Destination Earth

Destination Earth: Challenges in Federation of Compute and Data Resources

Utz-Uwe Haus, Craig Prunty, Hans-Christian Hoppe

2024-02-13

Funded by the European Union Destination Earth implemented by ECMWF esa EUMETSAT

ETP4HPC

Context/Disclaimer

- This presentation is based on work identifying challenges with respect to the federation of EuroHPC infrastructures provided for running DestinE twins.
- This is not a summary of a report commissioned by ECMWF in the DE_380 contract.
- We believe the points discussed here are more widely applicable than DestinE, but DestinE is the first project highlighting them.

- *This document has been produced in the context of the Destination Earth Initiative and relates to tasks entrusted by the European Union to the European Centre of Medium-Range Weather Forecasts implementing part of this initiative.*
- *This document is funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them. The European Centre of Medium-Range Weather Forecasts is not liable in respect of this document and gives no warranty for the information needed.*

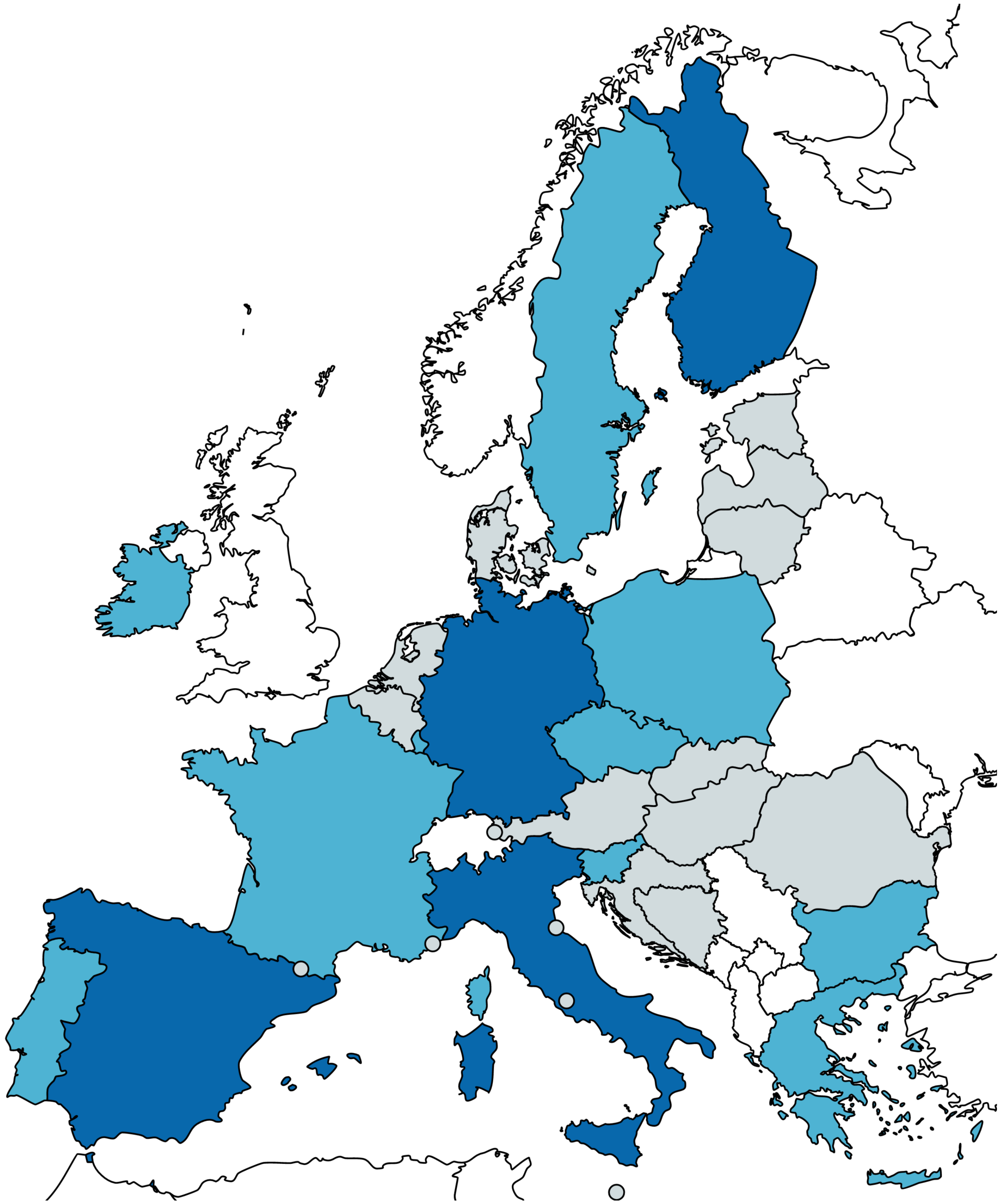
Funded by
the European Union

🟦 **DESTINE Target Federation Locations**:

- CSC, FI (Lumi)
- Cineca, IT (Leonardo)
- BSC, ES (Marenostrum5)
- JSC, DE (Jupiter)

🟦 **Potential Future Federation Locations**

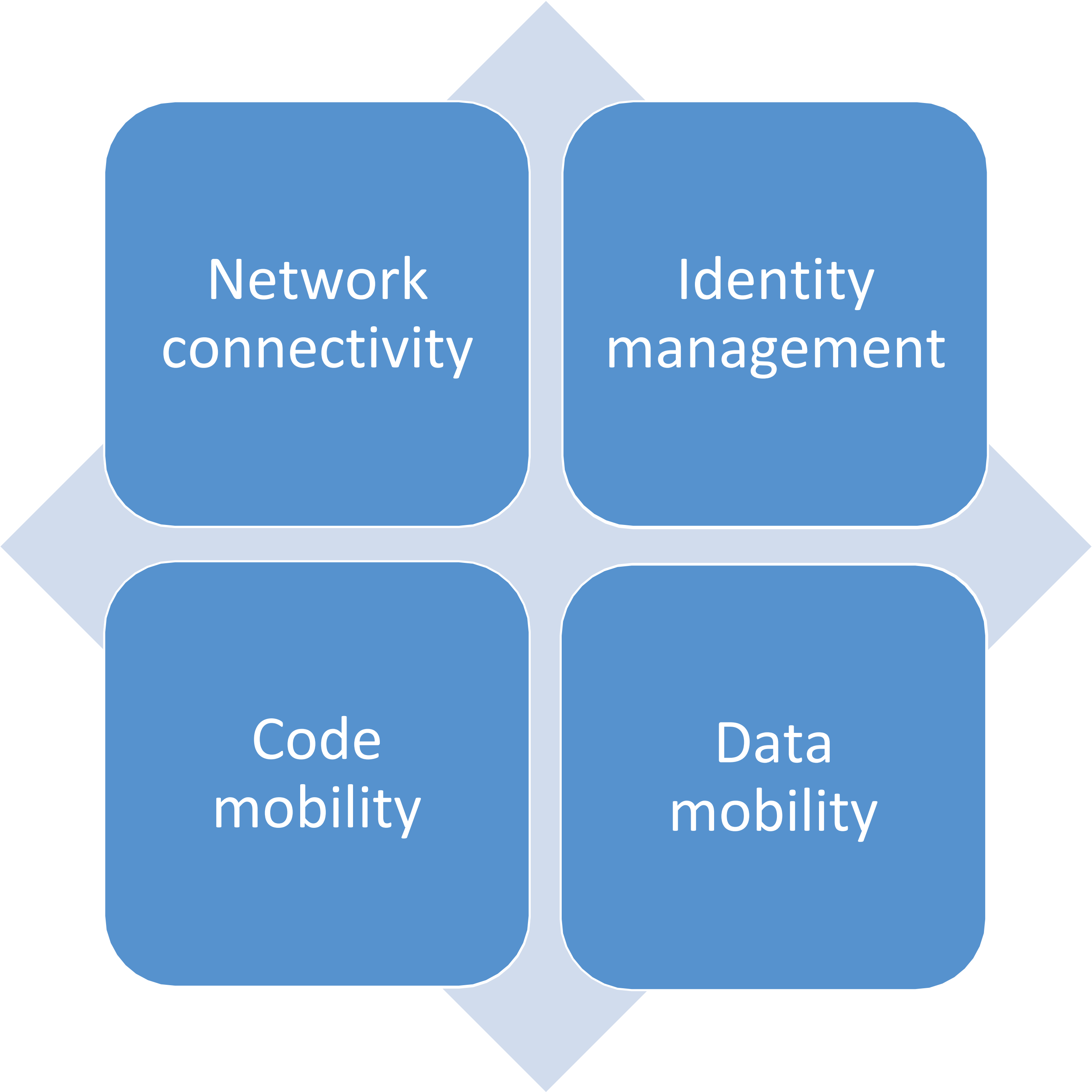
- Genci, FR (Jules Verne consortium)
- LuxConnect, LU (Meluxina)
- IT4Innovations, CZ (Karolina)
- MACC, PT (Deucalion)
- IZUM, SI (Vega)
- Sofia Tech Park, BG (Discoverer)
- GRNet, GR (Daedalus)
- Cyfronet, PL (EHPCPL)
- ICHEC, IE (CASPIr)
- Linköping University, SE (Arrhenius)



Funded by
the European Union

Fedederation

“The act of creating or becoming a union of organizations” (Merriam-Webster)



Network connectivity

Identity management

Code mobility

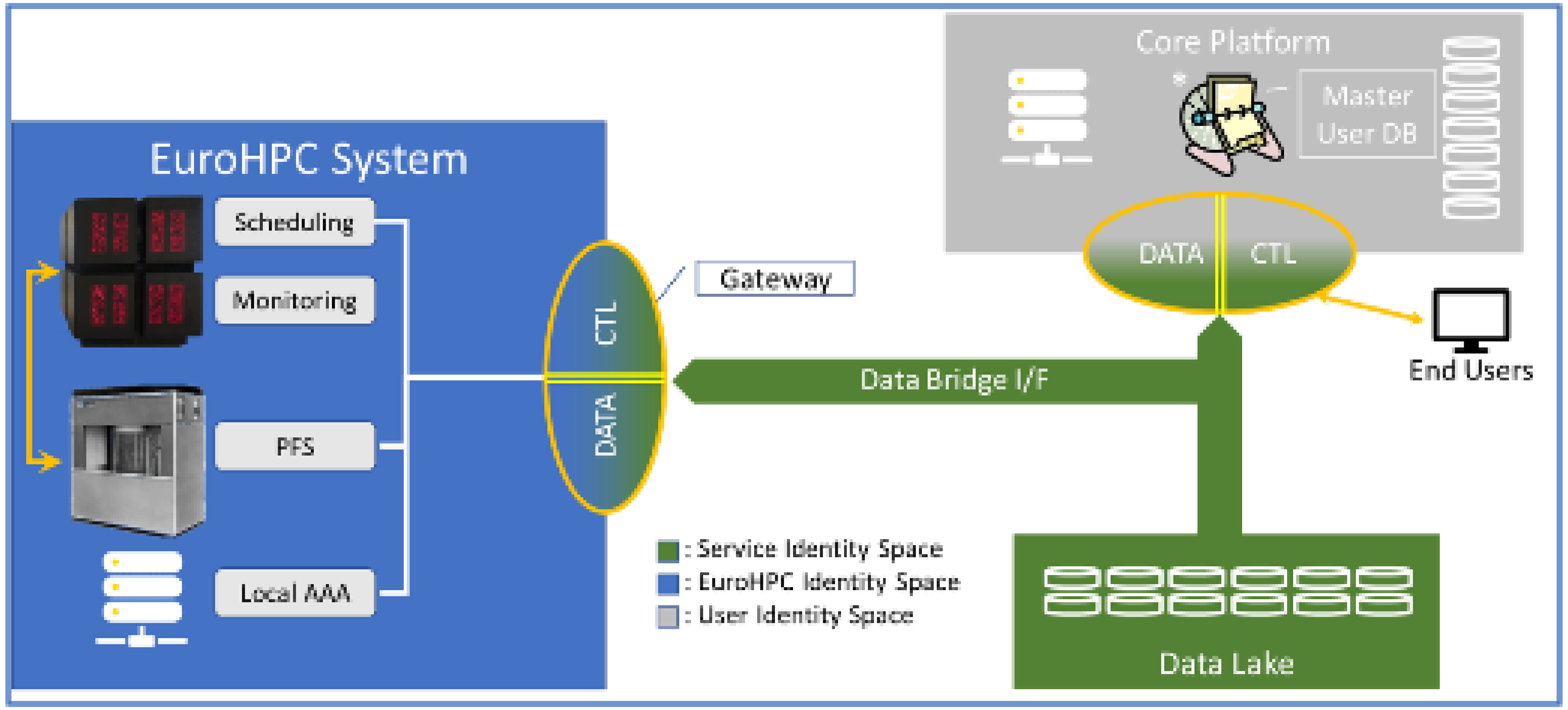
Data mobility

Funded by
the European Union

Conceptual federation architecture

Assumptions

- Users access via the core platform
- Invocation of DT services is automatic, using service identities
- HPC platforms = EuroHPC sites
- Data lake with common metadata schema



