Modeling and Execution of Data-Aware Choreographies

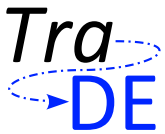
University of Stuttgart
Universitätsstr. 38
70569 Stuttgart
Germany

Michael Hahn[1], Uwe Breitenbücher[1],
Oliver Kopp[2] and Frank Leymann[1]

[1]Institute of Architecture of Application Systems
{hahnml, breitenbuecher, leymann}@iaas.uni-stuttgart.de

[2]Institute for Parallel and Distributed Systems
kopp@ipvs.uni-stuttgart.de

Phone +49-711-685 88416
Fax +49-711-685 88472



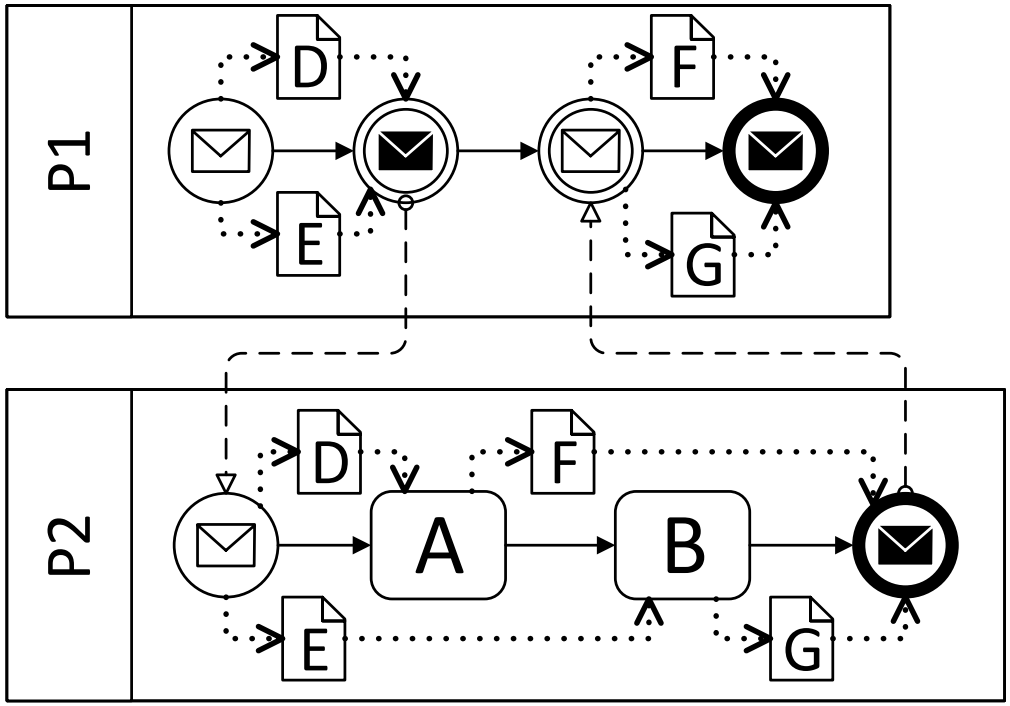
Motivation

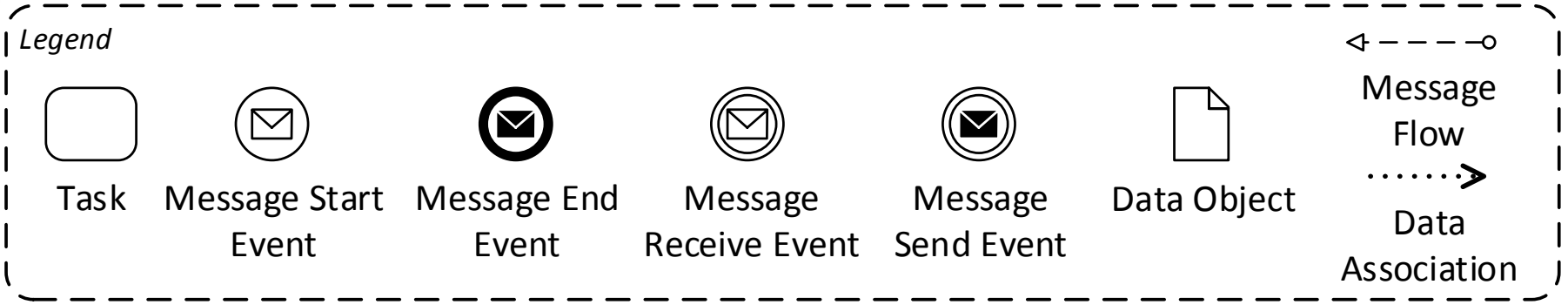
- Importance and value of data is increasing
- Service **choreographies** allow to specify **complex conversations** between multiple **interacting parties** from a **global perspective**
- … but do not fully reflect the paradigm shift towards data-awareness at the moment

Outline

- Shortcomings
- TraDE Approach
- Case Study
- System Architecture
- Conclusions and Outlook

