Partitionable groups

- New assumption: **network partitions**
- Some systems (VS) block processes and never deliver minority views
- Define majority from:
  - Configured vs operational servers
  - Alternative model: majority over previous view
- Other systems (EVS), such as Spread, deliver all views and it is up to the application to block if/when needed

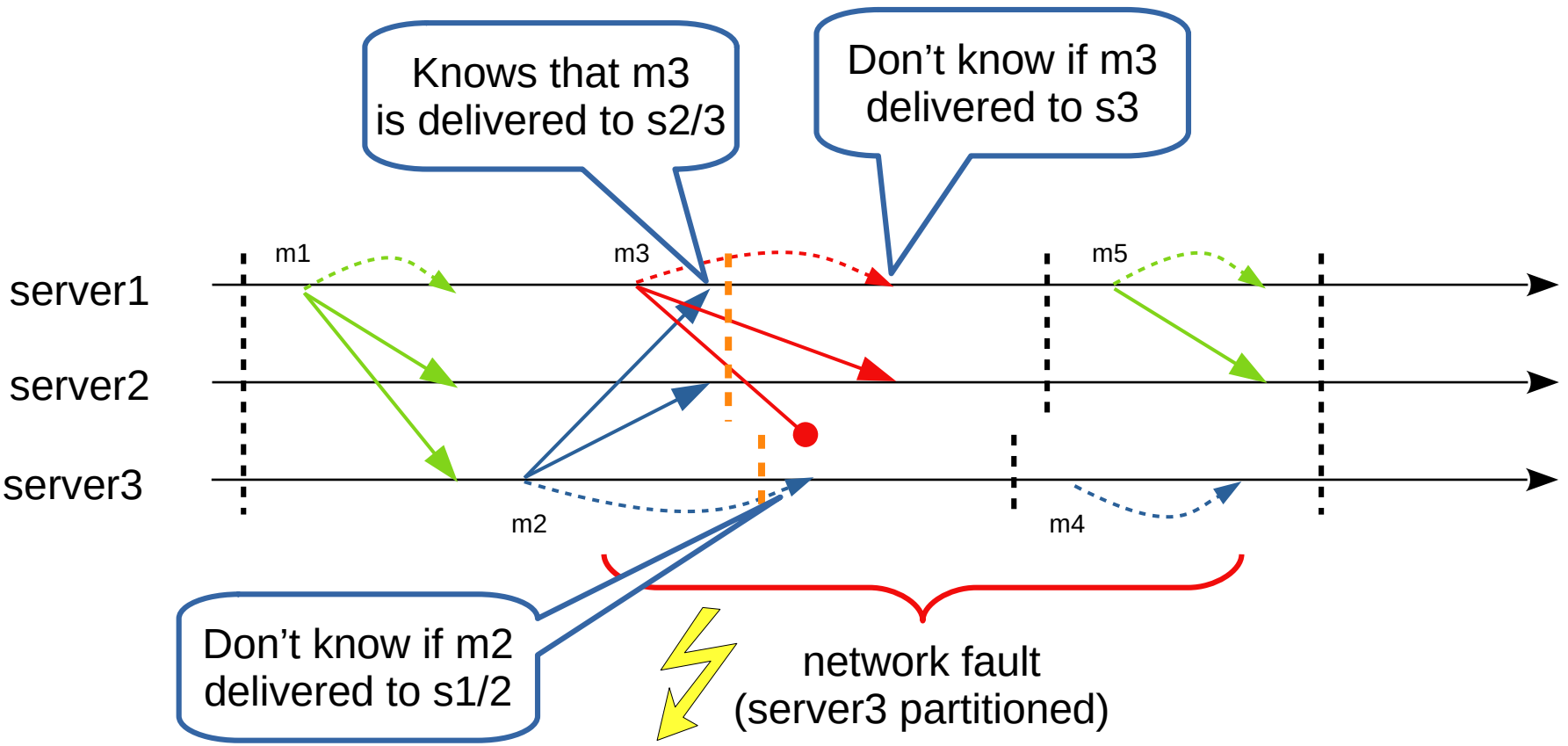
Partitionable groups

- A regular view describes the new group:
  - Informs why the view is changing all views being merged
  - The next safe messages will be delivered by all processes (in a regular or transitional view)
- A transitional view means:
  - A new regular view excluding some processes is about to be installed
  - The next safe messages might not be delivered by excluded processes

