

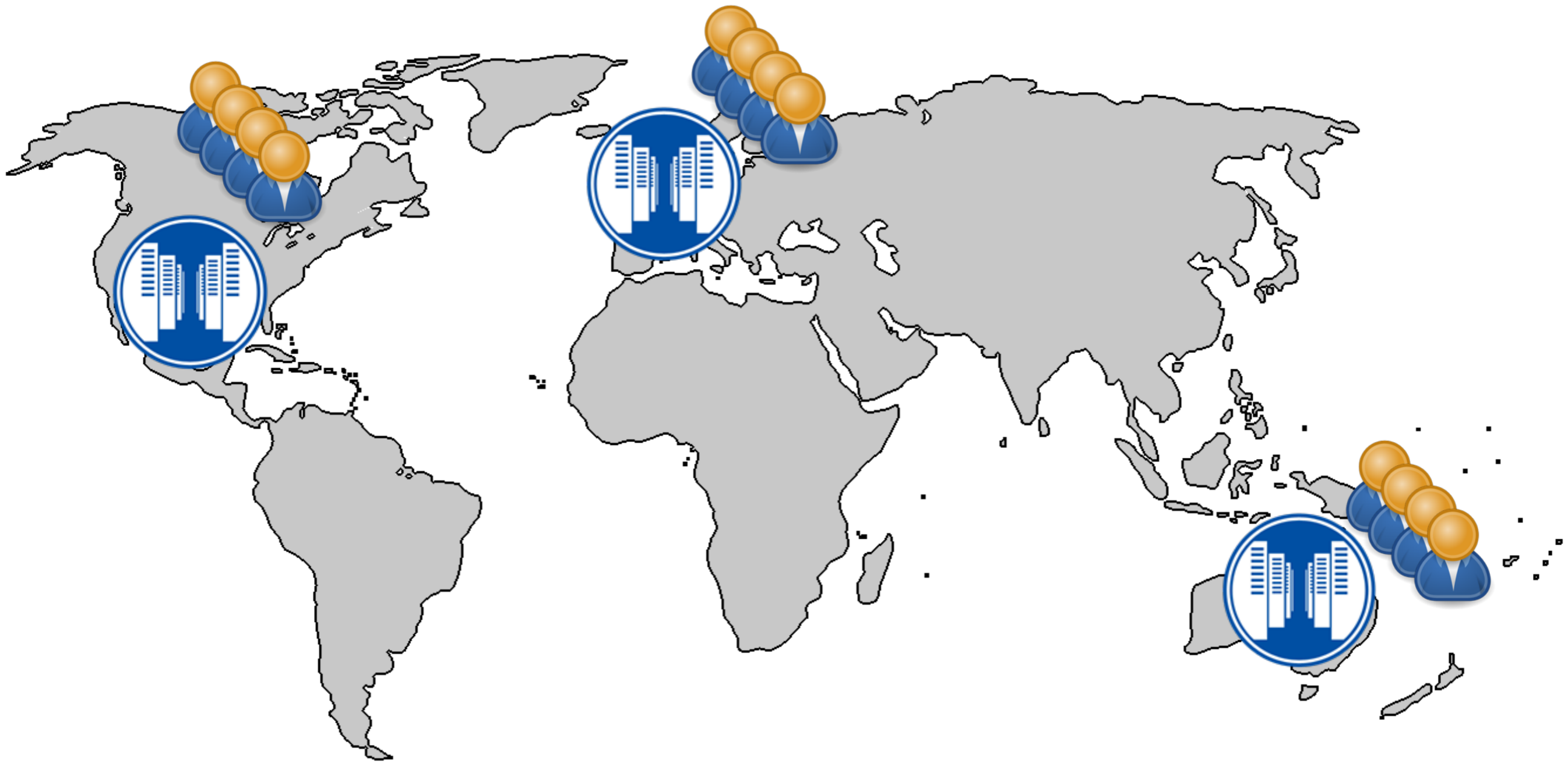
TOWARDS AFFORDABLE EXTERNALLY CONSISTENT GUARANTEES FOR GEO- REPLICATED SYSTEMS

Manuel Bravo
Luís Rodrigues

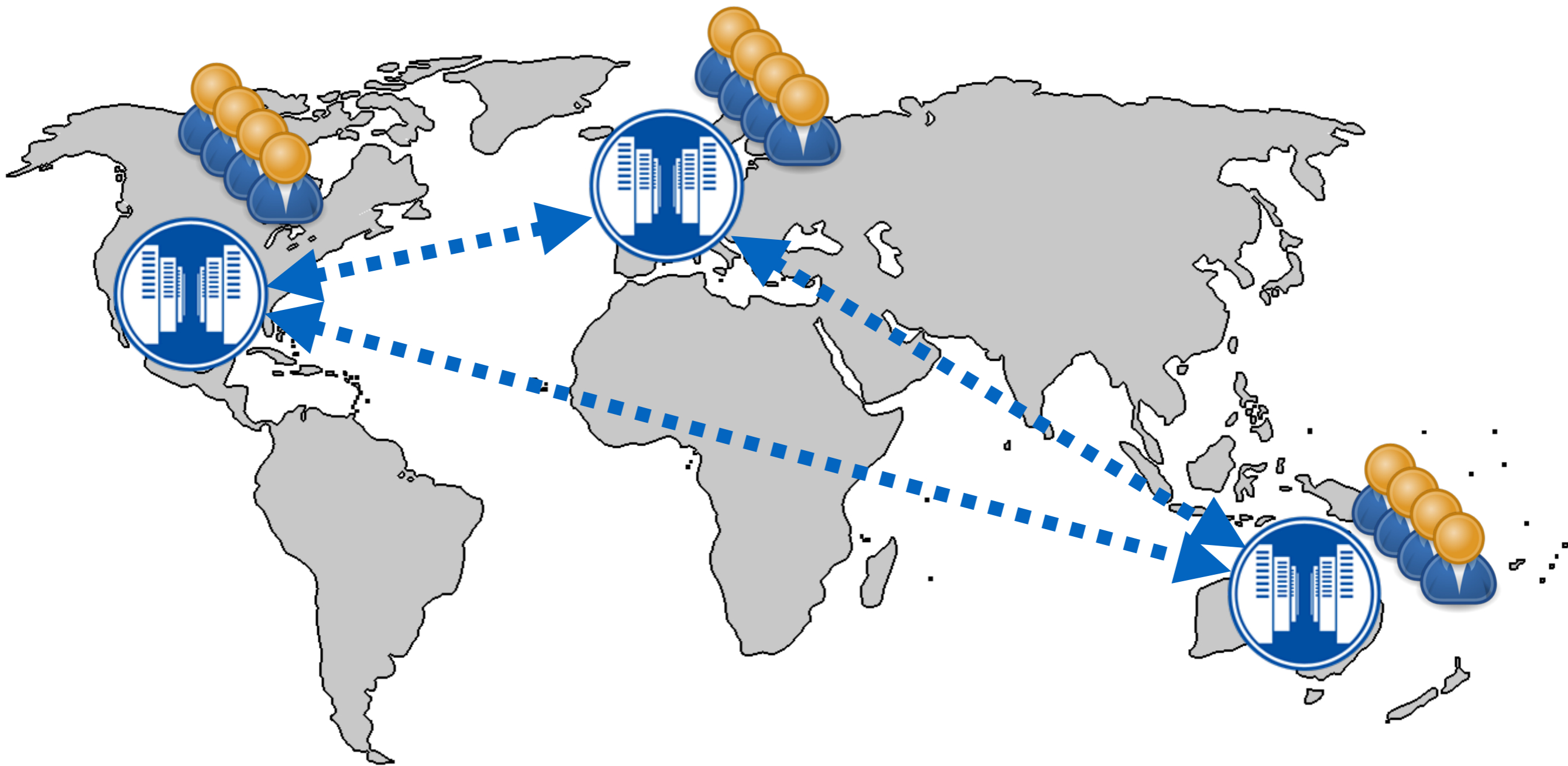
Geo-replication



Geo-replication



Geo-replication



Tension between performance and semantics



Tension between performance and semantics

strong
consistency

eventual
consistency

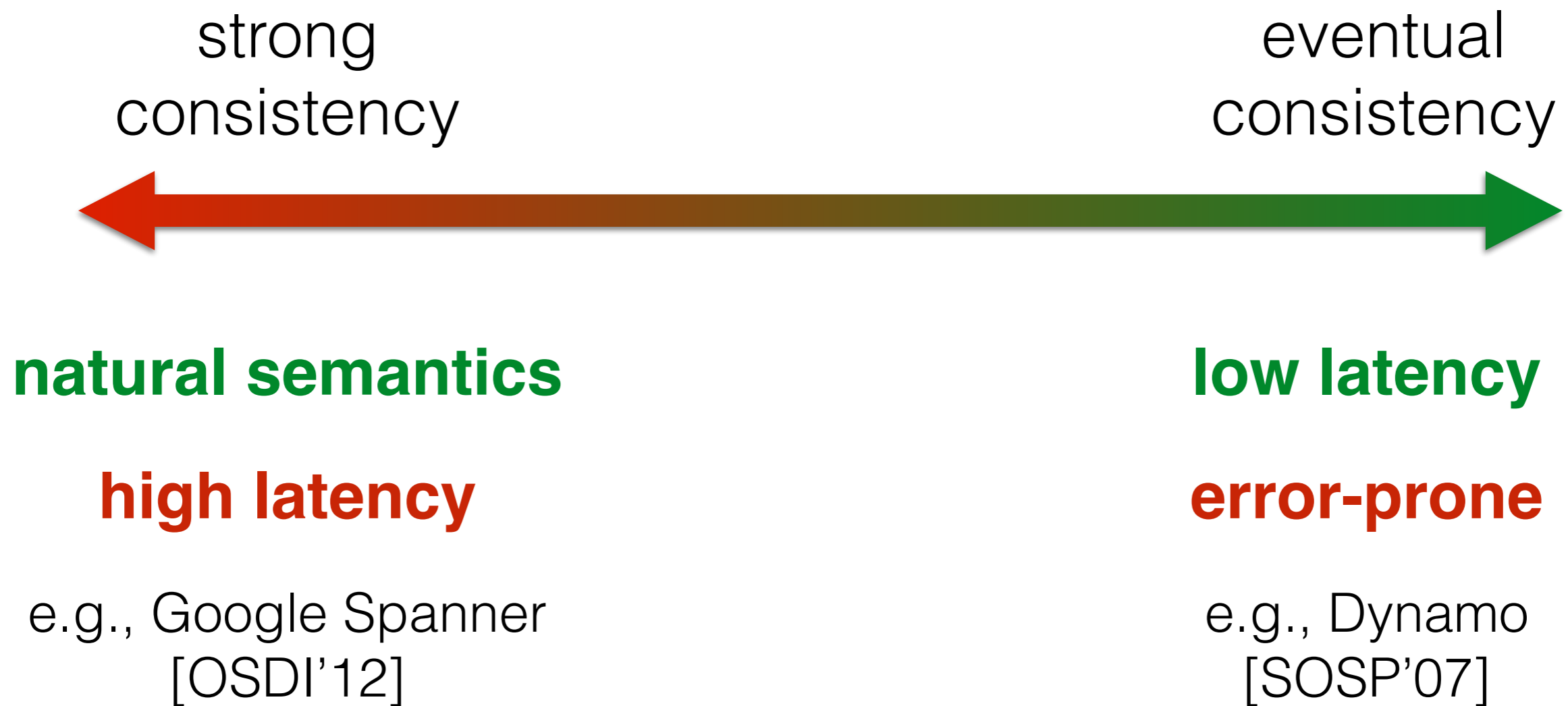


natural semantics

high latency

e.g., Google Spanner
[OSDI'12]

Tension between performance and semantics



Tension between performance and semantics



Tension between performance and semantics

strong
consistency

eventual
consistency



natural semantics

high latency

low latency

error-prone

Tension between performance and semantics

strong
consistency

eventual
consistency



A1

B1

A2

B2

A3

B3

A4

natural semantics

high latency

low latency

error-prone

Tension between performance and semantics

strong
consistency

eventual
consistency



linearizability

R1
↓
R2
↓
R3

A1
A2
A3
A4

B1
B2
B3

natural semantics

high latency

low latency

error-prone

Tension between performance and semantics

strong
consistency

eventual
consistency



linearizability sequential

R1
↓
R2
↓
R3

R1
↓
R3
↓
R2

A1
A2
A3
A4

B1
B2
B3

natural semantics

high latency

low latency

error-prone

Tension between performance and semantics

strong
consistency

eventual
consistency



linearizability sequential

R1
↓
R2
↓
R3

R1
↓
R3
↓
R2

natural semantics

high latency

causal

A1
↓
A2

B1
↓
B2
↓
B3

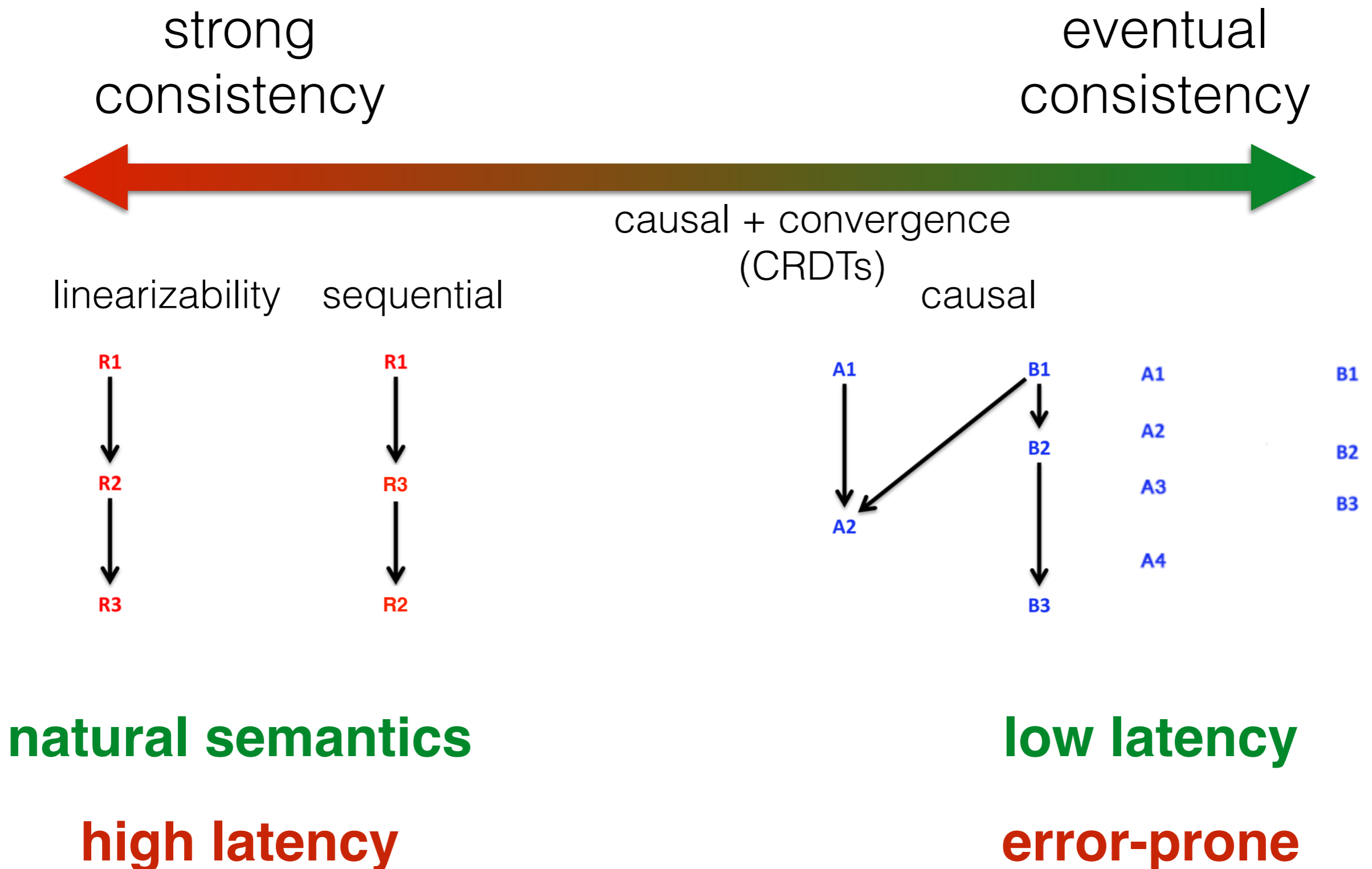
A1
A2
A3
A4

B1
B2
B3

low latency

error-prone

Tension between performance and semantics



Alternative

mixing semantics, consistent **when
necessary**

RedBlue Consistency
[OSDI'12]

Parallel Snapshot Isolation
[SOSP'11]

Explicit Consistency
[EuroSys'15]

Pileus
[SOSP'13]

Session guarantees
[SOSP'97]

Alternative

mixing semantics, consistent **when**

necessary

Interestingly, most
previous solutions do not

RedBlue Consistency
[OSDI'12]

consider **external—or**

global—guarantees

Parallel Snapshot Isolation
[SOSP'11]

Explicit Consistency
[EuroSys'15]

Pileus
[SOSP'13]

Session guarantees
[SOSP'97]

External consistency

clients are served with a view of the system **consistent to how an external observer** would witness the succession of events

CHAPTER 3: TRANSACTIONAL STORAGE

consider read and write actions, and we will assume that a transaction performs at most one read and one write action on a data element.

The question of serial consistency arises when a transactional storage system concurrently executes actions drawn from several transactions. Imagine two simple transactions:

T1	T2
(a11) Read Y	(a21) Read X
(a12) Write X	(a22) Read Y
(a13) Write Y	(a23) Write X

The order in which the actions of T1 and T2 are processed is called a *schedule*. A schedule is an arbitrary interleaving of the actions of a set of transactions into a single sequence. A *serial schedule* results when transactions are executed one at a time to completion. Thus, there are two possible serial schedules for our example:

Sa: {a11, a12, a13, a21, a22, a23}
Sb: {a21, a22, a23, a11, a12, a13}.

A schedule generates a *dependency* relation. A dependency relation describes transactions that depend on one another. If S is a schedule, then $\langle T_i, \dots, T_j \rangle$ is a member

Externally consistent guarantees

very **powerful semantics**

Externally consistent guarantees

very **powerful semantics**

an operation observes all operations
that were completed as when the operation began

Externally consistent guarantees

very **powerful semantics**

an operation observes all operations
that were completed as when the operation began

can cope with back-channeling!

Externally consistent guarantees

very **powerful semantics**

an operation observes all operations
that were completed as when the operation began

can cope with back-channeling!

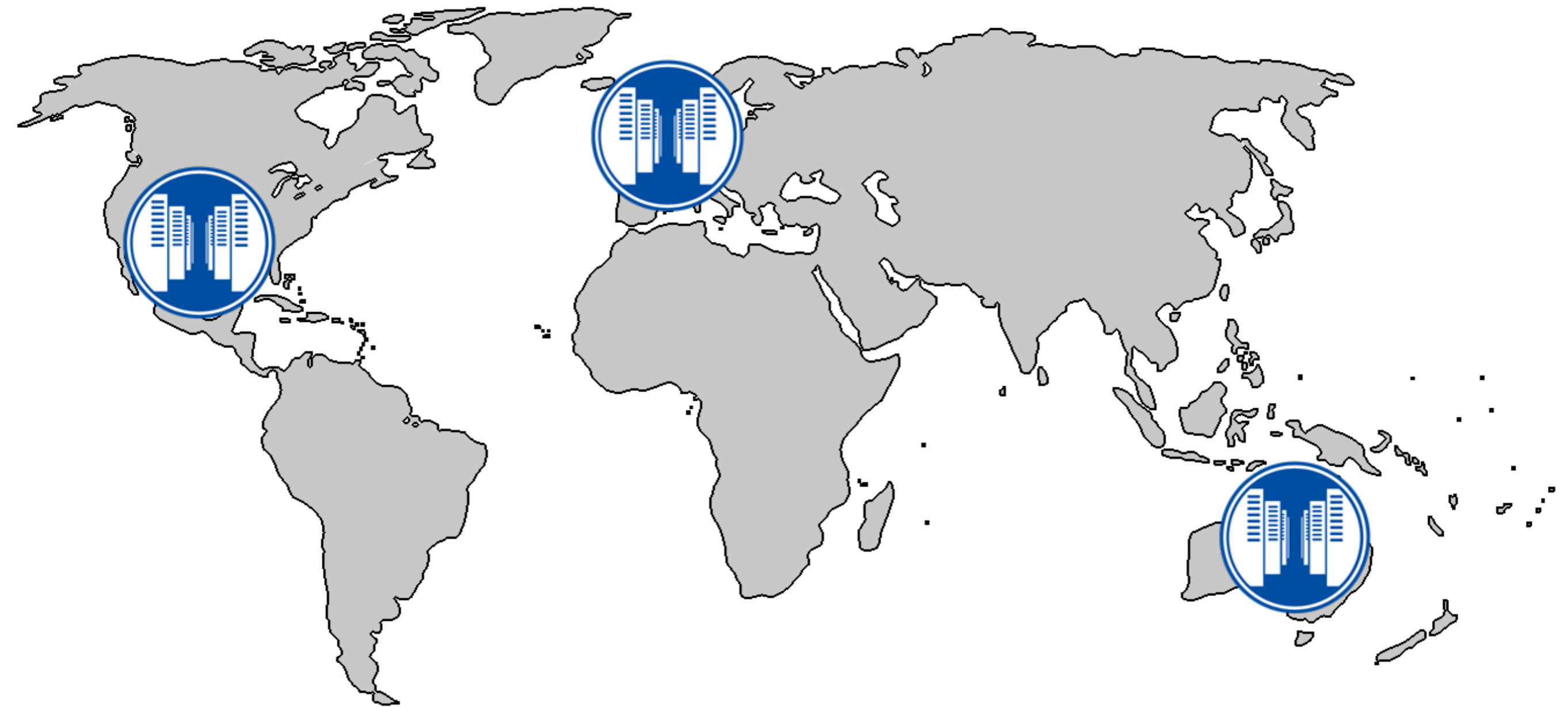
but, requires a lot of coordination
then, **very expensive**

Use case: an auction service

two roles: auctioneers and buyers

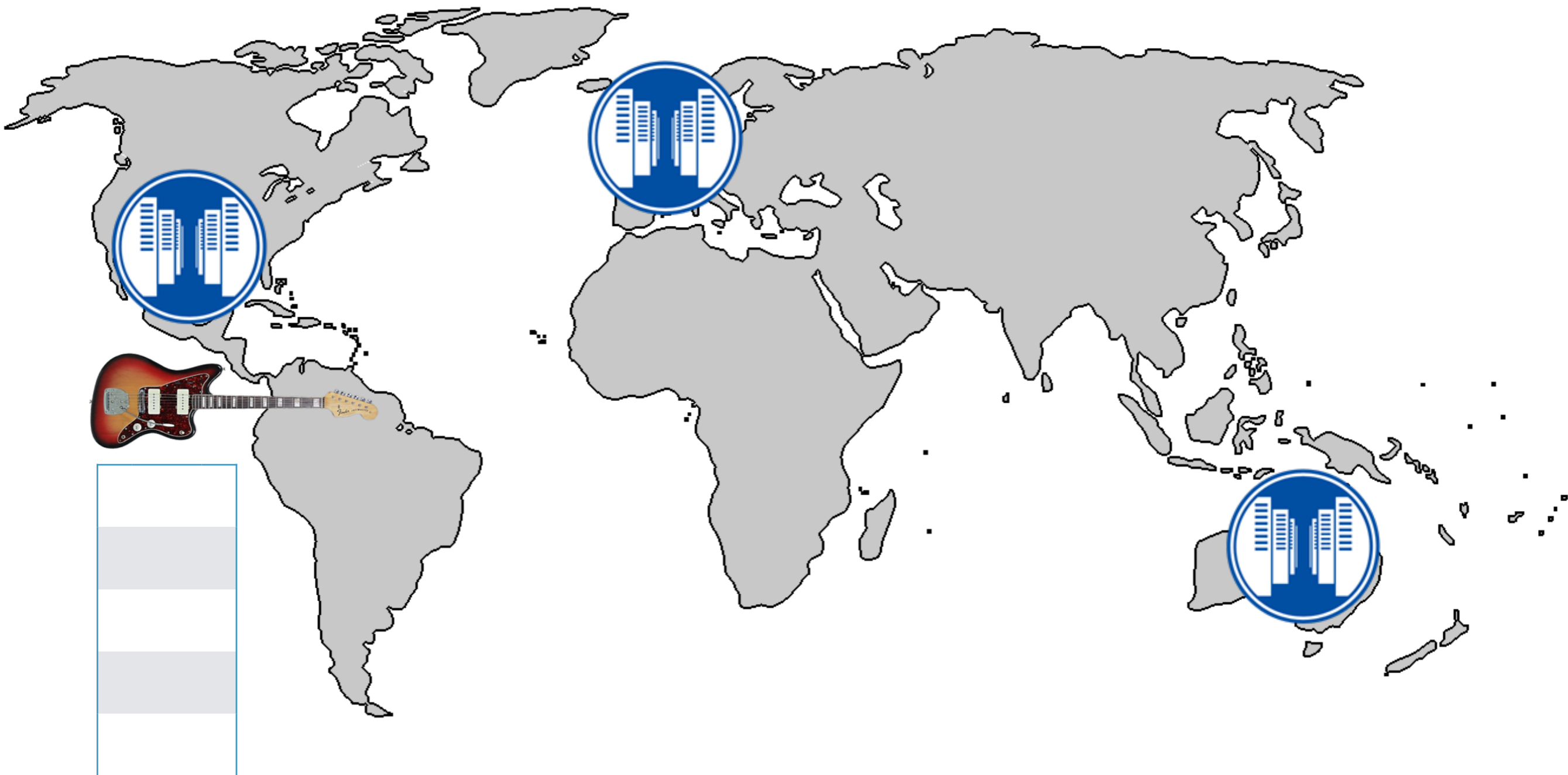
operations (among many): **start an auction, place a bid, close an auction, report the winner**

Use case: an auction service



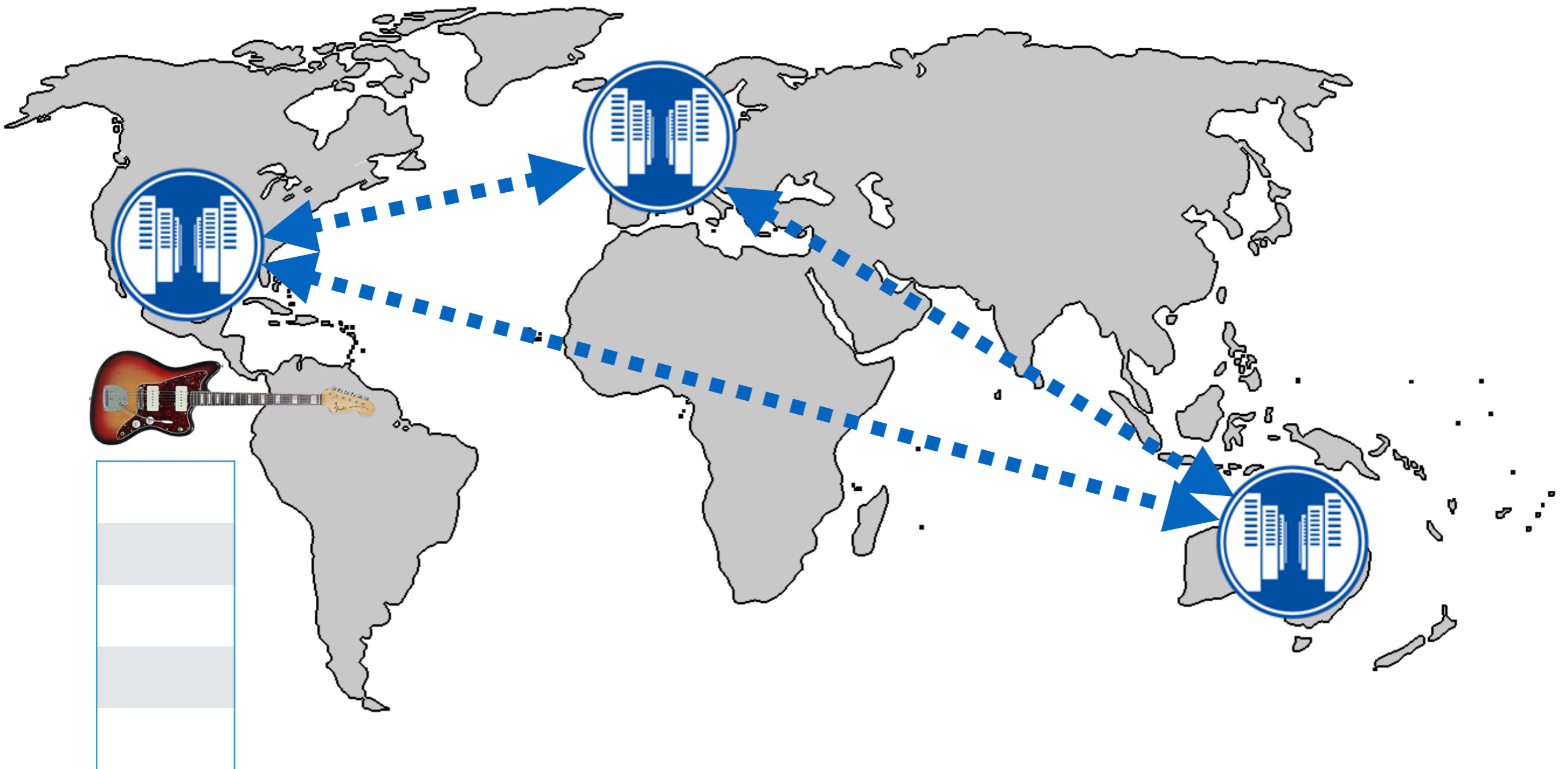
Use case: an auction service

start auction!