Funded by the European Union

Federation Challenges

User Management

User Policy

Authentication, Authorization & Accounting

Data Governance

Operational Handling of User Management System

Service Deployment / Operation, Orchestration, Scheduling

Service interfaces and Instantiation

Local Service Scheduling and Orchestration

Distributed Service Scheduling

Service Deployment and Updates

Quality of Service

System Integration

Independent systems, partially isolated,

Different architectures and programming I/F

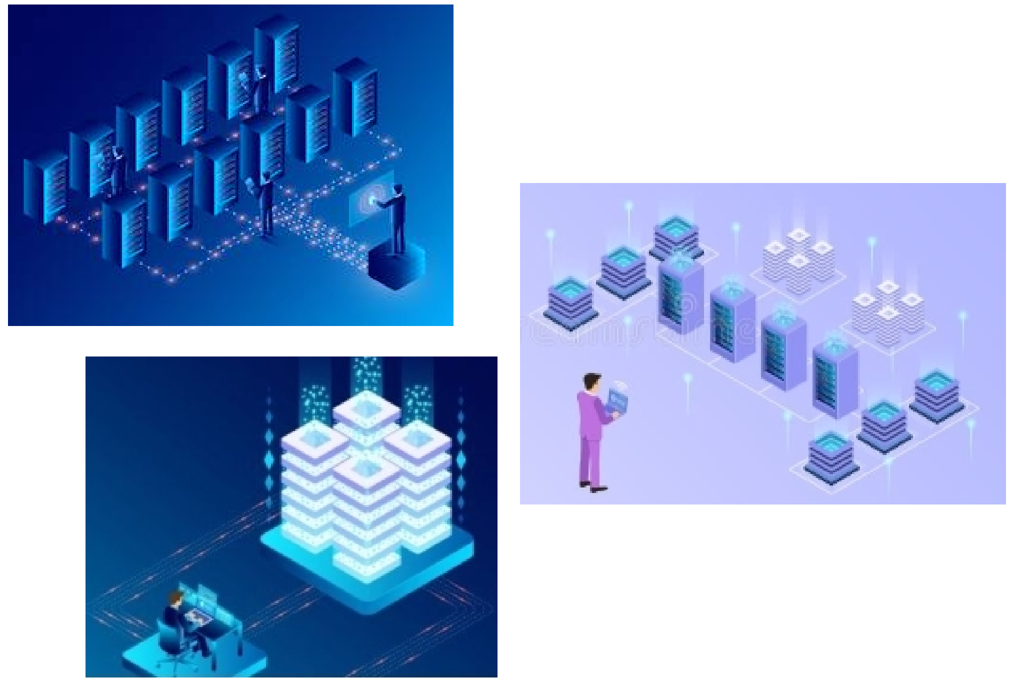
Mixed-use with other tenants

Funded by
the European Union

System Integration

System Integration

Independent systems, partially isolated
Different architectures and programming I/F
Mixed-use with other tenants



DestinE *systems are independent* (w *different architectures*) and *partially isolated*

- *mixed-use with local tenants* will be required
- creates *data pathing, scheduling*, and *workload portability/parametrization issues* in the federated solution
- DestinE Data Bridge design is a start for data pathing between sites but will not address all issues

Storage access and accounting are typically handled via the Unix permissions model on global shared storage. Storage access will need to implement:

- *guaranteed quality of service* (minimum or maximum bandwidth)

  or

- *extended data governance models*.

Funded by
the European Union

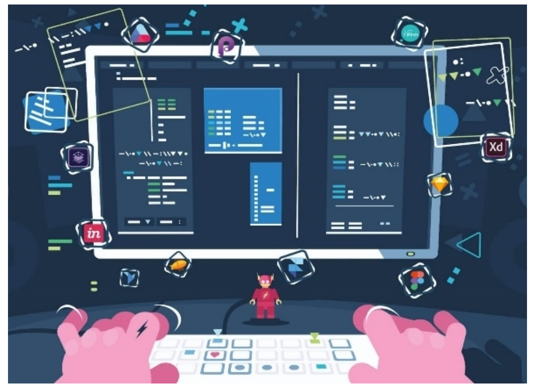
Service Deployment / Operation

Migrating traditional, session-based HPC systems to target a service model

Service Deployment / Operation, Orchestration, Scheduling

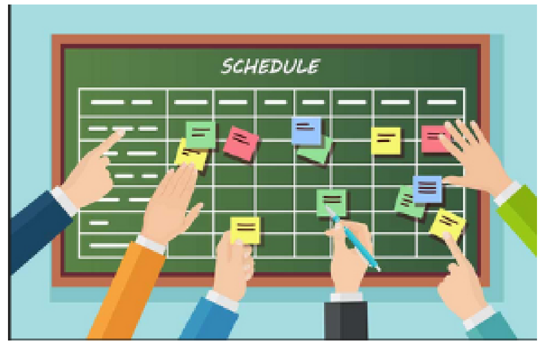
Service interfaces and Instantiation
Local Service Scheduling and Orchestration
Distributed Service Scheduling
Service Deployment and Updates
Quality of Service

Service interfaces and Instantiation



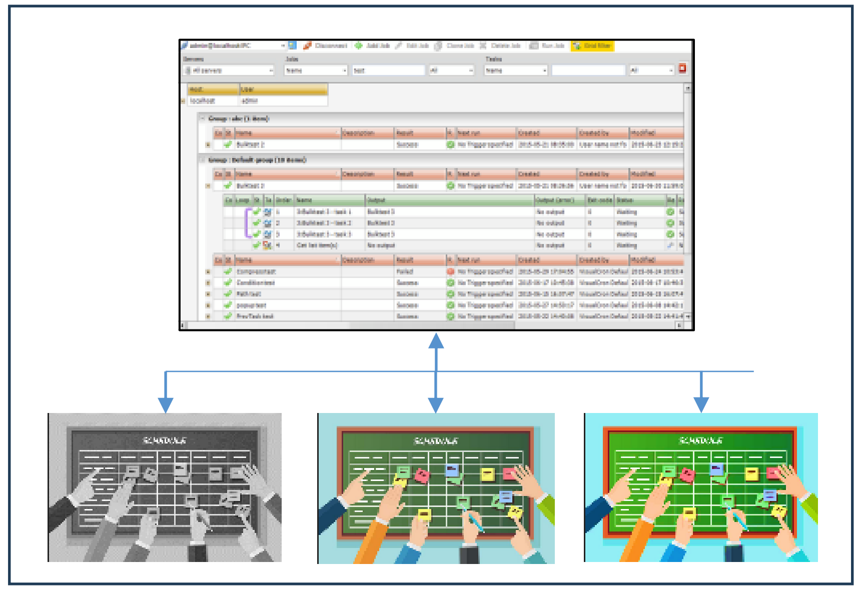
- Federated end-users → *centralized UI*. One user-side interface be imposed across centers?
  - HPC centres translate UI data internally to commands for the local systems?
  - Federated SW translates commands to interface to the HPC centres?

Local Service Scheduling and Orchestration



- *SLAs* → special mechanisms (data movement targeted networks, ...) or restriction of resources per job.
- *Prioritization* → pre-emption of running jobs
  - need for graceful degradation/migration and management of state information.
  - critical DT requests (like to aid disaster recovery) may exceed set quota limits, requiring a rebalancing.

Distributed Service Scheduling



- *Meta scheduling level* (longer term) → route service requests to the most suitable system (per resource requirements and status of systems).
  - Requires local status knowledge for each participating centre, including predicted time of execution of any newly invoked service.

Funded by the European Union

Service Deployment / Operation (2)

Migrating traditional, session-based HPC systems to target a service model

Service Deployment / Operation, Orchestration, Scheduling

Service interfaces and Instantiation
Local Service Scheduling and Orchestration
Distributed Service Scheduling
Service Deployment and Updates
Quality of Service

Software Environment Compatibility and Stability

- *Code porting and performance* are not uniform across systems
  - Ported codes can vary in execution across systems.
  - Many codes require porting and optimization (DT couplers, middleware, databases, deployment tools)
- *Software updates* → managed locally but federally aligned. Software containers could help, but
  - would require a common view on containerization policy, tools, and security
  - Container images must be adjusted with system changes.
- *Quotas* may be respected but still impact other applications through shared resources (network or memory).
  - At a minimum, the ability to monitor, diagnose and report these issues is required.
  - Longer term, QoS mechanisms to prioritize and police allocation.
- *Monitoring and observability* → address undetected observability of Destination Earth jobs by other users.

Service Deployment and Updates

- Local *system updates must be coordinated* to the federated system, and vice versa. Code changes must be reported to a central instance and should be automated where possible.
- *SW upgrades* may trigger a waterfall of service implementation updates, or at least recompilation, at the local and federated level.
- *continuous integration/continuous deployment (CI/CD)* should be planned at each center (longer term) to handle rebuilds and validation.

Funded by
the European Union

User Management

User Management

User Policy
Authentication, Authorization & Accounting
Data Governance
Operational Handling of User Management System

User Policy:

Central authentication of users is complex due to heterogeneity of the various hosting sites and compounded by the introduction of a federated service infrastructure

- *Nested user registration/authentication* needed
- *Definition and obsolescence of roles* will increase in type and number
- *Data governance* required due to proprietary (or at least restricted access) data

Authentication, Authorization & Accounting (AAA):

Federated solutions must comply with current local authentication processes or guarantee at least the same level of security.

- *Authorization* uses local data access control methods (based on access rights or “project ID”)
- *Accounting* also relies on user accounts or project IDs
  must account across federated systems
  will require fairness in allocating resources while managing user preferences and/or code limitations
  Accounting resolution may require higher precision to support nested allocation
- *Privacy* → Federation can introduce non-open data

- ***Authentication:*** validating identity
- ***Authorization:*** requested operation or resource is granted
- ***Accounting:*** compute resources used or data accessed vs quotas or billing
- **QoS**: description or measurement of performance of a service toward the user

Funded by the European Union

User Management (2)

User Management

User Policy
Authentication, Authorization & Accounting
Data Governance
Operational Handling of User Management System

Quality of Service (QoS)

Operational codes run in prescribed times. Regular HPC centre scheduling and resource management typically maximize total throughput, not completion times

- *Monitoring, prioritizing, and policing* needed, routing urgent service requests to HPC centres with capaci or desired compute. Existing jobs must either be migrated, downsized, or killed
- *Nested QoS* required

Data Governance

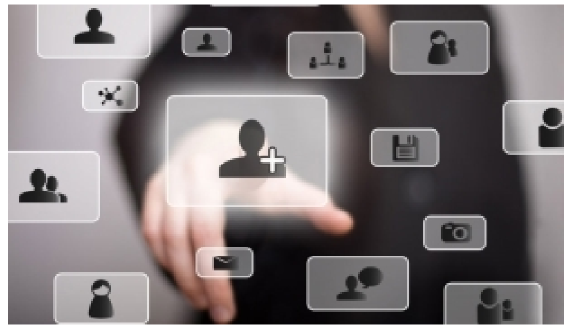
Express end-user privileges for creating/modifying, accessing and using DestinE data

- *Traditional access control* based on files will not be sufficient
- Need to *up-level* to capture/express *data usage patterns*

Operational Handling

Scaling of DestinE end-users and HPC centres will require a User Management Architecture which enables

- *Adding* and *managing end users* and their access privileges at the Core Platform
- *Setting up* and *managing identity mappings* for the participating HPC centres
- *Tracking usage patterns* and *data provenance*



- ***Authentication:*** validating identity
- ***Authorization:*** requested operation or resource is granted
- ***Accounting:*** compute resources used or data accessed vs quotas or billing
- **QoS**: description or measurement of performance of a service toward the user

Funded by
the European Union

ECMWF

ETP 4 HPC

Outlook

- Executing DestinE workloads will be a good challenge and proof point for EuroHPC's vision of a federated European HPC environment

- More challenges than outlined here exist
  - Data movement/Data streaming
  - Data acquisition/preprocessing/injection into twins
  - Security architecture/Data and code governance
  - Math and Algorithms

Thank you !

Funded by the European Union

Destination Earth:

DE_380: Data Streaming

Dumitru Roman, Valerio Frascolla,

Maria Perez, Gabriel Antoniu, Sai Narasimhamurthy

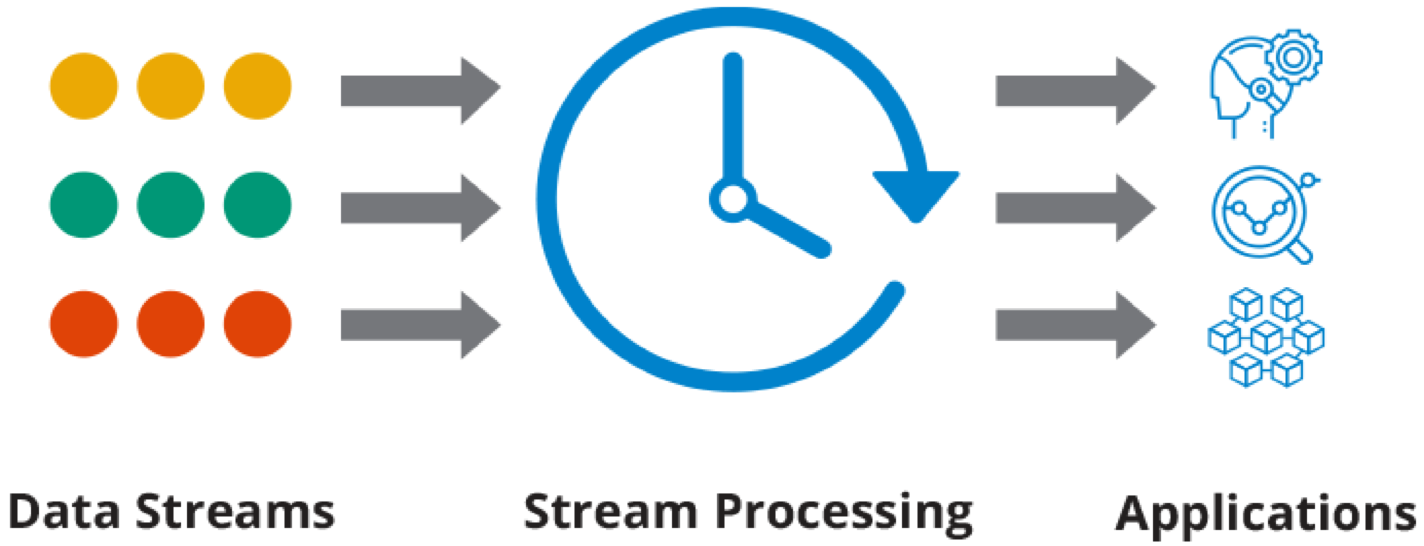
Context/Disclaimer

- This presentation is based on work identifying with respect to data streaming in EuroHPC infrastructures provided for running DestinE twins.
- This is not a summary of a report commissioned by ECMWF in the DE_380 contract.
- We believe the points discussed here are more widely applicable than DestinE, but DestinE is the first project highlighting them.

