Smart ALEC

Architecture for Learning Enabled Correlation

What are we trying to solve with ALEC?

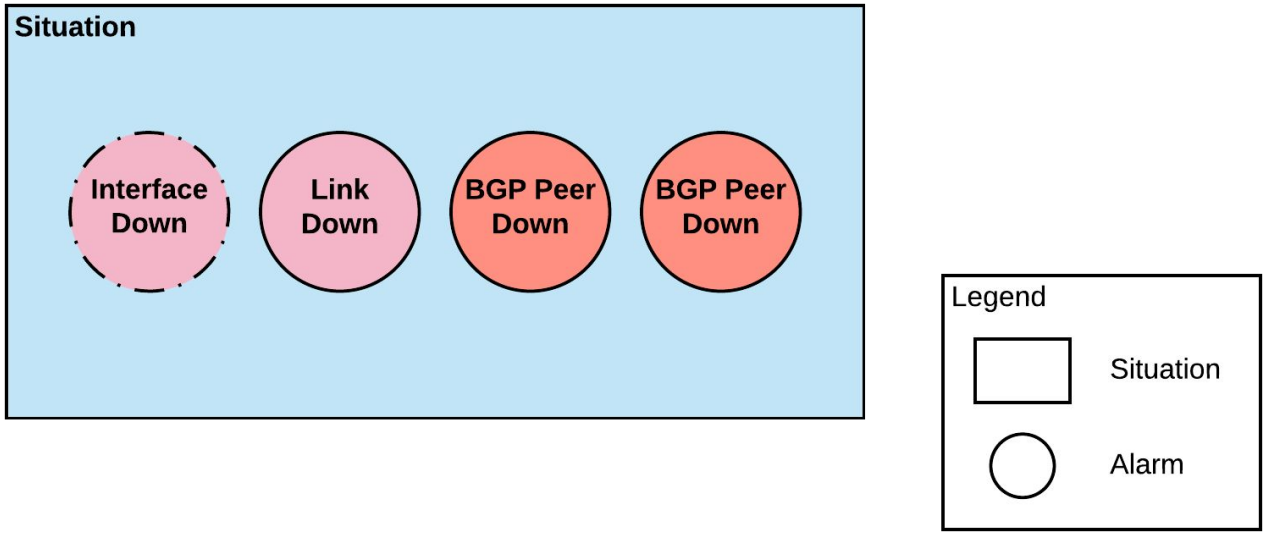
- Dealing with 100s, or even 1000s of alarms
  - Where to start?
- Alert fatigue

WHO?

- CTO of The OpenNMS Group
- Started as a user of the project
- Lead developer on ALEC

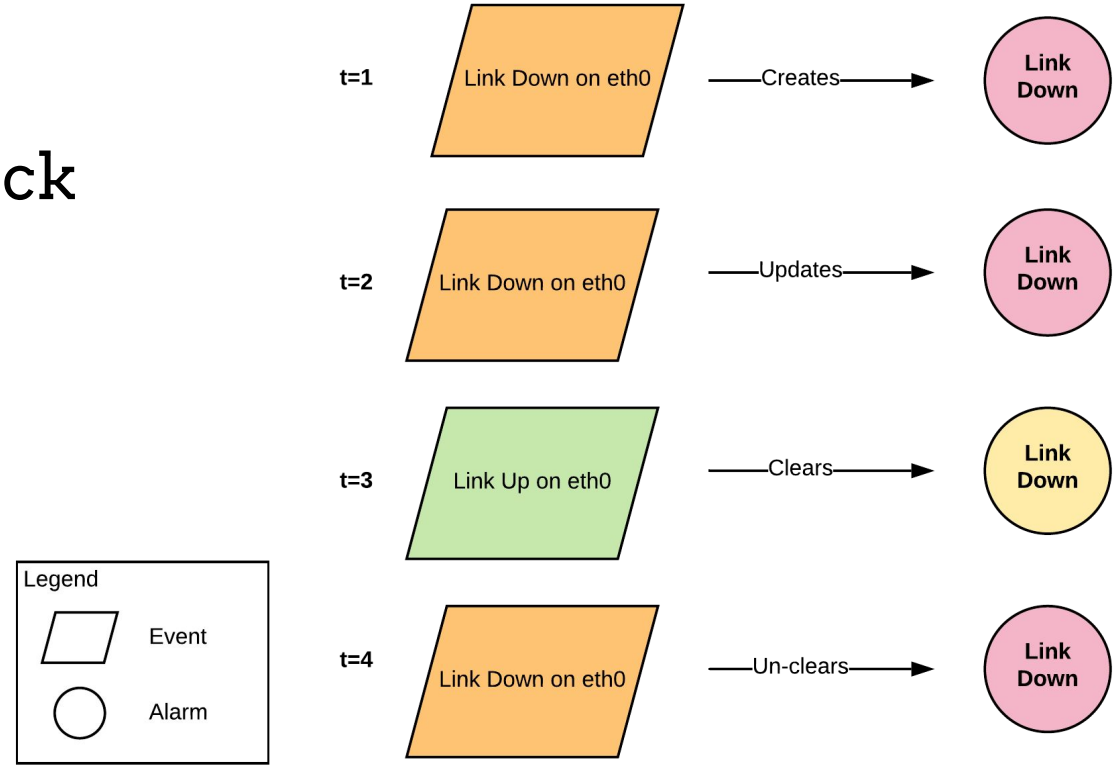
What is a situation?

