The IBM Common Data Provider for z Systems

“Transforming Operational Data in Insights”

Clyde E. Richardson
Executive IT Specialist
North America Technical Sales
richarcl@us.ibm.com
Disclaimer Statement

IBM’s statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM’s sole discretion.

Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision.

The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract.

The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user’s job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.
IBM z Operational Analytics strategy and vision*

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represents goals and objectives only.
Quick review of portfolio and capabilities
IT Operations Analytics Solutions on z Systems

“Insight to action in half the time”

Optimization
• IBM Capacity Mgmt Analytics.
• NEW - IBM z Operational Insights (SaaS)
• Capacity Management
• Software cost analysis
• Enterprise resource and workload optimization
• SME insights in cloud (SaaS)

Investigate and Automate
• NEW - IBM Operations Analytics for z Systems
• NEW - IBM Common Data Provider for z/OS
• Log Analytics
• Domain insights and expert advice
• Alert notification and automation
• Unified data collection for logs and SMF

Predict
• NEW - IBM zAware (Anomaly Detection)
• IBM Operations Analytics – Predictive Insights
• Pro-Active Outage Avoidance
• Predict Problems before occurrence
• Log anomaly detection

On Premise Hybrid Cloud
IBM Operations Analytics for z Systems v3.1

Reduce Operational Cost, Bring Analytics to your Data, Smarter IT Service Management with Rapid Time to Value

- New Problem Insights View
  - Consolidated view across the system for root cause analysis
- z/OS Security Insight Pack
  - See patterns of security incidents by user or resource
- Analyze critical operations data
- Includes zAware Software Appliance
  - Advanced machine learning to detect abnormal system behavior

Insurance industry client example

- Experienced an application outage that resulted in the team working **around the clock for 29 hours**
- After the issue was resolved, the logs were captured and sent to IBM lab for analysis using IBM Operations Analytics for z Systems
- **Within minutes**, the IBM team was able to focus in on the root cause of problem and find the relevant PTF to resolve the issue

[ibm.biz/ioazlivedemo](ibm.biz/ioazlivedemo)
IBM zAware v3.1

Cognitive infrastructure to detect anomalous systems behavior in near real-time

- New Software Appliance Delivery
  - Deploy and launch within minutes
- Enhanced Proactive Anomaly Detection
  - Act before there is a service outage via email notification
- New Historical View
  - See the history of an anomalous message in order to accelerate problem resolution
- Integrated with IBM Operations Analytics
  - Launch directly to logs in the context of an anomaly

ibm.biz/ioazlivedemo
IBM Common Data Provider for z Systems V1.1: Real Time Access to Analytics

A single source for all operational data streamed to the analytics platform of choice

- Simple to install, configure and use
  - Multiple data sources
  - Flexible output options
  - Write to any destination
  - Streaming SMF and log data
  - Built in filtering to control data volumes

- CDP provides consumable, near real time operational data
- Built to improve the ability to manage the growing complexity of data requests
- Tivoli Decision Support for z/OS customers can write their SMF data directly to IDAA

Reduce Risk to you Business:
Detect threats with your Security products using live streaming data

Optimize Costs and Efficiencies:
Feed all IT Operations data to analytical engines from a single product

Prevent Impact to Your Operation:
Proactive Analysis of data in near Real Time as an early warning

© 2015 IBM Corporation
Breaking new ground in the Cloud with IBM z Operational Insights (SaaS)

Expert cost & performance insights for z Systems in minutes. Just add operational data.

- IT Service Management software, delivered as SaaS, addresses key pain points:
  - z Systems performance & running costs
  - Pressure on staff availability & SME skills
  - Time-to-value of traditional on-premise tooling
- Analytics on z operational data, with embedded IBM expertise
- ‘CICS Essentials’, with intent to expend across subsystems
- Try IBM zOI for yourself at ibm.biz/try-zoi

- Potential benefits quantified upfront
- Tailored recommended actions
- First-in-kind comparisons to other users
- Built on expertise from z performance SMEs
- Clean, modern web browser-based interface
- Easy value assessment – free SaaS trial
Product Capabilities
IBM Operations Analytics for z Systems
IBM Operations Analytics for z Systems

