RSAConference2019
San Francisco | March 4–8 | Moscone Center

SESSION ID: STR-W03

Access Control for Multi-Vendor Big Data and BI Environments

Anmol Singh

Lead Analyst
KuppingerCole Analysts AG

Big Data & BI Environments

An Introduction

Big Data

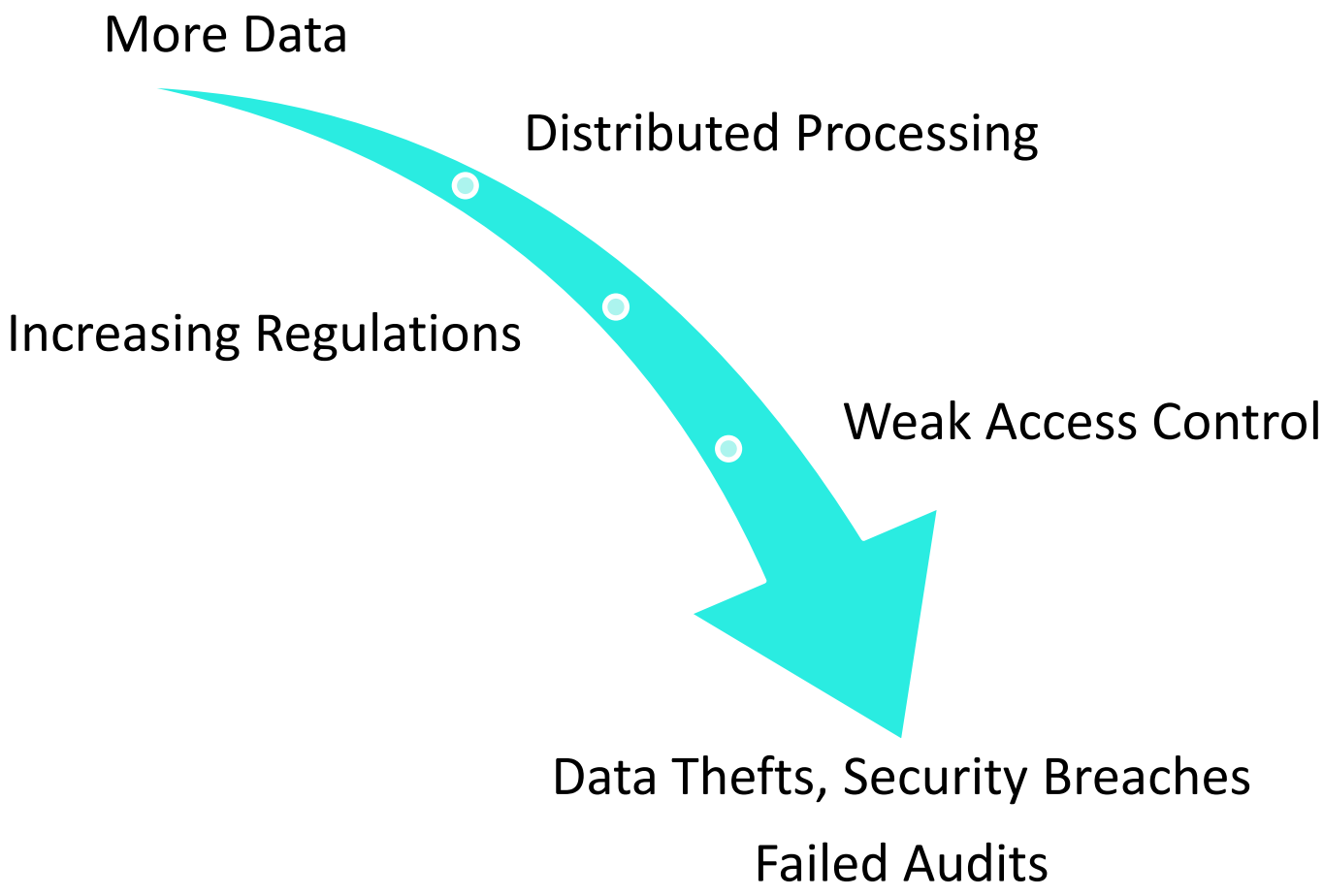
- ✓ Tons of Structured, Semi-unstructured and Unstructured Data
- ✓ Comprises of large and complex data sets that can't be processed by traditional database and software techniques

Business Intelligence

- ✓ Identify, extract and interpret business data using interactive tools for effective and accurate decision making
- ✓ Knowledge derived from discovering patterns and efficient data mining processes