A situation is a ***collection of alarms*** that share the same ***root cause***



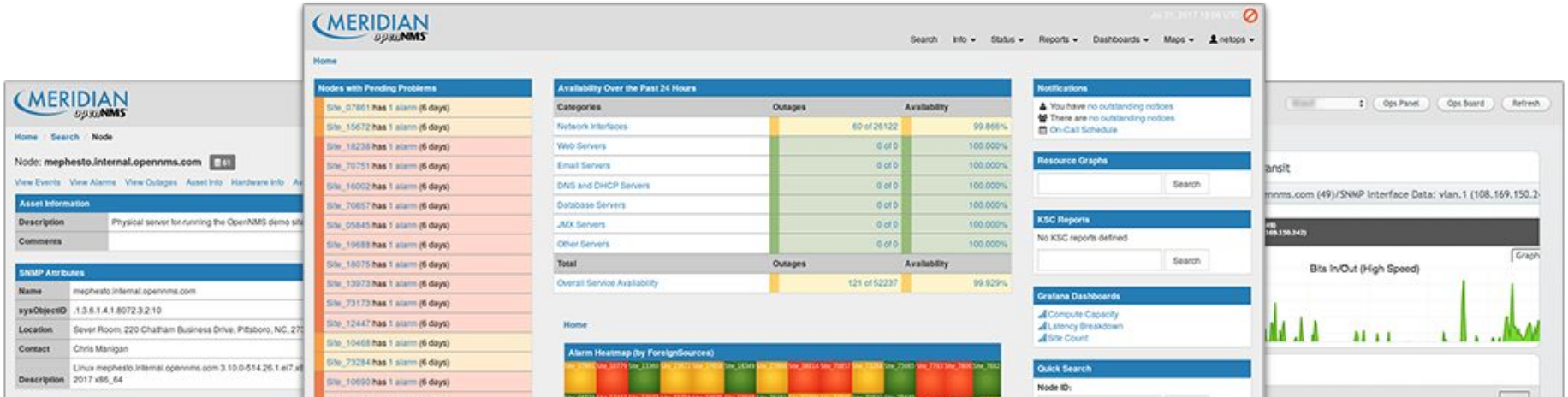
What is an alarm?

An alarm is a **state machine** used to track interesting **events**



ALEC leverages existing features in OpenNMS

- Fault Management
- Performance Management
- Network & Topology Discovery



Example 1 - Power Supply In/Out

Cisco Nexus Series 6000

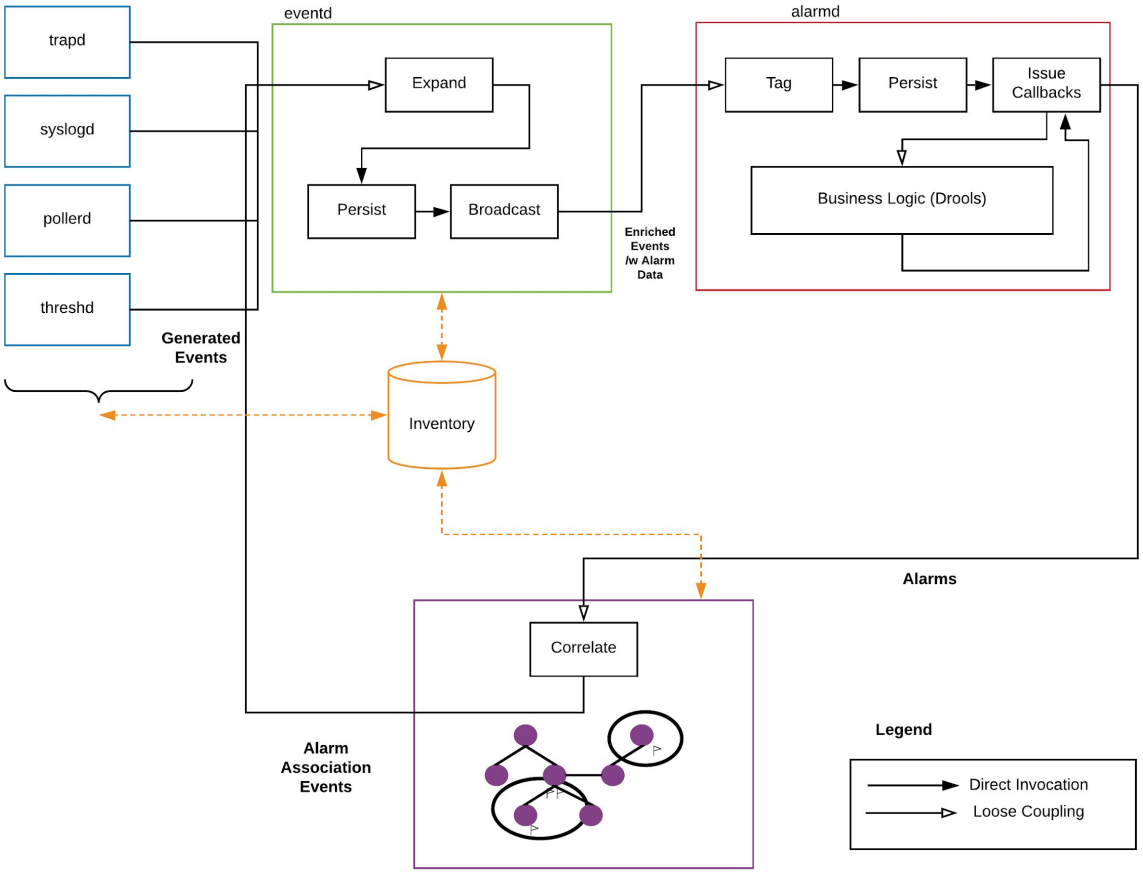
Time	Source	Event
21:37:48	Poll	PowerSupply Module 2- N2200-PAC-400W Out
21:38:43	Poll	PowerSupply Module 2- N2200-PAC-400W In
21:44:28	Syslog	%PFMA-2-FEX_PS_REMOVE: Fex 118 Power Supply 2 removed
21:44:48	Syslog	%PFMA-2-FEX_PS_FOUND: Fex 118 Power Supply 2 found
21:44:49	SNMP Trap	cefc power status down
21:44:49	SNMP Trap	cefc FRU inserted
21:44:49	SNMP Trap	cefc power status up

