BEZNEXT PERFORMANCE ASSURANCE SOLUTIONS FOR BIG DATA WORLD
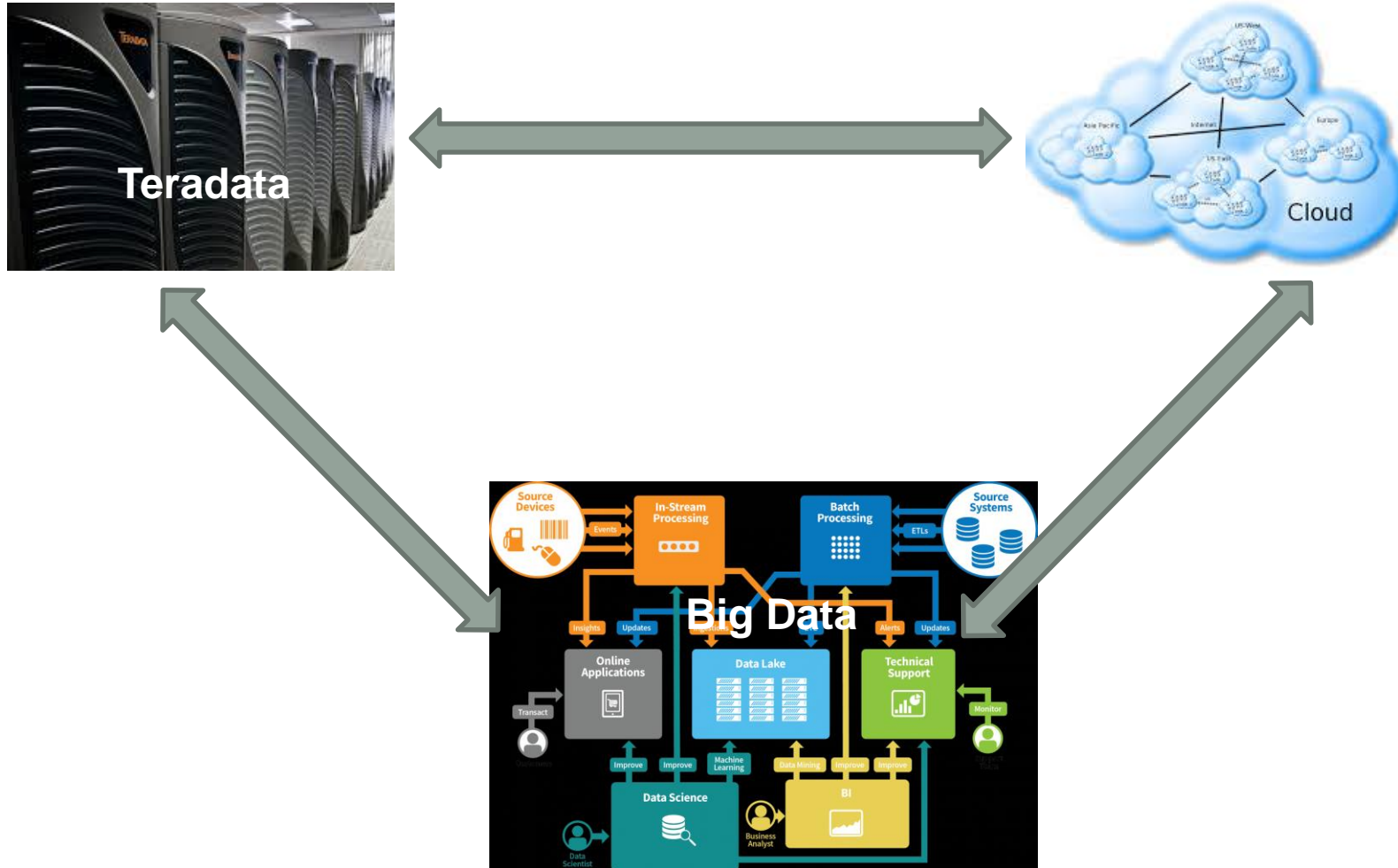
BEZNext

312-752-7766

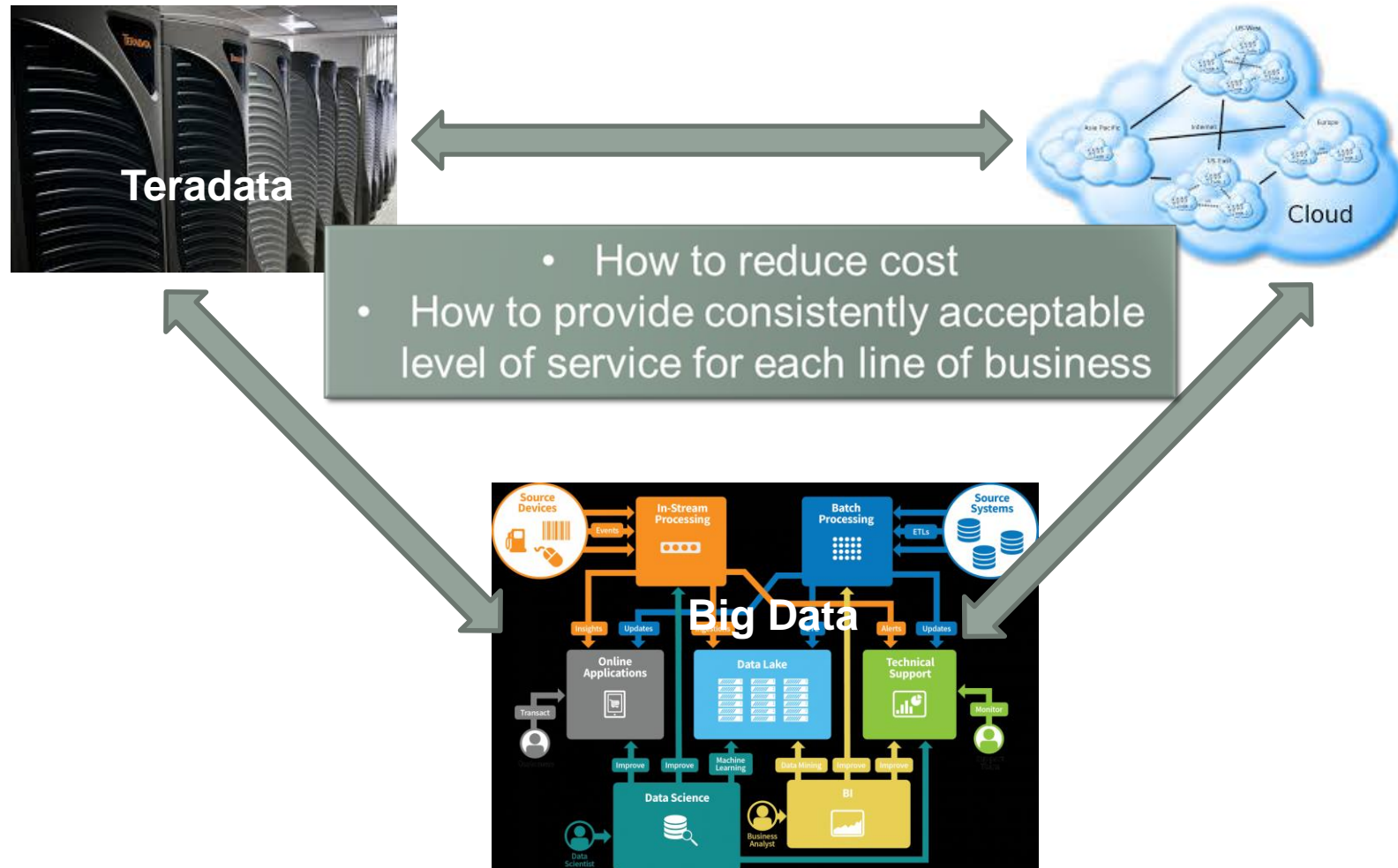bzibitsker@beznext.cor

www.beznext.com

Agenda

- Introduction
- BEZNext Performance Assurance Solutions
  - Performance Engineering
  - Dynamic Performance Management
  - Long Term Capacity Planning
- Demo
- Summary
- Q&A

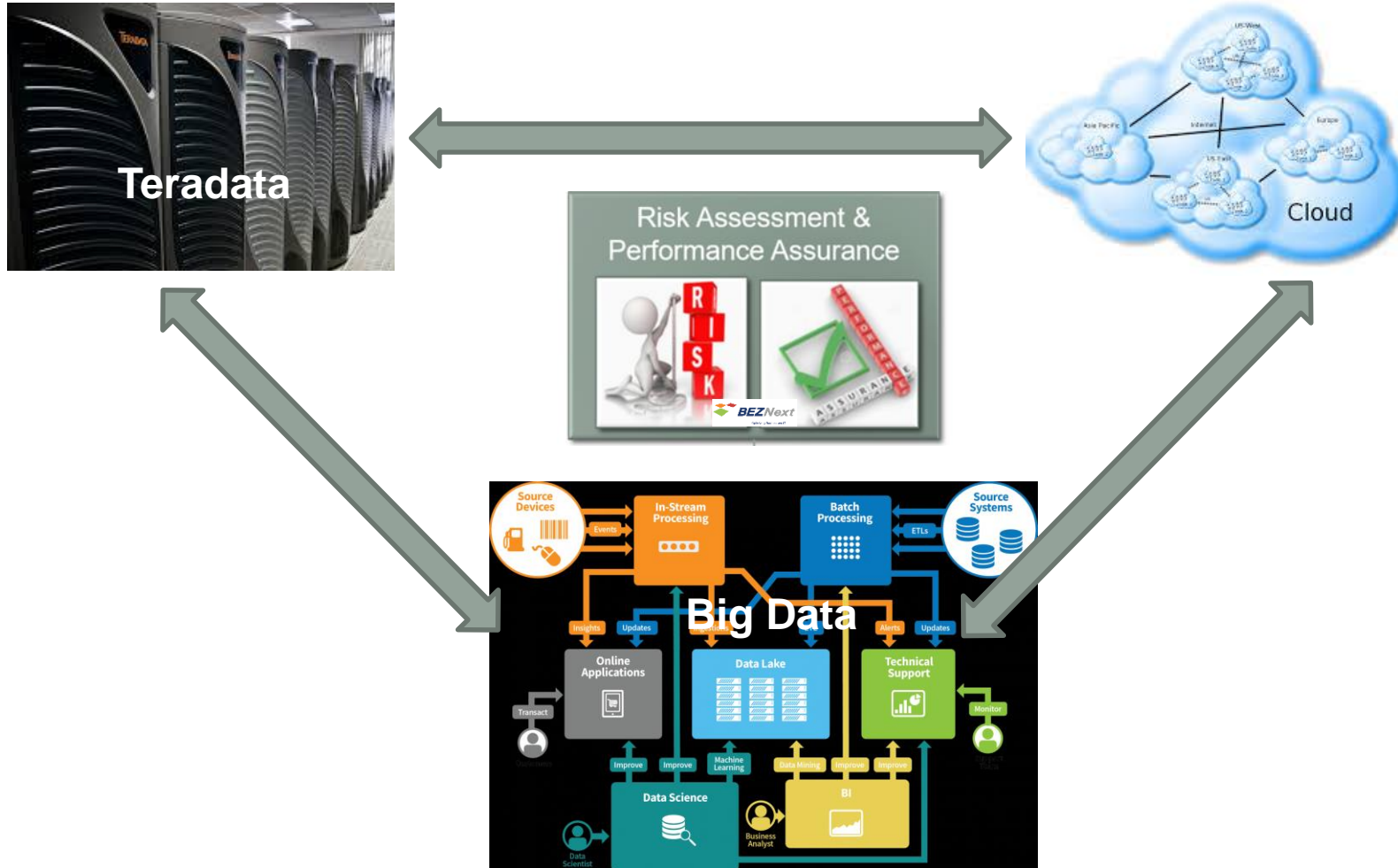# Typical Big Data, Teradata and Cloud Environment