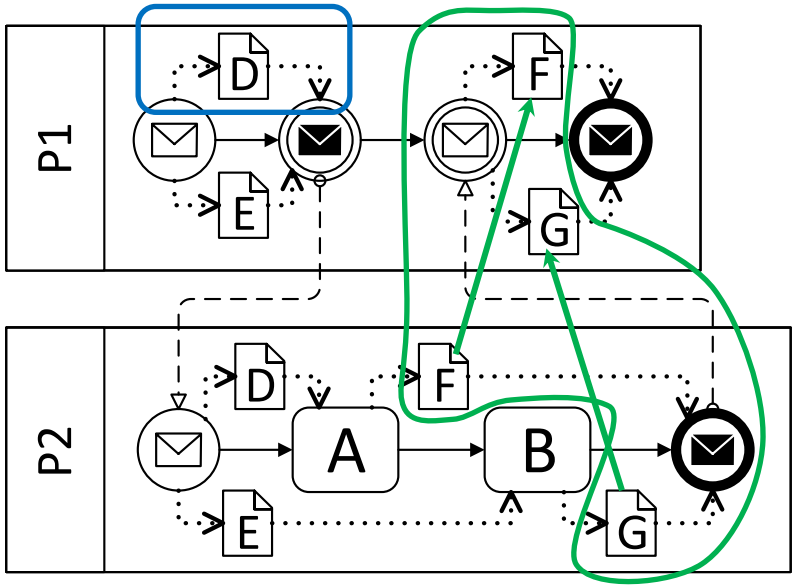
Motivation Example





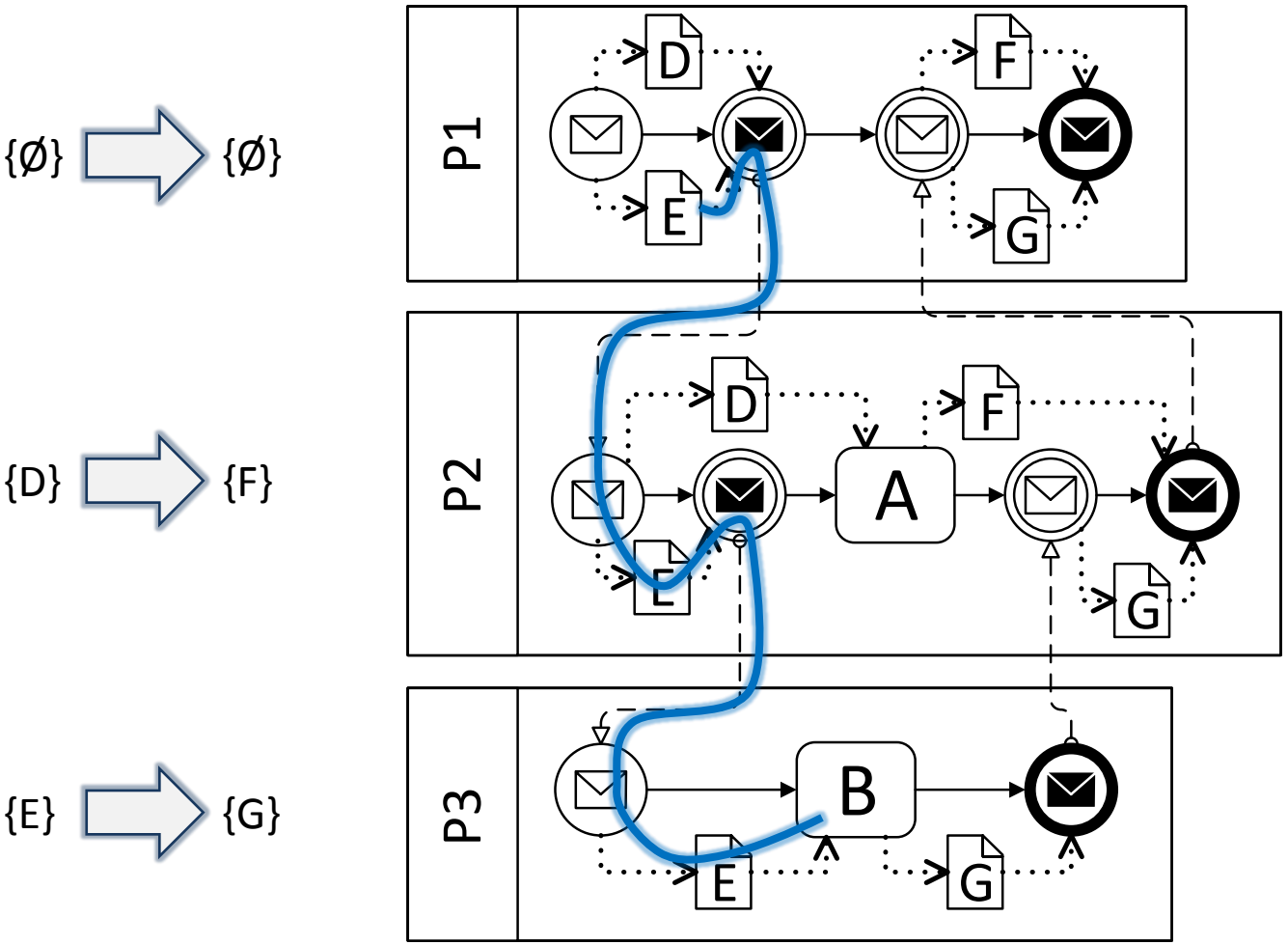
Shortcomings

- *Intra-participant* vs. *inter-participant* data flow
- Same data objects need to be specified in multiple participants
- Specification of a common, globally consolidated and agreed set of data objects (data contract) not supported in general



