECMWF further believes that higher availability should be designed into applications and based on deployment to multiple dissociated production cloud environments.

Responders are required at minimum to confirm that they "Will Comply" with the above.

Responders may wish to enhance their response by:

- providing guidance on *weak-link* areas to pay special attention to, including but not limited to smooth separation of responsibilities between suppliers and ECMWF and methods for measuring availability and avoiding disparity of measurements.

### 4.3.3 Three-Nines Planned Downtime (MHLR#3)

ECMWF believes that two-nines availability free from **planned** downtime is appropriate at the infrastructure level for production cloud computing environments; in the early stages of becoming a specialised cloud provider, ECMWF expects to need to carry out maintenance and development work on its production's environments for an aggregate one day on a quarterly basis with appropriate notice being given to users. Again, where higher availability is required, it should be designed into applications and based on deployment to multiple dissociated production cloud environments.

Responders are required at minimum to confirm that they "Will Comply" with the above.

Responders may wish to enhance their response by:

- confirming and/or caveating the extent to which this objective can be met. They are particularly encouraged to give their views on the complexity of upgrades and hot and/or interim fixes and on how these can best be factored into normal operations.

### 4.3.4 Two-Nines Disaster Recovery (MHLR#4)

OpenStack is a highly-interconnected and complex set of services employing databases, message buses, rich networking services, containers, processes, filesystems and much more. Usually, a failure results in a logged failure message and a service exiting in an orderly manner – a so-called fail-stop. But, this cannot be relied upon to happen with every failure. It is quite conceivable that a failure could result in irrecoverable corruption and/or disruption of an entire OpenStack + Ceph environment.

Equally, it is possible that a collection of servers, networking and storage could be damaged or destroyed by environmental effects such as fire, flooding or over-heating.

Where an action results in a completely destroyed production cloud environment, it must be possible to reconstruct that environment within four days for a single incident (assumes a destructive event has not destroyed the entire data centre, but at most one data hall).

Responders are required at minimum to confirm that they "Will Comply" with the above.

Responders may wish to enhance their response by:

- addressing particular areas where essential infrastructure data and/or configuration information must be backed-up and copied off-line in order to perform a successful disaster recovery.

### 4.3.5 Repeatably Deployable On-Premises Using Full Automation (MHLR#5)

ECMWF requires that every production cloud environment is deployable in a fully automatic manner. It further requires that such a deployment is repeatable – now and at any time in the future (subject to like-for-like equipment being available). Finally, it requires that every deployment is subjected to the fullest testing by automated use of OpenStack Tempest and associated tooling.

In order to achieve this, it is vital that:

- all artefacts involved in constructing a production cloud environment are harvested from the Internet and/or from local developments and saved in an on-premises, highly-available archived artefact repository,

- all processes, procedures, mechanisms, functions, scripts, tests, etc. are automated and that such automation artefacts are also archived in the above artefact repository,
- all documentation necessary to prepare for an automation run is professionally edited, rehearsed for correctness and archived,
- all post-deployment test-suite results are examined, approved (including allowed failures) and archived, and
- all future (re-)deployments are verified against the above archived post-deployment test-suite results for consistency.

By way of illustration, the artefact repository or repositories must be capable of handling **at least** the types of artefacts shown in Table 4.

| | | | | |
|---|---|---|---|---|
| git repositories | RPM packages | DEB Packages | PIP packages | Docker images |
| Wget fetched URLs | ISOs | QCOW2s | IMGs | RAW |

Table 4: Types of Artefacts Required to Build Production Clouds

Responders are required at minimum to confirm that they “Will Comply” with the above.

ECMWF believes that responders may wish to enhance their response by:

- giving a high-level discussion of how the above requirement might be achieved. They are encouraged to discuss:
    - all artefacts that must be available,
    - all repository types needed and tools available,
    - mechanisms for verifying newly proposed clusters without having to build them at full-scale, and
    - any DevOps toolchains that would simplify and accelerate such testing.

### 4.3.6 Repeatably Deployable Without Internet Access (MHLR#6)

Operating independently of Internet artefact sources improves build times in most cases – often quite substantially. Pre-caching also provides similar speedups, for example, building containers once and once only for each release level.

But, more importantly, it ensures (subject to thorough testing) that isolated operation is really possible – so-called sterile environment testing.

This is especially important in cases where some Internet sources are less than rigorous about versioning and retention. In other words, what works today may not work tomorrow.

More recently, we have seen examples where upstream providers are starting to question the sustainability of free access to all and are therefore limiting downloads quite severely.

This is not particularly onerous on storage requirements when compared with the size of the application data held in the European Weather Cloud. Accordingly, it is well worth the expense.

ECMWF recognises that isolated operation is a necessary condition for being able to build independently of Internet artefact changes, but that it is not a sufficient condition; version pinning is also essential and not fully enforced in every case where it is needed.

Responders are required at minimum to confirm that they "Will Comply" with the above.

ECMWF believes that responders may also wish to enhance their response by:

- describing how they would organise artefact repositories, version pinning and other associated technical aspects to ensure repeatable off-line building of production OpenStack environments.

### 4.3.7 Highly Scalable Without Performance Degradation (MHLR#7)

In Table 3, ECMWF shows the scale to which it expects to be able to expand any of its production cloud computing clusters, such clusters being a combination of OpenStack and Ceph.

In particular, it is important that up to the scale limit shown in Table 3, the addition of a further increment of resources does not cause the existing resources to degrade.

To illustrate by way of some examples:

- at 499 servers the addition of one more server must not cause an overall loss of performance across the full 500 servers. In other words, the addition of one more server should not render fewer aggregate resources than would be available were it not present,
- At 39 PB of HDD storage in Ceph, the addition of a further 1 PB must not cause an overall loss of performance across the full 40 PB of storage. In other words, there must be a further approximately 1 PB of storage available and the overall available I/O throughput must not be less than would be available were it not present.

Responders are required at minimum to confirm that they "Will Comply" with the above.

Responders may also wish to:

- comment on the achievability of this requirement
- indicate any choke points in such systems that may make reaching such scaling levels either impossible or prohibitively expensive.

### 4.3.8 Modern Networking Compliant (MHLR#8)

ECMWF, by moving its data centre to a completely new facility in Bologna, has invested heavily in achieving a modern, state-of-the-art network architecture which meets the following high-level requirements:

- **Virtualisation and cloud-native technologies:** ECMWF already provides services running in private and public clouds. It is, therefore, crucial to have an infrastructure that enables the use of and protects all services wherever they are hosted.
- **Scalability, reliability and performance:** it is essential to ensure that the network provides reliable connectivity with the highest possible bandwidth whilst being able to expand easily and quickly when required.

- **'Defence-in-depth':** the security challenges raised by modern IT environments require a different cybersecurity approach, in which defensive mechanisms are layered in order to protect valuable data and information.
- **Automation and orchestration:** the introduction of management tools will simplify the configuration and monitoring of the networks and security infrastructure, giving the capability to operate and configure the infrastructure remotely and enable faster deployment and operation of modern dynamic applications.
- **Modernity:** The use of a hybrid IPv4 and IPv6 will be the default. Omitting IPv6 will not be allowed.

The following are the main architectural elements of the new N&S design (see Figure 2 below):

- **'IP Fabric' architecture:** this is a state-of-the-art network architecture for medium- and large-scale data centres comprising two layers: leaf switches, to which systems connect, and spine switches, to which leaf switches connect. This architecture minimises delays and bottlenecks whilst offering greater scalability, reliability and performance.
- **Multi-site topology:** two physically segregated IP Fabric networks will be deployed in the new data centre: one in each data hall, thus creating two separate fault domains. This will significantly increase the availability of the resulting services as outages and maintenance sessions will impact only one hall at a time.
- **Security layer:** the segmentation of the data centre network into different security zones will offer higher control and visibility of data traffic. In addition, new security defence controls will be introduced to improve operational security and therefore, the ability to prevent and react to internal and external threats.

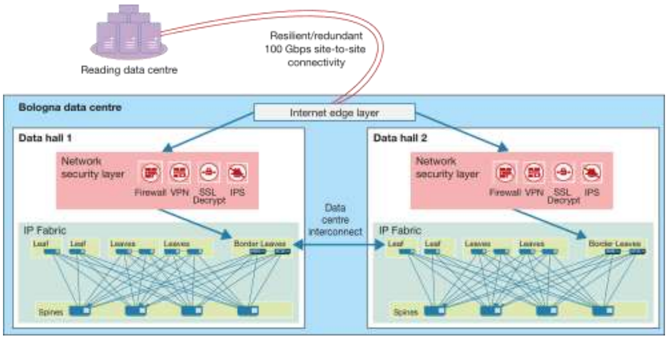


Figure 2: High-Level View – Bologna Network & Security Design

Responders are required at minimum to confirm that their offerings "Will Comply" with the above networking arrangements.

Responders may also wish to describe:

- how they would recommend integrating a large production OpenStack environment with the network architecture outlined above.

