18th Symposium and Summer School On Service-Oriented Computing

(June 24th - 29th, 2024) Computing Continuum – Session Chair: Schahram Dustdar

Distributed Intelligence in the Computing Continuum using Active Inference

Prof. Schahram Dustdar & Dr. Praveen Kumar Donta *Distributed Systems Group,* TU Wien, Vienna 1040, Austria. (26th June 2024)



Summer^{_}

Service Oriented Computing



Outline

- Distributed Systems and State-of-the-Art
- Distributed Intelligence
- Human Eco-systems vs Distributed Computing Continuum Systems
- Computing Continuum Monitoring and Self-healing
- Data Protocols for Computing Continuum
- Clustered Edge Intelligence through Active Inference

Evolution in Distributed Systems



Current State

- Distributed Systems are key to our society
- Underlie our critical infrastructures and applications (Smart cities, Healthcare, Autonomous vehicles,...)
- Interconnectedness (fabric) of components (HW, SW, People) induces complexity
- We increasingly see fundamental issues we need to address



Infrastructure

Computing continuum



- > Application performance highly dependent on the underlying infrastructure
 - Heterogeneity of resources & heterogeneous distribution
 - Resources diverse interconnection



Donta, P.K., et al. "Exploring the potential of distributed computing continuum systems." Computers 12.10 (2023): 198.

Distributed Computing Continuum Systems

Benefits

- Bandwidth Optimization
- Scalability & Availability
- Low latency
- Resource optimization & Load balancing
- Reliability or Flexibility

Challenges

- Interoperability
- Data Synchronization
- Governing resources
- Privacy

Distributed Intelligence in Computing Continuum

Casamayor Pujol, V., Donta, P. K., Morichetta, A., Murturi, I., & Dustdar, S. (2022, November). Distributed computing continuum systems–opportunities and research challenges. In International Conference on Service-Oriented Computing (pp. 405-407). Cham: Springer Nature Switzerland.

Traditional Approaches

Techniques used in Previous and Majority of Ongoing works



Resource constraint

Human Eco-system vs. Computing Continuum

The human body is comprised of a series of complex systems, including:

- Skeletal System
- ••• Nervous System
- · Cardiovascular System
 - Lymphatic System
- Endocrine System
 - Brain
 - Spinal Cord
 - Cranial Nerves
 - Spinal Nerves

- Oxygen
- White Blood Cells
- Hormones

Nutrients

Helping the body meet the demands (40k neurons)

Control Internal Environment, Memory and Learning (86 billion neurons)

Infrastructure Systems

Regulation Systems

Alshami, A. M. (2019). Pain: is it all in the brain or the heart?. Current Pain and Headache Reports, 23, 1-4

Human Ecosystem

Pujol, V., Donta, P.K., et. Al., 2022, November. Distributed computing continuum systems-opportunities and research challenges. In International Conference on Service-Oriented Computing (pp. 405-407). Cham: Springer Nature...

Human Eco-system vs. Computing Continuum

Infrastructure Systems

Regulation Systems

The human body is comprised of a series of complex systems, including:

- → Skeletal System
- · Nervous System
- · ► Cardiovascular System
 - Lymphatic System
- Endocrine System

Human Ecosystem



Human Ecosystem



Human Ecosystem vs. Distributed Continuum Systems

Human Ecosystem

Feedback Loops and adaptability

Communication and coordination with other parts

Decision-making Consciousness

Self-healing mechanism

Distributed Computing Continuum Systems

QoE & QoS Handle unpre

Handle *unpredictable* issues

Synchronization and Coordination

Design intelligent data protocols²

Efficient and Adaptable Resource management Explainability through Causal relations

Monitoring and self-healing devices¹

¹P.K., Donta, et al. "Governance and sustainability of distributed continuum systems: A big data approach." *Journal of Big Data* 10.1 (2023): 53. ²P.K., Donta and S. Dustdar "Intelligent Data protocols at the Edge", IEEE SERVICES 2023.

