

Session: 1292

Virtual Db2 Analytics Accelerator Experiences at Citigroup

Bob Perih – Citigroup

THINK 2018

March 21, 2018



# Agenda

---

- Citigroup.
- IBM IDAA on zSeries Introduction Slides.
- Why IDAA on Z?
- Planning for IDAA on Z.
- Installation of IDAA on Z.
- Testing IDAA on Z.
- Summary.




- Citigroup is a global bank that does business in more than 160 countries and jurisdictions.
- More than 200 million customer accounts. We have 219,000 employees world wide.
- The City Bank of New York, the earliest predecessor firm in the current company, was founded in 1812 in NY, NY.
- Citi pioneered the modern ATM in 1977 and was the first to offer 24X7 banking services.
- 2016 revenues of \$69.9 billion, net income of \$15 billion.
- Firm is broken up into to the following major operating units
  - Citi Global Consumer Banking - Retail Banking, Citi Branded Credit Cards, Citi Retail Services, Citi Commercial Bank and Citi Mortgage.
  - Institutional Clients Group – Citi Markets, Corporate and Investment Banking, Citi Private Bank, Citi Treasury and Trade Solutions and Citi Security and Fund Services.

- Citigroup Operations and Technology.
  - The technology needs of Citigroup are supported by the following four businesses:
    - Global Consumer Banking Technology – Application development, delivery, operations and support for all consumer banking and credit cards applications.
    - Institutional Technology - Application development, delivery, operations and support for all applications supporting all of the Citi Markets and Banking and Transaction Services applications.
    - Chief Technology Office.
    - Citigroup Technology Infrastructure.
- Db2 at Citigroup.
  - Db2 for z/OS is one of the core technologies that Citigroup relies on to store business data and process our time critical transactions.
    - Availability, scalability, performance.
    - Application enablement.
    - Db2 Analytics Accelerator.
    - Db2 ESP/Beta programs.
  - Application access includes:
    - CICS, batch COBOL, Websphere on z/OS, application servers running on UNIX or Windows using Type 4 drivers and desktop applications servicing both OLTP and analytics workloads.



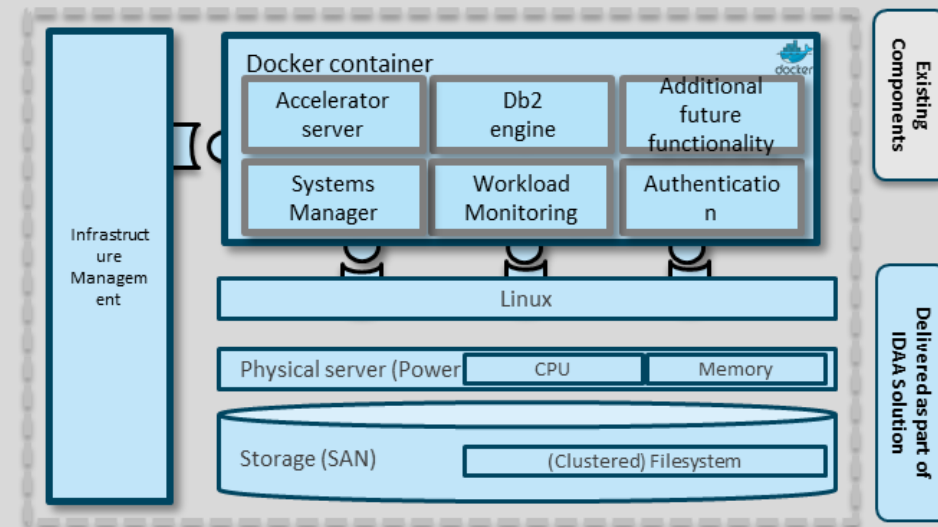
## Db2 Analytics Accelerator – Deployment options and naming

With the availability of Db2 Analytics Accelerator V7 clients may choose from three variants:

	<p>Long name: Db2 Analytics Accelerator for z/OS Version 5.1, deployment on PureData System for Analytics (PDA)</p> <p>Short name: Accelerator on PDA</p>
	<p>Long name: Db2 Analytics Accelerator for z/OS Version 7.1, deployment on IBM Integrated Analytics System (IIAS)</p> <p>Short name: Accelerator on Integrated Analytics System or Accelerator on IIAS</p>
	<p>Long name: <b>Db2 Analytics Accelerator for z/OS Version 7.1, deployment on IBM Z</b></p> <p>Short name: <b>Accelerator on IBM Z</b></p>



