Data Science & Engineering

Cahyo Listyanto
Head of Data Science and Engineering
Technology Department

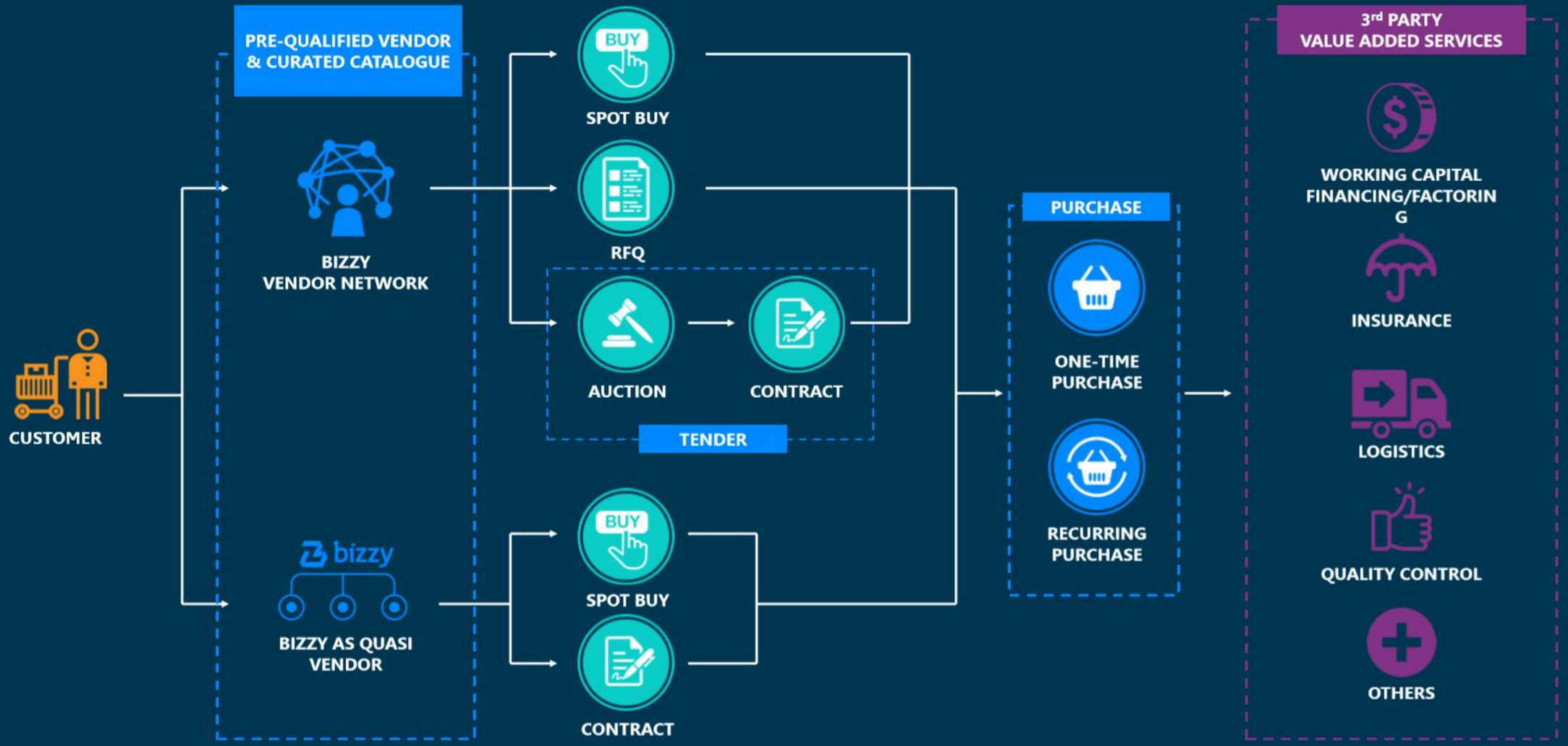
November 2018

- About Bizzy
- Data Case Study
- People
- Process
- Technology
- QnA

Bizzy has relentless MISSION to accomplish

To underpin a clean economy by powering an inclusive digital business ecosystem that enables transparent, efficient and accountable commerce for all stakeholders

