



2006 LHC DAYS IN SPLIT
2 - 7 October 2006



Application of Expert Systems Technology for the CMS Event Filter Farm



Kristina Marasović
University of Split, Croatia

E-mail: Kristina.Marasovic@pmfst.hr

Kristina.Marasovic@cern.ch

WWW: <http://www.pmfst.hr/~kim/>





Presentation Agenda

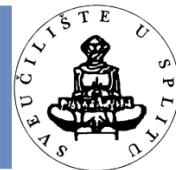


- Event Filter Farm – Towards Problem Solving Research
- Expert Systems Technology to the Rescue
 - When to use
 - Rule-Based Expert Systems
 - Jess: Expert System Shell
- Problem Solver
 - How Stuff Works
 - Performance Analysis
- Summary



Event Filter Farm

Towards Problem Solving Research



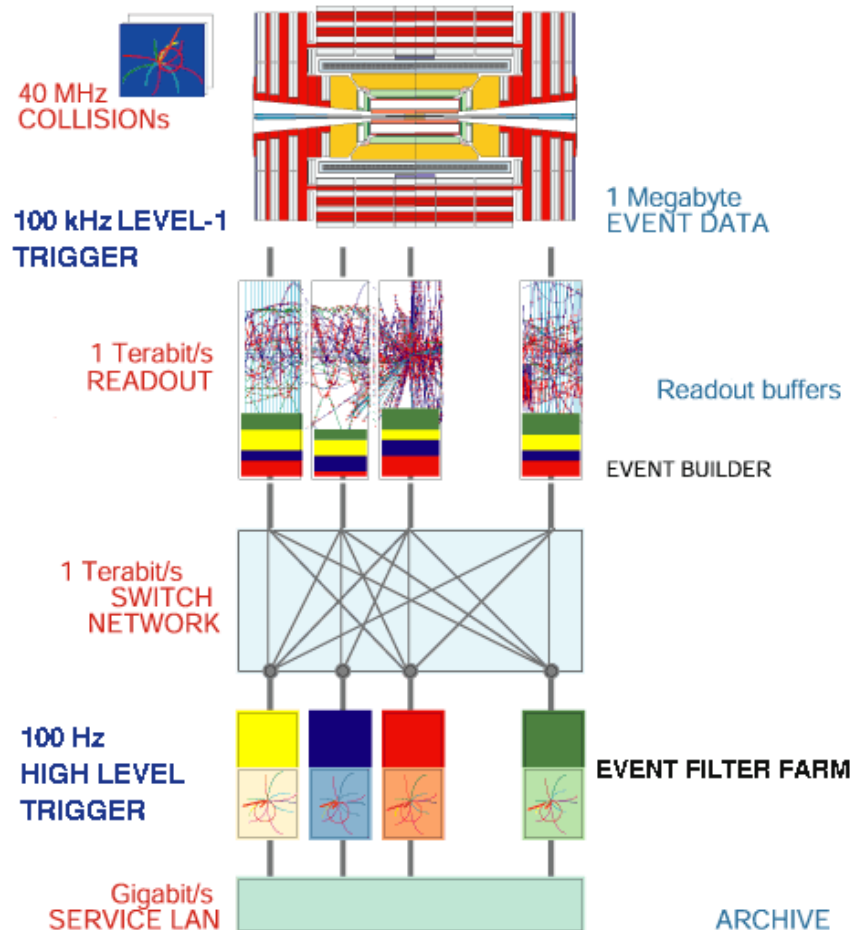
- **Trigger and Data Acquisition System (TriDAS)**

- Collects electrical signals from all subparts of the CMS detector
- Carries out the on-line data filtering process in two steps
 - Level-1 Trigger
 - High-Level Trigger (HLT)

- **Event Filter Farm**

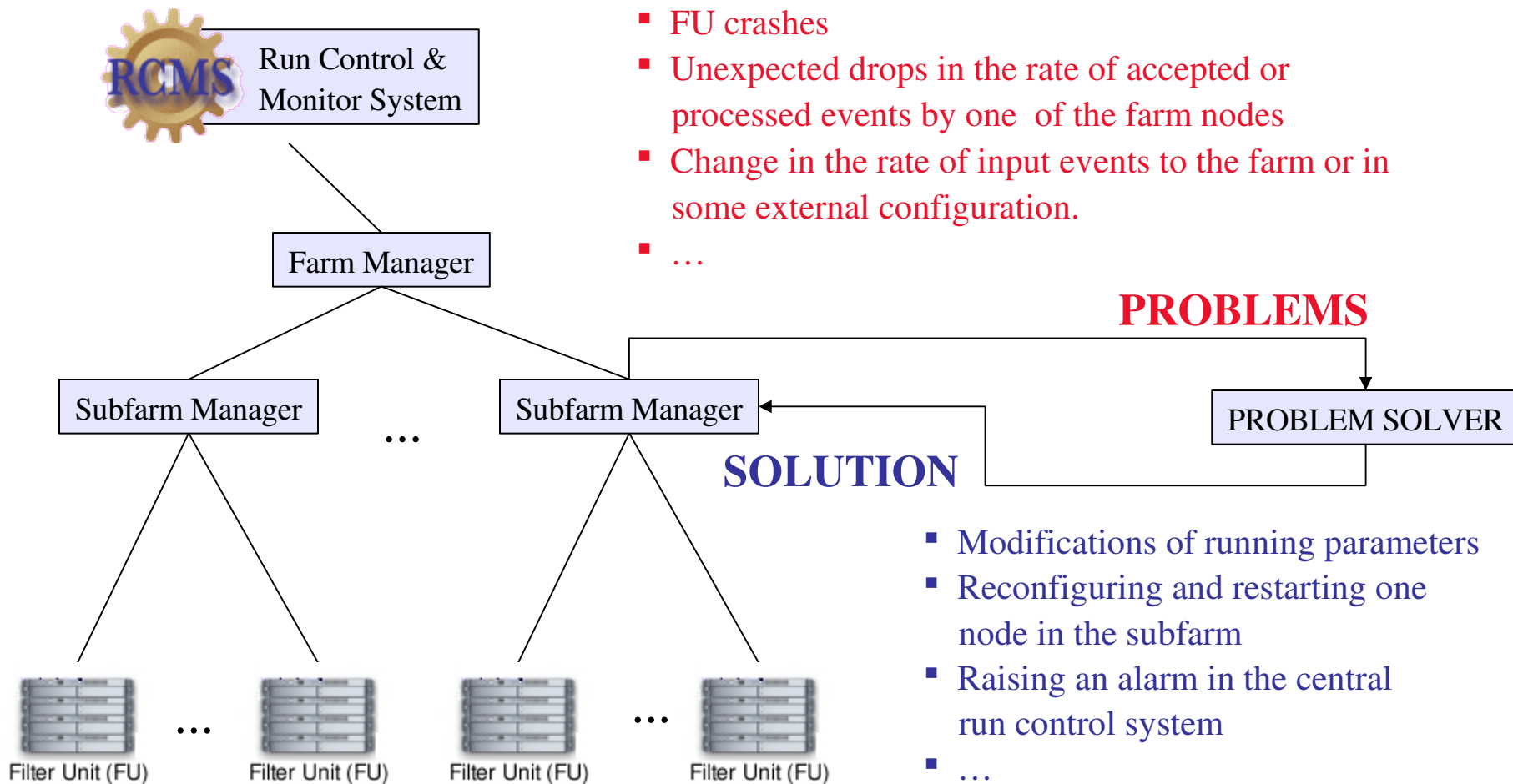
- Consists of 1000 dual-CPU PCs called Filter Units (FU)
- Carries out high level filtering of collision data
- Writes collision data to persistent storage for off-line analysis

Trigger and Data Acquisition Structure



Event Filter Farm

Towards Problem Solving Research



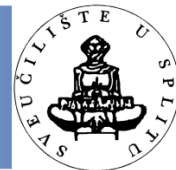
- FU crashes
- Unexpected drops in the rate of accepted or processed events by one of the farm nodes
- Change in the rate of input events to the farm or in some external configuration.
- ...

- Modifications of running parameters
- Reconfiguring and restarting one node in the subfarm
- Raising an alarm in the central run control system
- ...



Expert Systems Technology to the Rescue

When to use?