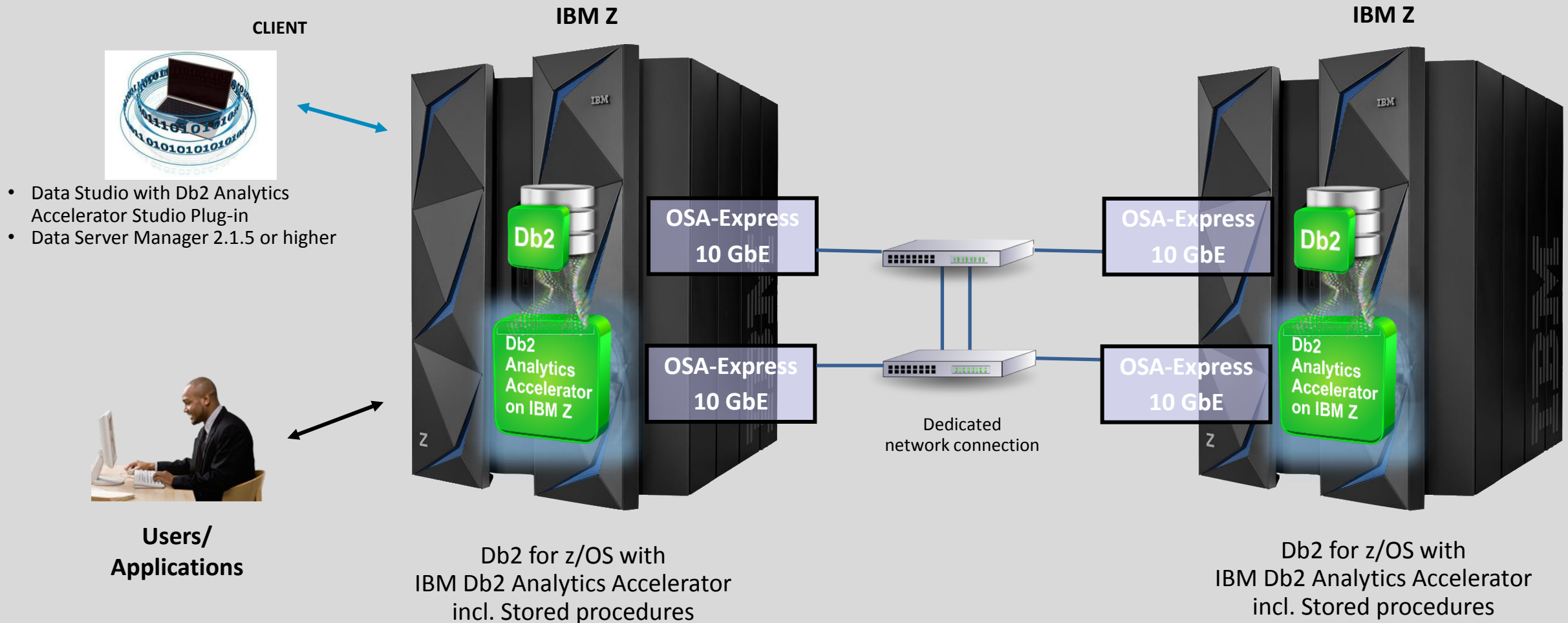
### Db2 Analytics Accelerator V7.1, deployment on IBM Z



- New generation software appliance
- A full solution “on Z”
- All components provided by IBM in a balanced, performance-optimized configuration
  - IBM Z: The world’s premium platform for reliability, stability, security, and more
  - SW stack including the Linux operating system, the docker software as well as the Docker container and the infrastructure management