Accelerate problem isolation and identification … Reduce mean time to repair

- **Analyze** various types of data (logs, metrics, events, trouble tickets) from multiple sources (mainframe and distributed).
- **Identify** unusual messages and message patterns in log data through integration with the zAware product.
- **Surface common issues** hidden in the data via the Problem Insights tab and take steps to problems resolution using the provided suggested actions.
- **Locate problems** from system, configuration, software logs and performance metrics using **rapid index search** and **pattern analysis**.
- **Isolate issues** across various domains including OS, Middleware, applications, etc..
- **Leverage Expert Advice** via links to support documentation and operations notes to resolve problems quickly.
- **Visualize** search results with analytic tools to **rapidly determine root cause**.
- **Out-of-the-box analysis and insights** for z/OS, WebSphere, DB2, CICS, IMS, MQ, Network, Security as well as distributed systems.
- **Enable early error detection** and broaden scope of automation with **event notifications**.
- **Fully customizable** to meet your needs.

Enhanced security insights from SMF 80
- Root cause insights through new Problem Insights views
- Integration of IBM zAware anomaly insights
- Common Data Provider for z Systems
- Support for industry-standard log management tools (Logstash, Apache Kafka)
- VSAM ESDS support
- Expanded WAS insights from SMF 120
- Client-side Expert Advice

in 2016
Skills & Agility: Integration with IT Service Management solutions

Experience superior, smarter problem diagnostics through integration with OMEGAMON, IBM’s premier performance management solution for z Systems, and other IT Service Management solutions

IBM Operations Analytics for z Systems
Transforms massive log messages into actionable insights

1. Reduce operational cost: Experience significant time reduction in problem identification, isolation and resolution

2. Bring analytics to your data: End-to-end z Systems solution available

3. Smarter IT service management: Integrate with your proven ITSM solutions

4. Rapid Time to Value: Deploy in a matter of a few hours; gain real-time insights instantaneously

Insurance industry client example
- Experienced an application outage that resulted in the team working around the clock for 29 hours
- After the issue was resolved, the logs were captured and sent to IBM lab for analysis using IBM Operations Analytics for z Systems
- Within minutes, the IBM team was able to focus in on the root cause of problem and find the relevant PTF to resolve the issue
Components included in the offering

5698-ABH V3 IBM Operations Analytics for z Systems (S&S 5698-AAQ)

- **Log Analytics server**
  - IBM Operations Analytics – Log Analysis (IOA-LA), included in the offering, provides the analytics engine as well as the web user interface for the solution.
    - At GA, IOAz v3.1 will include IOA-LA v1.3.3

- **IBM System z Advanced Workload Analysis Reporter (aka IBM zAware)**
  - Provides an integrated, self-learning analytics solution that helps identify unusual behaviors of workloads running on z/OS logical partitions (LPARs) as well as Linux on z Systems.

- **IBM Common Data Provider for z Systems**
  - Enables collection of IT operational data through a single interface, providing structured & unstructured data in near real time (live streaming) to a variety of analytics solutions; ops data can be provided both on & off platform in a consistent, consumable format.

- **z/OS Insight Packs and Extensions**
  - Designed as plug-ins for the Log Analytics server, these Insight Packs provide the intelligence to parse and annotate z/OS logs, as well as the out-of-the-box insights and application views included in the offering.

- **Logstash and Logstash custom plugin**
  - Logstash is an open-source data pipeline. It receives z/OS operational data from Common Data Provider for z Systems and makes them available for consumption by analytics platforms. An IBM-provided custom plugin manages forwarding of the operational data to the Log Analytics server.

- **zAware data gatherer**
  - Collects anomaly data from IBM zAware and forwards them to the Log Analytics server, enabling inclusion of this data in Problem Insights analysis.
The Common Data Provider (CDP) was driven by customer requests to address the growing Operational Analytics requirement.

The CDP provides:

- A single source for z/OS Operational Data in a flexible, consumable format both on and off platform
- Near Real time data feed of SMF data and log data
- Single interface that is easy to configure and use
- Read once – write many
- Multiple destinations in different formats for different consumers
- Batch data collection also available for deep dive analysis or to control CPU consumption
- Documented protocols and formats for sending and consuming data are provided, enabling data ingestion to widely used Industry Analytics Platforms or Enterprise-specific solutions for access and analysis
- One time charge based license means no limits on data volumes.

