



Asterics

European Data Provider Forum

Some updates from the VAMDC Infrastructure

C.M. Zwölf, N. Moreau, Y.A. Ba
Paris Observatory



Plan of this talk:

“IVOA”-part:

- convergence between XSAMS-VAMDC standard and SLAP-IVOA

“RDA Part”:

- Implementing the RDA recommendations on Data Citation
- RDA & IVOA symbiosis in VAMDC.

“Tools evolution”:

- Cf. Yaye-Awa’s talk.

Updated version of SLAP: Proposed data model

- Roadmap proposed during Trieste Interop (2016)
- Follow up in Shanghai & Santiago
- Last comment gathered in Victoria.

The updated data-model

- Can handle all use cases of previous SSLDM (radiative transitions)
- Can be extended (ionization)
- Closer to VAMDC Data Model, easy to go from one to the other
- But
 - Species definition is still lacking
 - Quantum number descriptions is not compatible
 - VAMDC scientific board is working on methods for serializing into compact unique string atomic states, with related QN.

Metadata Specification

- Adapted to VOSI specifications
- FORMAT=METADATA removed
- Addition of endpoints :
 - /capabilities
 - /availabilities

Service interfaces

- Two query modes are available :
 - Lines (list of lines according to given parameters)
 - Species (list of species queryable in the service)

Service interfaces

- Two query modes are available :
 - Lines (list of lines according to given parameters)
 - Species (list of species queryable in the service)
- Previously proposed version used doQuery parameter :

doQuery=[lines | request]

- Now replaced by two sync endpoints :
 - /lines
 - /species

Service interfaces

Resource type	Resource name	Required
DALI-sync	{lines}	yes
DALI-sync	{species}	no
DALI-examples	/examples	no
VOSI-availability	/availability	yes
VOSI-capabilities	/capabilities	yes

- Implementation of SLAP 2.0 in VAMDC node software
- Extension of the middleware used to deploy a VAMDC-TAP service on a database

Examples

All species in a service :

<http://serviceprovider.org/slap?request=querySpecies>

Replaced by :

<http://serviceprovider.org/slap/species>



```
<?xml version="1.0" encoding="ISO-8859-1"?>
<VOTABLE version="1.3" xmlns="http://www.ivoa.net/xml/VOTable/v1.3" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  - <RESOURCE type="results">
    <INFO value="OK" name="QUERY_STATUS"/>
    <INFO value="1530043425" name="REQUEST_COMPLETED_TIMESTAMP"/>
    <INFO value="2017-01-30" name="SERVICE_VERSION"/>
    <INFO value="sesam" name="SERVICE_NAME"/>
    - <TABLE>
      <FIELD name="SPECIES_NAME" utype="ssldm:Species.name" ucd="phys.atmol.element" arraysize="*" datatype="char"/>
      <FIELD name="SPECIES_TYPE" utype="ssldm:Species.type" arraysize="*" datatype="char"/>
      <FIELD name="INCHIKEY" utype="ssldm:Species.inChiKey" datatype="int"/>
      <FIELD name="INCHI" utype="ssldm:Species.inChi" datatype="int"/>
      - <DATA>
        - <TABLEDATA>
          - <TR>
            <TD>Molecular deuterium </TD>
            <TD>molecule</TD>
            <TD>UFHFLCQGNINYRNP-VVKOMZTBSA-N</TD>
            <TD>InChI=1S/H2/h1H/i1+1D</TD>
          </TR>
          - <TR>
            <TD>Hydrogen deuteride </TD>
            <TD>molecule</TD>
            <TD>UFHFLCQGNINYRNP-OUBTZVSYSA-N</TD>
            <TD>InChI=1S/H2/h1H/i1+1</TD>
          </TR>
          - <TR>
            <TD>Dihydrogen </TD>
            <TD>molecule</TD>
            <TD>UFHFLCQGNINYRNP-UHFFFAOYSA-N</TD>
            <TD>InChI=1S/H2/h1H</TD>
          </TR>
          - <TR>
            <TD>Carbon monoxyde </TD>
            <TD>molecule</TD>
            <TD>UGFAIRIUMAVXCW-UHFFFAOYSA-N</TD>
            <TD>InChI=1S/CO/c1-2</TD>
          </TR>
        </TABLEDATA>
      </DATA>
    </TABLE>
  </RESOURCE>
</VOTABLE>
```


Examples

All lines between 900 and 901 Angstroms in a service :

<http://serviceprovider.org/slap?request=queryLines&WAVELENGTH=9e-8/9.05e-8>

Replaced by :

<http://serviceprovider.org/slap/lines?WAVELENGTH=9e-8+9.01e-8>



```
<?xml version="1.0" encoding="ISO-8859-1"?>
<VOTABLE version="1.3" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:ssldm="http://www.ivoa.net/xml/SimpleSpectrumLineDM/SimpleSpectrumLineDM-v1.1.xsd"
xsi:noNamespaceSchemaLocation="xmlns:http://www.ivoa.net/xml/VOTable/VOTable-1.3.xsd">
- <RESOURCE type="results">
  <INFO value="OK" name="QUERY_STATUS"/>
  <INFO value="2.0" name="SERVICE_PROTOCOL">SLAP</INFO>
  <INFO value="queryData" name="REQUEST"/>
  <INFO value="1530045741" name="FILE_TIMESTAMP"/>
  <INFO value="2017-01-30" name="DATABASE_VERSION"/>
- <TABLE>
  <FIELD name="WAVELENGTH" utype="ssldm:Line.wavelength.value" ucd="em.wl" datatype="double" unit="m"/>
  <FIELD name="IDENTIFICATION" utype="ssldm:Line.title" ucd="em.line" arraysize="*" datatype="char"/>
  <FIELD name="LOWER_LEVEL_ENERGY" utype="ssldm:Line.lowerLevel.energy.value" ucd="phys.energy;phys.atmol.level" datatype="double" unit="J"/>
  <FIELD name="UPPER_LEVEL_ENERGY" utype="ssldm:Line.upperLevel.energy.value" ucd="phys.energy;phys.atmol.level" datatype="double" unit="J"/>
  <FIELD name="LOWER_LEVEL_ELEMENT" utype="ssldm:Line.lowerLevel.element.name" ucd="phys.atmol.element" arraysize="*" datatype="char"/>
  <FIELD name="LOWER_LEVEL_ELEMENT_INCHIKEY" utype="ssldm:Line.lowerLevel.element.inChiKey" ucd="phys.atmol.element" arraysize="*" datatype="char"/>
  <FIELD ucd="meta.bib" arraysize="*" datatype="char"/>
- <DATA>
  - <TABLEDATA>
    - <TR>
      <TD>901.0</TD>
      <TD>Hydrogen deuteride Upper energy : 128732.83 , Lower energy : 17745.45</TD>
      <TD>17745.45</TD>
      <TD>128732.83</TD>
      <TD>Hydrogen deuteride </TD>
      <TD>UFHFLCQGNINYRNP-OUBTZVSYSAN</TD>
    </TR>
  - <TR>
      <TD>901.01</TD>
      <TD>Hydrogen deuteride Upper energy : 112818.62 , Lower energy : 1831.96</TD>
      <TD>1831.96</TD>
      <TD>112818.62</TD>
      <TD>Hydrogen deuteride </TD>
      <TD>UFHFLCQGNINYRNP-OUBTZVSYSAN</TD>
    </TR>
  - <TR>
      <TD>901.01</TD>
      <TD>Hydrogen deuteride Upper energy : 114703.9 , Lower energy : 3717.53</TD>
      <TD>3717.53</TD>
      <TD>114703.9</TD>
      <TD>Hydrogen deuteride </TD>
      <TD>UFHFLCQGNINYRNP-OUBTZVSYSAN</TD>
    </TR>
  </TABLEDATA>
  </DATA>
</TABLE>
</RESOURCE>
</VOTABLE>
```

The Research Data & Data Citation new paradigms

About



Research Data Alliance

The Research Data Alliance (RDA) builds the social and technical bridges that enable open sharing of data.