Computing Continuum Monitoring, Diagnosis and Cure



Self-healing Agent



Data Protocols and Intelligence

- P. K. Donta and S. Dustdar (2023) Towards Intelligent Data Protocols for the Edge, IEEE intelligent EDGE 2023 (IEEE SERVICES 2023)
- Donta, P. K., Srirama, S. N., Amgoth, T., & Annavarapu, C. S. R. (2022). Survey on recent advances in IoT application layer protocols and machine learning scope for research directions. *Digital Communications and Networks*, 8(5), 727-744.

Data Protocols

CoAP (Constrained Application Protocol)



AMQP (Advanced Message Queuing Protocol)



MQTT (Message Queue Telemetry Transport)





DDS (Data Distribution Service)

P. K. Donta and S. Dustdar (2023) Towards Intelligent Data Protocols for the Edge, IEEE intelligent EDGE 2023 (IEEE SERVICES 2023)

Deterministic Rules – Data Protocols

- Protocols follow deterministic set of rules
- Benefits:
 - Low overhead
 - Limited computing or memory resources
 - Easy to track the misconducts
 - Predictability

Limitations

- Flexibility and Adaptability
- Handling uncertainty
- Scalability
- Dynamic decision making systems

Intelligence in Data Protocols

Benefits

- Scalability
- Adaptability
- Handling uncertainty
- Dynamic decision making, such as
 - Message filtering
 - Message expiry prediction
 - Resource usage (memory and CPU)
 - Congestion control
- Reliability

Implications

- Need of additional resources (to store history and perform analysis)
- Decide where learning to be performed
- Deciding run time accuracy
- Deciding suitable AI/ML for the protocols properties

Develop Novel Intelligent Data Protocol

- We intended to develop an lightweight intelligent protocol for Computing continuum systems by mitigating the following issues:
 - Message filtering
 - Message expiry prediction
 - Resource usage (memory and CPU)
 - Congestion control
- Adaptable and
- Interoperable features



Equilibrium in Computing Continuum through Active Inference

Sedlak, B., Pujol, V. C., Donta, P. K., & Dustdar, S. (2024). Equilibrium in the Computing Continuum through Active Inference. *Future Generation Computer Systems*.



Causality

Causality is the study of how things influence one other, how causes lead to effects.



Surprises are generally created by low-probability outcomes

Image sources: https://www.btelligent.com/en/blog/127801ea523cbf313725fcb60289fbf8-4/ https://ru.pinterest.com/pin/489555421981720364/ https://imgflip.com/gif/6keaol

Free Energy

The discrepancy (or uncertainty) between the agent's understanding of the process and reality is called *Free Energy* (FE)

Lower the Free Energy (FE), higher the accuracy.



This is commonly known as **Bayesian inference** and allows agents to use existing beliefs (widely known as priors) to calculate the probability of related events.

Markov Blanket

The Markov Blanket of a random variable is the subset of nodes that provide enough information to statistically infer its value. Concept from Judea Pearl [1].

In a Bayesian Network, the Markov Blanket of a node (N) is composed of the parents (P), the children (C) and the co-parents of the children (S).



A tool for *causal* filtering.

Causal Inference

Causal inference is a technique aiming at learning *causal relationships* between *variables*, and then being able to *model and predict* the system causal behavior.

- ➤ 3 Rungs on the ladder of causation. [2]
 - Observational
 - Interventional
 - Counterfactual
- Explainability capacity



Active Inference



Figure: Action – Perception Cycle

Service Level Objective - Type

whether a system is able to fulfill its intended function at a point in time

Three types of SLOs



Service Level Objective as system requirements

Components: Service + host + SLO + elasticity strategies Tailored adaptations to service and host



DeepSLOs

- ➤ A construct we envision relating SLOs
- Provides a complete view of DCC system
- Allows aggregation towards higher abstractions



Running Example

- Reflected in most of the architecture
- Use Case

Distributed video processing architecture where IoT streams are transformed on **edge devices** to preserve individual's privacy. After privacy enforcement, distribute streams over **cloud**.

