CPS: TTP Option: Small: Consistency vs. Availability in Cyber-Physical Systems

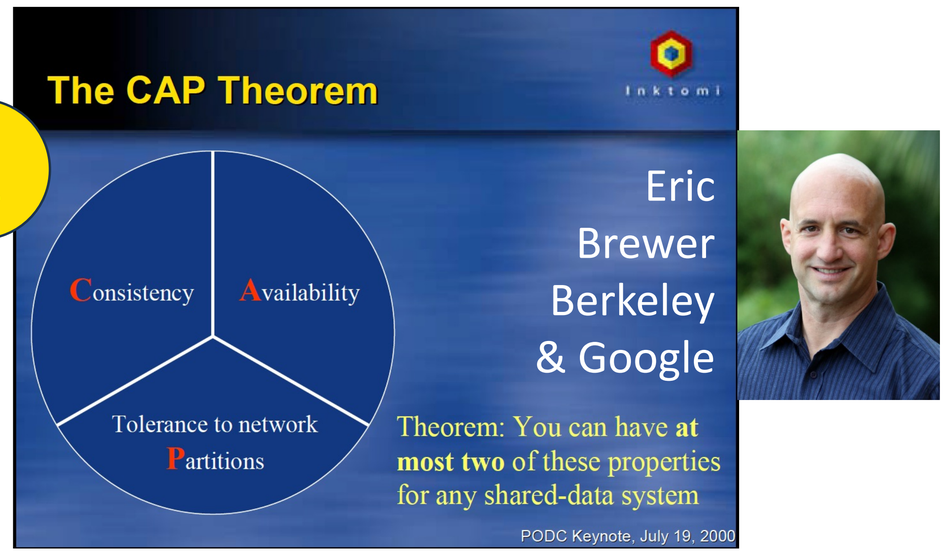
Edward A. Lee, UC Berkeley

http://lf-lang.org

This project is addressing the fundamentals of CPS, specifically, the fundamental limits in applications where the software components are distributed across networks or across parallel computers or cores. These applications include manufacturing, transportation systems, vehicular automation, medical devices, and many others. The goal is to develop a "system theory" for CPS, a rigorous mathematical framework that can be used to guide design decisions that are today ad hoc.

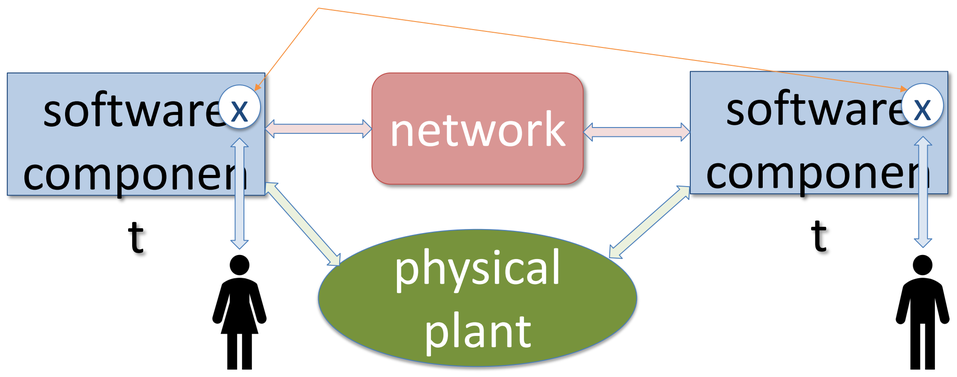
Background:

1



2

Consistency: agreement on the values of shared variables and the state of the world



Availability: ability to respond to reads and writes accessing those state variables

5

(In)Consistency

$$\bar{C}_{ij} = \max(\mathcal{T}(g_i) - \mathcal{T}(g_j))$$

- $\bar{C}_{ij}$: Inconsistency from node *j* to node *i*.
- g_i: Max over all updates on *j* that are sent to *i*.
- $\mathcal{T}(g_i)$: Logical time of notification of the update at node *i*.
- $\mathcal{T}(g_j)$: Logical time of an update of a shared variable on node *j*.