The RDA vision is researchers and innovators openly sharing data across technologies, disciplines, and countries to address the grand challenges of society.

Data Citation WG



Group details

Status: Recognised & Endorsed

Chair(s): Andreas Rauber, Ari Asmi, Dieter van Uytvanck

Case Statement: [Download](#)

The RDA Working Group on Data Citation (WG-DC) aims to bring together a group of experts to discuss the issues, requirements, advantages and shortcomings of existing approaches for efficiently citing subsets of data. The WG-DC focuses on a narrow field where we can contribute significantly and provide prototypes and reference implementations.

Recommendation is to store all the queries (with their metadata) into a Query Store (QS).

- The difficulty we have to cope with
- Handle a QS in the VAMDC distributed environment (VAMDC is a set of independent TAP services)
 - Integrate the QS with the existing VAMDC components

The Research Data & Data Citation new paradigms



[ABOUT RDA](#) [GET INVOLVED](#) [GROUPS](#) [RECOMMENDATIONS & OUTPUTS](#) [RDA FOR D](#)

RDA/WDS Scholarly Link Exchange (Scholix) WG

[Home](#) » [Working And Interest Groups](#) »

[Capture rectangulaire](#)

WG

i Group details

Status: [Recognised & Endorsed](#)

Chair (s): [Adrian Burton](#) , [Martin Fenner](#) , [Paolo Manghi](#) , [Wouter Haak](#)

Secretariat Liaison: [Lynn Yarmey](#)

TAB Liaison: [Larry Lannom](#)

Case Statement: [Download](#)



WGs Wrapping up (from ~12 months after RDA endorsement)

[History](#)

RDA WDS Scholarly Link Exchange (Scholix) WG
Co-chaired by: [Adrian Burton](#), [Martin Fenner](#), [Paolo Manghi](#), [Wouter Haak](#)

This WG is the follow up from the: [RDA/WDS Publishing Data Services WG](#)

The Scholarly Link Exchange Working group aims to enable a comprehensive global view of the links between scholarly literature and data. The working group will leverage existing work and international initiatives to work towards a global information commons by establishing:

Data model & protocols for interoperability of bibliographic and bibliometric data

- Article citing data
- Data citing article
- Data citing data
- Article citing article

QS in a nutshell:

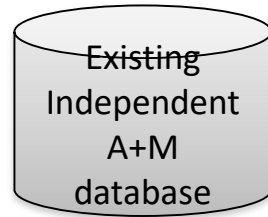
Implementation of the RDA recommendation for the VAMDC infrastructure was done as a part of a sub-contract funded by RDA-EU3 project.

- Made during the year 2017.
- The full source code + documentation is published at <https://github.com/VAMDC/QueryStore> (CC4ByNDNC license)
- Deployment in production for all the nodes is in progress.
- Technical presentation of the architecture and functioning
 - <https://youtu.be/OLe-qcqCcCw>
- The Query Store is a central real-time repository for all the queries served by any VAMDC TAP service
 - Independently by the used client for querying the infrastructure
 - Queries are stored together with their metadata:
 - Production/extraction context and timestamp + references + link to the generated XSAMS file + **Unique Identifier & DOI**
 - Provides live monitoring of the entire infrastructure
 - Data providers may measure their impact and have detailed statistics of usage
 - Increases the quality of the data
 - All the produced XSAMS are parsed for extracting references & validated

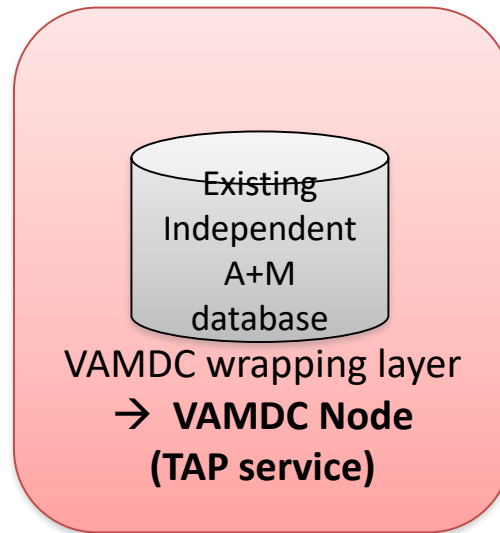
Advantage of getting a DOI:

- XSAMS files are long-lasting and safely stored on the Zenodo Repository (process triggered automatically on user demand).
 - Zenodo provides the uploaded query (data+metadata) with a DOI.
 - Easy to cite a dataset from its DOI (as currently done for papers).
 - Reverse link from Zenodo pointing to VAMDC Query-Store for enhancing trust & provenance in data stored into Zenodo.
 - Zenodo is indexed in OpenAire and linked with Scholix (tools used by publishers for computing bibliometrics and impact factors).
 - An author/paper cited by the data-set will get credits automatically when the dataset is cited (using the DOI) into a paper.
 - **Strong marketing argument: Put your data in VAMDC. You will get automatically credits each time your data is cited!**
- Demo of new DOI feature: https://youtu.be/CgC-KxOM_8g
- If it worked for VAMDC, this could work in the same way for any TAP service.

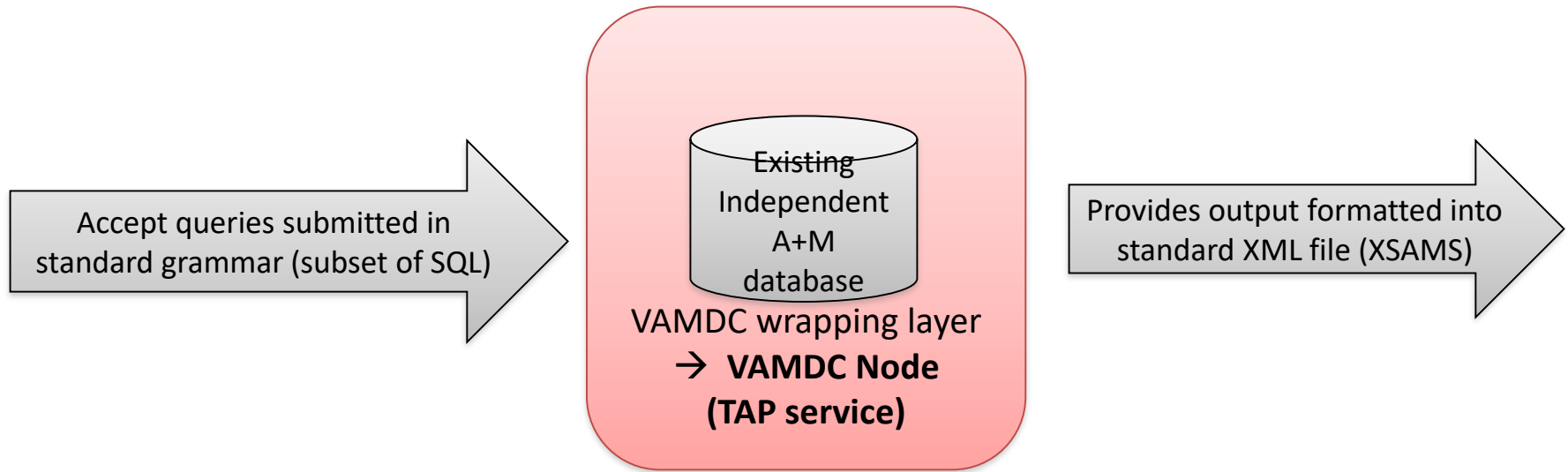
The VAMDC infrastructure technical architecture