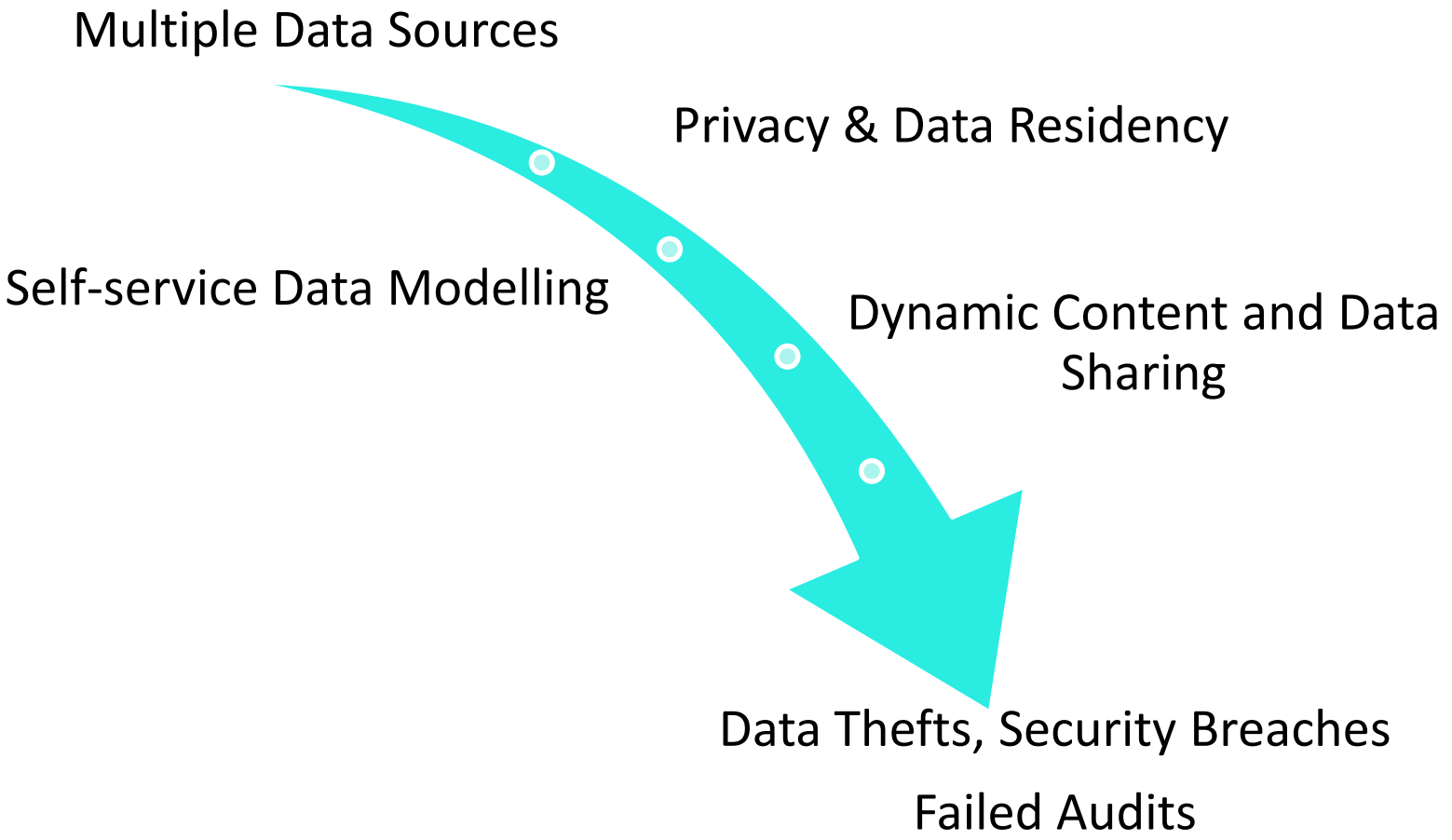
Big Data and Security

The evolving landscape..



BI and Security

The evolving landscape..



RSAConference2019

Security in Big Data & BI Environments

Key Challenges

Big Data & BI Security Challenges

Security remains an afterthought

- Security is not part of Design and Strategy
- Access is dependent on proprietary methods
- Existing IAM tools don’t support Big Data/ BI operations
- Access control for unstructured data is not a ‘thing’
- No data access governance for Big Data environments
- Big Data Security Solutions and Skills are a few and rare

