Off-host Processing: A Practical Approach

Colleen Gordon
Mainframe Data Integration Specialist, Luminex

Cathy Taddei
Infrastructure Systems Engineer

Discussion Topics

- Why process data off-host?
- How is it done?
- What are the challenges?
- A practical approach
- Use cases
- A User's Experience

Why Process Data Off-Host?

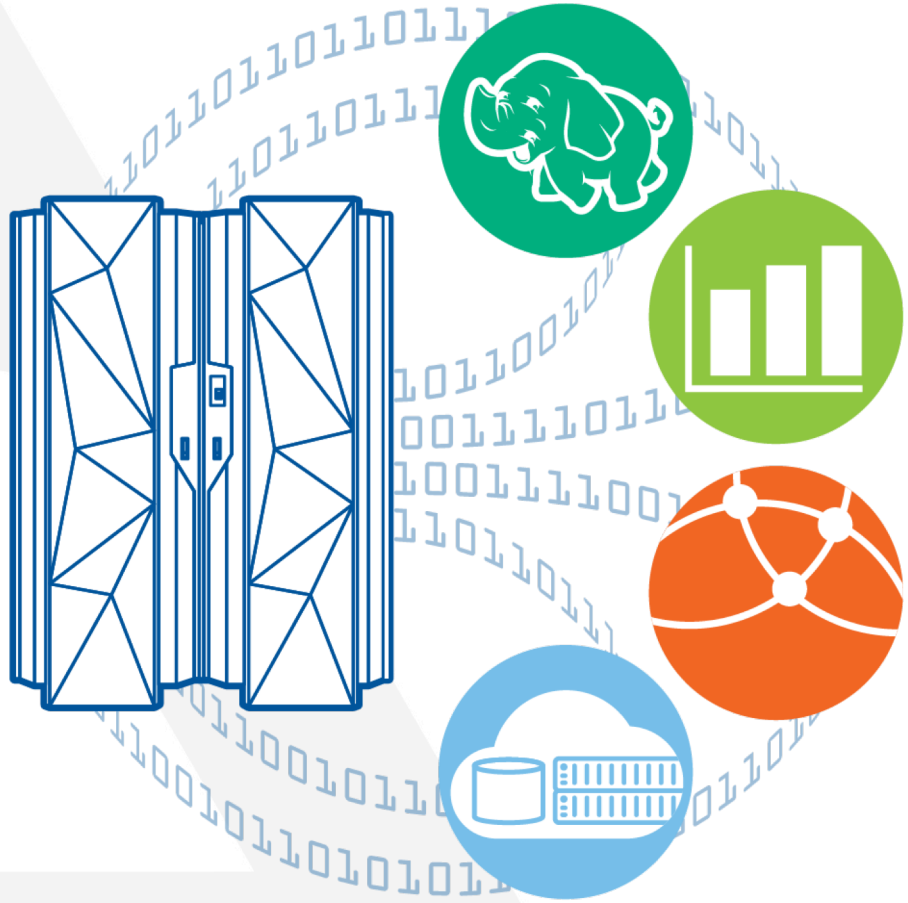
- Cost
  - MIPS
  - Licensing
  - DASD
- Performance
  - Cluster- and Cloud-compute
  - Purpose-built systems
- Flexibility
  - Fast deployment for new and/or temporary use cases
  - More choices and resources available off-host (who gets the biggest slice of the budget?)

Not all of these are true for all use cases, and there are challenges, so…

How Is It Done?

Steps for Off-Host Processing

- Identify applications or functions to off-host
- Move data from mainframe
- Process the data off-host
- Optionally, return results to the mainframe
- Continue processing of the data and/or report distributions



What are the Challenges?

- Data Movement
  - TCP/IP on the mainframe is slow and CPU-intensive
- Changes to Workflows
  - JCL
  - Rewriting applications
  - Learning new applications
- Security
  - Open ports
  - Digital certificates
  - Maintaining RACF-control
- Performance
  - Meeting current Service Level Agreements

A PRACTICAL APPROACH

Luminex Mainframe Data Integration (MDI) Platform

- An **off-host** data transfer and processing **platform**
- Leverages **mainframe-native FICON** channels for fast, efficient and secure transfer and communications
- **Profile-based architecture** enables highly-flexible deployments
  - Existing library of Profiles for popular use cases
  - Purpose-built Profiles
- Mainframe-centric design **keeps SAF in control**
- **Easy to implement** with minimal job changes
  - As simple as executing an IEBGENR

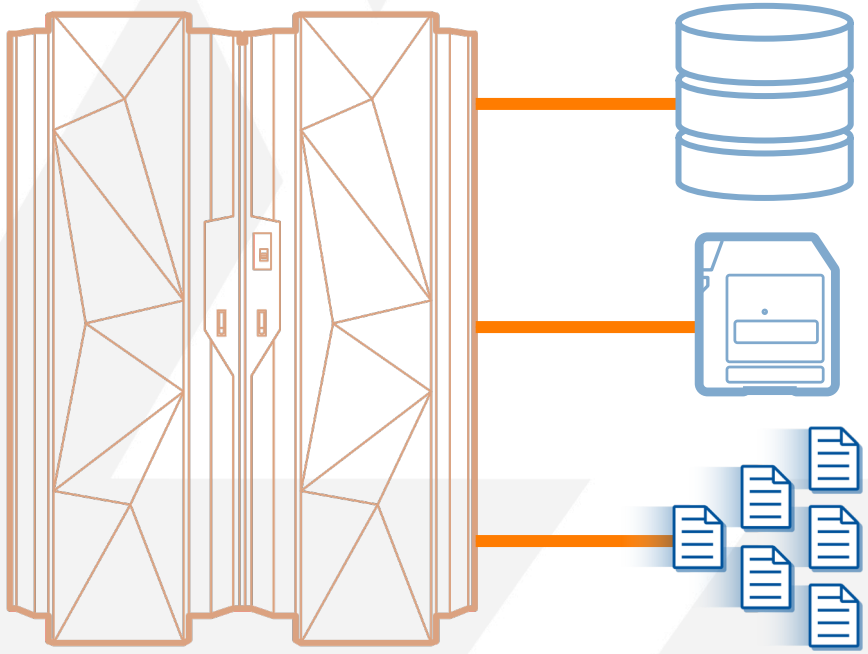
MDI enables mainframe data integration with enterprise-wide business applications and systems