Shortcomings

- Potentially unnecessary routing of data

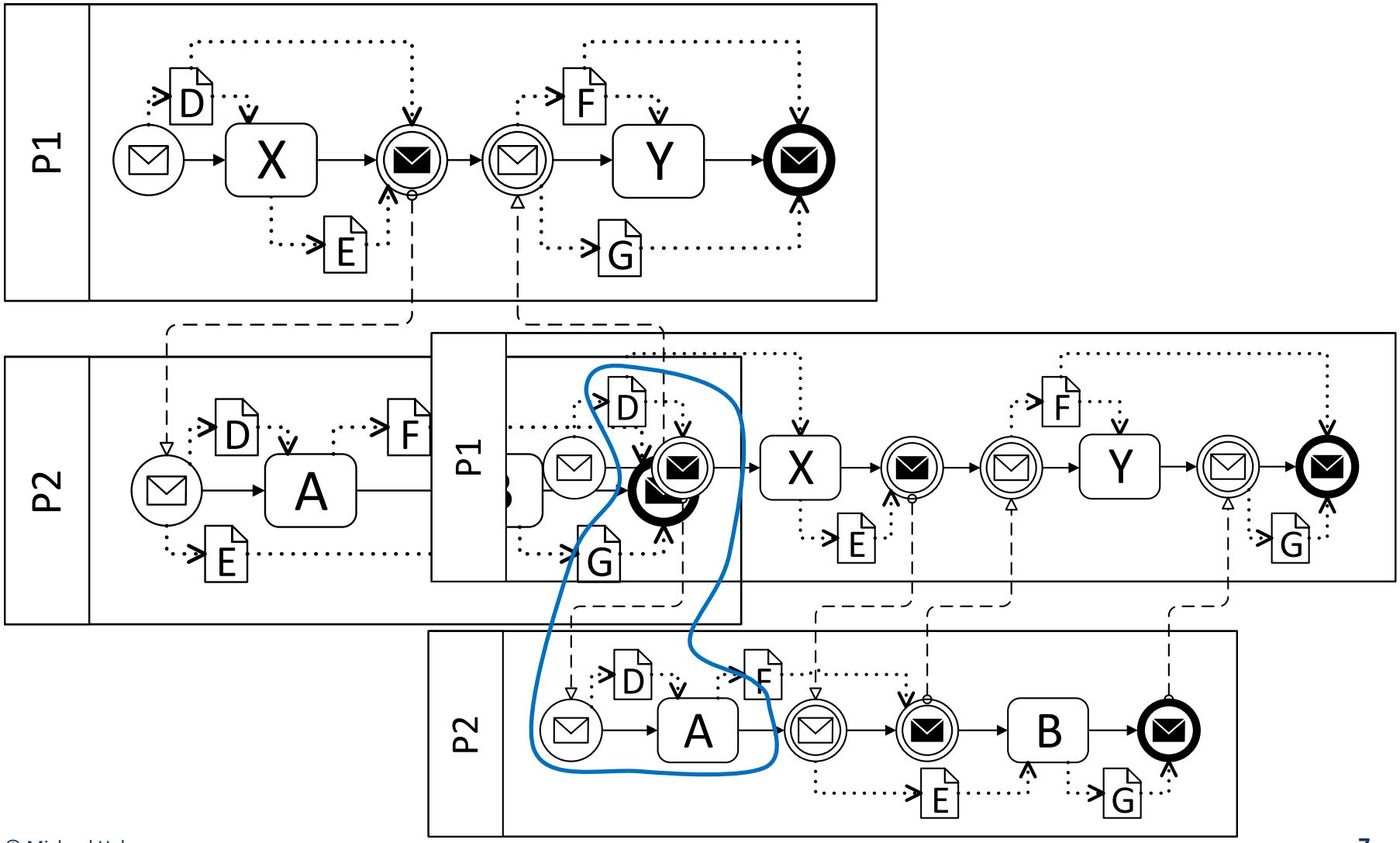


Legend

{requires} {produces}

Shortcomings

- … and blocking of control flow

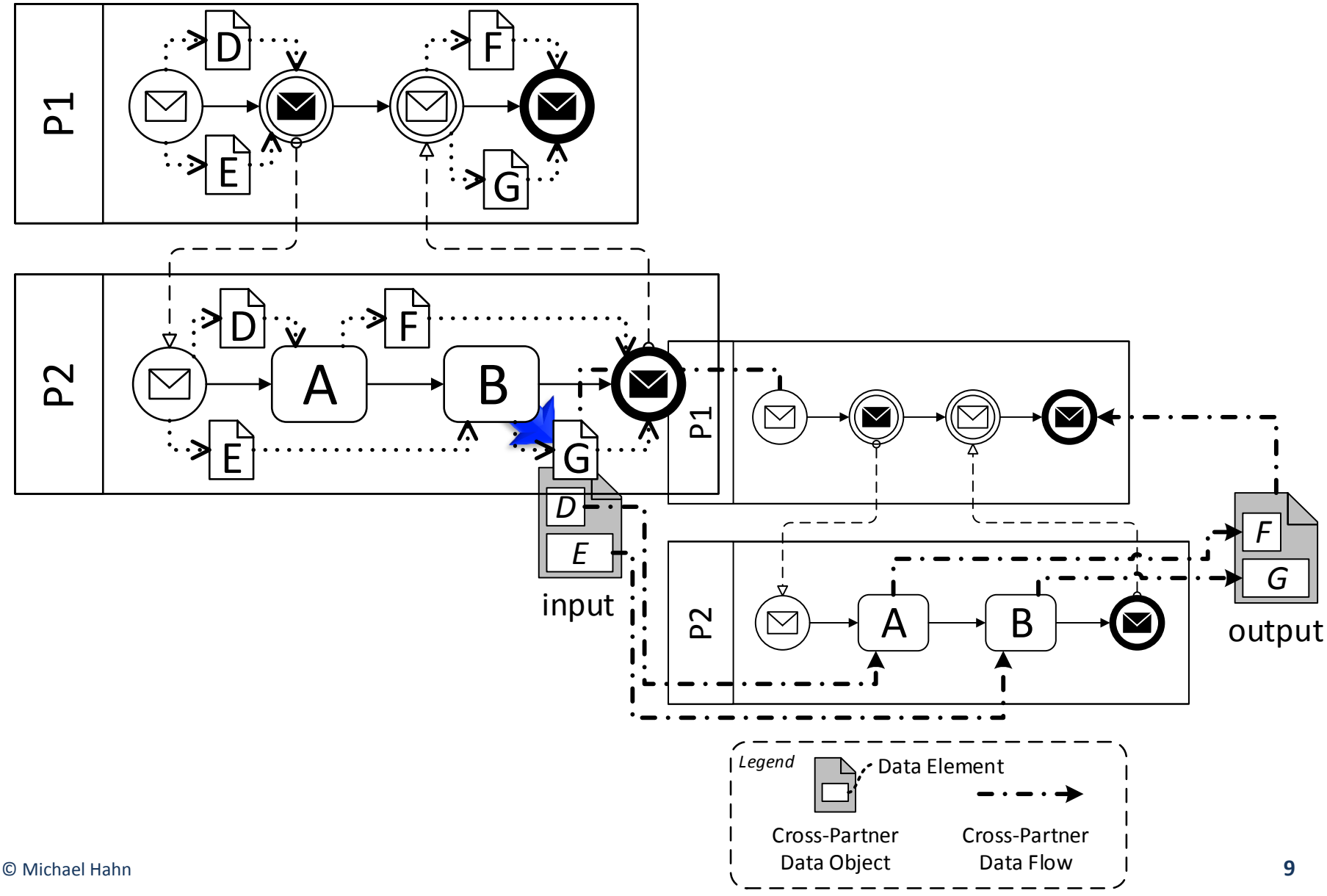


Shortcomings: Key Points to Take Away

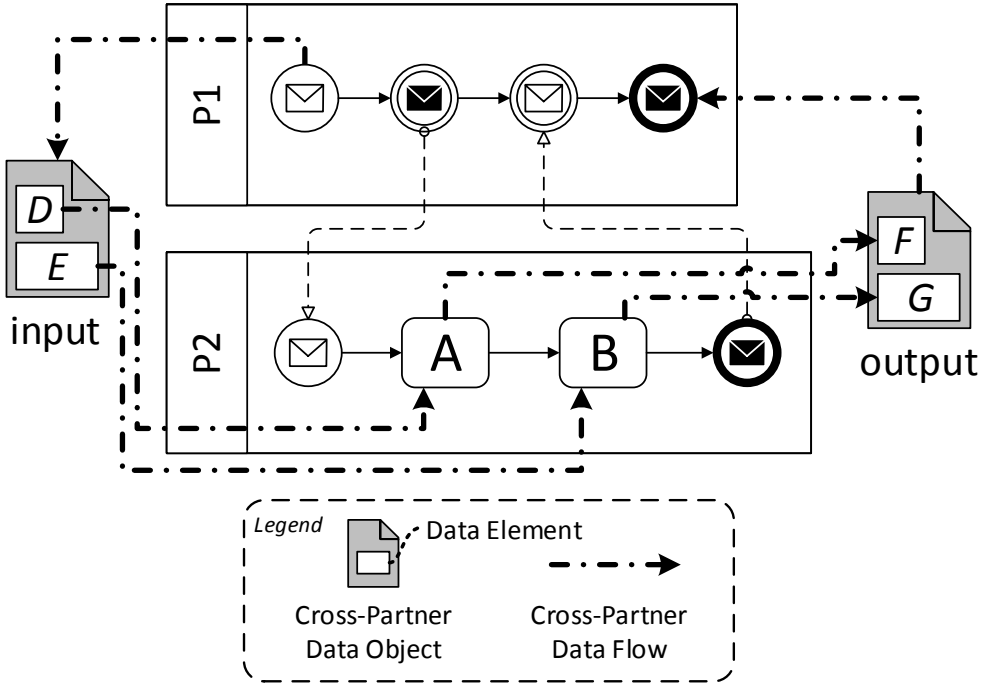
- Exchange of data during run time has to be specified completely upfront at modeling time
- Models become polluted with data management logic

➢ Results in more complex and rigid choreography models, making them also less flexible regarding their data perspective during run time

