End-to-End Security in Vertica

Fenic Fawkes (fae/faer) - Security Software Engineer

Chris Morris (he/him) - Security Software Engineer

MICRO FOCUS

A Secure Foundation

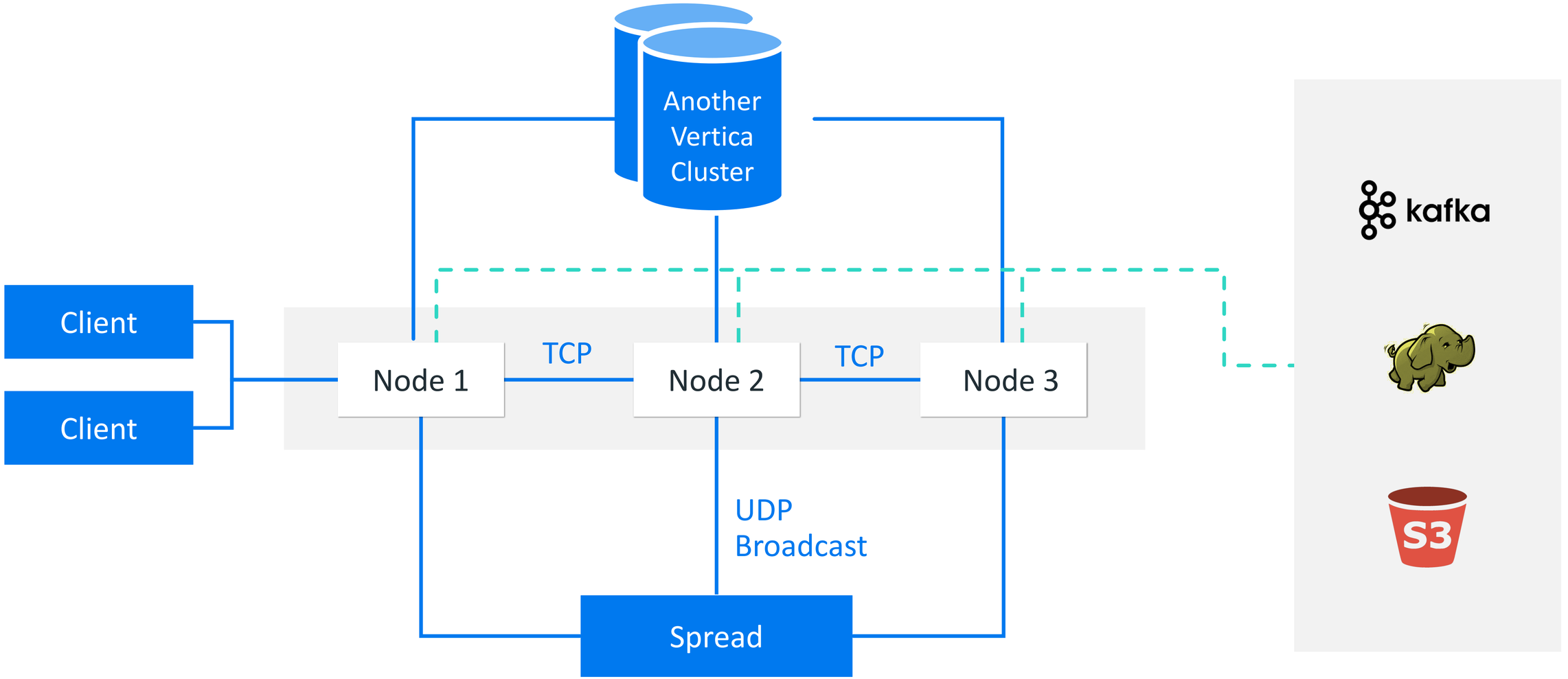
- Vertica has a ton of different security features
- Each feature answers a **key question**
- Use case + threat model determines security needs
- We'll discuss the building blocks of security in Vertica
- When crafting a secure setup, pick, assemble, and tweak the components according to your requirements

Overview

Topic	Question
Encryption	How do I protect my data at rest?
	How do I protect my data in transit?
Authentication	How do I prove who I am?
Identity	Who and what am I?
Authorization	What can I do?
Delegation/Impersonation	When talking to other systems, who is Vertica?
Auditing/Monitoring	What's happening in my database?

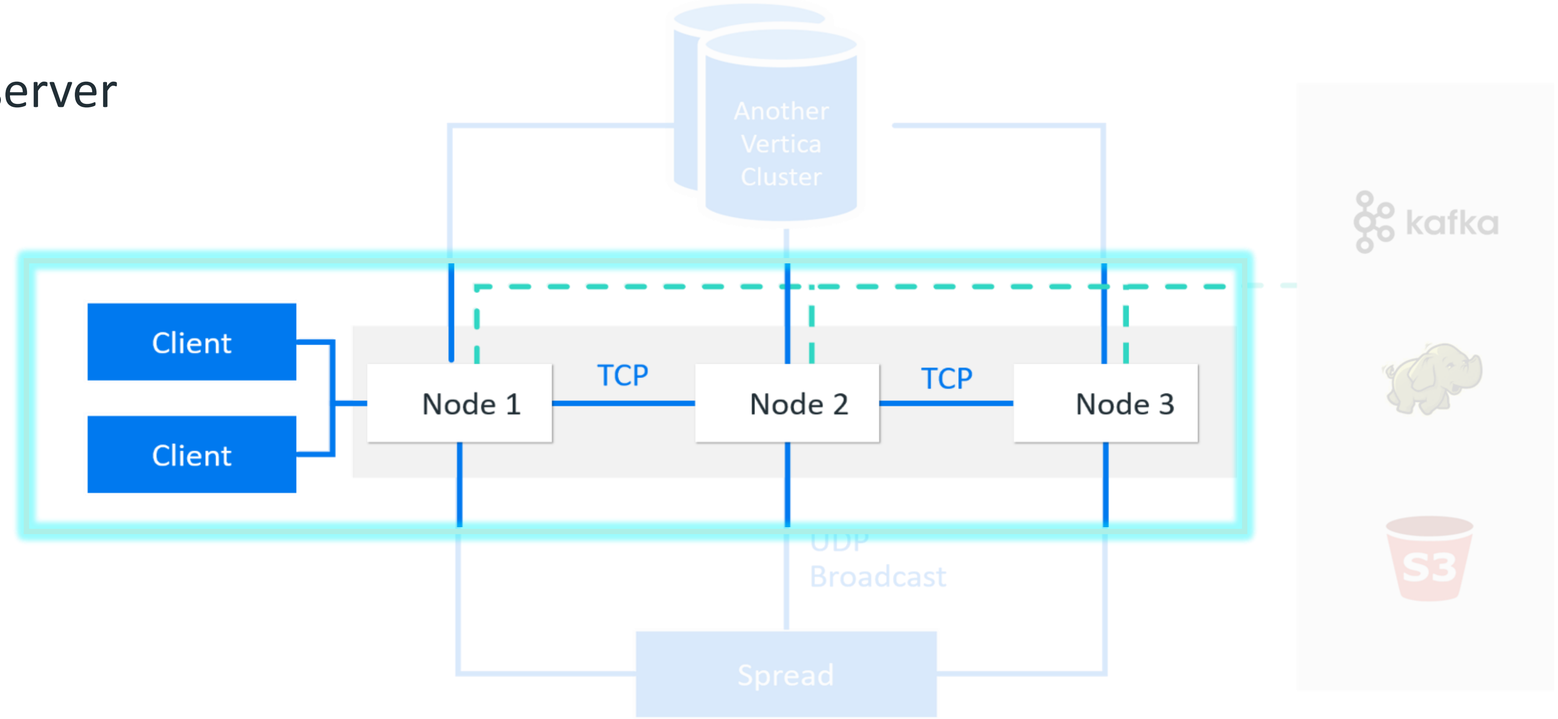
How Do I Protect My Data… In Transit?

What are the Network Connections That Need to be Secured?



