

# Process Mining

COMP9313: Big Data Management

# Tutorial / Demo: ProM Framework

The screenshot displays the ProM [5.2] software interface. The main window is titled "outpatientClinicExample.mxml" and features a dashboard with the following components:

- Dashboard Sidebar:** Contains icons for Dashboard, Funnel, Search, and a document icon.
- Key data:** A table showing simulation statistics.

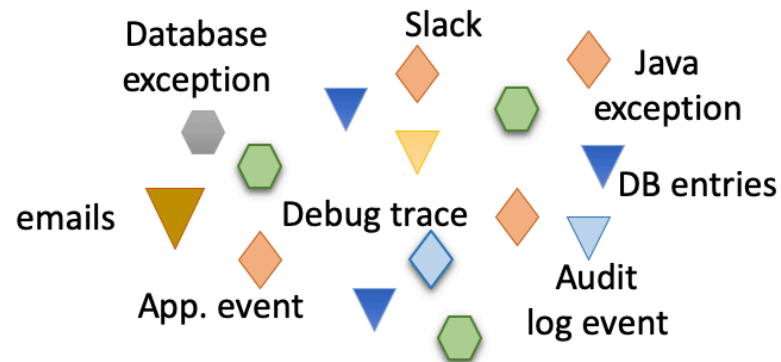
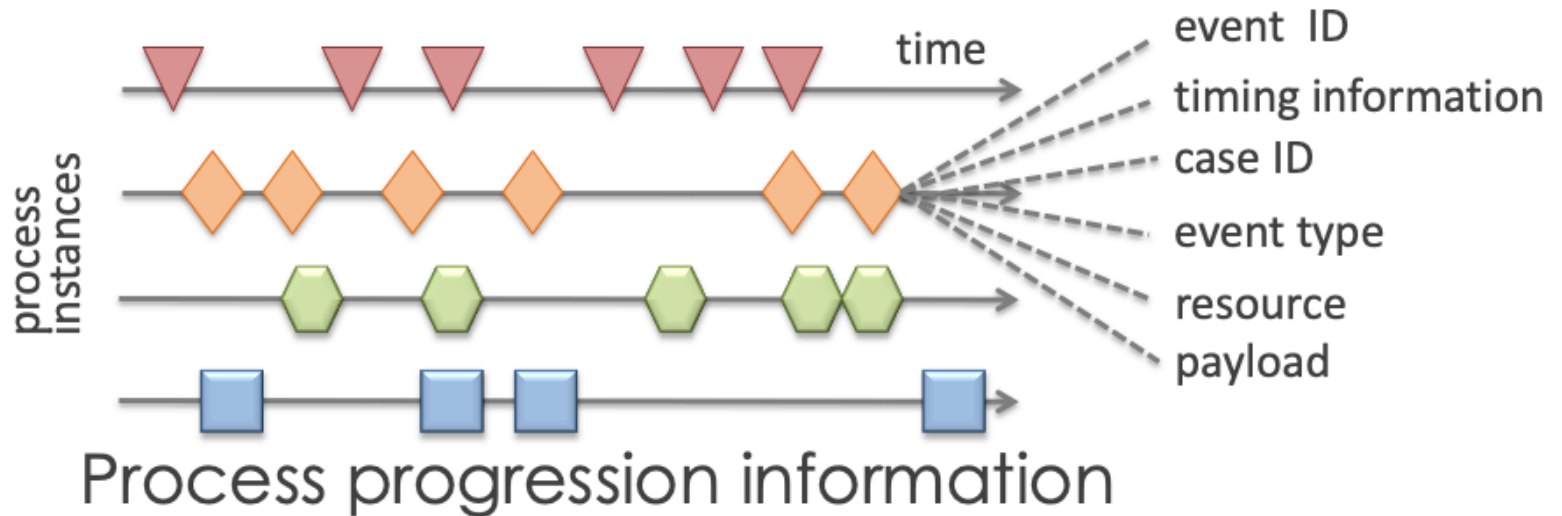
| Metric        | Value |
|---------------|-------|
| Processes     | 1     |
| Cases         | 1000  |
| Events        | 14000 |
| Event classes | 18    |
| Event types   | 2     |
| Originators   | 17    |
- Events per case:** A green rectangular area with summary statistics: Min 14, Mean 14, Max 14.
- Event classes per case:** A green rectangular area with summary statistics: Min 14, Mean 14, Max 14.
- Log info:** A panel providing details about the simulation log.
  - Source: CPN Tools simulation
  - Source program: CPN Tools simulation
  - Start date: 2007-01-01 11:00:00
  - End date: 2007-02-21 00:38:00
  - Description: CPN Tools simulation log
- Buttons:** A "start analyzing this log" button is located at the bottom right of the log info panel.