Teradata

Cloud

Big Data

Major Goals

Teradata

Cloud

- How to reduce cost
- How to provide consistently acceptable level of service for each line of business

Big Data

BEZNext Performance Assurance Solutions

Teradata

Risk Assessment & Performance Assurance
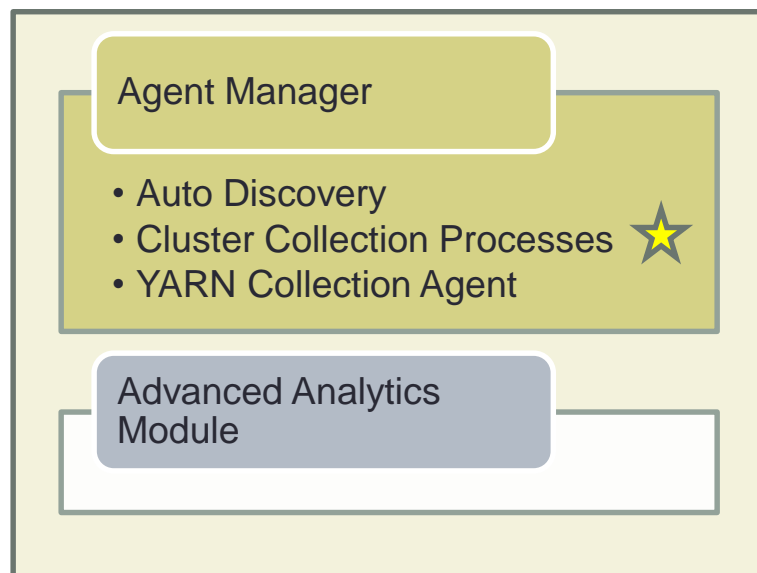
Cloud

Big Data

BEZNext Performance Assurance solutions
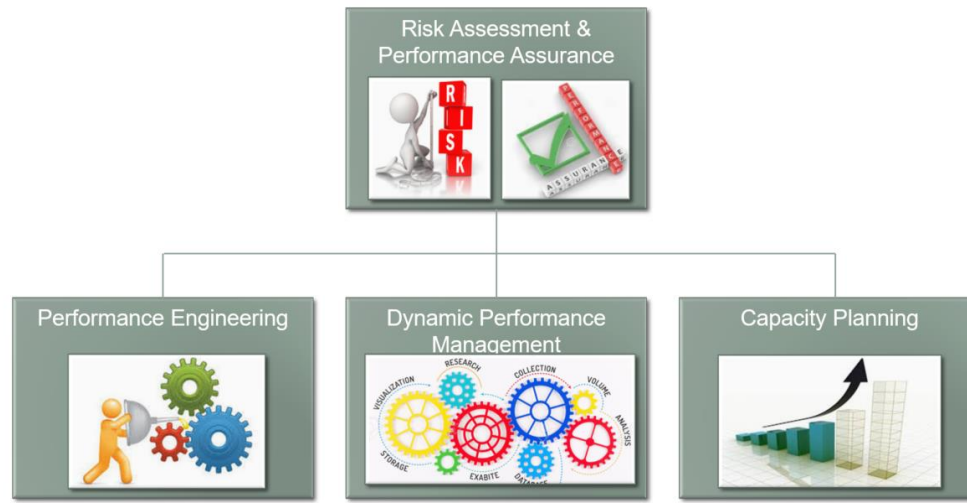
- We incorporate analytical models to automate evaluation of options, justification and verification of decisions

- We will illustrate how BEZNext Performance Assurance Solutions can help you to influence major IT and business decisions

Data Collection Solution Components
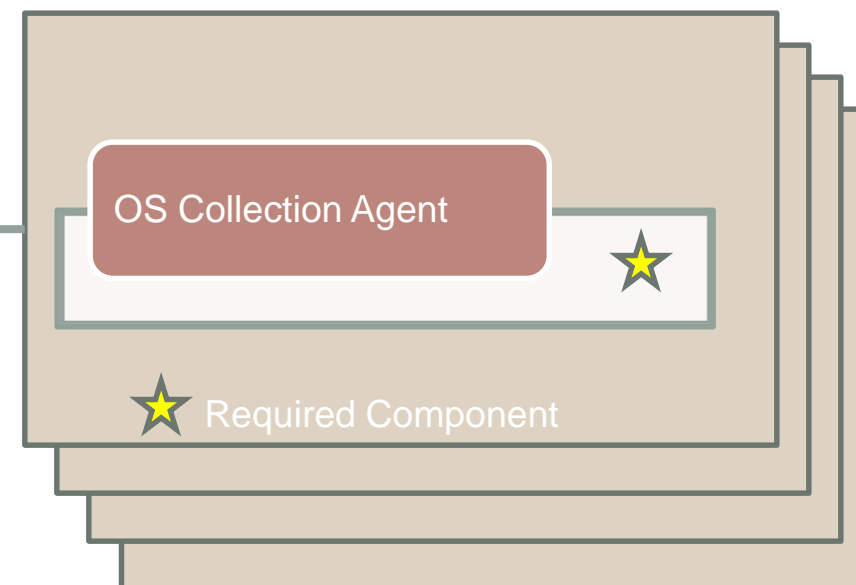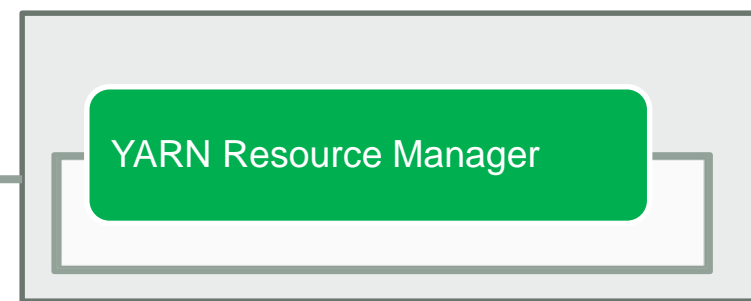
Automatic Configuration Discovery

- Storm
- Spark
- YARN
- Map/Reduce
- TEZ
- RunJar
- HBase
- HDFS
- Cassandra
- Kafka

- Process argument parsing
- Technologies identified with our OS process collector
- Identification rules can be expanded and customized
- to meet the needs of each installation

Identifying Sources of Activity

Advanced Analytics in BV 4.2
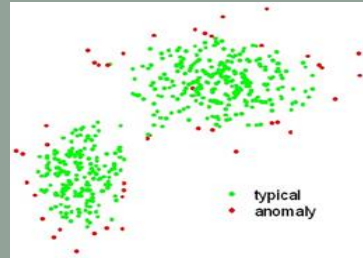
Anomaly Detection

Root Cause Identification

Problem Prediction

Peak Period Analysis

PERFORMANCE ENGINEE

New Application Design, Development and Testing

Selection of appropriate ML Algorithms

How new application will perform in production environment

Operationalization

DevOps

Infrastructure planning to support new applications

Example of Recommendation

- Response Time can vary between 0 and infinity. We transform the response time as 1 / (1 + RT) to make it as a number between 0 and 1, where 1 is better. In addition to calculating the score we check if predicted CPU Utilization and Memory Usage are less than 1.

- Value of score is used to recommend the appropriate ML algorithm and ML Library.

Algorithm	library	pred_score	pred_rank	true_score	true_rank
OLS	Python Sklearn	0.962057911	1	0.936165261	1
OLS	Pyspark ML	0.876712666	2	0.753752225	2
Ridge	Python Sklearn	0.781980143	3	0.725268522	3
Ridge	Pyspark ML	0.722426161	4	0.659234146	4
RF	Python Sklearn	0.476284999	5	0.429752013	5
RF	Pyspark ML	0.465422159	6	0.415271967	6
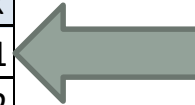
- ML OLS Algorithm using Python Sklearn ML library is the most appropriate algorithm to satisfy business requirements presented in example above.

