BigQuery Data Security & Compliance

访问控制

详细文档：https://cloud.google.com/bigquery/docs/access-control

这个非常直观，就不在这里赘述了。需要注意的一个最佳实践就是，避免直接给一个对象（人，组）授权，而是一个Role（角色）授权，然后把role赋予一个对象。

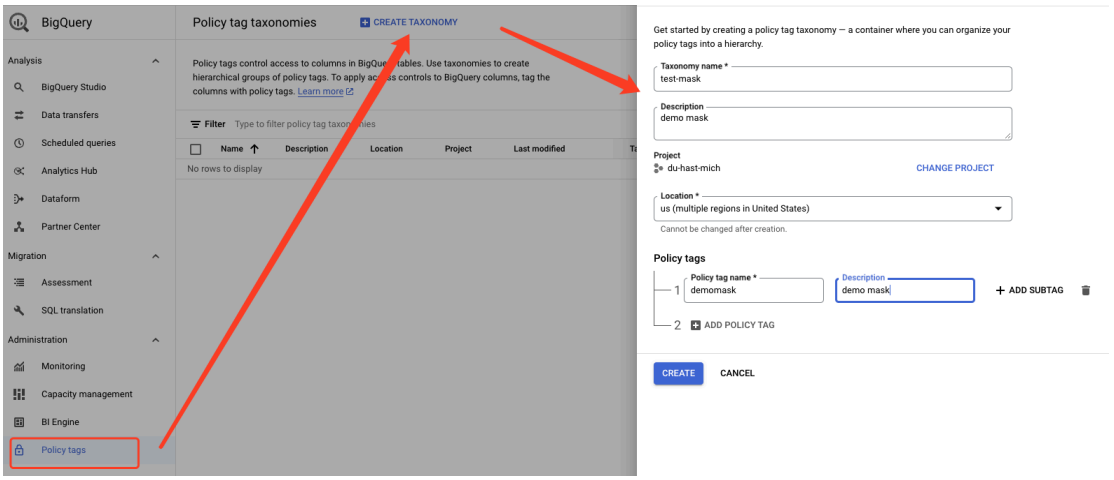
数据遮盖

BigQuery数据遮盖：https://cloud.google.com/bigquery/docs/column-data-masking-intro

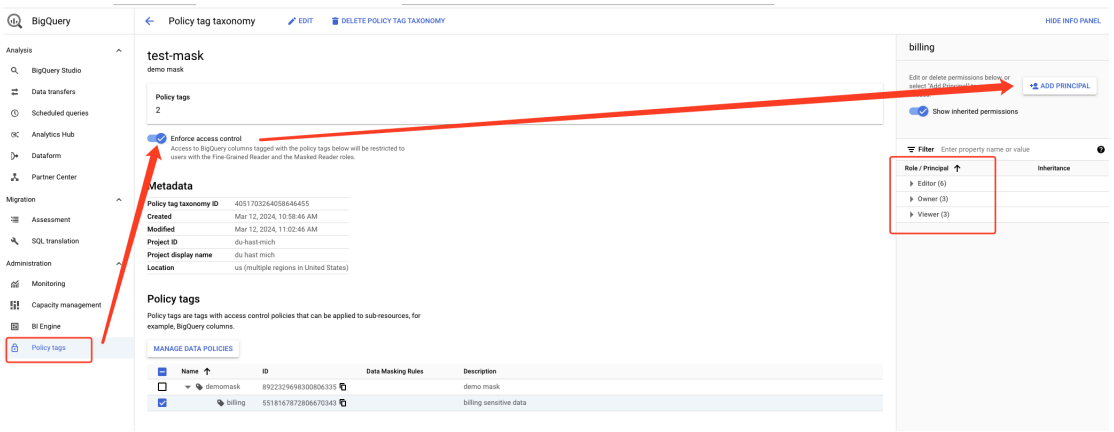
这块区别于访问控制，需要打开额外两个API

- Google Cloud Data Catalog API
- BigQuery Data Policy API

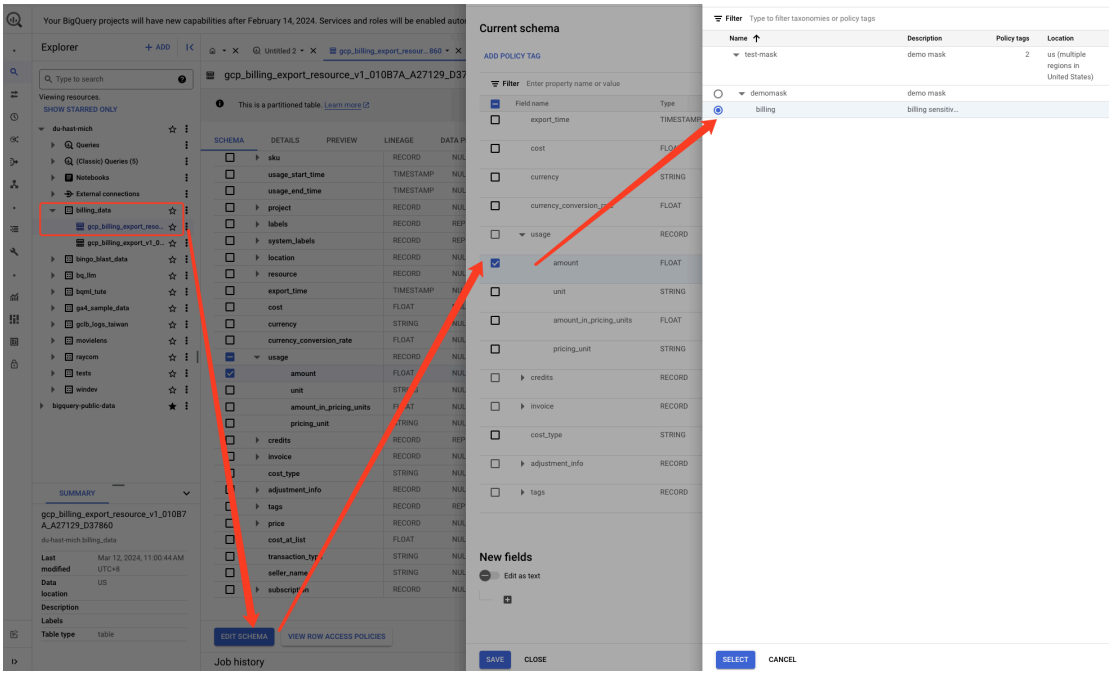
这里用通过policy tag（政策标签）来控制



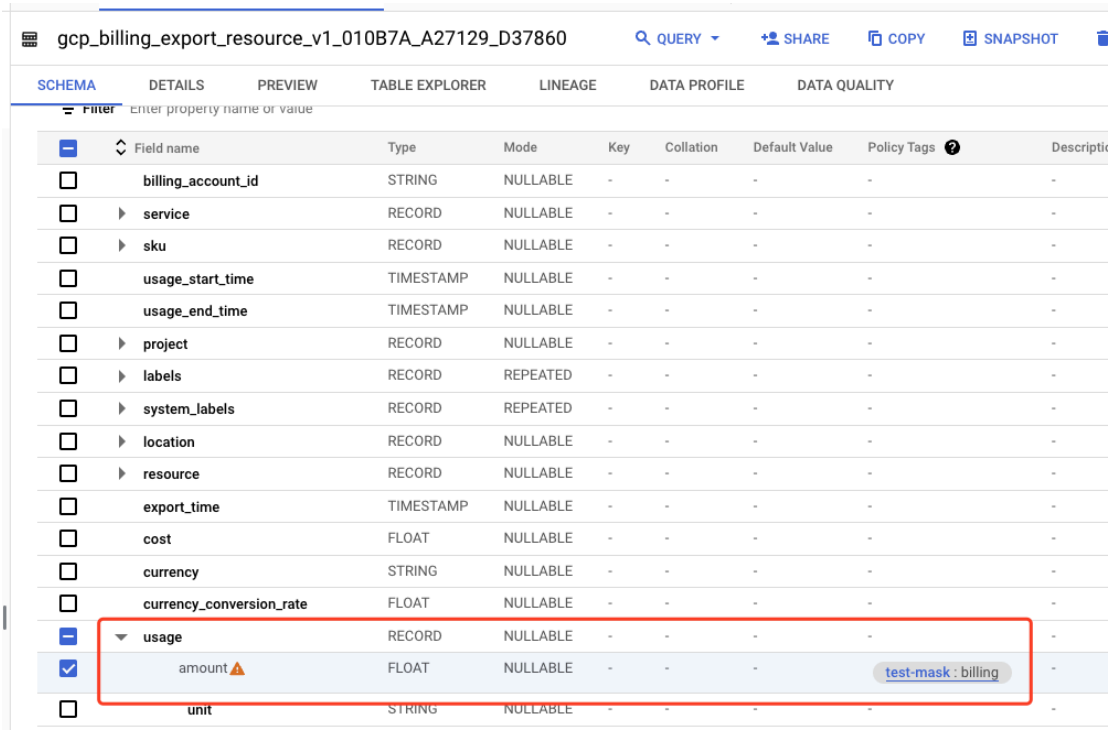
添加政策权限



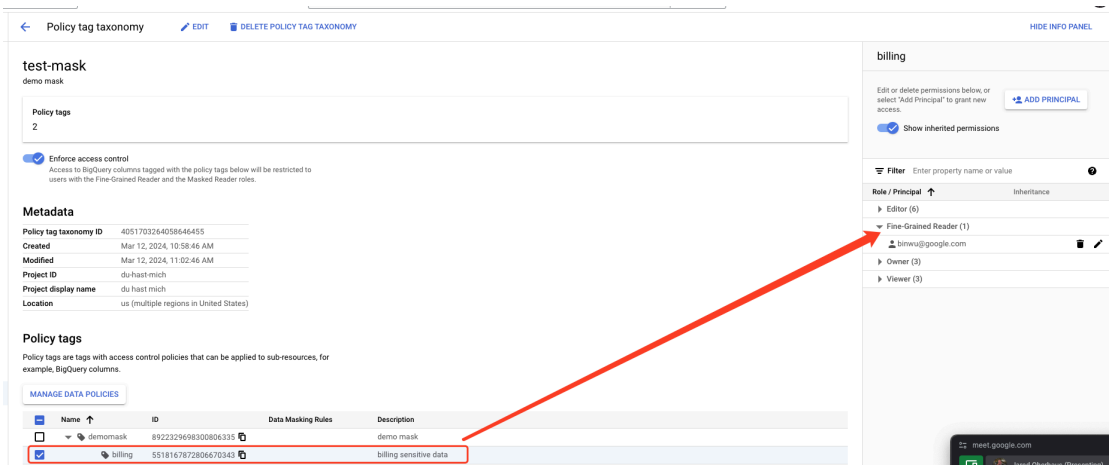
为列打上政策标签