### 4.3.9 Differing Hardware Flavours (MHLR#9)

As standard, OpenStack users are given a choice of sizes for each of the virtual machines that they wish to create. For example, these may range from *1 core with 512 MB of memory* to say *32 cores with 512 GB of memory*.

However, during the European Weather Cloud pilot, ECMWF has learned that there is a strong desire among the various use case evaluators for GPU capabilities within some of their OpenStack provisioned virtual machines. This adds a second dimension to the sizing of OpenStack *flavours.*

Over time, we can expect to incorporate compute servers of increasing capacity and performance. However, it is unlikely that this will match the timing of decommissioning of older generation servers. Accordingly, it is likely that a mature cluster will have a range of compute server capabilities and these too must be incorporated into the OpenStack *flavour*'ing scheme.

In the future, there may be further types of hardware injection into virtual machines that are not presently foreseen. These must be accommodated with relative ease.

Finally, some users may require a complete physical compute server (so-called bare metal) as one of their machines. This too must be possible within the *flavour*'ing scheme.

Responders are required at minimum to confirm that their offerings "Will Comply" with the above flavouring flexibility requirements.

Responders may also with to:

- describe how they would present such flavour variety to users, and how they would lifecycle manage it with guidelines on when to decommission and/or refurbish older equipment.
- comment on *sweet-spot* server configurations for use in general-purpose cloud environments in terms of CPU (cores and speed), memory (density, quantity and speed), NICs (quantity and speed) and rack-density against the cost.

### 4.3.10 Differing Storage Flavours (MHLR#10)

To date, ECMWF's experience of OpenStack cloud storage in **production** has been with Ceph as block, object and file storage for the CDS system, and in **pre-production** has been with Ceph as block and object storage for the pEWC. The overwhelming majority of the storage is taken up with climatological, atmosphere and meteorological data. Only a small proportion represents code and/or virtual machine images and instances. There is essentially one integrated development team and a tightly co-ordinated development process. There is little scope for historical data accretion.

The European Weather Cloud is expected, at least initially, to be dominated by developments from many different teams with otherwise unconnected interests. As time goes by, it is expected that this will develop into a balanced mix of further developments and collaborations accompanied by systems in production. The scope for accretion of (no longer used) historical information is somewhat greater than with the CDS.

Ordinarily, storage usage would be controlled by real-time charging. However, this may not be possible with the European Weather Cloud. The best we may be able to do is set caps on storage use using OpenStack Quotas or, for S3 storage, using Ceph Quotas.

Accordingly, we want to default as much storage use as possible to so-called ephemeral storage – that is to say, storage that is released once an application exits. We do not see this as a particular inconvenience to developers provided that they are using automated orchestration of their cloud applications. Once they have satisfied their organisation's criteria for applications to enter production, they can easily make persistent those parts of their applications that require it.

ECMWF expects to provide at minimum, compute server local ephemeral storage, Ceph-based HDD storage and Ceph-based SSD storage. ECMWF strongly prefers avoiding proprietary storage systems unless there is a compelling reason for not doing so.

Responders are required at minimum to confirm that their offerings "Will Comply" with the above storage flexibility requirements.

Responders may also wish to:

- argue the case for alternatives to Ceph
- argue the case for proprietary storage solutions
- give a description of how they would present a range of storage options to users, and how they would encourage least cost usage in a quota-only (no real-time charging) environment.

### 4.3.11 Software Defined Networking (MHLR#11)

ECMWF's new Data Centre in Bologna will provide a fully redundant Spine-Leaf Network topology across two fully independent data halls (DHs) interconnected using Data Centre Interconnect (DCI) and switch technology from Juniper networks. Each Data Hall has separate power sources and can independently be brought down for maintenance or otherwise.

When implementing the production European Weather Cloud in Bologna ECMWF wishes to take advantage of the above architecture to provide high reliability of the service. For this reason, we expect to architect our production cloud environment so that the equipment spans both data halls in a way that leaves us with ~50% capacity still running if one data hall goes off-line. It should be noted that we will also provide a third zone of separation for hosting quorum members, but not large enough to accommodate mass storage servers or extensive compute servers.

In this case a "Will Comply" response will not be sufficient.

Responders are asked to propose industry-recognised solutions for achieving the above objectives; multiple cluster solutions may be considered provided they offer sufficient cross-cluster integration of storage, authentication and authorisation. As this is an area where it is possible to spend a great deal of money for diminishing returns, responders are asked to propose a minimum of architectures showing in each case the approximate cost and the associated benefits.

### 4.3.12 Data Centre Components, Sub-Systems, Practices & Introspection (MHLR#12)

ECMWF has extensive data centre architecture, design and construction experience; however, this has, in the past, been concentrated on HPCs and on supporting statically managed server and storage farms.

ECMWF now requires to maintain its high data centre standards as it embarks on constructing production cloud environments. In recognition of the scale to which it wishes to go, ECMWF now wants to better understand how to:

- standardise server and storage components without undue vendor lock-in over long lifecycles,

- standardise server and storage connectivity so as to flexibly accommodate different suppliers and generations of equipment over long lifecycles and without the need for special cases or workarounds,
- apply on-line verification through introspection that systems are functioning correctly and are interconnected with the environment correctly (without disruption to production systems), and
- manage the corresponding configuration artefacts within the ECMWF corporate configuration management systems Opsview (CMDB) and NetZoom (DCIM).

Responders are required at minimum to confirm that their offerings "Will Comply" with the above requirements.

Responders may also wish to:

- describe how they would architect and implement a data centre intended to accommodate multiple clouds at the scale of the cloud clusters described in Table 3.
- describe all necessary virtual and physical automation. (Please note that ECMWF does not believe it is yet necessary for it to apply robotic data centre asset management; however, Responders are invited to discuss this point if they wish, with particular reference to when it becomes the appropriate method.)

### 4.3.13 Server Farms & Soft Clusters – Metal Management (MHLR#13)

ECMWF does not expect to be able to pre-determine the exact size and shape of the various production cloud computing clusters that it may be called upon to commission. This may become clearer in due course; however, in the early stages, maximum flexibility is called for.

To this end, ECMWF wishes to take a software-defined approach to constructing clusters[8]. In particular, it wishes to rack equipment in a standard way and then use software-defined networking to interconnect that equipment into separate clusters. In other words, physical proximity in a rack does not necessarily imply logical proximity. To achieve this, without a high risk of disruption to running production clusters, we would need accurate and highly-available state information and management software to keep track of the soft configuration of the environment. Furthermore, we require to be able to add to or remove from a cluster without completely destroying it as an intermediate step – so-called incremental change.

**To be absolutely clear, ECMWF requires to be able to add or remove compute nodes to or from a running cluster without cluster control-plane downtime or other user interruption save for possible purging of virtual machines running on a node destined for removal.**

Responders are asked to describe an architecture and associated software and management procedures to achieve this so-called server-farm and soft-cluster approach. They should take into account the existence of the ECMWF corporate configuration management systems Opsview (CMDB) and NetZoom (DCIM).

### 4.3.14 Logistics & Equipment Supply Chains (MHLR#14)

ECMWF requires to operate production cloud environments with a level of hot-spare capacity racked-and-stacked (on-line) and ready to go into service at the command of an automation

[8] Sometimes referred to as Metal Management or Metal-as-a-Service.

step. It further requires to keep some level of spares on-site ready for hot-replacement at a moment's notice. Finally, it requires a secure supply chain upstream.

In the extreme, a level of say 10% to 15% of the capacity shown in Table 3 represents a substantial investment that might better be used in a different way.

Responders are required at minimum to confirm that their offerings "Will Comply" with the above requirements.

Responders may also with to:

- propose a supply chain that ensures the availability requirements are met while cost optimising the level of in-rack, on-site and supply-chain spares.

### 4.3.15 1st, 2nd and 3rd Line Support - Personnel (MHLR#15)

ECMWF is composed of a number of categories of staff with differing lengths of service contract. These staff categories are augmented by consultants who, on average, serve for shorter periods than staff. And then there is support from suppliers to augment staff and consultants usually in relation to highly-specialised details of particular products or services.

ECMWF wishes to gain a better understanding of where the boundaries should be placed between staff, consultants and suppliers in respect of architecting, building and operating production cloud environments.

Responders are asked to describe a model in which production cloud environments can be architected, implemented and operated using staff, consultants and suppliers and to give numbers and educational/experience levels of said staff, consultants and suppliers.

### 4.3.16 Regulatory & Compliance (MHLR#16)

At the time of writing, ECMWF is expecting to be providing a production European Weather Cloud **to itself, Member States and Partners only.** The service is not expected to be available to all comers from the Internet – though it will be available over the Internet for authorised users.

ECMWF is soliciting legal opinions under separate cover on its duties and responsibilities as well as its potential exposures and liabilities as a general-purpose cloud computing provider – albeit one that is dedicated to the needs of the meteorological, atmospheric and climatological communities.