The Key to Practical Off-Host Processing: FICON

FICON was designed *specifically for the mainframe*

Attributes

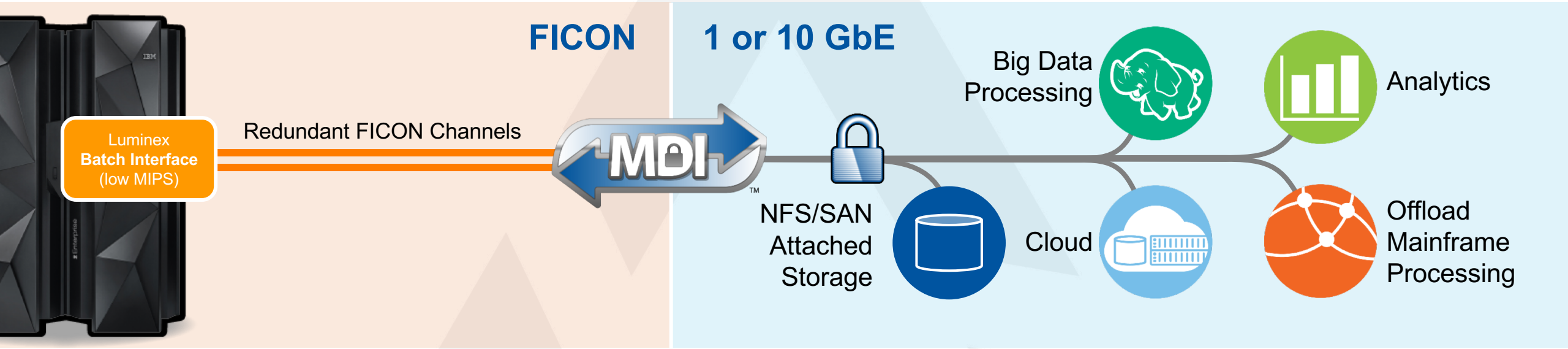
- Fast
- Efficient
- Secure



Uses

- DASD
- Tape
- … and now Off-Host Processing

MDI Overview



MDI Batch Interface

- Simple JCL
- Started Task (in some cases)
- SAF Security interface
  - RACF and others