Example 2 - BGP Peer Down

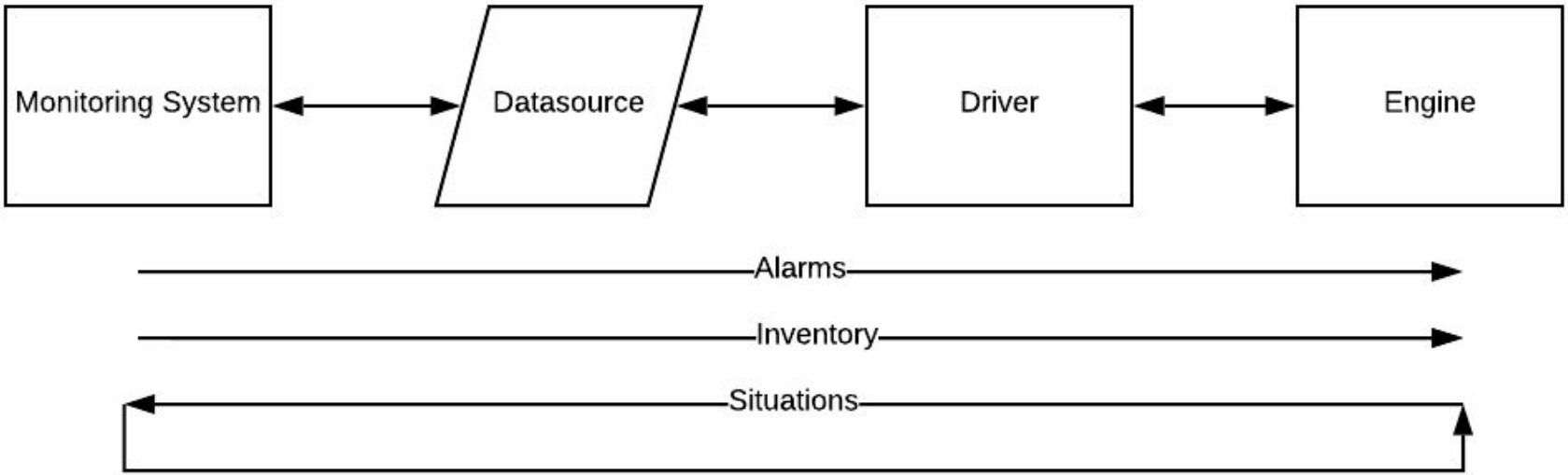
Cisco ASR 1002-X Router

Time	Source	Event
19:01:21	Syslog	%BGP-3-NOTIFICATION: sent to neighbor 10.0.0.1 4/0 (hold time expired) 0 bytes
19:01:21	Syslog	%BGP-5-ADJCHANGE: neighbor 10.0.0.1 vpn vrf REDACTED Down BGP Notification sent
19:01:22	SNMP Trap	BGP down trap
19:01:22	SNMP Trap	Cisco BGP backward transition trap
19:01:22	SNMP Trap	Cisco BGP down trap
19:01:43	Poll	BGP Neighbor connection lost between 10.0.0.1 and 10.0.0.2
19:02:12	Syslog	%BGP-5-ADJCHANGE: neighbor 10.0.0.1 vpn vrf REDACTED Up
19:02:14	Poll	BGP Neighbor connection restablished between 10.0.0.1 and 10.0.0.2