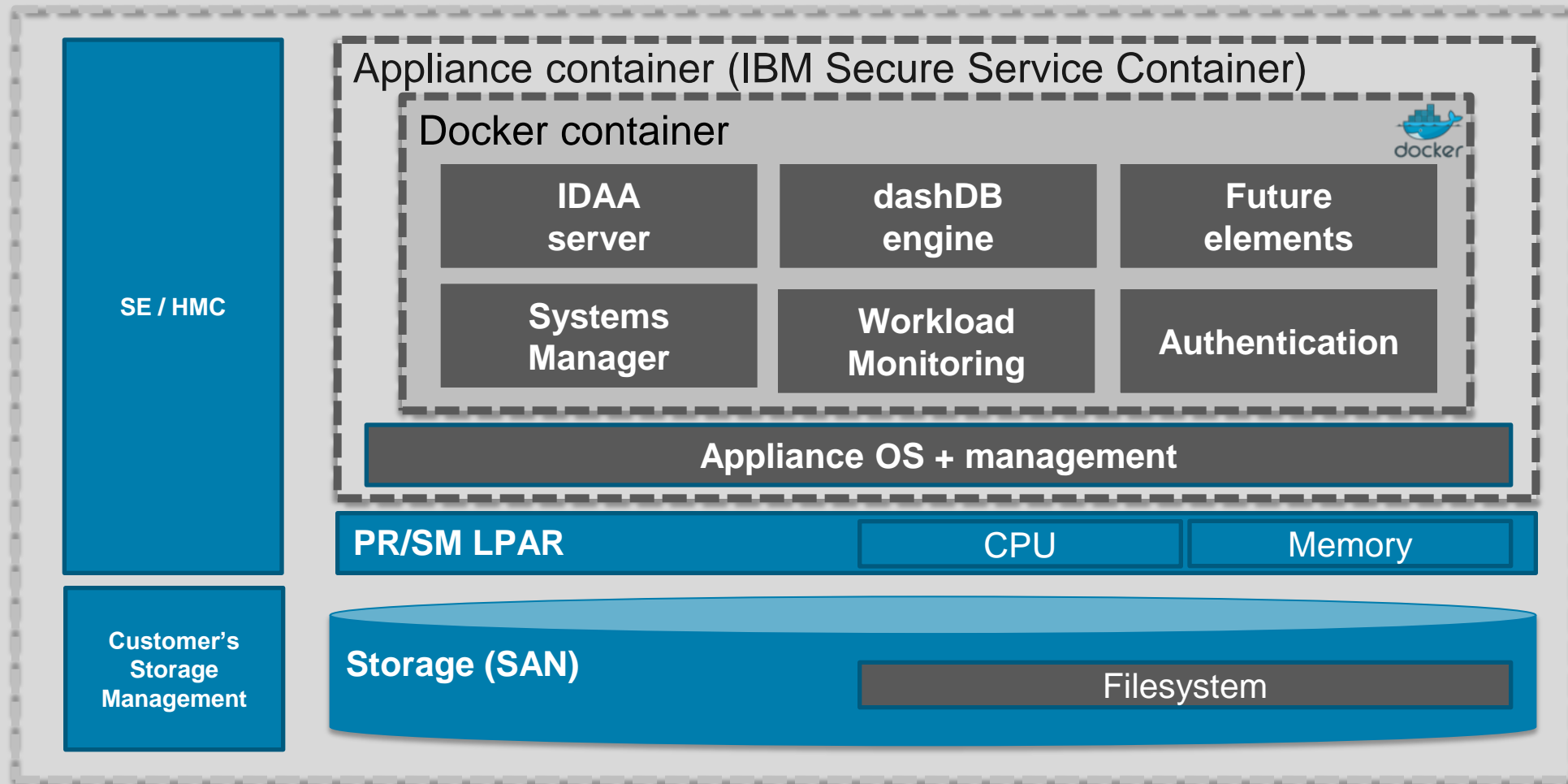


# IBM Db2 Analytics Accelerator – Product components





## Accelerator Architecture on Z



CPU, Memory, IO according to your requirements and infrastructure availability

Customer-provided

IBM-provided



Comparison

PDA N3001  
(Nettezza  
Technology)