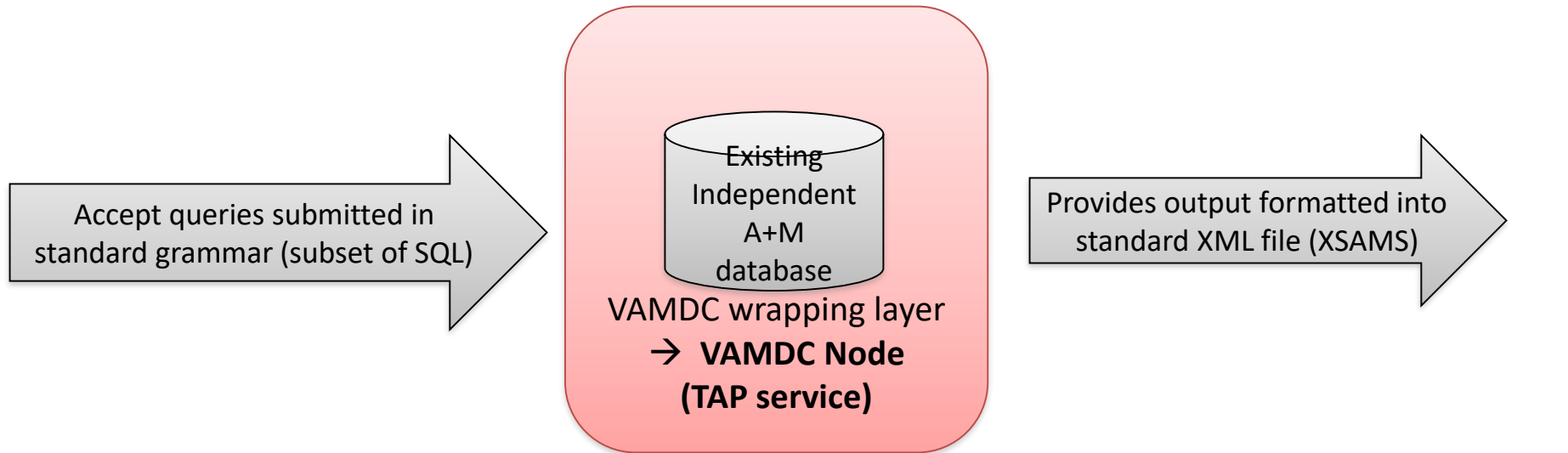
The VAMDC infrastructure technical architecture