The status bar at the bottom of the window displays the message: "11:17:19 [D] Processing data for buffered log reader completed."

# Challenges

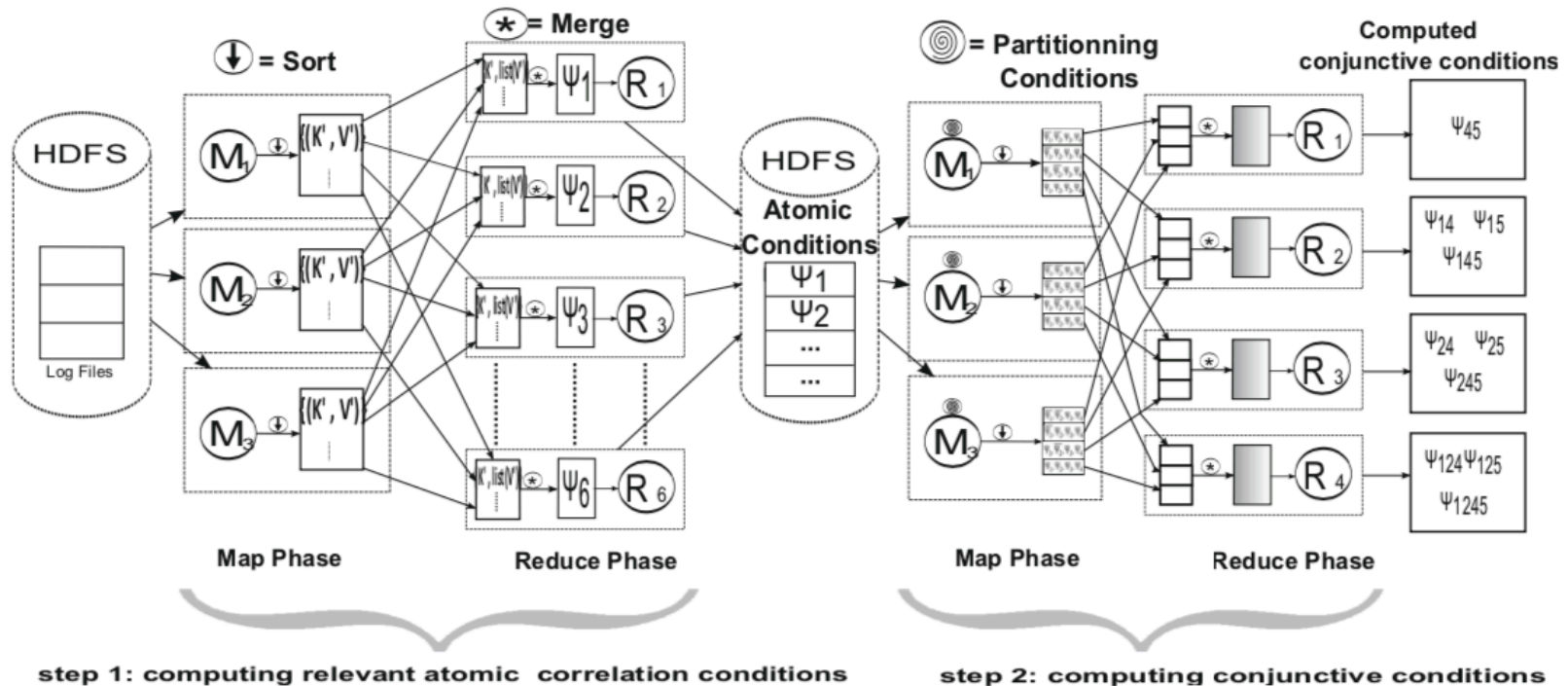
- Huge event logs generated by process-aware information systems (**Volume**)
- Events generated at very high speed (**Velocity**)
- Heterogenous events (**Variety**)
- Are the events in the log legit / precise? (**Veracity**)
- Do we get access to all events generated by process-aware information systems? (**Visibility**)
- Are all events really relevant in the context of process mining? (**Value**)

# Events and event correlation



What we actually get

# Event logs and event correlations



(\*) Reguieg et al., 2012, September. Using mapreduce to scale events correlation discovery for business processes mining. In *International Conference on Business Process Management*(pp. 279-284). Springer, Berlin, Heidelberg

Thanks