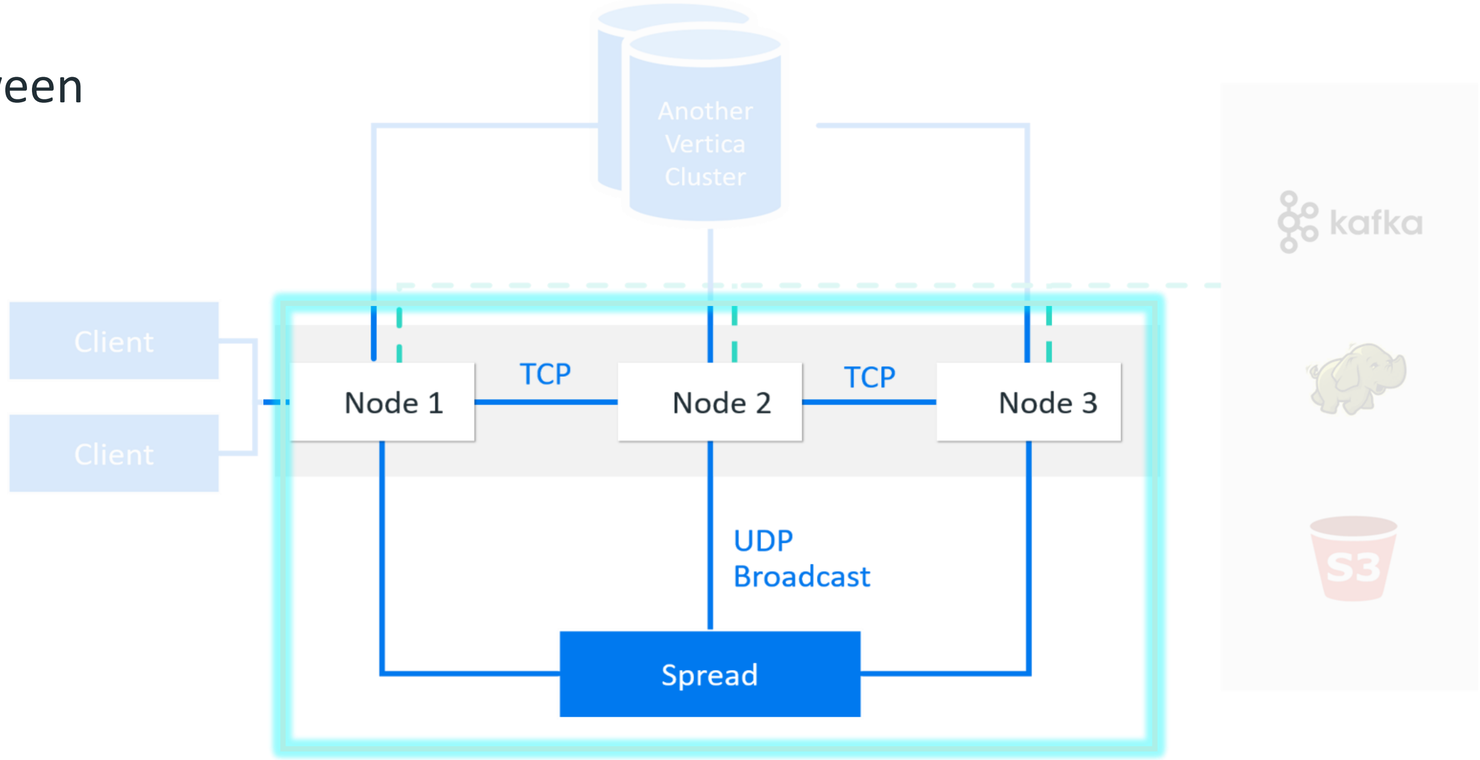
Client-Server Transport Layer Security (TLS)

- Set SSLCerticate and SSLPrivateKey configuration parameters
  - Takes effect for all new connections
  - Gives valuable warning messages
- Clients should have CA to check server certificate against
- Client SSLMode
- Authentication methods can disable non-TLS connections



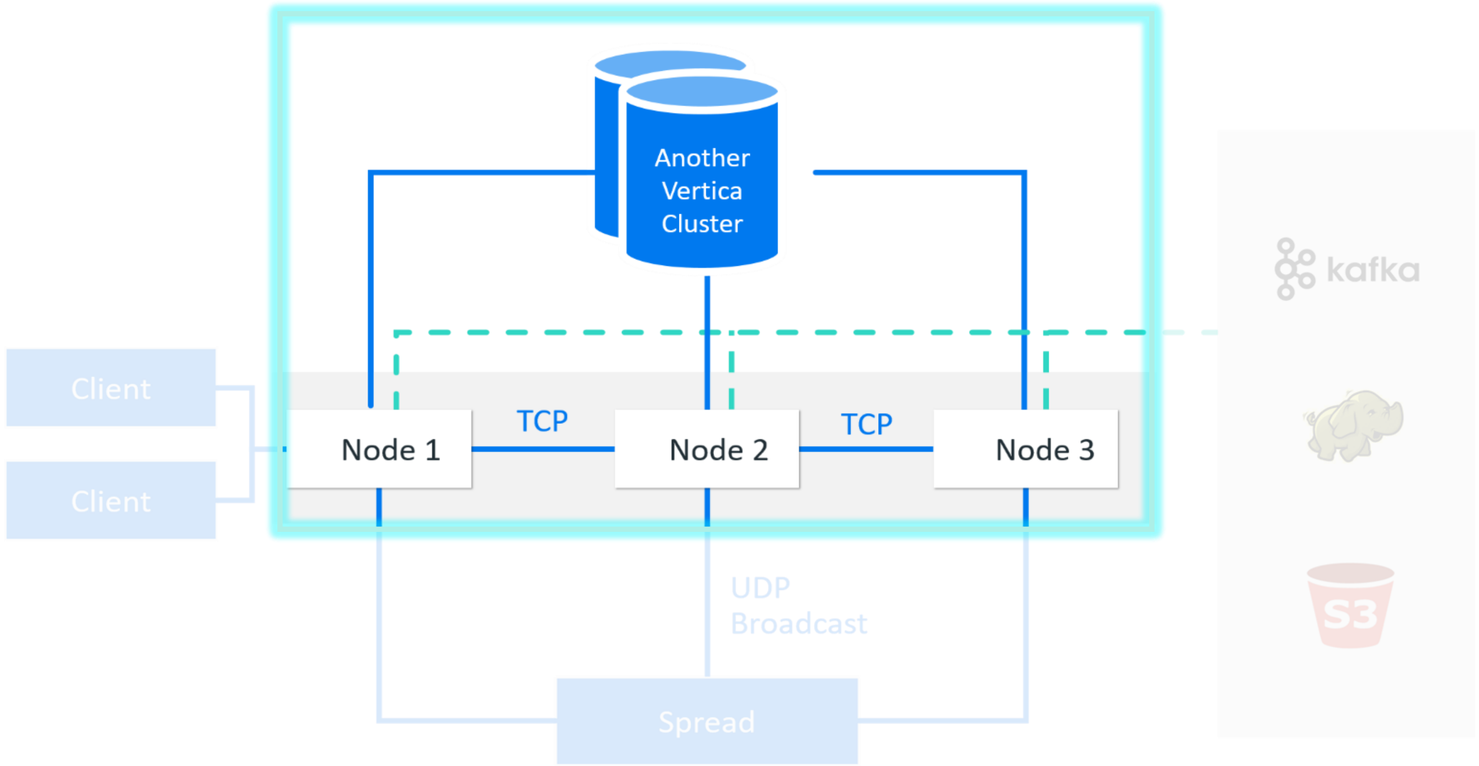
Intra-Cluster Encryption

- Less important if Vertica is running behind a strict firewall
- Very important on the public cloud and similar environments
- Encrypts all data and metadata between nodes and across spread