However, there are technical aspects to these issues that ECMWF wishes to address in this ItPQ.

Responders are asked to give an opinion on the main areas of concern for any such providers of services, including but not limited to:

- firewalling and intrusion detection of
  - the core infrastructure
  - users applications
  - inter-tenancy privacy
- storage and compute segregation of
  - data from differing jurisdictions

    - compute resources within differing jurisdictions
- logging and audit of user activity for compliance with
    - legal requirements
    - insurance requirements
    - post-event investigations

It must be stressed that the above list is not definitive; it is intended to give the flavour of the concerns that ECMWF wants to address. Responders are asked to point out any areas not mentioned and to comment on the duties and responsibilities implied.

### 4.3.17 Segmentable & Ring-Fenced (MHLR#17)

ECMWF wishes to be able to accommodate the needs of those of its Member States and Partners who wish to place dedicated resources within the ECMWF production cloud environment so as to have a guaranteed amount of compute & storage resources that are *networking-wise* near to the ECMWF data archive.

The Member States or Partners may wish to:

- buy a physical quantity of compute resources and storage resources,
- place them within the ECMWF data centre in Bologna,
- commission them as a part of one or more cloud platforms within the ECMWF cloud environment,
- have them operated as an integral part of the ECMWF production cloud environment,
- have them maintained as an integral part of the ECMWF production cloud environment,
- have them lifecycle managed as an integral part of the ECMWF production cloud environment.

The Member States or Partners may also wish to embed their own staff within ECMWF's cloud team(s) for periods of training and knowledge transfer for the purpose of establishing meteorological/atmospheric production cloud environments on their own territory.

Responders are asked to give their views on the following aspects of the above requirement:

- The technical mechanisms that would allow compute and/or storage resources to be added to a running cloud platform whilst at the same time ring-fencing the added resources to certain account holders and/or projects.
- The technical mechanisms that would allow the Member States and Partners to establish and manage their own range of user identities within a Keystone domain. This would be a delegated authentication mechanism whereby the user first obtains his general access rights to the full environment and then obtains his limited access rights to the ring-fenced resources – all done with one sign-on action.

- The appropriate mechanisms for performing knowledge transfer and validation for the purpose of establishing similar environments on Member State and Partner territories.
- Possible models for costing such arrangements, both in terms of manpower and equipment.

### 4.3.18 High-Integrity Data Storage (MHLR#18)

All data storage and supporting systems must meet well-recognised industry standards. It must be clear to users how their data assets are stored within the European Weather Cloud. It must also be clear how those storage systems are managed and the level of integrity they provide.

The purpose of this is not to achieve an unaffordable level of data integrity; instead, it is to make it possible for users to make an informed decision about what other off-line arrangements might be necessary for their most valuable data assets (if any).

Responders are required to confirm that their offerings "Will Comply" with the above requirements.

### 4.3.19 Resilient Operation Under non-OpenStack Failure Conditions (MHLR#19)

All interactions between OpenStack clusters and the hosting environment (network & power) shall be sufficiently redundant to permit continued operation under single-failure conditions. For example, all network paths shall be dual-redundant; all power sources shall be dual-redundant.

Responders are required to confirm that their offerings "Will Comply" with the above requirements.

### 4.3.20 Secure By Default –Secure Audited Operation & Misuse Detection (MHLR#20)

Upon deployment, all systems (e.g. OpenStack, Ceph, Kubernetes, the underlying server farm, the SDN switches and more) shall be secure by default and *wired-in* to the corporate monitoring and alerting systems.

ECWMF carries out security practices in accordance with its own security policies and procedures. These are written to cover components (for example servers, switches, storage) as well as major systems and are automatically applied on deployment. In many cases, they are being applied to systems that are already accessible to Member States and Partners over dedicated networks and the Internet. ECMWF assumes an attack profile commensurate with its mix of operational and research activities.

ECMWF wishes to ensure that in deploying large cloud environments of 100s of compute nodes and potentially 100s of PBs of mass data storage, it is extending its security practices in ways appropriate to such environments.

Furthermore, whilst at this stage, the users of the system will all be registered staff of the Member States and Partners, the service will, nevertheless, be primarily accessed over the Internet and open to the attacks that are typical of such cloud systems.

ECMWF wishes to put in place appropriate measures for cloud environments such as defences against:

- Data breaches

- Account hijacks
- Insider threats
- Malware injection
- Leakage between the cloud service layer and infrastructure and the control plane layers
- Adverse effects of processor sharing across multiple tenants
- Cloud service abuse and resource starvation
- Insecure APIs
- Denial-of-service attacks
- Due diligence especially in experimental contexts
- Shared responsibilities between ECMWF and its Member States and Partners
- Data loss by ECMWF of Member State and Partner data

Responders are asked to describe their experience of designing and implementing security of cloud components in general and cloud environments in particular.

### 4.3.21 Monitoring and Alerting (MHLR#21)

ECMWF is most concerned to protect its reputation built up over many years for running high-quality services on behalf of its Member States and Partners. To this end, it is essential that all aspects of the production cloud environment are engineered and orchestrated to the best-of-breed standards in the industry. As an integral part of this ECMWF wishes to implement monitoring, reporting and alerting that is:

- timely
- accurately targeted
- highly-available
- filtered and layered for quick and easy diagnosis of most issues
- automatically escalated where necessary to appropriate levels of support

ECMWF believes that the best of breed software in the cloud industry for achieving this is a combination of Prometheus, Thanos and Grafana. ECMWF also wishes to integrate these tools into its Operations Centre running Opsview & Splunk.

Responders are asked to give their view of how best to architect monitoring, reporting and alerting environment, which products to choose and how best to integrate them with operations, system support, consultancy support and the established monitoring environment at ECMWF mentioned above (OpsView & Splunk).

## 5 Technical Requirements - General Remarks

This section addresses a number of requirements that are considered consequent to achieving the mandatory high-level requirements (MHLRs). This does not mean that they are less important, rather that they reflect the various ways in which the MHLRs could be met.

In summary, they are as follows:

- choice of OpenStack distribution and deployment method (**TR#1**)
- choice of Ceph distribution and deployment method (**TR#2**)
- choice of Kubernetes distribution and deployment method (**TR#3**)
- containerisation using OpenShift / OKD (**TR#4**)
- operations and maintenance (**TR#5**)
- training (**TR#6**)
- documentation (**TR#7**)
- Failures, Fault Analysis, Fixing & Patching (**TR#8**)
- resource accounting and charging (**TR#9**)
- cloud environment development & supporting DevOps environment (**TR#10**)
- segmentation and zoning of ring-fenced resources (**TR#11**)

Each is considered further in the following sections.

### 5.1 Choice of OpenStack Distribution & Deployment Method (**TR#1**)

ECMWF is aware of the following OpenStack deployment mechanisms (in alphabetical order):

- Airship
- Canonical OpenStack
- Debian OpenStack
- Kayobe
- Kolla-Ansible
- Mirantis Cloud Platform (MCP)
- OpenStack-Ansible
- RedHat OpenStack Platform (RHOSP)
- StarlingX
- Triple-O (OoO)

Furthermore, ECMWF is aware that there is substantial variability in the extent to which each of the above is fully production hardened and backed-up by production-grade commercial support.

ECMWF wishes to stress that it has not, at this stage, committed to any particular deployment mechanism (or flavour of mechanism); it is keeping an open mind. That said, ECMWF has some experience of OpenStack deployment, which will now be described.

ECMWF has experience of using Triple-O to build the pilot European Weather Cloud. This project has met its mandate and so Triple-O is trusted by ECMWF for building such systems. ECMWF does not have experience of running a Triple-O OpenStack in a production environment.

Kolla-Ansible is believed to be in widespread use in production environments with Kolla in support for building the container images. ECMWF has some experimental experience of using Kolla-Ansible and Kolla.

OpenStack-Ansible is in use by a large on-premises cloud provider and as such is believed capable of delivering production-grade systems.

ECMWF would like to better understand the advantages and disadvantages of the various offerings mentioned above.

The following sections give further guidance on ECMWF's interests in these various offerings.

### 5.1.1 OpenStack Deployment using Triple-O / RedHat OpenStack Platform (RHOSP) (TR#1a)

ECMWF currently uses Triple-O, though at the start of the pilot project it gained some experience of RHOSP during a small-scale evaluation and during RedHat training courses.

ECMWF wants to better understand where Triple-O and RHOSP fit in the overall landscape of OpenStack deployment methods. It wants to better understand their advantages and disadvantages with respect to each other and with respect to the other offerings.