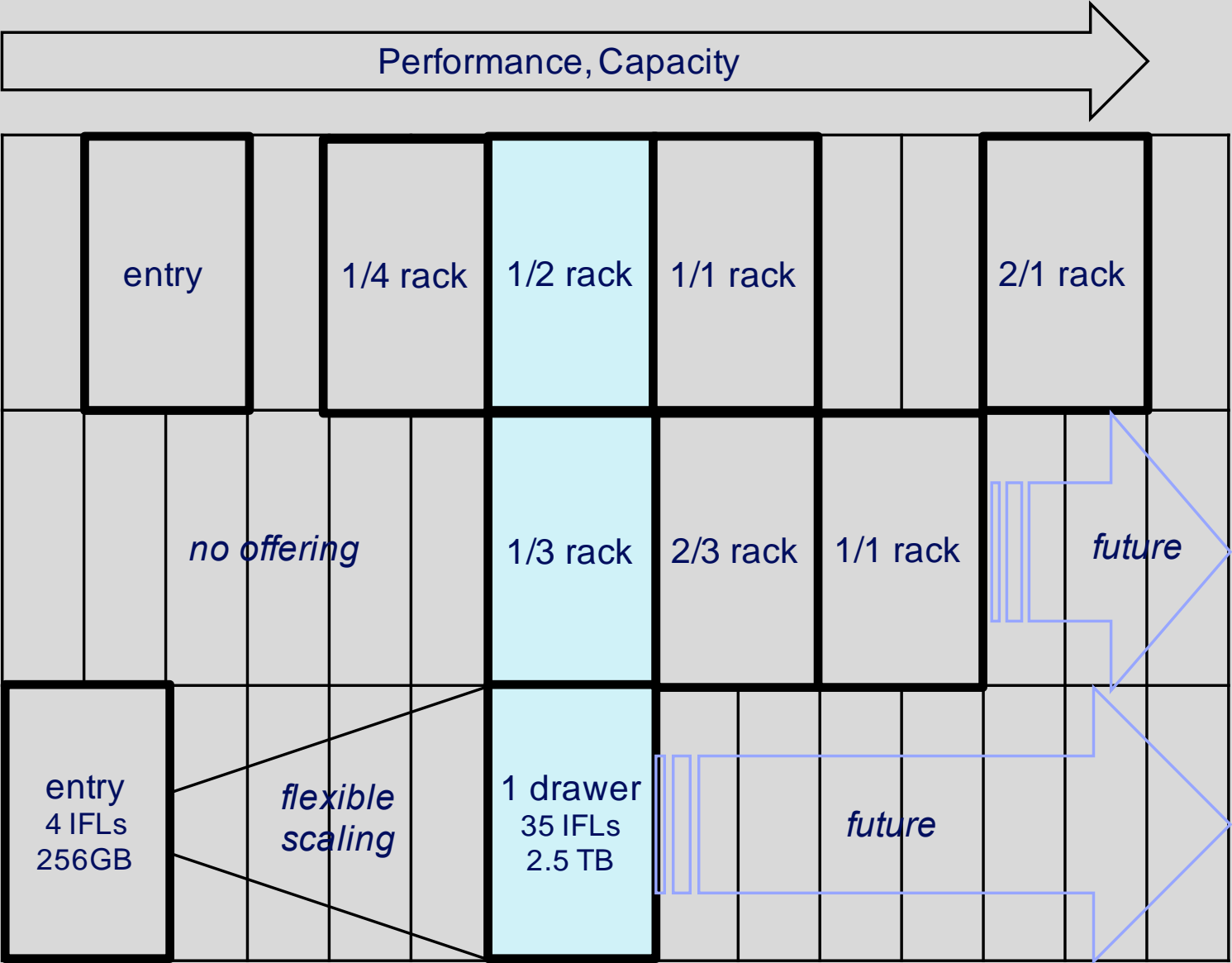
Integrated  
Analytics  
System M4001



Accelerator on  
IBM Z



New Generation Technology



## Why IDAA on z?

---

- A number of businesses in Citi are interested in testing IDAA, but many have been hesitant to commit to do it.
  - Large capital expense of Netezza hardware.
  - Most want an upfront guaranteed MIP saving.
  - IBM is hesitant to ship out Netezza boxes for POCs unless they have solid use cases from the business and a strong commitment to test. Citi generally will not agree to do a POC unless we secure funding to pay for whatever we are testing if the POC is a success.
- A Netezza box is not a mainframe.
  - There can be extensive outages to do hardware maintenance and some software upgrades. Not much can be done concurrently like on mainframes.
- Disaster recovery – Netezza based IDAAs have to be either fully reloaded at the DR site or they must be loaded and/or replicated to remotely from the production system. And you need to buy a Netezza box that will not be used except for a few days a year for a DR test.
- Our hope is that IDAA on z will help to address most of these issues and help us to increase the adoption of IDAA at Citi.

# Planning for IDAA on z

---

- Target testing environment is to use the disaster recovery site for a Db2 z/OS data warehouse that currently uses IDAA.
  - DASD is mirrored between the production site and the DR site. There are Netezza TwinFin-12 boxes connected to prod and DR LPARs.
- The data that we targeted for testing was 3.2 TB uncompressed. IBM sized the LPAR for the IDAA appliance as follows:
  - 33 IFLs, 610 GB of storage.
  - 111 Mod27 - 1 for the image, 2 for appliance internal, 108 for data (slightly over 2 TB).
  - Hipersocket connection between container LPAR to target z/OS LPAR.
- We needed to upgrade the microcode on the z13 CEC we are using to FW SE = Bundle S47 and HMC = Bundle H36. HMC was fine but the CEC needed to be upgraded.
- DB2 11 needed to be at PUT level 1707. We also needed a new pre-GA version of the IDAA stored procedures in order to access IDAA on z.

## Installation of IDAA on z.

---

- SW Appliance Configuration.
  - Define a Secure Service Container (SSC) LPAR in the HMC. Feedback from z/OS system programmer who built the LPAR.
    - “The preparation of the LPAR profile for SSC was an intuitive process, very similar to other profile configurations done with IODF with the exception of IP INFO and the requirement for a PASSWORD which would be later needed to logon to the appliance which would be used to further configure the network and storage. “
  - Next step was to receive the appliance image from IBM FTP site and upload it to the SSC LPAR. Again the z/OS system programmer said it was easy straight forward process.
  - We then formatted over 100 UCBs, which took about an hour and then initialized components of the LPAR, which took another few hours.

## Installation of IDAA on z.

---

- Feedback on installation process:
  - We would like the ability to install a new image and at the same time keep the currently working image off to the side so if we experience a problem we could easily fall back without having to re-initialize all the components.
  - It would be nice to see the version of the appliance indicated at the logon screen after each new image install.
  - It would be helpful to have an explanation in the appliance configuration panels that gives a brief summary on each of the processes and what is being done.
  - Currently when a function is running it does not give a status as to how far it has gone. It would help to have an expected time of completion.
  - It would be good to have a log after the install is completed that can be reviewed with timings for each step.
- We also have concerns on the two userids, one for the container administrator and the other for the IDAA administrator. These userids are managed by the appliance itself and are always active. A person who knows these userids can log on at any time and update settings or shutdown the IDAA or the SSC LPAR. I believe IBM is working on enhancements in this area.