Hierarchical network structure
 IoT devices provide streams to edge devices;
 streams processed locally at edge devices;
 video stream properties are configurable



Collaborative Edge Intelligence Framework

3 major contributions in interplay:

- 1. Continuous model accuracy and local SLO fulfilment
 - a. Static BNL and Inference
 - b. Continuous BNL and Inference (AIF)



Collaborative Edge Intelligence Framework

3 major contributions in interplay:

- 1. Continuous model accuracy and local SLO fulfilment
 - a. Static BNL and Inference
 - b. Continuous BNL and Inference (AIF)
- 2. Federation and combination of models



Collaborative Edge Intelligence Framework

3 major contributions in interplay:

- 1. Continuous model accuracy and local SLO fulfilment
 - a. Static BNL and Inference
 - b. Continuous BNL and Inference (AIF)
- 2. Federation and combination of models
- 3. Collaboration between cellular structures



1a – Static BNL and Inference

- Basic mechanism for assuring SLOs at individual devices
- Requires training data in upfront and is prone to data shifts
- Evaluates possible configurations through a 3-step method



1a – Static BNL and Inference (2)

Bayesian Network Learning (BNL)



Structure Learning Hill-Climb Search (HCS) Dir. Acyclic Graph (DAG)

Parameter Learning Max. Likelihood Estimation

Conditional Prob. Table (CPT)

Markov Blanket (MB) Selection



- □ Causality filter [1,4]
- Identify variables that have an impact on SLO fulfillment

Probability of SLO violations Ideal configuration

Knowledge Extraction

- P(SLO < x) for all
 variable combinations
- Find Bayes-optimal system configuration

1b – Active Inference Cycle

- AIF agent → Equilibrium-Oriented SLO Compliance (EOSC) model
- Agent uses SLOs as **preferences** during continuous adaptation
- BN trained incrementally from incoming observations
- Beliefs updated according to prediction errors



1b – AIF Agent Behaviour/Action

Determined by three factors:

- Pragmatic value (*pv*)
 Summarizes **QoE** SLOs (e.g., resolution)
- Risk assigned (*ra*)
 Summarizes **QoS** SLOs (e.g., network limit)

pv & *ra* calculated as **separate factors** from MBs; configurations rated according to SLO fulfilment; **interpolation** between known configurations

Information gain (*ig*)
 Continued on the next slide

 $u_c = pv_c + ra_c + ig_c$



CPU

pixel

1b – AIF Agent Behaviour/Action (2)

- Information gain (*ig*)
 - Favours configurations that promise **model improvement**
 - Summarizes surprise for observations included in the **MB**
 - Hyperparameter (e) allows exploring designated areas

AIF Action-Perception cycle:

- 1. Calculate **surprise** for current batch of observations
- 2. Retrain structure (or parameters) depending on surprise
- 3. Calculate behavioural factor for **empirically evaluated** configs 4. **Interpolate** between known configurations in 2D (or 3D) space
- 5. Choose the highest-scoring (device) configuration

Agent gradually develops **understanding** how to ensure SLOs

ig after 1 round *ig* after 1 round *ig* after 1 round *ig after 1 ig after 1 <i>ig after 1 ig after 1 ig after 1 <i>ig after 1 ig after 1 ig after 1 <i>ig after 1 ig*

 $ig(c) = e + \left(\frac{\tilde{\Im}_c}{\bar{\Im}}\right) \times 100$



2 – Knowledge Exchange

Extend from single devices to the CC

Heterogeneity among the Edge

- Impedes simple transfer learning of models
- Low model accuracy \rightarrow high surprise
- Requires a **cluster leader** (fog node or edge)
- EOSC models collected at a leader node
- Model selection according to hardware char.
- Merging models to provide tailor-fit one

Fast onboarding (= horizontal scaling) of devices



3 – Collaborative Scaling

Limited action scope of devices

- Individual devices restricted to local scope to resolve SLO violations
- Leader node collecting environmental metrics (e.g., network congestion)
- Incorporated to causal model, contrasted against local SLO fulfilment (AIF)
- Emerging structures allows optimizing cluster-wide SLO fulfilment
 - E.g., redistribute clients between impacted devices



Evaluation - Overview

• Use Case

Distributed video processing architecture where streams are transformed on **edge devices** to preserve privacy of individuals.