**Responders who propose Triple-O or RHOSP** are asked to answer the following questions:

- Are Triple-O and RHOSP of production quality? If so, what level of engineering skill and effort is required to deploy them?
- Are Triple-O and RHOSP (at their core) fundamentally the same product perhaps with differing release cycles?
- What are the essential differences between Triple-O and RHOSP? (consider such issues as coverage of OpenStack components and open source versus closed source)
- What is the support mechanism for Triple-O compared with the support mechanism for RHOSP?
- What is the full lifecycle cost of supporting a Triple-O deployment compared with the full lifecycle cost of supporting a RHOSP deployment keeping in mind the sizes to which ECMWF wishes to scale its cloud computing environments (see Table 3)?
- Do Triple-O and RHOSP comply with the core components highlighted in orange in Figure 1? If not, can the differences be reconciled with additional engineering and at what cost?
- What are your overall impressions of Triple-O and RHOSP?

### 5.1.2 OpenStack Deployment using Kolla-Ansible & Kolla (TR#1b)

ECMWF wishes to establish what would be involved in using Kolla-Ansible & Kolla directly to build the production European Weather Cloud environment.

**Responders who propose Kolla-Ansible & Kolla** are asked to answer the following questions:

- Are Kolla-Ansible & Kolla of production quality? If so, what level of engineering skill and effort is required to deploy them?
- Are Kolla-Ansible & Kolla available as commercially supported products for customer deployments on-premises?
- What community support is available for Kolla-Ansible & Kolla? What level of internal and/or consultancy engineering skill is required to interface between a production environment and the support community?
- What is the full lifecycle cost of supporting a Kolla-Ansible & Kolla deployment keeping in mind the sizes to which ECMWF wishes to scale its cloud computing environments (see Table 3)?
- Do Kolla-Ansible & Kolla comply with the core components highlighted in orange in Figure 1? If not, can the differences be reconciled with additional engineering and at what cost?
- What are your overall impressions of Kolla-Ansible & Kolla?

### 5.1.3 OpenStack Deployment using OpenStack-Ansible (TR#1c)

ECMWF has no operational experience of OpenStack-Ansible; however, it is considered one of the official deployment mechanisms for OpenStack.

ECMWF wishes to assess the merits of OpenStack-Ansible without the need for a full pilot activity.

**Responders who wish to propose OpenStack-Ansible** are asked to answer the following questions:

- Is OpenStack-Ansible of production quality? If so, what level of engineering skill and effort is required to deploy it?
- Is OpenStack-Ansible available as a commercially supported product for customer deployments on-premises?
- What community support is available for OpenStack-Ansible? What level of internal and/or consultancy engineering skill is required to interface between a production environment and the support community?
- What is the full lifecycle cost of supporting an OpenStack-Ansible deployment keeping in mind the sizes to which ECMWF wishes to scale its cloud computing environments (see Table 3)?
- Does OpenStack-Ansible comply with the core components highlighted in orange in Figure 1? If not, can the differences be reconciled with additional engineering and at what cost?
- What are your overall impressions of OpenStack-Ansible?

### 5.1.4 Other Strong OpenStack Deployment Candidates (TR#1d)

Where a responder believes there is an alternative deployment mechanism not mentioned above that ECMWF should consider in-depth as a possible contender for production use, they are asked to provide answers in the spirit of the questions posed in Sections 5.1.1 through 5.1.3.

## 5.2 Choice of Ceph Distribution & Deployment Method (TR#2)

ECMWF has built several Ceph pilot cells over several years with a reasonable degree of success. The most recent Ceph build of approximately 1PB of usable storage has been in operation for more than a year without outages. It is supporting the ongoing pilot European Weather Cloud use case evaluations.

On ECMWF's behalf, the supplier of the CDS infrastructure runs a 5 PB scale Ceph cluster in production, for which they have sole operational responsibility – i.e. hardware and software maintenance, load balancing, and monitoring and alerting.

ECMWF now wishes to build its own Ceph **production** clusters for general-purpose use, but at minimum sufficient to support OpenStack clusters and/or Kubernetes clusters.

In the past, ECMWF has used basic Ceph tools to perform deployments. However, this is labour intensive, time-consuming and error-prone, especially when a rich set of options is required to be switched on. It is possible to seek to automate these tools; however, this has already been done with Ceph-Ansible.

Accordingly, ECMWF is now seeking advice on the best deployment tools for deploying Ceph.

ECMWF has some limited experience of is Ceph-Ansible.

Responders who propose Ceph-Ansible are asked to answer the following questions:

- Is Ceph-Ansible of production quality? If so, what level of engineering skill and effort is required to deploy it?
- Is Ceph-Ansible available as a commercially supported product for customer deployments on-premises?
- What community support is available for Ceph-Ansible? What level of internal and/or consultancy engineering skill is required to interface between a production environment and the support community?
- What is the full lifecycle cost of supporting a Ceph-Ansible deployment keeping in mind the sizes to which ECMWF wishes to scale its cloud computing environments (see Table 3)?
- Does Ceph-Ansible comply with the core components highlighted in orange in Figure 1? If not, can the differences be reconciled with additional engineering and at what cost?
- What are your overall impressions of these Ceph-Ansible?

Responders who wish to propose an alternative deployment mechanism to Ceph-Ansible are asked to describe their proposed mechanism and to answer the previous six questions in the context of their proposed mechanism.

## 5.3 Choice of Kubernetes Distribution & Deployment Method (TR#3)

ECMWF has experience of building and operating Kubernetes clusters in support of development and production using **bare-metal** clusters of servers. This has been successful in the context of internal use cases.

Now that ECMWF is embarking on general-purpose compute service provision to all its Member States and Partners, the breadth of use cases has widened considerably.

Early indications from the European Weather Cloud pilot programme are that users want to be able to create their own small-scale single-tenant Kubernetes clusters and to have full administrative control of them.

In due course, we expect this to moderate in favour of a mix of self-managed small-scale single-tenant clusters and large-scale multi-tenant production clusters; the former running on virtual machines in the cloud, the latter running on bare metal managed as a part of the server farm.

ECMWF wishes to establish the appropriate technology for deploying Kubernetes in both the above use cases – small-scale single-tenant and large-scale multi-tenant.

ECMWF has some prototype experience of using Rancher for the deployment of Kubernetes on OpenStack.

ECMWF also has some development experience in the use of Kubespray as a standalone deployment tool for Kubernetes and with the OpenStack plugin as a cloud deployment tool for the same.

Furthermore, as a part of the CDS next phase, ECMWF will be assessing the use of OpenStack Magnum as an appropriate mechanism for deploying Kubernetes on OpenStack.

**Responders who wish to propose Rancher** are asked to answer the following questions:

- Is Rancher of production quality? If so, what level of engineering skill and effort is required to deploy it?
- Is Rancher available as a commercially supported product for customer deployments on-premises?
- What level of community support is available for Rancher? What level of internal and/or consultancy engineering skill is required to interface between a production environment and the support community?
- What is the full lifecycle cost of supporting a Rancher deployment keeping in mind the sizes to which ECMWF wishes to scale its cloud computing environments (see Table 3)?
- What are your overall impressions of these Rancher?
- What other deployment technologies in this class would you recommend we consider (for each such technology, please also give answers to the preceding five questions)?

**Responders who wish ro propose Kubespray** are asked to answer the following questions:

- Is Kubespray of production quality? If so, what level of engineering skill and effort is required to deploy it?
- Is Kubespray available as a commercially supported product for customer deployments on-premises?
- What level of community support is available for Kubespray? What level of internal and/or consultancy engineering skill is required to interface between a production environment and the support community?

- What is the full lifecycle cost of supporting a Kubespray deployment keeping in mind the sizes to which ECMWF wishes to scale its cloud computing environments (see Table 3)?
- What are your overall impressions of these Kubespray?
- What other deployment technologies in this class would you recommend we consider (for each such technology, please also give answers to the preceding five questions)?

**Responders who wish to propose Magnum** are asked to answer the following questions:

- Is Magnum of the same production quality as the OpenStack core components? If so, what level of engineering skill and effort is required to deploy and support it?
- What level of community support is available for Magnum? What level of internal and/or consultancy engineering skill is required to interface between a production environment and the support community?
- What are your overall impressions of Magnum?

**Responders who wish to propose an alternative to the above mentioned systems** should provide a brief description of their proposal and should answer questions about their proposed system in the spirit of the above questions.

## 5.4 Containerisation using OpenShift / OKD (TR#4) (OPTIONAL)

**Optionally,** responders who wish to are invited to give an analysis of the relative merits of using Kubernetes and its associated deployment tools versus using OpenShift / OKD and its associated deployment tools. The analysis should emphasise, but not be limited to:

- the common preferences of developers
- the common preferences of operators
- the richness of available tools and DevOps platforms
- the completeness as a container platform for application development and deployment
- the security of the platform
- the authentication and authorisation mechanisms
- the possibility to deploy OpenShift on OpenStack for small-scale individual containerisation clusters

## 5.5 Operations & Maintenance (TR#5)

ECMWF wishes to ensure that as much of the day-to-day operation of the production cloud environment is managed either fully automatically (perhaps with self-service) or by human initiation of automatic processes and procedures. Some day-to-day operations may be delegated to group focal points within internal development groups, Member State operations groups and/or Partner operations groups.