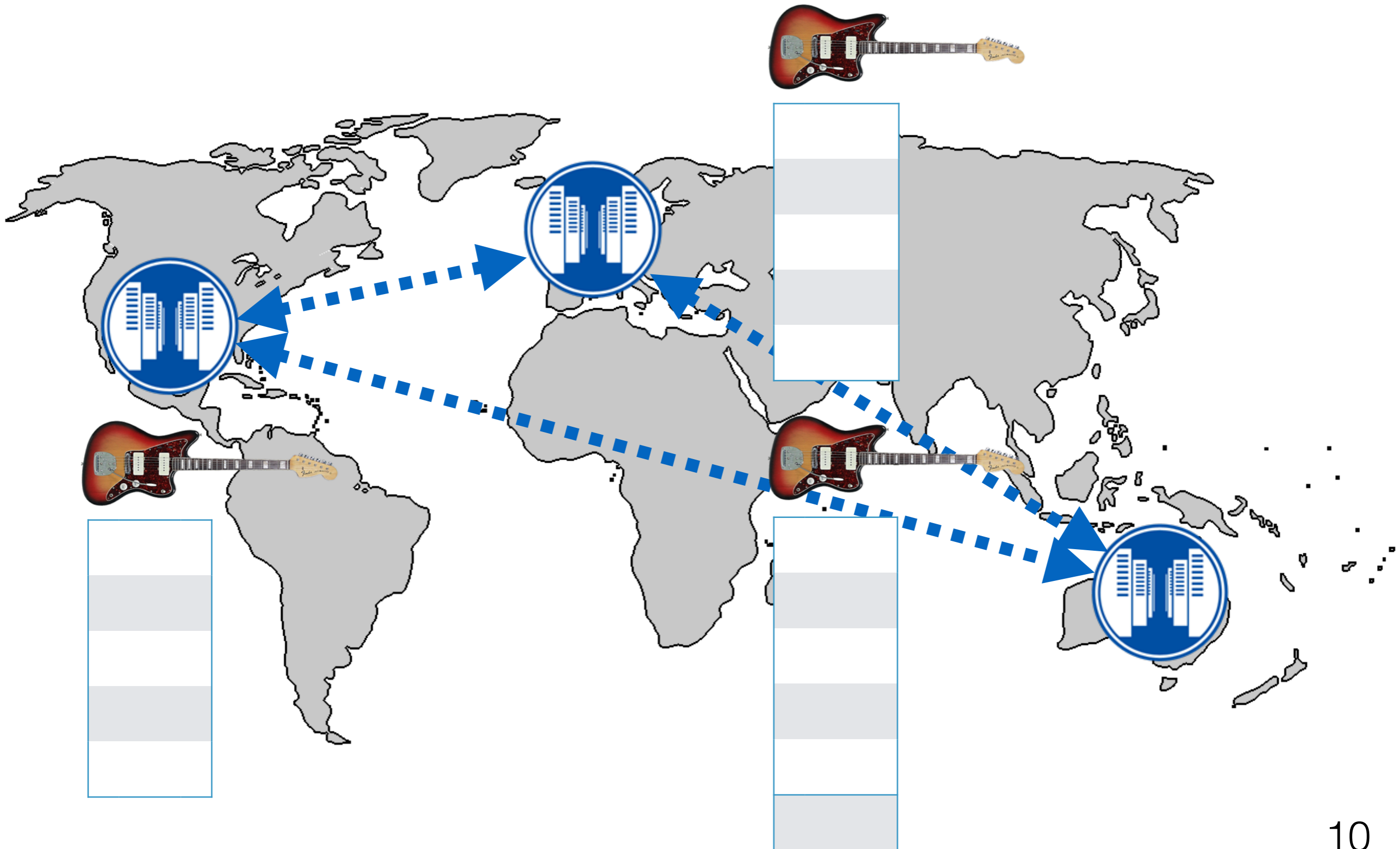
Use case: an auction service

start auction!



Use case: an auction service

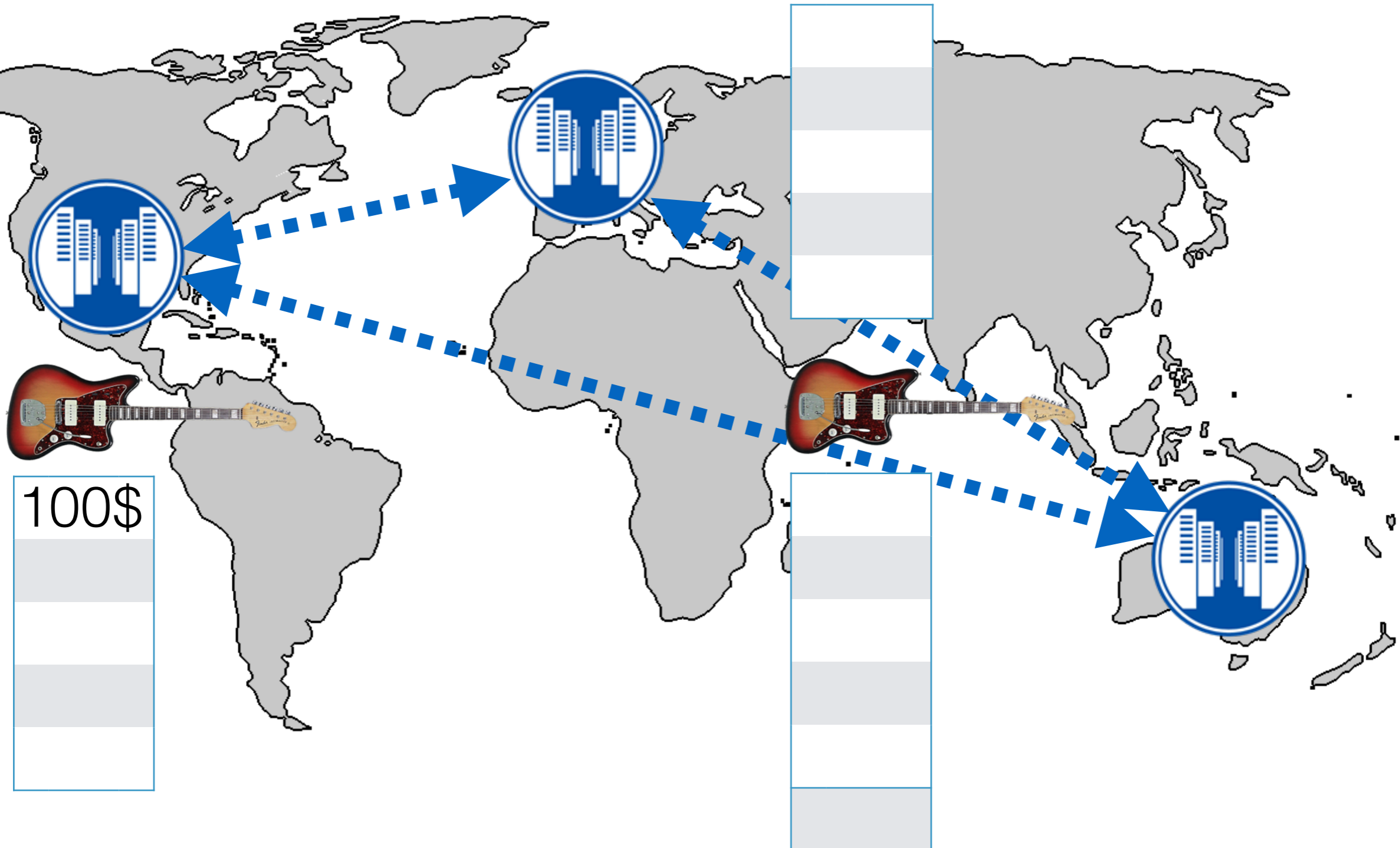
start auction!



Use case: an auction service

start auction!

bid!



Use case: an auction service

start auction!

bid!



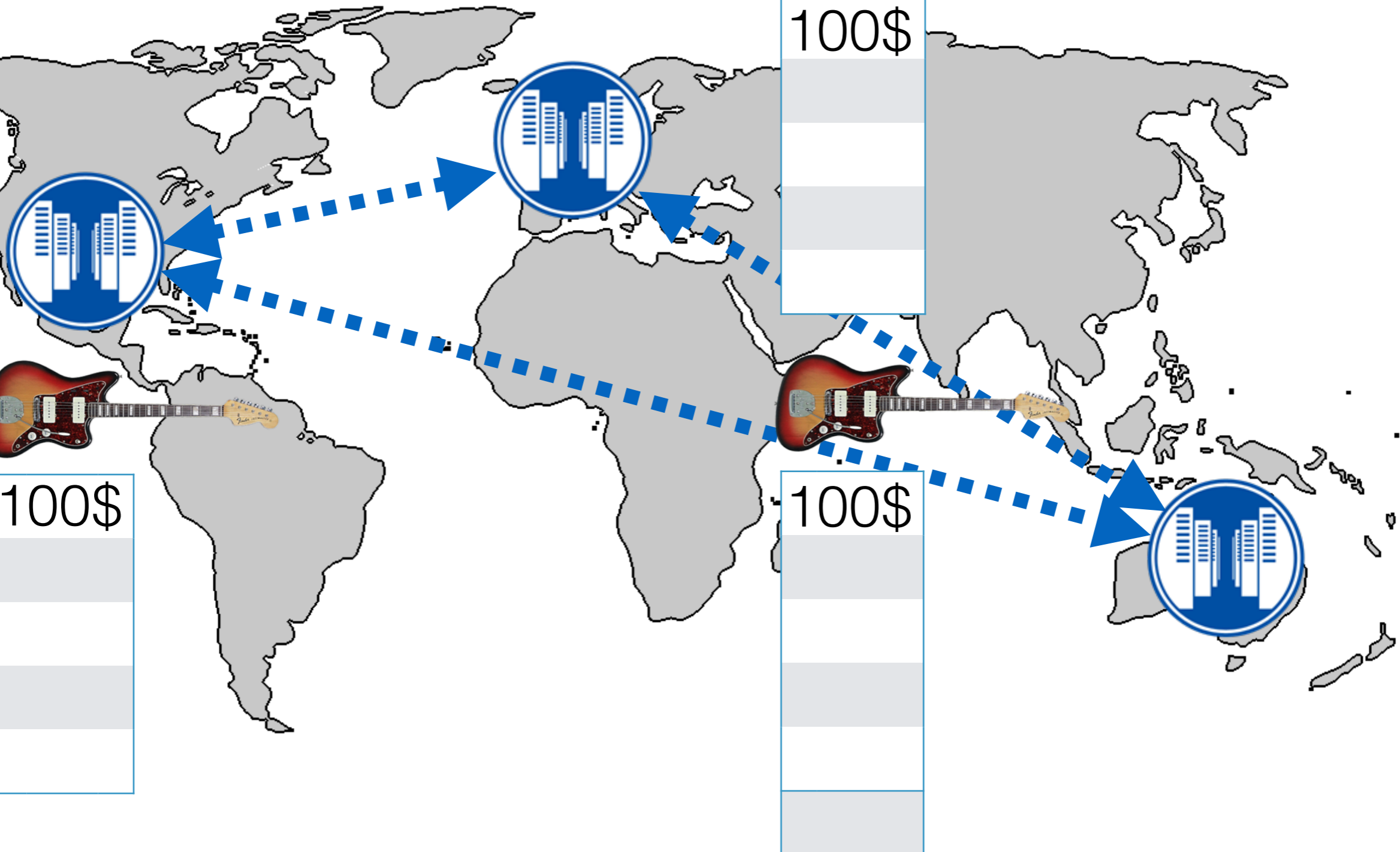
100\$



100\$



100\$



Use case: an auction service

start auction!

bid! bid! bid!



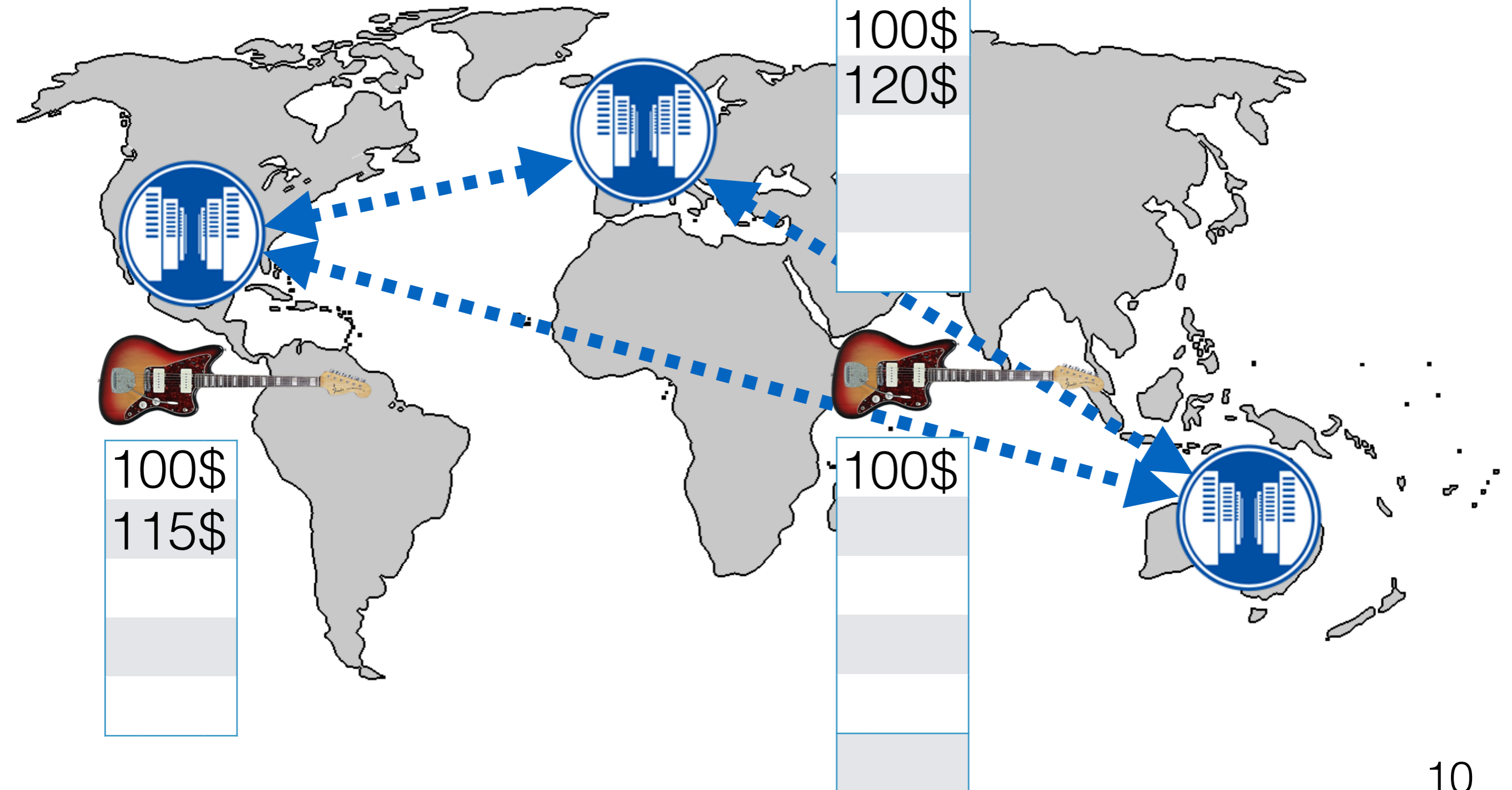
100\$
120\$



100\$



100\$
115\$



Use case: an auction service

start auction!

bid! bid! bid!



100\$
120\$
115\$



100\$
115\$
120\$



100\$
120\$
115\$



Use case: an auction service

start auction!

bid! bid! bid!
bid!



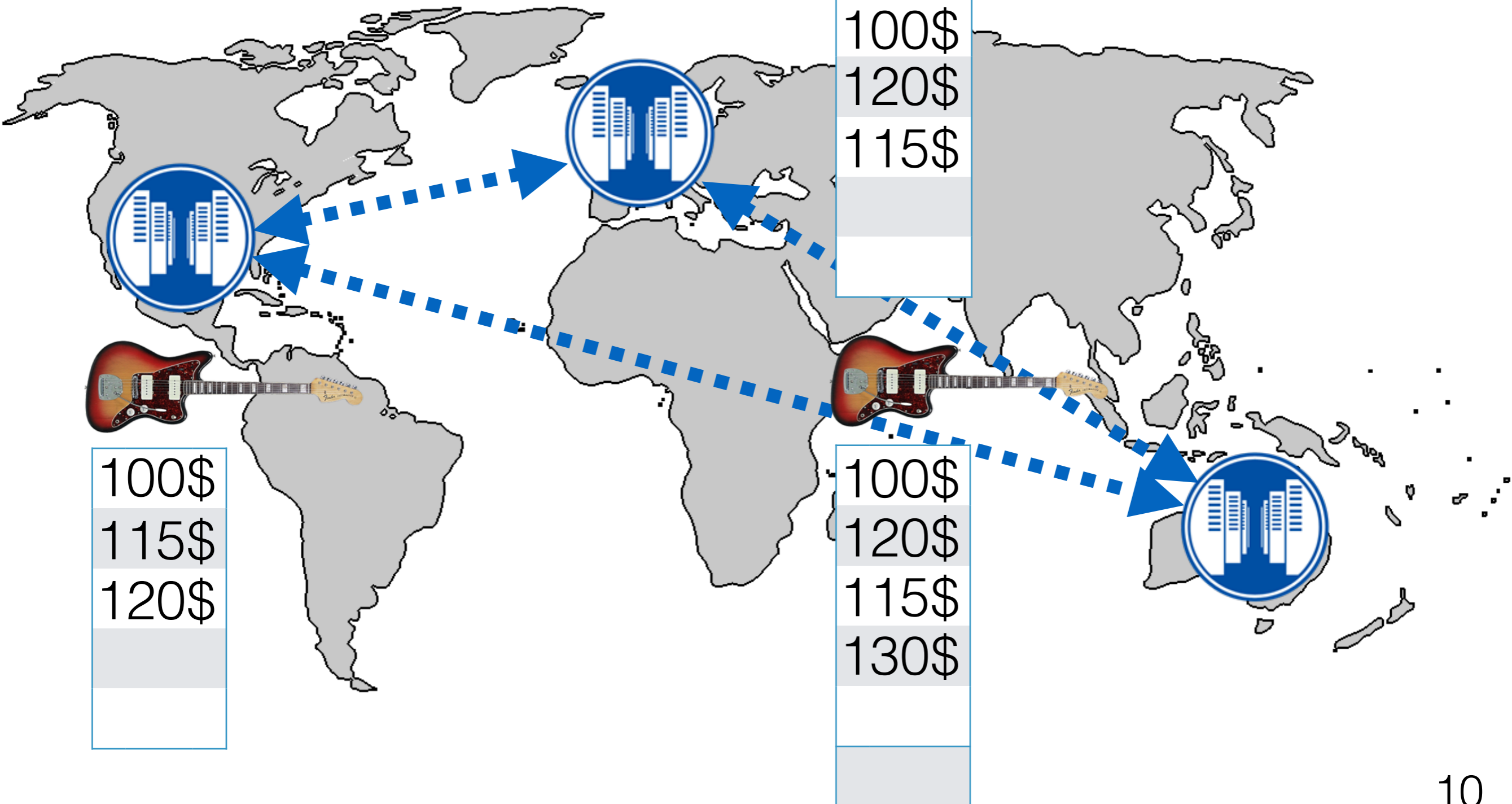
100\$
120\$
115\$



100\$
115\$
120\$



100\$
120\$
115\$
130\$



Use case: an auction service

start auction!

bid! bid! bid!
bid! bid!



100\$
120\$
115\$



100\$
115\$
120\$
125\$



100\$
120\$
115\$
130\$

Use case: an auction service

start auction!

bid! bid! bid!
bid! bid!



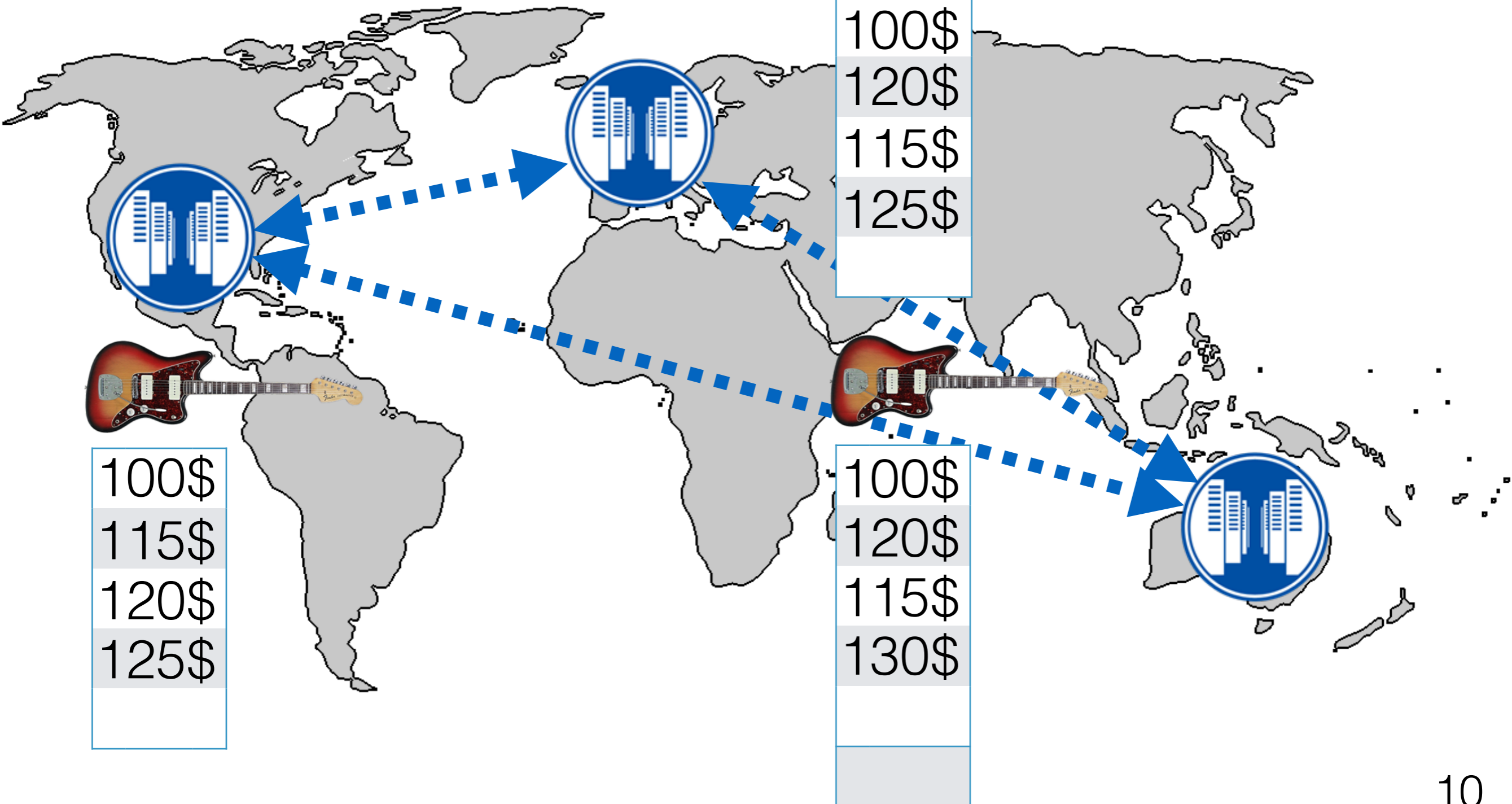
100\$
120\$
115\$
125\$



100\$
115\$
120\$
125\$



100\$
120\$
115\$
130\$



Use case: an auction service

start auction!

bid! bid! bid!
bid! bid!



100\$
120\$
115\$
125\$



100\$
115\$
120\$
125\$



100\$
120\$
115\$
130\$
125\$



Use case: an auction service

start auction!

bid! bid! bid!
bid! bid!

close auction!



100\$
120\$
115\$
125\$



100\$
115\$
120\$
125\$
✗



100\$
120\$
115\$
130\$
125\$



Use case: an auction service

start auction!

bid! bid! bid!
bid! bid!

close auction!



100\$
120\$
115\$
125\$
X



100\$
115\$
120\$
125\$
X



100\$
120\$
115\$
130\$
125\$
X

Use case: an auction service

start auction!

bid! bid! bid!
bid! bid!

close auction!



100\$
120\$
115\$
125\$
✗

130\$

130\$



100\$
115\$
120\$
125\$
✗



100\$
120\$
115\$
130\$
125\$
✗



Use case: an auction service

start auction!

bid! bid! bid!
bid! bid!

close auction!



100\$
120\$
115\$
125\$
X

130\$

100\$
120\$
115\$
130\$
125\$
X

130\$

report winner!



100\$
115\$
120\$
125\$
X

Use case: an auction service

start auction!

bid! bid! bid!
bid! bid!

close auction!



100\$
120\$
115\$
125\$
X

130\$



100\$
120\$
115\$
130\$
125\$
X

130\$

report winner!