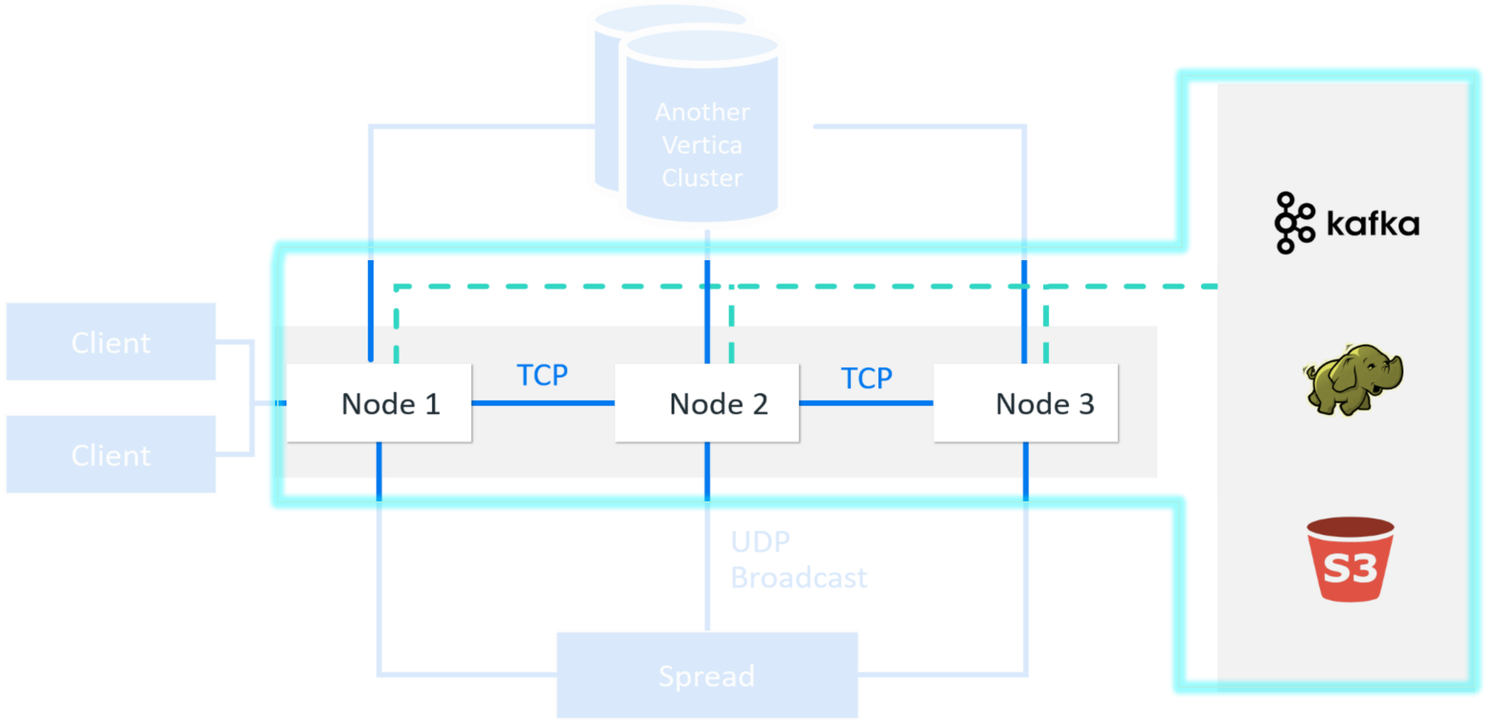
Import/Export

- Encryption between clusters just as important as within a cluster
- Both clusters need to have the same DataSSLParams CA certificate
- Cluster being connected to needs to have SSLCertificate set
- Password and Hash auth are supported



External Services

- Many of them
  - Kafka
  - Hadoop
  - S3
  - …
- How to configure differs between services
- See docs!



Network Security Roadmap

- Making Vertica Security Easier to Use
- Centralization and Simplification
- Improving how certificates and keys are managed
- Automating secure setup

How Do I Protect My Data… At Rest?

Lots of encryption choices

Different encryption approaches, different use cases:

Encryption type	Benefits
RAID controllers	Protects against media theft, hardware, invisible to Vertica
Disk encryption	Protects against media theft, software, invisible to Vertica
S3 encrypted buckets	Protects data in S3 object storage
GCP encryption	Protects data in Google Cloud Storage (incl. from Google!)
Field-level encryption	**All of the above, plus granular access control to specific values**

Lots of encryption choices

Different encryption approaches, different use cases:

Encryption type	Benefits
RAID controllers	Protects against media theft, hardware, invisible to Vertica
Disk encryption	Protects against media theft, software, invisible to Vertica
S3 encrypted buckets	Protects data in S3 object storage
GCP encryption	Protects data in Google Cloud Storage (incl. from Google!)
Field-level encryption	**All of the above, plus granular access control to specific values**

Voltage SecureData | FF1 Format-Preserving Encryption and Stateless Key Management

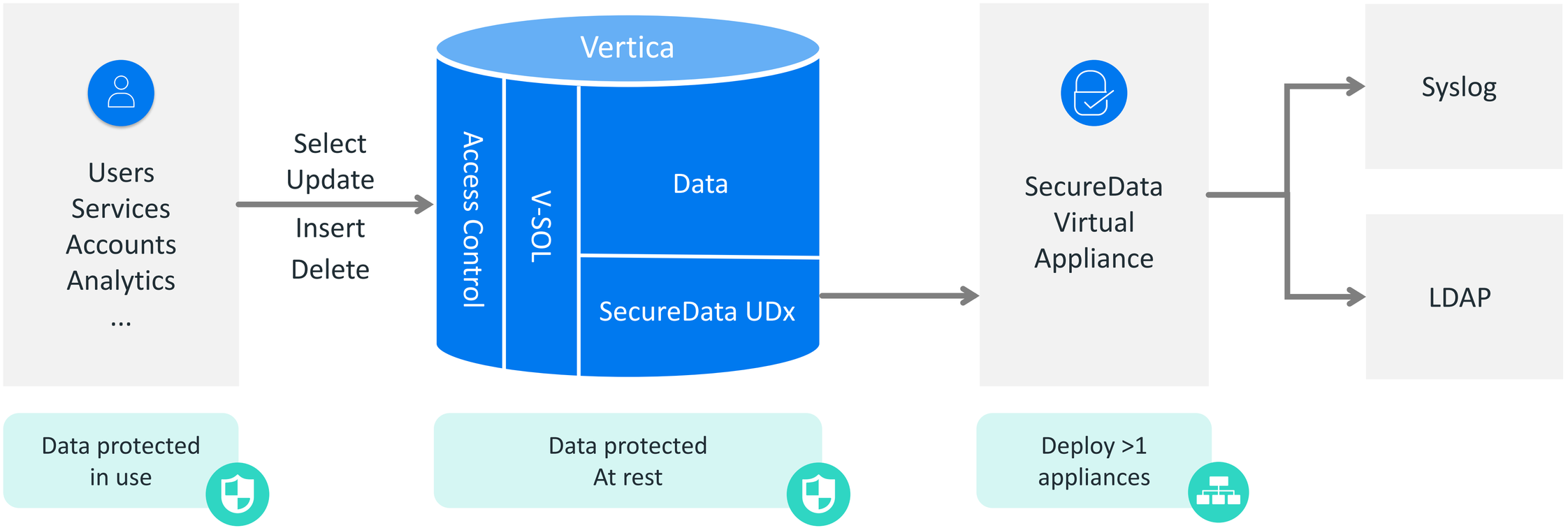
Voltage SecureData: Field-Level Encryption

Format-preserving encryption

Data element	Example data in the clear	Example protected data
Name	Härold Potter	5iW9VtS4ØlwpQ
Tax ID	532-09-1847	821-90-7385
Credit Card Number	4210-9735-8310-**4461**	9328-0218-7219-**4461**
Date of Birth	09/04/1979	05/01/1998
GPS location	37° 46' 26.2992'' N 122° 25' 52.6692'' W	91° 52' 05.7217'' N 731° 60' 21.6540'' W

Vertica + Voltage Integration

- SecureData UDx for Vertica integrates encryption layer with normal DB operations
- Vertica can flexibly store both protected and clear text data



Learn More!

Check out the Virtual BDC Talk!

Encryption at Rest Roadmap

- Built-in encryption at rest for Vertica
- Integrated with AWS KMS or standalone
- We’d love to hear your use cases!


VERTICA
Access Control
Authentication, Identity, Authorization
MICRO FOCUS

How Do I Prove Who I Am?

How Can I Log In?

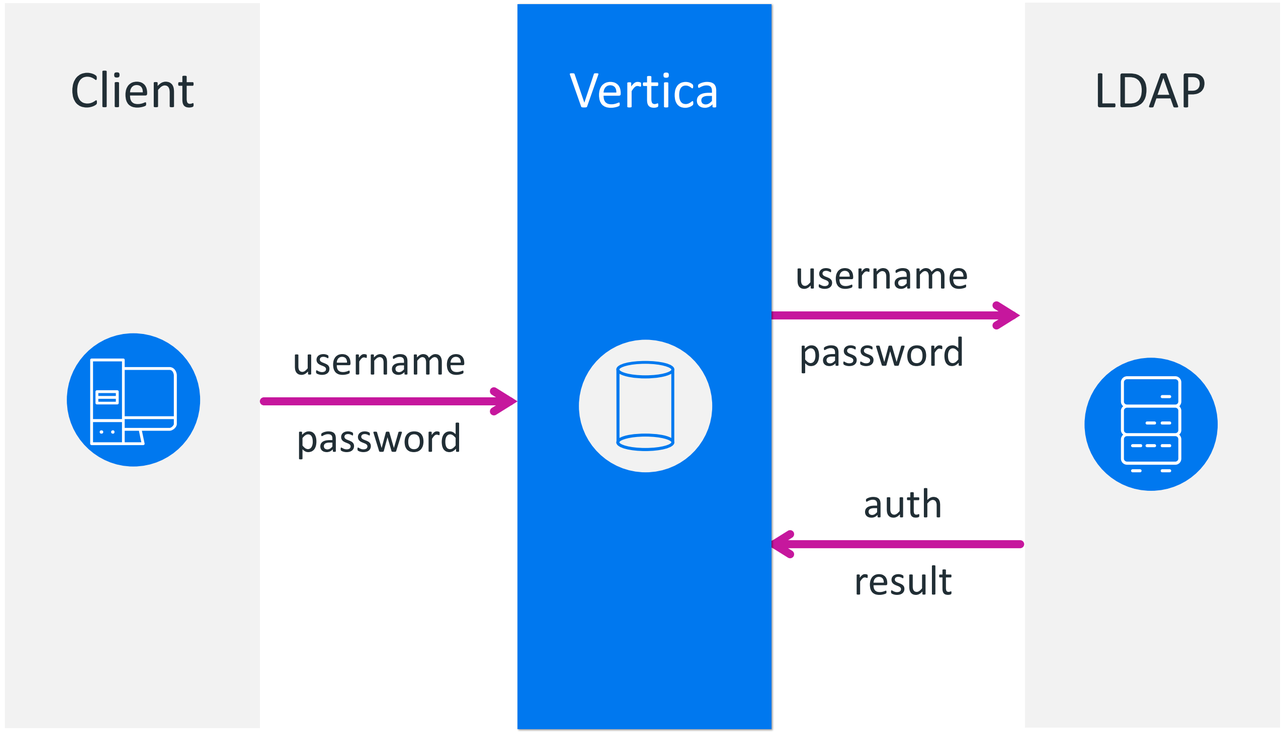
- Several authentication methods
  - Which one is best depends on deployment size/use case
- Order by priority & origin

