Welcome to the Virtual IMS user group newsletter. The Virtual IMS user group at itech-ed.com/virtualims is an independently-operated vendor-neutral site run by and for the IMS user community.
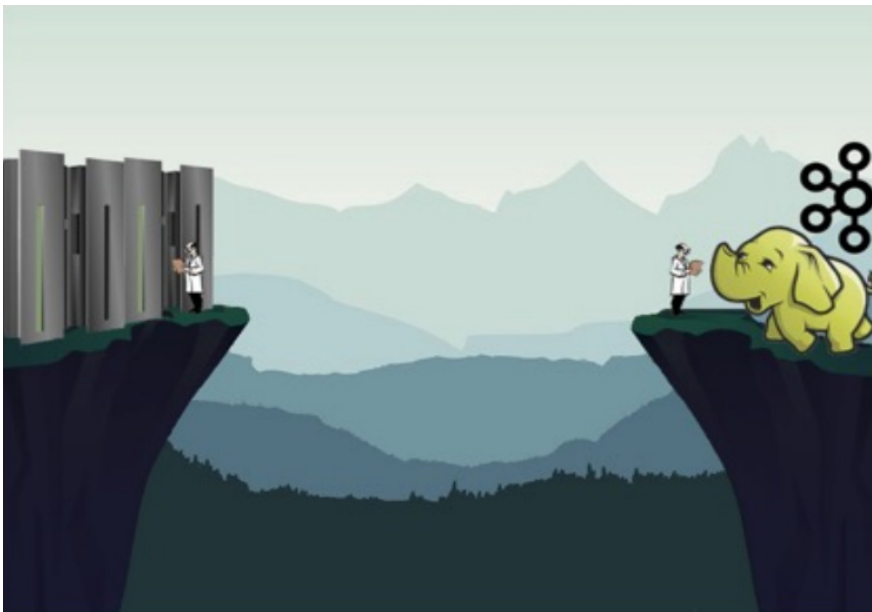

Figure 1: The 'Great Divide' is real

VSAM. Scott is considered a leading authority of database replication strategy and deployment, and stays actively involved with large IMS-related initiatives worldwide.

Scott Quillicy started his presentation by looking at what was meant by mainframe modernization. For some people, it was to migrate the workload from the mainframe to the Cloud. This means major core applications (like IMS, Db2z, and VSAM), and requires

Virtual IMS user group presentation

The latest webinar from the Virtual IMS user group was entitled, "Lessons Learned – IMS Application Modernization to the Cloud". It was presented by Scott Quillicy, Senior Director for Precisely.

Scott is a Senior Director for Precisely (formerly Syncsort) with just over 40 years' experience with IMS and z/OS data integration. Scott was a co-founder of SQData (acquired by Syncsort in 2019), known for its high-performance changed data capture (CDC) solutions for IMS, Db2z, and