See: http://www.cnds.jhu.edu/pub/papers/evs2vs.pdf (Chap. 3)

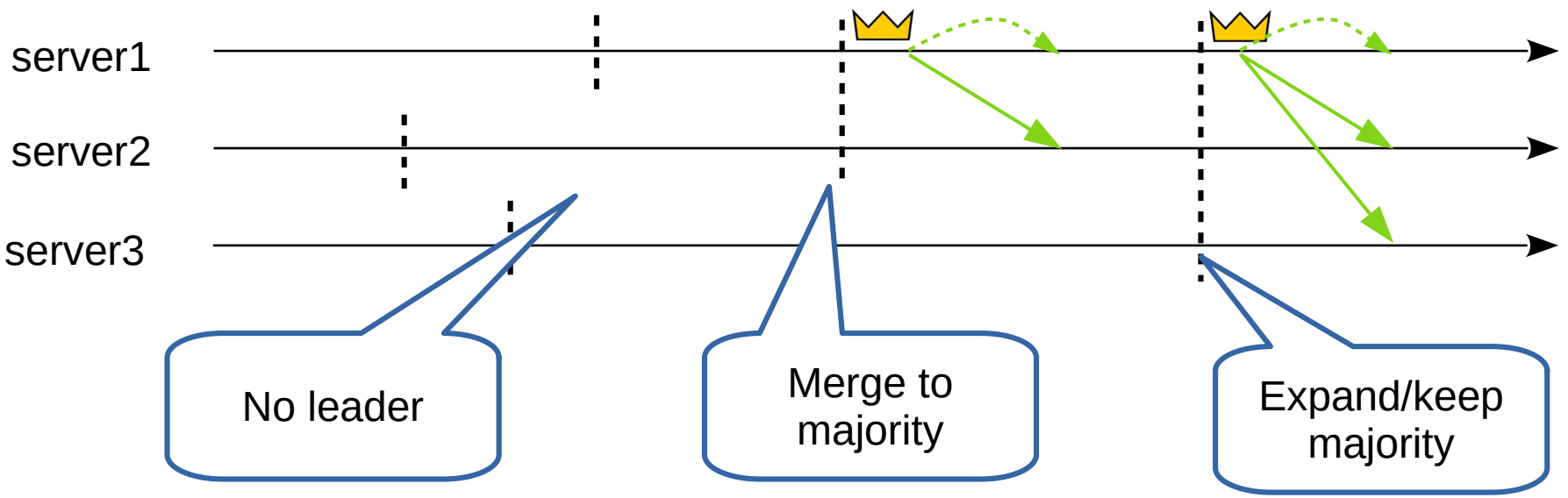
Partitionable groups

- Transitional views provide information about uncertainty:



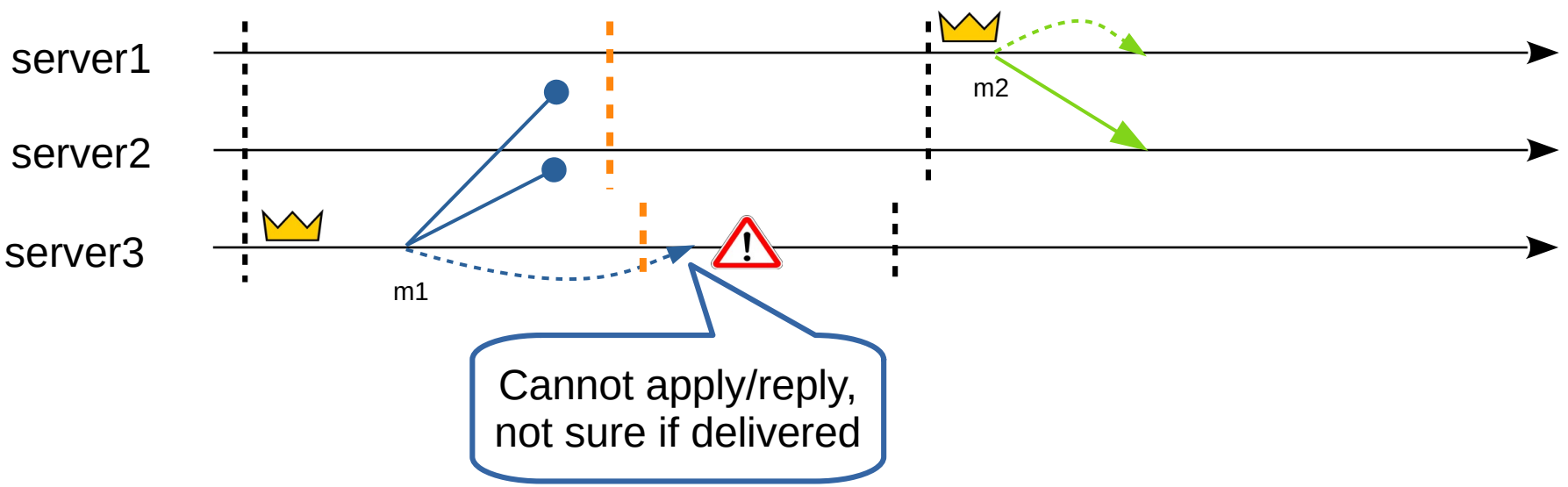
Primary-backup with partitions

- Assume: majority of configured servers and a single outstanding update
- Must select a leader only in majority views
- Should keep leader, if majority holds after view change



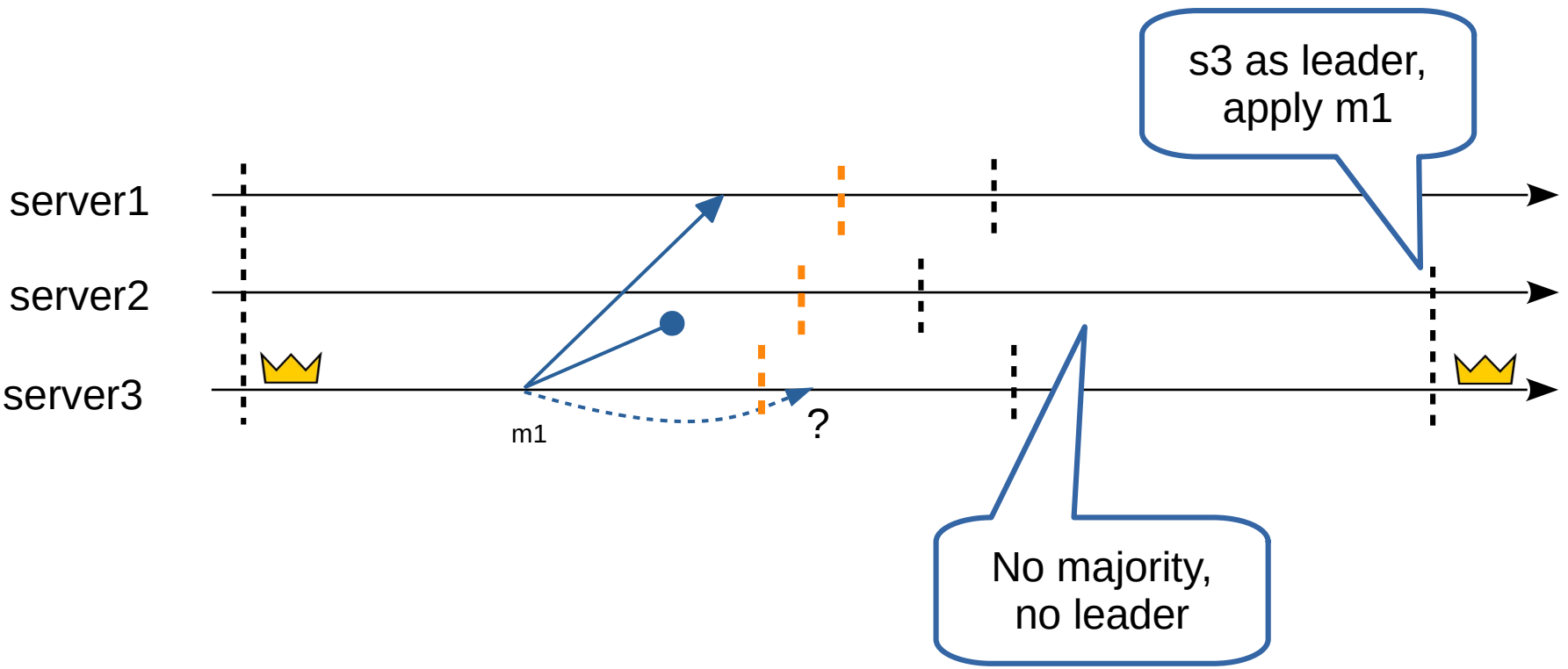
Partitioned replica (scenario 1)

