

# Observing the Consistency of Distributed Systems

Deepthi Akkoorath<sup>1</sup>   Viktória Fördös<sup>2</sup>   Annette Bieniusa<sup>1</sup>

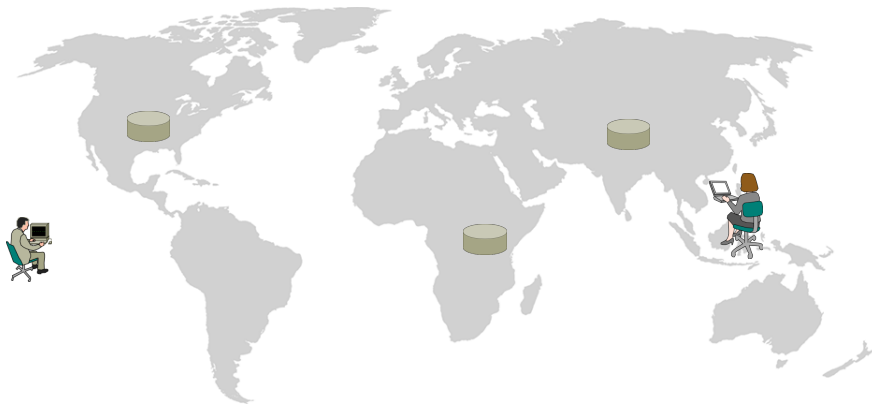
<sup>1</sup>University of Kaiserslautern

<sup>2</sup>Erlang Solutions

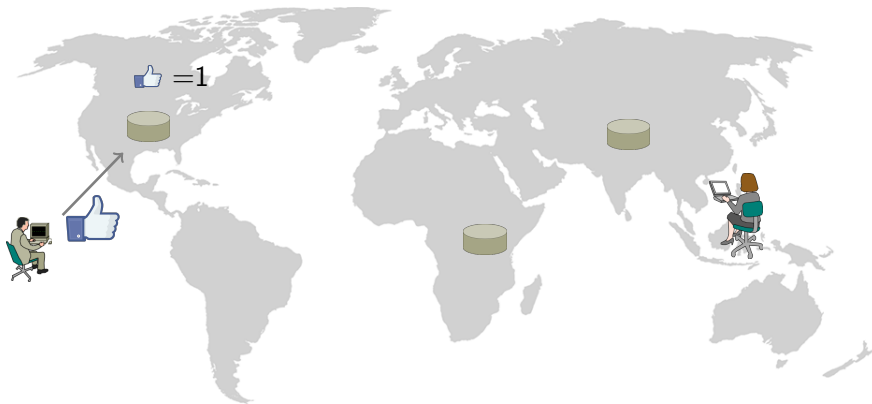
Erlang Workshop '16



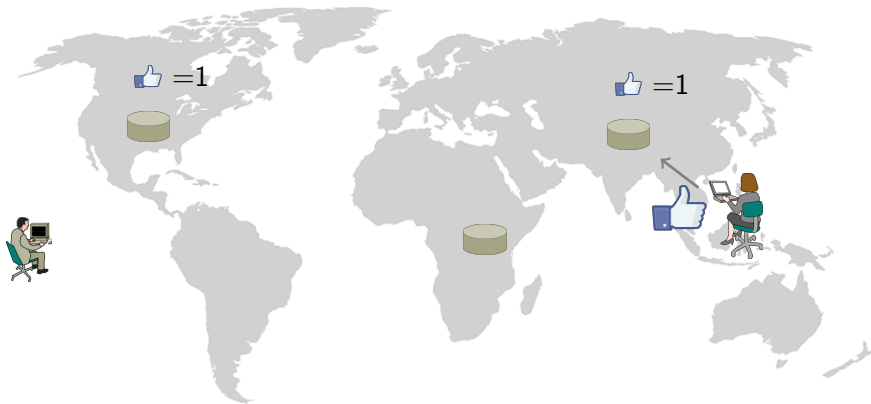
# Weakly Consistent Geo-replicated System



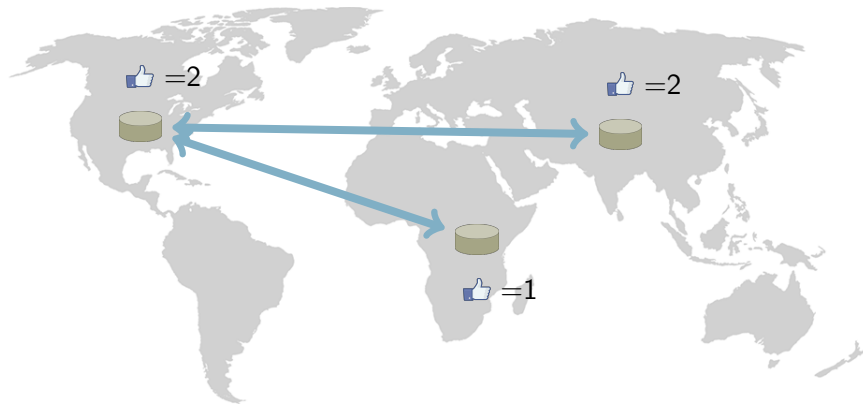
# Weakly Consistent Geo-replicated System