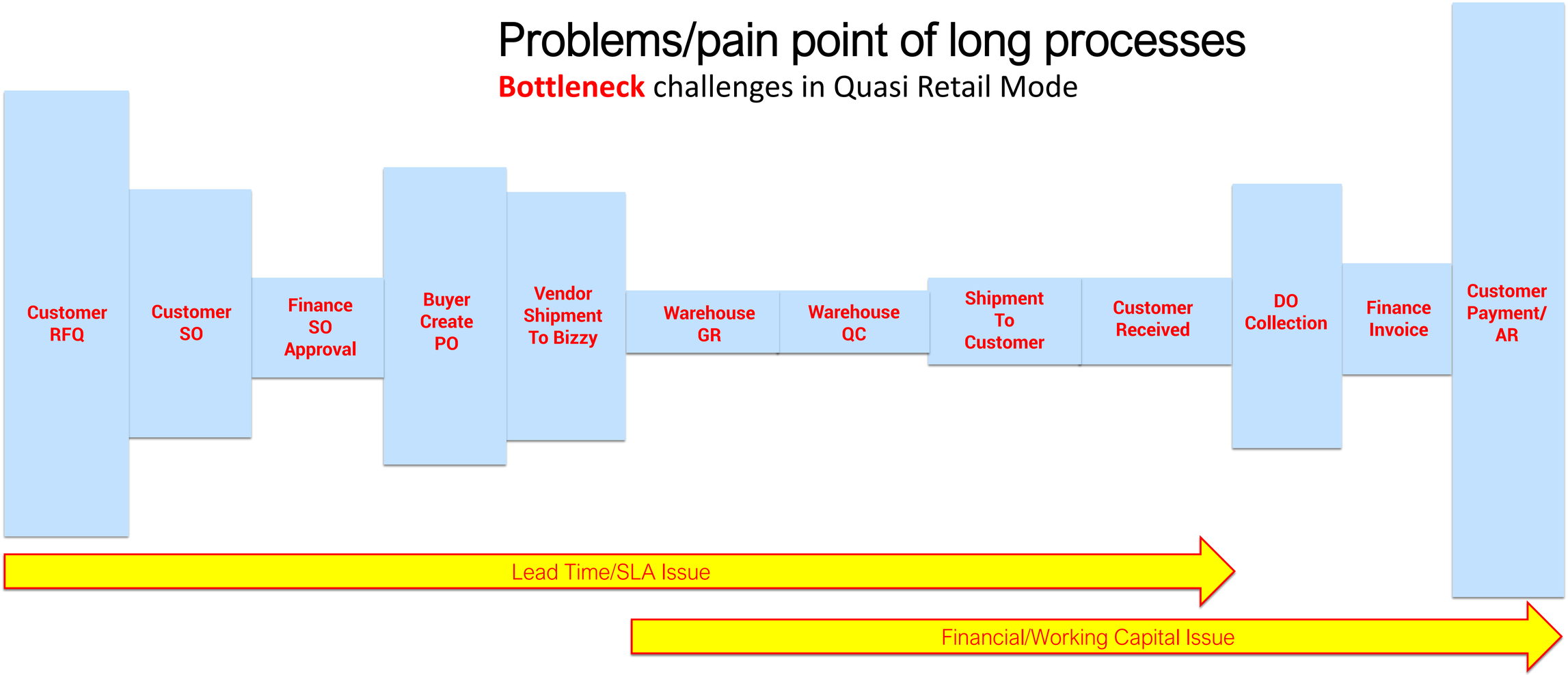
Bizzy enables execution of strategic, tactical and operational procurement activities



Data Case Study

Problems/pain point of long processes

Bottleneck challenges in Quasi Retail Mode



Data Case Study

Daily Notification of end to end process

Addressing bottleneck to increase SLA

Internal Notification:

- Finance Pending Approval
- Buyer Pending PO(Back Order)
- Vendor Pending GR
- Pending Shipment
- Pending DO Collection
- Finance Pending Invoice
- Vendor Pending Payment
- Customer Pending Payment
- Etc

Add Machine Learning prediction to detect potential bottleneck item

bizzy

Data Case Study

Daily Notification of end to end process

Addressing bottleneck to increase SLA

External Notification:

- Customer AR Statement Reminder
- Vendor Payment Notification
- Customer Weekly Spending Summary

Delivery Channel

- Chat Reporting
- Email Reporting
- Reports Web Portal
- Power BI Dashboard
- Excel Reporting

Leverage channel which the user **used** in day to day operation

Delivery Channel – Chat/Text Messages

Cahyo Listyanto 8/6 7:45 PM

Daily Performance Highlight

Period (D-1) : 6 Agustus 2018

	D-1	MTD	QTD	YTD
GMV - MP	Rp123.000	Rp200.000 (100%)	Rp200.000 (100%)	Rp200.000 (100%)
GMV - QR	Rp200.000	Rp200.000 (100%)	Rp200.000 (100%)	Rp200.000 (100%)
GMV - Total	Rp200.000	Rp200.000 (100%)	Rp200.000 (100%)	Rp200.000 (100%)
NMV - MP	Rp200.000	Rp200.000 (100%)	Rp200.000 (100%)	Rp200.000 (100%)
NMV - QR	Rp200.000	Rp200.000 (100%)	Rp200.000 (100%)	Rp200.000 (100%)
NMV - Total	Rp200.000	Rp200.000 (100%)	Rp200.000 (100%)	Rp200.000 (100%)
Revenue	Rp200.000	Rp200.000 (100%)	Rp200.000 (100%)	Rp200.000 (100%)
Invoice	Rp200.000	Rp200.000 (100%)	Rp200.000 (100%)	Rp200.000 (100%)

Finance (in Mio)