The VAMDC infrastructure technical architecture

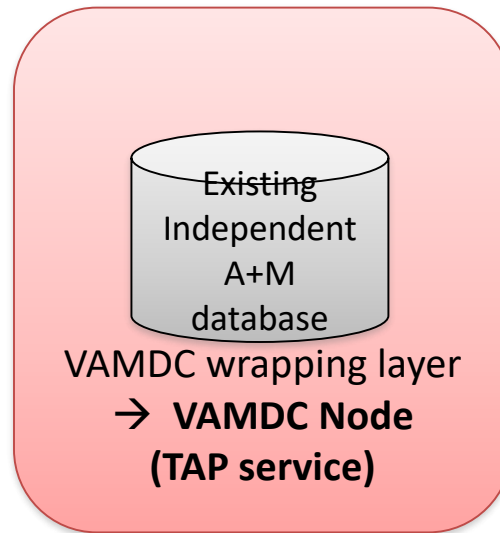


The VAMDC infrastructure technical architecture

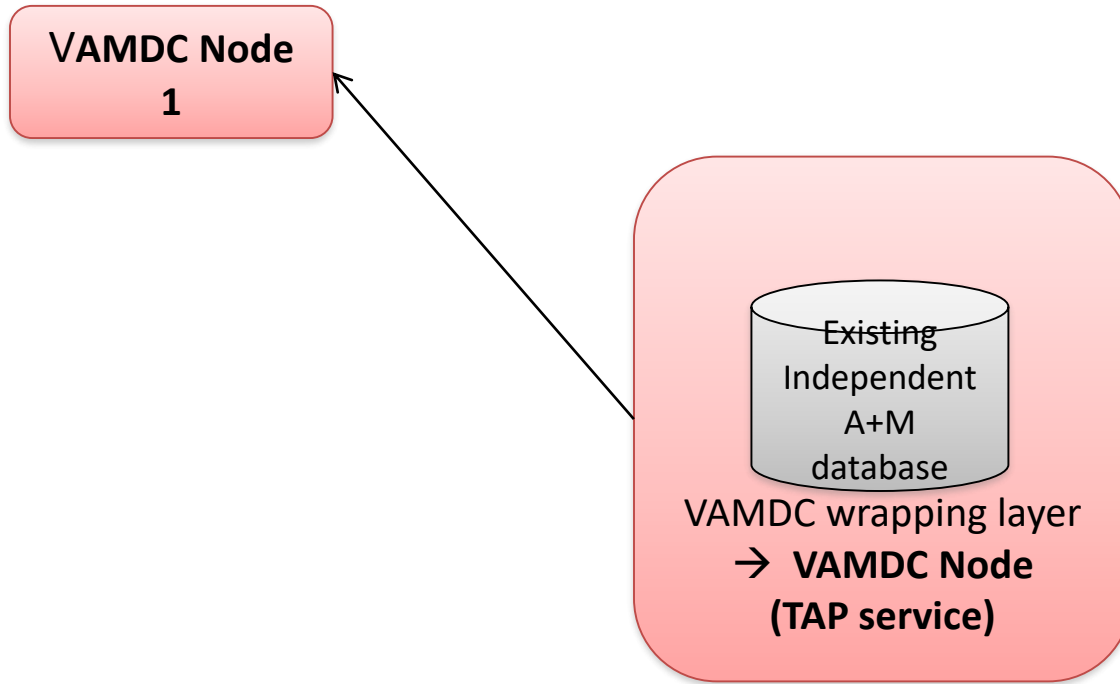


For further details, cf.
<http://standards.vamdc.eu>

The VAMDC infrastructure technical architecture



The VAMDC infrastructure technical architecture

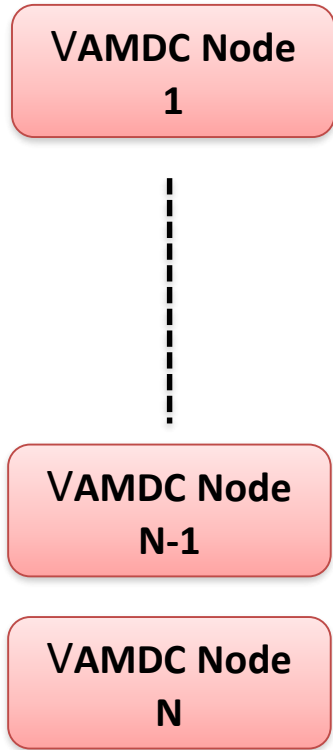


The VAMDC infrastructure technical architecture

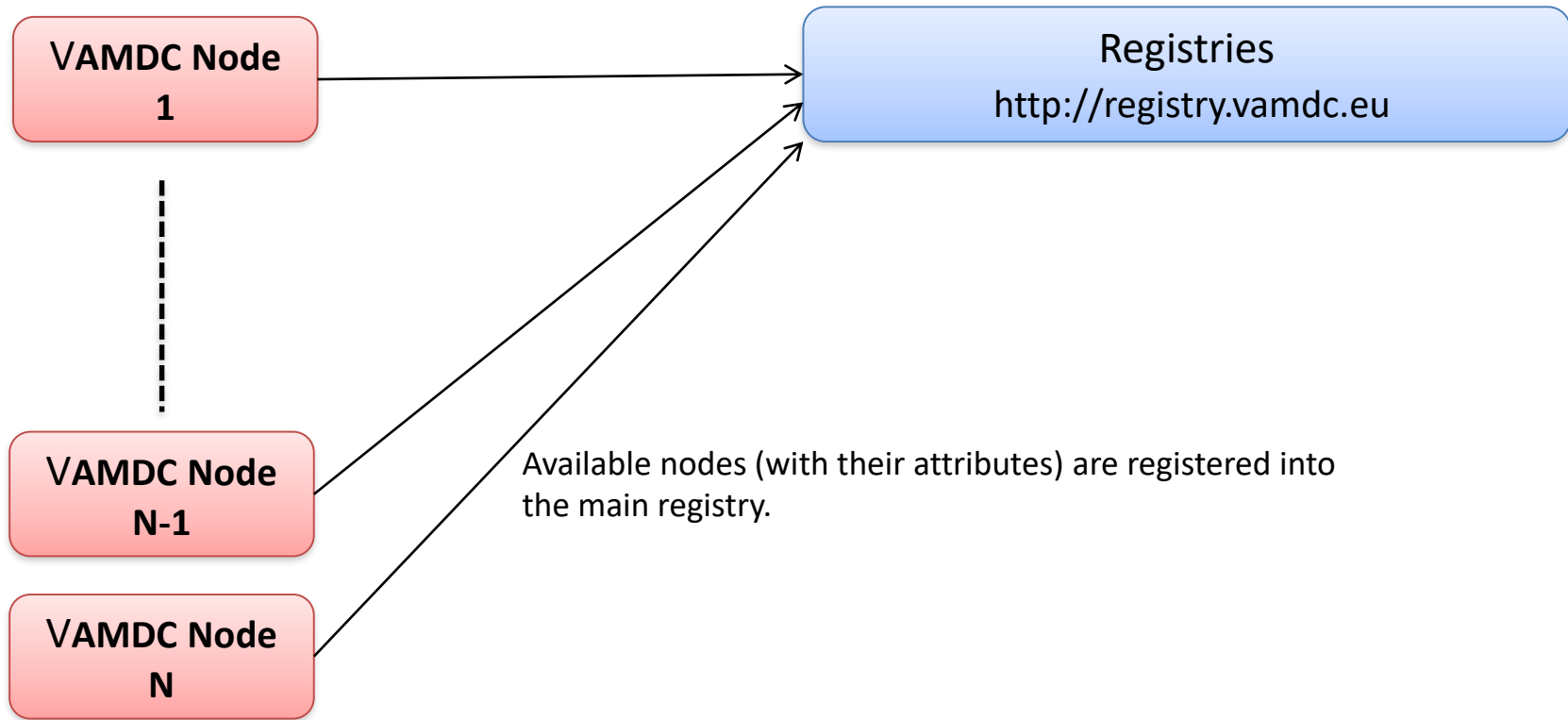
VAMDC Node

1

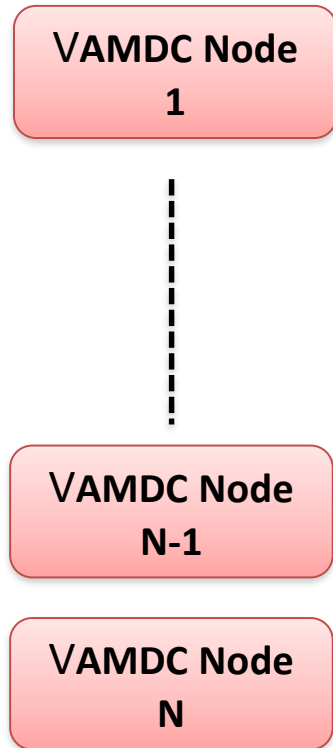
The VAMDC infrastructure technical architecture



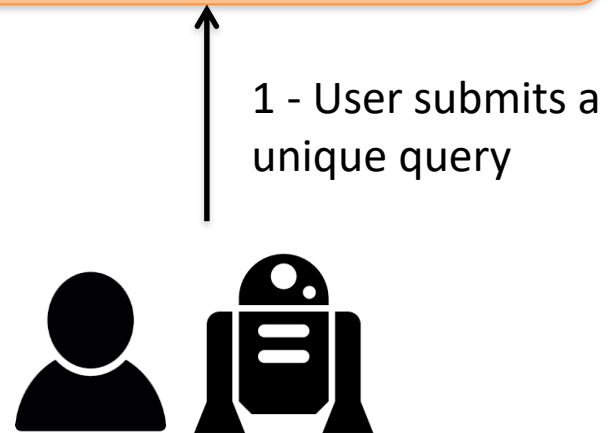
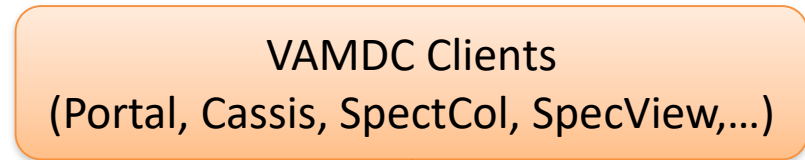
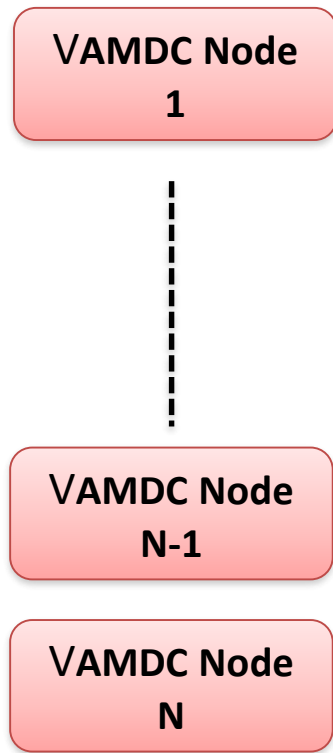
The VAMDC infrastructure technical architecture