- How would you write a computer program to solve the following problem?
[Friedman-Hille, E. 2003, *Jess in Action*, Manning Publications, Greenwich]

A foursome of golfers is standing at a tee, in a line from left to right. Each golfer wears different colored pants.

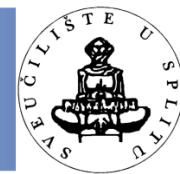
- One is wearing red pants.
- The golfer to Fred's immediate right is wearing blue pants.
- Joe is second in line.
- Bob is wearing plaid pants.
- Tom isn't in position one or four, and he isn't wearing the hideous orange pants.

In what order will the four golfers tee off, and what color are each golfer's pants?



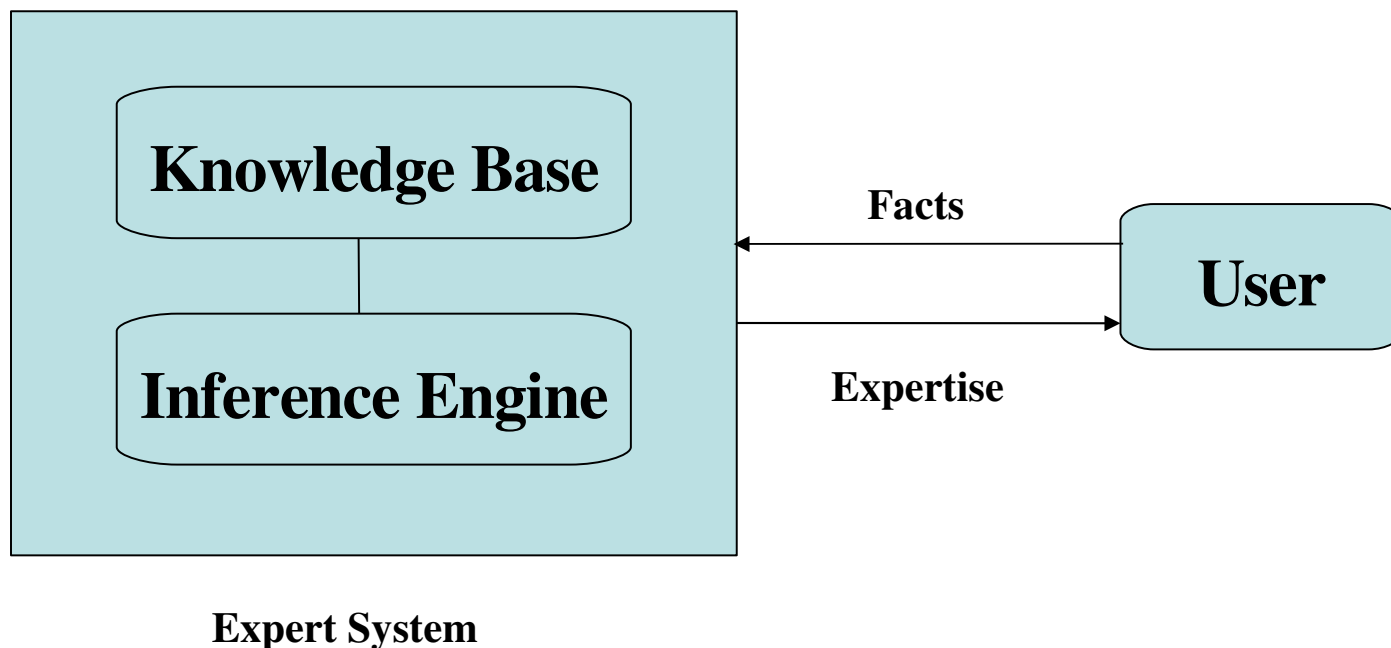
Expert Systems Technology to the Rescue

When to use?



Programs = Data Structures + Algorithms

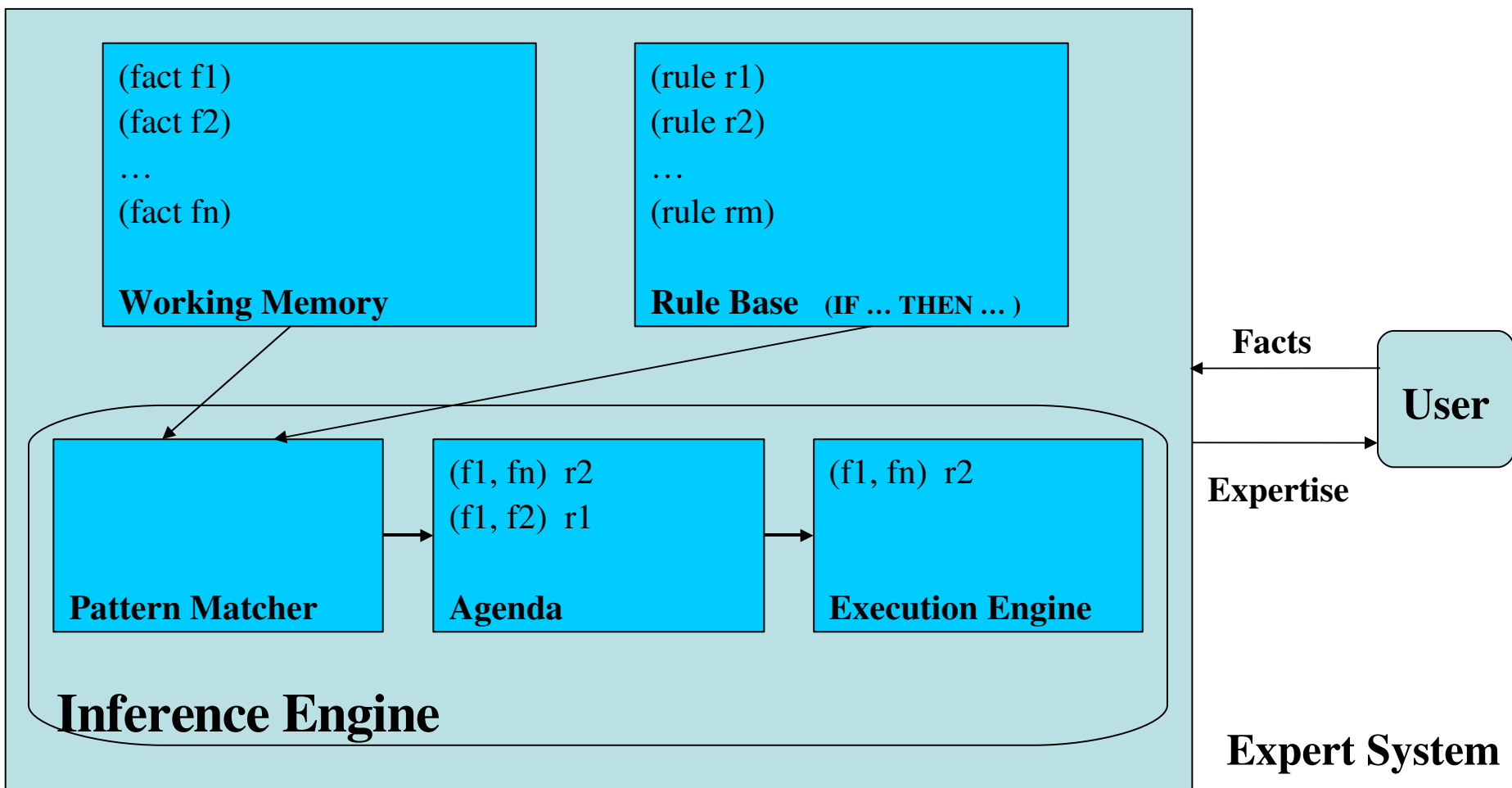
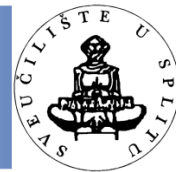
Expert Systems (ES) = Knowledge + Inference





