Enterprise Geodatabase: Topics in PostgreSQL Administration

Jim McAbee & Neeta Reddy

SEE WHAT OTHERS CAN'T

Intended Audience

You are…..

- A geodatabase administrator
- A deliberate DBA
- An accidental DBA
- Not sure what DBA means

And you…

- Store your data in a PostgreSQL database
- Are thinking about using PostgreSQL

This is your session!

Agenda

How do I …

- Configure PostgreSQL for enterprise geodatabases?
- Manage and maintain database security?
- Store and manage spatial data?
- Achieve optimal performance?

News since the last UC

How do I...

Configure PostgreSQL for enterprise geodatabases?

Installation Requirements

- Install a supported version of PostgreSQL
  - PostGIS Extension - *optional, except in AWS RDS and Azure*
  - Windows, Redhat Enterprise, SUSE Enterprise, Ubuntu LTS
  - **https://pro.arcgis.com/en/pro-app/help/data/databases/database-requirements-postgresql.htm**
- PostgreSQL client
  - Most ArcGIS applications that can connect to a PostgreSQL database include the required PostgreSQL client libraries
- Required libraries
  - Download st_geometry library from myesri.com
  - Install Visual C++ Redistributable Package (On Windows only)
- PostGIS
  - Use of PostGIS geography type with ArcGIS requires PostgreSQL 9.6.8 and later
- Patches
  - Make sure you are up to date with ArcGIS performance patches for your version

Connection Configuration

- PostgreSQL configuration files for connections
  - Located in PostgreSQL*version*\data
  - postgresql.conf for max_connections
  - pg_hba.conf to allow access

postgresql.conf

```
listen_addresses = '*'
port = 5432
max_connections = 100
shared_buffers = 25% Ram, 512MB
work_mem = 128MB
maintenance_work_mem = 64MB
checkpoint_segments = 64
effective_cache_size = 2/3* RAM, 4GB
log_line_prefix = '%m [p=%p, d=%d, u=%u, x=%x]'
log_statement = 'all'
```

pg_hba.conf

```
# TYPE  DATABASE        USER            ADDRESS                 METHOD

# IPv4 local connections:
host    all             all             127.0.0.1/32            md5
host    all             all             10.11.0.0/16            md5


# IPv6 local connections:
host    all             all             ::1/128                 md5
```

Configuring postgresql.conf

- Large data loads
  - shared_buffers
  - max_locks_per_transaction
- Adjust rows returned for query
  - cursor_tuple_fraction – what fraction of rows returned for query,
  - Set to 1.0 (100%)
  - Changing cursor_tuple_fraction in postgresql.conf will not apply
  - Must set cursor_tuple_fraction environment variable
- For third-party SQL spatial query performance
  - ST_GEOMETRY_OUTPUT_FORMAT environment variable (to switch from WKT)

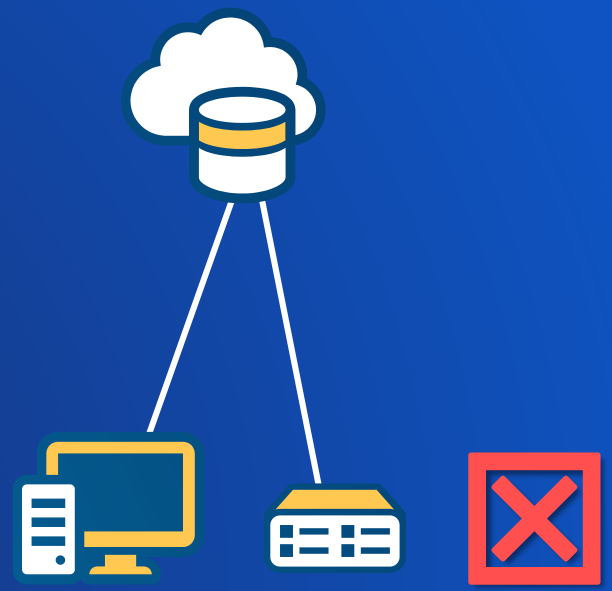
Memory Tuning in PostgreSQL

Creating Geodatabases

Create Enterprise Geodatabase	Enable Enterprise Geodatabase
Fast & simple, but…	Requires more planning, but…
Requires PostgreSQL administrator credentials	Needs fewer permissions
Cannot be used for DaaS	Can be used for DaaS