- *This document has been produced in the context of the Destination Earth Initiative and relates to tasks entrusted by the European Union to the European Centre of Medium-Range Weather Forecasts implementing part of this initiative.*
- *This document is funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them. The European Centre of Medium-Range Weather Forecasts is not liable in respect of this document and gives no warranty for the information needed.*

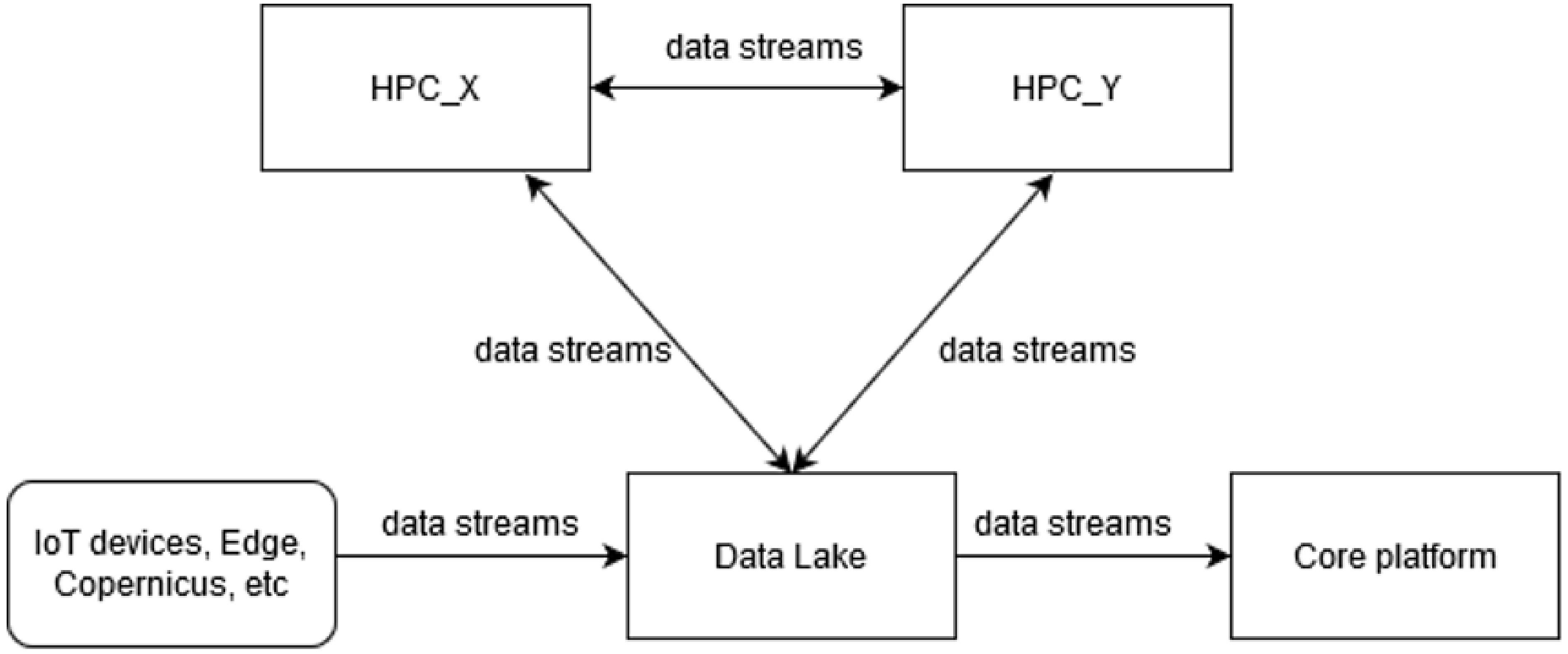
Funded by
the European Union

What is data streaming?



- Continuous transmission of data from a source to a destination
- Continuous generation of (small)
- Near real-time required to keep a current view of data
- Data consumers: human or applications
- Range from millisecond to hours
- Data streaming is used when real-time analysis is required
  - ○ E.g., retail, manufacturing, autonomous vehicles, cyber security, weather
- Challenges:
  - ○ Data consistency, scalability, reliability, etc.

Data streaming in Destination Earth

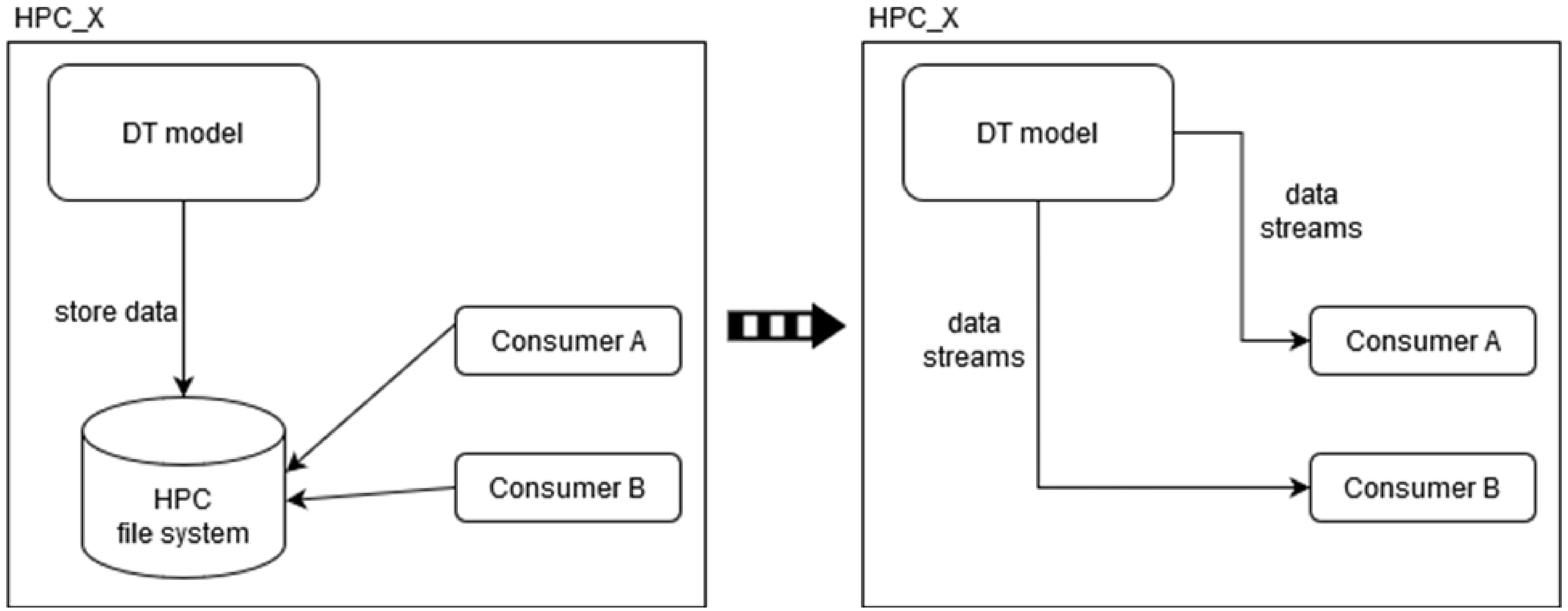


- Streaming may happen between the various entities involved
  - External sources
  - Digital twins running on HPC centers
  - Data Lake

- Streams may include meteorological data, results produced by DTs, and ephemeral data used to coordinate multiple DTs and pass information between them

- Data volumes are planned to be significant
  - ○ Operational data around 1 PB / day, DBr traffic 100s of TB / day, DL to CS 10s of GB / day

Data streaming in Destination Earth – within an HPC cluster



- Data will be transmitted continuously over the day, with bursts of activity at unpredictable times
- Notifications about availability of data are possible
- In general, data streams are not stored
- Data can be buffered for a short period of time
- Periodicity of stream creation can vary (> few seconds)

Data streaming in Destination Earth – within an HPC cluster (cont’)

DT producers

- Big models, **up to five**
- Forward data from ~~forecast or~~ simulations
- Stream data at a predictable rate
- Expected a homogeneous producer, with a standardized production protocol
- Data produced by a DT can be input to another DT
- The format of the raw data produced by the DT is open, the payload depends on the model

DT consumers

- **Larger number** of data consumers
- Cannot request data at certain rates
- Expected different consumption protocols (e.g., different frequency)
- Consumers may consume all data or partially; if they miss data and later need to access previously streamed data that was missed, they need to anticipate this
- Consumers can store the data if needed

Data streaming in Destination Earth – within an HPC cluster (cont’)

Fault tolerance

- Important data can be re-created
- Simulation models should include checkpointing
- Consumers behavior must not impact the producer
- Data is assumed highly reliable
- Gaps in data may exist
- Hardware failure will likely happen
- Requirements for the infrastructure are relatively strict

Throughput

- Data communication between DTs
- Expected latency is low for HPCs systems
- Throughput should be maximized for some resources

Security and access rights

- Dependent on the DT scenario
- Corruption of data in the file system
- Unauthorized access to data or malicious data corruption

Data stream challenges in Destination Earth

Architectural choices for the stream processing system

- Push or pull-based model
- Homogeneous or heterogeneous
- Control and orchestration
- Required capabilities

Data stream creation and access control

- Data source
- Data stream platform + integration
- Data producers
- Data consumers
- Monitoring
- Data archiving and storage

Funded by the European Union

Data stream challenges in Destination Earth

Data stream compression and encryption

- Dedicated HW accelerators vs. general purpose CPUs
- Alignment with other system-level security aspects and international regulations
- Balance energy consumption and performance
- Assess the need for both compression and encryption

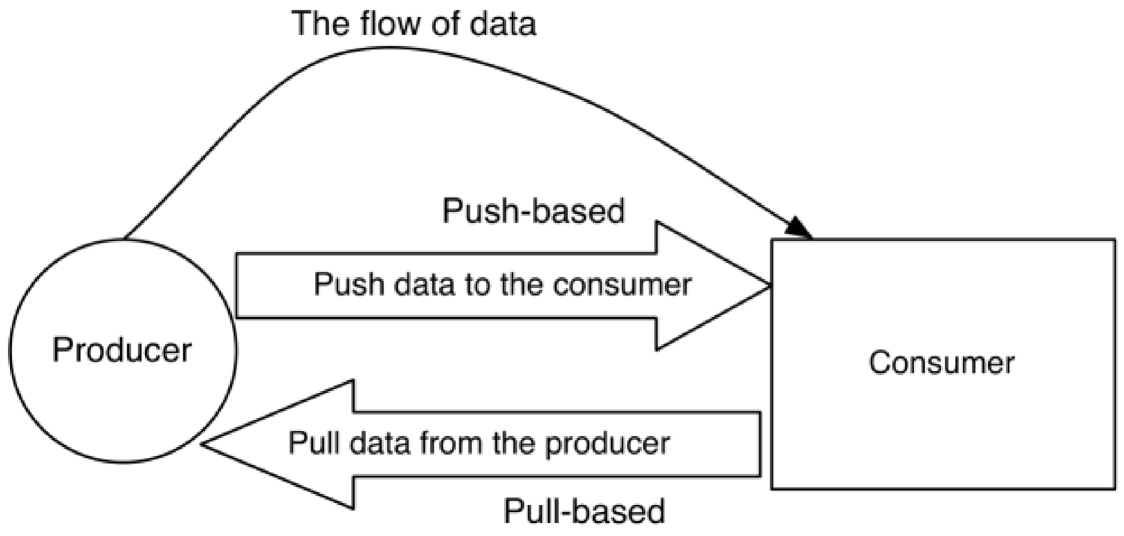
Streaming infrastructure considerations (HPC I/O)

- I/O services ephemerally spawned “Data Nodes”
- Use of NVMs as Burst Buffers in Data Nodes
- Use of FDB as object storage backend
- Use of hierarchical storage managers