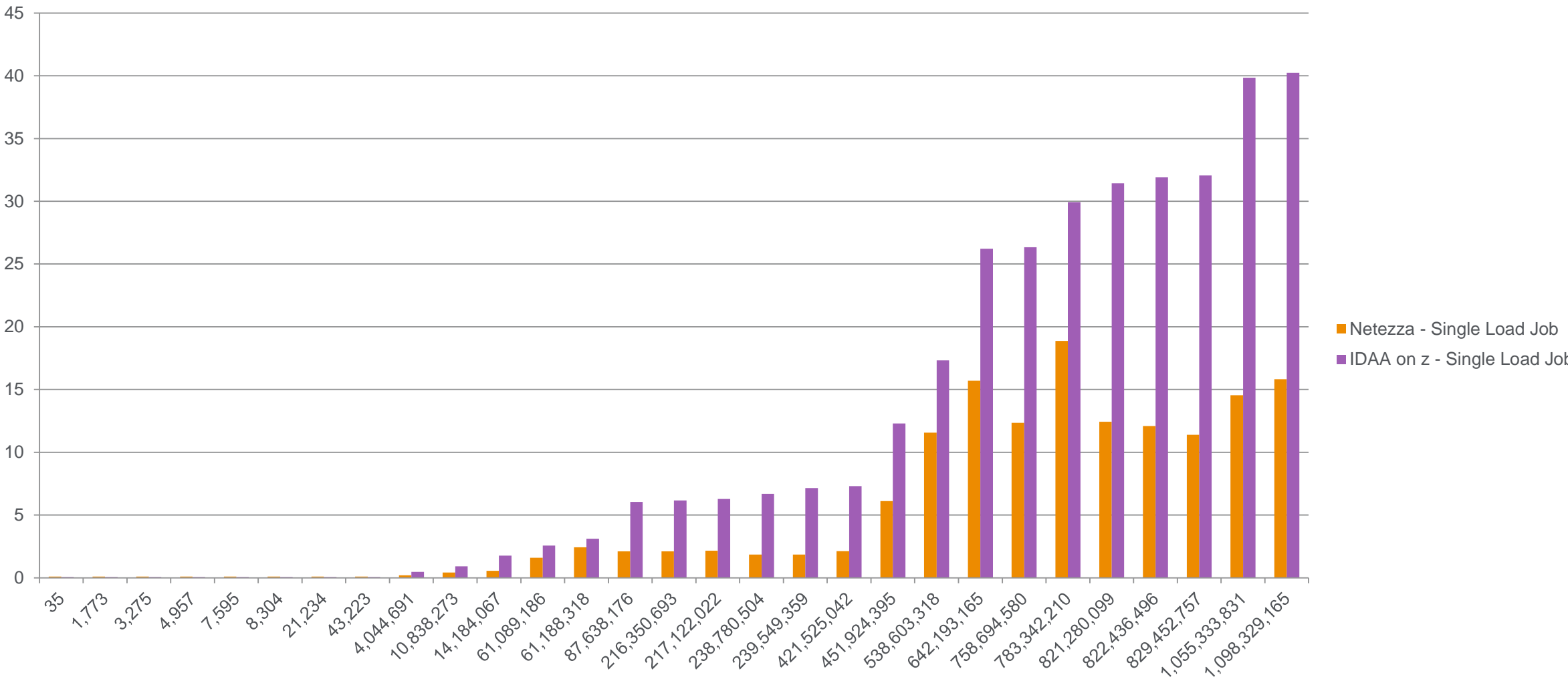
# Testing of IDAA on zSeries

---

- First step was to split the business continuance volumes at the DR site to get a consistent copy of the data on the production system. We then IPLed the DR LPAR.
- We then installed the new copy of the Db2 z/OS IDAA stored procedures into the DR DB2 subsystem and paired the subsystem with the IDAA running on the SSC LPAR.
- We also paired the subsystem with the DR Netezza IDAA using an older level of the IDAA stored procedures. The Netezza NPS level would have to be upgraded to make it compatible with version 7.1 of the IDAA stored procedures, which would make it unusable in the case of an actual disaster.
- After IDAA on z and Netezza where both paired we ran a series of load tests against both using the IDAA stored procedures called from batch jobs followed by a set of queries pulled from production also run from batch jobs.