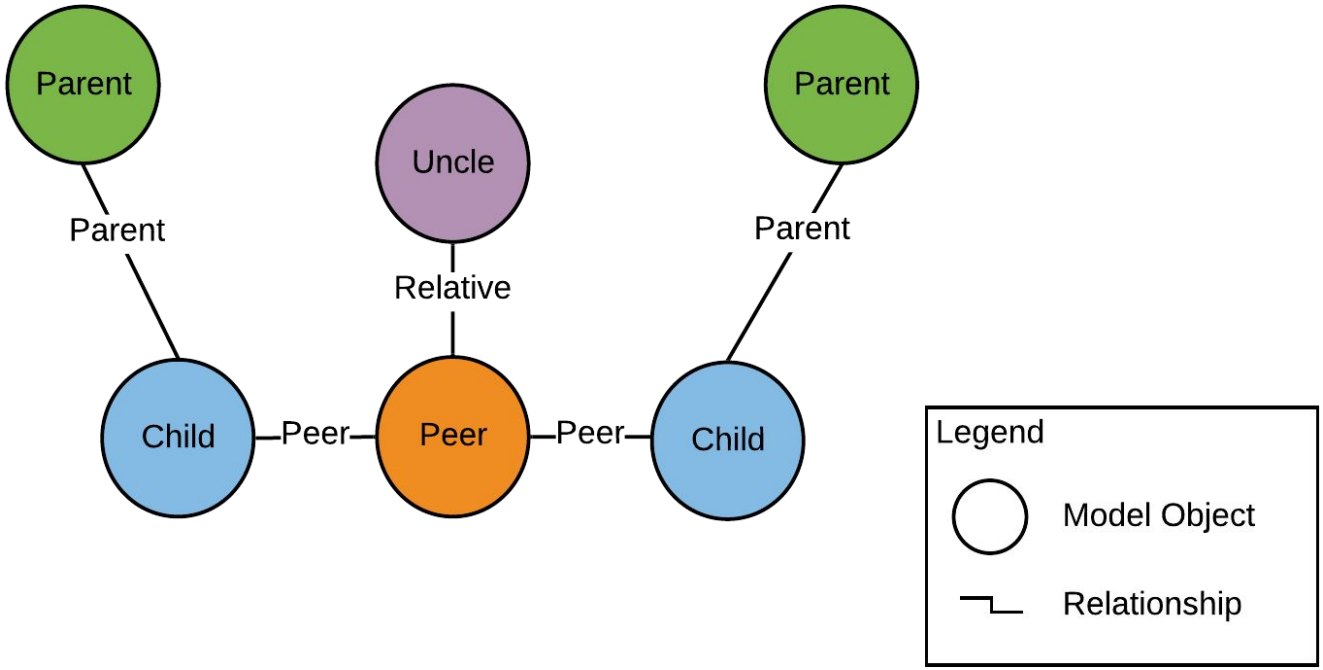
ALEC & OpenNMS Integration



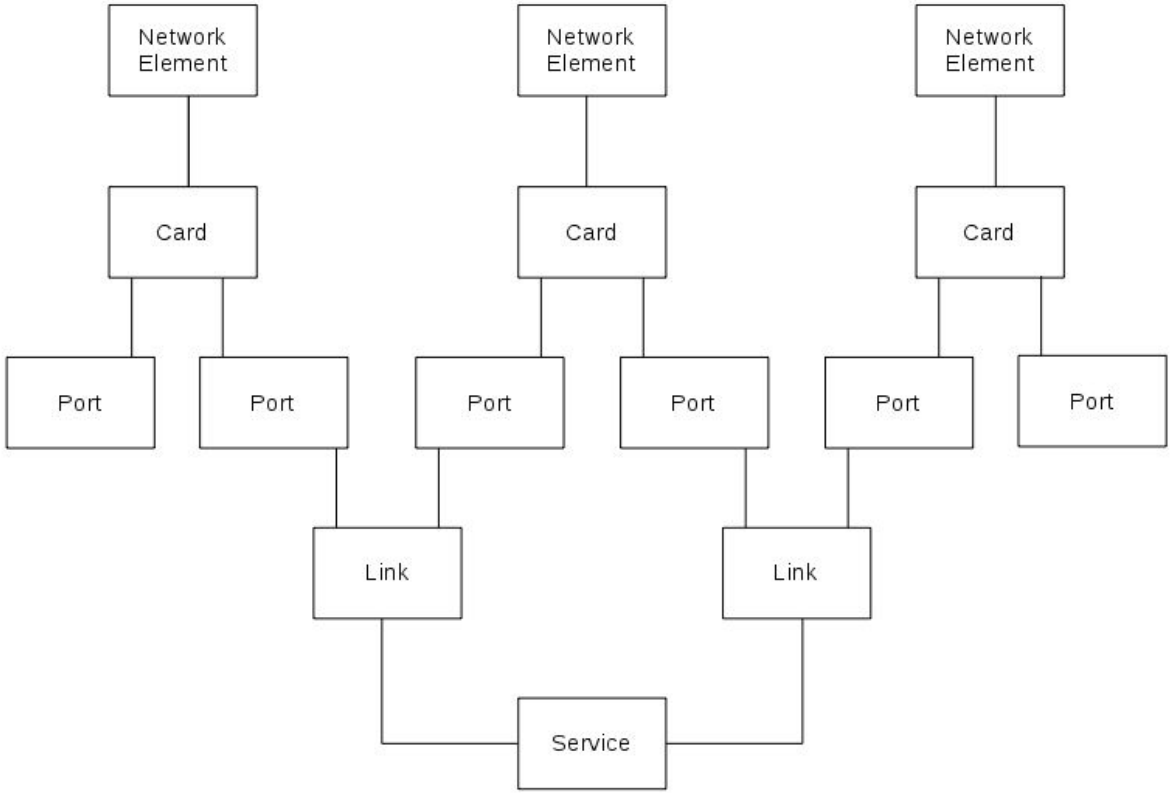
ALEC - High Level Architecture



Abstract Inventory Model

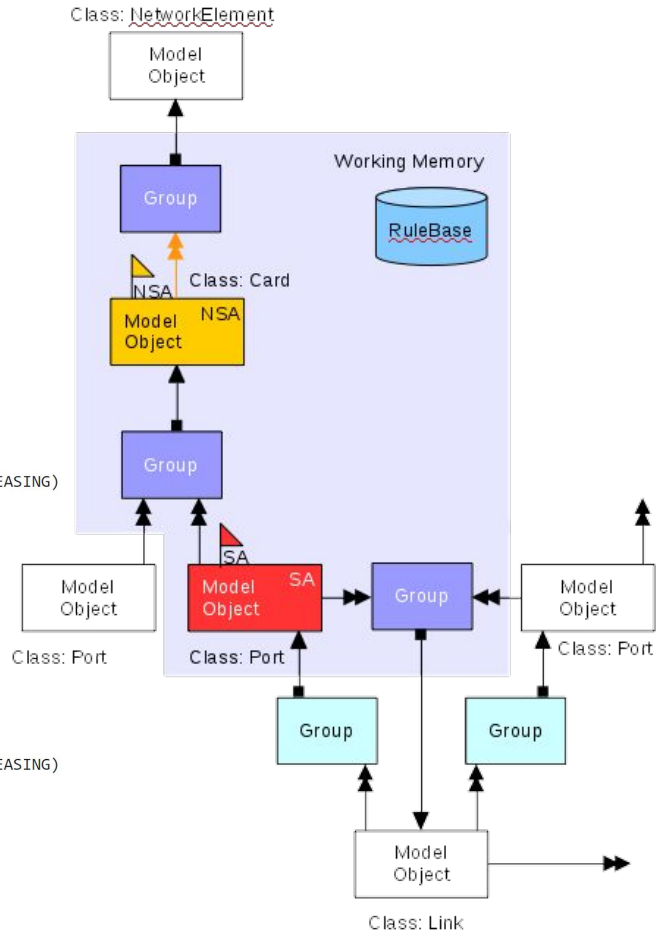


Simple Model Example



Rules Engine

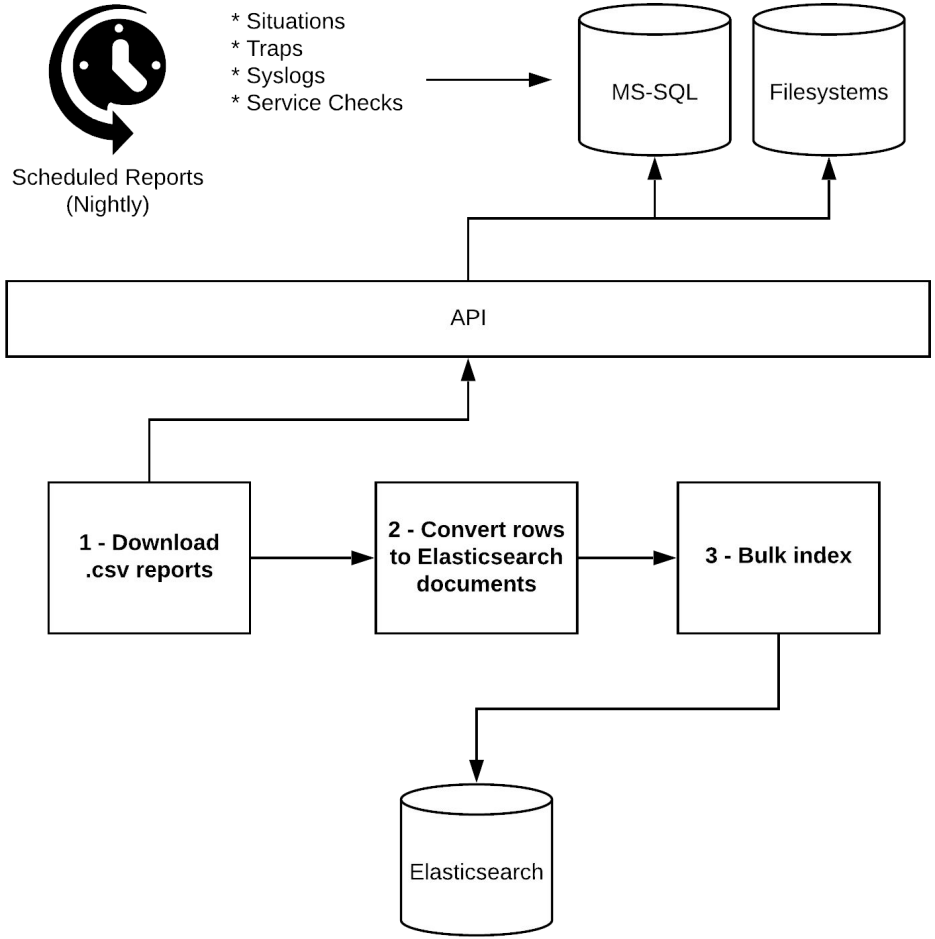
```
15  // RULE #1
16  rule "FirstPortDown"
17    when
18      $group : Group(owner.type == "Card", numberServiceAffecting == 1, serviceAffectingTrend == CountTrend.INCREASING)
19      not ModelObject(type == "Card", id == $group.getOwner().getId(), operationalState == OperationalState.NSA)
20    then
21      actionMgr.log("RULE #1");
22      actionMgr.synthesizeAlarm($group.getOwner(), OperationalState.NSA, Severity.MINOR, $group.getId());
23  end
24
25  // RULE #2
26  rule "CardDown"
27    when
28      $group : Group(owner.type == "Card", numberServiceAffecting == numberMembers, serviceAffectingTrend == CountTrend.INCREASING)