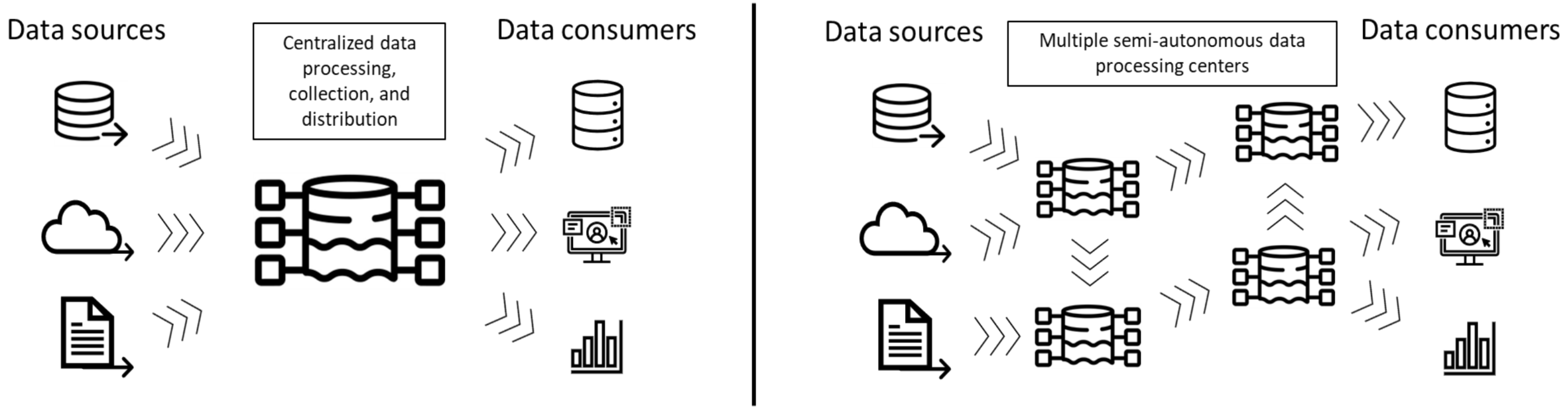
Funded by the European Union

Architectural aspects in data streaming systems

push-based vs. **pull**-based
stream processing model



homogeneous vs. **heterogeneous**
centralized vs. **federated**
data streaming systems


Data sources
Centralized data processing, collection, and distribution
Data consumers
Data sources
Multiple semi-autonomous data processing centers
Data consumers

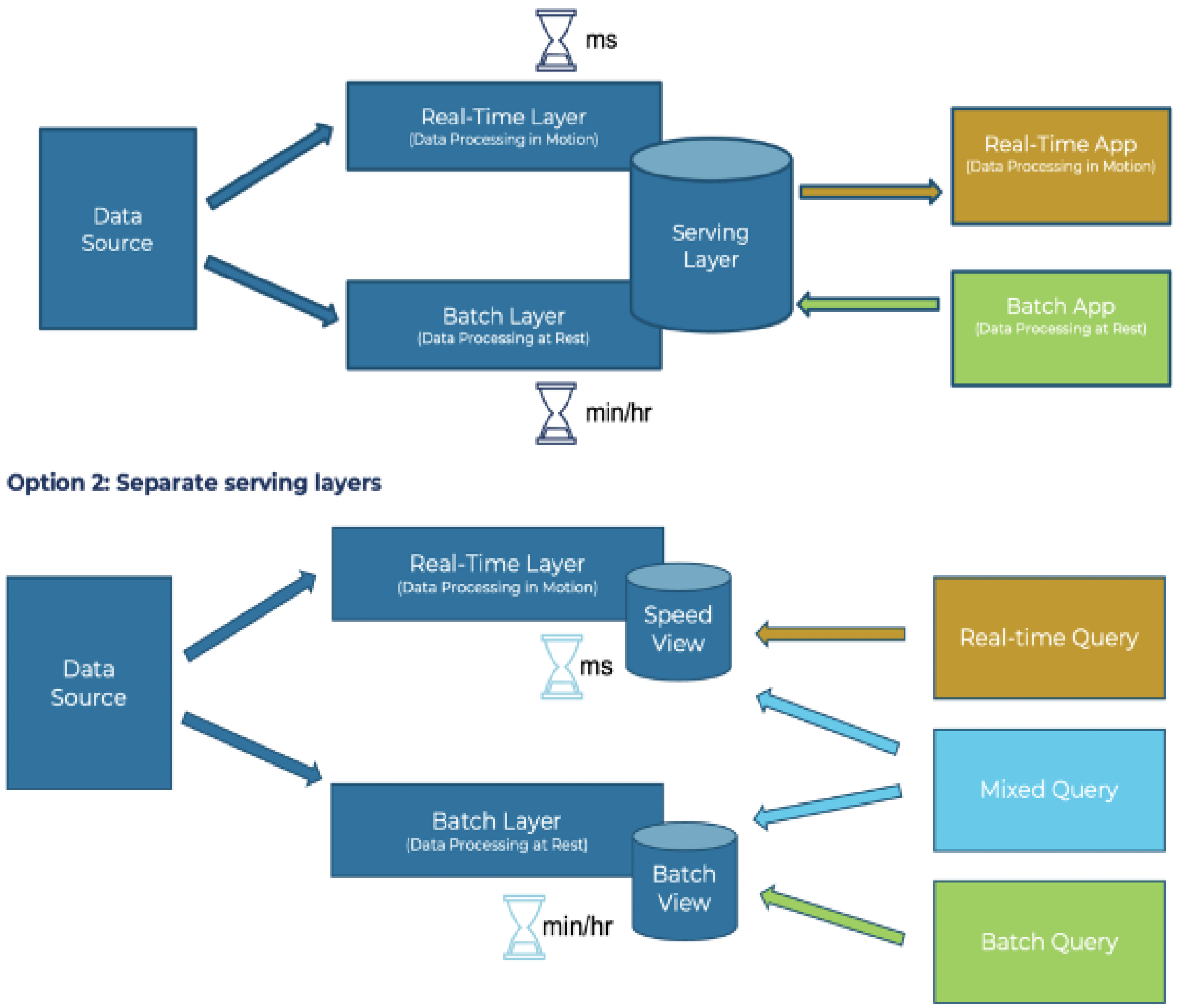
Funded by the European Union

Architectural aspects in data streaming systems (cont')

control and orchestration of streaming systems: Lambda, Kappa, Event-driven, Cloud-native architectures

Lambda Architecture

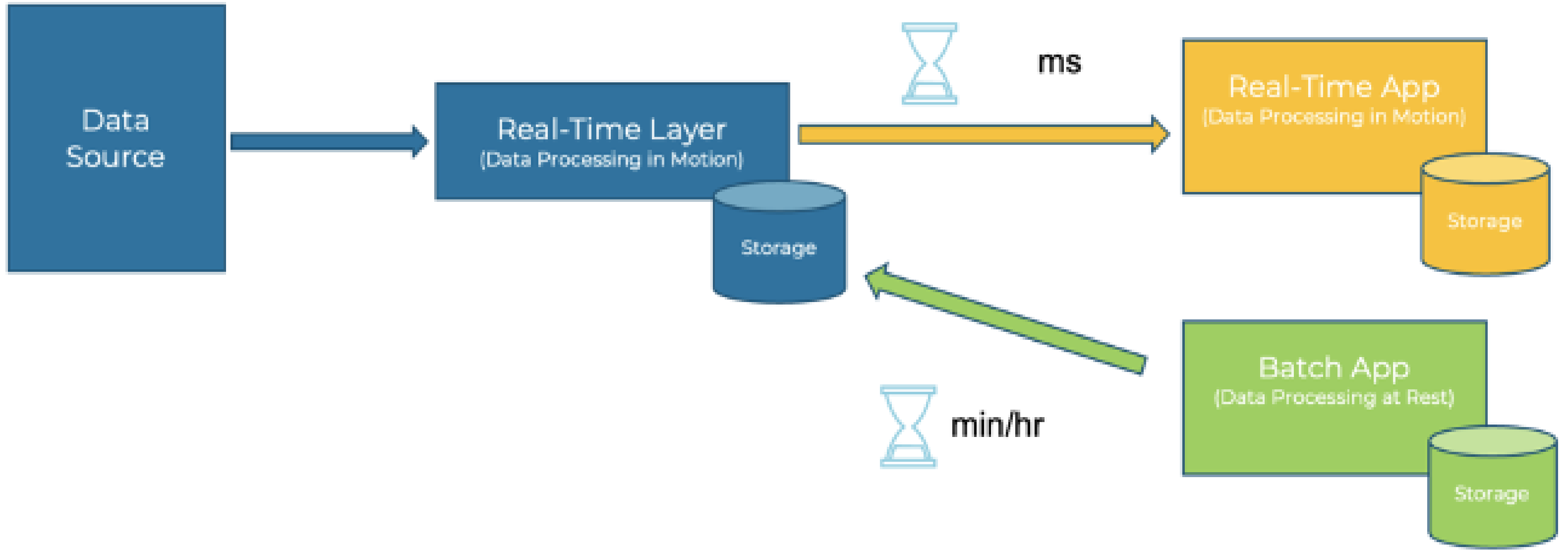
Option 1: Unified serving layer



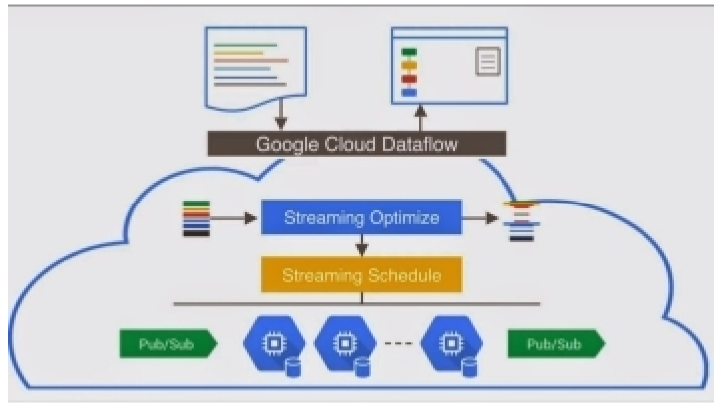
ms
Data Source
Real-Time Layer
(Data Processing in Motion)
Serving Layer
Batch Layer
(Data Processing at Rest)
min/hr
Real-Time App
(Data Processing in Motion)
Batch App
(Data Processing at Rest)
Option 2: Separate serving layers
Data Source
Real-Time Layer
(Data Processing in Motion)
Speed View
ms
Batch Layer
(Data Processing at Rest)
Batch View
min/hr
Real-time Query
Mixed Query
Batch Query

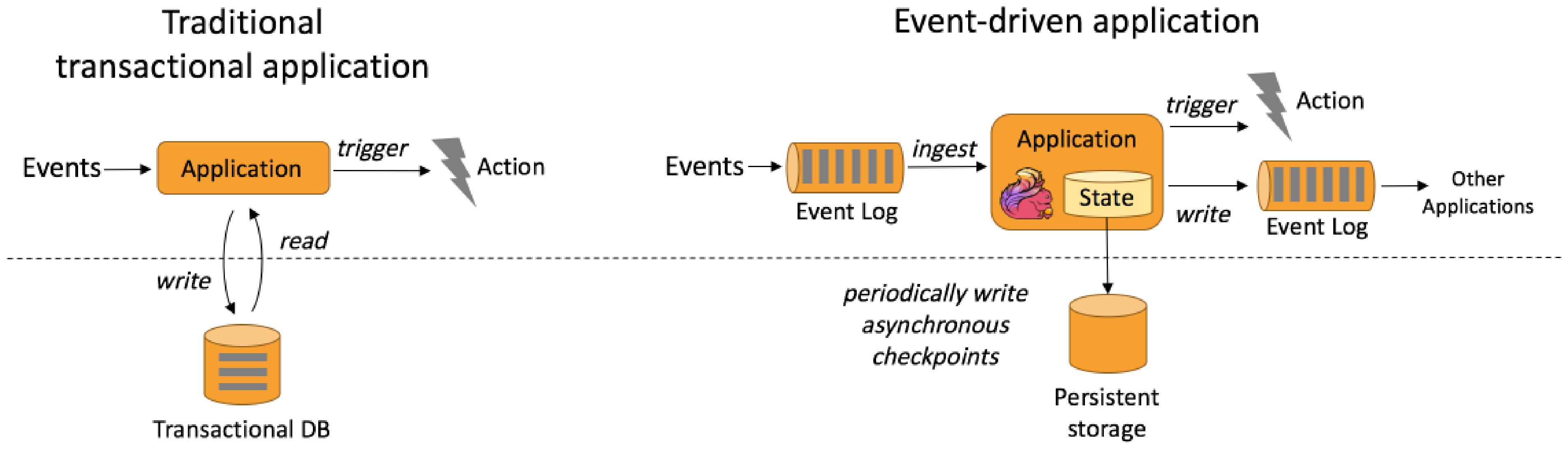
Kappa Architecture

One pipeline for real-time and batch consumers



Data Source
Real-Time Layer
(Data Processing in Motion)
Storage
ms
Real-Time App
(Data Processing in Motion)
Storage
min/hr
Batch App
(Data Processing at Rest)
Storage





Funded by the European Union

Architectural aspects in data streaming systems (cont’)
Capabilities of streaming systems

- **Scalability**: scale resource allocation up or down as needed to handle large amounts of data
  - Also to handle unpredictable spikes in streaming
- **Reliability**: cope with data loss, duplicates, or incorrect data processing
- **Real-time processing**: avoid not adding significant latency; data processed without any errors
- **Data integration**: integrate data potentially streamed from a variety of sources in different formats
- **Security**: effectively control access to data streams and prevent data loss
- **Fault tolerance**: operate the data stream even if some of its components fail

Funded by
the European Union

Architectural possibilities for the stream processing system in Destination Earth

Architectural choices for the stream processing system

Push or pull-based model
Homogeneous or heterogeneous
Control and orchestration
Required capabilities

Push or pull-based model

Push-based: data is actively pushed from the data source to the consumers as soon as it becomes available

- Possible scenario: continuously generated data from DTs to the Data Lake

Pull-base: clients request data from the source at their discretion, and data transmission only starts on request

- Possible scenario: consumer-initiated stream when end users invoke a DT service

Hybrid design may also be considered

Homogeneous or heterogeneous; centralized or federated design

- Federation naturally suggest a heterogeneous design, with distributed accountability and decentralized decisions
- Unified interfaces will be needed to make data findable