Vision: Reduce Complexity and Increase Flexibility

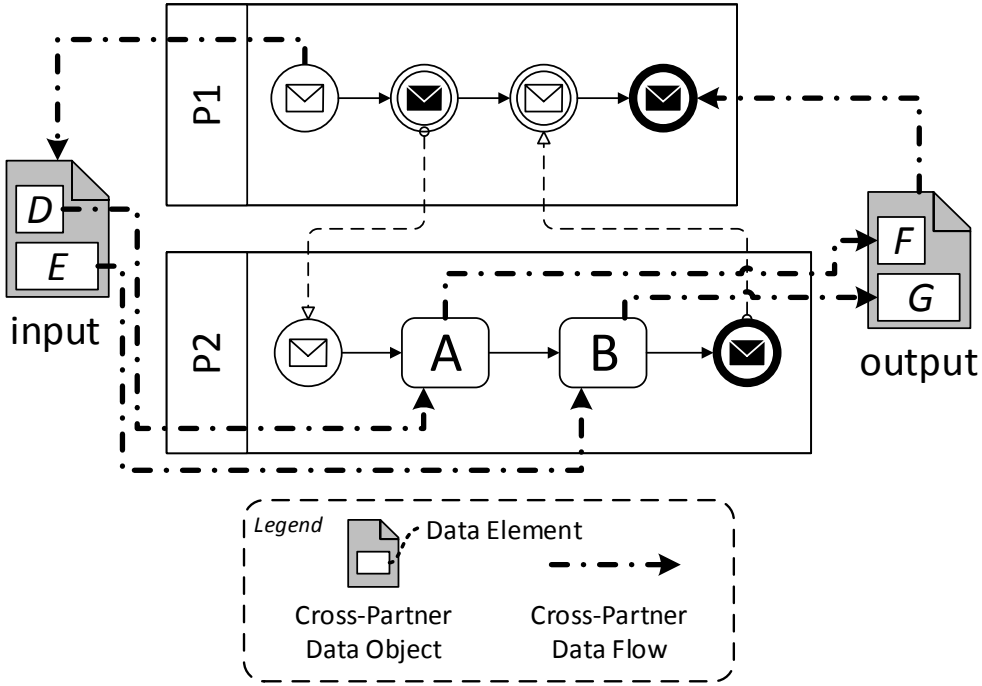


TraDE Approach: Modeling Data-aware Choreographies



- Collection of cross-partner data objects represents a choreography data model
- Choreography data model (CDM) allows to specify data in a self-contained and centralized manner (data contract)
- Cross-partner data objects
  - Express commonly agreed data structures
  - Shared between and accessible from all participants

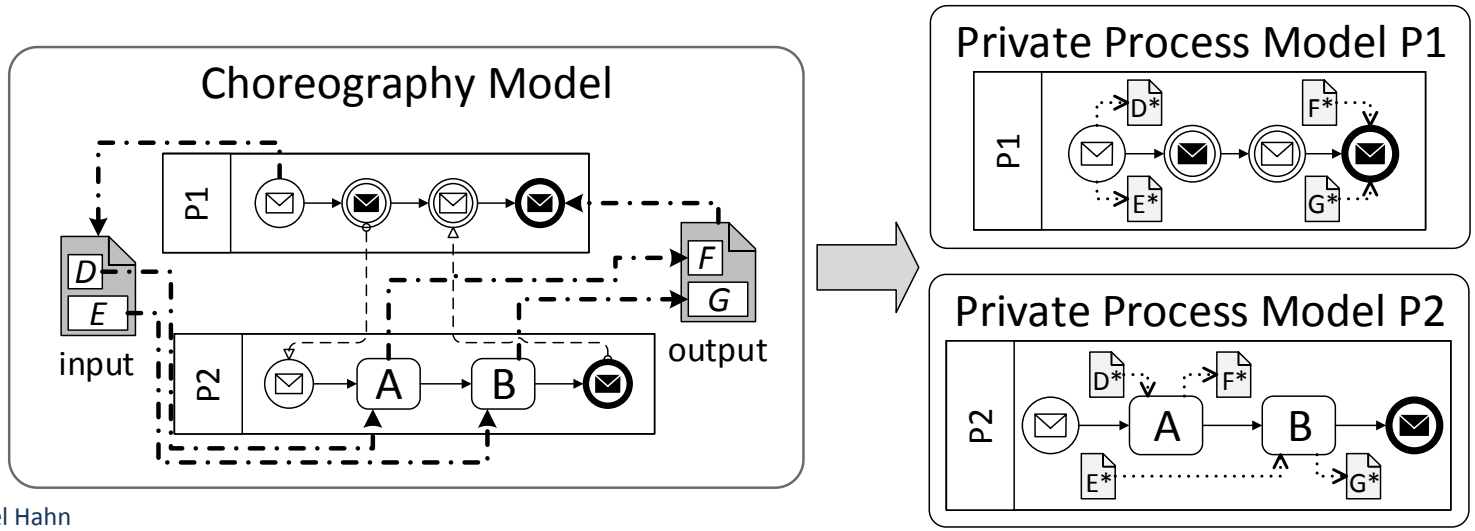
TraDE Approach: Modeling Data-aware Choreographies