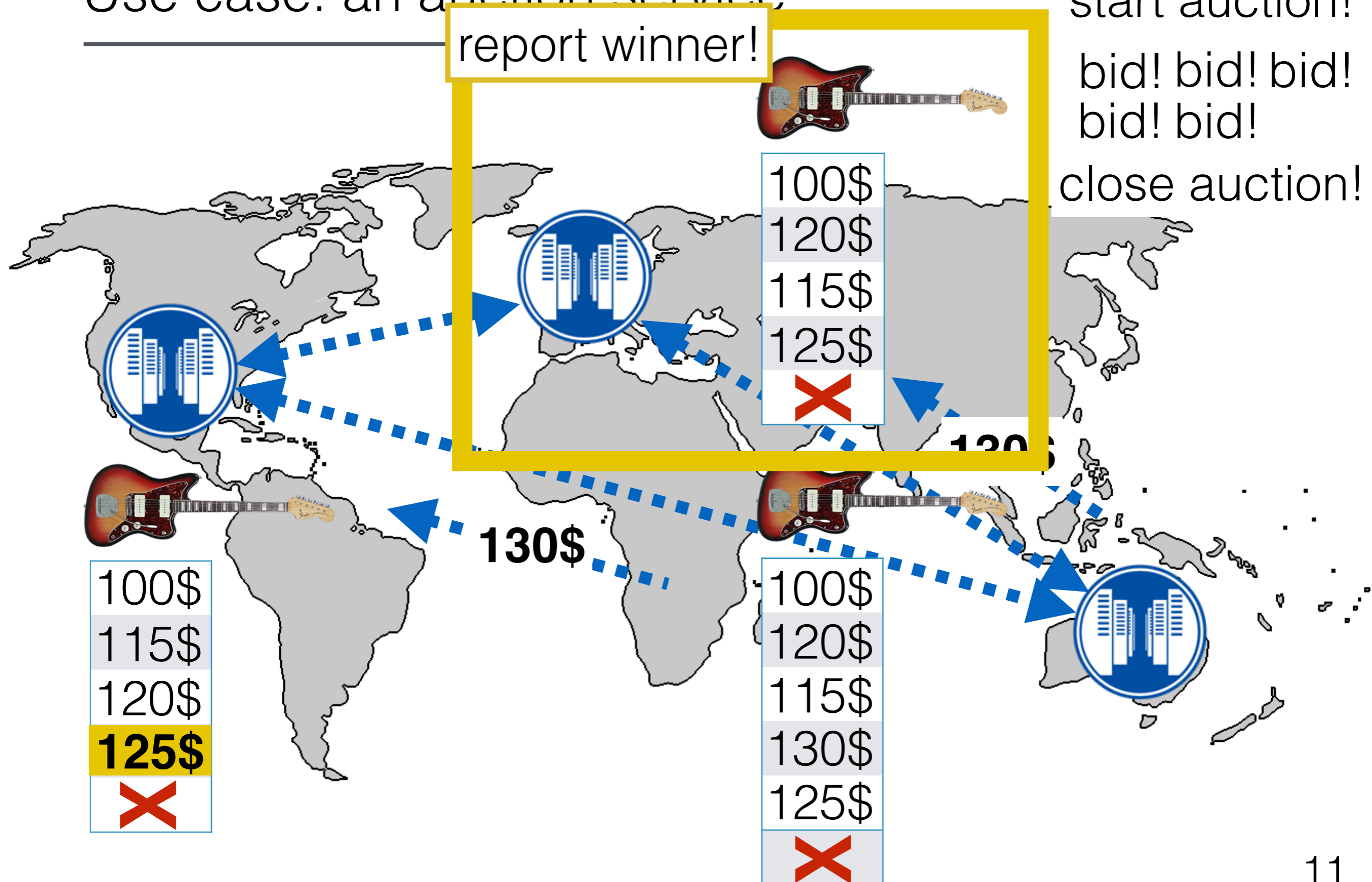
100\$
115\$
120\$
125\$
X

Use case: an auction service
report winner!

start auction!

bid! bid! bid!
bid! bid!

close auction!

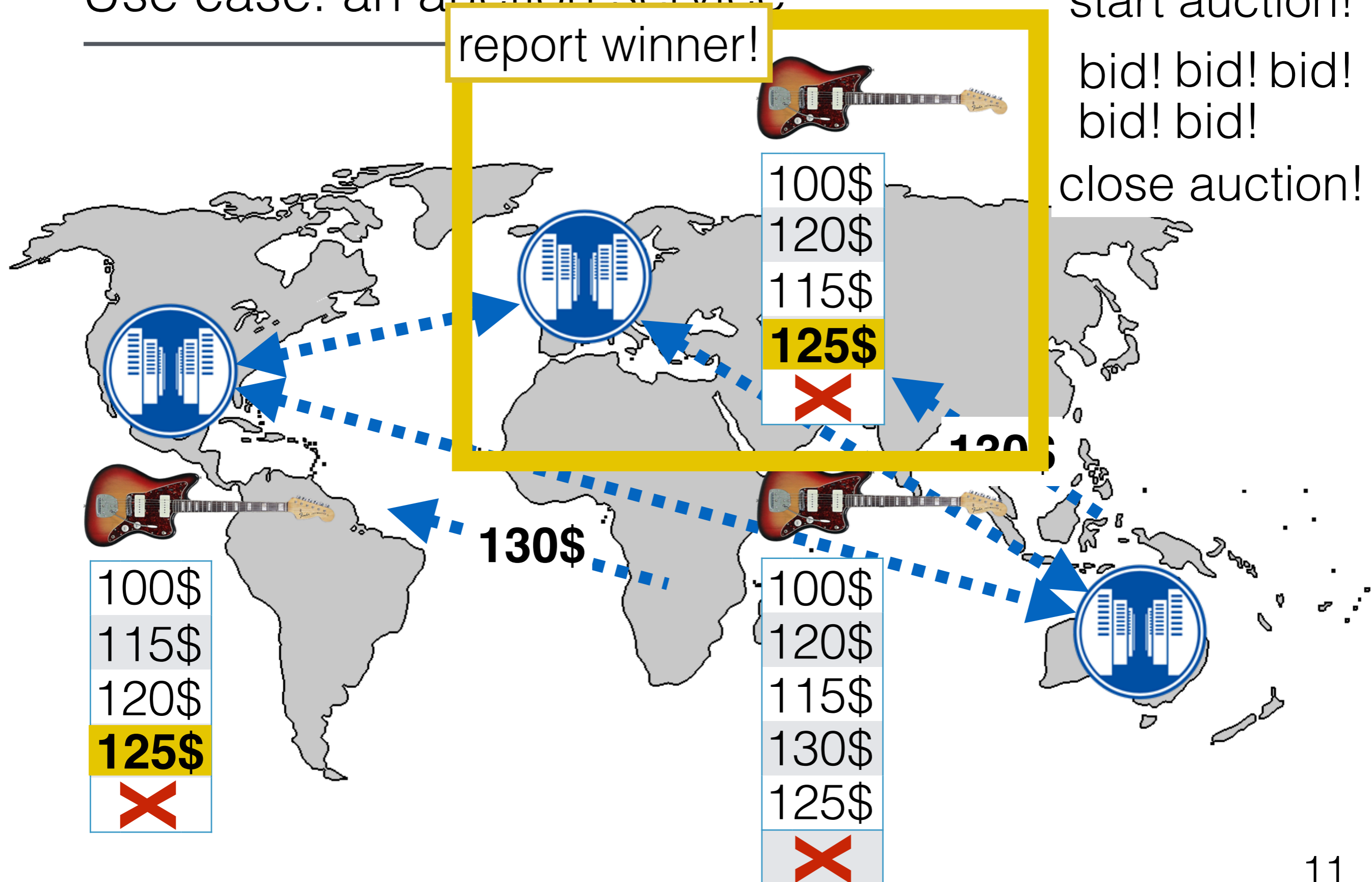


Use case: an auction service
report winner!

start auction!

bid! bid! bid!
bid! bid!

close auction!



Use case: an auction service

start auction!

bid! bid! bid!
bid! bid!

close auction!



100\$
120\$
115\$
125\$

130\$

report winner!

100\$
120\$
115\$
130\$
125\$

130\$

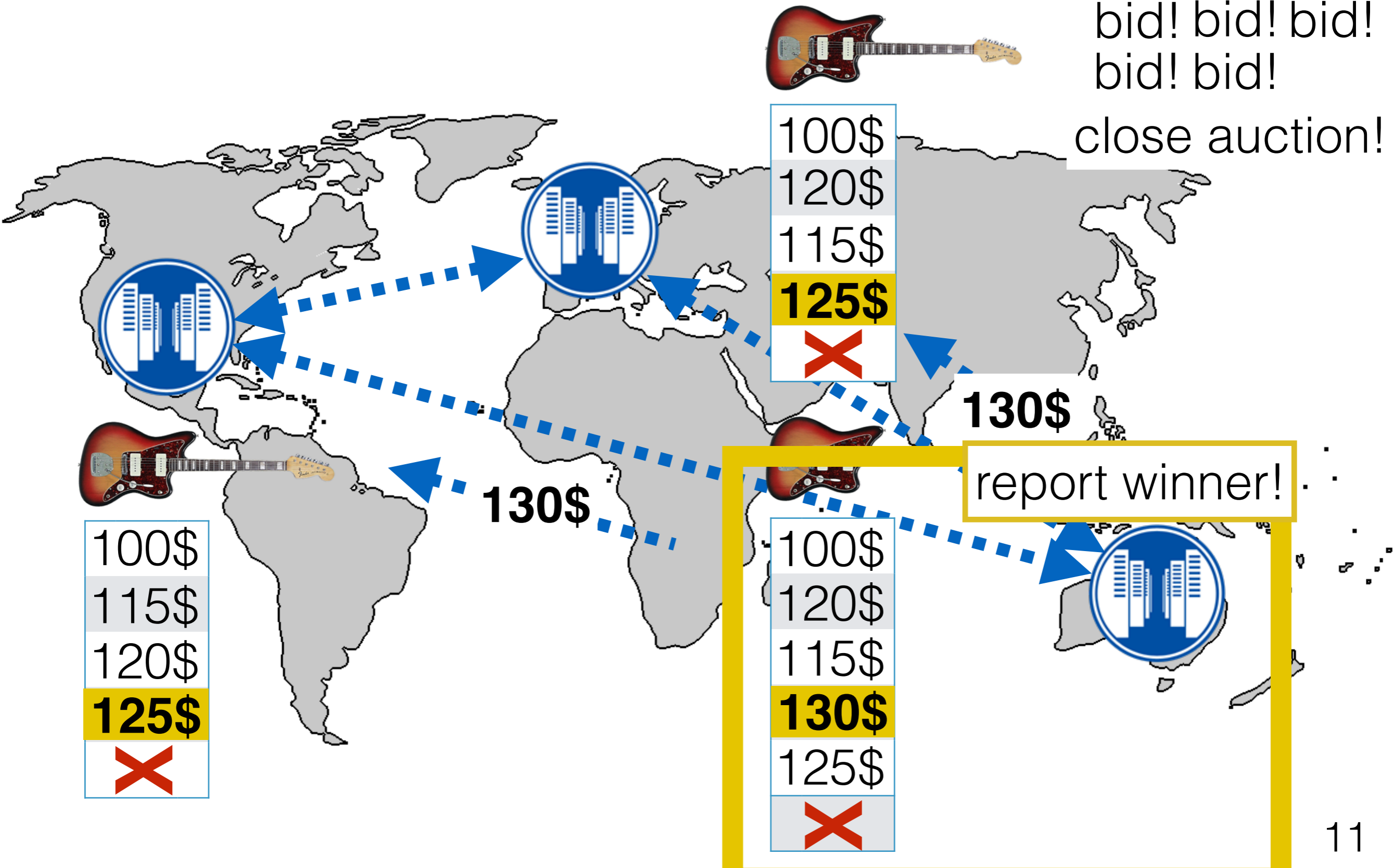
100\$
115\$
120\$
125\$

Use case: an auction service

start auction!

bid! bid! bid!
bid! bid!

close auction!



Use case: an auction service

How can we provide developers with mechanisms to overcome this type of problems without slowing down the system significantly?



100\$
115\$
120\$
125\$

100\$
120\$
115\$
130\$
125\$

Use case: an auction service

How can we provide
externally consistent
guarantees **on demand**?



100\$
115\$
120\$
125\$

100\$
120\$
115\$
125\$

100\$
120\$
115\$
130\$
125\$

Our work

addresses this inherent tension between
performance and meaningful semantics

we present a new consistency model:
external causality

takes causal consistency (strongest
available) and spice it up with external
consistency guarantees

External causality

internal operations: read from a causally consistent snapshot

external operations: read from a casual snapshot that includes latest updates as of the time when the operation began

Hypothesis

internal operations are **highly predominant** and **cheap** to implement (and we know how to)

external operations—although **expensive**—are **rarely required**

What about concurrency?

external causality **allows concurrency**, only guaranteeing that operations (both internal and external) are executed in causal order at each site

this is not ideal from the semantics point of view, but it has significant benefits in practice

