# SDE Performance Tips - Part 2

## SDE Spatial Index Tuning

# Introduction

- Welcome!
- Why this workshop?
- Basic terms
- Workshop goals

# Outline (1)

- Introduction
- SDE basics
- Table and index organization issues
- SDE grid indexing algorithm
- Layer I/O modes

# Outline (2)

- Performance testing guidelines and procedures
- Feature table statistics
- Spatial table statistics
- Case studies

# SDE basics

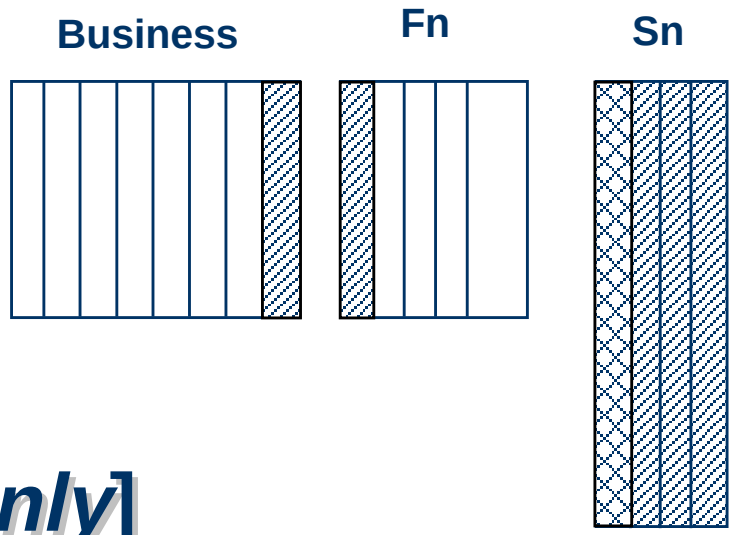
- What is SDE?
  - An enabling technology to implement spatial data within relational databases.
  - A client/server protocol for accessing vector data and attributes.
  - A tool for quickly locating small numbers of spatially distributed features.

# SDE basics

- What SDE is NOT:
  - GIS software
  - A graphics accelerator
  - A graphical *anything*

# SDE basics

- SDE tables
  - Business (A)
  - Feature (F)
  - Spatial Index (S)
  - Points (P) [*Sybase only*]



# Table and index organization issues

- Put the business, Fn, and Sn tables on different disk drives (*Pn, too*).
- Separate the table and indexes so they are on different disk drives.
- Beware of defeating striping!

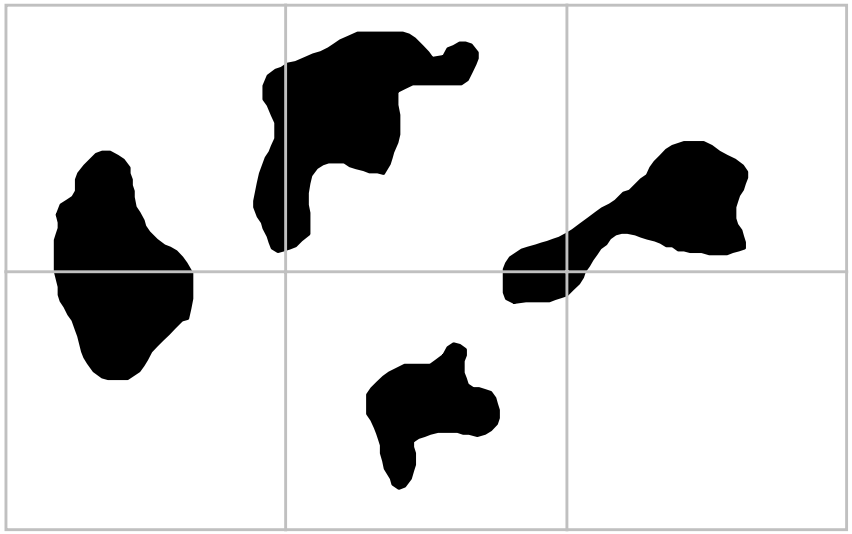
# Table and index organization issues

- Build indexes on likely attribute combinations.
- Add spatial column to index list if it will result in “covered” query.
- Indexes speed up queries, but slow down INSERT, UPDATE, and DELETE operations.