Vision and Purpose

An interactive framework for combining multiple views of the same data to provide a deeper understanding of the Enterprise
Logstash

- Logstash is an open source data collection engine
- CDPz cannot forward data directly to the Log Analysis server
  - Data gathered by CDPz must be forwarded to a Logstash instance
  - The Logstash instance then forwards the data to the Log Analysis server

- Once the data is in Logstash, you can also:
  - forward data from Logstash to other destinations
  - route data through a message broker such as Apache Kafka

- IBM® Operations Analytics for z Systems provides:
  - a copy of Logstash 2.3.4 that is executable on Linux for System z
    • Also executable on Linux for System x and Linux for System p
  - an IOAz Logstash output plugin that:
    • forwards data to the Log Analysis server for processing by the z/OS Insight Packs
    • defines required data sources on the Log Analysis server automatically
z/OS data ingestion pipeline in IOAz V3.1.0

- System Data Engine collects SMF data and forwards them to the CDPz Data Streamer.
- z/OS Log Forwarder collects log data and forwards them to the CDPz Data Streamer.
- CDPz Data Streamer transforms log and SMF data to UTF-8 and forwards them to Logstash using the TCP protocol.
- Logstash receives data using the tcp input plugin and forwards to the Log Analysis server using the IOAz output plugin.
IBM zAware V3.1

- IBM z Advanced Workload Analysis Reporter (IBM zAware)
- Delivers cutting edge pattern recognition analytics applied to system log messages with minimal impact to workloads
  - Helps diagnose major problems while they are occurring in near real time
  - Heightens awareness of small problems before they become big problems
  - Reduces mean time to recovery
  - Reduces the time and skill required to diagnose a problem
- A browser based view, which can show the entire zOS and Linux footprint in one window
- Version 1 (firmware) debuted in 2012 (zEC12) supporting zOS Operlog
- Version 2 (firmware) debuted in 2015 (z13) adding support for Linux syslog
- Version 3.1 (software) available 9/16/16
IBM Operations Analytics Architecture and Flows

Mainframe

- z/OS
  - WAS SYSPRINT
  - WAS SYSOUT
  - z/OS Syslog
  - CICS MSGUSR
  - CICS EYULOG
  - USS Log Files
  - Job logs
  - VSAM ESDS
  - NetView Netlog
  - Other Logs
  - SMF Data

Common Data Provider for z Systems

Linux on z Systems

- WAS SYSPRINT
- WAS SYSOUT
- DB2
- DB2 App
- Syslog
- Web Access Log
- Other Logs

Log collection

Secure Service Container

ITM Log File Agent

Operations Analytics server

- Logstash
- Generic Receiver
- z/OS Insight Packs
- Distrib. Insight Packs
- Custom Insight Packs
- zAware Data Gatherer
- Alerts
- Current Tier
- Archive Tier
- Hadoop Tier

Problem Insights
- Dashboards
- Search

Linux on x, p, z

- WAS SYSOUT
- z/OS Syslog
- CICS MSGUSR
- CICS EYULOG
- USS Log Files
- Job logs
- VSAM ESDS
- NetView Netlog
- Other Logs

- Other Logs

- USS Log Files
- Job logs
- VSAM ESDS
- NetView Netlog
- Other Logs

- Other Logs

- USS Log Files
- Job logs
- VSAM ESDS
- NetView Netlog
- Other Logs

- USS Log Files
- Job logs
- VSAM ESDS
- NetView Netlog
- Other Logs

Log collection*

* Possible log collection agents: syslogd, Elastic Filebeat, Logstash, etc.
Simple search interface EASY to customize

Log data is analysed and insights are surfaced as you search
Find problems you didn’t know existed
Simple search query interface EASY to customize

Search capabilities

• Based on the Apache Solr Search Engine.
• Query Syntax based on Indexing Engine query syntax –
  • https://wiki.apache.org/solr/SolrQuerySyntax
• The Log Analysis Server uses the Lucene query parser plug-in.
  • http://lucene.apache.org/core/5_1_0/queryparser/org/apache/lucene/queryparser/classic/package-summary.html
• Standard Search keywords and operators:
  • Or – Default operator. Either term or expression must be matched in the results
  • + - To exploit AND like functions
  • AND – Alternative to the + operator
  • “” – Group individual terms into phrases
  • () – Group expressions into phrases
  • * - Wildcard operator
  • ? – Wildcard operator
• There are other operators – see the IBM Operations Analytics – Log Analysis User’s Guide for more information…..
Dashboards, Information Links and Expert Advice

Visualize the data with Dashboards
Quick links to additional information and support documents.