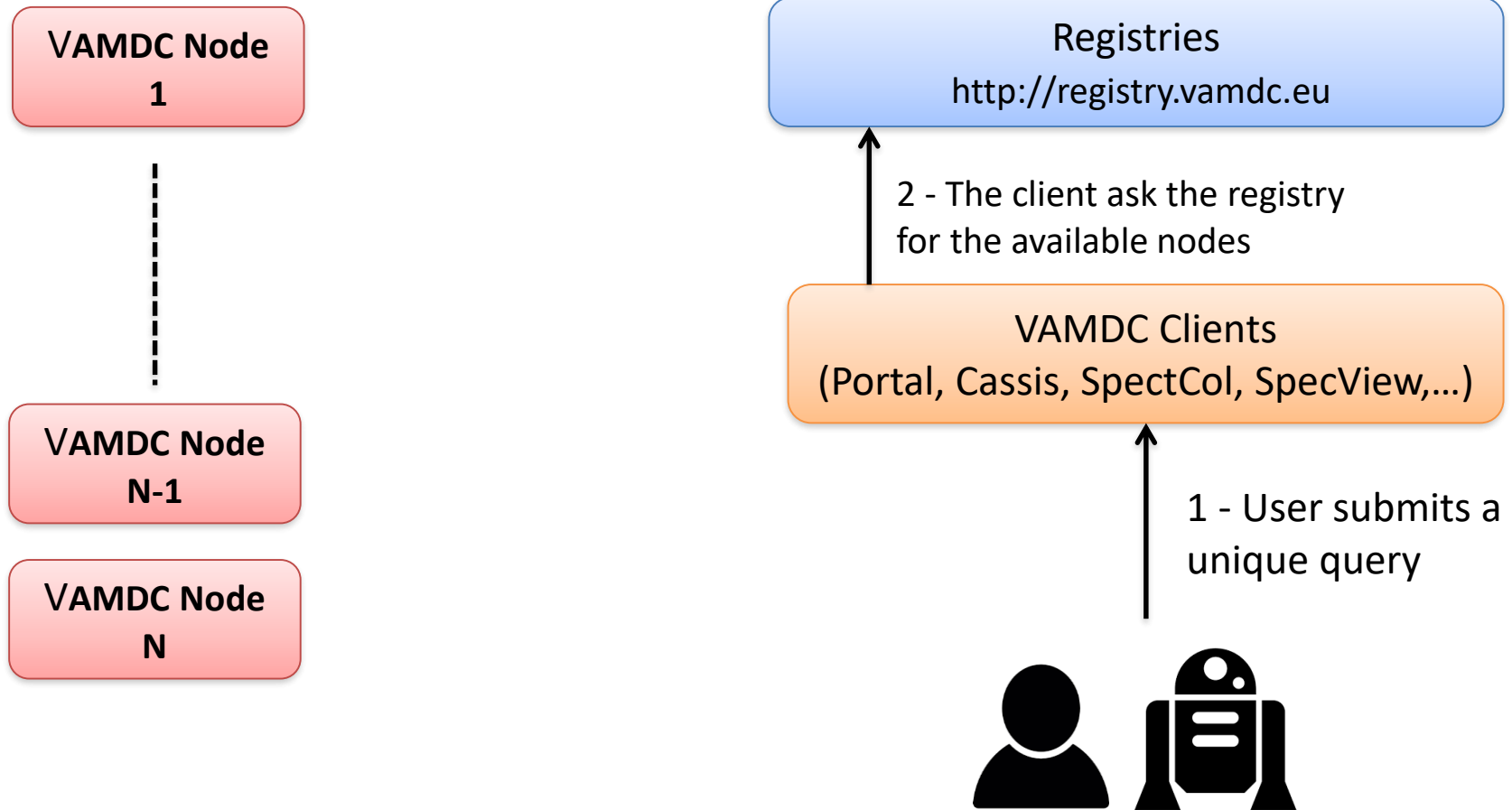
The VAMDC infrastructure technical architecture



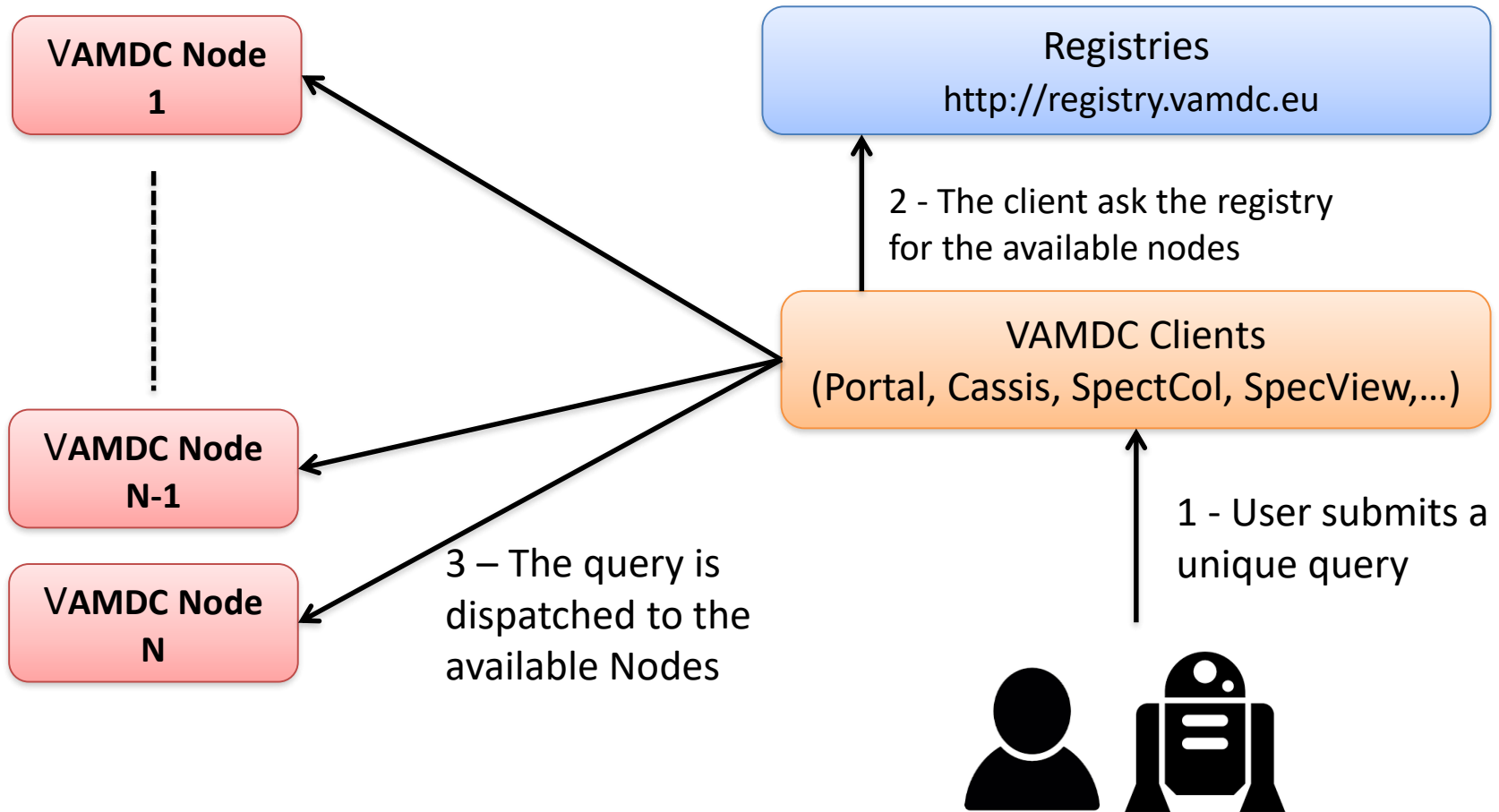
The VAMDC infrastructure technical architecture