Constraint: $\bar{C}_{ij} \geq 0$

Strong consistency: $\bar{C}_{ij} = 0$

(Un)Availability

$$\bar{A}_i = \max(T_i - \mathcal{T}(g_i))$$

- $\bar{A}_i$: Un-availability at node *i*
- g_i: Max over all accesses of the shared variable at *i*.
- T_i: Physical time (on a local clock) of the access of the updated shared variable at *i*.
- $\mathcal{T}(g_i)$: Logical time of the update of a shared variable on node *i*.

Constraint: $\bar{A}_i \geq 0$

Perfect availability: $\bar{A}_i = 0$

3

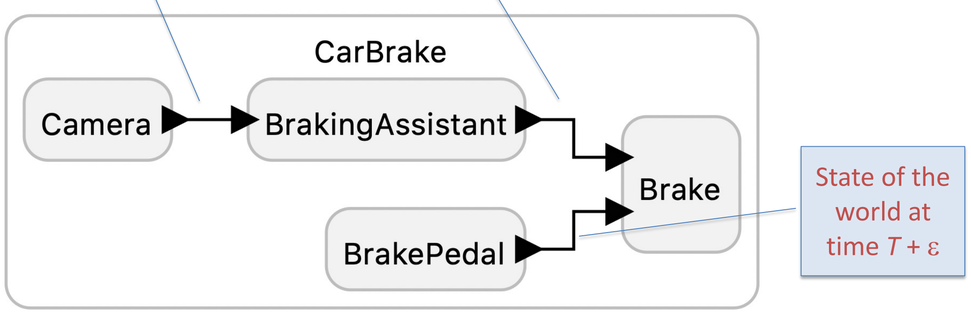
Availability or Consistency?

Denso autonomous braking demonstrating Advanced Driver-Assistance System (ADAS) in Oct. 2018 [Reported in The Daily Times]

Snapshot at time *T*

State of the world at time *T*

A software architecture:



6

Processing Offsets: $O_i = \max(T_i - \mathcal{T}(g_i))$

Apparent Latency:

$$\mathcal{L}_{ij} = \max(T_i - \mathcal{T}(g_j))$$

- $\mathcal{L}_{ij}$: Apparent latency for communication from *j* to *i*.
- g_j: Max over all messages from *j* to *i*.
- T_i: Physical time (on a local clock) of the received update message at *i*.
- $\mathcal{T}(g_j)$: Logical time of the original update at node *j*.

$$\mathcal{L}_{ij} = O_j + X_{ij} + L_{ij} + E_{ij}$$

- O_j: Processing offset at *j*.
- X_{ij}: Execution time overhead on path from *j* to *i*.
- L_{ij}: Network latency from *j* to *i*.
- E_{ij}: Clock synchro-nization error

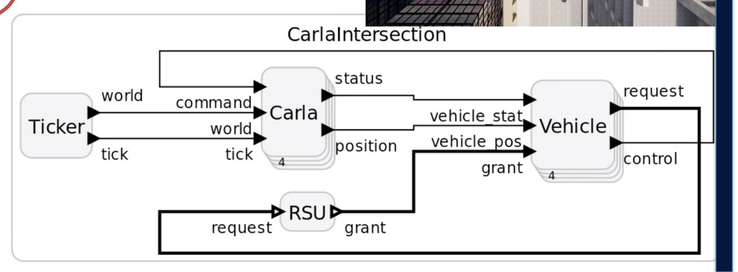
Note that this quantity is easily measured, unlike its components.

4

Availability or Consistency?

Consistency: agreement on the state of the intersection.

Availability: ability to enter the intersection.



Cal Theorem:

7

Theorem 1: Given a trace, the unavailability at process i is, in the worst case,

$$\bar{A}_i = \max\left(O_i, \max_{j \in N}(\mathcal{L}_{ij} - \bar{C}_{ij})\right), \qquad (6)$$

where O_i is the processing offset, $\mathcal{L}_{ij}$ is apparent latency (which includes O_j), and $\bar{C}_{ij}$ is the inconsistency.

In Max-Plus Algebra:

$$\boldsymbol{A} = (\boldsymbol{I} \oplus \Gamma)\boldsymbol{O}$$

LINGUA FRANCA

Award ID#: CNS-2233769