

INFO-H420
Management of Data Science and
Business Workflows

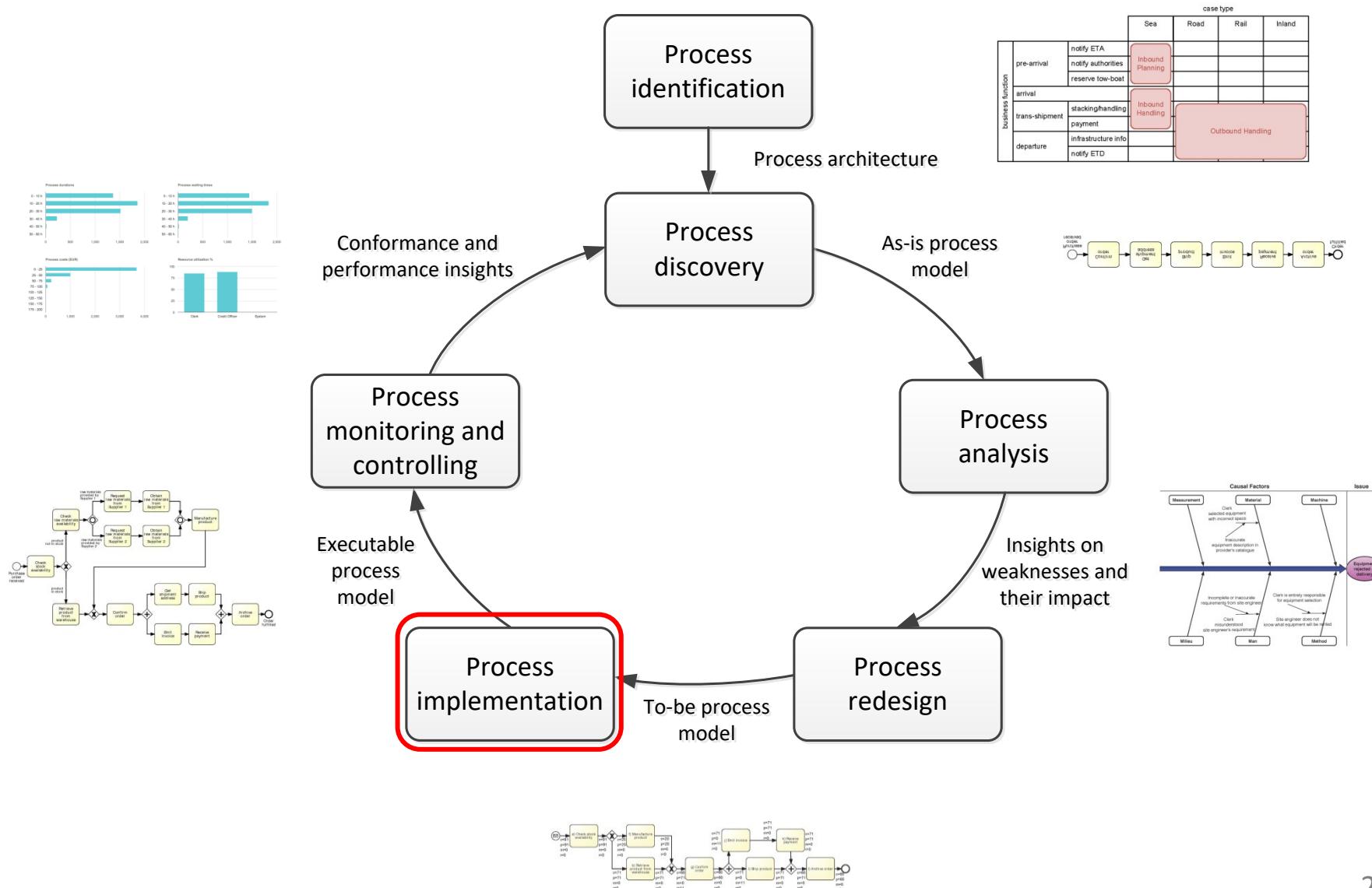
Part I

6. Process Automation and Mining

Dimitris SACHARIDIS

2023-2024

BPM Lifecycle



From “conceptual” to executable models

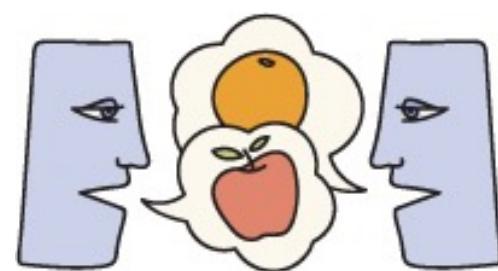
Conceptual “to-be” process models

- are made by domain experts
- provide a basis for communication amongst relevant stakeholders
- must be understandable
- must be intuitive and may leave room for interpretation
- contain purely a relevant set of process information



Executable process models

- are made by IT experts
- provide input to a process enactment system - BPMS
- must be machine readable
- must be unambiguous and should not contain any uncertainties
- contain further details that are only relevant to implementation



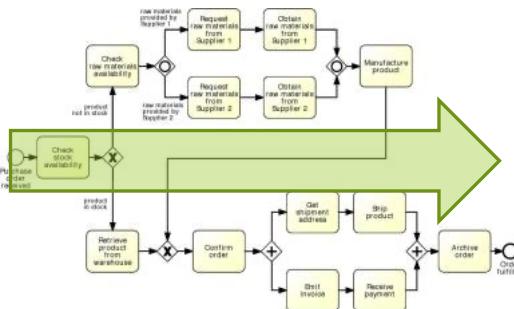
From “conceptual” to executable models

Conceptual “to-be” process models

- are made by domain experts
- provide a basis for communication amongst relevant stakeholders
- must be understandable
- must be intuitive and may leave room for interpretation
- contain purely a relevant set of process information

Executable process models

- are made by IT experts
- provide input to a process enactment system - BPMS
- must be machine readable
- must be unambiguous and should not contain any uncertainties
- contain further details that are only relevant to implementation



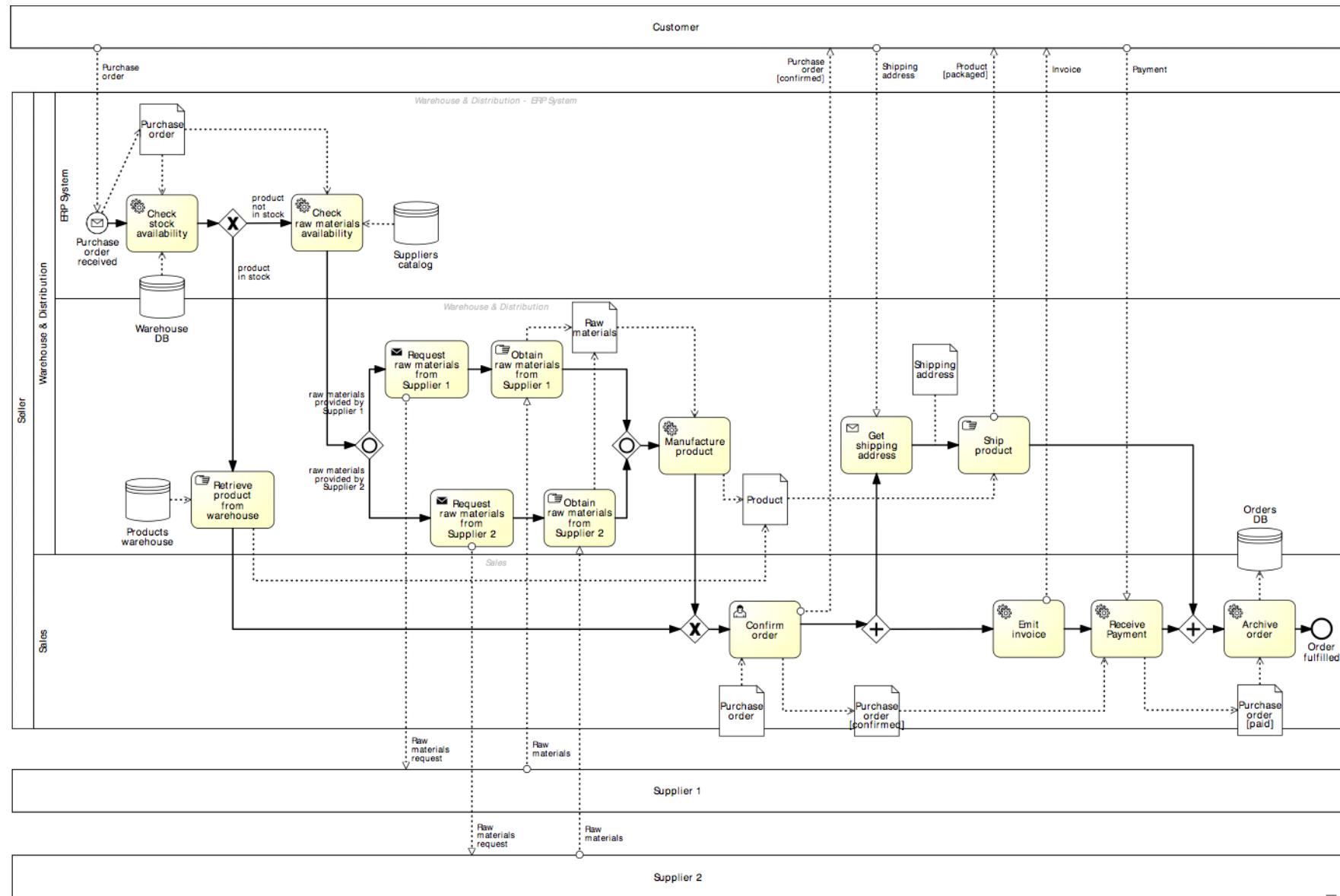
“to-be executed”
process model

Bridging the gap: A five-step method

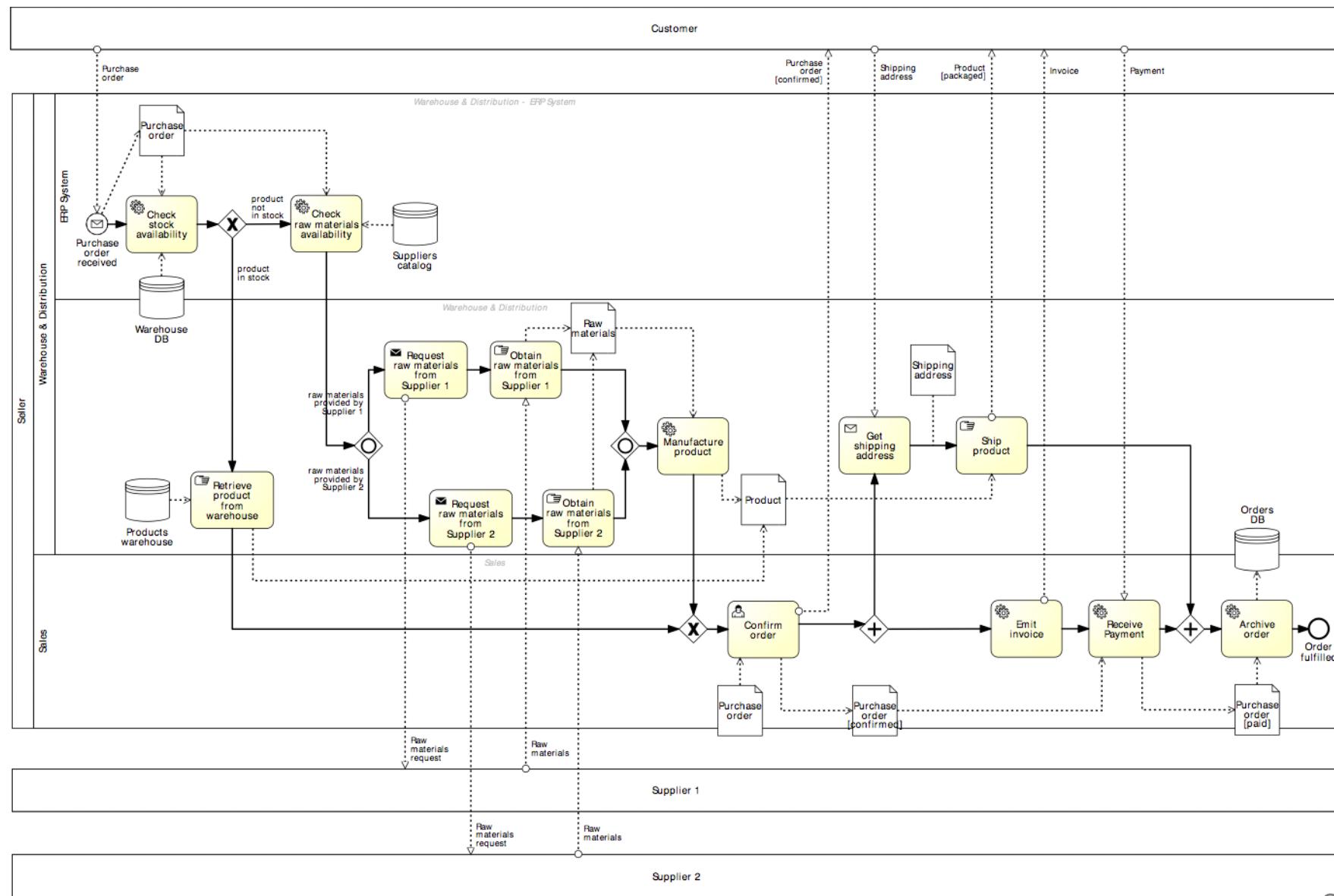
1. Identify the automation boundaries
2. Review manual tasks
3. Complete the process model
4. Adjust task granularity
5. Specify execution properties