Responders are asked to comment on the following example areas in respect of processes, procedures and tools:

- First-line:

  - automatic management of user accounts and projects (i.e. add, change, disable, delete)
  - automatic management of quotas (i.e. increase, decrease, diagnose, etc.)
  - automatic management of access to shared (but protected) resources (read, update, create, etc.)
- Second-line:
  - assist with cloud usage questions
  - provide cloud automation and orchestration tool support
- Third-line:
  - provision a cluster from the server farm
  - add or remove a compute server to/from a running production OpenStack cluster
  - assist with requirements capture for new capabilities

Responders may, **optionally**, wish to take a much broader view of the area of Operations & Maintenance and provide appropriate discussions of same.

### 5.6 Training (TR#6)

ECMWF wishes to place great emphasis on efficiency and correctness through automation and orchestration. However, it also wishes to provide a rich set of training resources so that the various levels of staff involved in providing the production cloud service are not just *pushing buttons*.

ECMWF also regards it as essential to provide a development path for all concerned from 1st line support at the helpdesk through to 3rd line support as Site Reliability Engineers (SREs).

Accordingly, we want to develop, and maintain up-to-date, a core set of training materials underpinning OpenStack, Kubernetes and Ceph, but customised to the ECMWF cloud environment.

Responders are asked to propose a combination of classroom, laboratory and self-training materials – together with proficiency verification – to support the production on-premises cloud environment.

### 5.7 Documentation (TR#7)

ECMWF wishes to establish a set of standards for the development of its on-premises production cloud capability. To this end, it wishes to focus documentation on those standards.

Documentation being notoriously difficult to keep up-to-date with the state of the software it supposedly relates to, we wish to develop documentation as far as possible directly from the source code for the running production environment. Again, this calls for standards of architecting, coding and testing.

Responders are asked to give their views on how best to maintain an up-to-date documented status of a running production cloud environment like the one planned by ECMWF.

### 5.8 Testing, Failures, Fault Analysis, Fixing & Patching (TR#8)

ECMWF wishes to:

- set a minimum standard for testing and verifying all systems that are planned for release into production. At a minimum, this will involve using OpenStack Tempest to the widest extent possible,

- monitor for failures to the fullest extent possible. Where reasonable, all sub-components of a complete cloud shall report good health rather than be detected as failed; where this is not possible, external failure detection shall be used,
- be in a position to perform most fault analysis on-premises. This calls for on-premises monitoring and reporting as described in Section 4.3.21 It also calls for sufficient training as described in Section 5.6,
- be able to fix and/or patch on the basis of the above fault analysis using on-premises resources; again, this heavily depends on the training described in Section 5.6.

Responders are asked to comment on the challenges presented by the above objectives, the likely timescale for implementation, and the level of external support required in the meantime, and eventually in the steady-state.

## 5.9 Resource Accounting & Charging (TR#9)

ECMWF has no plans at present to real-time charge for the use of production cloud environment resources. However, it is keen to start with basic resource accounting as a means of monitoring usage for the purposes of capacity planning. In addition, it may become necessary to more accurately account for individual/group use of resources.

Accordingly, ECMWF wishes to implement at an early stage of the project appropriate mechanisms within OpenStack and Ceph to allow resource accounting to be performed.

Responders are asked to comment on how they would perform resource accounting with particular regard to the following questions:

- What tools would you use?
- How would you aggregate and summarise the resulting data?
- What raw data would you archive for future analysis?
- What impact would such mechanisms have on the overall performance of the cloud?

## 5.10 Fully Automated Cloud & Supporting DevOps Environments (TR#10) (OPTIONAL)

Ideally, ECMW will arrive at a point where it can automatically command into existence an OpenStack cloud cluster simply by choosing from a menu of options[9].

For example, a small test environment may be chosen to look as follows:

- Non-HA Control Plane
- three compute nodes
- one network node
- one storage node on LVM
- Core-OpenStack + Magnum + Kuryr

On clicking *go:*

- a suitable set of hardware would be chosen and provisioned from the farm,

[9] Limited to Cloud Operations Staff and not end users

- a set of IP ranges would be chosen from pre-authorised ranges
- a software build would be started, and, in due course
- a small-scale OpenStack platform would be up and running and ready for whatever purpose it is intended.

Or, a large production environment may be chosen to look as follows:

- HA Control Plane with separate database nodes
- 30 compute nodes
- three network nodes
- a 20 server Ceph cluster
- Core-OpenStack with all add-ons, full monitoring, accounting, etc

On clicking *go:*

- a suitable set of hardware would be chosen and provisioned from the farm,
- a set of IP ranges would be chosen from pre-authorised ranges
- a software build would be started,
- a full tempest test and validation run be started,
- the cluster would pass/fail the testing gate, and if passed
- a large-scale OpenStack platform would be up and running and ready for production use.

Of course, such a capability pre-supposes a substantial inventory of equipment already pre-deployed and a rich automation environment to support it. For ECMWF, that is still some way off. However, the purpose of this section is to motivate a discussion of what it would take to achieve such a rich capability.

**Optionally**, responders are invited to blue-sky think through the general architecture and the high-level components that would need to be in place to achieve such a capability. They are also asked to comment on the financial viability of achieving such a capability.

## 5.11 Segmentation and Zoning of Ring-Fenced Resources (TR#11)

ECMWF operates a business model with its Member States whereby they perform a significant proportion of their computing needs on equipment housed on ECMWF's premises. For example, the HPCs are shared between ECMWF's needs and those of its Member States.

It is to be expected that this model will carry over to the European Weather Cloud environment and that it will be possible to guarantee each Member State, according to prior agreements, the amount of resource available to them uncontested by other user's demands.

It may even be necessary to dedicate segmented/zoned parts of cloud environments for use by particular Member States and not available to other Member States. Such a configuration would involve a Member state sharing the control plane, underlying data centre network and other data centre services, but having dedicated storage and compute resources that they have paid to have placed in the ECMWF data centres.

Responders are asked to advise on appropriate mechanisms for achieving this balance of shared and/or dedicated resources.

## 5.12 Architectures Types – Cloud Efficiency (TR#12) (OPTIONAL)

ECMWF wishes to ensure that it follows a path of maximum cloud efficiency. In this regard, ECMWF is concerned to ensure that it considers all possible server and technology options that may reduce costs, **help in** improving overall data centre efficiency and reduce environmental impact. Please note that this section does not address the wider **overall** efficiency considerations associated with modern data centre design.

**Optionally**, responders who wish to respond on this matter, are asked to consider whether there are alternative server/CPU architectures and/or storage architectures that ECMWF should be considered for construction of OpenStack clusters.

# 6 Indicative Pricing

## 6.1 General Approach to Pricing

**ECMWF wishes to gain a better understanding of the relative costs of the various options open to it**. For example:

- Should equipment be leased or purchased?
- What balance of manpower between on-payroll, consultant and supplier services is most appropriate?
- What are the best logistics options consistent with maintaining a production quality service?

**The exact scale of the production environment is not yet known** and so unit pricing is an appropriate approach.

**The distribution of work among on-payroll, consultant and supplier services is not known** and responder's opinions are sought; accordingly, general indications of cost are appropriate.

**The level of on-site spares is not known and responder's opinions are sought**; accordingly, indicative levels of spares and their associated cost are appropriate.

**The indicative pricing sought in this section will be used to:**

- validate budget planning within ECMWF and its partners
- compare different market offerings
- compare different technical solutions

## 6.2 Manpower Types, Amounts and Pricing

Responders should indicate the consultant and/or supplier service manpower costs as follows:

- Levels of individual(s)
- Numbers of individual(s)
- Purpose of individual(s) services in the context of the requirements stated in sections 2, 3, 4 & 5
- Likely duration for which individual(s) services will be required as a percentage of full-time
- Level of round-the-clock or on-call cover required
- Urgency with which individual(s) services may need to be called upon
- Most appropriate contracting terms for individual(s)

In addition, any costs arising from the external management of such individual(s) should be indicated, together with the relevant reporting structure.

## 6.3 Equipment Pricing

ECMWF recognises that different responders may wish to take different approaches to the provision of physical equipment. In particular, it recognises that some may wish to propose a lease[10] arrangement, whereas others may wish to propose outright purchase. Both approaches will be considered. That said, the basic approach to pricing must take account of the uncertain scale (at this time) of the requirement. Accordingly, ECMWF requires pricing to be based on units of resource including at minimum:

- Processing Capacity

  To be expressed in € / core in such a way that it is clear what constitutes a core. For example, responders may wish to take a well know model of central processing unit (CPU) (either Intel or AMD) as an example and express their definition of cores in terms of such a processor.

- Memory Capacity

  To be expressed in € / gigabyte. Responders may wish to mix memory of differing speeds and/or densities in which case each speed or density should be separately priced. Variations in performance and/or density are discourage unless there is a clear price differentiation to justify it.