- Provided with every Insight Pack
- Expert Advice to access white papers, tech notes, APARs, etc. for faster problem resolution
- Dashboard views created by subject matter experts, support teams and customers
- Immediate value out of the box
- Easy to modify or create and save your own
Analyze logs as you Search

Insights are surfaced automatically as you search. Patterns are surfaced based on the log type.

- Provided with every Insight Pack
- Logs are analysed automatically
- Log data is categorized by hostname, data source, message type, message source, etc.
- Patterns/Insights are surfaced to help you focus on the source of the problem.
  For example, log analysis automatically surfaces java exceptions in application logs.
- Perform searches and analyse multiple logs, organized per the needs of your enterprise.
- Create your own Insight Pack for any text logs with time stamps
IBM Operations Analytics for z Systems V3.1

Problem Insights and Anomaly Identification … Understanding the unknown

NEW in V3.1:
- Inclusion of IBM zAware v3.1 anomaly detection with Problem Insights
- Rapid analysis of vast amounts of data, and even more data types
- New z/OS Security Audit Events Insight Pack
- Use in conjunction with the NEW Common Data Provider for z/OS V1.1

Key Values:
- Built-in IBM expertise to predict issues instead of waiting for failure, for fewer outages
- Analyze & intelligently search ops data with anomaly detection, for faster root cause analysis and Mean Time to Recovery

http://ibm.biz/ioazlivedemo
Problem Insights

- Automatically surfaces important messages found in the log data.
- Provides easy to read problem summary and suggested actions for problem resolution.
- Displays Anomaly Interval scores from IBM zAware.
Suggested Actions

- The Suggested Actions are presented as a pop-up window.

Suggested Actions to investigate/resolve the issue

Link to the Knowledge Center for this message ID
IBM Support Portal based Expert Advice

Search for expert advice with the click of a button

Launch from client or server side

Launch to Tech Note

All IBM support site documents that reference messages from search results
Product Capabilities
IBM Common Data Provider for z Systems
IBM z Operational Analytics strategy and vision*

Data

- Common Data Provider
- Logs
- Metrics
- z/OS

Platform

- Platforms
  - IBM Operational Analytics - On-Prem
  - IBM Operational Analytics - SaaS
  - Non-IBM platforms
  - Other IBM consumers (IDAA, etc)

Insights

- z Insights – multi platform and Actionable Eco system enablement w/Partners
- Systems and Domain focus
- Services framework
  - Watson Analytics
  - Cognitive
  - Predictive
  - Anomaly Detection

* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represents goals and objectives only.
The Common Data Provider (CDP) was driven by customer requests to address the growing Operational Analytics requirement.

The CDP provides:

- A single source for z/OS Operational Data in a flexible, consumable format both on and off platform
- Near Real time data feed of SMF data and log data
- Single interface that is easy to configure and use
- Read once – write many
- Multiple destinations in different formats for different consumers
- Batch data collection also available for deep dive analysis or to control CPU consumption
- Documented protocols and formats for sending and consuming data are provided, enabling data ingestion to widely used Industry Analytics Platforms or Enterprise-specific solutions for access and analysis
- One time charge based license means no limits on data volumes.

Vision and Purpose

An interactive framework for combining multiple views of the same data to provide a deeper understanding of the Enterprise
Why the Common Data Provider for z Systems?

- **CDPz - OTC product** - Huge Benefit saves TCO as competitors based off amount of data CDPz allows unlimited data.

- **Supports over 400 SMF record types and can stream in Real-Time (near real-time)** which can save CPU cycles by lowering CPU lower overall costs.

- **Collects once and can write to many consumers** - Off platform including Splunk 15% of process is for ETL data processing, execute once and writing many saves added repeat ETL processing.

- **Supports sending z data to Splunk** and other third-party consumers (i.e. ELK) - I'm not sure how many we cover - covering more than one target provides more flexibility than some competitor products.

- **Solid research & development history has gone into development of this solution years of IBM z expertise & knowledge.**

- **Part of the solution runs in it's own JVM (Data Streamer),** and as such, can run on zIIP processors, saving CPU cycles saving CPU saves overall costs and provides better TCO.