DYNAMIC PERFORMANCE MANAGEM

Performance Management

- Seasonality Determination
- Anomaly Detection and Root Cause Determination
- Problem Prediction

Workload Management Optimization

- Priorities
- Concurrency
- Resource Allocation

Technology Stack Visibility

Total CPU Utilization by All Workloads
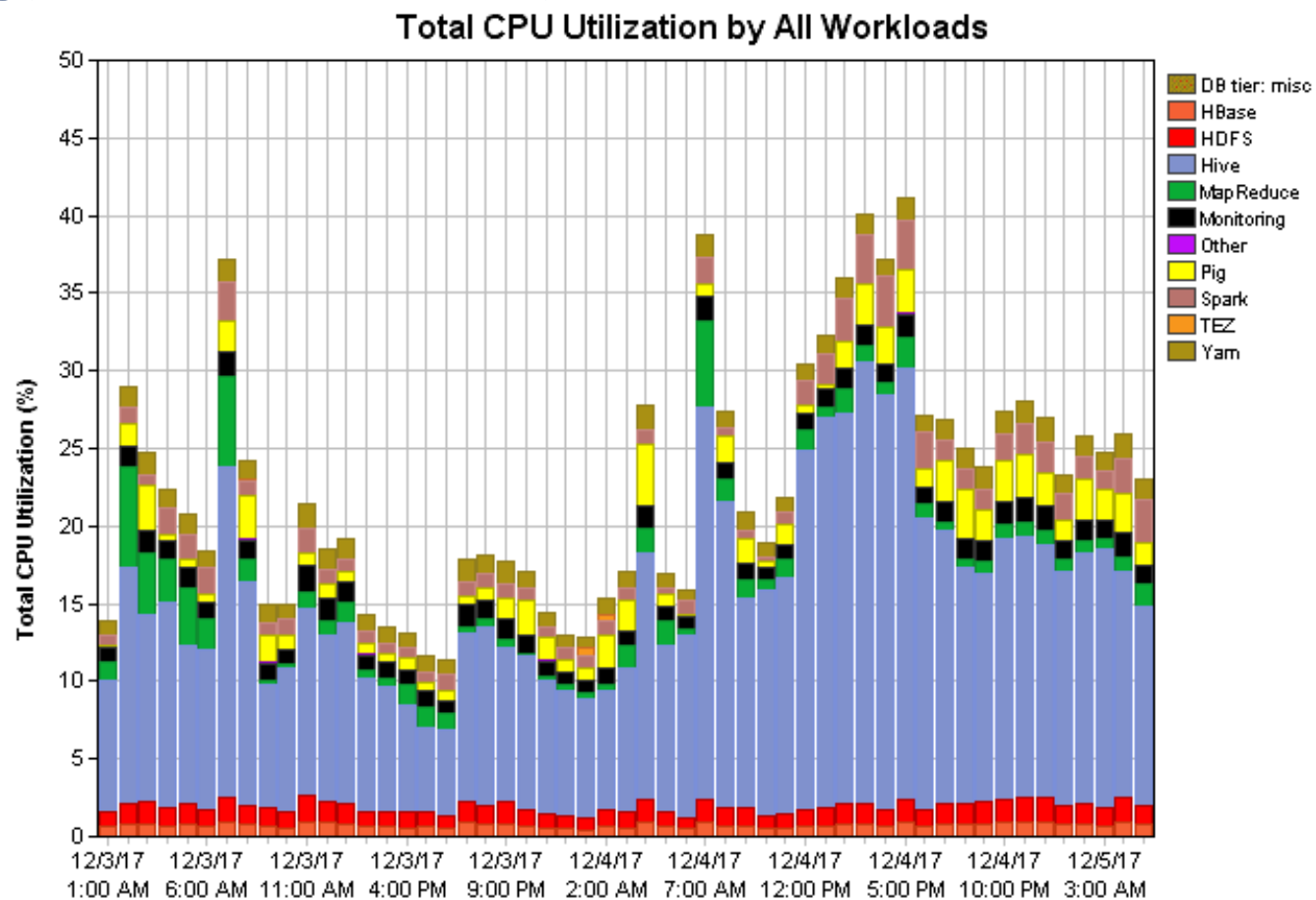
Technology Stack Profile Heatmap

Identification of detailed activity
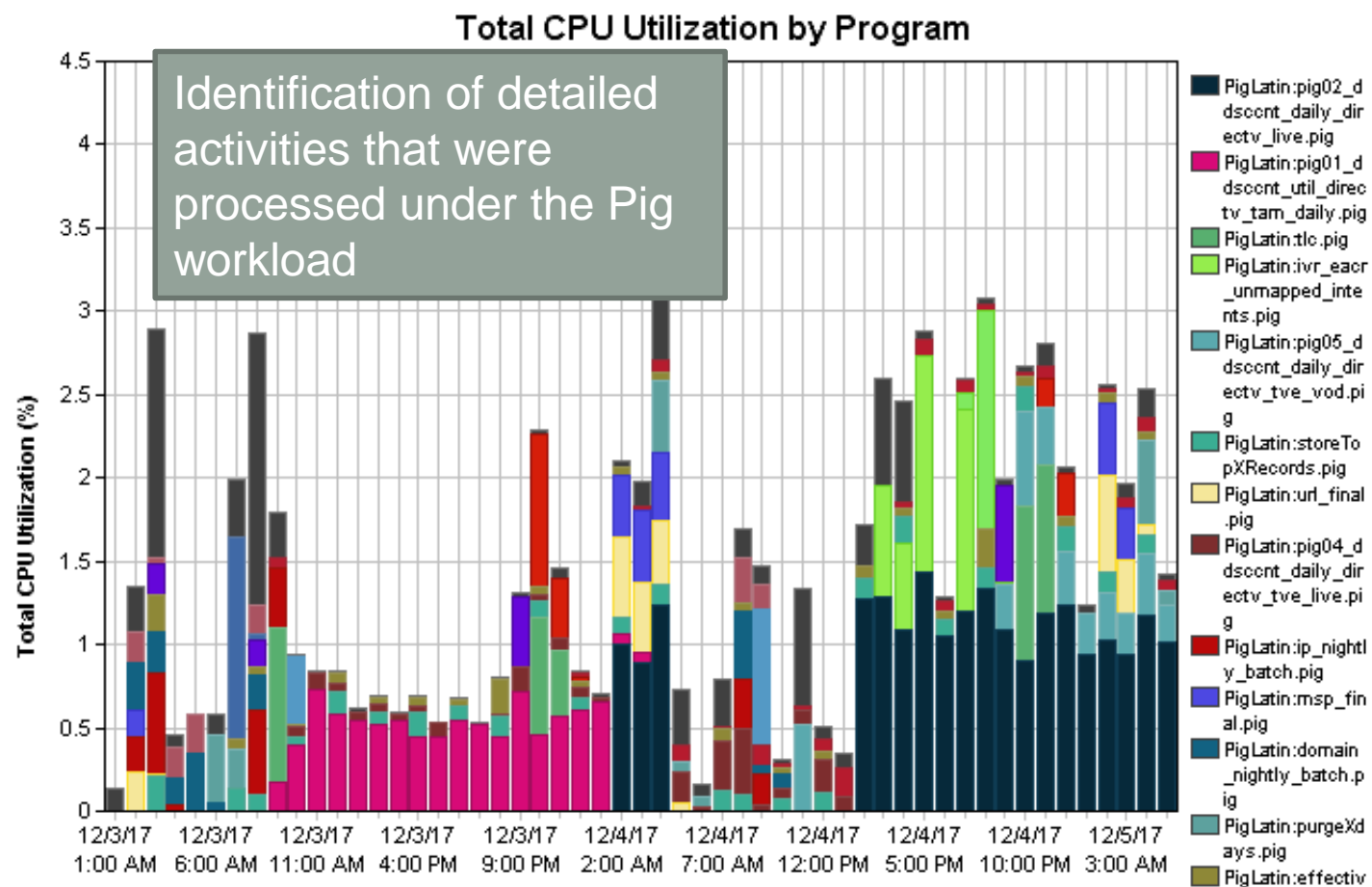
Identification of detailed activity

Total CPU Utilization by User

Identification of user level consumption for the Pig workload

YARN Queue Analysis

YARN Queue Activity

Investigation of activities within a single queue
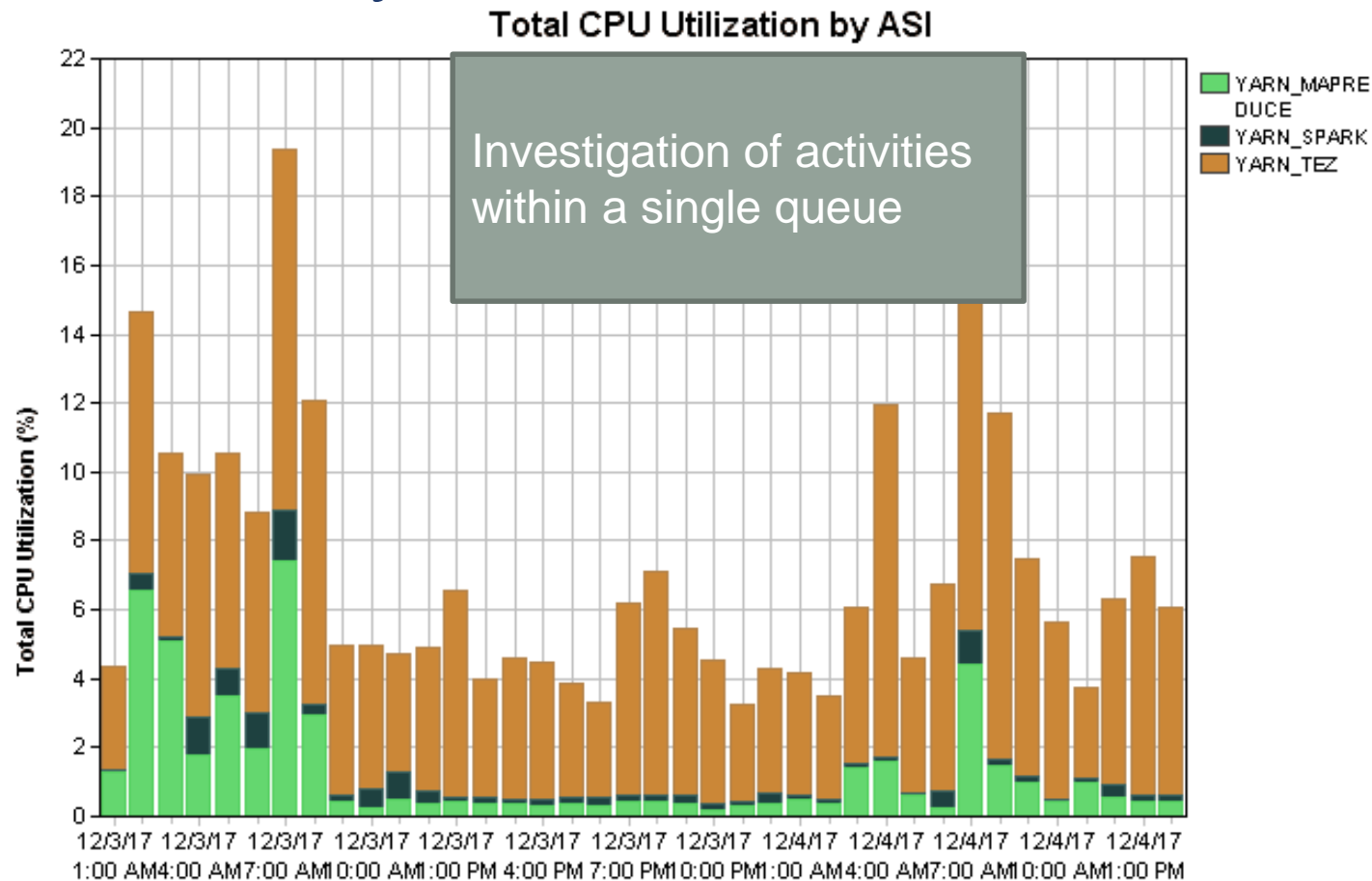
YARN Queue Profile Heatmap

Node Reporting: Cluster Summary

Cluster Analysis for XYZ123 Cluster — Timeframe: Last 24 hours

# of Nodes	Processors	Memory
60	24	48GB
6	48	36GB
1	48	64GB

Nodes in use
67

Balance Ratio
0.74

Averages per Node

CPU Utilization :	43%
I/O Ops per Hour:	256,912
Memory Usage:	24.98GB
etc ...	