Cloud Deployments

- Options
  - Database Instance on Virtual Server
    - *ST_Geometry required*
    - *ST_Geometry or PostGIS spatial types*
  - Database as a Service (DaaS)
    - *uses PostGIS spatial type*
    - *no ST_Geometry*
- Co-locate Clients (Desktop or Server)



Co-locate with any direct connecting client/server

Cloud Deployments

- Database as a Service (DaaS)
  - Alternative deployment option
  - Auto-scaling options
  - Always uses PostGIS storage
- Amazon Web Services (AWS)
  - Relational Database Service(RDS) for PostgreSQL
  - Aurora PostgreSQL
- Microsoft Azure
  - Azure Database for PostgreSQL

Cloud DaaS – Deployment

- AWS
  - ArcGIS Enterprise on AWS deployment tools - AWS CloudFormation Templates or ArcGIS Enterprise Cloud Builder CLI - for a preconfigured enterprise geodatabase.
  - Esri Amazon Machine Image (AMI)
  - Create Geodatabase if:
    - Need to customize the name or location of the geodatabase or
    - Require multiple geodatabases
- Microsoft Azure
  - Use the ArcGIS Enterprise Cloud Builder to add a Microsoft Azure Database for PostgreSQL instance to your deployment.
  - Cloud Builder registers the database with your ArcGIS Server site.
  - Choose to enable a geodatabase or leave as a database.

Configuring Geodatabases

- SDE_DBTUNE
  - Mostly legacy and defaults work well
  - Specify the **PG_GEOMETRY** configuration keyword when creating feature classes you want to store in the PostGIS geometry type.

