ArcSDE Performance Tips

Welcome

- Mansour Raad (mraad@esri.com)
- Rent-a-Tech
  - Design/Implement Enterprise Solutions
- ARCIMS
  - Developer
  - Technology Integrator

Outline

- Quick “what is ArcSDE ?”
- False X/Y and System Units.
- ArcSDE grid indexing algorithm.
- Table and index organization issues.

Outline

- Virtual Layers.
- Partitioning Very Large Datasets.
- ArcSDE Shape Table.
- ArcSDE for the World Wide Web.
- WebSDE demo.

Performance Tuning

- **Every System is unique.**
  - **CPU**
  - **Memory**
  - **Disk drives**
  - **Network**
  - **OS**
  - **Data**

Performance Tuning

- Few Magic Bullets with Significant Impact.
- Ramping process.
  - Shorten this ramp.
- Constant Over Time.

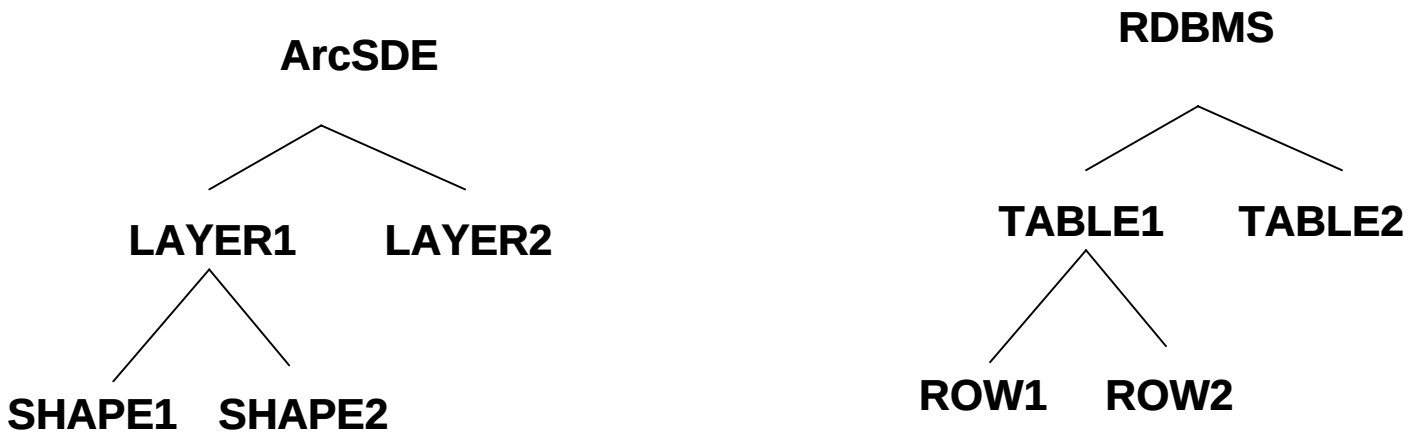
What Is ArcSDE ?

- Client/Server Architecture.
- Manages Very Large Seamless Database.
- Performs Spatial and Non-Spatial Searches.
- Has a Computational Geometry Library.
- Has a Projection Library.
- Inherits RDBMS qualities. (where applicable)
- “C” API.

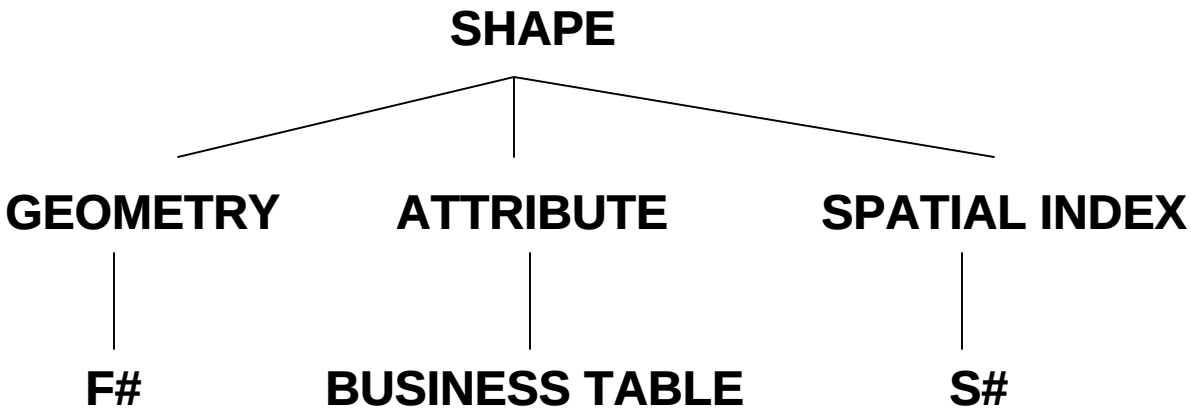
ArcSDE Is Not...