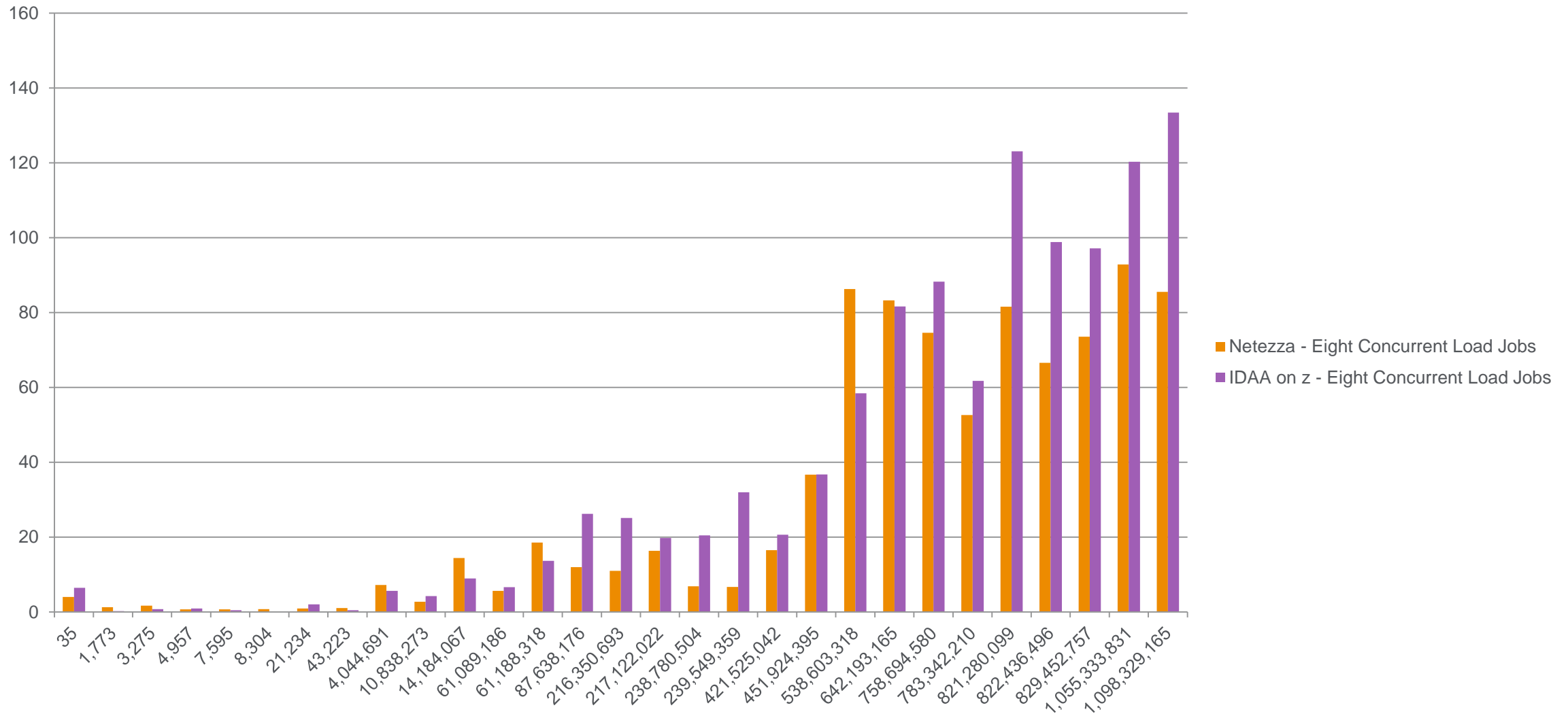
# Testing of IDAA on z – Table Loads Via Batch Jobs Using IDAA Stored Procedures

- Single batch job running. X axis is number of rows loaded. Elapsed time in minutes.



# Testing of IDAA on z – Table Loads Via Batch Jobs Using IDAA Stored Procedures

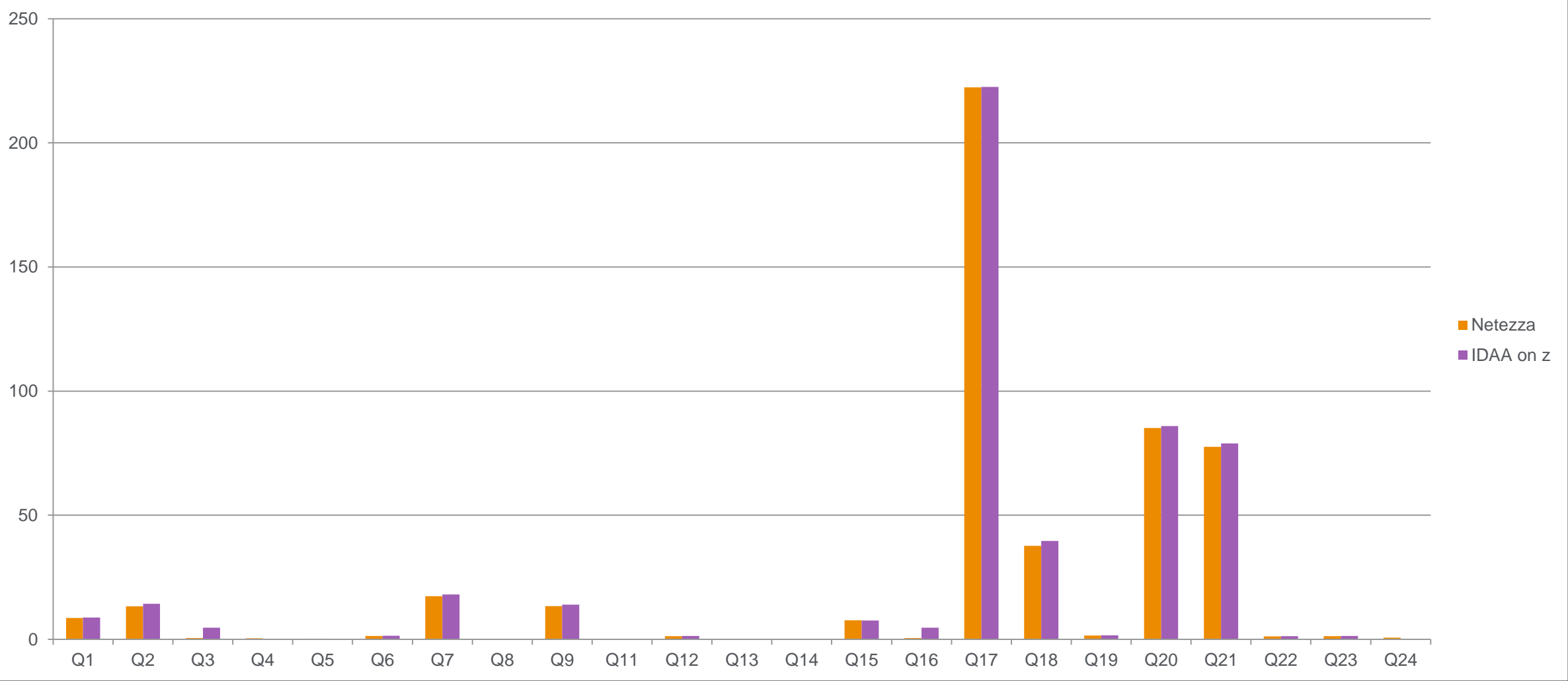
- Eight jobs submitted at same time. X axis is number of rows loaded. Elapsed time in minutes.