- **Can feed TDS z SMF data directly into IDAA.** Huge benefit and CPU savings if client as already has an investment in Tivoli Decision Support on z/OS when data is in IDAA, to save both on CPU process to access the data and storage size is also smaller footprint and you can store more detail data for better analysis of metrics.

- Can feed SMF data directly into IDAA even if client doesn't own TDSz.

- **Easy to configure via z/OSMF Web UI** - quick and easy configurations allows the admin to focus on more important tasks.
Why the Common Data Provider for z Systems?

- **Policy driven streaming of z Systems Operational Data**
  - SYSLOG
  - Supports over 400 SMF record types.
  - Send data from ANY Job DD
  - Send data from ANY zFS
  - CICS MSGUSR
  - WebSphere Application Server for z/OS Logs
  - WebSphere Liberty Profile Logs
  - DB2 on z/OS
  - MQ on z/OS
  - IMS on z/OS
  - Netview Log
Near Real Time SMF

OA49263: New function to enable SMF in-memory resources and the SMF real-time interface

- Robust, reliable access to data before it is written to the logstream.
  Note that logstream recording mode is a pre-requisite to be able to use the API
- Available for z/OS 2.1 and z/OS 2.2
- The CDP is one of the first exploiters of this new technology

In memory performance gains:

<table>
<thead>
<tr>
<th>Workload</th>
<th>Data Rate</th>
<th>Logstream CPU Utilization</th>
<th>SMF API CPU Utilization</th>
<th>Reduction in CPU Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMF30</td>
<td>86 MB per hour (2 GB per day)</td>
<td>0.19%</td>
<td>0.07%</td>
<td>63.2% reduction</td>
</tr>
<tr>
<td>SMF80</td>
<td>86 MB per hour (2 GB per day)</td>
<td>1.52%</td>
<td>0.686%</td>
<td>54.9% reduction</td>
</tr>
</tbody>
</table>

* Under test lab conditions
Data Gatherers

System Data Engine
- Based on 30+ years of engineering
- Designed to collect and process SMF data
- No DB2 prereq – installed and usable within hours
- Data can remain unprocessed or unpacked into a readable, consumable format
- Formats the data into multiple formats for ease of ingestion (e.g. CSV or DB2 LOAD format)
- Supports all standard IBM SMF types
- Supports multiple sources of SMF data – archive, logstream or direct from the new SMF Buffer api in near real time
- Has built-in filtering to control data types and volumes

Log Forwarder
- Gathers a variety of log data and some VSAM file formats for Analytics Engines
- Additional log support planned through ongoing continuous delivery
- Custom log types can be added and any dataset can be streamed giving great flexibility
The Data Streamer controls the destination and format of the Operations Data

- Receives data from the gatherers
- Splits and Annotates the log data into individual messages for ingestion into analytic engines
- Transforms data messages into the right format for the destination platform (e.g. UTF-8 and other code pages)
- Transport mechanism is TCP/IP – available as SSL for additional security
- Data sent in json wrapper for ingestion by Logstash for storage and analysis
- Extendable to other platforms like ELK and SPLUNK
- Streams data both on and off platform
- zIIP enabled for cost savings (pure Java)
IBM Operations Analytics Architecture and Flows

**Mainframe**
- **z/OS**
  - WAS SYSPRINT
  - WAS SYSOUT
  - z/OS Syslog
  - CICS MSGUSR
  - CICS EYULOG
  - USS Log Files
  - Job logs
  - VSAM ESDS
  - NetView Netlog
  - Other Logs
  - SMF Data

**Common Data Provider for z Systems**

**Linux on z Systems**
- WAS SYSPRINT
- WAS SYSOUT
- DB2
- DB2 App
- Syslog
- Web Access Log
- Other Logs

**Log collection**

**Logstash**

**Operations Analytics server**
- z/OS Insight Packs
- Distrib. Insight Packs
- Custom Insight Packs
- zAware Data Gatherer

**IBM zAware**

**Secure Service Container**

**Problem Insights**
- Dashboards
- Search

**Linux on x, p, z**

**Possible log collection agents:** syslogd, Elastic Filebeat, Logstash, etc.
Three main component types
Data Gatherers – flexible, customizable, efficient
Data Streamer – controls data formats and destinations
User Interface - simple intuitive configuration
Web Configuration Tool

The data sources, transformation and destinations are managed and controlled through a simple user interface

Plug-in for z/OSMF

Menu driven to configure:
- Data streams and their sources
- Any transformation requirements
- Output format
- Destination (on or off platform)