- Data dependencies of choreography and its participants become (explicitly) visible
- Exchange of data across participants can be expressed more intuitive using cross-partner data flow
- Allows decoupling the exchange of data from the exchange of messages

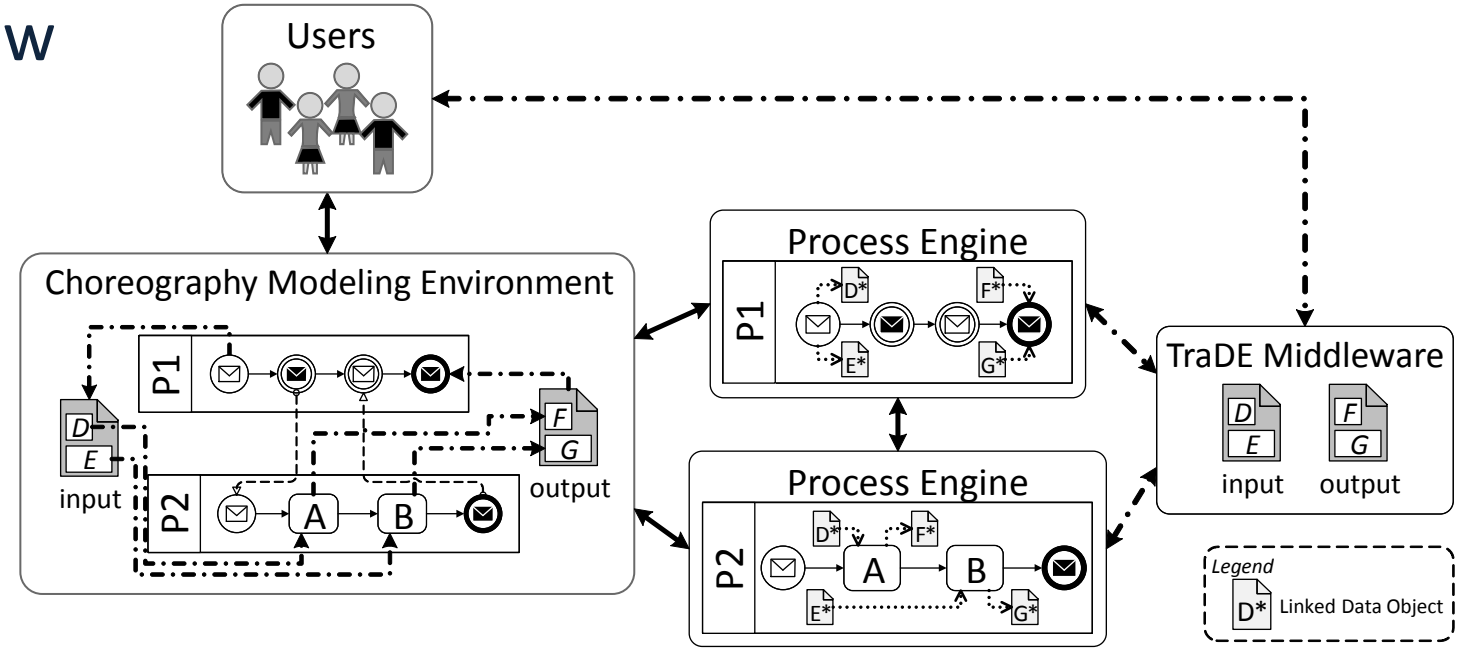
TraDE Approach: Refinement to Executable Processes

- Translate introduced cross-partner data objects into standards-based data containers in private process models (e.g., BPMN data objects or BPEL variables)
- CDM can be used as input for further automation and to reduce manual refinement efforts
- Modelers should not need to distinguish between local and globally shared data containers