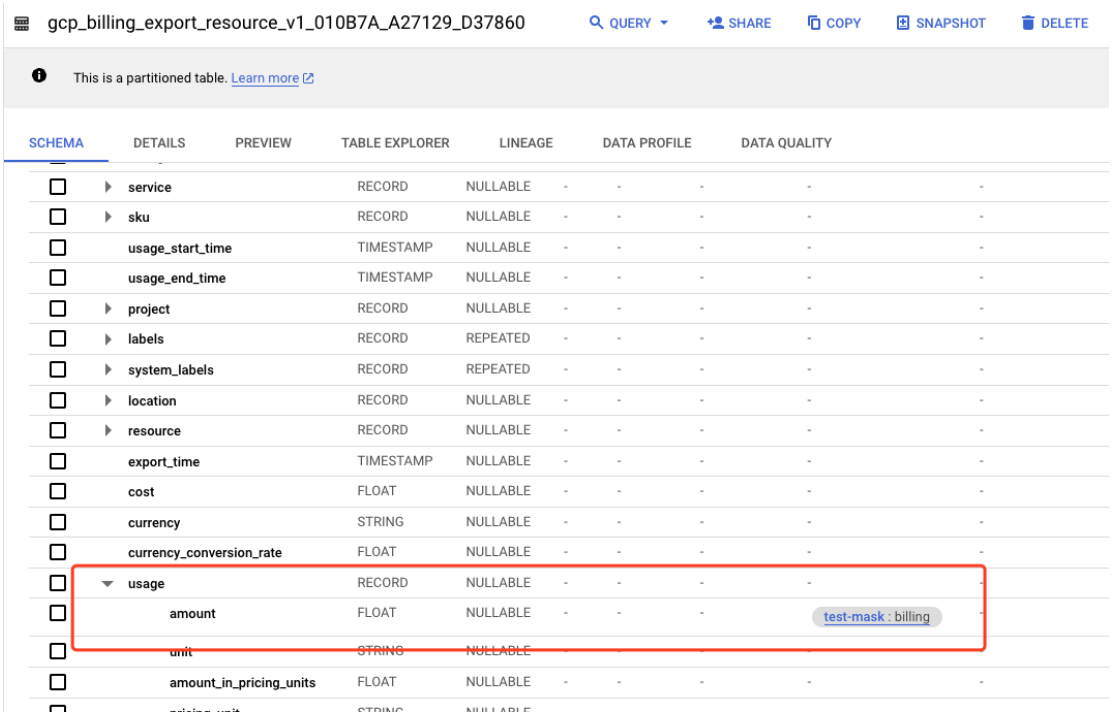
没有 fine-grained权限 用户访问受限



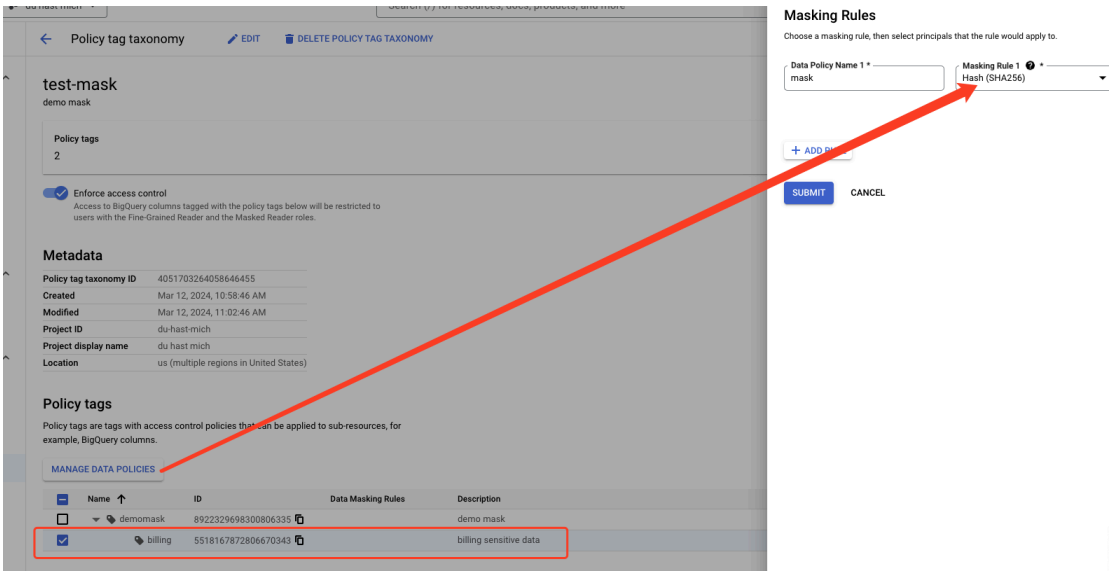
给对象添加可以访问数据的角色，角色具有相应权限



授权后可以正常访问



添加遮盖逻辑



授权给用户可以看到遮盖后数据

DLP

开通DLP API

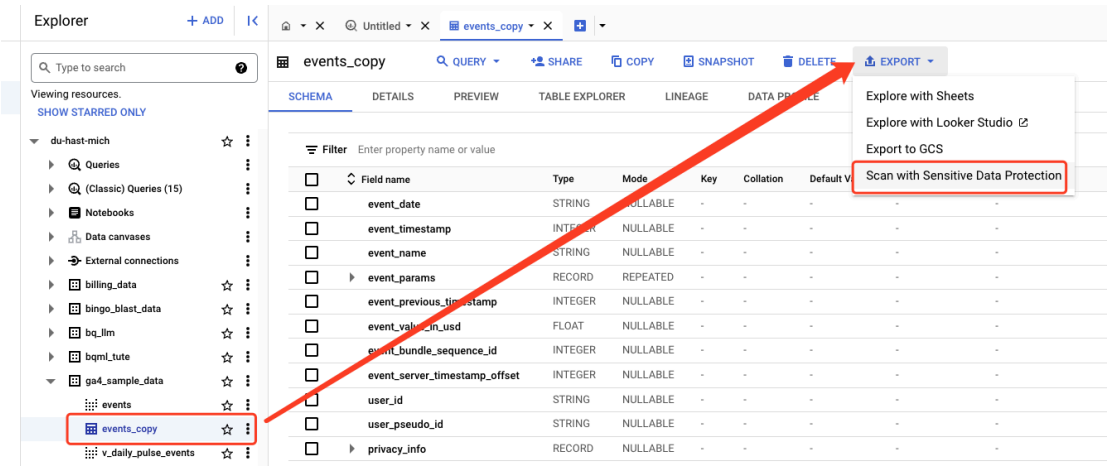
REST API: https://cloud.google.com/sensitive-data-protection/docs/reference/rest

这里理论上可以使用rest或者client lib来做线下或者实时检查

```
Python
client = bigquery.Client()
datasets = list(client.list_datasets(project=project_id))

if datasets:
    for dataset in datasets:
        tables = client.list_tables(dataset.dataset_id)
        for table in tables:
            #调用API 抽样检查数据 识别风险列