29      not ModelObject(type == "Card", id == $group.getOwner().getId(), operationalState == OperationalState.SA)
30    then
31      actionMgr.log("RULE #2");
32      actionMgr.synthesizeAlarm($group.getOwner(), OperationalState.SA, Severity.MAJOR, $group.getId());
33  end
34
35  // RULE #3
36  rule "CardDownReport"
37    when
38      $group : Group(owner.type == "Card", numberServiceAffecting == numberMembers, serviceAffectingTrend == CountTrend.INCREASING)
39      not ReportObjectImpl(owner.type == "Card", owner.id == $group.getOwner().getId())
40    then
41      actionMgr.log("RULE #3");
42      actionMgr.createReport($group);
43  end
44
```



JUST GIVE ME ALL THE DATA!

ETL Pipeline

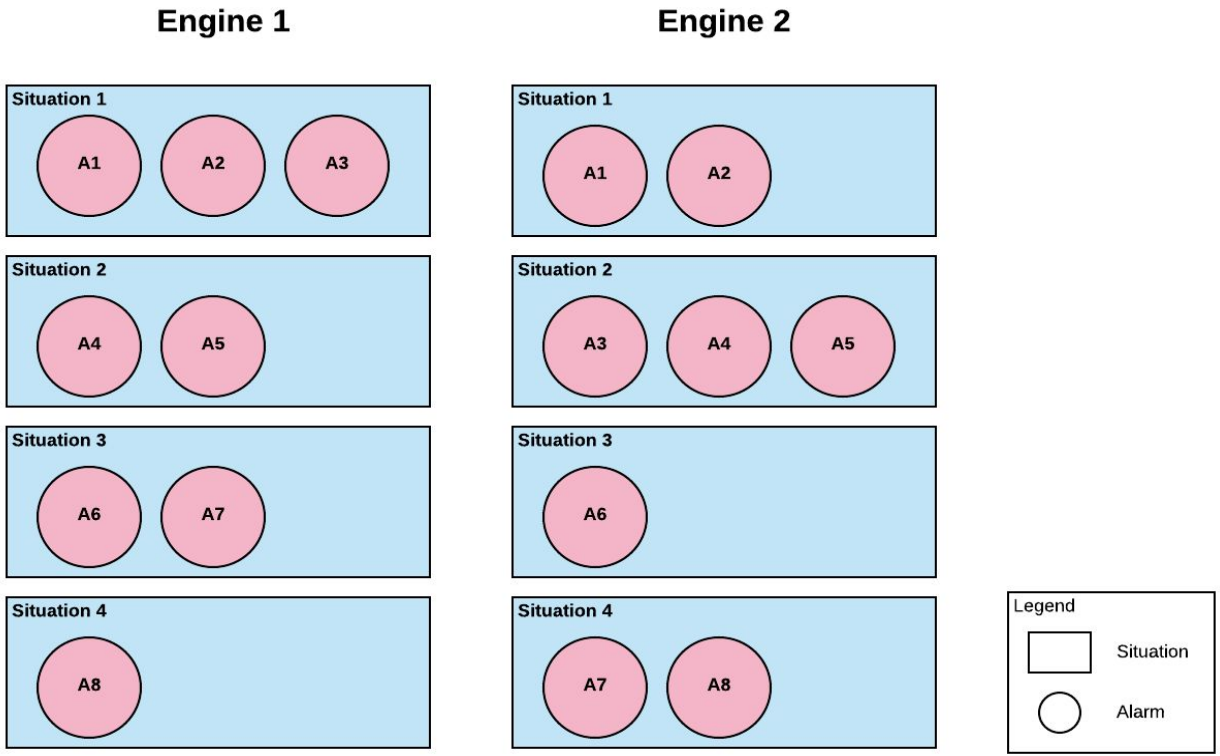
- Pull from existing solution using public APIs
- Make events, alarms & situations **easy** to access and analyze