Implementation

Prototype including video transformations and the collaborative edge intelligence framework.

• Evaluation Scope

Targeting each contribution with different aspects.





Evaluation - Use Case

BNL comprises metrics from various sources (e.g., IoT client or edge device); Extended with target conditions (i.e., SLOs) to create the **EOSC** model:

Model training takes 11 (3) metrics

Table 1: List of metrics captured by the devices, which are turned into variables by ACI

Name	Origin	Unit	Description	Param
pixel	IoT	num	number of pixel contained in a frame	Edge
fps	IoT	num	number of frames received per second	Edge
bitrate	IoT	num	number of pixels transferred per second	No
cpu	Edge	%	utilization of the device CPU	No
memory	Edge	%	utilization of the system memory	No
streams	Edge	num	number of IoT devices providing data	Fog
consumption	Edge	W	energy pulled by the device	No
network	Edge	num	data transferred over network interface	No
delay	App.	ms	processing time per video frame	No
success	App.	T/F	if a pattern (i.e., face) was detected	No
distance	App.	num	relative object distance between frames	No
slo_rate	Edge	%	combined SLO Fulfillment rate $(pv \times ra)$	No
device_type	Edge	enum	physical device type	No
congestion	Edge	num	network congestion that increases latency	No

SLOs from model variables

Table 2: Extracted SLOs and their classification.

SLO	Condition	Tier	Type
network	$network < 1.6~{\rm MB/s}$	Edge	QoS
in_time	delay < 1/fps	Edge	QoS
success	success = True	Edge	QoE
distance	distance < 50	Edge	QoE
slo_rate	$\max(slo_rate)$	Fog	Both

Parameters allow **configuring** a component's environment

Evaluation - Implementation

Python prototype for which we provide:

- <u>Github</u> repository¹
- **Docker** container²

Evaluation included a variety of edge devices:



Table 3: List of devices used for implementing and evaluating the presented methodology

bedded-systems/jetson-xavier-n

Full Device Name	ID	Price	CPU	RAM	GPU	<i>p</i> [1,4]	g [0,2]	Σ
ThinkPad X1 Gen 10	Laptop	1800 €	Intel i7-1260P (16 core)	32 GB		Very High (4)	None (0)	4
Nvidia Jetson Orin	Orin	500 €	ARM Cortex A78 (6 core)	8 GB	Volta (383 core)	High (3)	High (2)	5
Nvidia Jetson Nano	Nano	150 €	ARM Cortex A57 (4 core)	4 GB		Low (1)	None (0)	1
Nvidia Jetson Xavier	Xavier	300€	ARM Carmel v8.2 (6 core)	8 GB		Medium (2)	None (0)	2
Jetson NX GPU	NX	300 €	ARM Carmel v8.2 (6 core)	8 GB	Amp (1024 core)	Medium (2)	Low (1)	3

Devices combined within a cluster and classified relatively to each other

Evaluation - Aspects

We motivated, evaluated, and provided the results for **13** aspects:

A-1: Do MBs reduce the complexity of inference?
A-2: What is AIF's operational overhead?
A-3: How long require AIF agents to ensure SLOs?
A-4-1: Are the produced Bayesian networks interpretable?
A-4-2: Is the behaviour of AIF agents explainable?
A-5: What is the operational impact of including BNL in the AIF cycle?
A-6: Can changes in variable distribution be handled?
A-7: Can SLOs be modified during runtime?

K-1: What is the SLO fulfilment rate of transferred models? K-2: Can knowledge transfer achieve any speedup? K-3: Do tailored models have lover surprise compared to existing models?

S-1: How is the load distributed among resource-constrained devices? S-2: Can intelligent CC structures optimize local SLO fulfilment?