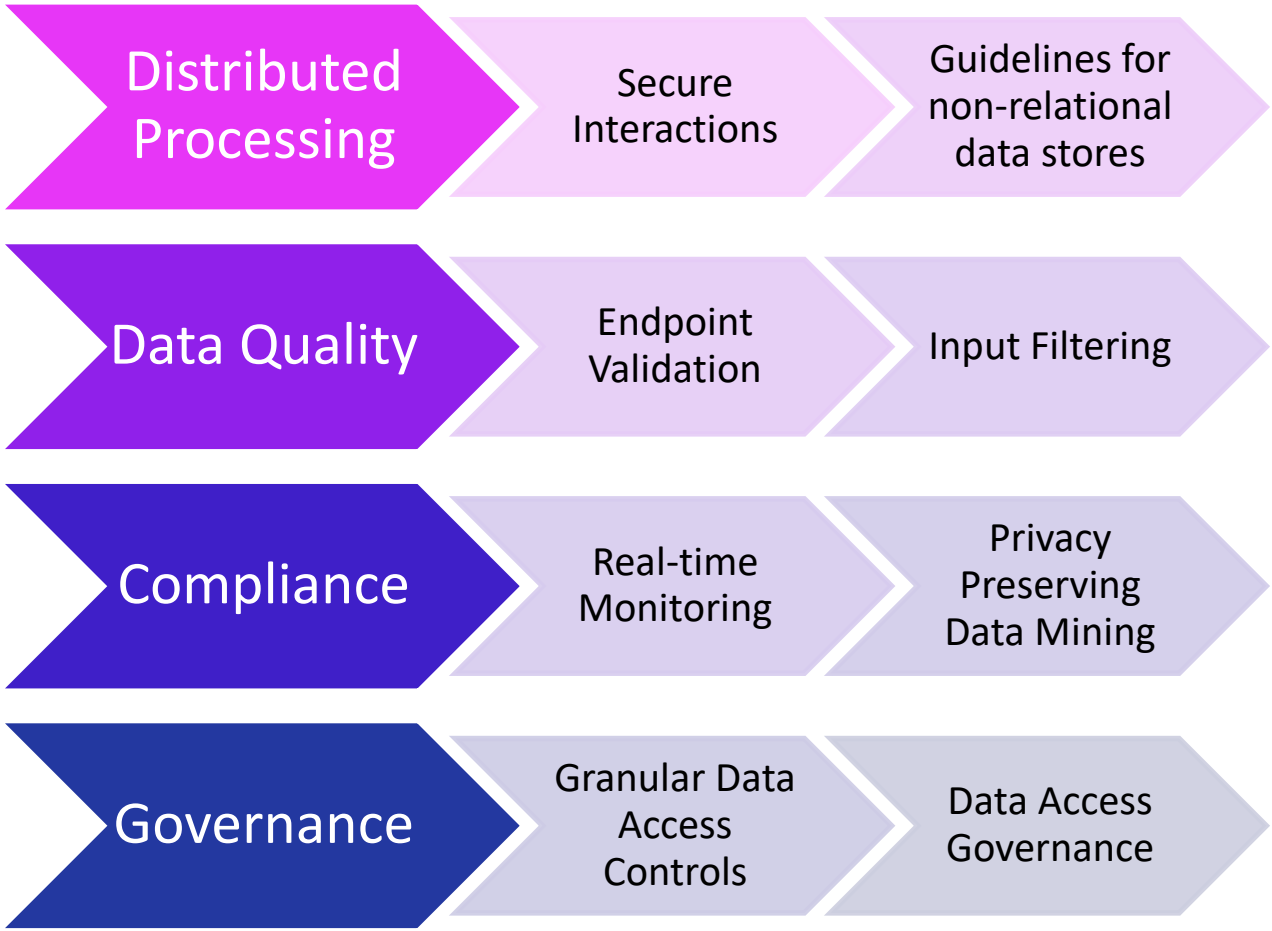
RSAConference2019

Security in Big Data & BI Environments

Primary Drivers

Primary Drivers for Big Data & BI Security

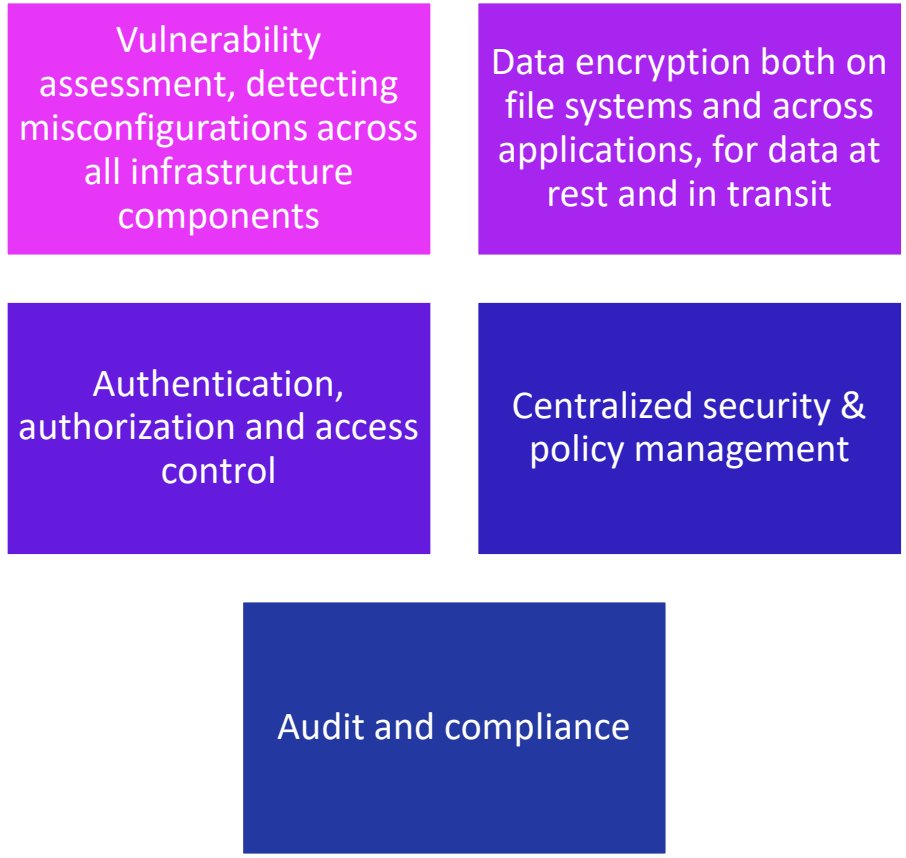
Demand a 'holistic' security approach



KuppingerCole‘s Defintion of Big Data Security

From the upcoming Leadership/Market Compass

- In our research for Big Data Security, we focus on products that follow a holistic approach towards protecting Big Data and BI platforms instead of the disconnected point security solutions. Although many generic security tools like firewalls or anti-malware may play an important role in securing parts of Big Data frameworks, we do not cover them as part of Big Data security to avoid the confusion of functionally distinct security products that are generally covered in other KuppingerCole’s reports under separate market segments.
- More precisely, we are not covering security solutions for protecting relational database management systems (RDBMS), since they are being reviewed in a separate Leadership Compass for Database Security.