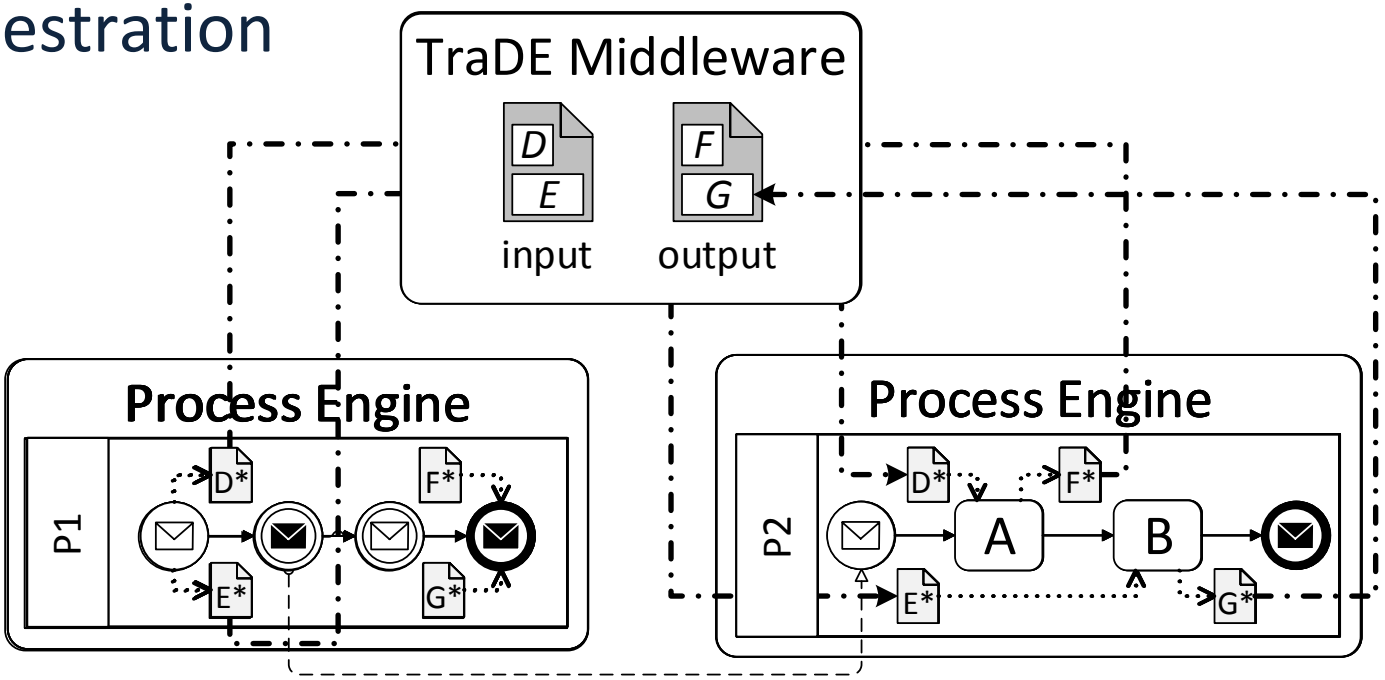
TraDE Approach: Refinement to Executable Processes

- Linking of a data container to a cross-partner data object by extension of underlying modeling constructs (using language extensibility features)
- TraDE extensions can then be parsed by process engine in order to communicate with the TraDE middleware to conduct modeled cross-partner data flow



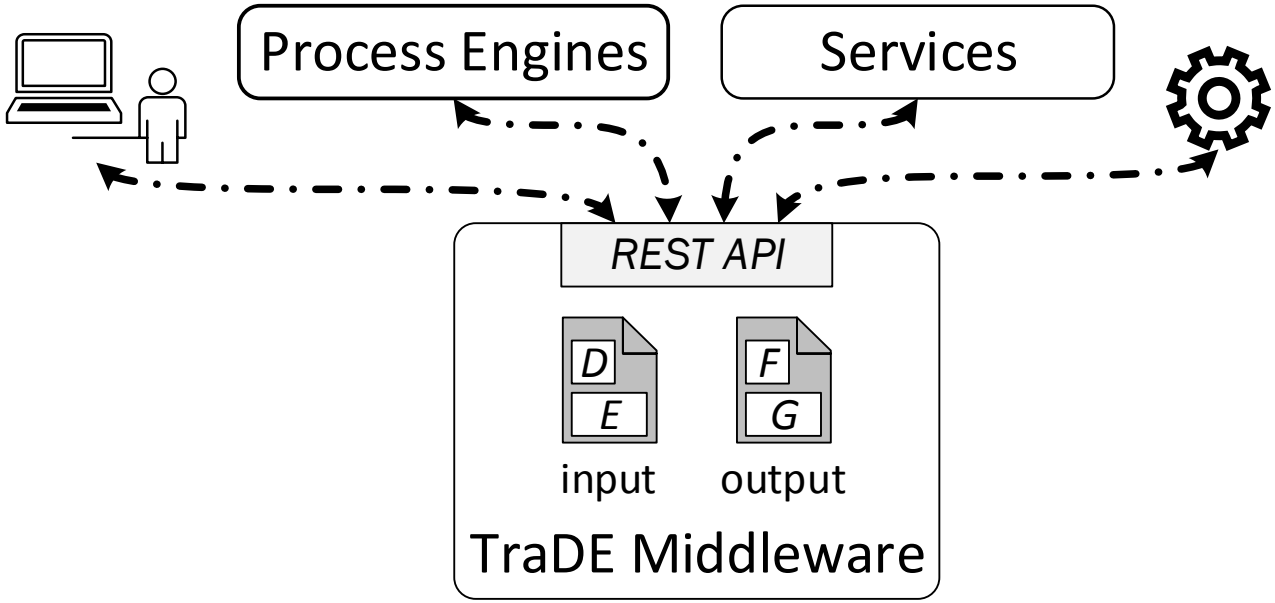
TraDE Approach: Middleware

- TraDE middleware acts as a data hub between process engines and users, services and other systems
- Goal is to enable an easier management, exchange, and provisioning of shared data independent of its processing within a service choreography or orchestration



TraDE Approach: Middleware

- Cross-partner data objects are exposed in a web-accessible manner through a REST API
- Data becomes easier accessible to other tools and systems for further processing, transformation or visualization
- eScience: Enables scientists to upload and provide simulation input data, inspect and observe intermediary results and reuse of existing data

