Google Research

Modeling and Control of Bittide Synchronization

Sanjay Lall, Google Research and Stanford University

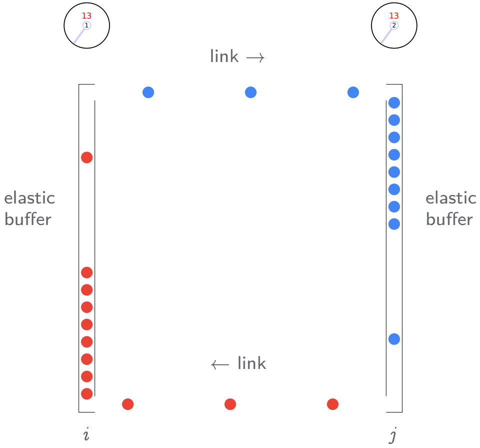
Călin Caşcaval, Martin Izzard, and Tammo Spalink, Google Research

Overview

- *Google research project*
  - originated at Princeton: Spalink 2006.
  - broad scope: applications, scheduling, simulation, hardware, theory
- *paper*
  - problem formulation
  - model well-posedness
  - simulation algorithm
- *outline:* the mechanism, logical synchrony, controlling frequency

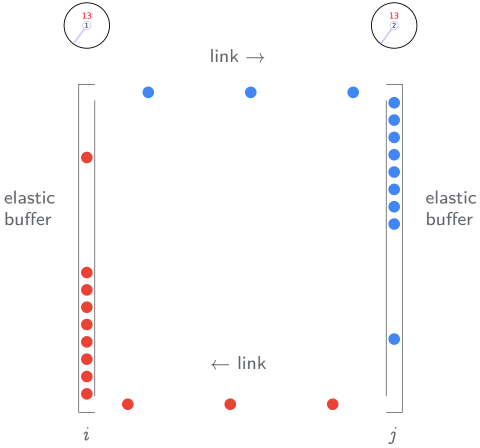
Overview

- at each node i
  - clock
  - each incoming link has a queue called the *elastic buffer*
- at each node, with each clock tick
  - a frame is removed from all elastic buffers
  - a frame is sent on all outgoing links

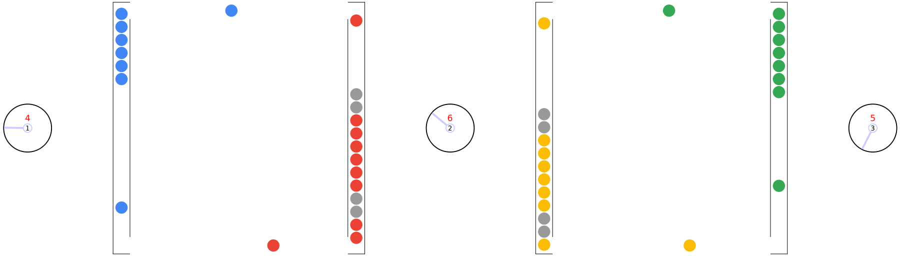


Mechanism

- if oscillator at node j is faster than that at i
  - j's elastic buffer will drain
  - i's elastic buffer will fill
- nodes observe elastic buffers, adjust frequency



Logical Synchrony



- marked frames from nodes 1 and 3 always arrive simultaneously at node 2
- an example of *logical synchrony*
- does not require clocks to be synchronized

Abstract frame model

Modeling frames and phase

$$\beta_{i \to j}(t) = \lfloor \theta_i(t - l_{i \to j}) \rfloor - \lfloor \theta_j(t) \rfloor + \lambda_{i \to j}$$

- history of clock phases θ determines location of every frame
- $\lambda_{i \to j}$ is a constant, determined by clock offsets at boot
- buffer occupancy is (roughly) phase difference between clocks at each end of the link

Control loop

- *dynamics*

$$\frac{d\theta_i}{dt} = \omega_i$$

$$\beta_{i \to j}(t) = \lfloor \theta_i(t - l_{i \to j}) \rfloor - \lfloor \theta_j(t) \rfloor + \lambda_{i \to j}$$

- *measurements*
  - at each node i, measure buffer occupancies $\beta_{j \to i}$ for all neighbours j (relative to a mid-point value β^{off})
  - let r_i be the sum $r_i = \sum_{j \mid j \sim i} (\beta_{j \to i} - \beta^{\text{off}})$

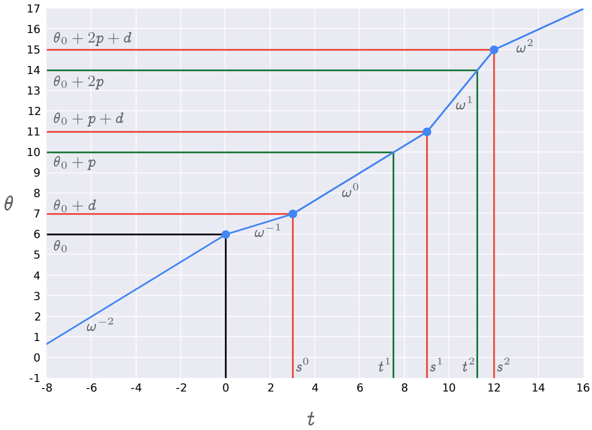
- *control*
  - apply *frequency correction:* $\omega_i = c_i + \omega_i^{\text{u}}$
  - ω_i^{u} is the *uncorrected frequency* of the oscillator, not known
  - *e.g.*, proportional control is $c_i = k_P r_i$