- Graphical Toolkit.
- Graphics Accelerator.
- Transaction Processor Middleware.

ArcSDE Data Layout (Logical)



ArcSDE Data Layout (Physical)



 # = SELECT LAYER_ID FROM SDE.LAYERS WHERE TABLE_NAME='<BUSINESS-TABLE-NAME>'

False X/Y System Units

- SDE converts and stores data in positive Integer format.
  - Greater or equal to zero
  - Less than 2^{31}
- Data is compressed using relative values into blobs in F Tables.
- Computational Geometry efficient.

False X/Y System Units

- Conversion is as follows:
  - SDE x = int (SU * (World x - False x))
  - SDE y = int (SU * (World y - False y))
- Choose False x/y and SU to preserve resolution of your data
- Keep the math “mentally-doable”

False X/Y System Units Example

- Given
  - World x = 123.45
- Having
  - False X = 100.0
  - SU = 100

False X/Y System Units

- Going Forward
  - int (SU * (World X - False X)
  - int(100.0 * (123.45 - 100.0)) = 2345
- Going Backward
  - SDE X / SU + False X
  - 2345 / 100.0 + 100.0 = 123.45

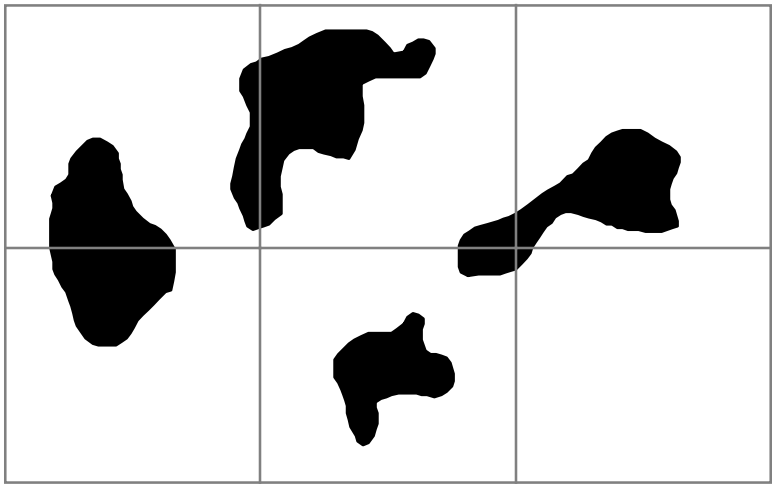
So What Values Should I Use Mansour ?

- Choose False X/Y to be the smallest values in your layer.
- Round Down to a “nice” number.
- Choose System Unit to be the number of decimals that you want to preserve.
- Use the same values for ALL the layers in your dataset.

False X/Y System Units

- For Layers With Latitude And Longitude Values, Choose:
  - FalseX = -180.0
  - FalseY = -90.0
  - System Unit = 1000000
- shp2sde example,
  - o create ….
  - x -180,-90,1000000

The Spatial Index (S table)