Expert Systems Technology to the Rescue

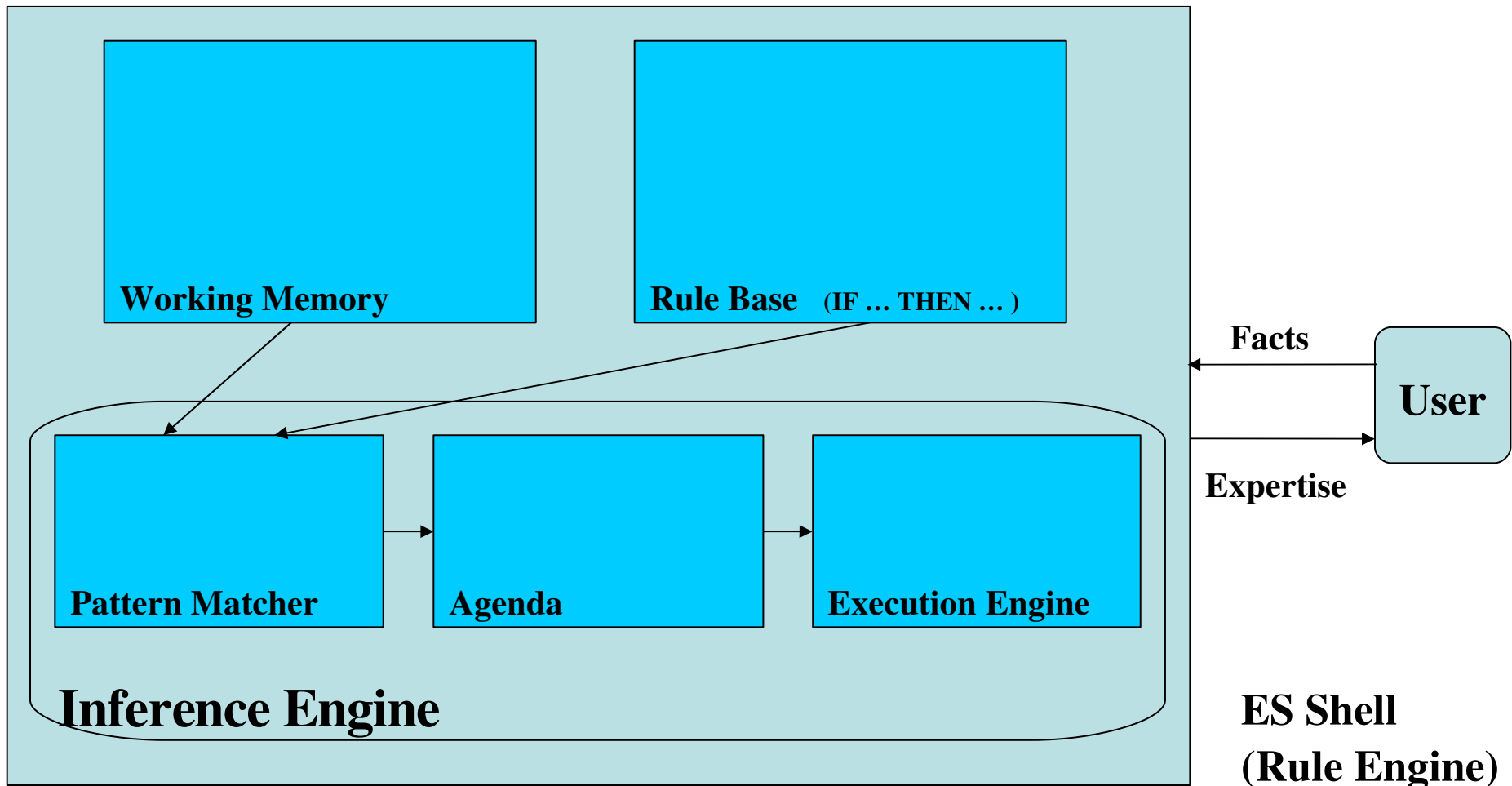
Rule-Based Expert Systems





Expert Systems Technology to the Rescue

Rule-Based Expert Systems





Expert Systems Technology to the Rescue

Jess - Java Expert System Shell



Rule engine for the Java platform by Ernest Friedman-Hill

at Sandia National Laboratories, Livermore [<http://herzberg.ca.sandia.gov/jess/>]

- Can access all Java classes and libraries from Jess
- Extendable by writing Java code
- Can be embedded in Java applications

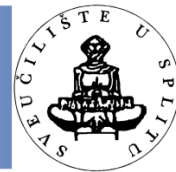
Can therefore be used in:

- Command-line applications
- GUI applications
- Servlets
- Applets



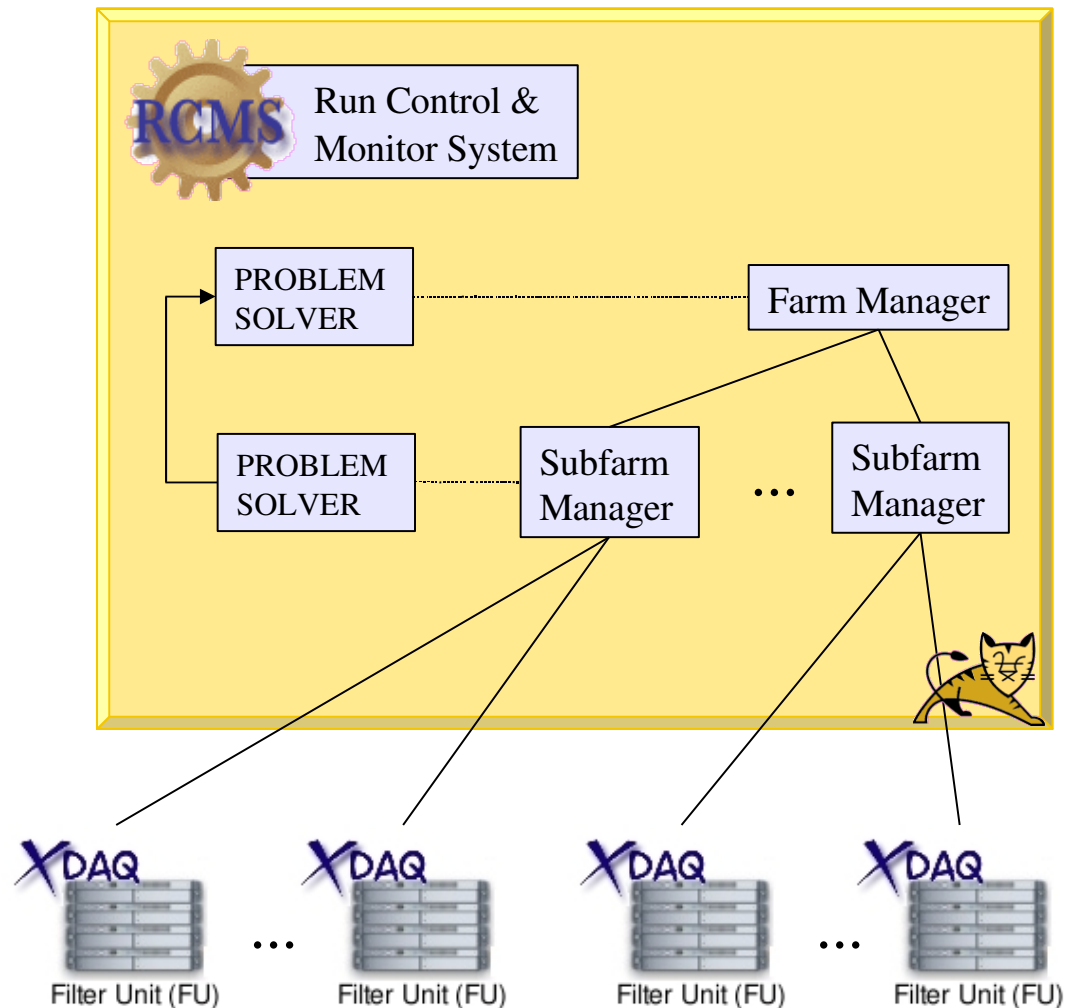
Problem Solver

How Stuff Works



Software Environment

- Apache Jakarta Tomcat
Servlet container
- RCMS
Responsible for controlling and monitoring CMS experiment during data taking
- CMSSW
Components for simulation, calibration and alignment, and reconstruction modules that process event data so that physicists can perform analysis
- XDAQ
Framework/middleware for local and remote inter-process communication, configuration, control, and data storage



Problem Solver

STARTUP

Monitor Data Collection Handler

Updater | Asserter | Factory

Jess Fact Handler

Working Memory (WM)