# SDE grid indexing algorithm

- Not applicable to normalized layers
- Utilizes KISS principle

# The spatial index (S table)

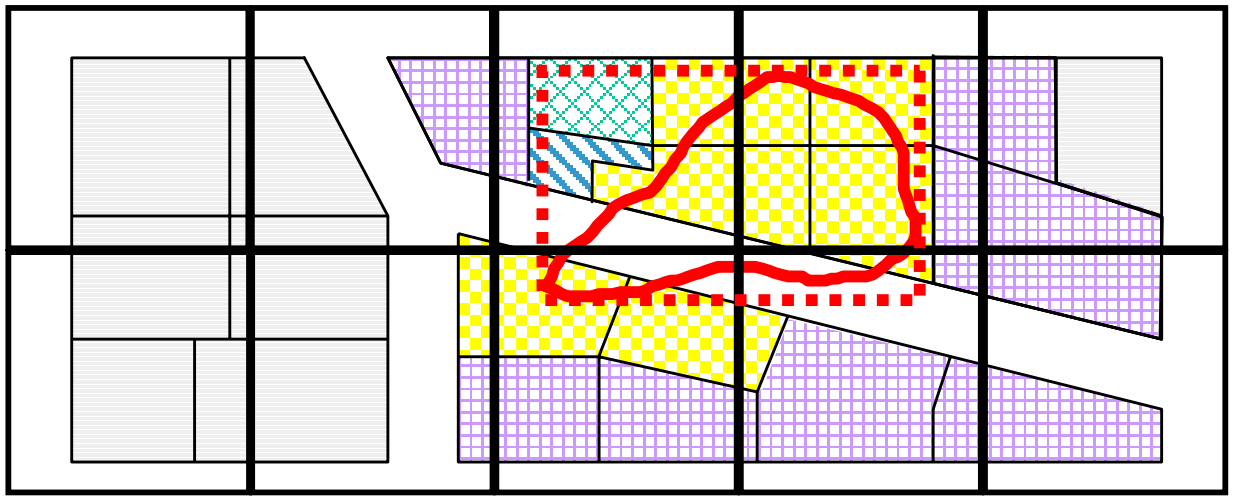


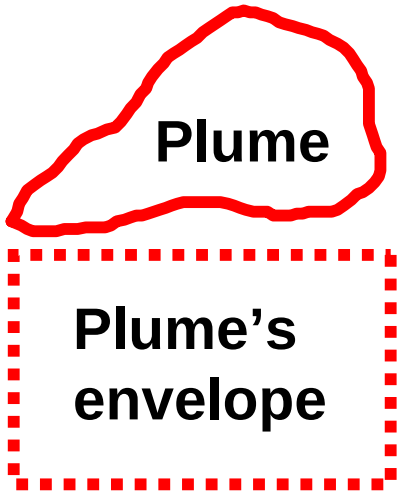
- A regularly-spaced square indexing grid
  - Each feature exists in one or more grids
  - Each grid may have multiple features

# The spatial index (S table) - Cont.

- Features are not split by grids or stored by grid
  - Grids are just used for fast envelope searches
- A spatial index is like a two-dimensional column index

# How the spatial index works






1. Features rejected by envelope search of plume's envelope against spatial index grid
2. Features rejected by envelope search of plume's envelope against individual feature envelopes in spatial index table
3. Features rejected by comparing the plume itself to the feature envelopes in the spatial index table
4. Features rejected by feature-to-feature overlap testing of plume against parcels from the feature table
5. Features selected by server and streamed to client

# Up to three spatial index grids

- Most layers have only one spatial index grid
  - Each grid requires a separate index search
  - Multiple grids are usually slower—try to use only one
  - Use when features are vastly different in size. This avoids needing huge numbers of grid cells to cover a large feature
  - SDE will not allow more than 1,000 cells/feature