Evaluation - Aspects (Filtered)

We motivated, evaluated, and provided the results for 13 aspects:

A-1: Do MBs reduce the complexity of inference?
A-2: What is AIF's operational overhead?
A-3: How long require AIF agents to ensure SLOs?
A-4-1: Are the produced Bayesian networks interpretable?
A-4-2: Is the behaviour of AIF agents explainable?
A-5: What is the operational impact of including BNL in the AIF cycle?
A-6: Can changes in variable distribution be handled?
A-7: Can SLOs be modified during runtime?

K-1: What is the SLO fulfilment rate of transferred models?K-2: Can knowledge transfer achieve any speedup?K-3: Do tailored models have lover surprise compared to existing models?

S-1: How is the load distributed among resource-constrained devices? S-2: Can intelligent CC structures optimize local SLO fulfilment?

A-1: Do MBs reduce the complexity of inference?

• Setup

Modify the AIF agent to calculate behavior factors (i.e., **surprise**, etc) for a reduced number of SLOs with or without MB

Result

Applying MBs reduced the median inference time of 4 SLOs from 197ms to 151ms

Implication

MB provided a decreased system view



A-4-1: Are the produced Bayesian networks interpretable?

• Setup

Train the EOSC model from scratch and extract the BN after X rounds

• Result

Dependencies gradually revealed:

• Implication

AIF can be used to identify **causal relations** according to current and upcoming observations. Results are intuitively comprehensible.



A-4-2: Is the behavior of AIF agents explainable?

• Setup

Train the EOSC model from scratch and extract the agent's behavioral factors after X rounds

• Result

Develops clear preferences

• Implication

Allows to **empirically debug** the behavior and **fine-tune** agent by adjusting hyperparameters



K-3: Do tailored models have lover surprise compared to existing models?

• Setup

Federate EOSC models within the cluster, select and *combine* models for joining edge device; track retraining.

Result

Tailor-made model reported the lowest **surprise**, although remaining models improved through **retraining**.

Implication

Surprise can be decreased by choosing a (best-)fitting device model .



S-2: Can intelligent CC structures optimize local SLO fulfillment?

• Setup

Clients distributed equally between **comparable** devices, introducing network *congestion* for one of them; rebalance load.

Result

Cluster-wide SLO fulfillment (Σ) improved from 1.03 to 1.53.

• Implication

Was able to **raise the scope** of elasticity strategies, but requires sufficient data to model the relation of *congestion* \rightarrow *slo_rate*.





Summary

- Human Eco-systems' intelligent bring Distributed Continuum systems more intelligent and adaptable.
 - Specially Self-healing
 - Higher QoS and QoE
 - Self-adaption
- Active Inference as key method for **self-adaptation**
 - Autonomous EOSC model training and updating
 - Fulfill SLOs through **continuous** reconfiguration

Future plans

- Pending comparison with other (ML) approaches
 - Evaluation of more complex use cases
- Composition of MBs for larger structures (DeepSLOs)
 - Constrain one MB depending on another's SLOs



Thank you

Dr. Praveen Kumar Donta

Distributed Systems Group TU Wien, Vienna, Austria (pdonta@dsg.tuwien.ac.at) https://dsg.tuwien.ac.at/team/pdonta/

