Machine Intelligence in Data Center Networks

Hariharan Krishnaswamy
Senior Principal Engineer
Dell EMC
Topics of discussion

- Data Center Network SLA Management
- Application Workloads
- Machine Intelligence Methods
Data characteristics

- Increasing amounts of data is produced
- From a variety of Devices
- Stored and Processed in various different ways
- Higher workloads on the data center
Workloads

- **Service Workloads (ONLINE)**
  - Web search
  - Web serving
  - Media streaming
  - Data Serving

- **Data Analysis Workloads (OFFLINE)**
  - Big Data Analytics
Why Machine Intelligence in Data Center Networks?

- Demands on data center networking infrastructure increasing
- Evolving Traffic Scenarios
- Uptime Service Level Agreement
Data Center Network Characteristics

- Growth in scale and complexity
- Addition and removal of system components
- Changing execution environments
- Changing workloads
- Updates and upgrades
Data Center Network Dynamics

- S/W upgrades are common
- Protocol evolution
- New applications and services
- Traffic fluctuations - routing anomalies, misconfiguration
- Network migrations, convergence
- Customer change/upgrade services
Data Center Network Outage

- 400+ network failures occur each year in data centers
- Network outages leads to extensive losses due to lack of responsiveness or availability
- Predictive intelligence LEADS TO Reduced downtime & Maximum efficiency
  - Predict Failure/outage in advance
  - Proactively mitigate ill-effects of the anomalous behavior
Applications Experience

- Connectivity disruptions
- Latency
- Throughput issues
- Application Failures
Outage Root Cause

- Device failures
  - Server Host, NIC, HBA, CNA
  - Router, Switch, Load Balancer, Firewall
  - Storage Controller, Disk Array
  - Cable/Optics/Media
- System/Protocol software failures
- Application failures
Machine Intelligence Framework

- Data architecture
- Blend & ingest a variety of structured, semi-structured and unstructured data
- Discover patterns & hidden correlations
- Detect anomalous behavior & Predict failure
- Root cause the anomalous behavior
- Suggest decision options & their implications
Source of Data

- Logs
- Alarms/Traps
- Traffic Statistics
- Counters
- Packet Traces
- IPC Messages
- Network Telemetry
- System Config Files
## Data Description

<table>
<thead>
<tr>
<th>Data Item</th>
<th>DataType</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logs, Traps, Alarms IPC messages</td>
<td>Text and Numerical</td>
</tr>
<tr>
<td>Environmental sensor statistics</td>
<td>Text and Numerical</td>
</tr>
<tr>
<td>Network port traffic Statistics</td>
<td>Numerical</td>
</tr>
<tr>
<td>Protocol packet statistics</td>
<td>Numerical</td>
</tr>
<tr>
<td>Flow level statistics</td>
<td>Numerical</td>
</tr>
<tr>
<td>Network ACLs data</td>
<td>Numerical</td>
</tr>
<tr>
<td>Quality of service (QOS) statistics</td>
<td>Numerical</td>
</tr>
<tr>
<td>System wide data collected by triggers</td>
<td>Text and Numerical</td>
</tr>
</tbody>
</table>
ANALYTICS MODEL BUILDING

- EVENT LOG
- TRACE LOG
- DEBUG LOG

Data Cleansing
Feature Extraction
Feature Representation
Learning Algorithm

- TRAPS ALARMS
- IPC MESSAGES
- TRAFFIC STATISTICS
- CONFIG FILES
- SYSTEM-WIDE DATA OBJECTS

ANALYTIC MODEL
Machine Intelligence Techniques

- Techniques used for analysis and intelligence extraction depends on the nature of data
- Same technique will not work on different type of data
- Data Characteristics needs to be taken into account when applying Machine Intelligence techniques
Sample Machine Intelligence use cases

- Network Traffic Prediction
  REGRESSION

- Session drop prediction
  CLASSIFICATION

- System failure prediction
  CLASSIFICATION
Network traffic prediction

- Time series
- More than one time-dependent variable
- Variables depend not only on its past values -
- But also has some dependency on other variable
- Multivariate Time Series Analysis problem
Time Series

- Trend over time (Ex: Gradual increase/decrease of activity over time)
- Seasonal trend or cycle (Ex: Traffic increases in the morning hours, peaks in the afternoon and declines late at night)
- Seasonal variability. (Ex: Application requests fluctuate wildly minute by minute during the peak hours of 4-8 pm, but at 1 am application requests hardly vary at all)
Anomaly prediction

- Protocols like OSPF/BGP/BFD establish neighbor-ship session as part of their state machine
- Session drops seen due to loss of hello packets
- Result: Outage in the network
- System and network parameters available with monitoring
- Can the outage be predicted?
Anomaly prediction Contd.

- Can be Modeled as a Classification Problem
- Predictor Variables?
  - CPU Utilization
  - Memory Usage
  - Over temperature alarm
  - Kernel Queue drops
  - Hardware Queue Drops
  - Interface Packet Errors/Discards
Machine Intelligence Methods

- Logistic Regression
- SVM (Support Vector Machine)
- Naive Bayes
- Decision Trees / Boosted Trees
- Random Forest
- Neural Networks
System failure prediction

- Log Analytics
- Recurrent Neural Networks
  - Capture implicit system behavior
  - Capture similarity in log entries
Recurrent Neural Networks

Recurrent network

input layer

hidden layers: “deep” if > 1

output layer (class/target)
System failure prediction

- Find the semantic relationship between the words
- Consists of a hidden layer acting as a memory that stores the internal state of the log data
- Training data is used to recognize the failure within the system.
- As new data arrived, the memory is updated, decisions are made according to the current and previous input
Hardware / Software

- CPU
- GPU
- FPGA
- TPU
- Custom Software
- Open Source Packages
Questions ??
THANK YOU