# Up to three spatial index grids - Cont.

- If a feature covers more than four grid cells, it is promoted to the next larger grid
- Use *sdelayer -o si_stats...* to see statistics on:
  - Number of features per grid cell
  - Number of grid cells per feature

# Up to three spatial index grids - Cont.

- Avoid high numbers of cells per feature, while tuning cell size to approximate the average query window
- Use large cell sizes with point layers.
  - Tiny cell sizes result in slow performance
  - Since the envelope of a point, IS the point, searches are fast

# Spatial index layout

| NAME | DATA TYPE | NULL? |
|---|---|---|
| sp_fid | SE_INTEGER | NOT NULL |
| gx | SE_INTEGER | NOT NULL |
| gy | SE_INTEGER | NOT NULL |
| eminx | SE_INTEGER | NOT NULL |
| eminy | SE_INTEGER | NOT NULL |
| emaxx | SE_INTEGER | NOT NULL |
| emaxy | SE_INTEGER | NOT NULL |

- sp_fid is the feature ID (FID)
  - The FID joins the spatial index to the feature table and business table
- gx and gy identify the cell's row and column
  - Two bits are reserved as flags to indicate whether this row contains a level 1, 2, or 3 size index grid cell
- eminx, eminy, emaxx, emaxy are the feature envelope

# RDBMS indexes on the spatial index

- SDE 3.0
  - S< layer_id>_IX1 - sp_fid
  - S< layer_id>_IX2 - gx, gy
- SDE 3.0.1
  - S< layer_id>_IX1 - gx, gy, eminx, eminy, emaxx, emaxy, sp_fid
- SDE 3.0.2
  - S< layer_id>_IX1 - gx, gy, eminx, eminy, emaxx, emaxy, sp_fid
  - S< layer_id>_IX2 - sp_fid

# Layer I/O modes

- There are two distinct phases of database access, loading and query.
- Most RDBMS vendors recommend different tuning configurations depending on the phase.

# Layer I/O modes

- SDE provides two layer access modes:
  - Normal I/O mode
  - Load-only I/O mode
    - Sn table truncated (Sn_IX* dropped)
    - An_IX1 dropped
    - Fn_IX1 retained!

# Layer I/O modes

- The commands to change I/O mode are:
  - sdelayer -o normal_io
  - sdelayer -o load_only_io
- Both shp2sde and cov2sde with -o create will first place a layer in load-only I/O mode before inserting rows.

# Layer I/O modes

- A layer in LOAD_ONLY_IO mode
  - Cannot be queried
  - Accepts data much faster (~10x)
- A layer cannot be put back into normal I/O mode if the spatial index parameters or layer keyword prevent spatial index creation.

# Layer I/O modes

- You can't test the performance of a layer in LOAD_ONLY mode.
- The first time you load a layer:
  - Use HUGE grid size
  - Use HUGE keyword parameters
  - Make measurements
  - Load it again

# Performance testing guidelines and procedures

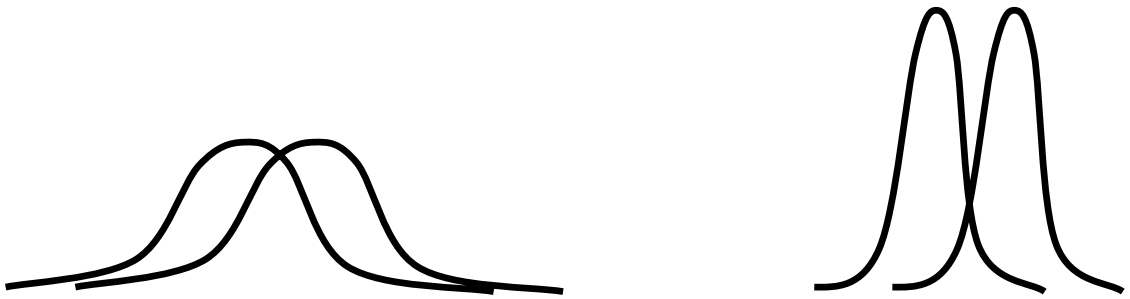
- Measurement & Sampling
  - There are many variables in measuring SDE performance:
    - CPU load
    - Network load
    - I/O load
    - Configured RAM
    - DBMS software
    - DBMS “warmth”

