



Guía de configuración de tripletas

Control de versiones	2
Introducción	2
Cómo se segmenta la infraestructura por tripletas	3
Capacidades de los plugins de Tot	3
Gobierno Activo	3
Sampleo de datos	5
Extracción de Metadata	5
Limitaciones o restricciones	6

Control de versiones

Versión	Fecha de modificación	Responsable	Aprobador	Resumen de cambios
1.0	28/10/2022	Anjana Producto	Anjana Producto	Creación del documento

Introducción

El presente documento es una guía para exponer:

- Cómo se segmenta la infraestructura por tripletas con el objetivo de calificarla de forma coherente a las capacidades de los plugins de tot
- Las capacidades de los plugins
- Las restricciones o limitaciones de los plugins

Cómo se segmenta la infraestructura por tripletas

Con vistas a un gobierno proactivo y preventivo en el que se pueda realizar extracción automática de metadata y gobierno activo (gestión de estructuras y gestión de permisos) se diseñan tripletas de valores para los atributos infraestructura, tecnología y zona.

Se conoce como **tripleta** a la conexión lógica indicada como parámetro en la configuración que plugins y agentes se encargan de resolver en conexiones físicas para conectar con la fuente.

La tripleta está compuesta por: infraestructura, tecnología y zona.

Para segmentar la infraestructura de la forma más práctica hacia el gobierno del dato pudiendo sacar el máximo partido a las capacidades de los plugins de tot se puede seguir el siguiente ejemplo:

- Particionamiento lógico de los sistemas de la organización (tabla izquierda)
- Tripletas obtenidas del análisis (tabla derecha)

Tipo	Infraestructura propia o proveedor	Tecnología	Localización		Tipo de entidad	Infraestructura propia o proveedor	Tecnología	Zona
Almacenamiento	Propia	Hadoop Hive	Cluster EU	→ TRIPLETA	Dataset	OnPrem	Hive	EU
		Hadoop HDFS	Cluster EU				HDFS	EU
		Hadoop Hive	Cluster USA				Hive	USA
		Hadoop HDFS	Cluster USA				HDFS	USA
	Google Cloud Platform	Bigtable	West EU			GCP	Bigtable	West EU
		Bigquery	West EU			Bigquery	West EU	
Procesamiento	Propia	Spark	Cluster EU	Procesos	OnPrem	Spark	EU	
	GCP	Apengine	West EU		GCP	Apengine	West EU	

Otro ejemplo de zona que suele usarse es: dev / pre / pro; en vez de la zona geográfica.

Capacidades de los plugins de Tot

Gobierno Activo

Permite delegar la provisión de grupos de usuarios a los que posteriormente se asignan permisos de acceso a recursos de datos gobernados por el producto que se explota.

`totplugin:`

```
location: http://totpluginazureadserver:15006/plugin/azure/ad/api/v1
server:
  url: http://totserver:15000/tot/
  keep-alive-seconds: 60
connection:
  pathSeparator: "/"
  clientId: "aaaaaaa-bbbb-cccc-dddd-eeeeeeeeeeee"
  tenantId: "aaaaaaaaaaa-bbbb-cccc-dddd-eeeeeeeeeeee"
  secret: "aaaaaaaaabbbbbbbccccccccceeeeeeeedddddddd"
  scopes: "https://graph.microsoft.com/.default"
aris:
  - ari: "anja:totplugin:im:/azure/ad/pre/"
```

Además, es capaz de crear recursos de datos en el producto, que se explotan si en Anjana se crea una entidad como gobernada en una triplete con la configuración de gobierno activo.

```
totplugin:
  location: http://totpluginazurefileserver:15007/plugin/azure/file/api/v1
  server:
    url: http://totserver:15000/tot/
    keep-alive-seconds: 60
  connection:
    pathSeparator: "/"
    sampleRows: 15
    storageAccount: totpluginazure
    key: "aaaaaaaaabbbbbbbccccccccccddddddeeeeeeeffffff=="
  aris:
    - ari: "anja:totplugin:extract:/azure/files/pre/"
    - ari: "anja:totplugin:sample:/azure/files/pre/"
    - ari: "anja:totplugin:im:/azure/files/pre/"
    imAri: "anja:totplugin:im:/azure/ad/pre/"
```

En el ejemplo anterior lo que sucede por etapas es:

1. El plugin en el que se va a delegar la gestión de identidades se pone la triplete igual que en su configuración pero comenzando por "imAri" debajo de la triplete que gestiona los permisos.
2. El plugin de gestión de identidades (origen de triplete imAri) creará un grupo, dando membresías a los usuarios y devolverá un identificador a la triplete "- ari: anja:totplugin:im:"
3. La triplete que gestiona los permisos (- ari: anja:totplugin:im:) con ese identificador recibido dará permisos mediante la tecnología específica (GRANT, ACL, ...)

Sampleo de datos

Mediante la resolución de la conexión lógica aportada por la tripleta, el plugin se conecta al recurso del dato y trae las primeras X filas del contenido de éste.

`sampleRows: 15` → Número de filas que vamos a poder ver.

```
totplugin:
  location: http://totpluginazurefileserver:15007/plugin/azure/file/api/v1
  server:
    url: http://totserver:15000/tot/
    keep-alive-seconds: 60
  connection:
    pathSeparator: "/"
    sampleRows: 15
    storageAccount: totpluginazure
    key: "aaaaaaaaabbbbbbbbbbccccccccccdddddddeeeeeeeffffff=="
  aris:
    - ari: "anja:totplugin:extract:/azure/files/pre/"
    - ari: "anja:totplugin:sample:/azure/files/pre/"
    - ari: "anja:totplugin:im:/azure/files/pre/"
    imAri: "anja:totplugin:im:/azure/ad/pre/"
```

Extracción de Metadata

A través de la resolución de la conexión lógica aportada por la tripleta el plugin es capaz de conectarse al recurso del dato y devolver el metadato de este.

```
totplugin:
  location: http://totpluginazurefileserver:15007/plugin/azure/file/api/v1
  server:
    url: http://totserver:15000/tot/
    keep-alive-seconds: 60
  connection:
    pathSeparator: "/"
    sampleRows: 15
    storageAccount: totpluginazure
    key: "aaaaaaaaabbbbbbbbbbccccccccccdddddddeeeeeeeffffff=="
  aris:
    - ari: "anja:totplugin:extract:/azure/files/pre/"
    - ari: "anja:totplugin:sample:/azure/files/pre/"
    - ari: "anja:totplugin:im:/azure/files/pre/"
    imAri: "anja:totplugin:im:/azure/ad/pre/"
```

Limitaciones o restricciones

1. Cada plugin sólo gestiona una conexión, de forma que si nuestra infraestructura requiere de conectar con dos SQLServer distintos, debemos tener dos plugins de SQLServer levantados.
2. Con ejemplo/restricción anterior hay que tener en cuenta de que no se pueden compartir tripletas entre distintos plugins
 - a. El primer plugin sqlserver puede tener la tripleta /OnPrem/sqlserver/dev/
 - b. El segundo plugin sqlserver podría tener la tripleta /Azure/sqlserver/dev/