External causality: example execution

Site **a**

Site **b**



internal
external
causality

Real
Time

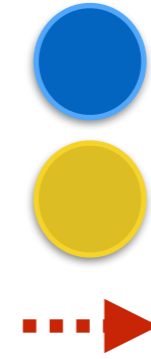


External causality: example execution

Real
Time

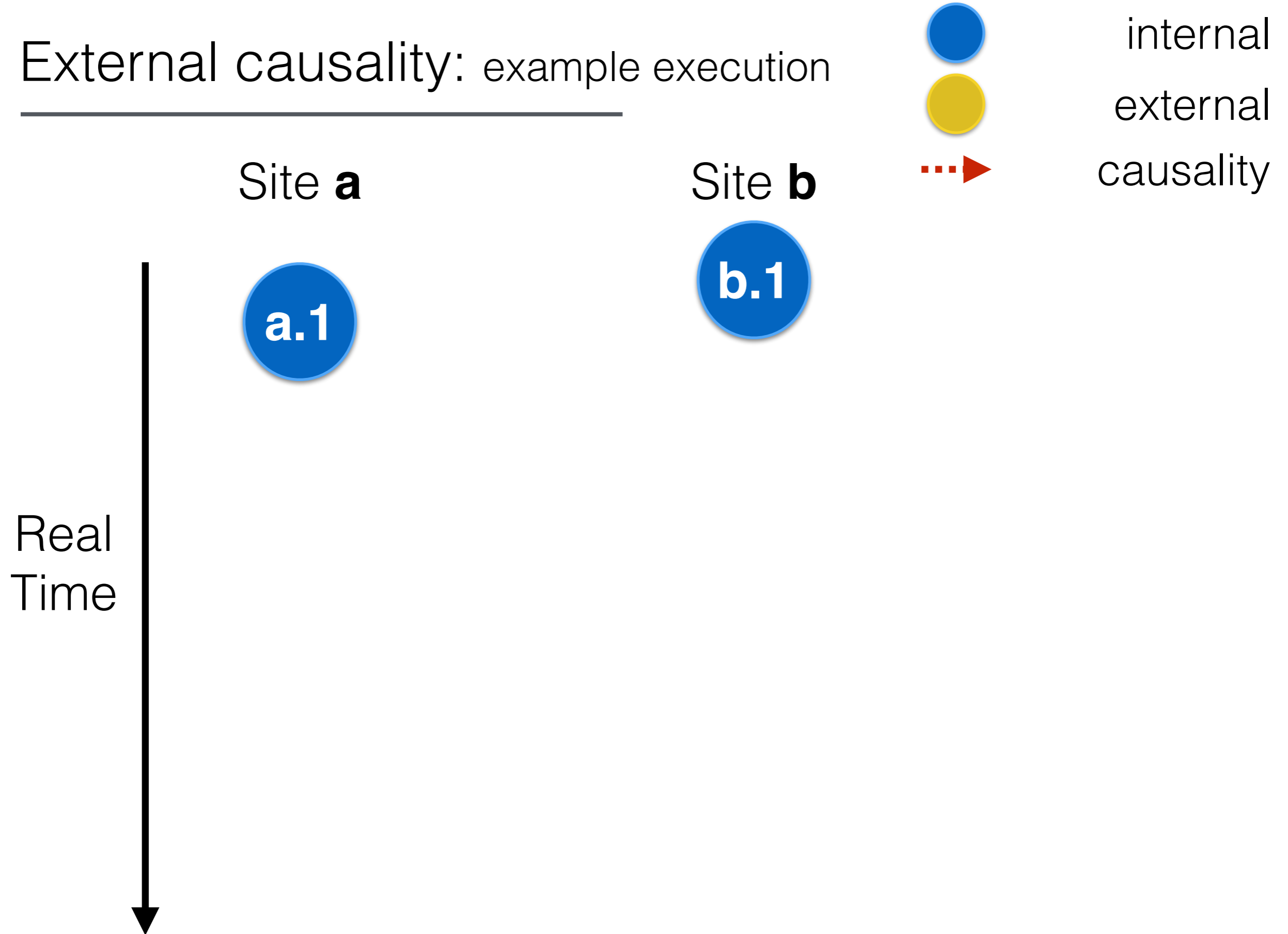
Site **a**

Site **b**

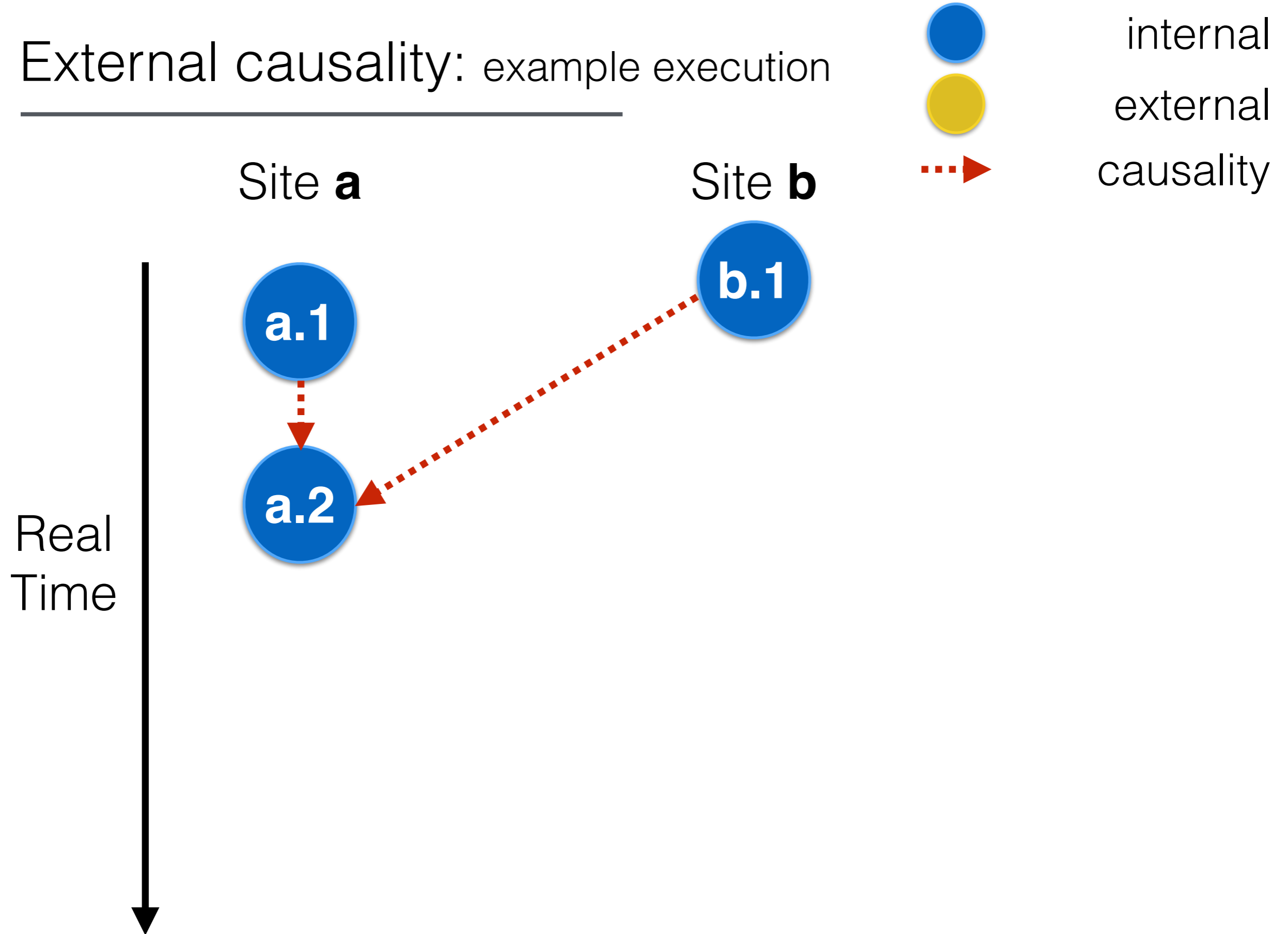


internal
external
causality

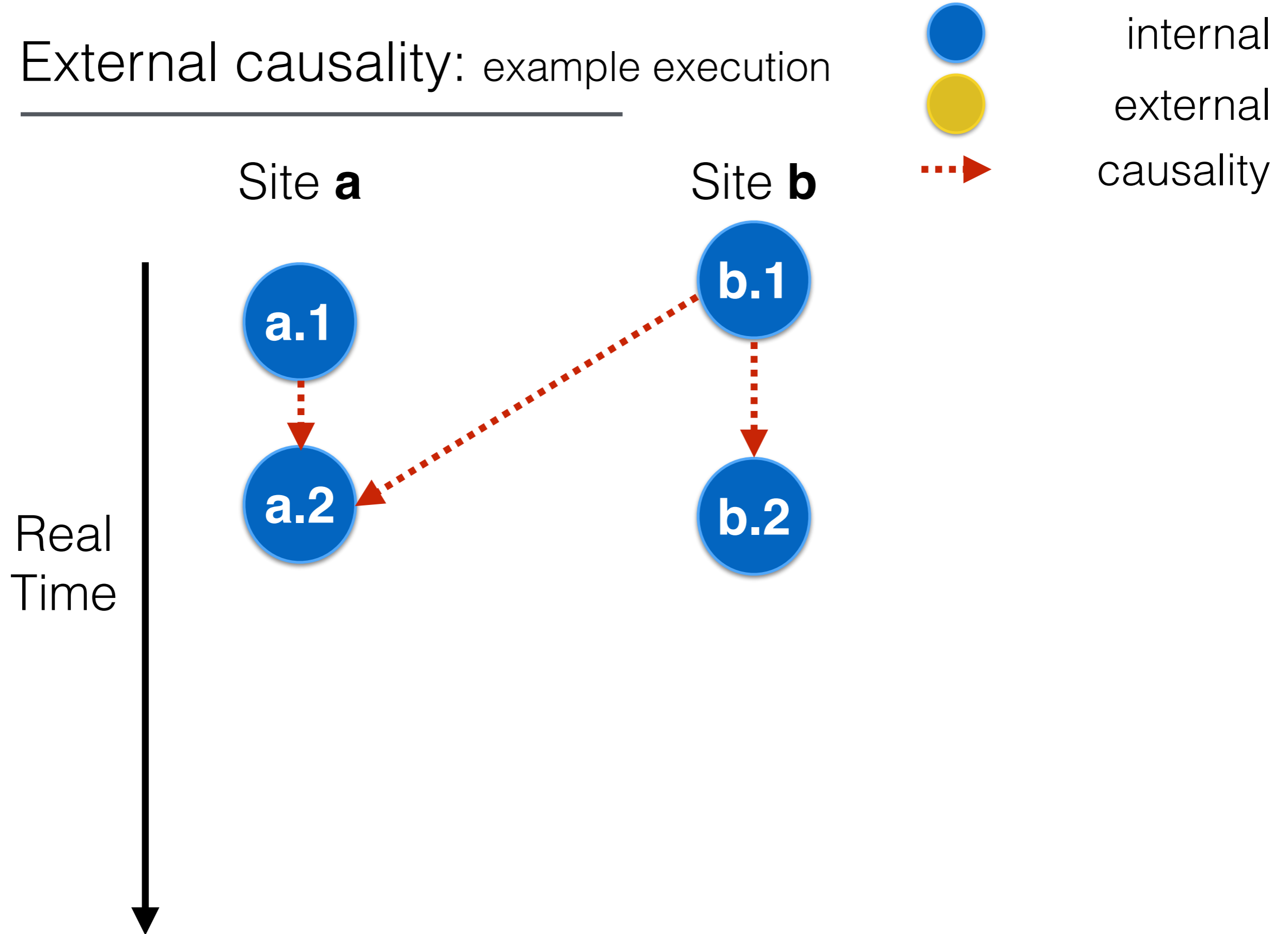
External causality: example execution



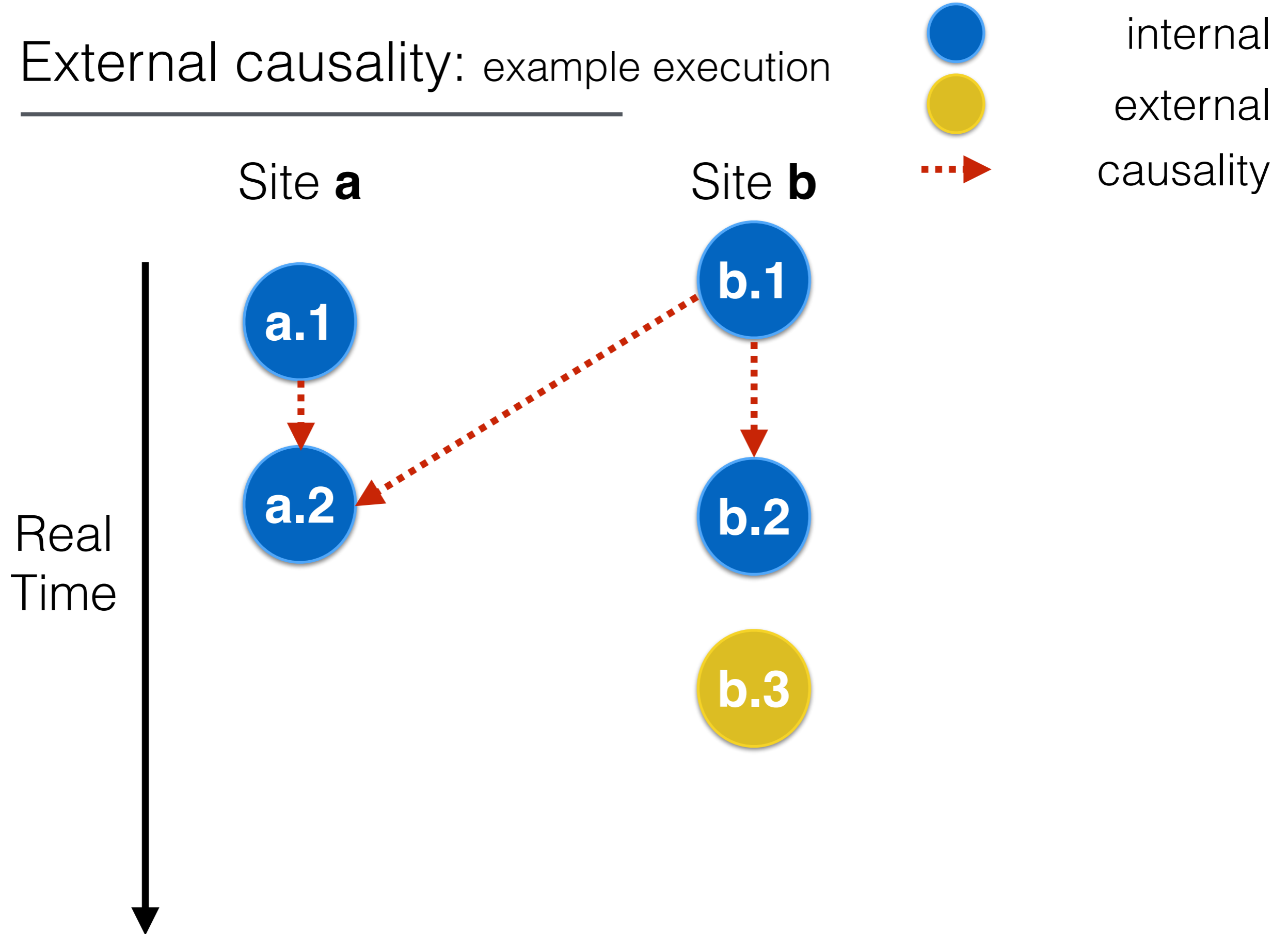
External causality: example execution



External causality: example execution

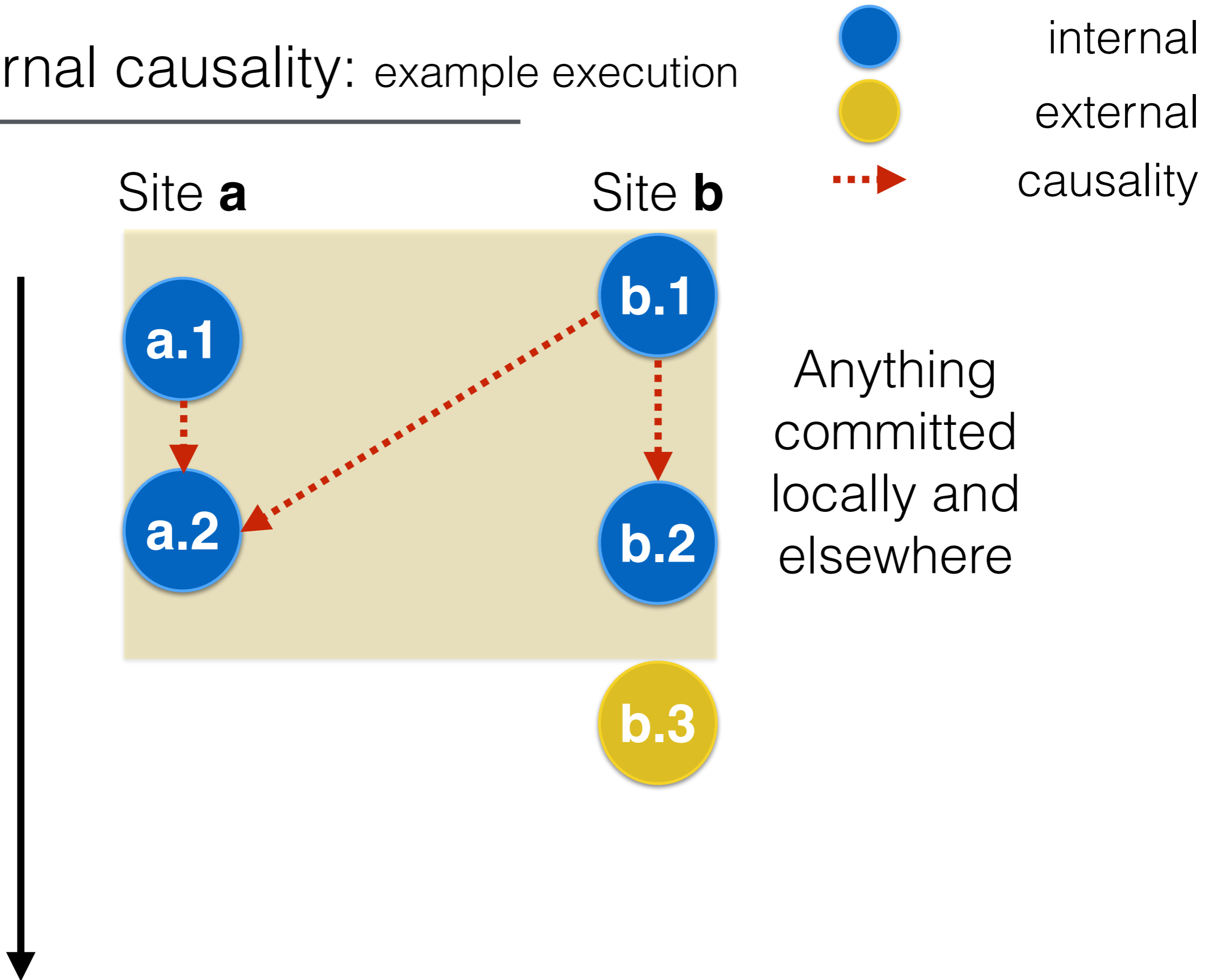


External causality: example execution



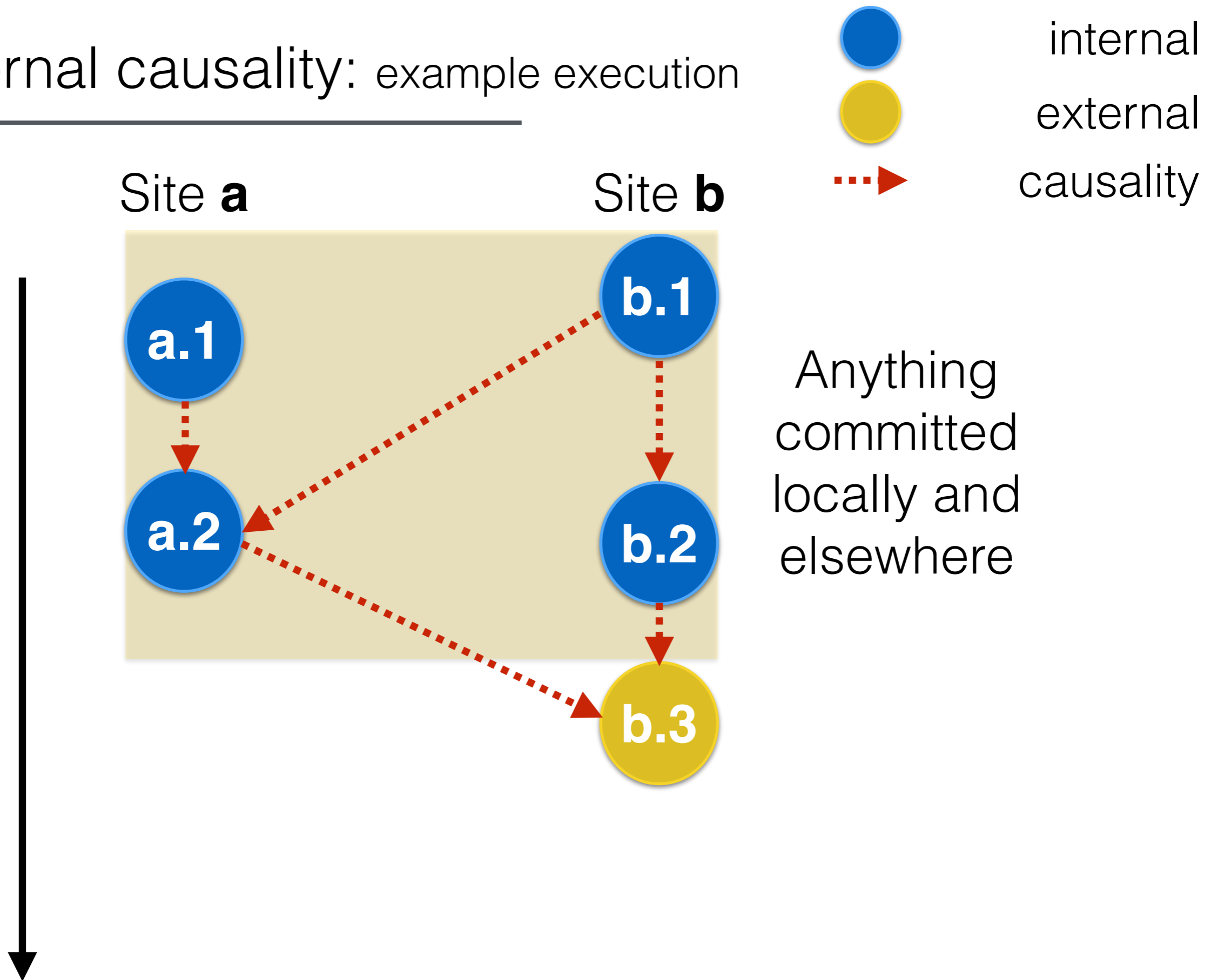
External causality: example execution

Real
Time

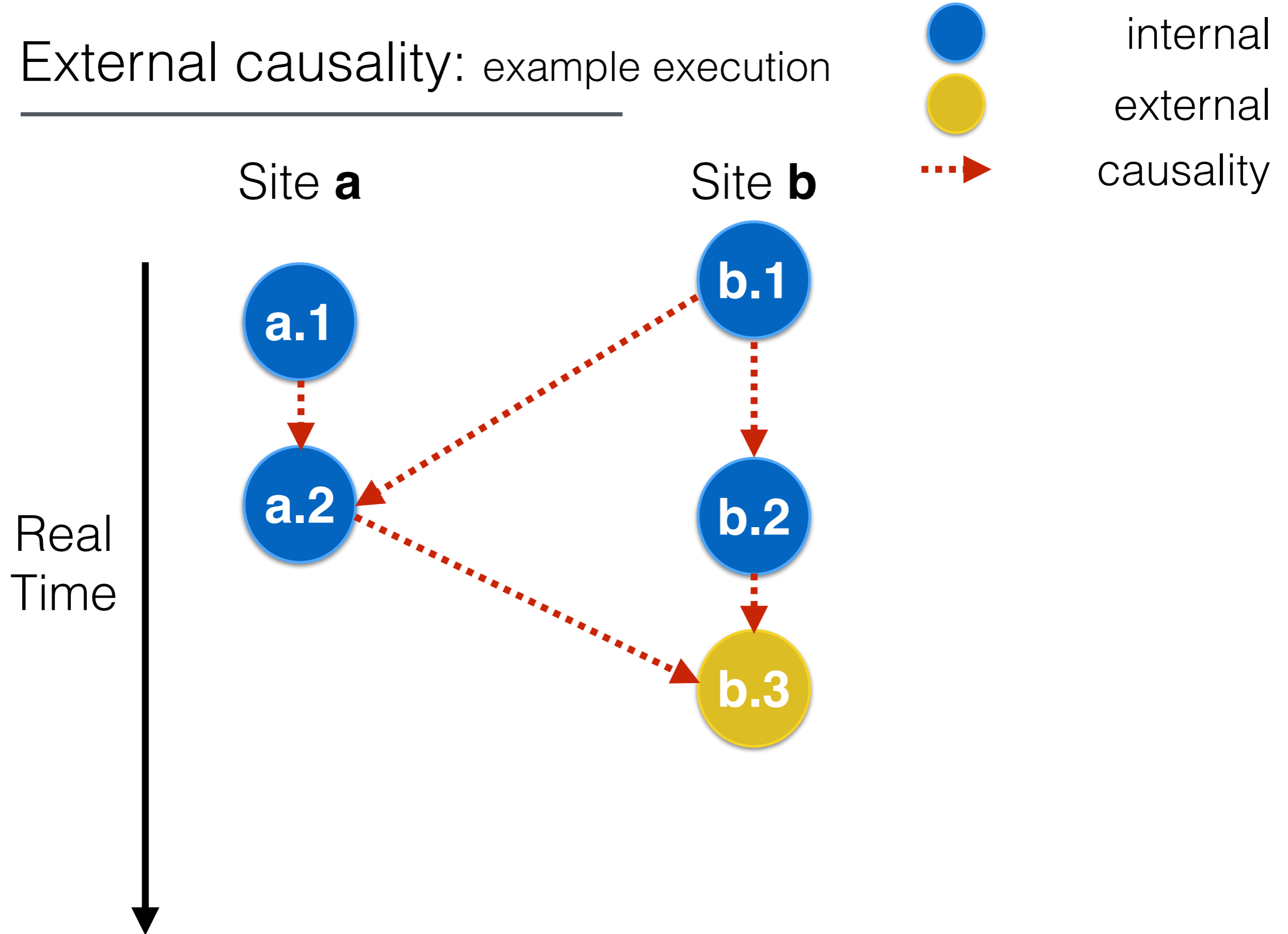


External causality: example execution

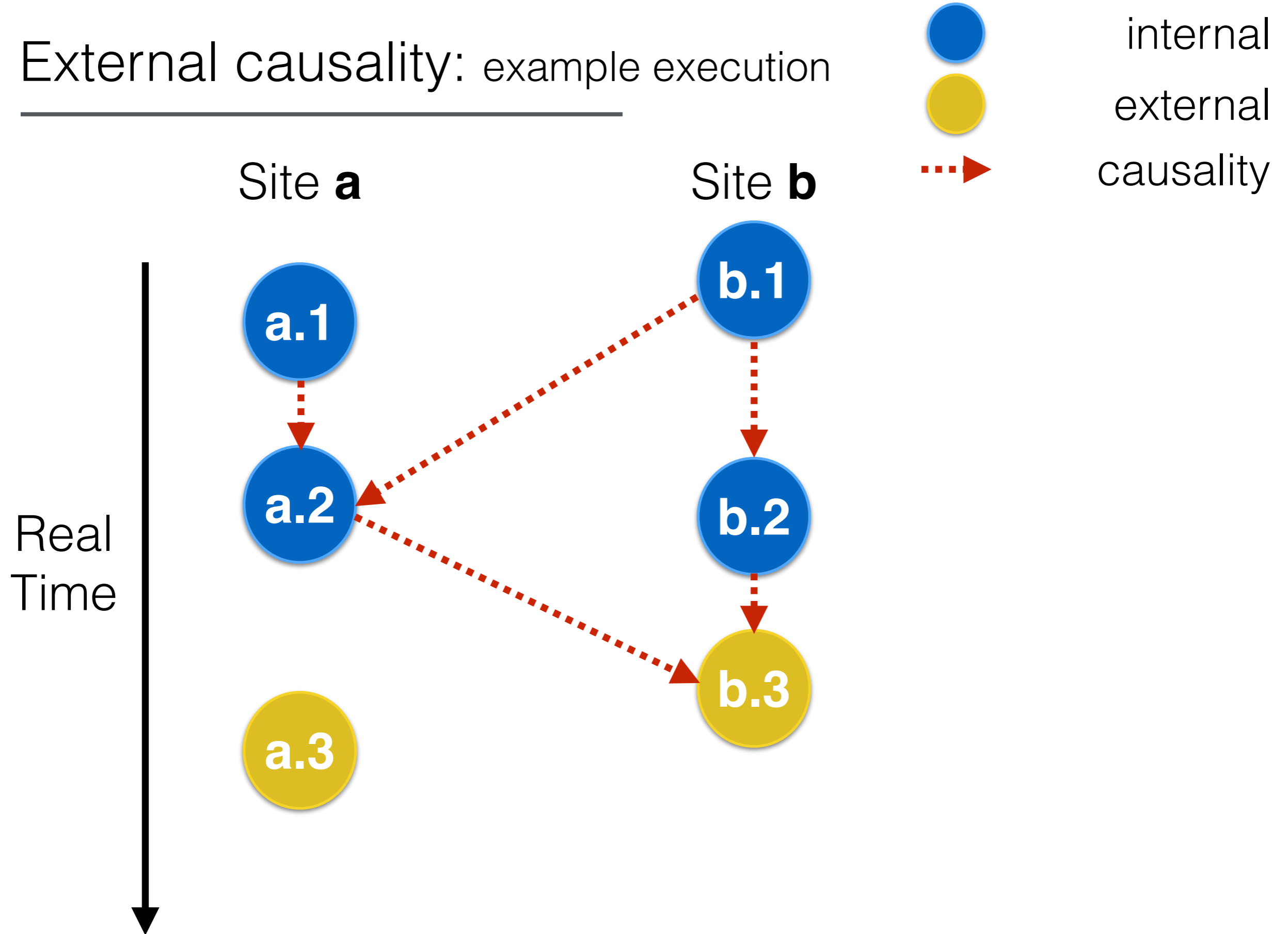
Real
Time



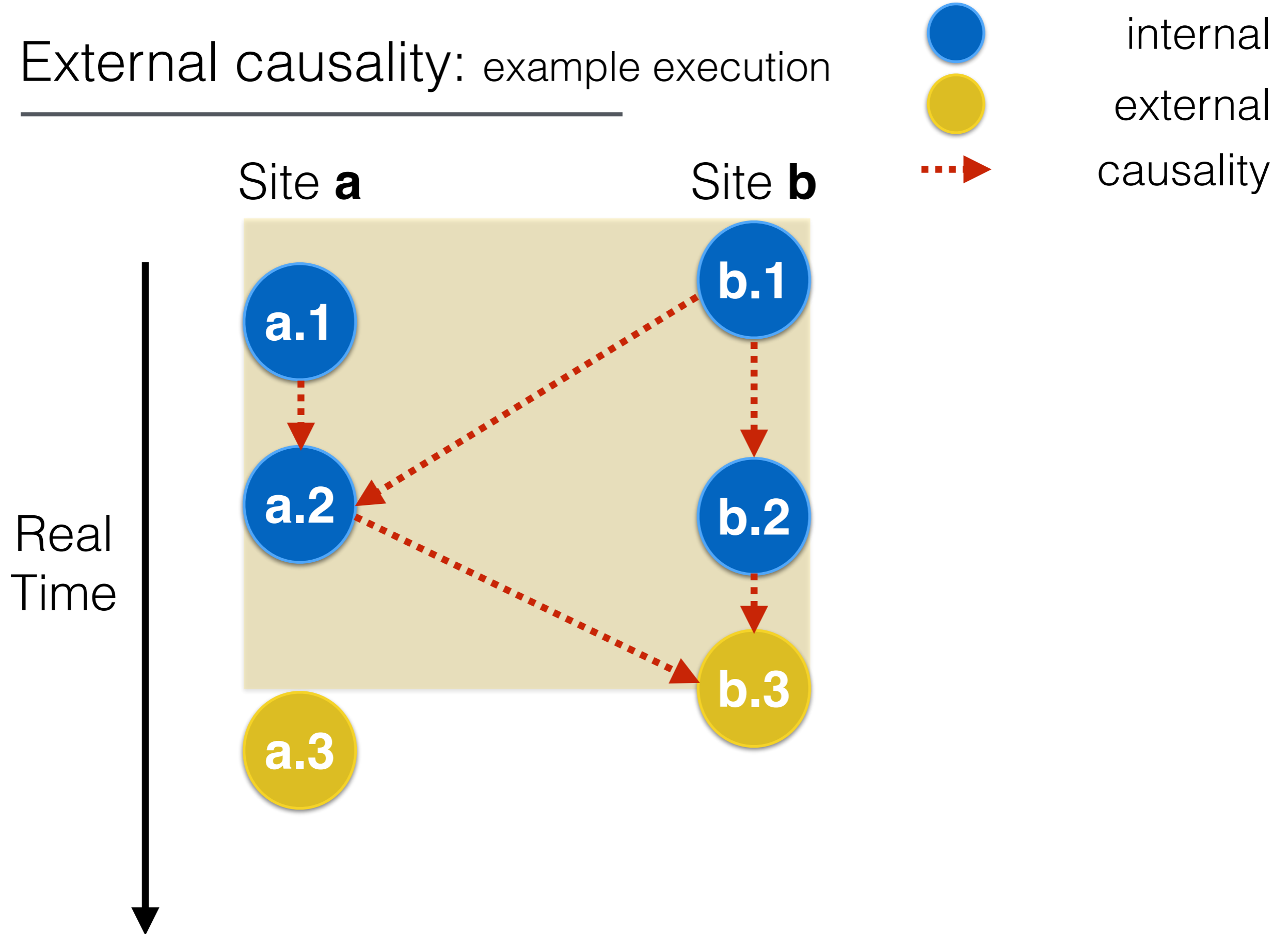
External causality: example execution



External causality: example execution

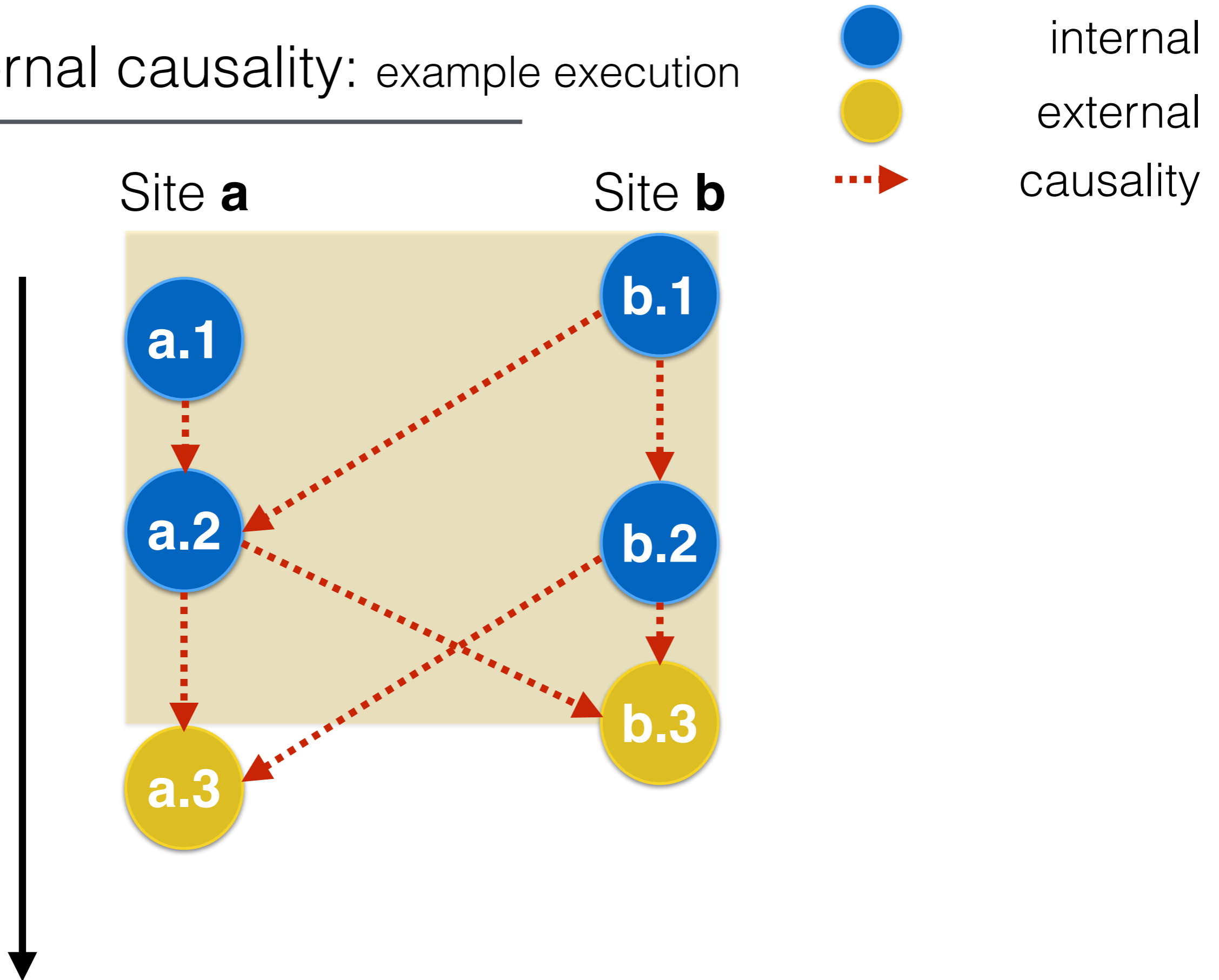


External causality: example execution

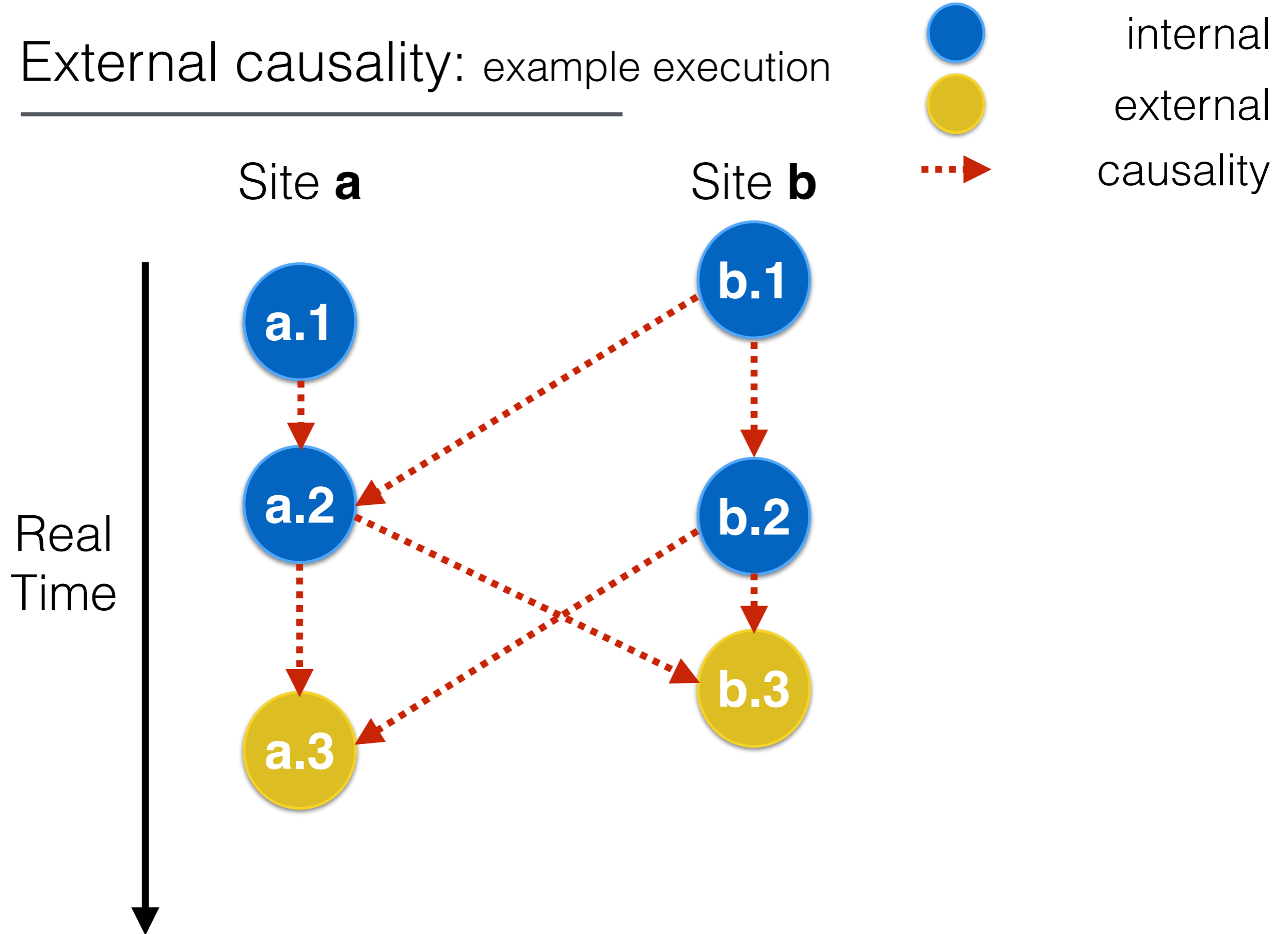


External causality: example execution

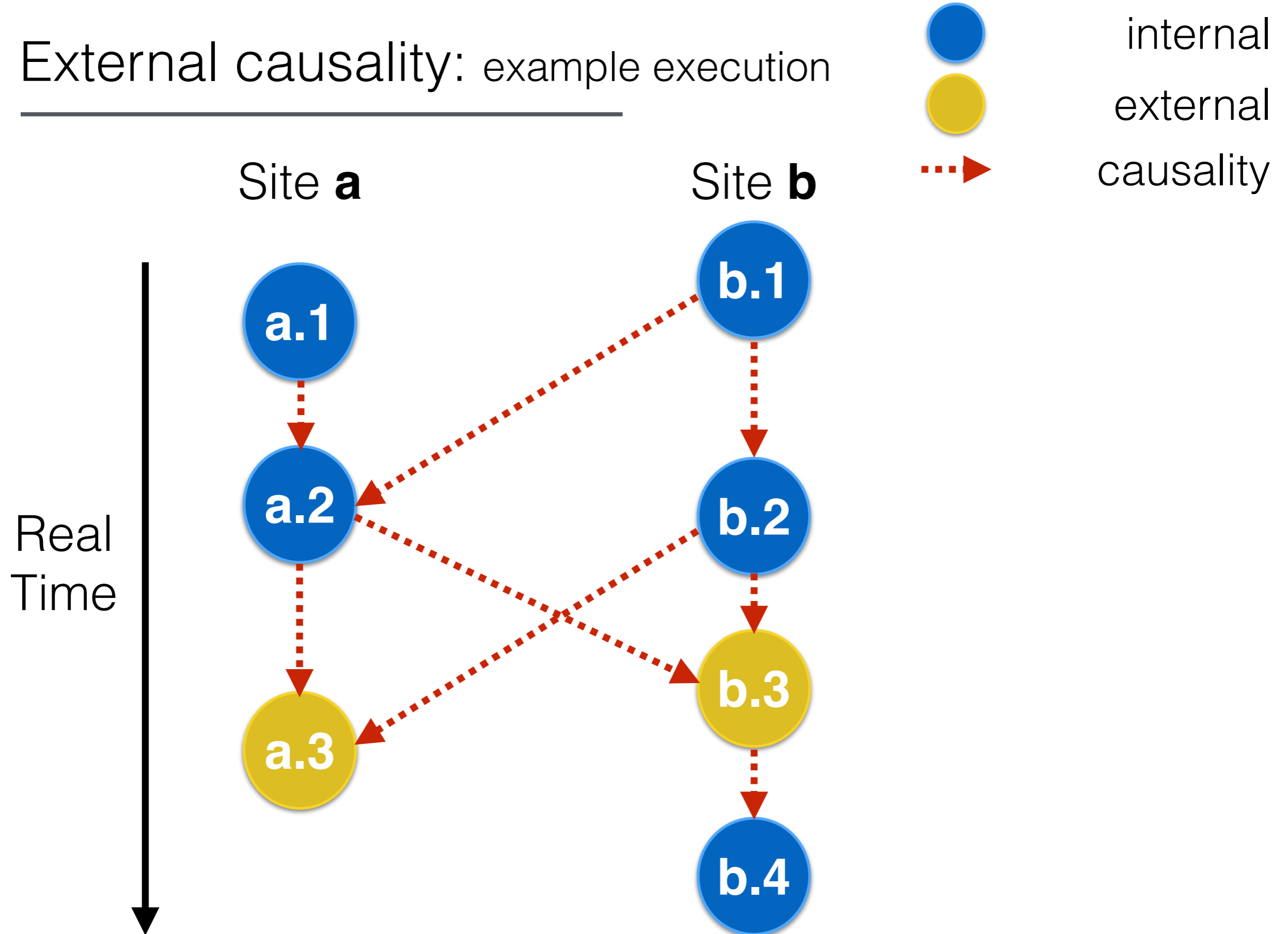
Real
Time



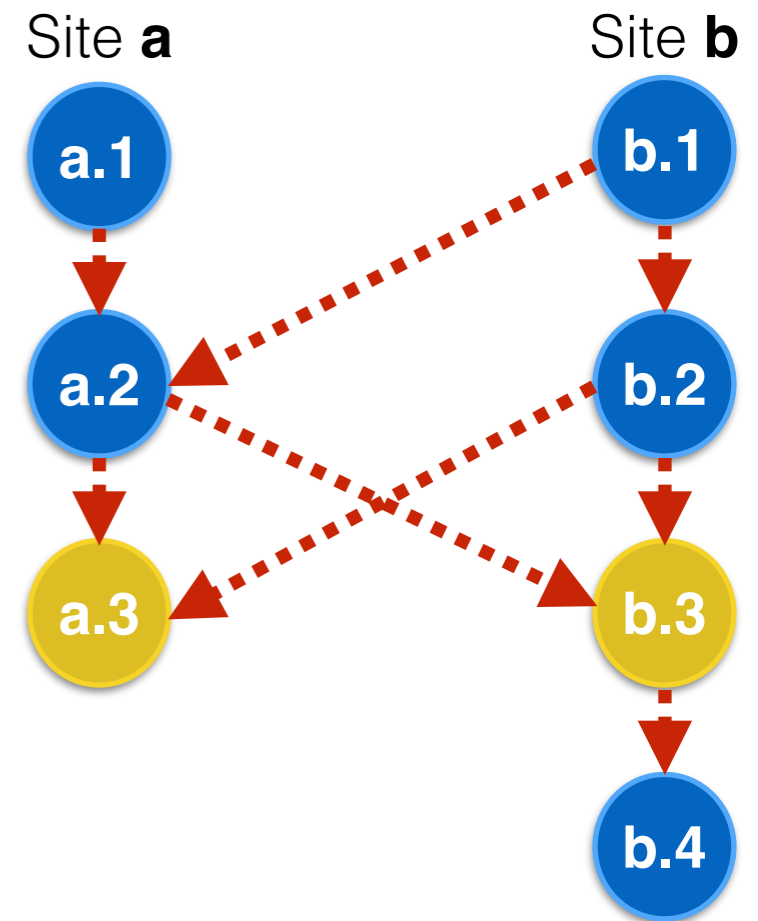
External causality: example execution



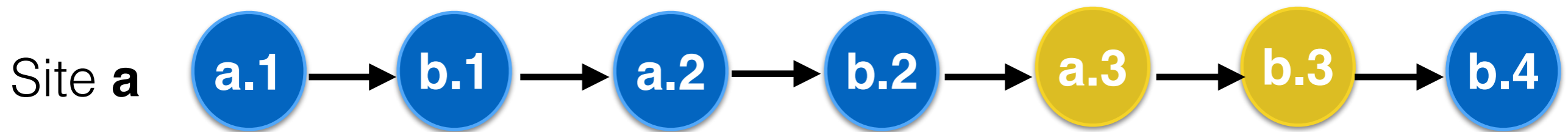
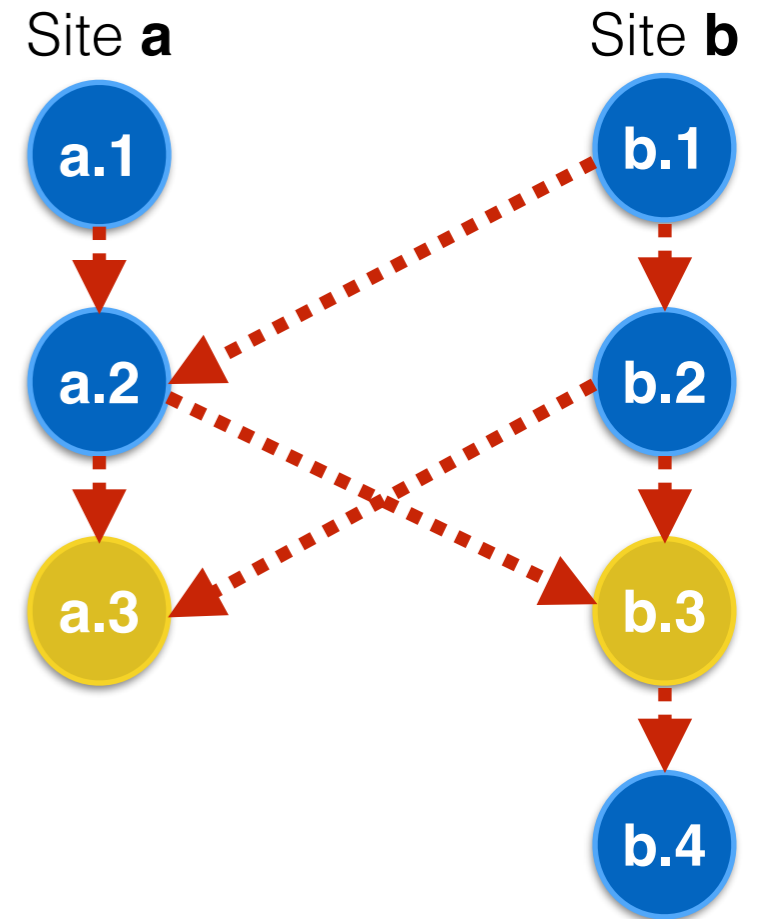