Balance Ratio
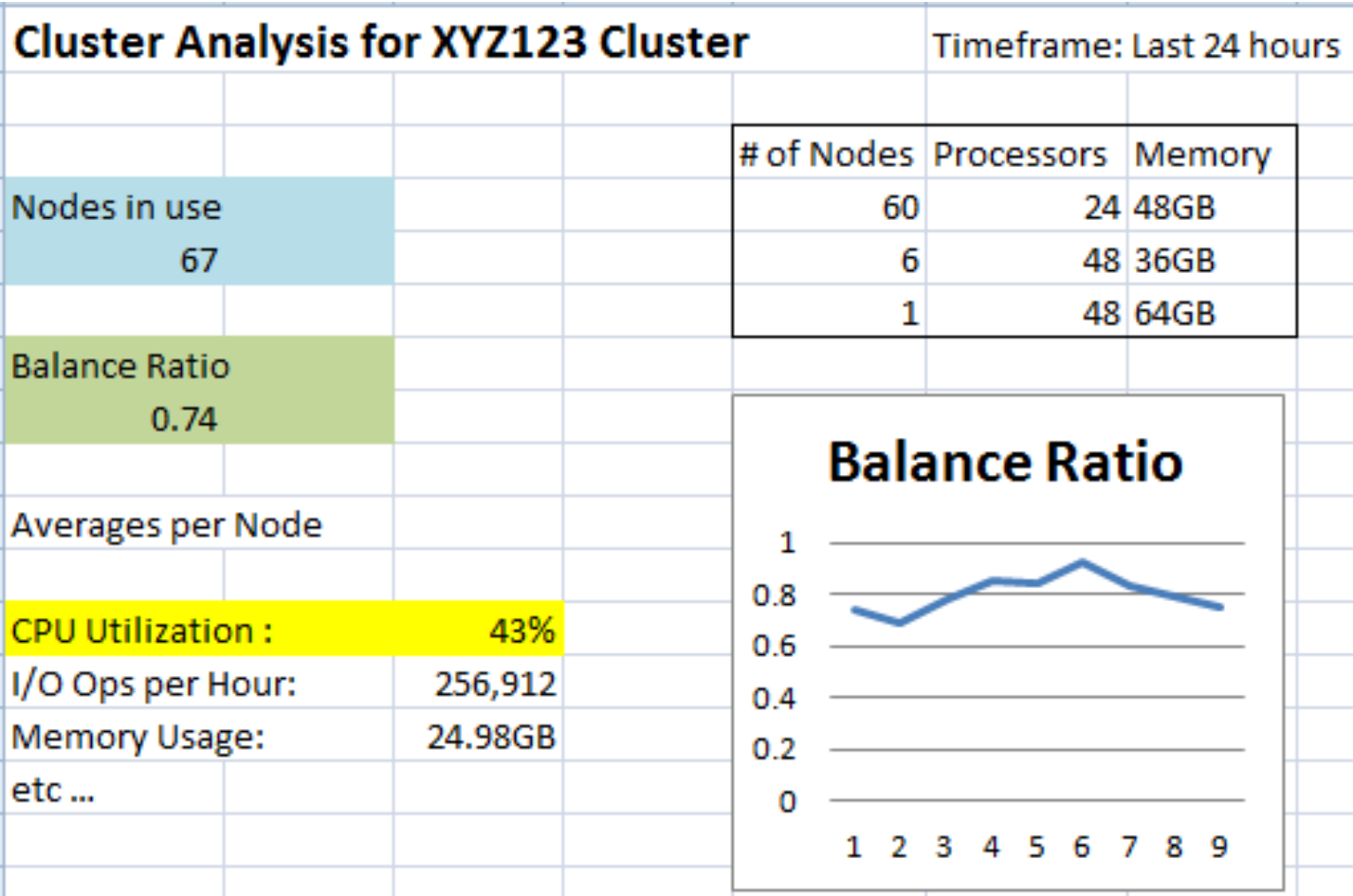
Node Reporting: Detail Analysis

Cluster Analysis for XYZ123 Cluster					Timeframe: Last 24 hours		
					Metric:	CPU Utilization	
Average	Std Dev	Min	Max		**Above Average**		
43	3.15	4%	73%		Node Id	Value	Average
					abc123klm	63	43
					dfs789wes	62	43
					uiy892klu	58	43
					sdf783mnb	58	43
					iop923ghy	56	43
					plu670ety	53	43

Node Breakdown

■ Above Average ■ Average Range ■ Below Average

10%   9%

81%
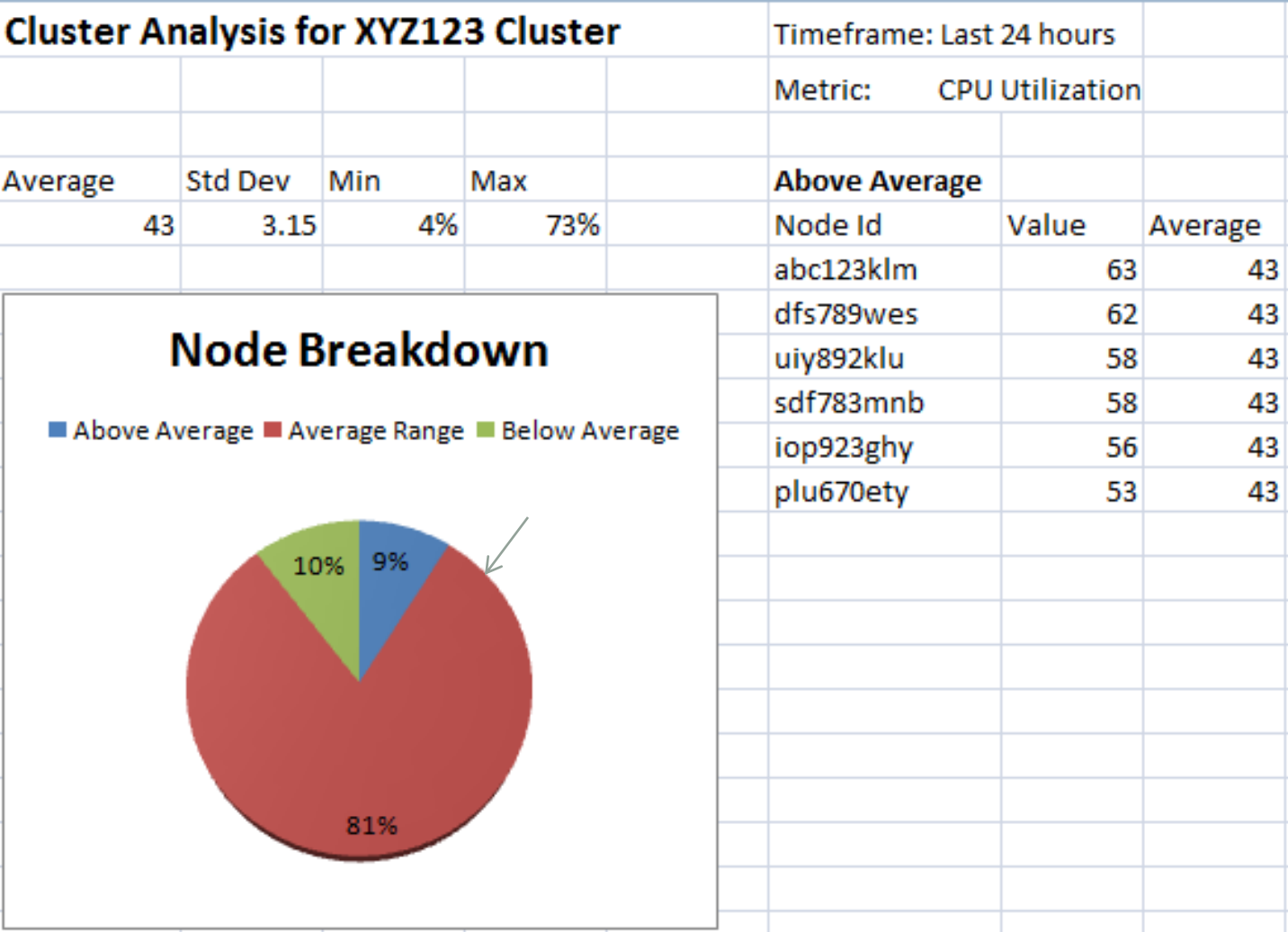
Node Reporting: Technology Usage

Cluster Analysis for XYZ123 Cluster   Timeframe: Last 24 hours

Node Contribution

Sub System	# of Nodes	Min	Max
MapReduce	67	64	67
Tez	25	21	28
Spark	45	40	52
Oozie	60	60	60
Yarn	67	67	67
XYZ	53	52	54

Node Trend for Tez

Series1

Values shown: 25, 26, 21, 20, 25, 28, 26, 22, 22, 21 (across points 1–10)

Node Comparisons for Tez

Node Id	CPU Utl%	I/O Ops	Memory	Requests	etc…
asd234kio	27	234,783	12	478	
sda123lmn	27	189,374	14	809	
etc …					