References

	On Causality in Distributed Continuum Systems	4082 IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL 35, NO 4, APRIL 2015	Tuture Generation Computer Systems 169 (2024) 92-108
	Victor Casamayor Pujol ¹⁰ , Boris Sedlak ¹⁰ , Praveen Kumar Donta ¹⁰ , and Schahram Dustdar ¹⁰ , Vienno University of Technology, Vienno, 1040, Austria	On Distributed Computing Continuum Systems	Contents lists available at ScienceOlrect
Systems with Markov Blaakets Active Inference on the Edge: A Design Study*		Schahram Dustdar ⁹ , Fellow, IEEE, Victor Casamaver, Puiol ⁹ and Praveon Kumar, Donta ⁹ , Member, IEEE	Future Generation Computer Systems
Bath Talkis, Yeare Country of Party Roma Rame Dates and Scholaus Dates	As distributed continuum systems (DCSs) are envisioned, they will have a massive		ELSEVIER journal homepage: www.elsevier.com/locata/tges
Brice Bricky Bricks Conserved Charles Provide Process & Second Conserved Conserve	impact on our future society. Hence, it is of utmost importance to ensure that their impact is socially responsible. Equipping these systems with causal models brings	contracting the an expression of water contracting to execute in multiple contracting to the track of the execution of company contracting the execution of the execution of the executed in multiple contracting times. Classify, Eqs and for This imple data develops manifold challenges due to the inherent complexity interrited from the underlying intractituations of these systems. This makes to execute the execution of a contract and the execution of an execution of a contracting the execution.	Equilibrium in the Computing Continuum through Active Inference
Description from Group TE Vare, Mell Talano, Inpresi In and public for description of the Vare of the	features such as explainability, accountability, and auditability, which are needed to provide the right level of trust. Furthermore, by combining causality with graph-based	impopulation in the direct control methods of the second s	Boris Sedlak *, Victor Casamayor Pujol, Praveen Kumar Donta, Schahram Dustdar
Khelmark, Visuparie Visionana, XVV ersone comprise a test many Khelmark, Visuparie Visionana, XVV ersone comprise a test many	service-level objectives, we can cope with dynamic and complex system requirements while achieving sustainable development of DCSs' capacities and applications.	equilibrium for these systems, providing a more fiscible management framework compand with the one based on thresholds, currently in use for finame based distributed systems. Finally, we also link the equilibrium with the development of adaptive mechanisms. However, we are seven that developing the endine methodology regions as big effort and the use of learning techniques, therefore, we line this is	ARTICLE INFO ABSTRACT
Ker of derive deschadung on componential time. Exclusing produces on the second deschadung of the derive of the second deschadung of the derive of the derive of the deschadung of the derive of th		article with an overview of the techniques required to develop this methodology index Terms—Distributed systems, computing continuum, edge computing, markov blenkerl	Kywerd: Comparing Continuum (02) systems are challenged to ensure the intricate requirements of each computational Ariter Inference Idea (Internet) and the system's stable, the Service Level Objectives (SLDA), which are prepared as these requirements, many the disassence of the system's track, the Service Level Objectives (SLDA), which are prepared as these requirements, many the disassence of the system's stable, the Service Level Objective (SLDA), which are prepared as these requirements, many the disassence of the system's stable, the system's SLDA (Internet) and the system's SLDA
Ingend to more constraint of hims controlly, the described in the second process in the second process in the second process in the second of the process in the second process in the second of the process in the second of the	OCSS: NOVEL CHALLENGES istributed systems, specifically cloud comput- istributed systems, specifically cloud comput-		Satability edge isatilignore, coubing individual edge devices to (1) devices a coust uncentrating at how to enforce their statisticate their SLO and (2) transfer knowledge to speed up the ordenting of heirogeneous devices. Through Transfer knowledge of the statisticate of the statistic
er (1) introduce a constitut titter familie funder. Unter the sector of	In a support many of our daily activities, span- ing from the home to the workspace and including many from the home to the workspace and more constrained, and, finally, at the bottom, we can	1 INTRODUCTION the software application itself. In general, these systems are being developed from a top-down perspective. Hence, each control of the software application is top-down and a perspective. The software application is the software applica	the clark in which at its . Spiritin is responsible for classing Quarky on hereive Quark and Quarky of Experience (QuA) should will be classically classically also well be classically classically classically consists to consum four SDA, furthermore, the undrying essent structures were also maintain the maintain of new types of devices can be data a pointering the functioned. There also them to more existing models, even though
endpaneties to EUEBag ULDs. No environment and execution they are not approximate to execution the execution of the exec	ment activities. These services have become ubiqui- tous due to their business model, performance, and	1 on Cloud infrastructures, however, during the recent years we are witnessing the appearance of new computing tiers, such as the Edge or the Fog. They provide valuable fea- tigen, but it falls short for the "computing on activity".	the device type had been unknown. Finally, rebalancing the lead within a device chater allowed individual edge devices to recover their SLD compliance after a network failure from Z2% to B2%.
Network for exchanges and a final accordingly is liven of incontrolled later lipses. If incontrolled accordingly is liven of incontrolled later lipses. If incontrolled accordingly is liven of incontrol later If incontrol lat	security guarantees. Computing infrastructure centrali- zation, i.e., the data center, has been a critical enabler for this tobeholderur urending an officient and academiater and to build it. The cloud does not govern all the devices in the figure, but each sub-	tures to applications, such as very low latency or privacy enhancements. Ultimately, new applications are emerging that take advantage of all the computing tiers available, entry the same autom. Also in some dimensional of the defined	Introduction and SLO fulfillment [5]. This promotes the usage of Active Inference (AD) [6], an emerging concept from neuroscience that describes how Computing Continuum (OC) systems, as envisioned in [1–5], are the buries continuum environment methods are information to