J
e
s
s

E
n
g
i
n
e

Inference Engine

Rule Base (RB)

Toolbox

Logger

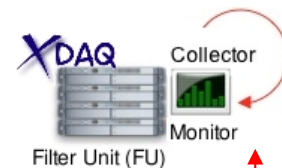
Network Analyzer

Batch

Retriever

Parser

Data Collection



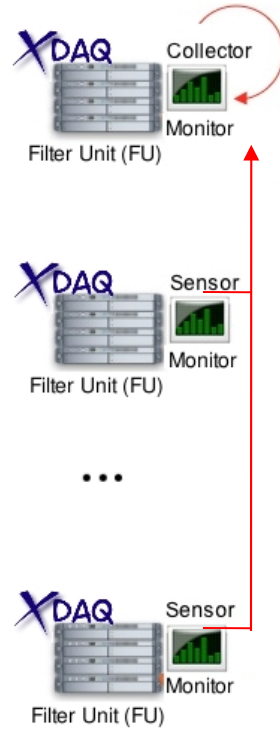
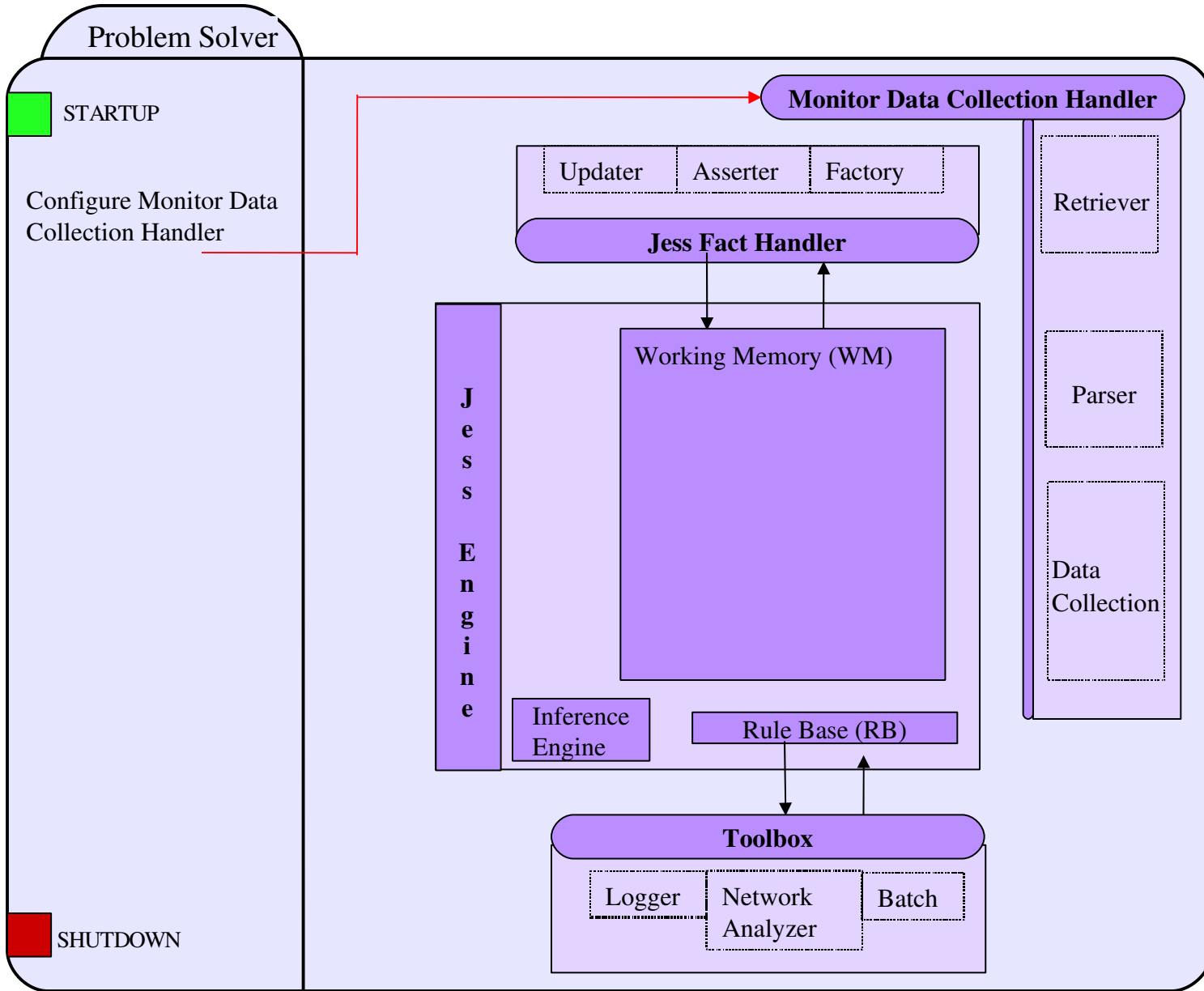
...

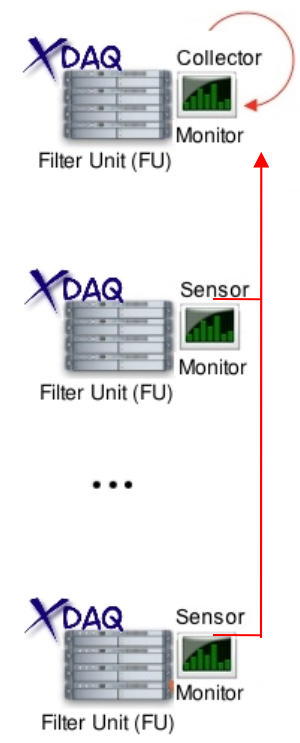
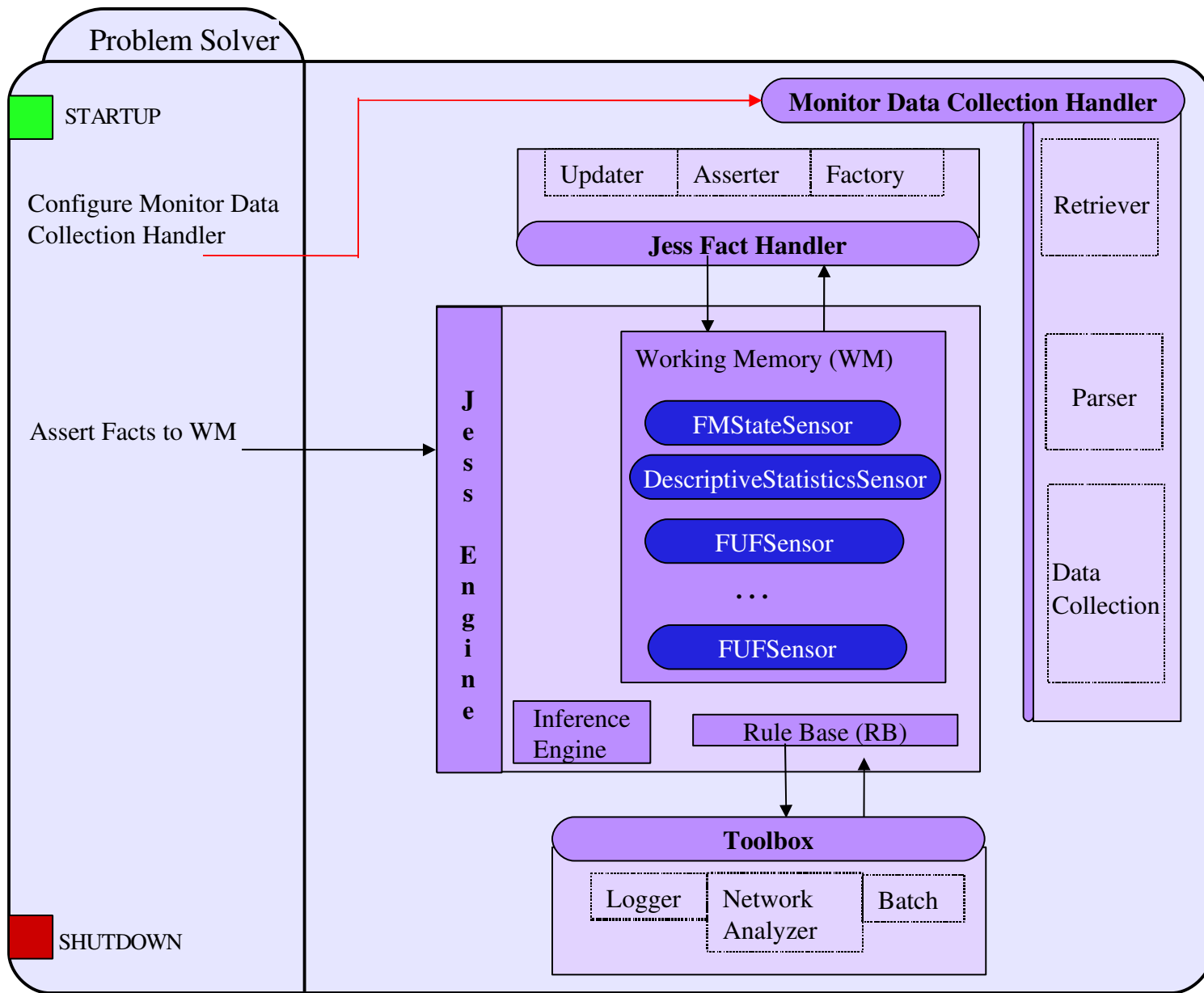