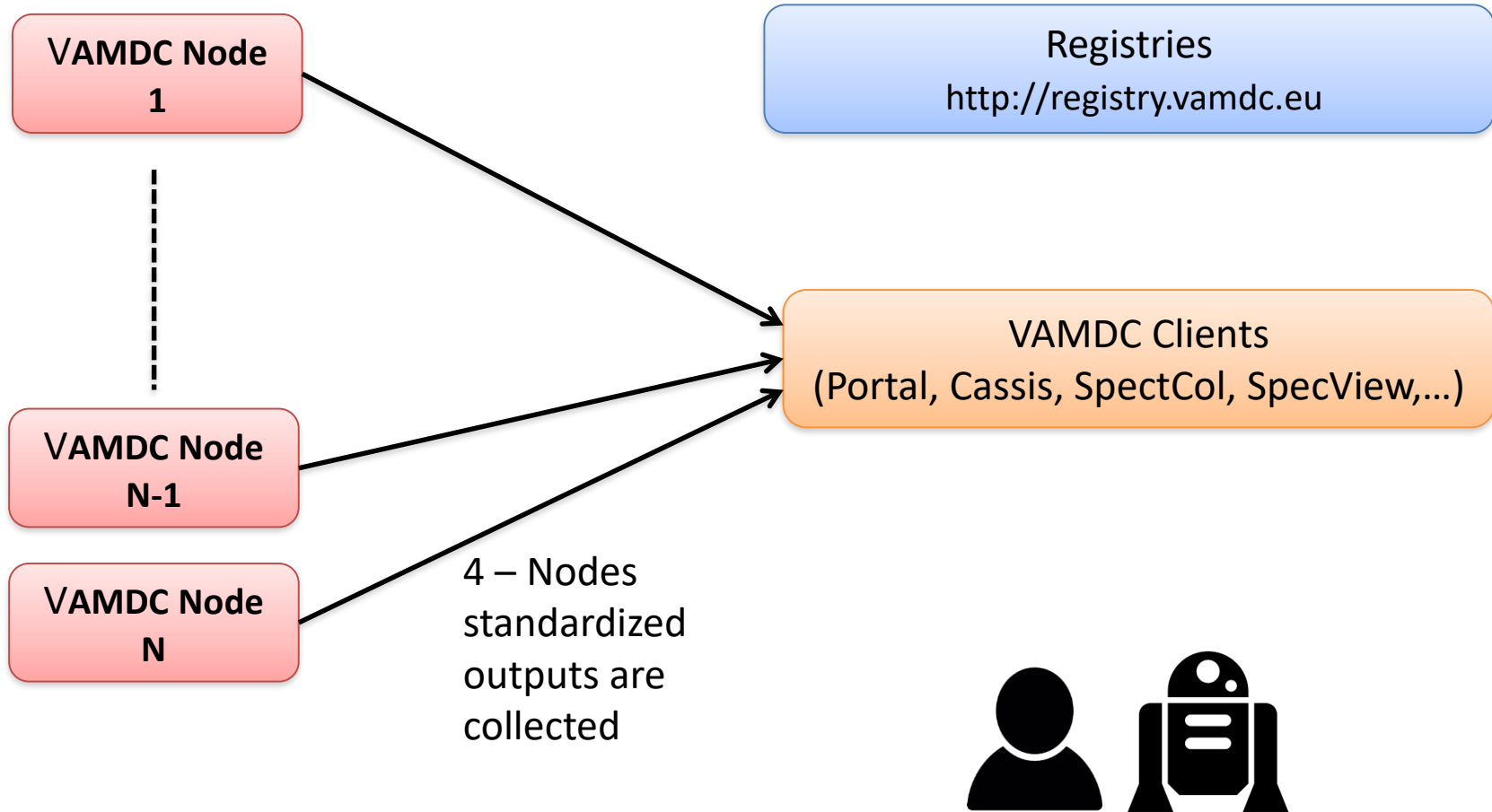
The VAMDC infrastructure technical architecture



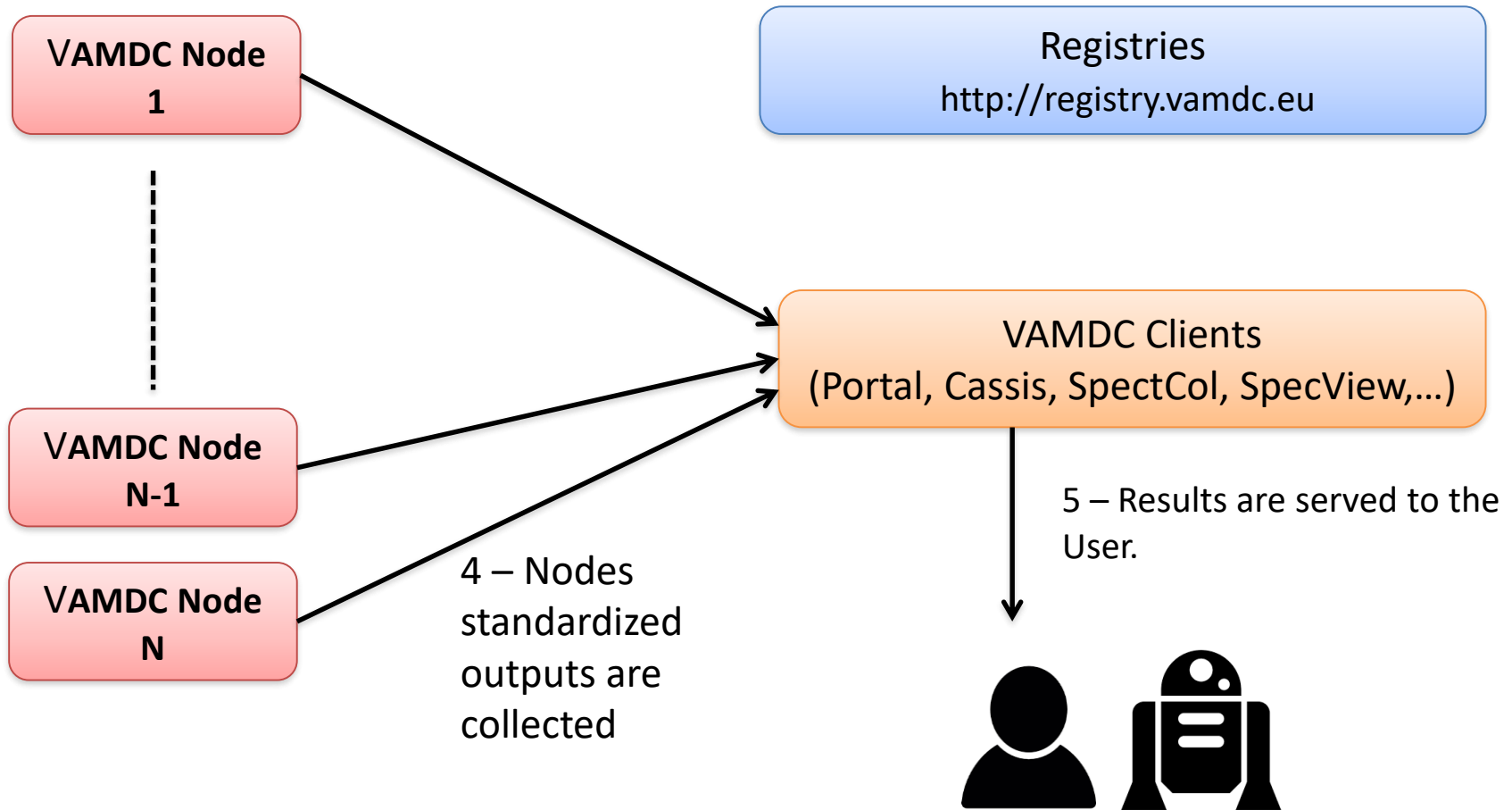
The VAMDC infrastructure technical architecture