# Performance testing guidelines and procedures

- Measurement & Sampling
  - The only way to be sure that the metrics you capture are meaningful is to take multiple samples
    - A single measure ignores the effect of caching on the production environment.
    - Samples provide access to statistics

# Performance testing guidelines and procedures

- Sampling statistics
  - Sample size is important, but don't overdo it!
  - Pay attention to variance and standard deviation.



# Performance testing guidelines and procedures

- Always perform measurements on a representative sample
- Prototype against 10-25% of the final feature count.

# Performance testing guidelines and procedures

- Limits on testing
  - Query performance measurements are only valid for comparison on a single system with a single data use, on a single layer,...
  - Changing more than one variable reduces the value of the results.
  - Results are only meaningful for the tested use.

# Feature table statistics

- Some useful measures for feature tables include:
  - Average number of vertices
  - Average shape storage compression
  - Average feature size
  - Average envelope size
  - Min & max of the above

# Feature table statistics

- SDE utilizes compression on shape geometry
  - "Small" features (envelope) compress better than "large" ones
  - "Large" features (vertices) compress better than "small" ones
  - Specifying a precision very much smaller than the accuracy wastes disk.

# Spatial table statistics

- Some useful metrics for spatial index tables include:
  - Average grids per feature
  - Average features per grid
  - Number of features in more than four grids

# Spatial table statistics

- The sdelayer -o si_stats report

```
Level 1,   Grid Size 5000
-------------------------------------------------------------------
  Grid Records: 9117
  Feature Records: 67
  Grids/Feature Ratio: 136.07
  Avg. Features per Grid:   4.78
  Max. Features per Grid: 14
  % of Features Wholly Inside 1 Grid:   0.00
-------------------------------------------------------------------
                Spatial Index Record Count By Group
  Grids:      <=4    >4    >10    >25    >50    >100   >250   >500
 ---------- ------ ------ ------ ------ ------ ------ ------ ------
  Features:      0     67     67     67     64     48      3      0
  % Total:       0%   100%   100%   100%    96%    72%     4%     0%
-------------------------------------------------------------------
```

# Spatial table statistics

- The sdelayer -o si_stats report

```
Level 1,   Grid Size 5000
-------------------------------------------------------------------
  Grid Records: 9117
  Feature Records: 67
  Grids/Feature Ratio: 136.07
  Avg. Features per Grid:   4.78
  Max. Features per Grid: 14
  % of Features Wholly Inside 1 Grid:   0.00
-------------------------------------------------------------------
                Spatial Index Record Count By Group
  Grids:      <=4    >4    >10    >25    >50    >100   >250   >500
 ---------- ------ ------ ------ ------ ------ ------ ------ ------
  Features:      0     67     67     67     64     48      3      0
  % Total:       0%   100%   100%   100%    96%    72%     4%     0%
-------------------------------------------------------------------
```

# Spatial table statistics

- The sdelayer -o si_stats report

```
Level 1,   Grid Size 5000
-------------------------------------------------------------------
  Grid Records: 9117
  Feature Records: 67
  Grids/Feature Ratio: 136.07
  Avg. Features per Grid:   4.78
  Max. Features per Grid: 14
  % of Features Wholly Inside 1 Grid:   0.00
-------------------------------------------------------------------
                Spatial Index Record Count By Group
  Grids:      <=4    >4    >10    >25    >50    >100   >250   >500
 ---------- ------ ------ ------ ------ ------ ------ ------ ------
  Features:      0     67     67     67     64     48      3      0
  % Total:       0%   100%   100%   100%    96%    72%     4%     0%
-------------------------------------------------------------------
```

