towards real time in wireless 6G communications: how latency gets complicated

Petar Popovski

AALBORG UNIVERSITY
DENMARK

THE VELUX FOUNDATIONS
VILLUM FONDEN VELUX FONDEN

GOALS

Danmarks Grundforskningsfond
Danish National Research Foundation

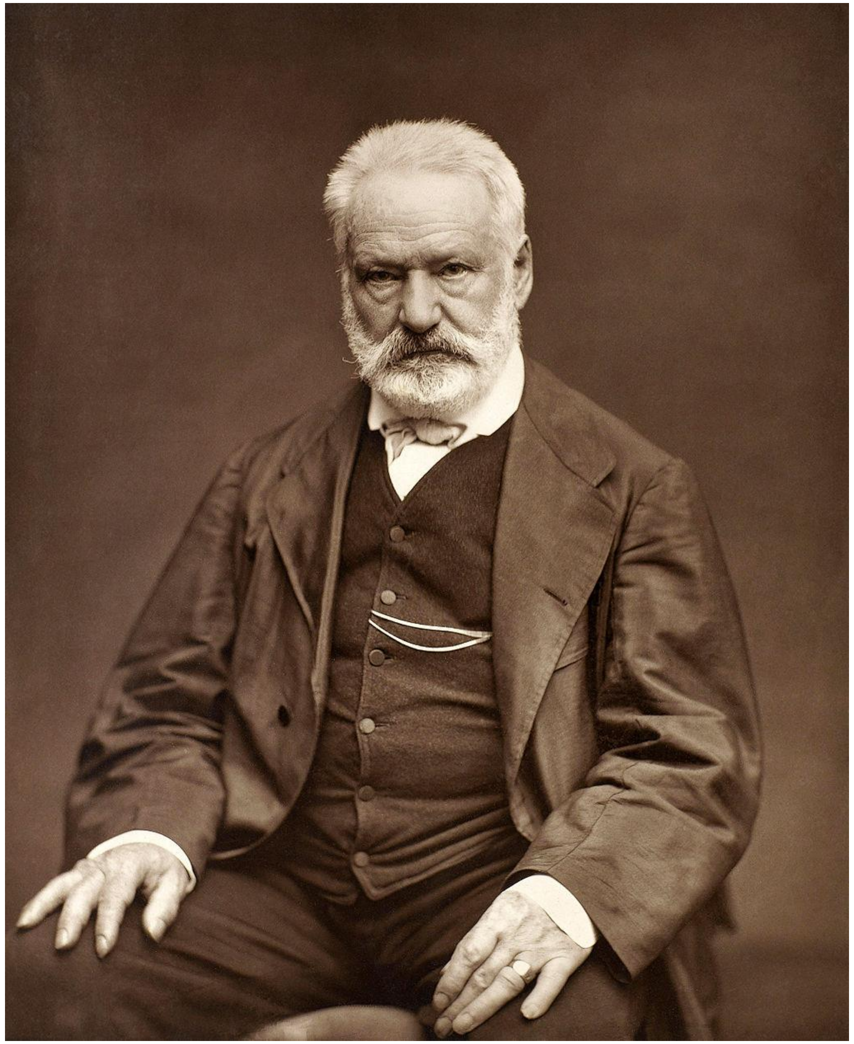
outline

- towards 6G
- modeling latency
- beyond latency
- time and intelligence
- physical vs. digital time, causality, and simultaneity

outline

- **towards 6G**
- modeling latency
- beyond latency
- time and intelligence
- physical vs. digital time, causality, and simultaneity

the shortest correspondence in history



in 1862, Hurst and Blackett

publishes *Les Misérables* by Victor Hugo

Hugo was on vacation and
sent a single-character telegram

?

the response from the publisher was

!

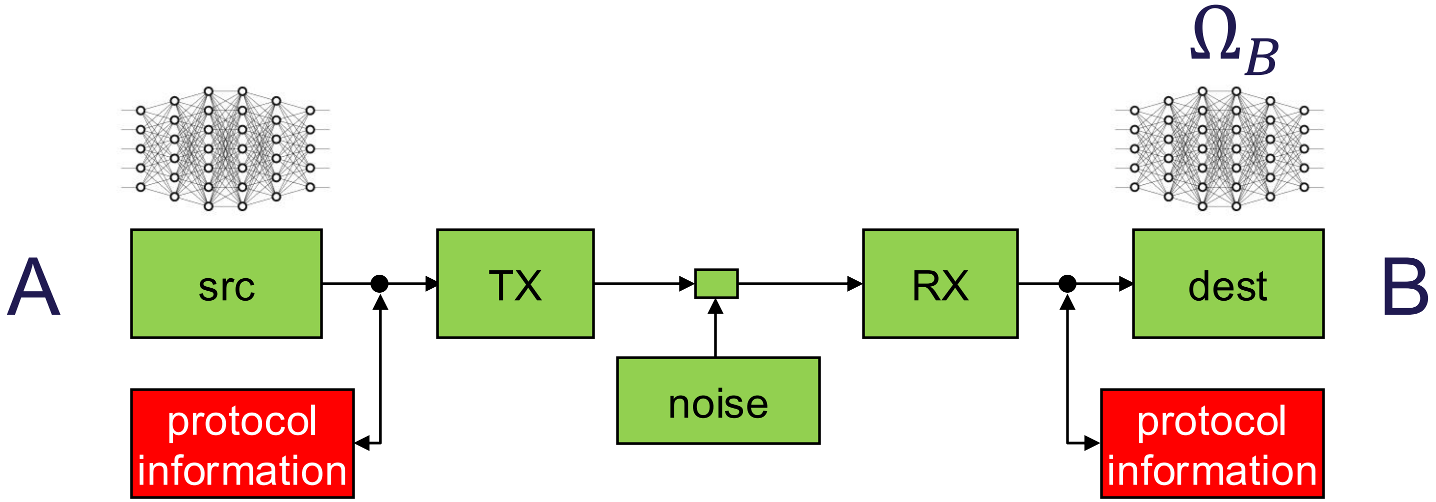
6G augments the connectivity horizon

AI/ML

how are the communication protocols affected
by the growing intelligence in the nodes?

$\cancel{H(X_A)}$

ready to send



$\hat{X}_A \cancel{= G}(Y_B)$

ready to receive

$H(\cancel{X_A}|\Omega_B)$

$\hat{X}_A = G(Y_B;\Omega_B)$

$H(X_A|\hat{\Omega}_B)$

the AI drive for redefining communications

the five language domains	communication systems	Shannon-Weaver
pragmatics	semantic and goal-oriented communications	level C: effectiveness problem
semantics		level B: semantic problem
syntax	control information	level A: technical problem of correct transmission
morphology	packets	
phonology	physical layer	

D. Gündüz, F. Chiariotti, K. Huang, A. E. Kalør, S. Kobus and P. Popovski, "Timely and Massive Communication in 6G: Pragmatics, Learning, and Inference," in IEEE BITS the Information Theory Magazine, doi: 10.1109/MBITS.2023.3322667, 2023.

early examples of token representation

3. THE SERIES OF APPROXIMATIONS TO ENGLISH

To give a visual idea of how this series of processes approaches a language, typical sequences in the approximations to English have been constructed and are given below. In all cases we have assumed a 27-symbol "alphabet," the 26 letters and a space.

1. Zero-order approximation (symbols independent and equiprobable).

   XFOML RXKHRJFFJUJ ZLPWCFWKCYJ FFJEYVKCQSGHYD QPAAMKBZAACIBZL-HJQD.

2. First-order approximation (symbols independent but with frequencies of English text).

   OCRO HLI RGWR NMIELWIS EU LL NBNESEBYA TH EEI ALHENHTTPA OOBTTVA NAH BRL.

3. Second-order approximation (digram structure as in English).

   ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TU-COOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE.

4. Third-order approximation (trigram structure as in English).

   IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID PONDENOME OF DEMONS-TURES OF THE REPTAGIN IS REGOACTIONA OF CRE.

5. First-order word approximation. Rather than continue with tetragram, … , n-gram structure it is easier and better to jump at this point to word units. Here words are chosen independently but with their appropriate frequencies.

   REPRESENTING AND SPEEDILY IS AN GOOD APT OR COME CAN DIFFERENT NAT-URAL HERE HE THE A IN CAME THE TO OF TO EXPERT GRAY COME TO FURNISHES THE LINE MESSAGE HAD BE THESE.

6. Second-order word approximation. The word transition probabilities are correct but no further structure is included.

   THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH WRITER THAT THE CHAR-ACTER OF THIS POINT IS THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE TIME OF WHO EVER TOLD THE PROBLEM FOR AN UNEXPECTED.

Seven Briefs on Semantic Communication and 6G

Petar Popovski 11 min read · Jun 10, 2022

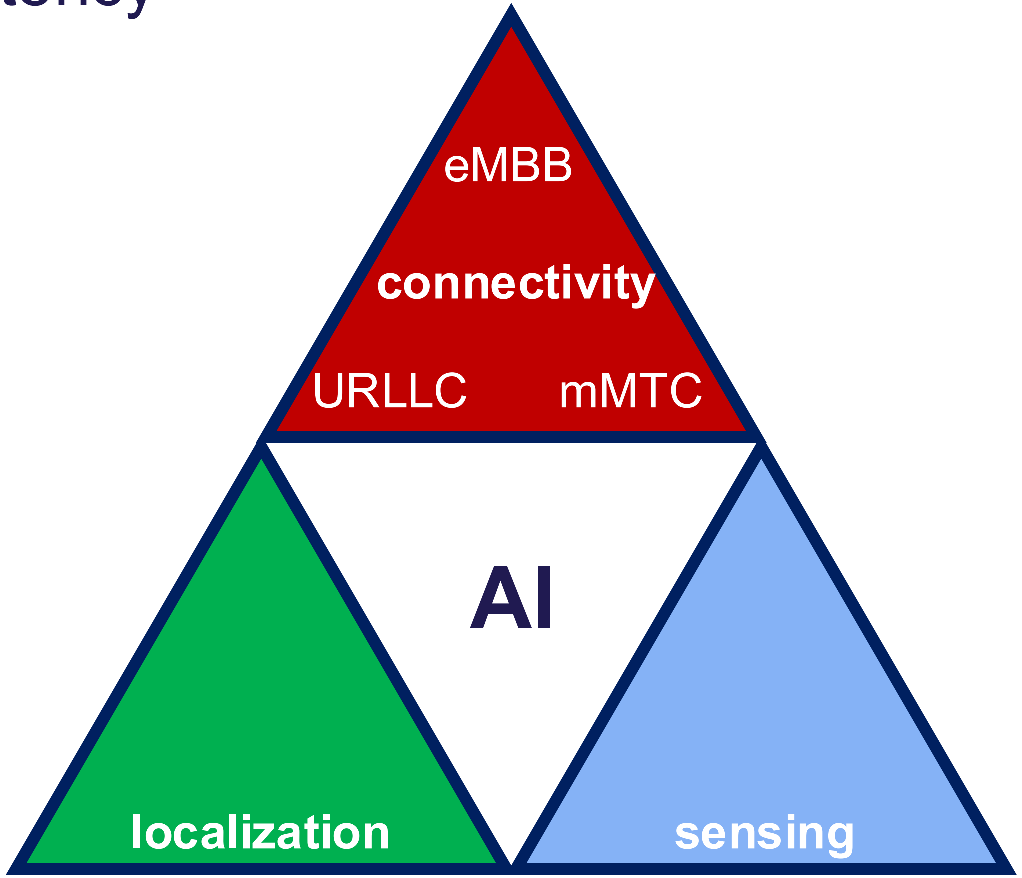
6G

5G had the grand ambition to become connectivity supermarket

- not as expected, but brought focus to latency
- there is no odd-generation-curse:
  iPhone was invented during 3G

emerging applications
towards increased **fusion**
of **physical** and **digital** world

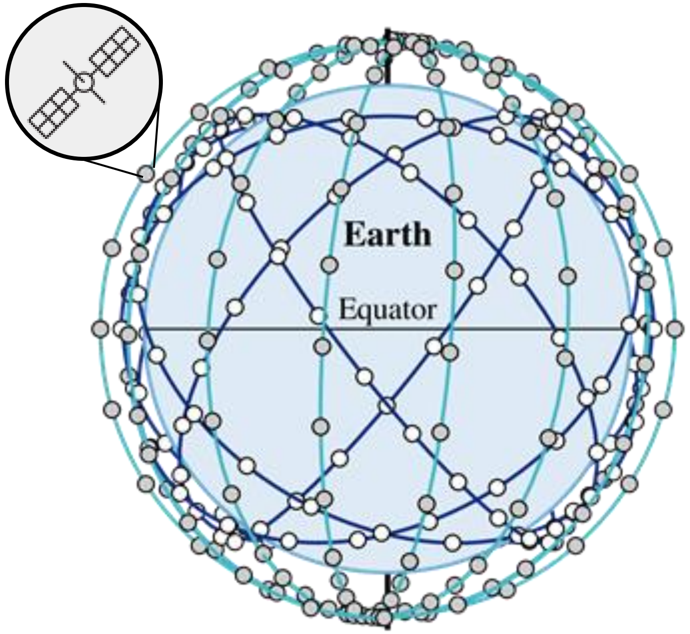
- digital twins
- metaverse
- extended reality (XR)



6G: augmenting further the connectivity horizon

Non-Terrestrial Networks

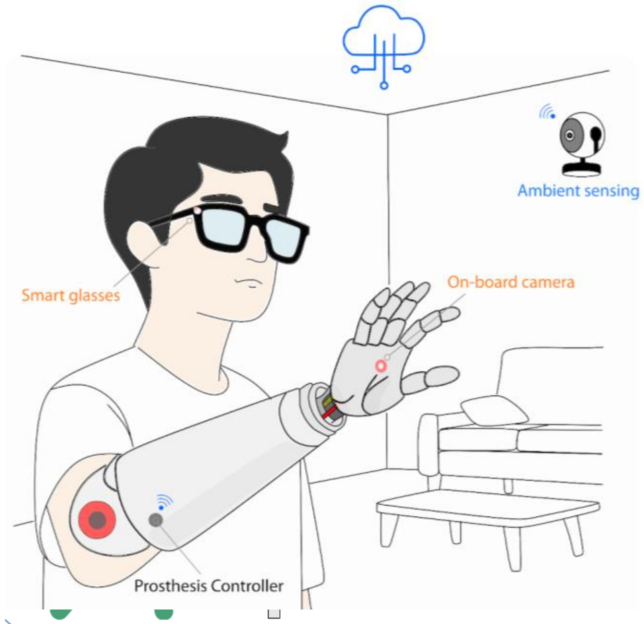
- augment the spatio-temporal horizon of wireless mobile networks



Base Stations get new role

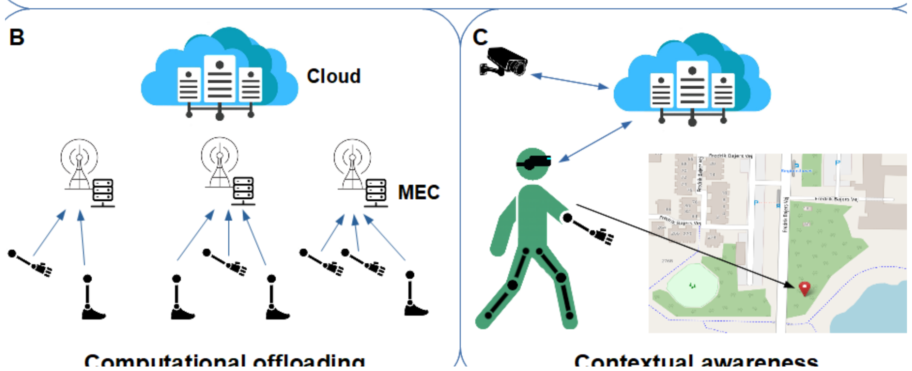
- Gateways between physical and digital
- ISAC and fueling real-time data about the physical world

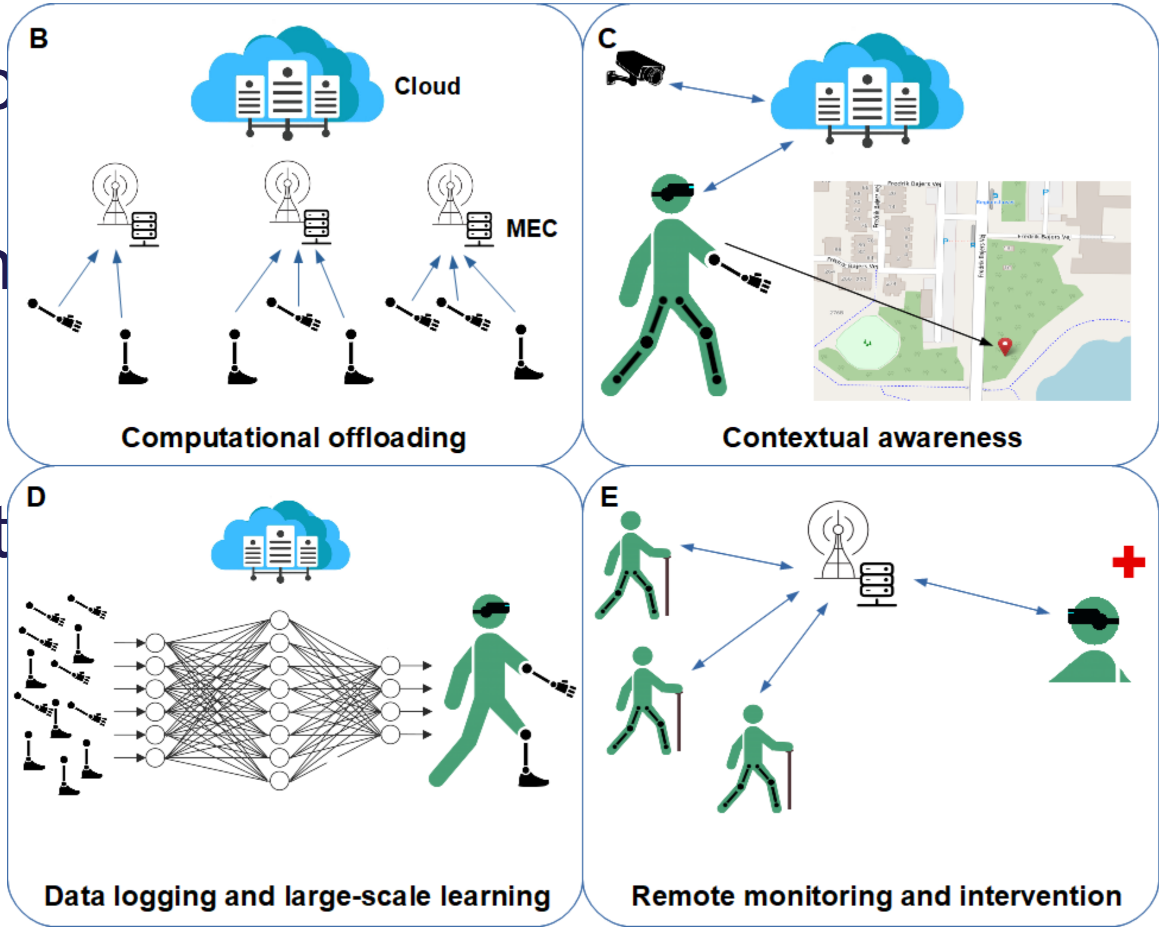
non-standard use cases with real-time wireless



bionic limb

1. mechan
2. analog
3. digital
4. connect

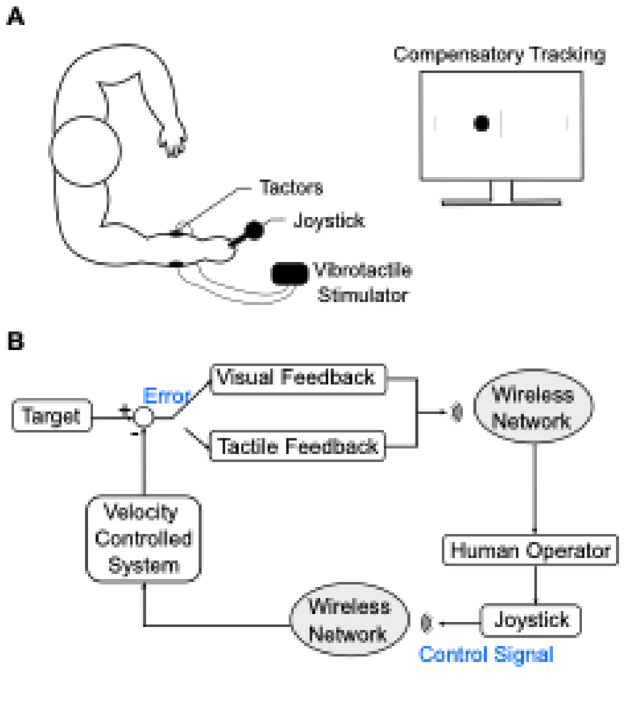




F. Chiariotti, P. Mamidanna, S. Suman, Č. Stefanović, D. Farina, P. Popovski, and S. Došen, "The Future of Bionic Limbs: The Untapped Synergy of Signal Processing, Control, and Wireless Connectivity," in IEEE Signal Processing Magazine, accepted, 2024.

what is a good enough latency?