AR Overdue : [redacted]
Total AR : **Rp.**
Total AP : **Rp.**

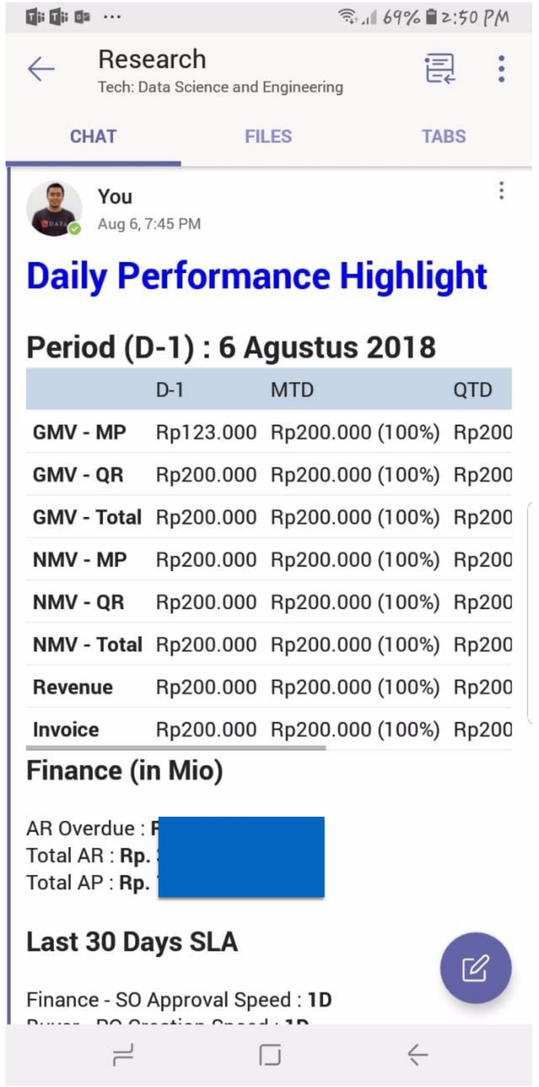
Last 30 Days SLA

Finance - SO Approval Speed : **1D**
Buyer - PO Creation Speed : **1D**
Vendor - PO Approval Speed : **1D**
Vendor - End To End Process Speed : **1D**
Finance - Invoice Delivery Speed : **1D**
Finance - Customer Payment Speed : **1D**

See less

Reply

Type a new message



Delivery Channel – Email Reporting

Summary Statistics

Daily Pending DO [redacted]

Total Pending DO
# of DO : 142
DO Value : Rp. [redacted]
# of SO : 103
# of SO Value : Rp. [redacted]

Total Pending DO ≥ 7 Days Jabodetabek
# of DO : 19
DO Value : Rp. [redacted]
# of SO : 19
# of SO Value : Rp. [redacted]

Total Pending DO < 7 Days
# of DO : 119
DO Value : Rp. [redacted]
# of SO : 82
# of SO Value : Rp. [redacted]

SO Pending Billing
# of DO : 100
DO Value : Rp. [redacted]

Total Pending DO ≥ 7 Days Outer Jabodetabek
# of DO : 4
DO Value : Rp. [redacted]
# of SO : 4
# of SO Value : Rp. [redacted]

SO Partially Fulfilled
# of DO : 42
DO Value : Rp. [redacted]

Total DO Collect

[redacted]

Shipping Area	Total DO Pending	Total DO Pending ≥ 7 Days	Total DO Pending < 7 Days	Total SO	Total SO Value
Jabodetabek	132	26	94	98	[redacted]
Outer Jabodetabek	31	16	49	23	
Total	163	42	143	121	

Customer Name	Total DO Pending	Total DO Pending ≥ 7 Days	Total DO Pending < 7 Days	Total SO	Total SO Value
189 PT. S[redacted]	18	2	16	7	[redacted]
5230 PT. [redacted]	10	1	9	5	
5599 PT [redacted]	9		9	7	
4239 PT. [redacted]	7		7	1	