G. Nove and two spaces or some twome of the second spaces in temptre (). Buywords includes index of the second internet (). Buywords includes internet (). Buywords interne	solution to manage work/wide scale services. Interestingly, the upcoming generations of Internet- interestingly, the upcoming generations of Internet-	hence, they are systems that are simultaneously executed on the Edge, Fog, and Cloud computing tiers. These systems are known as "computing continuum" systems.	large-cold distributed systems compared of multiple computational tiers. Each tierserves a unique purpose, e.g., providing latency-sensitive services (i.e., Edge), or an alonadance of virtual, scalable resources
Department for the control methods of targets of Methods of targets of Methods and targets Exception for equivalence of the department of a set of targets of targets of the department	distributed systems are moving away from these large and isolated data centers to be closer to the users at the off of the product Tbit treed is lustified business. In Figure 1.	The applications provided by the distributed "computing continuum" systems enable developments that belong to the infrastructure landscape for some years now, such a same We claim that new methodologies are required as	(i.e., Good). However, the requirements that each ther must falfill are equally diverse, as they pan a wide variety of edge devices and for nodes. Assume that requirements would be easured in the doad, or be available to the end of the post end end of the end of the post end end of the end of the post end end of the end of the post end end of the end of the post end end of the end of the post end end of the end of the post end end of the end of the end of the end of the end of the end of the end of the end of
Comparing Chailmann (C) entrings as continued in (2-1) are beginned in the second of a work particular formation	eral reasons, such as reducing network congestion, reducing service latency, or improving privacy guaran-	cities with greener and sustainable spaces; non-polluted ave- nues with autoenomus vehicles; sustainable and efficient manufacturing, with accurate tracoability of products;	using out and an econogram guarrant or weeks, name in the second seco
the CC gree and time requiring the second result from the second results of the second r	tees. The list is long and many scientific publications emphasize this change, e.g. Shi et al. and Satyanar- munor 2 subtromes and are mutual to several tenants. Given the open and shared environment these systems	health systems able to personalize anyone's treatment inde- pendently of their current location. Most research efforts have been devoted to solve specific of a complex systems, as cited in 111 a system is complex if its	Given the scale of the CC, requirements must be descentralized; this means that the logic to evaluate requirements must be transferred to the composent that they concern. Goad level requirements, i.e., Service devices in proximity are hundled into a device chater, administered
the tappy durational names of the CC. It is a choice adjustmently with the dynamically with the dynamical tappy and ta	ces closer to the user brings new applications for our society, such as autonomous cars, individualized health while keeping strict constraints fullible in terms of per-	challenges of these systems such as data caching, resource allocation or scheduling among others. However, general methodologies for design and management of these systems. The systems and, changes the	Level topectives (DAD), may must be enargergapties into similar parts by a log node; whenever the hope is scaled up with new devices the account of the state of
Is the reget, to version CC orders requiring documental and addigment, which addres sponse sub-which is used, a data memory which addres address sponse sub-which is used. A standard with a data memory address is used address addre	tracking and assistance, optimized resource manage- ment, holographic communications, and many other encontunities.	remain one of the fundamental challenges. This has been overlooked because the community has been using the tra- ditional methods for design and management develops infrastructure data, we will be able to manage these systems	While it is one challenge to segregate and disceminate SLDs, essuring them is another. Requirements are versate and may charge over time, every component must teld discover how its SLOs are related to the second device characteristics. The Code, at invasion of the second device characteristics of the Code, at
disk für eldt veren in om ender und ender en dener in om ender en ender e	The potential business opportunities are enormous. Hence, it is inevitable that a new generation of Internet- database components. We are studying the technologi- cal challences presented by these new setterms. ⁴ We	for the first Internet-based systems, those comprising a server and a client, which are completely specified through general methodology inhered from cloud computing show-	actions. For this to happen, the device could use Machine Learning (ML) techniques to discover causal relations between its environment
Teach of the Topologics Topologic Section (TALOA), (ECTUDE) © To Topologics Topologics Section (TALOA), (ECTUDE) © To Topologics Section (TALOA), (ECTUDE) © Topologics Section (TALOA	distributed systems, also known as distributed contin- <i>uum systems</i> (<i>DCSs</i>). ³ will eventually be a part of our <u>uum systems</u> (<i>DCSs</i>). ³ will eventually be a part of our	 The authors are unit Distributed Systems Group, TU Ware, 1040 Viewa, Second, to unravel that a completely new approach is fustria. E-mail: identify: causemper, p.denteMitg.humien.ac.at. 	¹ Corresponding andree. Kenad addresse: Iorizzediak@ag.zmeien.ae.at (N. Sodiak), v.zasamayor@htg.tmeien.ae.at (V.C. Pojal), pdosta@dug.zmeien.ae.at (P.K. Bosta), dasadar@big.tmeien.ae.at (S. B. Bastalar).
with the first state of the firs	1089-7801 to 2024 IEEE terms will become a reality. Digital Object Identifier 10.1109/MIC.2023.3344248 DCS applications will have many interactions with particle of current version IE Avril 2024	Measure there were a tool. 2012 (1996) and 11 tool. 2012 (accepted 2 pin. 2022. Date of publication 31 Jac. 2022, date of current tension? Mar. 2023. (Corresponding earlier: Victor Canonyair Phyle). Economised for exceptioner by V. Yang,	https://doi.org/10.1016/j.fumr.2004.03.056 Received 20 November 2020, Received in servined from 23 May 2004, Accepted 28 May 2024 Accelable Configure 20 Suprementational Configure 2014 (2014)