Resource Allocation & Management Optimization in YARN
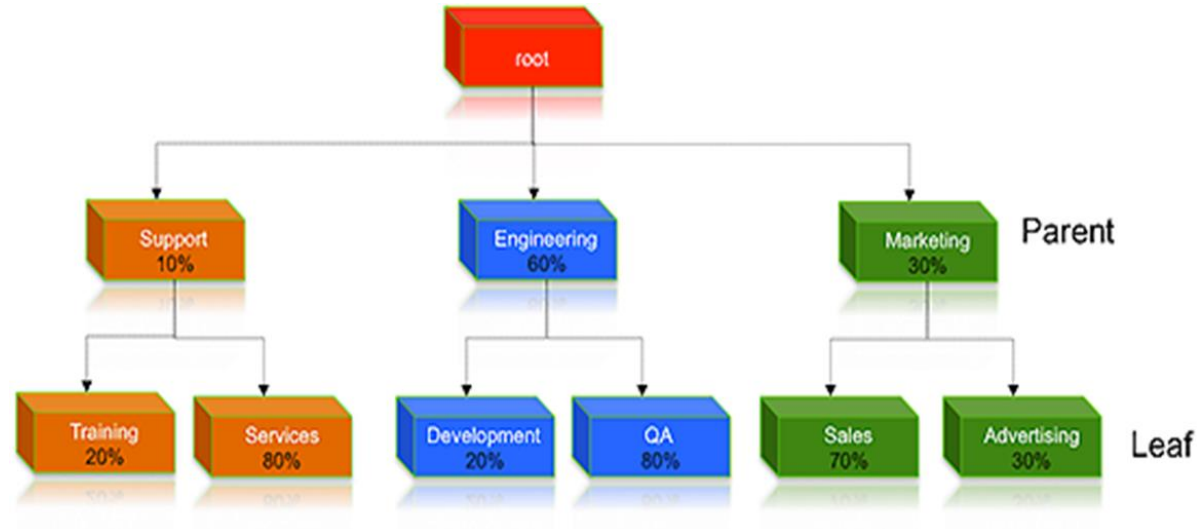
- YARN Capacity, Fair, and FIFO schedulers Rules
- Control tasks, execution and resource allocation
- The resources are divided by LOB or departments and their actual projects
- Incorporation elasticity into the YARN rules
- If resources are available a project that has a need for additional resources can allocate them

Anomalies Detection and Root Cause Determination
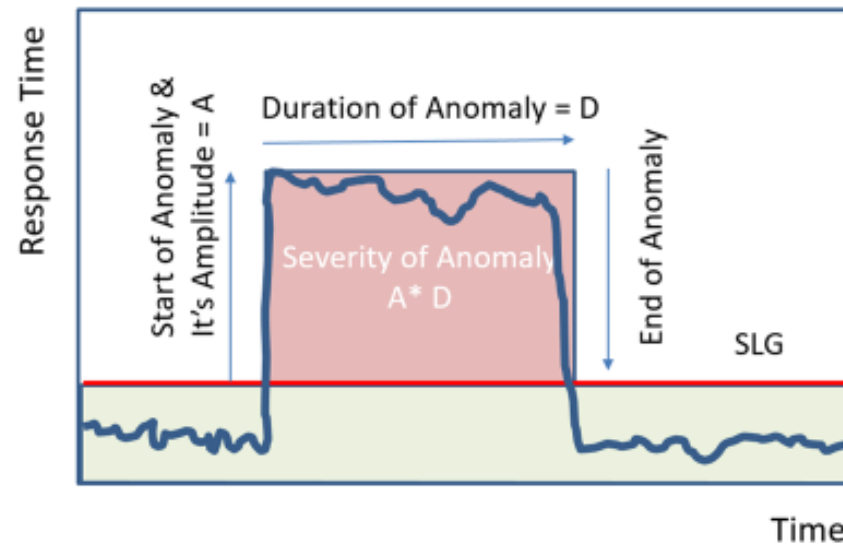
Anomaly Amplitude - A, Duration - D and Severity S = A * D.

Anomaly Detection (Summary View)

Anomaly Detection Parameters

System ID: 1

Ruleset Name: 1

Parameter name: Total Executions Count

Workload Name(s): Select options

Analyzed Hours: 24 selected

○ Interval   ○ Period

Start Date: 2016/01/08 19:00   Time Period: Last Day

End Date: 2016/01/20 15:00

Run

Anomaly Detection Results

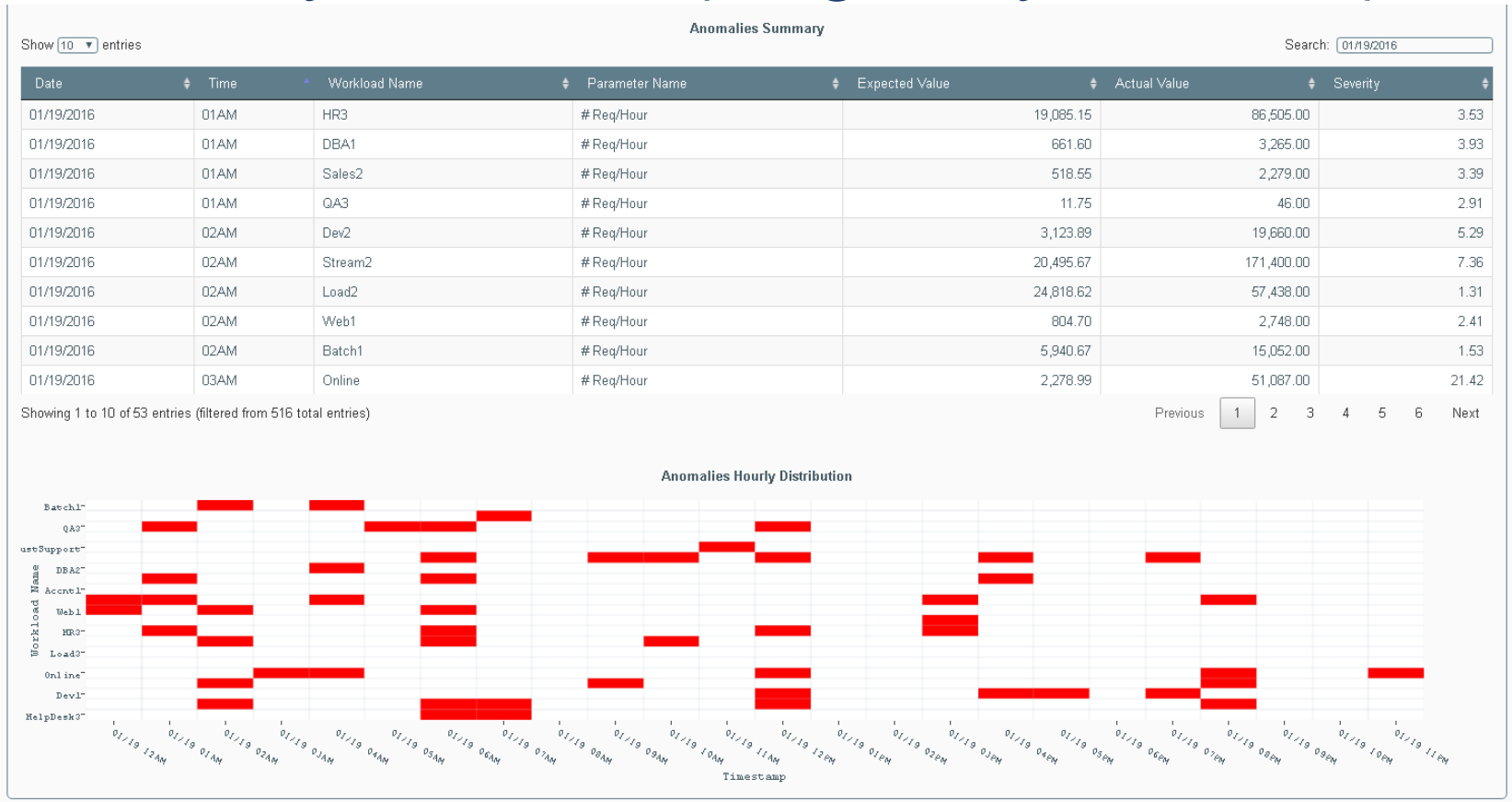Show 10 entries   Anomalies Summary   Search:

Workload Name	01/08/2016	01/09/2016	01/10/2016	01/11/2016	01/12/2016	01/13/2016	01/14/2016	01/15/2016	01/16/2016	01/17/2016	01/18/2016	01/19/2016	01/20/2016
	Details	Details	Details	Details	Details	Details	Details	Details	Details	Details	Details	Details	Details
Accnt1	0	0	1	3	1	2	0	0	0	1	2	0	2
Batch1	0	0	0	1	0	0	0	1	2	1	2	2	3
CustSupport	2	2	3	3	3	7	0	1	0	0	2	1	2
DBA1	1	5	1	3	3	1	0	1	4	1	3	5	4
DBA2	2	3	8	8	13	7	0	0	3	4	6	1	0
Dev1	0	2	7	2	4	2	0	1	1	2	2	4	1
Dev2	1	3	3	3	3	2	0	0	3	0	3	5	2
Finance1	1	6	5	12	6	6	0	1	1	1	2	6	3
HelpDesk3	0	0	2	1	3	1	0	0	0	0	2	2	3
HR3	0	1	1	0	5	2	0	0	2	0	4	4	3

Showing 1 to 10 of 21 entries   Previous 1 2 3 Next

Anomaly Detection (Single day overview)

Anomalies Summary

Show [10 ▼] entries   Search: [01/19/2016]

Date	Time	Workload Name	Parameter Name	Expected Value	Actual Value	Severity
01/19/2016	01AM	HR3	# Req/Hour	19,085.15	86,505.00	3.53
01/19/2016	01AM	DBA1	# Req/Hour	661.60	3,265.00	3.93
01/19/2016	01AM	Sales2	# Req/Hour	518.55	2,279.00	3.39
01/19/2016	01AM	QA3	# Req/Hour	11.75	46.00	2.91
01/19/2016	02AM	Dev2	# Req/Hour	3,123.89	19,660.00	5.29
01/19/2016	02AM	Stream2	# Req/Hour	20,495.67	171,400.00	7.36
01/19/2016	02AM	Load2	# Req/Hour	24,818.62	57,438.00	1.31
01/19/2016	02AM	Web1	# Req/Hour	804.70	2,748.00	2.41
01/19/2016	02AM	Batch1	# Req/Hour	5,940.67	15,052.00	1.53
01/19/2016	03AM	Online	# Req/Hour	2,278.99	51,087.00	21.42

Showing 1 to 10 of 53 entries (filtered from 516 total entries)   Previous 1 2 3 4 5 6 Next

Anomalies Hourly Distribution

Anomaly Detection (Root Cause Details)