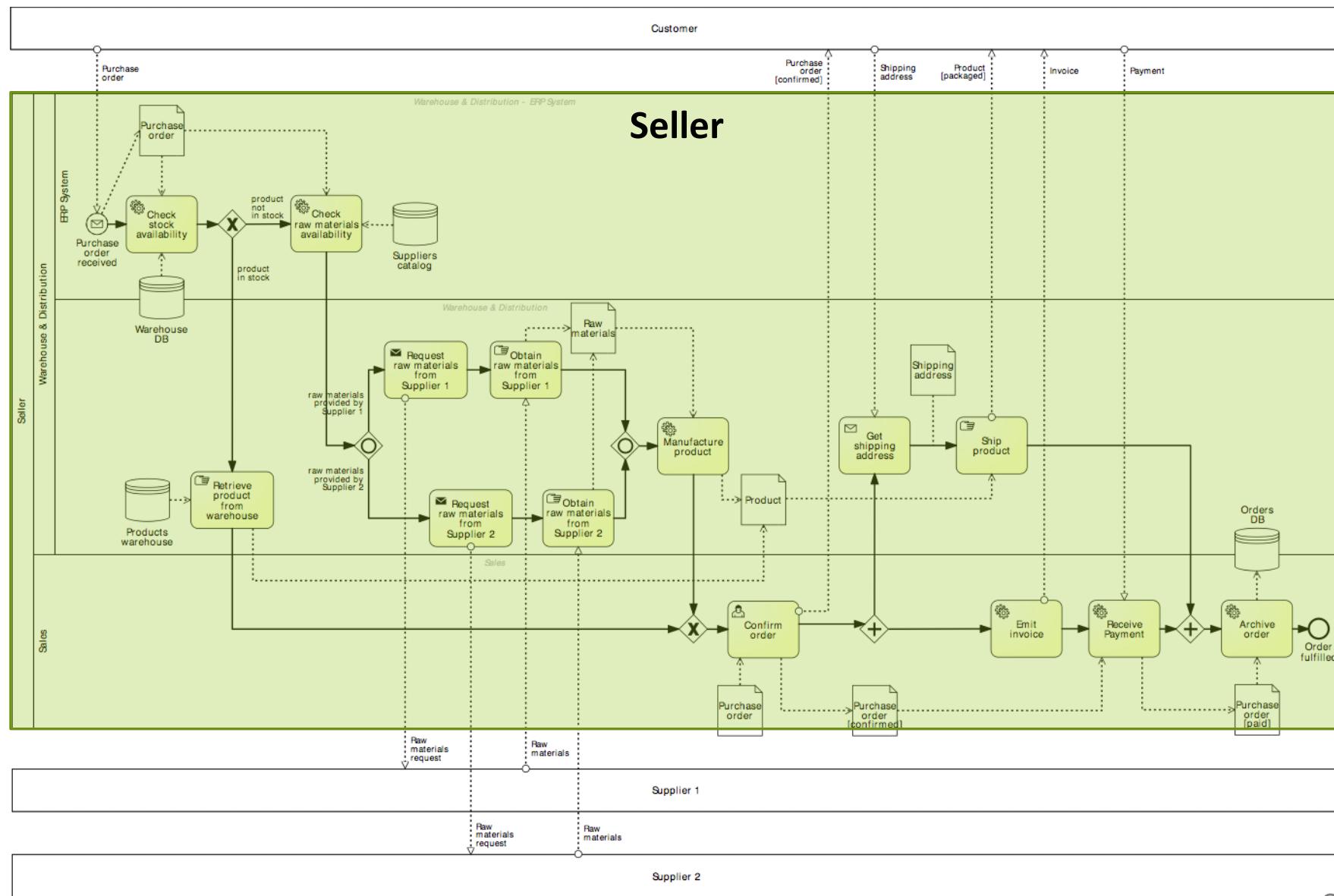
Running example



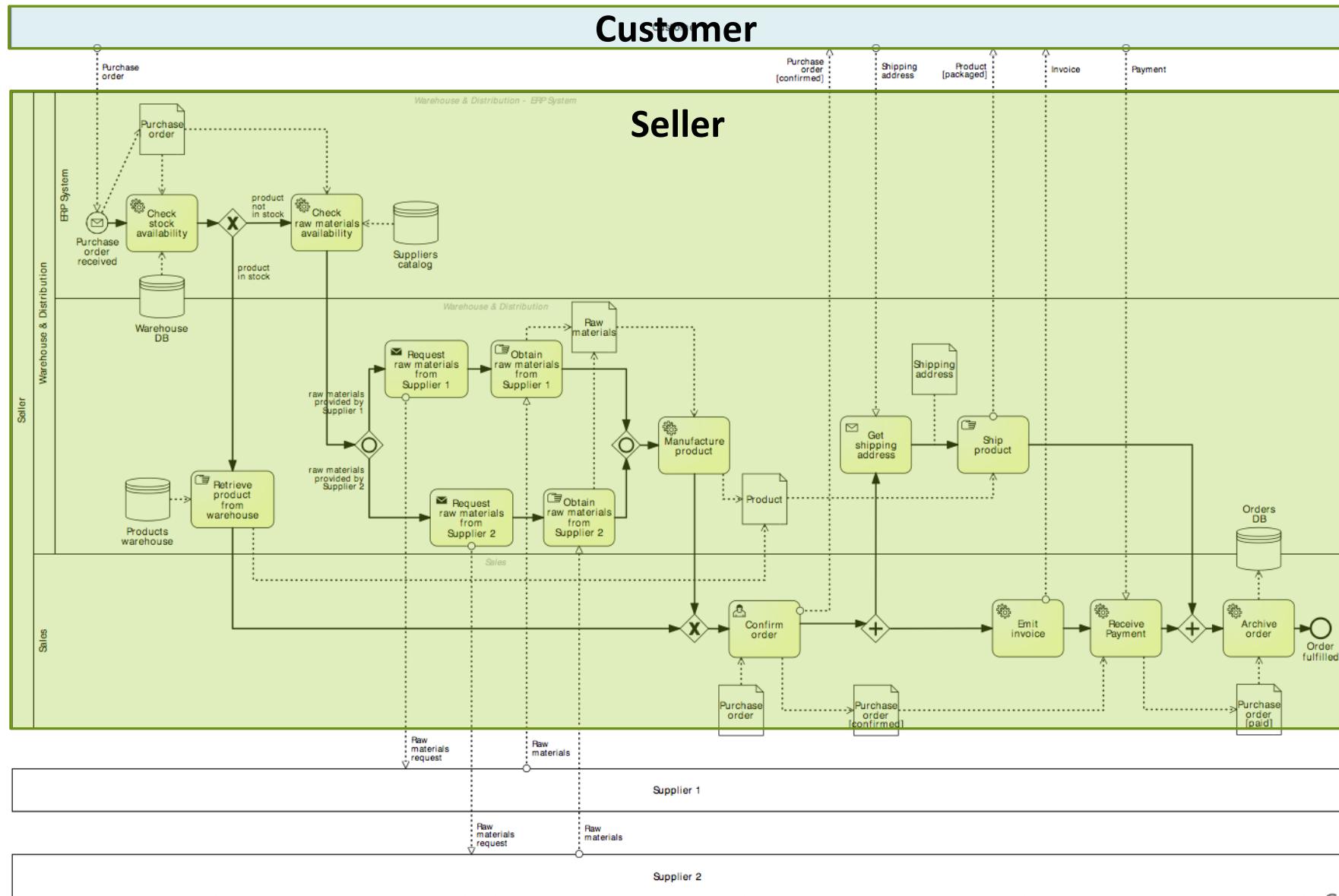
Running example



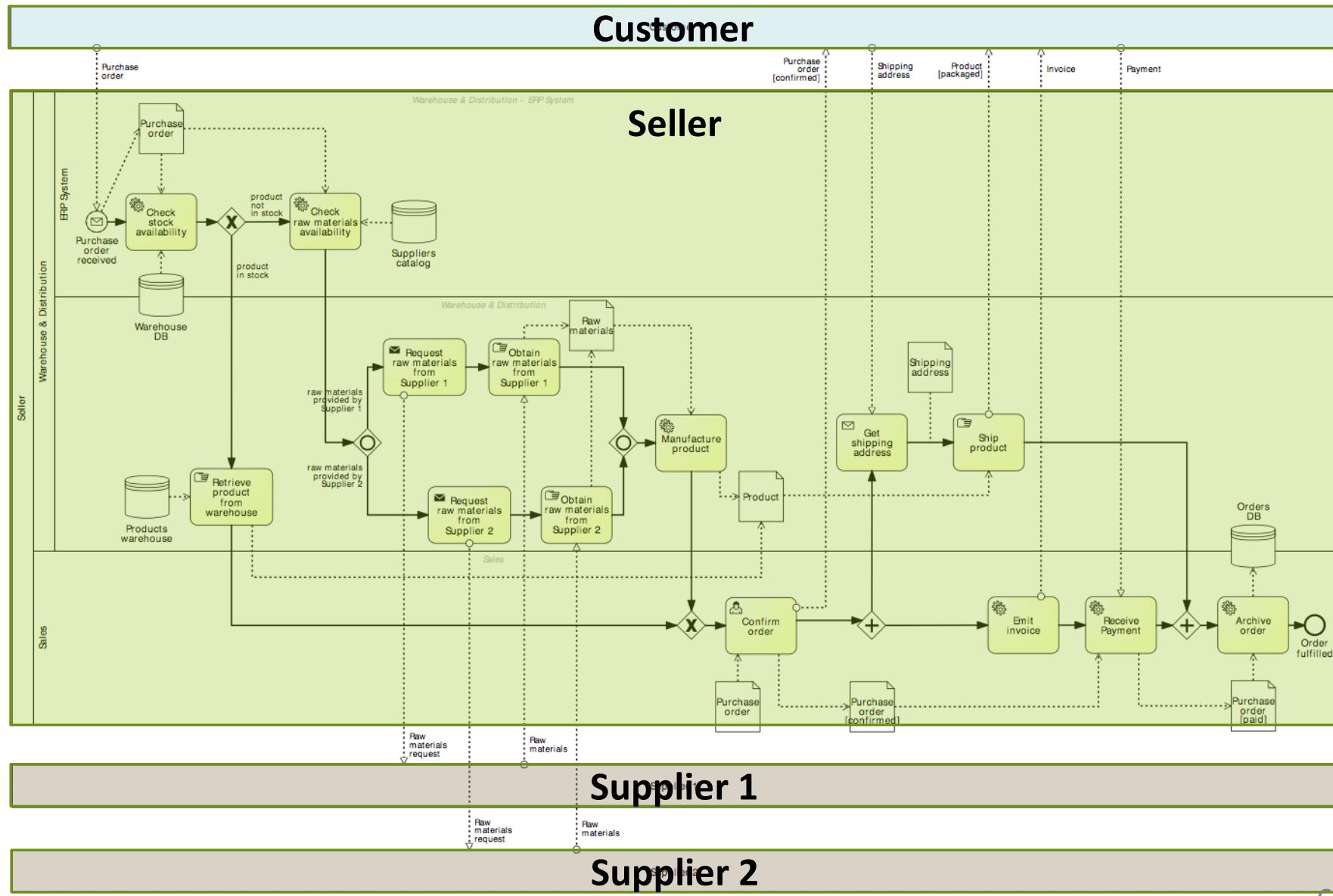
Running example



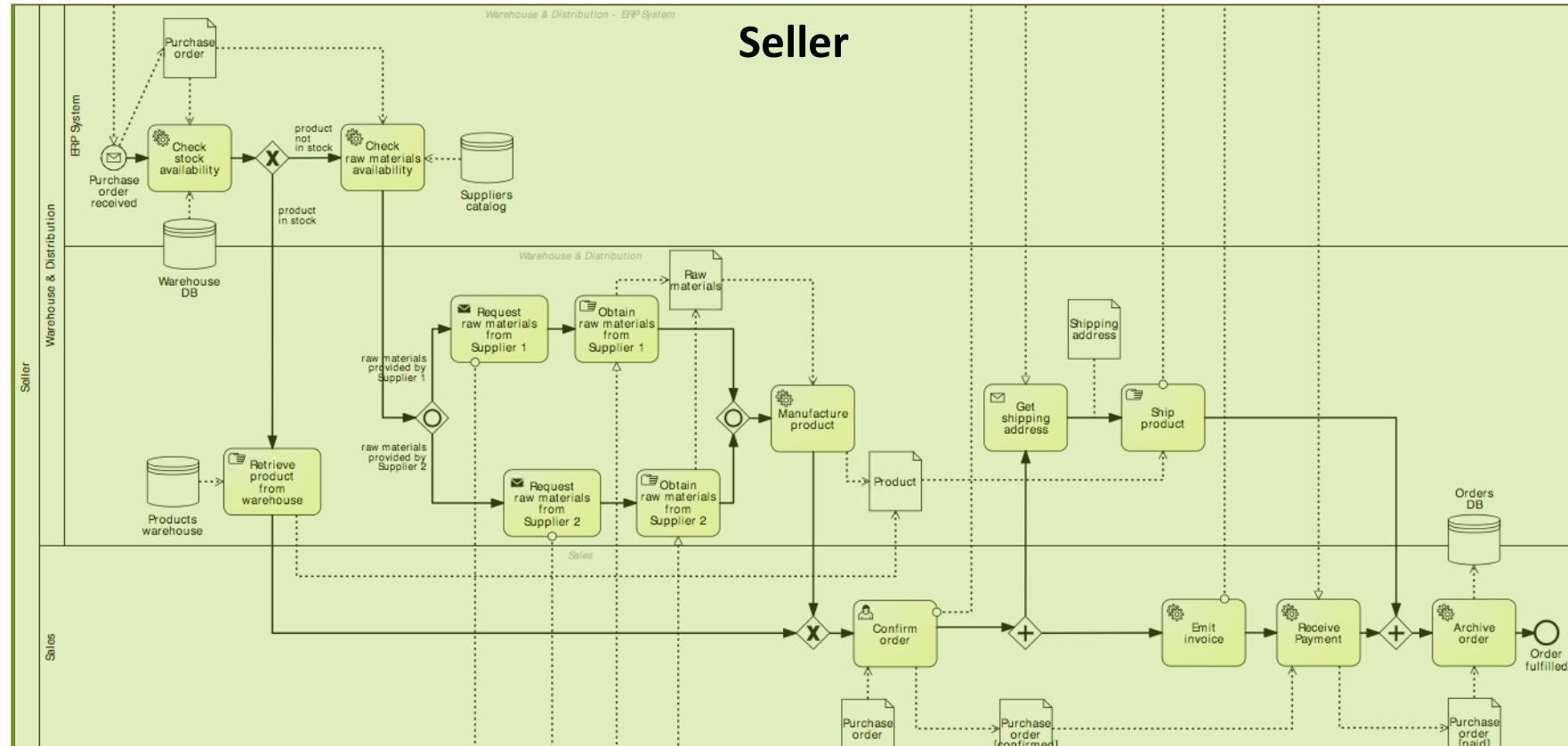
Running example



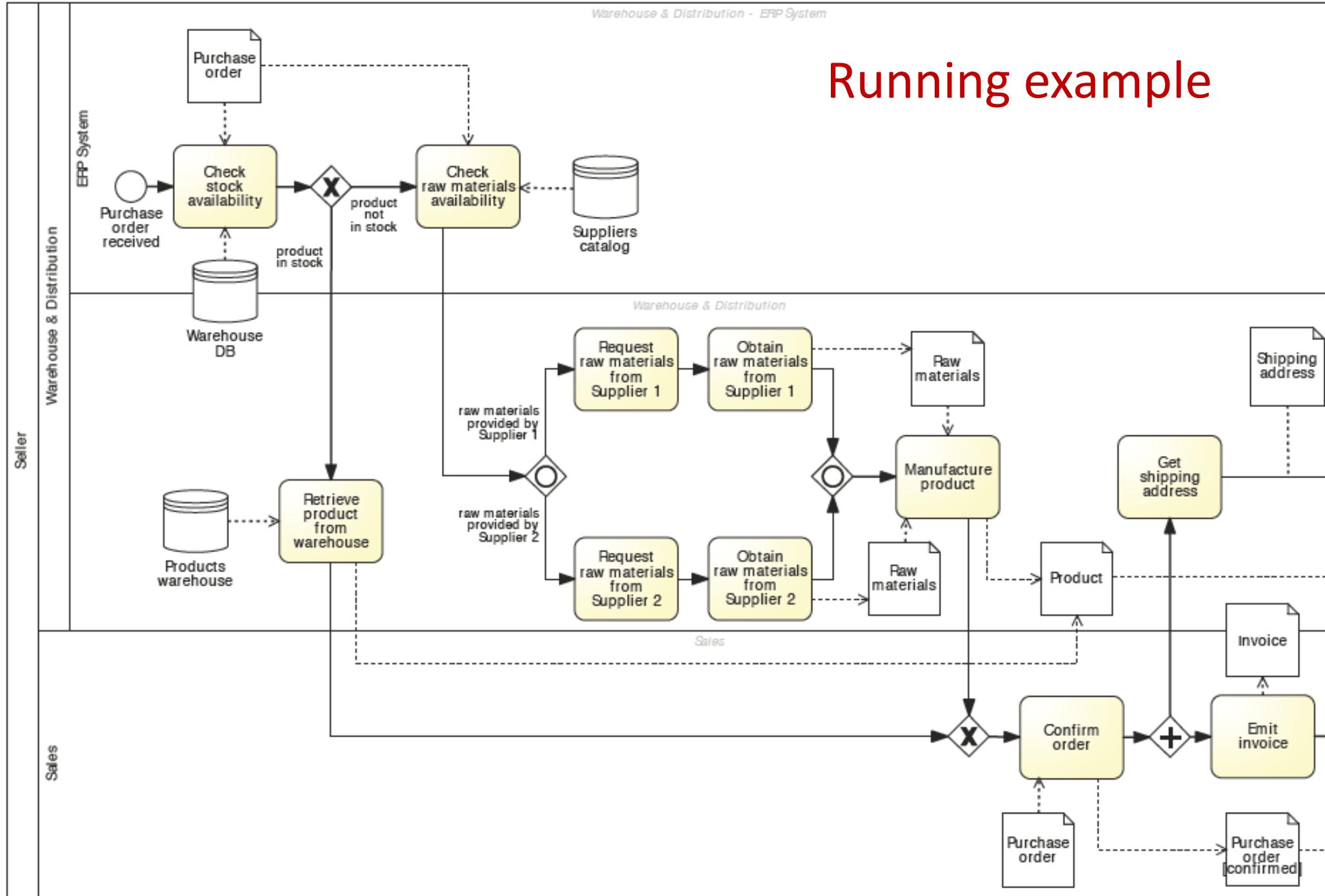
Running example



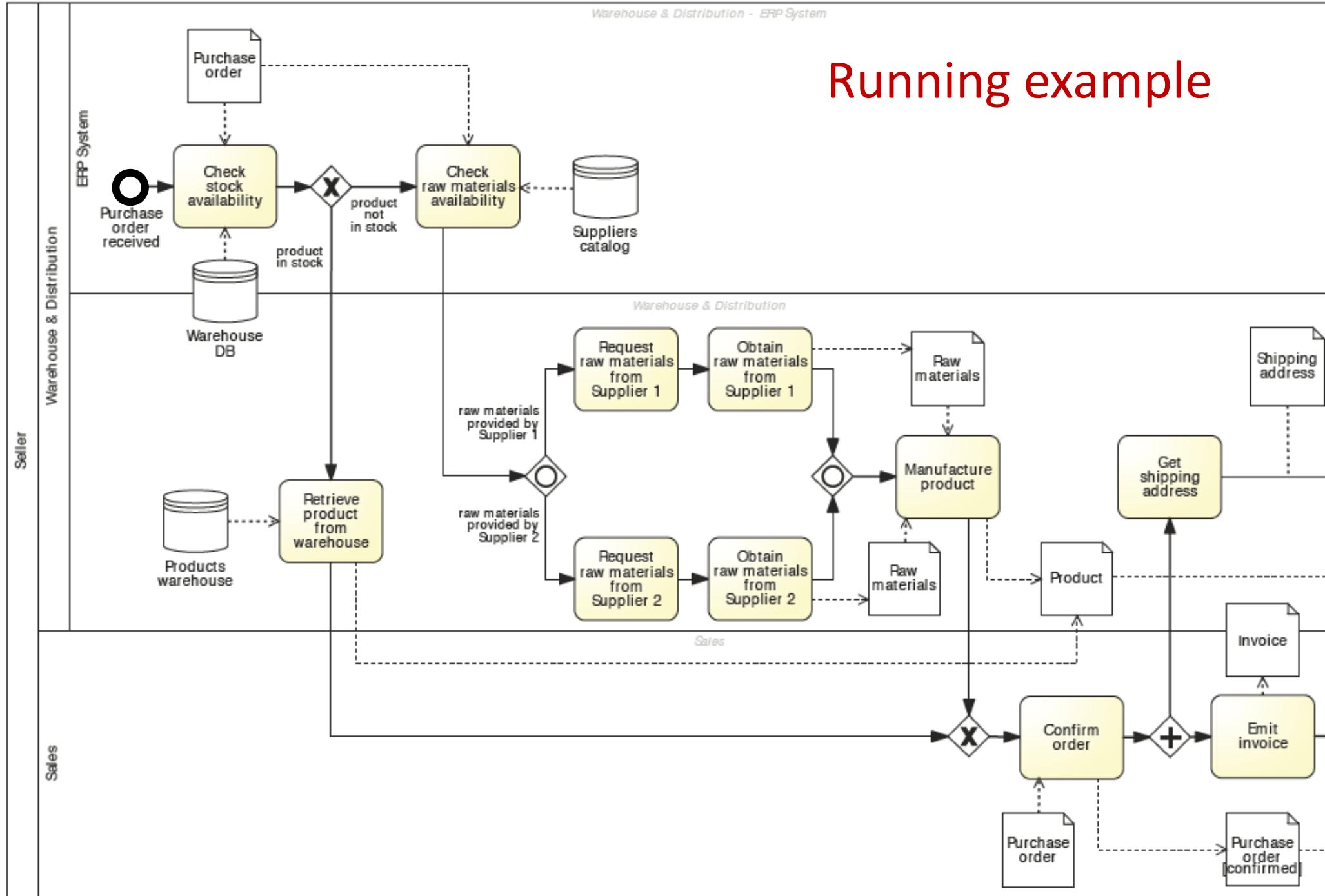
Running example



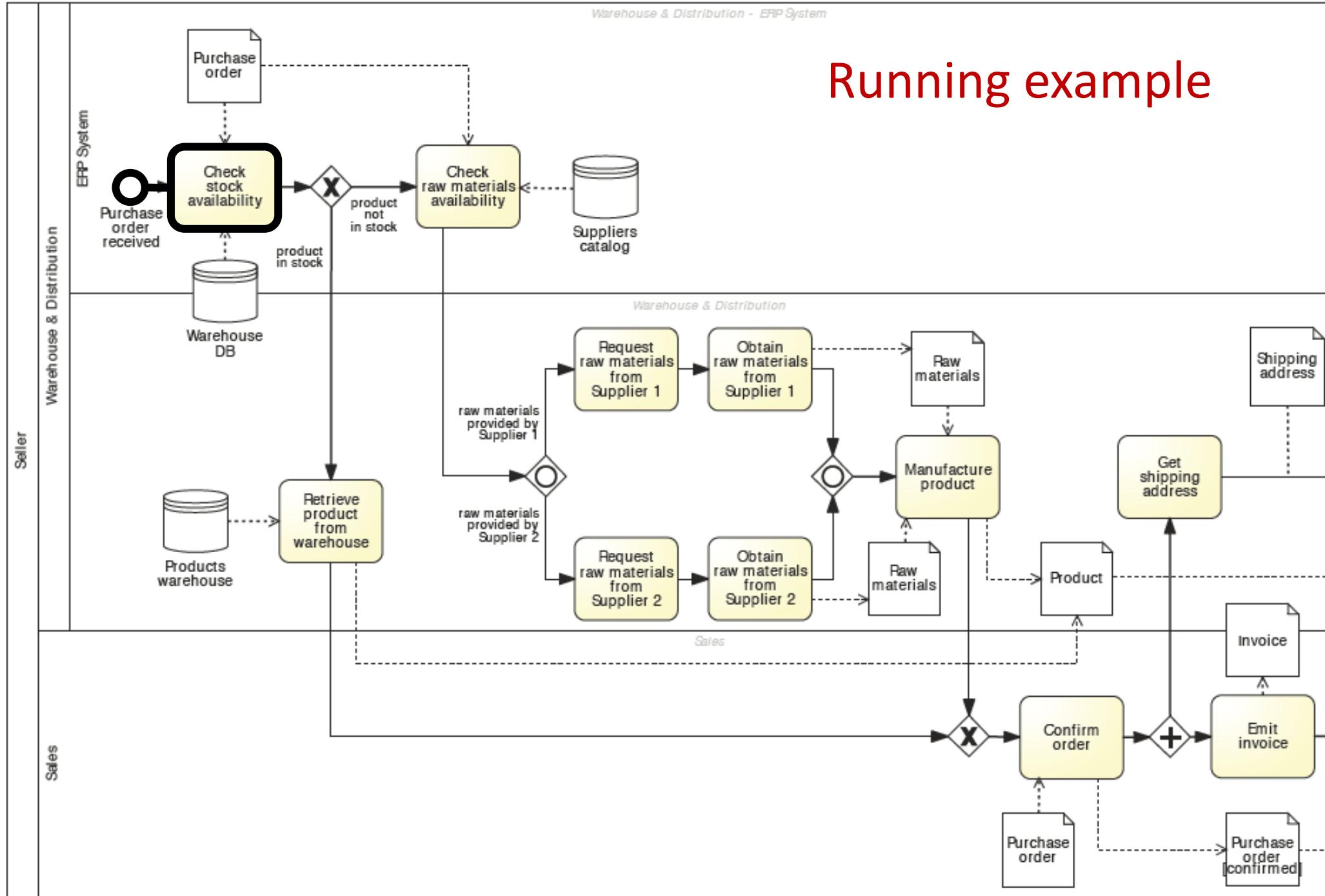
Running example



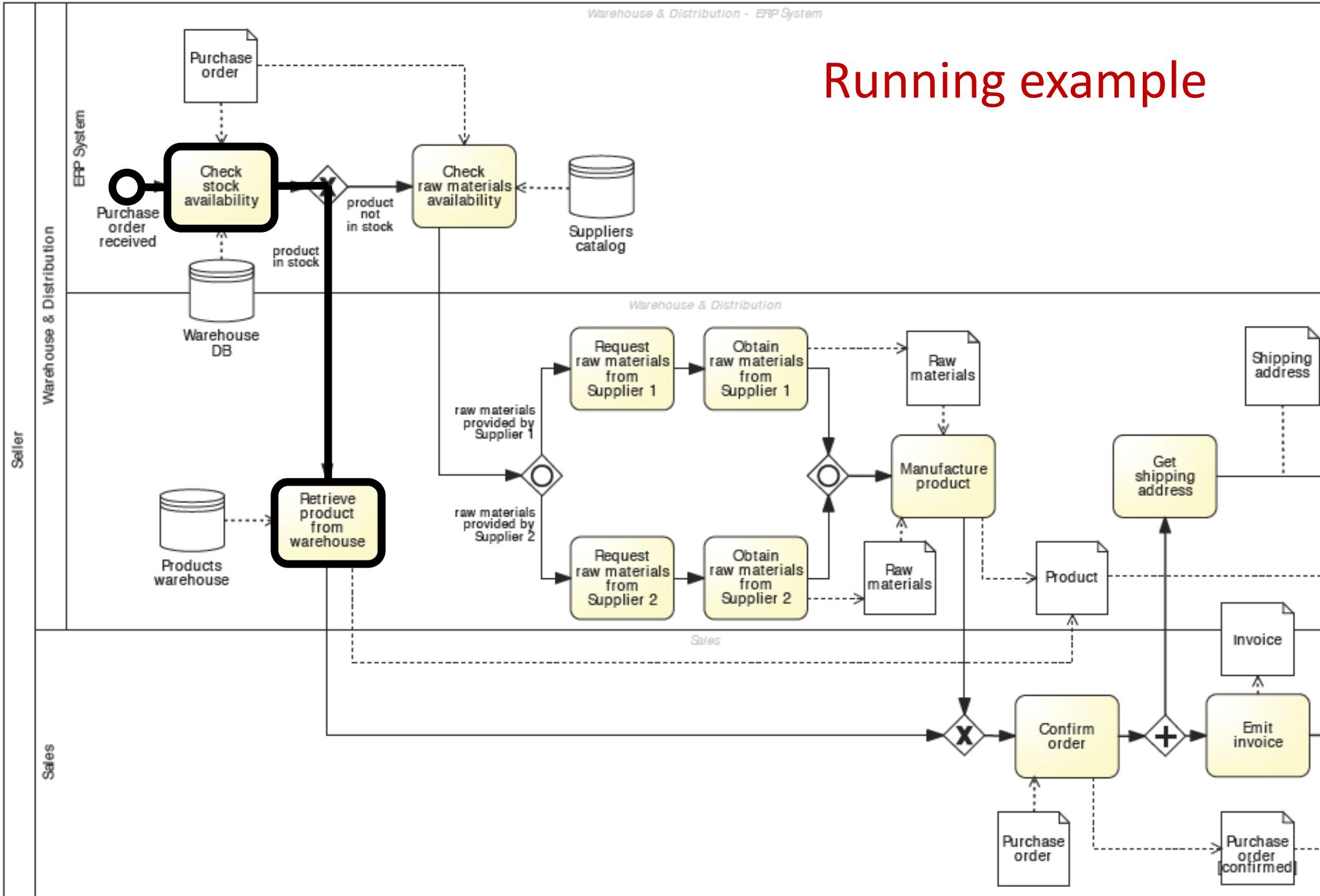
Running example



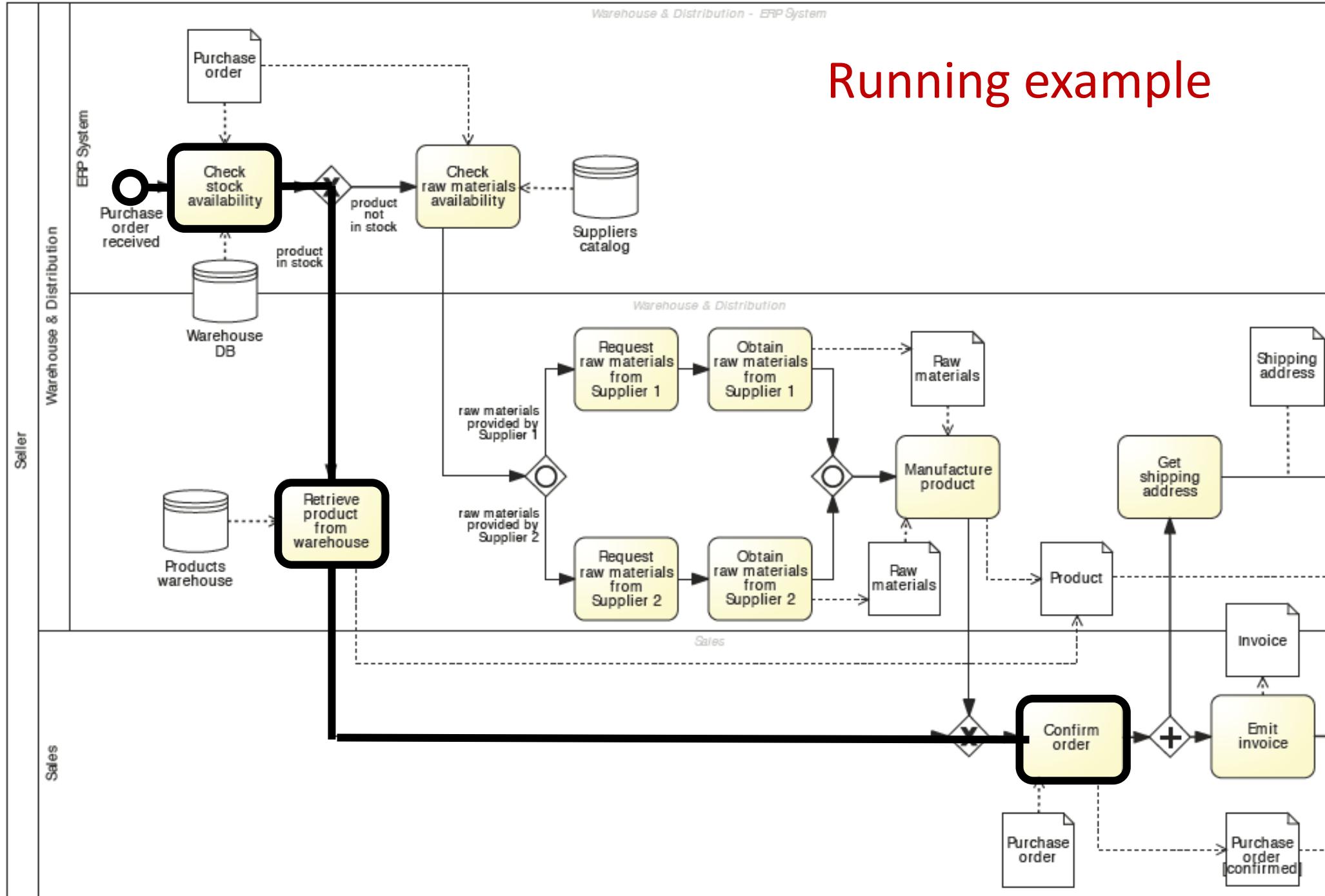
Running example



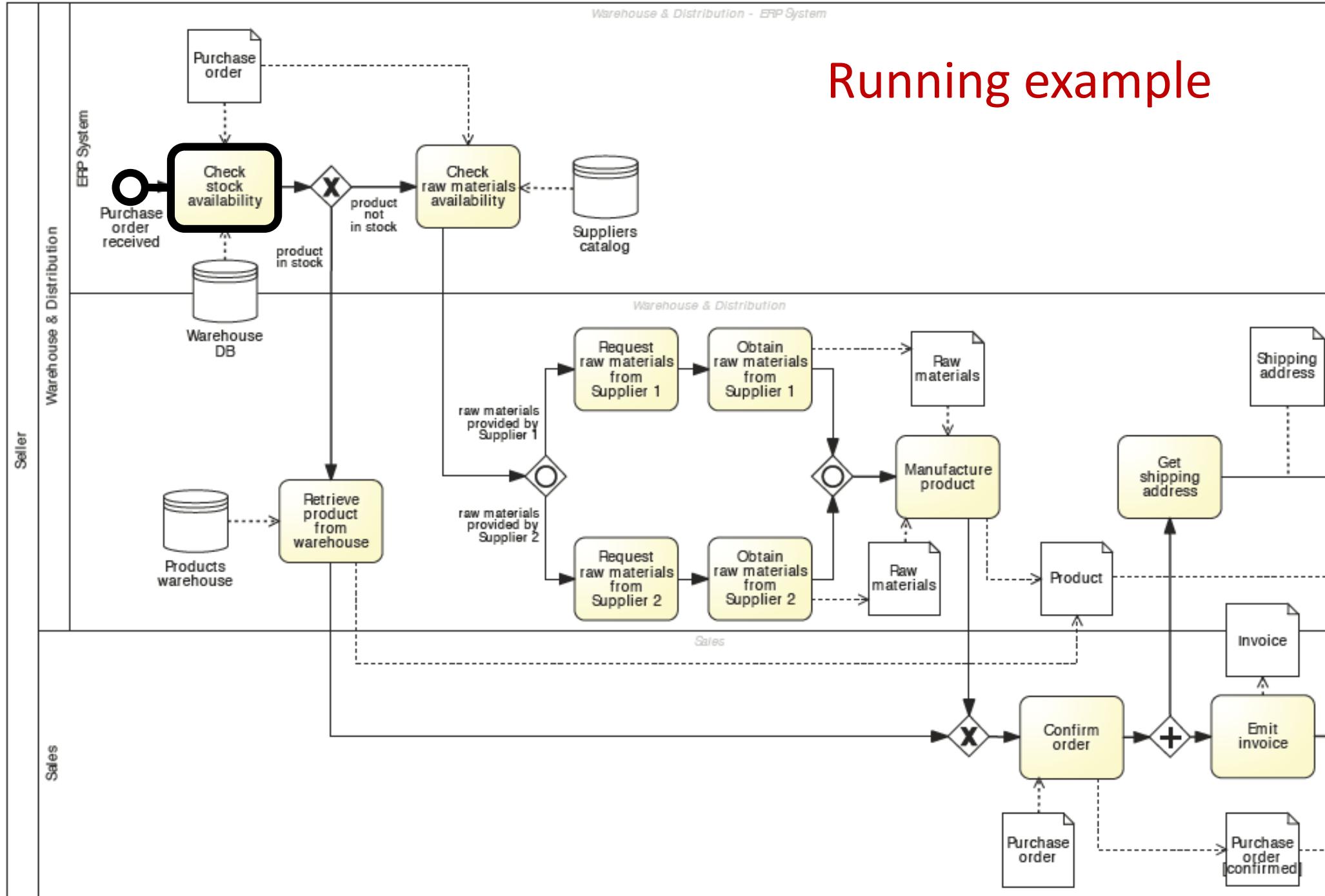
Running example



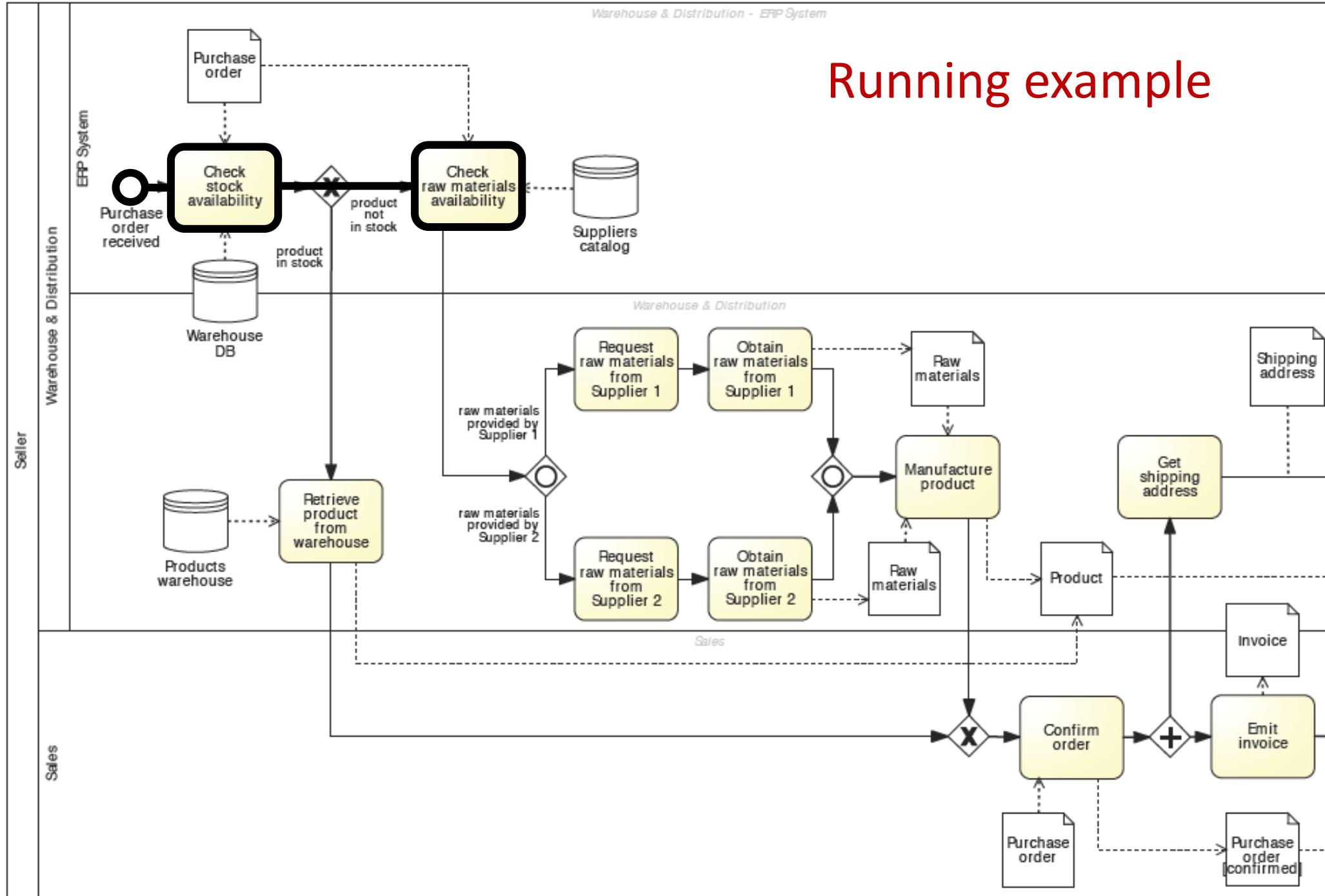
Running example



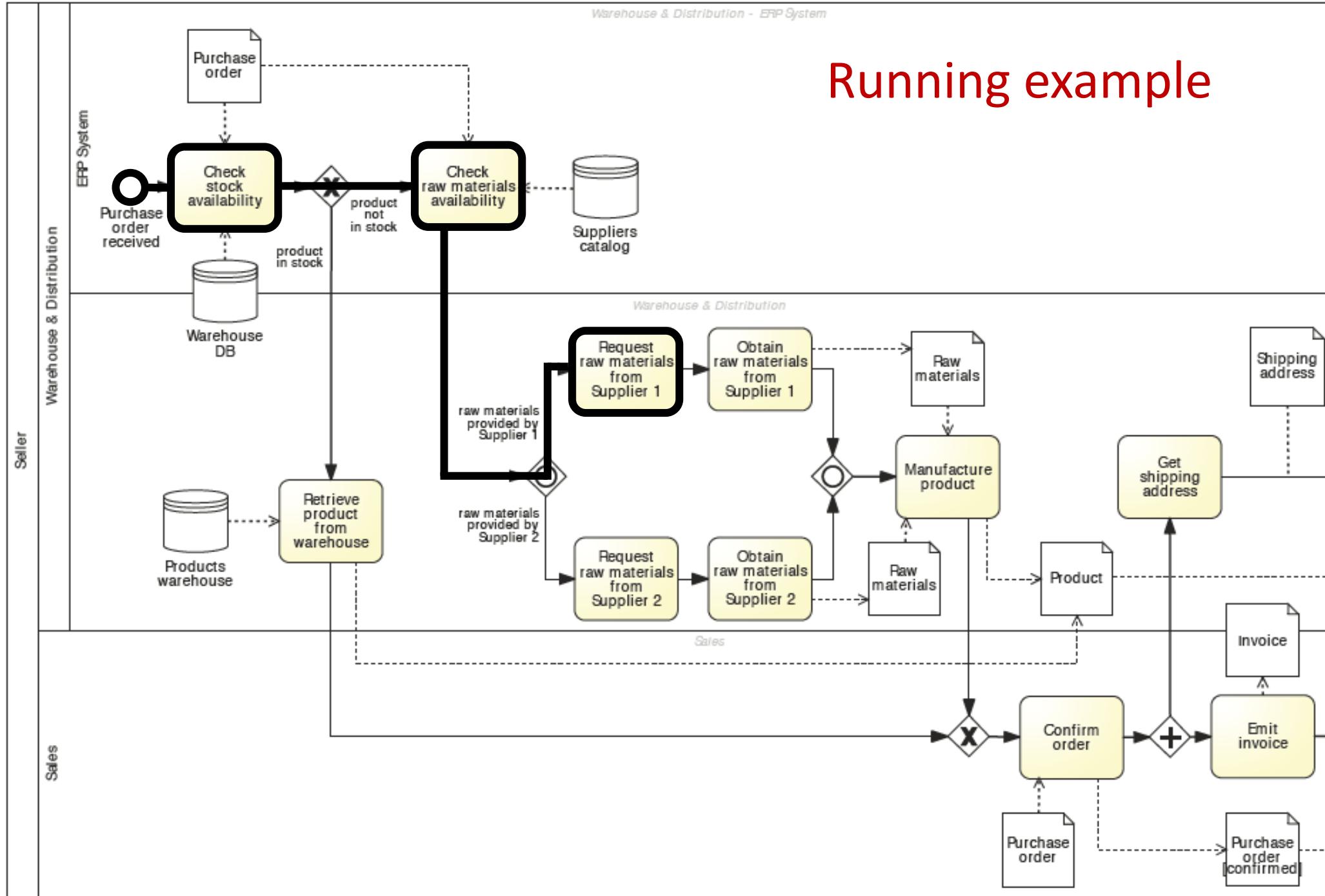
Running example



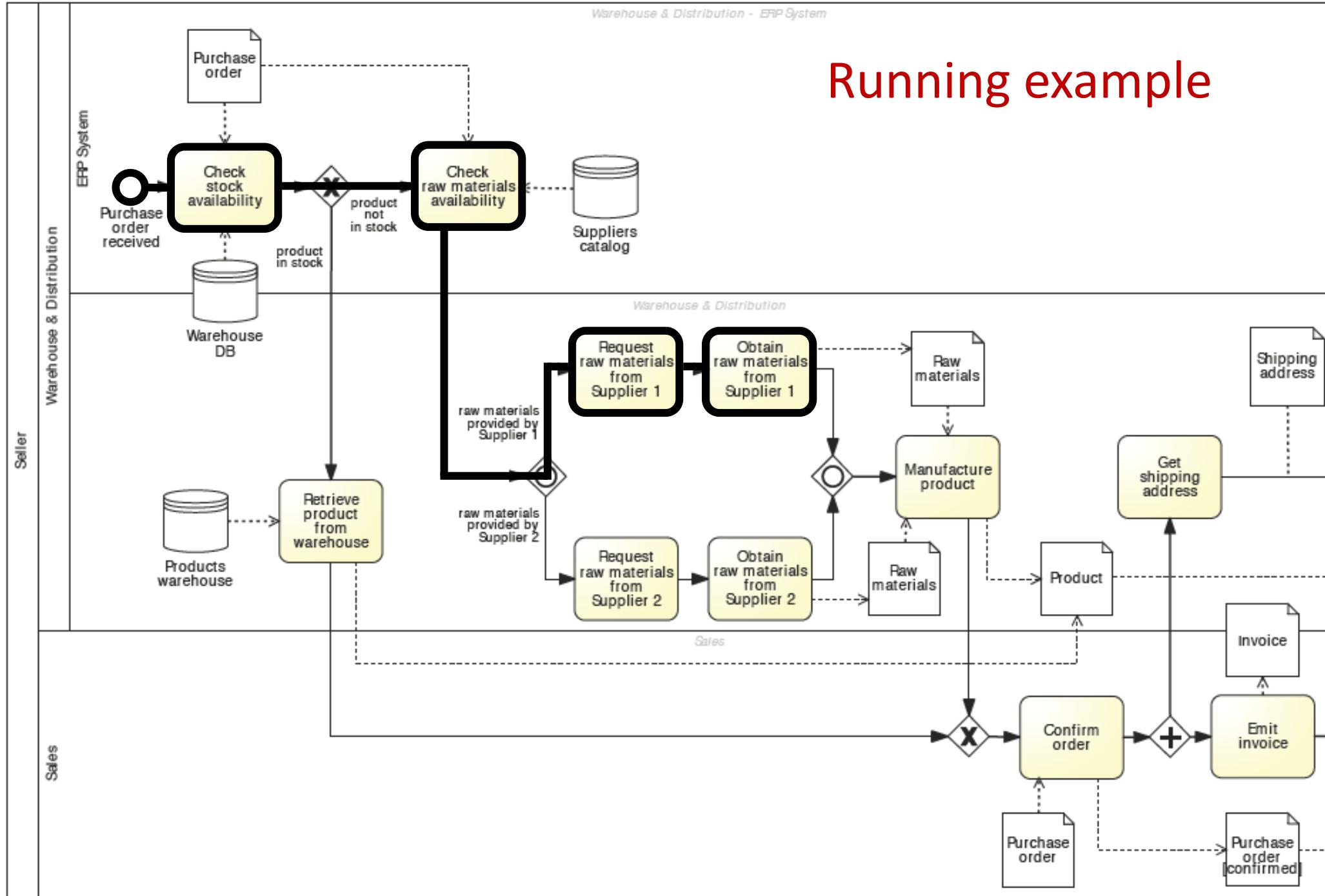
Running example



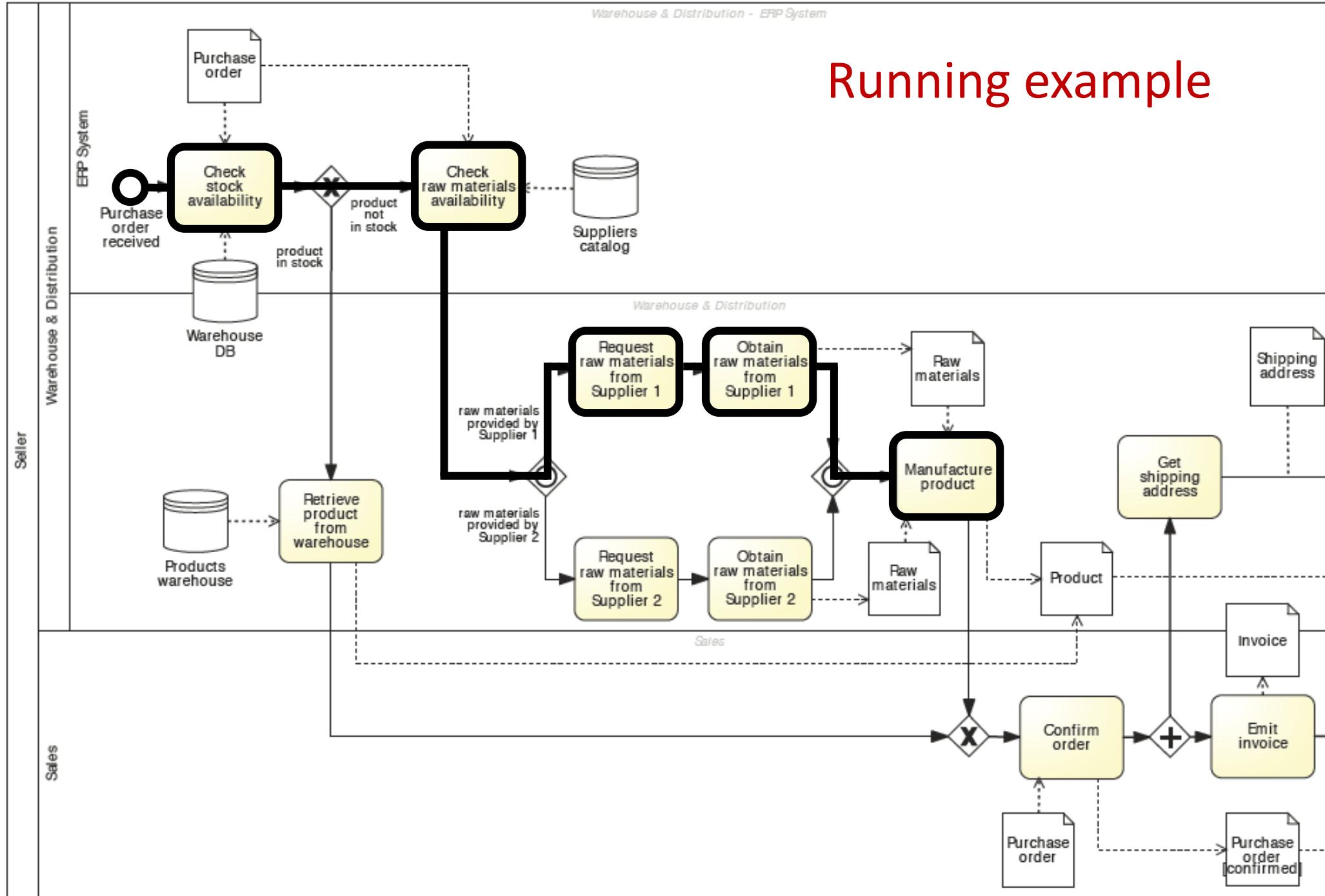
Running example



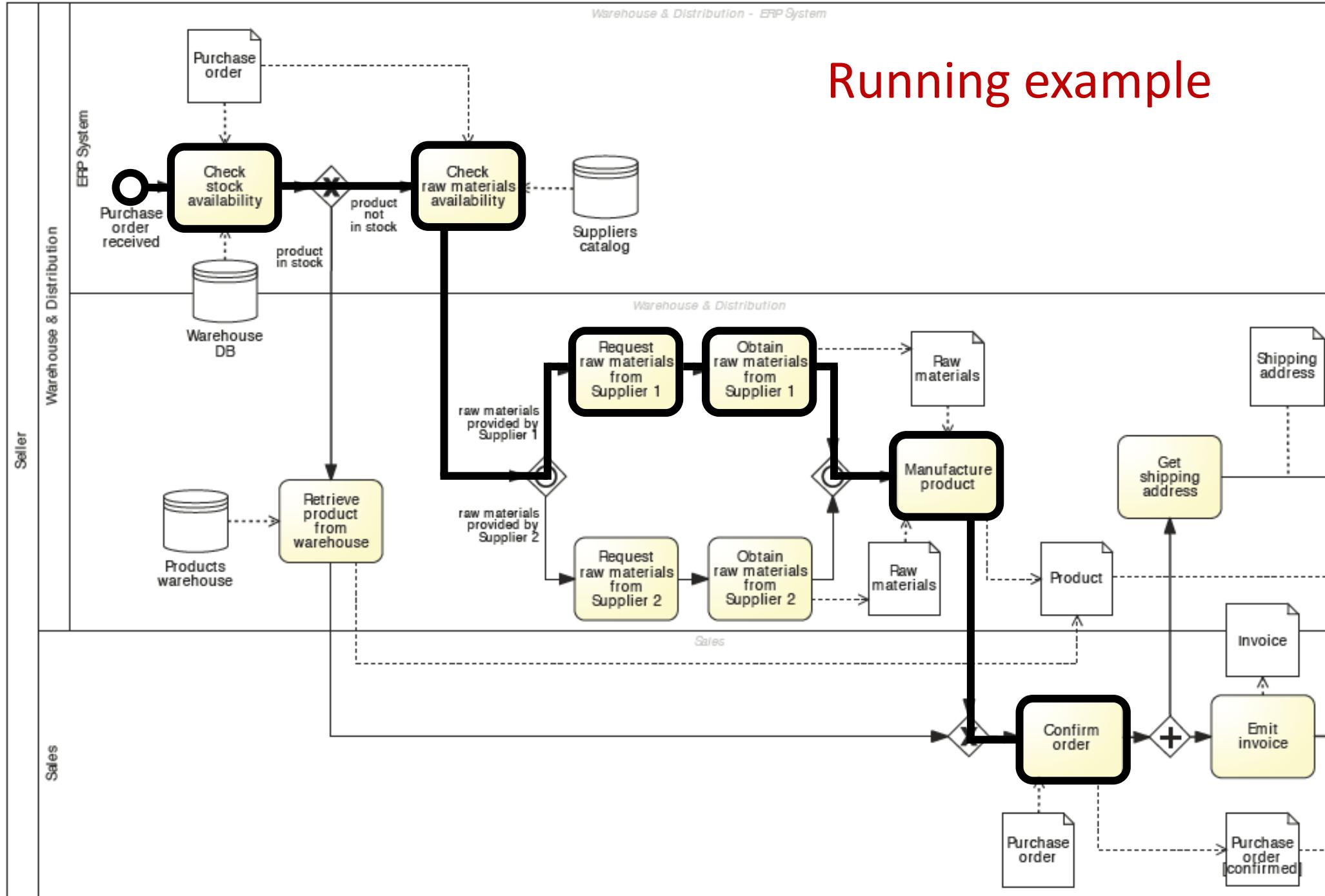
Running example



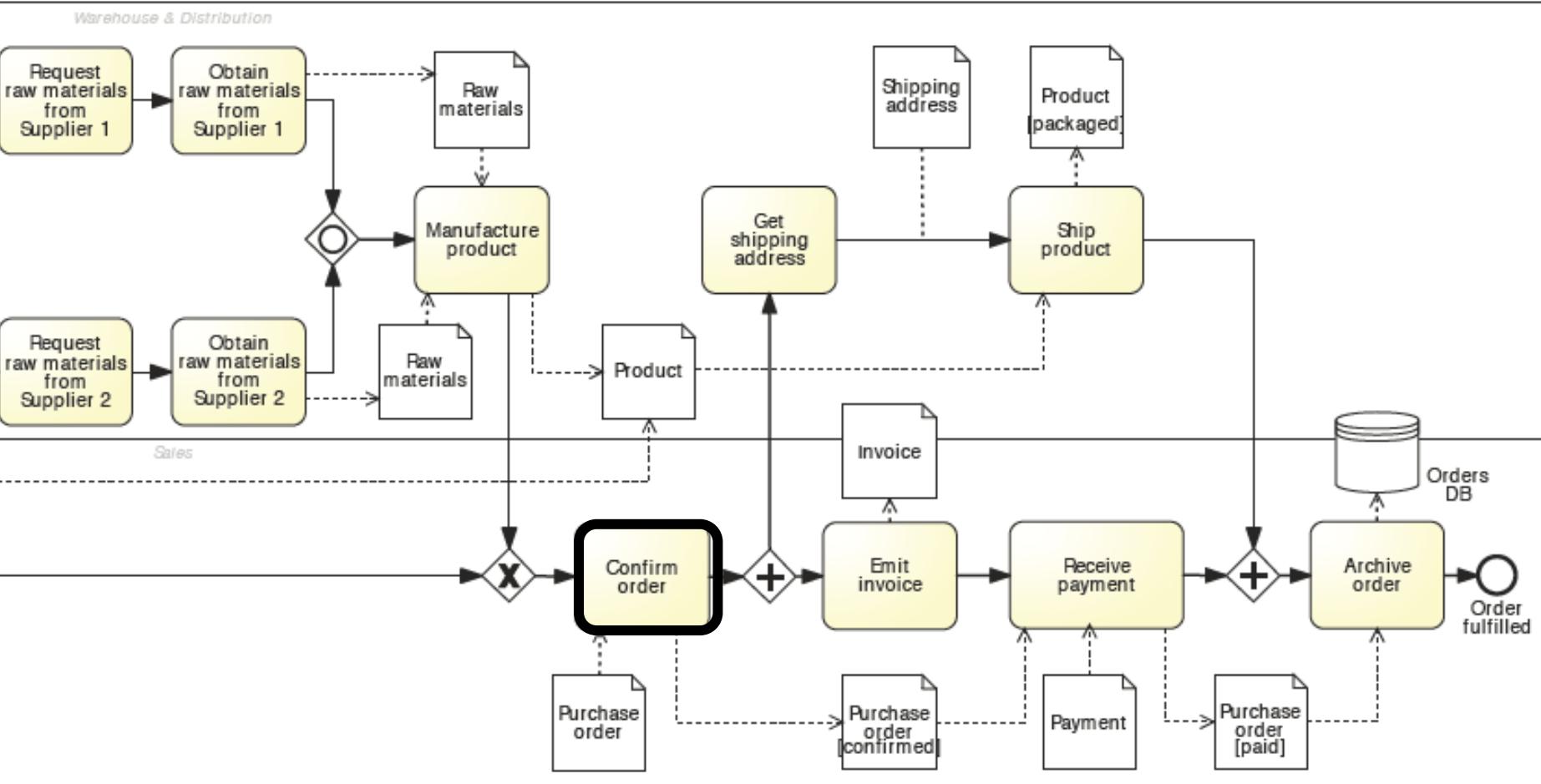
Running example



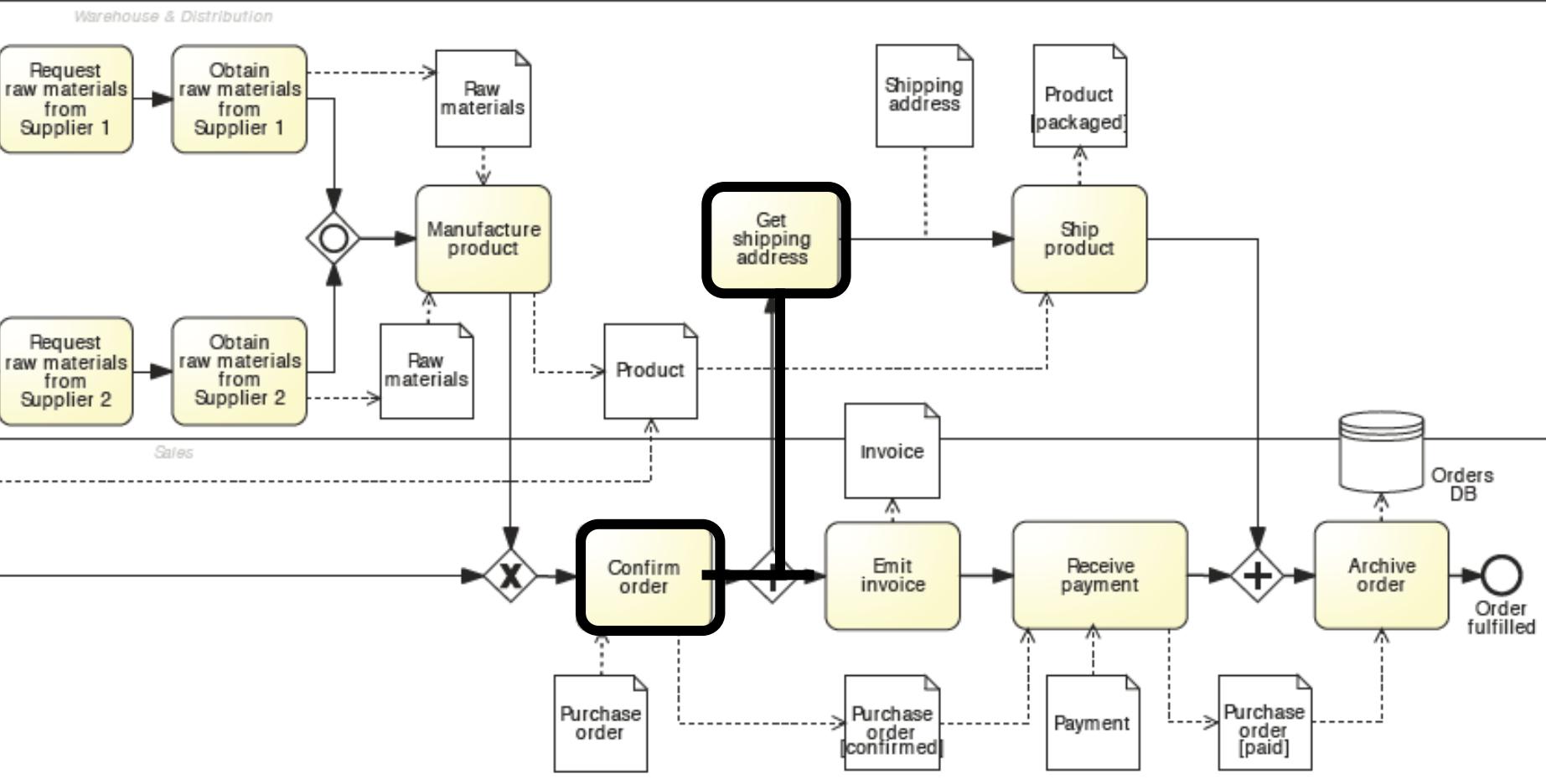
Running example



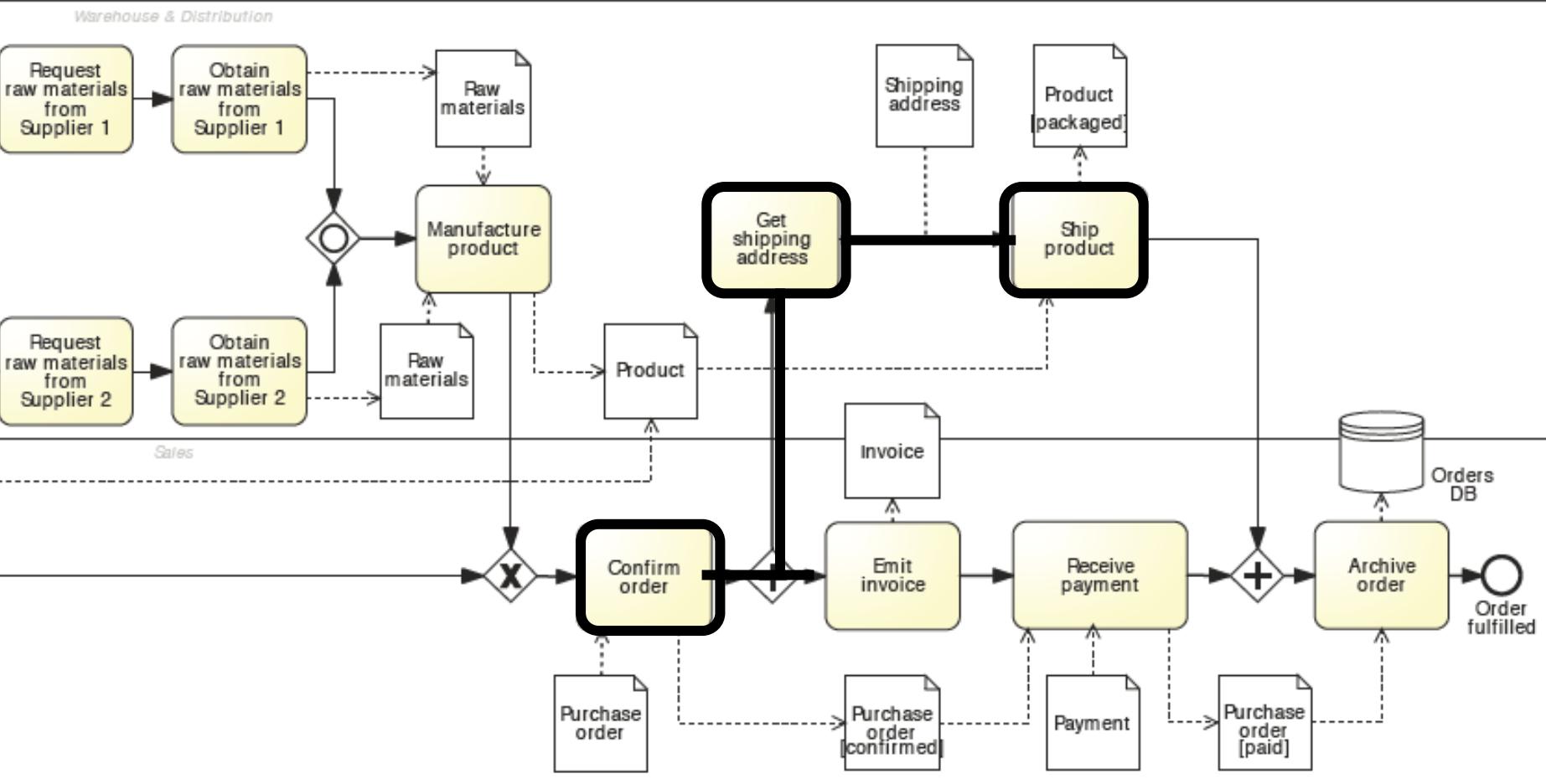
Running example



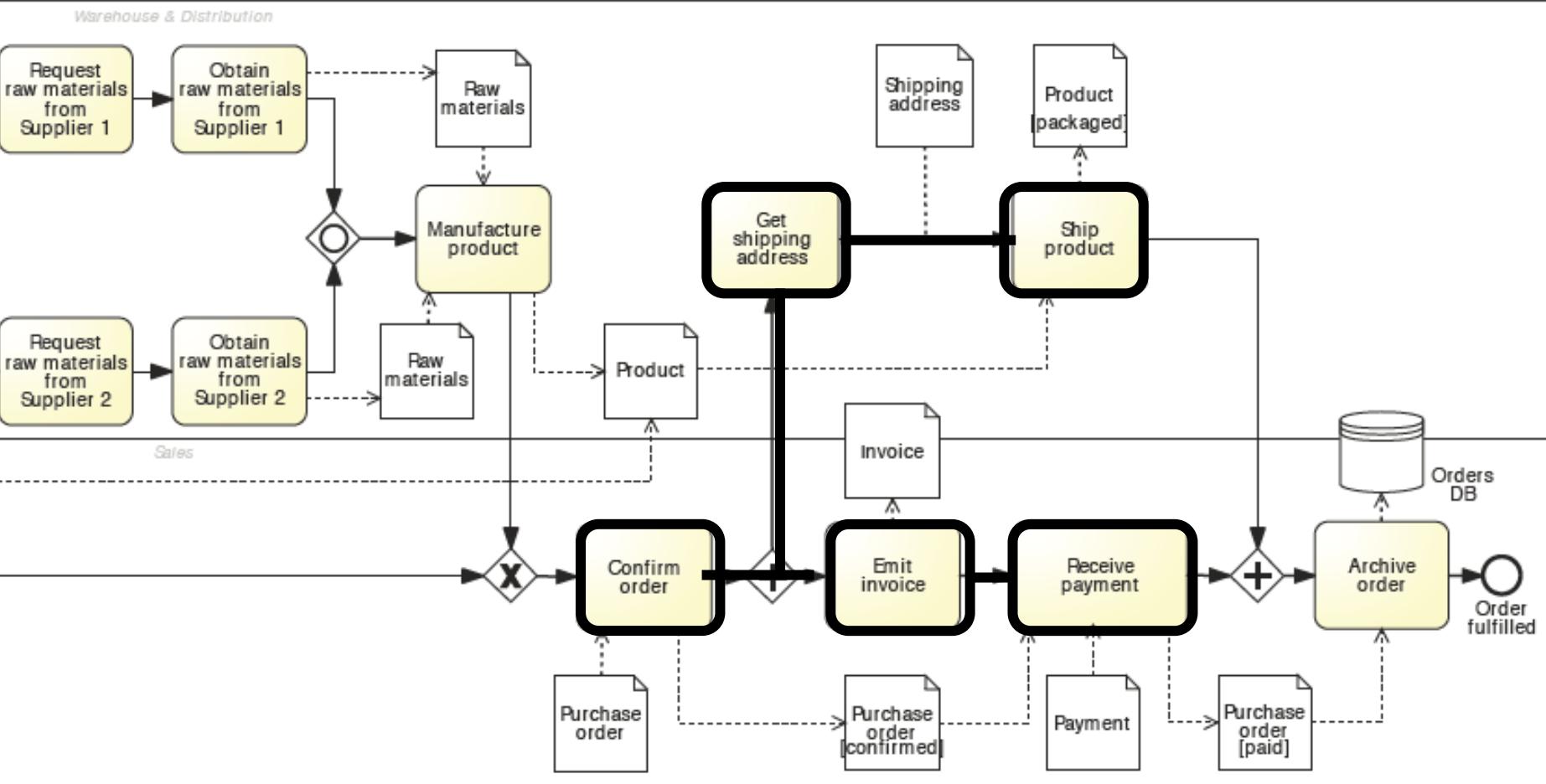
Running example



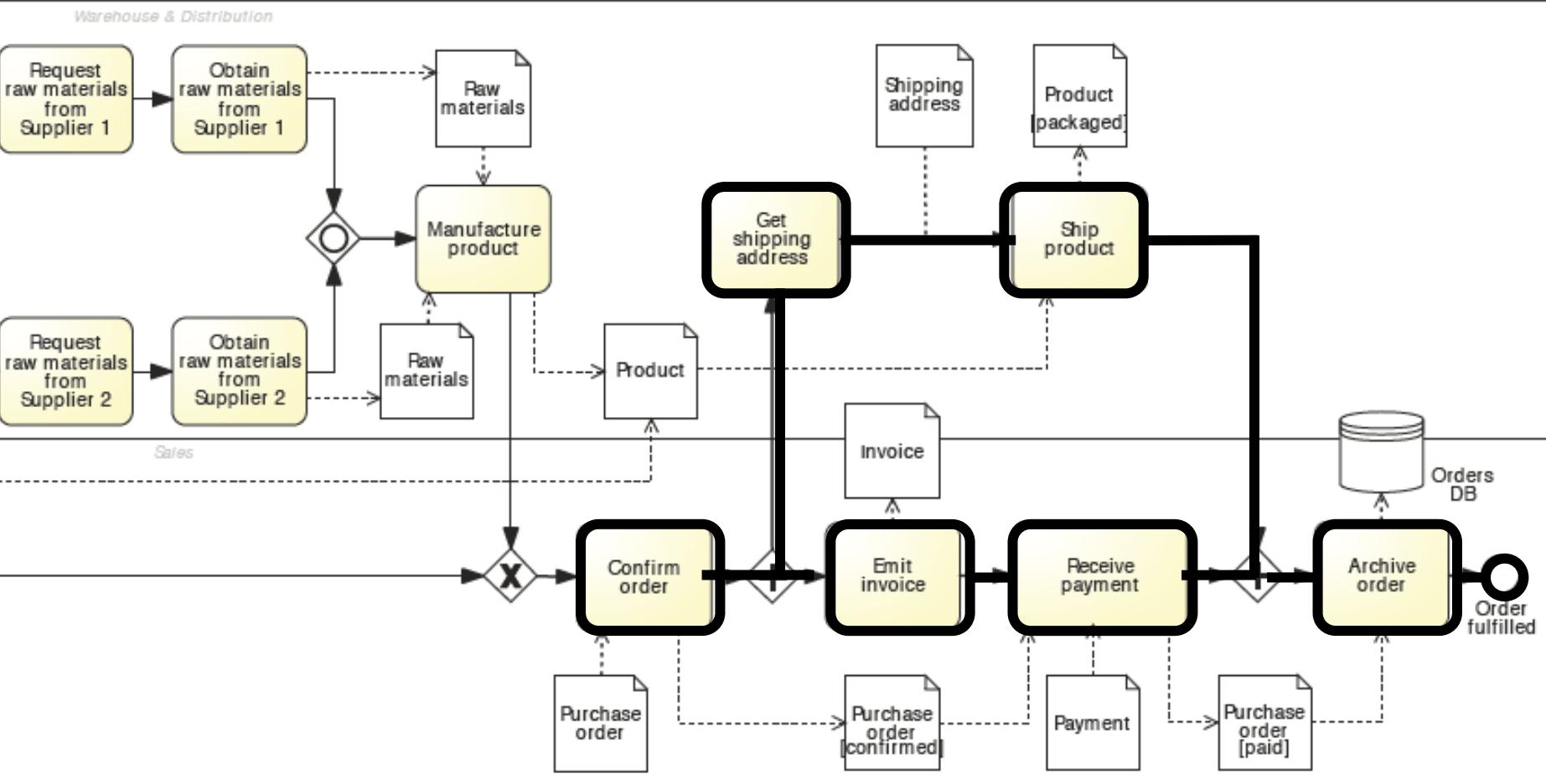
Running example



Running example



Running example



1. Identify the automation boundaries

Principle: not all parts of a process can be automated.

1. Identify the automation boundaries

Principle: not all parts of a process can be automated.

-> Start by identifying each task's type:

1. Identify the automation boundaries

Principle: not all parts of a process can be automated.

-> Start by identifying each task's type:

1



Automated tasks

1. Identify the automation boundaries

Principle: not all parts of a process can be automated.

-> Start by identifying each task's type:

1



Automated tasks

2



User tasks

1. Identify the automation boundaries

Principle: not all parts of a process can be automated.

-> Start by identifying each task's type:

1



Automated tasks

2



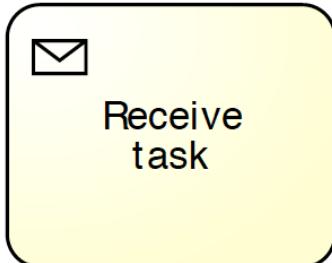
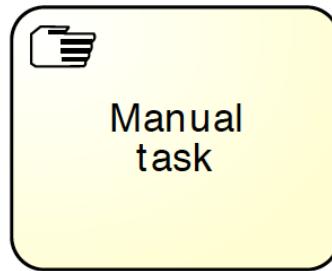
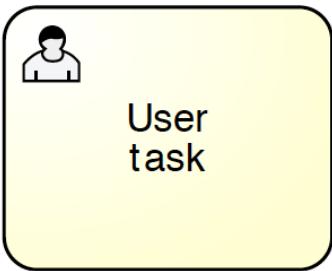
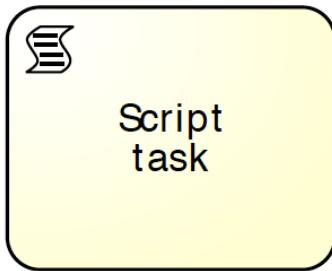
User tasks

3



Manual tasks

In BPMN: specify task markers



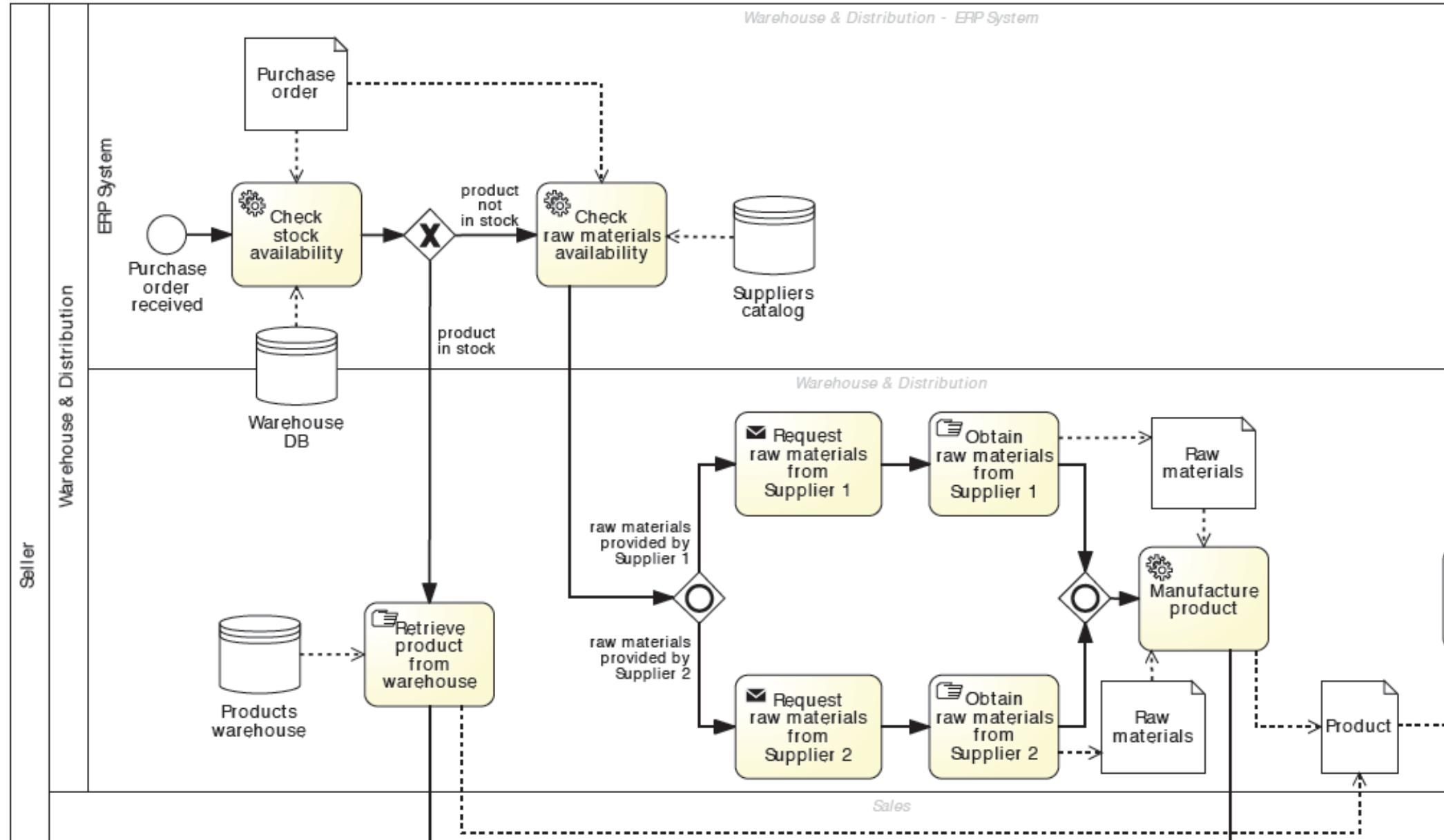
Automated tasks

User task

Manual task

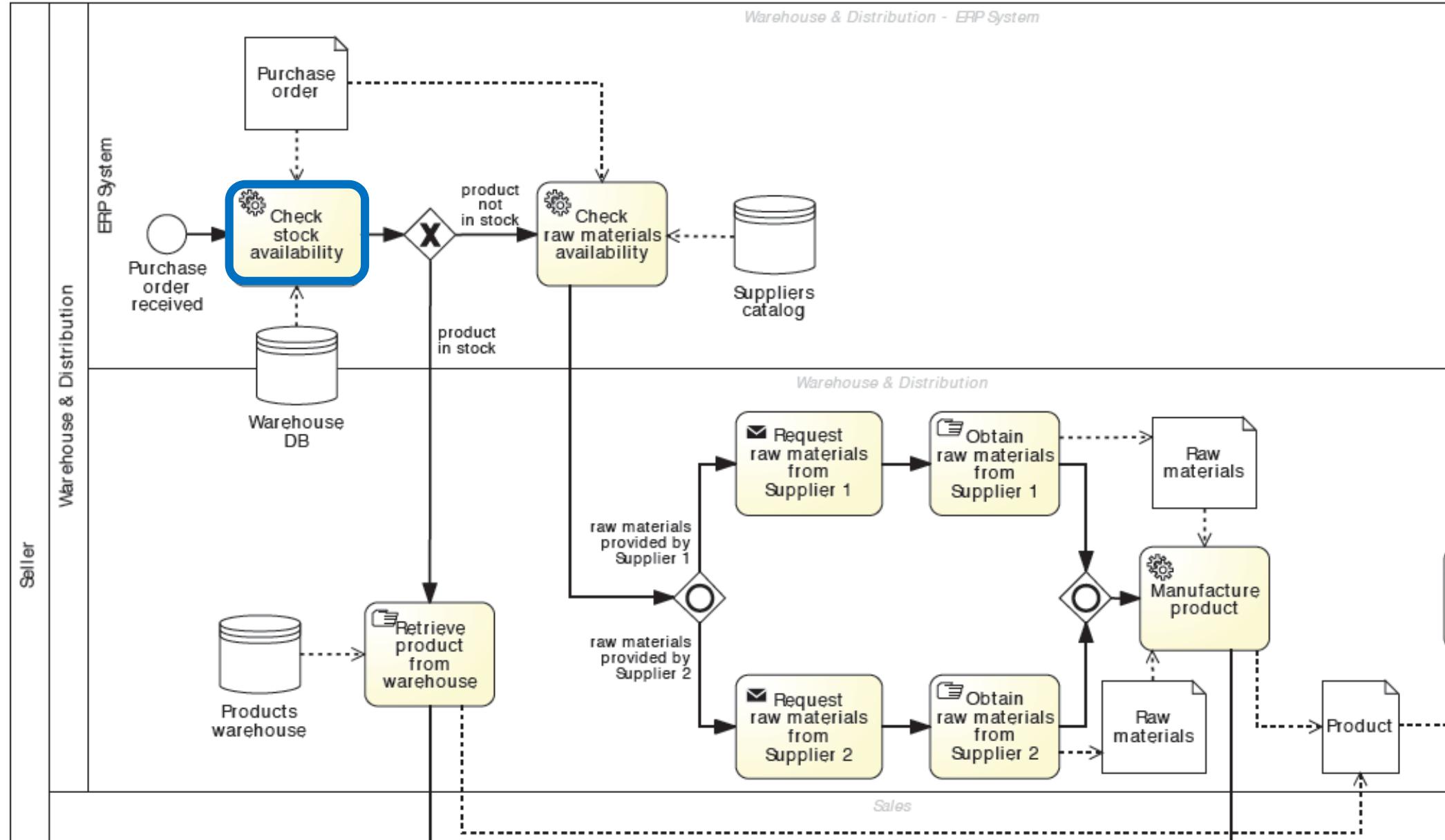
In our example...