Built-In Authentication Methods

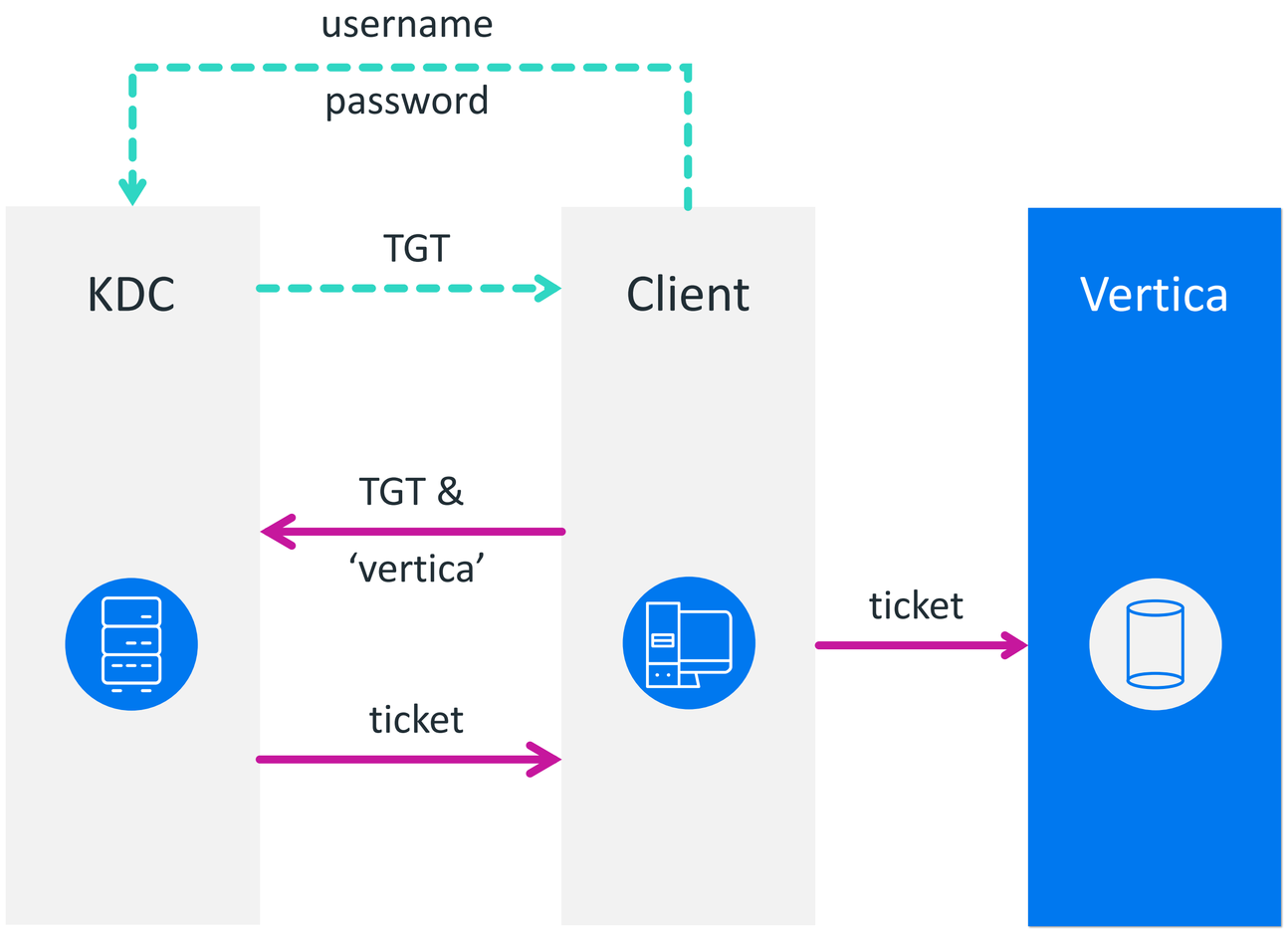
- Password auth
- Hash auth
- User Profiles
  - Password complexity requirements
- REJECT
  - Non-TLS connections
  - Connections from outside network
- Should only be used by small to medium deployments

LDAP Auth

- Allows users to be managed centrally, eg Active Directory
- Users authenticated against an LDAP server
- Supports TLS connections to LDAP server
- Create LDAP authentication and make it the default
  - CREATE AUTHENTICATION ldap_auth METHOD ‘ldap’ HOST ‘10.10.0.0/16’;
  - ALTER AUTHENTICATION ldap_auth SET host='ldap://172.16.65.177’, binddn_prefix='cn=', binddn_suffix=',ou=dev,dc=example,dc=com';
  - GRANT AUTHENTICATION ldap_auth TO PUBLIC;



Kerberos Auth



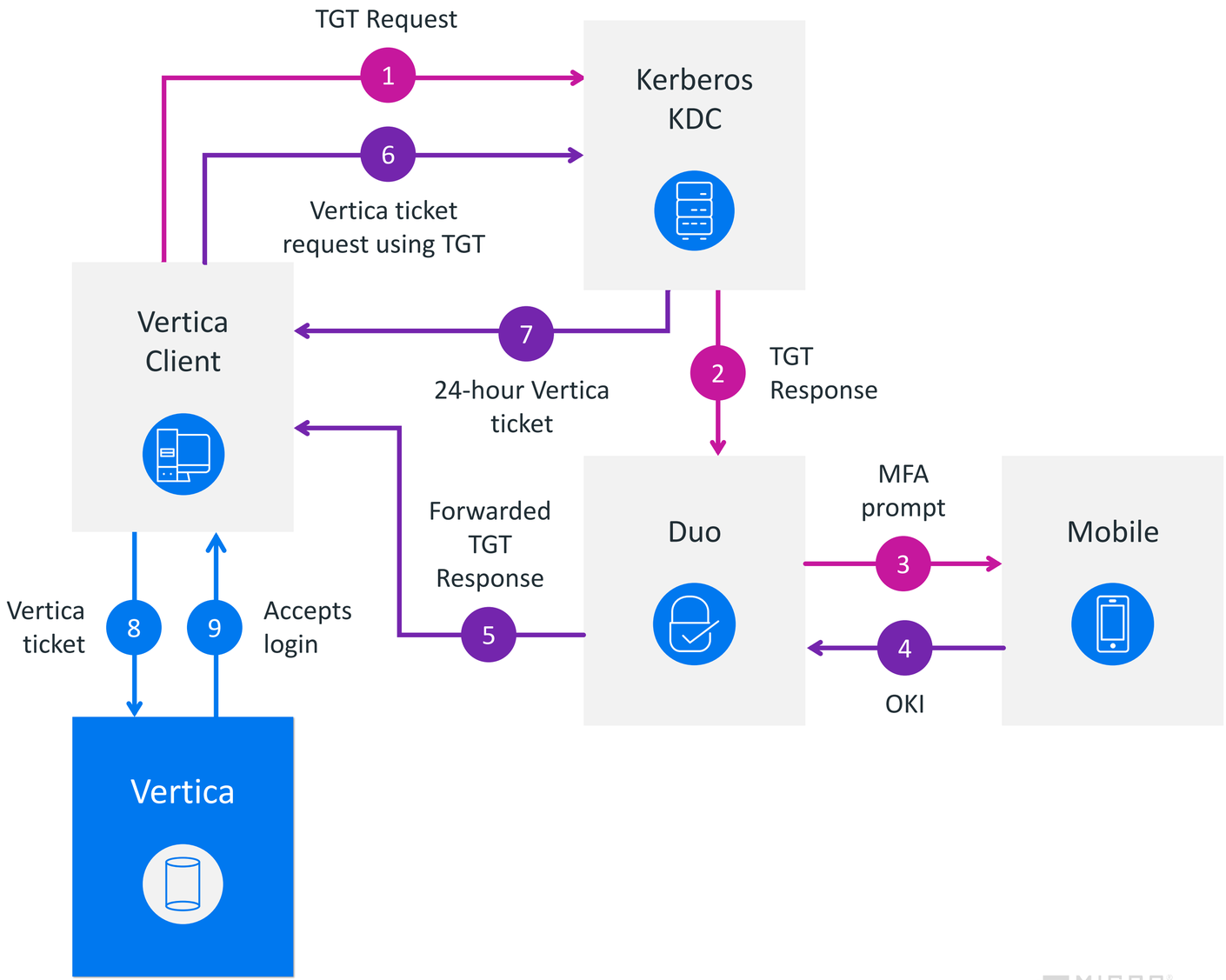
- Similar to LDAP
- Single sign on within organization
- Requires KDC
- Principals in the KDC
  - Vertica server
  - Users/applications that connect to vertica

