Beyond Diagnostics and Tuning: Using Machine Learning and Data Analytics to Maximize DB Performance

Mughees A. Minhas VP, Product Management Oracle Corp.



Copyright © 2017, Oracle and/or its affiliates. All rights reserved. | Oracle Confidential – Internal/Restricted/Highly Restricted

Age	nda	
/ 1	7 Database Management: Landscape and Challenges	
2	Oracle Management Cloud	
3	7 Real-Time Analytics	
4	Anomaly Detection and Machine Learning	
5		

DB Management and Monitoring Trends

- Heterogeneity and hyper-extensibility
 - Ability to monitor and manage open-source, relational, NoSQL databases running on traditional hardware, cloud, or software defined entities (e.g., VMs, Dockers)
- Hybrid/Multi-cloud management
 - Administer databases seamlessly across cloud boundaries offering single pane-of-glass for monitoring and management
- Demand for real-time analytic offerings
 - Problem avoidance (vs. fast resolution)
 - Need for more algorithmic decision-making









Management Challenges





Powered by AI & Machine Learning



Copyright © 2017, Oracle and/or its affiliates. All rights reserved. | Confidential – Oracle Internal/Restricted/Highly Restricted 6

Analytics and ML

Real-time analytics



Utilize DBA resources and time efficiently using big data analytics for smart systems management

Anomaly detection / machine learning

٩

Use machine learning and analytics to avoid problems by predicting them before they occur



#1 IT Trend: AI and Machine Learning widely deployed to make systems intelligent and autonomous Next generation of systems management



Real-Time Analytics



Copyright © 2017, Oracle and/or its affiliates. All rights reserved. | Confidential – Oracle Internal/Restricted/Highly Restricted

Real-Time Analytics: Find Needle in the Haystack Performance Analytics

- Analyze database performance based on long term performance data
- Get insights into database performance
 - Analyze performance degradation by response time
 - Identify inefficient databases by wait time
 - Analyze performance of highly variant SQLs and databases
 - Databases which are increasingly inefficient
 - Top SQLs



Maximize your cloud investments

- Achieve highest level of database performance
- Consolidate databases for better price to performance

Performance Analytics

Analyze performance degradation by response time

- Degradation with decrease in demand
 - Change in Response Time > +20% and Change in Demand < -10 %
- Degradation without change in demand
 - Change in Response Time > +20% and Change in Demand between -10% and +10%
- Degradation with increase in demand
 - Change in Response Time > +20% and Change in Demand < +10 %

Degradation with decrease in demand

1 DBs

Degradation without change in demand

 2_{DBs}

Degradation with increase in demand

1 DBs

Configurable thresholds for "Change in Response Time" and "Change in Demand"

- Default Values:
 - Change in Response Time 20%
 - Change in Demand 10%

Performance Analytics Analyze inefficient workloads

- Inefficient wait time is in Active Sessions which are not part of CPU Time, IO wait Time and idle wait events
- Inefficient (%) = Inefficient Wait Time / DB time
- Increasingly inefficient among the most inefficient databases = DBs with Inefficient (%) > 50 % and Inefficiency (%) Increasing over time

Inefficient databases to focus on



Performance Analytics Analyze workload stability

- Predict the relative variability of a SQL over a period of time
- Focus on databases with variant workload performance

Databases with variant workload performance to focus on



All databases with variant workload performance



Performance Analytics Analyze SQL performance

- Analyze SQL performance based on long term ASH data and across entire fleet of databases
- Helps find systemic root cause of SQL performance issues
 - Compare performance trends across databases, drill-down to execution plans for analysis
 - Key insights into SQL performance issues
 - Trends and forecast SQL response time and other key SQL performance metrics
 - Was there a plan change for the SQL?