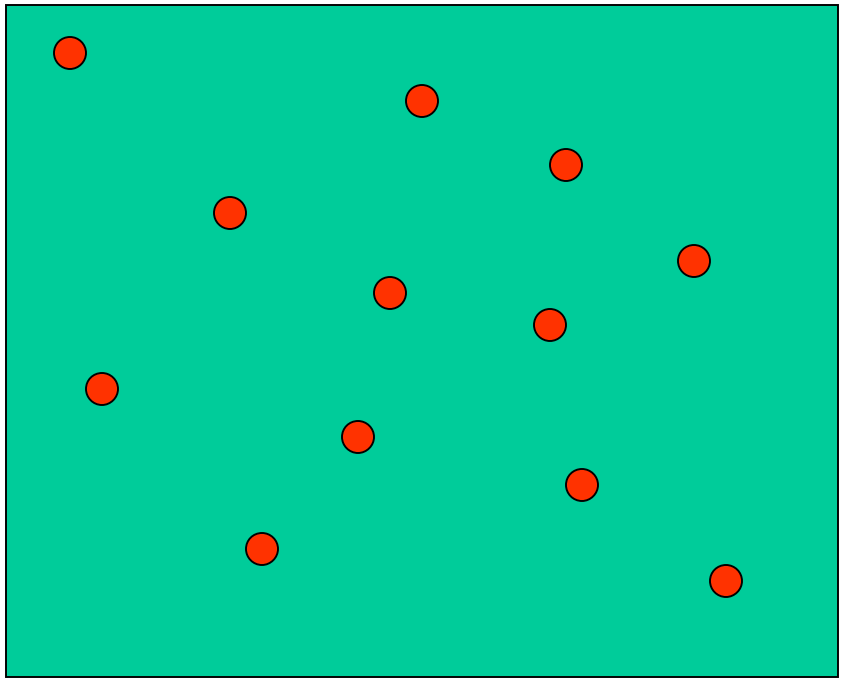


- A regularly-spaced square indexing grid
  - Each feature exists in one or more grids
  - Each grid may have multiple features

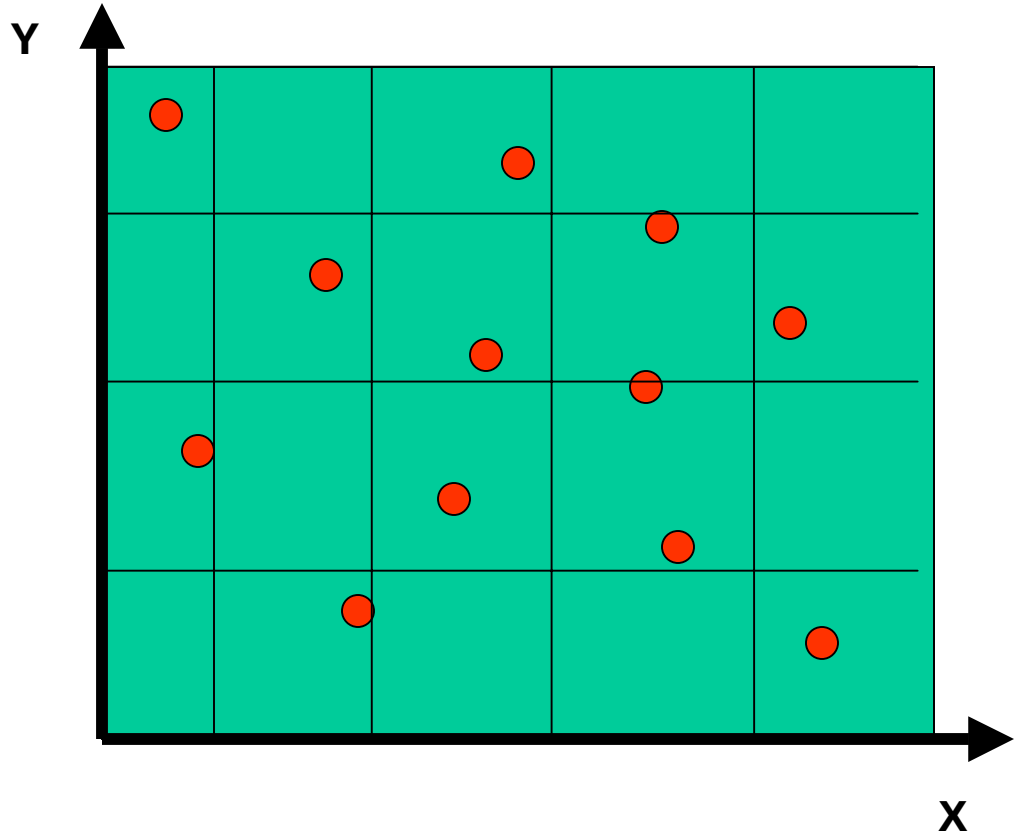
The Spatial Index (Cont.)

- 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

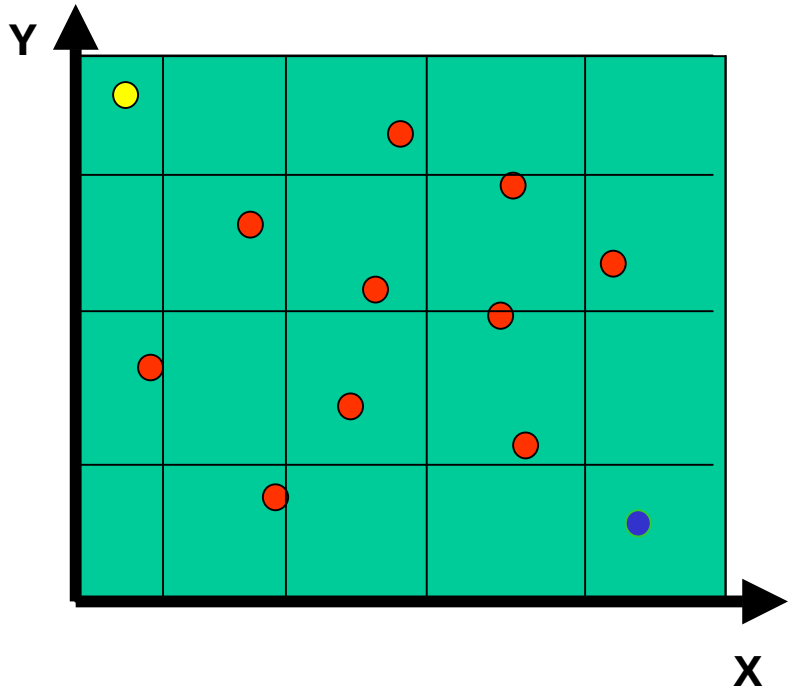
The Spatial Index (cont.)