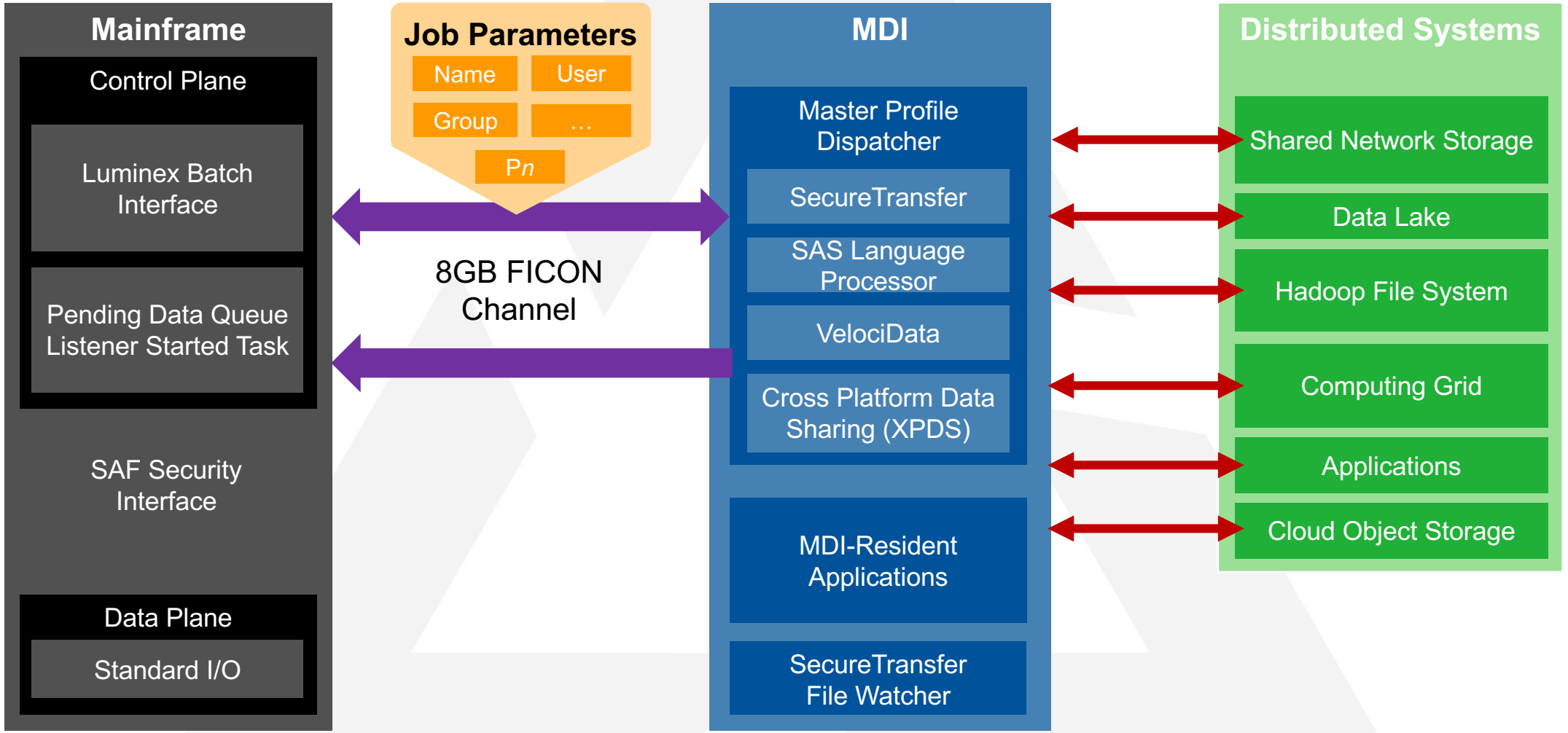
MDI Platform

- Linux server (2U)
- Looks like a virtual tape device to the mainframe
- MDI Profiles
- SAS language compiler
- File Watch

MDI Interfaces

- VelociData Vortex
- Apache Ignite
- Azure Data Lake
- Amazon AWS
- Cloud Object Storage
- NFS/SAN Storage
- SFTP
- And more…

MDI Architecture

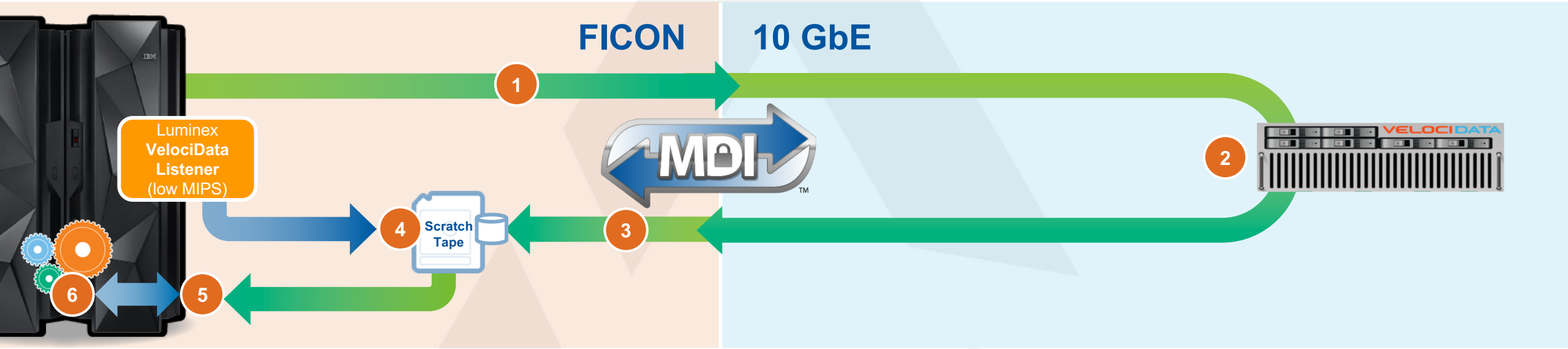


OFF-HOST PROCESSING USE CASES

Use Case: VelociData Vortex with MDI Stream Computing Interface (SCI)

- Real-time, in-line data transformation appliance
  - Custom hardware designed specifically to accelerate ETL operations, such as:
    - Sort
    - Join
    - Mask
    - Look Up and Replace
    - Schema Conversion
    - Field Content Validation
- Resulting data is made available to the host faster and more cost-effectively than native processing
  - Sort and aggregate data at 10 GbE line rates
  - Transform and cleanse millions of records per second
  - Secure (mask or encrypt) 10 million fields per second
  - Extract features from data feeds, supplying analytics with high value information

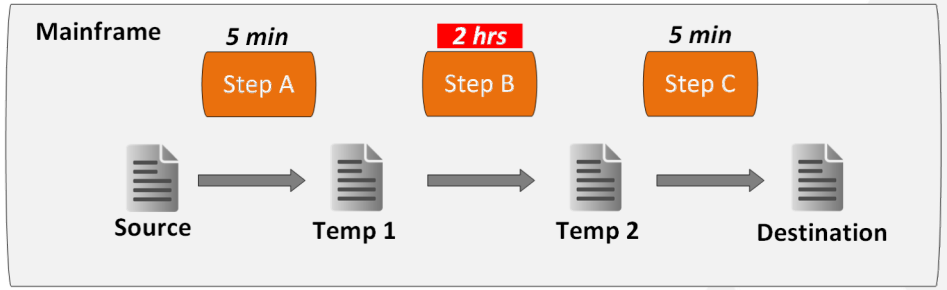
How It Works: SCI with VelociData



1. Data is transferred from the mainframe to VelociData
2. VelociData transforms the data in-line and streams it back to MDI
3. The VelociData Listener mounts a scratch tape and opens a new data set
4. SCI copies file to new output data set
5. The VelociData Listener closes the data set and catalogs it on mainframe
6. Catalog event “triggers” downstream batch processing to be initiated