- It is possible that the leader is partitioned but majority exists
- Thus it must stop after a minority transitional view



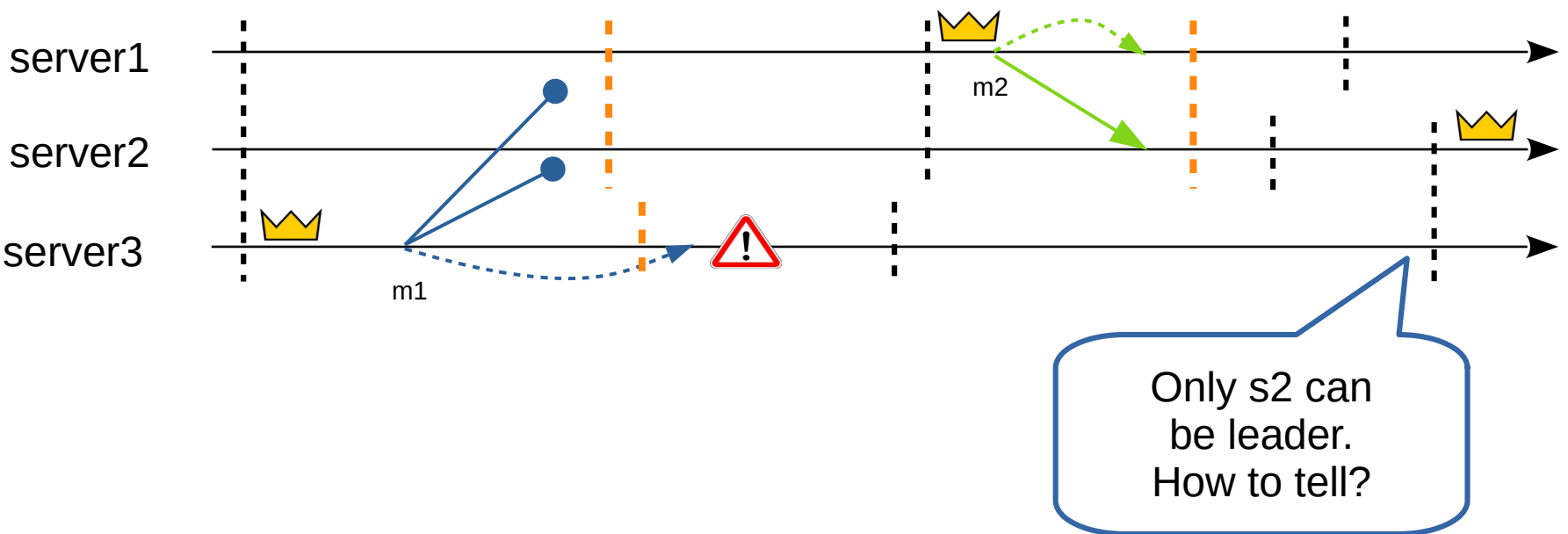
Partitioned replica (scenario 2)

- When the leader is partitioned but no majority exists
- The leader should stop after a minority transitional view



Partitioned replica (scenario 3)

- When excluding, any other in the majority can be leader
- When merging, how to tell where is current state?



Merge protocol

- When a new view is installed:
  - send latest update sequence numbers (applied and known) to group
  - collect sequence numbers from all (*)
  - order processes by (applied, known, name)
  - pick first as new leader

(*) a new view might be installed before complete…