Control and orchestration

- Lambda vs. Kappa vs. Event-driven vs. Cloud
- Depends on: type of data to be streamed, type of processing required, the need (or not) for low-latency processing or real-time processing

Required capabilities

- Aspects to be considered/prioritized: scalability, reliability, fault-tolerance, maintainability

Funded by the European Union

Data stream creation and access control in Destination Earth

Data stream creation and access control

Data source
Data stream platform + integration
Data producers
Data consumers
Monitoring
Data archiving and storage

Data source

- Data is typically binary data from HPC codes (generated in the Centre's ECMWF Integrated Forecasting System (IFS) or DTs)
- Challenge: interoperability, which may require the use of semantic technologies

Data stream platforms + integration

- A thorough evaluation and benchmarking of reference stream processing platforms will be required
- Assessment required in the light of the target capabilities
- Installation and configuration of the chosen data stream platform consistently with the operational policies of the data centres

Data producers and consumers

- While producers are limited (up to 5 DTs) and likely homogeneous, consumers can be heterogeneous, using different consumption protocols and their number is high

Testing and optimization

- It is relevant to provide different non functional properties, such as scalability, fault tolerance, reliability, security and compliance with existing regulations

Data storage

- Processing data streams must be able to interact with the storage, without affecting the overall performance

Funded by
the European Union

Data stream compression and encryption in Destination Earth

Data stream compression and encryption

Dedicated HW accelerators vs. general purpose CPUs
Alignment with other system-level security aspects and international regulations
Balance energy consumption and performance
Assess the need for both compression and encryption

Dedicated HW accelerators vs. general purpose CPUs

- GPUs vs. CPUs vs. dedicated HW accelerators vs. FPGAs
- Assess the possibility to use the same HW for compression and encryption
- The processing demand depends on the number of 'secure islands' identified

Alignment with other system-level security aspects and international regulations

- Several aspects related to security in a federation of HPCs scenario need to be considered together with compression and encryption at an early phase of the design of the system
- Streams of data may cross Member States' borders, potentially moving among different legislations related to security aspects

Balance energy consumption and performance

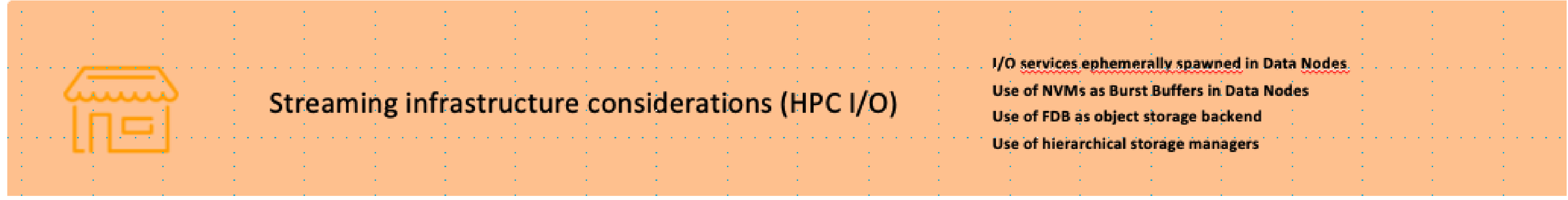
- Both encryption and compression, when dealing with a huge amount of data to be processed on-the-fly, imply a non-trivial amount of energy to perform the tasks; increased security vs. increased energy demand need to be assessed

Assess the need for both compression and encryption

- Sometimes compression may not be needed at all
- Sometimes encryption may not be needed at all

Funded by the European Union

Streaming infrastructure considerations in Destination Earth



Streaming infrastructure considerations (HPC I/O)

I/O services ephemerally spawned in Data Nodes
Use of NVMs as Burst Buffers in Data Nodes
Use of FDB as object storage backend
Use of hierarchical storage managers

Ephemeral services

- MultiIO (and its associated pipelines) and FDB services could be handled by separate Data Nodes
- “Burst buffers” based epheremal services such as GekkoFS working with NVMs

Use of NVMs as Burst Buffers in Data Nodes

- Very high throughput and low latency
- “Stage” the required data from the lower storage tiers and be used by workflows timely and efficiently

Use of FDB as object storage backend

- Final output (GRIB data) stored through FDB, such as DAOS or CEPH as objects in memory

Use of hierarchical storage managers

- E.g., Hestia
- Used to move data back and forth across the storage hierarchy between the Data nodes
- Better use of memory at the very top of the hierarchy in the data streaming context

Funded by
the European Union

Other streaming related topics that could be considered within Destination Earth

- Containerization and orchestration
- AI-Driven optimization
- Emerging technologies for parallel processing
- Edge and Fog Computing integration
- Data quality assurance
- Access rights management
- Preparing for Quantum Computing
- Interoperability standards
- Evaluation and benchmarking of existing data streaming platforms
- Compression/Encryption in HPC
- Energy consumption

TC I
TransContinuum Initiative

And now: Your questions!

DE_380 for Destination Earth:

Cyber Security

Evangelos Markatos, Roberto Cascella, Ana Ayerbe Fernandez-Cuesta, Fabio Martinelli, Artsiom Yautsiukhin

Funded by the European Union Destination Earth implemented by ECMWF esa EUMETSAT

Context/Disclaimer

- This presentation is based on work identifying Cyber Security challenges for running DestinE twins.
- This is not a summary of a report commissioned by ECMWF in the DE_380 contract.
- We believe the points discussed here are more widely applicable than DestinE, but DestinE is the first project highlighting them.

- *This document has been produced in the context of the Destination Earth Initiative and relates to tasks entrusted by the European Union to the European Centre of Medium-Range Weather Forecasts implementing part of this initiative.*
- *This document is funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them. The European Centre of Medium-Range Weather Forecasts is not liable in respect of this document and gives no warranty for the information needed.*

Funded by
the European Union

What is Cyber Security?

- Cyber Security provides any system with three properties:
- **Confidentiality**
- **Integrity**
- **Availability**

- Why do we need security?
- Without security
  - We **can not trust** the results of the system
    - We can not trust the inputs – we can not trust the outputs
    - We can not even trust that the system will be **available**

TCI
TransContinuum Initiative

Cyber Security Challenges in Destination Earth

- We have identified the following challenges:

1. **Asset** Identification and Risk Management
2. **Threat** Landscape
3. Cyber Secure **Communication** Systems
4. Secure **Data**
5. **Identity** Management
6. **Zero Trust**
7. Secure Management of **Models** and **Workflows**

- Some of these challenges apply to other systems as well
- But all of the above apply to DestinE

TC I
TransContinuum Initiative

1. Asset Identification and Risk Management

- **What** do we want to **secure**? What do we want to **protect**?
- Asset identification applies to all systems that need to be secured
- It is especially important for systems with lots of public data and lots of users
- Example Assets:
  - **Data** (at rest, in transit, and in computation)
    - Challenge: **which data** need to be protected? which data are restricted? which are public?
      - private data? password files? system data? personal data?
  - Digital twin **models**
    - trapdoors? bugs? viruses? third-party libaries?
- Directions and Options:
  - We need to provide an **initial detailed list** of assets to secure
  - Monitor and **update** the list regularly for new assets

TCI
TransContinuum Initiative

2. Threat Landscape

- **Who** is the adversary? Who is the enemy?
- **What** can they do to the system?
- Change the data (**integrity**)
  - without anyone noticing (we need integrity hashes...)
- Steal the data (the confidential ones - if any)
- Insider attacks
- DoS Attacks - make the system unavailable at critical times
- Ransomware (**encrypt selected important files** - passwords? parameters?)
- Un-authorized access
- Supply chain attacks (are all the third-party libraries trusted?)
- Directions and options
  - update and validate threat models
  - Engage a **Security Operations Center** to monitor threats and attacks

TCI
TransContinuum Initiative

3. Communication Systems

- Are there any communications that need protection?
- If so, which ones? (link with Identification of Assets)
- For those communications, we should worry about
  - **Snooping** on the data - **confidentiality**
  - Changing the data - **integrity**
  - Use of weak encryption protocols
    - That can be broken by modern computers
  - Man-in-the-middle attacks
- Directions and Options:
  - use **strong crypto** when confidential data are communicated
  - consider the benefits of **post-quantum crypto** as well
  - Consider **costs**: encrypting everything may be expensive - prioritize

TCI
TransContinuum Initiative

4. Secure data, data lakes and data centric policies

- Attacks against data at rest
- Access control
- Data lakes - decentralization: Who is responsible for the data in a distributed environment?
- Who has **access to the data**? How is this enforced?
- **Data usage control**: what operations are allowed on the data?
  - Goes beyond read/write access and includes:
  - Time to live: when should data be deleted?
  - Which nodes can use the data?
  - Obligation Management
    - e.g. compliance to policy requirements
  - Anonymization,
  - Physical spaces where data can reside, etc.

TCI TransContinuum Initiative

5. Secure and Trusted Digital Identity Management

- Authentication - Authorization
- User Management
- Challenge:
  - EuroHPC sites may have their own sets users
  - Data lakes may have a different set of users
  - It is not clear that there is a single authentication authority
  - Need to reconcile the **different sets of users**
    - some users have access to EuroHPC sites
    - some users have access only to the data and not to any HPC
    - How do you make the two (or several) sets inter-operate?
- Directions and Options:
  - Use several sets of users - at least two
    - **Privileged** users have access to HPCs and the models: they generate data
    - **Ordinary** users have read-only access to the data lakes

TCI TransContinuum Initiative

6. Zero Trust Architecture (I)

- Traditional approaches to security are **perimeter-based**:
  - Secure a perimeter
  - Focus on keeping enemies outside
  - Yet, once the perimeter passed, the way is free
- **New approach**:
  - Identify the most important sub-systems (e.g., servers)
  - Do not trust any other system elements/users
  - **Always verify access** to these sub-systems
  - **Monitor** access and the sub-system
  - Be ready to **revoke access**

- Zero Trust Moto:
  - Never Trust – Always verify

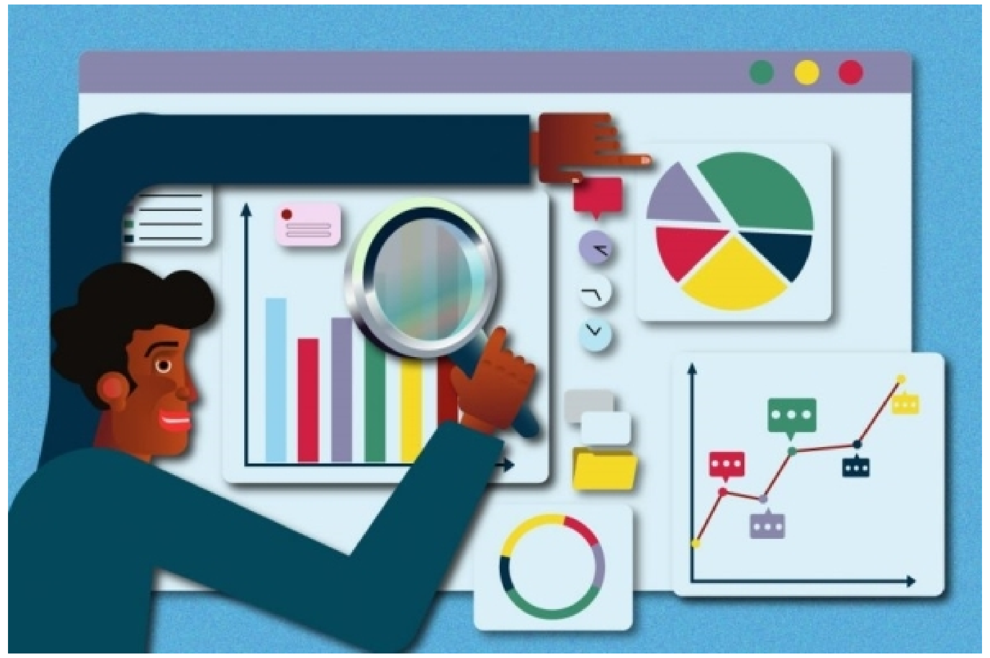
6. Zero Trust Architecture (II)

- Strengthen **access control**:
  - Consider the use of a single administration point to govern all enforcement points
  - Define and enforce access policies having possible risk in mind
    - e.g., additional credentials may be asked for risky operations
  - Ensure that only **least privileges** are granted
  - Authenticate all users (including other elements of DestinE system)
  - Use strong authentication mechanisms
  - Allow an **access only per session**