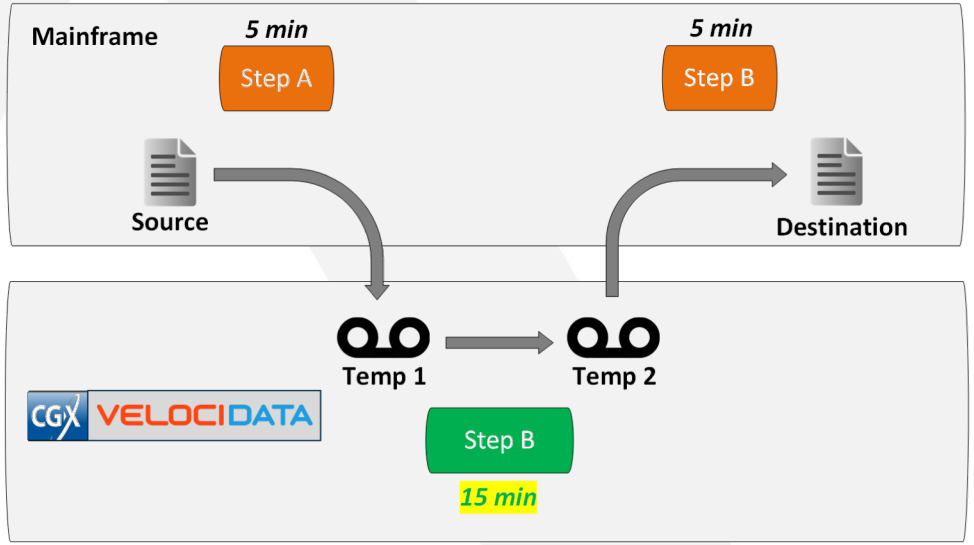
Usage Picture: Job / MIPS Offload & Acceleration

BEFORE: Typical Mainframe JCL



- Common multi-step mainframe JCL job
- *Note that middle task consumes majority of processing time / cost*

AFTER: VelociData Offload / Acceleration

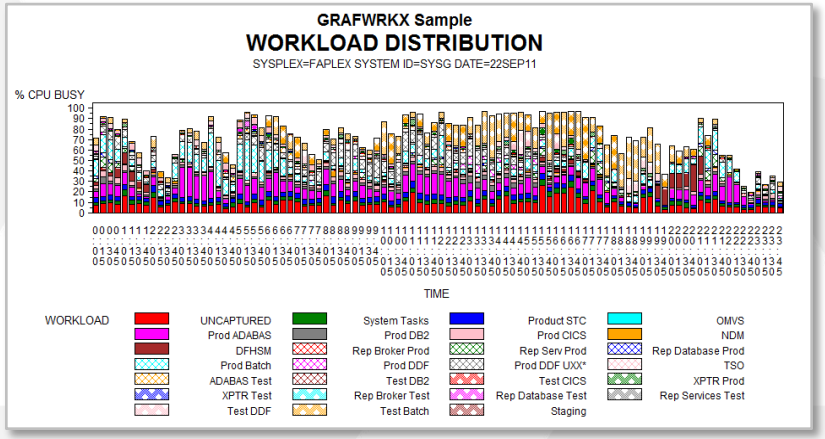


- Offload and accelerate critical steps of workflow
- Drastic cost savings ***and*** process time reduction
- Retain JCL control of operations

Use Case: SMF Reporting with MXG and MDI SAS Language Processor (SLP)

55.8% use mainframe batch processes to analyze collected performance data

- MXG has about 85% of the market share
- MXG uses SAS language to create data center reporting



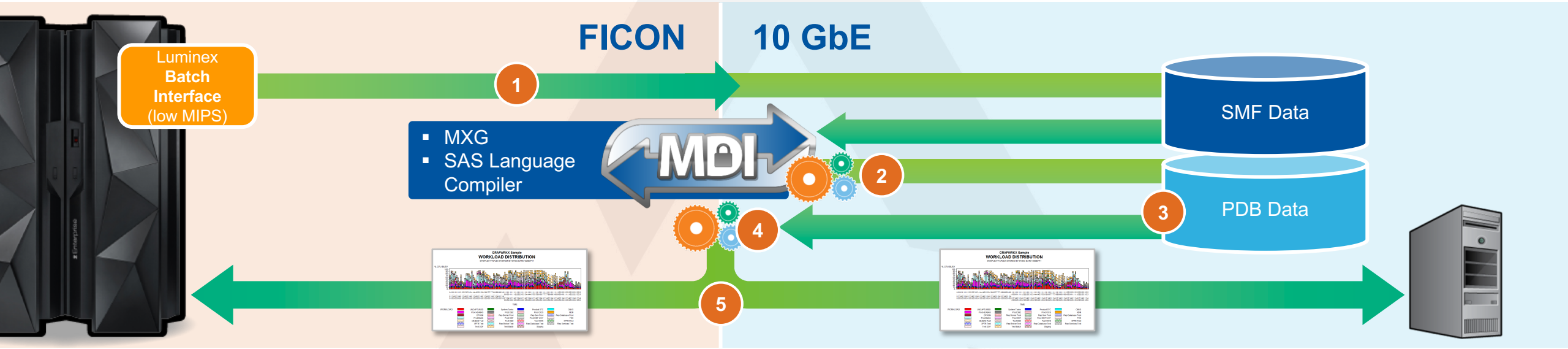
"State of the Mainframe for 2017" Syncsort Survey of IT Professionals *What Every Business Needs to Know About Big Iron and Big Data: Facts and Trends, January 2017*

MXG: SMF Reporting Tools

- One of the most widely used workload monitors for mainframes
  - Typically supported through costly SAS licensing
- MXG is a site license
  - If you are licensed for MXG on the mainframe, you are also licensed for the ASCII version
    - Merrill Consultants can provide services to help install and configure
- Minor differences between z/OS and ASCII version
  - No GDGs on ASCII version
    - MXG will create directories with dates embedded in the directory names
      - Age them off based on user supplied parameters
- On z/OS
  - Multiple jobs
    - Daily/Weekly/Monthly/Trend