X.509 Certificate Auth

- Clients use a valid certificate to authenticate
- Requires network-level TLS mutual authentication
- Especially useful for non-interactive use cases

Multi-Factor Authentication

- Not built into Vertica
- Can do it using Kerberos
- See blog https://www.vertica.com/kb/Vertica_Kerberos_and_Duo_Technical_Exploration/Content/Partner/Vertica_Kerberos_and_Duo_Technical_Exploration.htm



What Am I?

What Am I?

USER

Alice 

ROLES

Analyst 

Dbadmin 

Manager 

SET ROLE Dbadmin;

USER

Alice 

ROLES

Analyst 

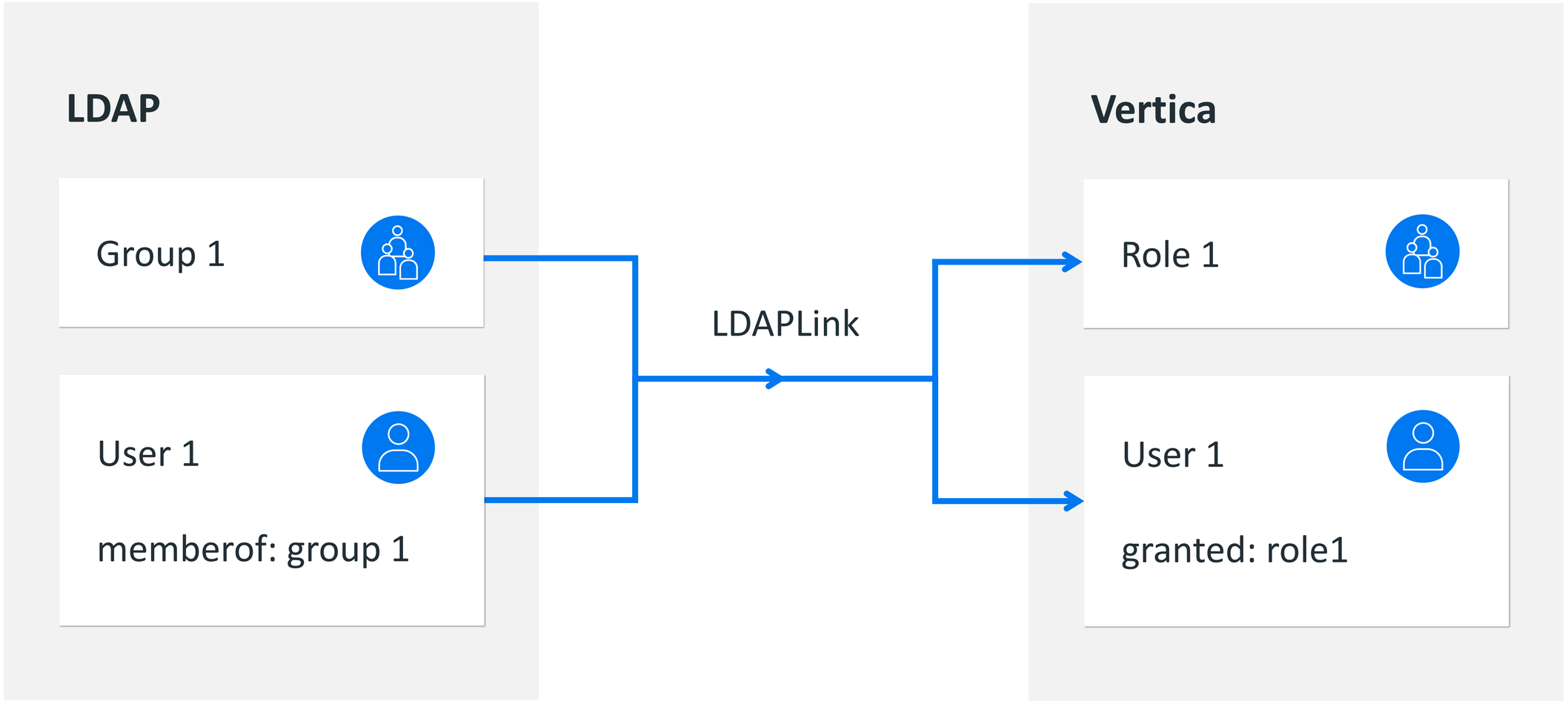
Dbadmin 

Manager 

How Are They Managed?

- We are a small organization, or only some people/services need Vertica access
  - Solution: manage with Vertica
    - CREATE/ALTER/DROP USER
    - CREATE/ALTER/DROP/GRANT/REVOKE ROLE
- We are a big organization, and we want Vertica to reflect what's in our centralized user management system
  - Use LDAPLink!

LDAPLink


LDAP
Group 1
User 1
memberof: group 1
LDAPLink
Vertica
Role 1
User 1
granted: role1

LDAPLink Dry Run

- Strictly isolate test configurations from the live LDAPLink system
  - Dry run uses explicitly separate versions of LDAPLink configuration parameters
- Easily access results of the dry run and parameters through a Data Collector table rather than multiple log files.
- Make sure to use the new dry run metafunctions!

```
=> SELECT event_timestamp, event_type, entry_name, ldapurihash, link_scope, search_base from LDAP_LINK_DRYRUN_EVENTS;
        event_timestamp          |    event_type    |       entry_name       | ldapurihash | link_scope | search_base
---------------------------------+------------------+------------------------+-------------+------------+-------------
2020-01-03 21:03:26.411753+05:30 | BIND_STARTED     | ---------------------- |           0 | sub        | dc=DC,dc=com
2020-01-03 21:03:26.422188+05:30 | BIND_FINISHED    | ---------------------- |           0 | sub        | dc=DC,dc=com
2020-01-03 21:03:26.422223+05:30 | SYNC_STARTED     | ---------------------- |           0 | sub        | dc=DC,dc=com
2020-01-03 21:03:26.422229+05:30 | SEARCH_STARTED   | **********             |           0 | sub        | dc=DC,dc=com
2020-01-03 21:03:32.043107+05:30 | LDAP_GROUP_FOUND | Account Operators      |           0 | sub        | dc=DC,dc=com
2020-01-03 21:03:32.04312+05:30  | LDAP_GROUP_FOUND | Administrators         |           0 | sub        | dc=DC,dc=com
2020-01-03 21:03:32.043182+05:30 | LDAP_USER_FOUND  | user1                  |           0 | sub        | dc=DC,dc=com
2020-01-03 21:03:32.043186+05:30 | LDAP_USER_FOUND  | user2                  |           0 | sub        | dc=DC,dc=com
2020-01-03 21:03:32.04319+05:30  | SEARCH_FINISHED  | **********             |           0 | sub        | dc=DC,dc=com
```