- Hard Disk Drive Capacity

  To be expressed in € / terabyte. Responders may wish to mix hard disks of differing speeds and/or densities in which case each speed or density should be separately priced. Variations in performance and/or density are discourage unless there is a clear price differentiation to justify it.

- Solid State Drive Capacity

  To be expressed in € / gigabyte. Responders may wish to mix solid state drives of differing speeds and/or densities in which case each speed or density should be separately priced. Variations in speed and/or density are discouraged unless there is a clear price differentiation to justify it.

  ***NB: Responders are also advised to keep a clear separation in speed and/or density between top-end Hard Disk Drives and bottom-end Solid State Drives.***

- Graphics Capacity

  To be expressed in € / VM-attached-GPU.

  Examples of acceptable specifications include: NVIDIA A100 TENSOR CORE GPU and AMD Instinct™ MI100 Accelerator

  Responders may wish to mix differing specifications of GPU and should individually price said mix. Excessive variations in specification are discouraged unless there is a clear price differentiation to justify it.

- Network Switches (Data & Out-of-Band)

  At this stage, responders are encouraged to absorb the cost of said equipment into the first five pricing categories above.

  At the eventual ItT stage ECMWF envisages allowing more flexible pricing of such equipment.

[10] Responders should indicate the kind of lease proposed.

- Racks, switchgear and power distribution

  At this stage, responders are encouraged to absorb the cost of said equipment into the first five pricing categories above.

  At the eventual ItT stage ECMWF envisages allowing more flexible pricing of such equipment.

- On-line, On-Site and Supply Chain Spare Parts
  Where it is appropriate to the operational model proposed in 4.3.14, responders are asked to price the recommended levels of spares to be held on-line, on-site and/or readily available within the supply chain. Responders should clearly price as a percentage of the first five pricing categories above.

## 6.4 Software Pricing

### 6.4.1 Open Source

ECMWF has a strong preference for constructing all of its cloud environemtns using open source software, prefereabply backed-up by an active support community. In all cases though, responders should indicate the licence under which the software is distributed. If possible, responders should estimate the manpower costs of using and supporting said software in an in-house production cloud environment.

### 6.4.2 Community Editions

Where community editions of products are available for non-profit use, these should be preferred over enterprise editions carrying licensing fees.

### 6.4.3 Enterprise Editions

Enterprise editions will be considered where there is functionality that is of demonstrable value that is not included in the community edition. In such cases, license fees shall be made clear in terms of licensing term, licensing volume and commercial support[11].

[11] Please note that ECMWF has an active Morpheus licence and may choose to continue this arrangement outside of this ItPQ and any subsequent ItT

# Appendix A Cross-Reference to Responders Documentation

This appendix contains a table listing each of the Mandatory High-Level Requirements **(MHLRs)** by section number and title and each of the Technical Requirements **(TRs)** also by section number and title. Reproduced with each is the response guidance contained in the body of the document. Responders are encouraged to annotate this table with references to the appropriate sections of their responses.

| |
|---|
| 4.3.1 OpenStack APIs to Ussuri Release Standard or Later (**MHLR#1**) |
| Responders are required at minimum to confirm that they "Will Comply" with the above.<br>Responders may wish to enhance their response by:<br>- verifying and/or clarifying ECMWF's understanding of OpenStack and supporting OS releases, and<br>- describing appropriate strategies for upgrading, bug fixing and/or backporting in OpenStack cloud environments. |
| Will Comply? YES / NO |
| Enhanced Response: |
| 4.3.2 Three-Nines Unplanned Downtime (**MHLR#2**) |
| Responders are required at minimum to confirm that they "Will Comply" with the above.<br>Responders may wish to enhance their response by:<br>- providing guidance on *weak-link* areas to pay special attention to, including but not limited to smooth separation of responsibilities between suppliers and ECMWF and methods for measuring availability and avoiding disparity of measurements. |
| Will Comply? YES / NO |
| Enhanced Response: |
| 4.3.3 Three-Nines Planned Downtime (**MHLR#3**) |
| Responders are required at minimum to confirm that they "Will Comply" with the above.<br>Responders may wish to enhance their response by:<br>- confirming and/or caveating the extent to which this objective can be met. They are particularly encouraged to give their views on the complexity of upgrades and hot and/or interim fixes and on how these can best be factored into normal operations. |
| Will Comply? YES / NO |
| Enhanced Response: |
| 4.3.4 Two-Nines Disaster Recovery (**MHLR#4**) |
| Responders are required at minimum to confirm that they "Will Comply" with the above.<br>Responders may wish to enhance their response by:<br>- addressing particular areas where essential infrastructure data and/or configuration information must be backed-up and copied off-line in order to perform a successful disaster recovery. |
| Will Comply? YES / NO |
| Enhanced Response: |
| 4.3.5 Repeatably Deployable On-Premises Using Full Automation (**MHLR#5)** |
| Responders are required at minimum to confirm that they "Will Comply" with the above.<br>ECMWF believes that responders may wish to enhance their response by:<br>- giving a high-level discussion of how the above requirement might be achieved. They are encouraged to discuss:<br>o all artefacts that must be available,<br>o all repository types needed and tools available,<br>o mechanisms for verifying newly proposed clusters without having to build them at full-scale, and<br>o any DevOps toolchains that would simplify and accelerate such testing. |
| Will Comply? YES / NO |
| Enhanced Response: |

4.3.6 Repeatably Deployable Without Internet Access (**MHLR#6**)

Responders are required at minimum to confirm that they “Will Comply” with the above.
ECMWF believes that responders may also wish to enhance their response by:

- describing how they would organise artefact repositories, version pinning and other associated technical aspects to ensure repeatable off-line building of production OpenStack environments.

Will Comply? YES / NO

Enhanced Response:

4.3.7 Highly Scalable Without Performance Degradation (**MHLR#7)**

Responders are required at minimum to confirm that they “Will Comply” with the above.
Responders may also wish to:

- comment on the achievability of this requirement
- indicate any choke points in such systems that may make reaching such scaling levels either impossible or prohibitively expensive.

Will Comply? YES / NO

Enhanced Response:

4.3.8 Modern Networking Compliant (**MHLR#8)**

Responders are required at minimum to confirm that their offerings “Will Comply” with the above networking arrangements.
Responders may also wish to describe:

- how they would recommend integrating a large production OpenStack environment with the network architecture outlined above.

Will Comply? YES / NO

Enhanced Response:

4.3.9 Differing Hardware Flavours (**MHLR#9**)

Responders are required at minimum to confirm that their offerings “Will Comply” with the above flavouring flexibility requirements.
Responders may also with to:

- describe how they would present such flavour variety to users, and how they would lifecycle manage it with guidelines on when to decommission and/or refurbish older equipment.
- comment on *sweet-spot* server configurations for use in general-purpose cloud environments in terms of CPU (cores and speed), memory (density, quantity and speed), NICs (quantity and speed) and rack-density against the cost.

Will Comply? YES / NO

Enhanced Response:

4.3.10 Differing Storage Flavours (**MHLR#10**)

Responders are required at minimum to confirm that their offerings “Will Comply” with the above storage flexibility requirements.
Responders may also wish to:

- argue the case for alternatives to Ceph
- argue the case for proprietary storage solutions
- give a description of how they would present a range of storage options to users, and how they would encourage least cost usage in a quota-only (no real-time charging) environment.

Will Comply? YES / NO

Enhanced Response:

4.3.11 Software Defined Networking (**MHLR#11**)

**In this case a “Will Comply” response will not be sufficient.**
Responders are asked to propose industry-recognised solutions for achieving the above objectives; multiple cluster solutions may be considered provided they offer sufficient cross-cluster integration of storage, authentication and authorisation. As this is an area where it is possible to spend a great deal of money for diminishing returns, responders are asked to propose a minimum of architectures showing in each case the approximate cost and the associated benefits.

Will Comply? YES / NO

Full Response:

4.3.12 Data Centre Components, Sub-Systems, Practices & Introspection (**MHLR#12**)

Responders are required at minimum to confirm that their offerings "Will Comply" with the above requirements.

Responders may also wish to:

- describe how they would architect and implement a data centre intended to accommodate multiple clouds at the scale of the cloud clusters described in Table 3.
- describe all necessary virtual and physical automation. (Please note that ECMWF does not believe it is yet necessary for it to apply robotic data centre asset management; however, Responders are invited to discuss this point if they wish, with particular reference to when it becomes the appropriate method.)

Will Comply? YES / NO

Enhanced Response:

4.3.13 Server Farms & Soft Clusters (**MHLR#13**)

Responders are asked to describe an architecture and associated software and management procedures to achieve this so-called server-farm and soft-cluster approach. They should take into account the existence of the ECMWF corporate configuration management systems Opsview (CMDB) and NetZoom (DCIM).

Will Comply? YES / NO

Full Response:

4.3.14 Logistics & Equipment Supply Chains (**MHLR#14**)

Responders are required at minimum to confirm that their offerings "Will Comply" with the above requirements.