- Designing Reconfigurable Intelligent Systems with Markov Blankets, ICSOC 2023, https://doi.org/10.1007/978-3-031-48421-6_4
- Active Inference on the Edge: A Design Study, IEEE PerconAI 2024, https://doi.org/10.1109/PerComWorkshops59983.2024.10502828
- On Causality in Distributed Continuum Systems, IEEE Internet Computing 2024. <u>https://doi.org/10.1109/MIC.2023.3344248</u>
- On Distributed Computing Continuum Systems, IEEE Transactions on Knowledge and Data Engineering, 2023. <u>https://doi.org/10.1109/TKDE.2022.3142856</u>
- Equilibrium in the Computing Continuum through Active Inference, Future Generation Computer Systems, 2024. https://doi.org/10.1016/j.future.2024.05.056

Additional readings:

- Deep SLOs in the Computing Continuum, ACM AppLIED 2024. <u>https://doi.org/10.1145/3663338.3663681</u>
- Edge intelligence—research opportunities for distributed computing continuum systems, IEEE Internet Computing 2023. https://doi.org/10.1109/MIC.2023.3284693
- Governance and sustainability of distributed continuum systems: a big data approach, Journal of Big Data, 2023. https://doi.org/10.1186/s40537-023-00737-0
- Fundamental Research Challenges for Distributed Computing Continuum Systems , Information, 2023. <u>https://doi.org/10.3390/info14030198</u>
- Exploring the potential of distributed computing continuum systems. Computers, 2023. https://doi.org/10.3390/computers12100198