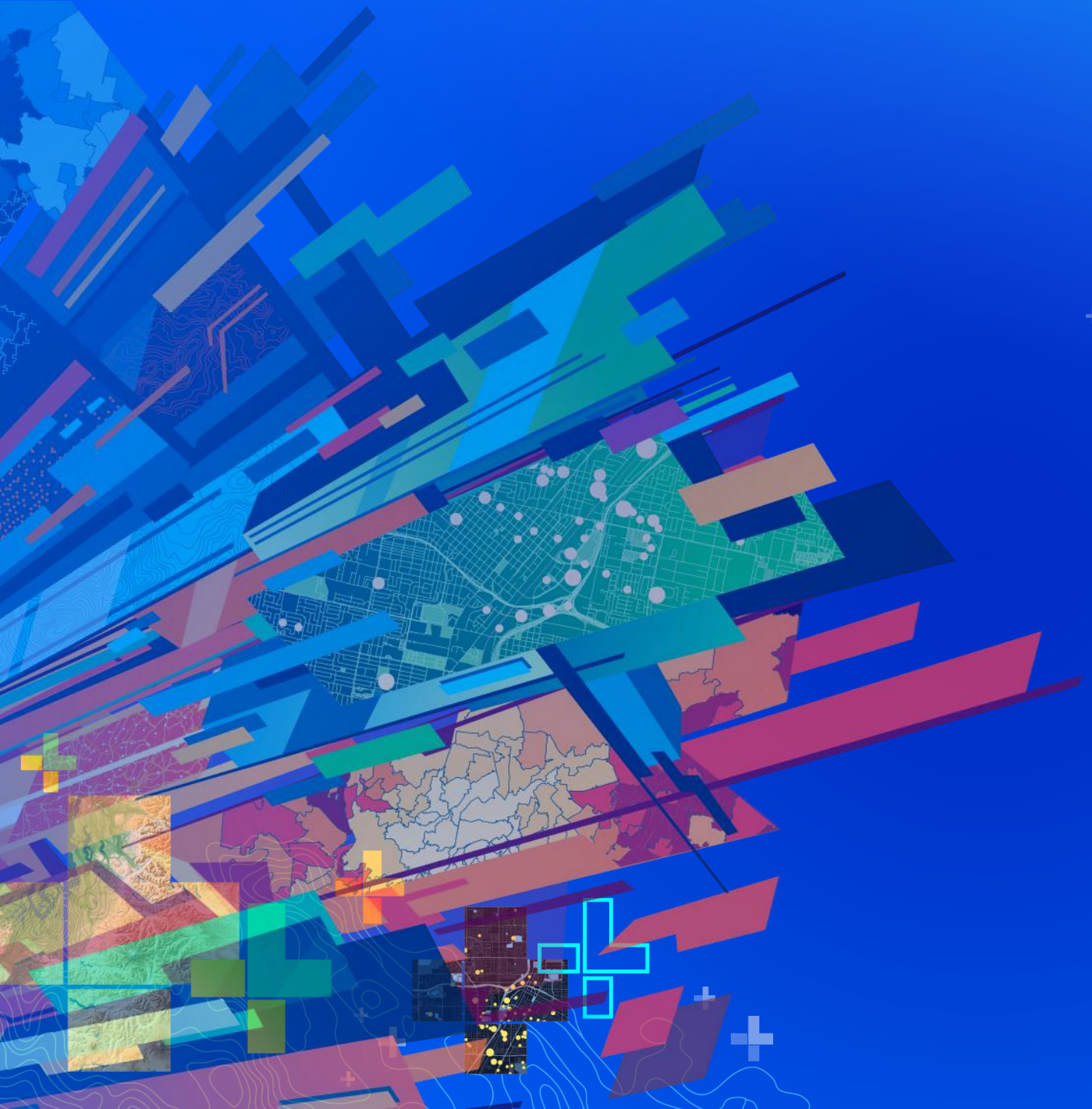
Case Sensitivity

- PostgreSQL is case sensitive
- It stores all object identifiers in lower case
  - Names of databases, schemas, users, tables, indexes, columns
  - Geodatabase also stores all identifiers in lower case in PostgreSQL
  - User data (attributes) can be in any case
- Do not use quotation marks to force different casing
  - ArcGIS does not quote strings; objects with upper or mixed case names will not be found
  - Be aware that pgAdmin quotes strings automatically

Upgrading a geodatabase

- The sde user requires superuser status
- Update the **st_geometry** library (on-premises only)
- When upgrading (ArcGIS, geodatabases, and PostgreSQL)
  - One at a time, step by step, incremental testing