Contents:
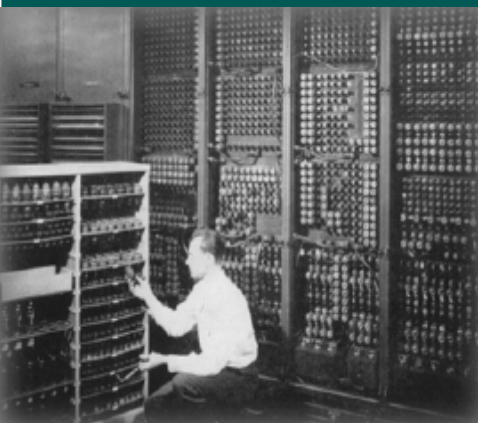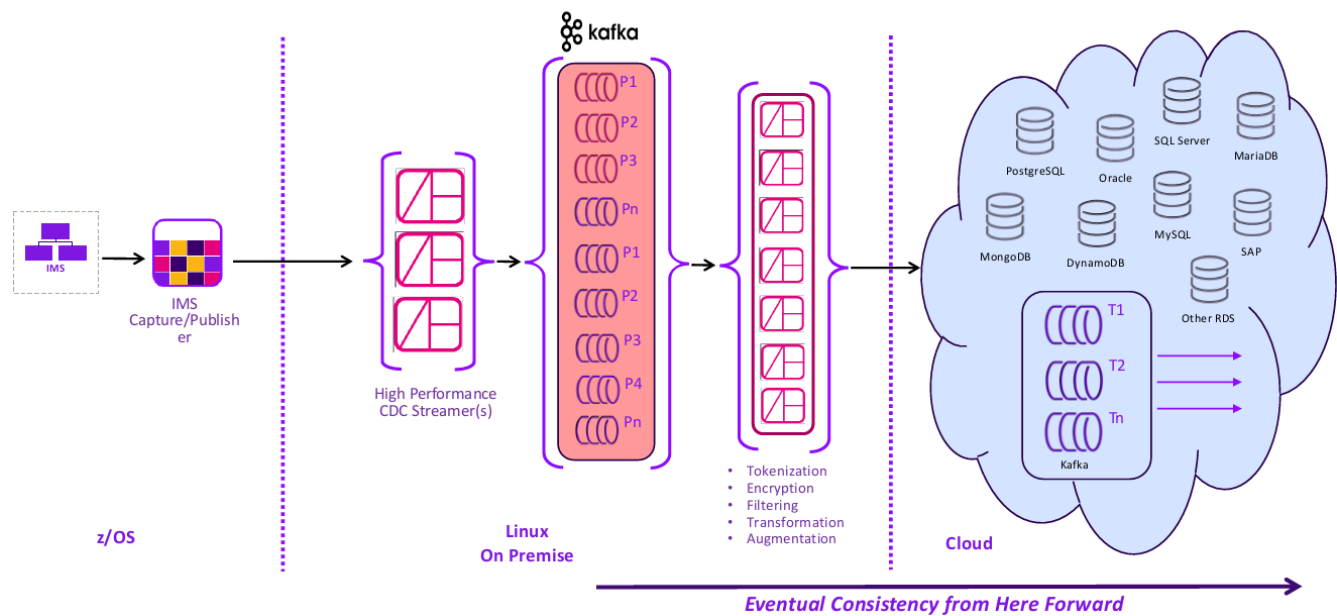
Figure 2: Popular forward synchronization architecture

very high transaction volumes (more than a billion updates/day).

There are different application strategies available. Firstly, companies can replatform (Lift and Shift), where they run IMS applications 'unchanged' on Linux using a framework like Micro Focus for IMS. Secondly, they could refactor, where they convert legacy COBOL/PLI to a Cloud-friendly language such as Java. Lastly, they can redevelop, where they completely rewrite and modernize their applications.

In terms of data, bulk data transfer and data replication are critical components. They have super-low latency requirements. They require two-way synchronization, ie mainframe to cloud and back.

Moving to the Cloud has a number of advantages. There's cost reduction because less physical space is needed, less hardware is required, there's a reduction in costly high-end software, and fewer support personnel are needed. There's a wide variety of access to the latest technology, plus a common framework across application infrastructure. In addition, 'shelfware' can be avoided because you use only what you need. There are advantages in terms of security. For example infrastructure, it must comply with industry standards, and data is less prone to employee theft. In terms of reliability, there's built-

in redundancy, and most providers guarantee 99.99% uptime. Plus, Cloud technical skills are plentiful.