█ automated
█ user
█ manual



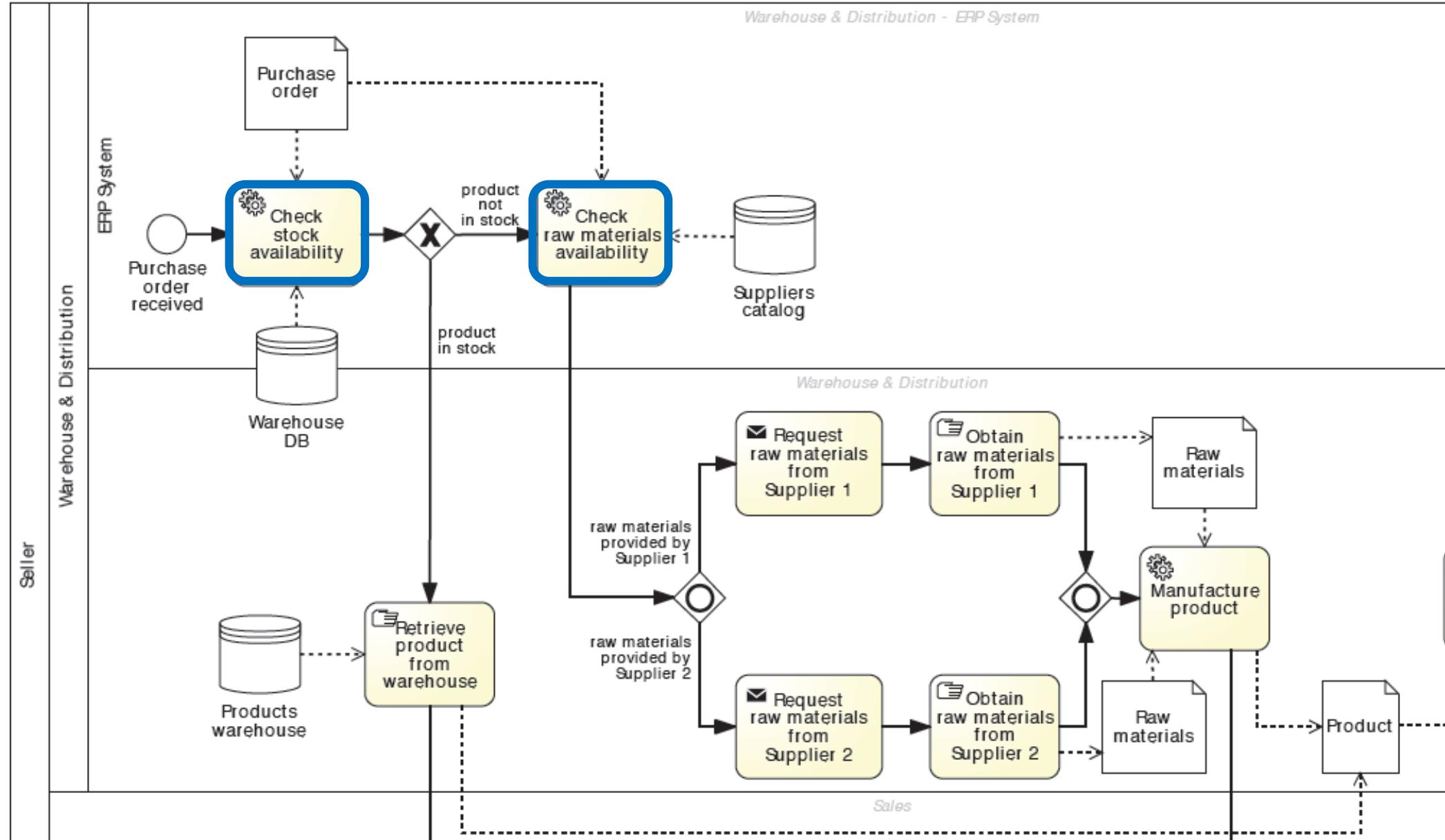
In our example...

█ automated
█ user
█ manual



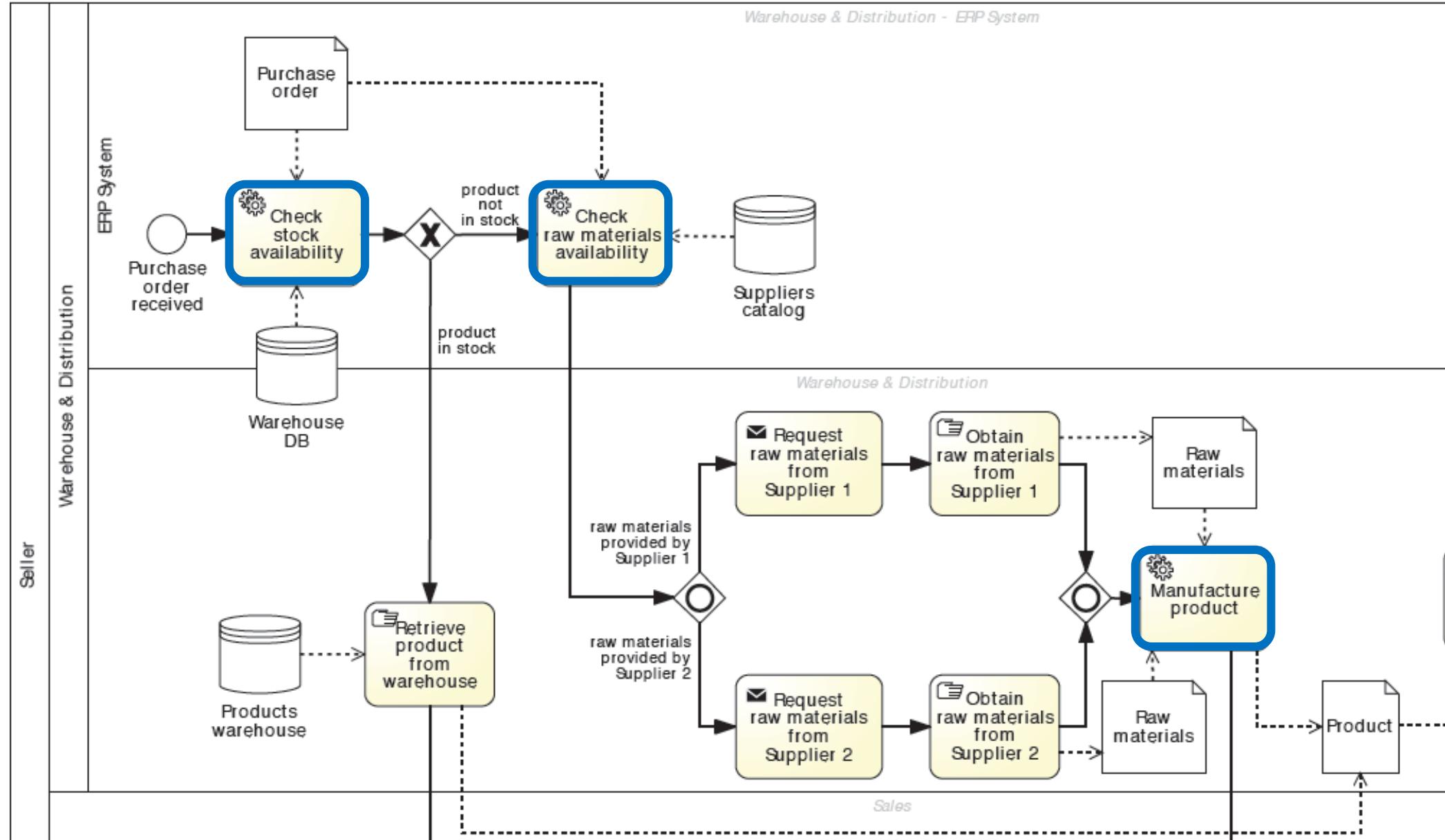
In our example...

█ automated
█ user
█ manual



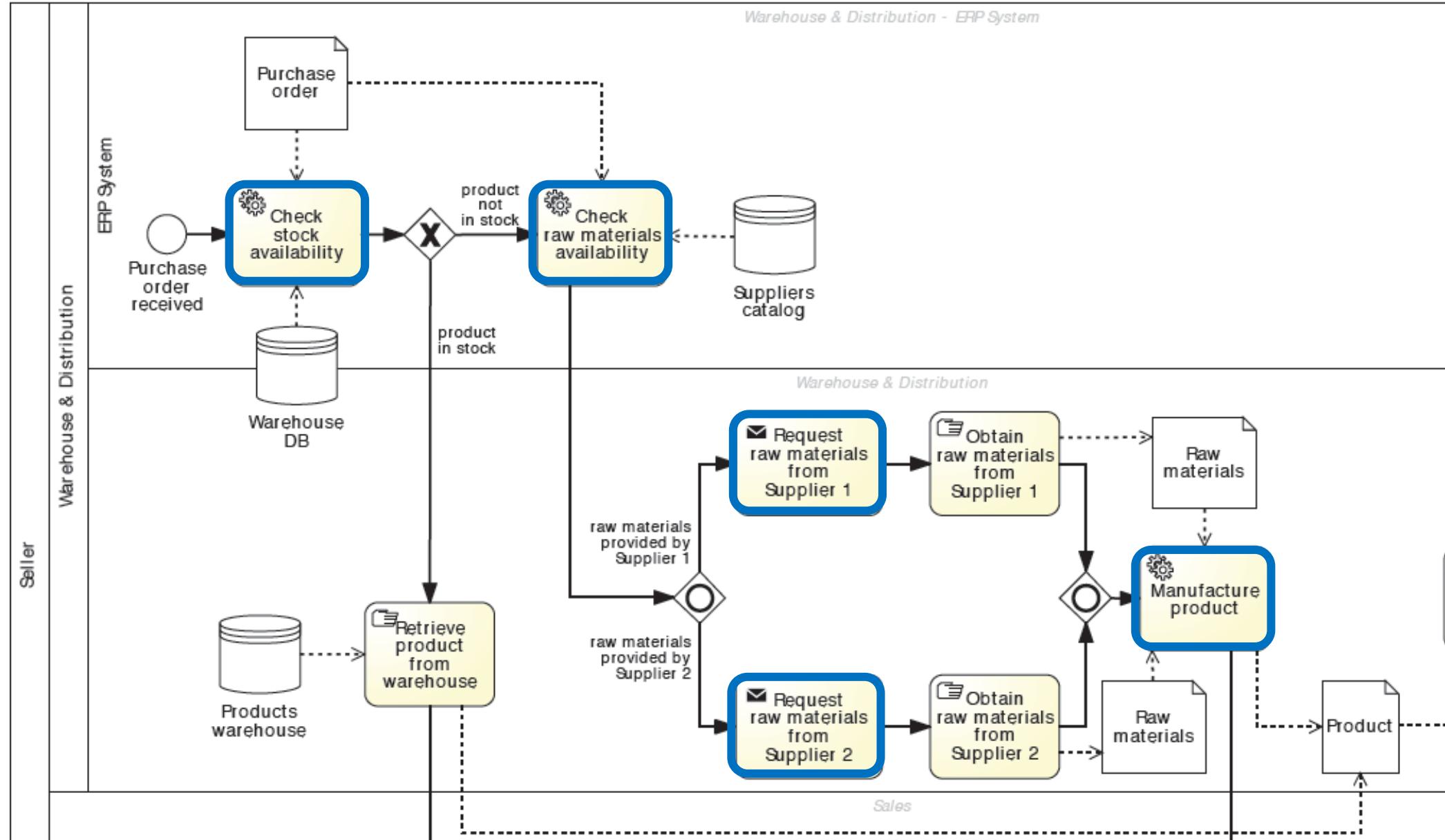
In our example...

█ automated
█ user
█ manual

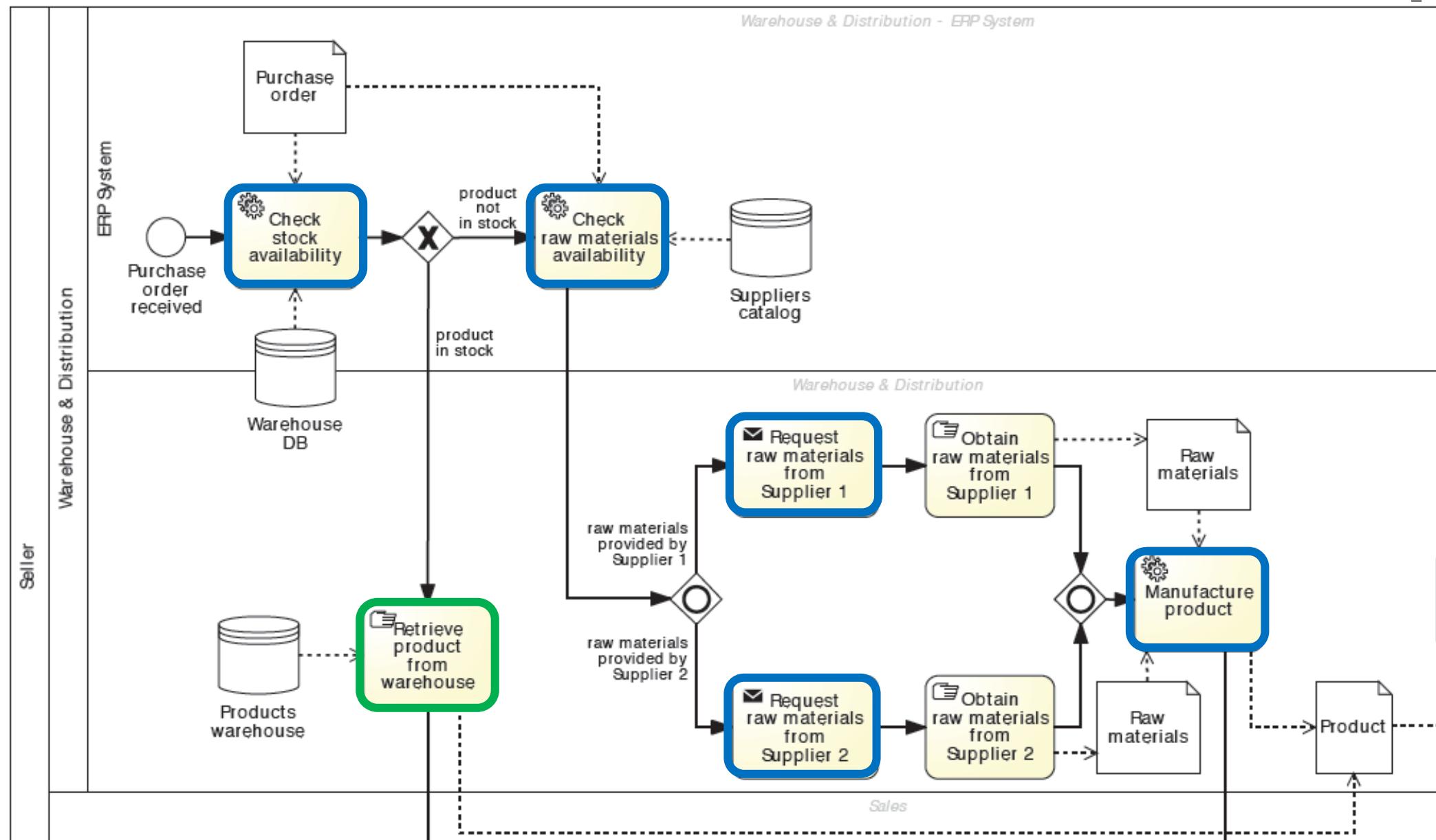


In our example...

█ automated
█ user
█ manual

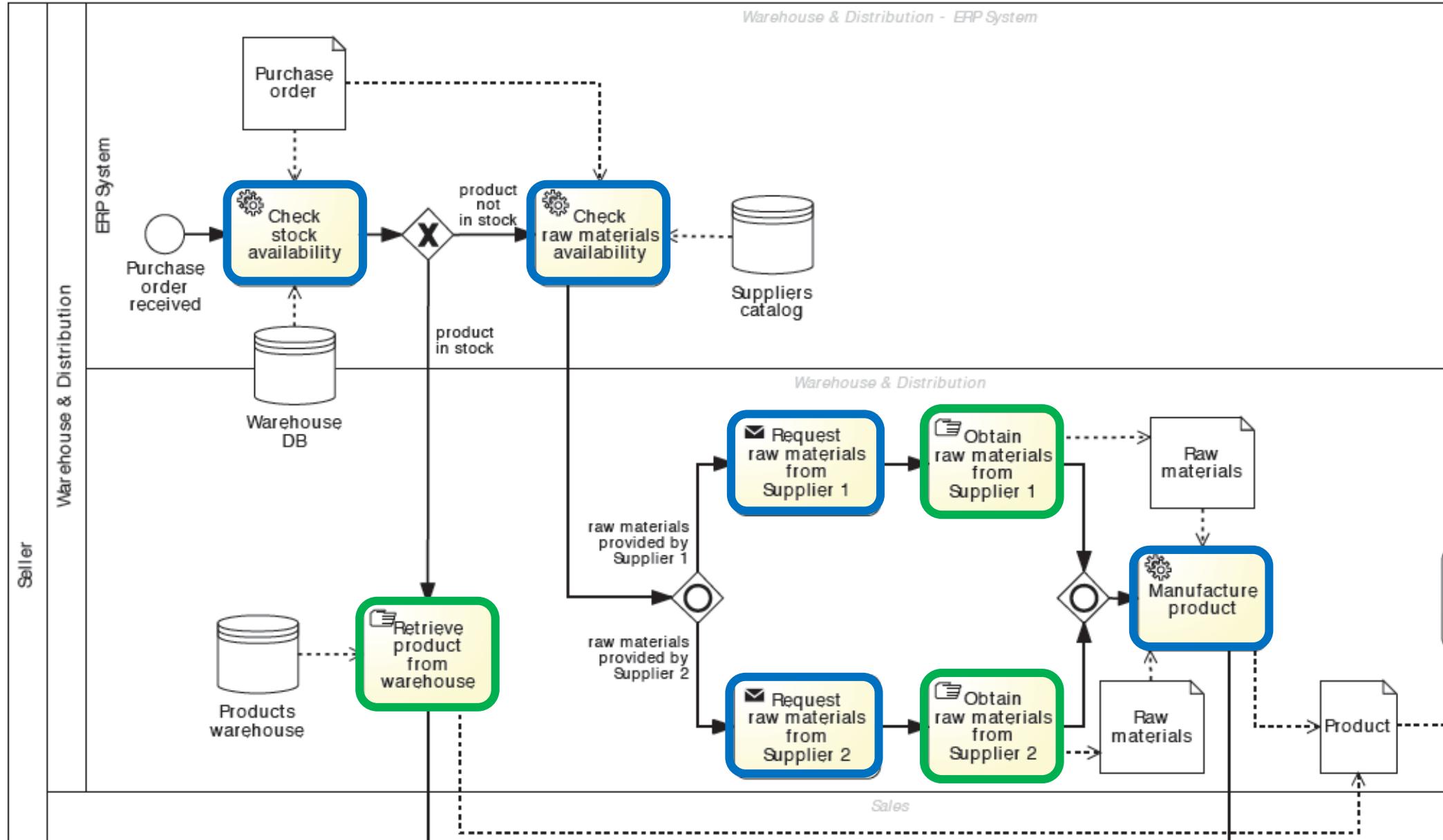


In our example...



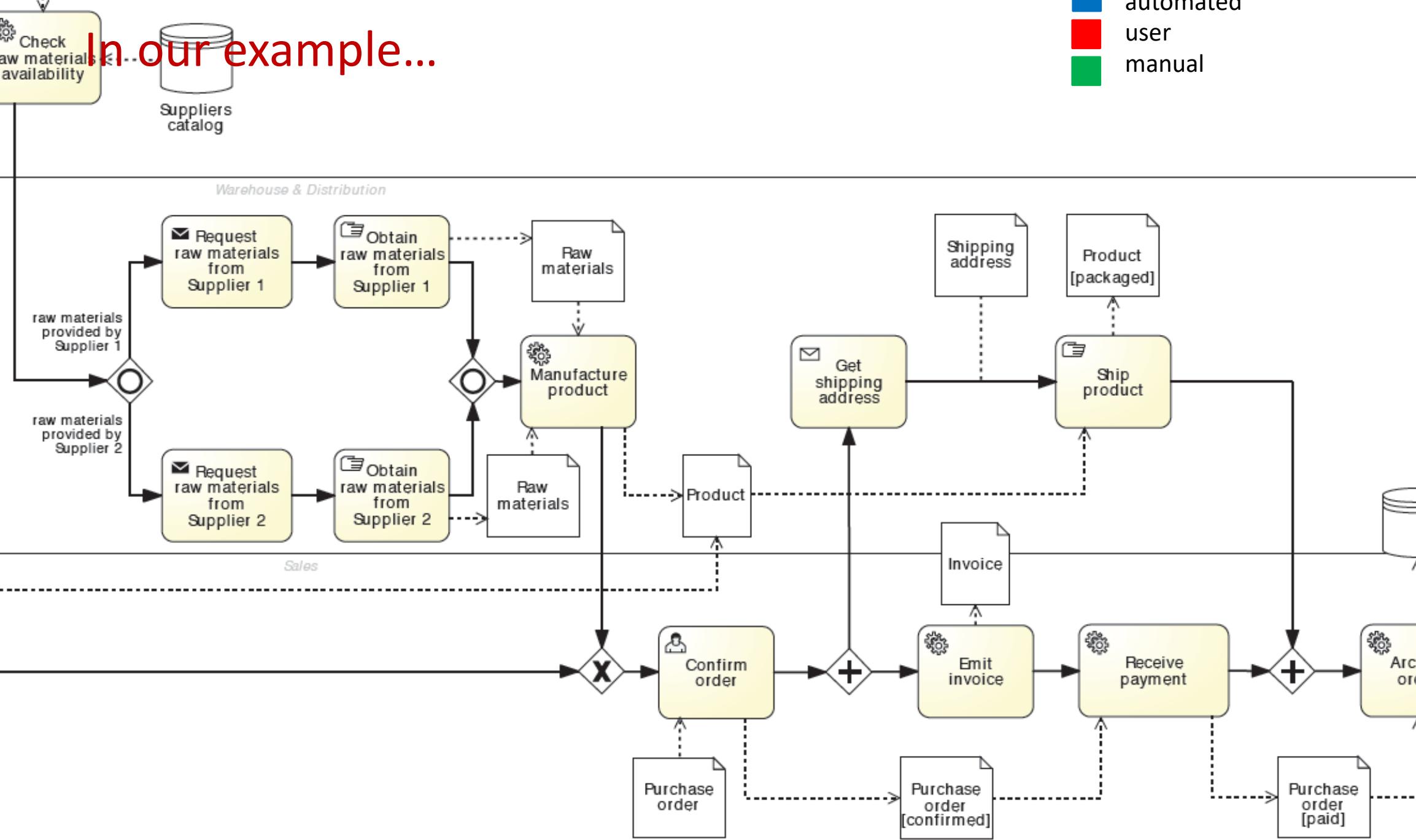
In our example...

█ automated
█ user
█ manual



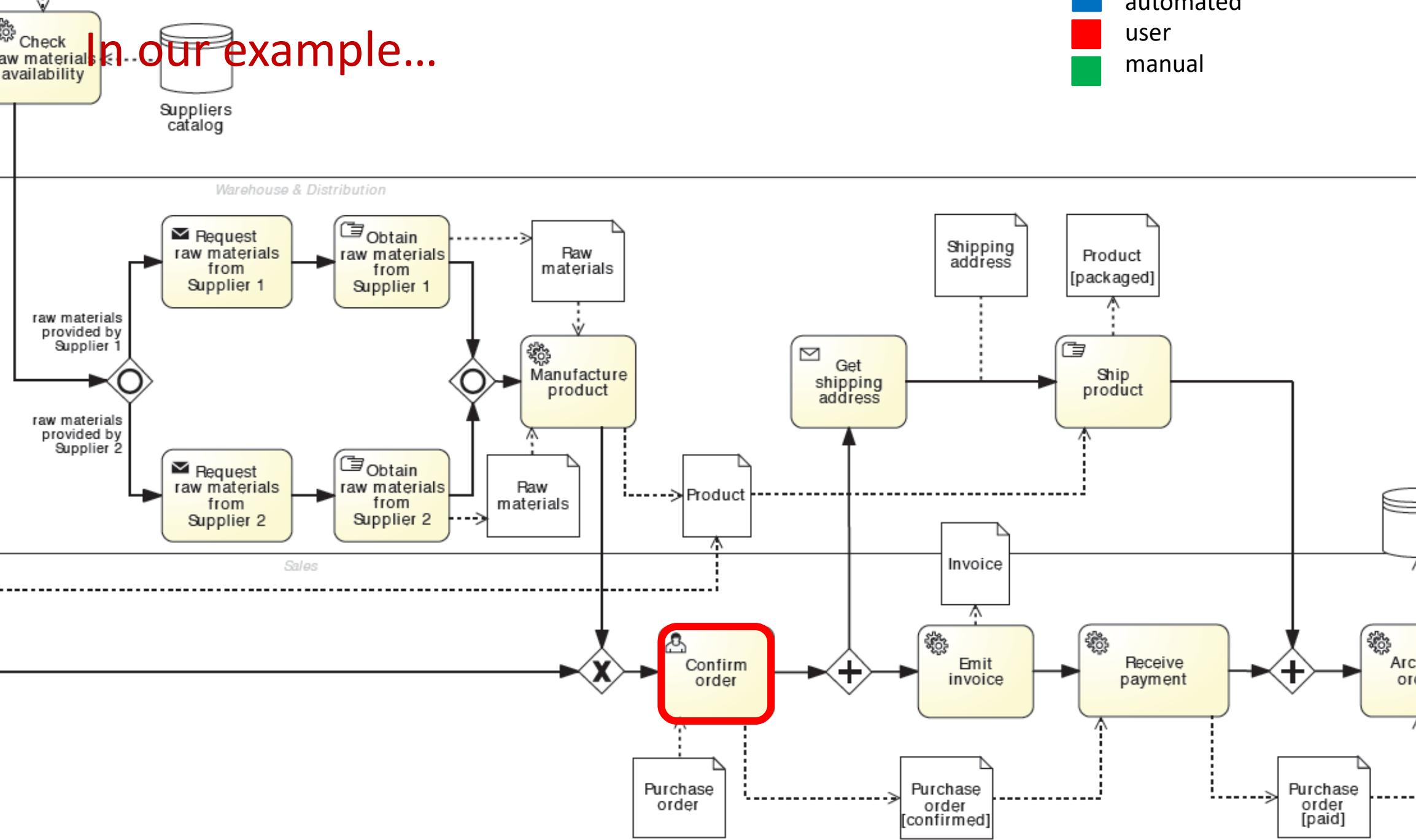
 automated
 user
 manual

In our example...



 automated
 user
 manual

In our example...



2. Review manual tasks

Principle: if it can't be seen by the BPMS, it doesn't exist.

-> Find ways to support manual tasks via IT:

2. Review manual tasks

Principle: if it can't be seen by the BPMS, it doesn't exist.

-> Find ways to support manual tasks via IT:

- via user task
- via automated task

2. Review manual tasks

Principle: if it can't be seen by the BPMS, it doesn't exist.

-> Find ways to support manual tasks via IT:

- via user task
- via automated task



 Retrieve
product
from
warehouse

2. Review manual tasks

Principle: if it can't be seen by the BPMS, it doesn't exist.

-> Find ways to support manual tasks via IT:

- via user task
- via automated task



2. Review manual tasks

Principle: if it can't be seen by the BPMS, it doesn't exist.

-> Find ways to support manual tasks via IT:

- via user task
- via automated task

2. Review manual tasks

Principle: if it can't be seen by the BPMS, it doesn't exist.

-> Find ways to support manual tasks via IT:

- via user task
- via automated task

-> Leave them out of the automated process



3. Complete the process model

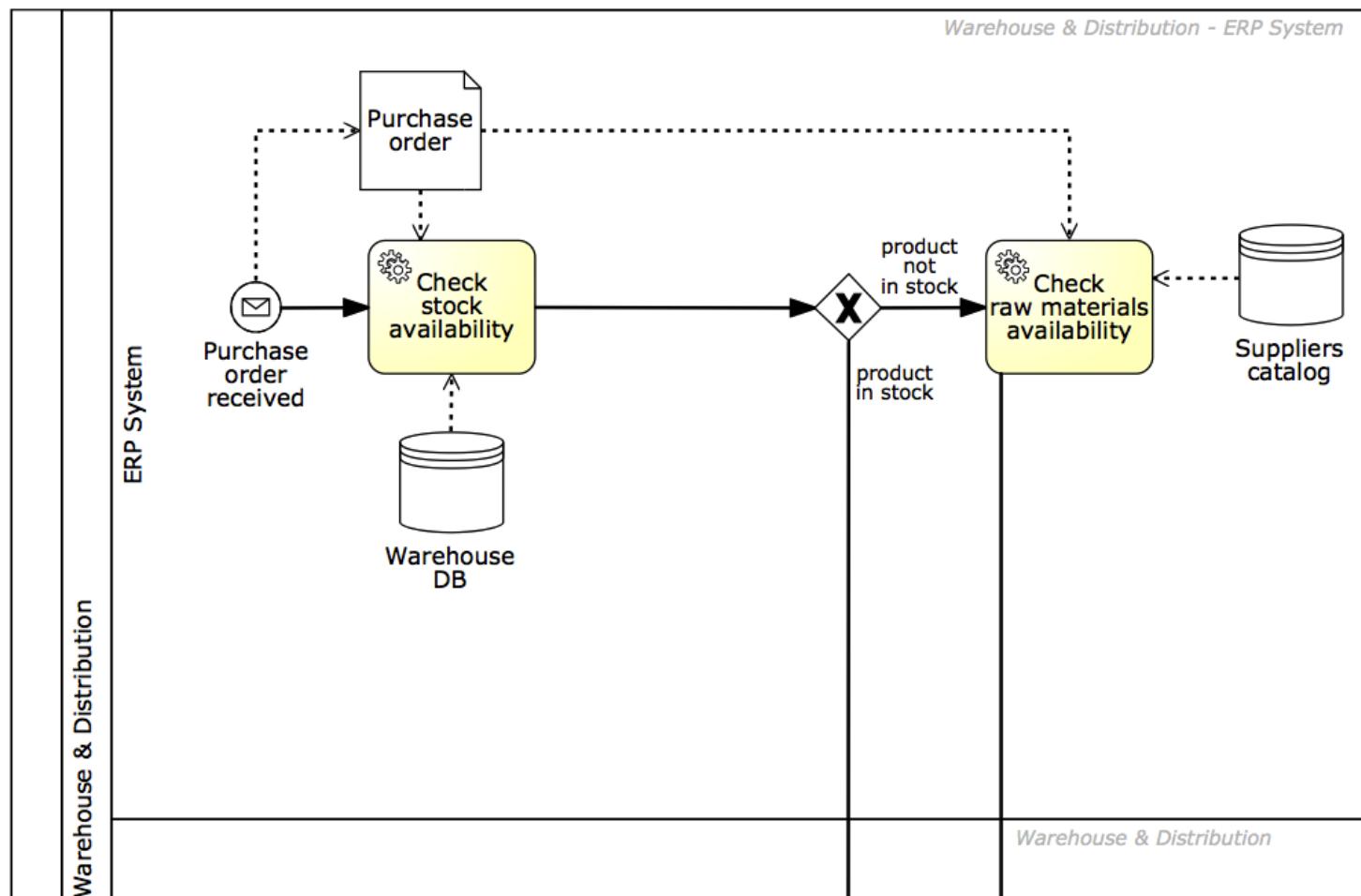
Principle: exceptions are the rule.

- Consider incomplete paths
- Rules of thumb
 - If we send something to another party, what happens if they do not respond? What happens if the response comes late? What happens if they do not respond the way we expect?
 - For each task: Can it go wrong and what happens if it goes wrong?
 - For each external party: Have we captured all messages or queries they might send us?

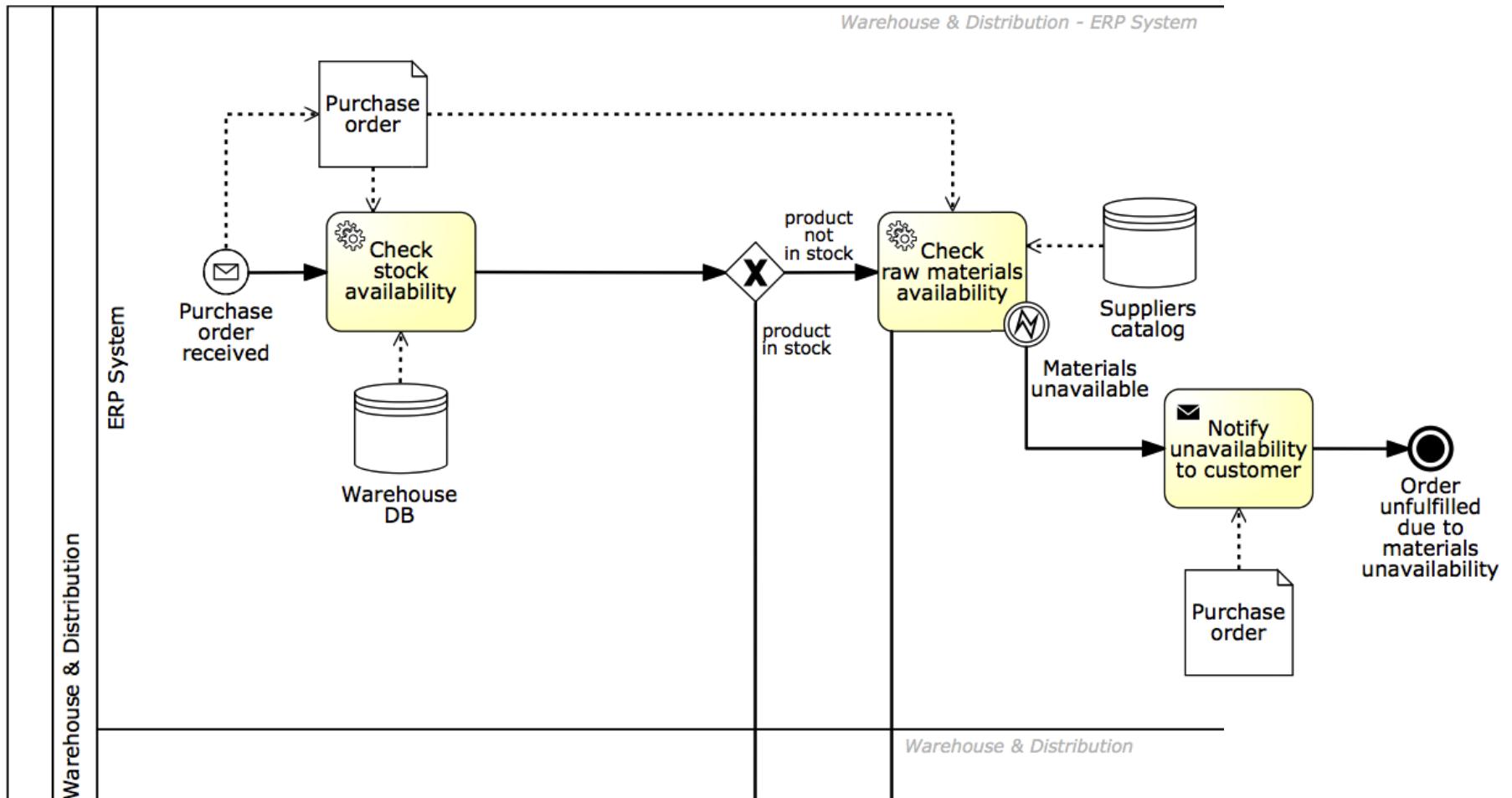
Principle: no data = no decisions, no tasks handover.

- Specify all (electronic) business objects
- For each task, determine which business objects it creates, reads, updates, delete (CRUD)
- For each decision, determine which objects it needs

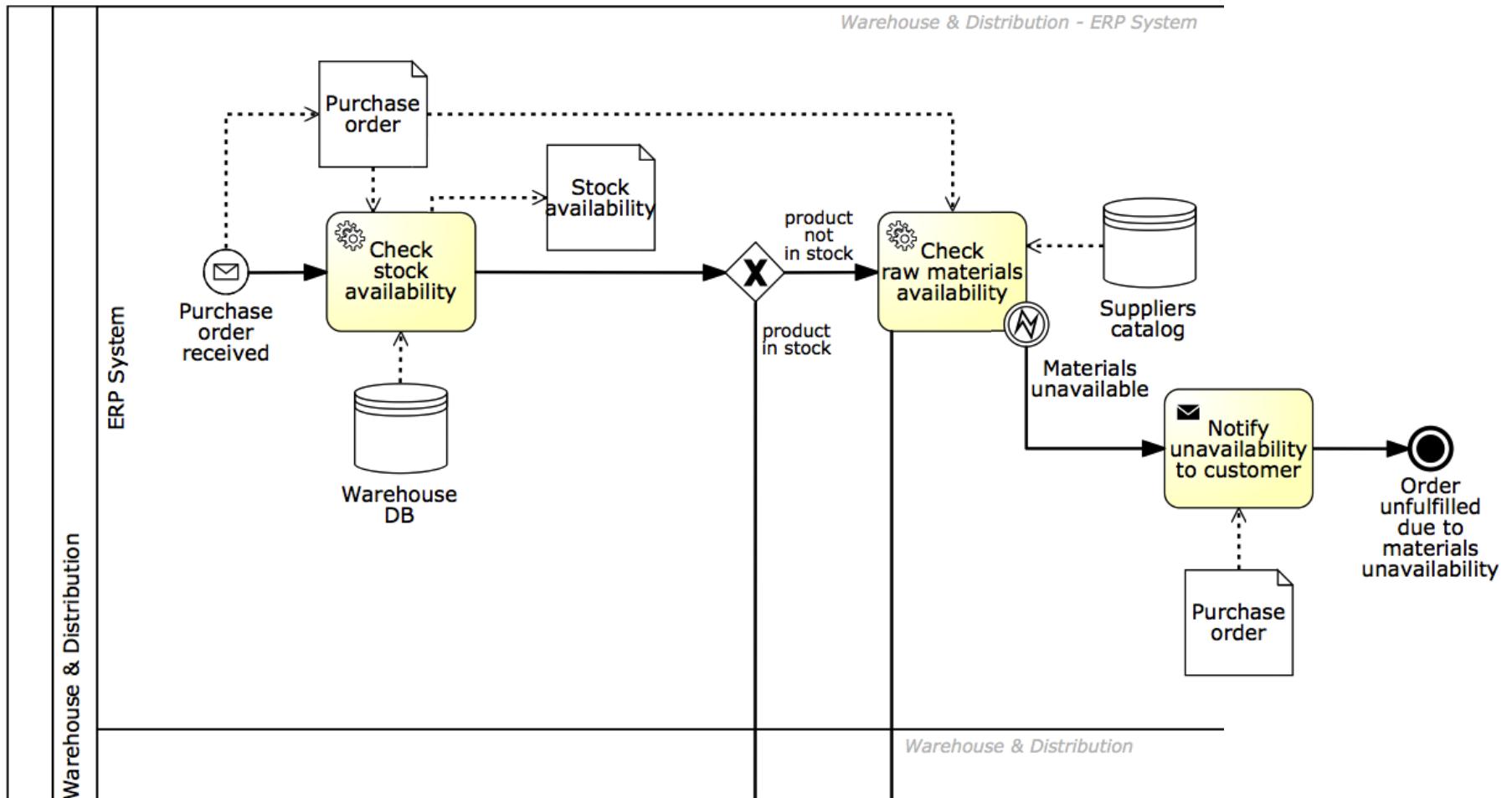
In our example...



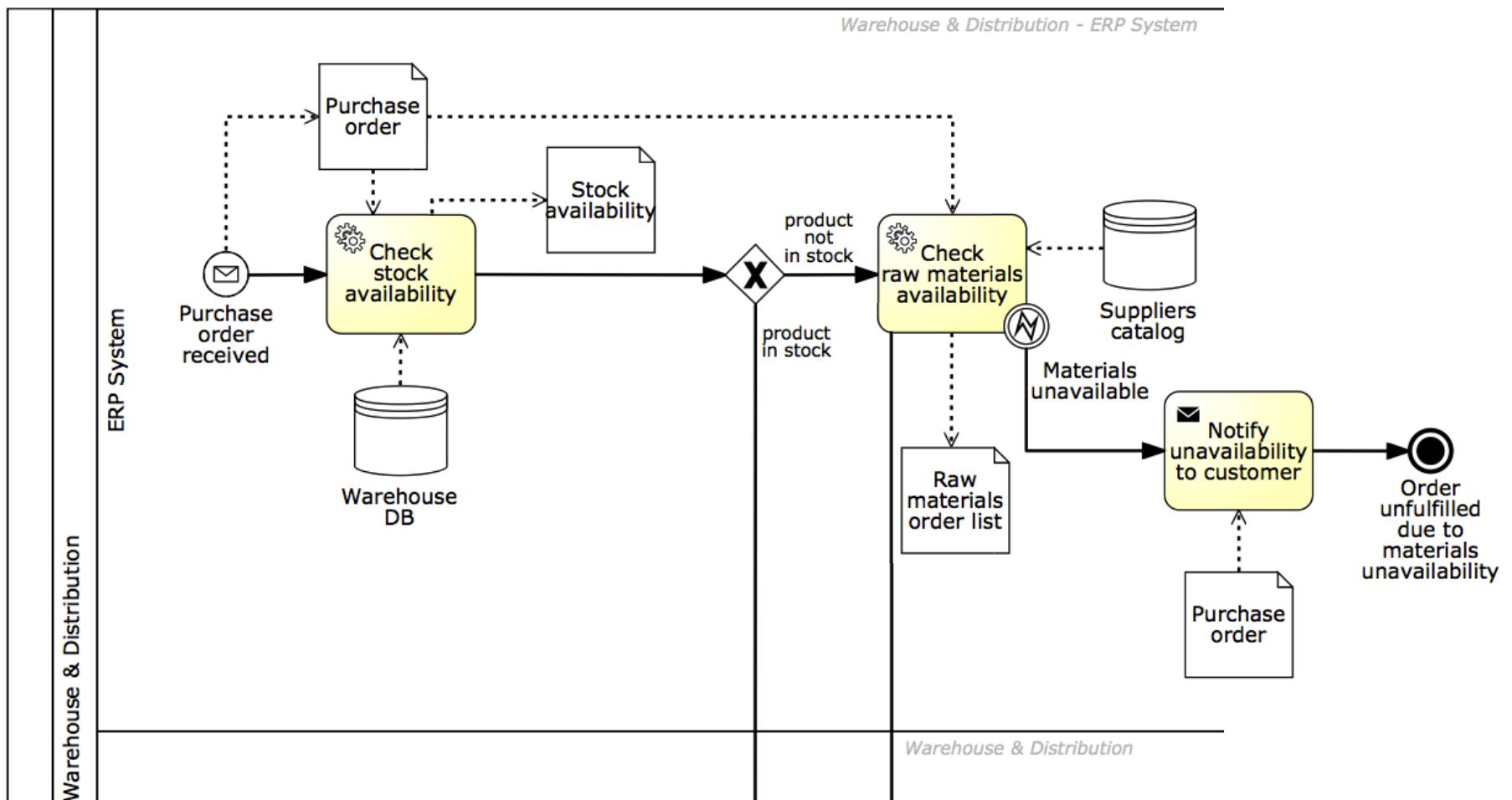
In our example...