- instead of a rigid requirement (1 ms),
  observe how information is processed and used (goal!)
- closed loop latencies over 150 ms
- objective: track a target trajectory with a joystick
  - we measure the correlation
- performance affected by **average latency**,
  not much by random latency fluctuations

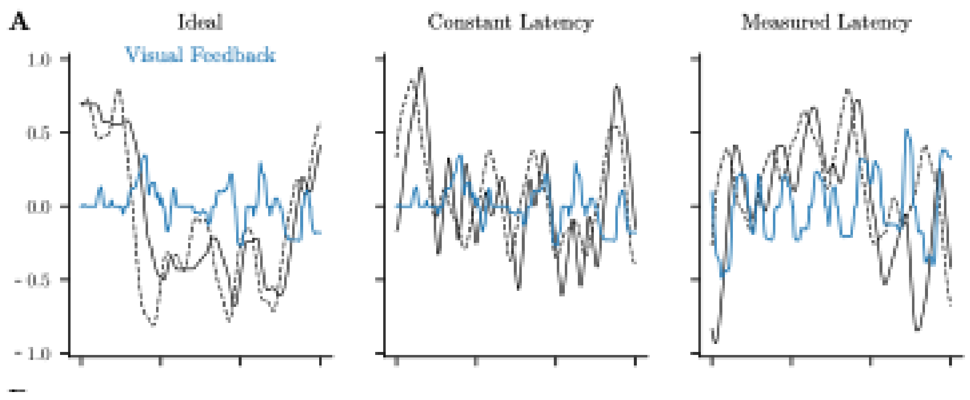


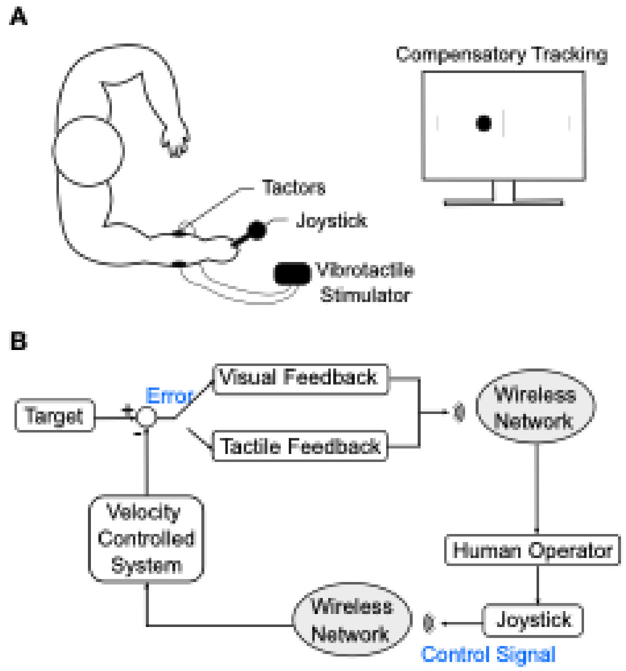
control with visual feedback always better than with tactile feedback

S. Suman, P. Mamidanna, J. J. Nielsen, F. Chiariotti, C. Stefanovic, S. Dosen, and P. Popovski, "Closed-loop Manual Control with Tactile or Visual Feedback under Wireless Link Impairments," in IEEE Transactions on Haptics, accepted, 2024.

lessons learned

- users used their intelligence to predict the trajectory and mitigate the errors and latency variations
- we had to design a deliberate random trajectory to measure the impact of the latency





latency and timing requirements should be coupled to the node intelligence

S. Suman, P. Mamidanna, J. J. Nielsen, F. Chiariotti, C. Stefanovic, S. Dosen, and P. Popovski, "Closed-loop Manual Control with Tactile or Visual Feedback under Wireless Link Impairments," in *IEEE Transactions on Haptics*, vol. 18, no. 2, pp. 352-361, April-June 2025.

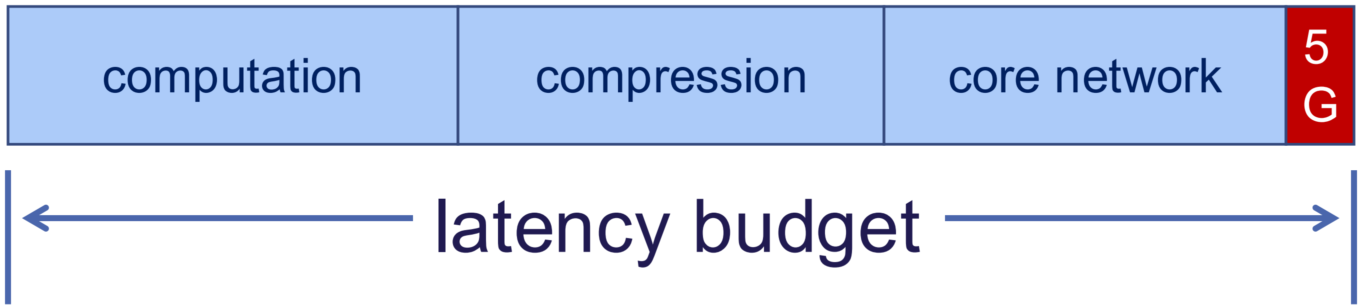
outline

- towards 6G
- **modeling latency**
- beyond latency
- time and intelligence
- physical vs. digital time, causality, and simultaneity

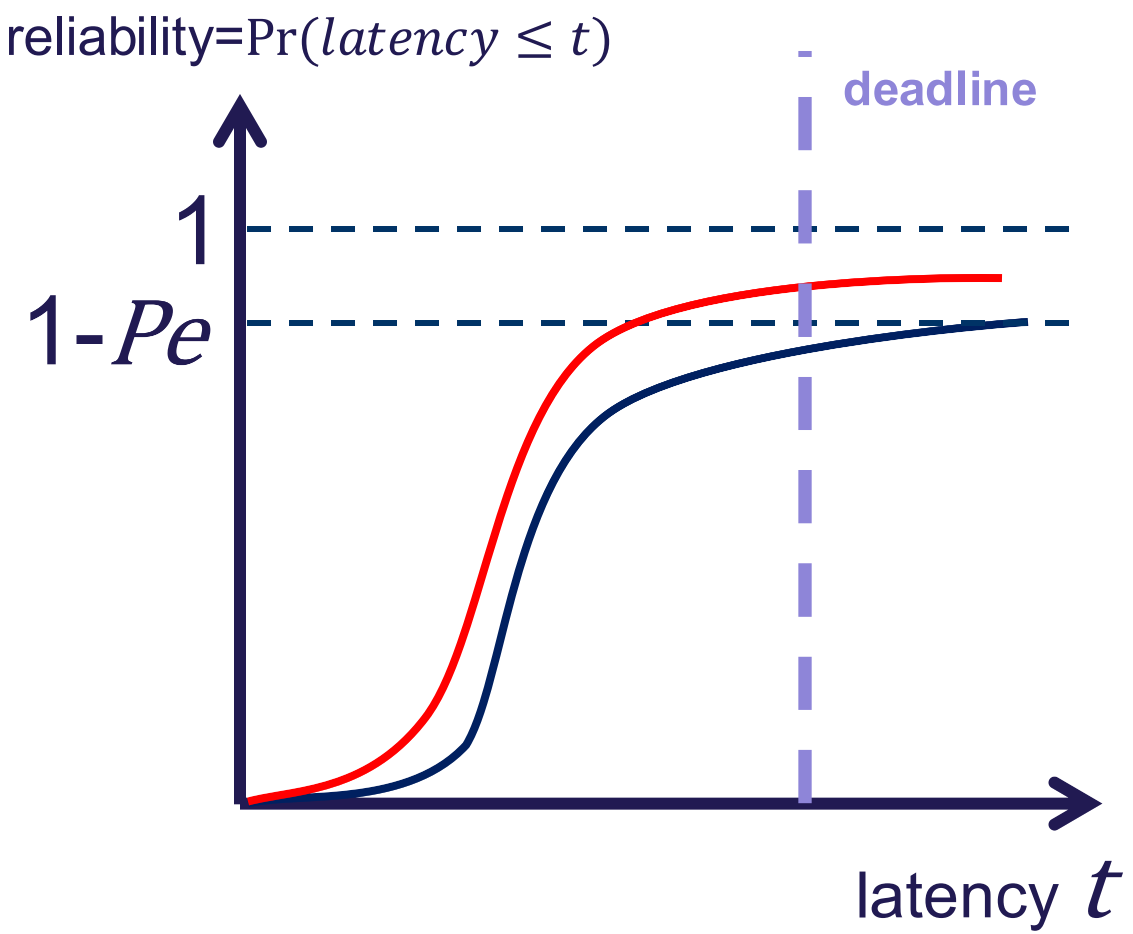
latency as inherited from 5G

- low latency heralded as one of the main features in 5G
  - coupled with high reliability into
    Ultra Reliable Low Latency Communication (URLLC)

- maximalist approach:
  cut a small piece of the overall latency



latency-reliability characterization for fixed data size



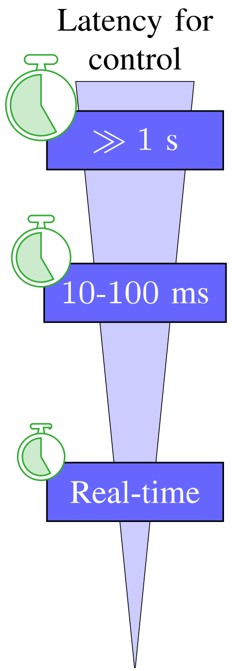
we move from the blue curve to the red curve by using diversity

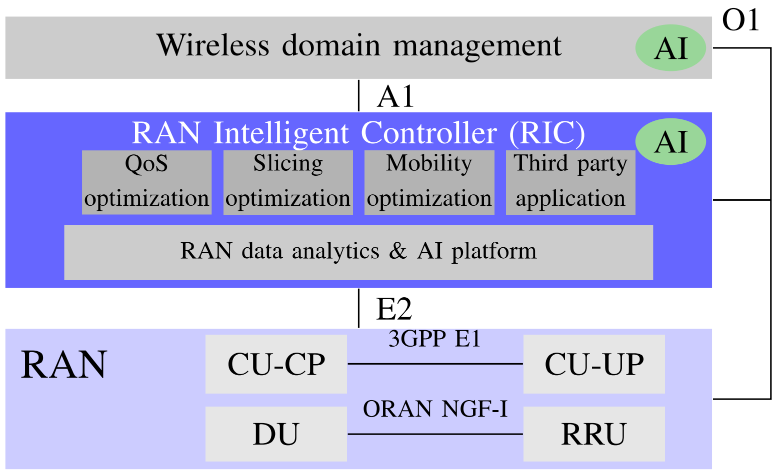
timing requirements at different scales

general tendency in 6G towards smaller timing scales

O-RAN (Open Radio Access Network) alliance defines 3 categories

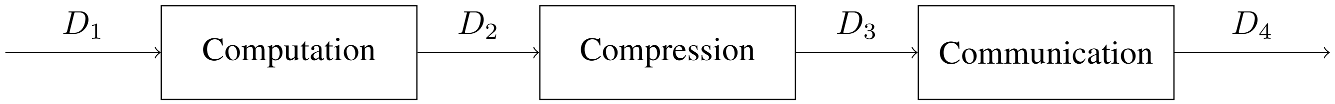
- real-time
- near-real time
- non-real time



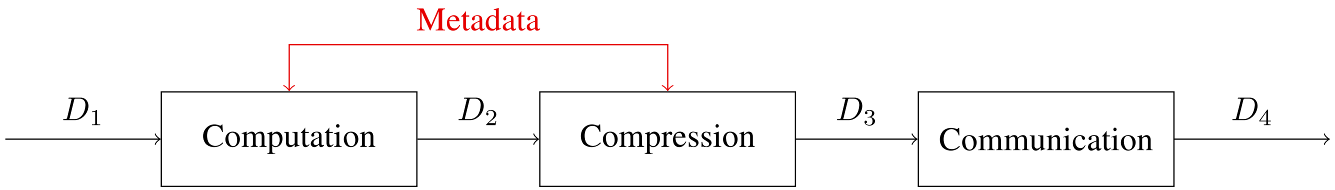


timing in a cascade of modules

$$T_1 \quad + \quad T_2 \quad + \quad T_3$$



processing time depends only on the data

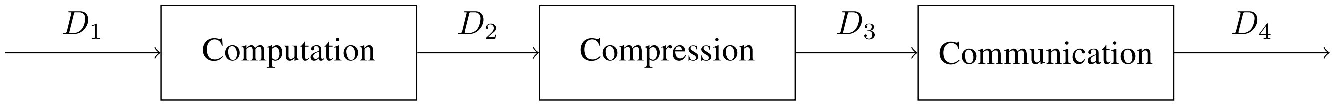


processing time of computation and compresion get coupled

processing time of computation and compresion inseparable

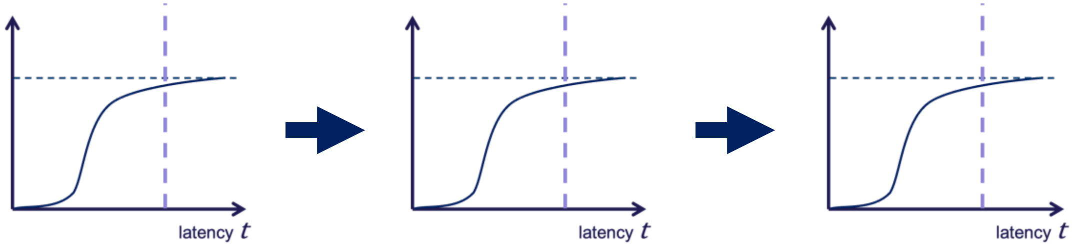
expected vs. tail

$$T \quad = \quad T_1 \quad + \quad T_2 \quad + \quad T_3$$

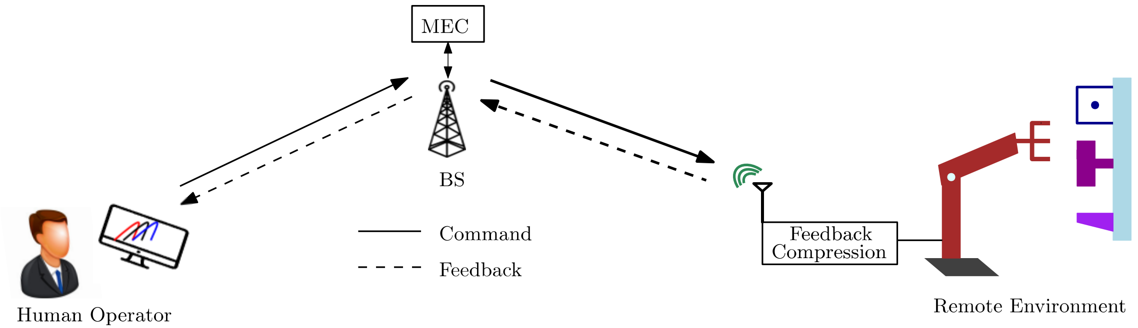


$$E_T[T] = E_T[T_1 + T_2 + T_3] = E_{T_1}[T_1] + E_{T_2}[T_2] + E_{T_3}[T_3]$$

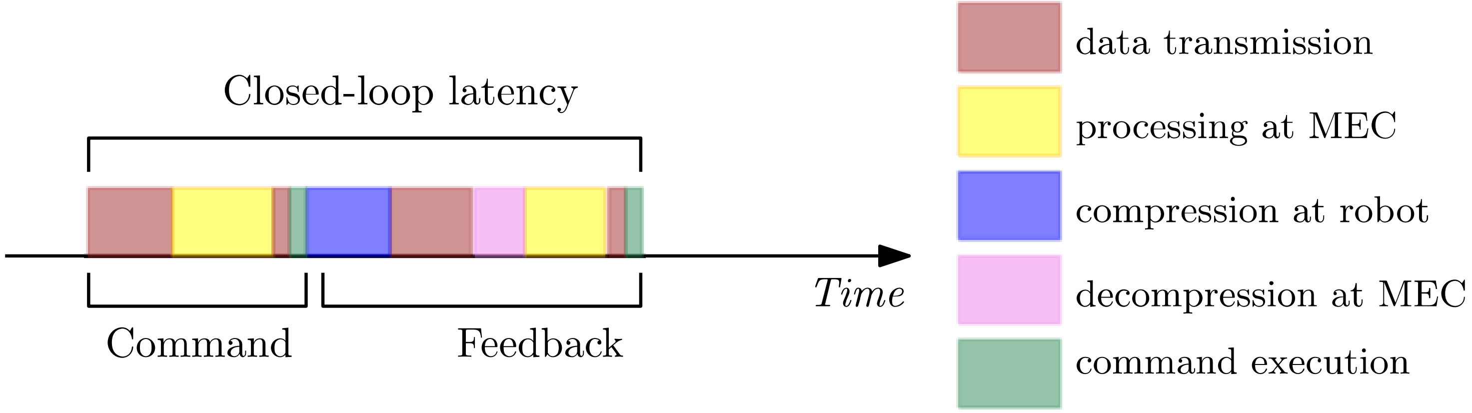
but $\Pr[T \leq \tau]$ has a nontrivial relation to the other CDFs.