The Spatial Index (cont.)



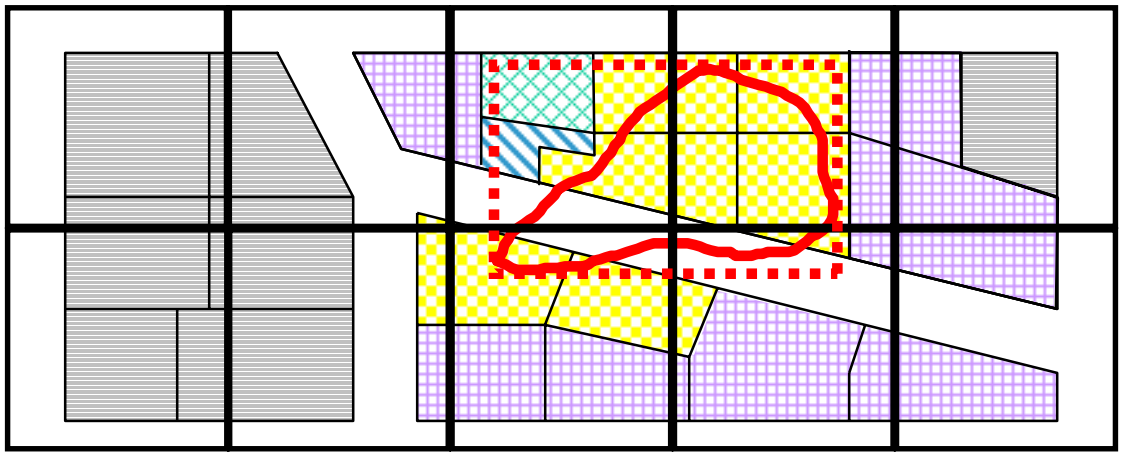
The Spatial Index (cont.)

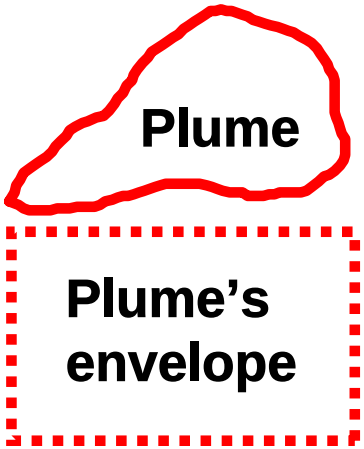


STABLE

FID	GX	GY
●	0	3
●	4	0

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

- If a feature covers more than four grid cells, it is promoted to the next larger grid.
- Most layers need only one spatial index grid.
- Each grid requires a separate index search.
- Multiple grids are usually slower—try to use only one per search.
- SDE will not allow more than 1,000 cells/feature

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

- S< layer_id>_IX1
  - gx, gy, eminx, eminy, emaxx, emaxy, sp_fid
  - Compound index to reduce IO by half
- S< layer_id>_IX2
  - sp_fid

What Should The Cell Size Be Mansour ?

- Very Simple Rule.
- 9 Grid Cells/Average Viewing Area.
- Use AV to measure average extent and divide by 3.
- Reduce the number of Stable scans.

What Should The Cell Size Be Mansour ?

- Create Higher Grid Levels Only If more than 20% of data has more than 4 grid cells.
- This is not absolute!
- Use sdelayer -o si_stats.