Security

Push Based model
- Host-based policy controls subscribers and data sources
- Policy can be secured by RACF for total control of data and subscribers
SMF Data Streaming Flow

1. Real time SMF data extracted over in memory API.

2. SMF data parsed, aggregated and formatted as CSV or JSON.

3. Data sent over TCPIP socket as tagged data packet.

4. Routes the data packet to all configured consumers.

5. Persistent TCPIP socket connection to consumer used to send traffic.

SMF Data Streaming

Data Sources

Data Gatherers

Web Configuration

Consumers
Streaming SMF Data Lifecycle

1) Application cuts SMF record
2) SMF accepts it and write to memory buffer
3) SDE reads it from the memory buffer, parsers and process it
4) SDE output in a CSV file that’s sent to the Data Streamer with a tag
5) Data Streamer looks up the subscriptions for the tag
6) Sends it to each consumer in the subscription
7) Consumer receives the data and processes it
Data Streamer capabilities

- **Responsible for the following:**
  - Receiving the data from the gatherers
  - Transforming the data to meet the subscribers needs
  - Pushing the data out to the subscribers

- **Data Transformations:**
  - Transcoding – usually to UTF-8 (but you can choose other code pages)
  - Splitting lumps of data into arrays of individual records
  - The input data is tagged with contents and format
    - **SMF30-CSV-EBCDIC**
    - Clients subscribe to the format they want
      - **SMF30-CSV-UTF8**
      - As the required transform is EBCDIC to UTF8 translation, the Data Streamer can perform it automatically

- **Pushing data:**
  - Feed to Logstash over TCPIP – SSL variant available
  - Data sent as CSV (or just raw data) wrapped in JSON
  - Split data sent as an array of CSVs wrapped in JSON

- Can leverage zIIP processors to offload its workload
**z/OS SYSLOG Data Streaming Flow**

1. **z/Log Forwarder** regularly scans the files and finds updates.

2. Data Sent to **Data Streamer** in tagged data packets.

3. **Data Streamer** transforms and sends to configured consumers.

**Streaming of SYSLOG and File Data**

- **Data Sources**
- **Data Gatherers**
- **Data Consumers**
- **Data Streamer**
- **Web Configuration**
Streaming Syslog file Lifecycle

1) Application issues a WTO
2) z/OS records WTO to Syslog
3) z/Log Forwarder exits catches the message
4) z/Log Forwarder collects all new messages, tags them and sends them to the Data Streamer as a single data packet
5) Data Streamer formats the data packet into an array of messages
6) Data Streamer sends formatted data to each subscribed consumer
Combined SMF & Syslog streaming flow

Combined Flows

Data Sources
- SMF
- Logs
- Files

Data Gatherers
- SDE
- z/Log Forwarder

Data Streamer
- Data Streamer

Data Consumers
- FTP Alternative
- Web GUI
- DB2 Analytics Accelerator

z/OSMF Consumers
- Web GUI

Config
- Config

FTP Alternative
- FTP Alternative

Web Configuration
Product Capabilities

Feed z Systems data to Splunk!
Use Case:
Better and more timely problem determination and resolution with real-time views of mainframe SMF data to identify real or potential problems earlier and to support triage repair, correction, or prevention more quickly.

Function: Streaming SMF data and log data to SPLUNK in near real time

- Customers get all z Systems type data
  - Customer can combine with distributed data for Enterprise view
  - One pane of glass for all operations data
  - Better visibility into the Mainframe operations

- Live real-time streaming of SMF data
  - Provides diagnostic & resource usage information saving time and money

- OTC pricing
  - Lower cost of product usage without paying for amount of data consumed

- Advanced filtering capability on SMF and logs
  - Reduced ingestion costs on SPLUNK
  - Faster processing of data
  - Customer can control the data
The DataReceiver sorts the streams as it writes the data to disk
- One set of files for each source type

The ingestion files have a pair of rules (input and props) for each source type
- Inputs.conf stanza reads it from disk and sets its source type
- Props.conf extracts timestamp and sets the field extraction rules

The DataReceiver rotates through a set of working files for each source type
- Manages its own files and disk space
CDP to SPLUNK – Setup and Configuration