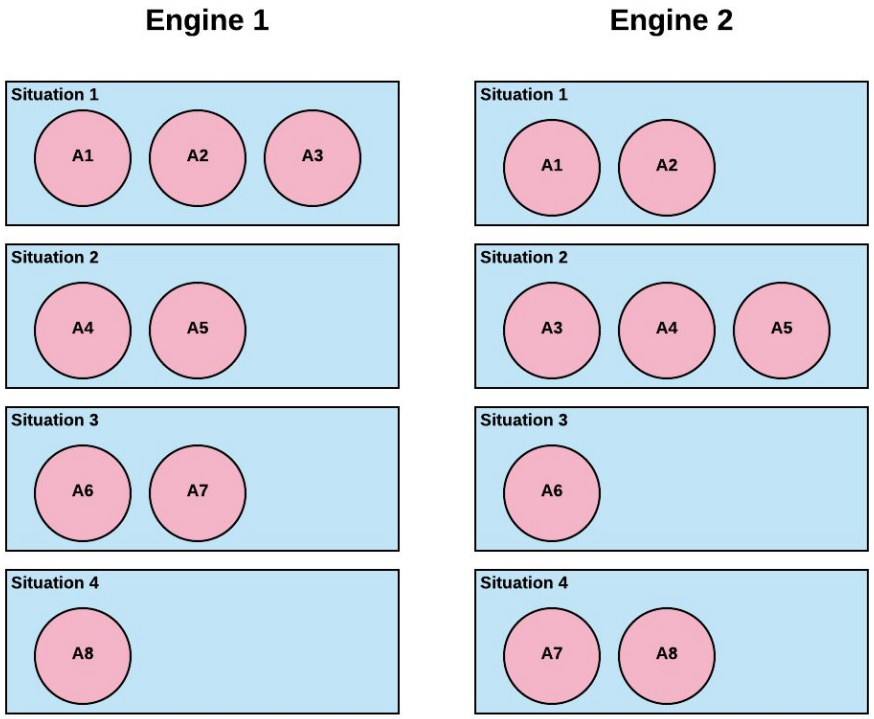
Given the same events, what situations would we produce?

```
karaf@root()> opennms-alec:process-alarms --alarms-in /home/jesse/labs/ato19/alarms.xml --inventory-in /home/jesse/labs/ato19/inventory.
xml --engine dbscan --situations-out /tmp/sim.situations.xml
Tick at Tue Dec 25 00:07:30 EST 2018 (1545714450000) - 0.00% complete - 0ms elapsed
Tick at Tue Dec 25 00:08:00 EST 2018 (1545714480000) - 0.03% complete - 2ms elapsed
Tick at Tue Dec 25 00:08:30 EST 2018 (1545714510000) - 0.07% complete - 5ms elapsed
Tick at Tue Dec 25 00:09:00 EST 2018 (1545714540000) - 0.10% complete - 6ms elapsed
Tick at Tue Dec 25 00:09:30 EST 2018 (1545714570000) - 0.14% complete - 8ms elapsed
Situation with id acf191cd-5643-493c-b6c4-e206071b91fc has 2 alarms.
Tick at Tue Dec 25 00:10:00 EST 2018 (1545714600000) - 0.17% complete - 9ms elapsed
Tick at Tue Dec 25 00:10:30 EST 2018 (1545714630000) - 0.21% complete - 11ms elapsed
Tick at Tue Dec 25 00:11:00 EST 2018 (1545714660000) - 0.24% complete - 12ms elapsed
Tick at Tue Dec 25 00:11:30 EST 2018 (1545714690000) - 0.28% complete - 15ms elapsed
Tick at Tue Dec 25 00:12:00 EST 2018 (1545714720000) - 0.31% complete - 23ms elapsed
Tick at Tue Dec 25 00:12:30 EST 2018 (1545714750000) - 0.35% complete - 25ms elapsed
Tick at Tue Dec 25 00:13:00 EST 2018 (1545714780000) - 0.38% complete - 27ms elapsed
Tick at Tue Dec 25 00:13:30 EST 2018 (1545714810000) - 0.42% complete - 29ms elapsed
Tick at Tue Dec 25 00:14:00 EST 2018 (1545714840000) - 0.45% complete - 31ms elapsed
Situation with id 2a793c1e-b32c-43c3-878f-929db14ac67d has 2 alarms.
Tick at Tue Dec 25 00:14:30 EST 2018 (1545714870000) - 0.49% complete - 40ms elapsed
Tick at Tue Dec 25 00:15:00 EST 2018 (1545714900000) - 0.52% complete - 42ms elapsed
Tick at Tue Dec 25 00:15:30 EST 2018 (1545714930000) - 0.56% complete - 44ms elapsed
Situation with id 67622d45-29a4-4336-bbad-1032cb47fbbb has 2 alarms.
Tick at Tue Dec 25 00:16:00 EST 2018 (1545714960000) - 0.59% complete - 46ms elapsed
Tick at Tue Dec 25 00:16:30 EST 2018 (1545714990000) - 0.63% complete - 48ms elapsed
```

How do we compare two differents sets of situations?