PostgreSQL must be at supported release for geodatabase AND ArcGIS clients



PostgreSQL Configurations for ArcGIS

Neeta Reddy

Remember

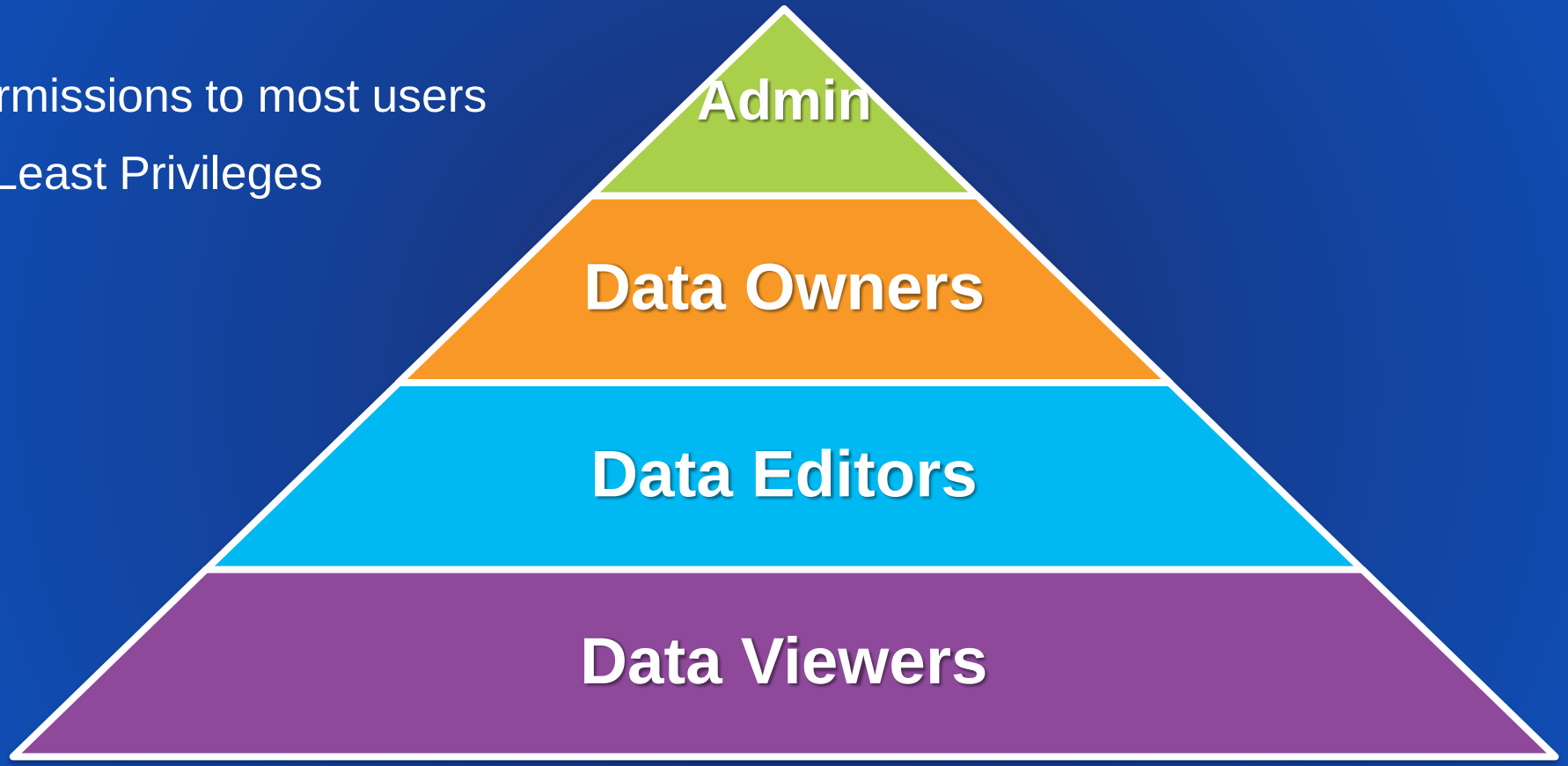
- Create Enterprise Geodatabase
  - Requires the postgres superuser credentials
  - Creates a database and required geodatabase objects
  - Useful for new databases
  - Used by dba and when database defaults are okay
- Enable Enterprise Geodatabase
  - Requires existing database and sde user and schema
  - Useful for databases with existing data
  - Used by geodatabase administrator and when custom database configuration is required

How do I...

Manage and maintain database security?