The abstract frame model

for $k \geq 0$ and $t \in \mathbb{R}$

$$\begin{aligned}
\dot{\theta}_i(t) &= \omega_i(t) \\
\beta_{j \to i}(t) &= \lfloor \theta_j(t - l) \rfloor - \lfloor \theta_i(t) \rfloor + \lambda_{j \to i} \\
\omega_i(t) &= \omega_i^k \quad \text{for } t \in [s_i^k, s_i^{k+1}) \\
t_i^k &= \theta_i^{-1}(\theta_i^0 + kp) \\
s_i^k &= \theta_i^{-1}(\theta_i^0 + kp + d) \\
\omega_i^k &= c_i^k + \omega_i^{\mathrm{u}}
\end{aligned}$$

$$\begin{aligned}
c_i^k &= \chi_i^k(m_i^0, m_i^1, \ldots, m_i^k) \\
m_i^k &= \{\beta_{j \to i}(t_i^k) \mid j \in \mathcal{V}, j \sim i\}
\end{aligned}$$



The abstract frame model

for $k \geq 0$ and $t \in \mathbb{R}$

$$
\begin{aligned}
\dot{\theta}_i(t) &= \omega_i(t) \\
\beta_{j \to i}(t) &= \lfloor \theta_j(t - l) \rfloor - \lfloor \theta_i(t) \rfloor + \lambda_{j \to i} \\
\omega_i(t) &= \omega_i^k \quad \text{for } t \in [s_i^k, s_i^{k+1}) \\
t_i^k &= \theta_i^{-1}(\theta_i^0 + kp) \\
s_i^k &= \theta_i^{-1}(\theta_i^0 + kp + d) \\
\omega_i^k &= c_i^k + \omega_i^{\mathrm{u}}
\end{aligned}
$$

$$
\begin{aligned}
c_i^k &= \chi_i^k(m_i^0, m_i^1, \ldots, m_i^k) \\
m_i^k &= \{\beta_{j \to i}(t_i^k) \mid j \in \mathcal{V}, j \sim i\}
\end{aligned}
$$

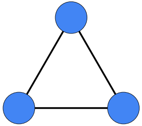
- features:
  - sample rate at node i depends on ω_i
  - no absolute time (integration/differentiation scaled by local clock rate)
  - hybrid: quantization
- existence of solutions? yes, see paper
- exact algorithm for simulation
- software: `bittide.googlesource.com`

Control

Control objectives

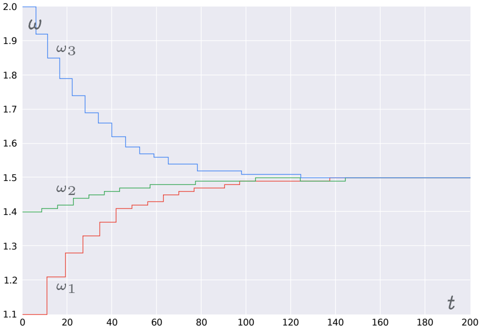
- frequencies cannot remain different for too long, otherwise buffers will over/underflow
- bittide performance requirement: maintain buffer occupancy within limits
- ideally buffer occupancies should be small, and frequencies large
- controller must be decentralized
- no *in-band* signalling
- failure handling, addition and removal of nodes, boot, *etc.*,

Proportional control



- $k_P = 0.01$
- $\texttt{latency} = 1.0$
- $\texttt{poll\_period} = 10$
- $\texttt{uncorrected\_frequency} = (1.1, 1.4, 2.0)$
- $\texttt{control\_delay} = 2$

- varying step width is a consequence of periodic sampling w.r.t. the local clock



Simpler models

- fluid model

$$\dot{\bar{\theta}} = \omega$$

$$\bar{\beta} = -B^{\mathsf{T}}\bar{\theta}$$

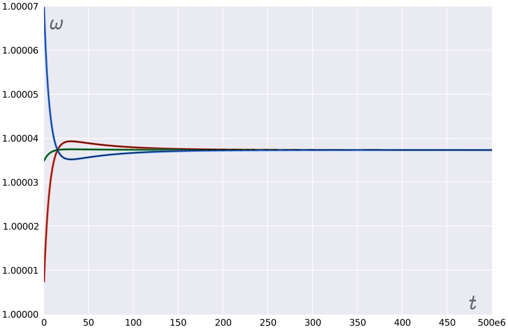
$$r = B\bar{\beta}$$

$$\omega = c + \omega^{\mathrm{u}}$$

- with proportional control

$$\dot{\bar{\theta}} = -k_P B B^{\mathsf{T}}\bar{\theta} + \omega^u$$

the well-known Laplacian dynamics



Thank you!