latency budget

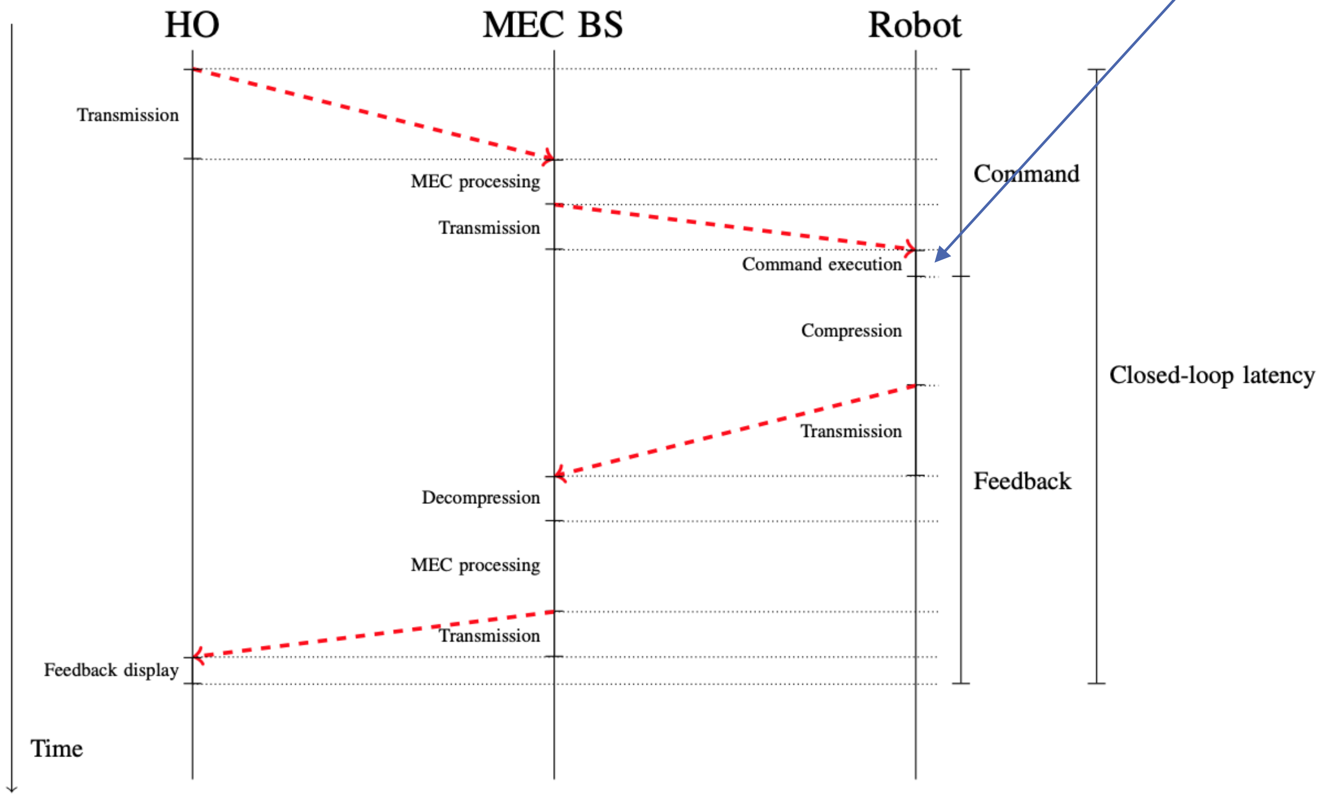


robotic hand controlled
via a Mobile Edge Computing



S. Suman, F. Chiariotti, Č. Stefanovic, S. Došen and P. Popovski, "Statistical Characterization of Closed-Loop Latency at the Mobile Edge," in IEEE Transactions on Communications, vol. 71, no. 7, pp. 4391-4405, July 2023,

system model



not included/modeled, heavily dependent on the application

all components of the closed-loop latency are random variables

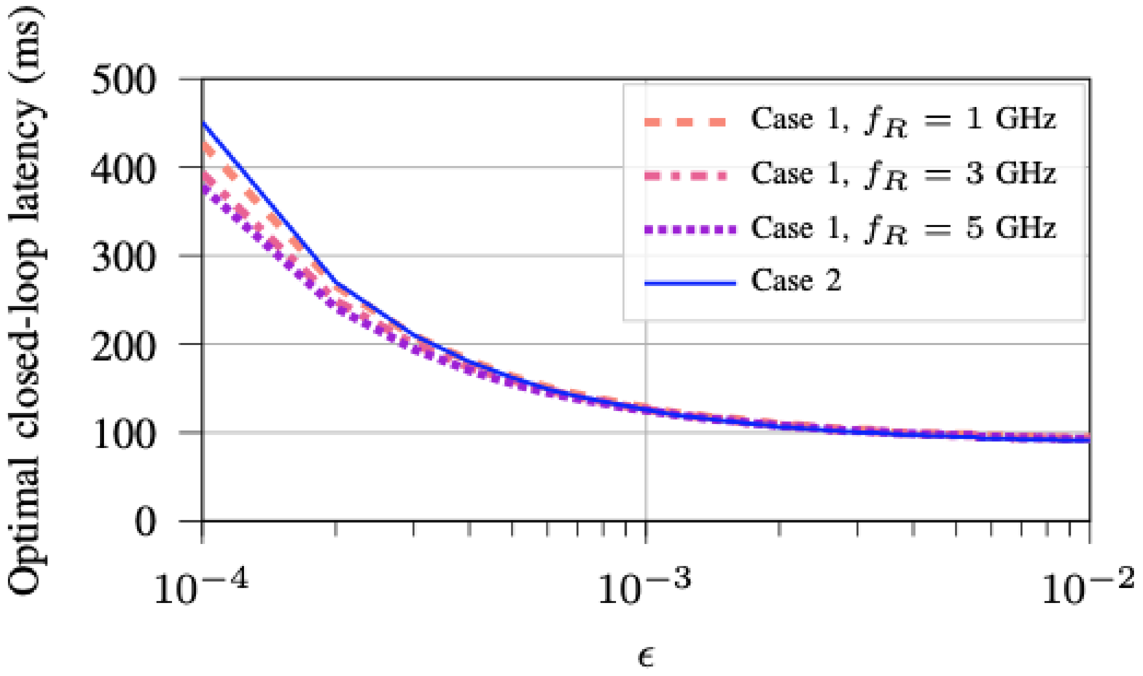
- assumed that they are independent – connected only through the data size
- may not work in ORAN with common computation resources

two cases

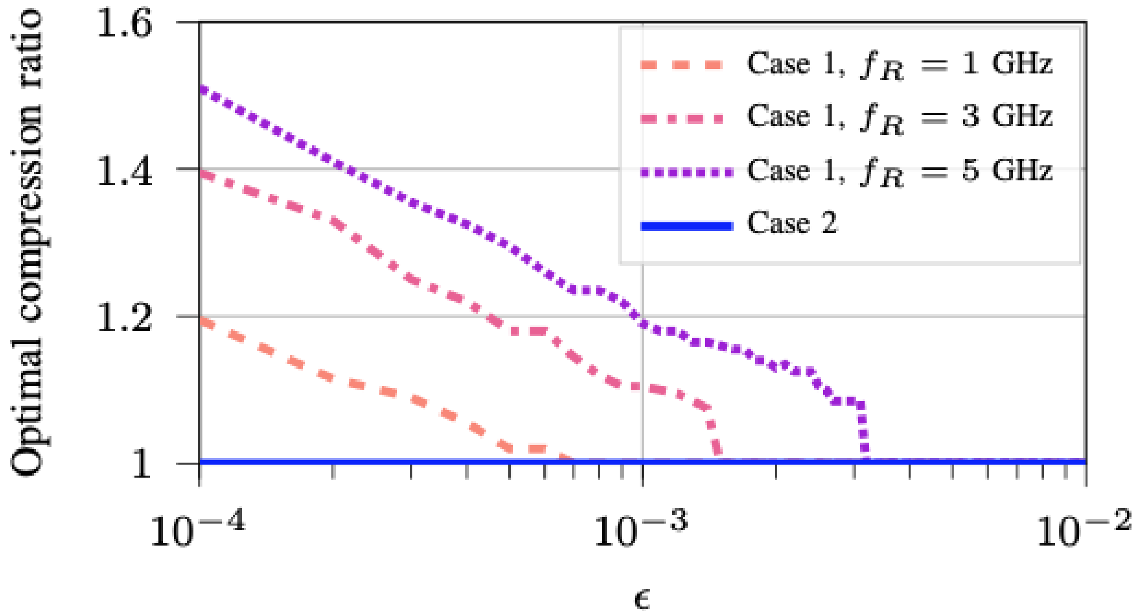
- case 1: robot compresses data
- case 2: robot does not compress

optimizing operations across components

$\rho_{\text{th}} = 0.95$



closed-loop latency



compression ratio

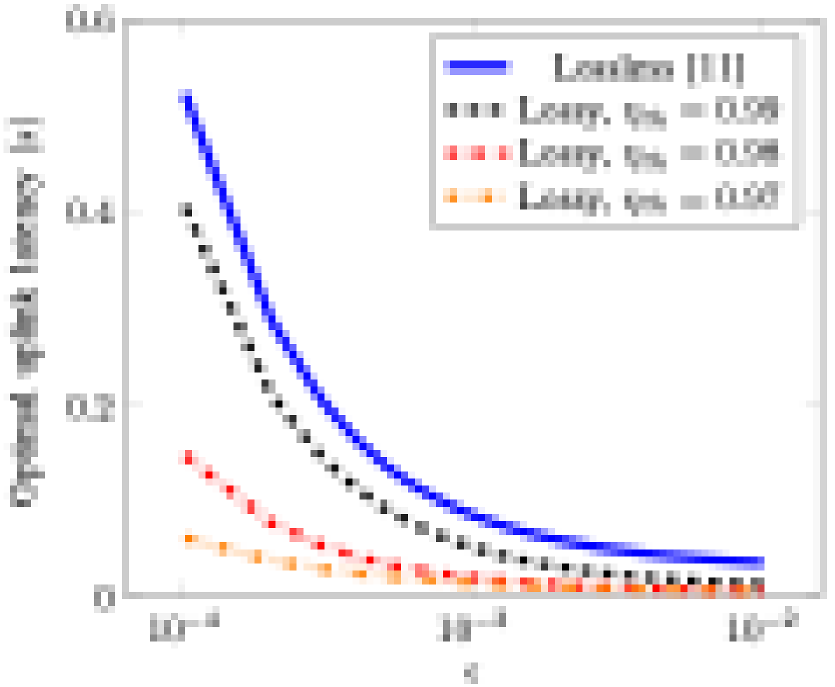
generalizing with lossy compression

accuracy with compression factor Q

$$\eta(Q) = \begin{cases} -0.00387Q + 0.99547 & \text{if } 1 < Q \leq 4, \\ -0.00083Q + 0.9833 & \text{if } 4 < Q \leq 16, \\ -0.00242Q + 1.009 & \text{if } 16 < Q \leq 49. \end{cases}$$

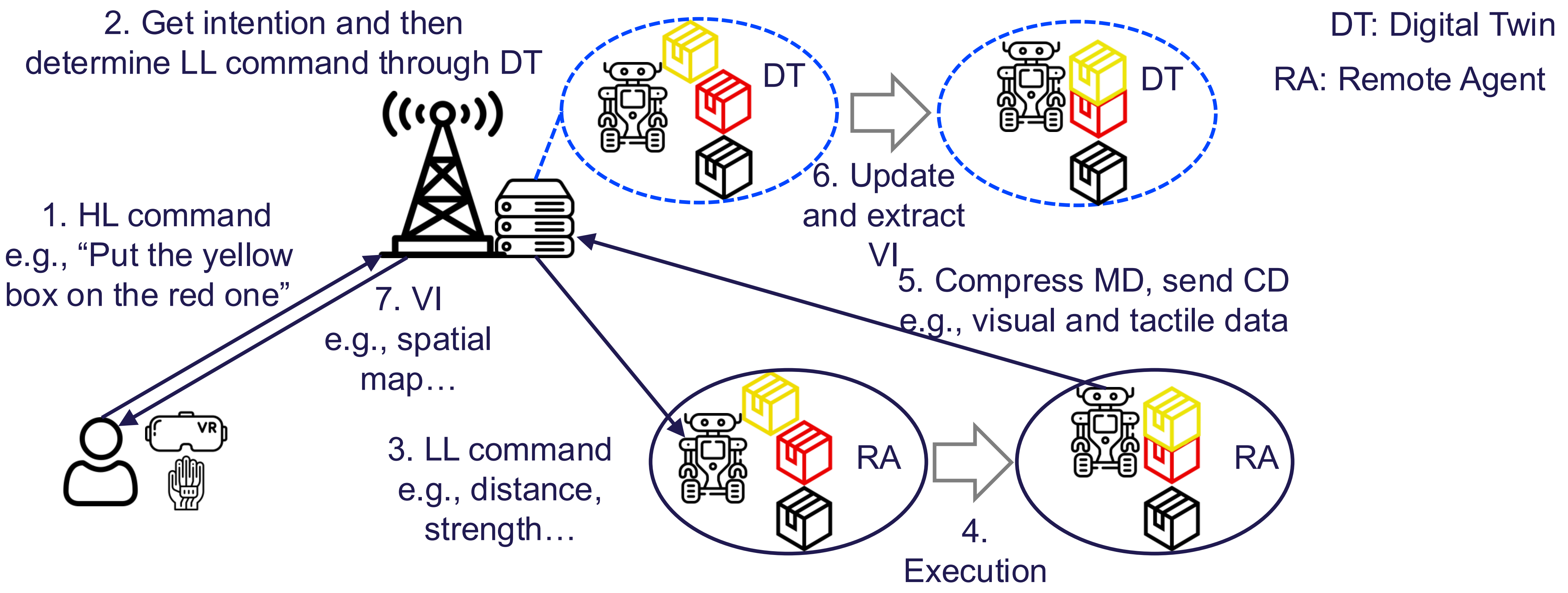
$$\min_{Q} \tau_u,$$
$$\textbf{(C1)}: F_{T_u}(\tau_u) \geq \varrho_{th}, \quad \textbf{(C2)}: \eta(Q) \geq \eta_{th}$$
$$\textbf{(C3)}: 1 < Q \leq Q^{\max}.$$

find minimal latency that satisfies the reliability and accuracy requirements



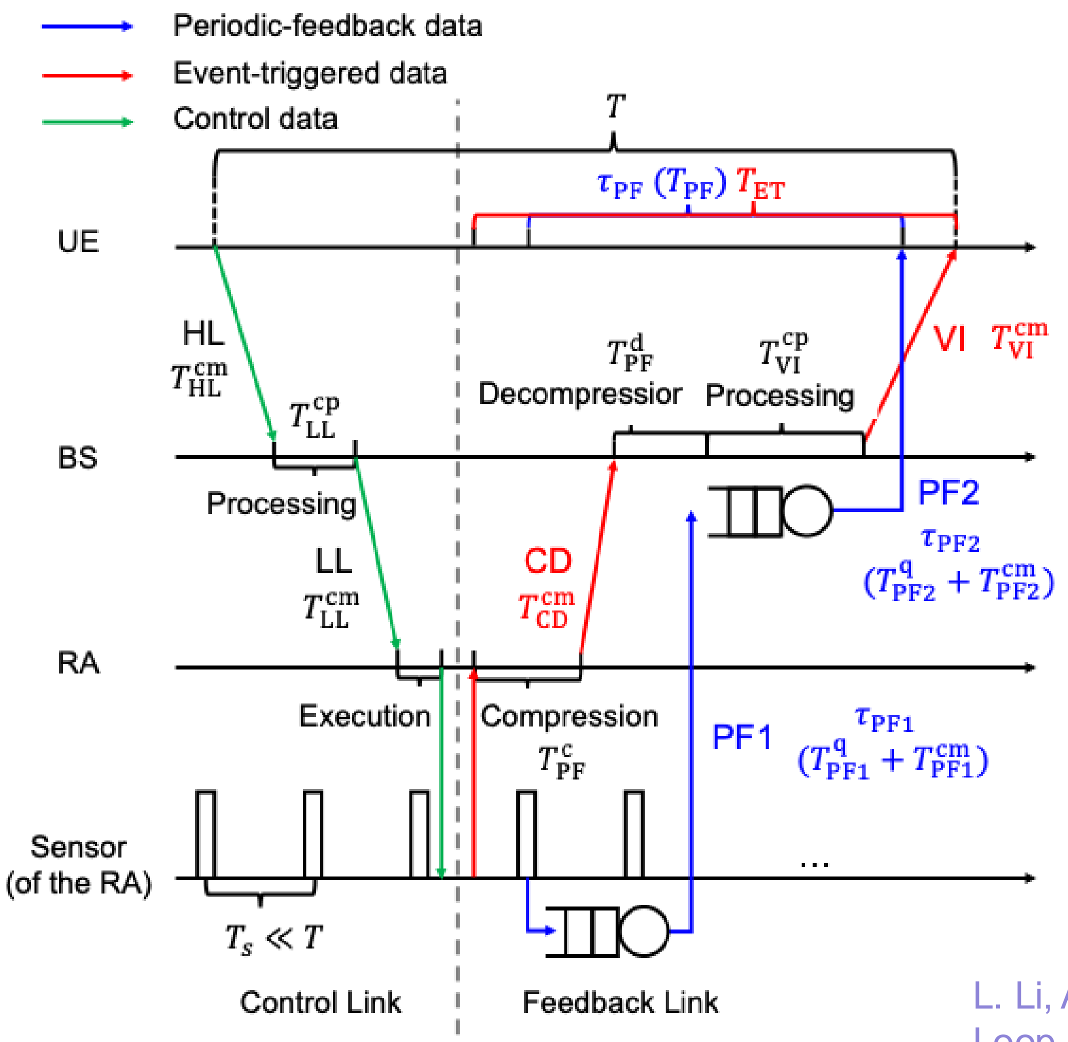
outage probability

generalized closed-loop latency 1/2



L. Li, A. E. Kalør, P. Popovski and W. Chen, "Unified Timing Analysis for Closed-Loop Goal-Oriented Wireless Communication," in IEEE Transactions on Wireless Communications, doi: 10.1109/TWC.2024.3525087.

generalized closed-loop latency 2/2



large number of latency components

saddlepoint approximation to characterize the closed loop latency

L. Li, A. E. Kalør, P. Popovski and W. Chen, "Unified Timing Analysis for Closed-Loop Goal-Oriented Wireless Communication," in IEEE Transactions on Wireless Communications, doi: 10.1109/TWC.2024.3525087.