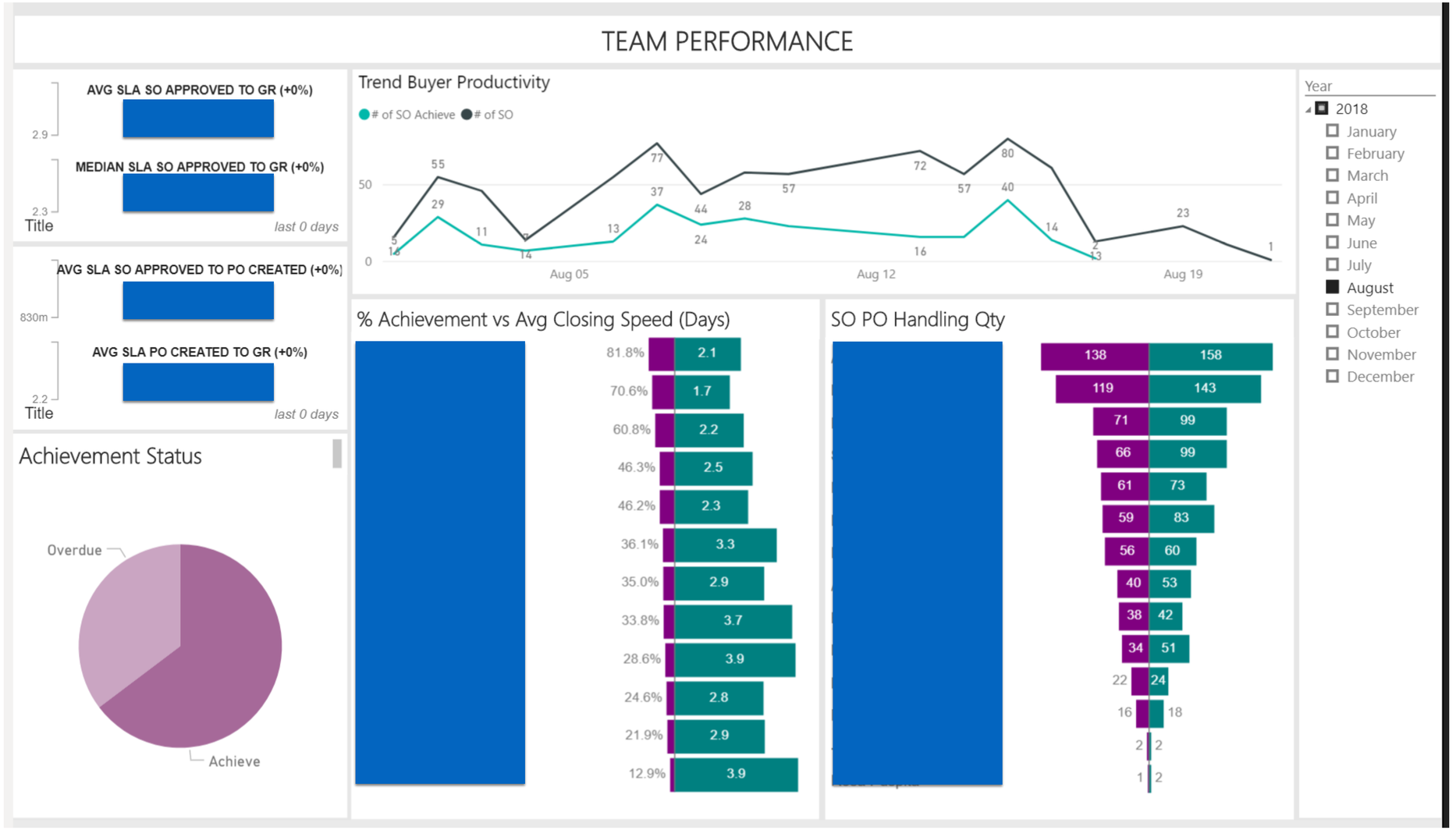
Detail Action Items

JABODETABEK

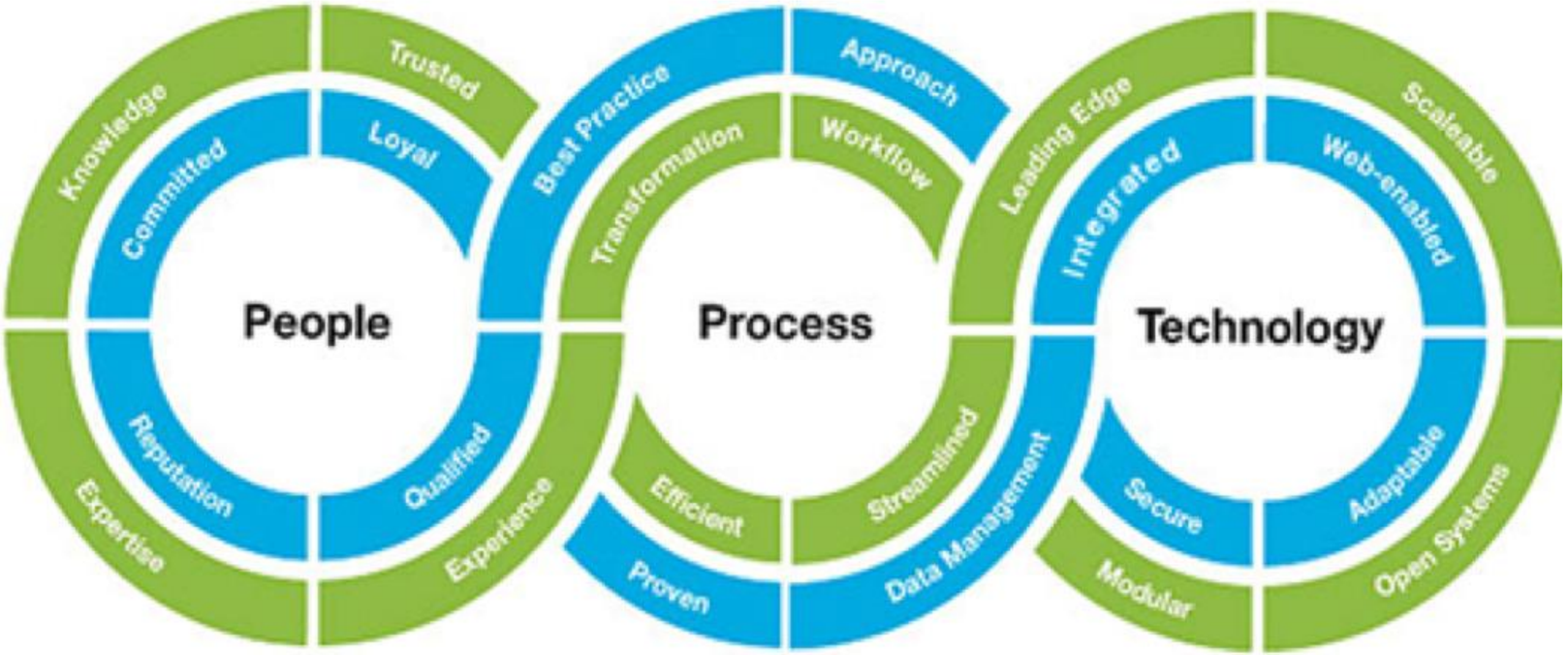
DO	Customer Name	Date Received	# Days Pending	SO	Status SO	Segment	Ship City	Billing City	Tanggal Terima DO dari Kurir	Tanggal serah terima DO ke Finance	Total Amt	Memo
13218	5444 PT. PABRIK [redacted]	2018-01-16	42	12443	Pending Billing	Offline Transaction	Jakarta Pusat	Jakarta Pusat			[redacted]	[redacted] No PO : MJK-45719286 Note : Mohon barang dapat dikirimkan segera ke Customer (URGENT) sebelum tgl 05 Maret 2018 - DROPSHIP ** Pengiriman di Lampirkan PO **
13517	[redacted]	2018-01-31	27	12461	Pending Billing/Partially Fulfilled	Offline Transaction	Jakarta Pusat	Jakarta Pusat				[redacted] 005/TVS/GEN/I/2018 Note : Mohon barang dapat dikirimkan segera ke Customer (URGENT) ** Pengiriman di Lampirkan PO **
13583	[redacted]	2018-02-01	26	12745	Pending Billing	Offline Transaction	Jakarta Pusat	Jakarta Pusat				[redacted] No PO : PRW-49763031 Note : Mohon barang dapat dikirimkan segera ke Customer (URGENT) sebelum tgl 03 April 2018 ** Pengiriman di Lampirkan PO - DO Asli harus kembali ke Jakarta untuk Penagihan **