TCI
TransContinuum Initiative

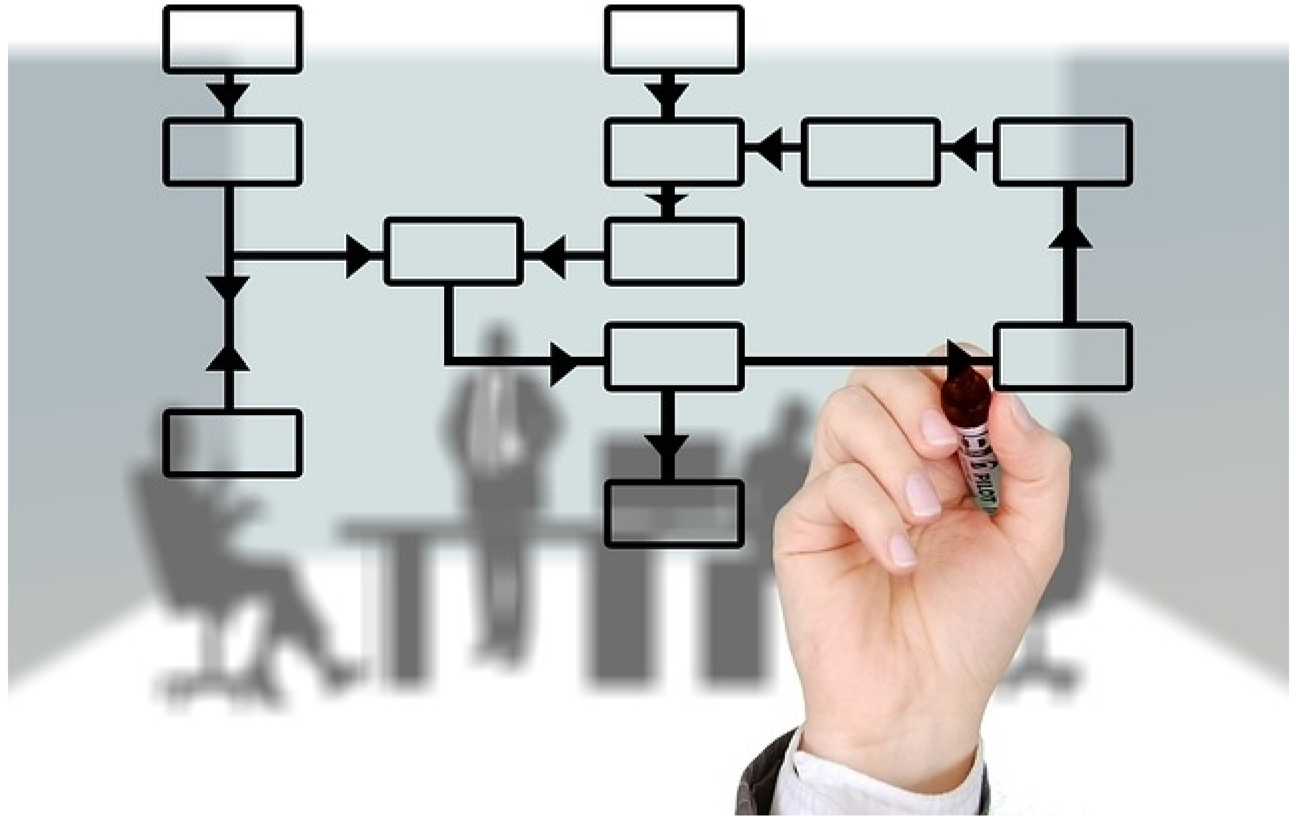
6. Zero Trust Architecture (III)

- Do not trust even after granting access:
  - **Monitor** access parameters
  - Monitor system health (e.g., patching level)
  - Monitor ongoing events (e.g., attacks and suspicious actions)
  - Monitor user actions
  - Enforce dynamic policies (or usage control)
    - **Re-evaluate access decision** when required



7. Secure Management of Models and Workflows

- Models as software are susceptible to all software attacks as **vulnerabilities, backdoors**, weak crypto, …
- It would be useful to adopt malware and vulnerability **scanning tools** to sanitize the models and avoid possible malicious behaviors.
- One should also consider **auditing and tracking** approaches for model evolution inside the system, and ensuring integrity with signatures
- This also entails data tampering, i.e. unauthorized modification of data within the DTs

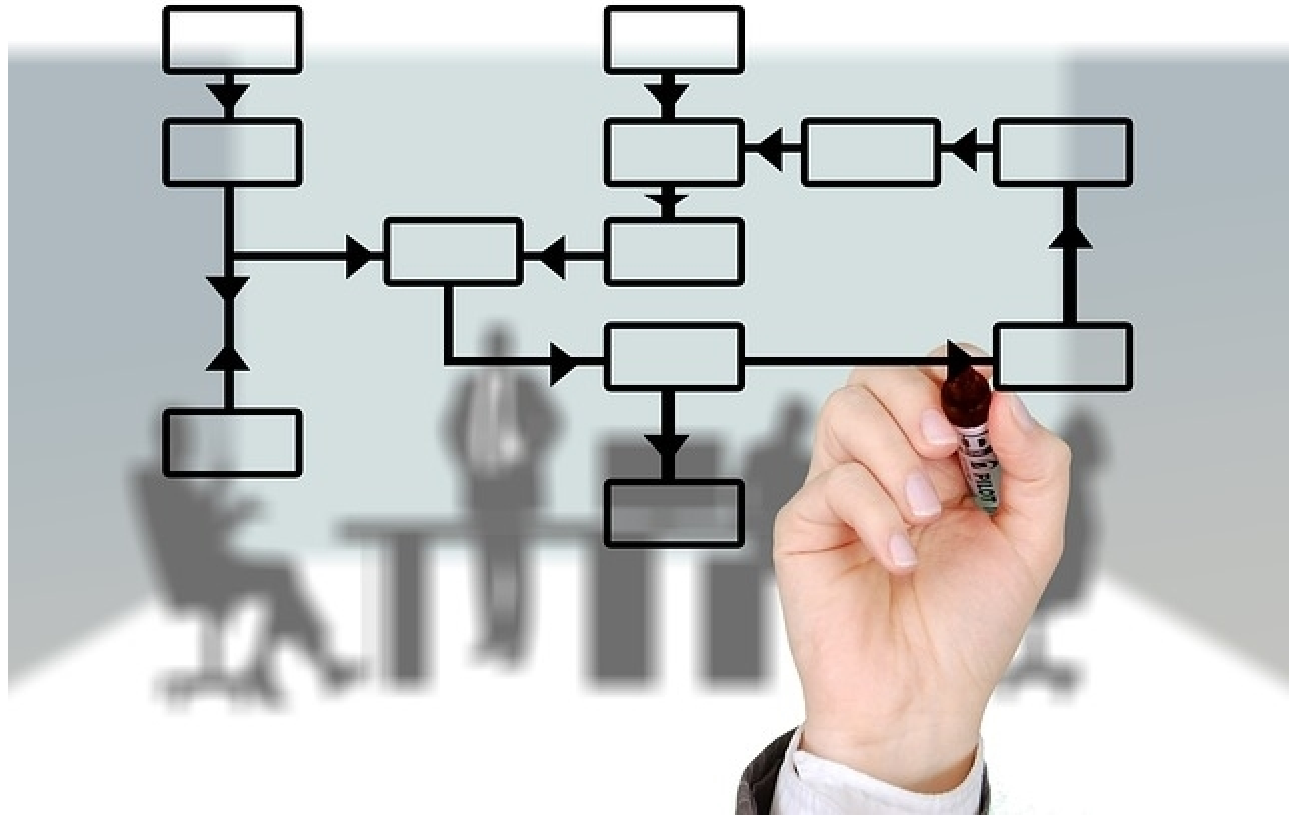


7. Secure Management of Models and Workflows (II)

- We thus need to ensure that models are managed by **privileged users**, allowing only those users to install, modify and execute the models.

  User profiling could be also useful.

Considering holistic workflow management for model creation, storage, modification, execution and termination would be necessary in the long term with their specific authorization frameworks



Summary - Prioritization

- Identify what (**assets**) you want to protect and against what (**threats**)
  - Assess risks and identify relevant security measures
- **Identify** and authenticate your **users**
  - without authentication, protection has little meaning
- **Secure** all processing of the assets
  - Storage and access (**secure data**)
  - Transfer (**communication**)
  - Computation (**secure models and workflows**)
- **Adopt zero-trust** attitude
  - never trust - always verify

TCI
TransContinuum Initiative

TC I
TransContinuum Initiative

And now: Your questions!

ECMWF

ETP 4 HPC

Acks

- Images from Pixabay

Destination Earth

Destination Earth: Challenges in data assimilation, weather simulations, and streaming data with Machine Learning

Z. Horváth, G. Louppe, J.A. Lozano, V. Monbet, on behalf of EU-MATHS-IN

20/03/2024

EU MATHS IN

Funded by the European Union Destination Earth implemented by ECMWF esa EUMETSAT

ETP 4 HPC

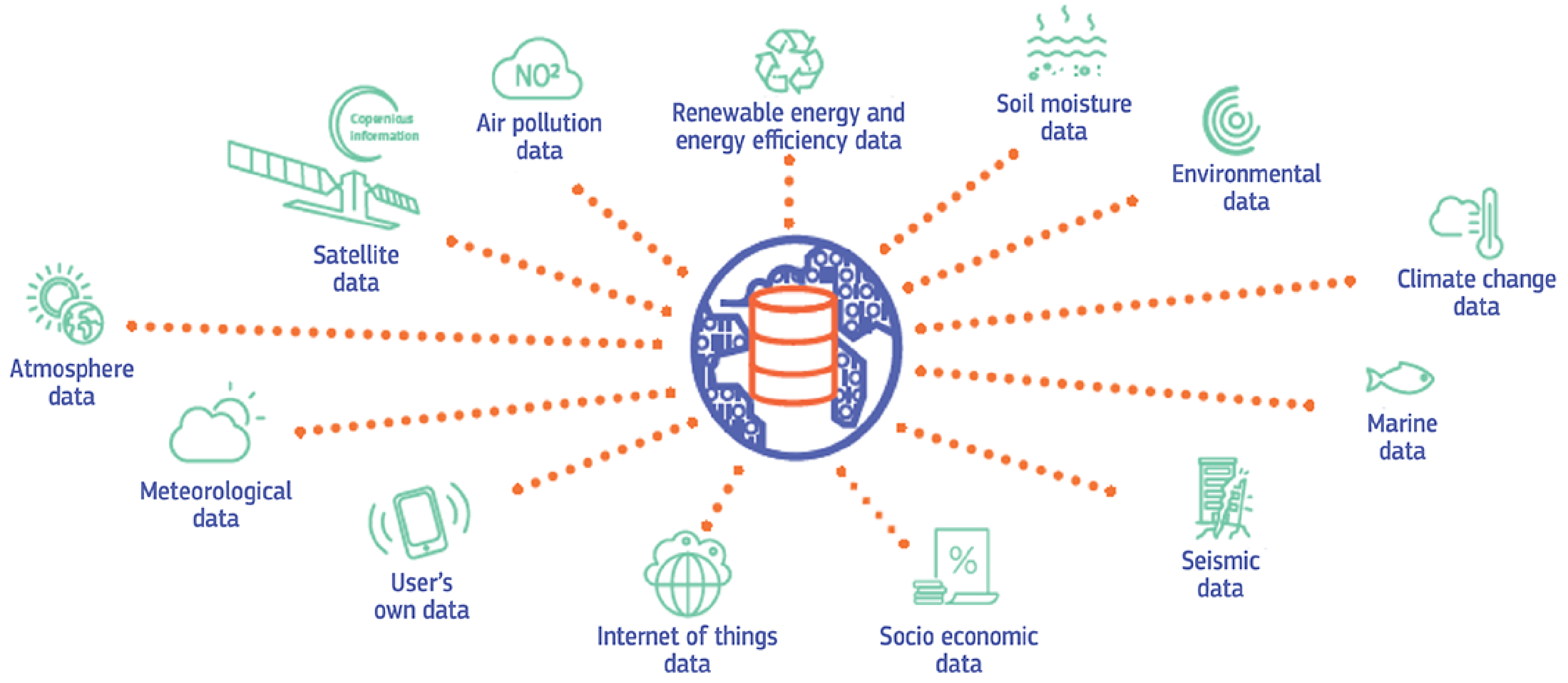
Context/Disclaimer

- This presentation is based on work identifying mathematical and algorithmic challenges of machine learning for running DestinE twins.
- This is not a summary of a report commissioned by ECMWF in the DE_380 contract.
- We believe the points discussed here are more widely applicable than DestinE, but DestinE is the first project highlighting them.

- *This document has been produced in the context of the Destination Earth Initiative and relates to tasks entrusted by the European Union to the European Centre of Medium-Range Weather Forecasts implementing part of this initiative.*
- *This document is funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them. The European Centre of Medium-Range Weather Forecasts is not liable in respect of this document and gives no warranty for the information needed.*

Funded by
the European Union

Introduction



- DestinE will produce 1 PB of data per day to the DestinE Data Lake.

- Machine Learning (ML) technology, incl. methods, software, and hardware is developing explosively. ML-based emulators are much faster than traditional simulation models.

- Main questions:

- Shall ML change the operational simulation methodologies for DestinE? To what extent is that cost-effective?
- What new math & algorithms are needed for the ML-based software of the DestinE digital twins?