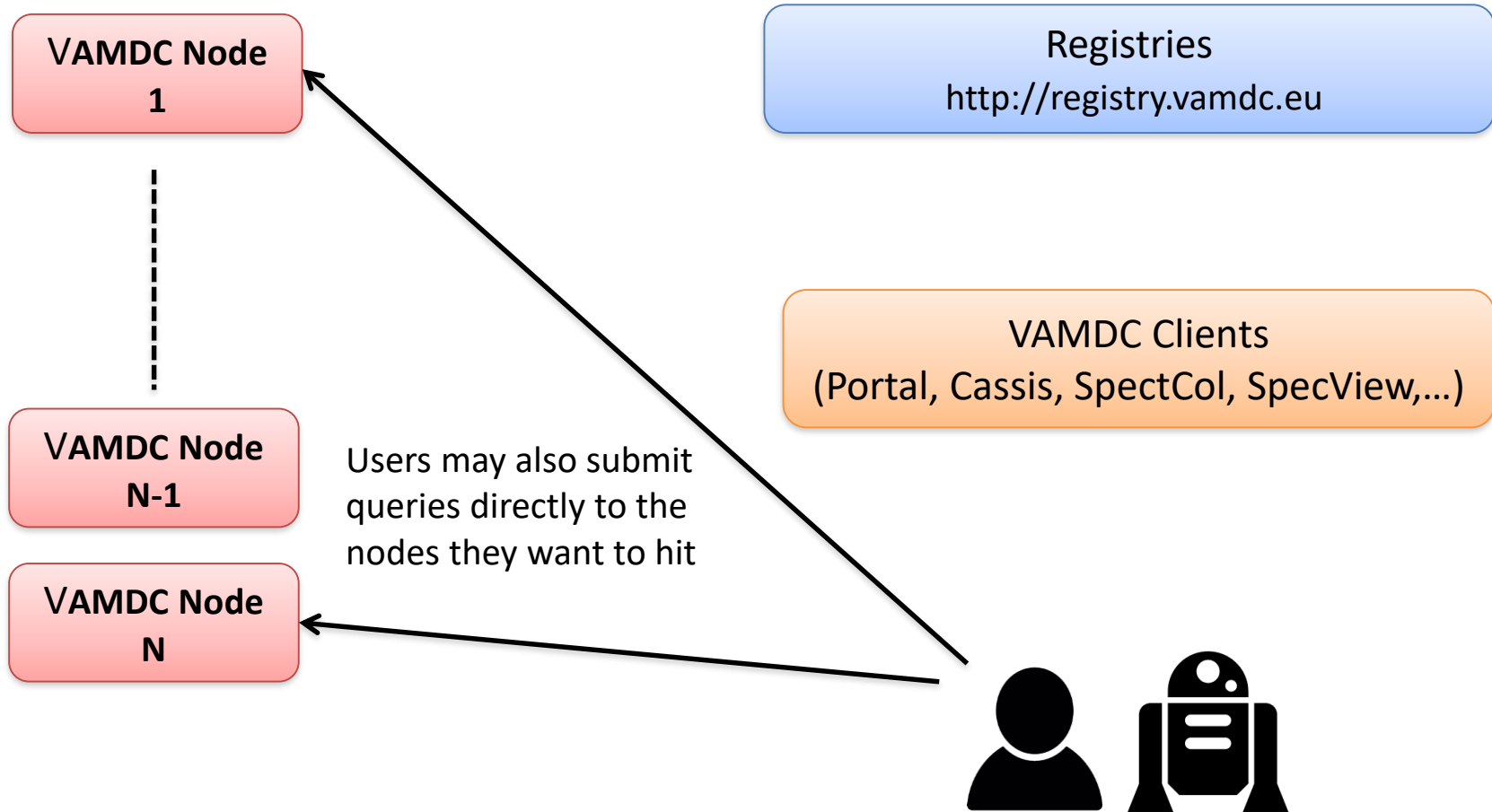
The VAMDC infrastructure technical architecture



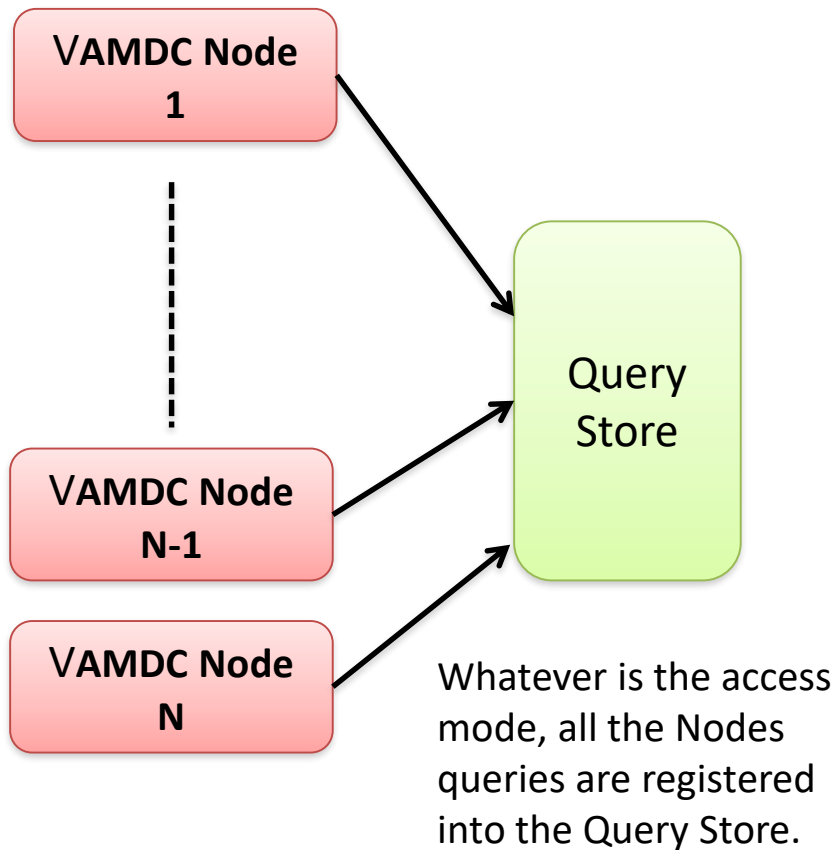
The VAMDC infrastructure technical architecture