External causality: example execution



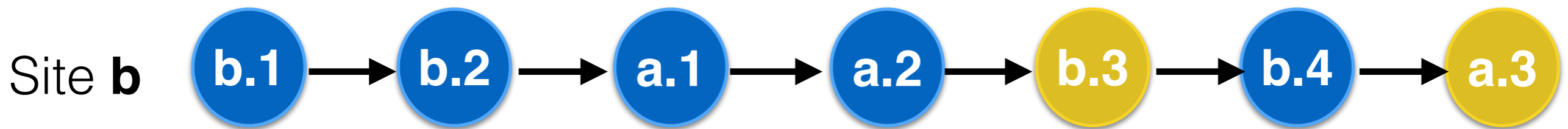
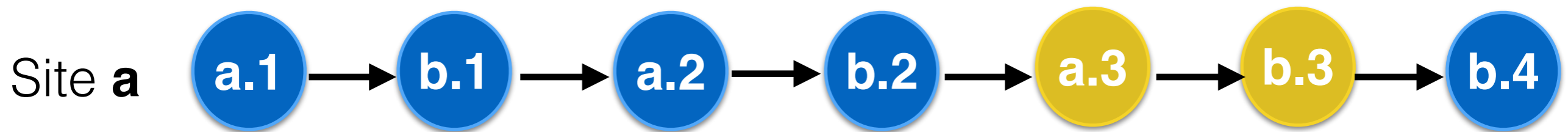
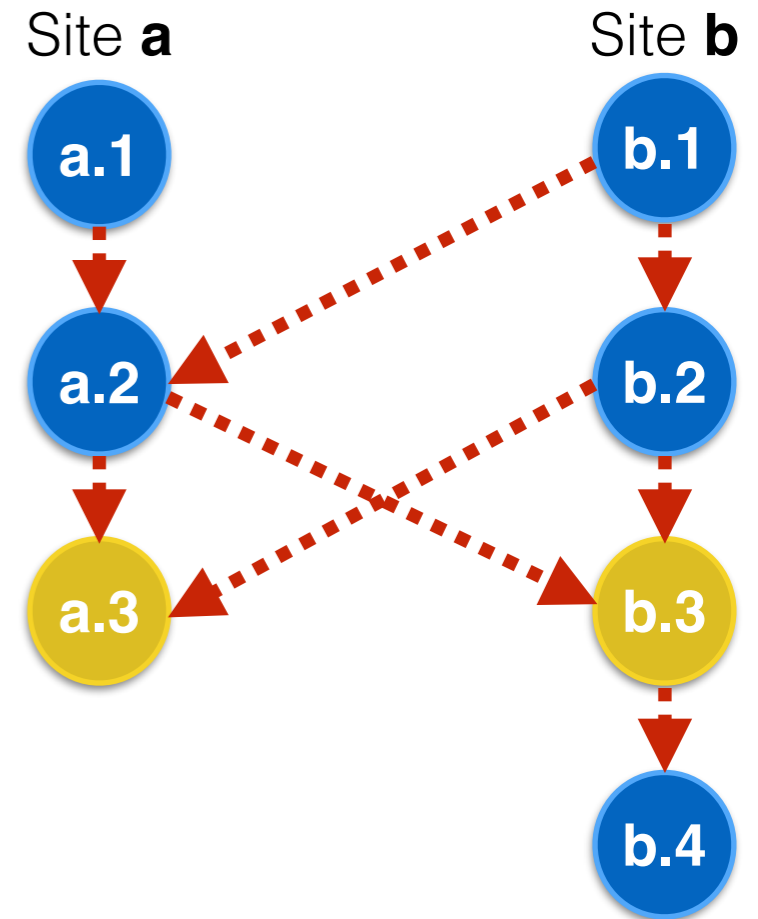
External causality: causal serialisations of



External causality: causal serialisations of



External causality: causal serialisations of



Practical issues

internal operations can read from the local site,
without prior communication

external operations require prior communication—
prepare phase—to compute the external snapshot
from which they are gonna read from

to ensure external guarantees, the write quorum of
both internal and external operations **must overlap**
with the quorum use by external operations in its
prepare phase

Let's recap: an auction service

two roles: auctioneers and buyers

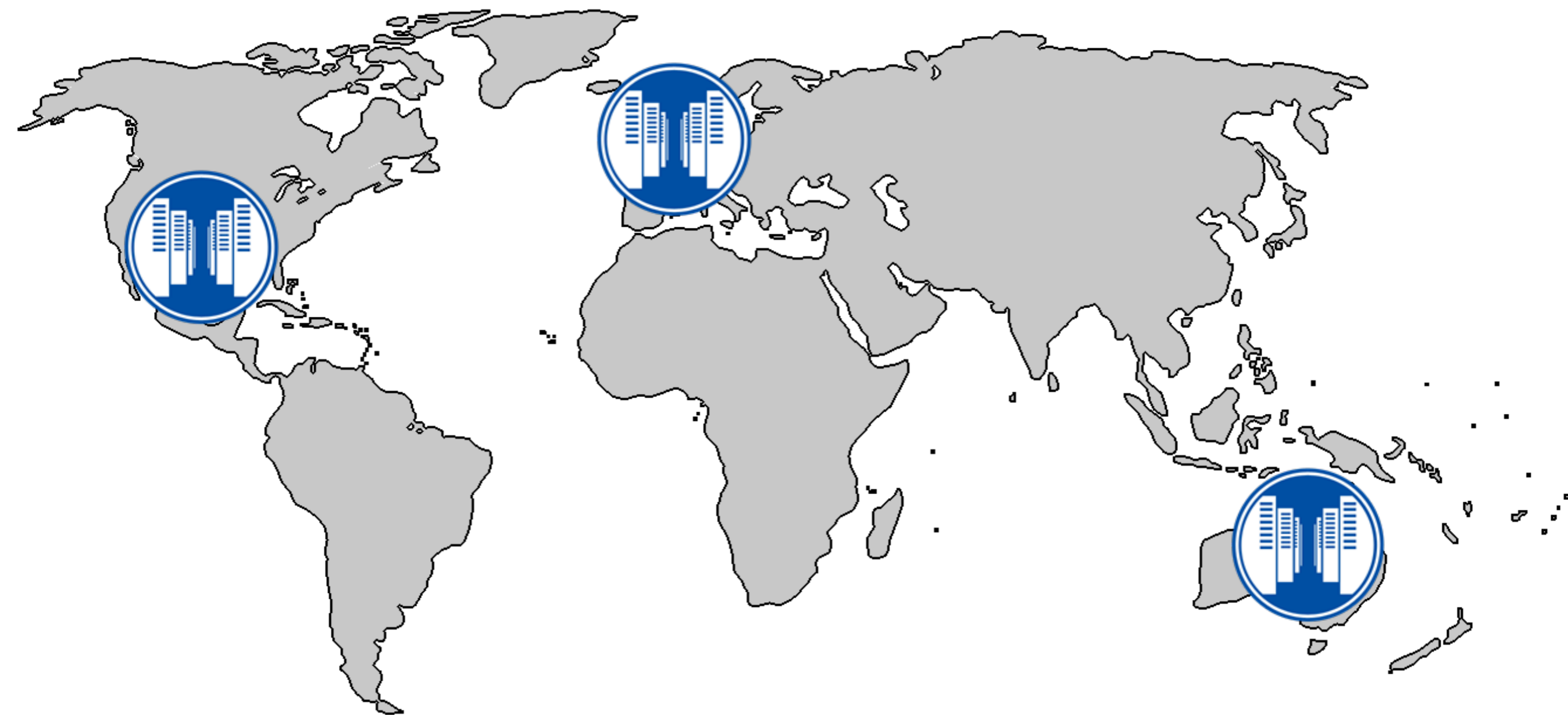
operations (among many): start an auction, place a bid, close an auction, report the winner

Let's recap: an auction service

two roles: auctioneers and buyers

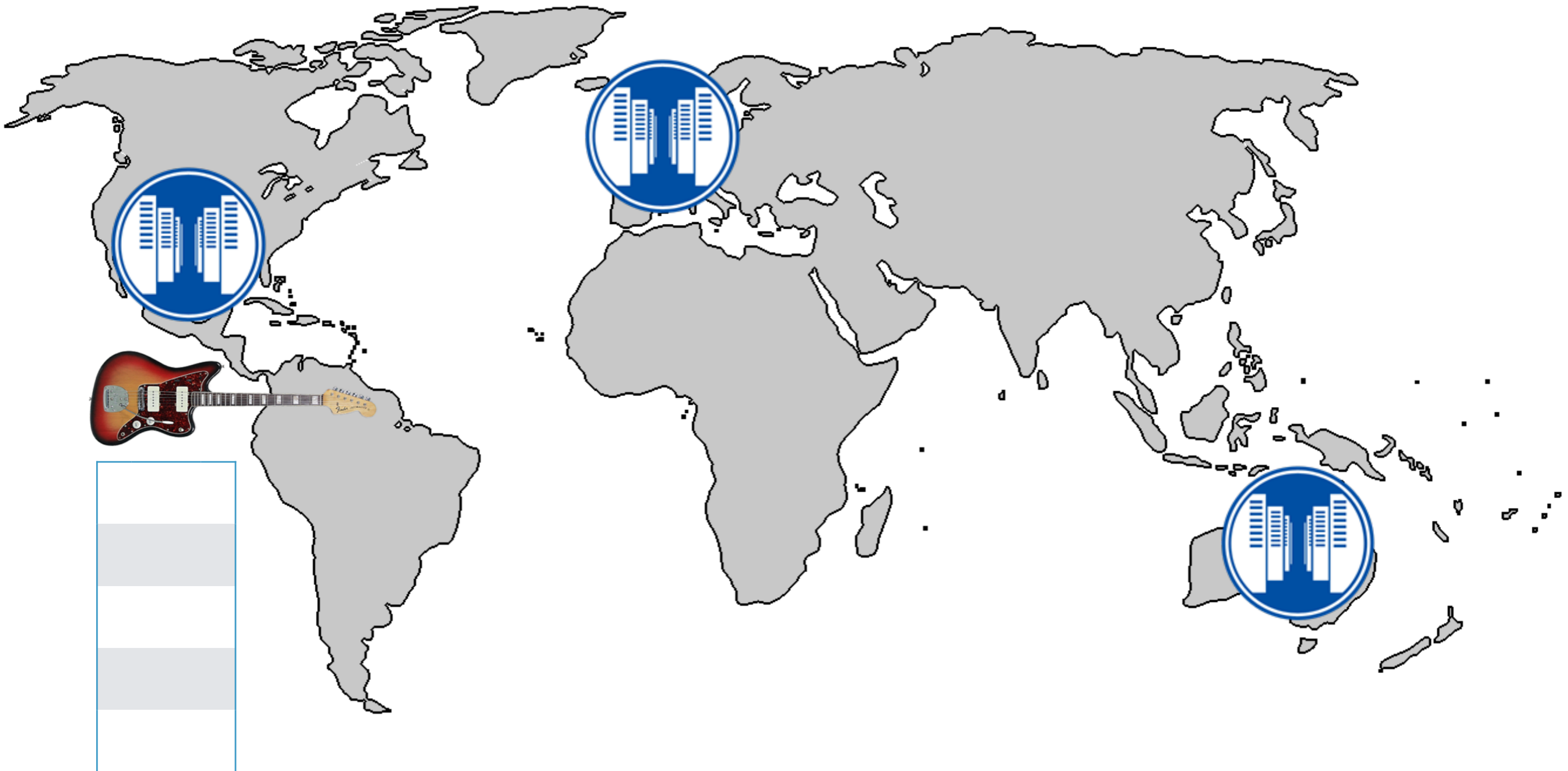
operations (among many): **start an auction**,
place a bid, **close an auction**, **report the winner**

Let's recap: an auction service



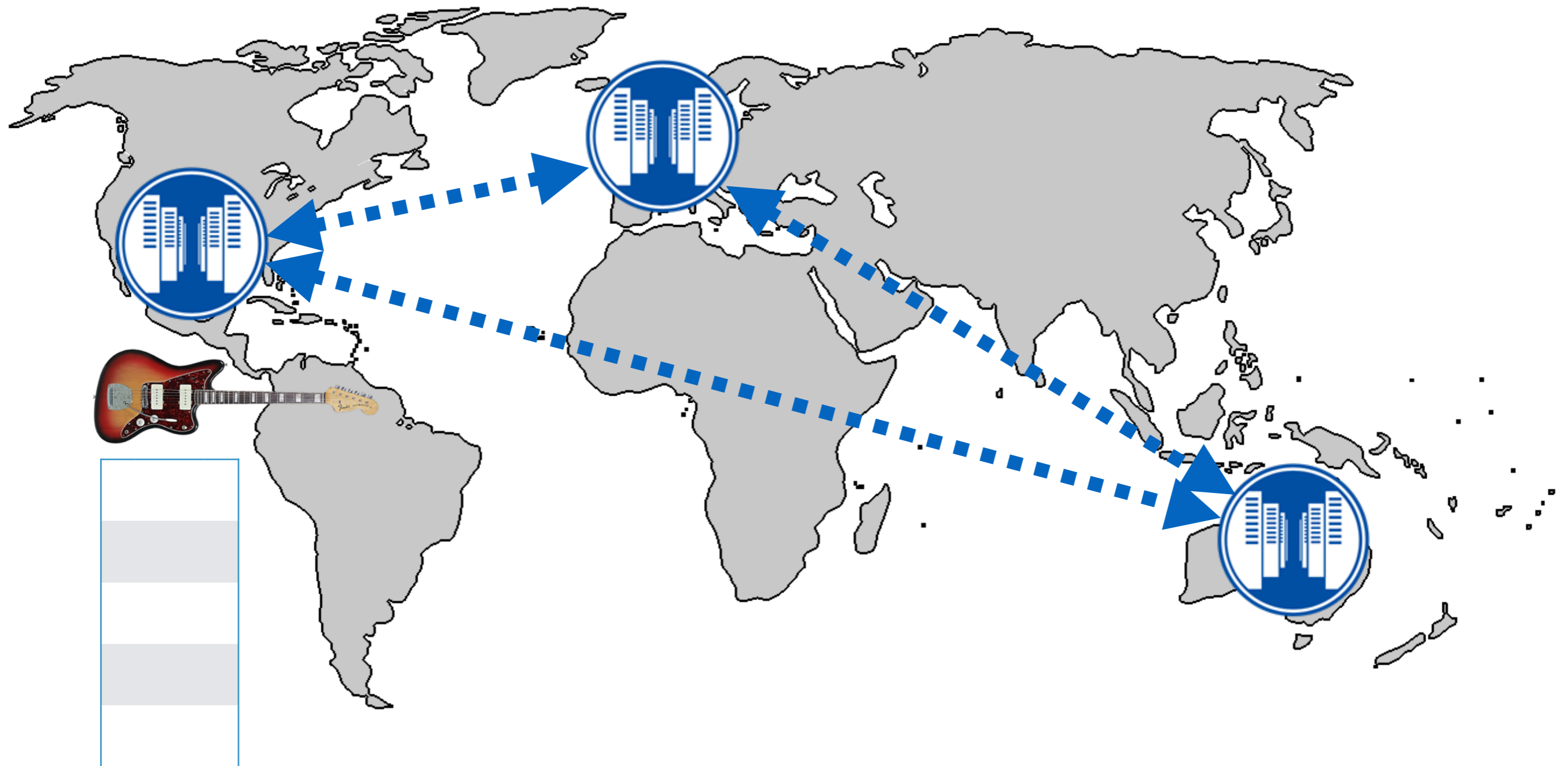
Let's recap: an auction service

start auction!



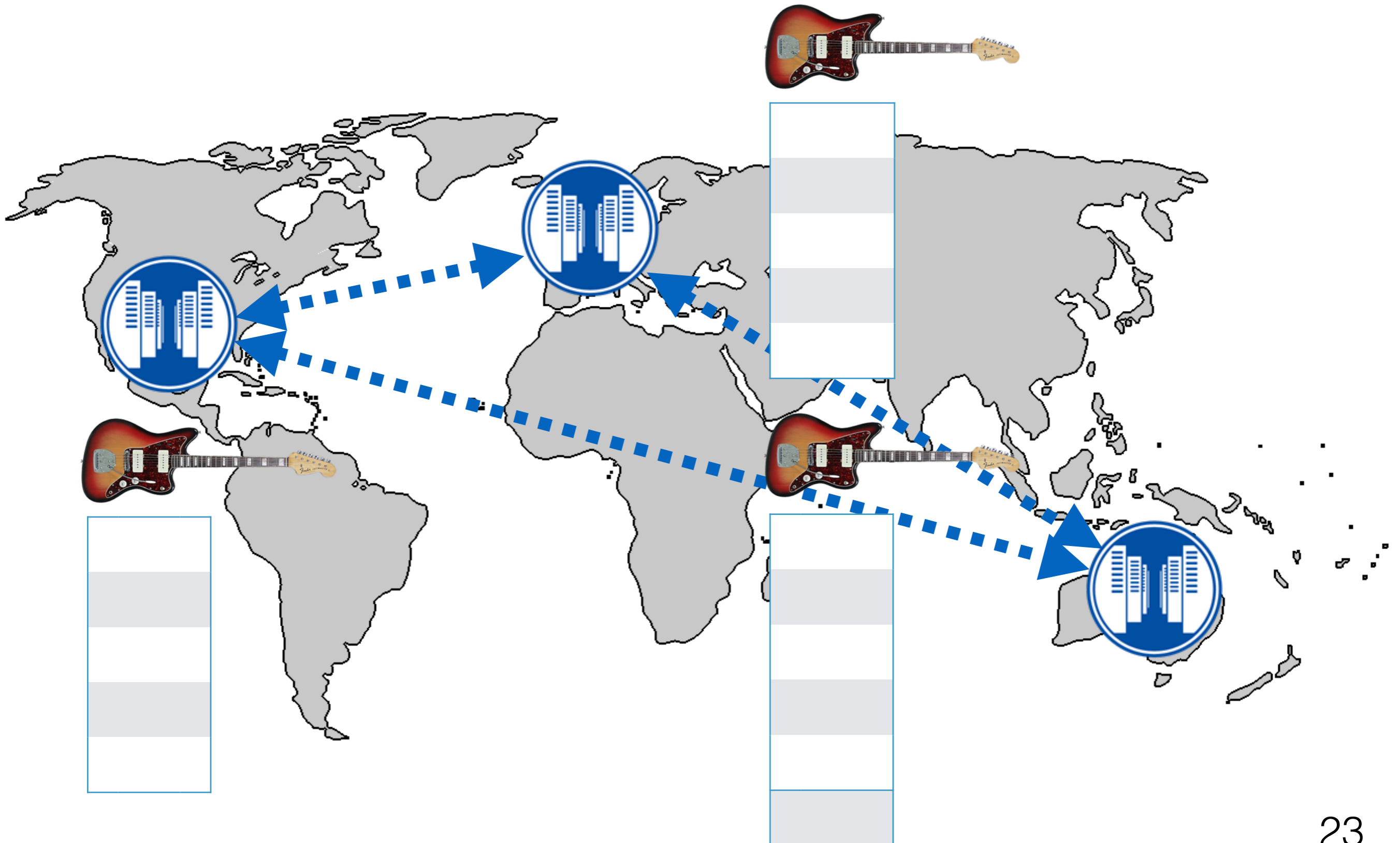
Let's recap: an auction service

start auction!