Responders may also with to:

- propose a supply chain that ensures the availability requirements are met while cost optimising the level of in-rack, on-site and supply-chain spares.

Will Comply? YES / NO

Enhanced Response:

4.3.15 1st, 2nd and 3rd Line Support - Personnel (**MHLR#15**)

Responders are asked to describe a model in which production cloud environments can be architected, implemented and operated using staff, consultants and suppliers and to give numbers and educational/experience levels of said staff, consultants and suppliers.

Will Comply? YES / NO

Full Response:

4.3.16 Regulatory & Compliance (**MHLR#16**)

Responders are asked to give an opinion on the main areas of concern for any such providers of services, including but not limited to:

- firewalling and intrusion detection of
  - the core infrastructure
  - users applications
  - inter-tenancy privacy
- storage and compute segregation of
  - data from differing jurisdictions
  - compute resources within differing jurisdictions
- logging and audit of user activity for compliance with
  - legal requirements
  - insurance requirements
  - post-event investigations

It must be stressed that the above list is not definitive; it is intended to give the flavour of the concerns that ECMWF wants to address. Responders are asked to point out any areas not mentioned and to comment on the duties and responsibilities implied.

Will Comply? YES / NO

Full Response:

4.3.17 Segmentable & Ring-Fenced (**MHLR#17**)

Responders are asked to give their views on the following aspects of the above requirement:

- The technical mechanisms that would allow compute and/or storage resources to be added to a running cloud platform whilst at the same time ring-fencing the added resources to certain account holders and/or projects.
- The technical mechanisms that would allow the Member States and Partners to establish and manage their own range of user identities within a Keystone domain. This would be a delegated authentication mechanism whereby the user first obtains his general access rights to the full environment and then obtains his limited access rights to the ring-fenced resources – all done with one sign-on action.
- The appropriate mechanisms for performing knowledge transfer and validation for the purpose of establishing similar environments on Member State and Partner territories.
- Possible models for costing such arrangements, both in terms of manpower and equipment.

Will Comply? YES / NO

Full Response:

4.3.18 High-Integrity Data Storage (**MHLR#18**)

Responders are required to confirm that their offerings “Will Comply” with the above requirements.

Will Comply? YES / NO

4.3.19 Resilient Operation Under non-OpenStack Failure Conditions (**MHLR#19**)

Responders are required to confirm that their offerings “Will Comply” with the above requirements.

Will Comply? YES / NO

4.3.20 Secure By Default –Secure Audited Operation & Misuse Detection (**MHLR#20**)

Responders are asked to describe their experience of designing and implementing security of cloud components in general and cloud environments in particular.

Full Response:

4.3.21 Monitoring and Alerting (**MHLR#21**)

Responders are asked to give their view of how best to architect monitoring, reporting and alerting environment, which products to choose and how best to integrate them with operations, system support, consultancy support and the established monitoring environment at ECMWF mentioned above (OpsView & Splunk).

Full Response:

5.1.1 OpenStack Deployment using Triple-O / RedHat OpenStack Platform (RHOSP) (**TR#1a**)

**Responders who propose Triple-O or RHOSP** are asked to answer the following questions:

- Are Triple-O and RHOSP of production quality? If so, what level of engineering skill and effort is required to deploy them?
- Are Triple-O and RHOSP (at their core) fundamentally the same product perhaps with differing release cycles?
- What are the essential differences between Triple-O and RHOSP? (consider such issues as coverage of OpenStack components and open source versus closed source)
- What is the support mechanism for Triple-O compared with the support mechanism for RHOSP?
- What is the full lifecycle cost of supporting a Triple-O deployment compared with the full lifecycle cost of supporting a RHOSP deployment keeping in mind the sizes to which ECMWF wishes to scale its cloud computing environments (see Table 3)?
- Do Triple-O and RHOSP comply with the core components highlighted in orange in Figure 1? If not, can the differences be reconciled with additional engineering and at what cost?
- What are your overall impressions of Triple-O and RHOSP?

Answers:

5.1.2 OpenStack Deployment using Kolla-Ansible & Kolla (**TR1#b**)

**Responders who propose Kolla-Ansible & Kolla** are asked to answer the following questions:

- Are Kolla-Ansible & Kolla of production quality? If so, what level of engineering skill and effort is required to deploy them?
- Are Kolla-Ansible & Kolla available as commercially supported products for customer deployments on-premises?
- What community support is available for Kolla-Ansible & Kolla? What level of internal and/or consultancy engineering skill is required to interface between a production environment and the support community?
- What is the full lifecycle cost of supporting a Kolla-Ansible & Kolla deployment keeping in mind the sizes to which ECMWF wishes to scale its cloud computing environments (see Table 3)?
- Do Kolla-Ansible & Kolla comply with the core components highlighted in orange in Figure 1? If not, can the differences be reconciled with additional engineering and at what cost?
- What are your overall impressions of Kolla-Ansible & Kolla?

Answers:

5.1.3 OpenStack Deployment using OpenStack-Ansible (**TR1#c**)

**Responders who wish to propose OpenStack-Ansible** are asked to answer the following questions:

- Is OpenStack-Ansible of production quality? If so, what level of engineering skill and effort is required to deploy it?
- Is OpenStack-Ansible available as a commercially supported product for customer deployments on-premises?
- What community support is available for OpenStack-Ansible? What level of internal and/or consultancy engineering skill is required to interface between a production environment and the support community?
- What is the full lifecycle cost of supporting an OpenStack-Ansible deployment keeping in mind the sizes to which ECMWF wishes to scale its cloud computing environments (see Table 3)?
- Does OpenStack-Ansible comply with the core components highlighted in orange in Figure 1? If not, can the differences be reconciled with additional engineering and at what cost?
- What are your overall impressions of OpenStack-Ansible?

Answers:

5.1.4 Other Strong OpenStack Deployment Candidates (**TR#1d**)

Where a responder believes there is an alternative deployment mechanism not mentioned above that ECMWF should consider in-depth as a possible contender for production use, they are asked to provide answers in the spirit of the questions posed in Sections 5.1.1 through 5.1.3.

Answers:

5.2 Choice of Ceph Distribution & Deployment Method (**TR#2**)

Responders who propose Ceph-Ansible are asked to answer the following questions:

- Is Ceph-Ansible of production quality? If so, what level of engineering skill and effort is required to deploy it?
- Is Ceph-Ansible available as a commercially supported product for customer deployments on-premises?
- What community support is available for Ceph-Ansible? What level of internal and/or consultancy engineering skill is required to interface between a production environment and the support community?
- What is the full lifecycle cost of supporting a Ceph-Ansible deployment keeping in mind the sizes to which ECMWF wishes to scale its cloud computing environments (see Table 3)?
- Does Ceph-Ansible comply with the core components highlighted in orange in Figure 1? If not, can the differences be reconciled with additional engineering and at what cost?
- What are your overall impressions of these Ceph-Ansible?

Responders who wish to propose an alternative deployment mechanism to Ceph-Ansible are asked to describe their proposed mechanism and to answer the previous six questions in the context of their proposed mechanism.

Answers:

5.3 Choice of Kubernetes Distribution & Deployment Method (**TR#3**)

**Responders who wish to propose Rancher** are asked to answer the following questions:

- Is Rancher of production quality? If so, what level of engineering skill and effort is required to deploy it?
- Is Rancher available as a commercially supported product for customer deployments on-premises?
- What level of community support is available for Rancher? What level of internal and/or consultancy engineering skill is required to interface between a production environment and the support community?
- What is the full lifecycle cost of supporting a Rancher deployment keeping in mind the sizes to which ECMWF wishes to scale its cloud computing environments (see Table 3)?
- What are your overall impressions of these Rancher?
- What other deployment technologies in this class would you recommend we consider (for each such technology, please also give answers to the preceding five questions)?

**Responders who wish ro propose Kubespray** are asked to answer the following questions:

- Is Kubespray of production quality? If so, what level of engineering skill and effort is required to deploy it?
- Is Kubespray available as a commercially supported product for customer deployments on-premises?
- What level of community support is available for Kubespray? What level of internal and/or consultancy engineering skill is required to interface between a production environment and the support community?
- What is the full lifecycle cost of supporting a Kubespray deployment keeping in mind the sizes to which ECMWF wishes to scale its cloud computing environments (see Table 3)?
- What are your overall impressions of these Kubespray?
- What other deployment technologies in this class would you recommend we consider (for each such technology, please also give answers to the preceding five questions)?

**Responders who wish to propose Magnum** are asked to answer the following questions:

- Is Magnum of the same production quality as the OpenStack core components? If so, what level of engineering skill and effort is required to deploy and support it?

- What level of community support is available for Magnum? What level of internal and/or consultancy engineering skill is required to interface between a production environment and the support community?
- What are your overall impressions of Magnum?