Performance Trend

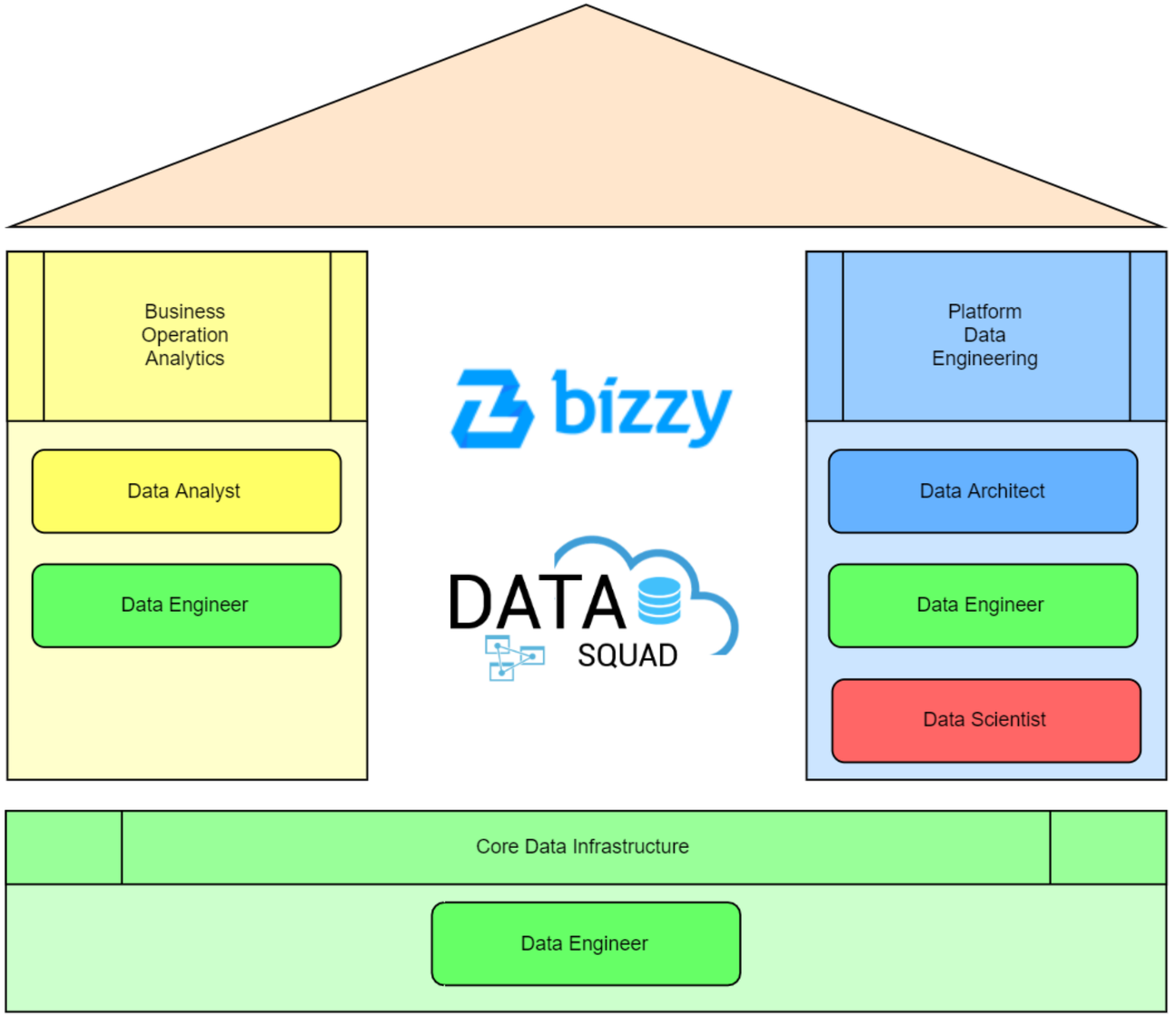
Delivery Channel – Dashboard



Framework for Business/Digital Transformation



People



bizzy

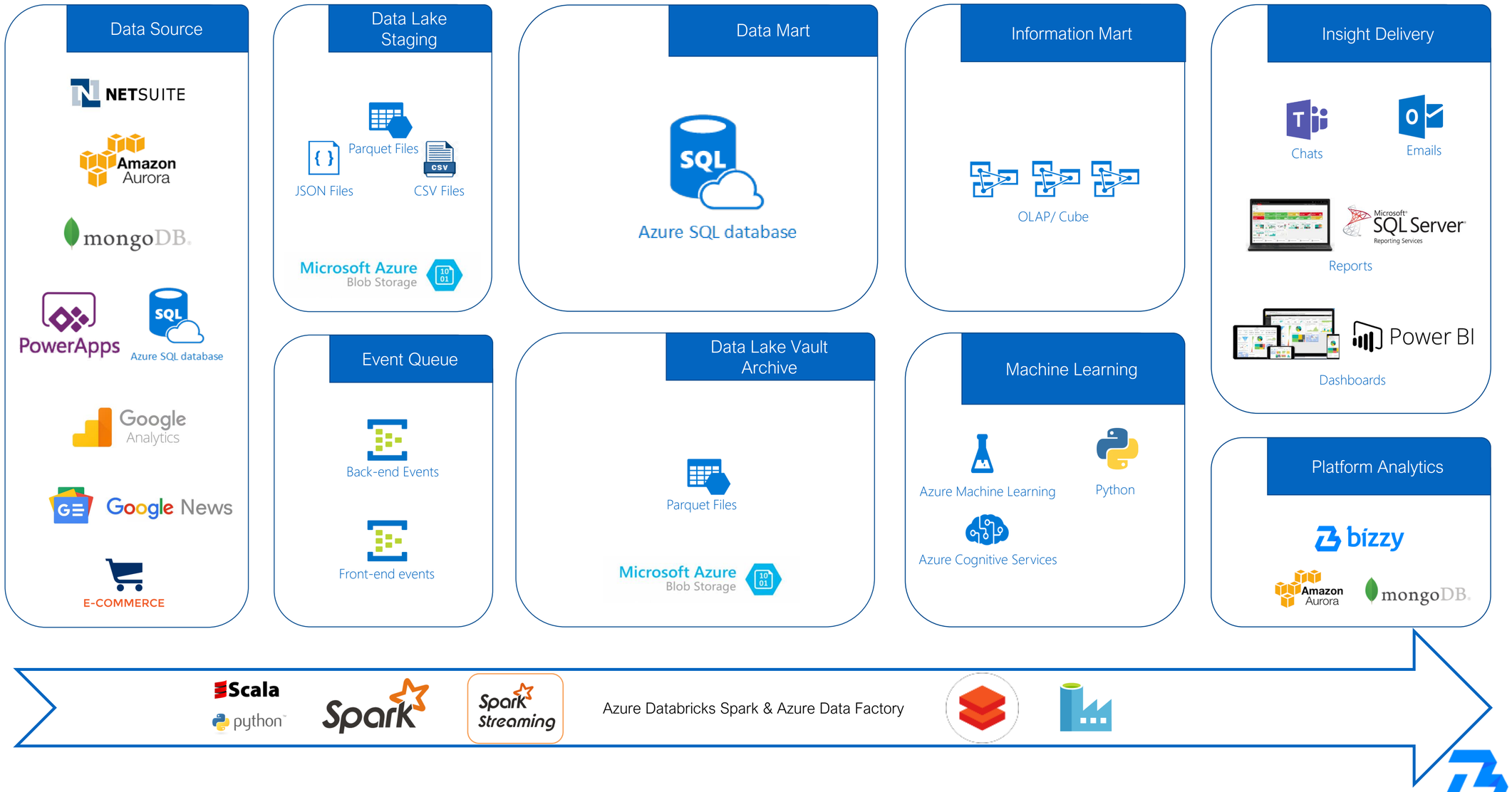
Process

- JIRA (Agile Methodology)
  - Scrum
  - Kanban

- Confluence (Wiki/Documentation)

- Bitbucket (Repository)