energy, information processing, and latency

data processing time

$$T_D = \frac{D \cdot X_c}{f}$$

D data size

$X_c \sim Gamma(\kappa, \beta)$

f processor frequency

power consumption

$$P \sim f^3$$

energy consumption

$$E \sim f^3 \cdot \frac{D}{f} = f^2 D$$

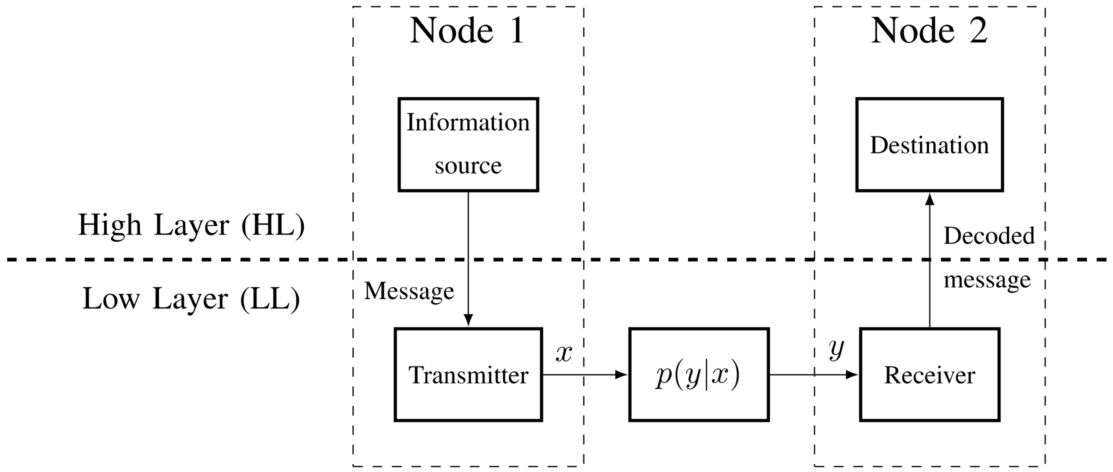
low **latency** requires
fast **information processing**
and induces **energy** cost

outline

- towards 6G
- modeling latency
- **beyond latency**
- time and intelligence
- physical vs. digital time, causality, and simultaneity

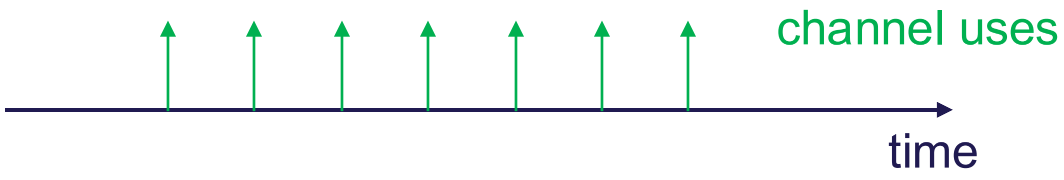
Shannon's communication model and time

- the model relies on causal relationships among channel uses



- layered representation

- timing is NOT inherently a part of Shannon's model
  - need to specify a bandwidth B and timing structure underneath the channel uses
  - precondition to connect the mathematical model to the physical world

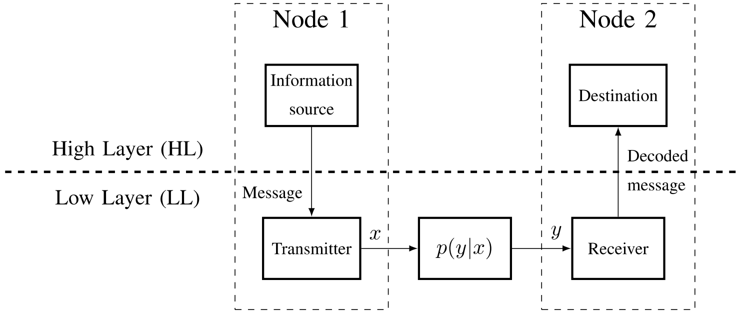


$$T_n = \frac{n}{2B}$$

timing in Shannon's communication model

- tacit assumptions

1. instant message decoding
2. no additional latency due to protocol interactions
3. sender and receiver ready
4. data always available.



- the model is suitable to treat the correctness of communications
  - for full "syntax", we need to account for the control information
- adding semantics/pragmatics requires change of the model, e.g. knowledge about the state of the receiver.

latency vs. age

- latency performance historically characterized with packet delays
- tracking applications and sense-compute-actuate cycles
  are not sensitive to packet delay,
  but to the freshness of the information at the receiver

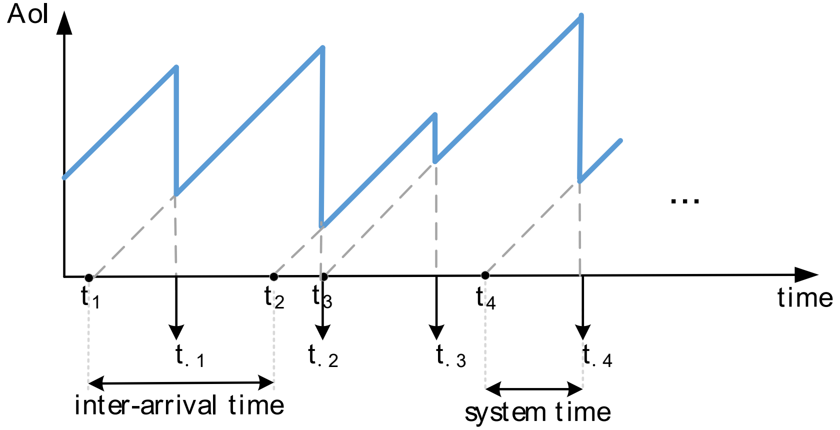
example:
satellite-based tracking

Age of Information

- Age of Information (AoI) and its byproducts
are better metrics to capture the freshness of the information
- suitable for tracking applications and sense-compute-actuate cycles
- exogeneous vs. controlled sampling (generate-at-will)

key assumption

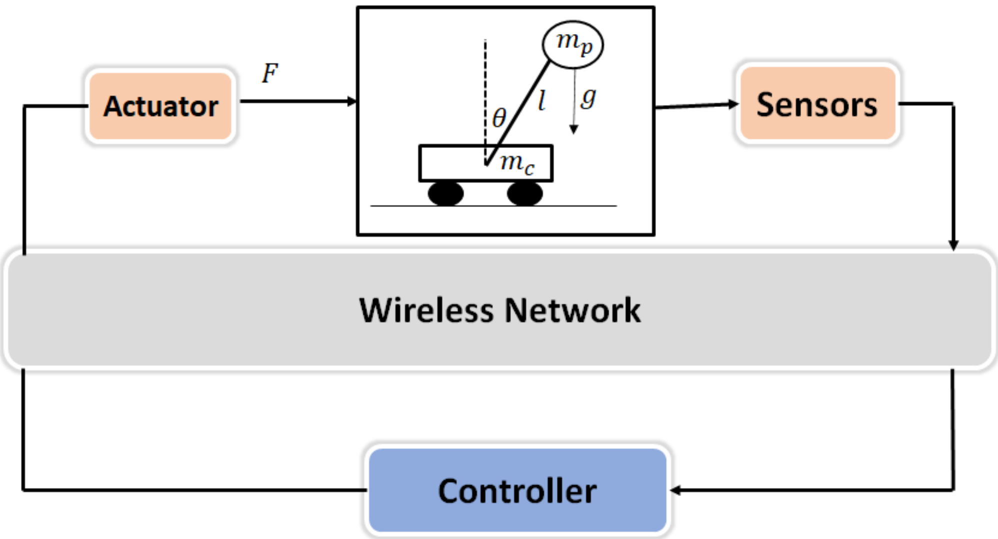
- fresher information carries more relevance, has a higher effectiveness (e.g. stability) and or presents a more meaningful picture of the current state of affairs.

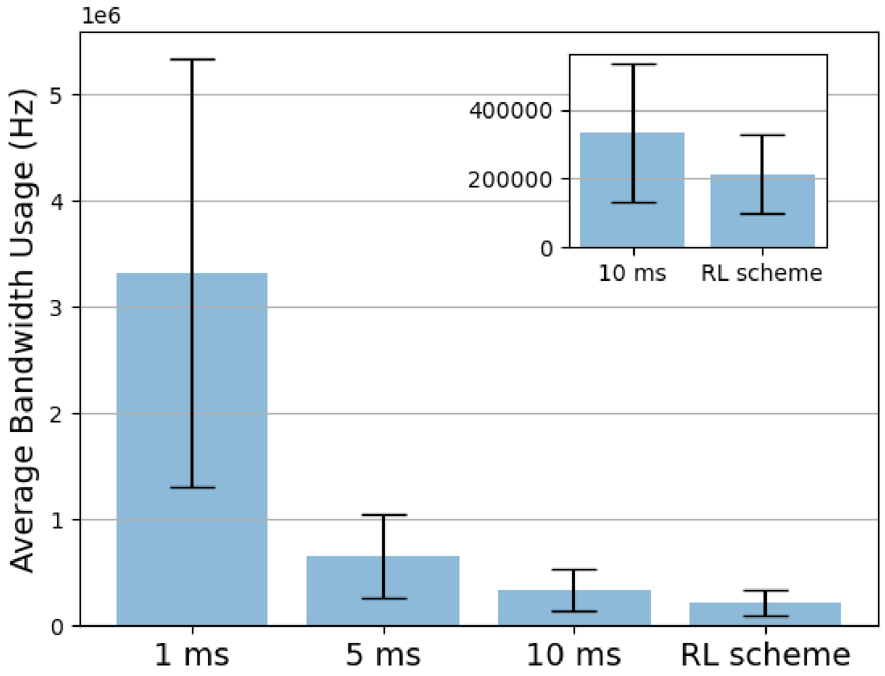


[1] S. Kaul, R. Yates, and M. Gruteser, “Real-time status: How often should one update?” in International Conference on Computer Communications. IEEE, Mar. 2012, pp. 2731–2735.

[2] A. Kosta, N. Pappas, and V. Angelakis. "Age of information: A new concept, metric, and tool." Foundations and Trends in Networking 12.3 (2017): 162-259.

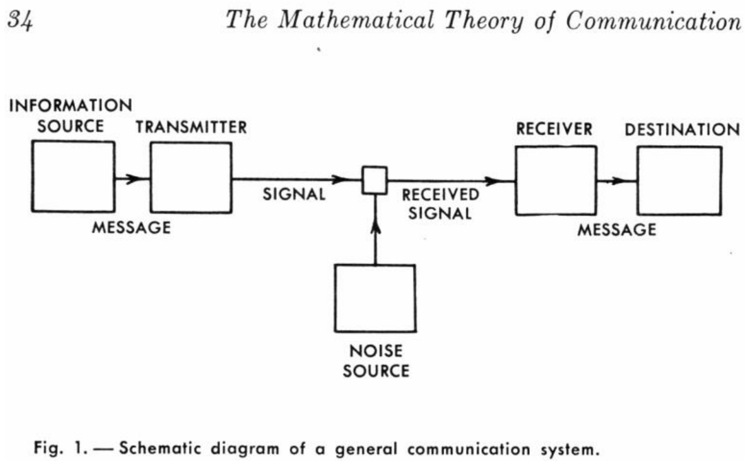
Age of Loop for goal-oriented communications





P. M. de Sant Ana, N. Marchenko, P. Popovski and B. Soret, "Age of Loop for Wireless Networked Control Systems Optimization," 2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2021
P. M. de Sant Ana, N. Marchenko, B. Soret and P. Popovski, "Goal-Oriented Wireless Communication for a Remotely Controlled Autonomous Guided Vehicle," in IEEE Wireless Communications Letters, vol. 12, no. 4, pp. 605-609, April 2023

who initiates communication?



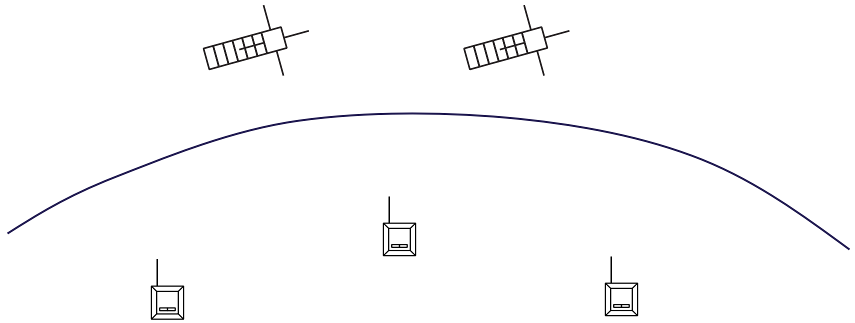
push-based communication

- example: Shannon’s model
  - the transmitter determines what to send
  - the receiver is always ready to receive
  - the receiver is *permanently subscribed* to the sender
  - the receiver is always interested in what the transmitter has to say
  - difficult to define semantics, unless we define what the receiver does with the message

who initiates communication?

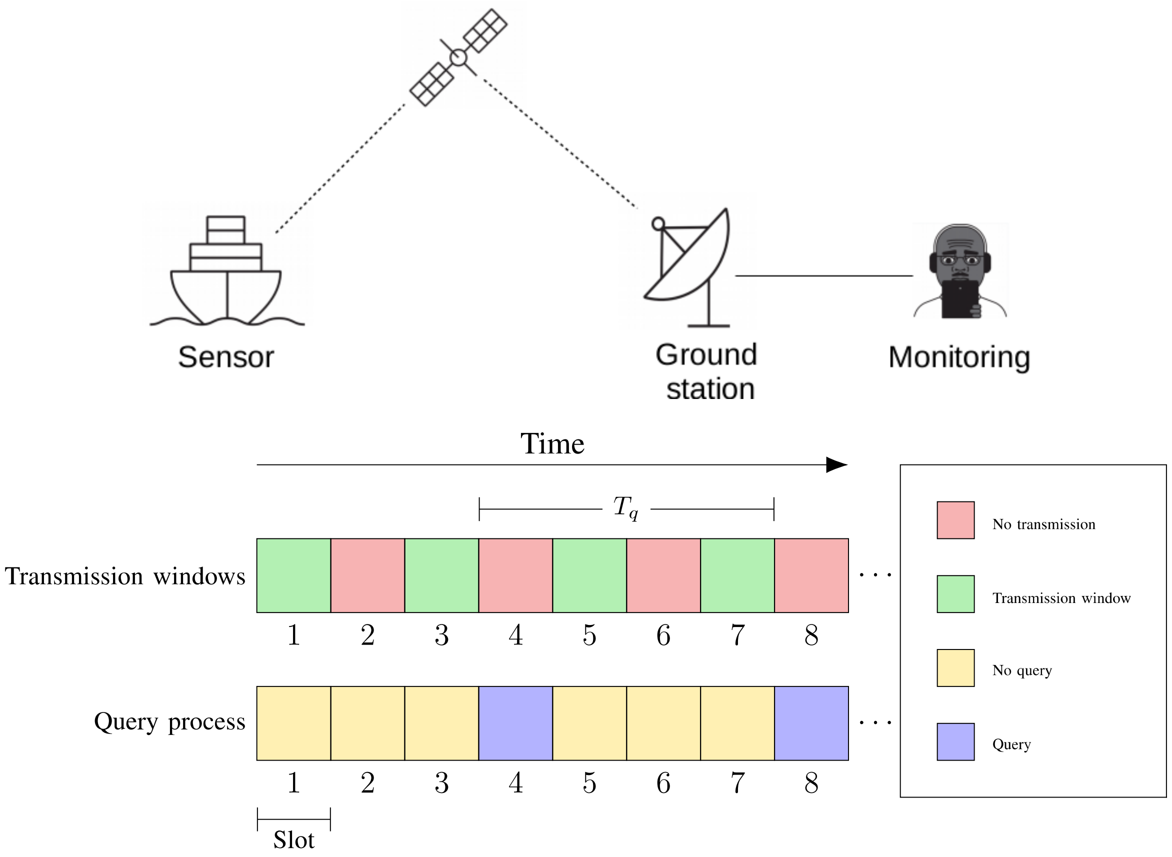
pull-based communication

- reading process initiated by query
- satellites, cloud-based queries to the edge devices, data fetching in a control loop
- semantics related to the query
- predictive query response



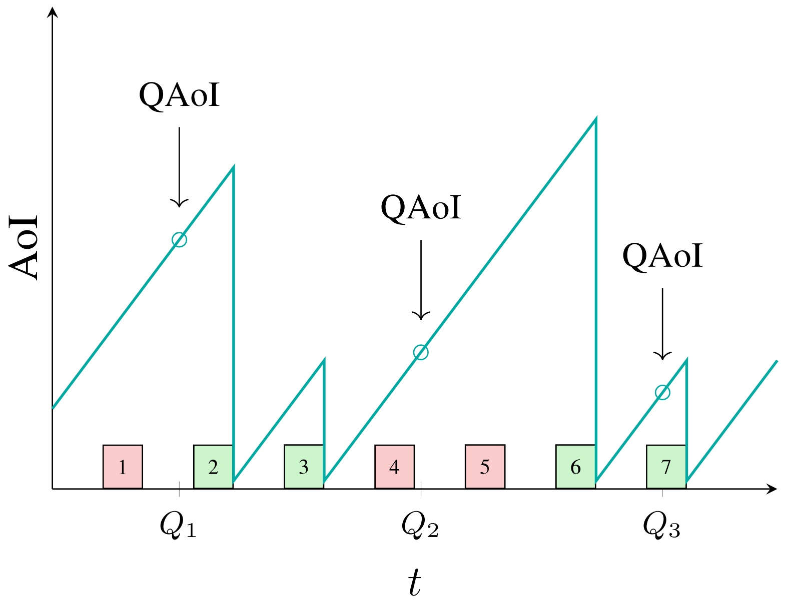
J. Holm, A. E. Kalør, F. Chiariotti, B. Soret, S. K. Jensen, T. B. Pedersen, and P. Popovski, "Freshness on Demand: Optimizing Age of Information for the Query Process", in Proc. IEEE ICC, Montreal, Canada (Virtual), June 2021.

Query Age of Information (QAoI)

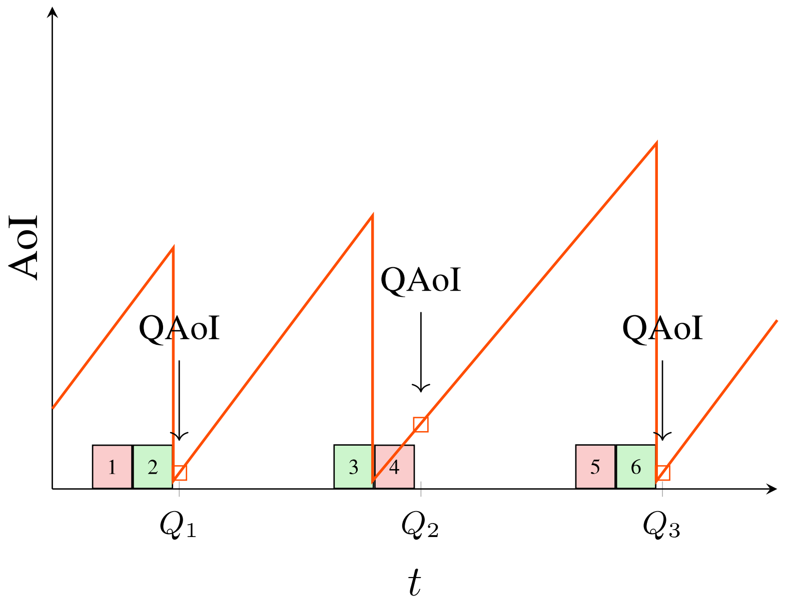


[*] F. Chiariotti, J. Holm, A. E. Kalør, B. Soret, S. K. Jensen, T. B. Pedersen, and P. Popovski,, "Query Age of Information: Freshness in Pull-Based Communication," in IEEE Transactions on Communications, vol. 70, no. 3, pp. 1606-1622, March 2022

Query Age of Information (QAoI)



permanent subscription



query-aware transmission

key question: how to organize the transmissions
in order to offer freshest data at the moment of query?