Funded by the European Union

Main math&algo challenges

ECMWF has identified 3 sets of challenges in the DE_380 project:

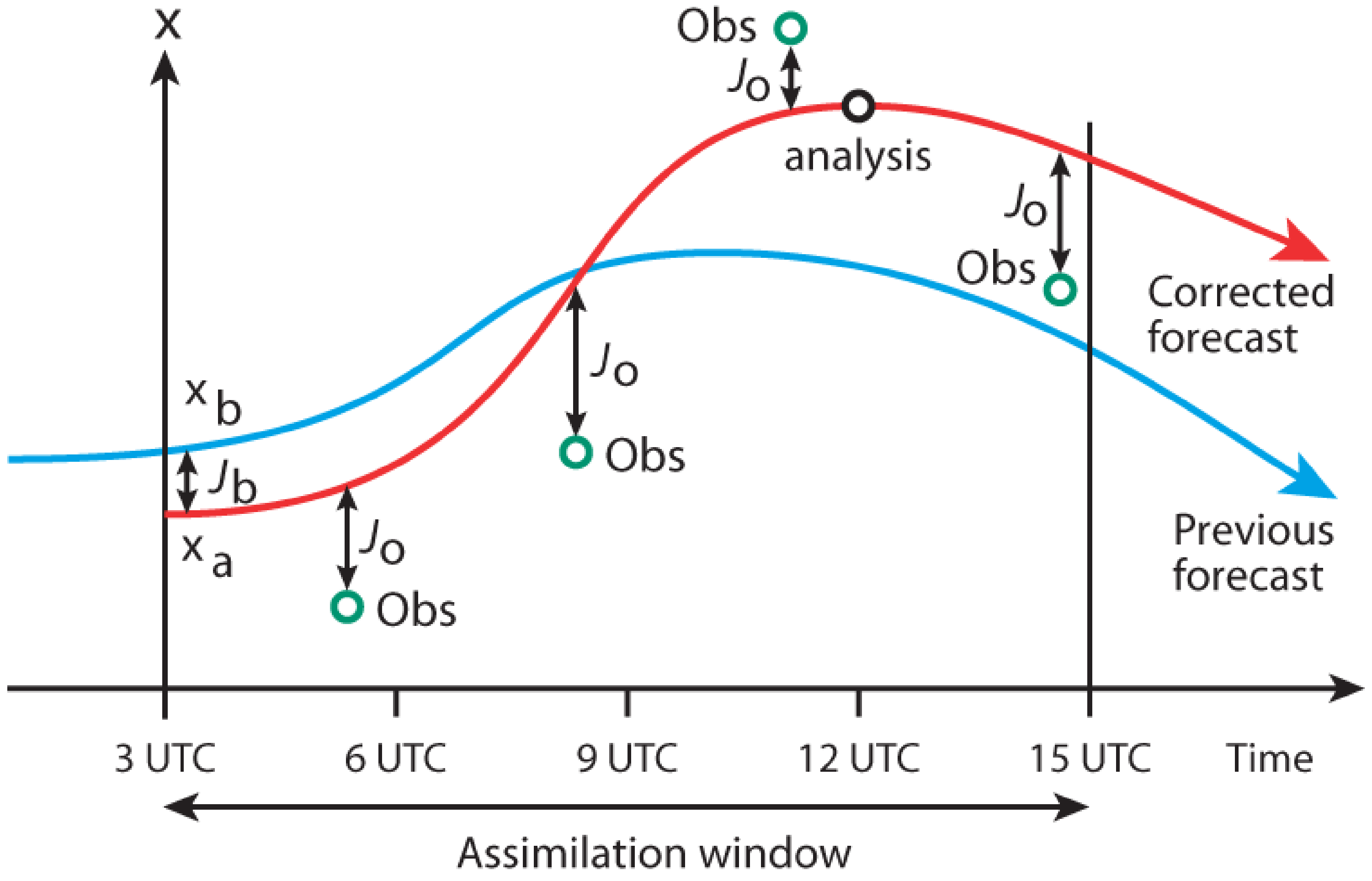
How to use ML cost-effectively for

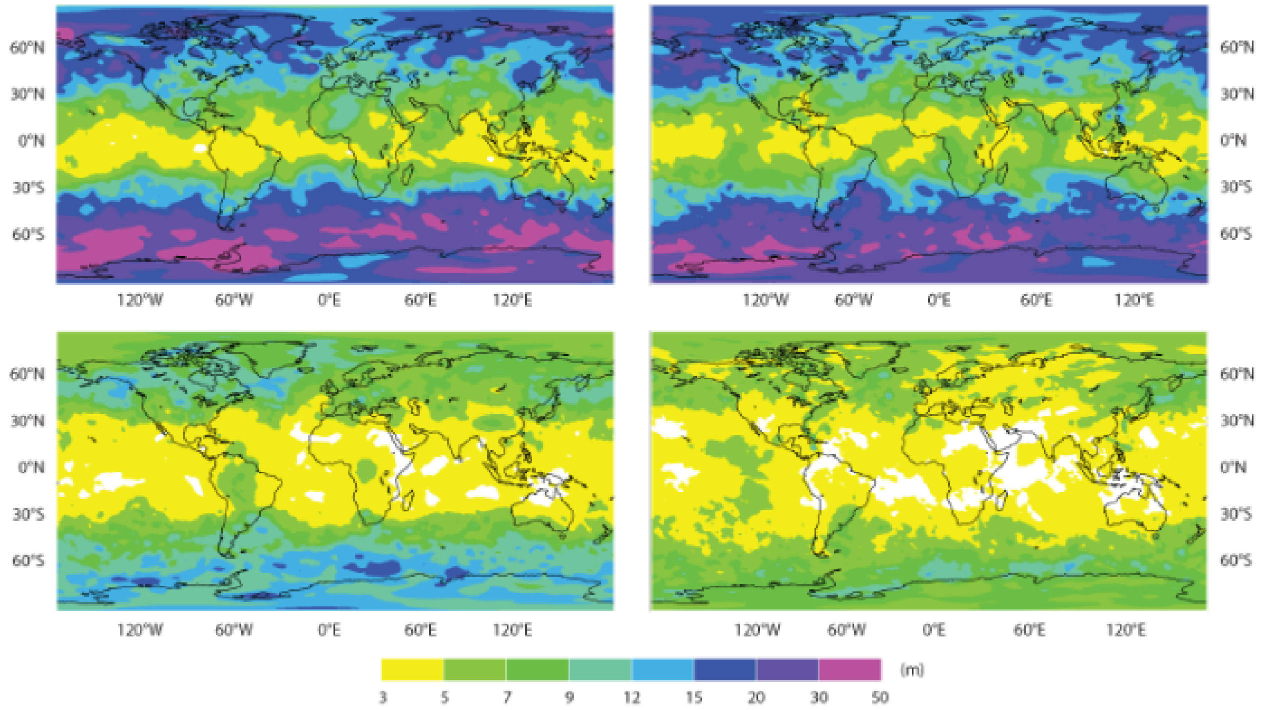
1. reliable data assimilation,
2. numerical weather forecasting,
3. streaming data,

in the context of DestinE digital twins?

Funded by the European Union

Challenge 1: Reliable data assimilation with ML – **State-of-the-Art**





The charts show the root-mean-square error for 24-hour forecasts of 500 hPa geopotential height in ECMWF's operational assimilation system for the month of October 1997 with 3D-Var (top left), October 1998 with 4D-Var (top right), October 2007 with 4D-Var (bottom left), and October 2016 with 4D-Var and the Ensemble of Data Assimilations (EDA).

https://www.ecmwf.int/en/about/media-centre/news/2017/20-years-4d-var-better-forecasts-through-better-use-observations

Data assimilation is used to estimate the state of the Earth system from observations.

Funded by
the European Union

Challenge 1: Reliable data assimilation with ML – **State-of-the-Art**

$$p(x_{1:t}|y_{1:t}) = \frac{p(y_{1:t}|x_{1:t})p(x_{1:t})}{p(y_{1:t})}$$

- Due to operational constraints, the posterior is often approximated with a point estimate of the current state.

- The full Bayesian posterior would instead provide a more **complete, principled, and reliable description of the uncertainty** in the current state, accounting for
  - Prior uncertainty in the state
  - Observational uncertainty
  - Nuisance parameters in the physical model
  - Stochastic variability in the system.

- Recent advances in **deep generative models** offer a promising approach to approximating the **full posterior distribution** of the states.

Funded by
the European Union

Challenge 1: Reliable data assimilation with ML – **Research perspectives**

Short-term:

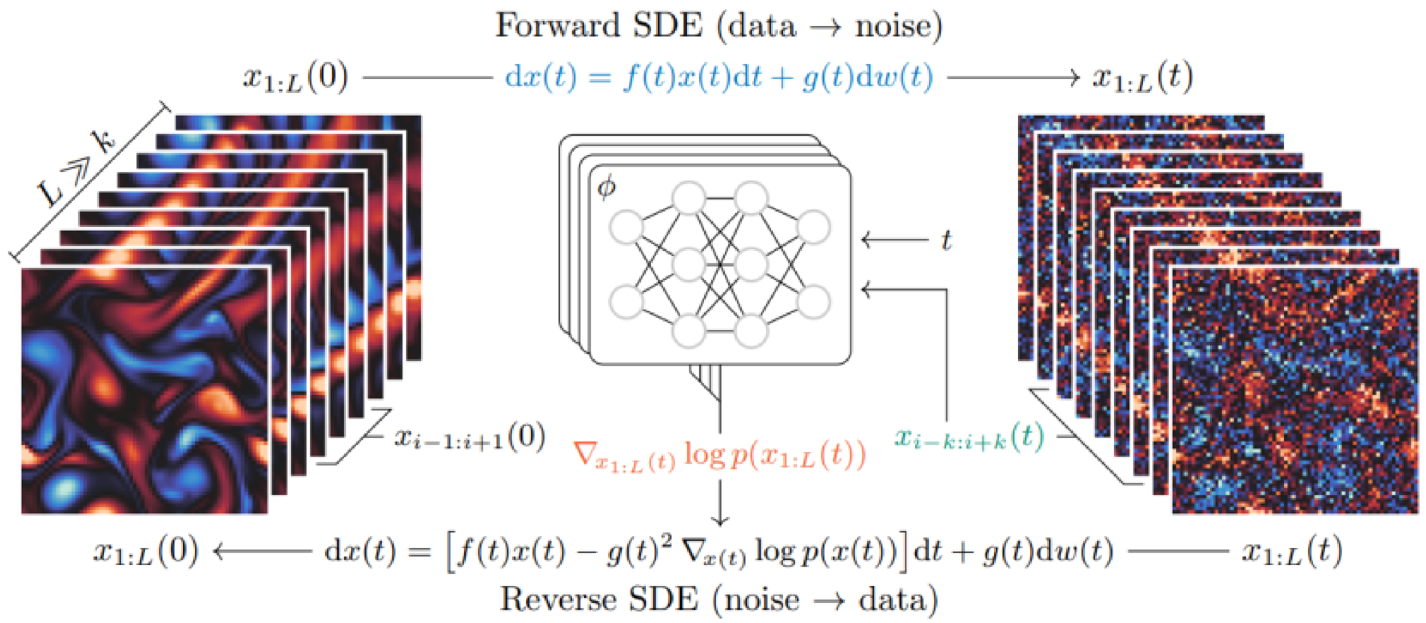
- Accelerate current DA systems with data-driven emulators and automatically differentiable models.
- Design benchmarks and validation platforms to encourage the scientific community to innovate.

Mid-term:

- Change of paradigm: full posterior reconstruction with deep generative models (e.g. diffusion models).

Long-term (speculative):

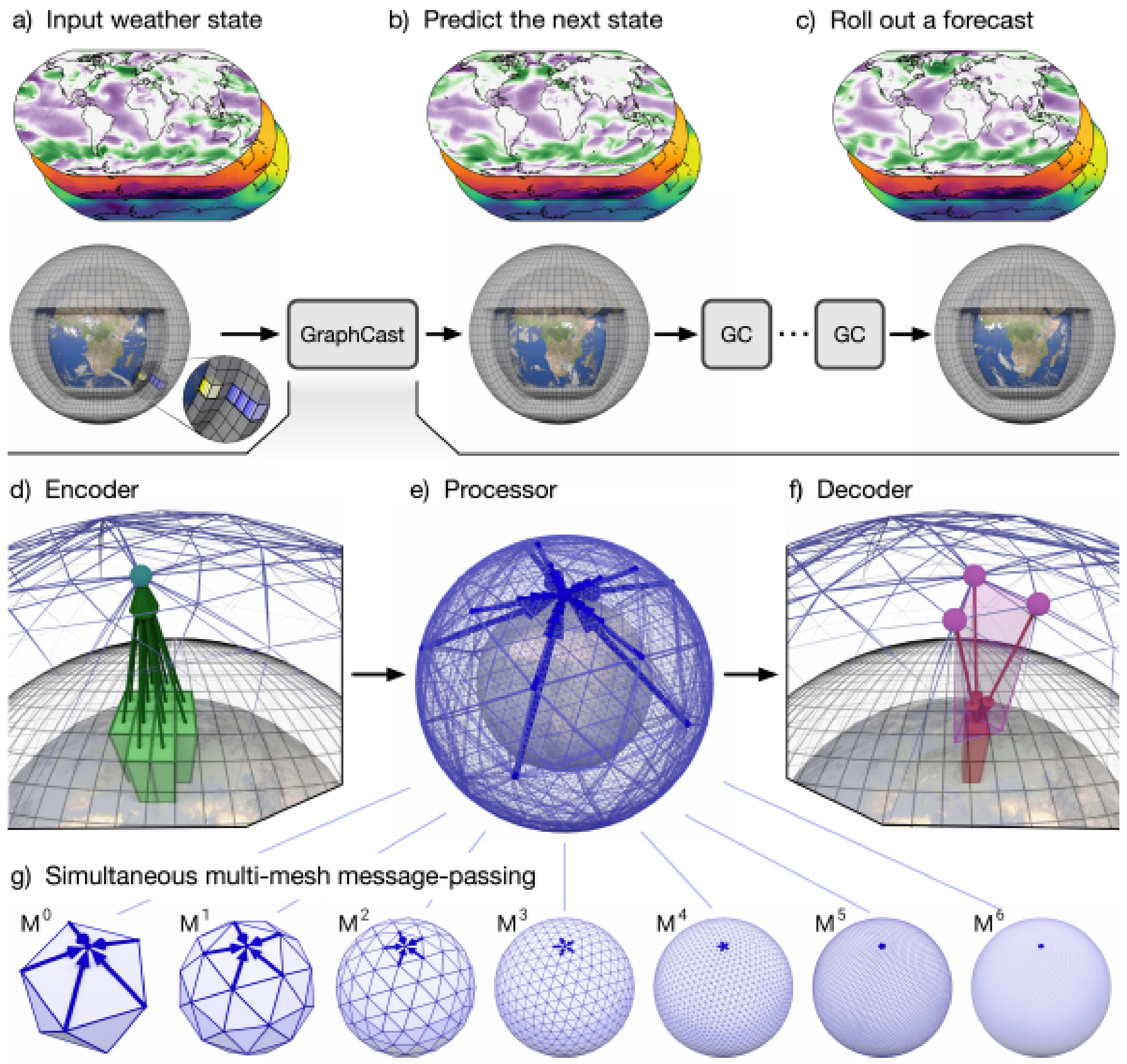
- Transition to foundation models of the Earth system.