Peer-based scoring strategy



Alarms	Peers Matched	Divider
A1	1	2
A2	1	2
A3	0	2
A4	1	2
A5	1	2
A6	0	1
A7	0	1
A8	0	1
Total	**4**	**13**

Score: 30.77%

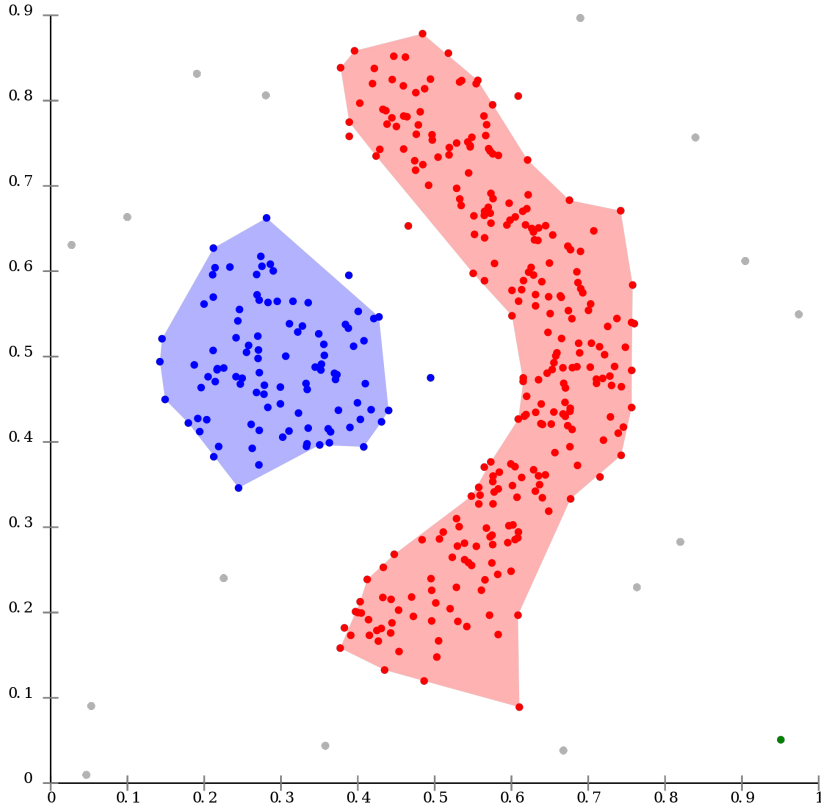

AH-HA!

Observation

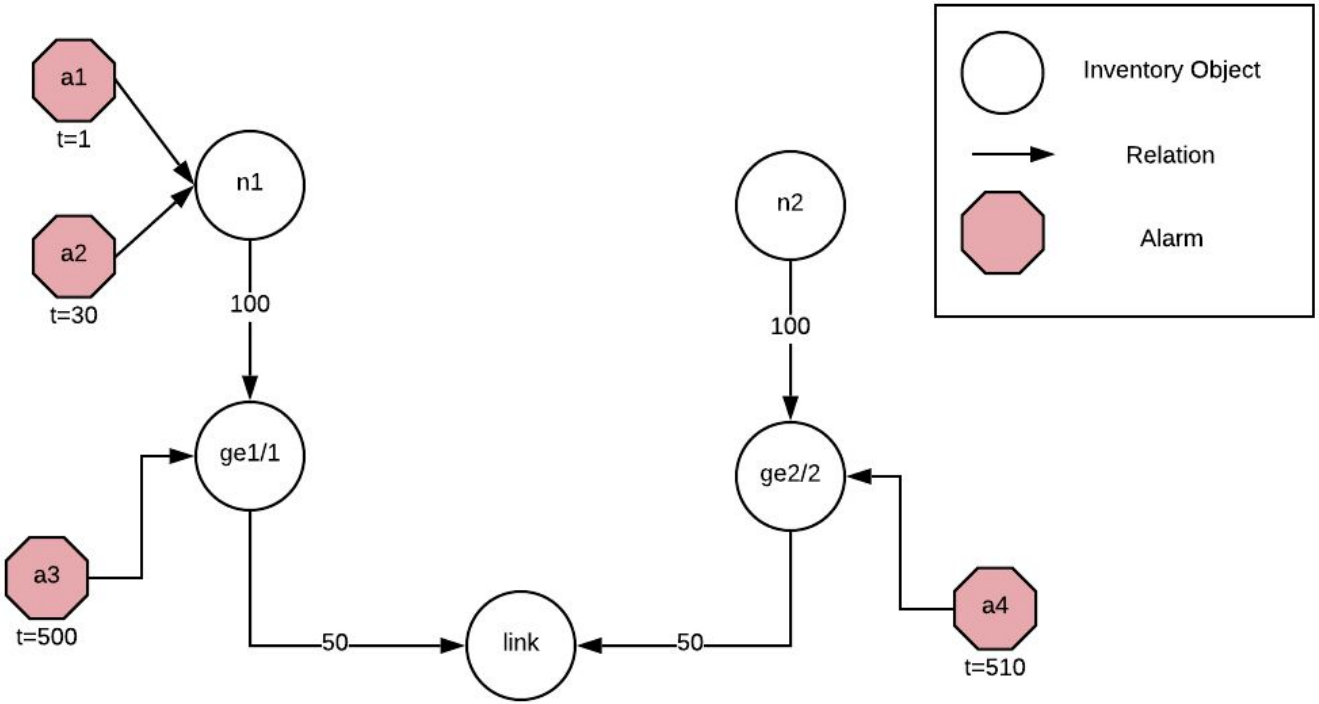
Situations contain alarms that are close in both **space** and **time**