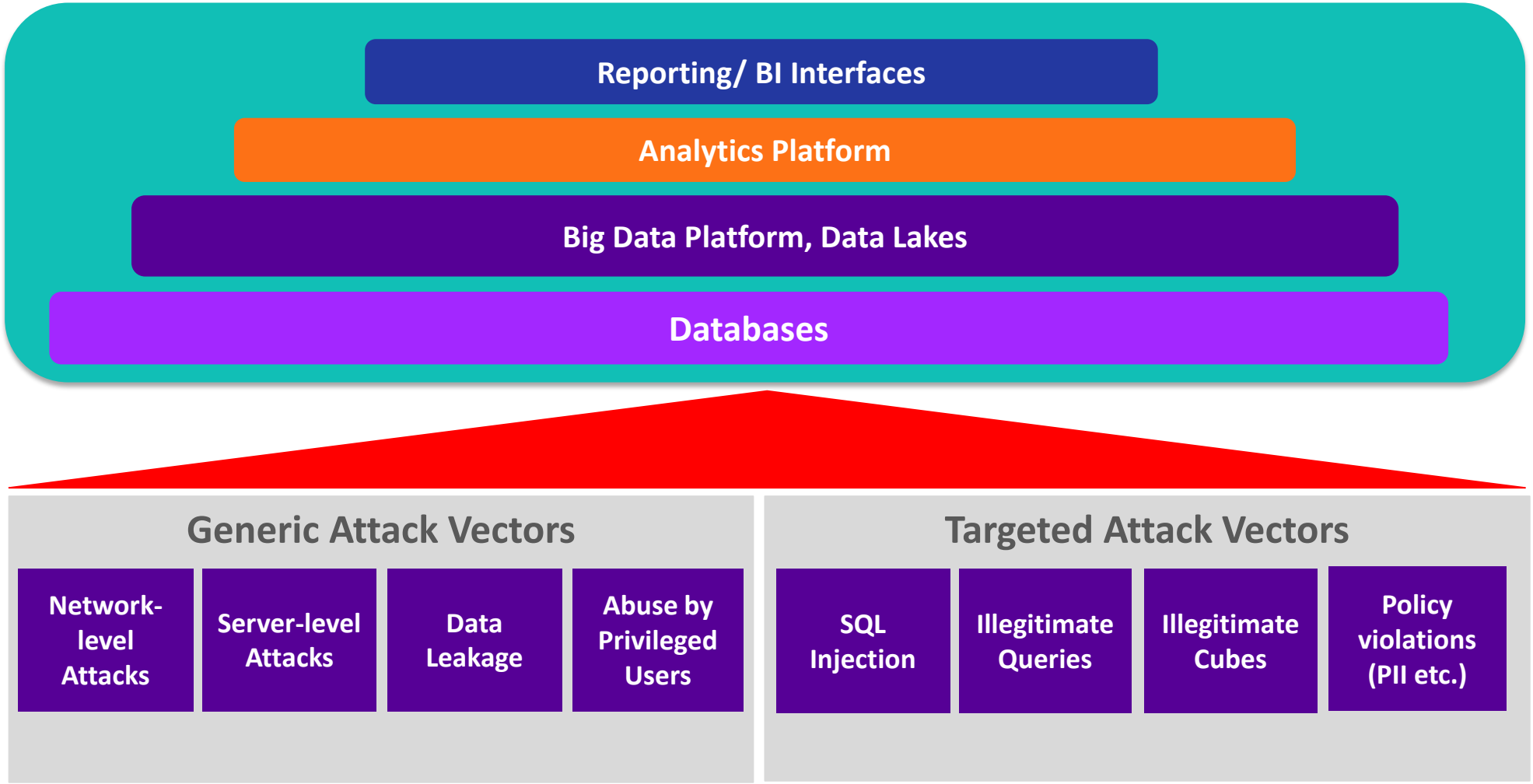
RSAConference2019

Security Threats in Big Data and BI Environments

Generic and Targeted Attack Vectors

Big Data & BI Security Risks & Threats

Multiple generic and targeted attack vectors



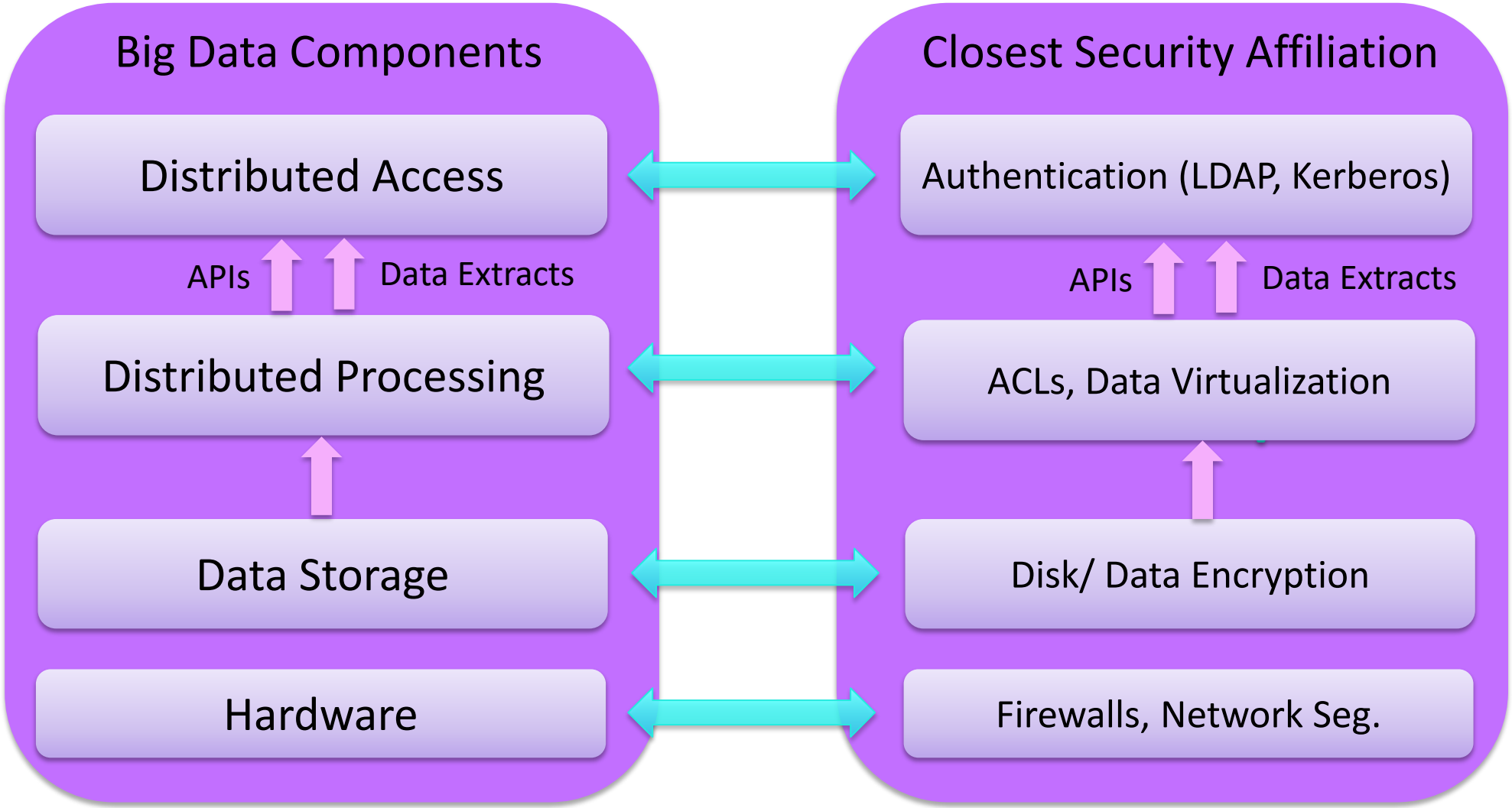
RSAConference2019

Big Data Security: Existing Technologies

Are these sufficient?

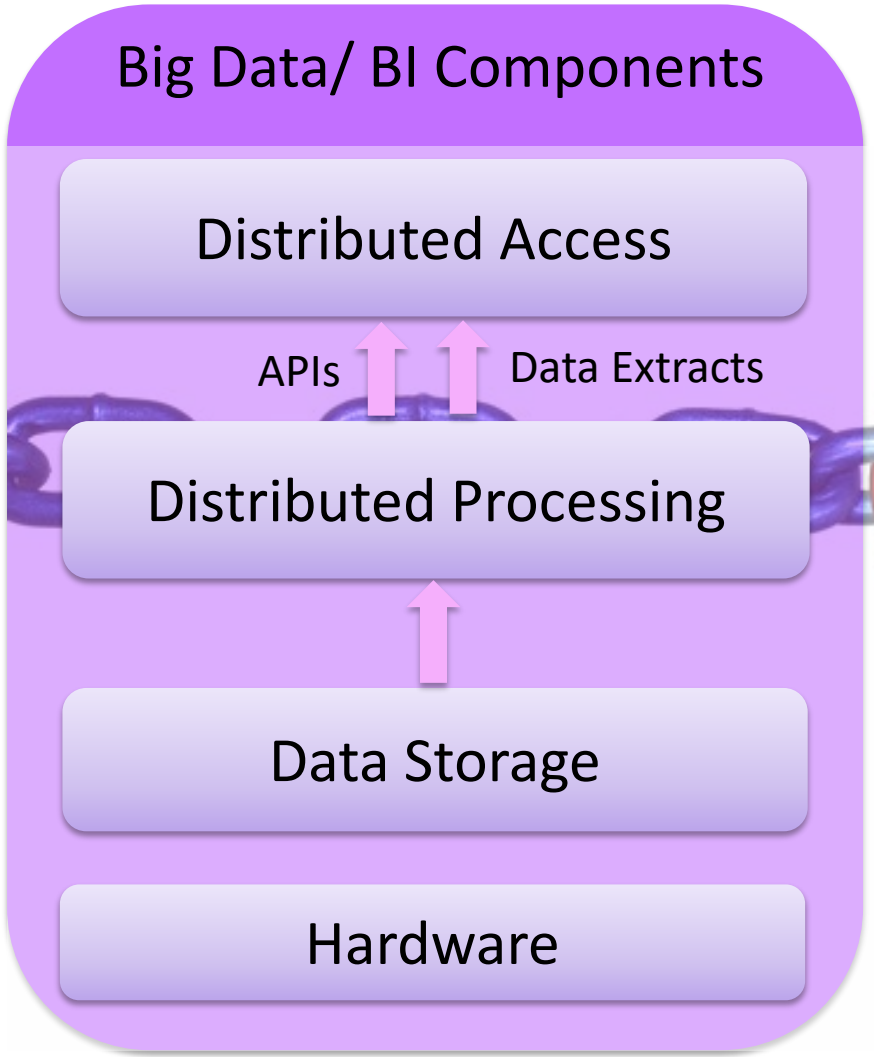
Current State of Big Data Security in Organizations