SHUTDOWN

Chainsaw







Problem Solver

STARTUP

Configure Monitor Data Collection Handler

Assert Facts to WM

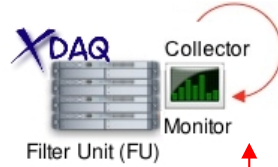
Load RB

Monitor Data Collection Handler

Updater Asserter Factory

Jess Fact Handler

Retriever



J
e
s
s

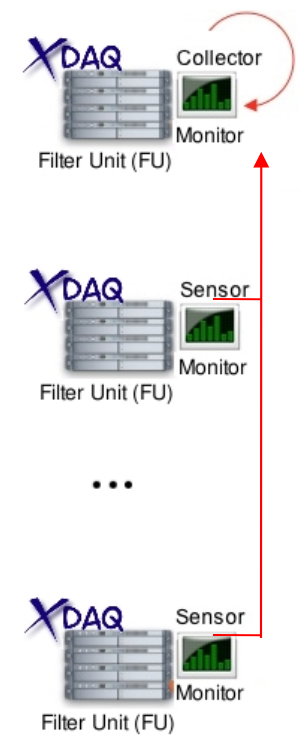
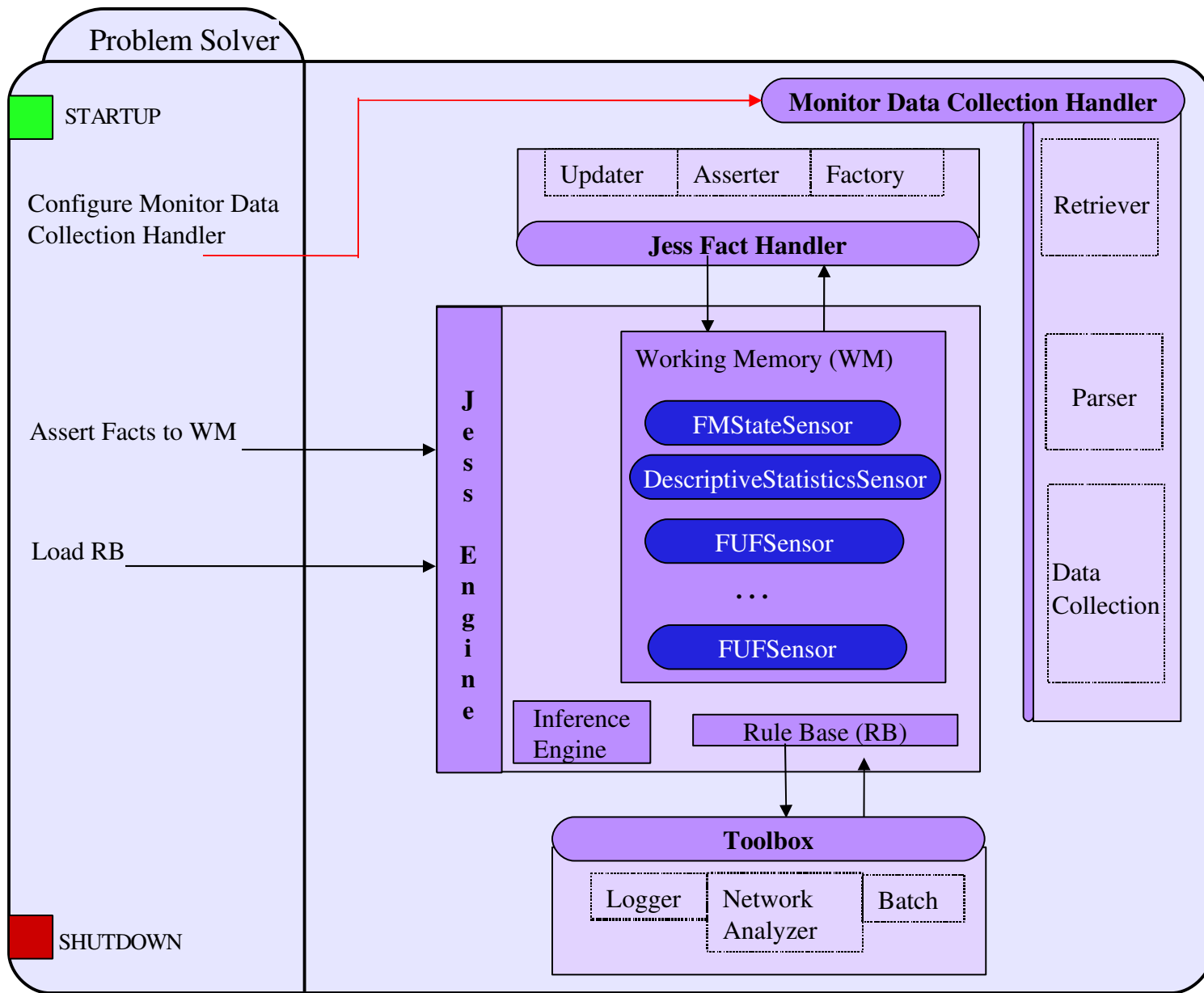
E
n
g
i
n
e