Clustering

- Build generic **graph**
- Define a **distance** function
- Leverage known **clustering** algorithms



Graph-based Model



Distance

$$d(a_1, a_2) = \alpha((\beta|t(a_1) - t(a_2)|\frac{1}{60} + (1 - \beta)dg(a_1, a_2))$$

where:

- a_1 and a_2 are the points representing the alarms
- $\alpha \in (0, \infty)$ is a scaling constant (directly related to ϵ)
- $\beta \in [0, 1]$ is a weighting constant
  - When β is closer to 0, more weight is given to the temporal component
  - When β is closer to 1, more weight is given to the spatital component
- $t(a_k)$ returns the time (timestamp in seconds) of the last occurence of the given alarm
- $dg(a_i, a_j)$ returns the normalized distance on the shortest path between the vertices for a_i and a_k
  - If both alarms are on the same vertex, then the distance is 0
  - If there is no path between both alarms, then the distance is ∞

DBSCAN-based Engine - Unsupervised ML

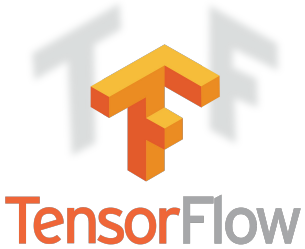
- Use DBSCAN against the graph with our distance function to build clusters
- Tune parameters to maximize score on training set
  - Epsilon, alpha, beta
  - BOBYQA
- Test against data outside of training set

Field Results w/ DBSCAN Engine

- ~80% score when compared to existing solution
- Good at clustering intra-node alarms
- Mixed results with inter-node alarms
- Difficult to tune for specific scenarios
    - Can only adjust model & weights, constants are global

Next Iteration

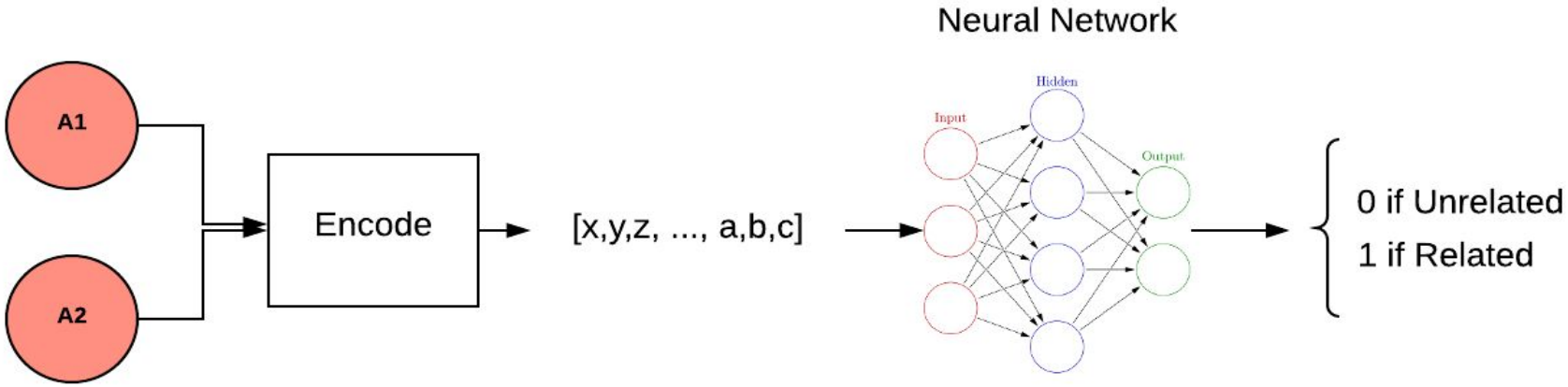
Can we use ***deep learning*** to improve results?




AH HA!!!!!!
imgflip.com

Binary Classification Problem

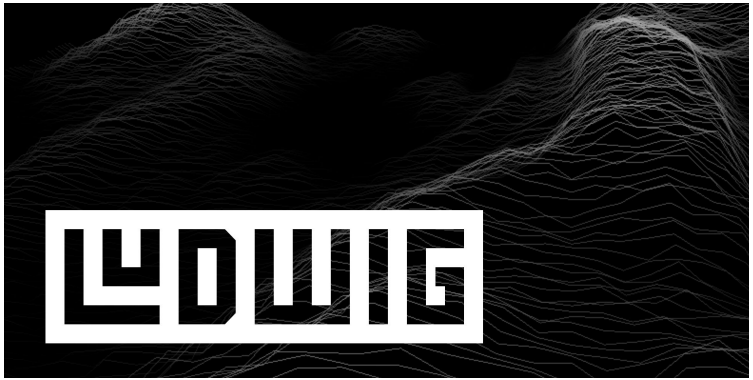
- Use existing **graph**
- Compare alarms to others **nearby**
- For each pair, ask if these are **related** or not



Deep Learning Model

Input Features	Type
A Type	Categorical
B Type	Categorical
Same Instance?	Binary
Same Parent?	Binary
Share Ancestor?	Binary
Time Delta	Numerical
Distance On Graph	Numerical
ID Similarity (Levenshtein)	Numerical
Label Similarity (Levenshtein)	Numerical

Output Features	Type
Related?	Binary



Field Results w/ Deep Learning Engine

- ~70% score when compared to existing solution
  - Currently inferior to DBSCAN engine - need more neurons!
- Possible to train for specific scenarios with labeled data (user feedback)
- Can be improved with more R&D


LET US PRAY
TO THE DEMO GODS
memegenerator.net

Thank You

Learn more at:

alec.opennms.com