Reflect a state that is deficient in Security



Missing Security Links

Several links missing to a 'Secured Big Data Estate'

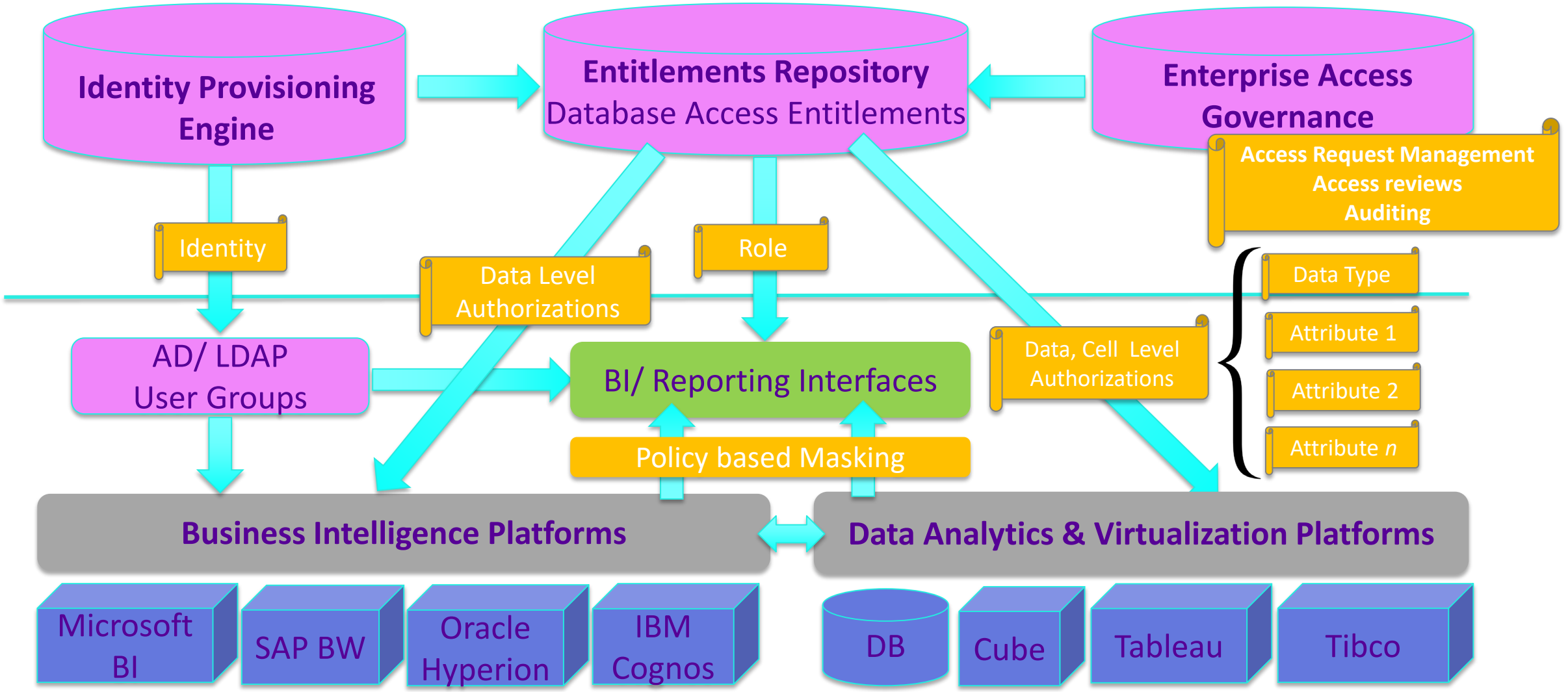


- ✘ *Contextual Authentication*
- ✘ *Granular Authorization*
- ✘ *Dynamic Authorization*
- ✘ *Service-to-service Authentication*

- ✘ *API Security*
- ✘ *Data Filtering/ Data validation*
- ✘ *Data Virtualization*
- ✘ *Data Masking*

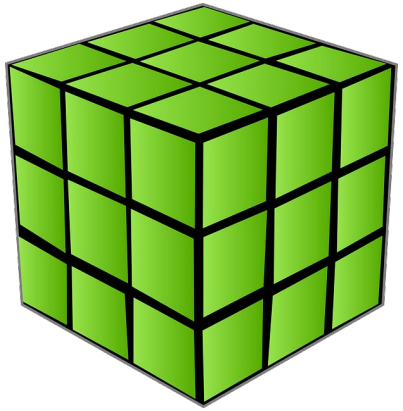
IAM for Big Data & BI Environments

Existing IAM tools do not support the complexity



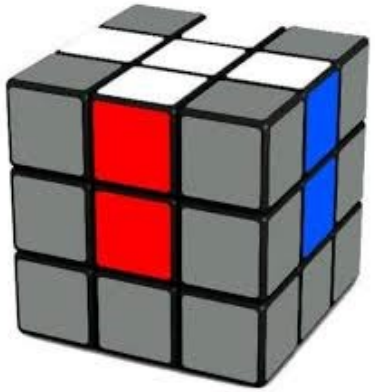
Access Governance in Big Data & BI Environments

The variety of data authorizations create complexity



Access at the Cube Level
- Data per source
- No further splits

Multi-dimensional access
- Several access combinations
- Granular role splits
- Multiple access restrictions