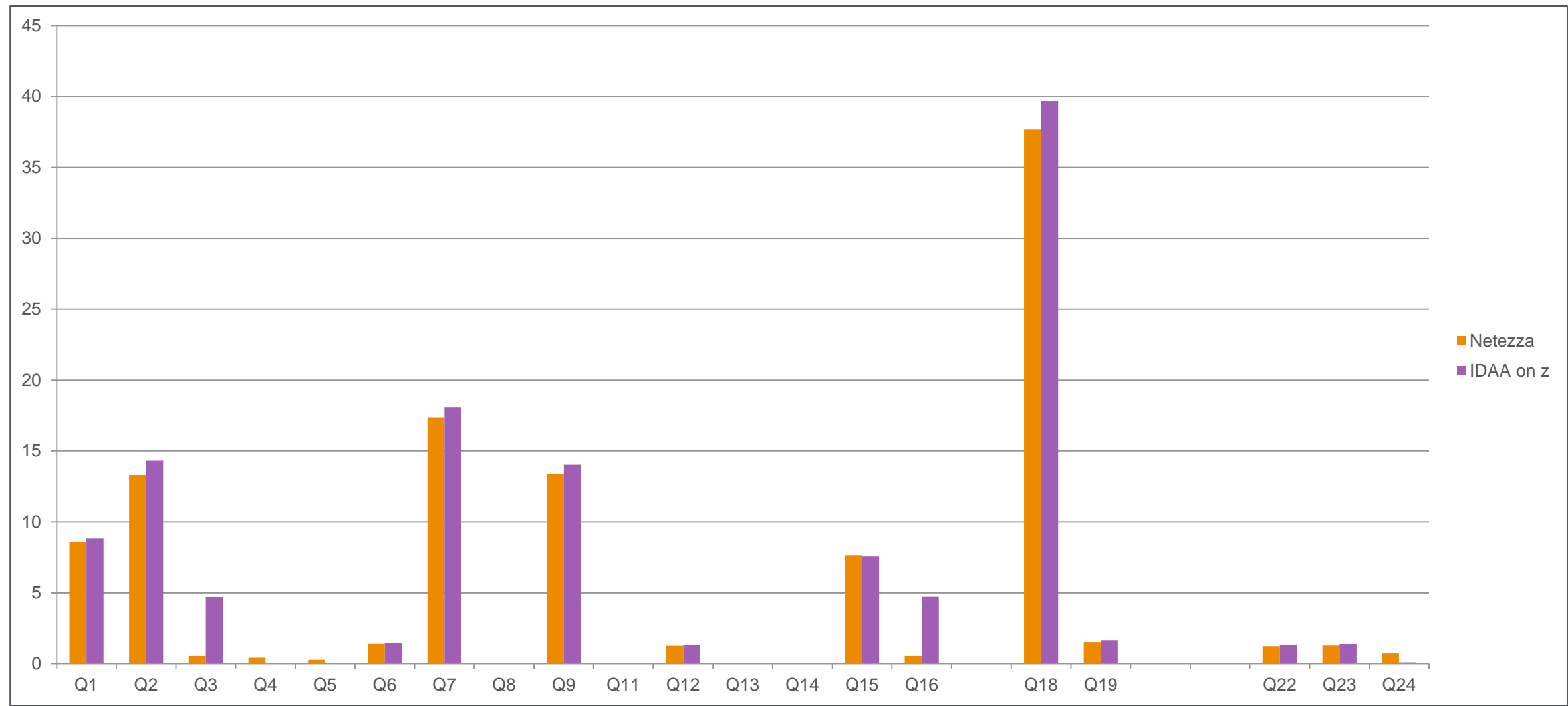
# Testing of IDAA on z – Query Tests

- Stand alone query testing in minutes.