Setup

Pick the machine to use as your Splunk> Forwarder
Install and set up Splunk>
Download the CDPDataReceiver and CDPSplunkBuffer application from z/OS
Install and set up the CDPDataReceiver
Install the CDPSplunkBuffer application into Splunk>
Set up additional rules for the forwarder to send the data to the correct index servers (suggested routing by Sysplex)

Configuration

The supplied Splunk> Apps can ingest any standard CDPz stream
Simply link the stream to the Splunk> Receiver in the CDPz dialogs - the data will turn up in Splunk>

Custom data requires addition configuration:
  – Define and link it in the CDPz Dialogs
  – On the Splunk> server you need to create your own ingestion rules
    • Use the stanzas in inputs.conf and props.conf from the IBM Apps as samples
CDP to SPLUNK – data and reporting

CDP can supply log data (e.g. Syslog) or SMF data

Can use native Splunk> tools for reporting
No local buffering on Splunk> Forwarder end
- If the forwarder is down the data is buffered on the z/OS systems (maximum 2GB)
- Data will be backed up on z/OS if it arrives faster than Splunk> can ingest it

Simpler to install
- Download and install the CDPSplunkDirect app

Under development
- Needs a new protocol to send the data in a flat CSV or JSON
Product Capabilities

*Feed* SMF data to **IDAA!**
Use Case:
More timely and cost effective way for loading and storing data used for analysis, problem determination and view of resource usage.

Function: Streaming CDPz Enhanced SMF Data direct to IDAA

- Reduced DB2 footprint on z/OS ($10 per GB per week)*
  - Reduced data storage requirements provides reduced cost

- Streamline loading process without DB2 updates or aggregation during collect
  - Over 30% improvement on CPU for collecting and loading data

- Provides live access to data compared to batch static load from TDSz
  - Access data sooner

- Similar database scheme to TDSz
  - No need to change existing reporting

- Cheaper storage on IDAA (90c per GB per week)*
  - Months of data stored instead of weeks provides more data for analysis
  - Able to store 40x times the amount of data compared to TDSz DB2 storage
Tivoli Decision Support for z/OS

Tivoli Decision Support for z/OS (TDSz) is a Performance Reporting and Capacity Management tool that collects SMF data for long terms trending and analysis. Data is collected, aggregated and stored in a DB2 database on z/OS.

The data can also be used for chargeback purposes, SLA validation and deep dive analysis.

With the growing complexity and volume of SMF data there are some opportunities to reduce the CPU consumption and storage requirements.
Leveraging the IDAA with TDSz

TDSz customers are looking for ways to exploit the IBM DB2 Analytics Accelerator (IDAA) to take advantage of the speed and analytical capability.

The current option is to load into DB2 and copy the data into the IDAA

Pros
✔ Longer storage time for SMF data (months instead of week)
✔ Access to the lightening fast IDAA queries
✔ No need to change reporting system

Cons
✘ DB2 footprint remains the same
✘ Addition CPU cycle on the copy step
✘ No detail for most of the transactional records (DB2, CICS, IMS)
Leveraging the IDAA with CDPz

The CDP has the ability to LOAD the TDSz data directly into IDAA_ONLY tables, bypassing the DB2 tables. Without the storage restrictions, timestamp records can be collected and loaded.

Pros

✔ Longer storage time for SMF data (months instead of weeks)
✔ Access to the lightening fast IDAA queries
✔ No need to change reporting system

✔ DB2 footprint dramatically reduced
✔ 100% reduction on copy step
✔ Timestamp detail for deep dive analysis
✔ No need for aggregation – additional CPU savings
Summary
Common Data Provider for z Systems v1.1

The Common Data Provider is a combination of SMF data provider and log provider and is the data transport layer for Analytics

The open and flexible design makes it useful across the spectrum of z Systems software

Function

• A single source for all Operation Data for use in Analytics, Security, Monitoring, DevOps, Automation, Performance and Capacity Planning
• Simple to configure and use
• An interactive framework for combining multiple views of the same data to provide a deeper understanding of the Enterprise

GA: 16th September 2016
Ongoing Continuous Delivery
IBM team is read to help!

- Request a Demo
- Join the Beta Program
- Contact Development

Alan Place, IBM
IBM Senior Development Manager
zSystems zOperational Analytics
alandp@au1.ibm.com

- Contact Offering Management

Barry T Klutz
IBM Senior Offering Manager
zSystems - ITOAz
bklutz@us.ibm.com
Thank You