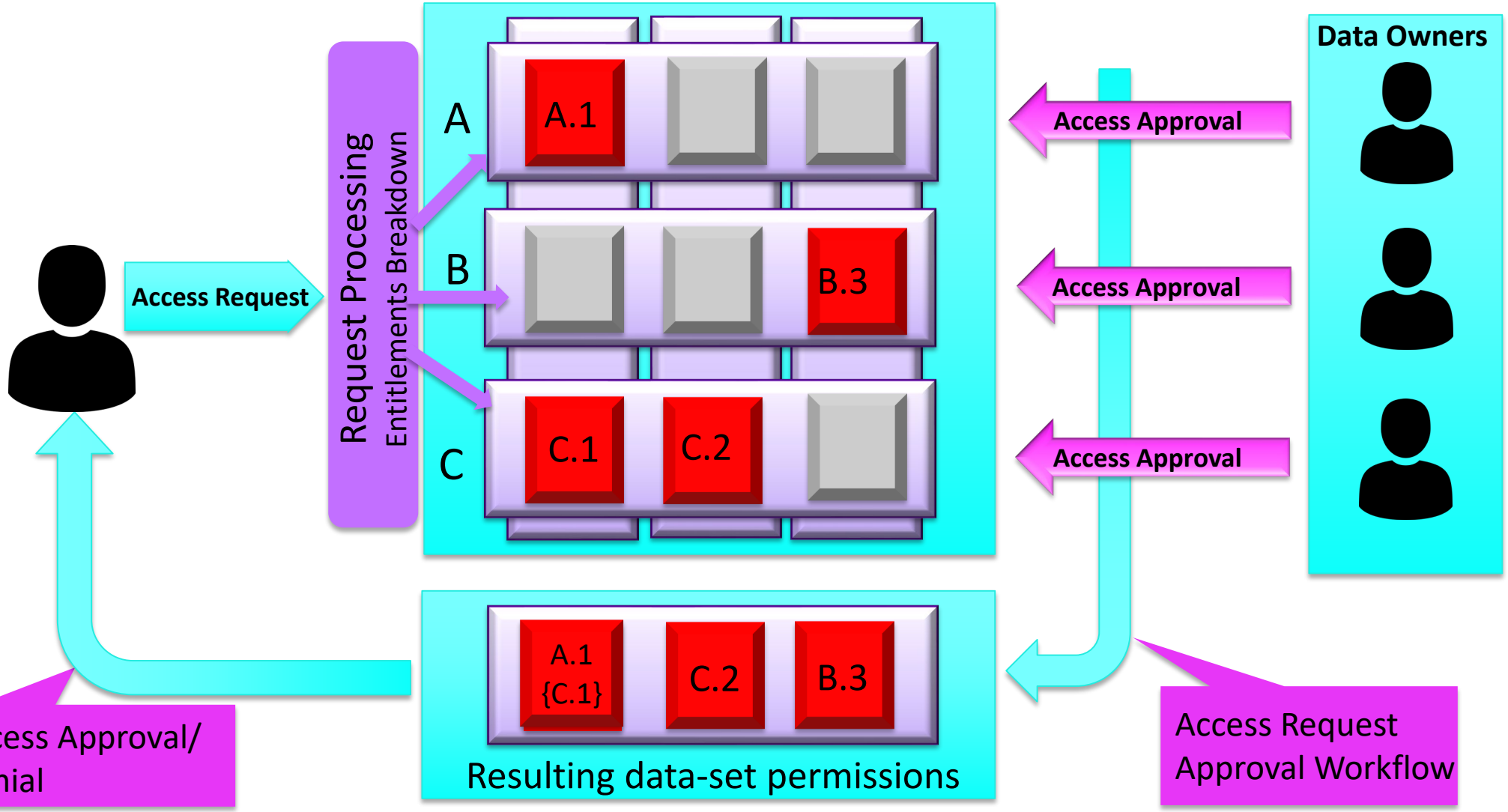


Access at the Cube splits
- One dimensional access
- Fewer access combinations

Disjointed Access Patterns
- Access across multiple data sets & providers
- Complex role combinations

Access Approvals: Granular Permissions

Governing access down to data-level permissions



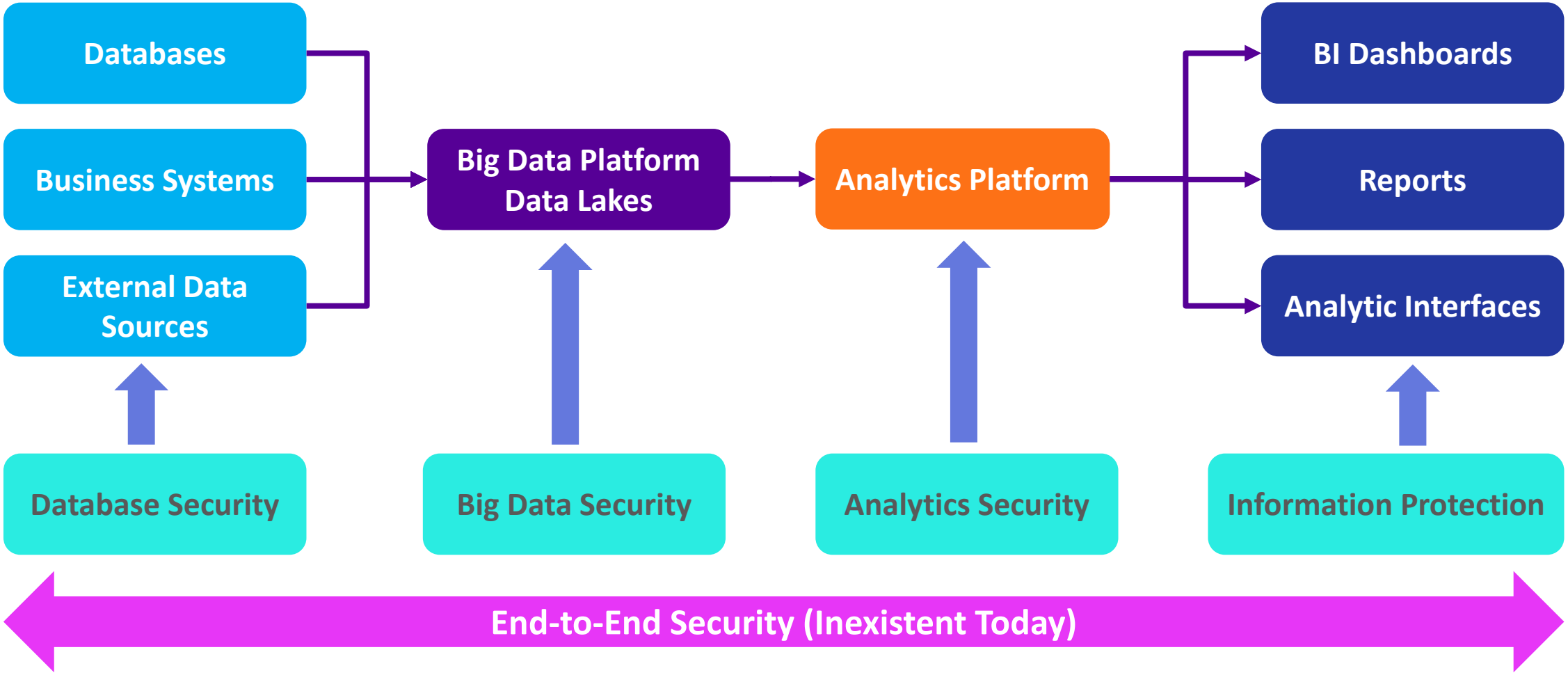
RSAConference2019

Security in Big Data & BI Environments

The BIG Question..