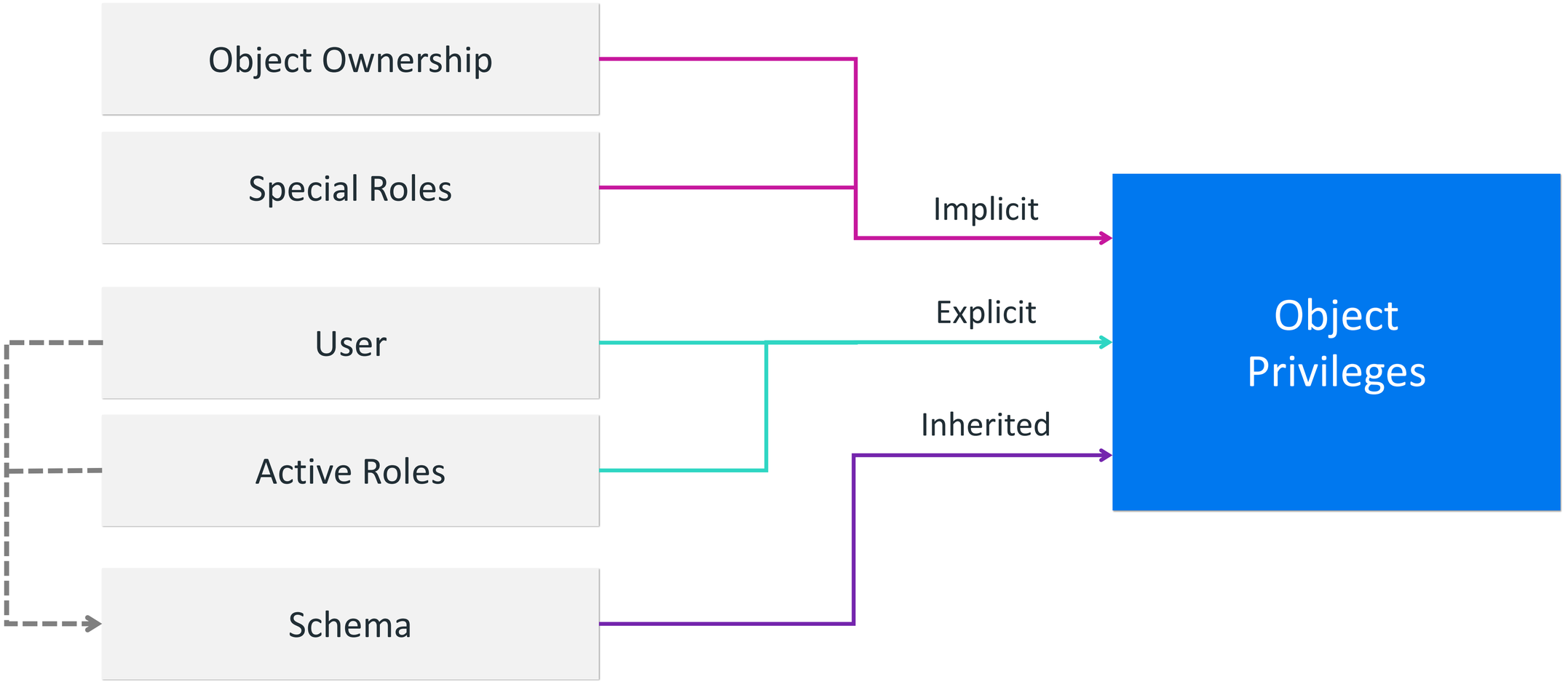
What Can I Do?

What Can I Do?

- I want to control who can do what. What is the question I must ask?
- Sometimes the question is just “Who are you?”
  - CREATE USER => am I the superuser?
  - See an active session => is it my session?
- Sometimes more complex: does the principal have the required privilege(s)?
  - SELECT FROM schema.table=> USAGE on schema, SELECT on table
  - SELECT udxFunction() => USAGE on library, EXECUTE on udxFunction
  - See a table => USAGE on schema, any privilege on table

ALTER/DROP used to be simple actions (owner only). Now they are complex actions which map to ALTER/DROP privileges.

Where Do Privileges Come From?



Special Roles

Role Name	Permissions
DBADMIN	Specific elevated privileges (refer to documentation)
PSEUDOSUPERUSER	Anything the real superuser can do
DBDUSER	Can run Database Designer functions
SYSMONITOR	Elevated auditing permissions (more on this later)
PUBLIC	Anything you want to be allowed for everyone

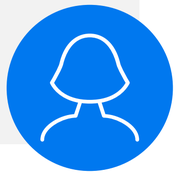
Scenario: Lots of Relations and Principals

- I have a really big schema with lots of relations
- For each principal, I want the privileges for all tables and views in the schema to be the same
- Approach 1: remember to run grants every time a new table/view is created (a pain!)
- Approach 2: use schema-inherited privileges!

Schema

We need full access to all tables and views.

Analyst

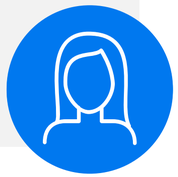


We need read access only.

Salespeople

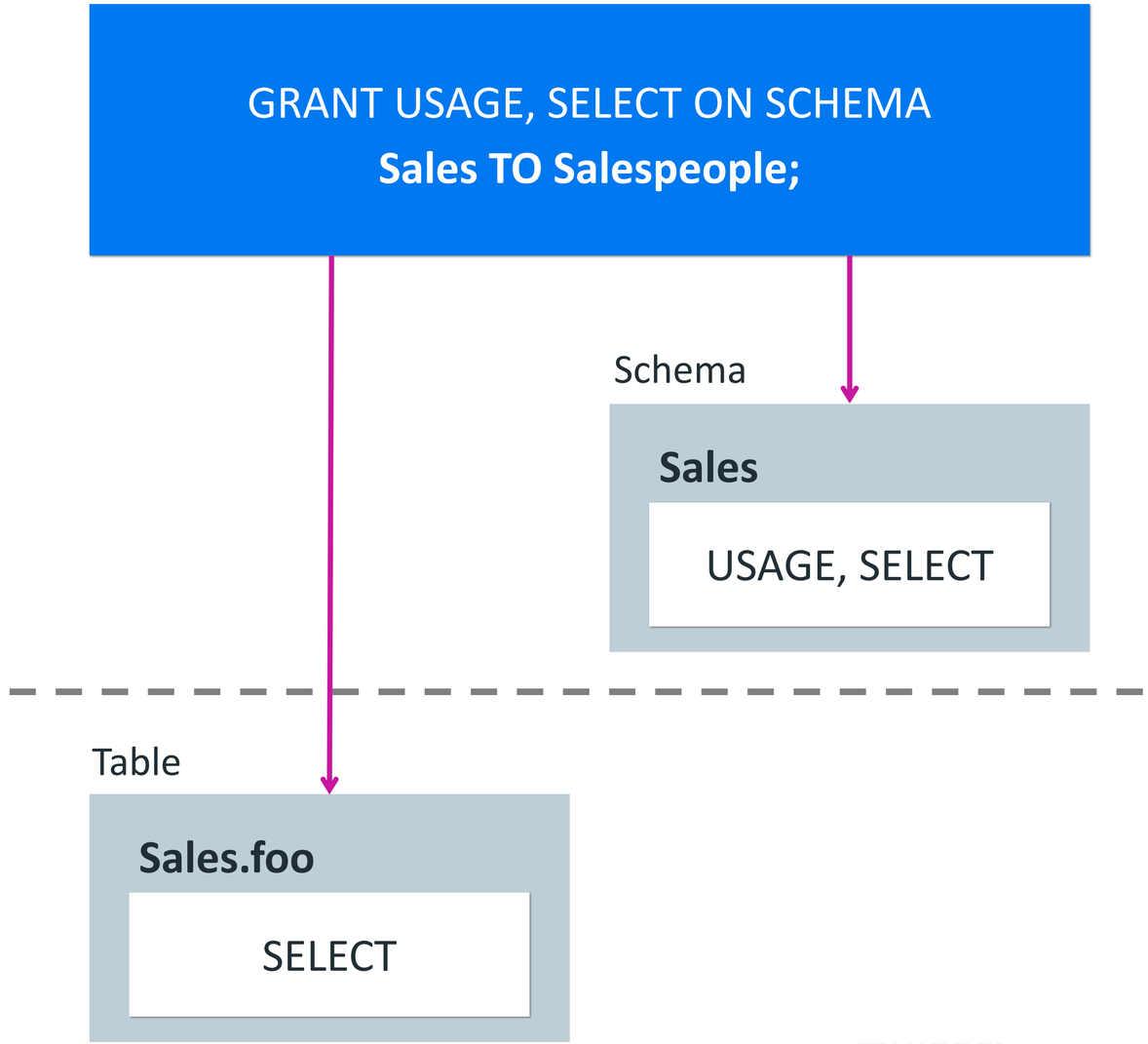
We need read access, and to clean up other people's tables and views.

Manager



Inherited Privileges

- Schema grants can include relational privileges
- If a relation is marked as inheriting, schema grants to a principal also apply to the relation!
- Our scenario: instead of lots of grants, now we run one ALTER SCHEMA statement and three GRANT statements
  - ALTER SCHEMA sales DEFAULT INCLUDE PRIVILEGES;
  - GRANT SELECT ON SCHEMA sales TO salespeople;
  - GRANT ALL EXTEND ON SCHEMA sales TO analysts;
  - GRANT SELECT, DROP on SCHEMA sales TO manager;



Monitoring Privileges

- Three system tables
  - *grants*
  - *inherited_privileges*
  - *inheriting_objects*

```
=> SELECT grantee, privileges_description FROM grants WHERE object_name='myschema';
 grantee | privileges_description
---------+------------------------
 Bob     | USAGE, CREATE
 Alice   | CREATE
 (2 rows)
```

- Show privileges that are active for you
  - Your user
  - Active roles (and their roles, and theirs, etc.)
- Shows tables and views which inherit privileges
- For the current user: sum up ALL privileges with a single metafunction!
  - *get_privileges_description('type', 'qualifiedName');*

Fine-Grained Access Control

I should see a full SSN so I can insert and update them.