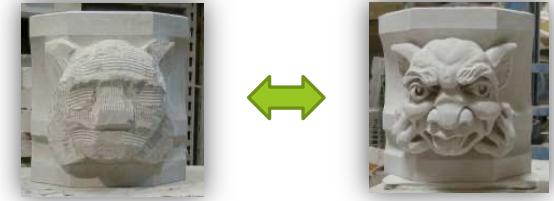
In our example...



In our example...

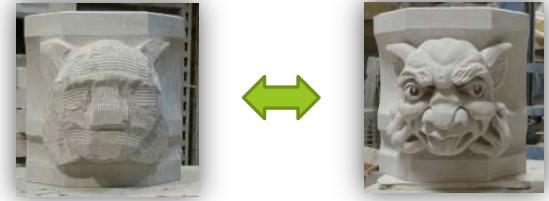


4. Adjust task granularity



The goal is to coordinate handovers of work between resources, therefore:

4. Adjust task granularity



The goal is to coordinate handovers of work between resources, therefore:

- > Aggregate consecutive tasks assigned to the same performer

4. Adjust task granularity



The goal is to coordinate handovers of work between resources, therefore:

- > Aggregate consecutive tasks assigned to the same performer

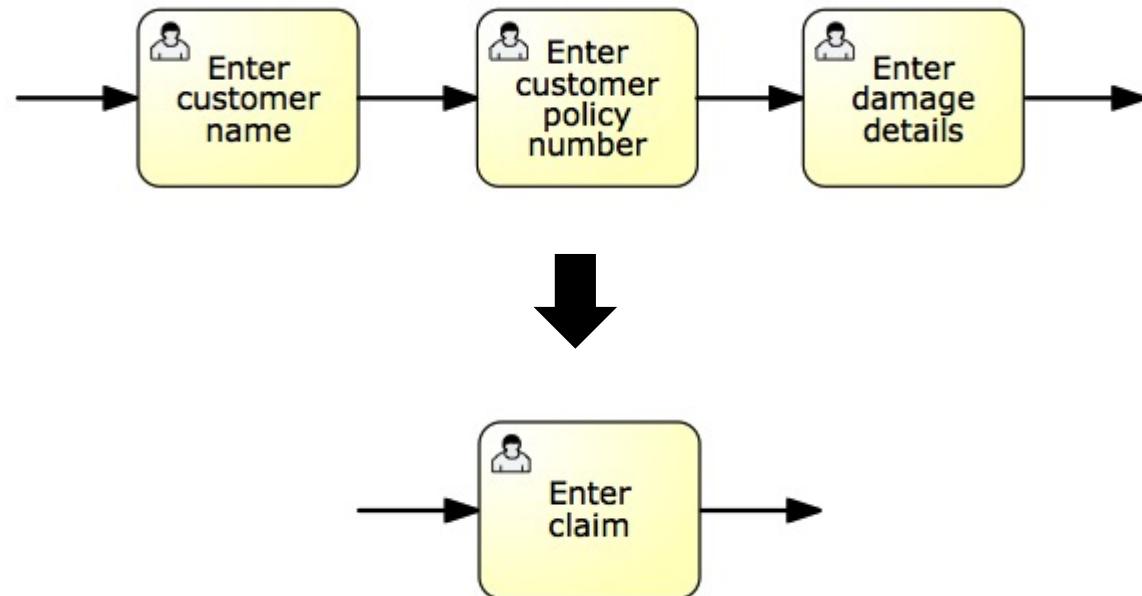


4. Adjust task granularity



The goal is to coordinate handovers of work between resources, therefore:

- > Aggregate consecutive tasks assigned to the same performer



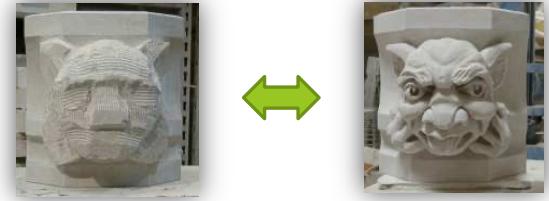
4. Adjust task granularity



The goal is to coordinate handovers of work between resources, therefore:

- > Aggregate consecutive tasks assigned to the same performer
- > Split tasks if they require different performers

4. Adjust task granularity



The goal is to coordinate handovers of work between resources, therefore:

- > Aggregate consecutive tasks assigned to the same performer
- > Split tasks if they require different performers

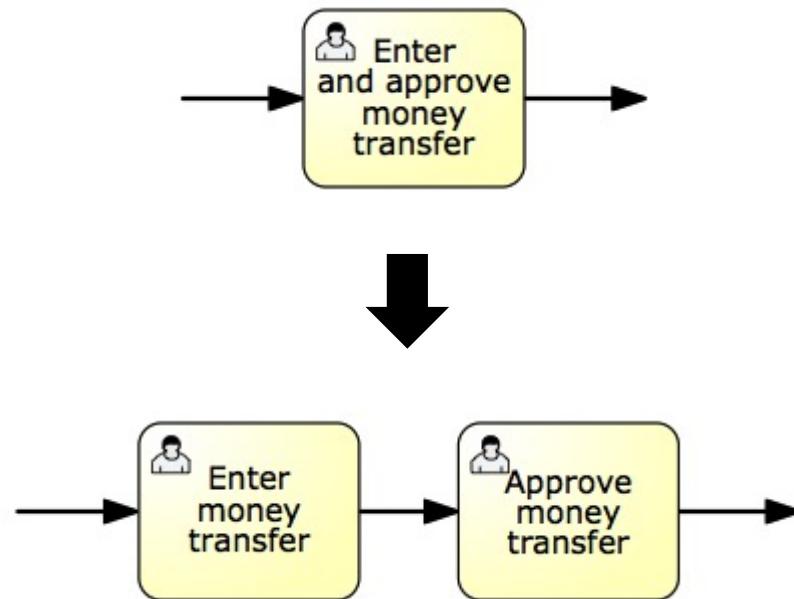


4. Adjust task granularity

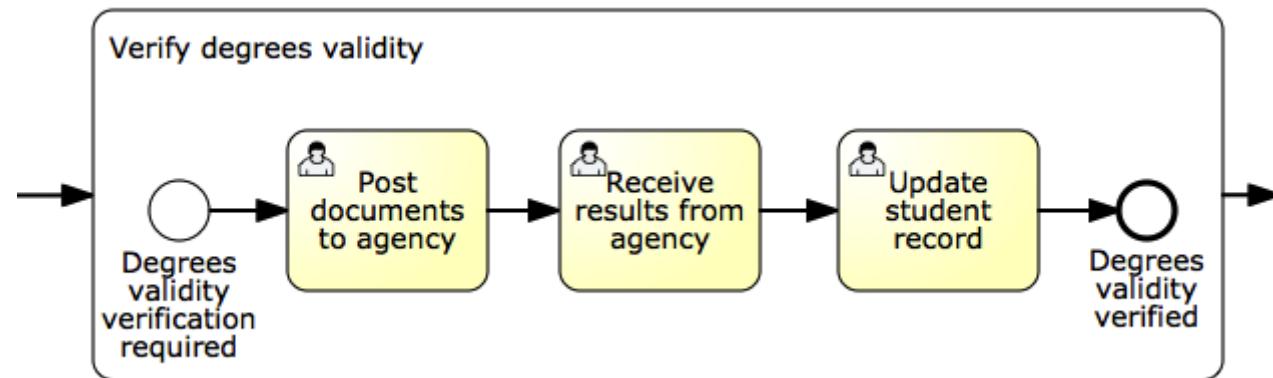


The goal is to coordinate handovers of work between resources, therefore:

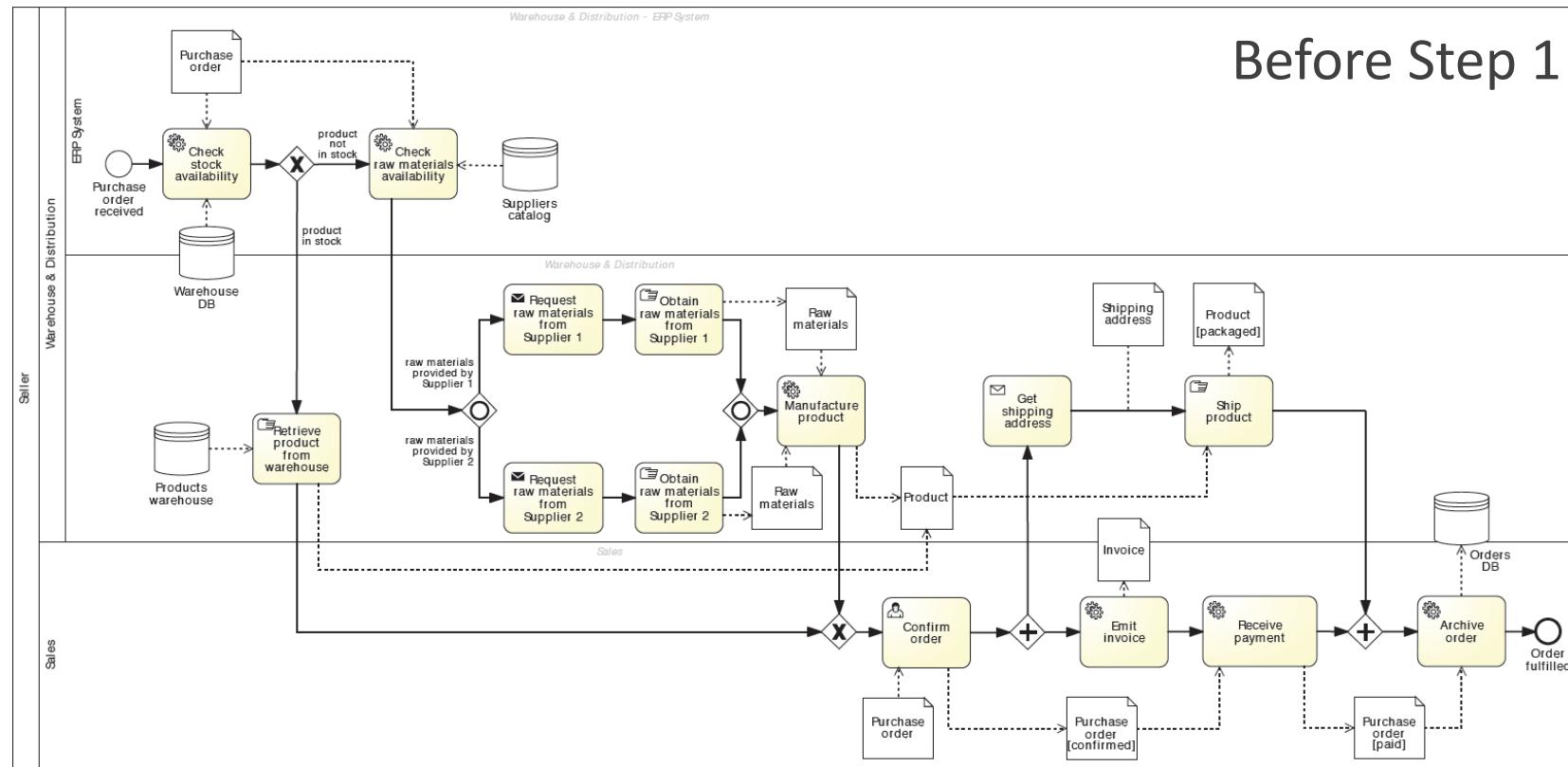
- > Aggregate consecutive tasks assigned to the same performer
- > Split tasks if they require different performers



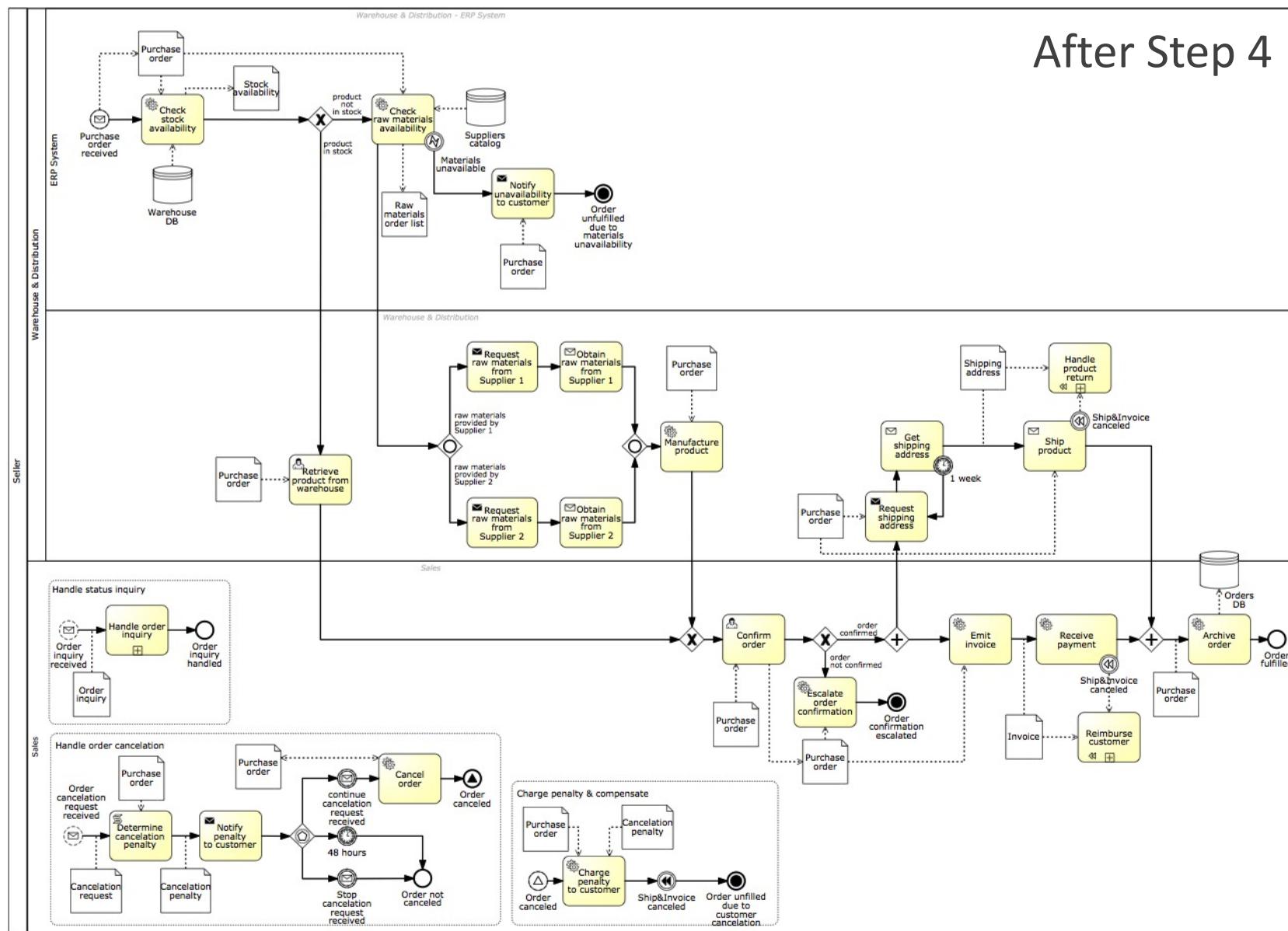
An exception to the rule



Our example...

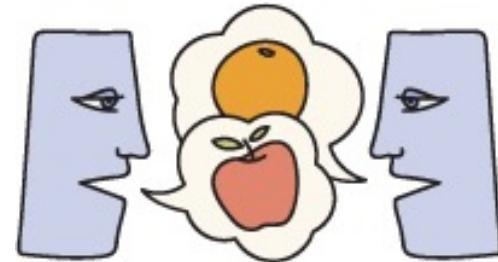


Our example...



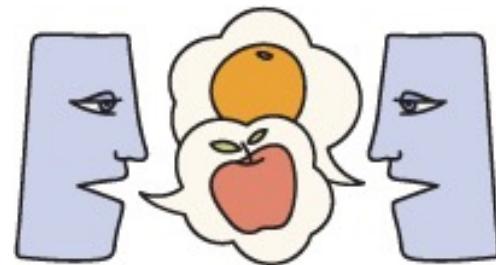
Bridging the gap: one task at a time

1. Identify the automation boundaries
2. Review manual tasks
3. Complete the process model
4. Adjust task granularity
5. Specify execution properties



Bridging the gap: one task at a time

1. Identify the automation boundaries
2. Review manual tasks
3. Complete the process model
4. Adjust task granularity
5. Specify execution properties

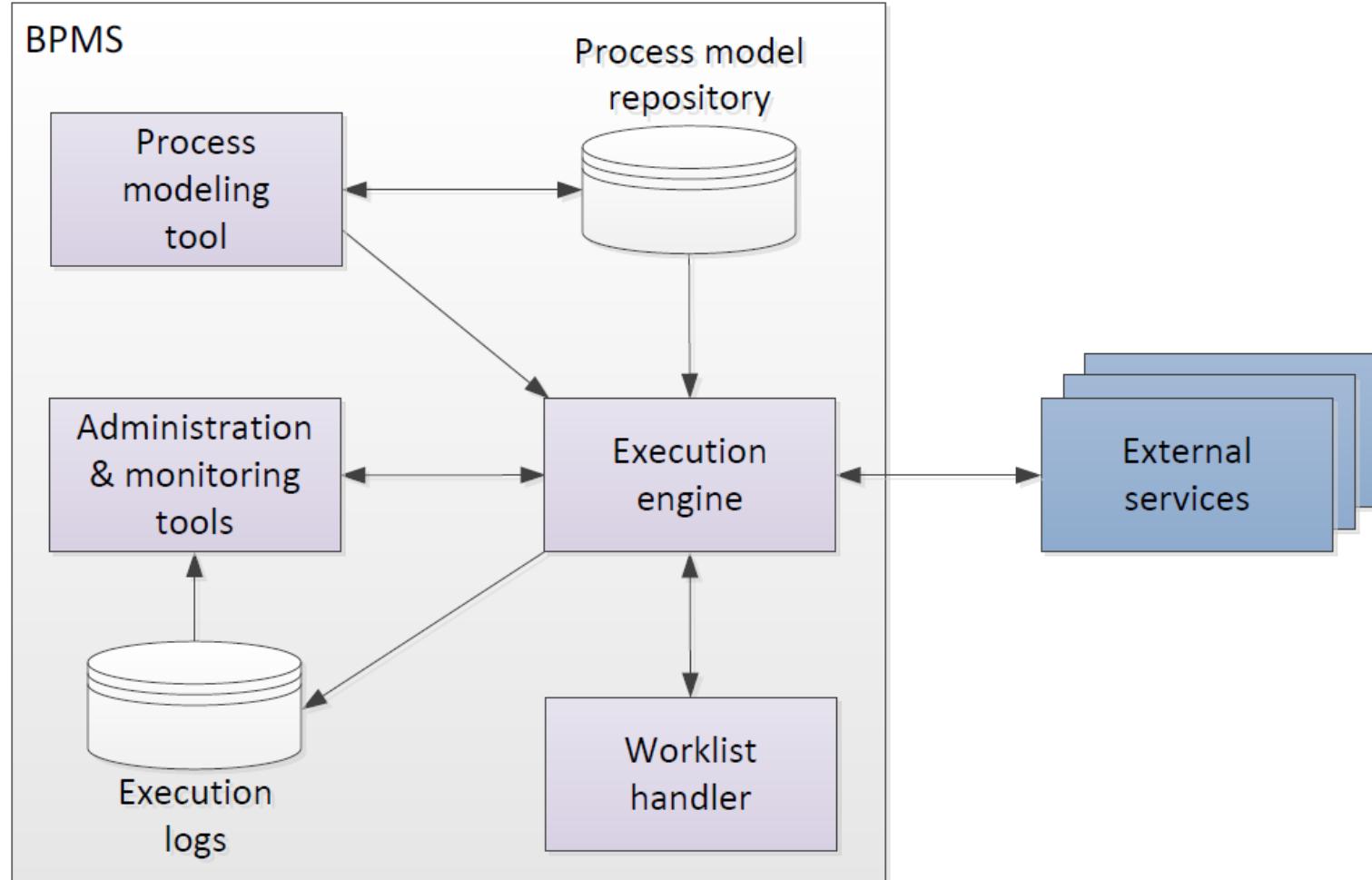


5. Specify execution properties

5. Specify execution properties

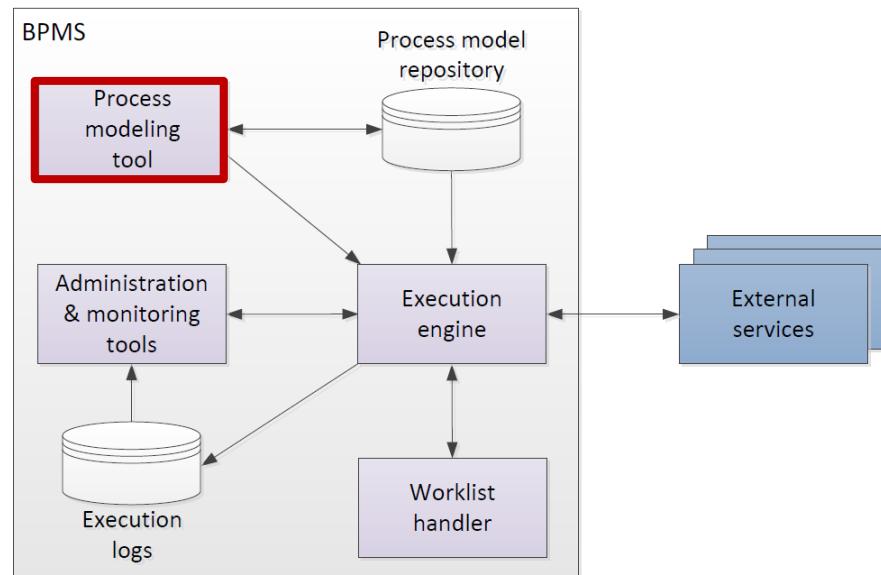
- Process variables, messages, signals, errors
- Task and event variables and their mappings to process variables
- Service details
- Code snippets
- Participant assignment rules
- User interface specification (e.g. forms)
- Task, event and sequence flow expressions
- BPMS-specific: work queues, forms, connectors...

Business Process Management System

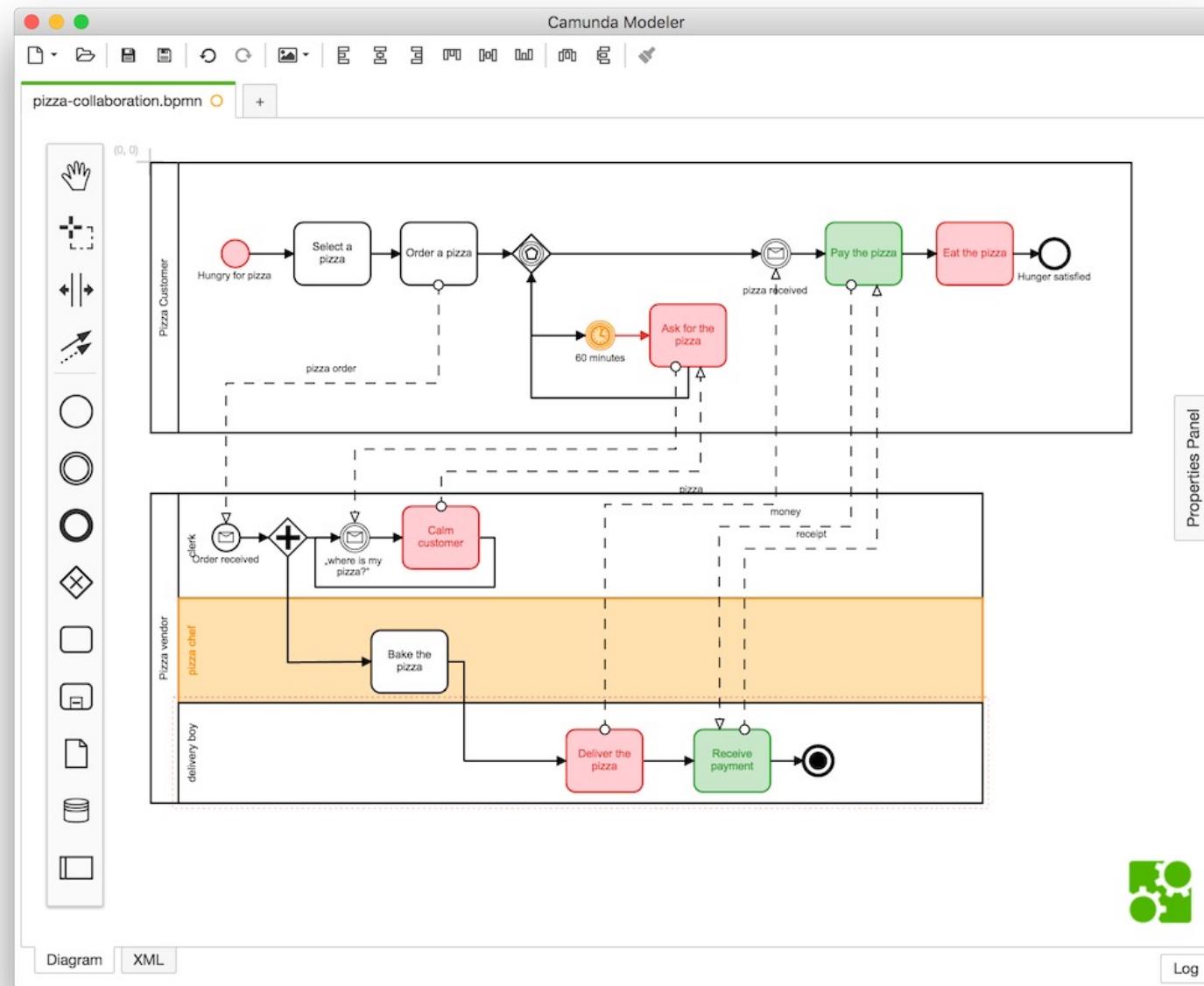


Process modeling tool

- To create and modify executable process models (by specifying execution properties)
- To store and retrieve automation solutions from a process model repository

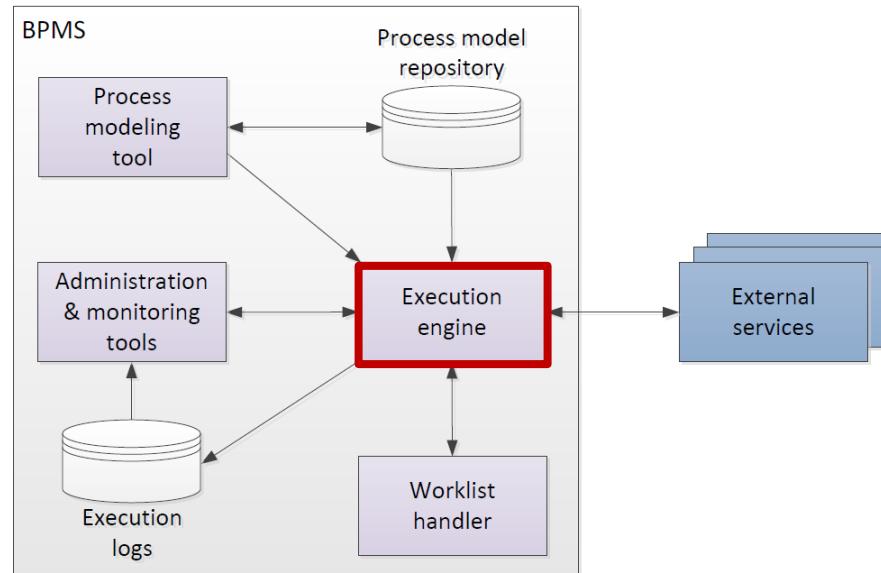


Example of a Process modeling tool



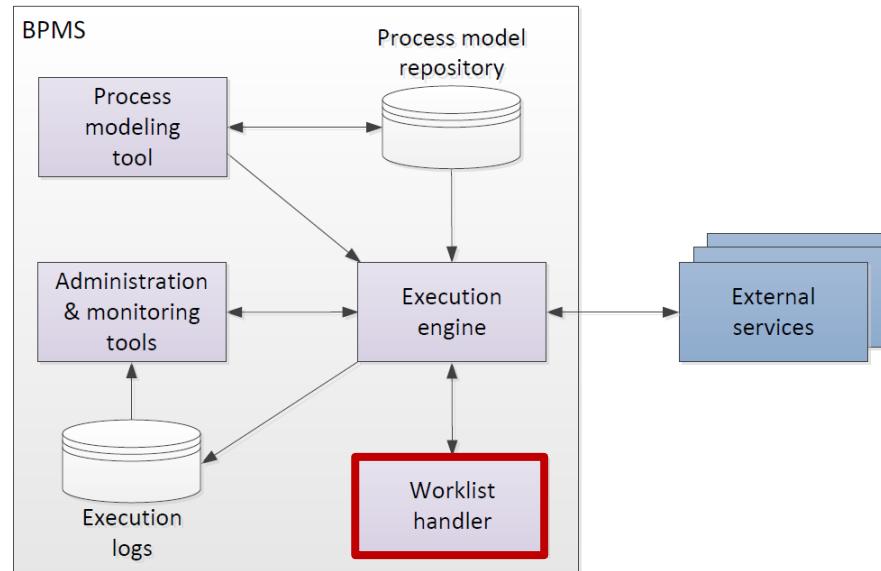
Execution Engine

- Instantiates executable process models (also called “cases”)
- Orchestrates distribution of work items to process participants and software services in order to execute a business process from start to end
- Logs execution data



Worklist Handler

- Imagine it as an “inbox”
- Offers work items to process participants and allows participants to commit to these work items
- Handles participants’ work queues and work item priorities
- May provide social network capabilities



Example of a Worklist Handler

The screenshot shows the Bonita Portal interface. On the left, a sidebar titled "To do" lists "My tasks", "Available tasks", "Hidden", "Done", and two sections under "Apps": "Help Desk" and "Travel Request". The main area is titled "Validate Travel Request". It displays a list of tasks:

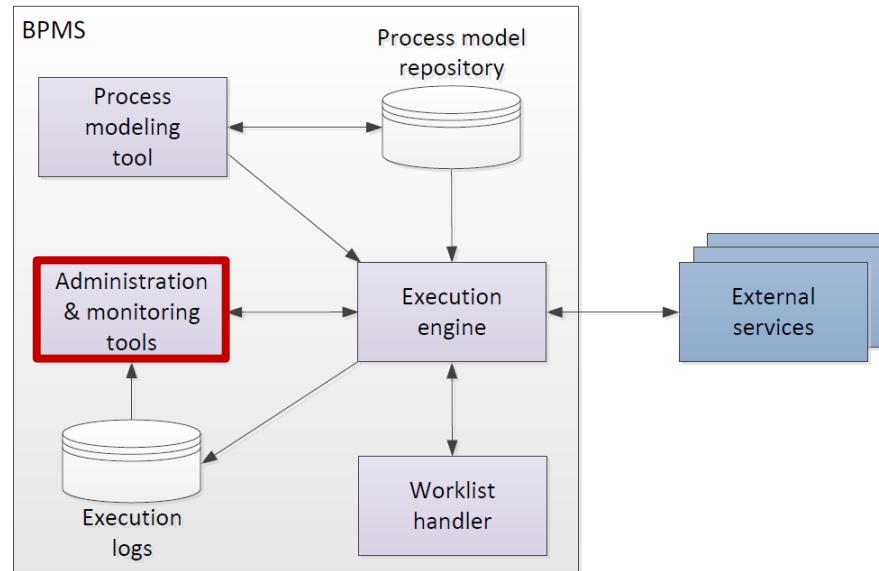
	Name	Due date
<input type="checkbox"/>	Resolve Case : Need a new ... Help Desk	In 36 min
<input type="checkbox"/>	Resolve Case : Provide VPN ... Help Desk	In 36 min
<input type="checkbox"/>	Resolve Case : Add member ... Help Desk	In 36 min
<input checked="" type="checkbox"/>	Validate Travel Request Travel Request	In 51 min

Below the tasks, there are sections for "Subtasks" (empty), "Done subtasks" (empty), "Technical details" (Ready since May 22, 2013), and "Comments" (a text input field labeled "Type new comment" and a "ADD COMMENT" button). A green "DO IT" button is prominently displayed at the top right of the main content area.

Bonita Soft Bonita Open Solution

Administration & Monitoring Tools

- To manage automation solutions
- To configure access to system components
- To monitor participants availability and performance of process cases



Example of an Administration & Monitoring Tool

Lombardi | Process Portal

Welcome, kolban | Help | Preferences | Log Off IBM.

My Tasks

- Inbox
- History
- Help Requests
- Alerts
- All Task

My ScoreBoards

- My Performance
- My Team Performance
- Process Performance
- SLA Overview
- Ad-Hoc Reports

PM Process Analysis : My Performance

Save | Manage Exposed Process Values | Export

Team Filter: All the Teams I belong to | Process Filter: All

On Track, At Risk, and Overdue Task Percentages.

Task Status

100.00% overdue

When will Tasks go Overdue

No Data

Axis Label

9/23 9/24 9/25 9/26 9/27 9/28 9/29 9/30

Click Pie Slice to see the tasks in that status. Or click All. << < > >> You can scroll left and right with the arrows below the chart

Showing all tasks for AdHoc1 process, Alert BPD process, Fulfillment Process process, P1 process, Quik Success Process process, Service Request process

Status	Id	Subject	Priority	Due Date	Current User	Team	Run
Overdue	67	Task: Submit Request	Normal	09/14/2010		All Users	▶
Overdue	120	Task: Escalation	Normal	09/14/2010		All Users	▶
Overdue	285	Step: Untitled1	Normal	09/21/2010		All Users	▶

Your Performance

Performance over the last 7 days vs average performance of the other members.

This report shows your performance trend for the last 14 days.

Your Performance vs Team Average

Task Count

Completed Active

kolban Team Average

Your Performance Trend

Task Count

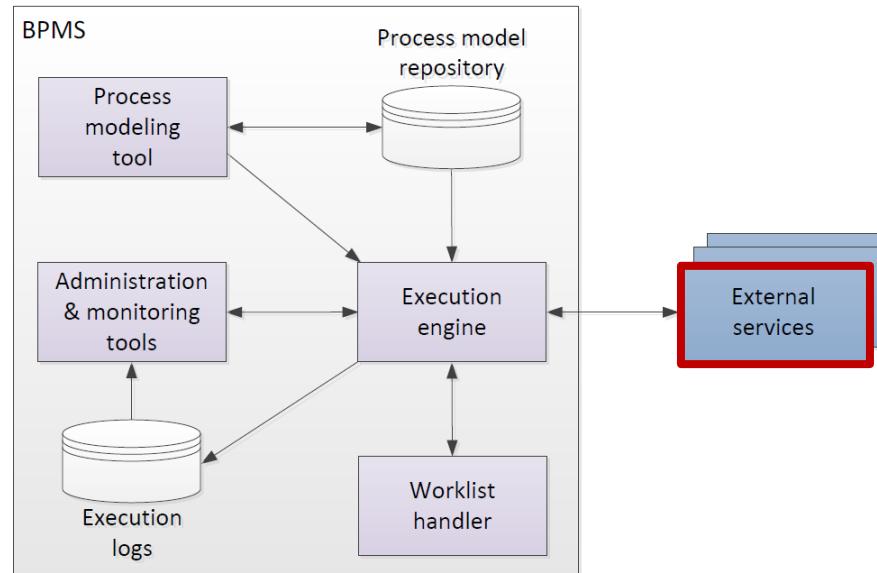
Completed Active

9/8 9/9 9/10 9/11 9/12 9/13 9/14 9/15 9/16 9/17 9/18 9/19 9/20 9/21 9/22

IBM BPM Process Portal

External Services

- Expose a service interface with which the engine can interact
- The engine provides the invoked service with the necessary data it will need to perform the activity for a specific case
- Examples: rules engine, email or Twitter notification, DB connector, CRM connector...



BPMS Landscape

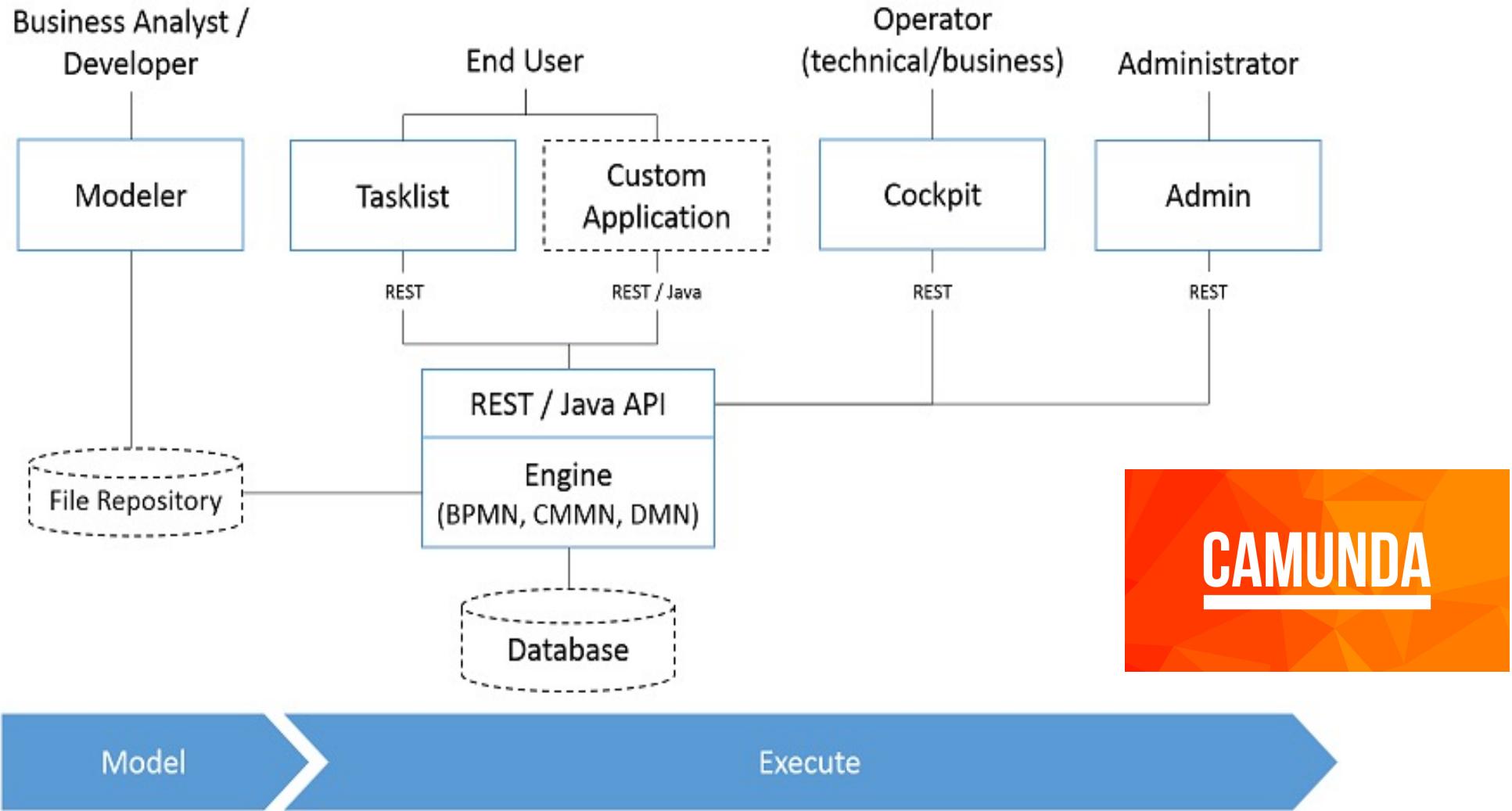
Closed source

- BIC Platform
- BizAgi
- IBM BPM
- OpenText BPM
- Oracle BPMS
- Pega Platform
- Software AG webMethods

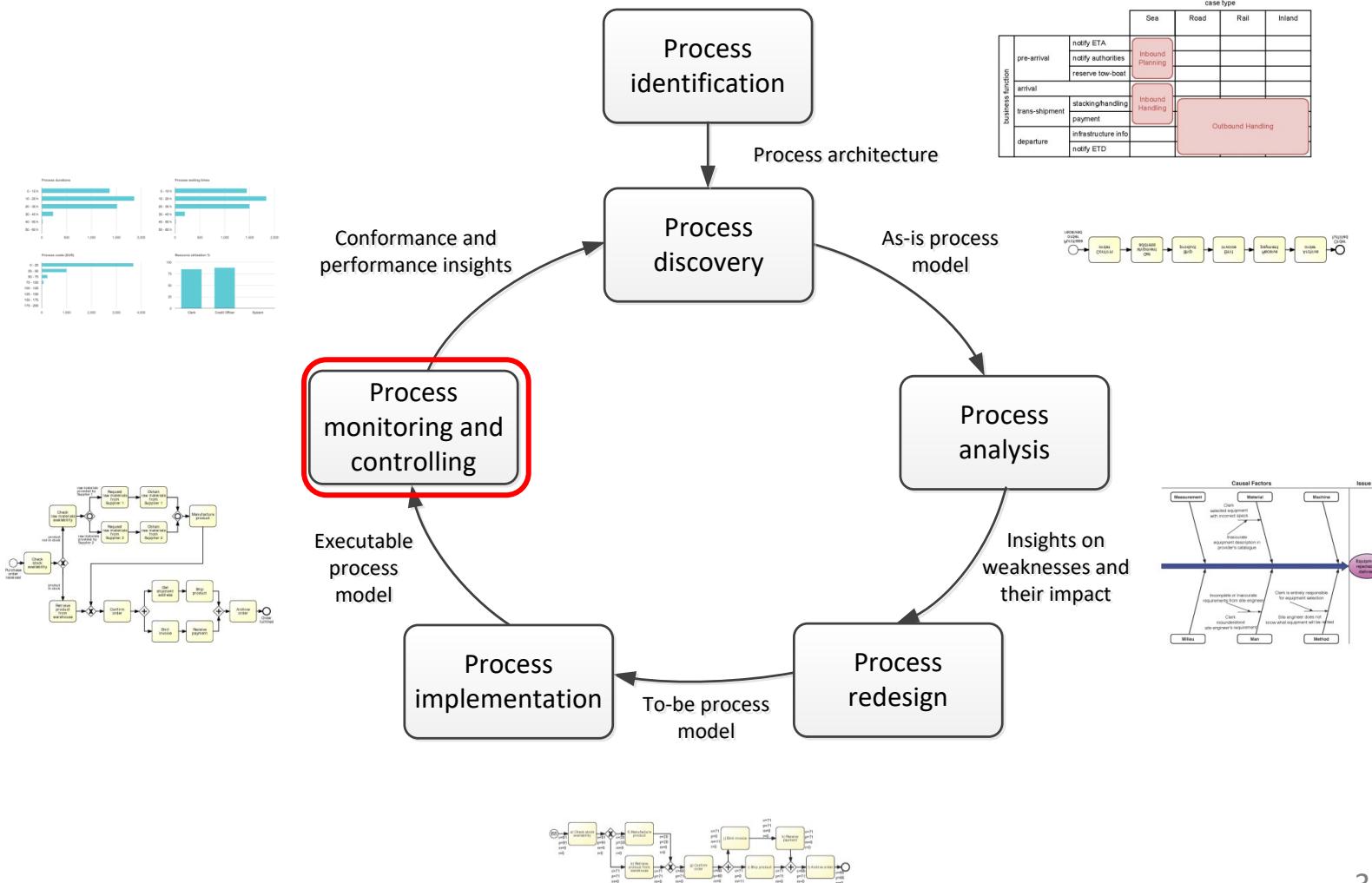
Commercial open-source

- Bonita Platform
- Camunda
- Flowable
- JBoss jBPM
- ProcessMaker

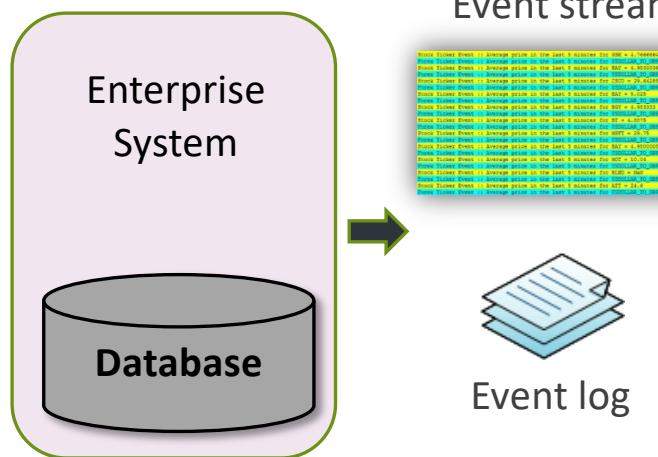
Camunda Architecture



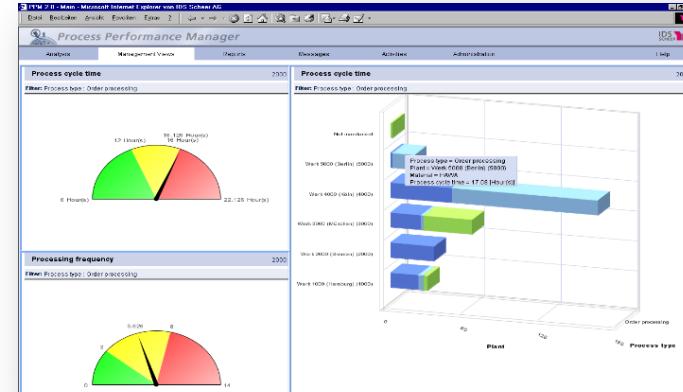
Where are we?