Let's recap: an auction service

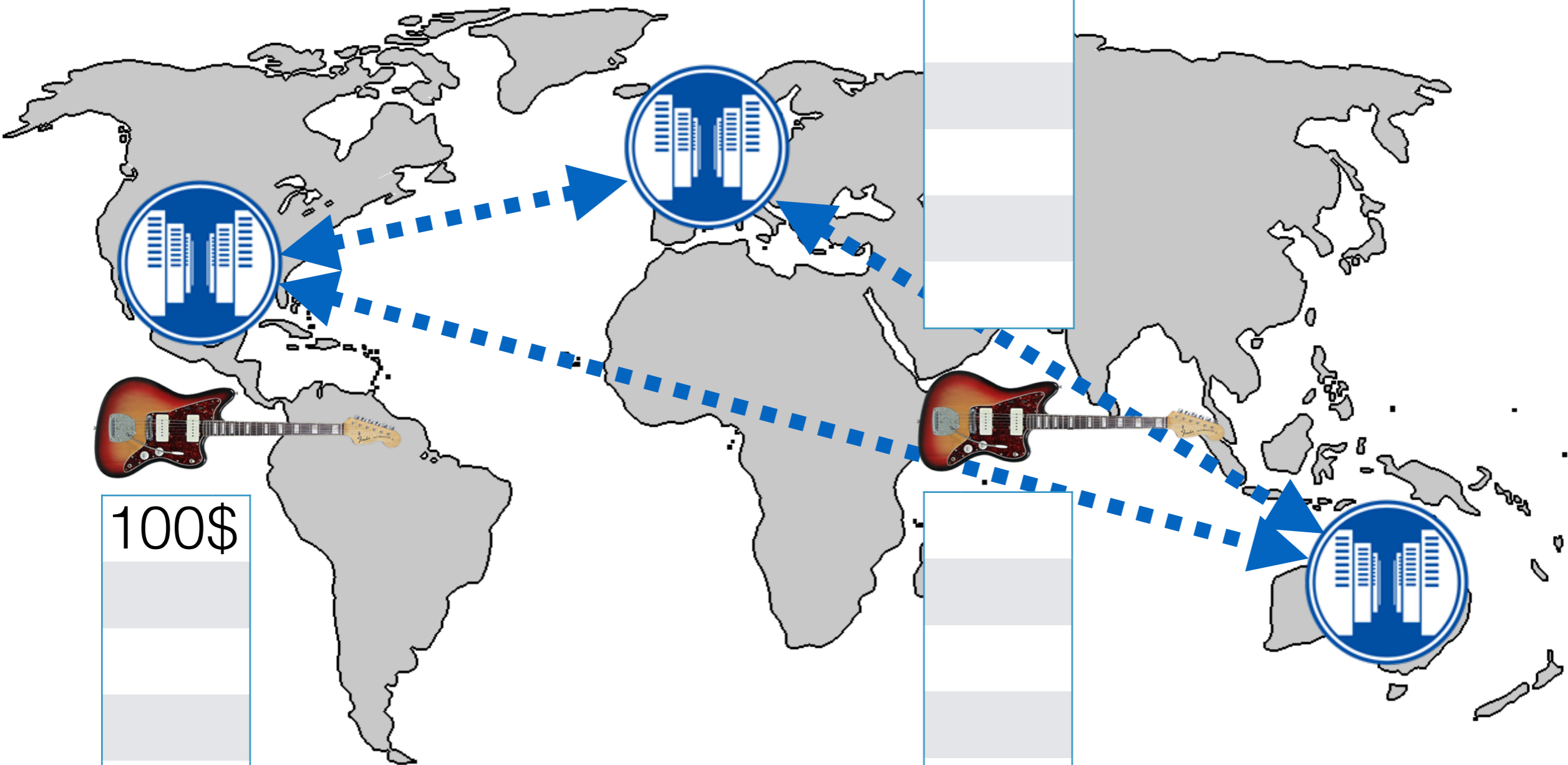
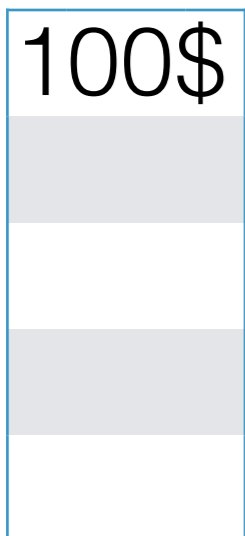
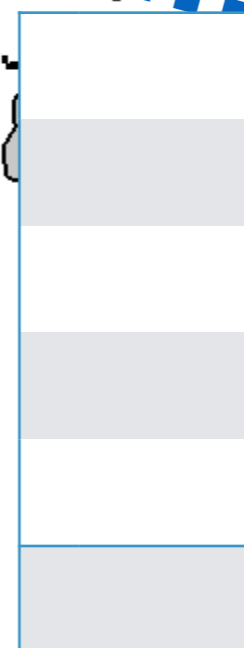
start auction!



Let's recap: an auction service

start auction!

bid!



Let's recap: an auction service

start auction!

bid!



100\$



100\$



100\$



Let's recap: an auction service

start auction!

bid! bid! bid!



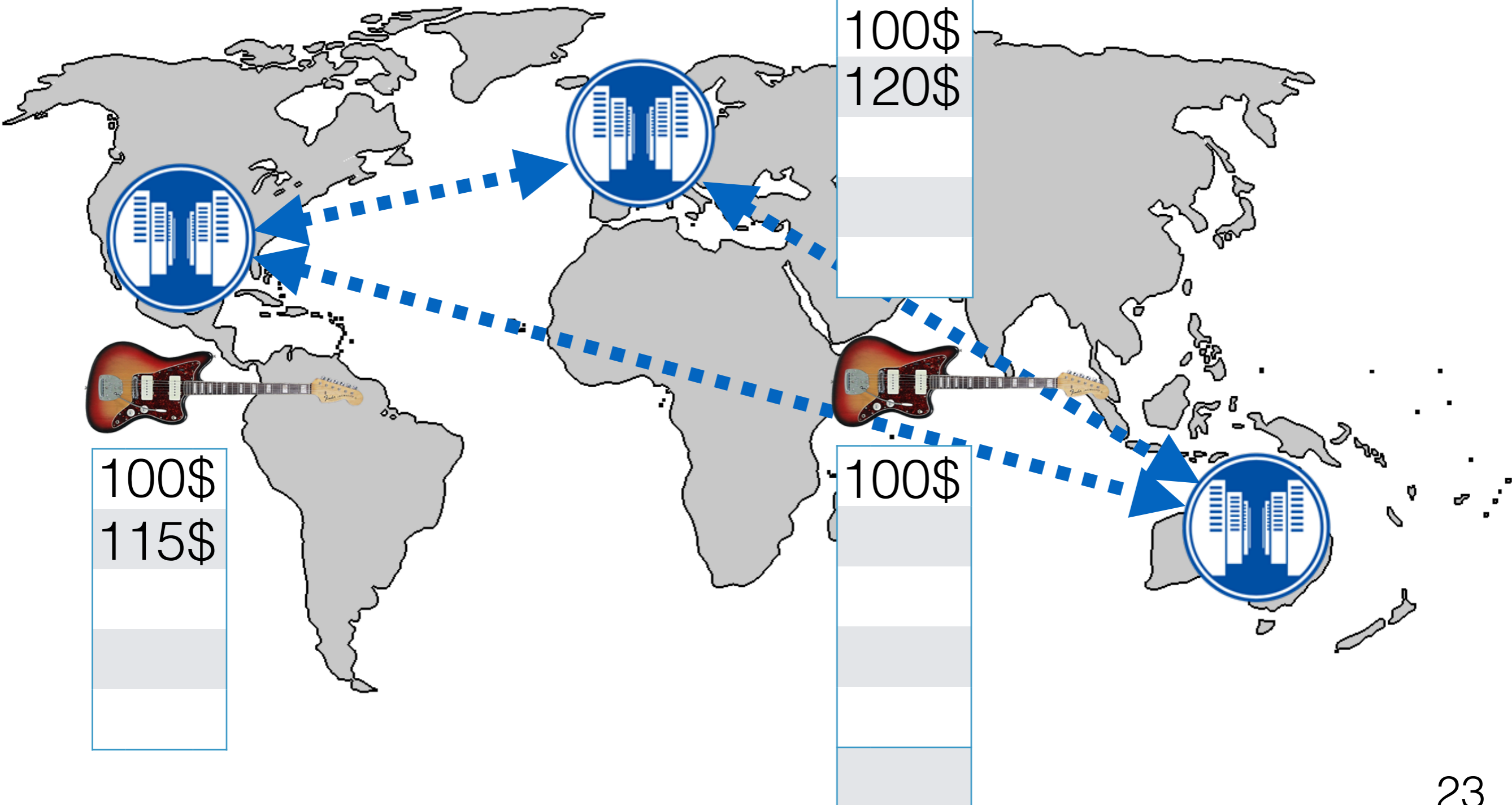
100\$
120\$



100\$



100\$
115\$



Let's recap: an auction service

start auction!

bid! bid! bid!



100\$
120\$
115\$



100\$
115\$
120\$



100\$
120\$
115\$



Let's recap: an auction service

start auction!

**bid! bid! bid!
bid!**



100\$
120\$
115\$



100\$
115\$
120\$



100\$
120\$
115\$
130\$



Let's recap: an auction service

start auction!

**bid! bid! bid!
bid! bid!**



100\$
120\$
115\$



100\$
115\$
120\$
125\$



100\$
120\$
115\$
130\$



Let's recap: an auction service

start auction!

**bid! bid! bid!
bid! bid!**



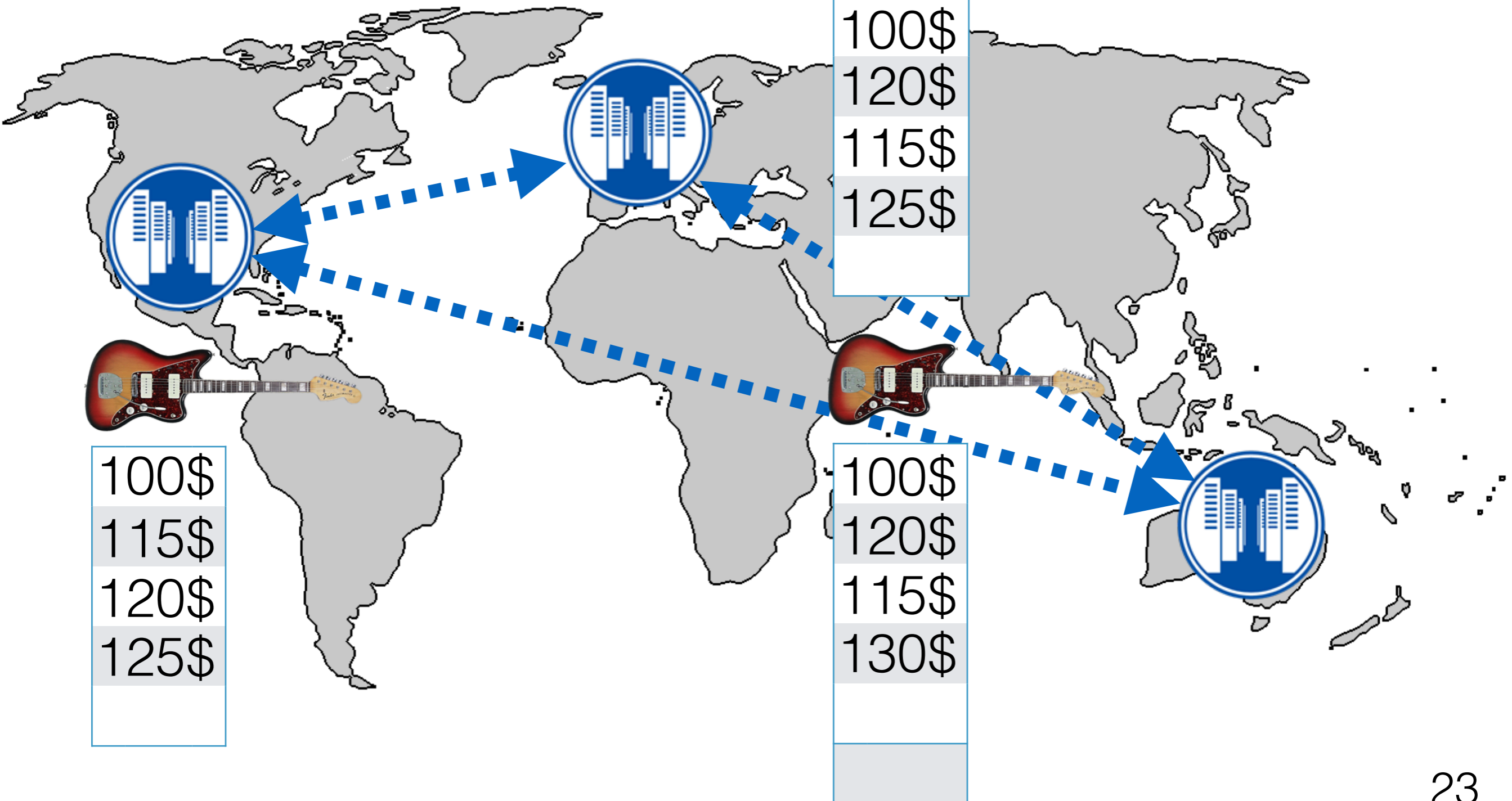
100\$
120\$
115\$
125\$



100\$
115\$
120\$
125\$



100\$
120\$
115\$
130\$



Let's recap: an auction service

start auction!

**bid! bid! bid!
bid! bid!**



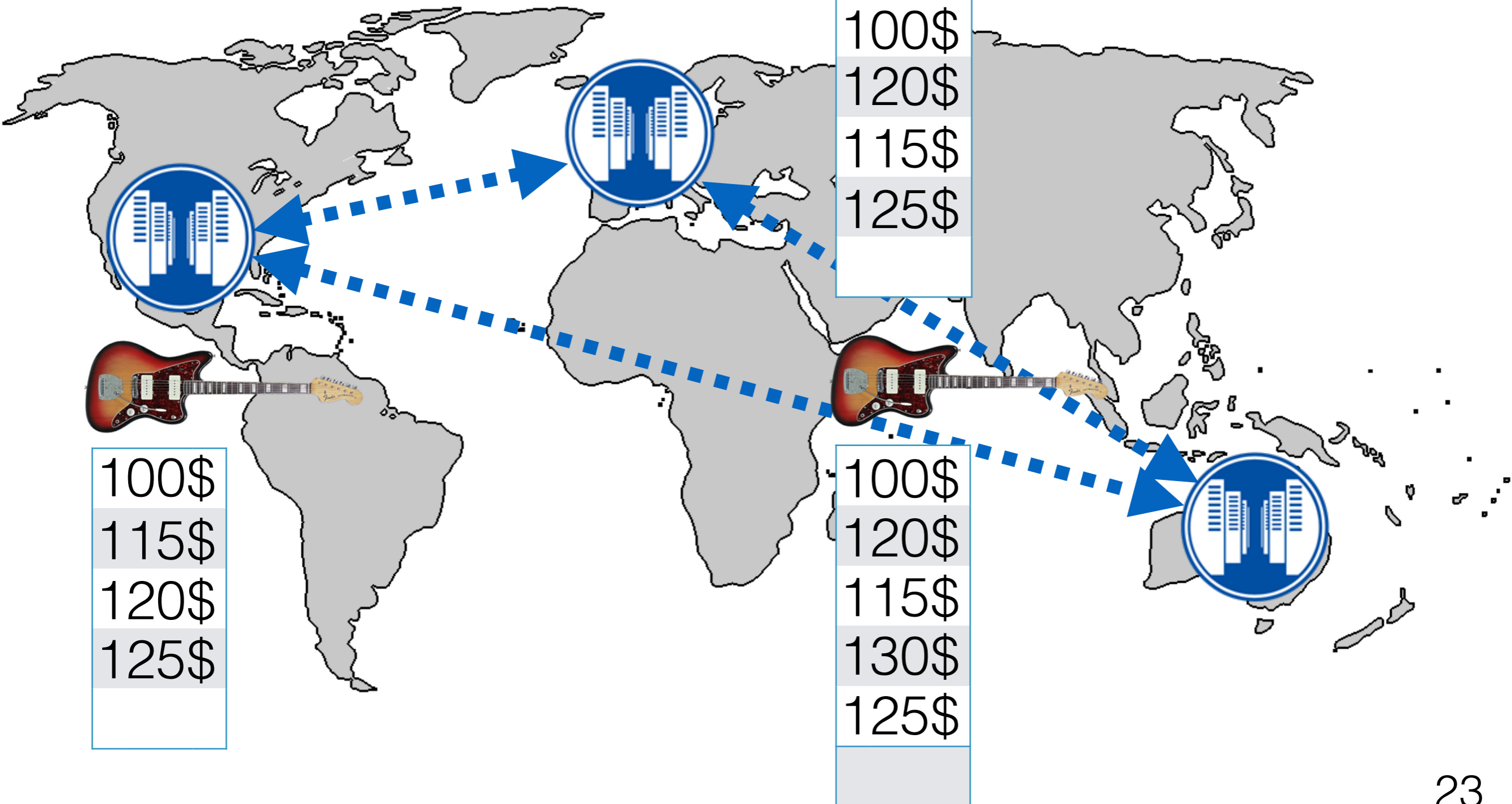
100\$
120\$
115\$
125\$



100\$
115\$
120\$
125\$



100\$
120\$
115\$
130\$
125\$



Let's recap: an auction service

start auction!

**bid! bid! bid!
bid! bid!**

close auction!



100\$
120\$
115\$
125\$



100\$
115\$
120\$
125\$
✗



100\$
120\$
115\$
130\$
125\$

Let's recap: an auction service

start auction!

**bid! bid! bid!
bid! bid!**

close auction!



100\$
120\$
115\$
125\$
X



100\$
115\$
120\$
125\$
X



100\$
120\$
115\$
130\$
125\$
X



Let's recap: an auction service

start auction!

**bid! bid! bid!
bid! bid!**

close auction!



100\$
120\$
115\$
125\$
X

130\$



100\$
120\$
115\$
130\$
125\$
X

130\$



100\$
115\$
120\$
125\$
X

Let's recap: an auction service

start auction!

**bid! bid! bid!
bid! bid!**

close auction!



100\$
120\$
115\$
125\$
X

130\$



100\$
120\$
115\$
130\$
125\$
X

130\$



100\$
115\$
120\$
125\$
X

Let's recap: an auction service

start auction!

**bid! bid! bid!
bid! bid!**

close auction!

bid!



100\$
120\$
115\$
125\$
✗

130\$



100\$
115\$
120\$
125\$
✗

130\$



100\$
120\$
115\$
130\$
125\$
✗



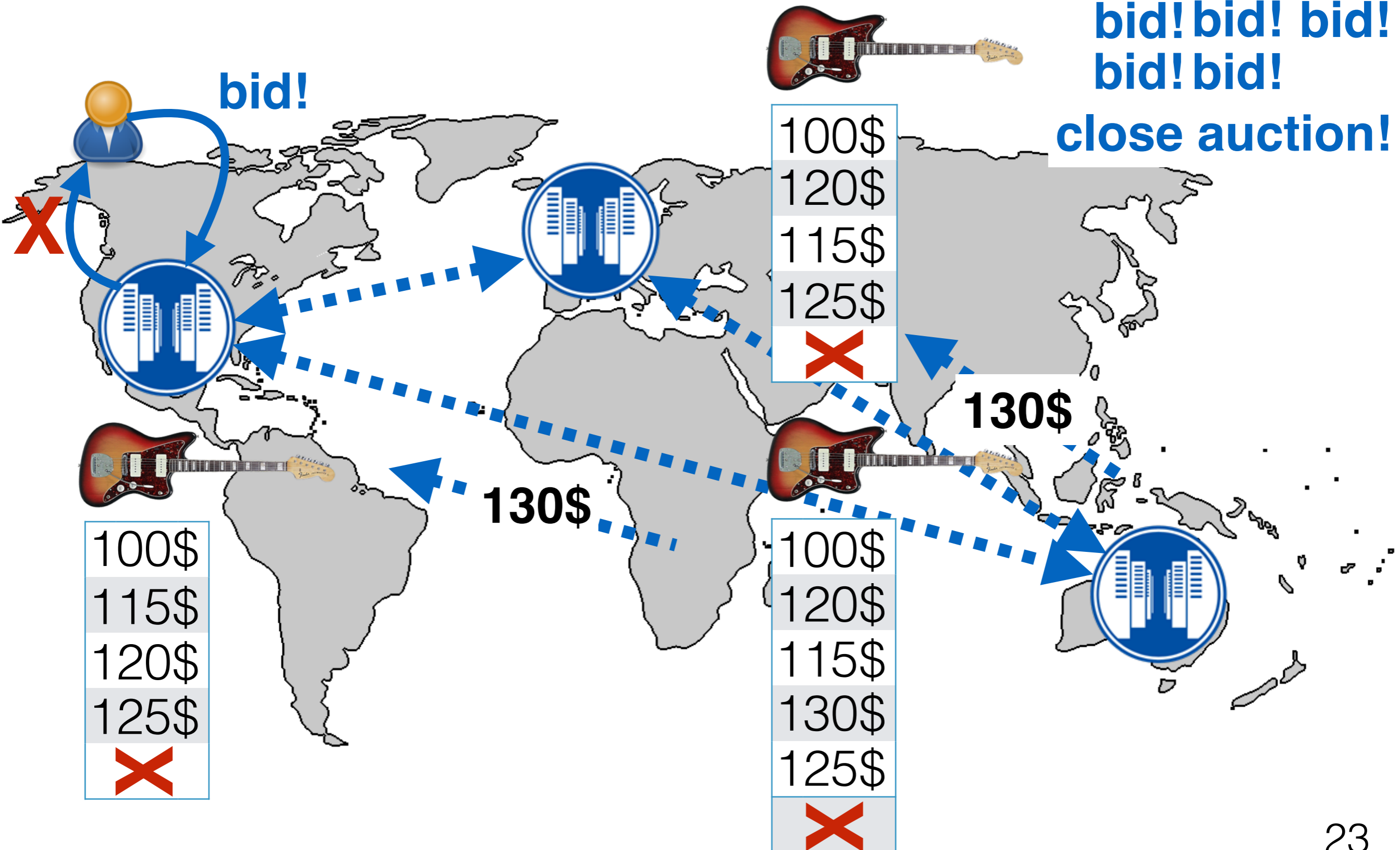
Let's recap: an auction service

start auction!

**bid! bid! bid!
bid! bid!**

close auction!

bid!



Let's recap: an auction service

start auction!

**bid! bid! bid!
bid! bid!**

close auction!



100\$
120\$
115\$
125\$

130\$



100\$
120\$
115\$
130\$
125\$

130\$

report winner!



100\$
115\$
120\$
125\$

Let's recap: an auction service

start auction!

bid! bid! bid!
bid! bid!

close auction!



100\$
120\$
115\$
125\$
X

130\$



100\$
120\$
115\$
130\$
125\$
X

130\$

report winner!



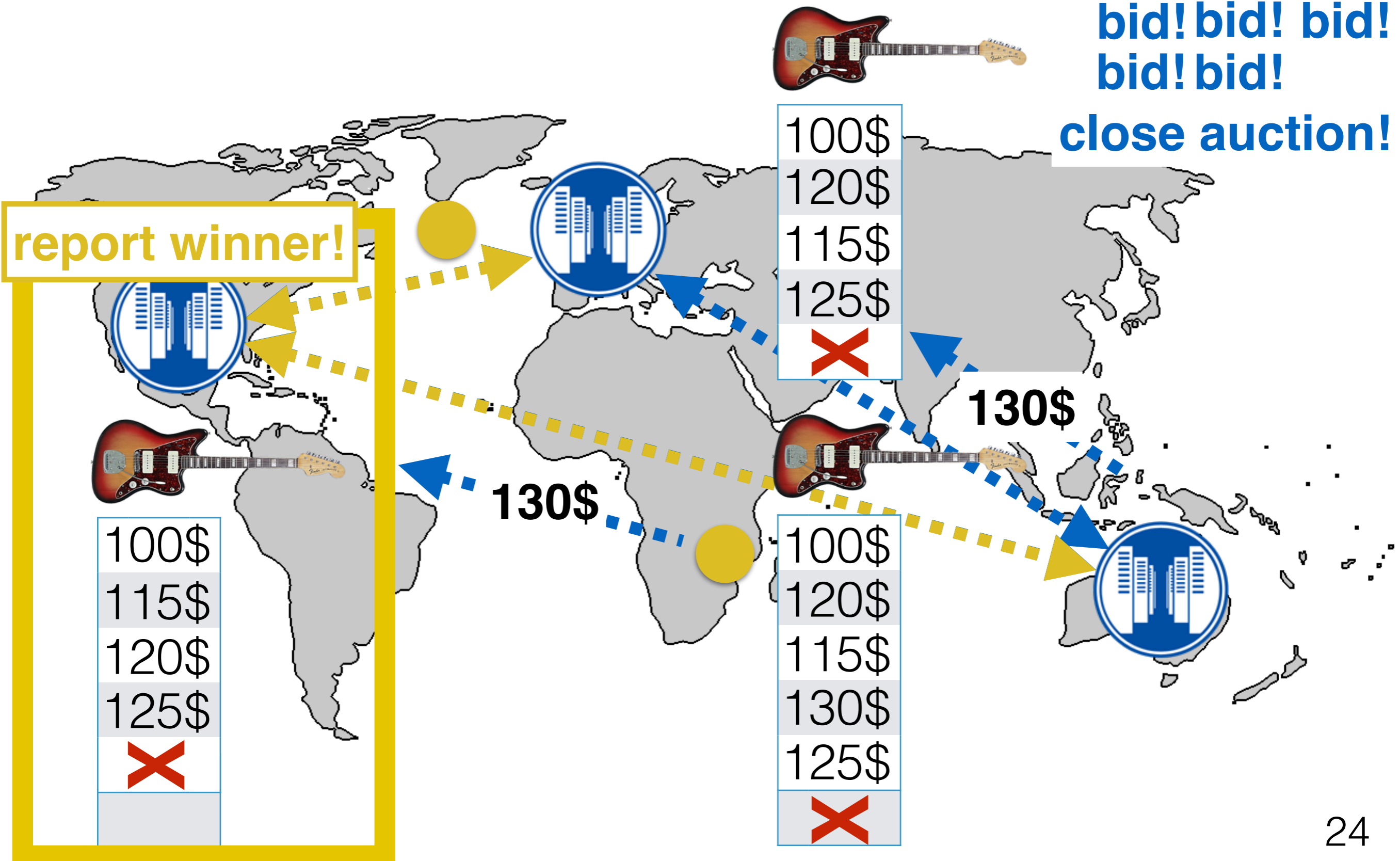
100\$
115\$
120\$
125\$
X

Let's recap: an auction service

start auction!

**bid! bid! bid!
bid! bid!**

close auction!

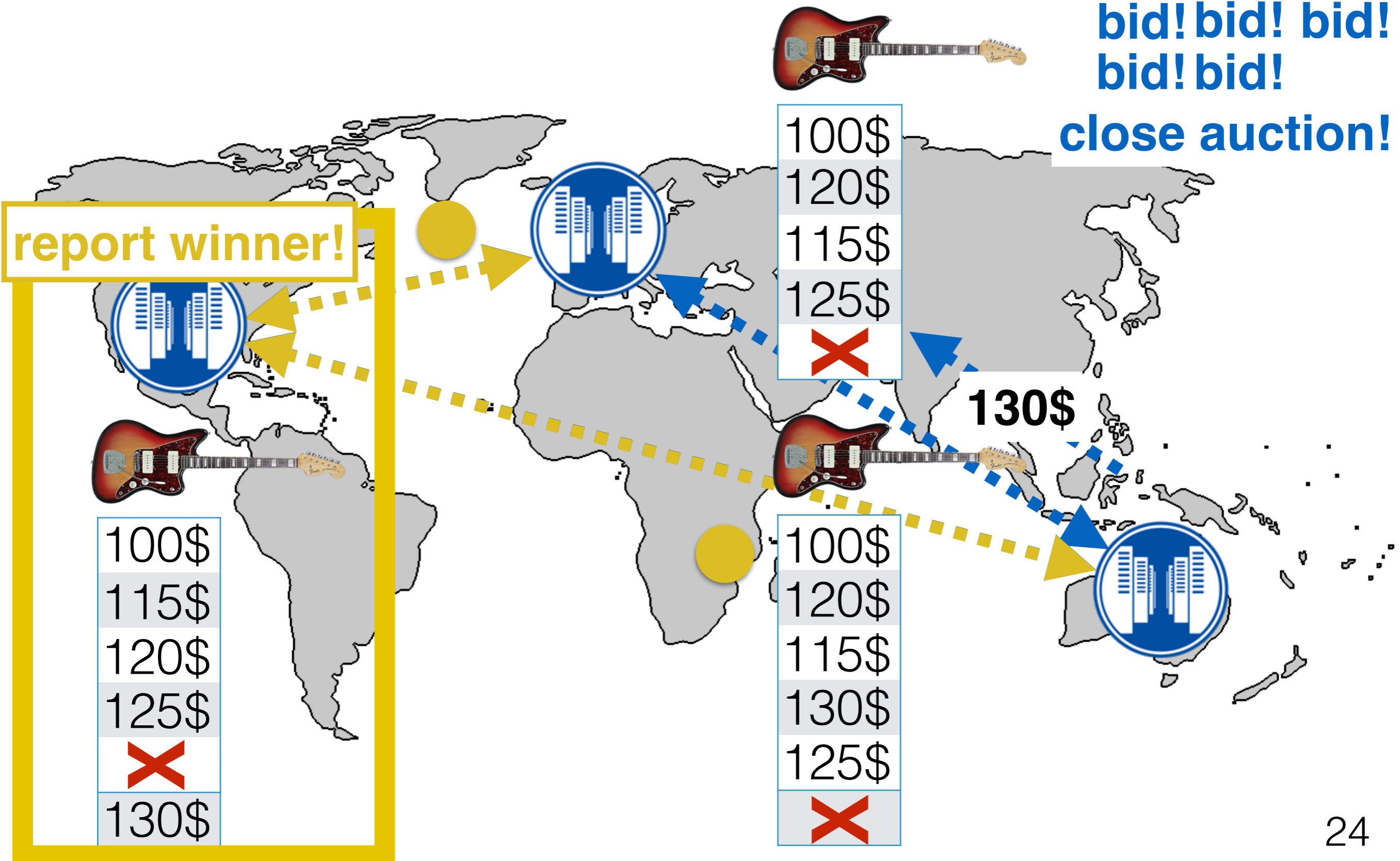


Let's recap: an auction service

start auction!