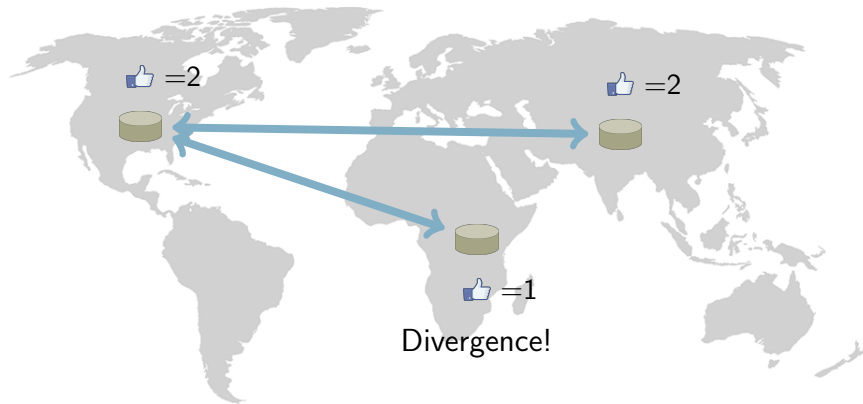
# Weakly Consistent Geo-replicated System



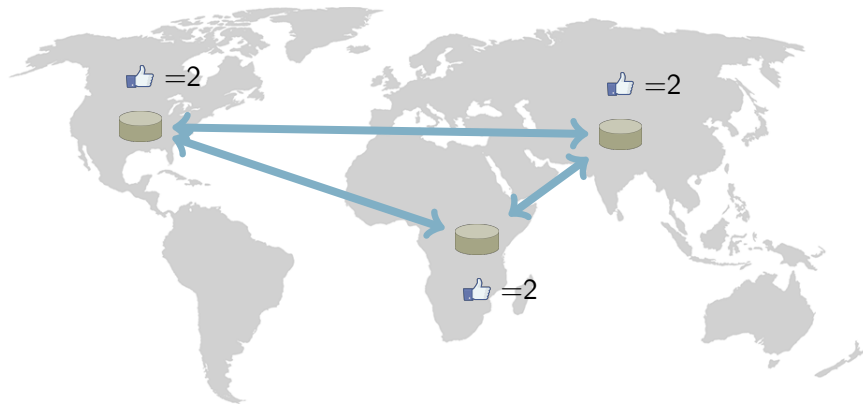
# Weakly Consistent Geo-replicated System



# Weakly Consistent Geo-replicated System



# Weakly Consistent Geo-replicated System



# Problem

Application built on EC databases may **tolerate some divergence**.  
But large divergence may result in SLA violations or unsatisfactory client experiences.

- How to quantify divergence?
- How to monitor divergence?



# Divergence

- Measure of how different the state of a replica compared to others
- Number of missing updates
- Require information about all updates in other replicas
- Not practical to measure

# Divergence

- Measure of how different the state of a replica compared to others
- Number of missing updates
- Require information about all updates in other replicas
- Not practical to measure

## Goal

- Estimate divergence from local knowledge

# Staleness

- How old is the data?
- $\approx$  Time elapsed from last synchronisation
- Only **potential** staleness

# Antidote

- Georeplicated Key-value store
- Written in Erlang
- Weak consistency (Eventual/Causal)
- Conflict-free Replicated Data Types for convergence

# WombatOAM

- An operation and maintenance tool
- For systems running on BEAM virtual machines

# Measuring Staleness in Antidote

Consistency protocol: Cure

- Based on timestamps from real clocks such as NTP
- Maintains a vector of timestamps
- Eg:-  $[\{dc_1, 10\}, \{dc_2, 11\}]$ 
  - Received all updates with timestamps  $\leq 10$  from  $dc_1$
  - Received all updates with timestamp  $\leq 11$  from  $dc_2$

# Measuring Staleness in Antidote

Consistency protocol: Cure

- Based on timestamps from real clocks such as NTP
- Maintains a vector of timestamps
- Eg:-  $[\{dc_1, 10\}, \{dc_2, 11\}]$ 
  - Received all updates with timestamps  $\leq 10$  from  $dc_1$
  - Received all updates with timestamp  $\leq 11$  from  $dc_2$
- Staleness = difference between local timestamp and the observed

# Measuring Staleness in Antidote

- An erlang process running in each node
- Periodically probes vector clock and calculates staleness
- Store result in Exometer histogram
  - Exometer - package for instrumenting Erlang code



# Monitoring Divergence

- WombatOAM plugin for antidote
  - Read metrics stored in Exometer
  - Communicates to Wombat
- WombatOAM
  - Configure threshold values
  - Exposes the maximum and median values
  - Raise alarms when divergence exceeds the threshold

DEMO

localhost:8080/#/topo

WombatOAM™

Topology Metrics Notifications Alarms Services Help

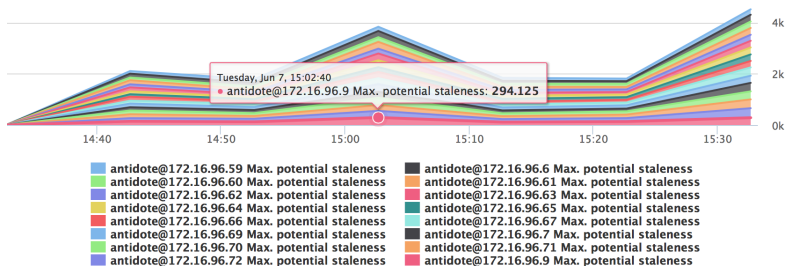
admin

Node Info General Information Add Node Node Graph

OTP APN 181 01 R16B03

You can choose between different functionalities using the top bars.  
Click on a family or a node to view its details.

# Results



# Questions?

- Antidote <https://antidotedb.org>
- WombatOAM <https://www.erlang-solutions.com/products/wombat-oam.html>
- akkoorath@cs.uni-kl.de

Thank You!