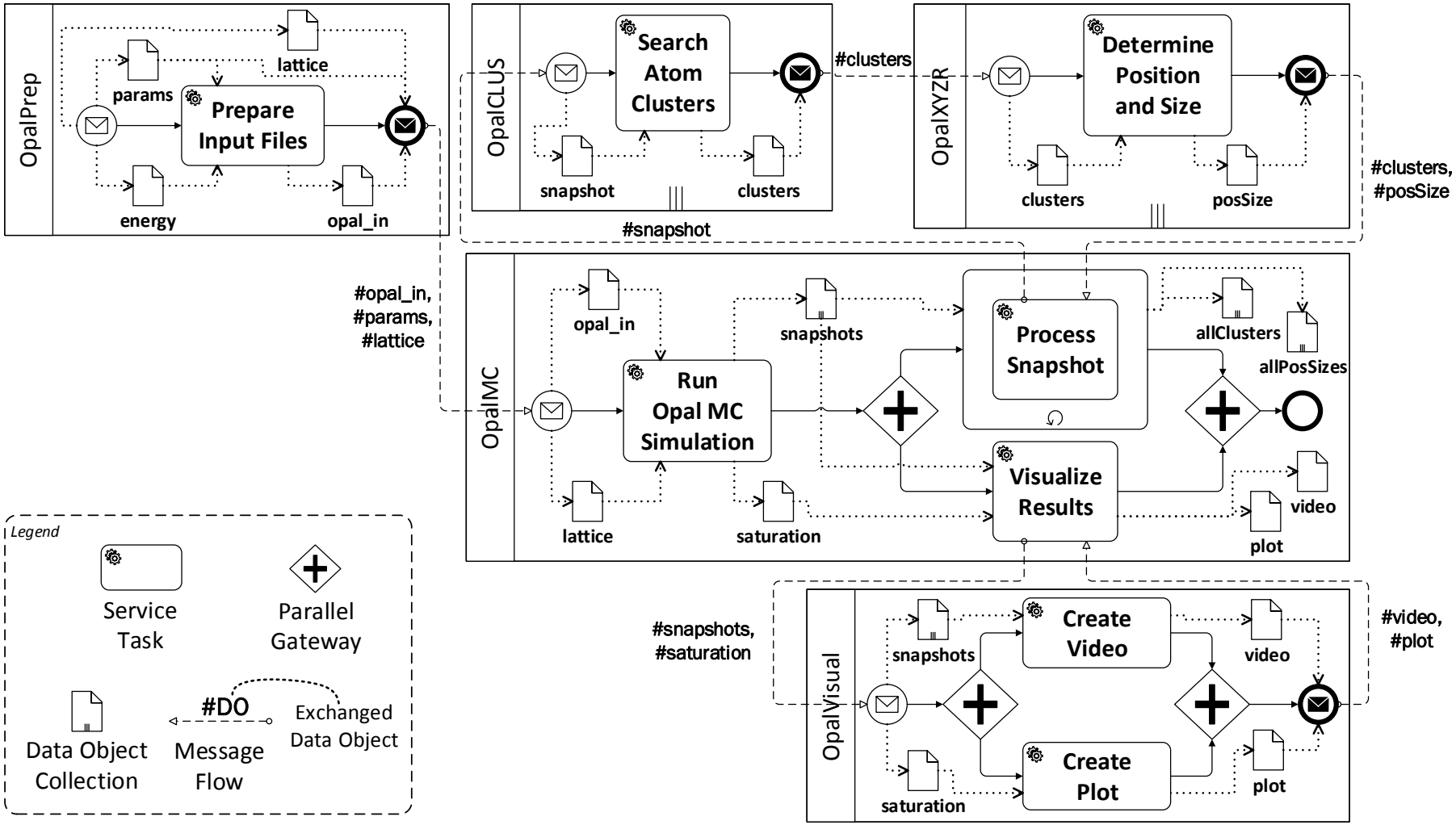


TraDE Approach: Negative Side Effects

- Sharing of data across participants (and with external clients) might lead to new issues regarding aspects such as concurrency, security, etc.
- Probability for concurrent access of shared data from different, potentially not synchronized participants, is much higher than in classical scenarios
- Modelers have to pay attention when specifying cross-partner data flow

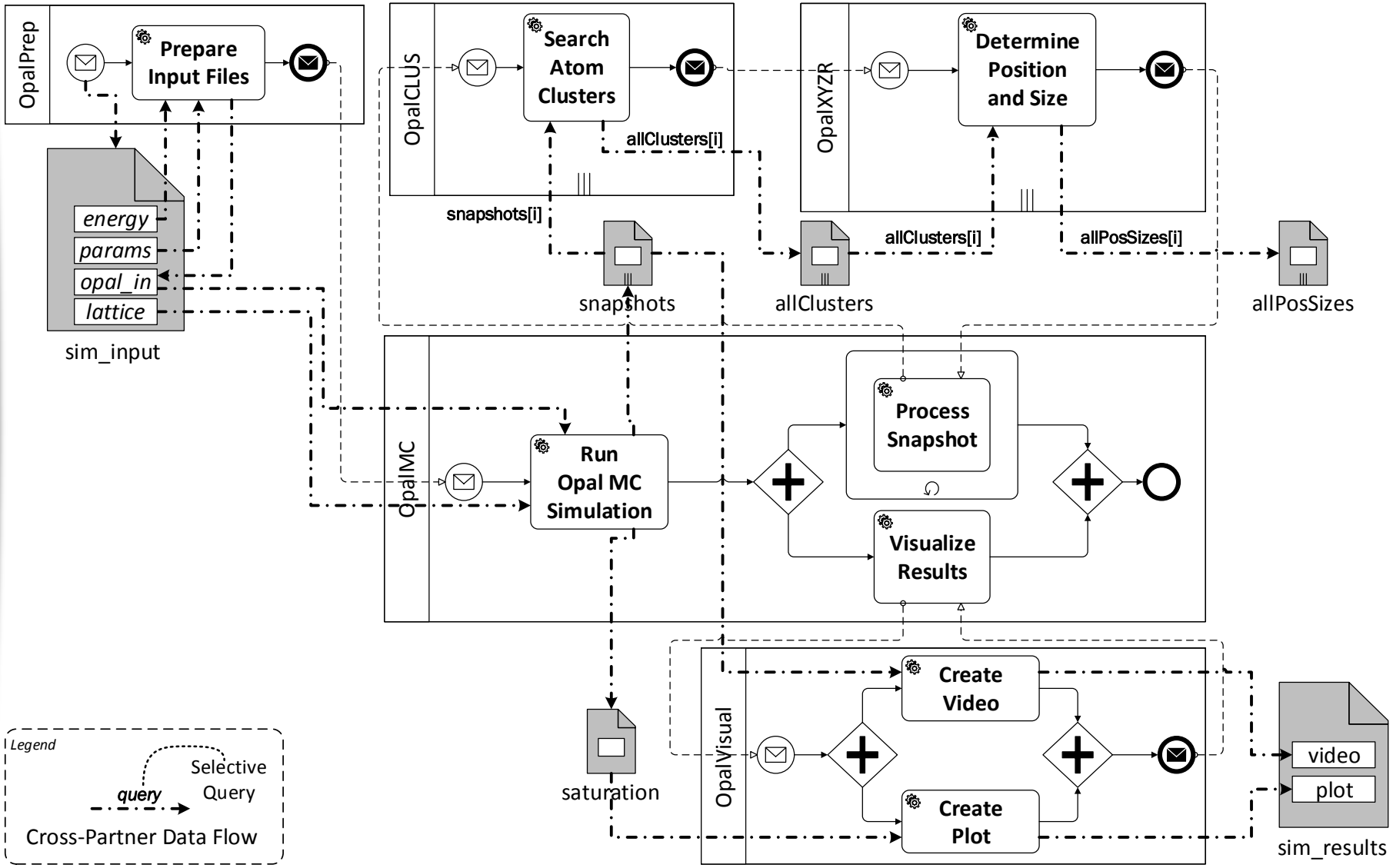
- Future work on identification and analysis of such issues to apply corresponding mechanisms based on state of the art