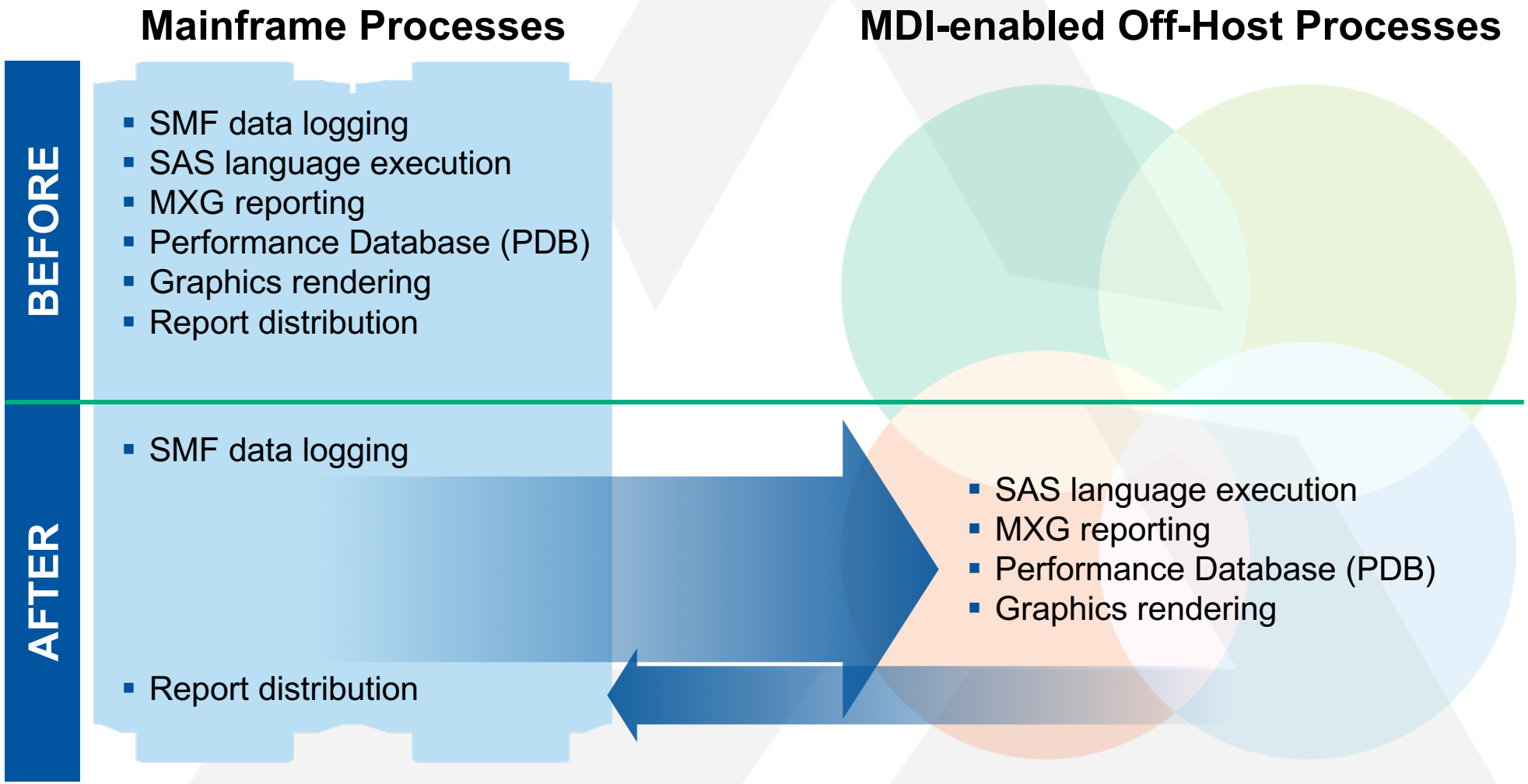
- On ASCII Version
  - 1 job does it all
  - Significantly more efficient graphics rendering

How It Works: SLP with MXG

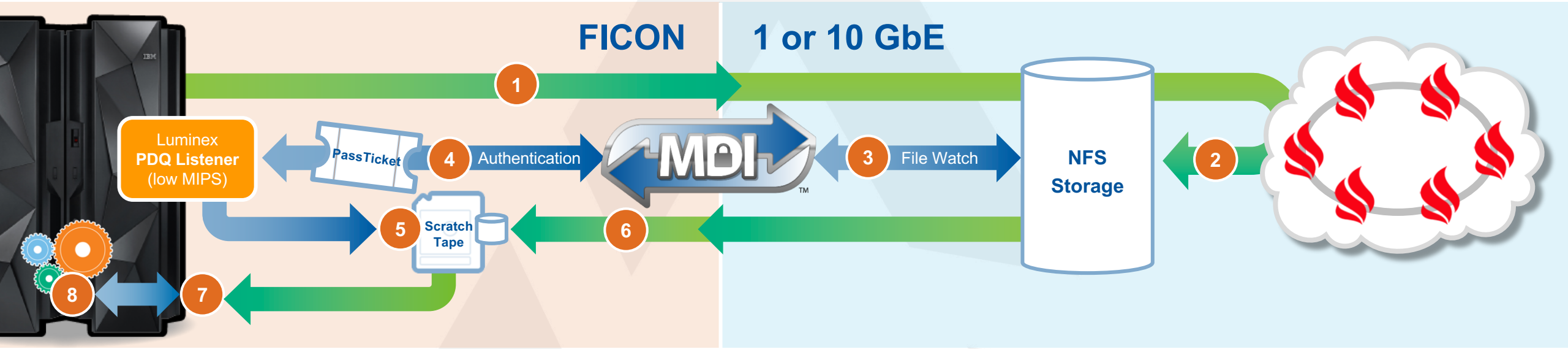


1. MDI Batch Interface moves SMF data off the mainframe
2. ASCII version of MXG is executed on MDI off-host processor saving MIPS
3. MXG PDBs written to storage attached to the MDI Platform
4. Client executes all MXG/SAS programs with data in the PDBs
5. Output returned to the mainframe; distributed locally; emailed, CSV files, etc.