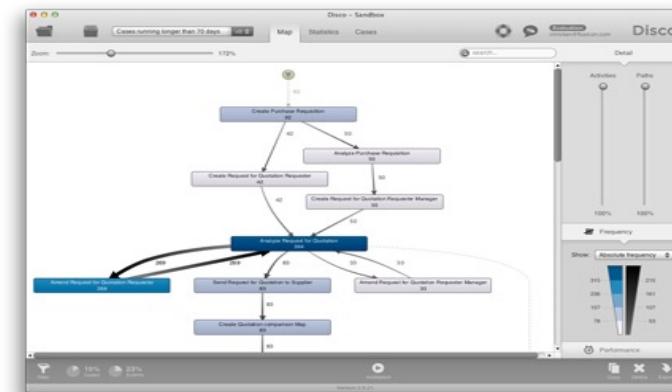
Business Process Monitoring



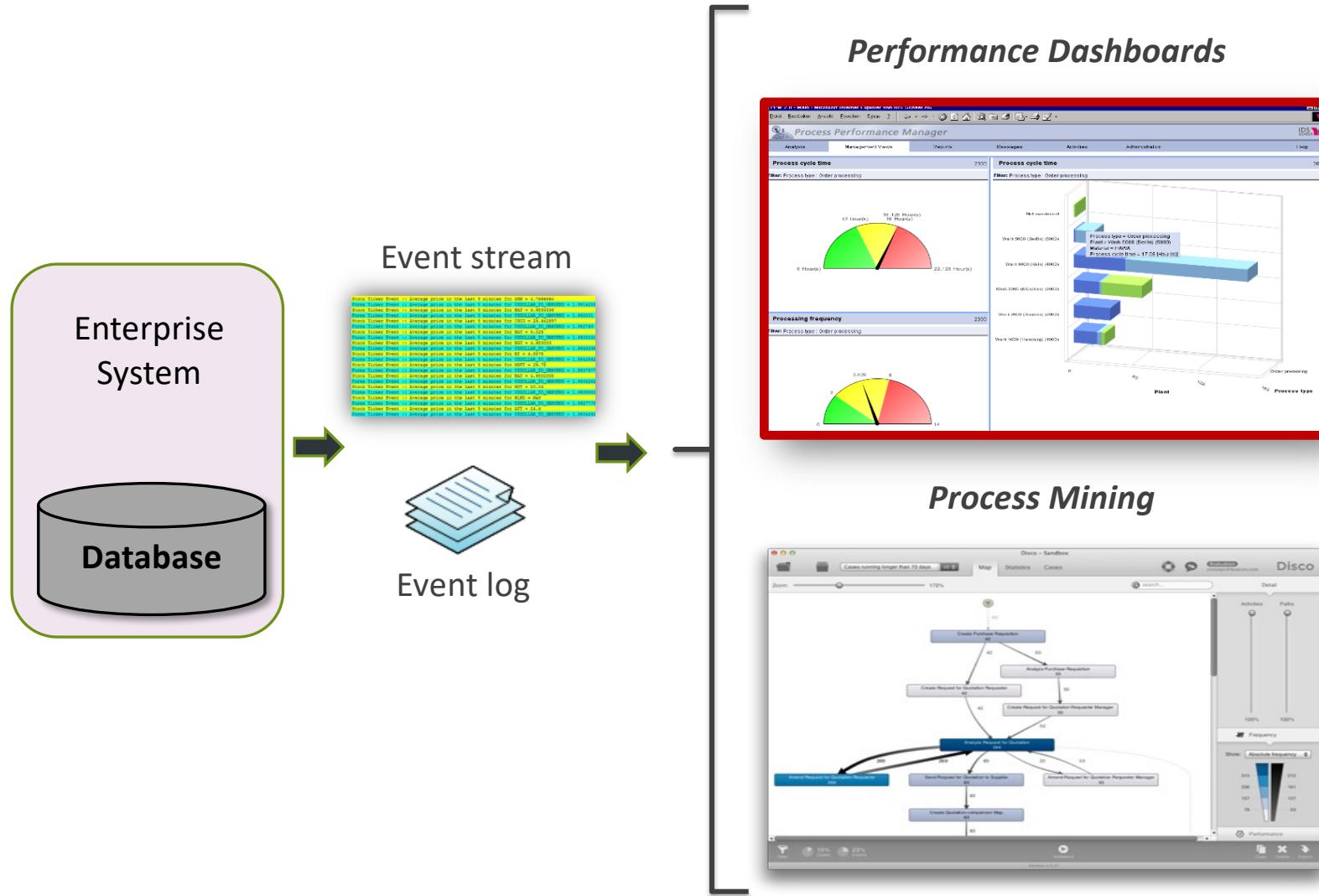
Performance Dashboards



Process Mining



Business Process Monitoring



Structure of a Business Process Event Log

Case ID	Timestamp	Activity	Resource	Loan goal	Requested amt	Offered amt
C001	18-10-2016	Check completeness	Sue	Mortgage	100 000	-
C001	19-10-2016	Check credit history	Sue	Mortgage	100 000	-
C001	19-10-2016	Calculate risk score	Bob	Mortgage	100 000	-
C001	20-10-2016	Make offer	Mike	Mortgage	100 000	70 000
C001	25-10-2016	Make offer	Mike	Mortgage	100 000	80 000
C002	20-10-2016	Check completeness	Sue	Car	15 000	-
C002	20-10-2016	Check credit history	Sue	Car	15 000	-
C002	22-10-2016	Calculate risk score	Elsa	Car	15 000	-
C002	24-10-2016	Reject application	Elsa	Car	15 000	-
C003	02-11-2016	Check completeness	Maria	Mortgage	30 000	-
C003	04-11-2016	Ask for additional data	Maria	Mortgage	30 000	-
C003	10-11-2016	Check credit history	Maria	Mortgage	30 000	-
...

Structure of a Business Process Event Log

Case ID	Timestamp	Activity	Resource	Loan goal	Requested amt	Offered amt
C001	18-10-2016	Check completeness	Sue	Mortgage	100 000	-
C001	19-10-2016	Check credit history	Sue	Mortgage	100 000	-
C001	19-10-2016	Calculate risk score	Bob	Mortgage	100 000	-
C001	20-10-2016	Make offer	Mike	Mortgage	100 000	70 000
C001	25-10-2016	Make offer	Mike	Mortgage	100 000	80 000
C002	20-10-2016	Check completeness	Sue	Car	15 000	-
C002	20-10-2016	Check credit history	Sue	Car	15 000	-
C002	22-10-2016	Calculate risk score	Elsa	Car	15 000	-
C002	24-10-2016	Reject application	Elsa	Car	15 000	-
C003	02-11-2016	Check completeness	Maria	Mortgage	30 000	-
C003	04-11-2016	Ask for additional data	Maria	Mortgage	30 000	-
C003	10-11-2016	Check credit history	Maria	Mortgage	30 000	-
...

Structure of a Business Process Event Log

Case ID	Timestamp	Activity	Resource	Loan goal	Requested amt	Offered amt
C001	18-10-2016	Check completeness	Sue	Mortgage	100 000	-
C001	19-10-2016	Check credit history	Sue	Mortgage	100 000	-
C001	19-10-2016	Calculate risk score	Bob	Mortgage	100 000	-
C001	20-10-2016	Make offer	Mike	Mortgage	100 000	70 000
C001	25-10-2016	Make offer	Mike	Mortgage	100 000	80 000
C002	20-10-2016	Check completeness	Sue	Car	15 000	-
C002	20-10-2016	Check credit history	Sue	Car	15 000	-
C002	22-10-2016	Calculate risk score	Elsa	Car	15 000	-
C002	24-10-2016	Reject application	Elsa	Car	15 000	-
C003	02-11-2016	Check completeness	Maria	Mortgage	30 000	-
C003	04-11-2016	Ask for additional data	Maria	Mortgage	30 000	-
C003	10-11-2016	Check credit history	Maria	Mortgage	30 000	-
...

Structure of a Business Process Event Log

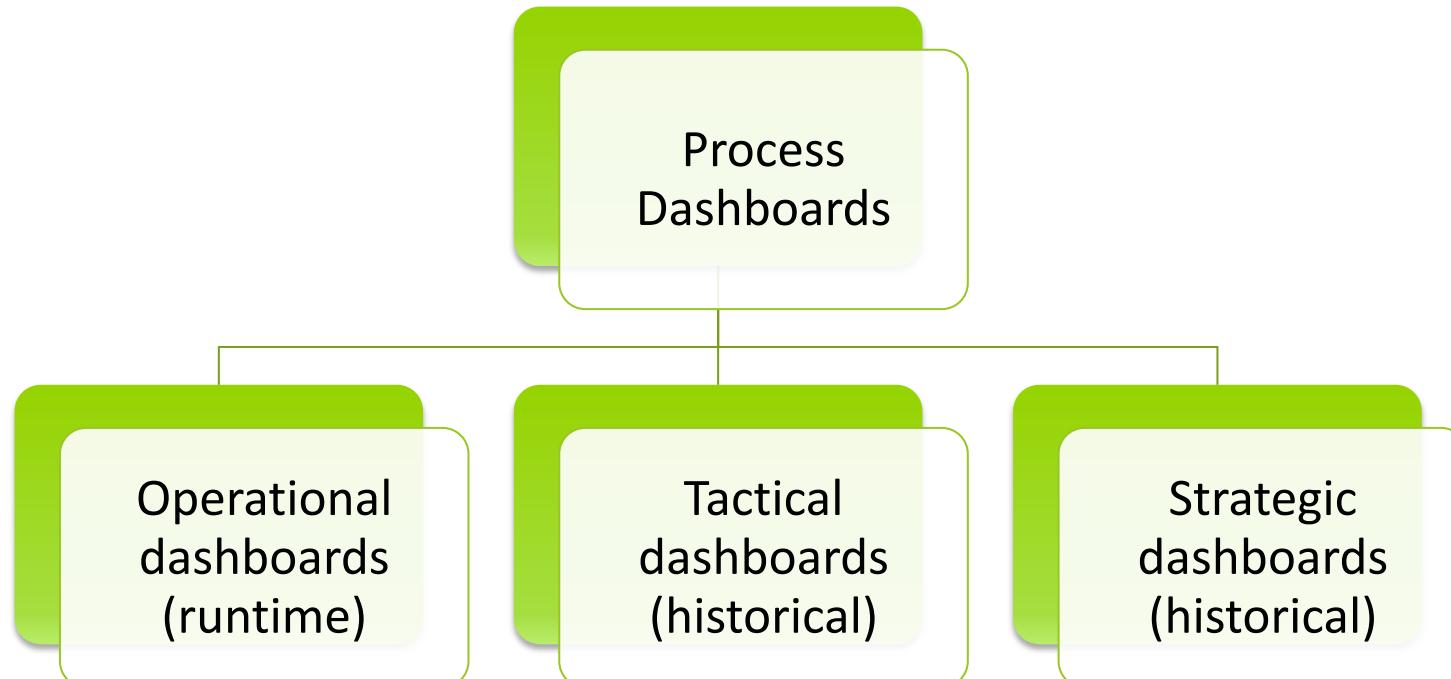
Case ID	Timestamp	Activity	Resource	Loan goal	Requested amt	Offered amt
C001	18-10-2016	Check completeness	Sue	Mortgage	100 000	-
C001	19-10-2016	Check credit history	Sue	Mortgage	100 000	-
C001	19-10-2016	Calculate risk score	Bob	Mortgage	100 000	-
C001	20-10-2016	Make offer	Mike	Mortgage	100 000	70 000
C001	25-10-2016	Make offer	Mike	Mortgage	100 000	80 000
C002	20-10-2016	Check completeness	Sue	Car	15 000	-
C002	20-10-2016	Check credit history	Sue	Car	15 000	-
C002	22-10-2016	Calculate risk score	Elsa	Car	15 000	-
C002	24-10-2016	Reject application	Elsa	Car	15 000	-
C003	02-11-2016	Check completeness	Maria	Mortgage	30 000	-
C003	04-11-2016	Ask for additional data	Maria	Mortgage	30 000	-
C003	10-11-2016	Check credit history	Maria	Mortgage	30 000	-
...

Process performance dashboard

A visualization of a collection of process performance measures to put into evidence information and patterns to help business users answer one or more business questions

Process performance dashboard

A visualization of a collection of process performance measures to put into evidence information and patterns to help business users answer one or more business questions



Operational process dashboards

- Aimed at process workers & operational managers
- Emphasis on monitoring (detect-and-respond), e.g.:
 - Work-in-progress
 - Problematic cases – e.g., overdue/at-risk cases
 - Resource load

Operational process dashboards

- Aimed at process workers & operational managers
- Emphasis on monitoring (detect-and-respond), e.g.:
 - Work-in-progress
 - Problematic cases – e.g., overdue/at-risk cases
 - Resource load



Tactical performance dashboards

- Aimed at process owners / managers
- Emphasis on analysis and management
 - E.g., detecting bottlenecks
- Typical process performance indicators
 - Cycle times
 - Error rates
 - Resource utilization

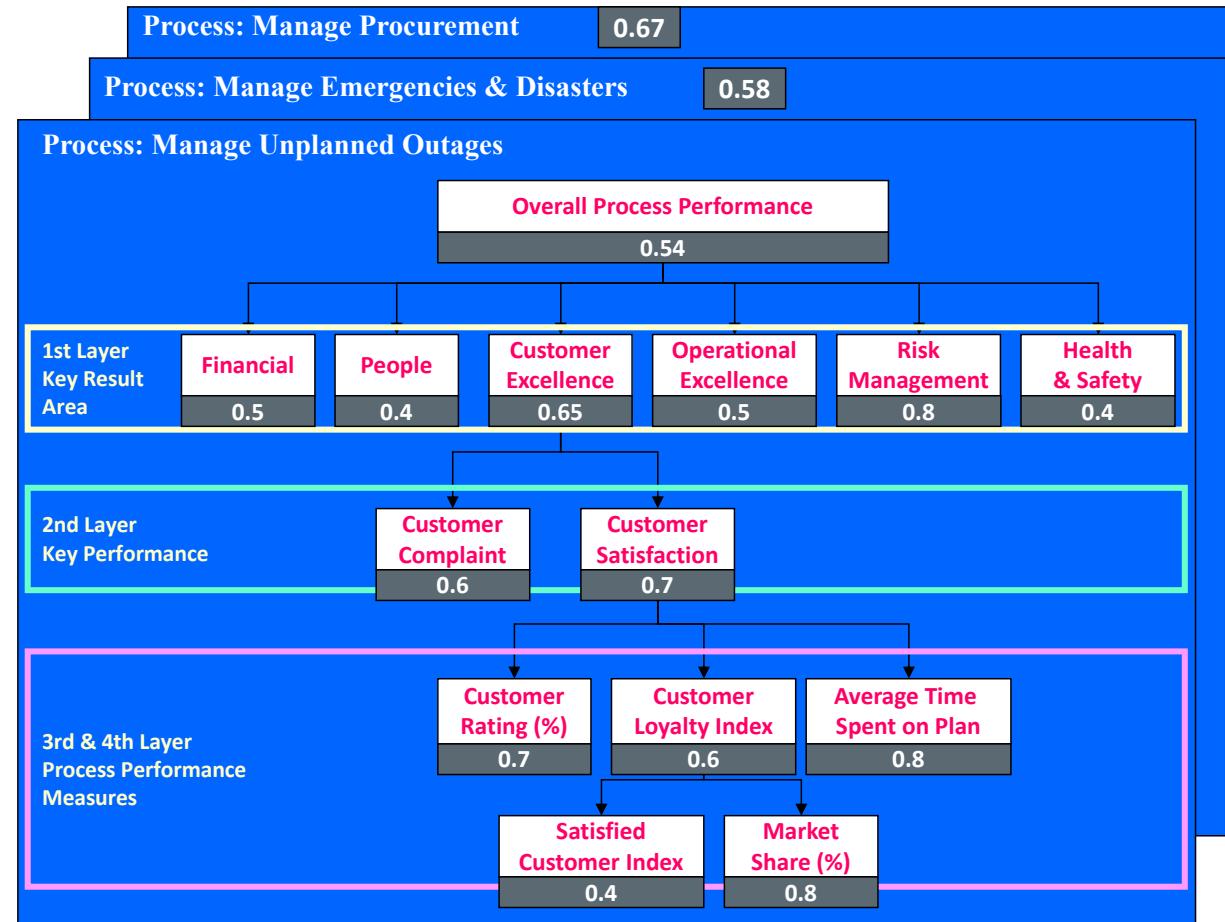
Tactical performance dashboards

- Aimed at process owners / managers
- Emphasis on analysis and management
 - E.g., detecting bottlenecks
- Typical process performance indicators
 - Cycle times
 - Error rates
 - Resource utilization

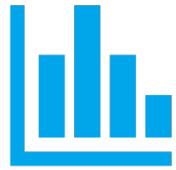


Strategic performance dashboards

- Aimed at executives & managers
- Emphasis on linking process performance to strategic objectives



Process performance dashboards



A dashboard is a collection of widgets



Each widget displays at least one performance measure

E.g. The average case duration



Four types of widgets:

Single-dimensional widgets

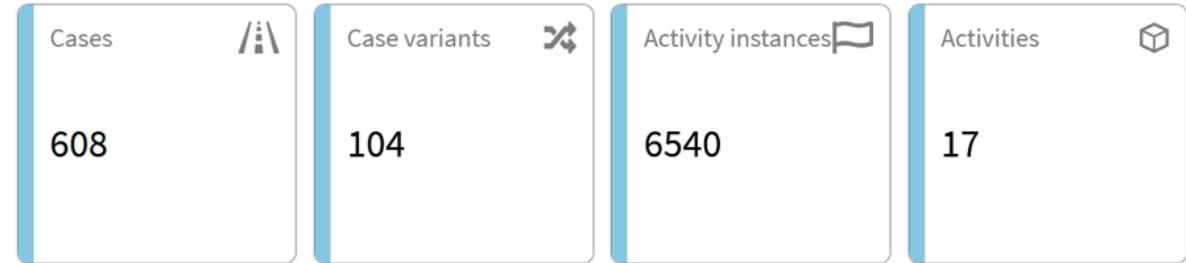
Two-dimensional widgets

Three-dimensional widgets

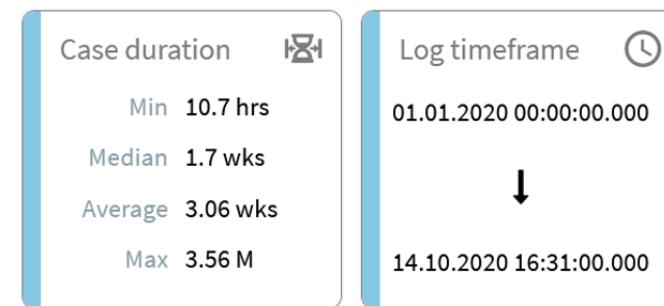
N-dimensional widgets (tables)

Process performance dashboard widgets

- Single-dimensional widgets display one variable



- We can group multiple measures in one widget for convenience.



Process performance dashboard widgets

Two-dimensional widgets display two variables:

- Independent variable (x-axis) – variable with respect to which we want to analyze
- Dependent variable (y-axis) – the variable that we wish to analyze

Usually, a two-dimensional widget is a chart

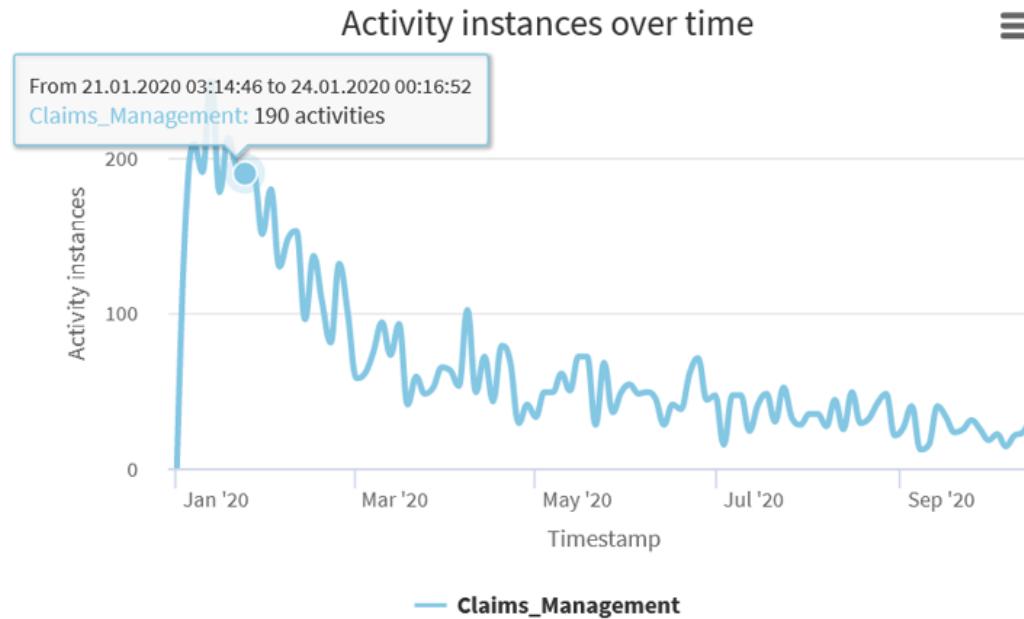
- Bar chart, pie chart, etc.

Types of two-dimensional charts:

- Longitudinal charts: x-axis is time, y-axis is a performance measure
- Cross-sectional chart: x-axis is an attribute (e.g. activity, resource, country), y-axis is a performance measure
- Histogram: x-axis is a performance measure, y-axis is a “count”

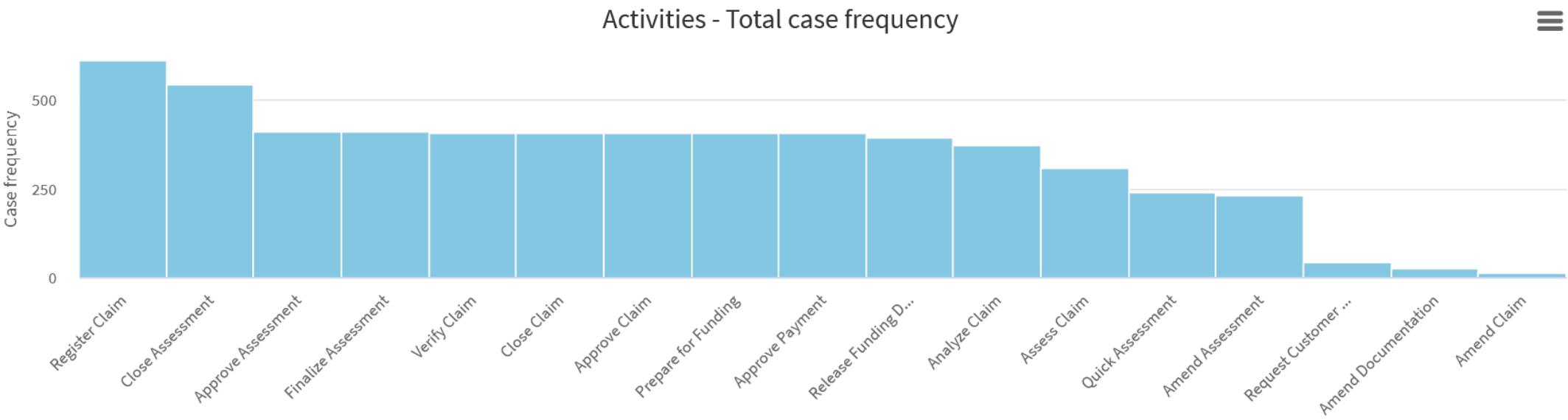
Process performance dashboard widgets

Longitudinal chart



Process performance dashboard widgets

Cross-sectional chart



Process performance dashboard widgets

Cross-sectional chart

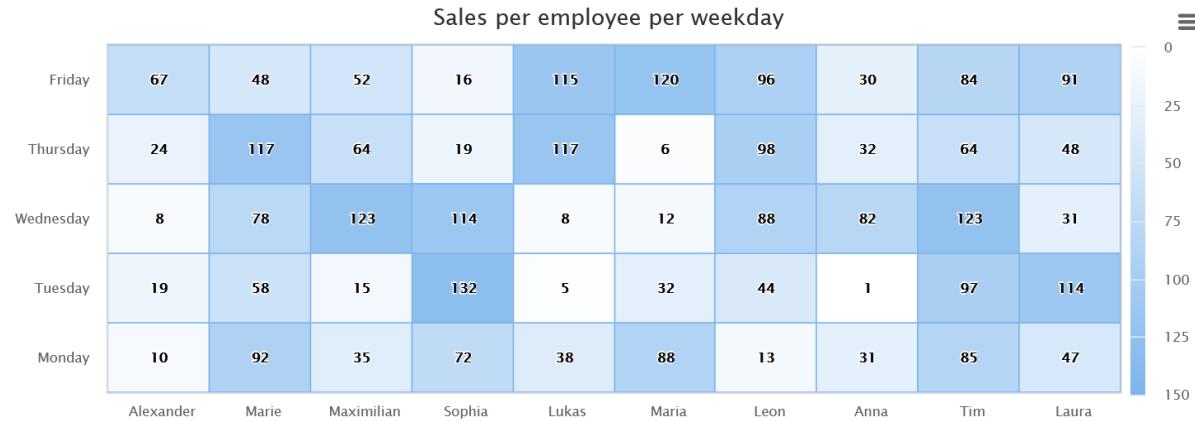


Process performance dashboard widgets

- Three-dimensional widgets display three variables, usually:

- Two cross-sectional variables
- A performance measure

- The third dimension can be coded in different ways:
 - Heat map: intensity and/or color
 - Bubble chart: size and/or intensity
 - Scatter plot: useful when the third dimension is Boolean



Process performance dashboard widgets

- N-dimensional widgets display N dimensions, usually:
 - A cross-sectional variable (e.g. the customer, the case identifier, etc.)
 - Several performance measures
- In general, N-dimensional widgets are tables

Cases ▾	Case ID	Activity instances	Variant	Start time	End time	Duration
	361	38	69	20.02.2020 20:38:00	13.05.2020 00:42:00	2.7 months
	249	32	103	03.02.2020 04:57:00	25.04.2020 23:31:00	2.72 months
	223	29	74	29.01.2020 08:22:00	13.05.2020 05:05:00	3.44 months
	637	28	96	02.04.2020 15:24:00	27.06.2020 21:53:00	2.83 months
	872	25	99	09.05.2020 05:06:00	20.08.2020 20:01:00	3.4 months
	438	24	65	05.03.2020 01:49:00	13.06.2020 15:20:00	3.3 months
	929	24	48	16.05.2020 08:41:00	09.08.2020 18:03:00	2.81 months
	1011	23	62	28.05.2020 13:20:00	03.09.2020 21:13:00	3.23 months
	459	23	51	07.03.2020 22:12:00	28.05.2020 17:36:00	2.69 months
« «	1 / 61	> >				[1 - 10 / 608]

Designing Process Performance Dashboard

Identify a user or users and define a clear question or set of questions that users will answer with this question



Identify the type of dashboard elements based on the insights required to answer the question



Identify a type of visualization (e.g. type of chart) for the element



Determine x-axis (independent variable): longitudinal time), cross-sectional (attribute)



Determine y-axis (dependent variable): performance measure or attribute and the aggregation function

Tools for performance dashboarding

Operational-Tactical – Business Activity Monitoring (BAM)

- Axway AMPLIFY
- VITRIA Operational Process Intelligence
- Oracle BAM
- SAP Operational Process Intelligence (OPI)

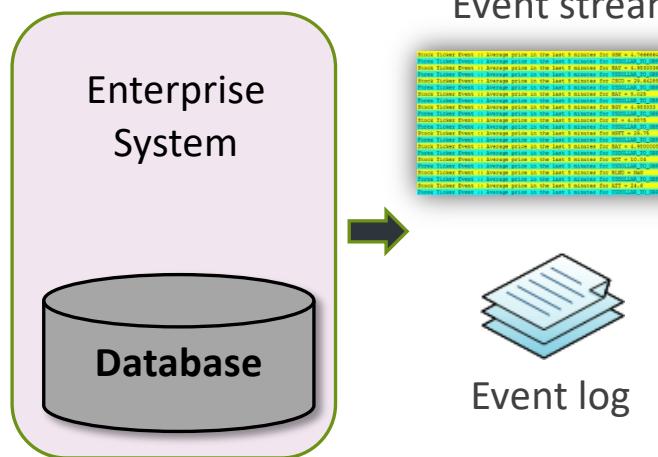
Tactical

- Business Intelligence (BI) tools: PowerBI, Qlikview, Tableau...
- Process Mining (dashboard modules): Apmomore, Celonis, QPR, ...

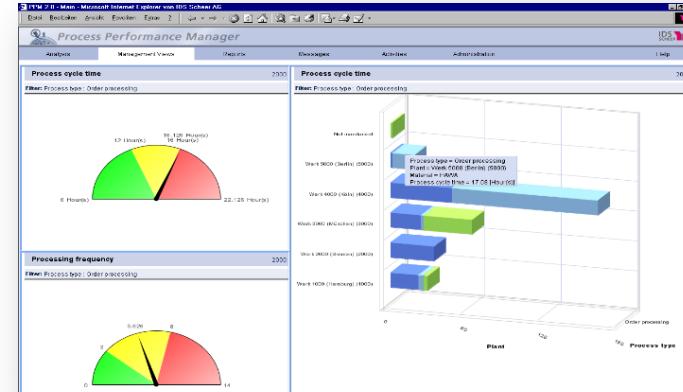
Strategic Level – Balanced Scorecard tools

- BSC Designer Online, Quickscore, Sisense, etc.

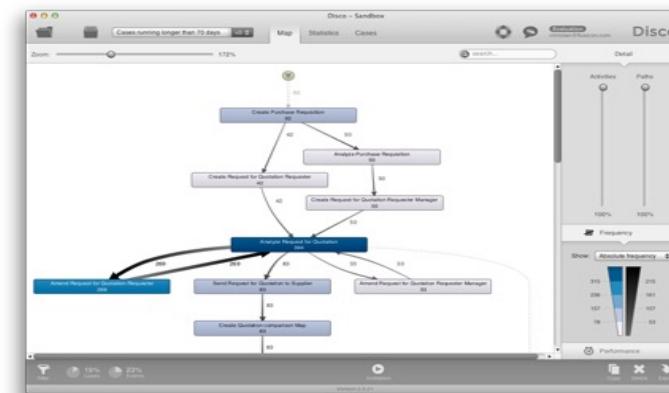
Business Process Monitoring



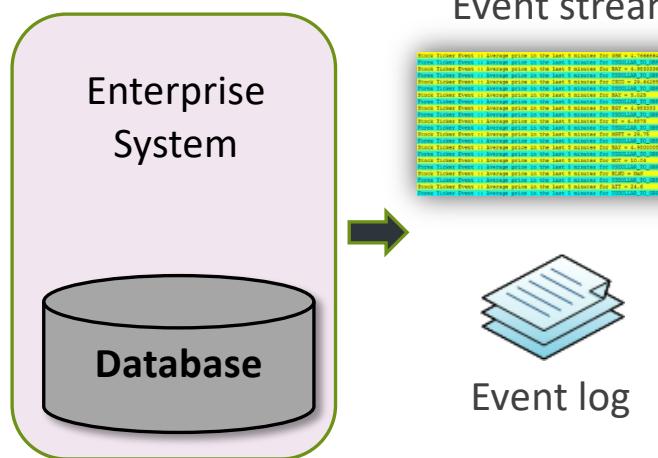
Performance Dashboards



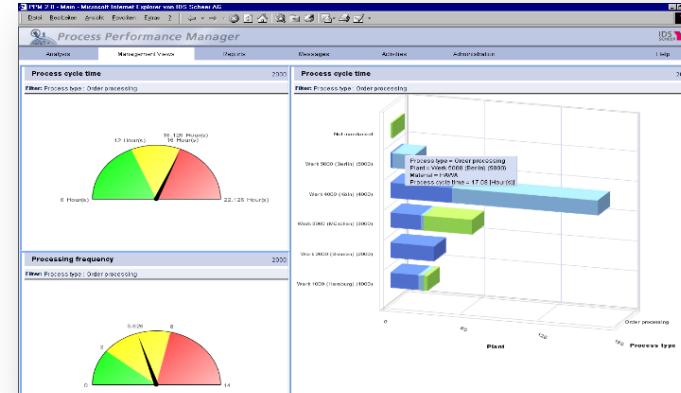
Process Mining



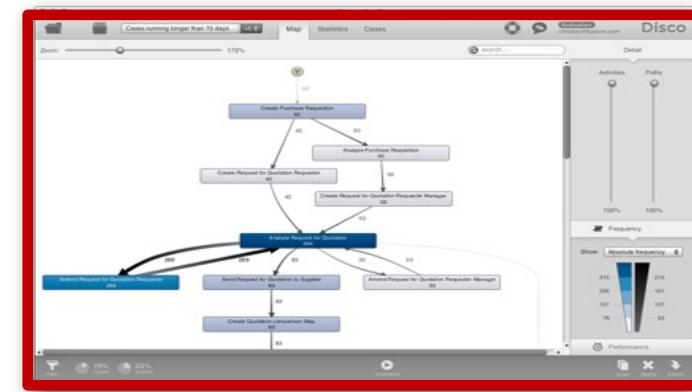
Business Process Monitoring



Performance Dashboards



Process Mining

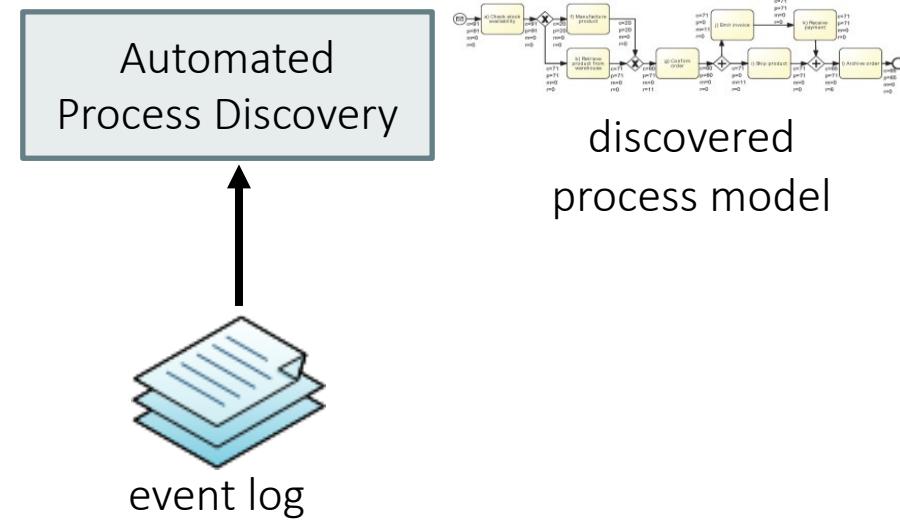


Process Mining



event log

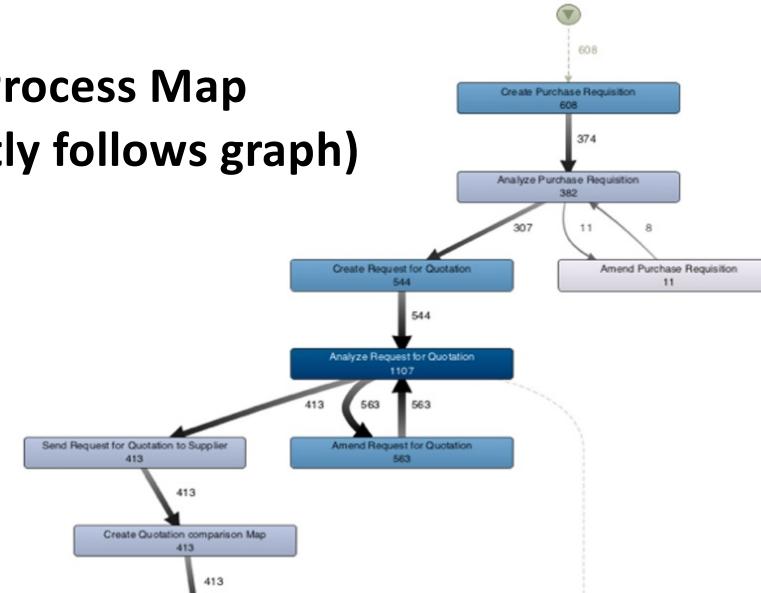
Process Mining



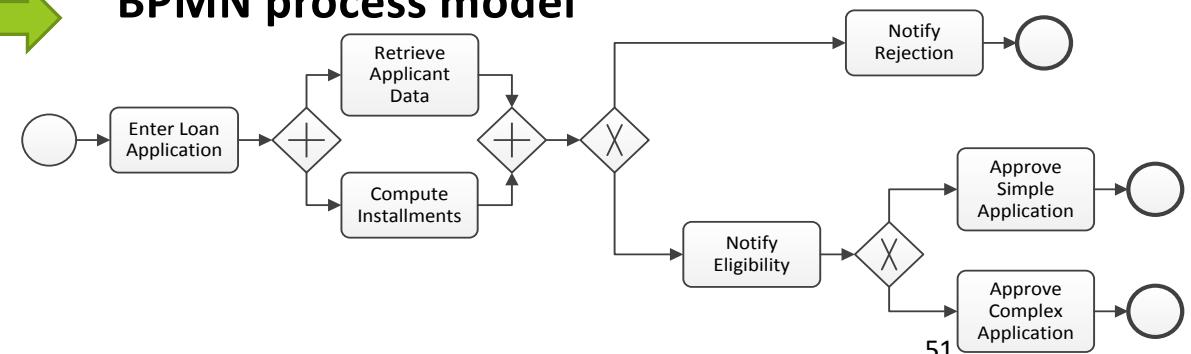
Automated Process Discovery

CID	Task	Time Stamp	...
13219	Enter Loan Application	2007-11-09 T 11:20:10	-
13219	Retrieve Applicant Data	2007-11-09 T 11:22:15	-
13220	Enter Loan Application	2007-11-09 T 11:22:40	-
13219	Compute Installments	2007-11-09 T 11:22:45	-
13219	Notify Eligibility	2007-11-09 T 11:23:00	-
13219	Approve Simple Application	2007-11-09 T 11:24:30	-
13220	Compute Installements	2007-11-09 T 11:24:35	-
...

Process Map
(directly follows graph)



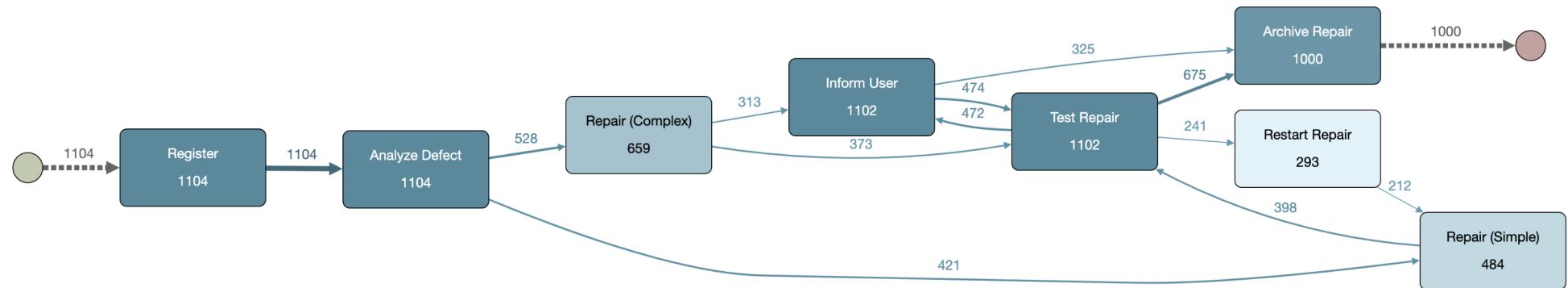
BPMN process model



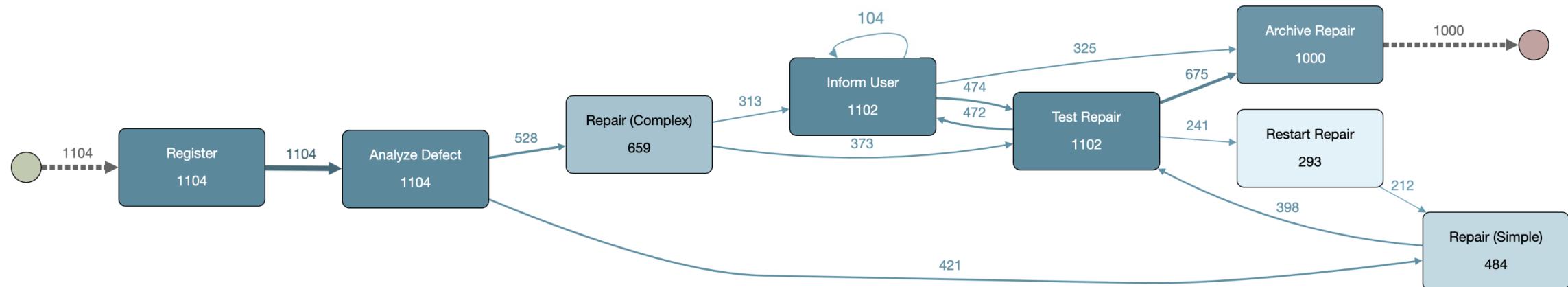
Process maps (aka Directly-follows graphs)

A process map of an event log is a graph where:

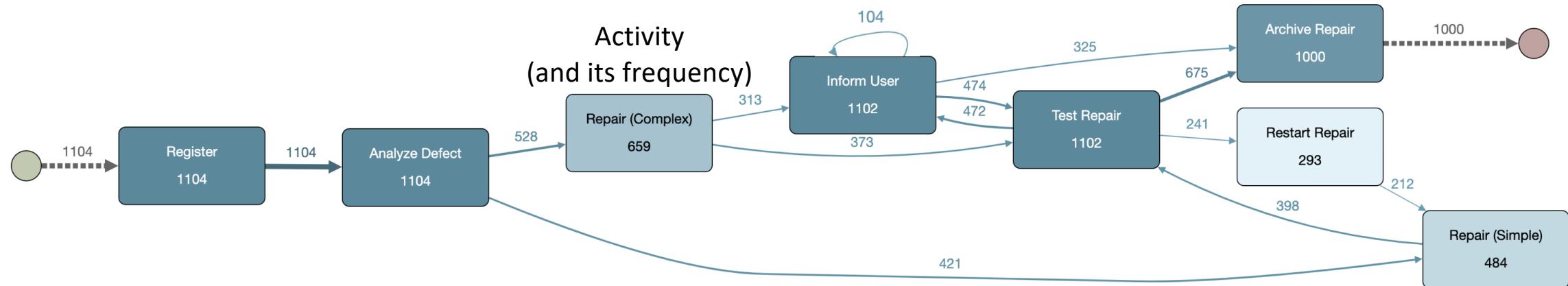
- Each activity is represented by one node
- An arc from activity A to activity B means that B is directly followed by A in at least one case in the log