Users

- Limit permissions to most users
- Rule of Least Privileges

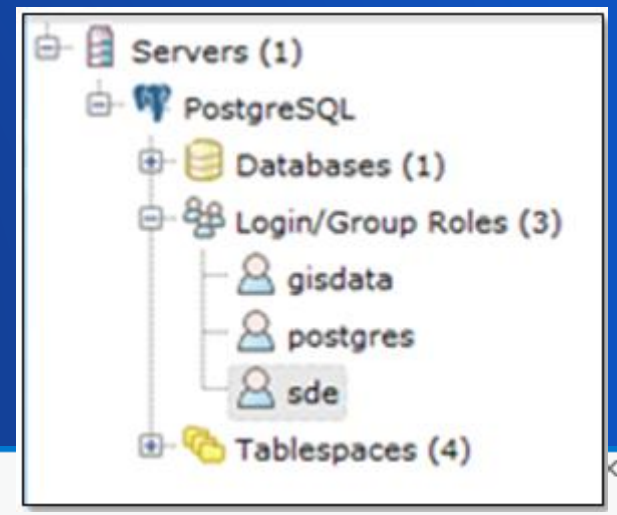


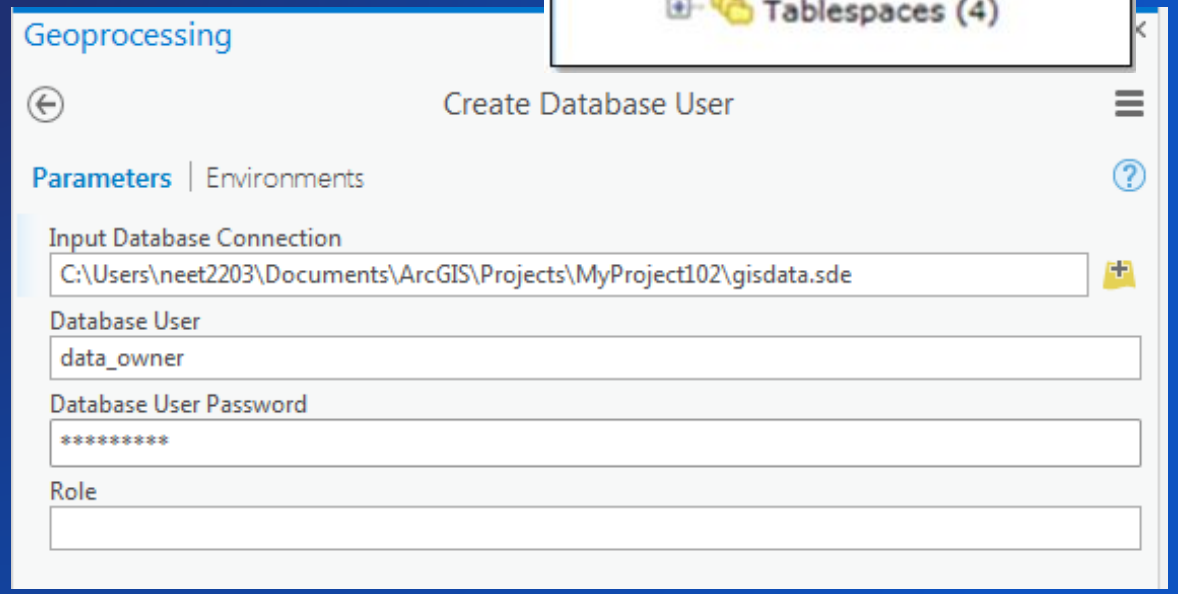
PostgreSQL Users - Roles

- Roles are used to log in to database cluster and databases
- Individual users called login roles
- Roles can also represent groups of users
- Create different types of users
- Data owner login roles must own a schema with same name
  - Data owner group role would still need individual data owner schemas

PostgreSQL Users - Roles

- Data owners
  - Can use Create Database User tool
    - Creates a login role
    - Creates a matching schema in the specified database only
    - Grants usage privileges on the new schema to public
      usage privilege can be restricted if required
- Data editors / viewers
  - Use SQL to create these users
  - Use Change Privileges tool to grant data access
    - Grants usage on owner schema
    - Grants privileges on datasets





BACKUP YOUR DATA NOW!!!!