Security for Big Data & BI platforms

Is there a ‚holistic' security approach?



What are some of the tools in use today

Can quickly turn into a 'zoo' of technologies!

- ✓ Database Security Tools
- ✓ Data Discovery & Classification (for structured and unstructured data)
- ✓ Database & Data Encryption
- ✓ UBA (User Behaviour Analytics) for Data Access
- ✓ Data Masking & Tokenization
- ✓ Data Virtualization
- ✓ IGA (Identity Governance & Administration)
- ✓ PAM (Privileged Access Management)
- ✓ Dynamic Authorization Management
- ✓ DLP (Data Leakage Prevention)
- ✓ API (Application Programming Interface) Security

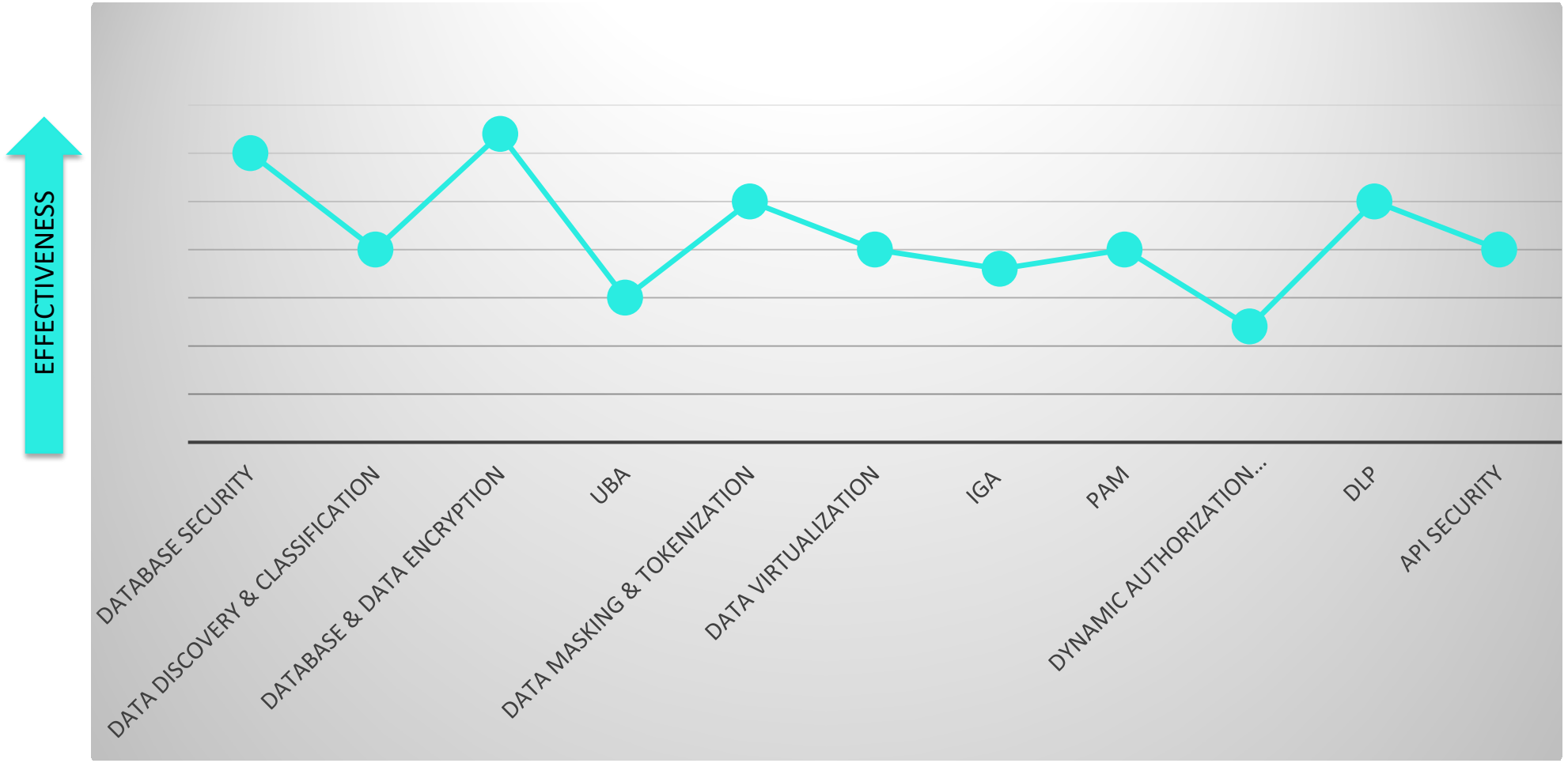
Limitations of existing security technologies

There's no *perfect* solution!

Technology	Limitations
Database Security	Commonly limited to RDBMS, not built for today's Big Data and BI/analytics
Data Discovery & Classification (for structured and unstructured data)	Only identifies the critical data, might require significant manual effort – helps to target protection but does not protect by itself
Database & Data Encryption	Encryption works at rest (and, in other form such as TLS, in motion), but not or only very limited for data in use, and it creates additional challenges for „use of data"
UBA (User Behavior Analytics) for Data Access	Helps in identifying critical use, but does not limit the access to data or the ability to combine certain sets of data
Data Masking & Tokenization	Potentially good protection also when it comes to exporting and recombining data, but applications might need access to full set of data
Data Virtualization	An efficient approach from a data protection perspective, but can create massive amounts of transient (insecure) information views , affects performance
IGA (Identity Governance & Administration)	Relatively few out-of-the-box connectors for managing users and, in particular, fine-grain access entitlements in these environments. Might require massive customization and suffer from complexity due to complex entitlement structures of multi-level/multi-dimensional data models
PAM (Privileged Access Management)	Focused on securing administrative access, not the fine-grained access control for business users
Dynamic Authorization Management	Very few out-of-the-box solutions, very limited support for environments, currently only a point solution. Potential performance impact
DLP (Data Leakage Prevention)	Focus on files, i.e. the results, not their creation
API Security	Limited to APIs access only