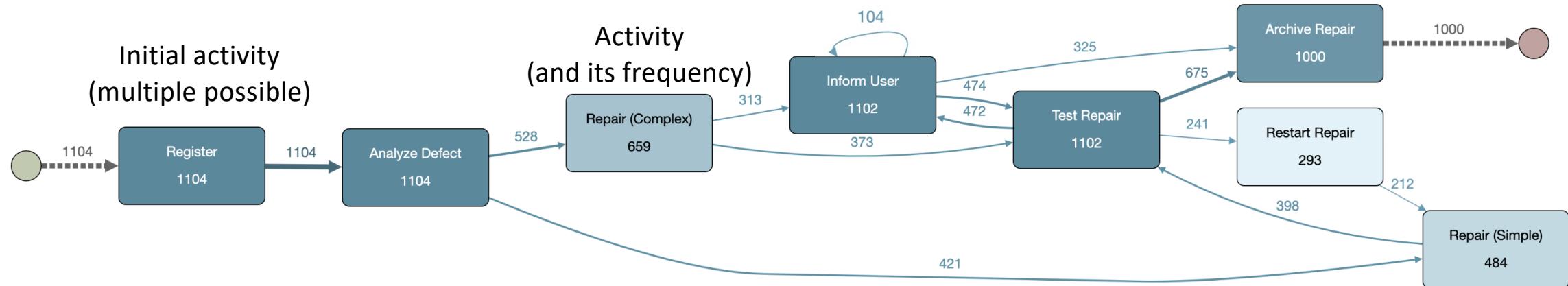
Anatomy of a Process map



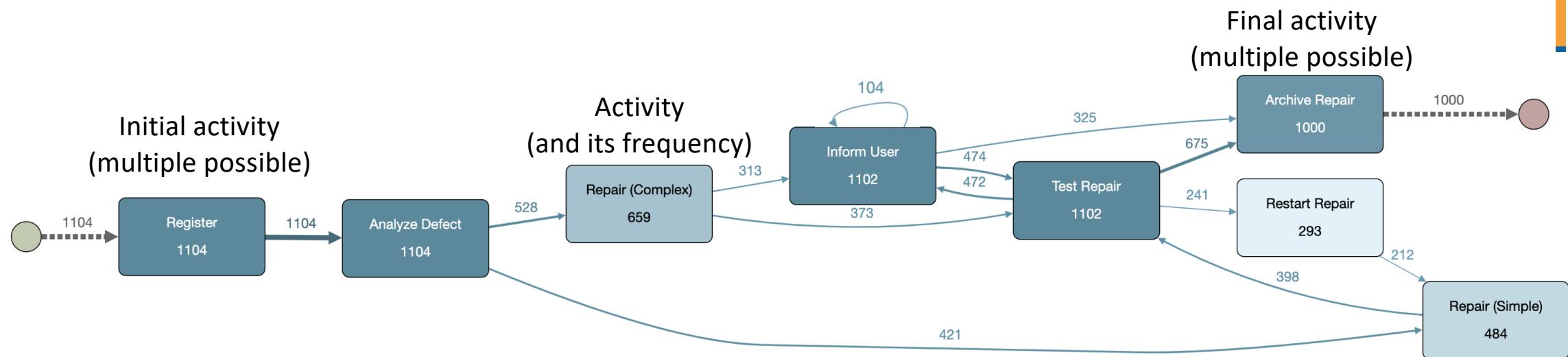
Anatomy of a Process map



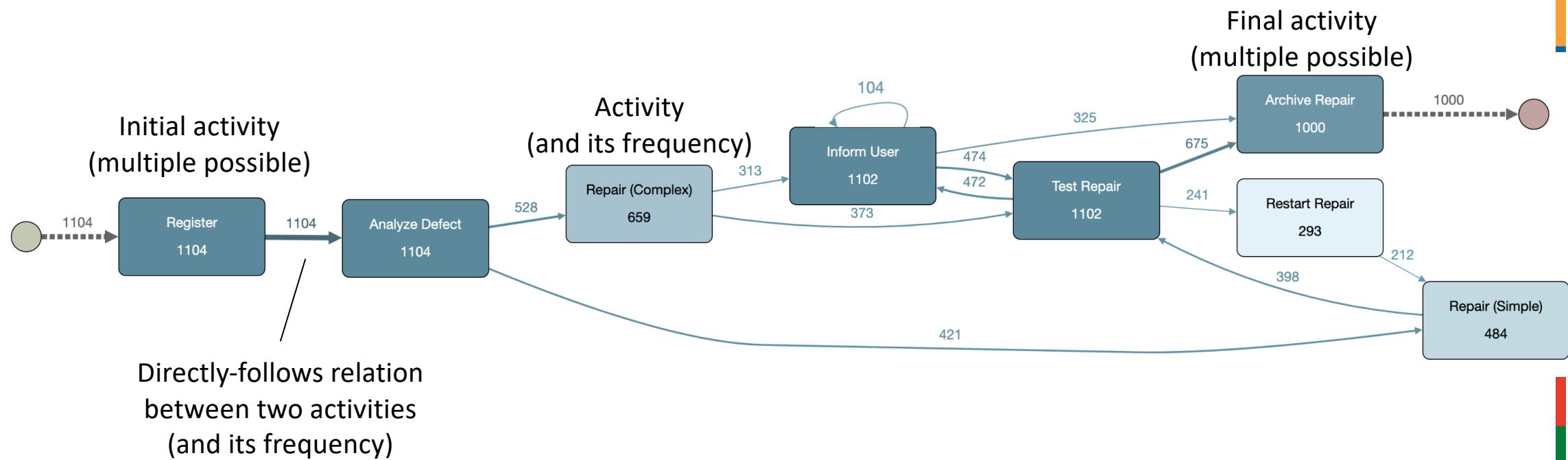
Anatomy of a Process map



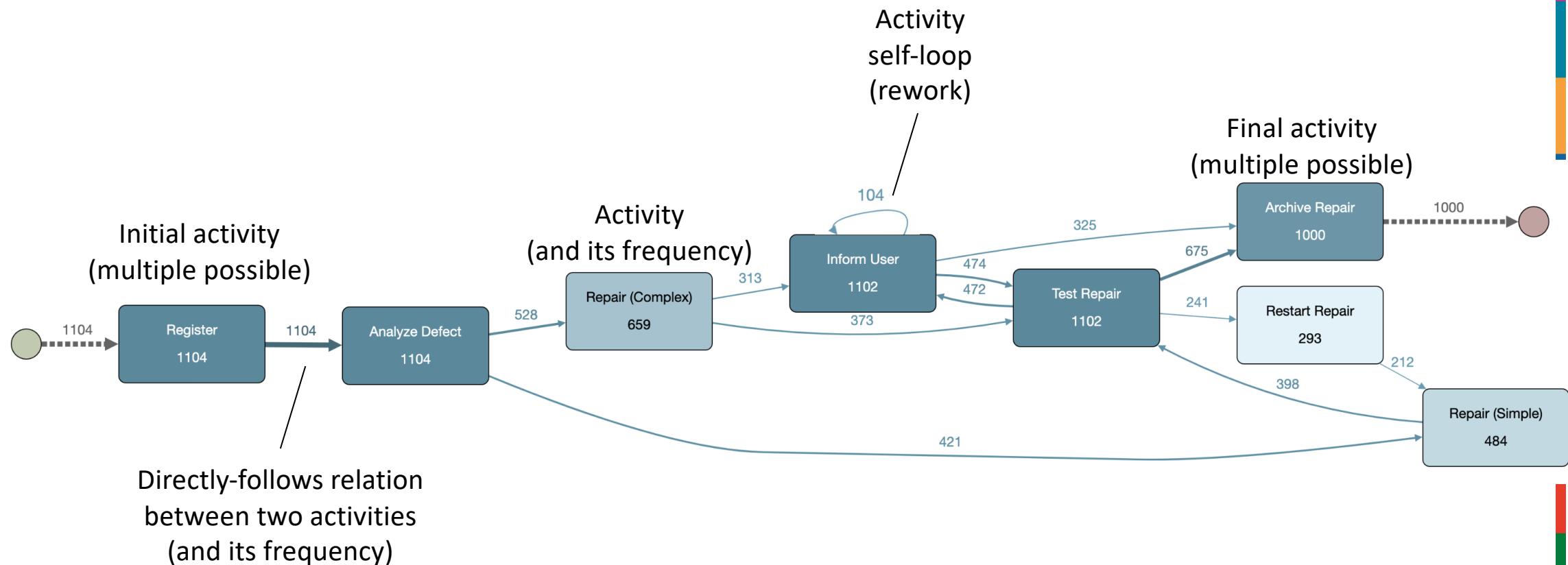
Anatomy of a Process map



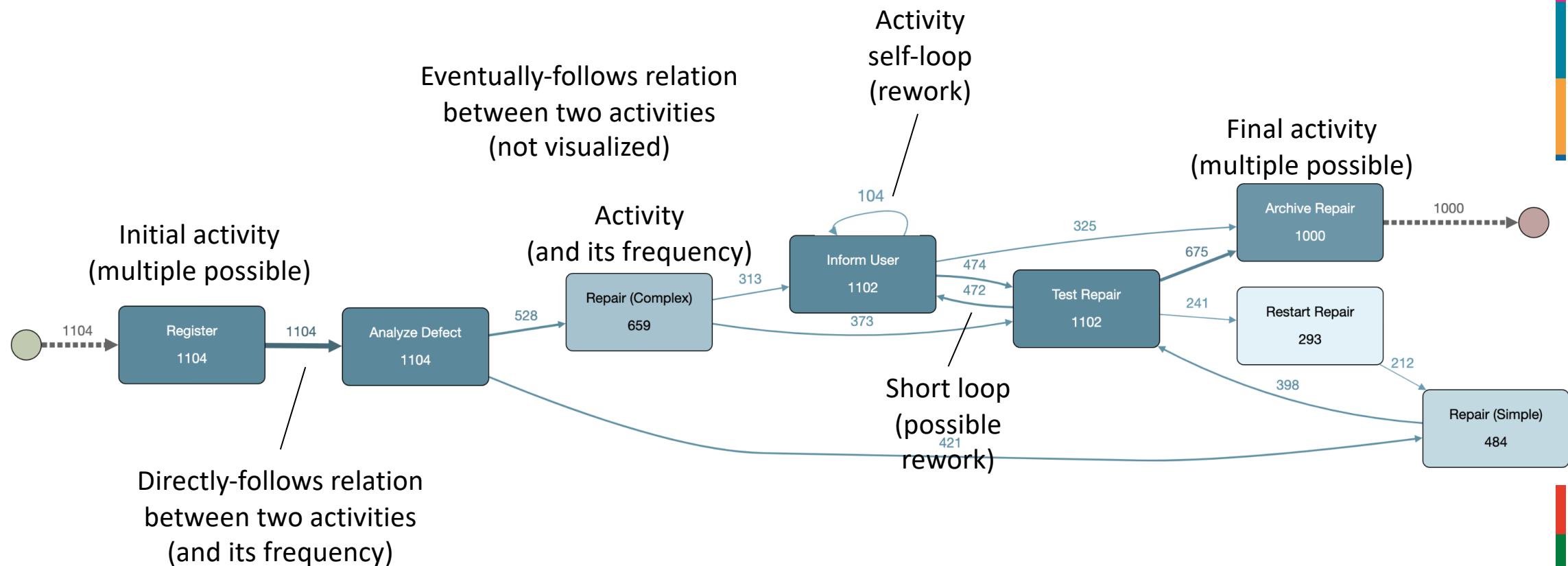
Anatomy of a Process map



Anatomy of a Process map



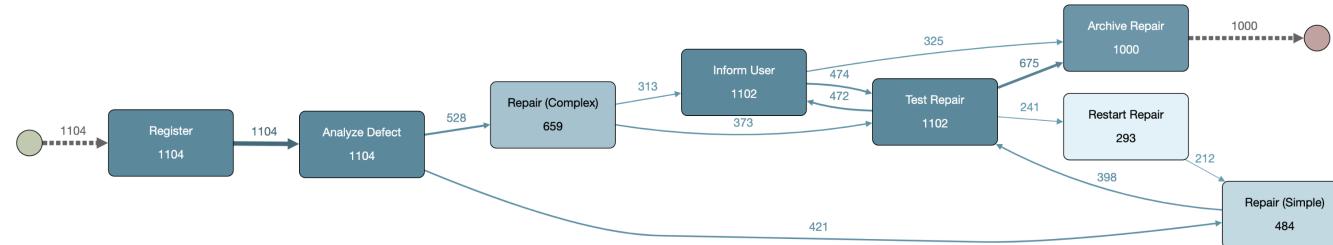
Anatomy of a Process map



Process maps: overlays

Nodes in a process map can be coloured, and arcs' thickness can be changed, to capture:

- **Frequency:** How often a given activity or directly-follows relation occurs?
- **Duration:** processing times and cycle times for activities, waiting times for directly-follow relations
- **Other attributes:** some tools support enhancement by other attributes, e.g., cost, revenue, sales volume, if these data attributes are available



Example: Process Map

with frequency

Event log

5: a,b,c,g,e,h

4: a,b,c,f,g,h

3: a,b,d,g,e,h

3: a,b,d,e,g,h

Example: Process Map

with frequency

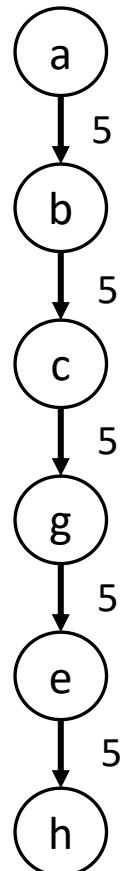
Event log

5: a,b,c,g,e,h

4: a,b,c,f,g,h

3: a,b,d,g,e,h

3: a,b,d,e,g,h



Example: Process Map

with frequency

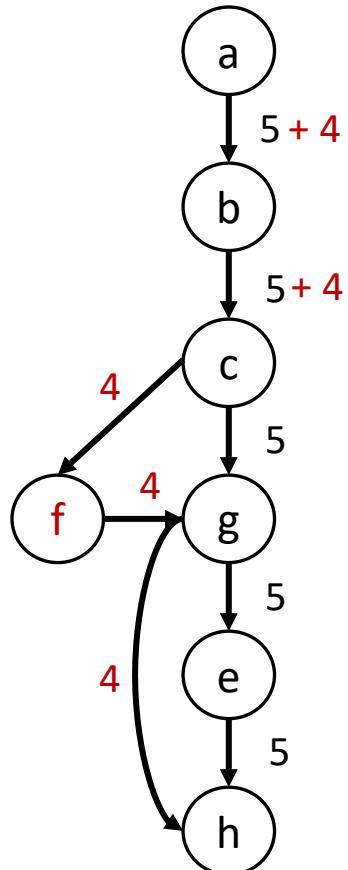
Event log

5: a,b,c,g,e,h

4: a,b,c,f,g,h

3: a,b,d,g,e,h

3: a,b,d,e,g,h



Example: Process Map

with frequency

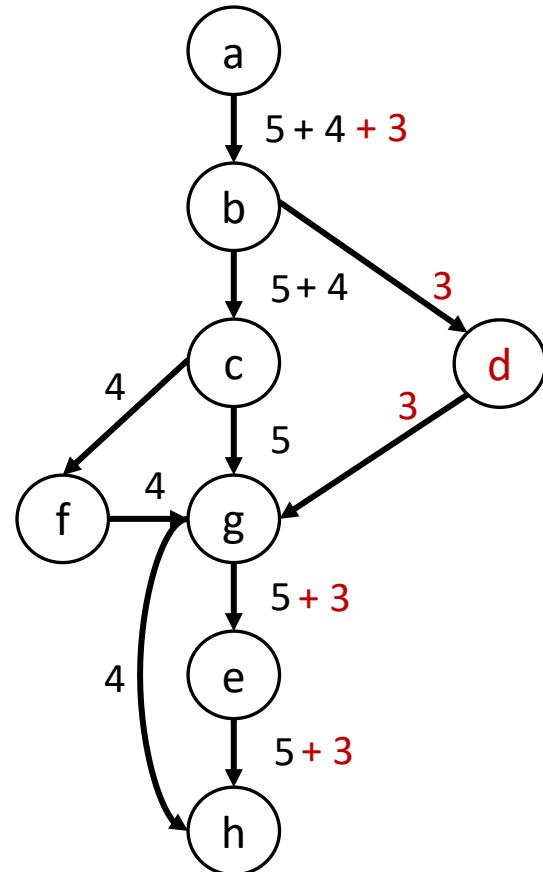
Event log

5: a,b,c,g,e,h

4: a,b,c,f,g,h

3: a,b,d,g,e,h

3: a,b,d,e,g,h



Example: Process Map

with frequency

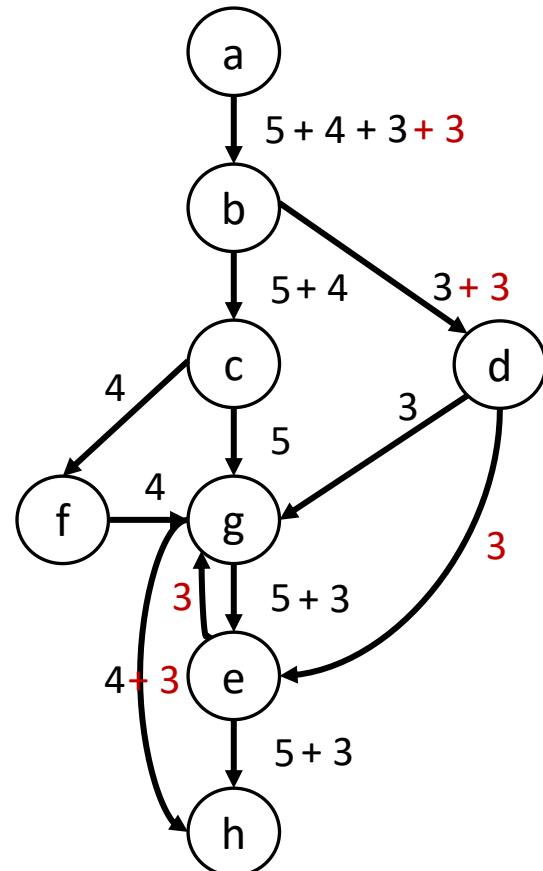
Event log

5: a,b,c,g,e,h

4: a,b,c,f,g,h

3: a,b,d,g,e,h

3: a,b,d,e,g,h



Example: Process Map

with frequency

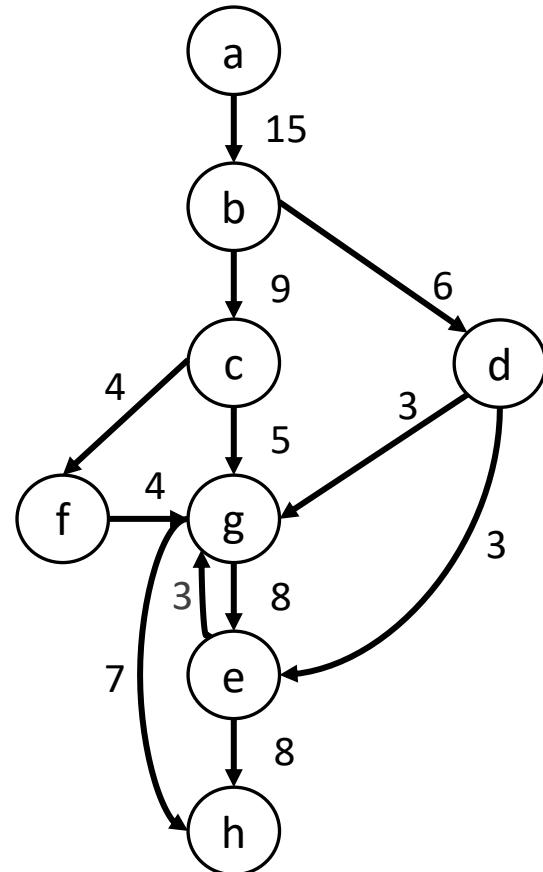
Event log

5: a,b,c,g,e,h

4: a,b,c,f,g,h

3: a,b,d,g,e,h

3: a,b,d,e,g,h

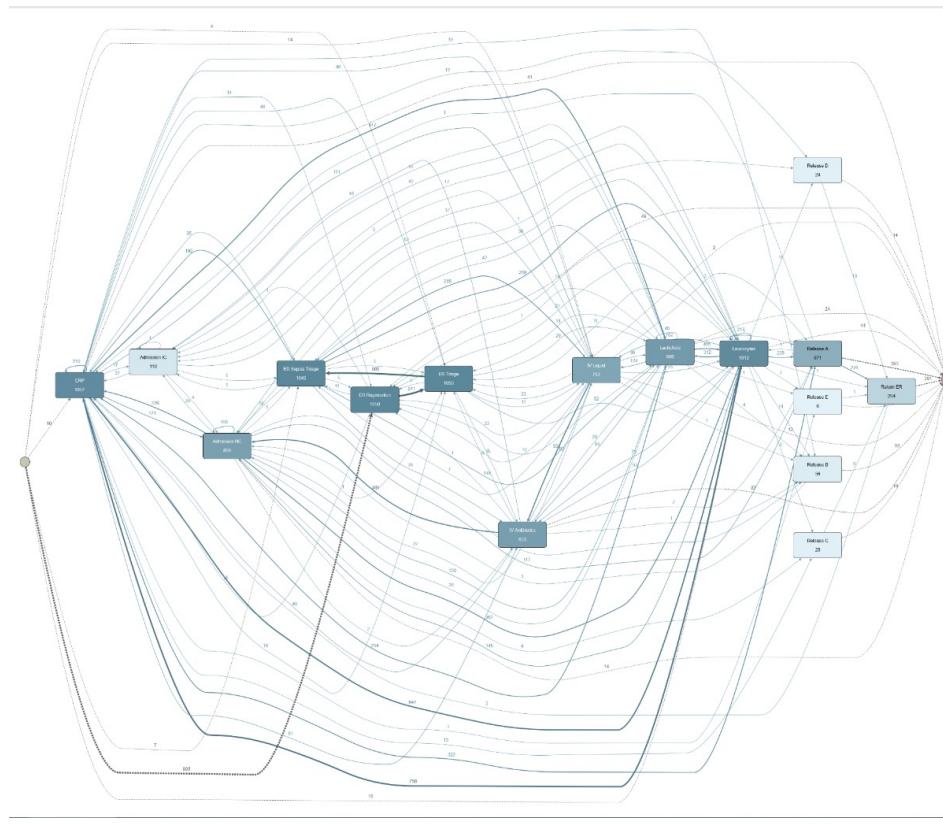


Process Maps

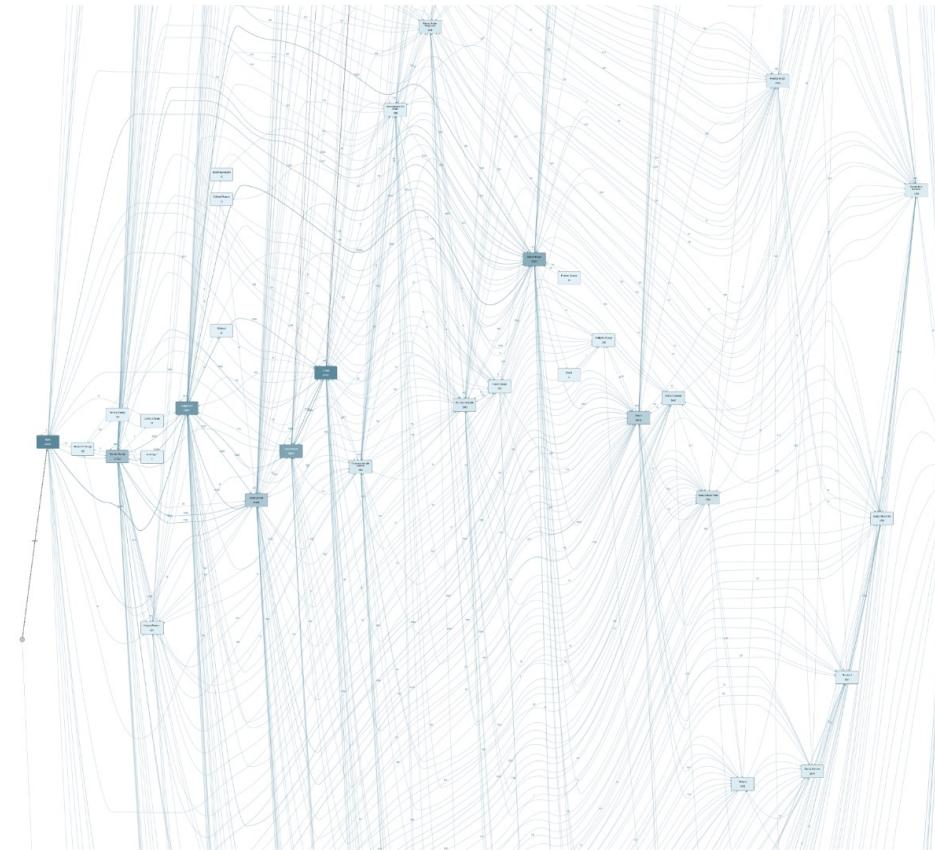
- Commercial process mining tools use process maps as the main visualization technique for event logs
- These tools also provide three types of operations:
 1. **Abstract** the process map:
 2. **Filter** the traces in the event log...

Abstraction and Filtering

- Real-life processes are really messy



Patient Treatment Process @ Hospital (Sepsis infections)



IT Incident Management @ Bank

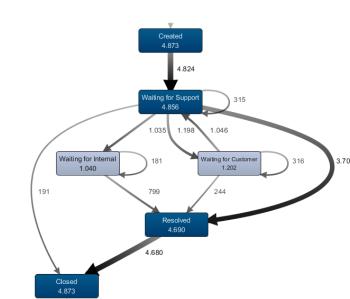
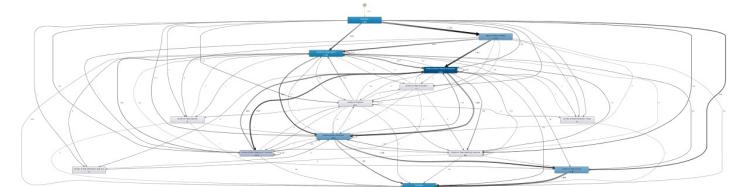
Process map: Abstraction

Common process mining tools provide abstraction capabilities on top of process maps, to:

- Show only **most frequent** activities
- Show only **most frequent** arcs

Some tools offer further abstraction capabilities to:

- Show only **least frequent** activities
- Show only **least frequent** arcs
- Show only **fastest/slowest** activities
- Show only **fastest/slowest** arcs



Log filters

- Case-Level Log Filters
 - Retain (or remove) all cases that fulfill a given condition, for example all cases that take more than 3 days
- Event-Level Filter
 - Retain or remove individual events (*activity instances*) that fulfill a given condition, e.g., all activity instances that were performed by “John Smith”
- A case-level filter either retains an entire case or it removes the whole case while an event-level filter may remove some events in a case and remove others.

Case-Level log filters

Attribute filters

- Retain only cases such that at least one event fulfils a condition over one or more attributes, for example: *at least one event where Activity = "Close order" and Resource = "Anti Alman"*

Timeframe filter

- Retain traces that start, end, or are contained in a time period (e.g. all cases that started in January)

Performance filter

- Retain only cases that have a duration above or below a given value (e.g. 3 days)

Path filter

- Retain traces where there is a pair of events that fulfil a given condition (e.g. "Create invoice" eventually followed by "Create purchase order")

Repetition filter

- Retain traces that start with or finish with an event that fulfils a given condition

Event-Level log filters

Attribute filters

- Retain only events that fulfil a condition, for example: *at least one event where Activity = “Close order”*

Timeframe filter

- Retain all events that happen during a certain time period (e.g. all events in January)

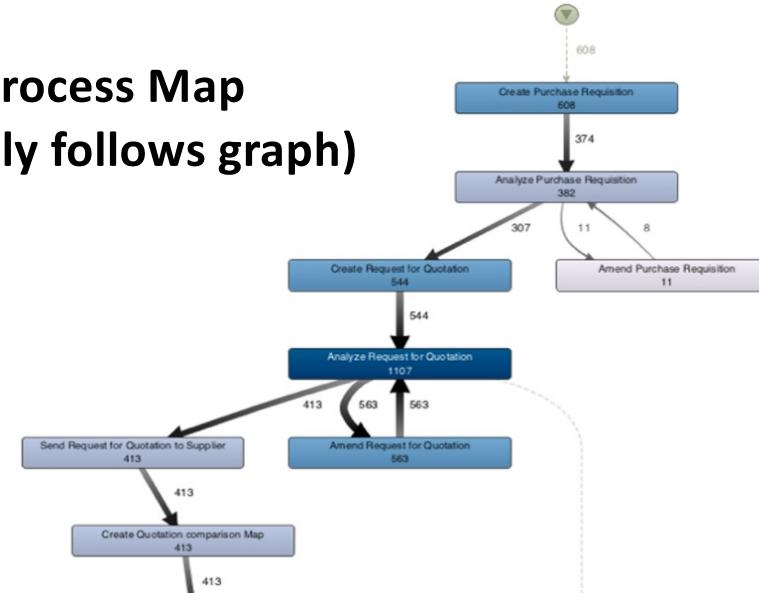
Between filter

- For each case, retain all events between the (first or last) occurrence of an activity and the (first or last occurrence) of another activity (e.g. between “Create PO” and “Approve PO”)

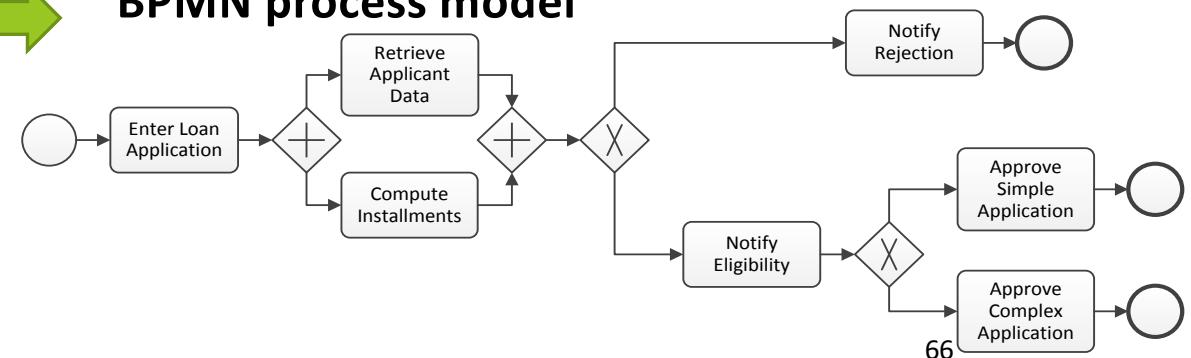
Automated Process Discovery

CID	Task	Time Stamp	...
13219	Enter Loan Application	2007-11-09 T 11:20:10	-
13219	Retrieve Applicant Data	2007-11-09 T 11:22:15	-
13220	Enter Loan Application	2007-11-09 T 11:22:40	-
13219	Compute Installments	2007-11-09 T 11:22:45	-
13219	Notify Eligibility	2007-11-09 T 11:23:00	-
13219	Approve Simple Application	2007-11-09 T 11:24:30	-
13220	Compute Installements	2007-11-09 T 11:24:35	-
...

Process Map
(directly follows graph)



BPMN process model



Limitations of process maps

Process maps over-generalize: some paths of a process map might actually not exist and might not make sense

- Example: Draw the process map of [abc, adc, afce, afec] and check which cases it can recognize for which there is no support in the event log.

Process maps make it difficult to distinguish conditional branching, parallelism and loops

- See previous example... or a simpler one: [abcd, acbd]

Solution: automated discovery of BPMN models

Limitations of process maps

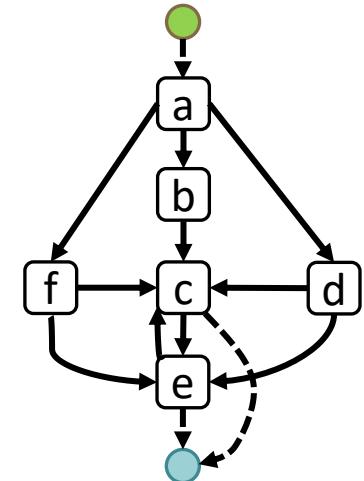
Process maps over-generalize: some paths of a process map might actually not exist and might not make sense

- Example: Draw the process map of [abc, adc, afce, afec] and check which cases it can recognize for which there is no support in the event log.

Process maps make it difficult to distinguish conditional branching, parallelism and loops

- See previous example... or a simpler one: [abcd, acbd]

Solution: automated discovery of BPMN models



Limitations of process maps

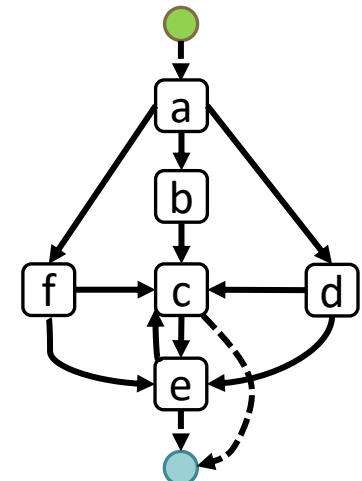
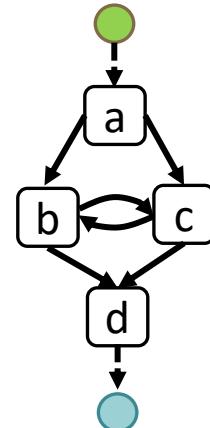
Process maps over-generalize: some paths of a process map might actually not exist and might not make sense

- Example: Draw the process map of [abc, adc, afce, afec] and check which cases it can recognize for which there is no support in the event log.

Process maps make it difficult to distinguish conditional branching, parallelism and loops

- See previous example... or a simpler one: [abcd, acbd]

Solution: automated discovery of BPMN models



Limitations of process maps

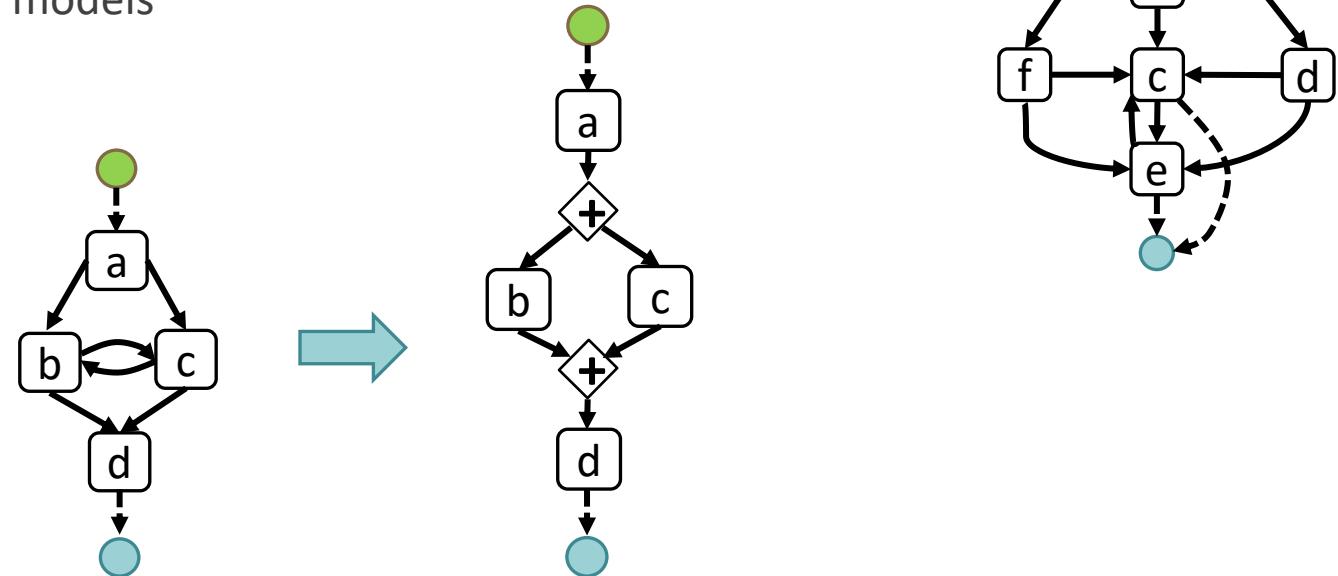
Process maps over-generalize: some paths of a process map might actually not exist and might not make sense

- Example: Draw the process map of [abc, adc, afce, afec] and check which cases it can recognize for which there is no support in the event log.

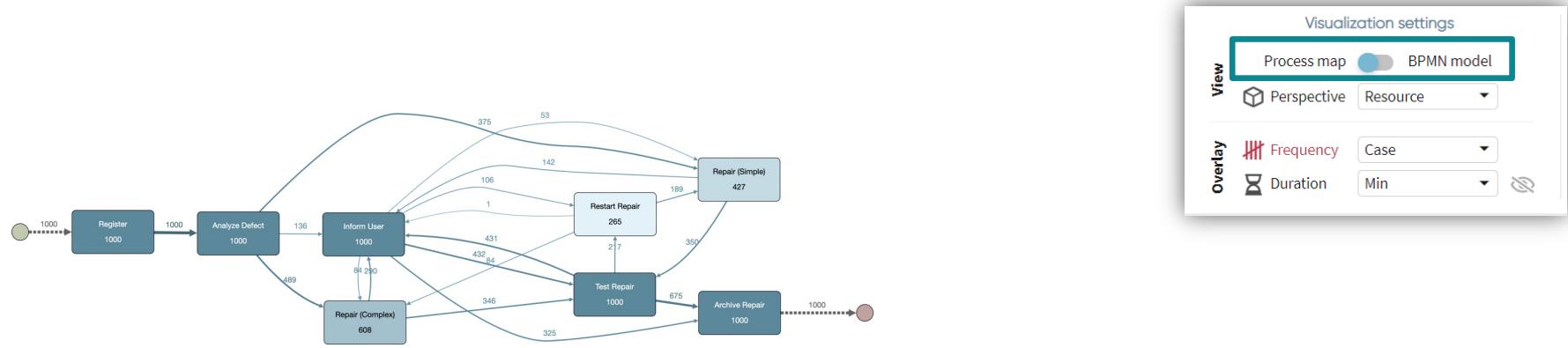
Process maps make it difficult to distinguish conditional branching, parallelism and loops

- See previous example... or a simpler one: [abcd, acbd]

Solution: automated discovery of BPMN models

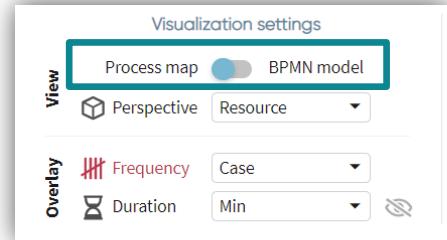
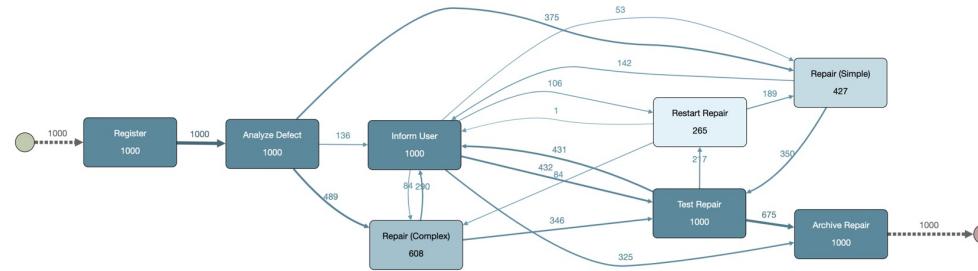


Discovery of BPMN models

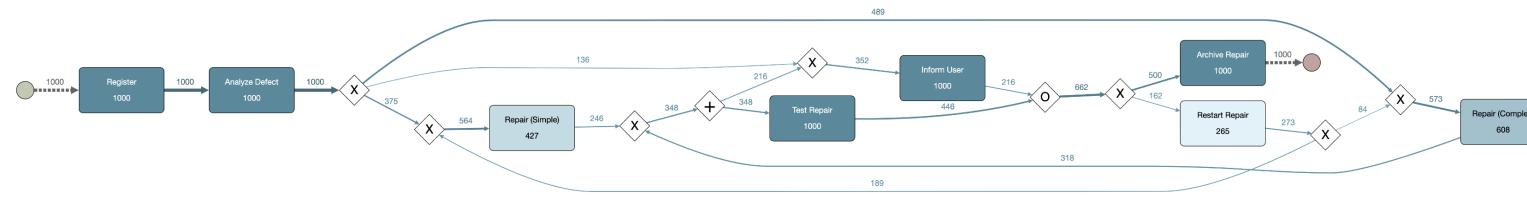


Note: Apromore uses an algorithm called **Split Miner** to turn process maps into process models

Discovery of BPMN models

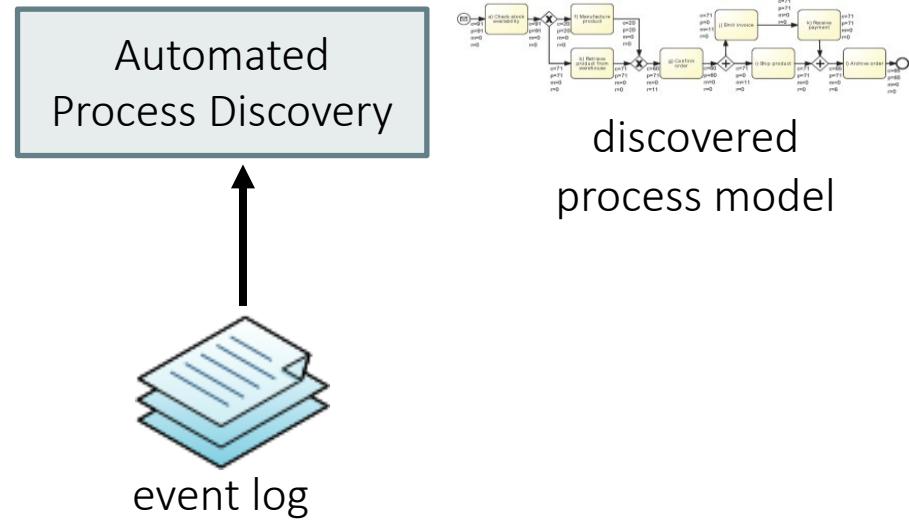


A given configuration of node/arc sliders, and parallelism slider

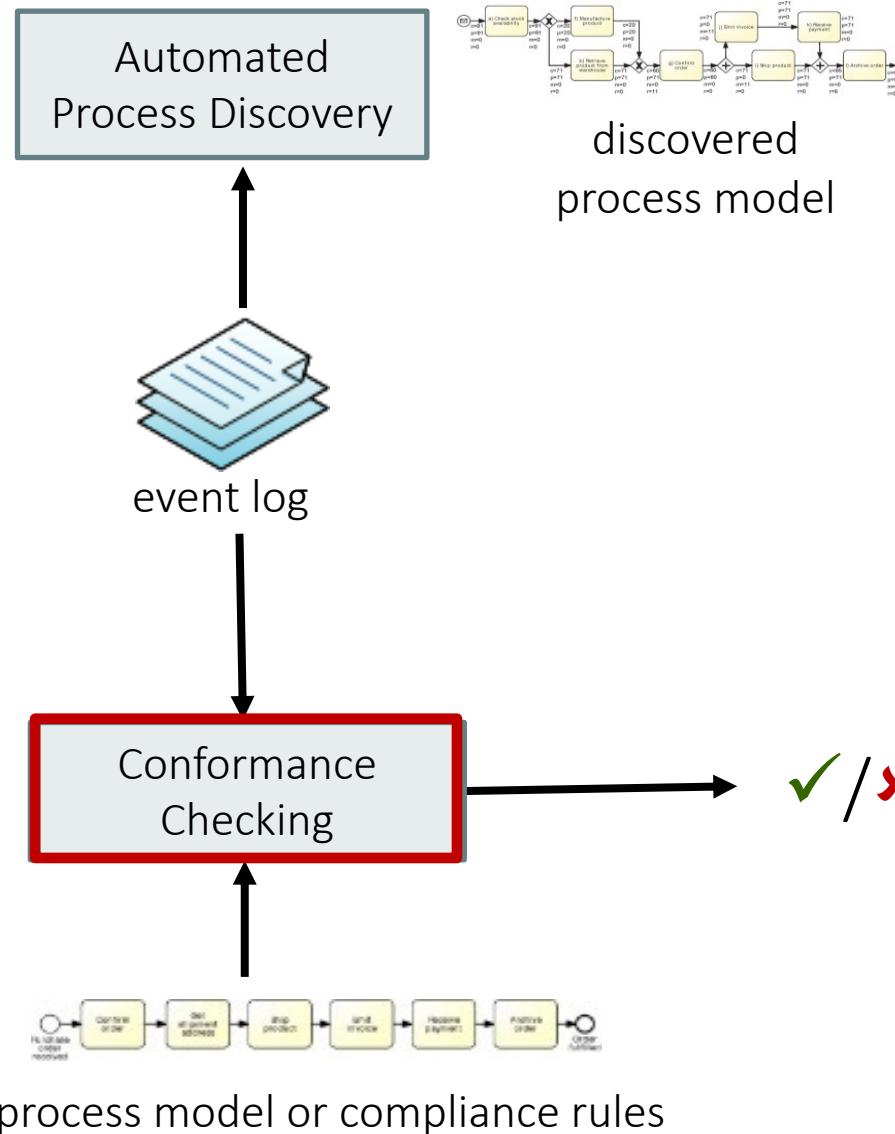


Note: Apromore uses an algorithm called **Split Miner** to turn process maps into process models

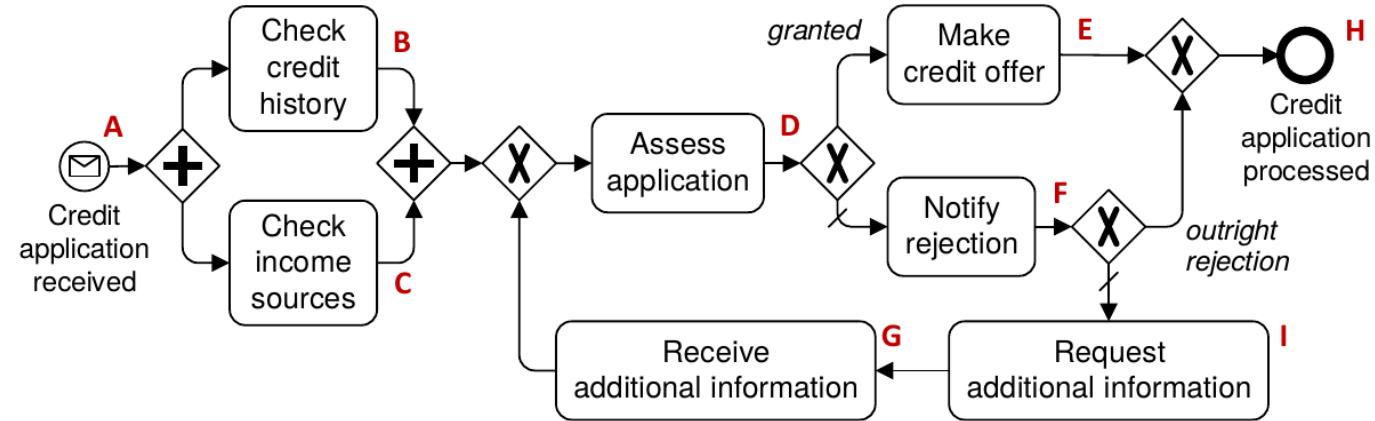
Process Mining



Process Mining



Conformance Checking (model-based)



Event log:

ABCDEH

ACBDEH

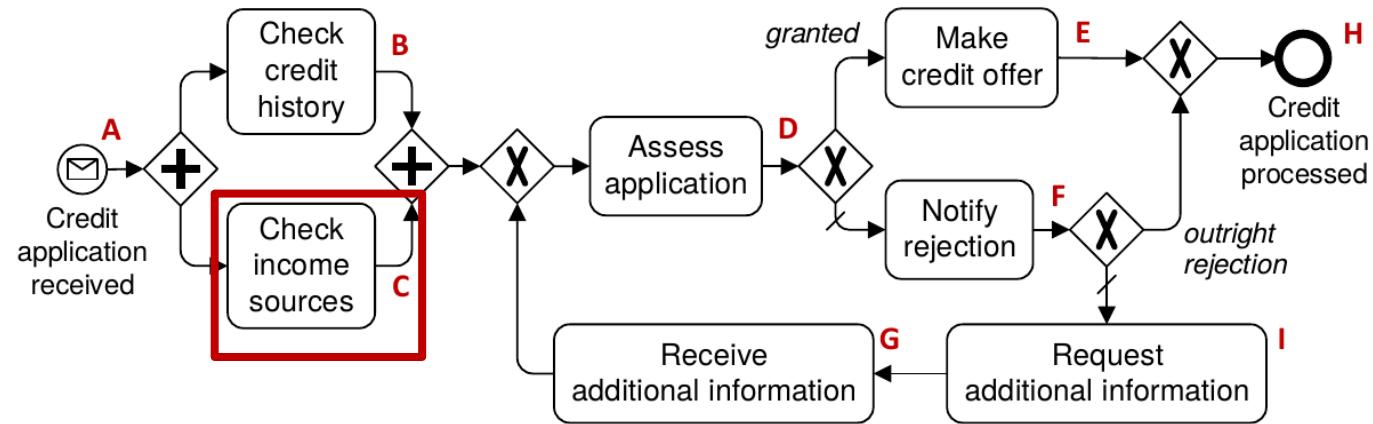
ABCDFH

ACBDFH

ABDEH

ABDFH

Conformance Checking (model-based)



Event log:

ABCDEH

ACBDEH

ABCDFH

ACBDFH

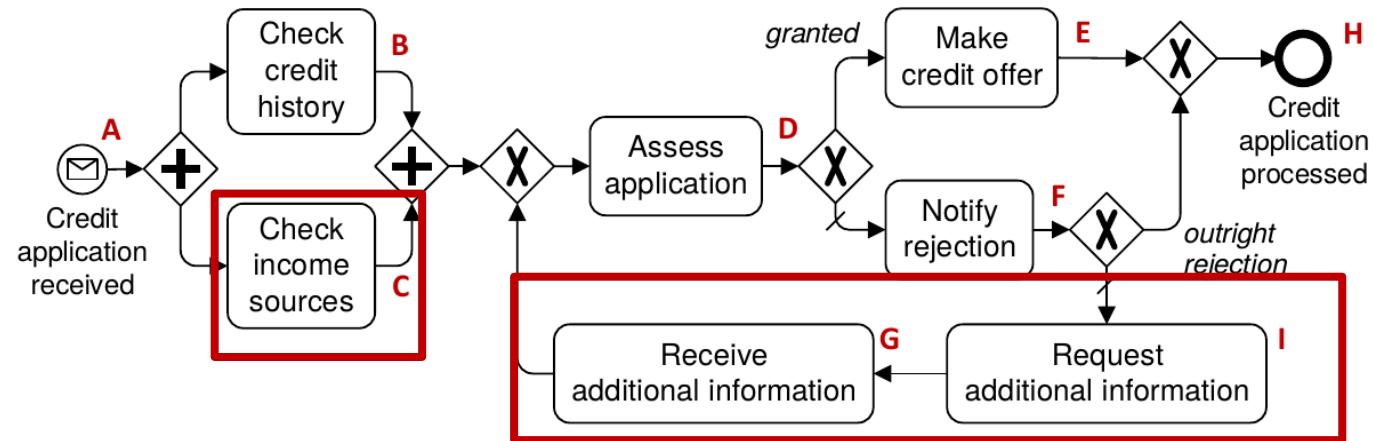
ABDEH

ABDFH

Unfitting behaviour:

- Task C is *optional* (i.e. may be skipped) in the log

Conformance Checking (model-based)



Event log:

ABCDEH

ACBDEH

ABCDFH

ACBDFH

ABDEH

ABDFH

Unfitting behaviour:

- Task C is *optional* (i.e. may be skipped) in the log

Additional behavior:

- The *cycle* including IGDF is not observed in the log

Conformance checking (rule-based)

- Oftentimes, we don't have a full process model available
- Instead, we have a set of compliance rules, e.g.
 - Certain tasks MUST be executed in every case
 - Certain tasks can be executed at most once
 - When a task A occurs, another task B must also occur (within e.g. 2 hours)
 - Two tasks A and B CANNOT be executed by the same person (four-eyes principle)

Conformance Checking Levels

Accordingly, we can identify the following *compliance checking levels*:

- **Flow:** have all mandatory activities been performed? Have activities been performed in the right order and with the right frequency?
- **Temporal:** has a given timeframe between activities, process milestones or for the end-to-end process been adhered to?
- **Resource:** did everyone obey the segregation of duties?
- **Exception:** explore highly-infrequent behavior to identify possible deviations from the standard process

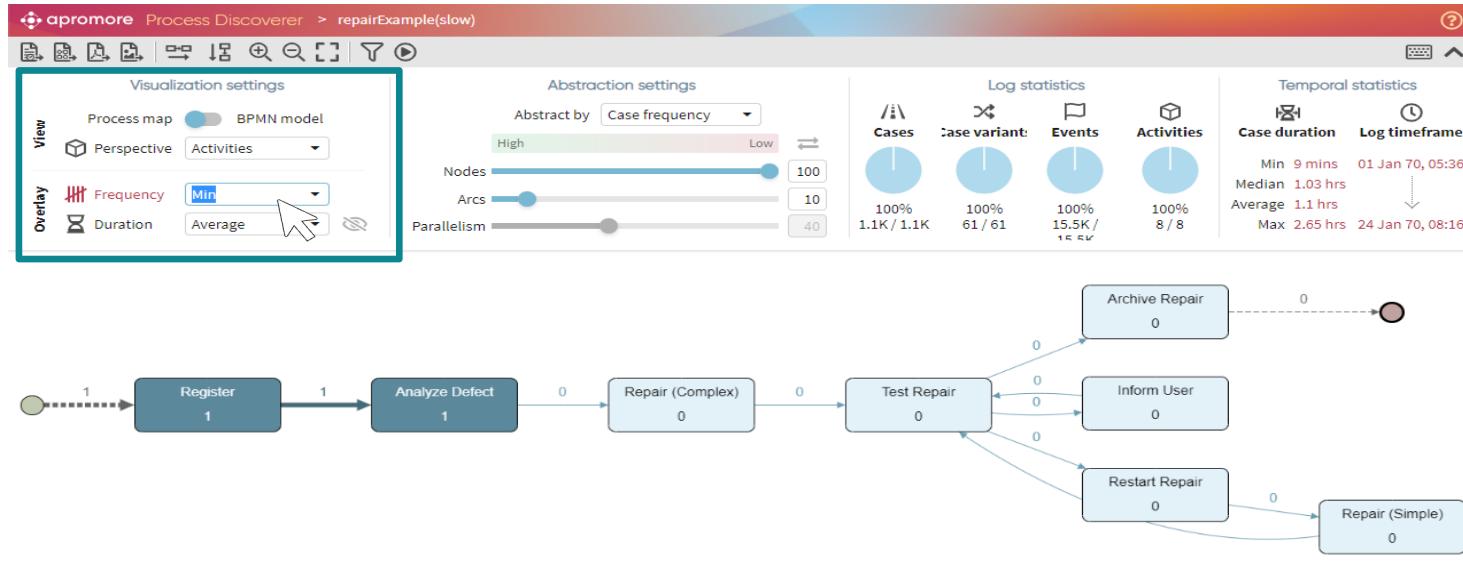
Procedure:

1. Define compliance rules (based on attribute-value pairs, activity orders or performance targets)
2. Apply one or more filters to keep cases that violate the rules in question
3. If needed, switch perspective to highlight non-compliance issues (e.g. show resources that violate a segregation of duties principle)
4. Assess impact on process performance

Flow Compliance Checking

1

Mandatory tasks: Have all mandatory tasks been completed?