(and practice restoring it)

Backup and Restore

- Backup methods
  - Database – pg_dump (typical method)
  - File based (cold) backup
  - VM backup

```
pg_dump -h localhost -p 5432 -U postgres  -F c -v -f E:\backups\prod_090616.bak prod
```

- Database restore - pg_restore
  - Restore schemas in this order: public, sde, data owners
  - If sde and remaining schemas are restored together, spatial indexes will be missing on user data

```
pg_restore -n public  -p 5432 -U postgres -d db_name –c -v E:\backups\db_060716.bak
```

Points to remember

- Backups are the only way to reliably protect your data
  1. How much time/data can you afford to lose?
  2. Create restore plan that will achieve #1
  3. Create backup plan that supports #2
  4. Implement your plans
  5. Test your recovery plan regularly by using real backup media to restore to a system capable of being used in production

Can’t restore from non-DaaS Geodatabase to DaaS Geodatabase

How do I...

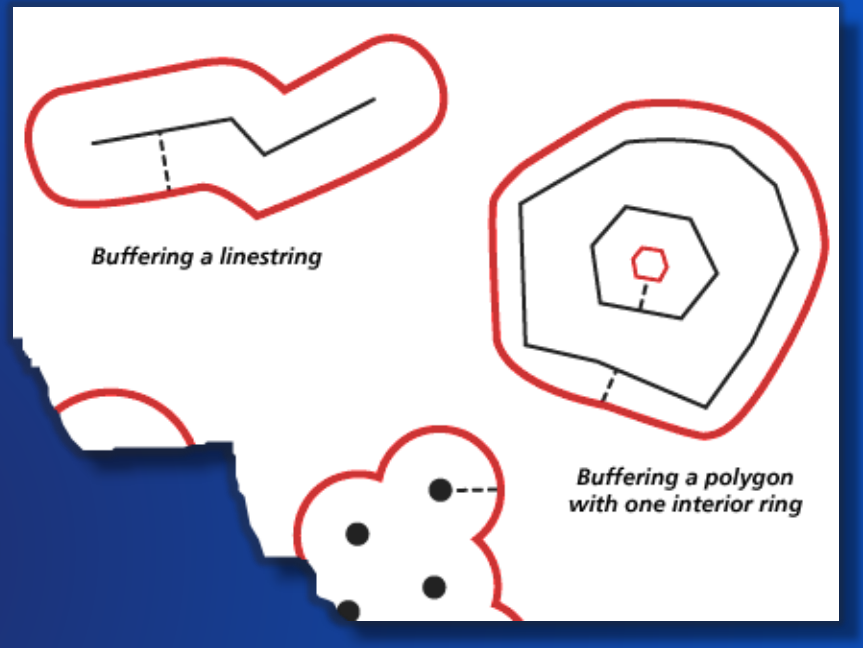
Store and manage spatial data?

Geodatabase Considerations

- Which spatial type to use
  - ST_Geometry (Esri default) or PostGIS (Geometry & Geography)
- Multiple geodatabases
  - One instance or multiple instances
- Workload separation
  - Editing architecture, security, SLA, administrative requirements
- Environment isolation
  - Development, staging, production
  - Refreshing/Moving data between environments

SQL Support

- Creation of Features
- Analysis through SQL
- Geodatabase behavior not supported
- SQL Types:
  - ST_Geometry (default)
  - PostGIS – geometry and geography
  - sde.st_intersects *(geometry1 sde.st_geometry, geometry2 sde.st_geometry)*



Data Types and Name Limits

- PostgreSQL has over 100 data types, ArcGIS supports a subset
  - PostgreSQL Data Types supported in ArcGIS
- Geodatabase name limits
  - Enterprise Geodatabase Size and Name Limits
- Important to review if migrating to PostgreSQL from another RDBMS

ArcGIS 10.3.x

	DB2	Informix	Oracle	PostgreSQL	SQL Server
Database name	8	31	30	31	31
Field (column) name	31	31	30	31	31
Index name	128	128	30	63	128
Password	31	31	31	31	31
Table or feature class alias	31	31	30	31	31
Table or feature class name	128	128	30	63	128
User or role name	30	31	30	31	31