Offloaded Processes for MXG Reporting

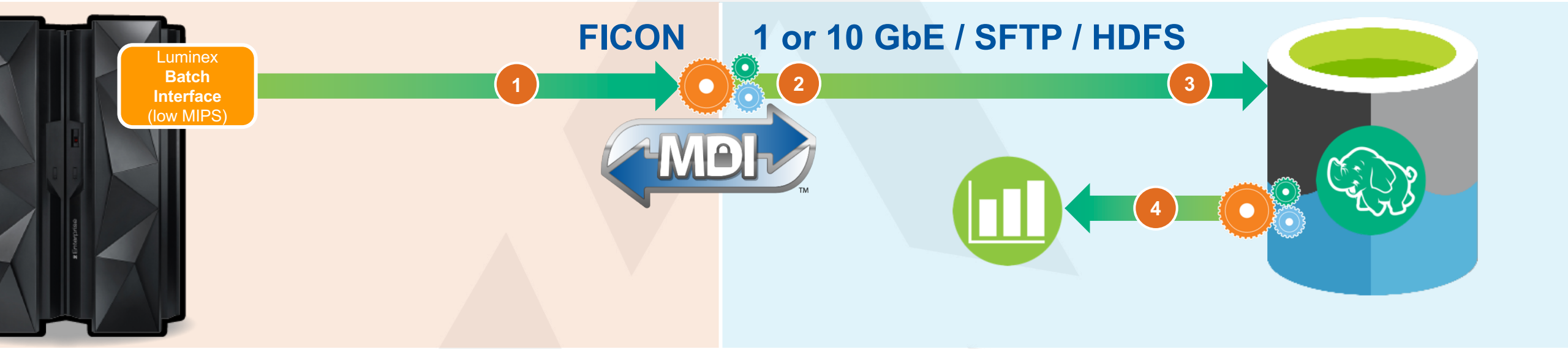


Use Case/How It Works: Apache Ignite Compute Grid with MDI Cross Platform Data Sharing (XPDS)



1. File is transferred from the mainframe to Apache Ignite NFS
2. Ignite picks up the file and processes; places the file in a designated folder on the NFS
3. XPDS File Watch detects the file and communicates to PDQ on the mainframe
4. XPDS constructs a security PassTicket and passes to PDQ for validation
5. PDQ mounts a scratch tape and opens a new data set
6. XPDS copies file to new output data set
7. PDQ closes the data set and catalogs it on mainframe
8. Catalog event “triggers” downstream batch processing to be initiated

Use Case/How It Works: Azure Data Lake and Hadoop with MDI SecureTransfer



1. MDI Batch Interface moves data off the mainframe
2. Data is transformed from EBCDIC to ASCII
3. Data is sent to HDFS for ingestion
4. Client uses tools to manipulate data and create outputs

JCL Interface to MDI

```
//FTP8     JOB (EARL),'8   TAPE',CLASS=W,MSGCLASS=W,
//*  RESTART=XPROC.X,
//   NOTIFY=&SYSUID,
//   REGION=0M
// JCLLIB ORDER=(LUMENG.MDI.JCL)
// SET DEV=MDI9900
// SET F1DSN=TEST.SMF.T1.COPY1.DD1
// SET F2DSN=TEST.SMF.T1.COPY1.DD2
// SET F3DSN=TEST.SMF.T1.COPY1.DD3
// SET F4DSN=TEST.SMF.T1.COPY1.DD4
// SET F5DSN=TEST.SMF.T1.COPY1.DD5
// SET F6DSN=TEST.SMF.T1.COPY1.DD6
// SET F7DSN=TEST.SMF.T1.COPY1.DD7
// SET F8DSN=TEST.SMF.T1.COPY1.DD8
// SET PROFILE=SFTP
```

```
//XPROC1  EXEC  LUMXPROC,PROFILE=&PROFILE
//XPROCLOG DD  SYSOUT=*
//DOWN1    DD  DISP=OLD,DSN=&F1DSN,
//             UNIT=(&DEV,,DEFER)
//DOWN2    DD  DISP=OLD,DSN=&F2DSN,
//             UNIT=(&DEV,,DEFER)
//DOWN3    DD  DISP=OLD,DSN=&F3DSN,
//             UNIT=(&DEV,,DEFER)
//DOWN4    DD  DISP=OLD,DSN=&F4DSN,
//             UNIT=(&DEV,,DEFER)
//DOWN5    DD  DISP=OLD,DSN=&F5DSN,
//             UNIT=(&DEV,,DEFER)
//DOWN6    DD  DISP=OLD,DSN=&F6DSN,
//             UNIT=(&DEV,,DEFER)
//DOWN7    DD  DISP=OLD,DSN=&F7DSN,
//             UNIT=(&DEV,,DEFER)
//DOWN8    DD  DISP=OLD,DSN=&F8DSN,
//             UNIT=(&DEV,,DEFER)
```