Query Age of Information (QAoI)

QAoI for the i-th query

$$\tau(i) = \Delta(t_{q,i})$$

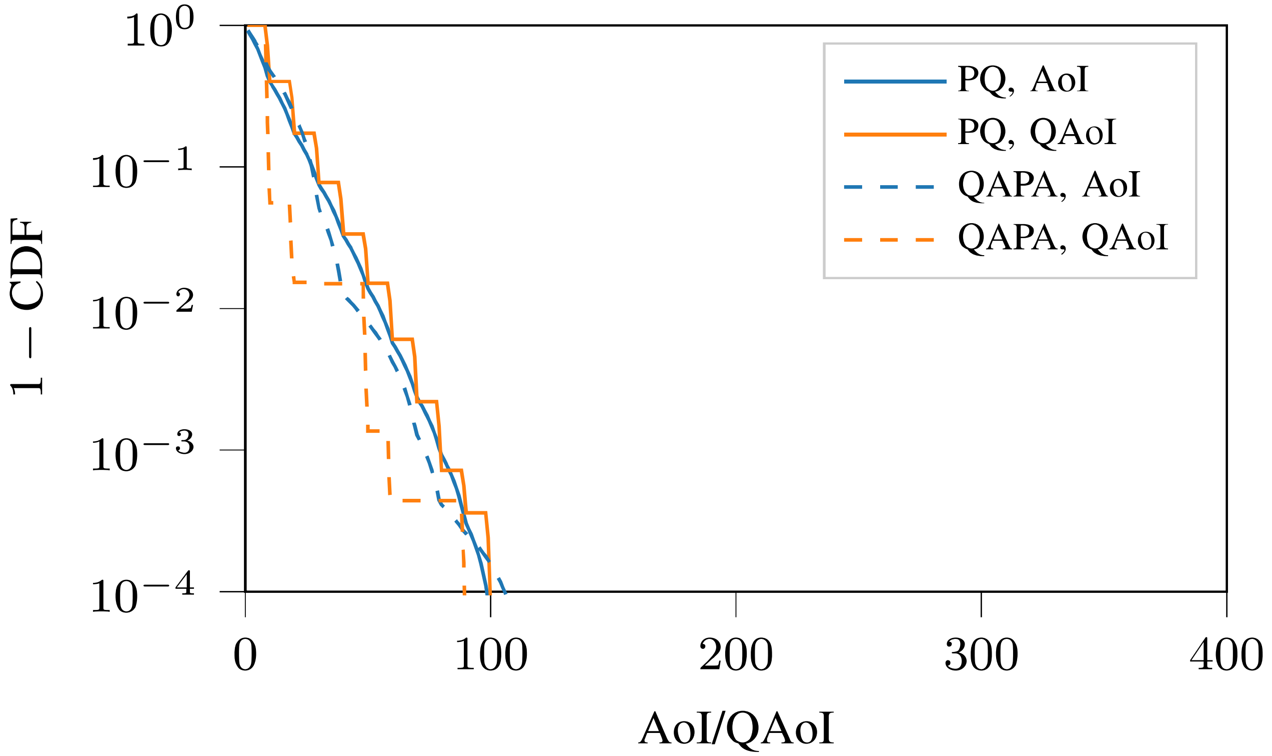
$$\tau_\infty = \limsup_{T \to \infty} \frac{1}{T} \mathrm{E}\left[\sum_{i:t_{q,i} \leq T} \Delta(t_{q,i}) \right]$$

- channel assumed to be based on a Markov model
- formulated as Markov Decision Process (MDP)

Query Age of Information (QAoI)

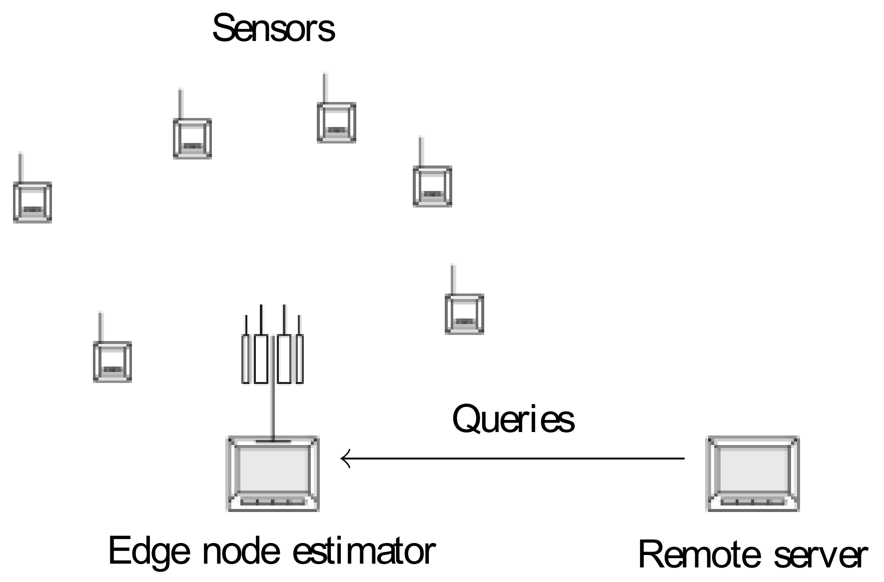
Query Arrival Process Aware (QAPA) vs. Permanent Query (PQ)

- periodic query arrival



looking into the query content

the edge node
estimates a nonlinear function
based on sensor readings



each sensor gets a
noisy observation of a process

$$\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t-1) + \mathbf{v}(t),$$

the BS estimates the state based on a Kalman filter
and tries to schedule the sensor that is believed to have the most meaningful information

F. Chiariotti, A. E. Kalør, J. Holm, B. Soret and P. Popovski, "Scheduling of Sensor Transmissions Based on Value of Information for Summary Statistics," in IEEE Networking Letters, vol. 4, no. 2, pp. 92-96, June 2022.

looking into the query content

instead of miminizing the MSE of $\hat{\mathbf{x}}(t)$
calculate the summary statistics

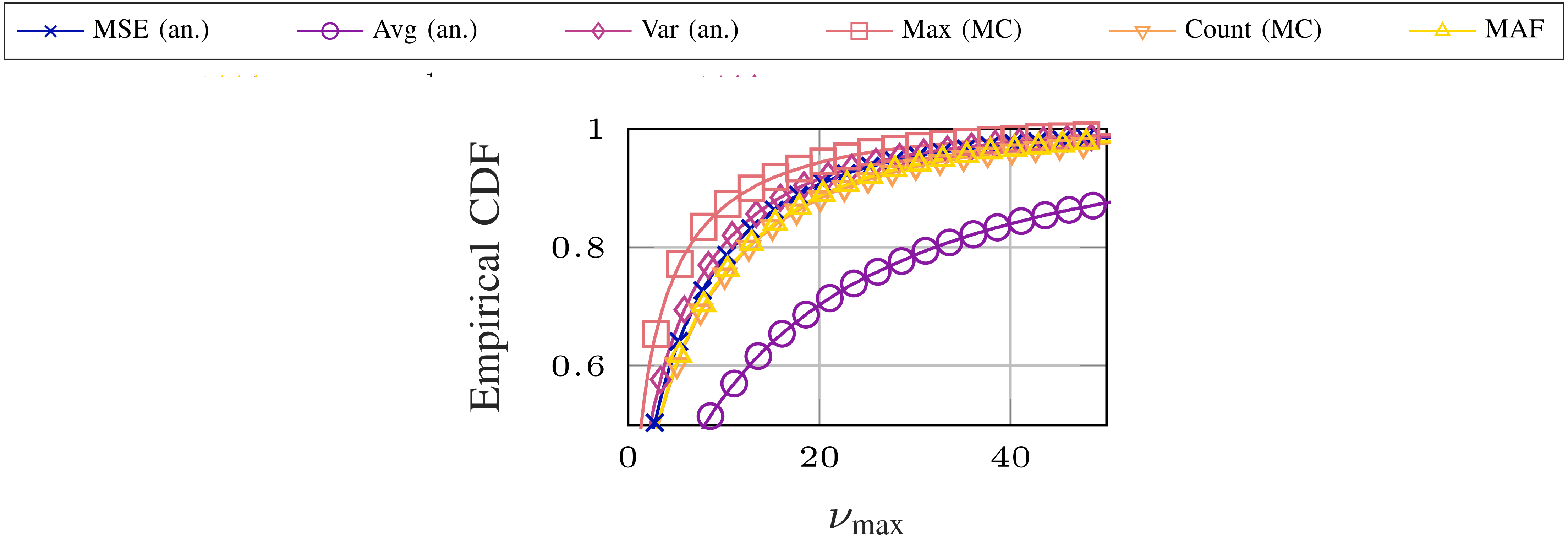
$$\hat{z}(t) = \mathbb{E}_{\mathbf{x}\sim\mathcal{N}(\hat{\mathbf{x}}(t),\boldsymbol{\psi}(t))}\left[z(\mathbf{x})\right]$$

statistics-aware Monte Carlo scheduling

- the server runs a Monte Carlo simulation on the potentially scheduled sensors
- takes into account the communication failures

F. Chiariotti, A. E. Kalør, J. Holm, B. Soret and P. Popovski, "Scheduling of Sensor Transmissions Based on Value of Information for Summary Statistics," in IEEE Networking Letters, vol. 4, no. 2, pp. 92-96, June 2022.

looking into the query content



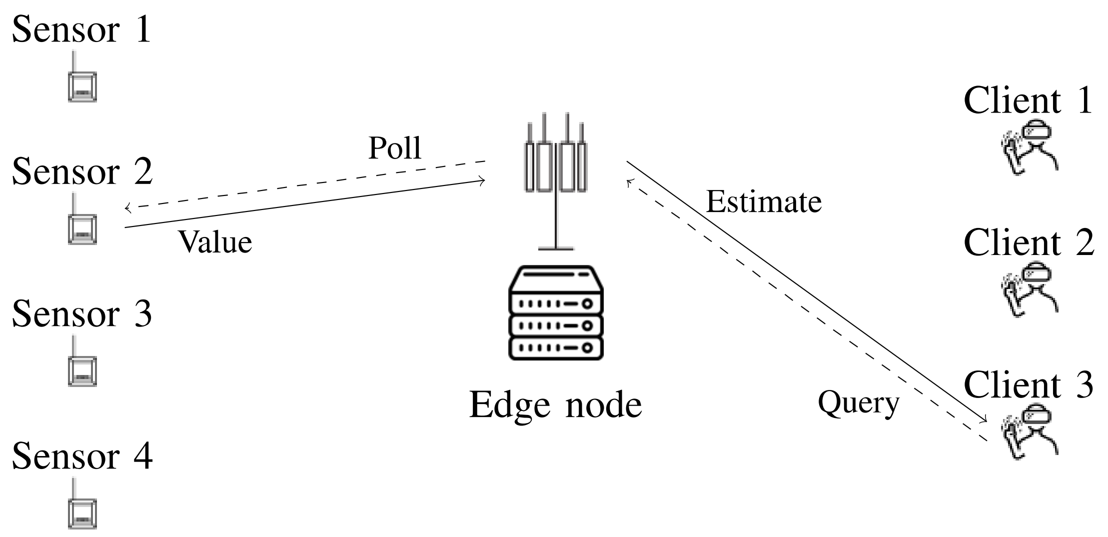
F. Chiariotti, A. E. Kalør, J. Holm, B. Soret and P. Popovski, "Scheduling of Sensor Transmissions Based on Value of Information for Summary Statistics," in IEEE Networking Letters, vol. 4, no. 2, pp. 92-96, June 2022.

generalizing to multiple queries

different clients may be interested in different queries

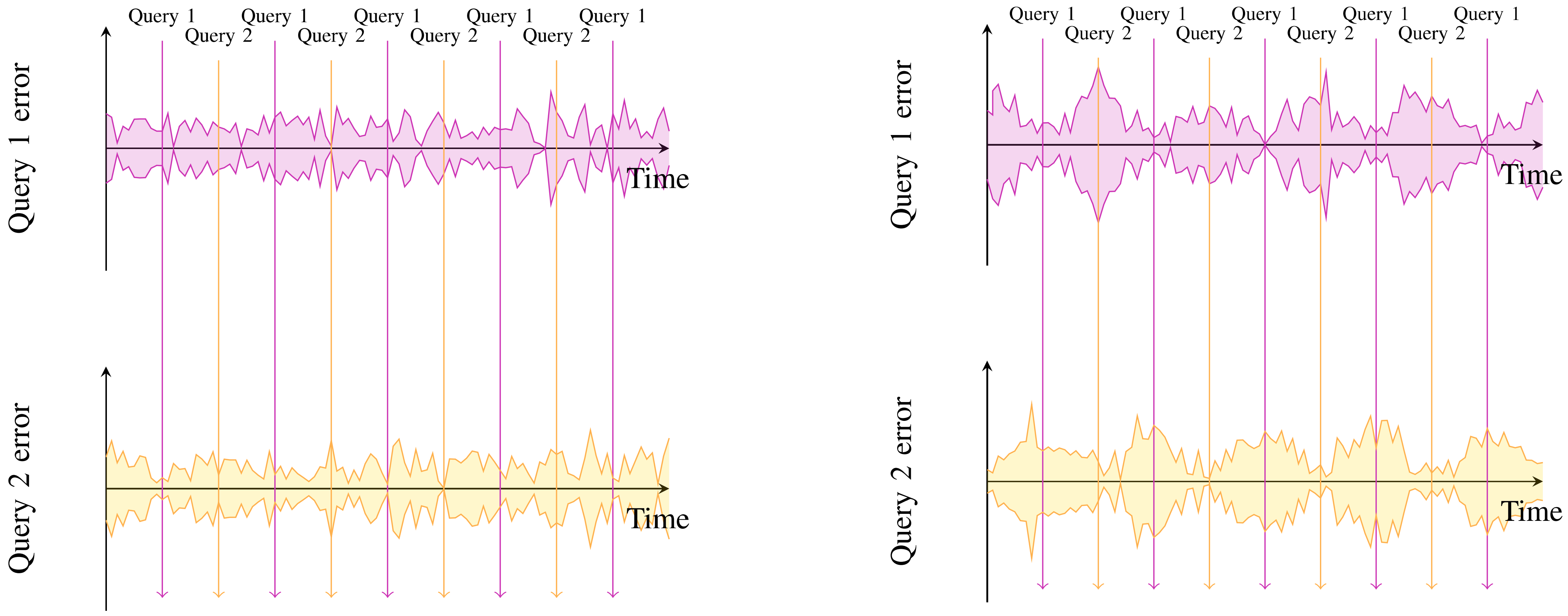
- a query is a specific function of the sensor states
- queries arrive at different times

how to collect data to be able to respond to all queries in a satisfactory manner?



J. Holm, F. Chiariotti, A. E. Kalør, B. Soret, T. B. Pedersen and P. Popovski, "Goal-Oriented Scheduling in Sensor Networks With Application Timing Awareness," in IEEE Transactions on Communications, vol. 71, no. 8, pp. 4513-4527, Aug. 2023, doi: 10.1109/TCOMM.2023.3282256.

pragmatic sensor scheduling



(a) VoI-based sensor scheduling.

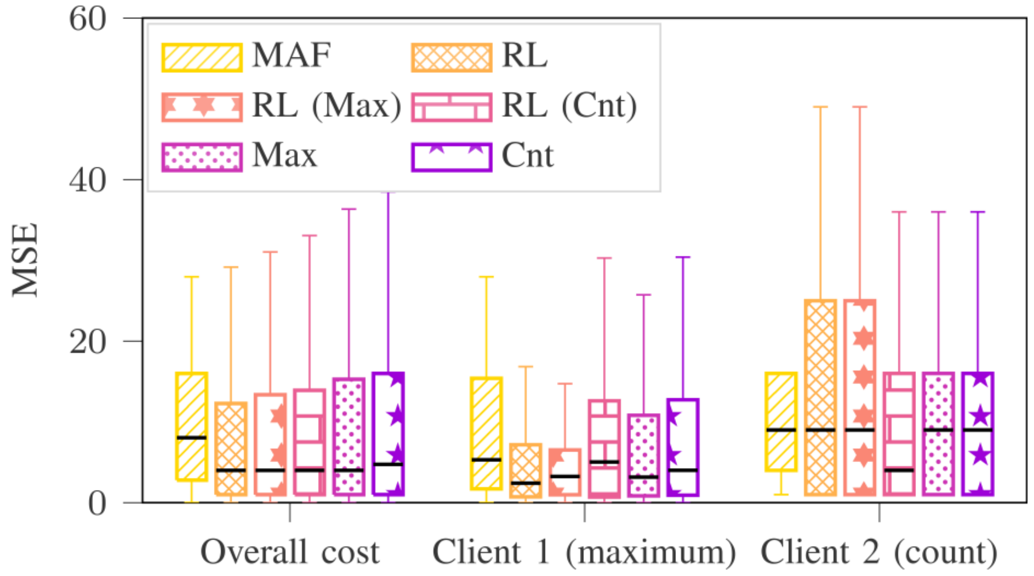
(b) Pragmatic sensor scheduling.