To check whether all your tasks have been executed at least once, change the frequency view to “Min.” in the visualization settings of your Process Discoverer.

The numbers in your activities now indicate the minimum number of times it has been completed in all process instances.

Temporal Compliance Checking

2

Prescribed time frame: Has the given time frame been adhered to?

Select the “Path” filter and define the steps that must be performed in a defined time frame. Use “Time interval” to then set your temporal figures.

Example:

Compliance rule:
Waiting time between “Inform User” and “Repair (Simple)” should not exceed 1 hours.

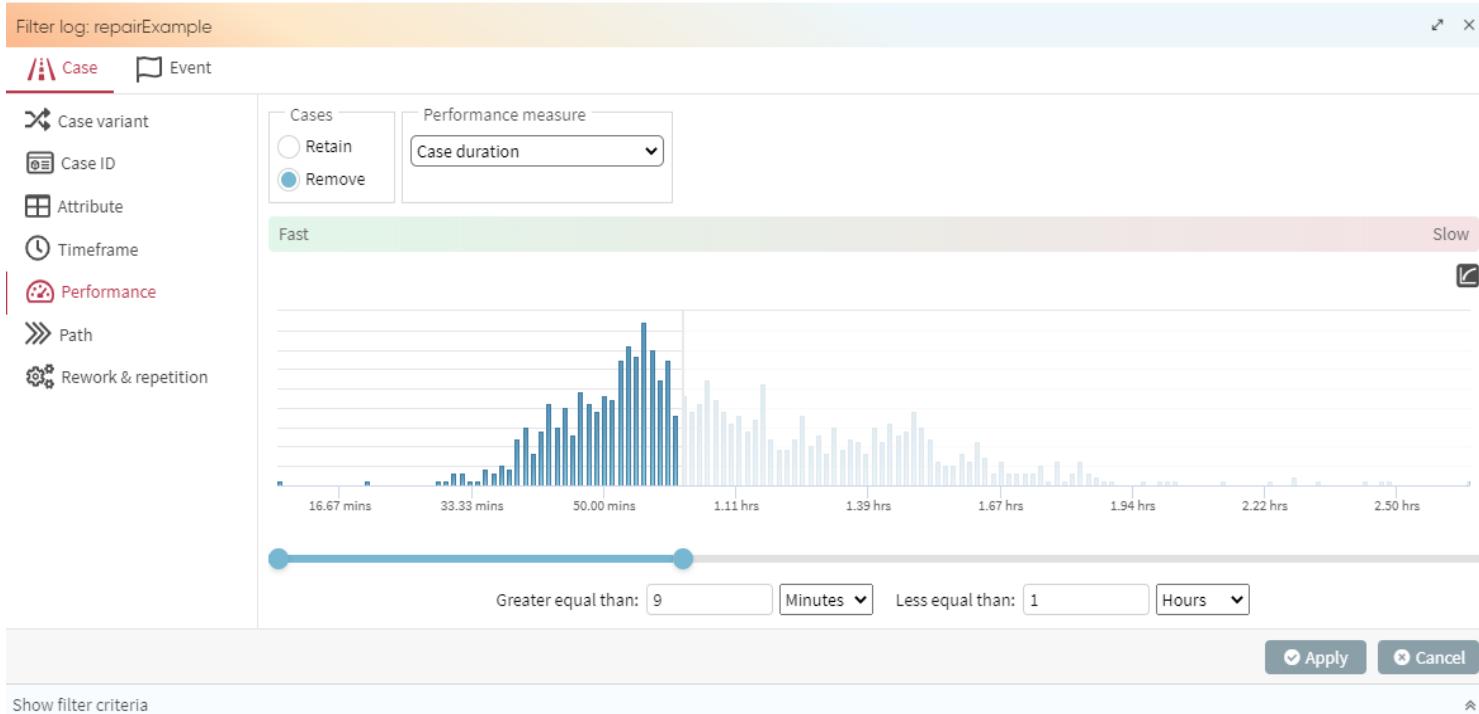


“Time Interval” Filter:
> 1 hour, so that only those cases are retained that do not comply with the compliance rule.

Temporal Compliance Checking

1

Prescribed time frame: Has the given time frame been adhered to?



Performance filter can help to check for end-to-end SLA violations.

For example, retain only those cases that gets completed in at most 1 hour.

Resource Compliance Checking

1

Segregation of duties: Did everyone obey the segregation of duties?

The screenshot shows a process mining interface with a filter dialog open. The filter criteria are set to 'Path' and 'Same Resource'. The 'Same Resource' checkbox is checked. The 'From' and 'To' fields are set to '1 / 9'. The 'Time interval' dropdown is set to 'hours' with values '≥ 1' and '≤ 1'. The 'Case' tab is selected, showing options like 'Case variant', 'Case ID', 'Attribute', 'Timeframe', 'Performance', and 'Path'. The 'Path' tab is selected, listing activities: Analyze Defect, Archive Repair, Inform User, Register, Repair (Complex), Repair (Simple), Restart Repair, Test Repair, and Start!. The 'Event' tab is also visible. At the bottom, there are 'Apply' and 'Cancel' buttons.



To check whether the segregation of duties has been breached, select a “Path” filter. Now, set another filter which indicates that the same resource must have completed the tasks (e.g. creating & approving an invoice)

The close-up shows the 'Same Resource' checkbox checked. Below it is a dropdown menu set to 'Resource'.

Resource Compliance Checking

1

Unexpected resource assignments

Filter log: repairExample [?](#)

Case Event

Case variant

Case ID

Attribute

Timeframe

Performance

Path

Rework & repetition

Cases

Primary attribute

Activity

Secondary attribute

Event attribute Case attribute

Resource

Matching

Any value All values

Activity	Cases	Frequency	Total
Analyze Defect	1104	100%	1104
Register	1104	100%	1104
Inform User	1102	99.819%	1102
Test	1102	99.819%	1508

Resource (1 / 6)	Occurrence	Frequency
Tester6	195	17.663%
Tester1	193	17.482%
Tester3	191	17.301%
Tester5	164	14.855%
Tester4	159	14.402%

Show filter criteria

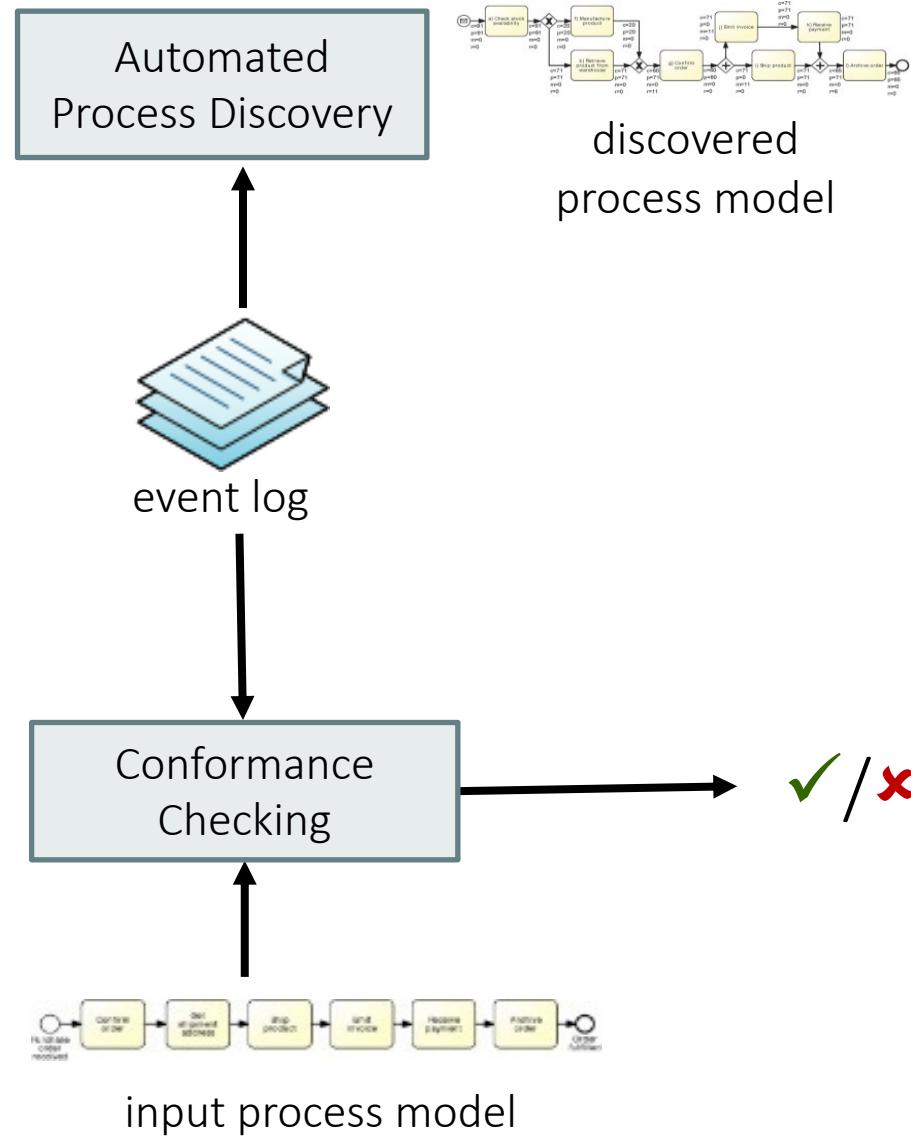
Apply Cancel

User the Attribute filter with two attributes (primary and secondary) to retrieve cases where a given activity was performed by a given resource (or any other combination of attributes)

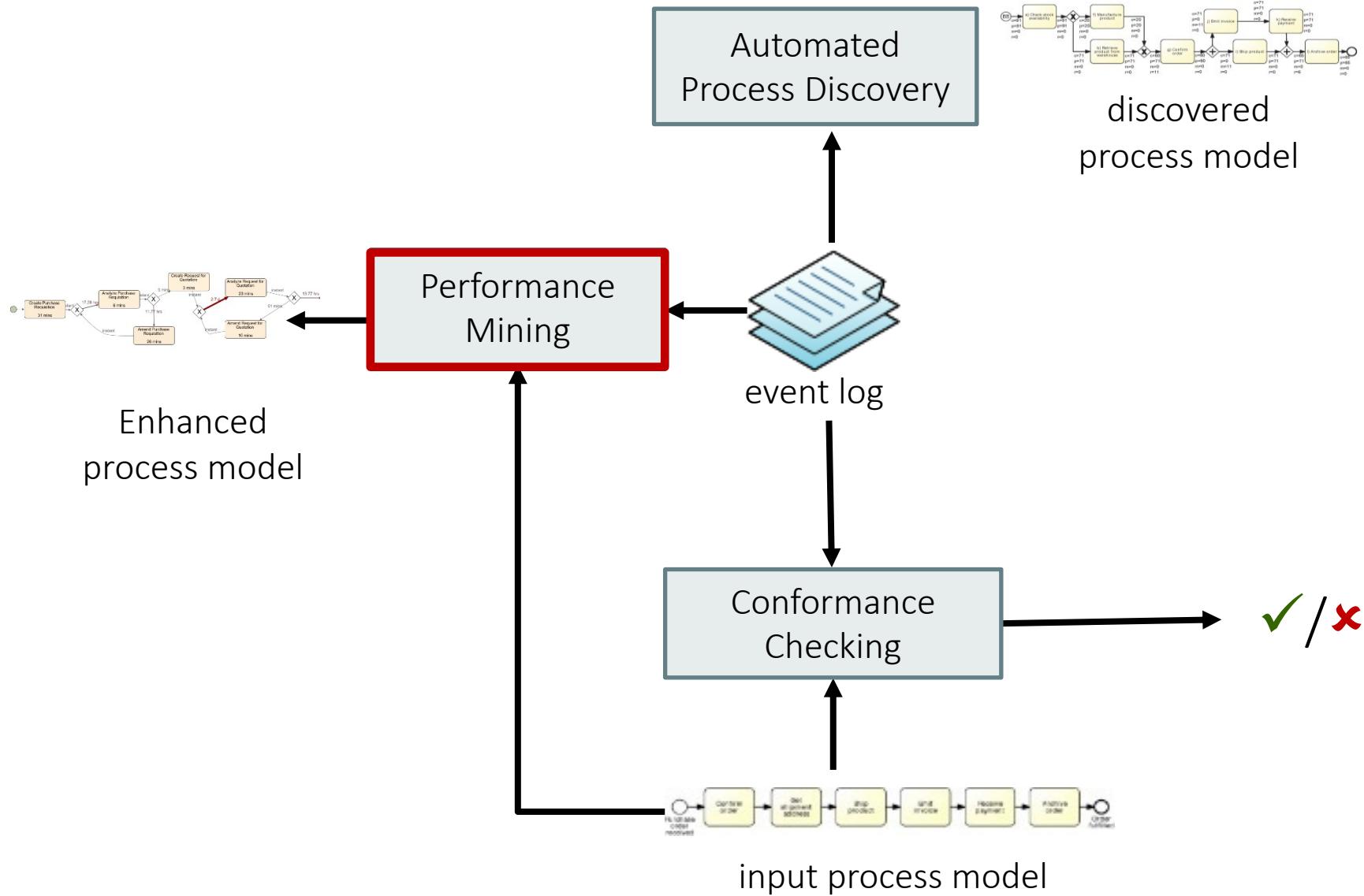
Conformance Checking: Analysis Template

What?	How?
Flow compliance checking <ul style="list-style-type: none"> 1. Check skipped mandatory activities 2. Check forbidden repetitions. 3. Check activity exclusion or co-occurrence relations 	Identify skipped activities in the process map or BPMN model Visualize process map with the min/max frequency metrics Visualize min/max statistics in Activity tab of the dashboard (EE) Use Path filter (directly-follows or eventually-follows)
Temporal compliance checking Check temporal constraints violations	Use performance filter (to check for end-to-end SLA violations) Use Path filter (eventually-follows) with a duration constraint to identify violations of temporal constraints between two activities
Resource compliance checking Check four-eyes principle or same-case-handner principle Check if an activity has been performed by a resource who should not do so.	Use Path filter (eventually-follows) and require the same Resource (or different resources) Use the Attribute filter and search with “Activity” as the primary attribute and “Resource” as secondary attribute to check if a resource performed an activity they were not supposed to do.

Process Mining



Process Mining



Process performance mining: Analytics & Analysis types

Extract process execution performance in an aggregate or detailed level (i.e. at the level of individual activities and directly-follows relations)

Analytics:

- Duration-enhanced process maps (activity, resource perspective...)
- Duration-enhanced BPMN models
- Process performance dashboards

Process performance mining: Analytics & Analysis types

Extract process execution performance in an aggregate or detailed level (i.e. at the level of individual activities and directly-follows relations)

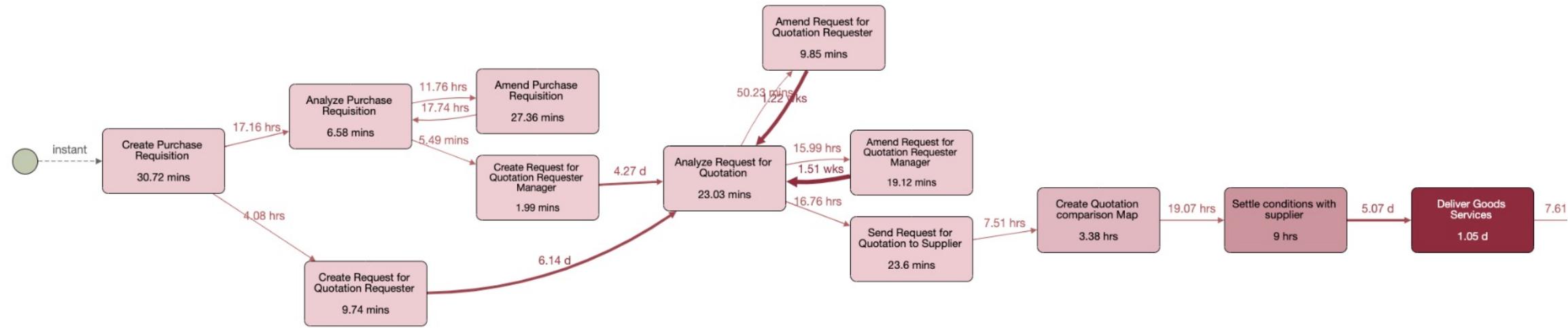
Analytics:

- Duration-enhanced process maps (activity, resource perspective...)
- Duration-enhanced BPMN models
- Process performance dashboards

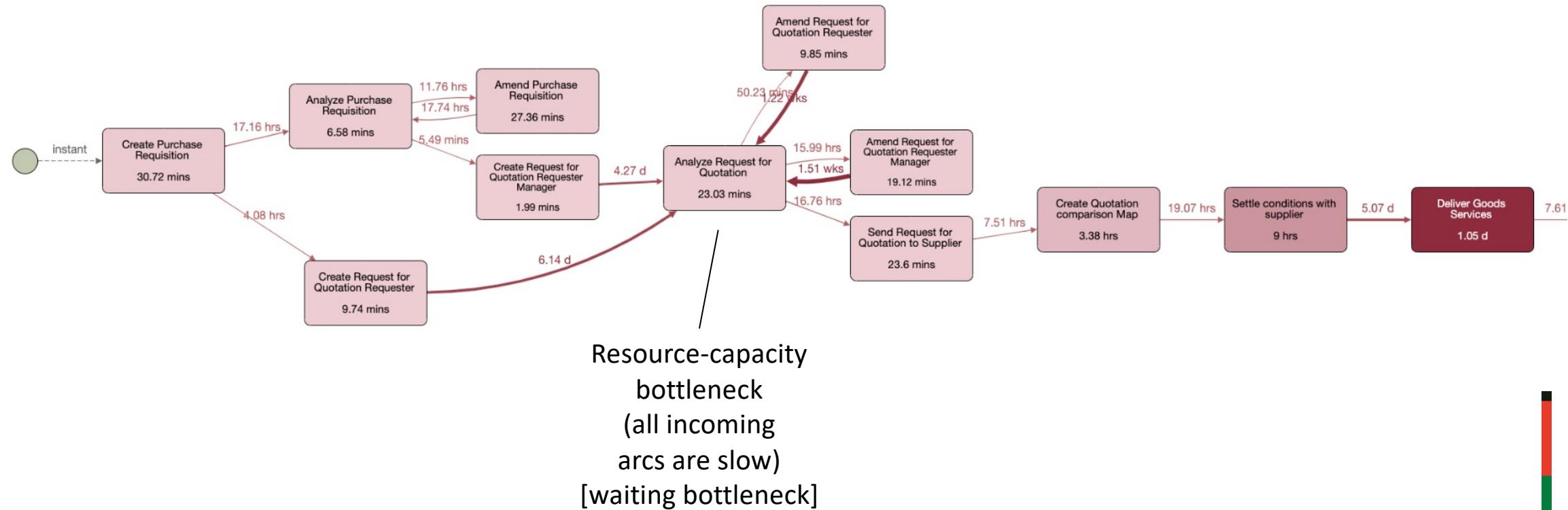
Analysis types:

- Bottleneck analysis
- Workload analysis
- Rework analysis

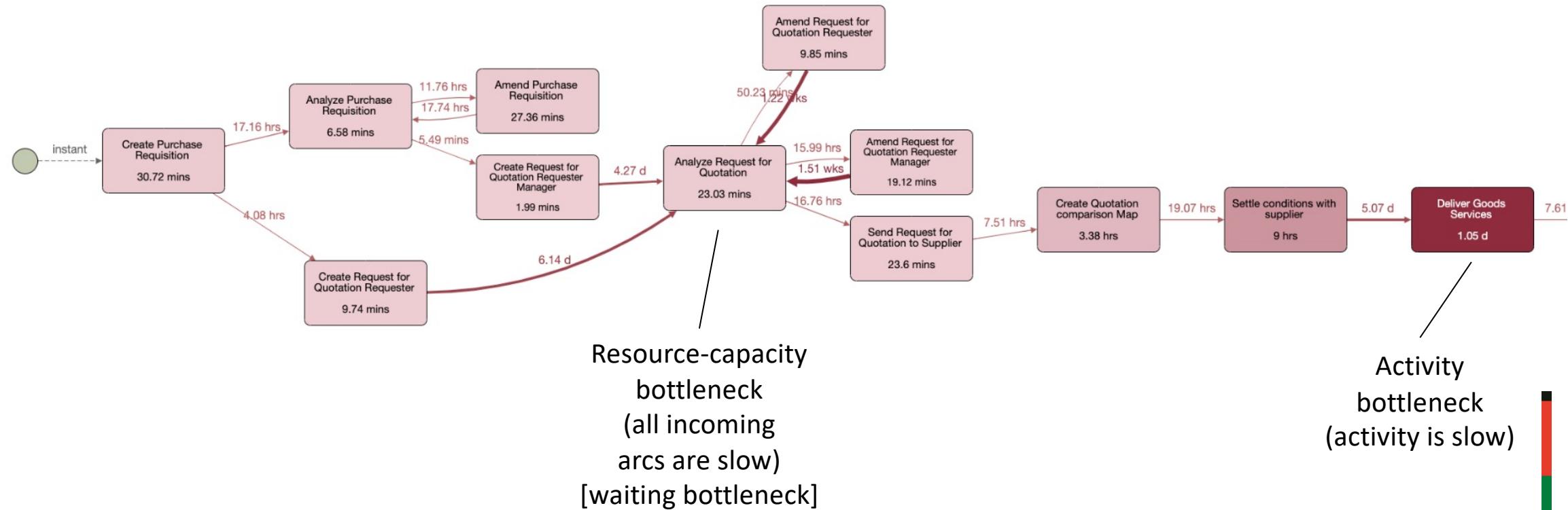
Bottleneck analysis in the Activity perspective



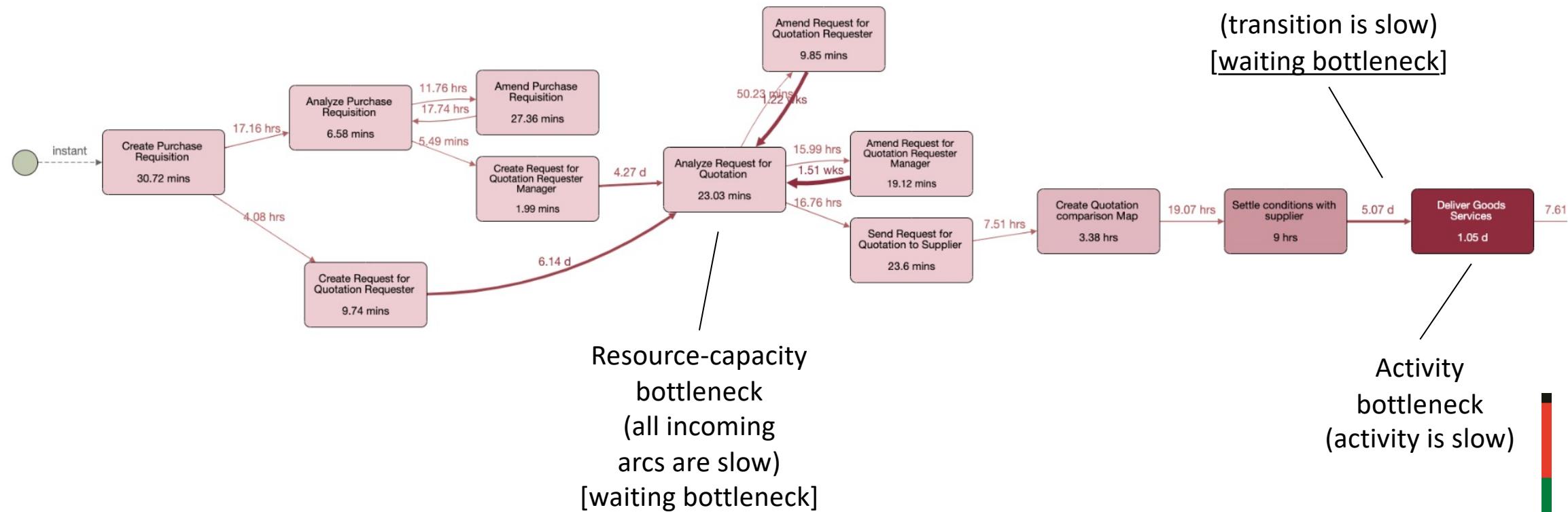
Bottleneck analysis in the Activity perspective



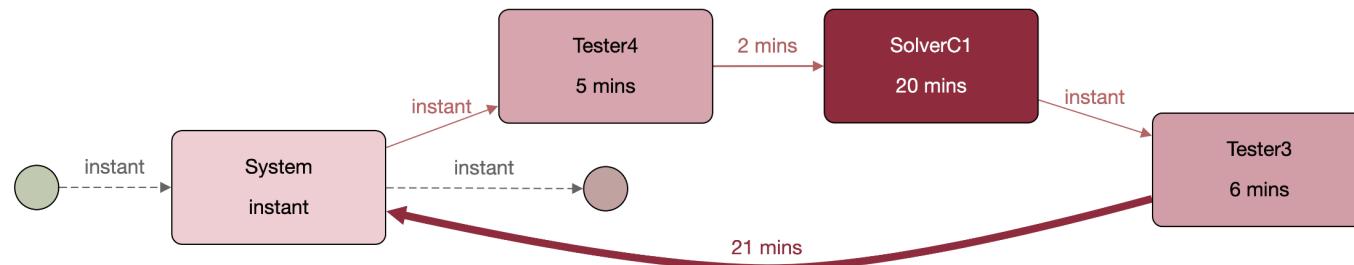
Bottleneck analysis in the Activity perspective



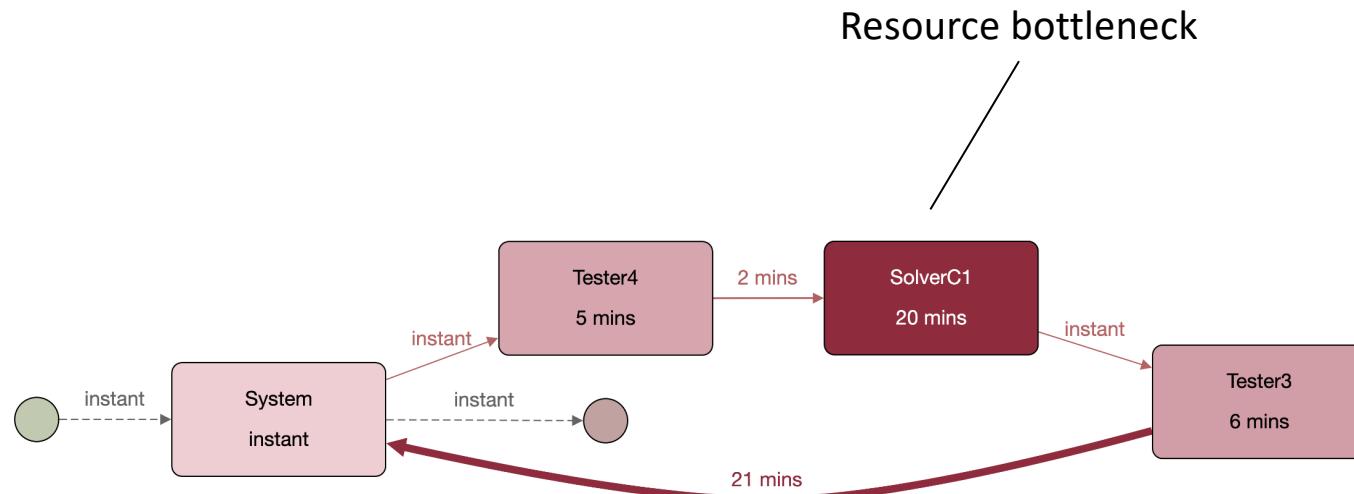
Bottleneck analysis in the Activity perspective



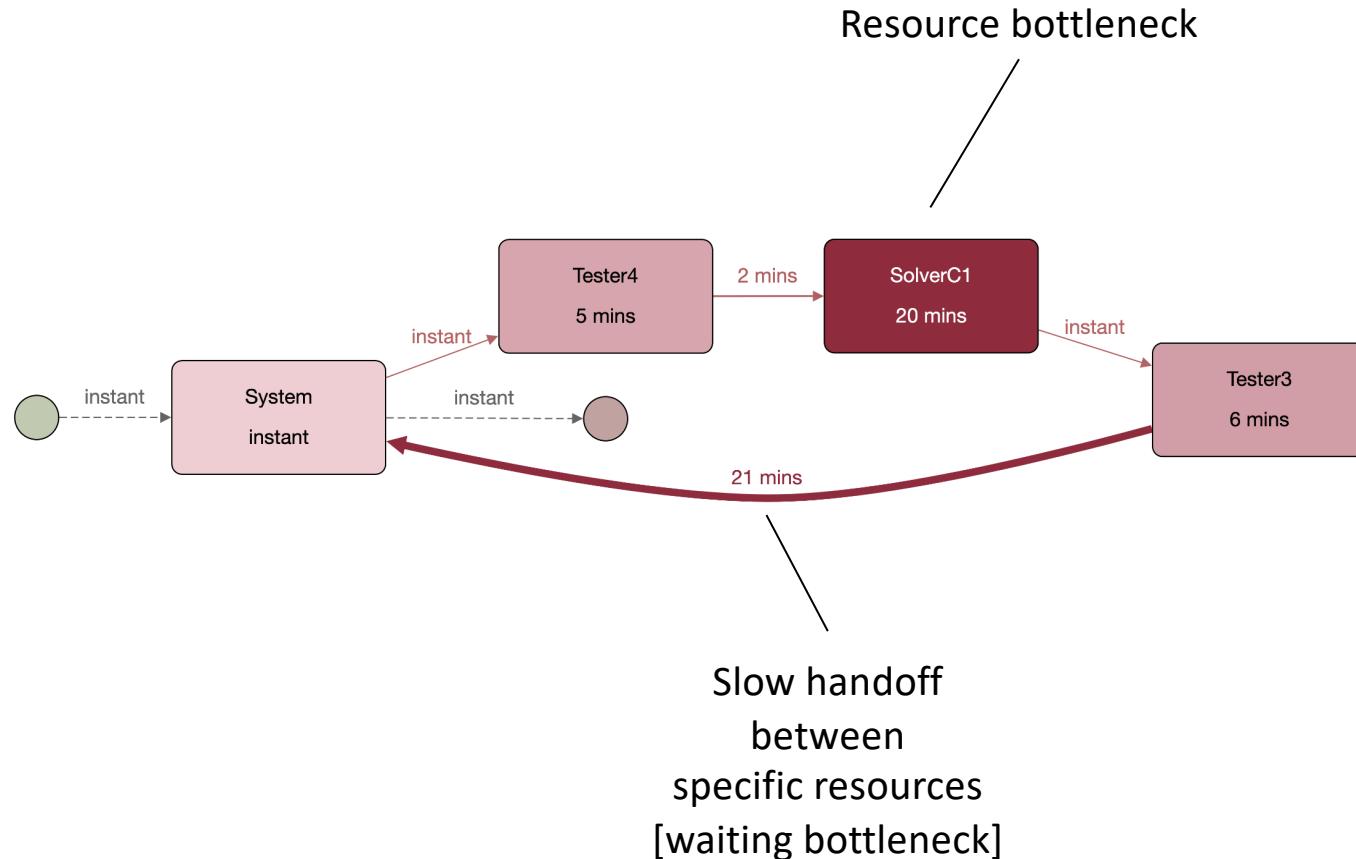
Bottleneck analysis in the Resource perspective



Bottleneck analysis in the Resource perspective

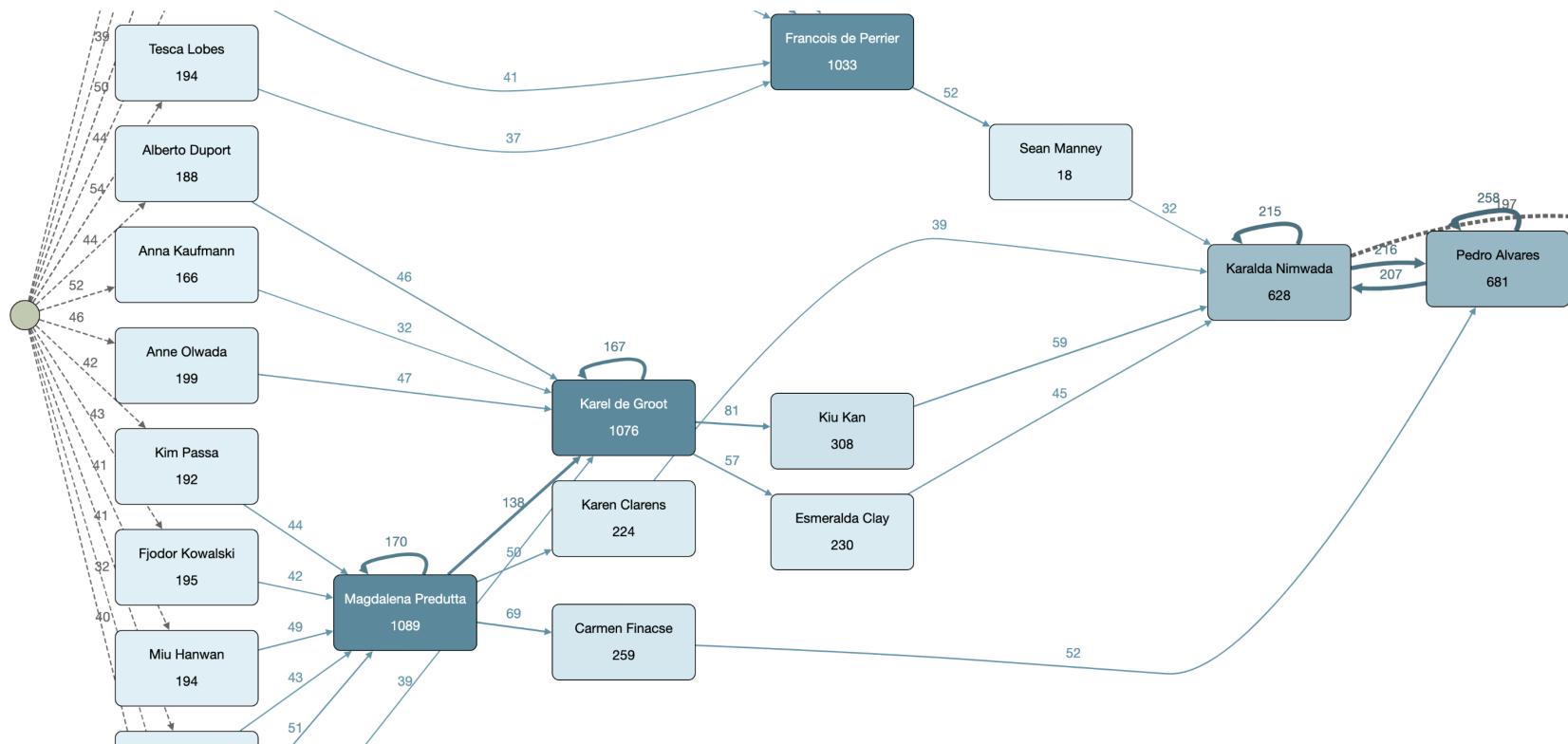


Bottleneck analysis in the Resource perspective



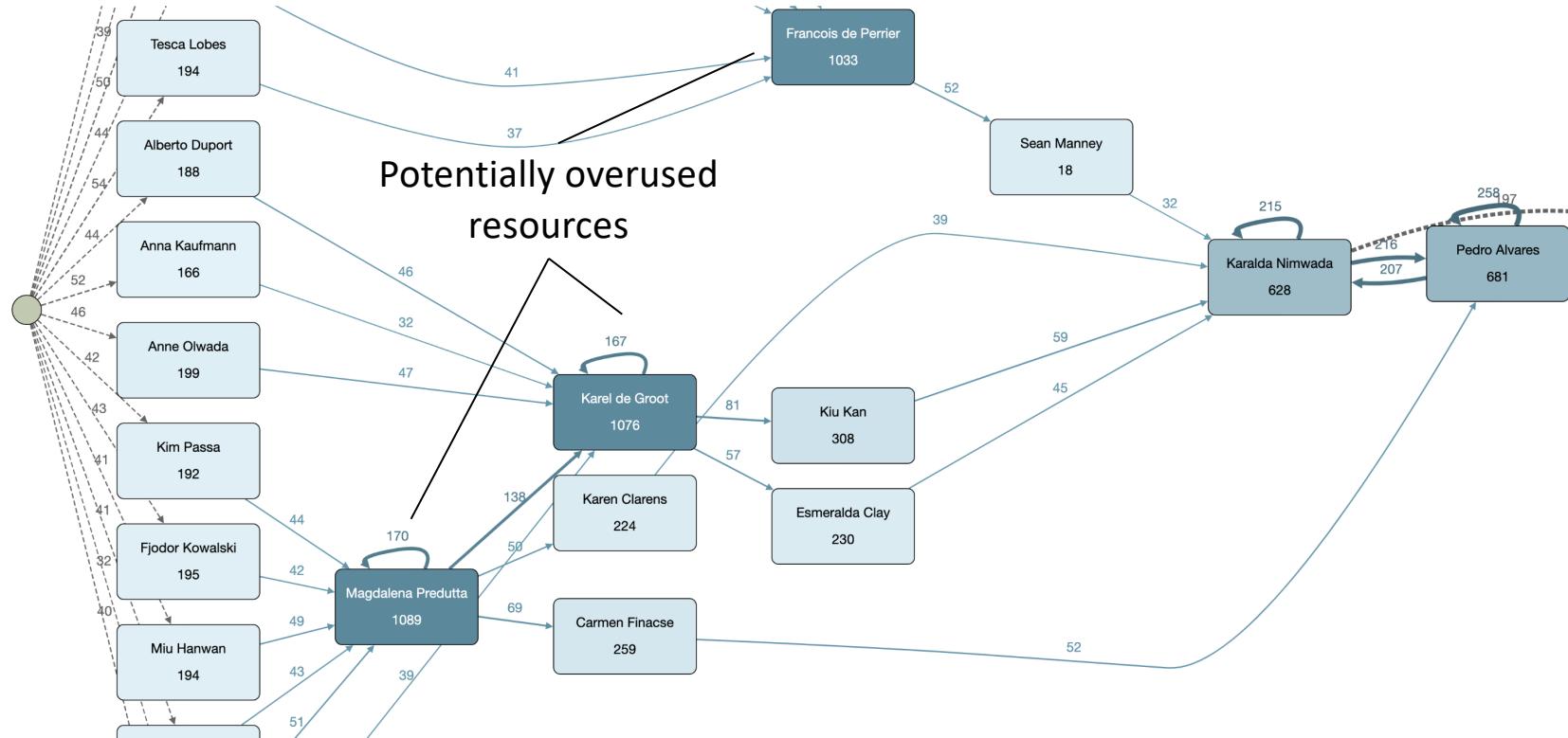
Workload analysis: overused vs underused resources

- **Potentially overused resources** are those with high total frequency and/or high total duration
 - **Potentially underused resources** are those with low total frequency and/or low total duration



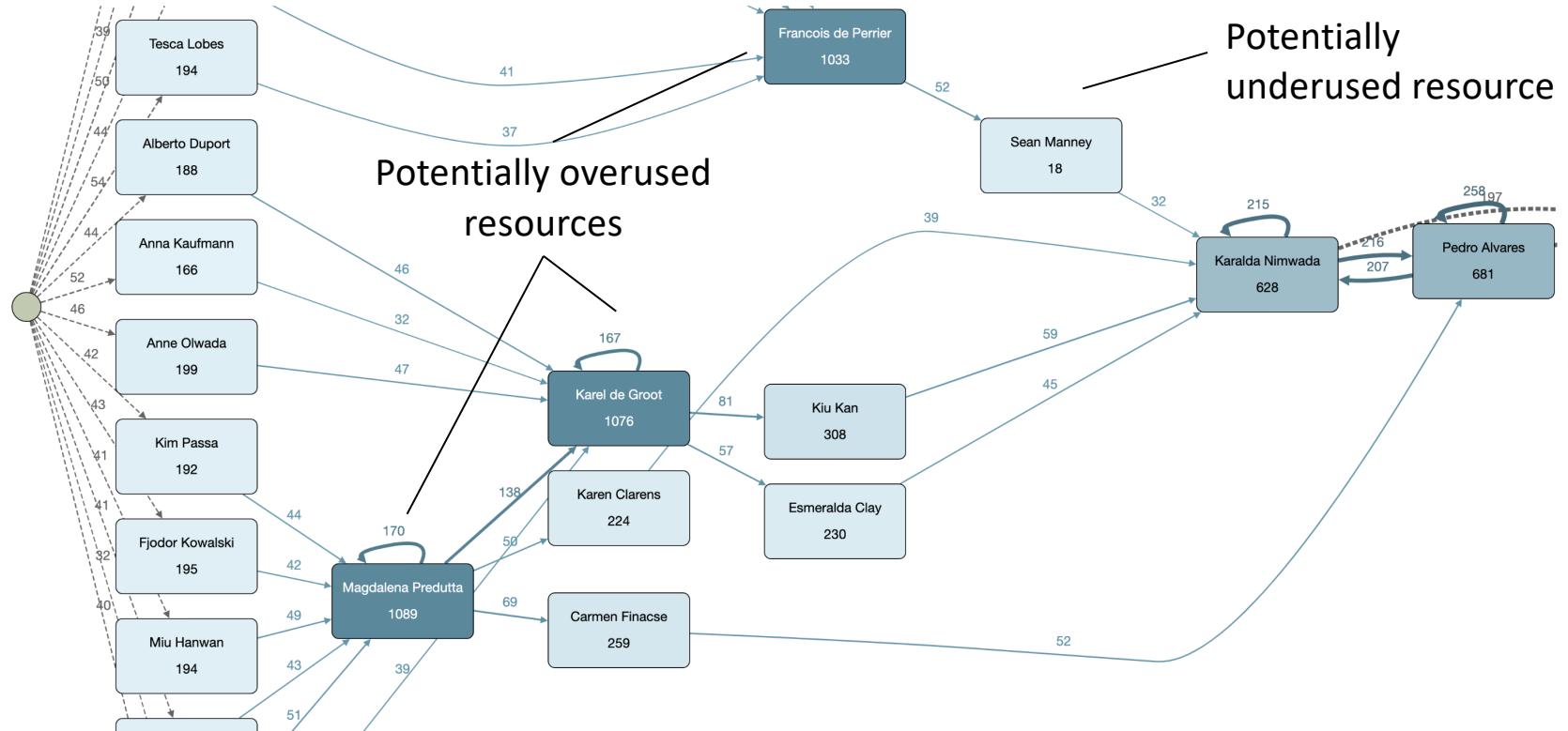
Workload analysis: overused vs underused resources

- Potentially overused resources are those with high total frequency and/or high total duration
- Potentially underused resources are those with low total frequency and/or low total duration



Workload analysis: overused vs underused resources

- Potentially overused resources are those with high total frequency and/or high total duration
- Potentially underused resources are those with low total frequency and/or low total duration



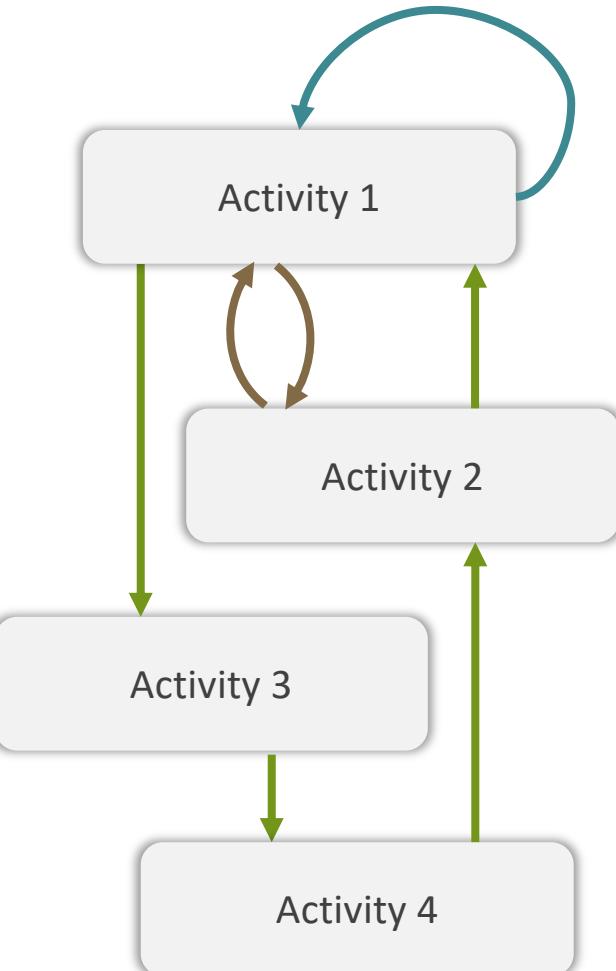
Rework analysis

Rework means repetition of activities within the same case:

- **Self loop:** an activity takes place n times in a row
- **Short loop (“ping-pong effect”):** two activities are repeated one after the other. Note: this may be confused with parallelism
- **Indirect repetition (loop):** several activities are executed several times in sequence.