ArcGIS 10.7.x

	Db2	Informix	Oracle	PostgreSQL	SQL Server
Field (column) name	31	31	30	31	31
Password	31	31	31	31	31
Geodatabase table or feature class alias	255	255	255	255	255
User or role name	30	31	30	31	31
Version name	62	62	62	62	62

Methods to Move Geodatabases

- Move data to another geodatabase on same PostgreSQL cluster
  - Geodatabase names must be unique on same cluster
  - Copy data into new Geodatabase in ArcGIS using copying and pasting, using the Export tool, using the Extract Data wizard, or using XML workspace documents.
- Move data to new PostgreSQL cluster
  - Use pg_dump command restoring schemas in correct order - Restore a geodatabase to PostgreSQL
  - If you use object identifiers (OIDs) in your user-defined data, be sure to specify the –o option to include the OIDs in the dump.
  - Re-create tablespaces and run ANALYZE after restoring
  - Connect as a superuser to run pg_dump.
  - **Methods to Move a PostgreSQL Geodatabase**

Good performance starts with good design

- Only store what you need to
  - Don’t use big fields for small values
- Only ask for what you need
  - Don’t query for fields you don’t need to see
- Add appropriate indexes
  - Joins
  - Frequently queried columns
- Feature datasets are not like folders
  - Control spatial relationships: networks, topologies etc

How do I maintain good performance?

- Standard maintenance
  - Traditional versioning – compress
  - Index fragmentation
  - Stale statistics
- Vacuum
  - Runs by default
  - Highly configurable
- Consider re-indexing
  - If high volume of change occurred to a table
- Memory allocation
  - The postgres.log will give suggestions

News

ArcGIS and PostgreSQL changes since last year

What’s new in ArcGIS

- ArcGIS supports installing the PostGIS extension in any schema
- Feature binning
- Support for PostgreSQL 11
- Microsoft Visual C++ 2017 Redistributable needs to be installed: PostgreSQL database requirements
- Shape Validation
  - At 10.6.1 - Minimum shape validation for PostGIS spatial types thereby improving ArcGIS draw/query performance.

Supported PostgreSQL versions

ArcGIS 10.3.x

9.1.4 PostGIS 2.0
9.2.9 PostGIS 2.0
9.3.5 PostGIS 2.1

ArcGIS 10.4.x

9.2.14 PostGIS 2.1
9.3.10 PostGIS 2.2
9.4.5 PostGIS 2.2

ArcGIS 10.5.x

9.3.13 PostGIS 2.2
9.4.8 PostGIS 2.2
9.5.3 PostGIS 2.2

ArcGIS 10.6.x

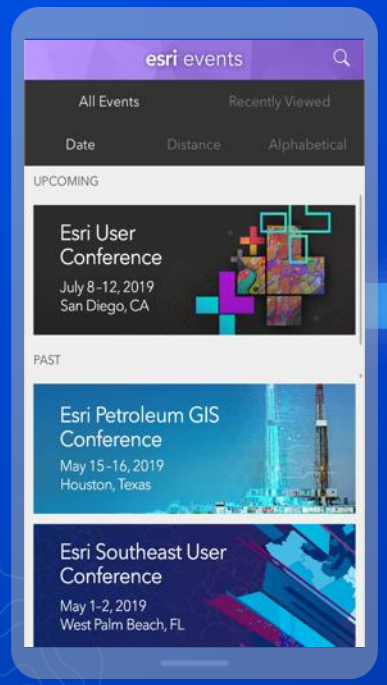
9.4.8 PostGIS 2.2 - *only 10.6.0 Pro 2.1*
9.5.12 PostGIS 2.2
9.6.8 PostGIS 2.3
10.3 PostGIS 2.4 - *only 10.6.1 Pro 2.2*

