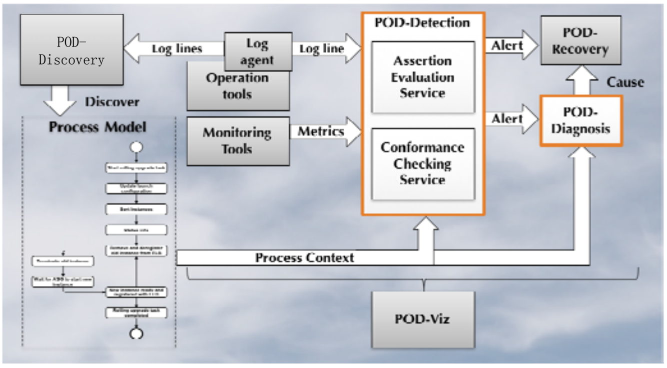
# POD-Detection & POD-Diagnosis

Detecting and diagnosing errors during Cloud operations

www.data61.csiro.au

Detecting and diagnosing errors during Cloud operations based on *process knowledge*:

- Runtime detection of deviations from log behavior & expected resource states
- Optimized diagnosis procedure to identify root causes
- All with high accuracy, even when noise from interfering operations is present



## POD-Detection

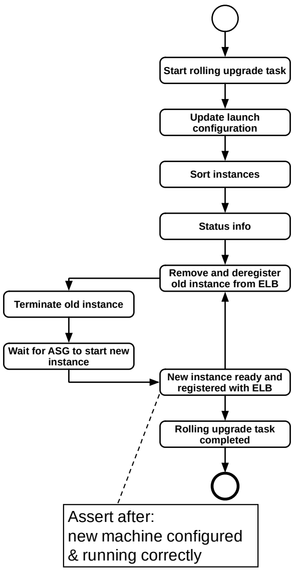
### Problem

When treating a system under observation as a black box, how can we non-invasively:

- detect deviation from normal log behaviour?
- detect abnormal resource state changes?

### Solution

- Conformance checking of log lines
  - Based on discovered process model and mapping of log events to process activities
- Assertion checking of resource states, using process context
  - Check if expected state matches actual state
  - Relies on open APIs, like AWS
- Anomaly detected → trigger diagnosis


Start rolling upgrade task
Update launch configuration
Sort instances
Status info
Remove and deregister old instance from ELB
Terminate old instance
Wait for ASG to start new instance
New instance ready and registered with ELB
Rolling upgrade task completed

Assert after:
new machine configured
& running correctly

## Next Steps & Impact

- Larger-scale evaluations
- Use it for security policy checking and intrusion detection
- POD-Detection:
  - Assertion part released as OSS
  - Next: automatic derivation of assertions
- POD-Diagnosis:
  - optimization to be parallelized
- Research publications:
  - DSN 2014 & MW4NextGen 2013
  - DSN 2015 paper in review

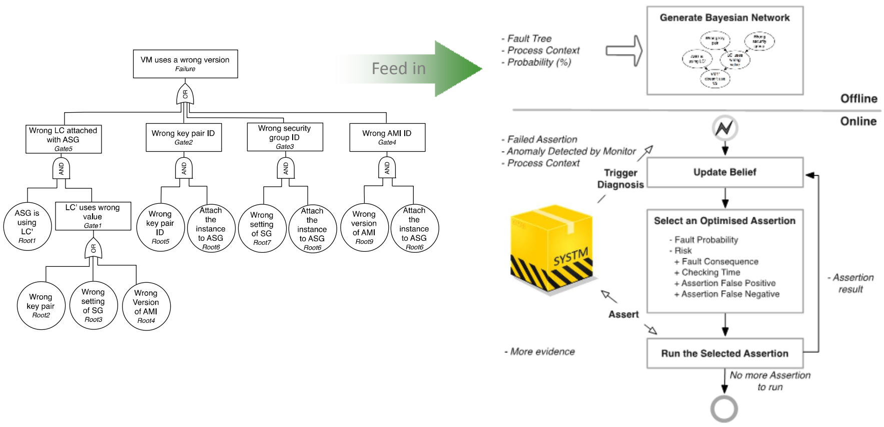
## POD-Diagnosis

### Problem

- Distinguish errors from operation effects
  - Sporadic operations often create “signals” similar to errors and faults
- How to diagnose root causes?

### Solution

- Use process context to:
  - Distinguish legitimate ops from errors
  - Narrow down possible causes
  - Update probabilities of possible causes
- Perform diagnostic tests to drill down
  - Optimized selection of test sequence, based on likelihood, impact and time
  - Partly relies on open APIs
- Data models:
  - Fault tree: causes and diagnostic tests
  - Bayesian network: probabilities and their updates, basis for optimization algorithm
- Error diagnosed → trigger recovery



FOR FURTHER INFORMATION
Ingo Weber
e Ingo.Weber@data61.csiro.au