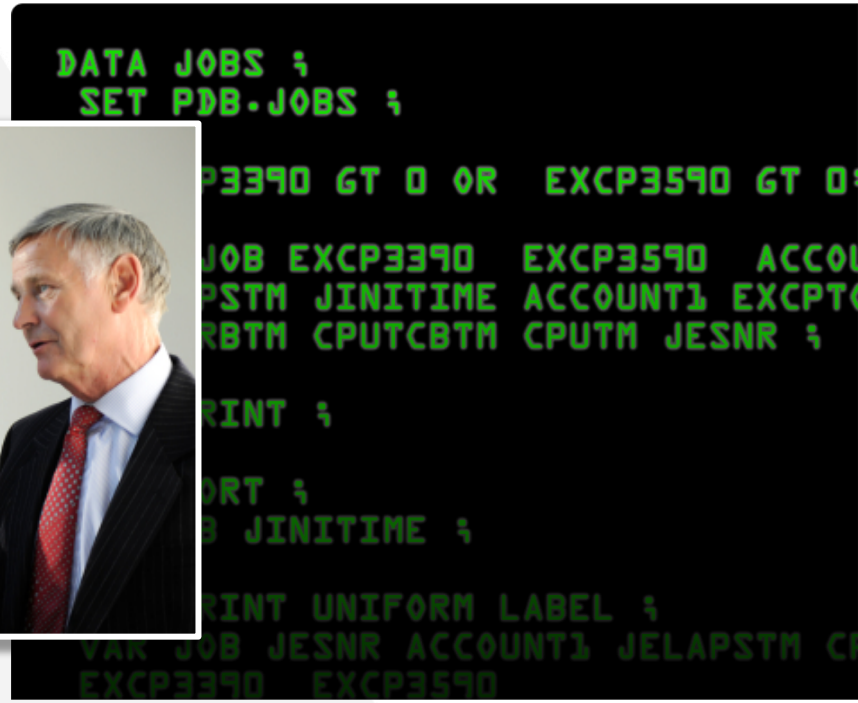
Large File JCL Example and Timings

```
//SYSIN    DD *
-PARM=destination=SERVER; parallel=8;
login=UID;
emailall=notify@luminex.com;
preaction="LISTFILEL *";
postaction="LISTFILEL *";
-DD_DOWN1=TEST.SMF.T1.COPY1.DD1 oformat=ascii;
-DD_DOWN2=TEST.SMF.T1.COPY1.DD2 oformat=ascii;
-DD_DOWN3=TEST.SMF.T1.COPY1.DD3 oformat=ascii;
-DD_DOWN4=TEST.SMF.T1.COPY1.DD4 oformat=ascii;
-DD_DOWN5=TEST.SMF.T1.COPY1.DD5 oformat=ascii;
-DD_DOWN6=TEST.SMF.T1.COPY1.DD6 oformat=ascii;
-DD_DOWN7=TEST.SMF.T1.COPY1.DD7 oformat=ascii;
-DD_DOWN8=TEST.SMF.T1.COPY1.DD8 oformat=ascii;
//
```

- Large SMF File Transfer
  - 1.5 Terabytes of data
    - 8 concurrent transfers
  - Total CPU
    - 3 minutes 56 seconds
- FTP of the same 8 SMF files
  - Consumed 23.3 hours of CPU
- Processing Times
  - 333GB SMF data
  - 3 PDBs per LPAR
  - 10 Concurrent processes
    - Processing complete in 1 hour 10 minutes

Luminex Professional Services

- Discovery
- Analysis
- Planning for appropriate migration strategy
- Implementation



Benefits of Off-Host Processing with MDI

SHARE
EDUCATE • NETWORK • INFLUENCE

 **Secure**	▪ More secure than TCP/IP on the mainframe ▪ Reduce/eliminate open ports on the mainframe ▪ Maintain mainframe-centric job control and security
 **Fast**	▪ Unmatched transfer rates, up to 800 MB/s per MDI Platform ▪ Take advantage of high-performance purpose-built systems
 **Efficient**	▪ Reduce mainframe CPU overhead for mainframe (applications, TCP/IP, etc.) ▪ Selectively move applications off the mainframe ▪ Minimal changes to current workflows
 **Cost-Effective**	▪ Reduce software licensing costs ▪ Free up expensive DASD storage by moving data to commodity storage ▪ Licensing not based on MIPS/MSUs
 **Ease of Implementation**	▪ Can be as simple as executing an IEBGENER ▪ JCL Conversion Utility and Professional Services

CATHY TADDEI
INFRASTRUCTURE SYSTEMS ENGINEER

About Me

- First time speaker at SHARE!
- Live near Portland, Oregon
- Started working with mainframes in 1982
- MVS Systems Programmer for 11 years, DB2 for 4, MVS for 2, DB2 for 12, z/OS Engineer for 5, Freelance
- Successfully outsourced myself by getting all mainframe hardware and software components upgraded, documented, 5 year capacity plan, and turned over to outsourcer teams

Mainframe Environment

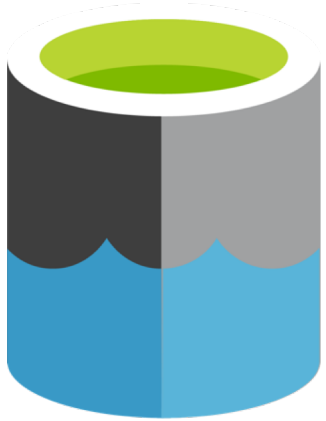
- Major Applications
  - CICS / VSAM
  - SAS Batch
- Primarily Batch; fewer than 1 million transactions/day
- Tightly coupled with Oracle on Unix and Windows apps via ETL, FTP, MQ, Classic Federation

- IBM z13s
  - 851 MIPS
    - 3 LPARS
- IBM DS8884
  - 11 TB, backed up to virtual tape
- Luminex CGX virtual tape
  - Replicated to DR

Business Environment

- “Boutique” Insurance Company
  - Specializing in LTD and Life Insurance via Employee Benefits
  - 3,000 employees
- “Bread and butter” applications still run on the mainframe
  - The mainframe is a poorly understood “black box”
  - Data of record trapped in VSAM.
  - Very few developers brave the mainframe. For those who do, SFTP is “too hard” = “not available”, so they work around it.