J. Holm, F. Chiariotti, A. E. Kalør, B. Soret, T. B. Pedersen and P. Popovski, "Goal-Oriented Scheduling in Sensor Networks With Application Timing Awareness," in IEEE Transactions on Communications, vol. 71, no. 8, pp. 4513-4527, Aug. 2023, doi: 10.1109/TCOMM.2023.3282256.

pragmatic sensor scheduling

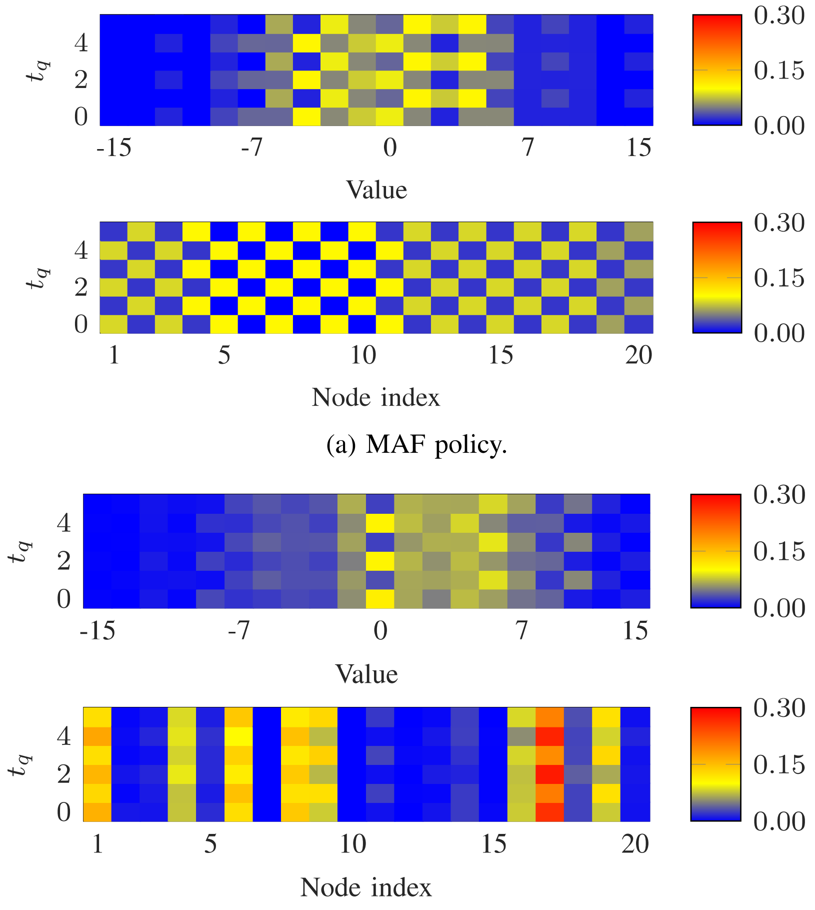
problem defined as a POMDP

- at each time slot decide from which sensor to pull information
- solved by reinforcement learning
- results for M=20 sensors
- C=2 clients
    - client 1 asks for count query
    - client 2 asks for max query



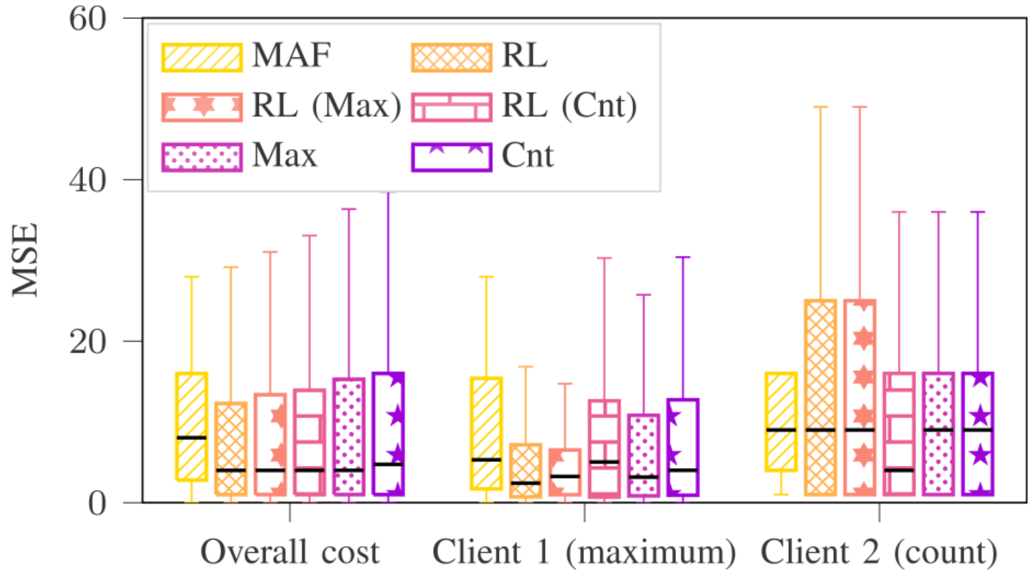
J. Holm, F. Chiariotti, A. E. Kalør, B. Soret, T. B. Pedersen and P. Popovski, "Goal-Oriented Scheduling in Sensor Networks With Application Timing Awareness," in IEEE Transactions on Communications, vol. 71, no. 8, pp. 4513-4527, Aug. 2023, doi: 10.1109/TCOMM.2023.3282256.

pragmatic sensor scheduling



(a) MAF policy.

(b) RL policy.



outline

towards 6G

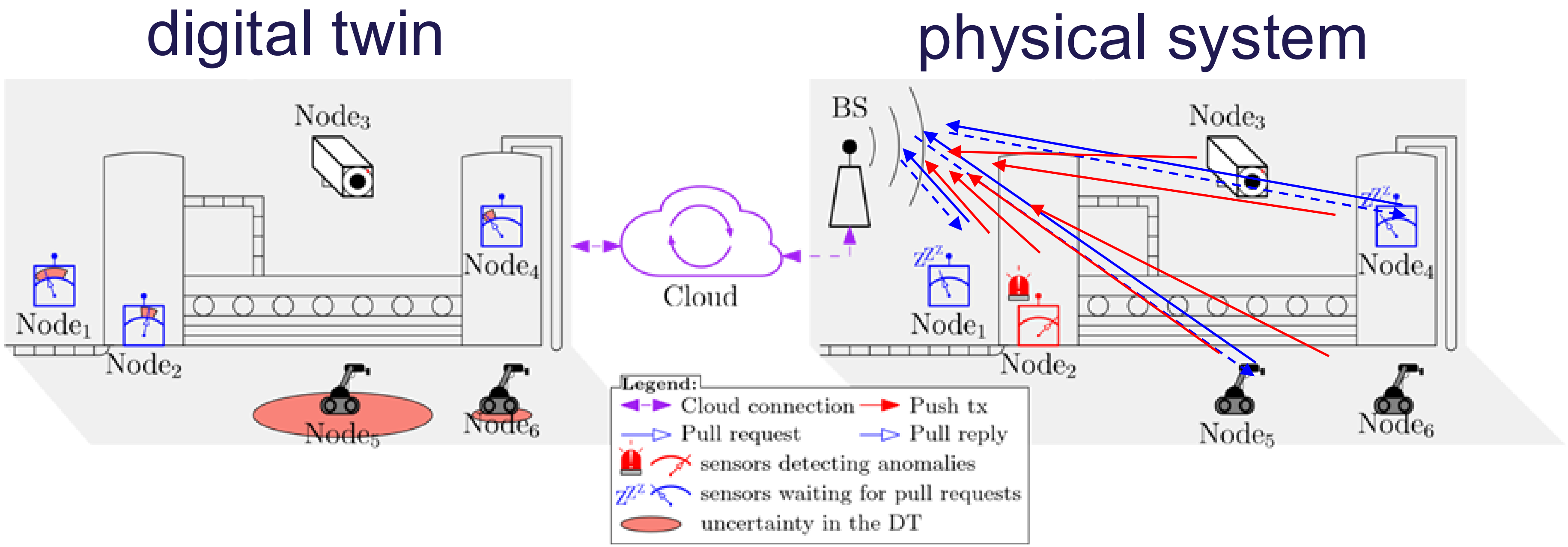
modeling latency

beyond latency

time and intelligence

physical vs. digital time,
causality, and simultaneity

AI-driven wireless of sensors and actuators



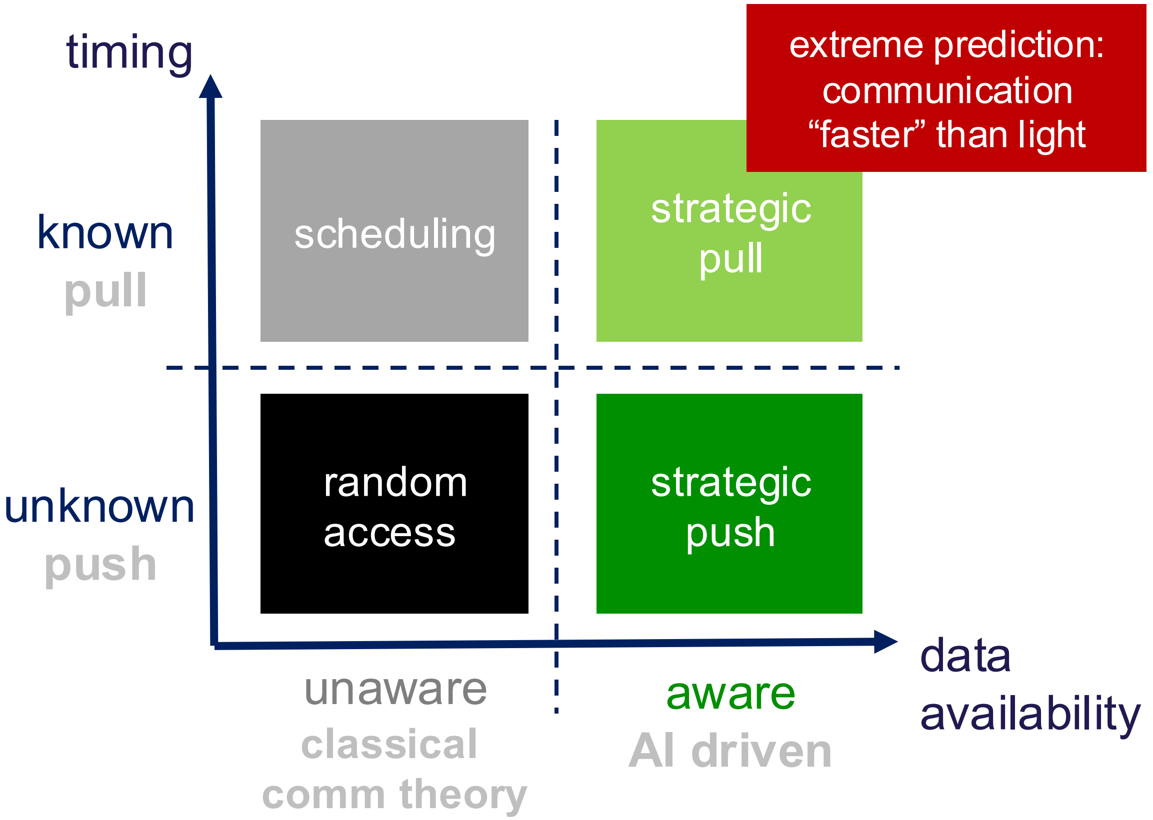
how to run the digital twin updated in real time?

- push communication by nodes that transmit at their own will
- pull communication per BS scheduling

push-pull coexistence over a shared wireless medium

- in **classical networks**,
  push and pull distinction given by *certainty of data collection timing*

- in **AI-driven networks**,
  push and pull distinction depends on the knowledge about the *data availability* (meaning and value)



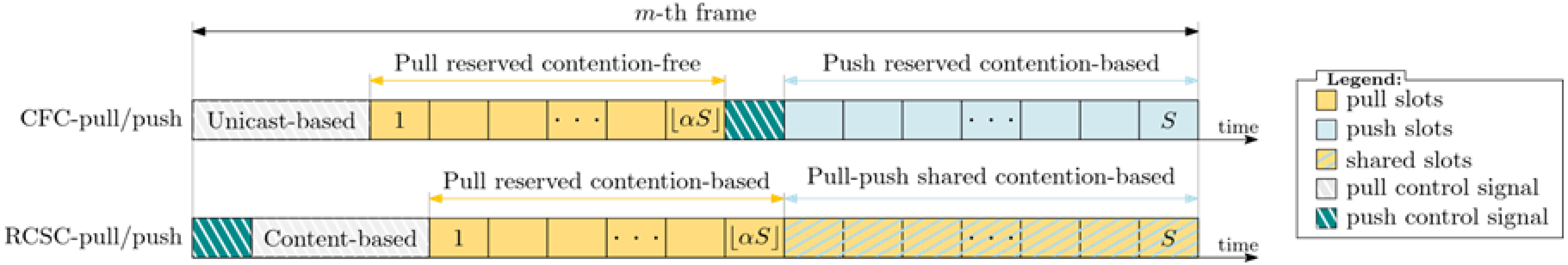
some options for MAC design

Contention-free pull and contention-push (CFC-pull/push)

- ***pull unicast scheduling:***
  retrieve data from *specific devices at specific times*
- ***push contention:***
  devices transmit *locally relevant* data by contention

Reserved pull-contention and shared pull-push contention (RCSC-pull/push)

- ***pull content-based scheduling:***
  semantic query pulls data from unknown set of relevant devices
- ***pull–push contention:***
  devices transmit *locally relevant data* by contending for the channel with all the devices

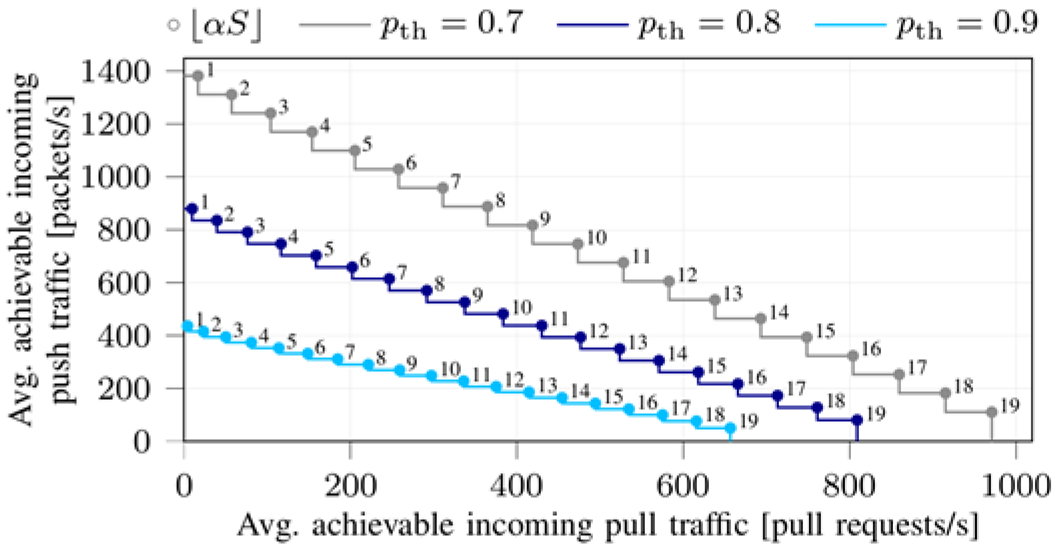


challenge

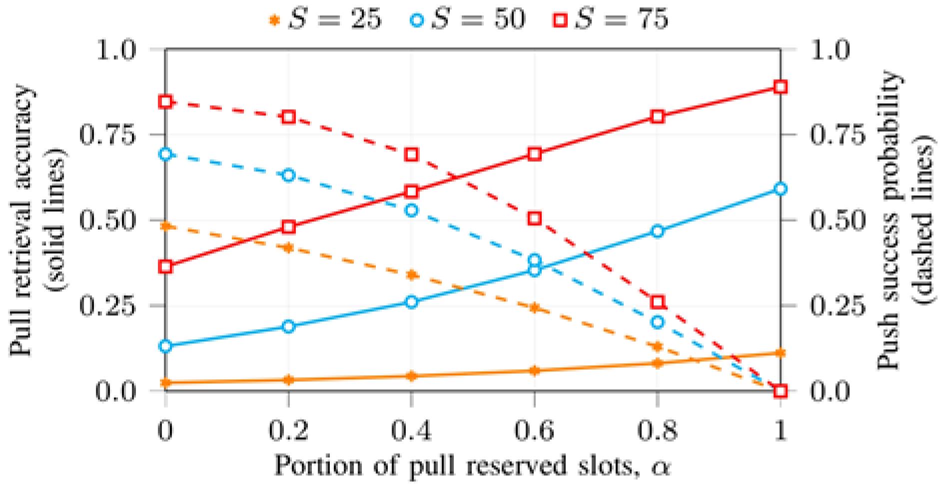
managing the available resources to maximizes the goal-oriented utility for both traffic

performance tradeoffs

Contention-free pull and contention-push (CFC-pull/push)



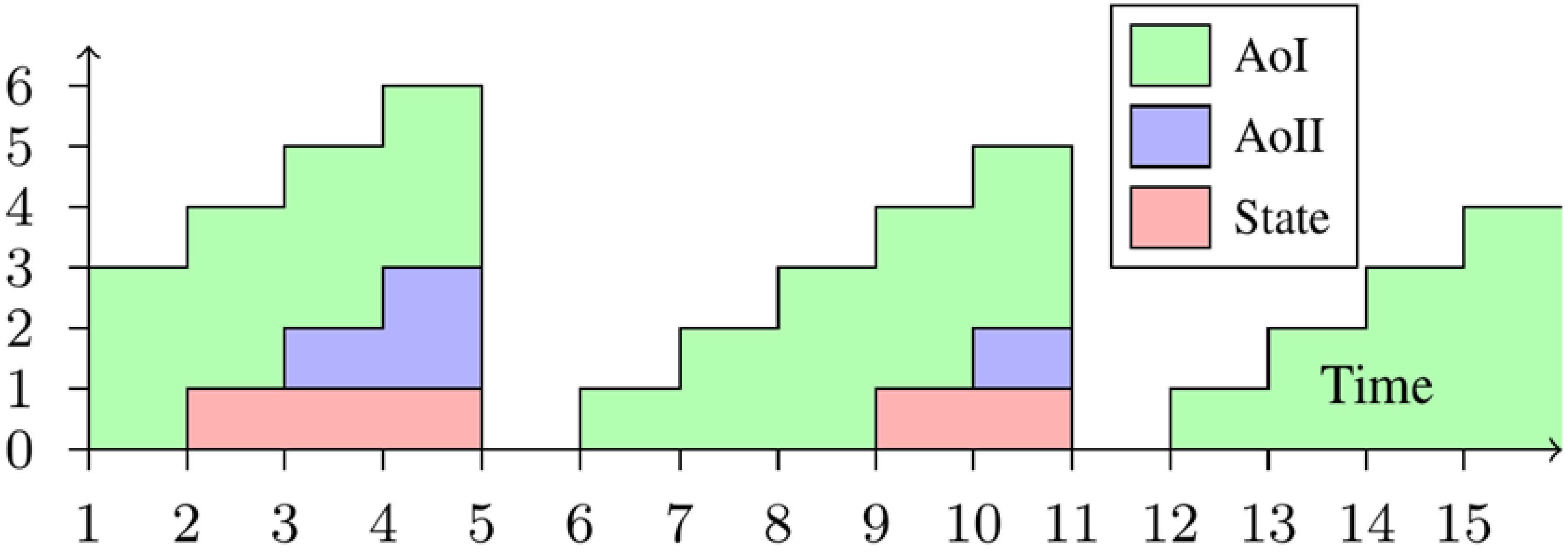
Reserved pull-contention and shared pull-push contention (RCSC-pull/push)



Age of Incorrect Information (AoII)

anomaly reporting is suitable for push-based communication:

- if anomalies are rare, being proactive is good and round robin-like methods can be highly suboptimal



F. Chiariotti, A. Munari, L. Badia, and P. Popovski, "Distributed Optimization of Age of Incorrect Information with Dynamic Epistemic Logic", in Proc. IEEE INFOCOM, London, UK, May 2025.

exploiting common knowledge

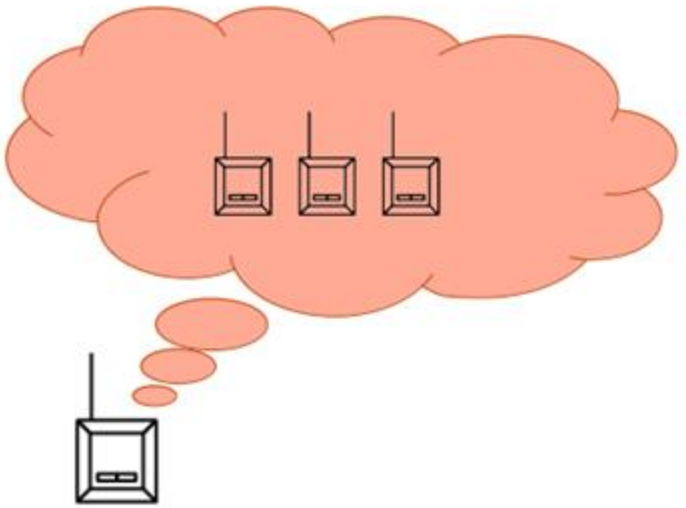
acknowledgments represent public announcements:

- ACK: someone transmitted successfully
- NACK: someone had something to transmit, but the transmission failed

silence: no one decided to transmit

design a protocol based on

- common knowledge (public information)
- private knowledge (sensor state and local AoII)
- inferred AoII of the other nodes

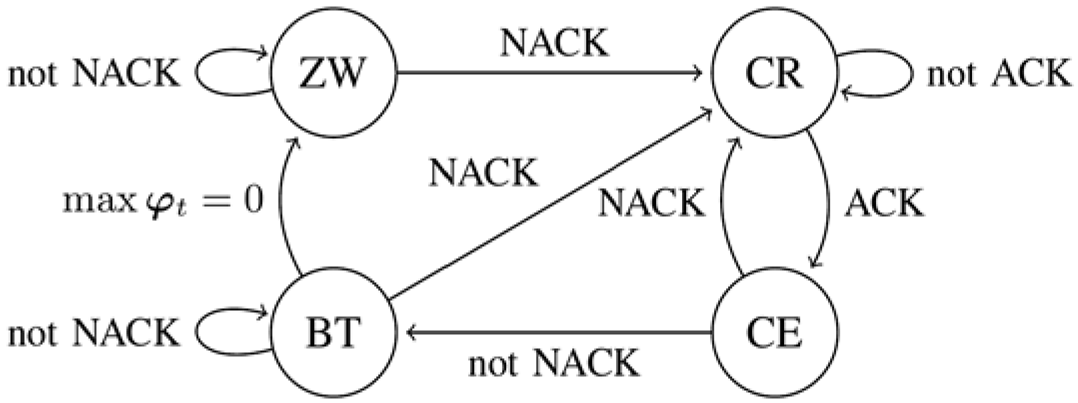


F. Chiariotti, A. Munari, L. Badia, and P. Popovski, "Distributed Optimization of Age of Incorrect Information with Dynamic Epistemic Logic", in Proc. IEEE INFOCOM, London, UK, May 2025.