**bid! bid! bid!
bid! bid!**

close auction!

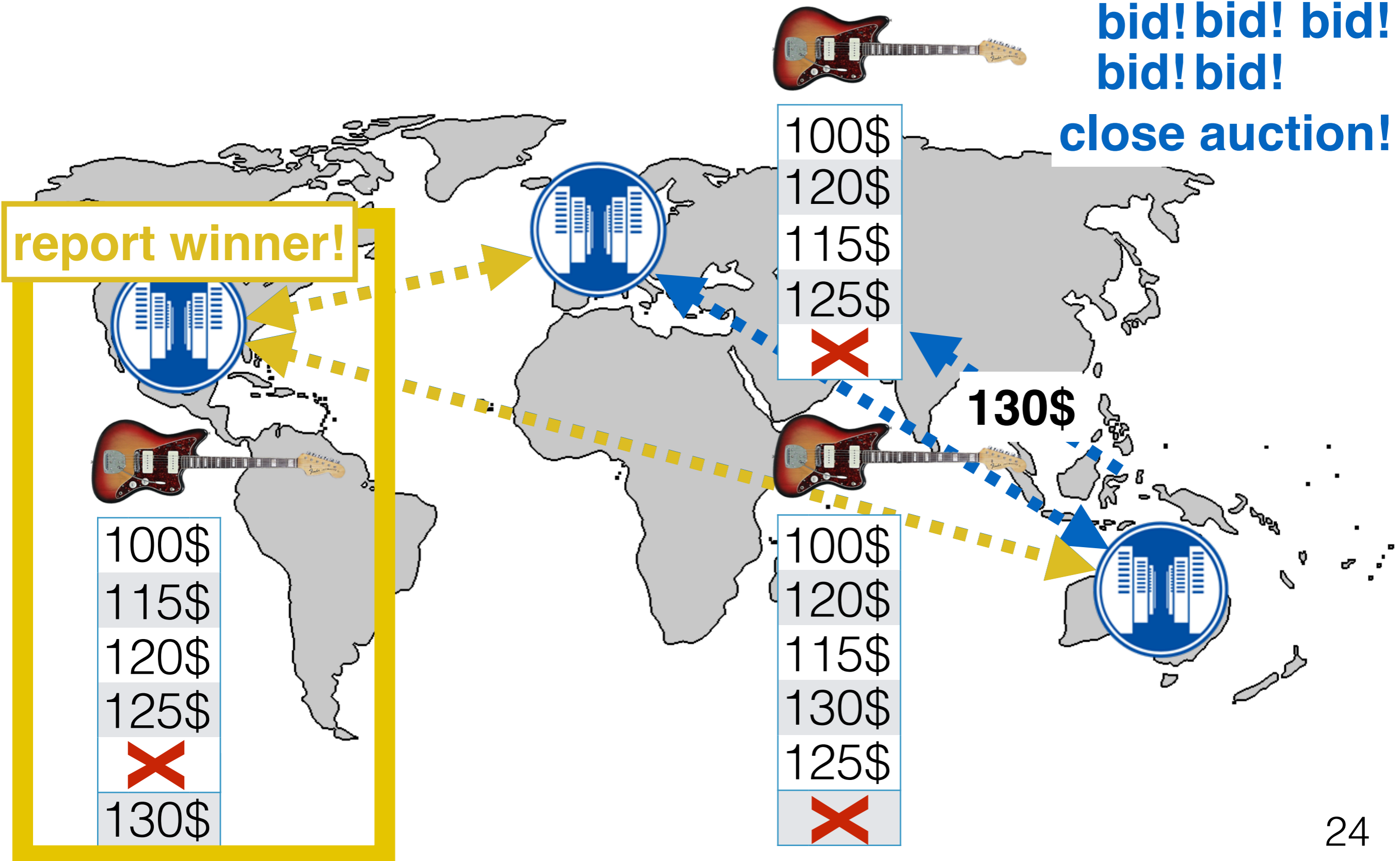


Let's recap: an auction service

start auction!

**bid! bid! bid!
bid! bid!**

close auction!

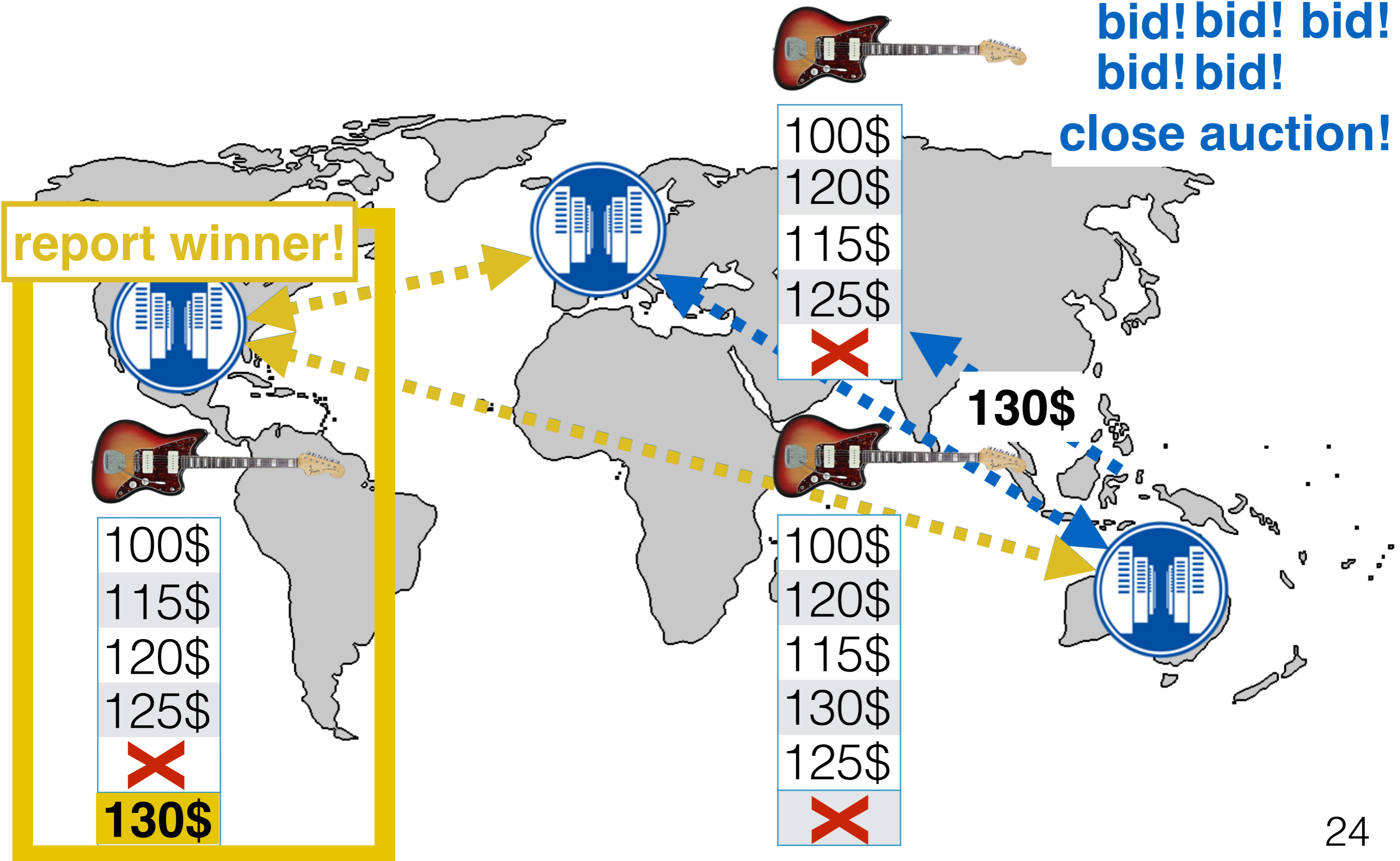


Let's recap: an auction service

start auction!

**bid! bid! bid!
bid! bid!**

close auction!





Other use cases: a cache

SPI 16-8 DEX | Spirou Charleroi - Belfius Mons | Match Summary

Secure | <https://www.flashscore.com/match/YgpQWNdk/#match-summary>

BELGIUM: EuroMillions Basketball League - Round 12

 **Spirou Charleroi** **88 - 88** **Belfius Mons** 

Match Odds Comparison H2H Standings

Match Summary Lineups

Score					
Spirou Charleroi	16	16			
Belfius Mons	8	8			

☒ **LIVE betting** is available for this match. [Bet LIVE with bet365.com](#)

Live odds

bet365 **BET** 1 1.18 2 4.40 **BONUS €50**

Pre-match odds

bet365 **BET** 1 1.50 2 2.50 **BONUS €50**

Other use cases: a cache

When users access the **website**, this is **loaded from** the **cache**

SPI 16-8 DEX | Spirou Charleroi - Belfius Mons | Match Summary

Secure | <https://www.flashscore.com/match/YgpQWNdk/#match-summary>

BELGIUM: EuroMillions Basketball League - Round 12

Spirou Charleroi 88 - 88 Belfius Mons

Match Odds Comparison H2H Standings

Match Summary Lineups

Score					
Spirou Charleroi	16	16			
Belfius Mons	8	8			

LIVE betting is available for this match. [Bet LIVE with bet365.com](#)

Live odds

bet365 1 1.18 2 4.40 **BONUS €50**

Pre-match odds

bet365 1 1.50 2 2.50 **BONUS €50**

Other use cases: a cache

When users access the **website**, this is **loaded from** the **cache**

The screenshot shows a web browser window with the address bar displaying 'https://www.flashscore.com/match/YgpQWNdk/#match-summary'. The page title is 'SPI 16-8 DEX | Spirou Charleroi - Belfius Mons | Match Summary'. The main content area shows the match score 'Spirou Charleroi 88 - 88 Belfius Mons'. Below the score, there are tabs for 'Match', 'Odds Comparison', 'H2H', and 'Standings'. The 'Match Summary' tab is selected, showing a 'Score' table with columns for 'Spirou Charleroi' and 'Belfius Mons'. The table shows a final score of 16-8. Below the score table, there is a green banner with a checkmark icon and the text 'LIVE betting is available for this match.' and a button 'Bet LIVE with bet365.com'. Below the banner, there are sections for 'Live odds' and 'Pre-match odds'. The 'Live odds' section shows a 'bet365' logo, a 'BET' button, and odds for '1' (1.18) and '2' (4.40), along with a 'BONUS €50' button. The 'Pre-match odds' section shows a 'bet365' logo, a 'BET' button, and odds for '1' (1.50) and '2' (2.50), along with a 'BONUS €50' button.

Score	Spirou Charleroi	Belfius Mons
16	16	8

LIVE betting is available for this match. Bet LIVE with bet365.com

Live odds

bet365 BET 1 1.18 2 4.40 BONUS €50

Pre-match odds

bet365 BET 1 1.50 2 2.50 BONUS €50

This **cache** could be **updated** periodically, by means of an **internal operation**

Other use cases: a cache

The screenshot shows a web browser window displaying a basketball match summary for 'Spirou Charleroi - Belfius Mons' in the 'BELGIUM: EuroMillions Basketball League - Round 12'. The score is 88-88. The page includes navigation tabs for 'Match', 'Odds Comparison', 'H2H', and 'Standings'. Below these, there are tabs for 'Match Summary' and 'Lineups'. A 'Score' table shows the current score: Spirou Charleroi 16, Belfius Mons 8. A green banner indicates 'LIVE betting is available for this match.' with a button to 'Bet LIVE with bet365.com'. Below this, there are sections for 'Live odds' and 'Pre-match odds', both featuring the bet365 logo and betting options for '1' and '2' with odds of 1.18 and 4.40 for live, and 1.50 and 2.50 for pre-match. A 'BONUS €50' is also displayed.

Score	
Spirou Charleroi	16
Belfius Mons	8

Live odds	
1	1.18
2	4.40

Pre-match odds	
1	1.50
2	2.50

When users access the **website**, this is **loaded from** the **cache**

This **cache** could be **updated** periodically, by means of an **internal operation**

There could be an alternative option to refresh the cache with the **latest information** by means of an **external operation**

Other use cases: a cache

SPI 16-8 DEX | Spirou Charleroi - Belfius Mons | Match Summary

Secure | <https://www.flashscore.com/match/YgpQWNdk/#match-summary>

BELGIUM: EuroMillions Basketball League - Round 12

Spirou Charleroi **88 - 88** Belfius Mons

Match Odds Comparison H2H Standings

Match Summary Lineups

Score					
Spirou Charleroi	16	16			
Belfius Mons	8	8			

☒ **LIVE betting** is available for this match. [Bet LIVE with bet365.com](#)

Live odds

bet365 **BET** 1 1.18 2 4.40 **BONUS €50**

Pre-match odds

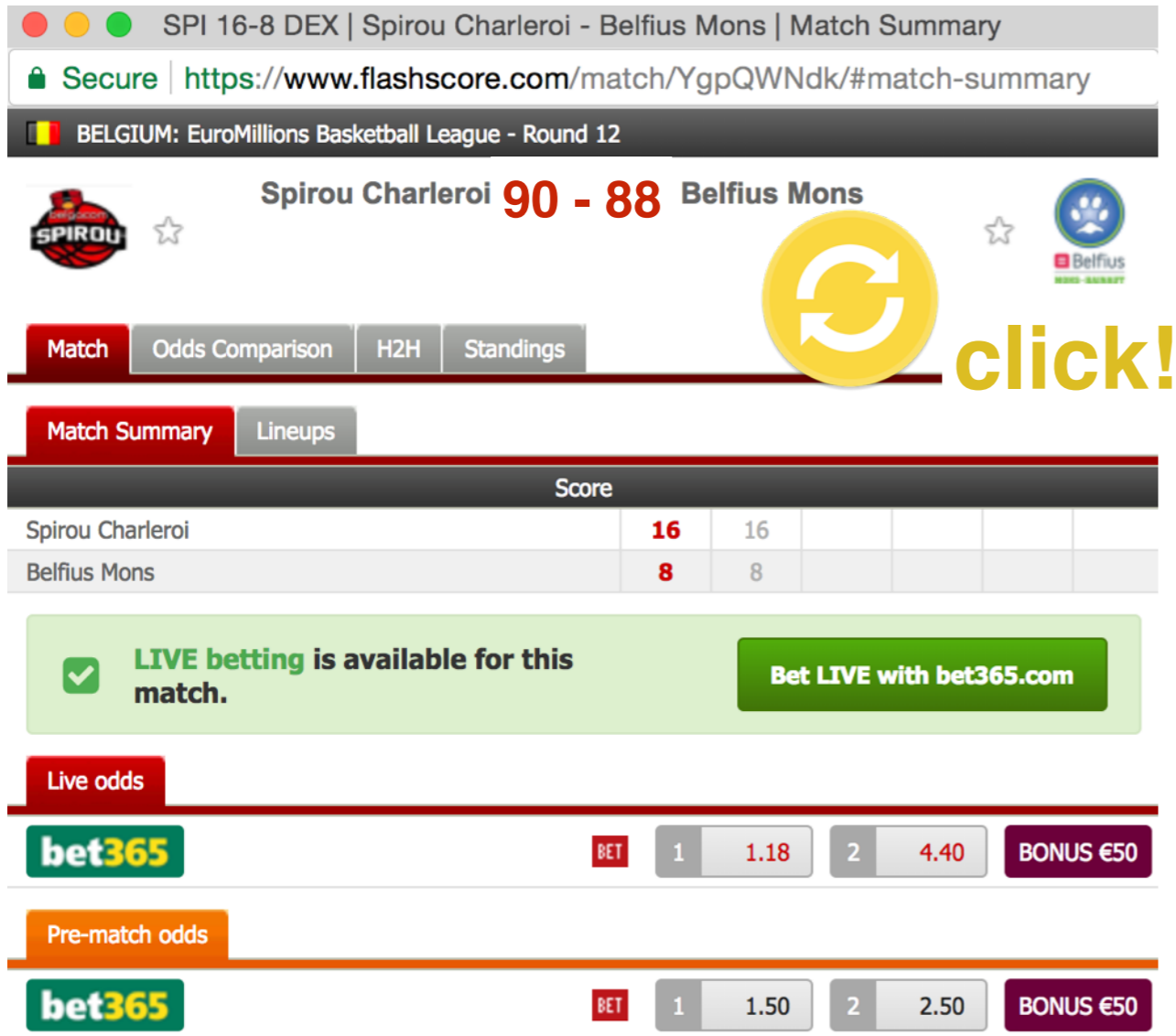
bet365 **BET** 1 1.50 2 2.50 **BONUS €50**

When users access the **website**, this is **loaded from** the **cache**

This **cache** could be **updated** periodically, by means of an **internal operation**

There could be an alternative option to refresh the cache with the **latest information** by means of an **external operation**

Other use cases: a cache



SPI 16-8 DEX | Spirou Charleroi - Belfius Mons | Match Summary

Secure | <https://www.flashscore.com/match/YgpQWNdk/#match-summary>

BELGIUM: EuroMillions Basketball League - Round 12

Spirou Charleroi **90 - 88** Belfius Mons

Match Odds Comparison H2H Standings

Match Summary Lineups

Score					
Spirou Charleroi	16	16			
Belfius Mons	8	8			

☒ **LIVE betting** is available for this match. [Bet LIVE with bet365.com](#)

Live odds

bet365 **BET** 1 1.18 2 4.40 **BONUS €50**

Pre-match odds

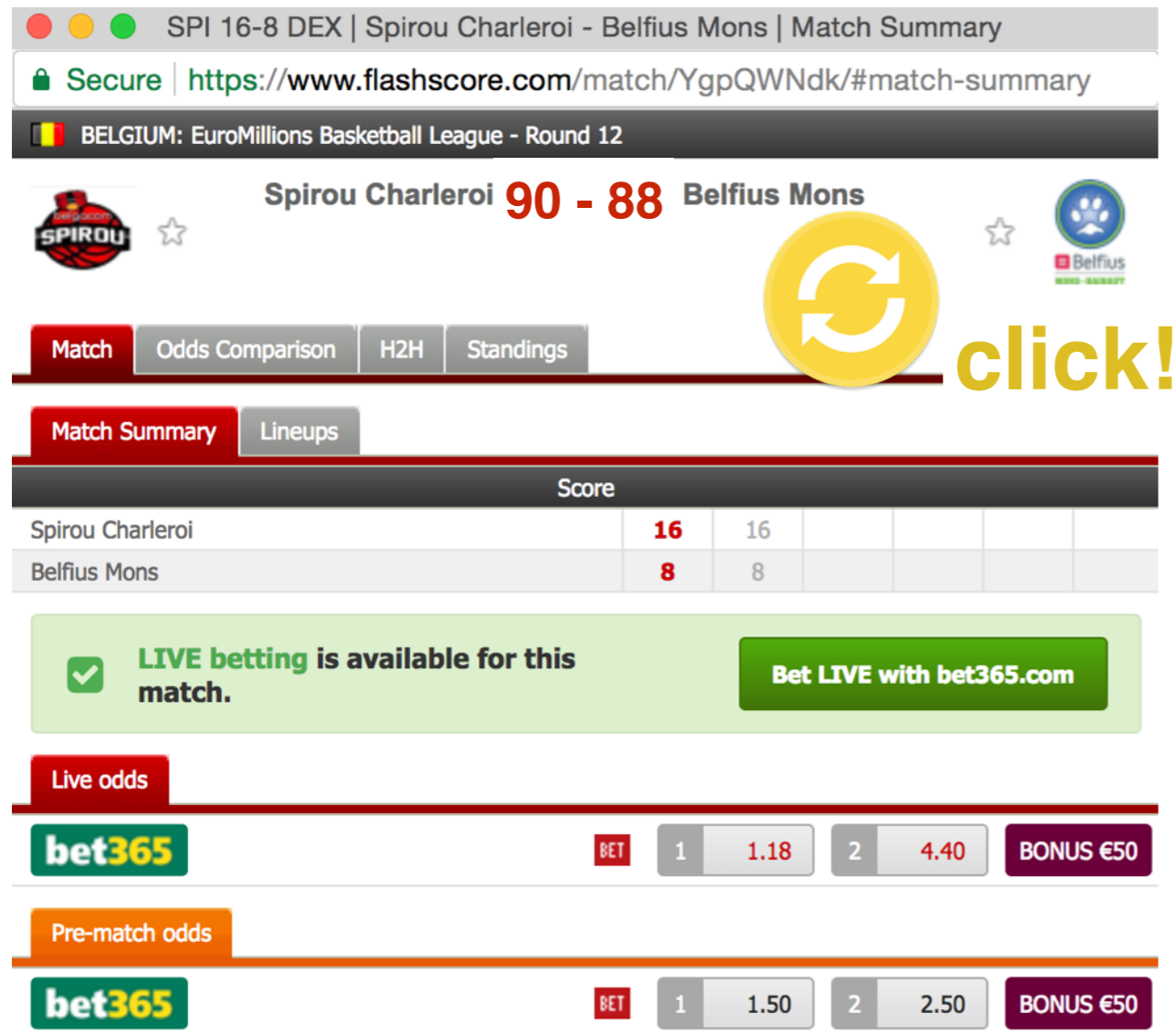
bet365 **BET** 1 1.50 2 2.50 **BONUS €50**

When users access the **website**, this is **loaded from** the **cache**

This **cache** could be **updated** periodically, by means of an **internal operation**

There could be an alternative option to refresh the cache with the **latest information** by means of an **external operation**

Other use cases: a cache



The screenshot shows a web browser window displaying a basketball match summary for "Spirou Charleroi - Belfius Mons" in the "BELGIUM: EuroMillions Basketball League - Round 12". The score is 90 - 88. A yellow circular refresh button with a circular arrow icon and the text "click!" is overlaid on the page. The page includes tabs for "Match", "Odds Comparison", "H2H", and "Standings", with "Match Summary" and "Lineups" sub-tabs. A "Score" table shows Spirou Charleroi with 16 points and Belfius Mons with 8 points. Below the score, there is a "LIVE betting is available for this match." banner with a "Bet LIVE with bet365.com" button. The "Live odds" section shows betting options for "bet365" with odds of 1.18 for option 1 and 4.40 for option 2, plus a "BONUS €50". The "Pre-match odds" section shows similar betting options with odds of 1.50 for option 1 and 2.50 for option 2, also with a "BONUS €50".

Team	Score
Spirou Charleroi	16
Belfius Mons	8

Section	Option	Odds	Bonus
Live odds	1	1.18	BONUS €50
	2	4.40	
Pre-match odds	1	1.50	BONUS €50
	2	2.50	

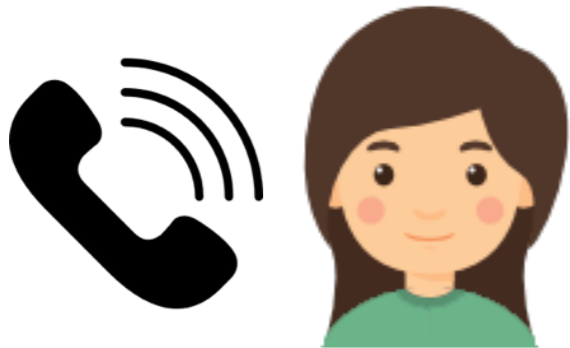
When users access the **website**, this is **loaded from** the **cache**

This **cache** could be **updated** periodically, by means of an **internal operation**

There could be an alternative option to refresh the cache with the **latest information** by means of an **external operation**