Technology – Data Architecture



bizzy

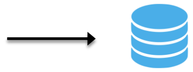
Various Data Pipelines & Technology

BATCH

Data Sources

Analytical Model DB (DSS)

Data Sources



Staging DB

Star Schema Data Warehouse

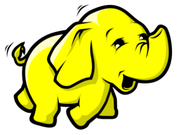
Data Sources

Staging DB

Data Vault (Relational EDW)

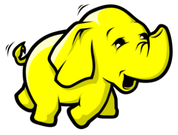
Star Schema Data Mart

Data Sources



Data Lake Big Data

Data Sources

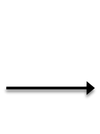


Data Lake Big Data

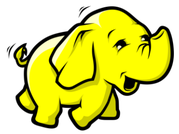
Star Schema Data Warehouse

STREAM

Application API



Event/Message Queue



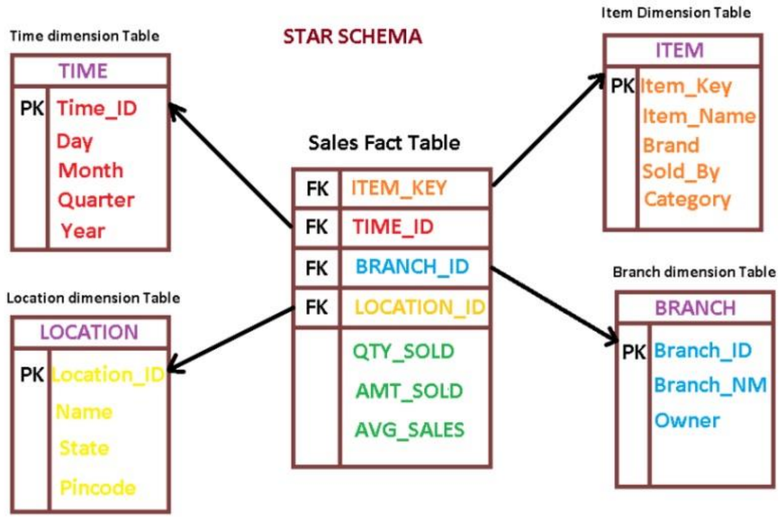
Data Lake Big Data

Data Warehouse

Real-time Dashboard

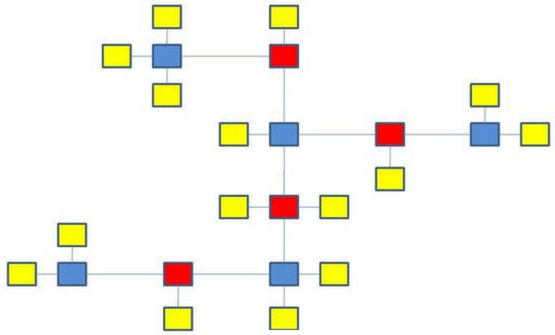
bizzy

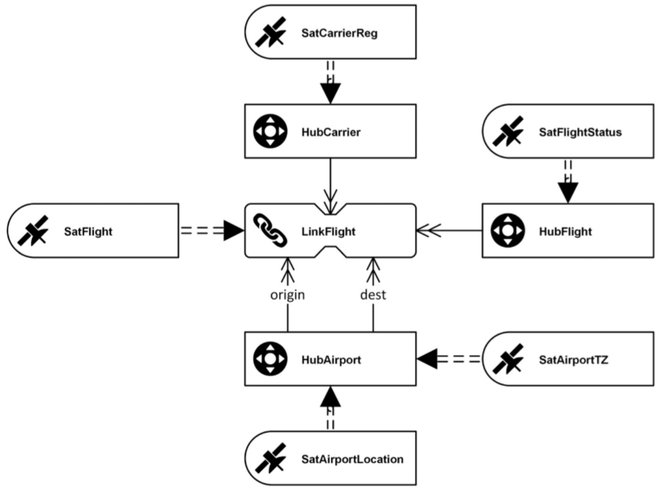
Data Warehouse – Star Schema



Data Warehouse – Data Vault

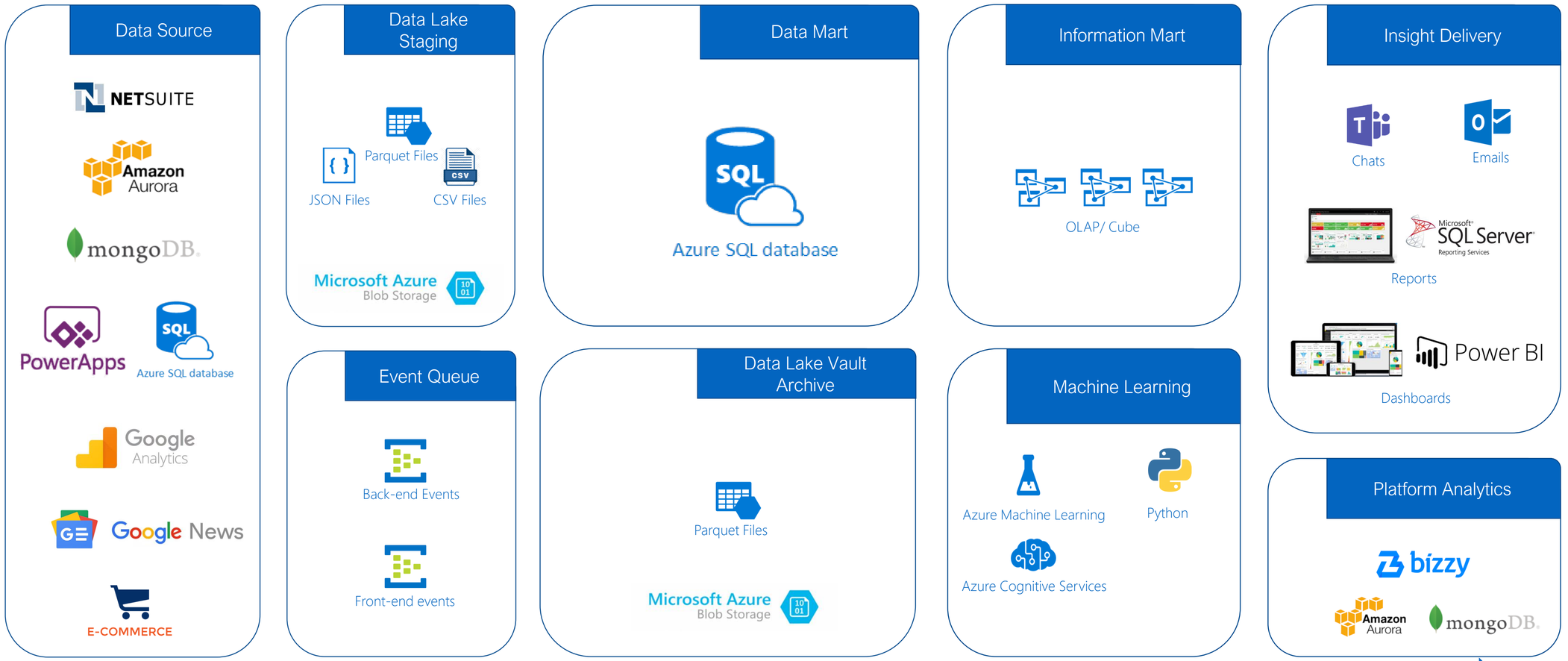
Data Vault – Hubs / Links / Satellites





bizzy

Technology – Data Architecture



bizzy

Big Data Infrastucture

Cloud

- Leasing cost model with easier chargeback representation.
- Can be cheaper for low uptime workloads(midnight ETL). Separation of compute and storage.
- Elasticity. For example, the ability to spin up huge numbers of instances to completely run a new release while running the old release at the same time.
- Software based managed services as an option.
- Integrates well with data sources already stored in the cloud.

- No control of hardware, and limited control of software – especially if you use a cloud vendor's distribution.
- Latency when interoping with on-premise resources. E.g. warehouse.
- Software level privacy, rather than hardware.
- Lock-in. This can be alleviated some if you use something like Cloudera, Hortonworks, or MapR.

On - Premise

- Typically cheaper for non-elastic workloads – at the moment.