But there are disadvantages. There can be cost creep in terms of compute resources and departmental use/abuse. There will be downtime resulting from communication failure and Internet drops. There are security issues. It does not guard against weak digital security methods, and can be like leaving your laptop open at Starbucks! There are performance issues when it comes to traditional batch workloads and complex transactions. There's also Cloud vendor lock because vendors highly encourage 'their stuff' making it difficult to move off, once you are there.
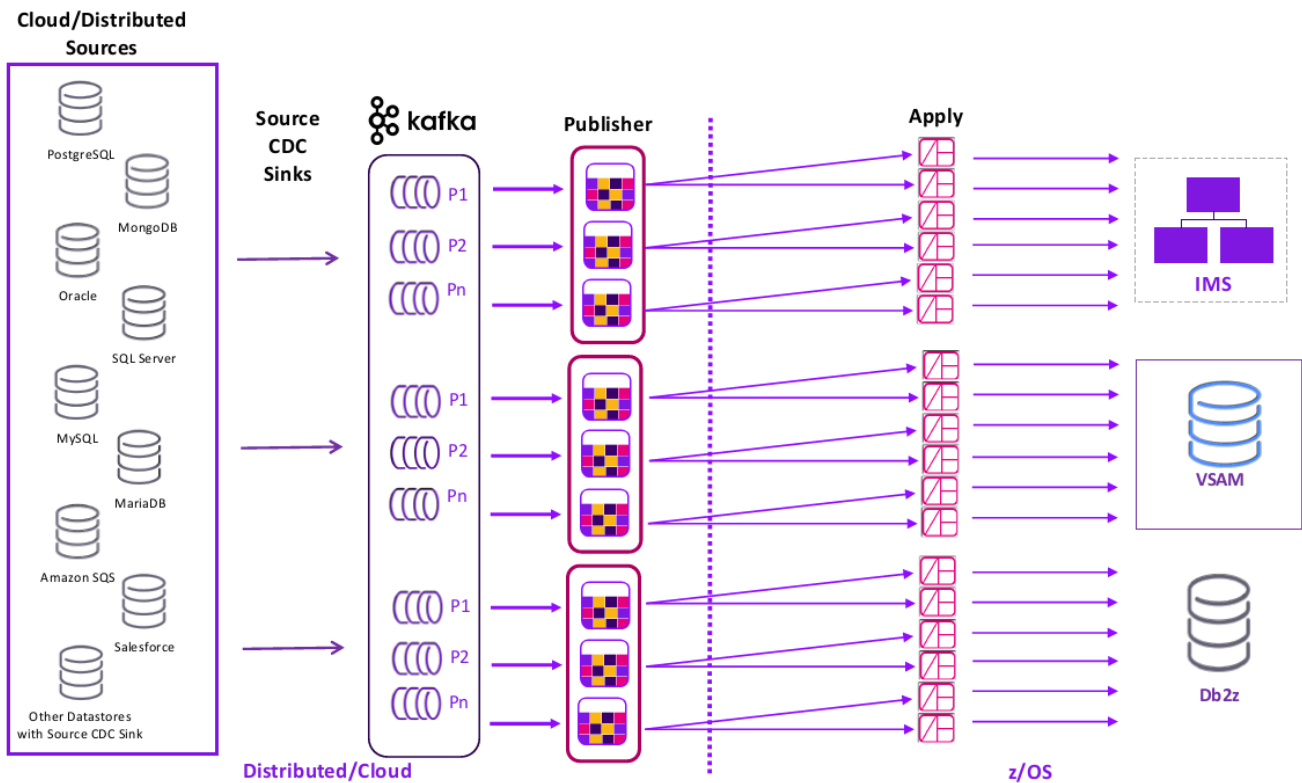
Figure 3: Sample reverse synchronization setup

Scott informed the user group that Amazon (AWS) has the biggest market share, followed by Microsoft and Google.

Scott said that the great divide between the mainframe world and Cloud and distributed is real (see Figure 1). There's a significant language barrier. There are different levels of discipline. And both sides have pride of ownership. There is a shrinking talent pool for IMS and mainframes in general. So, working together requires patience, communication, and collaboration.

When it comes to lessons learned, Scott confirmed that in his experience the 'Great Divide' is a real thing, and that mainframe expertise is in short supply and IMS skills are rapidly becoming scarce. He said that Cloud is king, and on-premise targets represent less than 10% of deployments in 2020-21. Application modernization is super-hot, and moving off the mainframe has gained serious momentum. He did warn that Cloud databases can be painfully slow for batch-type workloads. Scaling is often an afterthought, and bi-directional replication (cloud

to IMS) is becoming a more common use case.

Looking at databases in the Cloud, they are highly scalable for 'normal' transactions, ie quick in / quick out, but they are much slower than traditional on-premise databases. Ping time is the major factor, and can be 10 times slower than on-prem. It's also very important to consider for app modernization that the batch window will increase. Security concerns are much higher because of increased data tokenization. High-volume customers need to include a streaming

component. This allows for a high degree of scaling to relational (and other) targets. However, it must be implemented carefully.

Scott's key considerations were:

- Performance / throughput depends on the speed of the target.

- Application modernization requires lower latencies.

- Higher-volume customers will need to scale

  – Peak IMS transaction arrival rate (CDC records/second) – usually batch

  – Compare against speed of target for a single process/ thread. Relational DB on-prem: 2K rows/ second, whereas relational DB in Cloud: 250 - 500 rows/second, depending on ping time. A streaming platform can have over 100K+ messages/second

  – Goal: target keeps up with source CDC record arrival rate with acceptable latency

- Slow apply rate means: increased mainframe back-pressure; higher CPU usage; and unhappy business users.