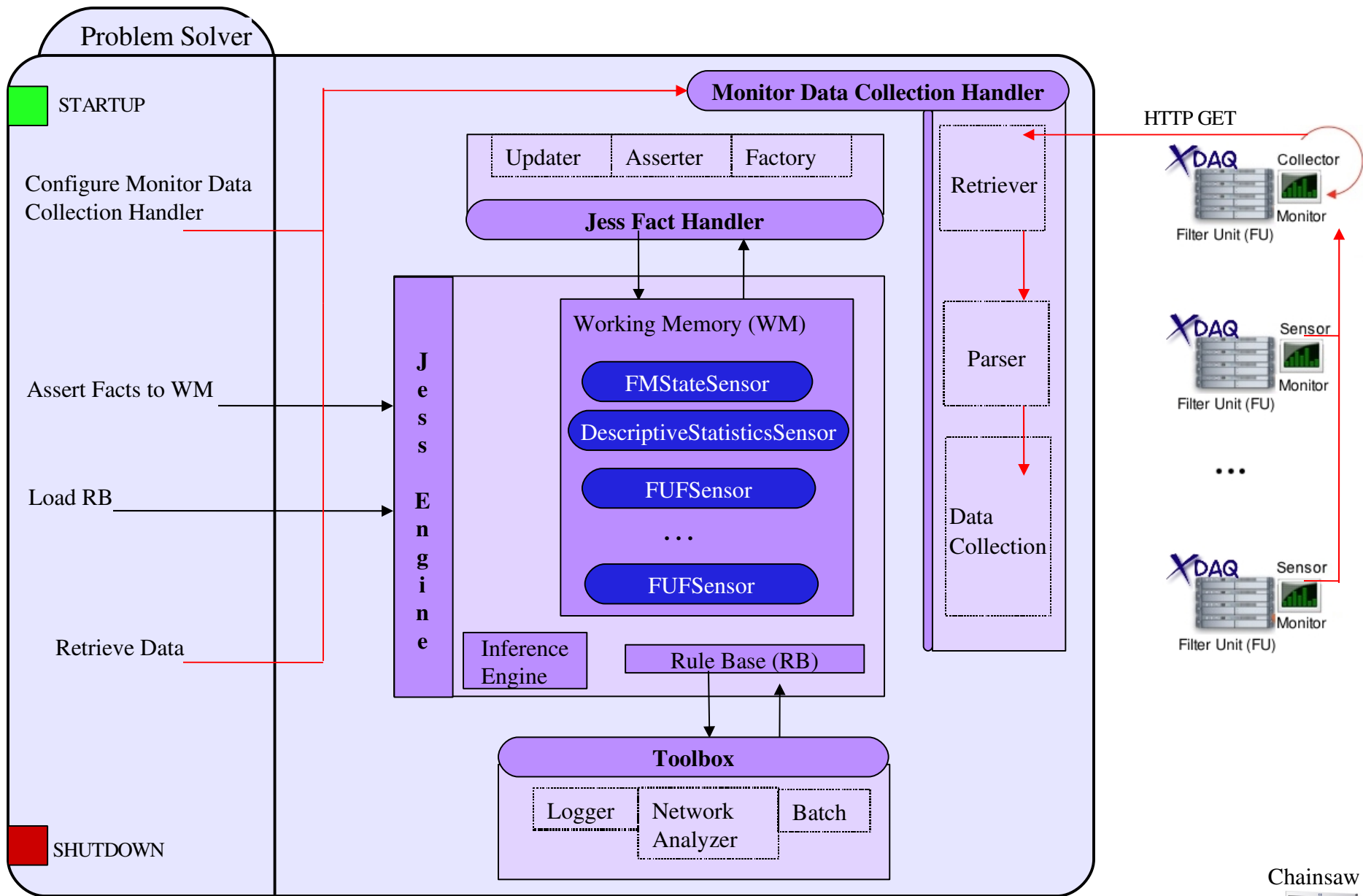
```
...  
(defrule SENSOR::network-analyzer "Tests host reachability"  
  ?f <- (SENSOR::FUFSensor (xdaqURL ?a) (stateName "Faulty") (problemSolving  
    FALSE))  
  =>  
  (modify ?f (problemSolving TRUE))  
  (modify ?f (stateName "Problem solving"))  
  (facts)  
  (bind ?string (str-cat "stateName=Faulty on host running FUF (xdaqURL = " ?a "  
    NEXT STEP: Testing host reachability"))  
  (call SubfarmProblemSolverLogger fatal ?string )  
  (printout t ?string )  
  (bind ?action (new NetworkAnalyzer ?a))  
  (assert (SENSOR::NetworkAnalyzer  
    (localHostReachable (call ?action getLocalHostReachable))  
    (remoteHostReachable (call ?action getRemoteHostReachable))  
    (message (call ?action getMessage))  
    (xdaqURL ?a)))  
  (facts))  
...  
)
```