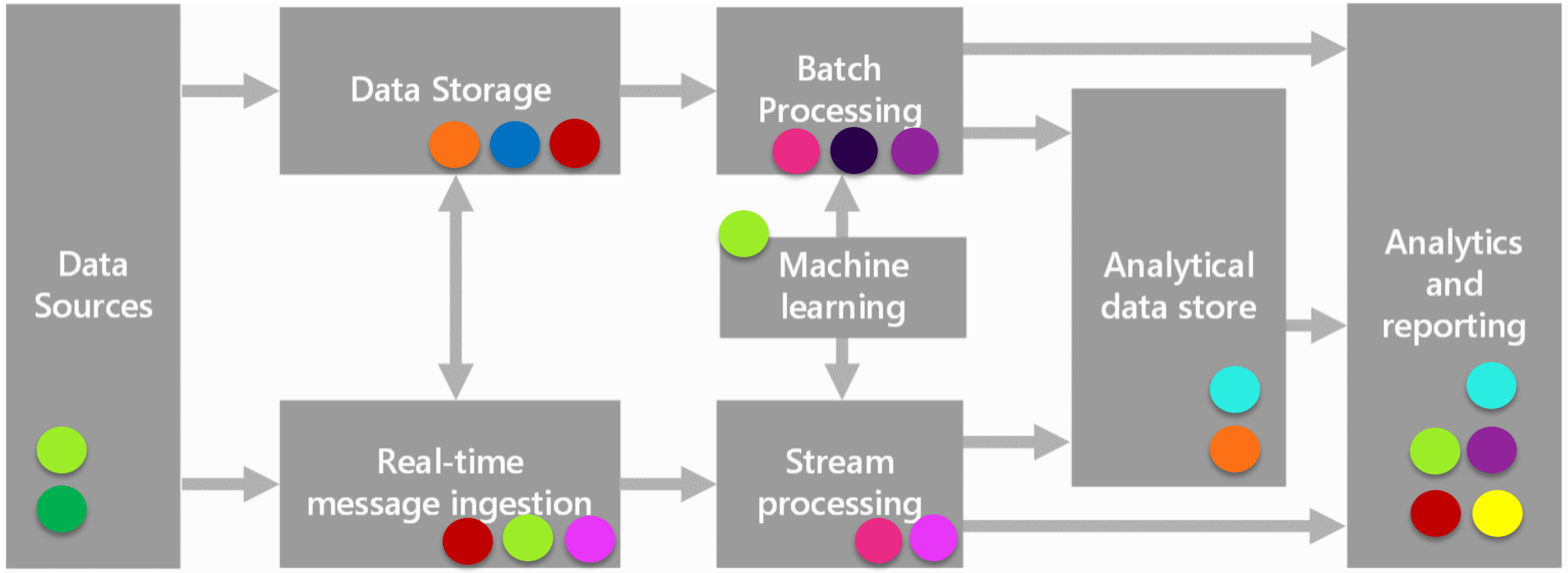
Efficacy of existing technologies to Big Data Security

Few technologies are more effective than others



Where do they fit in Big Data & BI?

How to map security technologies to Big Data/ BI Components



- Database Security
- Data Discovery & Classification
- Database & Data Encryption
- DLP (Data Leakage Prevention)
- User Behaviour Analytics
- Data Masking & Tokenization
- Data Virtualization
- API Security
- IGA
- PAM
- Dynamic Authorization Management

RSAConference2019

Security in Big Data & BI Environments

Priority Framework

Key Questions on Big Data Security & Governance

How to identify your priorities?

Question	Database Security Tools	Data Discovery & Classification	Database & Data Encryption	User Behaviour Analytics	Data Masking & Tokenization	Data Virtualization	Identity Governance & Administration	Privileged Access Management	Dynamic Authorization Management	Data Leakage Prevention	API Security
Is your data adequately protected during storage and distributed processing?	2	1	3	0	0	0	0	1	0	1	1
Is governance and security consistently enforced across the entire Big Data ecosystem, from source to the target?	2	1	3	0	1	2	2	2	1	2	1
Is there adequate insight and governance over data combinations?	1	0	2	0	3	2	3	1	2	2	0

Key Questions on Big Data Security & Governance

How to identify your priorities?

Contd...1

Question	Database Security Tools	Data Discovery & Classification	Database & Data Encryption	User Behaviour Analytics	Data Masking & Tokenization	Data Virtualization	Identity Governance & Administration	Privileged Access Management	Dynamic Authorization Management	Data Leakage Prevention	API Security
Do you know where sensitive data such as PII and credit card data resides?	1	3	2	0	2	1	0	0	0	1	0
Is there a centralized solution for managing and protecting that data?	3	2	2	0	1	1	1	2	1	2	1
Is your current approach for data protection performing well for the BI use cases?	2	1	2	1	3	2	2	2	1	2	1

Key Questions on Big Data Security & Governance

How to identify your priorities? *Contd...2*

Score	*Priority*	*Time for action*
0-3	Low	No urgency
3-5	Medium	1 to 2 years
>5	High	Next 3-6 months

RSAConference2019

Recommendations

Plan to succeed with Big Data & BI Security

Recommendations

Plan to succeed with Big Data & BI Security

✓ Implement Data Discovery and Classification

- Establish an enterprise data catalogue and keep it updated
- Identify and classify data: Harvest and maintain metadata

✓ Make Data Sanitization & Encryption a Continuous Process

- Enterprise Information Protection: Classify and encrypt the documents you create
- Network Security: Protect data transfer, Define micro segments of the sensitive areas in your network
- Infrastructure & Server Security: Server Hardening
- Encryption: Encrypt data in transit and at rest, Use a KMS (preferably HSM)
- Database Design: Design databases with confidentiality in mind (Separate fields for easy filtering and encryption)

Recommendations

Plan to succeed with Big Data & BI Security *Contd..*

✓ Implement Access Control and Dynamic Authorization

- Understand data flows and authorization requirements of your Big Data & BI environments
- Implement fine grained access controls: Define authorizations at file, service and data levels to implement an ABAC Model
- Implement policy based dynamic masking and row filtering
- Use PAM controls to prevent rouge administrative access to sensitive data
- Use Application to Application Password Management (AAPM) for A2DB (Application to DB) authentication

Recommendations

Plan to succeed with Big Data & BI Security

- ✓ Enforce Monitoring Controls: Track user access details for activity reviewing, logging and auditing purposes
- ✓ Implement Access Governance: Conduct regular and periodic data access certifications
- ✓ Implement API security and input validation: Use API Gateways and device authentication
- ✓ Consider a ‘holistic approach’ - traditional security controls only address parts of Big Data Security
- ✓ Use the matrix discussed to assess your Big Data security state and prioritize your technology investments

Related KuppingerCole Research

Where to find more relevant research on the topic?

- ✓ Leadership Compass: Database Security - 70970
- ✓ Advisory Note: Big Data Security, Governance, Stewardship - 72565
- ✓ KuppingerCole and BARC Joint Study: Big Data and Information Security - 74001
- ✓ Advisory Note: Enterprise Big Data IAM – Challenges and Opportunities - 71207

RSAConference2019

Thank you