Scalable Storage Analytics

Sushant Sawant  
Principal Consultant, Hitachi Data Systems  

May 20, 2015  

© Hitachi Data Systems Corporation 2015. All rights reserved.
1) Big Data Analytics Platform
- MARS is the scalable performance analysis product which enables analysis, assessment and recommendations on how to improve infrastructures. Storage performance,

2) Intelligent Storage Analytics Solution

Measure
Configuration, Performance and Event information from
- Storage

Analyze
- Performance analysis of Storage environments at granular level
- Long term (3+ years) performance trending of Enterprise environments

Report
- Based on industry best practice guides and our own base-lining techniques

Solve
- The most difficult performance issues in a timely manner

Key Features

Common Data Platform
- One place for Hitachi Data
- Extensible ETL Support

Custom Reports
- Best Practice Canned Reports
- Extensible BI Platform

Volume, Variety, Velocity
- Volume – number of metrics, frequency, retention time
- Variety – Open Systems server + storage, mainframe + storage, applications to storage
- Velocity – optimized for storing data and running analytics.
Hitachi Data Center Analytics (HDCA)

- **HDCA is an integral component of the Hitachi Data Systems (HDS) Big Data Strategy and Portfolio:**
  - **Volume, Granularity and Variety:** Discover, store and analyze a variety of data from HDS platforms and Software Suites, for extended periods of historical data.
  - **Agile, Ad-Hoc, Basic Reporting:** Provide agile, flexible reporting framework with standard reports and built in custom report capability.
  - **Data Latency (1 day old):** Because it takes time to load and process data into the warehouse, HDCA data, especially performance data, will be up to one day old.
  - **Reuse by HDS Software Suites:** Stored Big Data sets can also be “reused” by existing HDS software management suites for more value-add features.
  - **BI Integration:** As well as providing standard “canned” reports and a custom report capability, it is also be possible to use BI tools such as Actuate BIRT OSS for reporting and visualization capability for customers already using BI tools.
  - **Big Data and Hitachi Command Suite (HCS) Integration:** Ability to integrate with both HCS and other Big Data datasets.
Scalable Storage Analytics

Are you challenged monitoring storage performance for your critical business applications as your data infrastructure continues to grow? Hitachi Data Center Analytics provides a new efficient approach to analyzing historical performance trends across your storage infrastructure.

- Build a scalable and flexible data repository for your storage analytic needs
- Collect highly granular performance statistics
- Report on storage performance trends spanning multiple years
MARS has two components, both are installed as virtual machines

**Probe**
- Gathers performance and configuration data from targets
- Transfers data to the MARS server

**Server**
- Receives the data from the probe
- Processing, analysis and reporting
HDCA Server Architecture

- Web Browser / Applications (Bird’s Eye)
- FTP Server
- DB (Proprietary)
- Web Request Handler
- Probe Data Importer
- Query Processor
- Data Access API
- Result Cache
- Report Engine
- Report Scheduler
- Import Scheduler
- Cache (Configuration/Performance)
- Result Data (XML / JSON)
- Excel, Word
- Uploads compressed and encrypted data

REST API (async)
HDCA Probe Architecture

- Web Browser
- FTP
- Management Wizard
- Data
- Uploader
- SMI-S Storage Probe
- NTAP Probe
- vCenter Probe

SNIA CIM API
OnTAP API

Solutions Enabler with SMI-S

© Hitachi Data Systems Corporation 2015. All rights reserved.
Flexible Deployment Model

Private Cloud

- Server
- Probe

Public Cloud

- Server
- Probe

Hybrid Cloud

- Server
- FTP
- Probe

Customer’s Infrastructure
BIRT is an open-source flexible reporting system that integrates with existing Java application to produce compelling reports.

BIRT integration with MARS allows users the flexibility to create their own (custom) reports using data available in MARS.
Navigation

• Different types of Storage Systems are grouped separately (since some features are specific to certain types of storage systems)

• Relevant Reports are displayed at both the Node level (e.g., Pools) and at the specific Instance level (e.g., Pool 1)

When there is a large number of items to be displayed, HDCA intelligently displays only the first few items. This minimizes the unnecessary processing for displaying items that may not be required. Other items are displayed when users clicks on it.
The above Report shows the IOPS vs. Response Time for a DP Pool.

The Reports & Charts are interactive charts to make it the most useful for an Analytics tool.
The above Report shows the IOPS vs. Response Time for a DP Pool.

The Reports & Charts are interactive charts to make it the most useful for an Analytics tool.
Creating your own Queries

Users can create their own queries by any of the methods below:

a. Start with a ‘pre-defined’ query and extend it as required
b. Build the query using the Graphical Query Builder
   
   a. Use any of the pre-defined queries as a ‘starting point’ and then make required changes
   
   b. Build the Query using the Query Builder
   
   c. Directly enter the MQL in the Query window and ‘Find’
Query Builder enables the user to write complex queries using simple point-and-select drop-down options and harness the power of MARS!
Save and re-use Query

Regardless of the method used to create the query, you can save the query for re-use later, even by other users.

c. Write your queries directly using MARS Query Language...
Advanced Analytics

MARS retains minute-level metrics without having to roll up. This allows the user to specify any date range for which data is collected ("go back in time") for analysis.
Add new reports

- User can select any pre-defined reports that are available for ready use.
- The selected report(s) will be displayed on Dashboard and scheduled.

1. Select “Add Report”
2. Select from one of many pre-defined reports
3. Configure report parameters
4. Schedule report
Add new reports (continued)

Newly added report is displayed on the Dashboard and is automatically scheduled to be generated at
Alerting

Users can set up Alerts to get notified when certain threshold violation happens. MARS can correlate multiple metrics across time periods to create meaningful Alerts.

1. Set criteria for Alert
2. Specify who should receive the alert
3. Send Alerts by e-mail
Hitachi Data Center Analytics provides a new efficient approach to analyzing historical performance trends across your storage infrastructure.

- Build a scalable and flexible data repository for your storage analytic needs
- Collect highly granular performance statistics
- Report on storage performance trends spanning multiple years
Questions and Discussion

Hitachi Data Systems
Upcoming WebTechs

- **WebTechs**, 9 a.m. PT, 12 p.m. ET
  - From Monolithic design to Software Defined building block: The Evolution of Converged Infrastructure, June 10
  - Hybrid Enterprise to Cloud Services, June 24

- **Check** [www.hds.com/webtech](http://www.hds.com/webtech) for
  - Links to the recording, the presentation, and Q&A (available next week)
  - Schedule and registration for upcoming WebTech sessions

Questions will be posted in the HDS Community: [http://community.hds.com/groups/webtech](http://community.hds.com/groups/webtech)
Thank You