# Spatial table statistics

- The sdelayer -o si_stats report

```
Level 1,   Grid Size 5000
-------------------------------------------------------------------
  Grid Records: 9117
  Feature Records: 67
  Grids/Feature Ratio: 136.07
  Avg. Features per Grid:   4.78
  Max. Features per Grid: 14
  % of Features Wholly Inside 1 Grid:   0.00
-------------------------------------------------------------------
                Spatial Index Record Count By Group
  Grids:      <=4    >4    >10    >25    >50    >100   >250   >500
 ---------- ------ ------ ------ ------ ------ ------ ------ ------
  Features:      0     67     67     67     64     48      3      0
  % Total:       0%   100%   100%   100%    96%    72%     4%     0%
-------------------------------------------------------------------
```

# Spatial table statistics

- **The sdelayer -o si_stats report**

```
Level 1,   Grid Size 5000
-------------------------------------------------------------------
  Grid Records: 9117
  Feature Records: 67
  Grids/Feature Ratio: 136.07
  Avg. Features per Grid:   4.78
  Max. Features per Grid: 14
  % of Features Wholly Inside 1 Grid:   0.00
-------------------------------------------------------------------
                Spatial Index Record Count By Group
Grids:      <=4    >4    >10    >25    >50    >100   >250   >500
--------- ------ ------ ------ ------ ------ ------ ------ ------
Features:      0     67     67     67     64     48      3      0
% Total:       0%   100%   100%   100%    96%    72%     4%     0%
-------------------------------------------------------------------
```

# Spatial table statistics

- The sdelayer -o si_stats report

```
Level 1,   Grid Size 500000
-------------------------------------------------------------------
  Grid Records: 74
  Feature Records: 67
  Grids/Feature Ratio:   1.10
  Avg. Features per Grid:  37.00
  Max. Features per Grid: 42
  % of Features Wholly Inside 1 Grid:  89.55
-------------------------------------------------------------------
                Spatial Index Record Count By Group
  Grids:      <=4    >4    >10    >25    >50    >100   >250   >500
 ---------- ------ ------ ------ ------ ------ ------ ------ ------
  Features:     67      0      0      0      0      0      0      0
  % Total:     100%     0%     0%     0%     0%     0%     0%     0%
-------------------------------------------------------------------
```

# Spatial table statistics

- The sdelayer -o si_stats report

```
Level 1,   Grid Size 50000
-------------------------------------------------------------------
  Grid Records: 273
  Feature Records: 67
  Grids/Feature Ratio:   4.07
  Avg. Features per Grid:   4.79
  Max. Features per Grid: 14
  % of Features Wholly Inside 1 Grid:   1.49
-------------------------------------------------------------------
                Spatial Index Record Count By Group
  Grids:      <=4    >4    >10    >25    >50    >100   >250   >500
 ---------- ------ ------ ------ ------ ------ ------ ------ ------
  Features:     52     15      0      0      0      0      0      0
  % Total:      78%    22%     0%     0%     0%     0%     0%     0%
-------------------------------------------------------------------
```

# Spatial table statistics

- Conflicting inclinations:
  - Make the grid
    - BIG
    - SMALL
  - Index on the
    - QUERY
    - FEATURE

# Case studies

- “We didn’t change nothin’.”
- Query in the right order
- Time series can be trouble

# Case studies

- Lessons learned
  - Restart the software after every configuration change.
  - Run ANALYZE ... DELETE STATISTICS.
  - Organize physical data records in spatial index order.
  - Upgrade to 3.0.2 the moment it arrives.

# Case studies

- Free utilities
  - sdequery
  - sdestats
  - sdeanalyze
  - sdesort
  - asc2sde

# Parting advice

- Be ready with details when you call for help:
  - OS Version
  - SDE Version
  - DBMS Version
  - Synopsis of data & use
  - Exact error code(s)/messages

# Questions?

- Here & Now
- SDE Island / Doctor's Office

Geography Matters

# UC98

San Diego, California

ESRI

EIGHTEENTH ANNUAL INTERNATIONAL USER CONFERENCE