Anomaly Detection Results

← Back

Root Cause Details

Workload Cause:	HelpDesk3	First Occurrence:	01/19/2016 6AM
Parameter Cause:	Avg CPU Time	Number of Occurrences:	1
User Cause:	075DA95	Max Severity:	1.0696
Program Cause:	QR1		

Anomalies Caused

Show 10 ▾ entries   Search: [   ]

Date	Time	Workload Name	Parameter Name	Expected Value	Actual Value	Severity
01/19/2016	06AM	HelpDesk3	# Req/Hour	14.98	31.00	1.07

Showing 1 to 1 of 1 entries   Previous   1   Next

Seasonal Peaks

Seasonal Peaks Determination Parameters

System ID: 1   Start Date 2016/01/02 00:00

Ruleset Name: 1   End Date 2016/01/20 23:00

Parameter name: Total Executions Count

Workload Name(s): Select options

Threshold Value: 0.55

Run

Seasonal Peaks Determination Results

Seasonal Peaks Summary

Show 10 entries   Search:

Workload Name	Parameter Name	Peak Type	Peak Start Date	Duration	Avg Amplitude	Standard Deviation	Min Value	Max Value	95 Percentile	Peak Length STD
Accnt1	# Req/Hour	Daily	01/02 04AM	6	451.67	171.56	21.00	656.00	656.00	1.85
Admin1	# Req/Hour	Daily	01/02 05AM	1	8,740.43	3,256.4	1,225.00	11,909.00	11,909.00	0.00
BusDev1	# Req/Hour	Daily	01/02 02AM	1	10,393.75	6,801.83	16.00	23,335.00	20,568.60	0.00
CustSupport	# Req/Hour	Daily	01/02 03AM	1	6,687.75	2,974.55	27.00	15,504.00	7,235.60	0.00
Design	# Req/Hour	Daily	01/02 05AM	1	1,783.20	396.7	420.00	2,461.00	1,693.00	0.00
Dev1	# Req/Hour	Daily	01/02 04PM	1	90.75	30.49	27.00	129.00	129.00	0.00
HelpDesk3	# Req/Hour	Daily	01/02 11PM	1	91.00	30.62	9.00	108.00	108.00	0.50
HR3	# Req/Hour	Daily	01/02 05PM	1	66,391.88	47,305.03	64.00	205,571.00	146,633.30	0.50
HR4	# Req/Hour	Daily	01/02 11AM	1	711.33	387.09	14.00	1,095.00	1,095.00	0.00
Legal2	# Req/Hour	Daily	01/02 12AM	1	1,347.75	538.28	211.00	2,137.00	2,137.00	0.00

Showing 1 to 10 of 24 entries   Previous 1 2 3 Next

Seasonal Peaks

HR3	# Req/Hour	Daily	01/02 05PM	1	66,391.88	47,305.03	64.00	205,571.00	146,633.30	0.50
HR4	# Req/Hour	Daily	01/02 11AM	1	711.33	387.09	14.00	1,095.00	1,095.00	0.00
Legal2	# Req/Hour	Daily	01/02 12AM	1	1,347.75	538.28	211.00	2,137.00	2,137.00	0.00

Showing 1 to 10 of 24 entries   Previous   1   2   3   Next

Daily Peaks
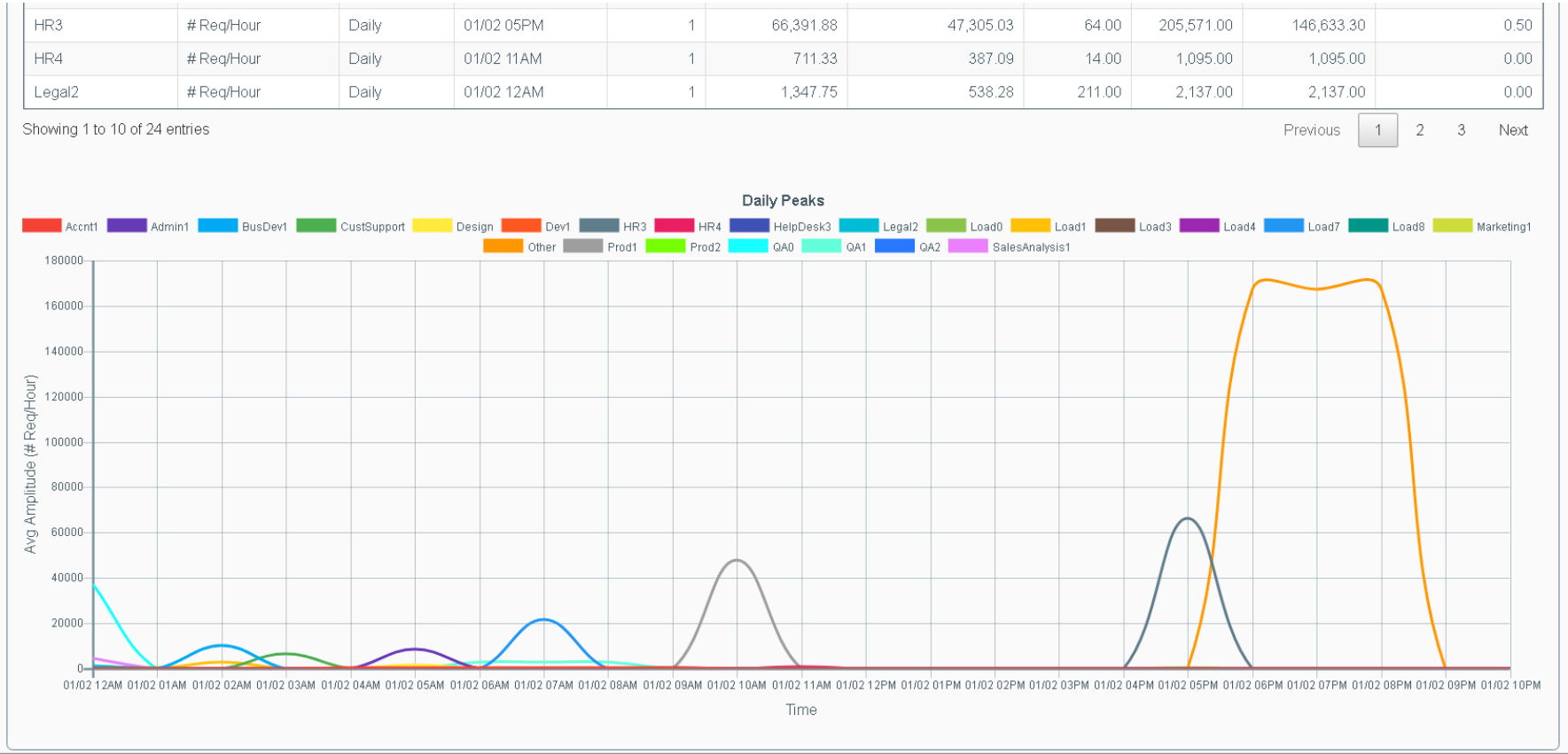
Prediction when SLGs will not be met
What should be changed to meet SLGs?
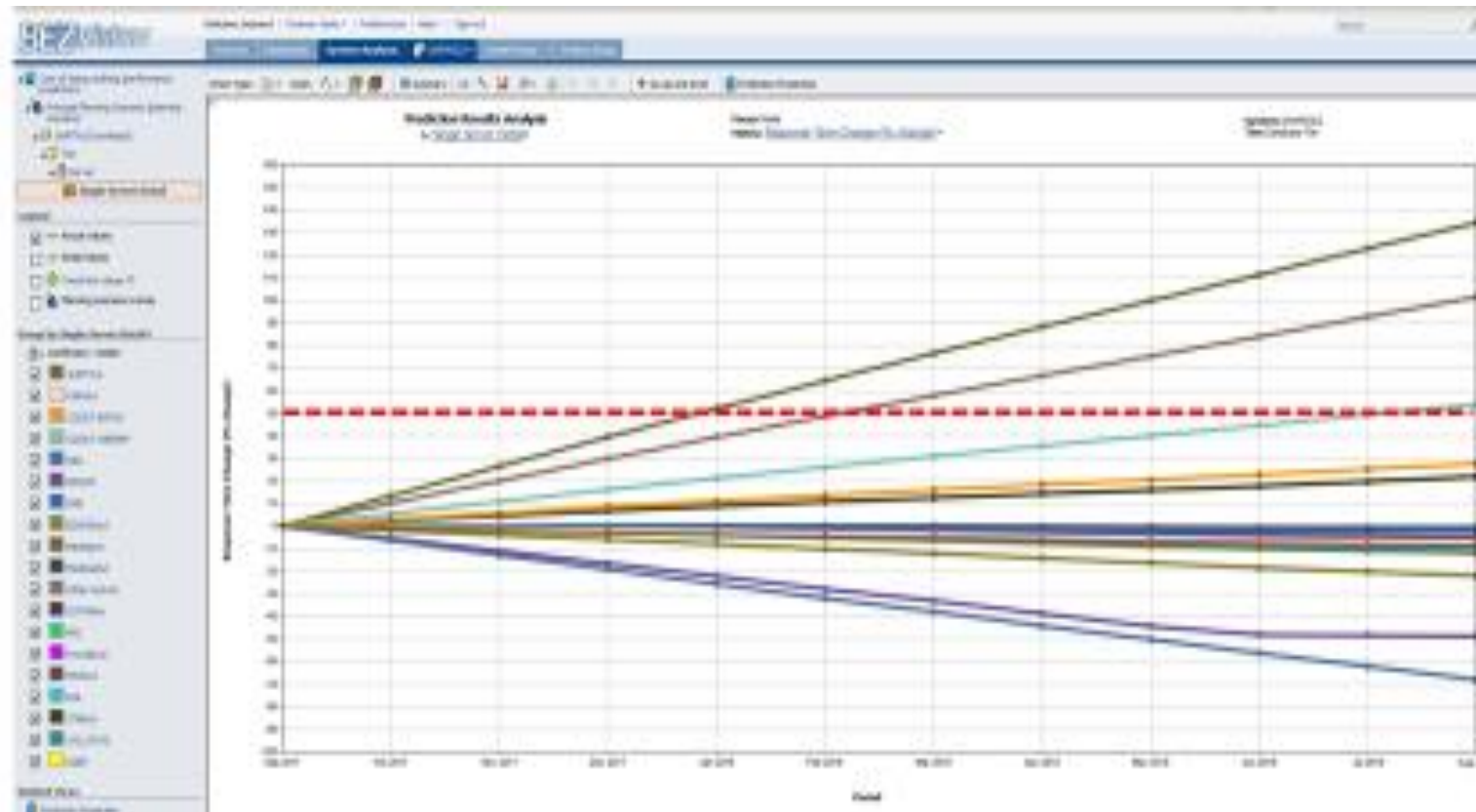
Option 1: Reduce Priority for "Yellow" workload
It will allow to meet SLG for "Blue", but "Green" and "Brown" workloads will not meet SLGs