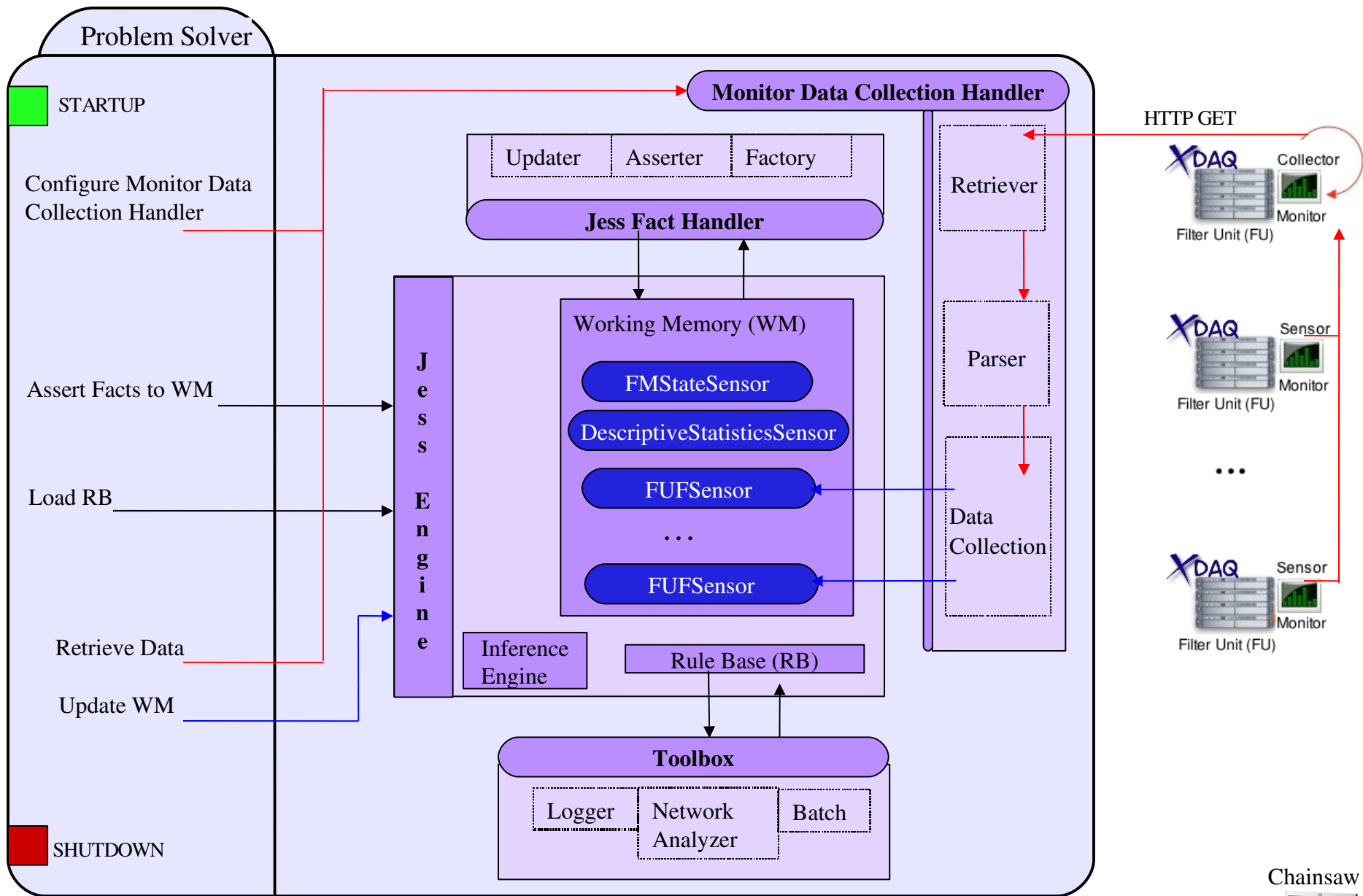
SHUTDOWN

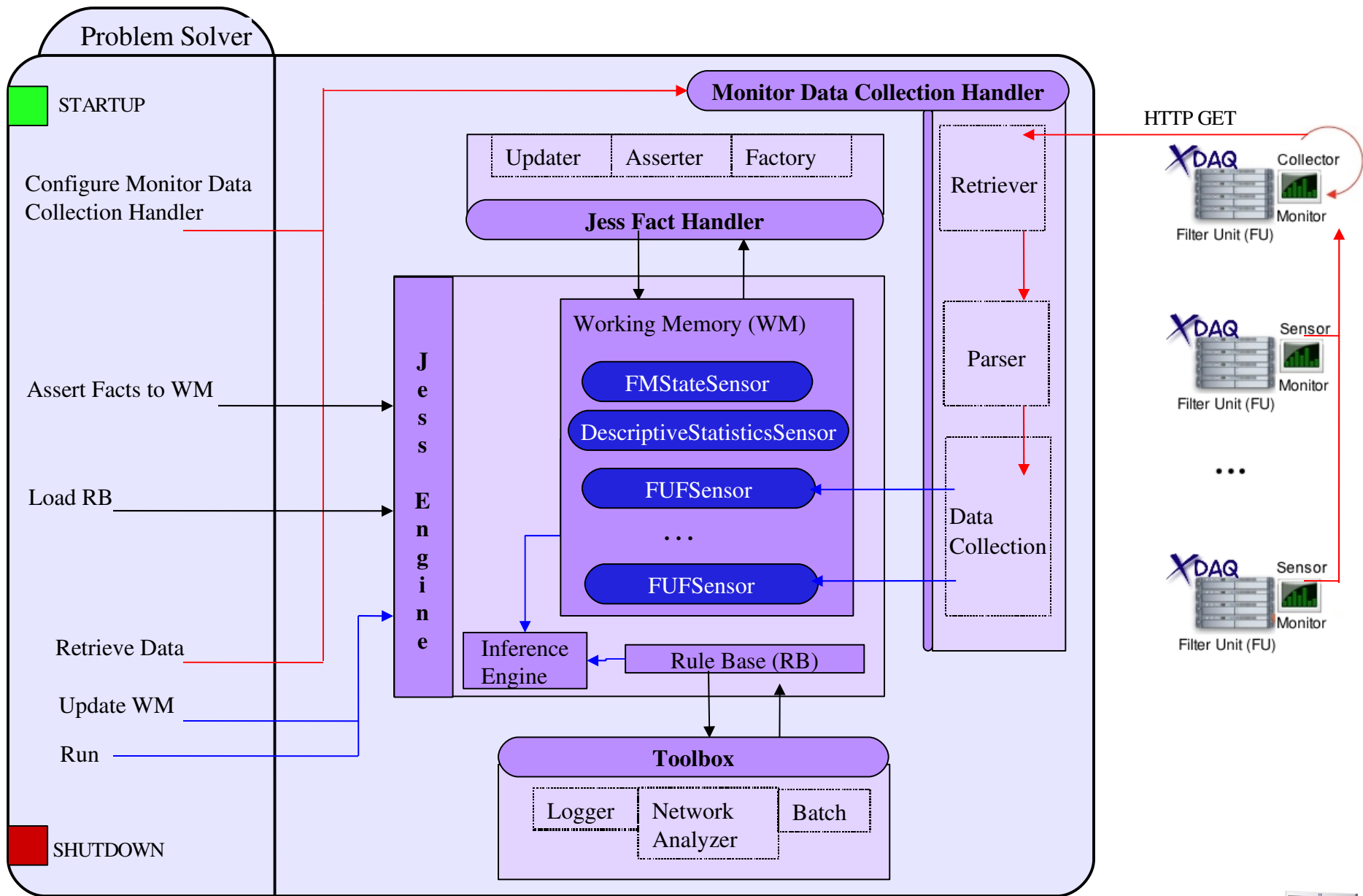
Chainsaw







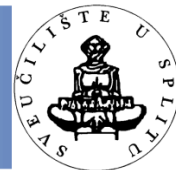




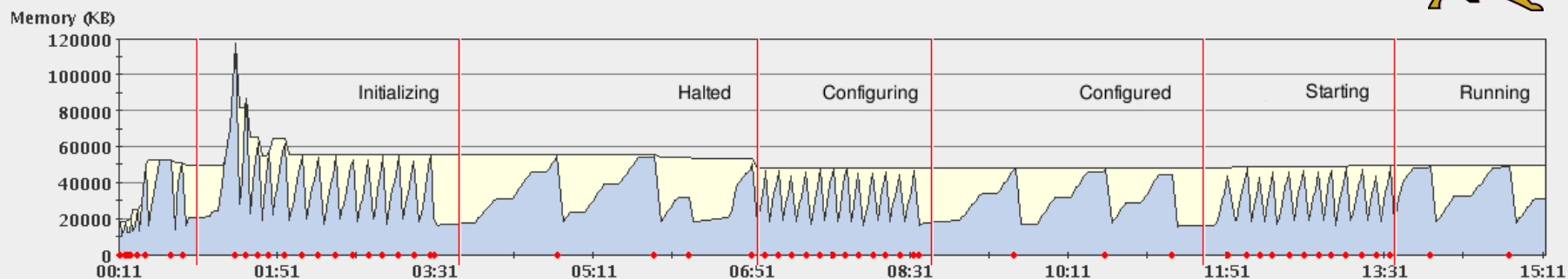


Problem Solver

Performance Analysis



- Heap Usage Chart of the Tomcat server hosting RCMS & Problem Solver



Maximum heap size: 256MB
Problem Solver's cumulative size: 9.4MB
Jess's WM (1050 facts -> 1050 FUs): 409.5KB

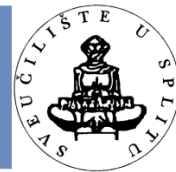
State transition: steep spikes in the Heap usage chart - a lot of SOAP Messages (short-lived objects) are being sent by a Subfarm Manager to Filter Units

Steady state: increasing growth – retrieval and processing of the monitor data collection