**Responders who wish to propose an alternative to the above mentioned systems** should provide a brief description of their proposal and should answer questions about their proposed system in the spirit of the above questions.

Answers:

### 5.4 Containerisation using OpenShift / OKD (**TR#4**)

**Optionally,** responders who wish to are invited to give an analysis of the relative merits of using Kubernetes and its associated deployment tools versus using OpenShift / OKD and its associated deployment tools. The analysis should emphasise, but not be limited to:

- the common preferences of developers
- the common preferences of operators
- the richness of available tools and DevOps platforms
- the completeness as a container platform for application development and deployment
- the security of the platform
- the authentication and authorisation mechanisms
- the possibility to deploy OpenShift on OpenStack for small-scale individual containerisation clusters

Optional Response:

### 5.5 Operations & Maintenance (**TR#5**)

Part 1: Responders are asked to comment on the following example areas in respect of processes, procedures and tools:

- First-line:
  - automatic management of user accounts and projects (i.e. add, change, disable, delete)
  - automatic management of quotas (i.e. increase, decrease, diagnose, etc.)
  - automatic management of access to shared (but protected) resources (read, update, create, etc.)
- Second-line:
  - assist with cloud usage questions
  - provide cloud automation and orchestration tool support
- Third-line:
  - provision a cluster from the server farm
  - add or remove a compute server to/from a running production OpenStack cluster
  - assist with requirements capture for new capabilities

Part 2: Responders may, **optionally**, wish to take a much broader view of the area of Operations & Maintenance and provide appropriate discussions of same.

Part 1: Full Response:

Part 2: Optional Response:

### 5.6 Training (**TR#6**)

Responders are asked to propose a combination of classroom, laboratory and self-training materials – together with proficiency verification – to support the production on-premises cloud environment.

Full Response:

### 5.7 Documentation (**TR#7**)

Responders are asked to give their views on how best to maintain an up-to-date documented status of a running production cloud environment like the one planned by ECMWF.

Full Response:

### 5.8 Testing, Failures, Fault Analysis, Fixing & Patching (**TR#8**)

Responders are asked to comment on the challenges presented by the above objectives, the likely timescale for implementation, and the level of external support required in the meantime, and eventually in the steady-state.

Full Response:

5.9 Resource Accounting & Charging (**TR#9**)

Responders are asked to comment on how they would perform resource accounting with particular regard to the following questions:

- What tools would you use?
- How would you aggregate and summarise the resulting data?
- What raw data would you archive for future analysis?
- What impact would such mechanisms have on the overall performance of the cloud?

Full Response:

5.10 Cloud Environment Development & Supporting DevOps Environment (**TR#10**)

**Optionally**, responders are invited to blue-sky think through the general architecture and the high-level components that would need to be in place to achieve such a capability. They are also asked to comment on the financial viability of achieving such a capability.

Optional Response:

5.11 Segmentation and Zoning of Ring-Fenced Resources (**TR#11**)

Responders are asked to advise on appropriate mechanisms for achieving this balance of shared and/or dedicated resources.

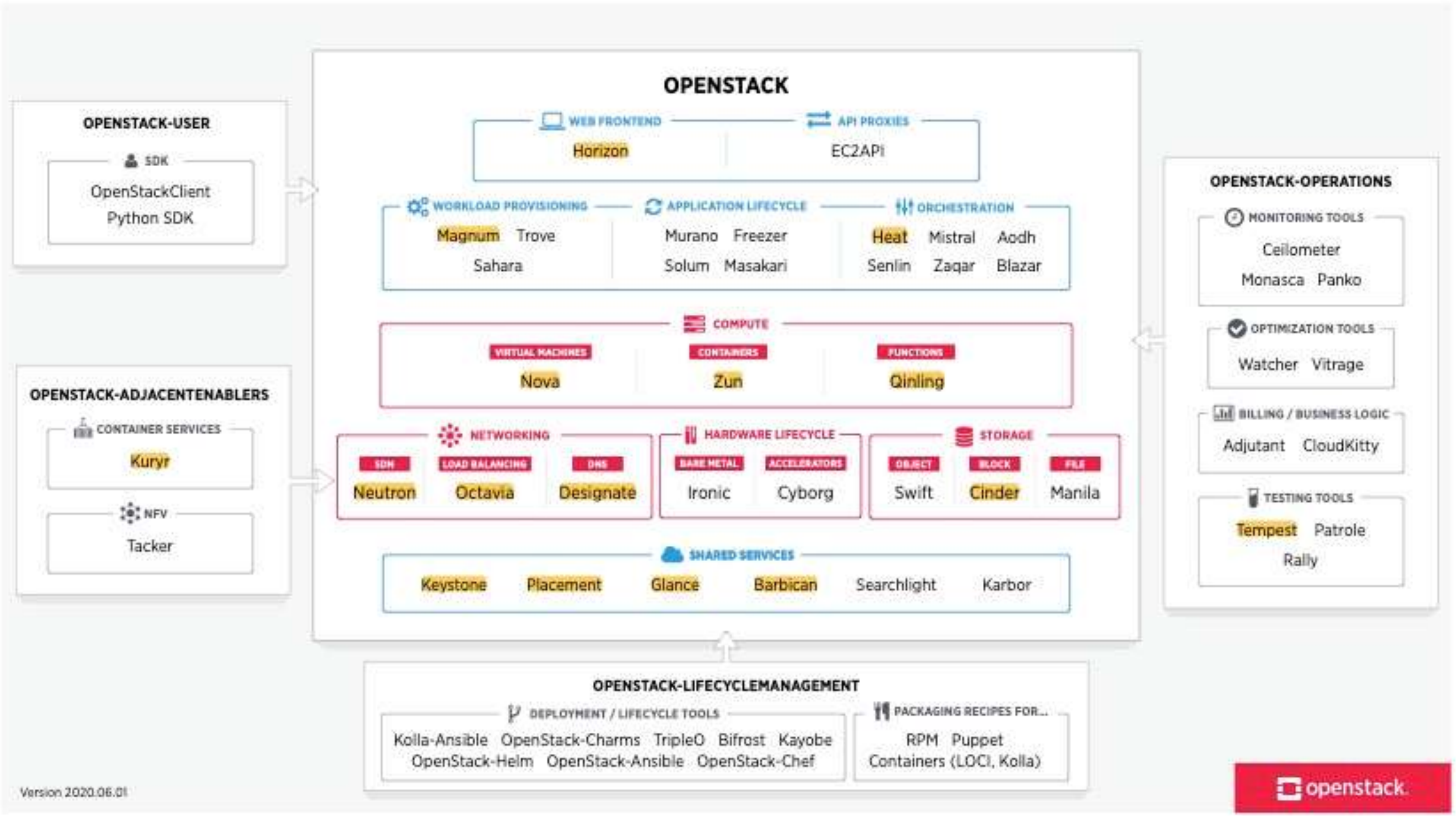
Full Response:

5.12 Architectures Types – Cloud Efficiency (**TR#12**) **(OPTIONAL)**

**Optionally**, responders who wish to respond on this matter, are asked to consider whether there are alternative server/CPU architectures and/or storage architectures that ECMWF should be considered for construction of OpenStack clusters.

Optional Response:

# Appendix B OpenStack Components Diagram



OPENSTACK-USER

SDK

OpenStackClient
Python SDK

OPENSTACK-ADJACENTENABLERS

CONTAINER SERVICES

Kuryr

NFV

Tacker

OPENSTACK

WEB FRONTEND

Horizon

API PROXIES

EC2API

WORKLOAD PROVISIONING

Magnum Trove
Sahara

APPLICATION LIFECYCLE

Murano Freezer
Solum Masakari

ORCHESTRATION

Heat Mistral Aodh
Senlin Zaqar Blazar

COMPUTE

VIRTUAL MACHINES

Nova

CONTAINERS

Zun

FUNCTIONS

Qinling

NETWORKING

SDN

Neutron

LOAD BALANCING

Octavia

DNS

Designate

HARDWARE LIFECYCLE

BARE METAL

Ironic

ACCELERATORS

Cyborg

STORAGE

OBJECT

Swift

BLOCK

Cinder

FILE

Manila

SHARED SERVICES

Keystone Placement Glance Barbican Searchlight Karbor

OPENSTACK-OPERATIONS

MONITORING TOOLS

Ceilometer
Monasca Panko

OPTIMIZATION TOOLS

Watcher Vitrage

BILLING / BUSINESS LOGIC

Adjutant CloudKitty

TESTING TOOLS

Tempest Patrole
Rally

OPENSTACK-LIFECYCLEMANAGEMENT

DEPLOYMENT / LIFECYCLE TOOLS

Kolla-Ansible OpenStack-Charms TripleO Bifrost Kayobe
OpenStack-Helm OpenStack-Ansible OpenStack-Chef

PACKAGING RECIPES FOR...

RPM Puppet
Containers (LOCI, Kolla)

Version 2020.06.01

**---------- END OF INVITATION TO PRE-QUALIFY ----------**