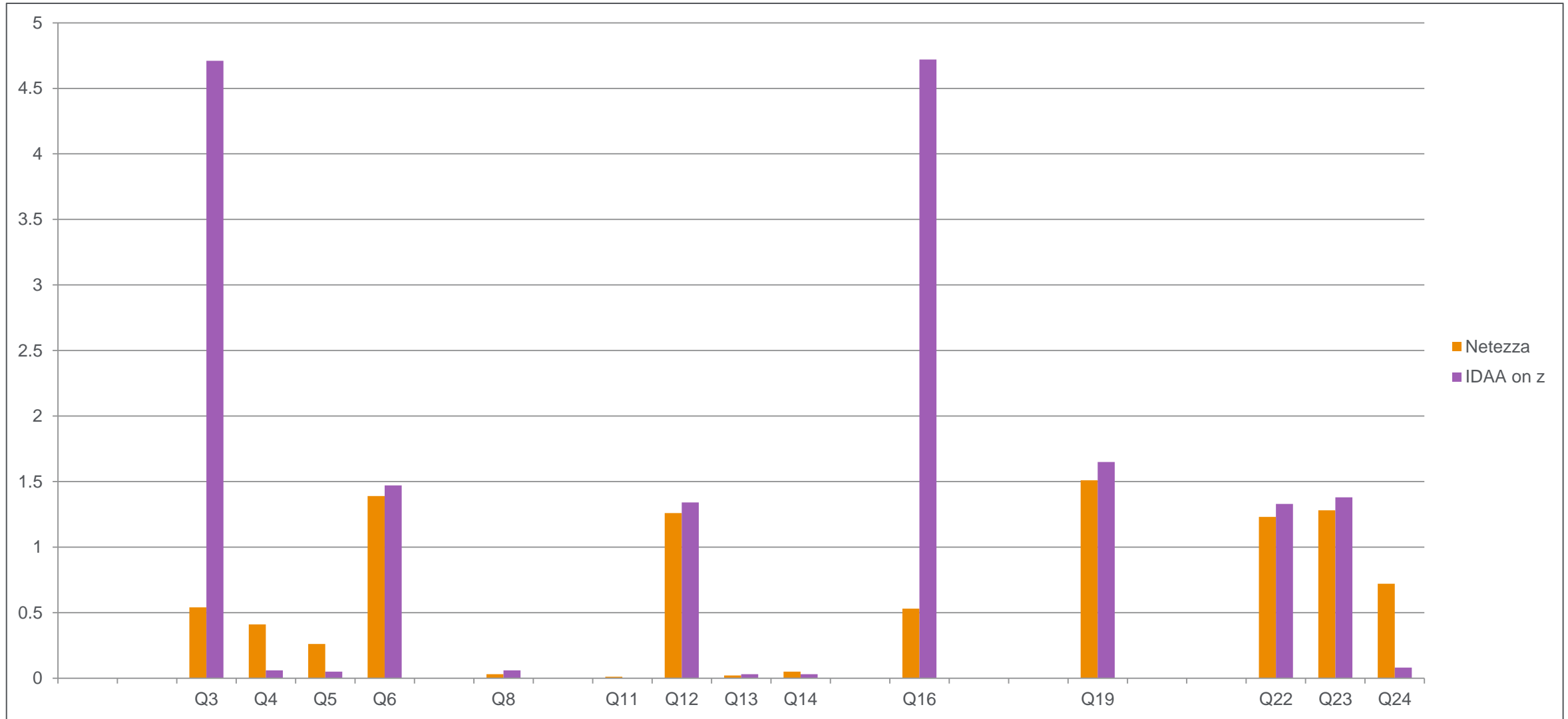
# Testing of IDAA on z – Query Tests

- Stand alone query testing in minutes – queries running longer than 60 minutes removed.



# Testing of IDAA on z – Query Tests

- Stand alone query testing in minutes– queries running longer than 5 minutes removed.



# Summary

---

- Query performance is in line with what we get from our current Netezza box. Most queries perform the same. Some are a bit slower and others faster. There has been a large improvement in query performance from the beginning of the POC to the GA code.
- As we started to run more concurrent workload on the IDAA on zSeries performance improved quite a bit.
- Loads via the DB2 supplied stored procedure are still a bit slower than on the Netezza but as you run more concurrent work the difference gets significantly smaller.
- We feel we have learned a great deal in a very short time regarding SSC LPARs and the IDAA on zSeries capabilities. We look forward to exploring future options at Citi to take advantage of this technology.
- As IDAA on zSeries evolves through the end of this year and into 2019 we look forward to enhancements including GDPS support that will be a major differentiator between this product, the Netezza and Integrated Analytics System offerings.