Forward SDE (data → noise)
$x_{1:L}(0)$
$\mathrm{d}x(t) = f(t)x(t)\mathrm{d}t + g(t)\mathrm{d}w(t)$
$x_{1:L}(t)$
$L \gg k$
ϕ
t
$x_{i-1:i+1}(0)$
$\nabla_{x_{1:L}(t)} \log p(x_{1:L}(t))$
$x_{i-k:i+k}(t)$
$x_{1:L}(0)$
$\mathrm{d}x(t) = \left[f(t)x(t) - g(t)^2 \nabla_{x(t)} \log p(x(t))\right]\mathrm{d}t + g(t)\mathrm{d}w(t)$
$x_{1:L}(t)$
Reverse SDE (noise → data)

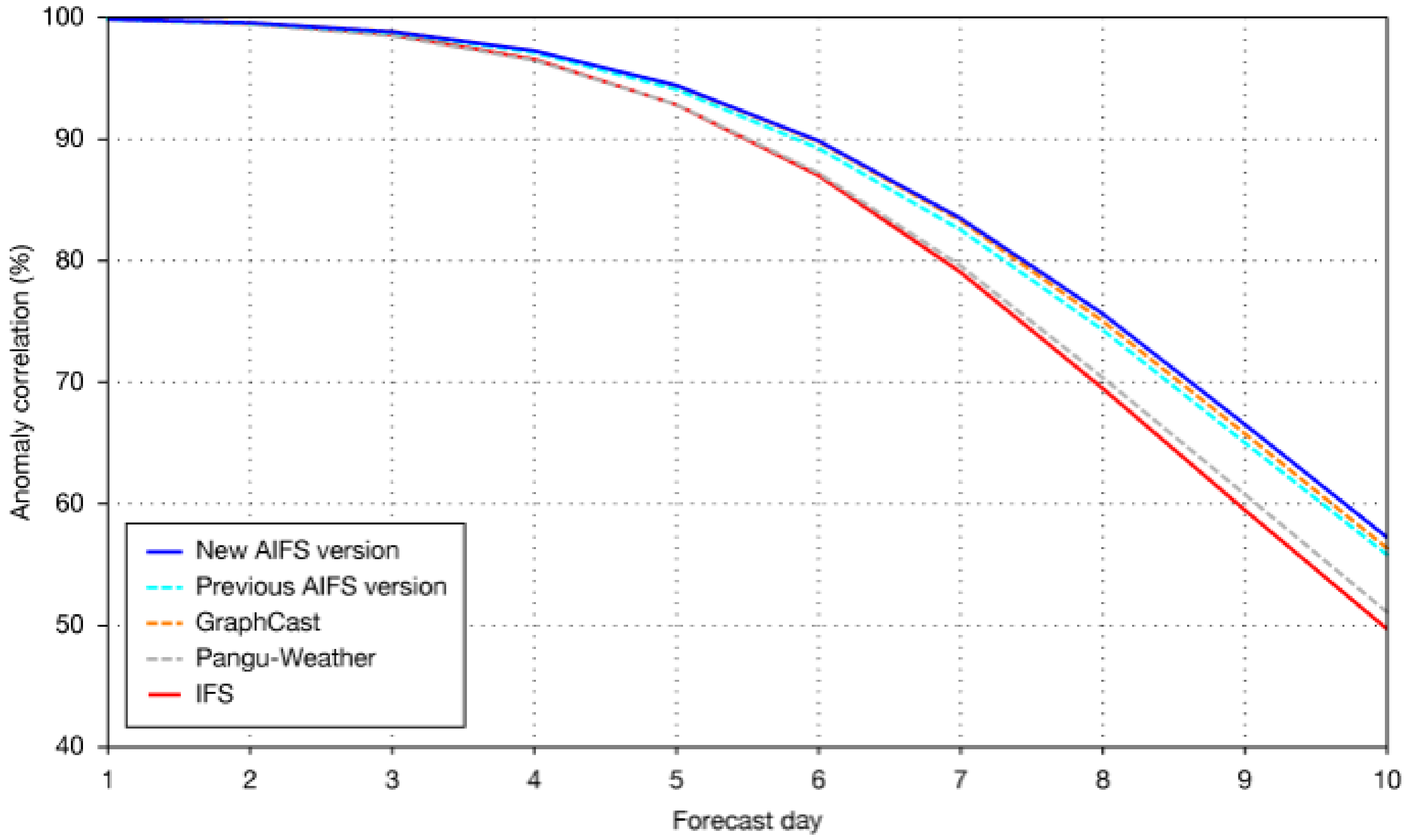
Rozet and Louppe, "Score-based data assimilation", 2023.

Funded by the European Union

Challenge 2: ML and hybrid-based weather forecasting – **State-of-the-Art**



Doi: 10.1126/science.adi2336



The chart shows the anomaly correlation for 500 hPa geopotential in the northern hemisphere. The latest version of the AIFS has a grid spacing of 28 km, while the previous AIFS version had a grid spacing of 111 km. The graph shows that the new version performs better than other machine-learning forecasts (GraphCast and Pangu-Weather) and ECMWF's IFS.

ML-based forecasting has become as skillful as NWP.

Funded by
the European Union

Challenge 2: ML and hybrid-based weather forecasting – **Research perspectives**

- **Short-term:**
  - Extend and promote design benchmarks and validation platforms to encourage the scientific community to innovate.
  - Provide easy access to real observations, in addition to reanalysis data.
  - Continue exploring AI-based forecast systems. Focus on the proper account of uncertainties and the robustness to distribution shifts.

- **Mid-term:**
  - Investigate hybrid AI-physics models to improve the interpretability, generalization and extrapolation capabilities compared to pure AI-based forecasting systems.

- **Long-term (speculative):**
  - Move to a complete digital twin in the form of a deep network combining physics with neural networks. This digital twin should resolve various scales and phenomena (weather, ocean, chemistry, etc).
  - Furthermore, we propose to explore causality in deep neural networks to get easier interpretability for pure or hybrid data-driven solutions.

Funded by
the European Union

ECMWF

ETP 4 HPC

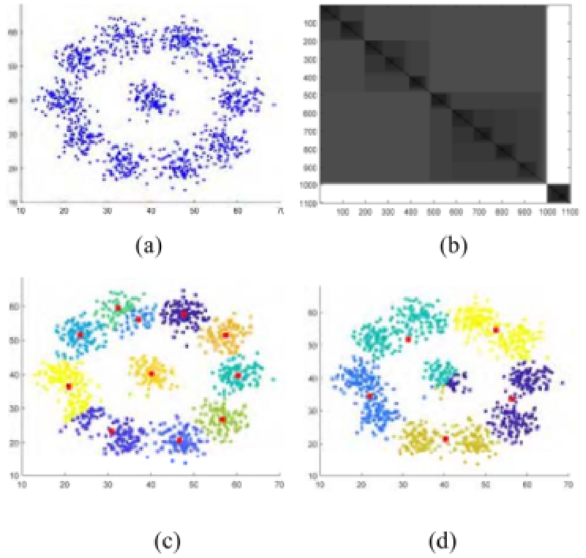
Challenge 3: ML for streaming data – **State-of-the-Art**

Streaming data: continuously arriving data to memory/storage of computers from observations, simulation models, etc.

Due to operational constraints, streaming data is requested to be processed only within a short time window, often at one-time processing.

→Dynamic/incremental training for streaming data, which is much less explored than in batch mode.

Illustration: data clustering in batch mode (left) and streamed mode (right)



$$DB(U_{11}) = 0.57 < DB(U_5) = 0.77 \quad \Rightarrow U_{11}$$
$$RI\ (U_{11}) = 0.96 > RI\ (U_5) = 0.87 \quad \Rightarrow U_{11}$$

(e)

Fig. 2. Three Components of Batch Cluster Analysis. (a) Data Set X. (b) iVAT image of X: SCA1. (c) $U_{11} (c = 11)$: *SCA*2. (d) $U_5 (c = 5)$: *SCA*2. (e) SCA3: Both CVIs prefer U_{11}.

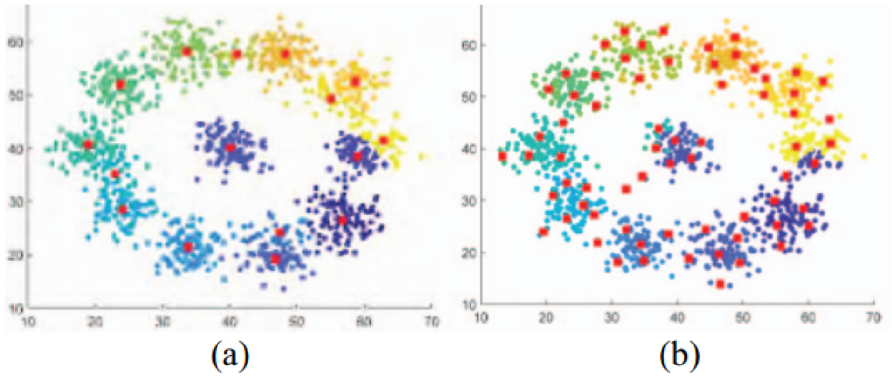


Fig. 7. PrS applied to the data set X. (a) PrS with $T = 8.57$. (b) PrS with $T = 4.29$.

Streaming data analysis: Clustering or classification?

JC Bezdek, JM Keller - IEEE transactions on systems, man, and cybernetics …, 2020

Funded by the European Union

Challenge 3: ML for streaming data – **Research perspectives**

- **Short-term:**
  - New methodologies for **multiscale** problems. Benchmark against “classical” models such as Markov Switching auto-regressive models for the large scale and ML models (such as a k-nearest neighbor) for the fine scale models.
  - Methods for **highly imbalanced data** should be developed.
  - Concept drift localization methodologies have to be investigated and introduced to digital twin solutions. Avoid retraining and allow efficient corrections of previous decisions.

- **Mid-term:**
  - Investigate **multi-label classification** of streaming data where labels are **highly imbalanced**.
  - Try new methodologies, e.g. transform the data into spectrogram images and process images with Deep Learning.
  - Enforce learning methods to **fulfill the physical constraints**.

- **Long-term (speculative):**
  - Develop efficient ML algorithms dedicated to **novel HW architectures**, e.g.,
    - quantum computers that may provide **continuous learning** from data streams,
    - neuromorphic chips (for very complex situations in a fraction of energy and a reduced amount of time)

Funded by
the European Union

Conclusions

- ML may change the cost-effective operational simulation methodologies for DestinE.
- For reliable modeling, and also for predicting rare and uncertain events, hybrid models (e.g. physics-informed NNs) should be investigated.
- ML for multiscale, imbalanced, uncertain streaming data gives many challenges for creating and implementing appropriate methods for real-time digital twins.

Funded by
the European Union

Thank you!

Funded by
the European Union

Funded by the European Union
Destination Earth
implemented by ECMWF esa EUMETSAT

Destination Earth

Destination Earth: Challenges when implementing Digital Twins on EuroHPC systems

20/03/2024

ECMWF
ETP 4 HPC

Setup of DE_380 project

TC I
TransContinuum Initiative

- Scope: Produce a **Technology Agenda for DestinE:**
  - A series of technical white papers addressing specific DestinE implementations challenges on EuroHPC systems
  - Authors come from different European associations and projects collaborating in the " Transcontinuum Initiative"
- **Domains addressed:**
  - **Federation of compute and data resources**
  - **Data streaming**
  - IOT and networking
  - **Cyber security**
  - **Mathematical methods and algorithms**

ETP 4 HPC EUROPEAN TECHNOLOGY PLATFORM FOR HIGH PERFORMANCE COMPUTING

BDV BIG DATA VALUE ASSOCIATION

Alliance for IoT and Edge Computing Innovation

6G SNS IA

EU MATHS IN

ECS EUROPEAN CYBER SECURITY ORGANISATION

HiPEAC

Funded by the European Union

ECMWF

ETP 4 HPC

Agenda

16:30	Welcome
16:35	**Destination Earth – Concept and digital twins implementation**
16:50	**Federation of Compute and Data Resources**
17:05	**Data Streaming**
17:20	**Cyber Security**
17:35	**Mathematical Methods and Algorithms**
17:50	**Q&A**

Funded by the European Union

Context/Disclaimer

- The following presentations are based on work identifying challenges in various IT domains in the context of running DestinE twins on EuroHPC compute and data infrastructures.
- The presentations are not a summary of a report commissioned by ECMWF in the DE_380 contract.
- We believe the points discussed here are more widely applicable than DestinE, but DestinE is the first project highlighting them.

- *The documents have been produced in the context of the Destination Earth Initiative and relate to tasks entrusted by the European Union to the European Centre of Medium-Range Weather Forecasts implementing part of this initiative.*
- *The documents are funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them. The European Centre of Medium-Range Weather Forecasts is not liable in respect of the documents and gives no warranty for the information needed.*

Funded by
the European Union