ArcGIS 10.7.x

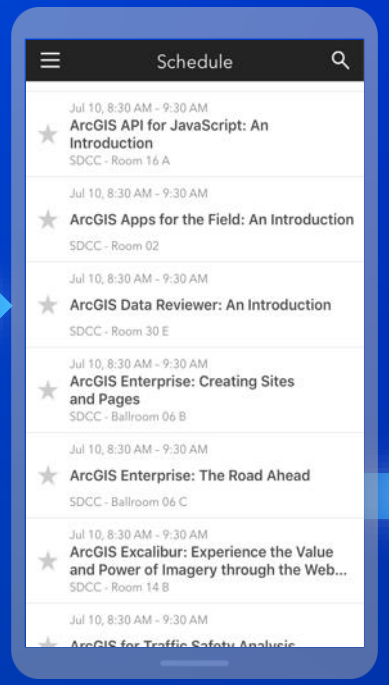
9.5.12 PostGIS 2.2 - *only 10.7.0 Pro 2.3*
9.6.12 PostGIS 2.3
10.7 PostGIS 2.4
11.2 PostGIS 2.5.1 - *only 10.7.1 Pro 2.4*

Please Share Your Feedback in the App

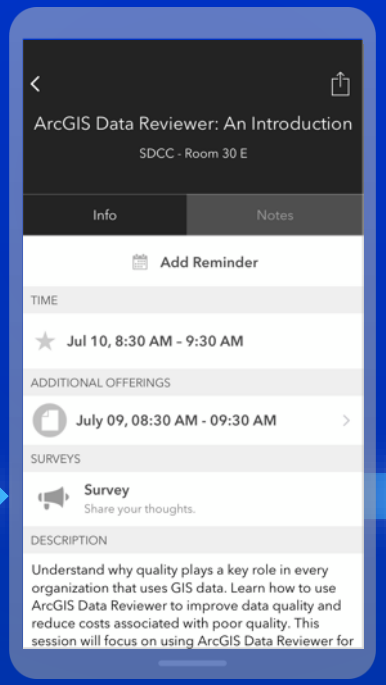
Download the Esri Events app and find your event



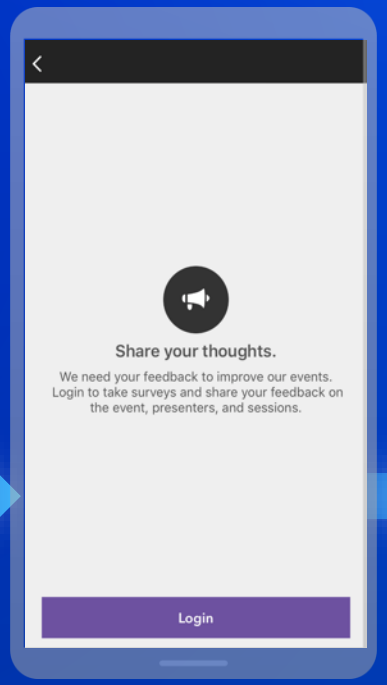
Select the session you attended



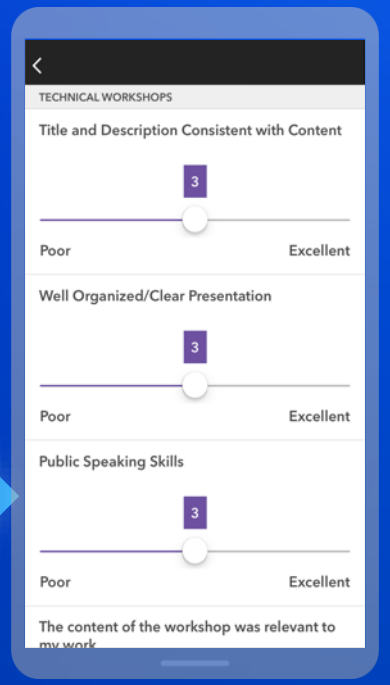
Scroll down to “Survey”



Log in to access the survey



Complete the survey and select “Submit”



Thank you

Questions?


esri
THE
SCIENCE
OF
WHERE