Case Study

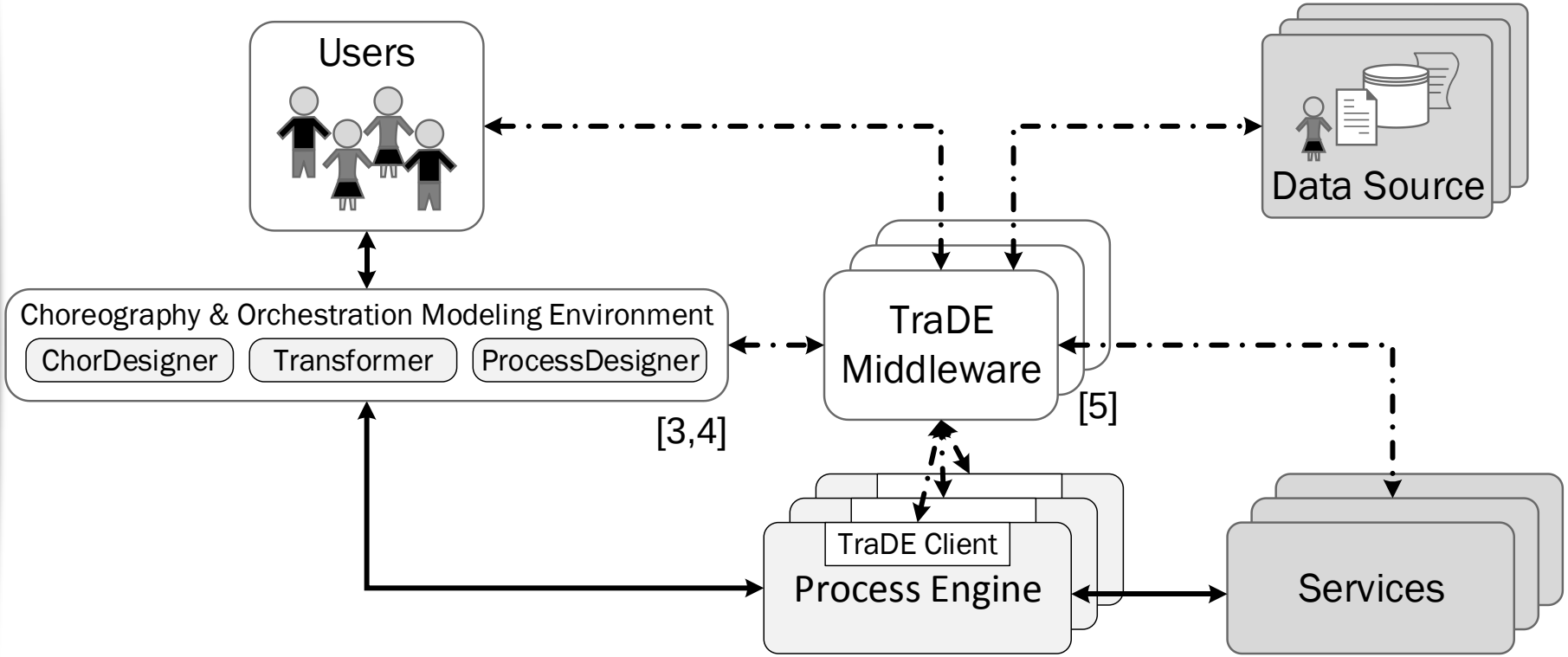


[1] Binkele and Schmauder: An atomistic Monte Carlo simulation for precipitation in a binary system. In: International Journal for Materials Research, 94, pp. 1-6, 2003.
[2] Sonntag et al.: Using Services and Service Compositions to Enable the Distributed Execution of Legacy Simulation Applications. In: Proc. of ServiceWave'11.

Case Study


OpalPrep
Prepare Input Files
energy
params
opal_in
lattice
sim_input
OpalCLUS
Search Atom Clusters
allClusters[i]
snapshots[i]
snapshots
OpalXYZR
Determine Position and Size
allClusters[i]
allPosSizes[i]
allClusters
allPosSizes
OpalMC
Run Opal MC Simulation
Process Snapshot
Visualize Results
saturation
OpalVisual
Create Video
Create Plot
video
plot
sim_results
Legend
Selective Query
query
Cross-Partner Data Flow

System Architecture



[3] Weiß et al.: Modeling Choreographies using the BPEL4Chor Designer: an Evaluation Based on Case Studies, Technischer Bericht Nr. 2013/03, 2013.
[4] Sonntag and Karastoyanova: Model-as-you-go: An Approach for an Advanced Infrastructure for Scientific Workflows. In: Journal of Grid Computing. Vol. 11(3), Springer Verlag, 2013.
[5] Hahn et al.: A Management Life Cycle for Data-Aware Service Choreographies. In: Proc. of ICWS, 2016.

Conclusions and Outlook

- Shortcomings of choreograph modeling languages regarding their data modeling capabilities
- TraDE approach: modeling, refinement & middleware
- Case Study & System Architecture

- Support distributed, multi-node deployments of TraDE middleware
- Analysis of potential concurrency issues & application of synchronization and scheduling mechanisms
- Formal framework for data-aware choreographies & algorithms to enable their transformation into standards-based models and vice versa

Michael Hahn
michael.hahn@iaas.uni-stuttgart.de
http://www.iaas.uni-stuttgart.de