Option 2: Reduce Priority for "Yellow" workload and increase for "Green" and "Brown"

It will be sufficient to meet SLG for all workloads
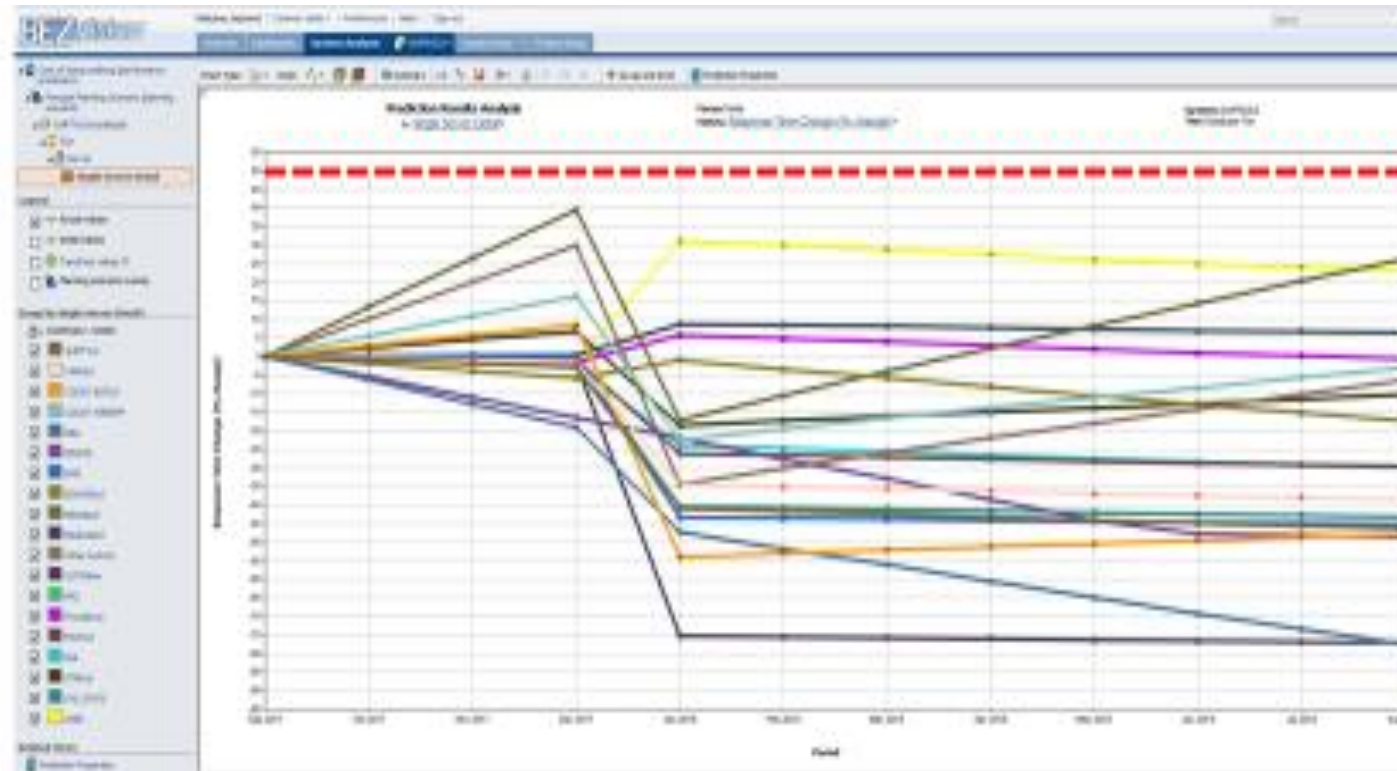
LONG TERM CAPACITY PLANNING

Increase in number of Users and Volume of Data

New application implementation

Moving workloads and Data between Big Data clusters and other subsystems or different processing windows

Moving workloads to another system or different processing window

Enterprise IT Capacity Management

Workload and Data Consolidation

Hardware upgrade justification

Automatic results verification

Capacity Planning Process
Focus on supporting growing and changing demand

- Management
- Business Representatives
- Capacity Planners

Methodology
- Data collection
- Workload characterization
- Auto-discovery
- Seasonality
- Anomaly and Root Cause
- Workload Forecasting
- Scenario planning
- Modeling
- Recommendations
- Verification

According to Model SLGs will not be met in 2 months
How much additional capacity will be required to meet SLGs for all workloads
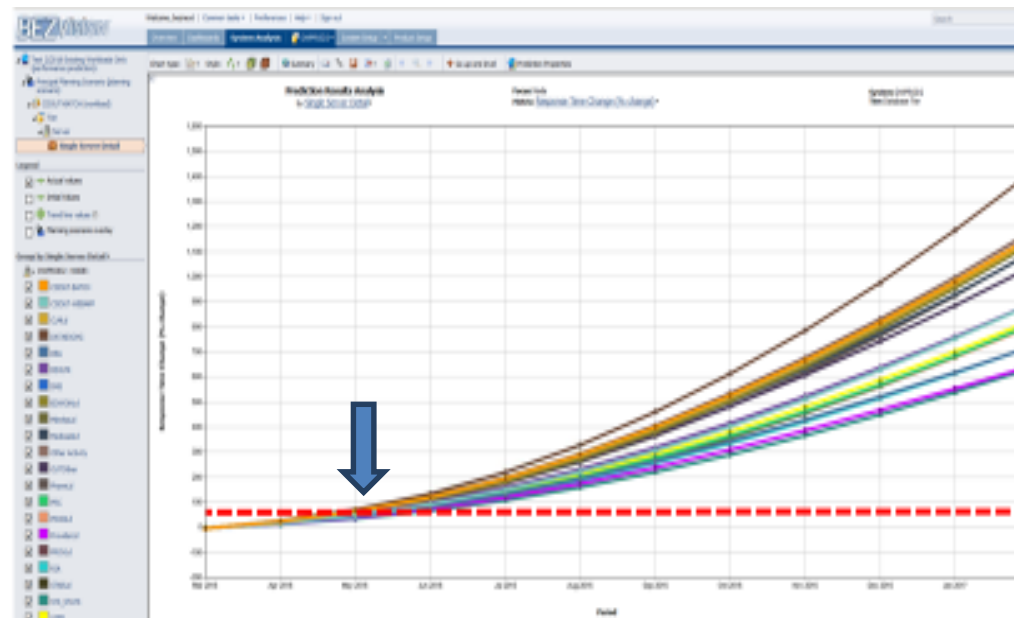
Predicting Impact of New Application Implementation and Development Recommendations
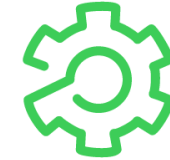
Predicting Analytics

- ### Long Term
  - Queueing Network Models
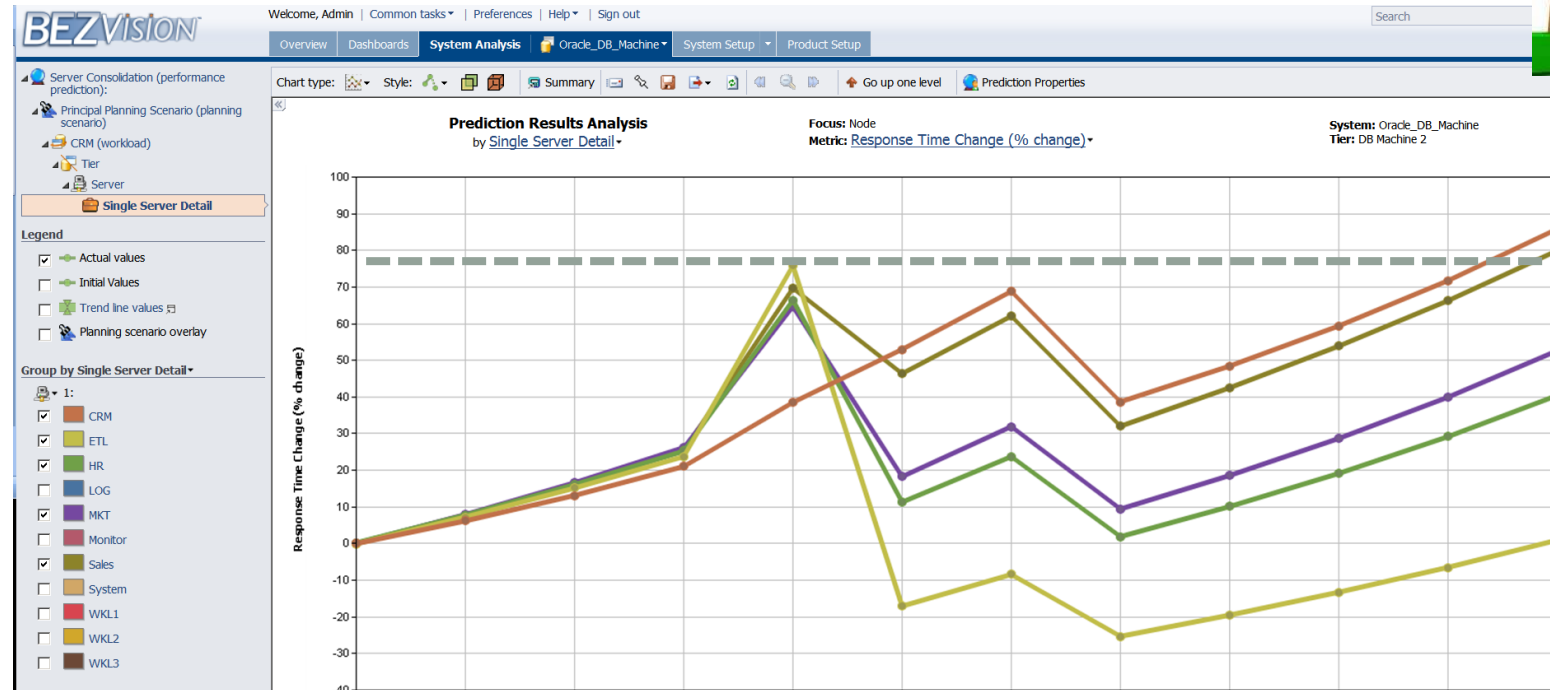- ### Short Term
  - Machine learning algorithms
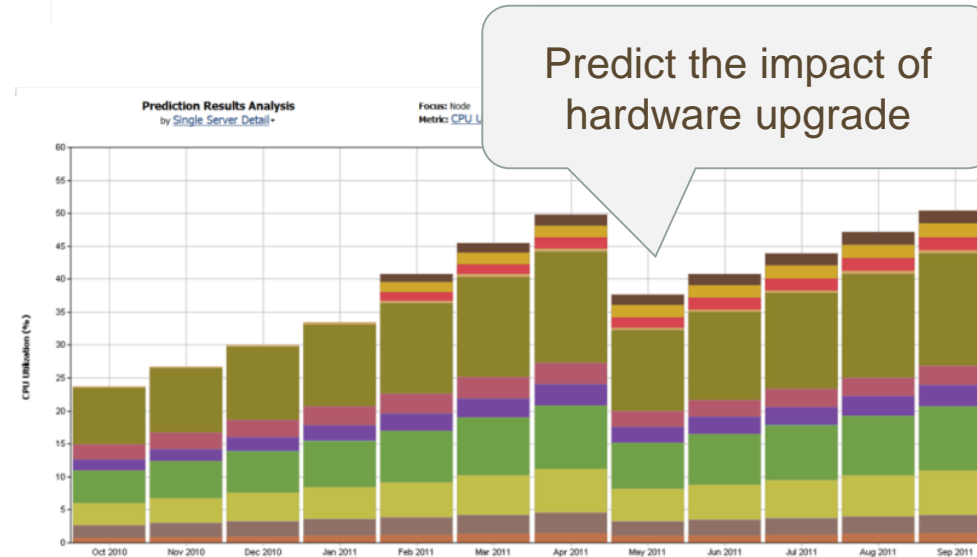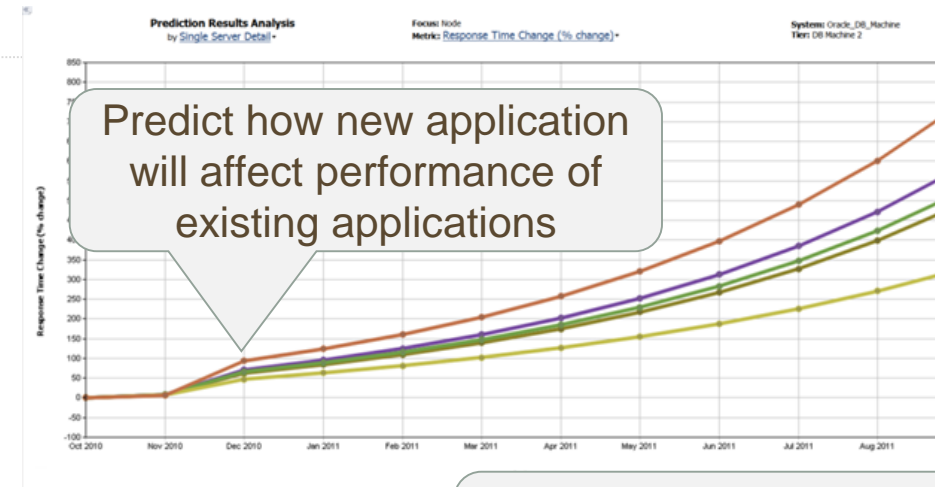
Automation