```

BQ里可以相对简单的扫描已有数据



用程序读取Pub/sub，并且定期轮换密码/key

用程序（dataflow）读取pubsub：
https://cloud.google.com/pubsub/docs/stream-messages-dataflow

pubsub 也有丰富的SDK：https://cloud.google.com/pubsub/docs/reference/libraries

SA key rotation：https://cloud.google.com/iam/docs/key-rotation

附录

1，BQ目前有三个版本，Data masking等功能在Enterprise以上版本中提供支持

如下图所示：（link）

	BigQuery editions			
	Standard	Enterprise	Enterprise Plus	On-demand pricing
Pricing model	Slot-hours (1 minute minimum)	Slot-hours (1 minute minimum)	Slot-hours (1 minute minimum)	Pay per query with free tier
Compute model	Autoscaling	Autoscaling + Baseline	Autoscaling + Baseline	On-demand
Fine-grained security controls	No access to fine-grained security controls	Column-level access control Row-level security Dynamic data masking	Column-level access control Row-level security Dynamic data masking Custom data masking	Column-level access control Row-level security Dynamic data masking Custom data masking

因此，为了进行详细的权限管理和Masking需要采用Enterprise edtion，虽然在成本方面Enterprise高于Standard，但通过合理规划可以避免成本明显增长，

- 在测试评估阶段，建议采用enterprise - pays as you go模式，这阶段查询较少因此不会产生过多费用
- 将来上了生产环境，建议采用enterprise-commit（1yr/3yr）模式，成本相较于目前的standard - pay as you go模式不会明显增长

同时，目前的测试建议在单独project进行，不影响生产系统采用的计费模式。

Standard和Enterprise editions的价格说明：(link)

Standard Edition

The following table shows the cost of slots in Standard edition.

US (us)

Commitment model	Hourly cost	Details
Pay as you go	$0.04 / slot hour	No commitment. Billed per second with a 1 minute minimum

Enterprise Edition

The following table shows the cost of slots in Enterprise edition.

US (us)

Commitment model	Hourly cost	Details
Pay as you go	$0.06 / slot hour	Billed per second with a 1 minute minimum
1 yr commit	$0.048 / slot hour	Billed for 1 year
3 yr commit	$0.036 / slot hour	Billed for 3 years