Scott had some helpful tips to share. He said to set realistic expectations in terms of latency and throughput. These depend on transaction arrival rate, scaling factor, and infrastructure capacity. It's best to lower expectations (ie sub-second) and avoid unreasonable commitments. Batch transactions/UOWs will not likely be sub-second. Of course, results may vary depending on your environment. The speed of the target generally dictates latency and throughput. Cloud can be significantly slower than on-premise. Network bandwidth becomes a factor. Kafka or other fast streaming components are highly recommended. They allow for target-side scaling and multiple consumer groups for a single source. The final solution will consist of software and APIs from multiple vendors, even competing offerings, and a streaming component (if moderate/high volume with low latency).

Figure 2 shows a popular forward synchronization architecture.

Looking at the idea of eventual consistency, data

will eventually be in synch. Data arrival can (and will) arrive out-of-order across topic partitions, although order is maintained within a single partition. So, be consistent with the partitioning key. Some people try using 'commit events' to consolidate transactions. This complicates things. It doesn't always work if a transaction is spread across multiple partitions because a commit record can arrive before all the transaction data has been written. It also slows the data flow. Best practice is to process data by physical key, making sure a key is used for topic partitioning. Combine transaction data before writing to a streaming target. And carry source update timestamps to the target. This keeps things in order and avoids overlaying newer data with older data.

Reverse synchronization (see Figure 3) is bi-directional, ie Cloud to mainframe. It keep legacy databases in synch with updates from the Cloud. It requires initial planning (an absolute requirement). The rules are simple:

1 There can be one and only one system of record (SOR)

2 SOR cannot be shared across platforms (see #1)

3 Collision detection is required, in case #1 and #2 are violated

4 Data context must be complete because legacy applications may not function without enough data. Ask, "what happens on an insert?"

Partial updates will slow things down by at least 50%.

Scott stressed the need for telemetry and the use of the OpenTelemetry Collector (OTel Collector). This gathers metrics, logs, traces, etc from all critical agents. It provides real-time analysis/debugging, it can explore events not defined in advance, and avoids / minimizes outages. In terms of monitoring, it analyses pre-defined metrics. There's a Prometheus backend (time series database) and Grafana, CloudWatch front ends. These come with cool dashboards, real-time debugging, and common framework.

A copy of Scott Quillicy's presentation is available for download from the Virtual IMS user group website at: https://itech-ed.com/virtualims/presentations/IMSCloudFeb23.pdf

You can see and hear the whole user group meeting at: https://youtu.be/igtM2cVTL7E

Meeting dates

The following meeting dates have been arranged for the Virtual IMS user group:

• On 18 April, Dusty Rivers, Senior Director, US Systems Z at Mainline Information Systems, will be discussing "Mainframe Systems (IMS,CICS etc) As Clients in the New World".

• The following meeting will be on 6 June, when James Martin, Senior zSolutions Advisor, System Z, at Rocket Software will be presenting.

IMS articles and blogs

Curious how IMS 15.3 performs on the new IBM z16 platform? by Amanda Stephens and Justin Soe in the IMS part of the IBM Z and LinuxONE Community (9 February 2023). You can find the article at: https://community.ibm.com/community/user/ibmz-and-linuxone/blogs/amanda-stephens/2023/02/07/curious-how-ims-153-performs-on-the-new-ibm-z16-pl

Sponsorship opportunity

Our user group has been in existence since 2007 and is well-respected among

users of IMS. Currently, co-sponsorship is available for the user group. Please contact me (trevor@itech-ed.com).

Arcati Mainframe Yearbook

The 2023 edition of the Arcati Mainframe Yearbook is now available. The Yearbook stands out as an excellent reference work for all IBM mainframe professionals. It includes a mainframe evolution section, a glossary of terminology, a vendor directory section, a mainframe strategy section, and a mainframe user survey It is completely **FREE**. You can download it from https://itech-ed.com/arcati/

About the Virtual IMS user group

The Virtual IMS user group was established as a way for individuals using IBM's IMS hierarchical database and transaction processing systems to exchange information, learn new techniques, and advance their skills with the product

The Web site at https://itech-ed.com/virtualims provides a central point for coordinating periodic meetings (which contain technically-oriented topics presented in a webinar format), and provides articles, discussions, links, and other resources of interest to IBM IMS practitioners. Anyone with an interest in IMS is welcome to join the Virtual IMS user group and share in the knowledge exchange.

To share ideas, and for further information, contact trevor@itech-ed.com.

The Virtual IMS user group is free to its members.