increase # CPUs from 1456 to 1642 and increase # Disks from 3432 to 3436 are required to meet SLGs for all workloads during next 12 months
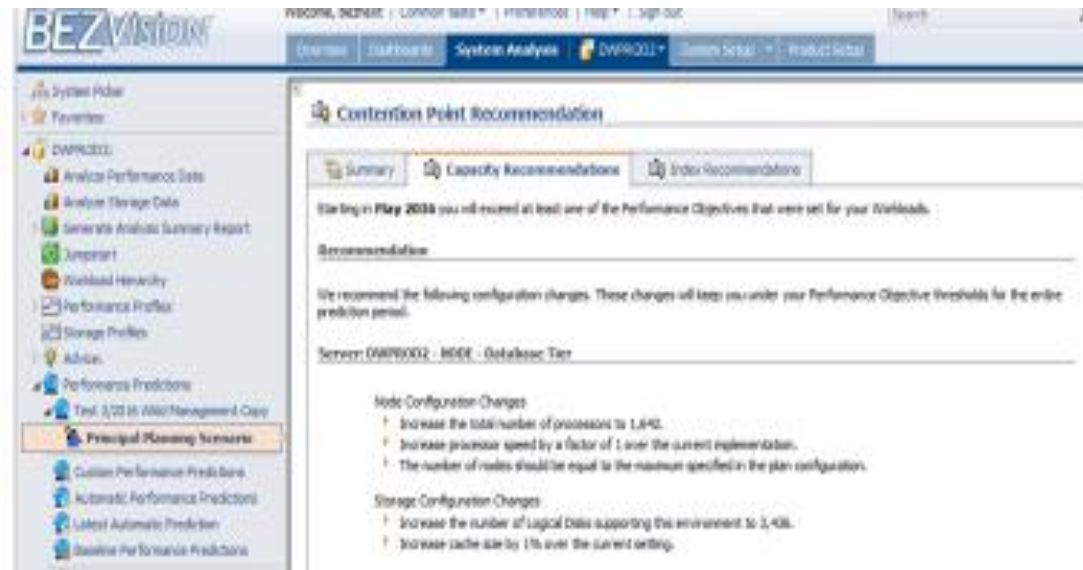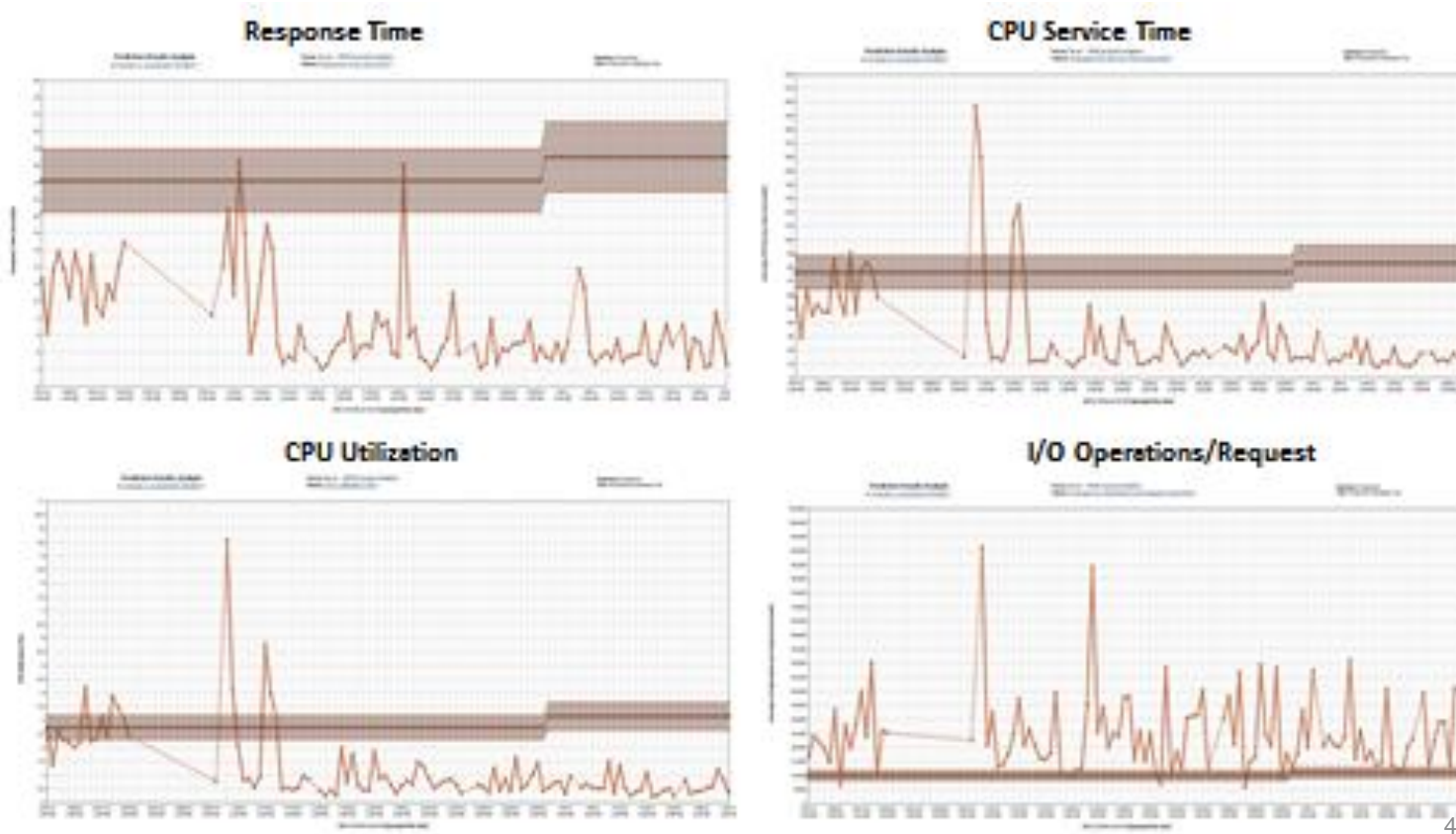
Plan of Actions and Predicted Expectations

Automatic Verification and Feedback Control

Comparing Actual Results vs Expected (A2E)

Uniqueness of our Solutions

- Application of Advanced Analytics on reducing cost and improving level of service to individual Line of Business/Workload
- Hiding complexity
- Reducing uncertainty and Risk while making decisions during Applications Life Cycle
- Flexible Workload profiling by Line of Business, Application, Users, etc.
    - Predicting impact of new workload/application implementation and development proactive recommendations
    - Automatic anomaly detection and root cause analysis
    - Seasonal peaks determination and adjustment of workload management by changing TASM and YARN software parameters, including Priorities, Concurrency and Resource Usage Limitation
    - Hardware upgrade recommendation
- Enterprise vide solution
- Continuous validation of results

Summary

- We Review BEZNext Performance Assurance solutions including use cases for
    - Performance Engineering
    - Dynamic Performance Management
    - Capacity Planning

THANK YOU!

For more information:

- visit our web site : www.beznext.com
- send us inquiries: bzibitsker@beznext.com