SQL Performance Analytics



Deeper SQL performance analysis

SQL Performance Analytics



Compare performance trends across databases, drill-down to execution plans for analysis



SQL Performance Analytics

Identify Top SQLs by CPU and IO and GROUP BY various dimensions

< Top SC	QLs By CP	U											Searc	h By SOL Id
													ocure	in by o'de la
Select a SQL t	o get a quick viev	v of the performa	ince											
Group By	atabases	▼ Siz	e Active Sessie	ons 🔻	Color Perce	ntage Change	-	in CPU Time			-			
Color Key -5	50%	00%												
CRM-SD	CRM-SD CR		CRM-CP CRM-ET			CRM-ED		HR-PROD		FINT	_		SALES-ED	
26jnxb4amf95.	25jnxb4amf95	44jnxb4amf95	43jnxb4amf95	38jn×b4amf95	37jnxb4amf95	23jnxb4amf95	22jnxb4amf95	g3md0paw4vb76		4dptv	1050s 4dptw6yxtqd4m		f3md0p:1w4vb76	15bv3ug0137wf
								ba771d78×th5	1 dn3hkd8	з	2b5p9butuzbqx			1bx4d5jpdfphs
24jn×b	4amf95d24	42jnxb4a	amf95d42	36jn×b4amf95d36		21jnxb4amf95d21				SALE	S-CT		INVCP	
FINP		SALES-NT		INVSD		PROLLT		PAYD				2ad9fnk42p4ba	49jnxb4amf95d49	48jnxb4amf95d48
									. 31jnxb4amf95	epdz	86t2×qw09			
		41jnxb4amf95	40jnxb4amf95	35jnxb4amf95	34jnxb4amf95	29jnxb4amf95	28jnxb4amf95					9sztf6dz1p957		
47jnxb4amf95.	46jnxb4amf95							32jnxb4amf95		5			OrderDB	CRM-ST
										SALE	S-SD			
													tj5uv3muk0122	9kda1k5mzvwzw
										Gir	Siguita tem Disatu 20jn×b	20jnxb4amf95d20		
45jn×b	45jnxb4amf95d45		amf95d39	33jnxb4amf95d33		27jnxb4amf95d27		30jnxb4amf95d30			ChixovaniisSuv		6b5p9butuzbqx 7d	·· 0nvg7s 5wz
< Tor	SQLS B	v I/O												
		J												SQL Search
Select a s	SQL to get a q	uick view of th	ne performan	ce										
Group E	By Database	s	 Size 	Active Sess	ions	Color	Percentage	Change	-	in	I/O Time		-	
Color Ke	ey -50%	100%												
PAYD	-		SALES-NT			CRM-ED			FINT			SALE	S-ED	
32jn	32jnxb4amf95d32 31jnxb4amf95d31		41jnxt	xb4amf95d41 39jnxb4am xb4amf95d40		23jnxb4amf95d23					4dptw6y×t	qd4m	15t	v3ug0137wf
31jn			40jn×t					21jnxb4am	g3md0paw4v		2b5p9butt	uzbqx	1b:	<4d5jpdfphs
											SALES CT		INVCE	

Utilization Analytics

- Analyze current utilization of the hardware & software
- Understand future resource requirements for host, VM based on current consumption
- Forecast future database resource requirements based on current consumption
- Consolidate additional databases or workloads on hosts



Ad-hoc Performance Analytics ASH Analytics

- Performance activity data gathered sampling technique
 - Implemented in the database kernel for Oracle
 - Similar sampling technique for non-Oracle Databases
 - Efficiently handles volumes of data by down sampling appropriately
- Data sampled every 1 second
- DB time as a key measure and common currency cross databases for SQL statistics
- Active sessions for DB Activity based stats



Ad-hoc Analytics Demo



Ad-hoc Analytics

- I have a CDB with several PDBs in it
- DBA gets complaint from some users of poor performance
- Uses ASH Analytics to diagnose problem root cause





Reactive Performance Management

Analyzing transient performance problems using ASH Analytics





ASH Analytics identifies User I/O as the problem

SQL ID	SQL ID v by Wait Class Schedule SQL Tuning Advisor Create SQL Tuning Set Identify SQL's that are subject to User I/O									
Select	SQL ID	Activity (Average Active Sessions)	User Session							
	15zzdzyvmp3tv	2.66	2:1101,85	.03						
	ghmpac93jgwbv	1.94	2:1184,4573	.03						
	anmbr78ay04w9	1.86	2:1193,139	.03						
	6duq7b:6onpa	1.68	2:120,49	.03						
	8yyadhv3bnn78	1.57	2:1296,89	.03						
	5m48s8x3mfhd7	1.45	2:1383,365	.03						
	fnauxa95hvm9u	1.27	2:1393,79	.03						
	cdh3v5ss7w00b	1.21	2:19,111	.03						
	3mysk9479s1d8	1.12	2:207,325	.03						
	gngdmdc70fqqf	1.04	2:298,3995	.03						