The VAMDC infrastructure technical architecture



The VAMDC infrastructure technical architecture

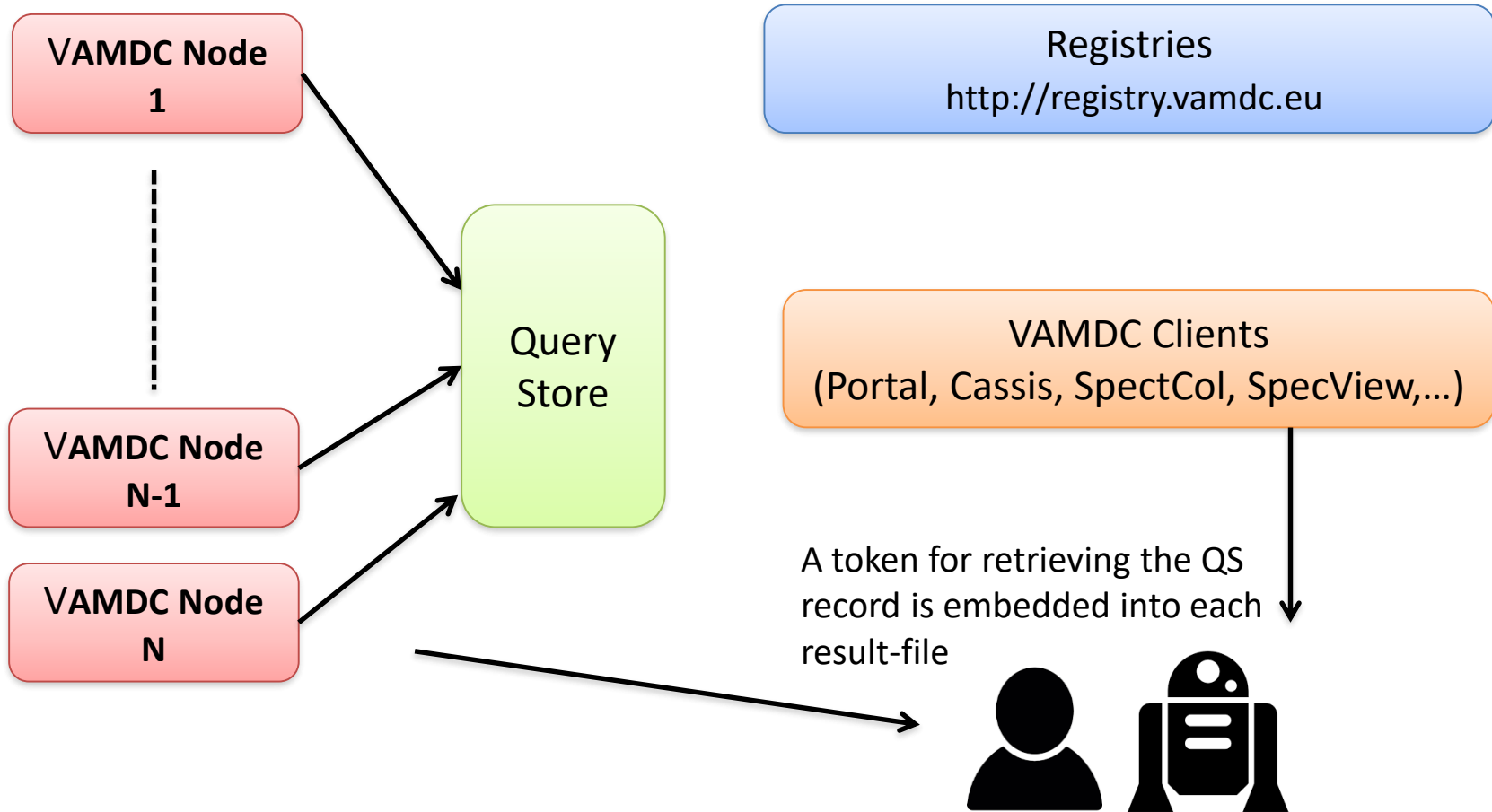


Registries
<http://registry.vamdc.eu>

VAMDC Clients
(Portal, Cassis, SpectCol, SpecView,...)

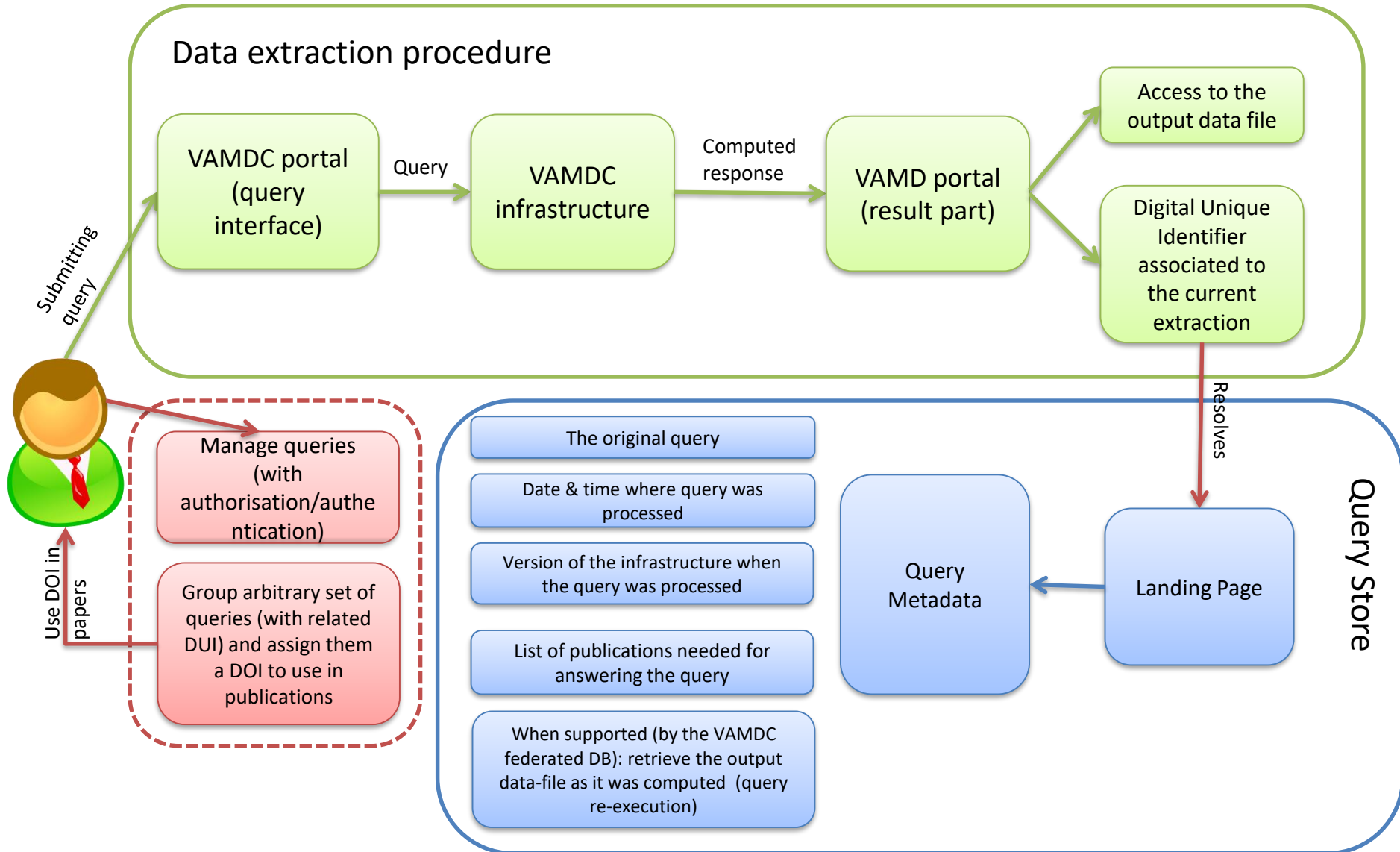


The VAMDC infrastructure technical architecture



The same picture shown two years ago here in Heidelberg

Where are we today ?



Concluding remarks:

- The current workflow shown how to remove all the technical barriers linked with the automatic data-citation & delegation of credits in case of TAP services.
 - From our initial plans, only query clustering is missing (no particular technical issues linked with this part).
 - The RDA recommendations on data citation are implemented as an overlay of IVOA-based services
 - Scholix recommendation is not implemented directly, but obtained through the linking with Zenodo:
 - Example of interoperability virtuous circle: One could say that the interoperability-capabilities propagation speed is greater than the interoperability-protocols adoption speed.