DELTA protocol

Dynamic Epistemic Logic for Tracking Anomalies (DELTA)

1. **Zero Wait**: everything is fine
2. **Collision Resolution**: we need the collided nodes to transmit
3. **Collision Exit**: we need to make sure no collided nodes are left
4. **Belief Threshold**: we need to figure out a way to reduce the backlog

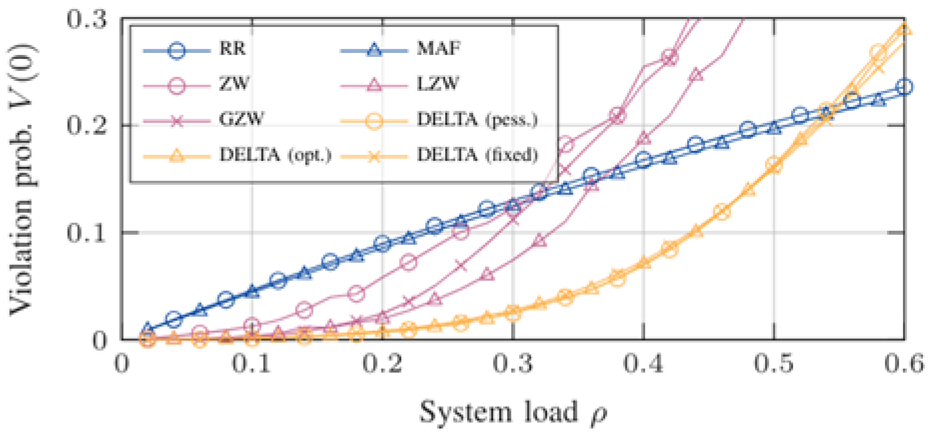


F. Chiariotti, A. Munari, L. Badia, and P. Popovski, "Distributed Optimization of Age of Incorrect Information with Dynamic Epistemic Logic", in Proc. IEEE INFOCOM, London, UK, May 2025.

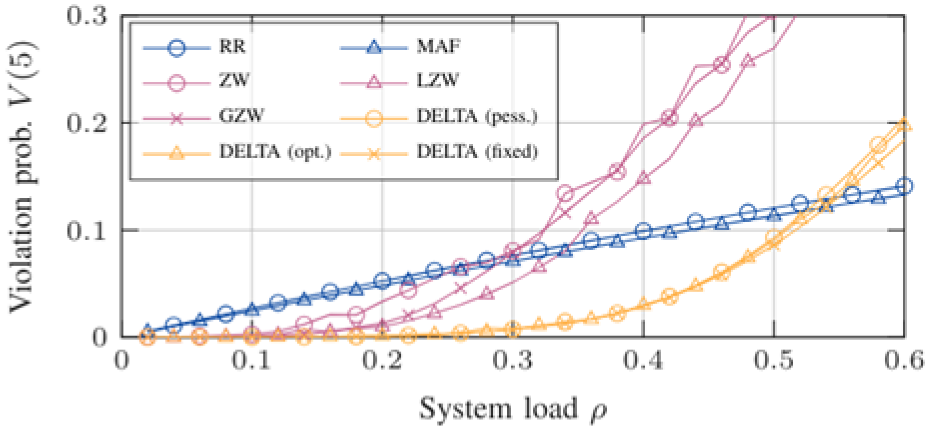
AoII threshold Θ_{max} violation

DELTA significantly outperforms ALOHA-like random access protocols and round robin-like scheduled schemes

deals with a relatively high load (up to 50%) with 20 nodes



(a) AoII violation probability ($\Theta_{max} = 0$).



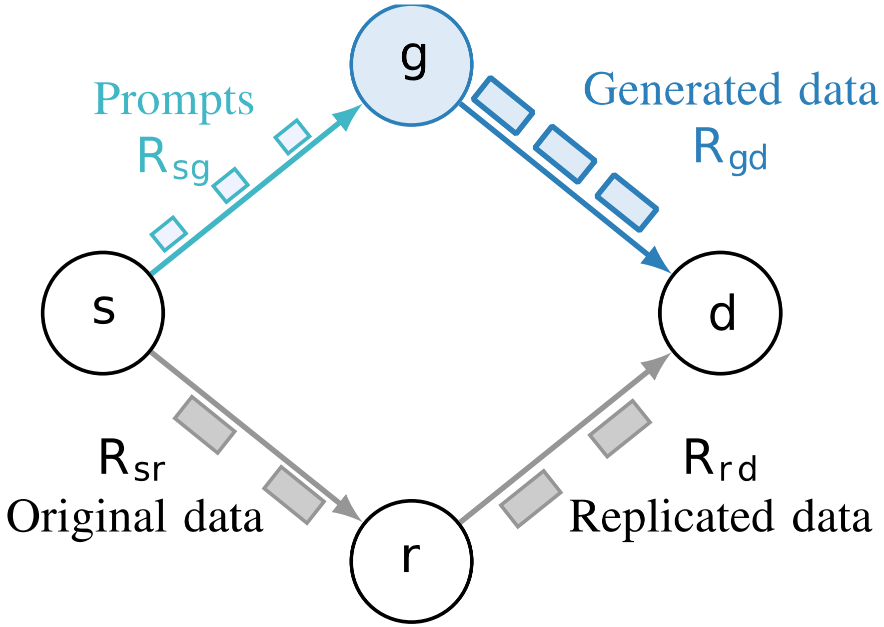
(b) AoII violation probability ($\Theta_{max} = 5$).

F. Chiariotti, A. Munari, L. Badia, and P. Popovski, "Distributed Optimization of Age of Incorrect Information with Dynamic Epistemic Logic", in Proc. IEEE INFOCOM, London, UK, May 2025.

latency and generative AI

traditional communication systems:

- destination needs a packet replica
- maximum sending rate given by min-cut of network
- caching can help but all data is still transmitted



communication systems with generative AI:

- prompts help alleviate congestion in weak links
- destination receives approximations of data

**if it can be predicted,
does not need to be communicated**

M. Thorsager, I. Leyva-Mayorga, B. Soret, and P. Popovski, "Generative Network Layer for Communication Systems with Artificial Intelligence", in IEEE Networking Letters, accepted, 2024.

outline

towards 6G

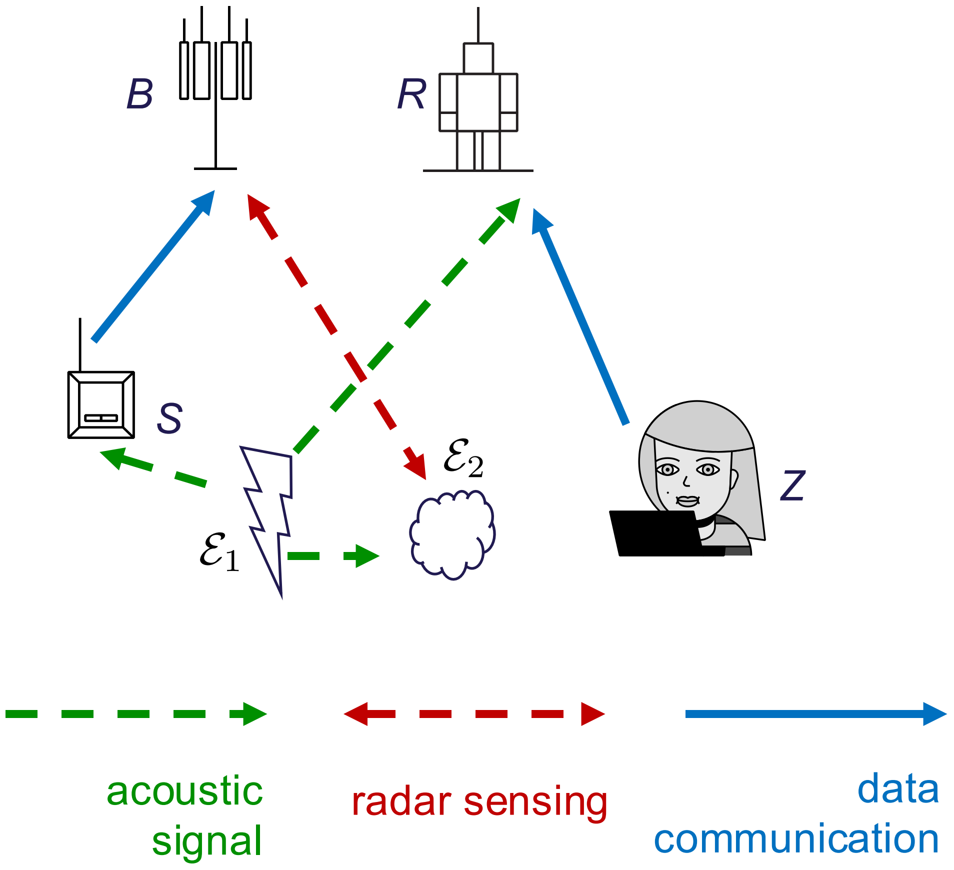
modeling latency

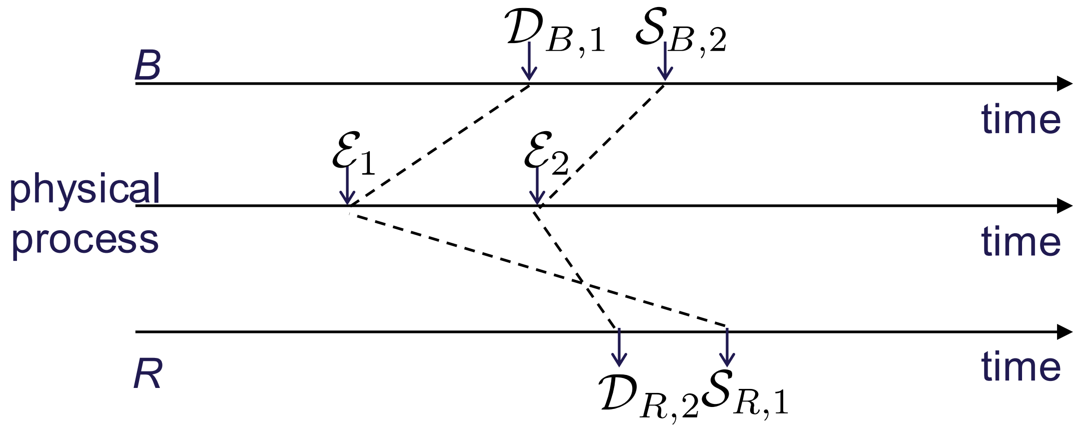
beyond latency

time and intelligence

physical vs. digital time, causality, and simultaneity

time in fused physical-digital world





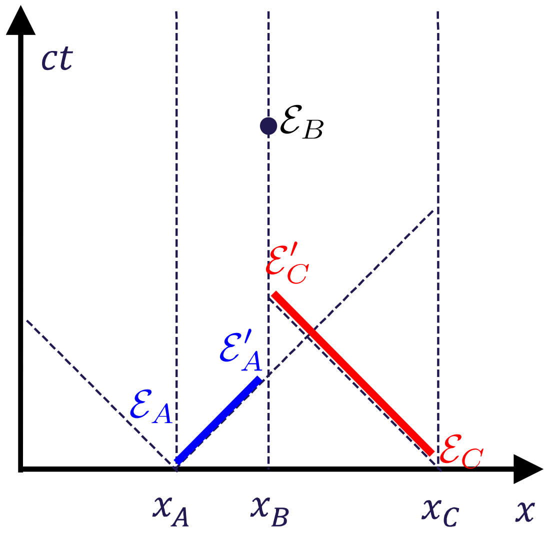
digital and physical time get intertwined

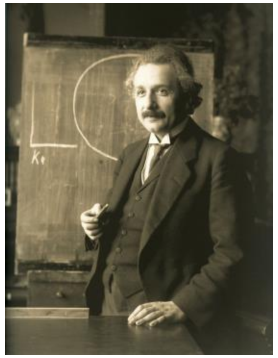
chronology depends on the observer and data processing

physical vs. digital time

physical time and its relativity

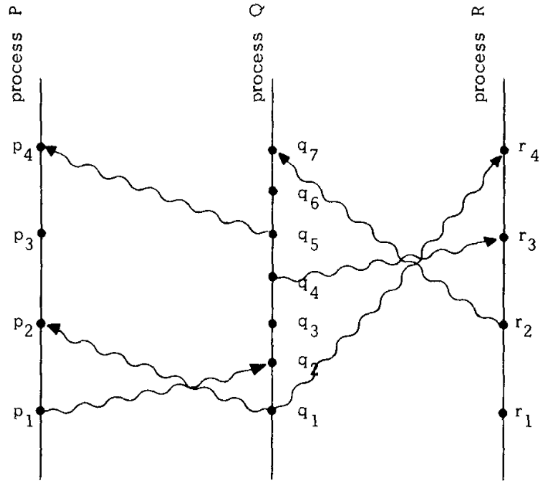
A. Einstein, “On the electrodynamics of moving bodies,” Annalen der Physik, vol. 17, no. 10, pp. 891–921, 1905.

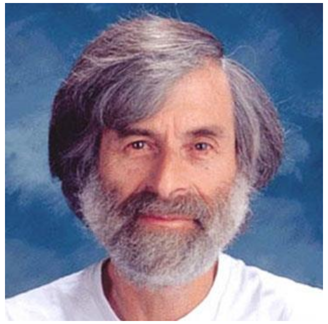




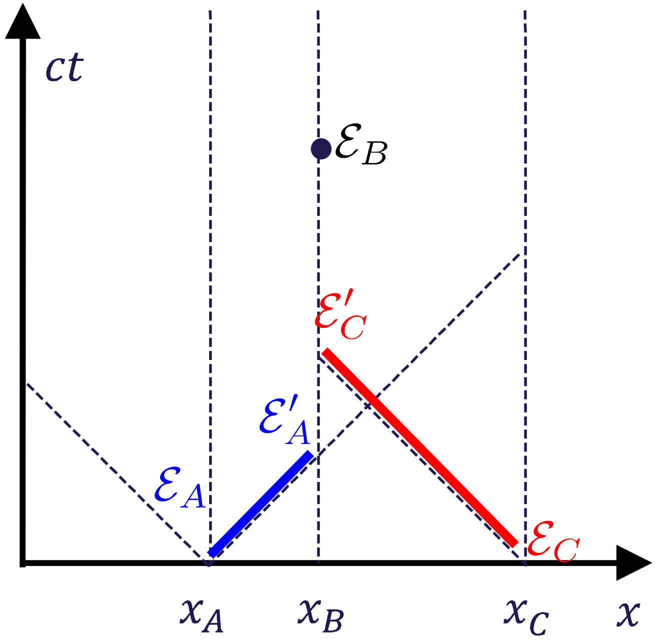
digital time in computer systems

L. Lamport, “Time, clocks, and the ordering of events in a distributed system,” Commun. ACM, vol. 21, no. 7, pp. 558–565, Jul. 1978.

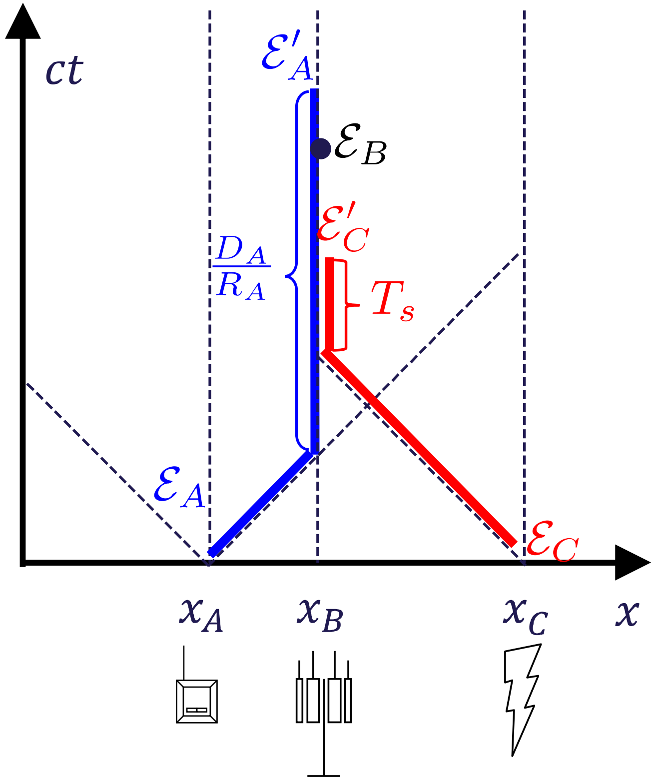




time and information processing



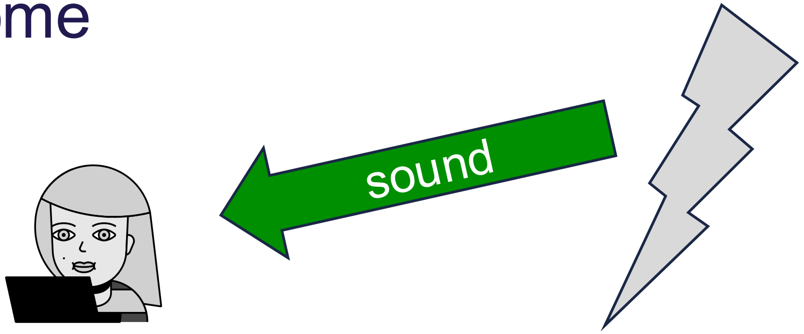
light cones and causality
in special relativity
with instantaneous
information processing



light cones and causality
with the effects of
information processing

how do we make sense of a spacetime?