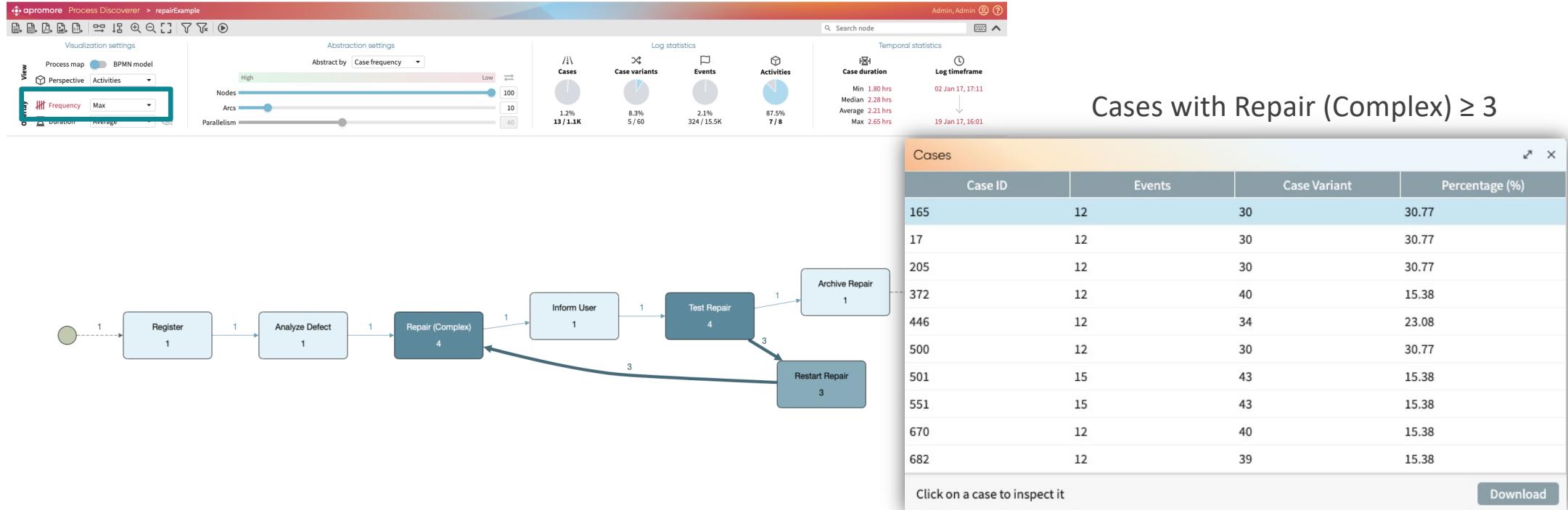
To analyze repetitions, one can answer the following questions:

- (1) What type of repetition is present?
- (2) How often is an activity / series of activities repeated?
- (3) How many cases showcase this repetition? How often does it occur?
- (4) Which cases are affected exactly? And how do they differ from cases where there is no repetition?



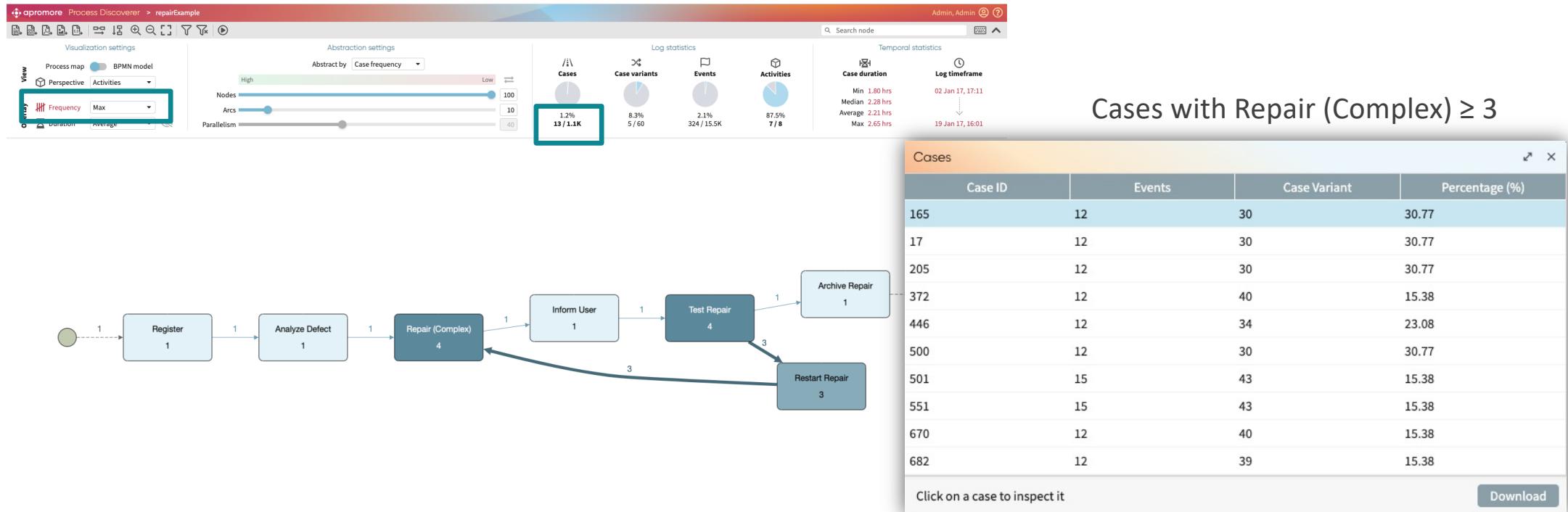
Rework analysis

To assess the impact of rework on process performance (time and cost), we should apply filtering to isolate cases that are affected by rework.



Rework analysis

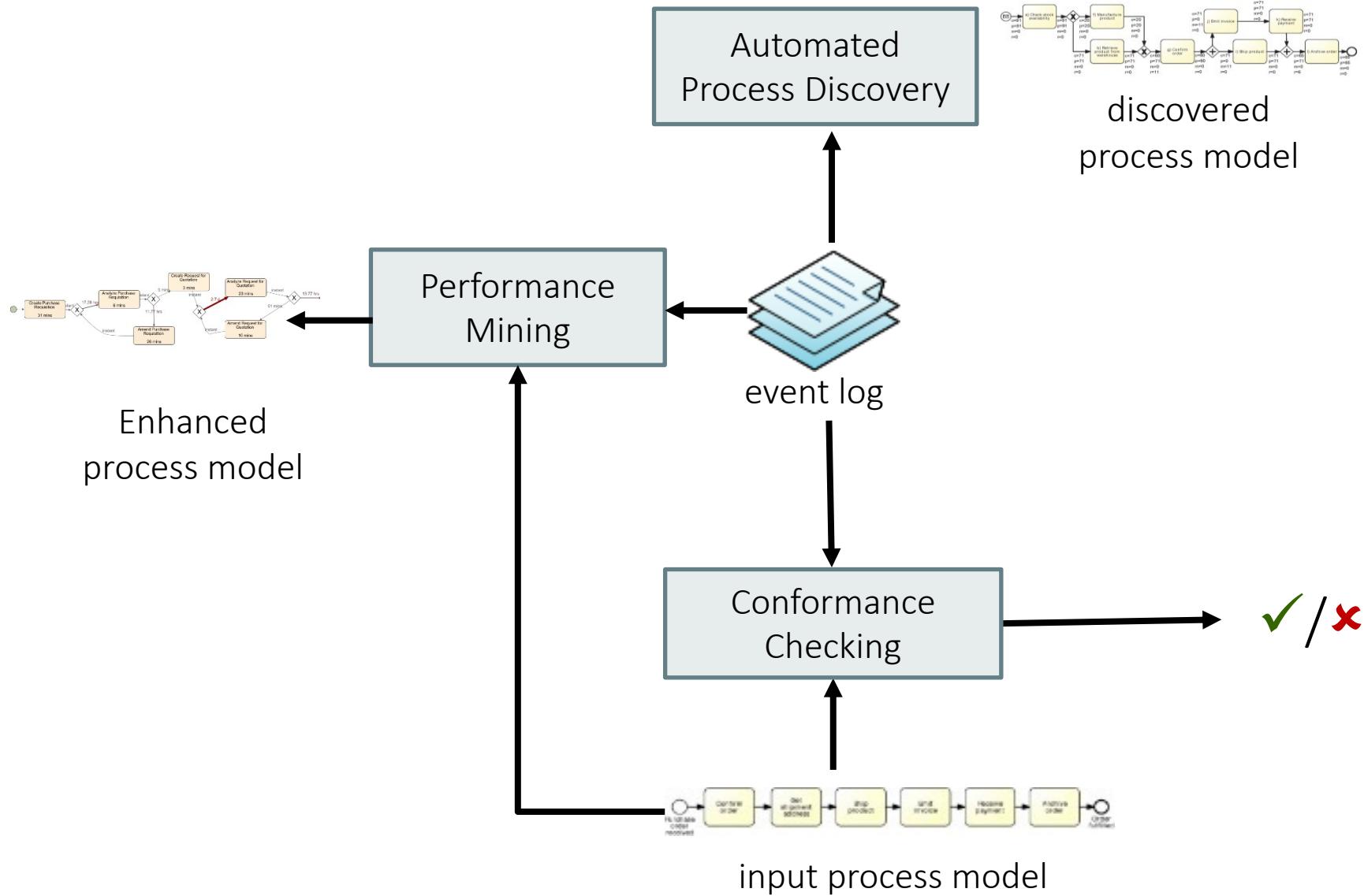
To assess the impact of rework on process performance (time and cost), we should apply filtering to isolate cases that are affected by rework.



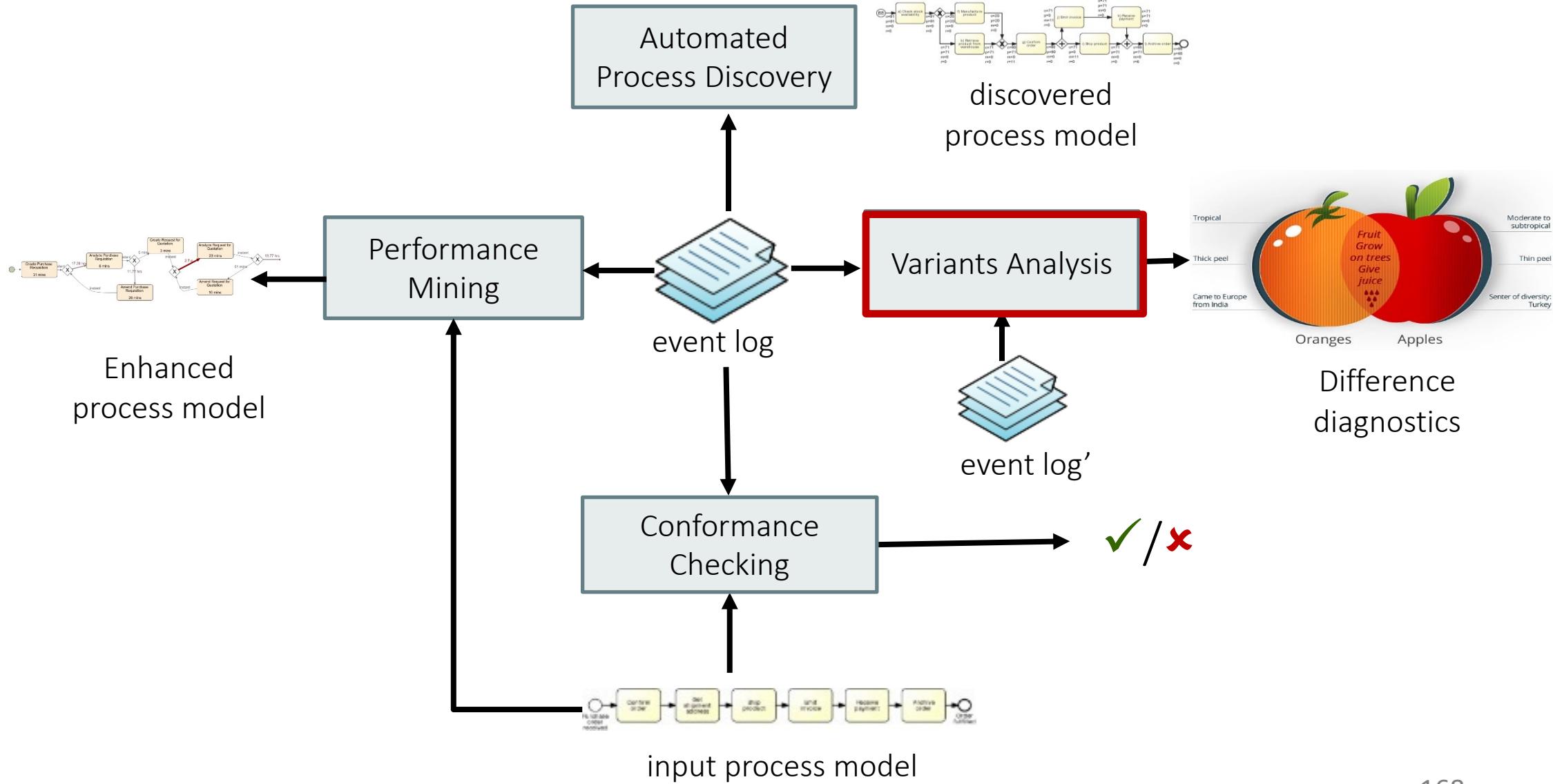
Performance Mining: Analysis Template

What?	How?
Bottleneck analysis <ul style="list-style-type: none"> Identify activity bottlenecks (high-effort activities) Identify waiting bottlenecks (slow handoffs and capacity bottlenecks) Identify resource bottlenecks 	<ul style="list-style-type: none"> In the process map, select the duration overlay (average or total duration): <u>activity bottlenecks</u> have long processing time; <u>slow handoffs between activities</u> are arcs with long waiting times; <u>resource-capacity bottlenecks</u> relate to activities with all/most incoming arcs with long waiting times Switch to the Resources perspective in the map: <u>slow handoffs between resources</u> are arcs with long waiting times; <u>resource bottlenecks</u> are resources with long avg duration (slow resources) Use slider with Average duration to focus on slowest or fastest nodes/arcs Use Activities tab in Dashboard to visualize activity processing times Use Animation plugin to visualize the build-up of bottlenecks.
Workload analysis <ul style="list-style-type: none"> Identify overused vs underused resources/idleness waste Analyze resource workload 	<ul style="list-style-type: none"> In the Resources perspective: <u>overused resources</u> have high total frequency or high “total duration”; <u>underused resources</u> have low total frequency and low <u>total duration</u> Use also the Resources tab in Dashboard to visualize resource workload distribution Use Filtering to focus on cases where the resource(s) with the highest workload are involved.
Rework analysis Identify defects/errors and rework waste	<ul style="list-style-type: none"> Use the Max frequency overlay in the process map to visualize the max amount of repetitions of each activity / transition Use the Case > Rework & repetition filter to retain cases with activity repetition; use the Case > Path filter with eventually-follows relation between an activity and itself to isolate cases with specific loops Use the chart in the Activities Dashboard to visualize the distribution of max activity frequencies per case.

Process Mining

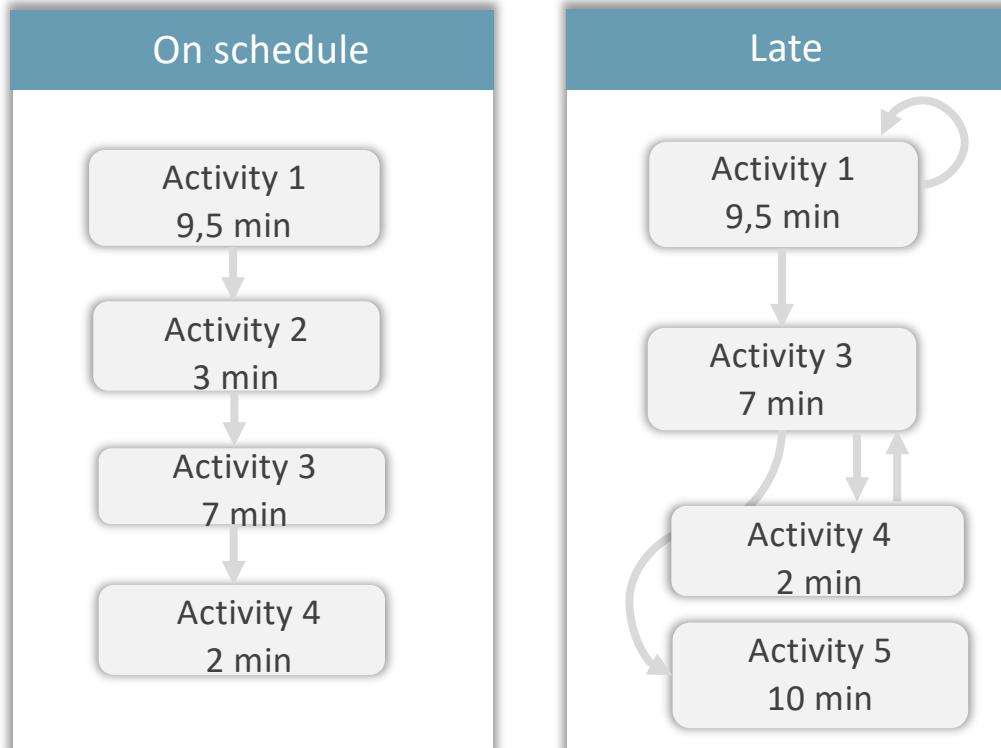


Process Mining



Variant analysis

Given two or more logs representing variants of the same process, find the *differences* and *root causes* for variation or deviance between the logs.



The comparison of process variants offers insights into the the “why”: Why do certain process cases differ from others?

- Why some take longer than others? (fast vs. slow cases)
- Why some end successfully, while others don’t? (positive vs. negative outcomes)
- Why is employee XY more efficient and more successful when it comes to their activities than others? (resource XY vs. remaining resources)

Variant analysis: drivers

Performance

Identify and compare variants based on performance measures and their targets (e.g. slow vs fast cases)

Logical

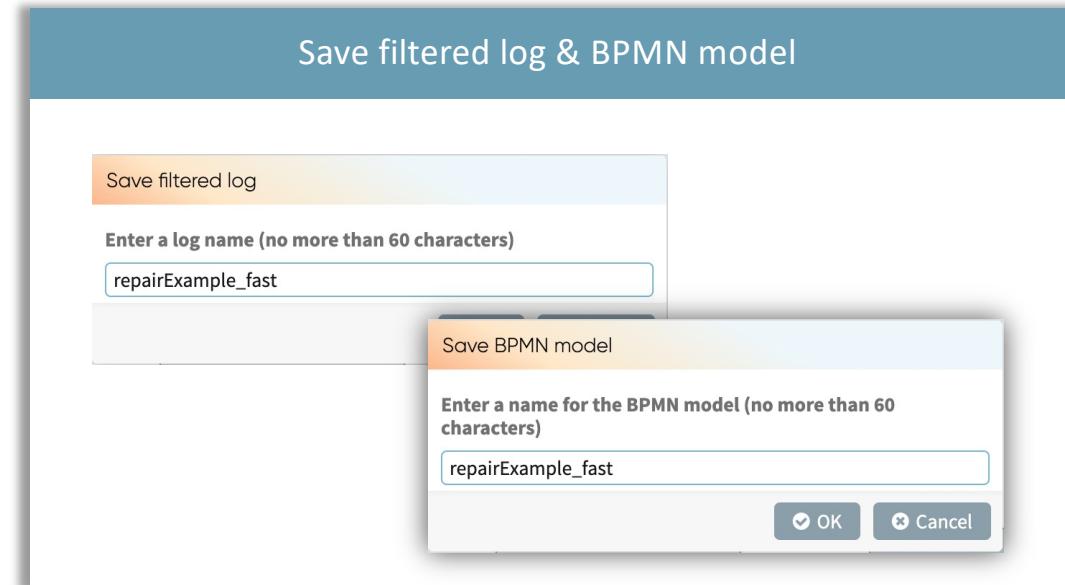
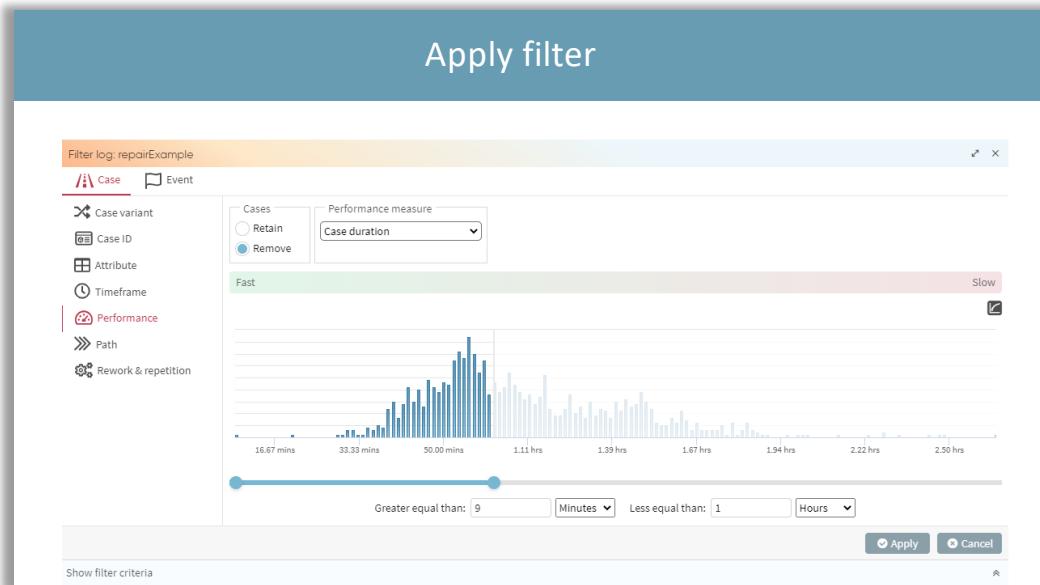
Identify and compare variants based on case attributes (e.g. product type, customer segment, geographic area, loan amount, claim outcome)

Temporal

Identify and compare variants based on different log timeframes (e.g. before and during COVID)

Variant analysis: approach

1. Apply the same filter one per variant, to isolate the respective sublog:
 - **Performance:** use Case > Performance filter
 - **Logical:** use Case > Attribute filter
 - **Temporal:** use Case > Timeframe filter
2. Save filtered process map and BPMN model back to the repository
3. Analyze the variants using one or more analysis techniques.



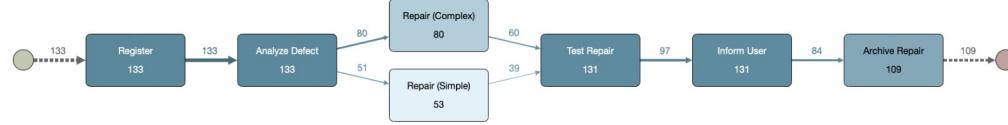
Variant analysis: comparison via maps/models

Visual comparison of process maps/BPMN models, to identify structural differences:

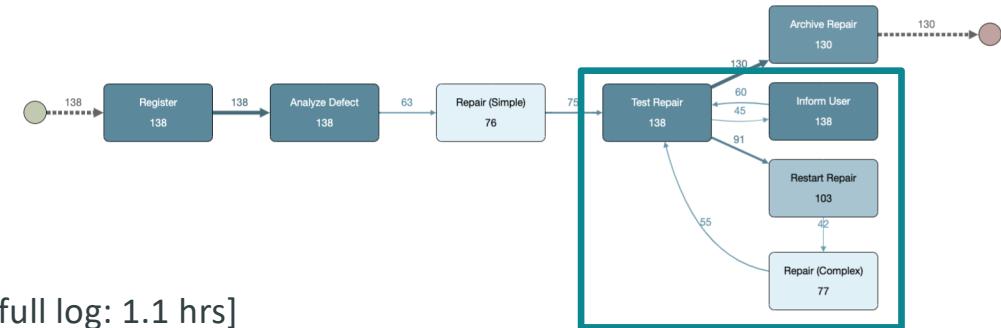
- Flow
- Frequency & rework
- Bottlenecks
- etc.

Consider different perspectives (Activities, Resources...)

Variant 1: “fast” (up to 45 min)



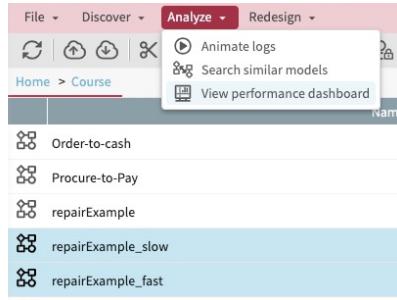
Variant 2: “slow” (more than 1.5 hrs)



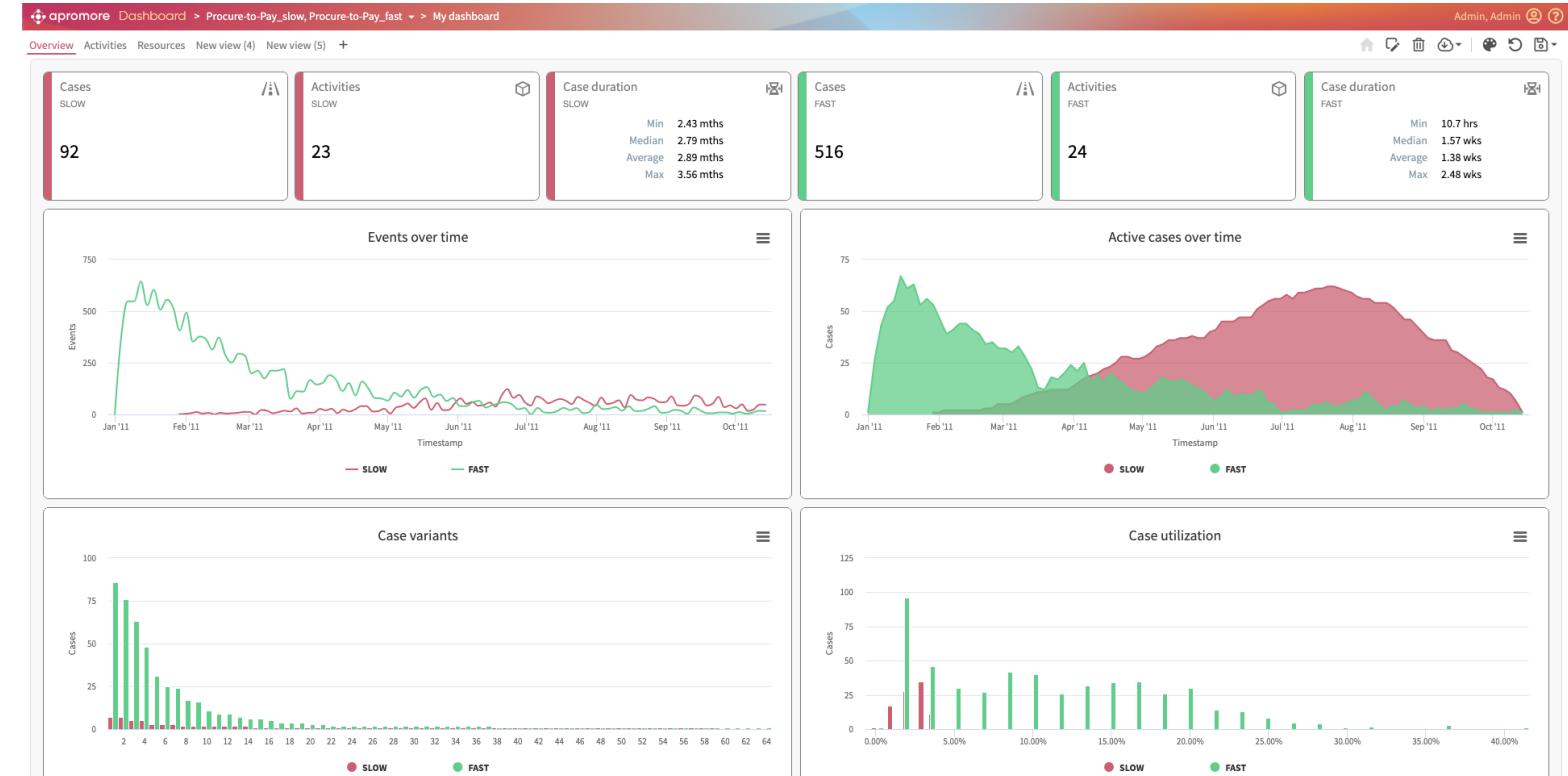
[Avg duration for full log: 1.1 hrs]

Variant analysis: comparison via dashboard

Statistical comparison via dashboard charts and tables. Variants are color-coded (e.g. red for slow, green for fast cases)



Select your filtered logs and launch the Dashboard. All dashboard functionalities are also available when comparing multiple variants.



Variant Analysis Template

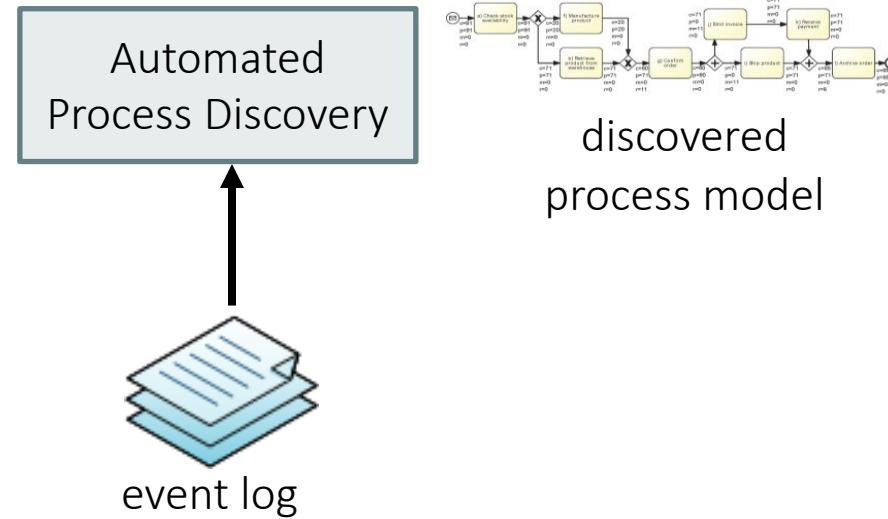
What?	How?
Flow comparison Identify notable diffs in activity flow	<ul style="list-style-type: none">Discover a BPMN process model from the log of each variant. Visualize them side-by-side. Look for the gateways around each task to detect for example situations where two tasks are in parallel in one variant, but sequential in the other.
Frequency & rework comparison Compare activity or handoff frequency and rework	<ul style="list-style-type: none">Compare the maps of the process variants side-by-side (as above), but using the frequency overlay instead of the duration overlay. This allows us to detect the most frequent transitions between tasks. Look also for “thick” loop-backs.Use a multi-log dashboard to compare frequencies of activities, resources or other attributes.
Bottleneck comparison Compare the location and magnitude of bottlenecks across variants	<ul style="list-style-type: none">Use filtering interface to retain all traces that fulfill the condition to belong to a variant (then repeat for the other).Open the maps of the process variants side-by-side and compare the dependencies (arcs) with the highest waiting times. If the maps are too complex, use the abstraction slider with the abstraction metric “average duration” and the ordering from slow (right) to fast (left). Use the arc slider to retain the slowest dependencies.Compare activity durations using side-by-side comparison of process maps or a multi-log dashboard (open the logs of 2+ variants in a single dashboard).Consider comparing the two variants using the “resource” perspective to identify hand-offs between resources.

Summary

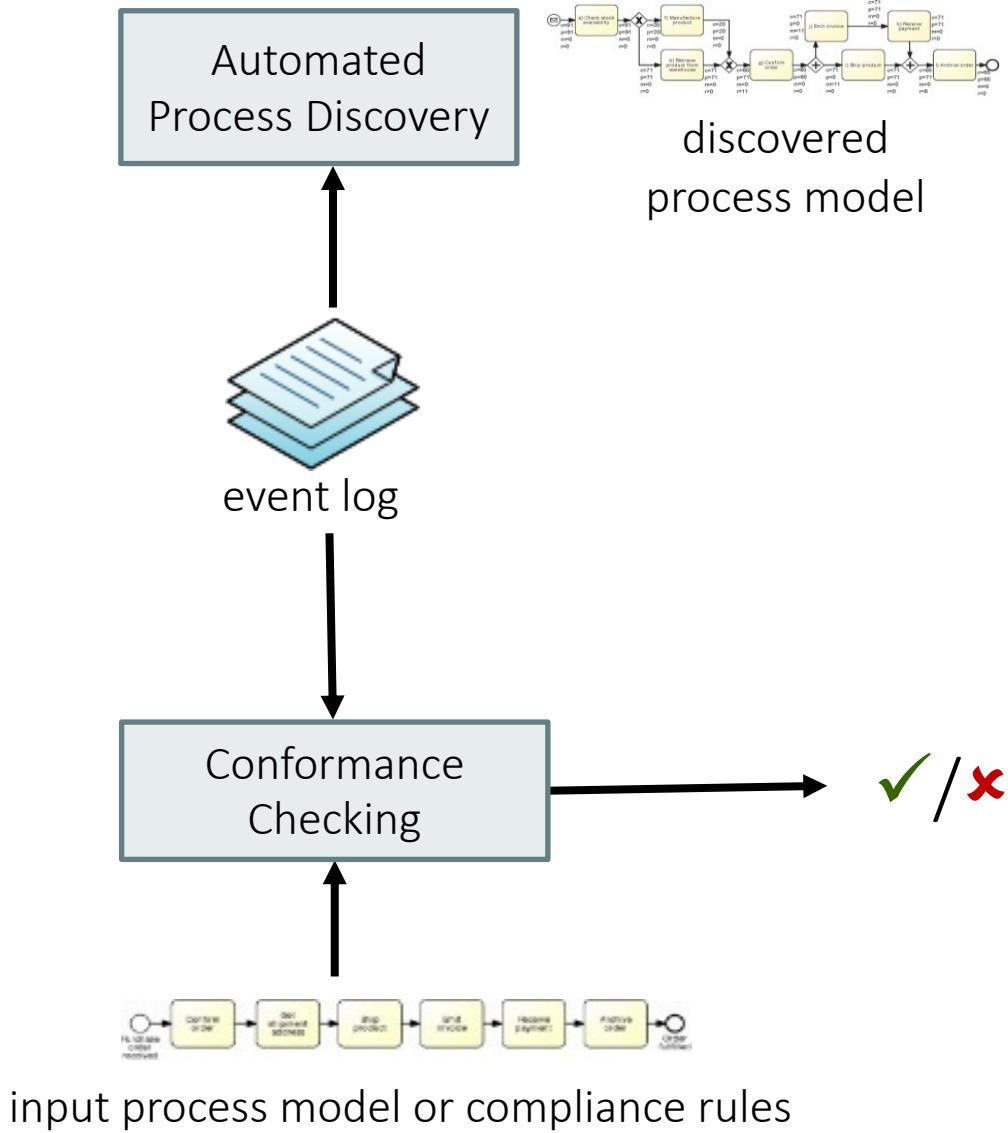


event log

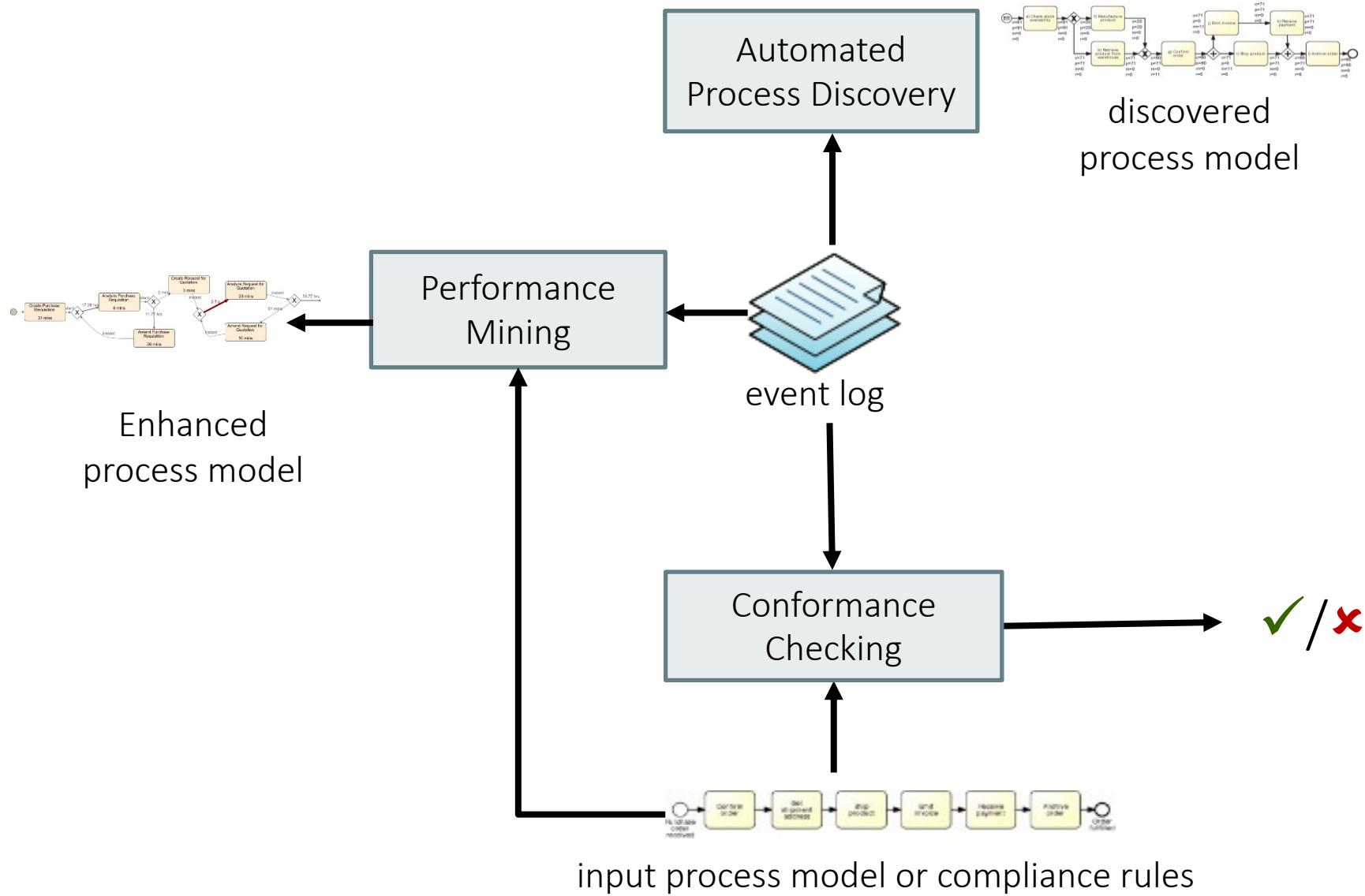
Summary



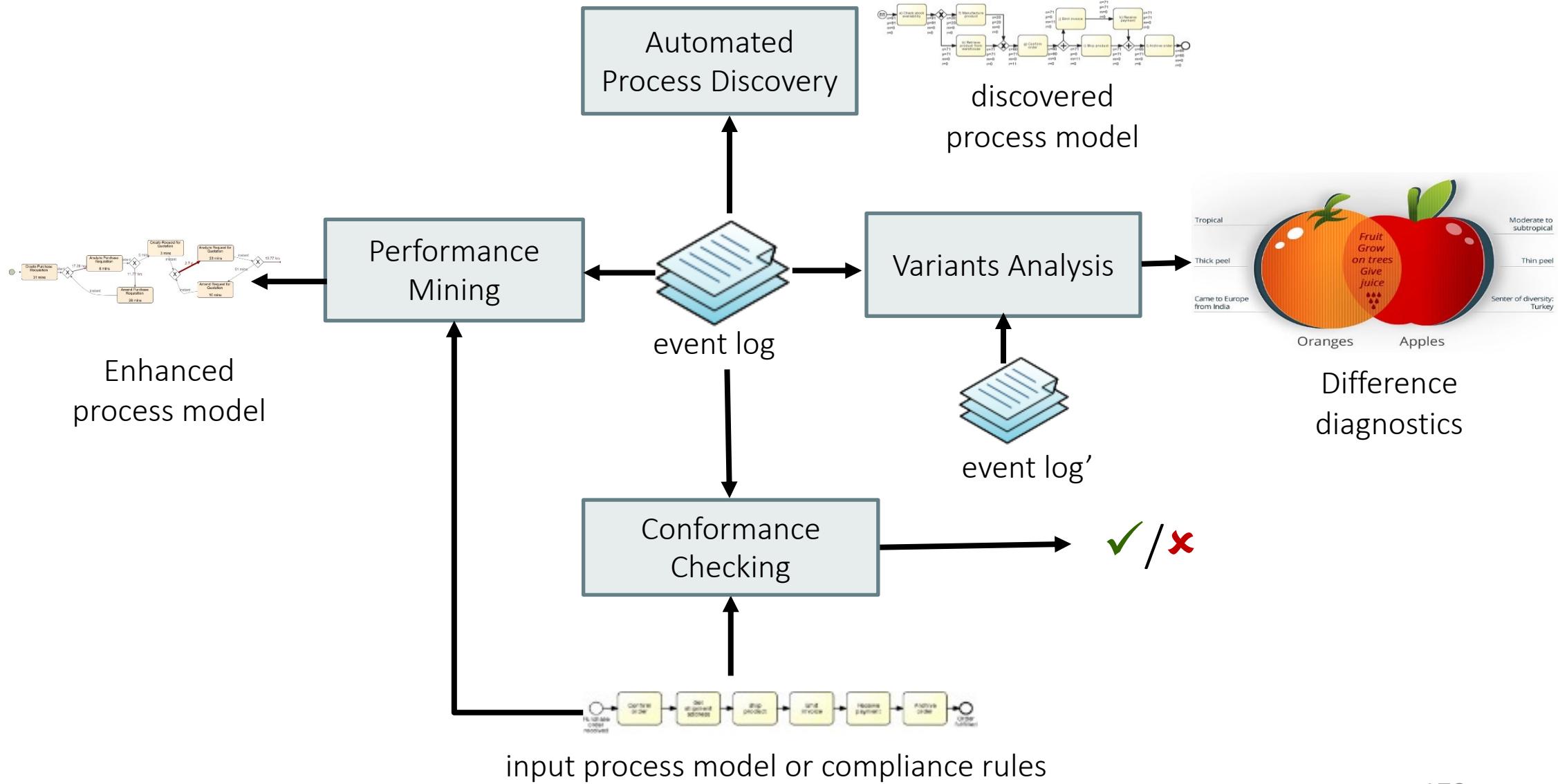
Summary



Summary



Summary



Summary: Four Process Mining Capabilities

Summary: Four Process Mining Capabilities

The bulk of process mining techniques can be divided into four capabilities.

EVENT LOGS

Summary: Four Process Mining Capabilities

The bulk of process mining techniques can be divided into four capabilities.

EVENT LOGS

Automated Process Discovery

By using the automated process discovery, an actual process model is created from the event log. This represents the current sequence of work steps and does not necessarily have to correspond to the target process.



Summary: Four Process Mining Capabilities

The bulk of process mining techniques can be divided into four capabilities.

EVENT LOGS

Automated Process Discovery

By using the automated process discovery, an actual process model is created from the event log. This represents the current sequence of work steps and does not necessarily have to correspond to the target process.



Conformance Checking

Conformance checking provides a comparison of the actual and the target process and thus enables the detection of differences and deviations.



Summary: Four Process Mining Capabilities

The bulk of process mining techniques can be divided into four capabilities.

EVENT LOGS

Automated Process Discovery

By using the automated process discovery, an actual process model is created from the event log. This represents the current sequence of work steps and does not necessarily have to correspond to the target process.



Conformance Checking

Conformance checking provides a comparison of the actual and the target process and thus enables the detection of differences and deviations.



Variant Analysis

Process outcomes often differ in practice. But why? The variant analysis enables a simple comparison of several process variants and serves as the basis for process optimization.



Summary: Four Process Mining Capabilities

The bulk of process mining techniques can be divided into four capabilities.

EVENT LOGS

Automated Process Discovery

By using the automated process discovery, an actual process model is created from the event log. This represents the current sequence of work steps and does not necessarily have to correspond to the target process.



Conformance Checking

Conformance checking provides a comparison of the actual and the target process and thus enables the detection of differences and deviations.



Variant Analysis

Process outcomes often differ in practice. But why? The variant analysis enables a simple comparison of several process variants and serves as the basis for process optimization.



Performance Mining

By applying performance mining techniques, e.g. nodes or cycle times can be highlighted in color to simplify the subsequent analysis.



Acknowledgements

- All material comes from Marlon Dumas, Marcello La Rosa, Jan Mendling, Hajo A. Reijers, authors of the “Fundamentals of Business Process Management” book.