The Azure Data Lake Use Case

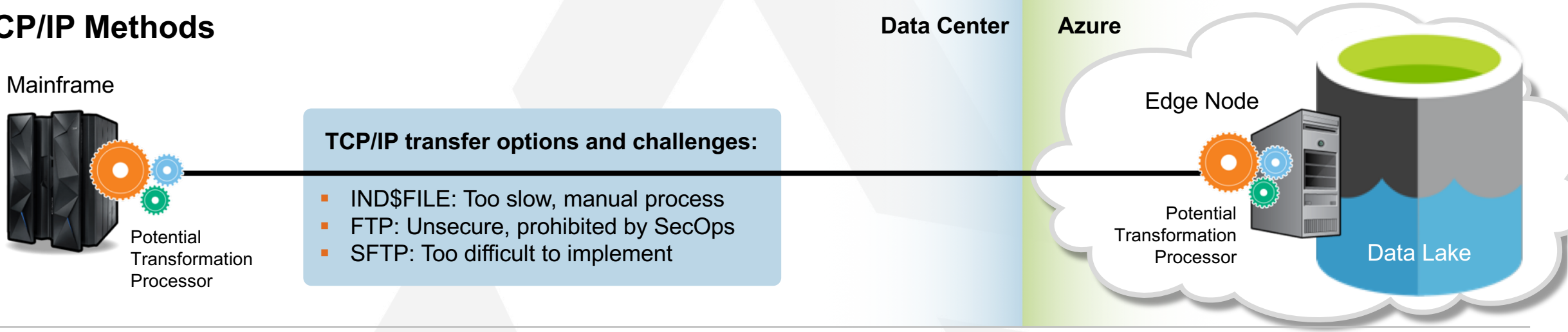


- The Data Lake
  - Pump in data from all sources
  - Analyze to discover hidden correlations and questions you didn't know you had
- The actuaries had a question:
  - When are you going to die? Actually, a bit more complex…
- How had we been answering this question?
  - Let's check…4 mainframe VSAM applications plus Oracle
  - Put the data in a spreadsheet, manually maintain for 20 years
- Azure / Hadoop – a better way?
  - How do we get the data there?

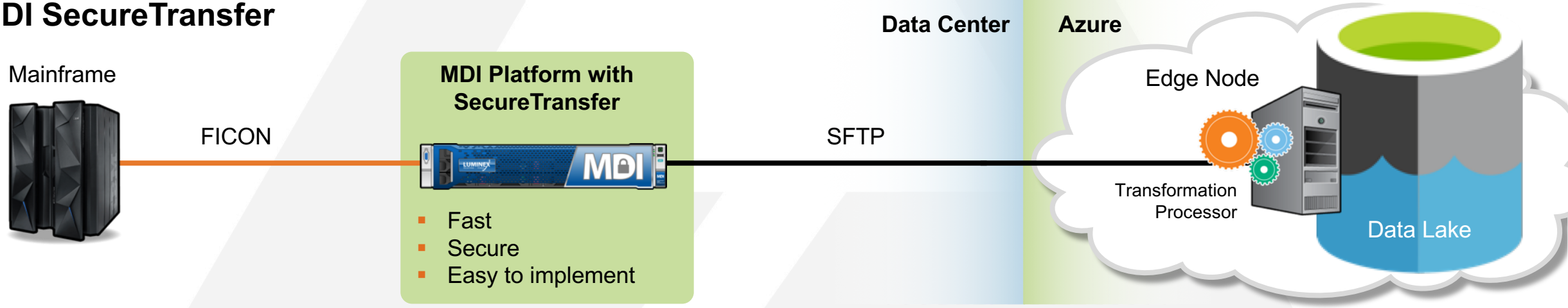
Options for Moving Mainframe Data to the Data Lake

SHARE
EDUCATE • NETWORK • INFLUENCE

TCP/IP Methods



MDI SecureTransfer



The MDI Proof of Concept

- Criteria
  - Security
  - Speed
- Process
  - Use MDI to transmit VSAM data to Azure “edge node”
  - Transform and put into Azure Data Lake
  - Analyze with U-SQL and Power BI

The MDI Proof of Concept Results

More than met the actuaries' requirements in 4 week-long sprints

- Security
  - Data access controlled via the MDI Profile and RACF, and via permanent SSH keys between MDI and the edge node
- Speed
  - Faster than regular, unsecure FTP from z/OS
- BONUS: Easy to Use!
  - Luminex support created profiles and generated SSH keys
  - Took 30 minutes to automate JCL creation
  - Took another 30 minutes to automate output reporting

Other MDI Use Cases

- SIEM
  - Expose Security Information and Events
- Lights-out operation
  - Ship all logs to global monitor to generate alerts
- EDI
  - Convert unsecure FTP to SFTP
- … These are just the use cases identified for this company
  - Since each environment is different, it's easy anticipate others envisioning even more new use cases

SHARE
EDUCATE • NETWORK • INFLUENCE

Thank You for Attending!

Please remember to complete your evaluation of this session in the SHARE mobile app.

Off-Host Processing: A Practical Approach

Session 21742