Anomaly Detection and Machine Learning



Smart Monitoring with Anomaly Detection

- Baselines automatically calculated to determine normal behavior
 - Tracks and adjusts for daily and weekly seasonality
 - For example, daily seasonality: Load from 9-10 am is expected to be higher than 6 pm – 7 pm
 - Weekly seasonality: Load on Mondays 9 am 10 am is expected to be higher than Fridays 9 am – 10 am
- Auto-detection of performance anomalies
 - Anomaly = value outside expected baseline
 - Alert when performance is abnormal
 - Metrics from related entities, highly correlated with metric alert



Diagnose Errors Using Machine Learning Analyze and explore log data efficiently

- Specialized machine learning algorithms enable smart clustering of log events based on common signature
 - Reduce millions of log events into smaller set
 - Enables rapid troubleshooting by quickly identifying outliers and anomalies
 - Identify constants and variables in log records; bucketize variables for analysis
 - Topology-aware log exploration
 - Example: Cluster log records from application server and associated databases
 - Show trend of each clustered group; correlate clustered events that show similar trends
 - Example: Storage block corruption log records and DB write errors rising at the same time

Aug 18 11:00:57 Unable to create file ABC.txt for user root Aug 18 23:07:26 Unable to create file DEF.txt for user larry Aug 18 23:08:30 Unable to create file XYZ.ppt for user moe Aug 18 23:08:22 Unable to create file ZZ.top for user curly Aug 18 23:08:01 Succeeded authorizing right 'system.priv.admin'



Reduce the above log entries to a pattern and an anomaly **Pattern**:

<Date> Unable to create file <File-Name> for user <User-Name> Outlier:

Aug 18 23:08:01 Succeeded authorizing right 'system.priv.admin'





Copyright © 2017, Oracle and/or its affiliates. All rights reserved. |

Hybrid Cloud Management



Fleet Management

- Single pane of glass management view for all databases on-premise and cloud
- Highlight active services, services with problems (alerts, failed backups), workload spikes, workflow status, etc.
 - Identify needle in the haystack outliers based on CPU, I/O, memory and other important statistics
- Drill-down to respective single service pages for active management operations, troubleshooting and diagnostics
- Support for active operations at fleet level
 - E.g., set an init.ora parameter for all databases supporting an application



High Availability

- High availability fleet page
 - Protect your data by enabling business rules
 - Availability status of all managed databases
 - Backup/recovery related alerts for drill down
 - Identify databases that need attention if meeting recovery window and unprotected data window goals
 - Supports Active operations at fleet level
 - Example: Backup all development databases as a single operation



Data Movement

Managed Workflows facilitate workload movement from on-premise to cloud



Move To Oracle Cloud

Lift and Shift workloads to Oracle Cloud with continuous and scheduled refresh of data from test master



Clone production systems to cloud, track cloning lineage, data masking and TDE support for sensitive data

Lift and Shift from On-premise to Cloud





Test/Dev Creation and Data Refresh



ORACLE

Copyright © 2017, Oracle and/or its affiliates. All rights reserved.

Modern DB Management Using ML and Analytics

Comprehensive Monitoring



Hybrid cloud management

Gain 360-degree insight into the Oracle DBCS performance, availability & capacity planning



Designed to support IT operations, DevOps, LOB, and C-level executives

Real-time Analytics



Anomaly Detection through ML



Hybrid Cloud Management



DB Management: Solution Overview





New Capabilities for Autonomous Database

- Move databases from on-premise to cloud using automated managed workflows
- Monitor key application metrics with alerting and anomaly detection
- Diagnose and improve application SQL logic
- Analyze logs, performance and utilization trends based on historical data