Loader Service

123-45-6789
234-56-7890
345-67-8901

I should see the last four digits of an SSN.

Human Resources

XXX-XX-6789
XXX-XX-7890
XXX-XX-8901

I shouldn’t be able to see SSNs.

Analyst

XXX-XX-XXXX
XXX-XX-XXXX
XXX-XX-XXXX

Access Policies

- Row and column access policies
  - Hide or transform data in a row/column depending on who is SELECTing
- If access policies let you see the raw data, you can still modify the data
  - Access policies should be used to refine access for principals with read-only access
- Monitorable by system table *access_policy*

VERTICA

Delegation and Impersonation

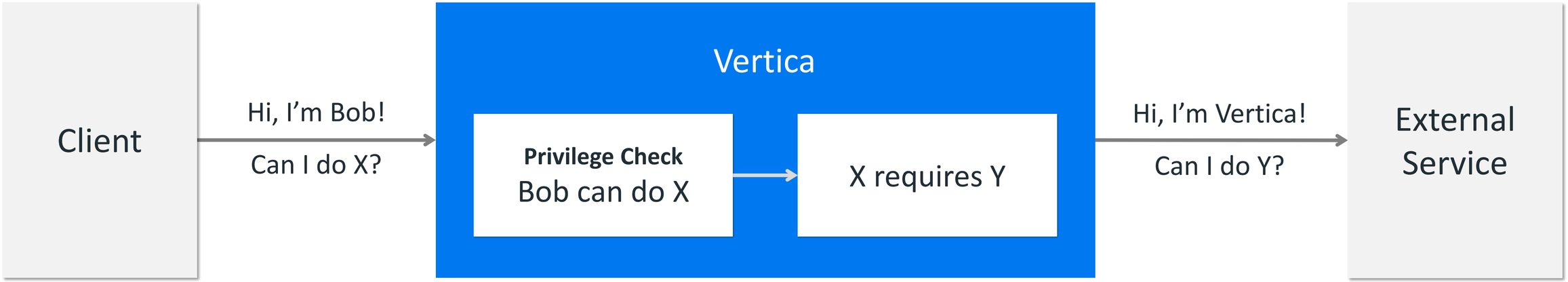
MICRO FOCUS

Who is Vertica?

Who Is Vertica?

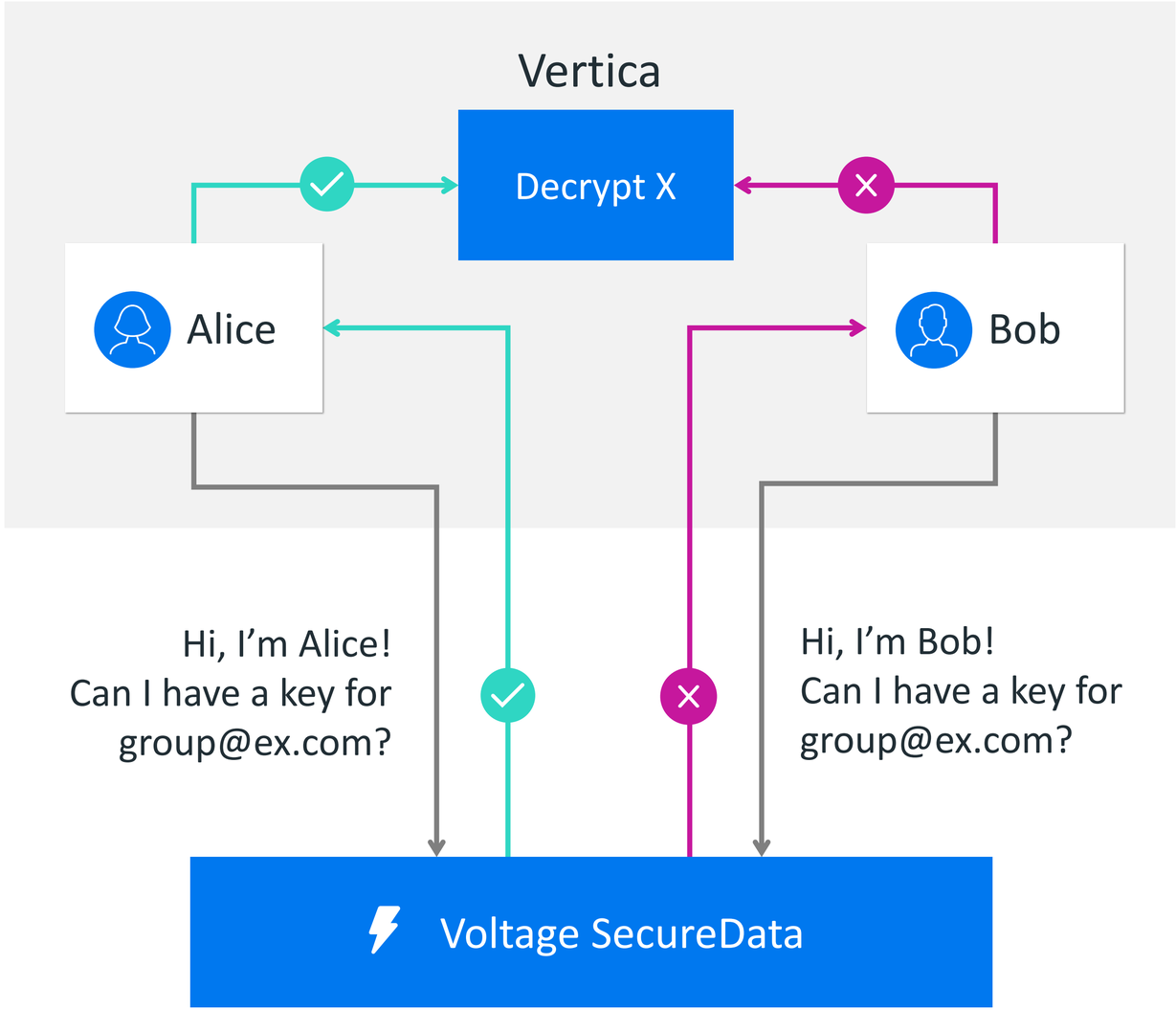
- When Vertica connects to a downstream service, how should Vertica identify itself?
- Most services: we connect as Vertica, and make sure to do so only if the Vertica principal is properly authorized

- Some services support a better model of delegation
  - Vertica can connect on behalf of the upstream user that initiated the action, using some identity which corresponds to that user
- Let the service decide who can do what, so Vertica isn't the only line of defense!
- Big benefits for auditing