- Since most Cloud providers use some type of network storage, performance is typically better for a bare-metal based deployment.
- Full control of the Hadoop hardware + software.
- Latency to and from systems you integrate with can be minimized.
- Physical data isolation and privacy.

- Infrastructure managed services. E.g. smart hands to install servers and network.
- Barrier to entry is typically higher since there is commonly “hardware friction" in getting new infrastructure on the floor and operating. (Private Clouds can help reduce “hardware friction" for proof of concepts and development/testing environments.)
- Chargeback models for Hadoop can be complex.

Spark ETL Generator – Source to Staging

```
val srcDb = "catalog"
val tgtDb = "stg_phx_catalog"
val tables = List("attribute_code","attribute_set","attribute_value","brand", "category", "category_temp", "misc", "uom",
"product_group", "product_group_attribute", "product_variant", "product_variant_migration", "product_vendor",
"product_vendor_migration", "stocking_uom")

import com.bizzy.dbrk.etl.JdbcEtlProcess

val notebookPath = dbutils.notebook.getContext.notebookPath.get
val jdbc = new JdbcEtlProcess(notebookPath, spark, jdbcUrl, connectionProperties)
jdbc.process(srcDb, tgtDb, tables, "overwrite", 3, 1000)
```


bizzy

Spark ETL Generator –Staging to Archive

```
val staging = "stg_phx_catalog"
val archive = "dlv_phx_catalog"
val tables  = List("brand", "category", "category_temp", "product_group", "product_variant", "product_vendor")

for(table <- tables ) {
  Archive.archive(s"$staging.$table", s"$archive.$table", s"$tmp.$table", id, date, "phx")
}
```


bizzy

THANKS!

QnA

PT Bizzy Commerce Indonesia
Dimo Space Building, 2nd Floor
Jl. Timor No 6, Menteng Jakarta Pusat, 10350

Bizzy.co.id
021-278 999 55
support@bizzy.co.id
Bizzy Indonesia
Bizzy-Indonesia
Bizzy.co.id/blog