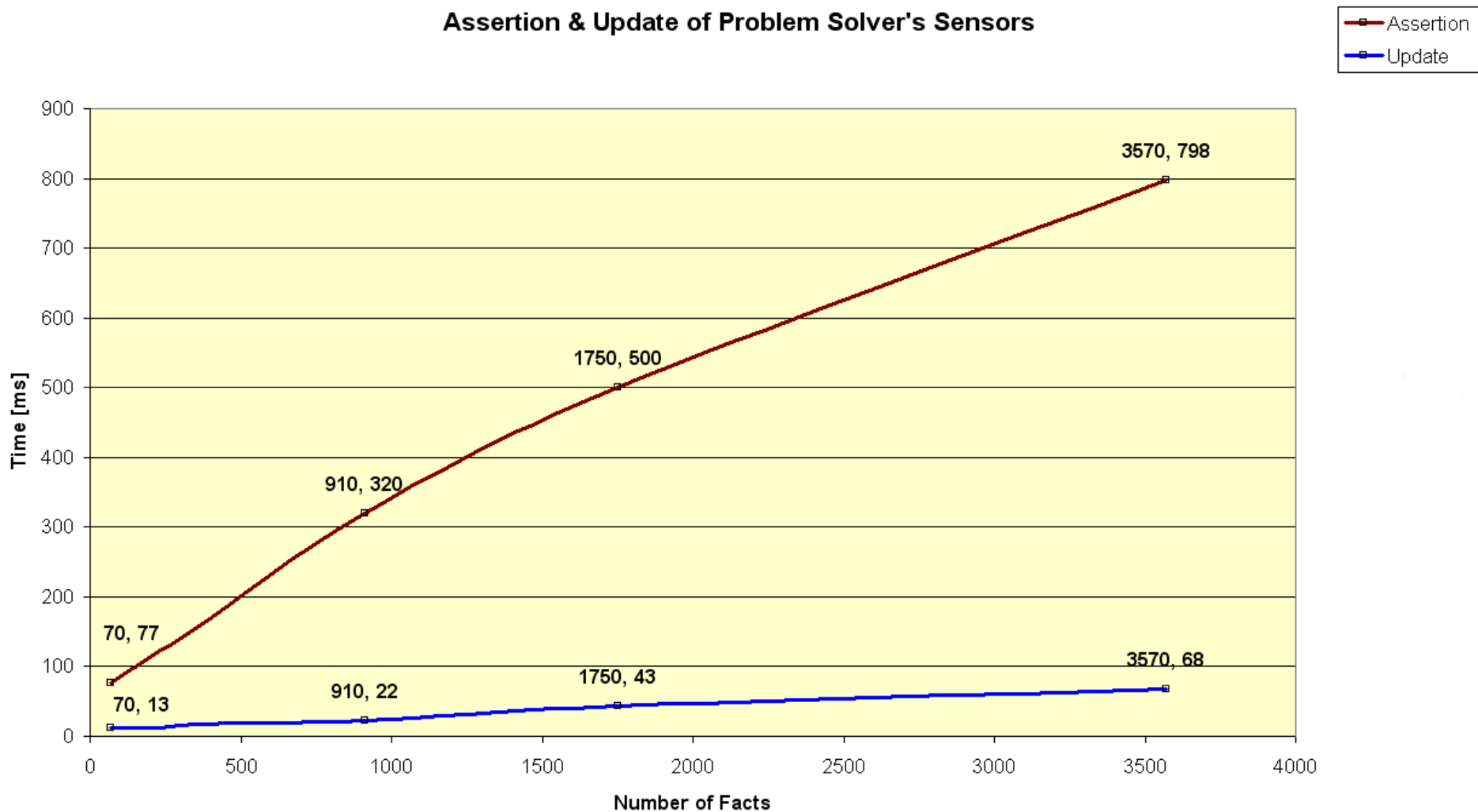


Problem Solver

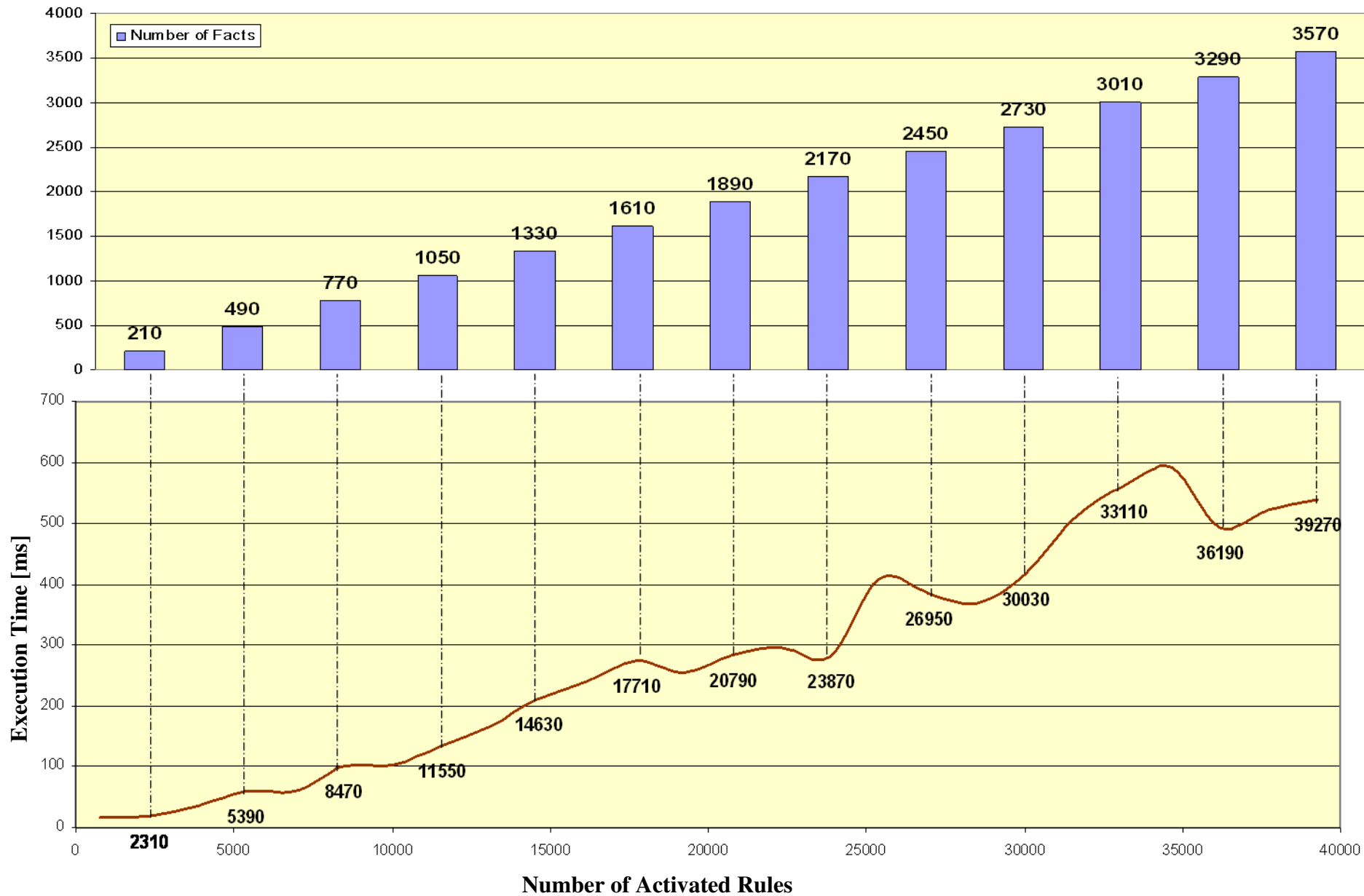
Performance Analysis



Assertion & Update of Problem Solver's Sensors



Inference Engine Performance



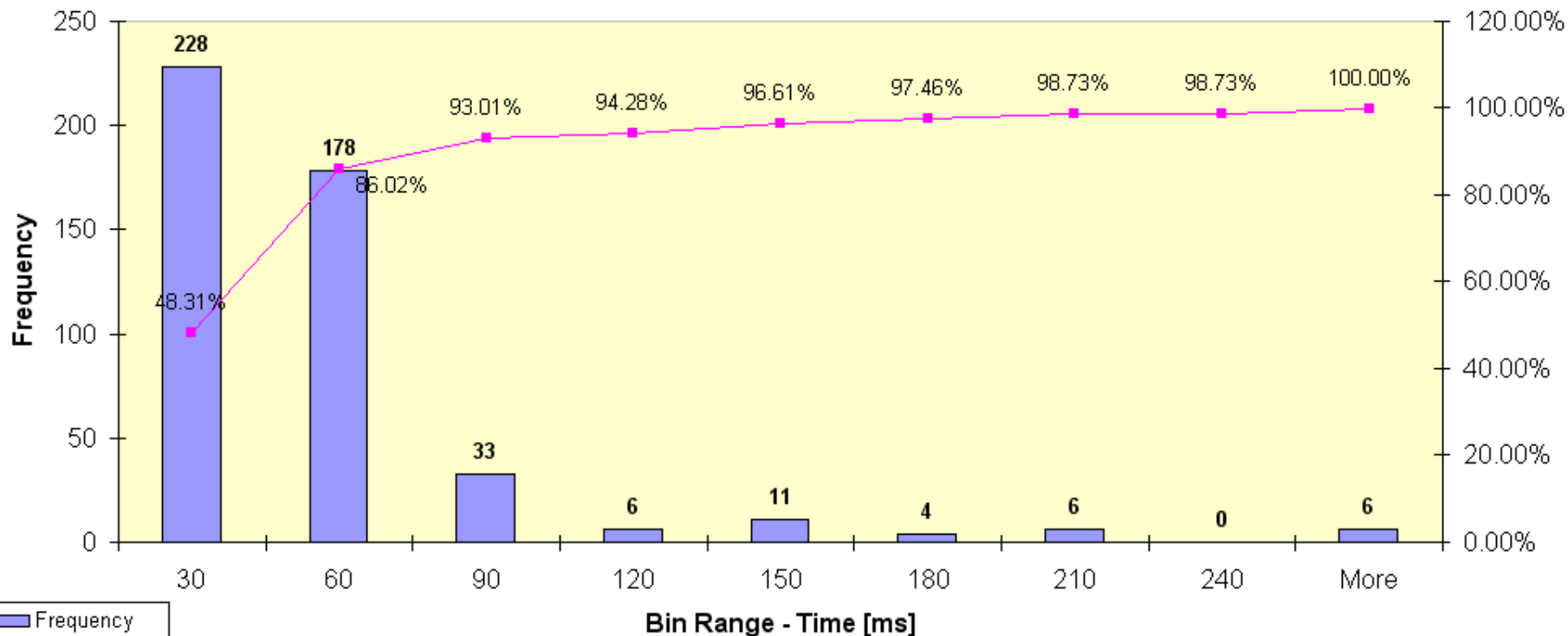


Problem Solver

Performance Analysis



Monitor Data Collection Retrieval & Processing



Confidence Level(95.0%) 72.669

Count 472

Mean 122.479

Minimum 17

Median 31

Maximum 9201

Mode 19

Other test cases: Count 6056

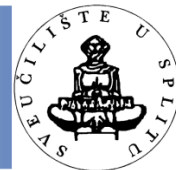
Data Loss 1.8%

Count 6919

Data Loss 3.37%



Summary



- Application of the Problem Solver with a large quantity of information in its working memory does not slow the whole system down.
- Inferencing which includes pattern matching with thousands of facts and forming an agenda with more than ten thousands of activated rules is not a time consuming task. Execution time of each activated rule depends on the rule's action part.
- Retrieval and processing time of Monitor's data collection is under 1s, and depends on the network load.