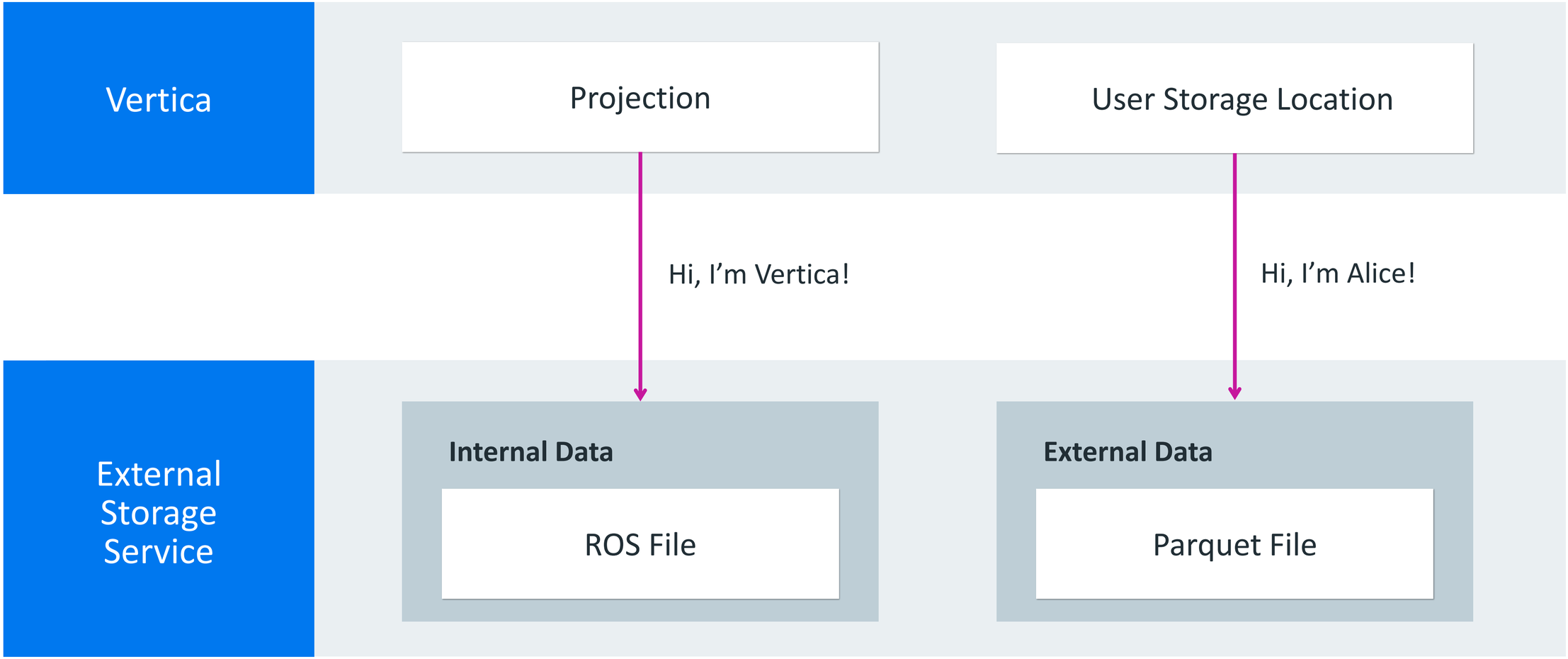


Voltage SecureData

- Use case 1: I'm just encrypting for compliance/anti-theft reasons
  - Solution: Use one "global" identity to encrypt/decrypt with Voltage
- Use case 2: I want to control which users can decrypt which data
  - Solution: give Voltage users access to appropriate *identities*, which control encryption for sets of data
  - A Voltage user can access multiple identities (like groups)
  - Vertica users can set their Voltage username/PW in Vertica
  - Vertica talks to Voltage as that Voltage user!



The Storage Problem



AWS S3

- What if I want to COPY FROM/EXPORT TO PARQUET my own bucket?
- Option 1:
  - Give Vertica as a whole access to the bucket
  - Problematic – least access, zero trust
- Option 2:
  - Use an ID/secret key pair to an AWS (IAM) principal that does have access to the bucket
  - Set at session level – more isolation
- Coming up:
  - Support for “keyless” delegation using assumable IAM roles

Hadoop

Option 1
Kerberos Delegation

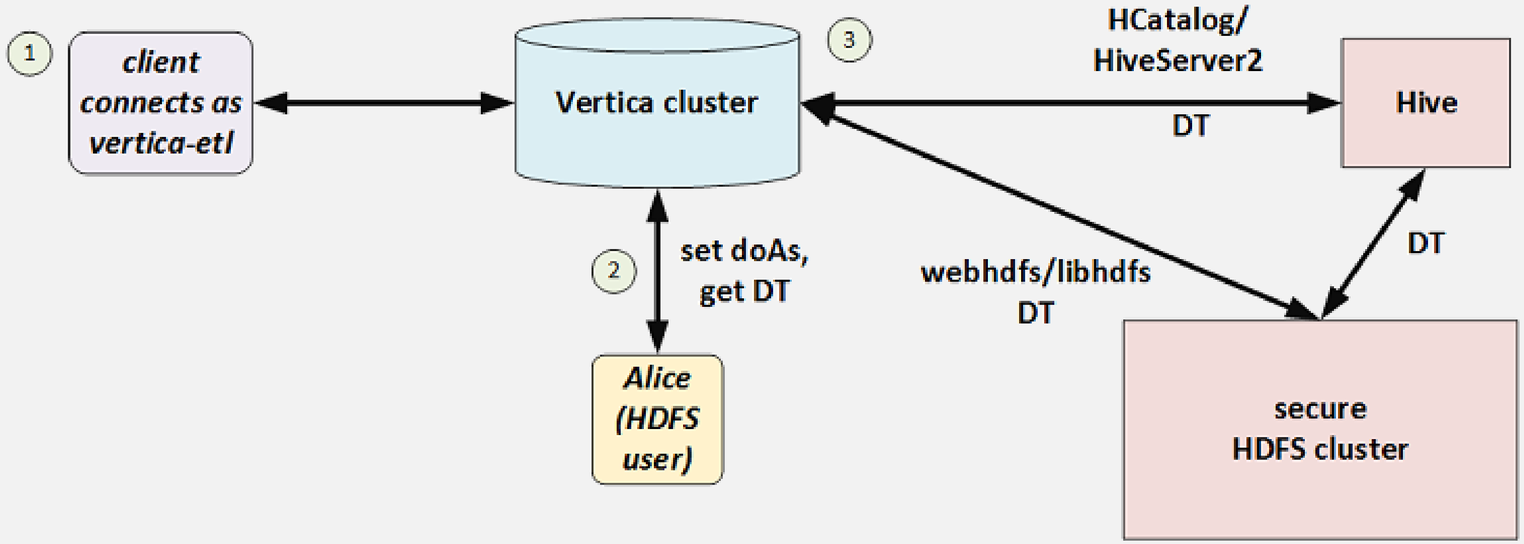
- Most secure, tightest access control, most Kerberos/HDFS configuration required

Option 2
Secure Impersonation

- Simple, best with highly trusted Vertica userbase, best when auditing is your primary concern

Option 3
“Bring Your Own Delegation Token”

- Flexible but manual, similar to AWS



Option 2: Secure Impersonation

VERTICA

Auditing and Monitoring

MICRO FOCUS

What’s happening?

System Tables

- Collection of information about events, system state, performance, etc.
- SELECT only

- Two types:
  - Metadata (V_CATALOG)
    - E.g. *users*
  - Monitoring (V_MONITOR)
    - E.g. *disk_storage*

```
cmorris=> select name, memorysize, maxmemorysize, priority, runtimepriority, executionparallelism from resource_pools;
   name    | memorysize | maxmemorysize | priority | runtimepriority | executionparallelism
-----------+------------+---------------+----------+-----------------+----------------------
 general   |            | Special: 95%  |        0 | MEDIUM          | AUTO
 sysquery  | 1G         |               |      110 | HIGH            | AUTO
 sysdata   | 0M         | 0M            |          |                 |
 tm        | 13G        |               |      105 | MEDIUM          | AUTO
 refresh   | 0%         |               |      -10 | MEDIUM          | AUTO
 recovery  | 0%         |               |      107 | MEDIUM          | AUTO
 dbd       | 0%         |               |        0 | MEDIUM          | AUTO
 jvm       | 0%         | 2G            |        0 | MEDIUM          | AUTO
 blobdata  | 0%         | 10%           |          |                 |
 metadata  | 0%         |               |          |                 |
(10 rows)
```

Who Can See This Stuff?

- Metadata can be sensitive, too
  - Principle of least privilege = only show auditing/monitoring info to those who need it
- [Pseudo]superuser sees all
  - Only role that can see sensitive config parameters
- How to show other users information they need?
  - SYSMONITOR
  - [RESTRICT|RELEASE]_SYSTEM_TABLES_ACCESS()

More Control: Roles and Grants

- SYSMONITOR and RESTRICT/RELEASE metafunctions are inflexible
- Willing to do a little more setup? Use your own grants and roles!
  - System tables support GRANT/REVOKE statements like regular relations
- Flexible and fine-grained approach:
  - Revoke access to system tables of your choice from PUBLIC
  - Create new roles and grant them to users and other roles
  - Grant access to appropriate system tables to the new roles
- More fine-grained access required?
  - Create views!

Logs and the Data Collector

- Superusers (and those with filesystem/OS access) have more ways to see events
- Logging:
  - vertica.log (dbLog during startup)
  - SNMP
  - Syslog
  - *active_events* (a sensitive system table)
- Data Collector
  - Sorts events by subject (“component”) and logs to rotating files
    - E.g. *AnalyzeStatistics*
  - Files can be exported into a monitoring database: **Management Console Extended Monitoring**

Putting It All Together

Topic	Question
Encryption	How do I protect my data at rest?
	How do I protect my data in transit?
Authentication	How do I prove who I am?
Identity	Who and what am I?
Authorization	What can I do?
Delegation/Impersonation	When talking to other systems, who is Vertica?
Auditing/Monitoring	What's happening in my database?


VERTICA
Q&A
MICRO FOCUS

Vertica Academy

Your source for new comprehensive Vertica on-demand training

- Self-Paced Learning
- Online Training
- Knowledge Check
- Certificate of Completion & Certifications

Sign up for free today at **academy.vertica.com**

VERTICA

Vertica Forums

Vertica Engineers are available to answer your questions and moderate discussions right now

For more information visit **forum.vertica.com**