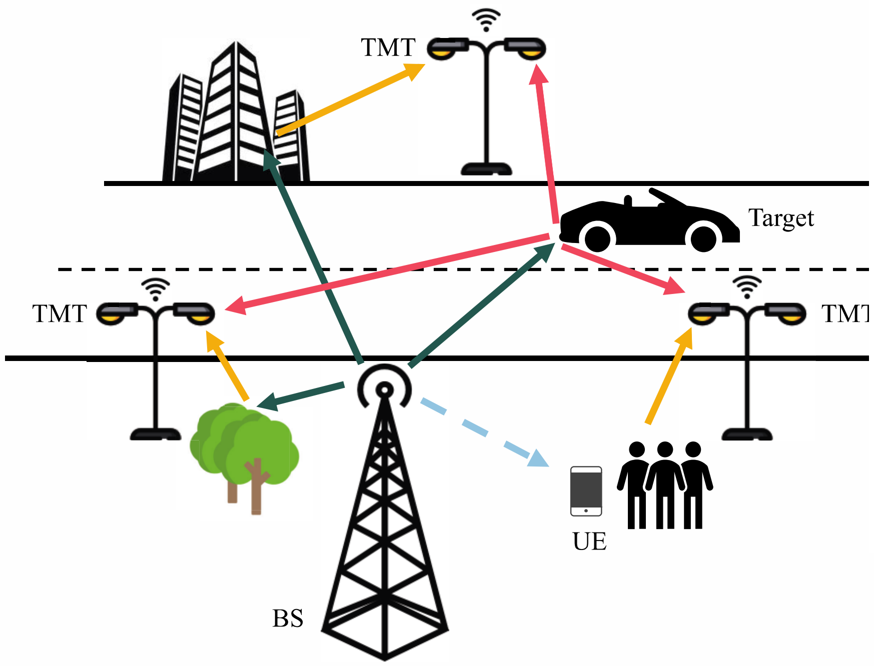
- auditory and visual stimuli arrive at different speeds
- the brain plays them out after synhronizing their outputs
- now is defined by this processing outcome
- simultaneity determined by what was evolutionary important



P. Popovski, F. Chiariotti, K. Huang, A. E. Kalør, M. Kountouris, N. Pappas, and B. Soret, “A Perspective on Time towards Wireless 6G,” in Proceedings of the IEEE, vol. 110, no. 8, pp. 1116-1146, Aug. 2022

towards perceptive mobile networks

L. Xie, S. Song, Y. C. Eldar, and K. B. Letaief, “Collaborative sensing in perceptive mobile networks: Opportunities and challenges,” IEEE Wireless Commun., vol. 30, no. 1, pp. 16–23, Feb. 2023.



Time, Simultaneity, and Causality in Wireless Networks With Sensing and Communications

PETAR POPOVSKI (Fellow, IEEE)

Department of Electronic Systems, Aalborg University, 9220 Aalborg, Denmark

CORRESPONDING AUTHOR: P. POPOVSKI (e-mail: petarp@es.aau.dk)

The work was supported by the Villum Investigator Grant “WATER” from the Velux Foundation, Denmark.

ABSTRACT Wireless systems beyond 5G evolve towards embracing both sensing and communication, resulting in increased convergence of the digital and the physical world. The existence of fused digital-physical realms raises critical questions regarding temporal ordering, causality, and the synchronization of events. This paper addresses the temporal challenges arising from the fact that the wireless infrastructure becomes an entity with multisensory perception. With the growing reliance on real-time interactions and applications such as digital twins, extended reality, and the metaverse, the need for accurate timestamping and temporal forensics becomes crucial. The paper introduces a model that incorporates Temporal Window of Integration (TWI) to emulate human multisensory perception and discusses the implications for setting timing constraints in real-time applications and enabling temporal forensics. The analysis explores trade-offs, probabilities, and bounds for simultaneity and causality violation in the context of wireless systems evolving towards perceptive networks. This work underscores the significance of timestamping in the evolving wireless landscape, provide insights into system-level implications, and points out new research avenues for systems that combine sensing and communications.

temporal window of integration (TWI) 1/2

inspired by multisensory perception

Current Biology, Vol. 13, R519–R521, July 1, 2003, DOI 10.1016/S0960-9822(03)00445-7

Multisensory Integration: Maintaining the Perception of Synchrony

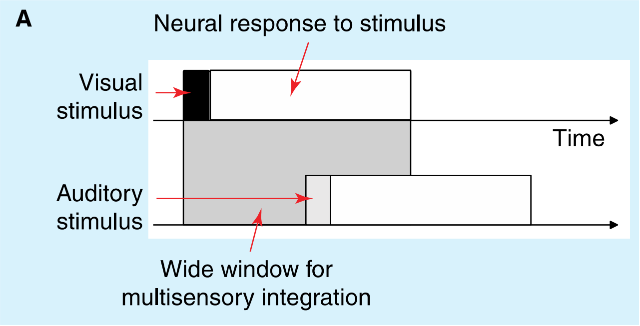
Dispatch

Charles Spence and Sarah Squire

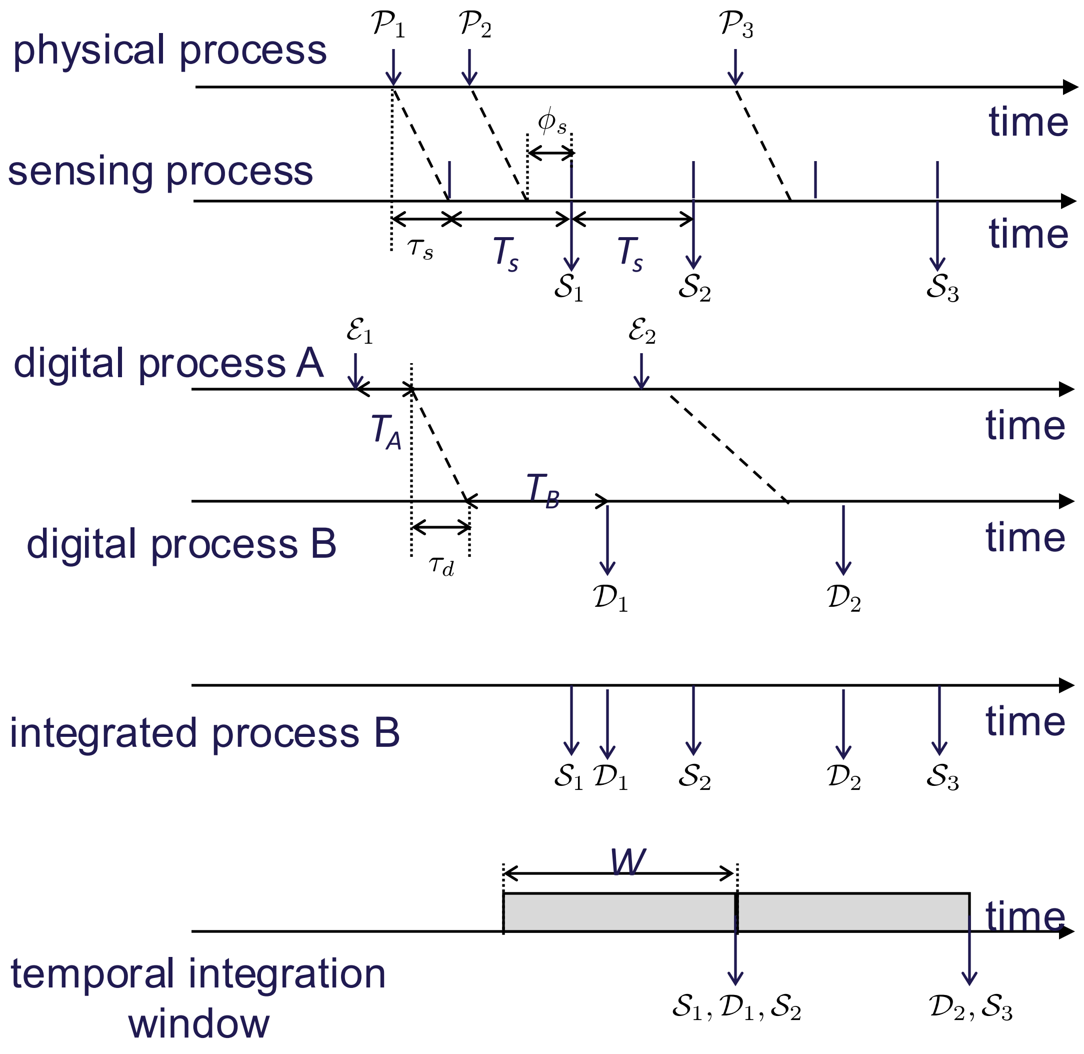
We are rarely aware of differences in the arrival time of inputs to each of our senses. New research suggests that this is explained by a ‘moveable window’ for multisensory integration and by a ‘temporal ventriloquism’ effect.

We are all familiar with the experience of seeing lightning before hearing the associated thunder. This discrepancy in our perception of a synchronous, albeit distant, multisensory event is caused by physical differences in the relative time of arrival of stimuli at the eye and ear. Light travels through air far more rapidly than sound: 300,000,000 *versus* 330 metres per second.

us, in order to accommodate the fact that sound will increasingly lag behind vision with increasing distance (Figure 1b). The existence of a moveable window for multisensory simultaneity, to accommodate the typical delay in auditory arrival times, is consistent with previous studies showing that people (and animals) are more



temporal window of integration (TWI) 2/2



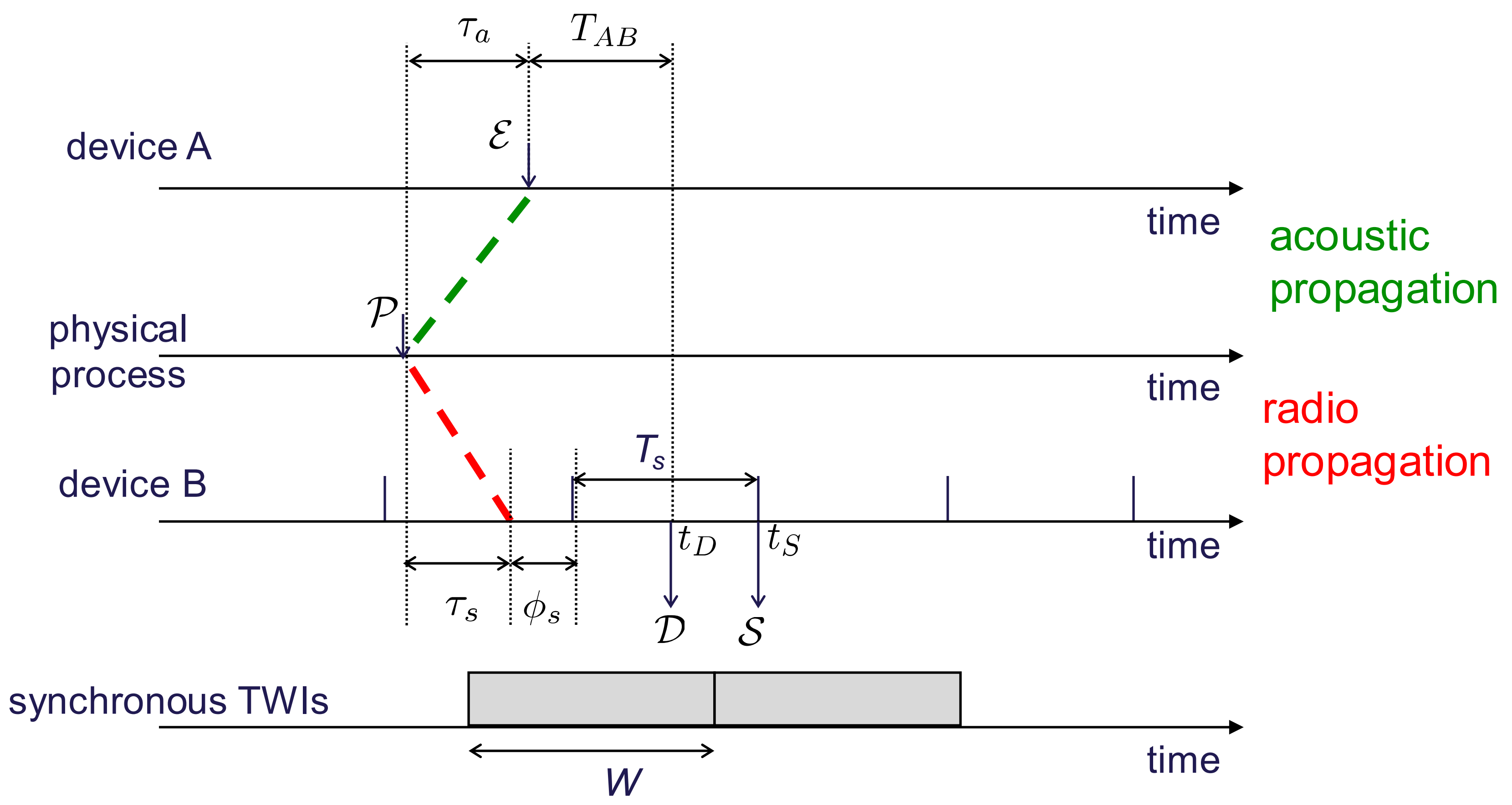
synchronous sensor

communication data

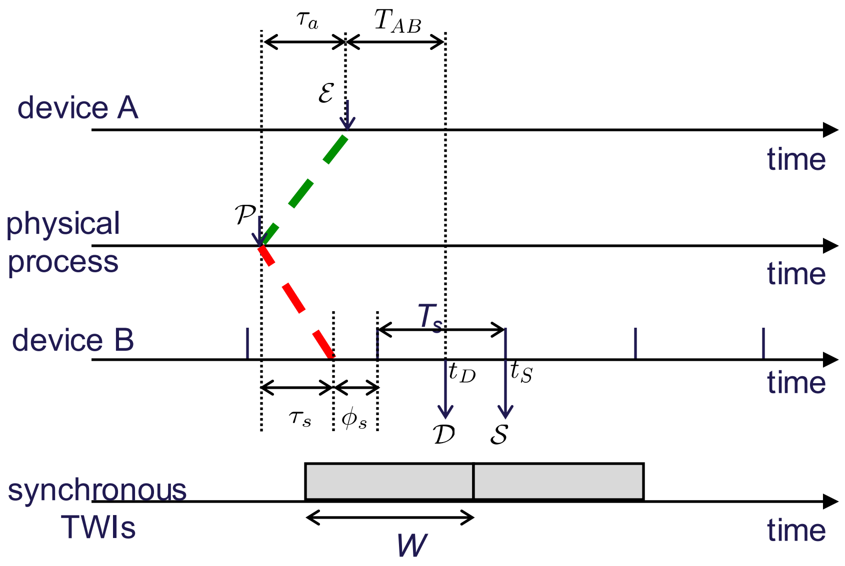
integrated process

with temporal window of integration

causality violation: physical causes digital



causality violation: physical causes digital



how to choose the window size W?

$$W_{\min} \geq 2T_s + \tau_{s,\max} - T_{\min} - \tau_{a,\min}$$

if device A uses prediction and/or generative AI, then $\tau_{a,\min}$ can be negative.

larger TWI leads to lower probability of causality violation, but it also decreases the event throughput.

we have seen similar dimensioning for guard interval in OFDM.

timestamping in future base stations

base stations as the guardians of the intertwined physical and digital time

- setting multi-input constraints for real-time applications
  - orchestrated timing performance of communication and sensing
- temporal forensics
  - certifying the order of events in the wake of powerful AI, preventing manipulation of causality and chronology

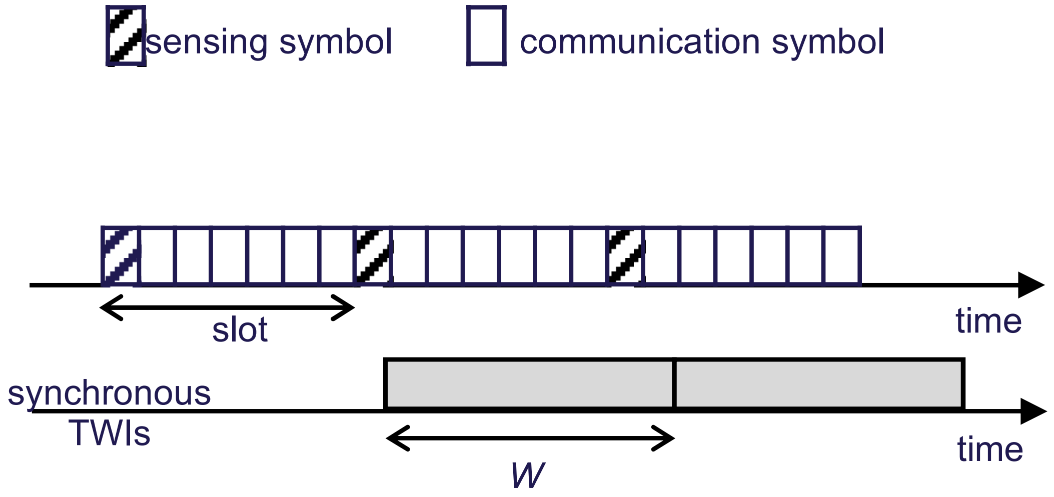
what's in for the base stations?

setting the latency/timing requirements for a service

- the application choses a temporal window based on the goal-oriented fusion/inference process
- fundamental tradeoff latency-energy-accuracy
- timestamp the inputs for inference further in the network

multiuser systems

- redefine resource allocation based on heterogeneous temporal windows
- allocate sensing resources for correct timestamping



intelligence for causality and simultaneity

- brain uses deeply rooted knowledge to compensate for external factors
  - causal inference at edge-cloud
- brain might be flexible and change its temporal window
  - applications with varying levels of criticality
- brain actively shift the time to compensate for gaps
  - use generative AI to replace data that has not arrived

J. Vroomen and M. Keetels, “Perception of intersensory synchrony: A tutorial review,”
Attent. Percept. Psychophys., vol. 72, no. 4, pp. 871–884, 2010.

AI and time

IEEE Spectrum | How AI's Sense of Time Will Differ From O... | Type to search

ANALYSIS | AI

How AI's Sense of Time Will Differ From Ours › Machines may see cause and effect in problematic new ways

BY PETAR POPOVSKI | 13 AUG 2025 | 5 MIN READ

Petar Popovski is a professor of communications engineering at Aalborg University, Denmark.

conclusion and outlook

- 6G: on a route beyond communications
  - fusion of physical and digital
  - increased role of intelligence
- timing performance depends on multiple components
- latency just an instance of more general timing measures
  - interplay with intelligence
- intertwining of physical and digital time,
  flow of time depends on information processing
  - base stations will need to act as trustworthy multisensory perceptors
  - temporal integration window for controlling timestamping,
    as well as (non)violation of causality and simultaneity

CLASSIQUE

CENTER OF EXCELLENCE
CLASSICAL COMMUNICATION IN THE QUANTUM ERA