Useful to **read time-sensitive** information

Other use cases: a cache



Other use cases: a cache



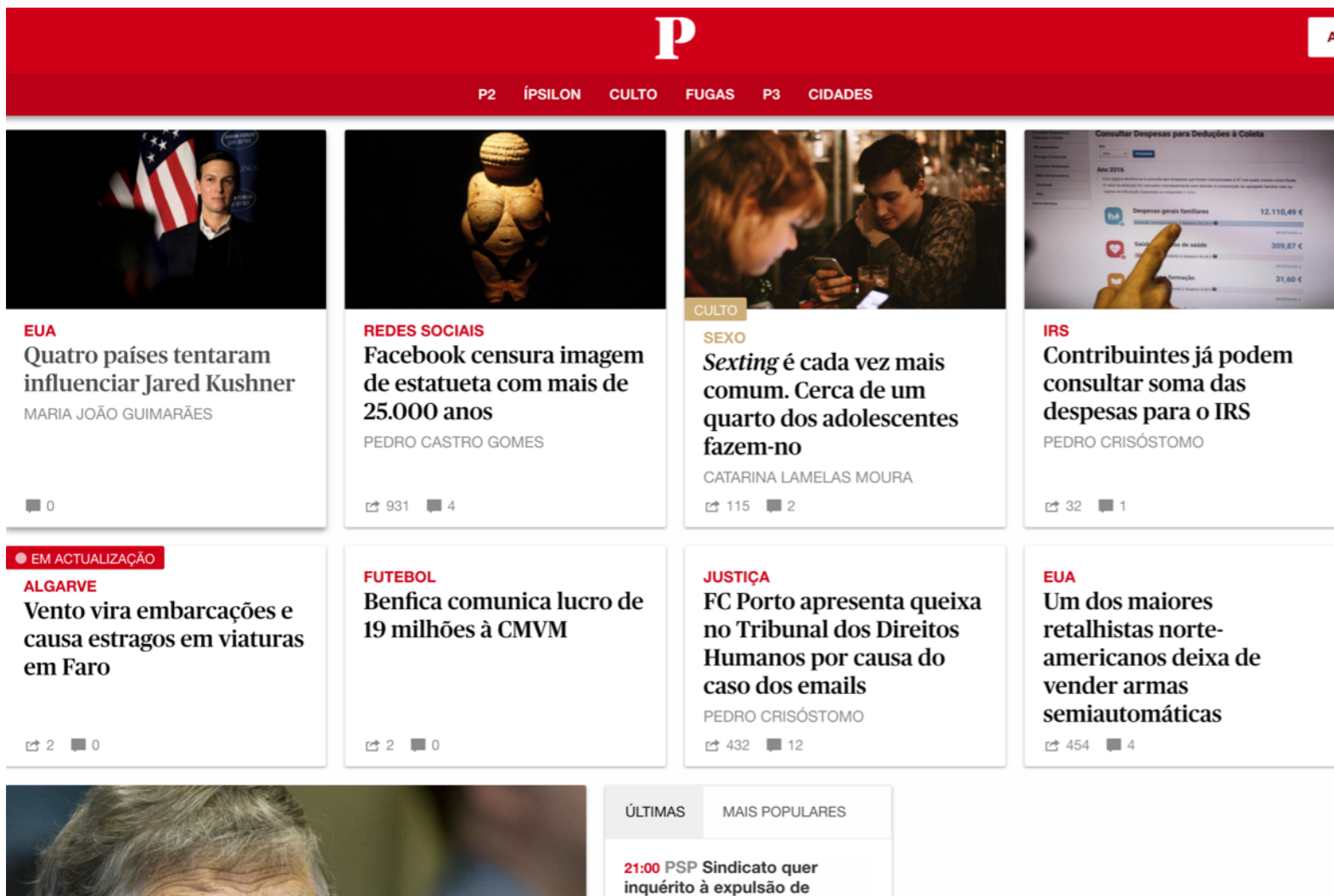
Other use cases: a cache



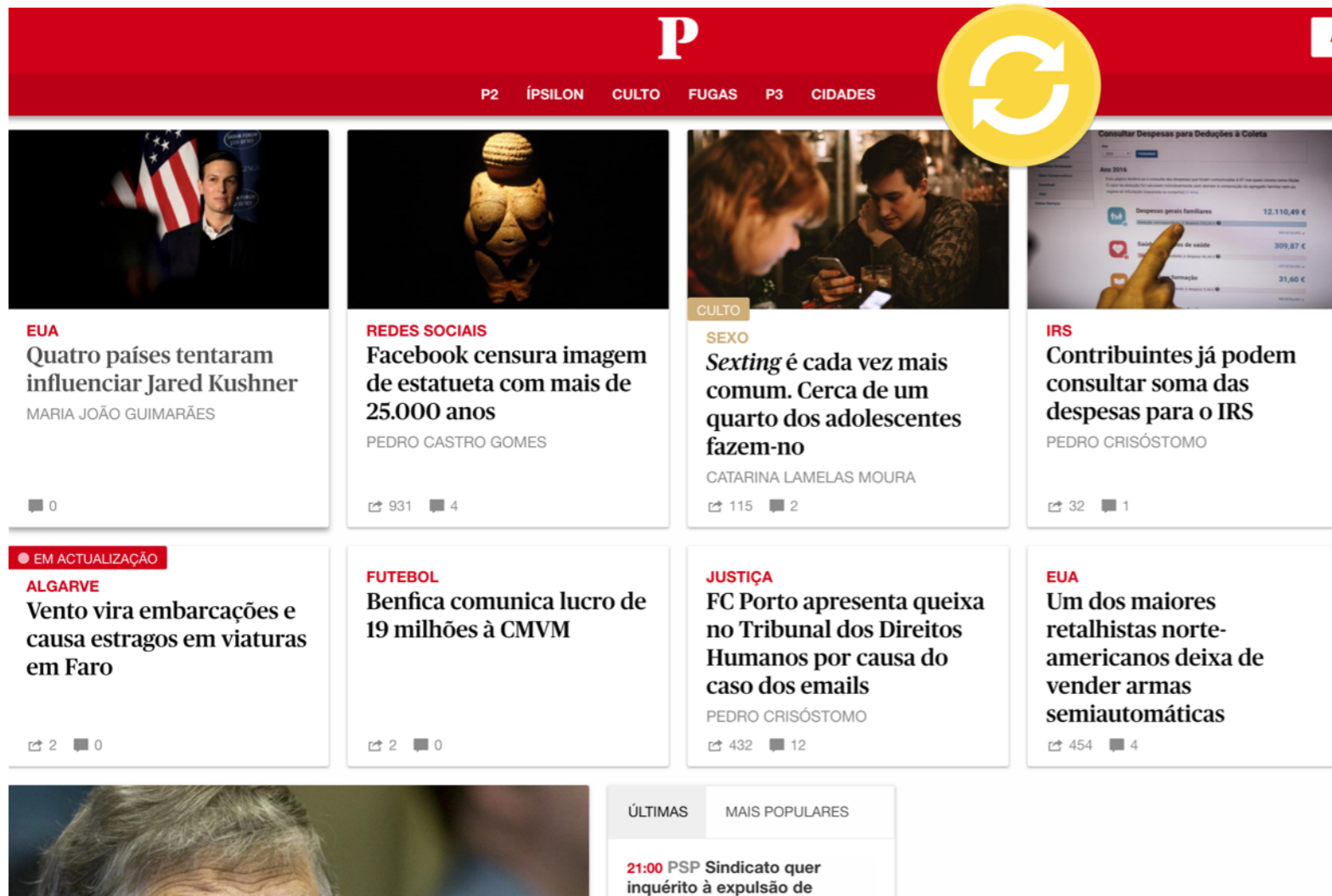
Other use cases: a cache



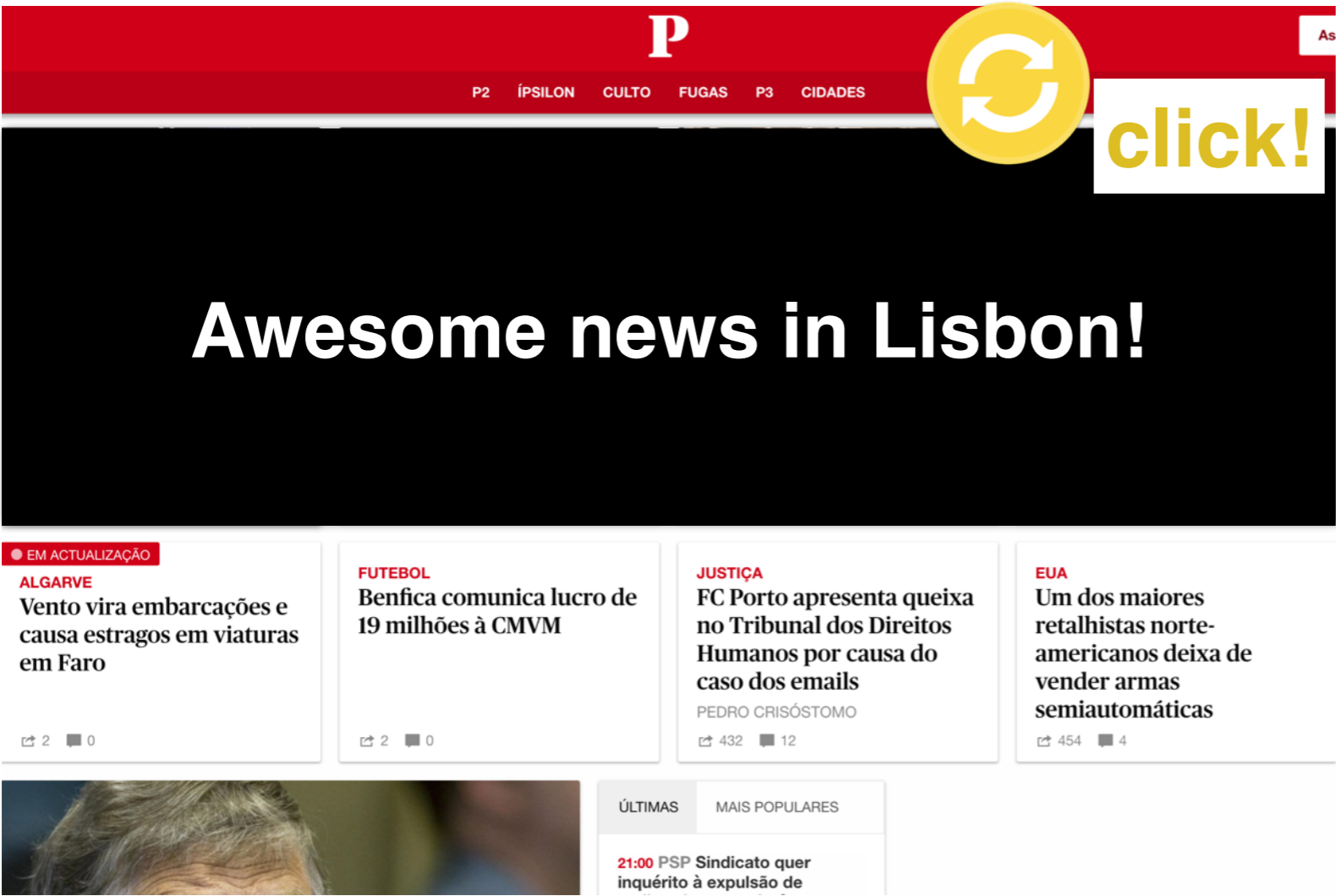
Other use cases: a cache



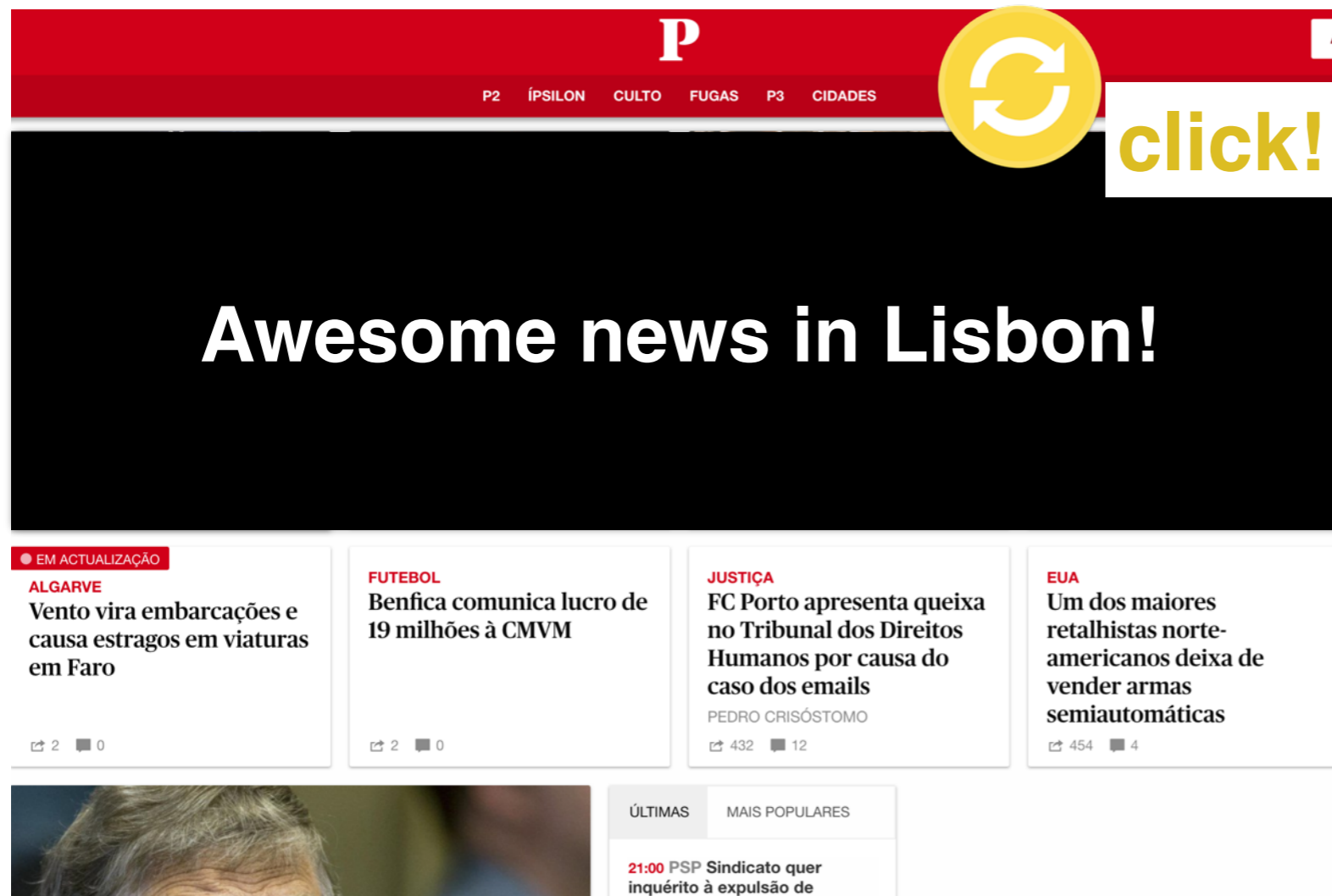
Other use cases: a cache



Other use cases: a cache



Other use cases: a cache



Useful to **cope**, at least partially, with **back-channelling**

A word on fault tolerance

to ensure external guarantees, the write quorum of both internal and external operations **must overlap** with the quorum use by external operations in its prepare phase

A word on fault tolerance

to ensure external guarantees, the write quorum of both internal and external operations **must overlap** with the quorum use by external operations in its prepare phase

we plan to only **ack** write operations when these have been installed in **two sites**, in order to tolerate the failure of one site

A word on fault tolerance

to ensure external guarantees, the write quorum of both internal and external operations **must overlap** with the quorum use by external operations in its prepare phase

we plan to only **ack** write operations when these have been installed in **two sites**, in order to tolerate the failure of one site

thus, the **prepare phase** of external operations does not require contacting all sites

A word on fault tolerance

bad things: we are making writes slightly more expensive

A word on fault tolerance

bad things: we are making writes slightly more expensive

good things: reads of internal operations still require a single site

A word on fault tolerance

bad things: we are making writes slightly more expensive

good things: reads of internal operations still require a single site

more good things: we enhance availability when compared to causal consistency, which otherwise is **sticky available**

Related work

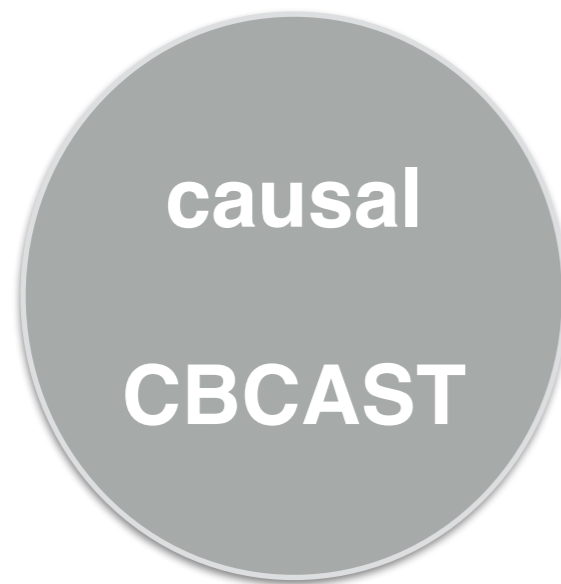
closest to our work:

lazy replication [TOCS' 92] and the ISIS work [TOCS' 87]

Related work

closest to our work:

lazy replication [TOCS' 92] and the ISIS work [TOCS' 87]



Related work

closest to our work:

lazy replication [TOCS' 92] and the ISIS work [TOCS' 87]



Related work

closest to our work:

lazy replication [TOCS' 92] and the ISIS work [TOCS' 87]



Related work

closest to our work:

lazy replication [TOCS' 92] and the ISIS work [TOCS' 87]

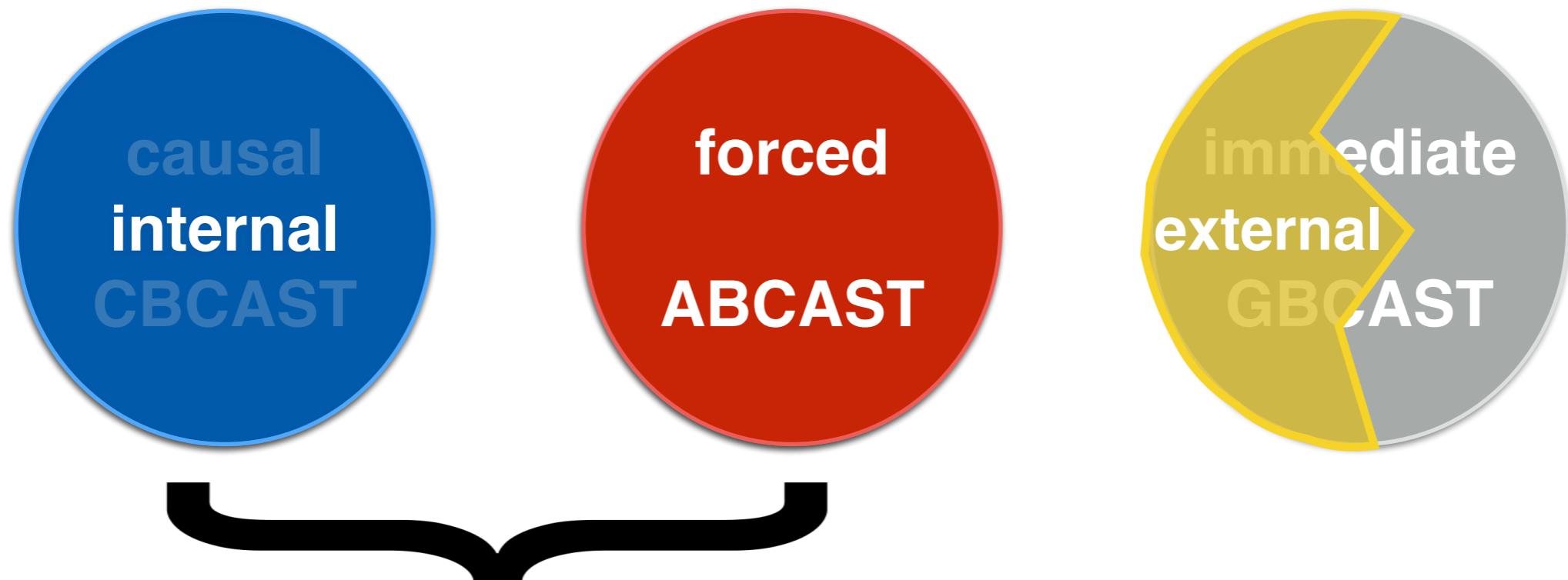


RedBlue consistency [OSDI' 12]

Related work

closest to our work:

lazy replication [TOCS' 92] and the ISIS work [TOCS' 87]



RedBlue consistency [OSDI' 12]

Related work

immediate

GBCAST

Related work



Performed at all replicas in the same order relative to all operations

Related work



Performed at all replicas in the same order relative to all operations

Ordered consistently with external events

Related work



Performed at all replicas in the same order relative to all operations

Ordered consistently with external events

The operation becomes all-present after completion

Related work



Performed at all replicas in the same order relative to all operations

Ordered consistently with external events

The operation becomes all-present after completion

Related work



Performed at all replicas in the same order relative to all operations

Ordered consistently with external events

The operation becomes all-present after completion

Related work



Performed at all replicas in the same order relative to all operations

Ordered consistently with external events

~~The operation becomes all-present after completion~~

Related work



~~Performed at all replicas in the same
order relative to all operations~~

Ordered consistently with external
events

~~The operation becomes all-present
after completion~~

Related work



Performed at all replicas in causal order

Ordered consistently with external events

~~The operation becomes all-present
after completion~~

Related work



Performed at all replicas in causal order

Ordered consistently with external events

~~The operation becomes all-present
after completion~~

Open questions

shall we add a new operation that becomes
all-present after completion but it is not
order consistently with external events?

Open questions

shall we add a new operation that becomes
all-present after completion but it is not
order consistently with external events?



Open questions

shall we add a new operation that becomes
all-present after completion but it is not
order consistently with external events?



Open questions

shall we add a new operation that becomes
all-present after completion but it is not
order consistently with external events?

Performed at all replicas in
causal order

Ordered consistently with
external events



Open questions

shall we add a new operation that becomes
all-present after completion but it is not
order consistently with external events?

Performed at all replicas in
causal order

Ordered consistently with
external events



Open questions

shall we add a new operation that becomes
all-present after completion but it is not
order consistently with external events?

Performed at all replicas in
causal order

Ordered consistently with
external events



The operation becomes all-
present after completion

Open questions

shall we add a new operation that becomes **all-present after completion** but it is not order consistently with external events?

Performed at all replicas in causal order

Ordered consistently with external events



The operation becomes all-present after completion

this could further help to **beat back-channeling**

Open questions

shall we **add** mechanisms to ensure **convergence** (CRDTs) and to preserve invariants (forced/**red operations**) in order to build a complete system?

Open questions

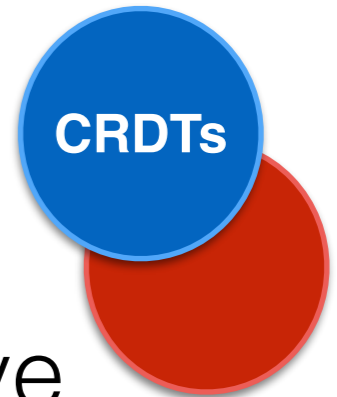


CRDTs

shall we **add** mechanisms to ensure **convergence** (CRDTs) and to preserve invariants (forced/**red operations**) in order to build a complete system?

Open questions

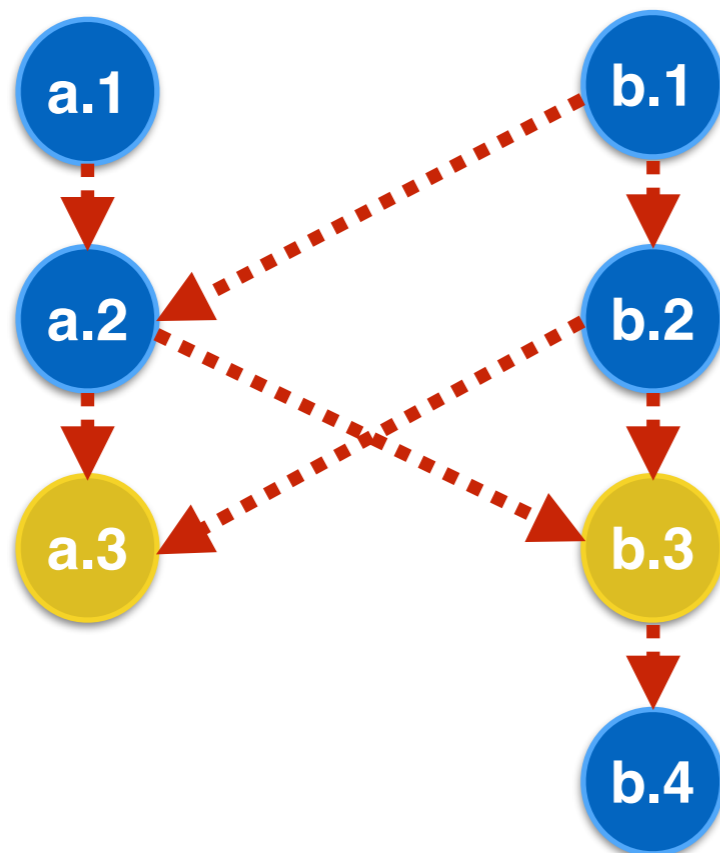
shall we **add** mechanisms to ensure **convergence** (CRDTs) and to preserve invariants (forced/**red operations**) in order to build a complete system?



Open questions

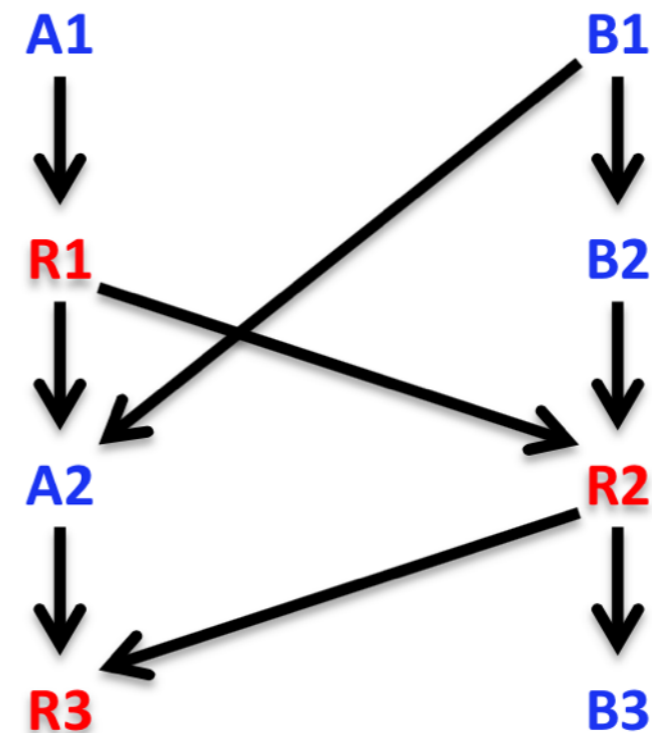
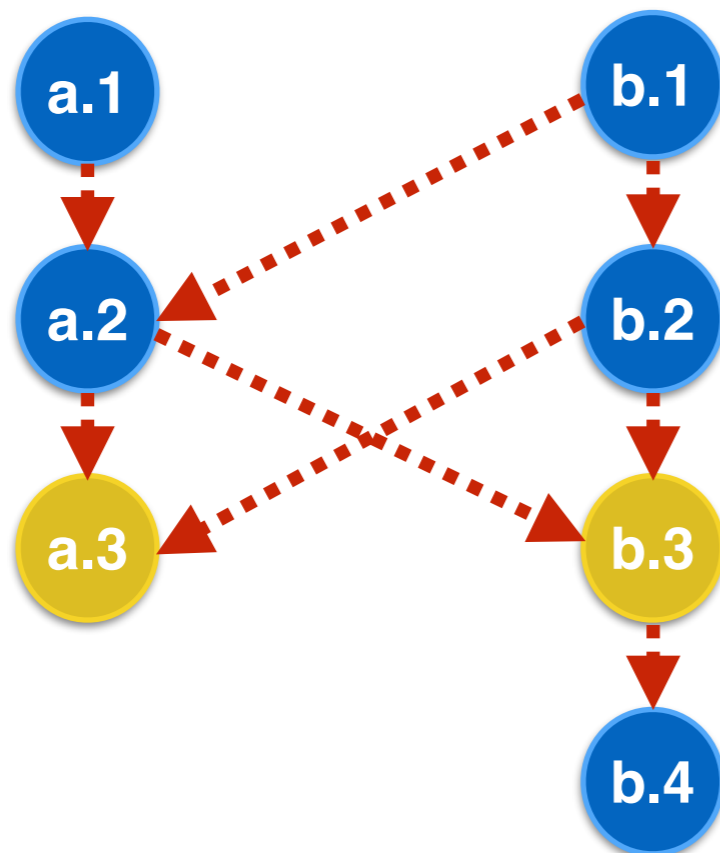
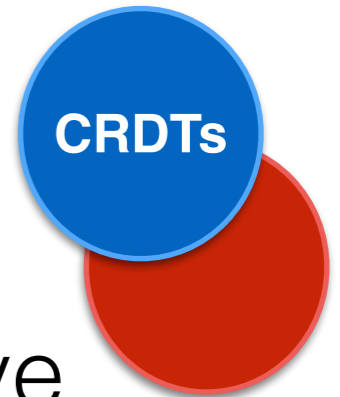
CRDTs

shall we **add** mechanisms to ensure **convergence** (CRDTs) and to preserve invariants (forced/**red operations**) in order to build a complete system?



Open questions

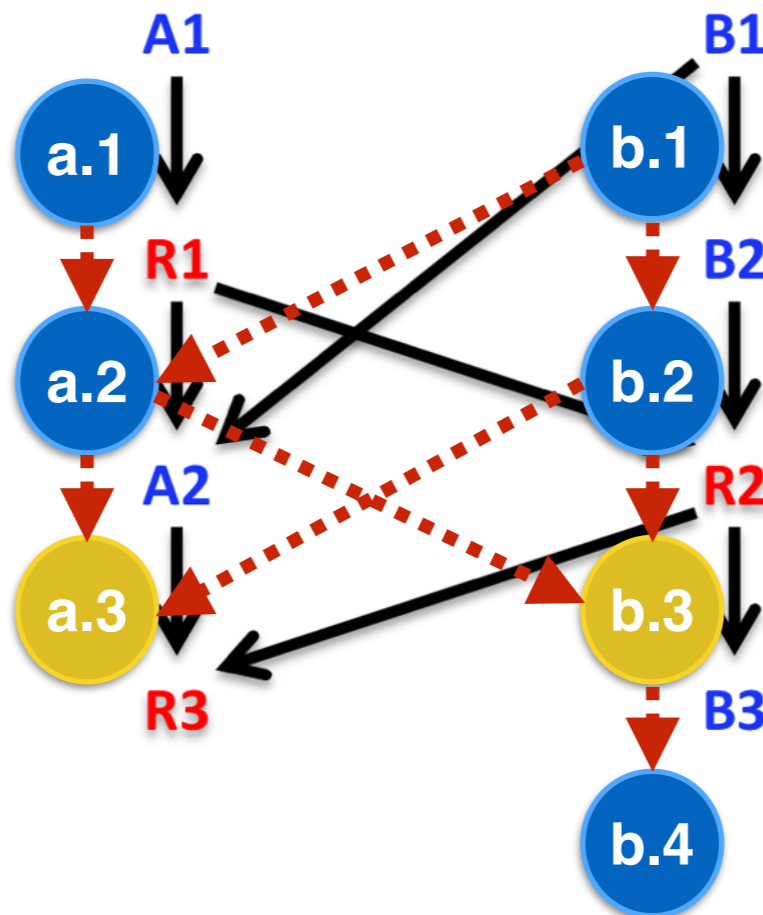
shall we **add** mechanisms to ensure **convergence** (CRDTs) and to preserve invariants (forced/**red operations**) in order to build a complete system?



Open questions

CRDTs

shall we **add** mechanisms to ensure **convergence** (CRDTs) and to preserve invariants (forced/**red operations**) in order to build a complete system?



Open questions

how can we **efficiently implement** external causality in a geo-replicated, distributed system?

Thanks! Questions?

we present a new consistency model:
external causality

addresses the inherent tension between
performance and meaningful semantics

combines causal consistency (strongest
available) and external consistency
(strongest semantically)