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 Index Order

- Load the data in Spatial Index Order
- Object that are close to each other in the real world, put then close to each other on the disk.
- Rtree Or Hilbert data reload.
- Siorder.exe will be available on ftp

$SDEHOME/etc/dbtune

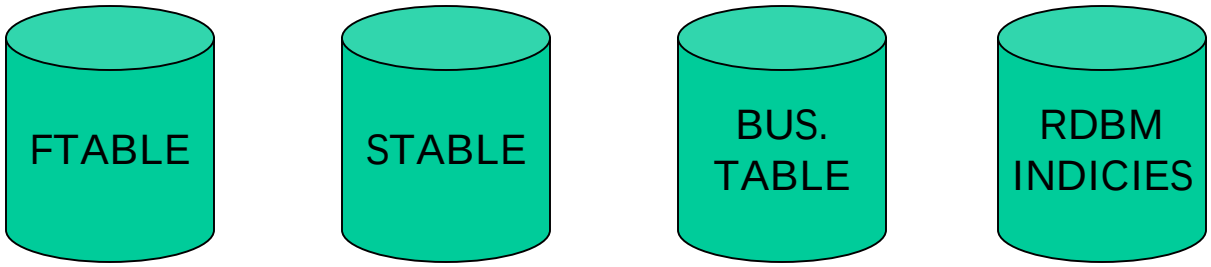
- Create 8 or 16 K Data Block Size.
- Create HUGE Entry.
- For Oracle users with Read-Only Data
  - Set Percent Increase to Zero.
  - Set Percent Used to 95.
  - Set Percent Free to 1 or Zero.

Double Load Your Data

- Load Data using HUGE dbtune entry.
- Analyze Newly Loaded Data.
  - Extents Count and Fragmentation.
- Create new entry in dbtune.
- Drop the layer.
- Reload Data With New Entry.

Disk Layout

- IO is the slowest part of your system.
- Put each table type in its own spaces (tablespaces/datafile in Oracle lingo)
- For theme on theme analysis, create TWO tablespace types.



Oracle Optimal Configuration

- DISK 0 - Oracle/App Software
- DISK 1 - SYSTEM, Control File 1
- DISK 2 - RBS, TEMP, Control File 2
- DISK 3 - REDO 1,2,3, Export Files
- DISK4 - Feature Data (F# tables)
- DISK5 - Spatial Index Data (S# tables)
- DISK6 - Attribute Data (Business tables)
- DISK7 - Oracle Indexes

RDBMS Indexes

- Ask The Users About The Data Fields That They Will Need.
- Monitor SQL Queries.

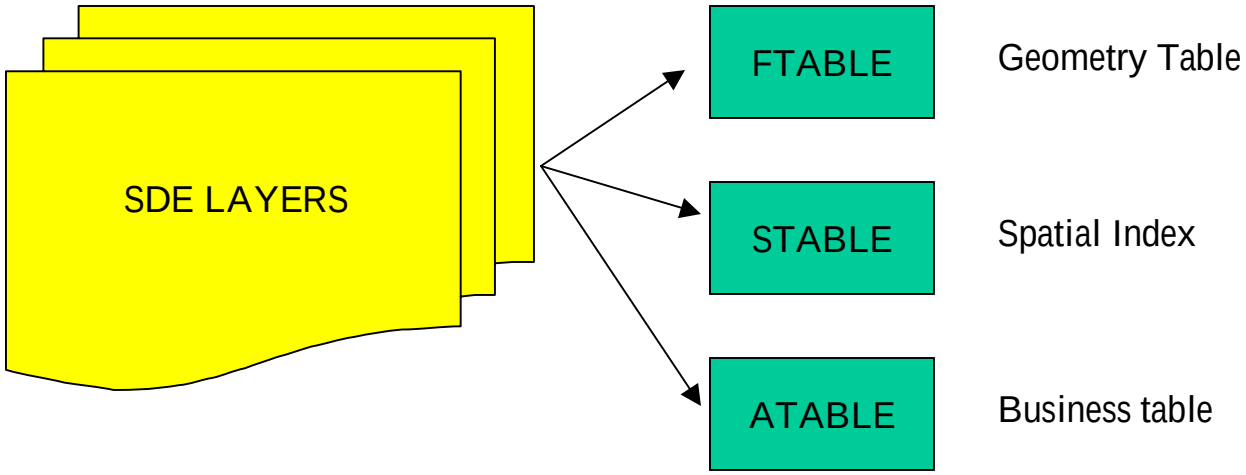
RDBMS Indexes

- Create Compound Indexes where Primary Key is the field in demand followed by the shape column.
  - CREATE INDEX NAME_IX2 (NAME,SHAPE) ON STREETS;
- Put indexes in own tablespace.

Virtual Layers

- One Layer of Data With Different Views.
- Views Are Very Different in Sizes.
- Different “Sweet Spot” Grid Sizes.
- Do Not Want To Duplicate Data.
- Performance issue.

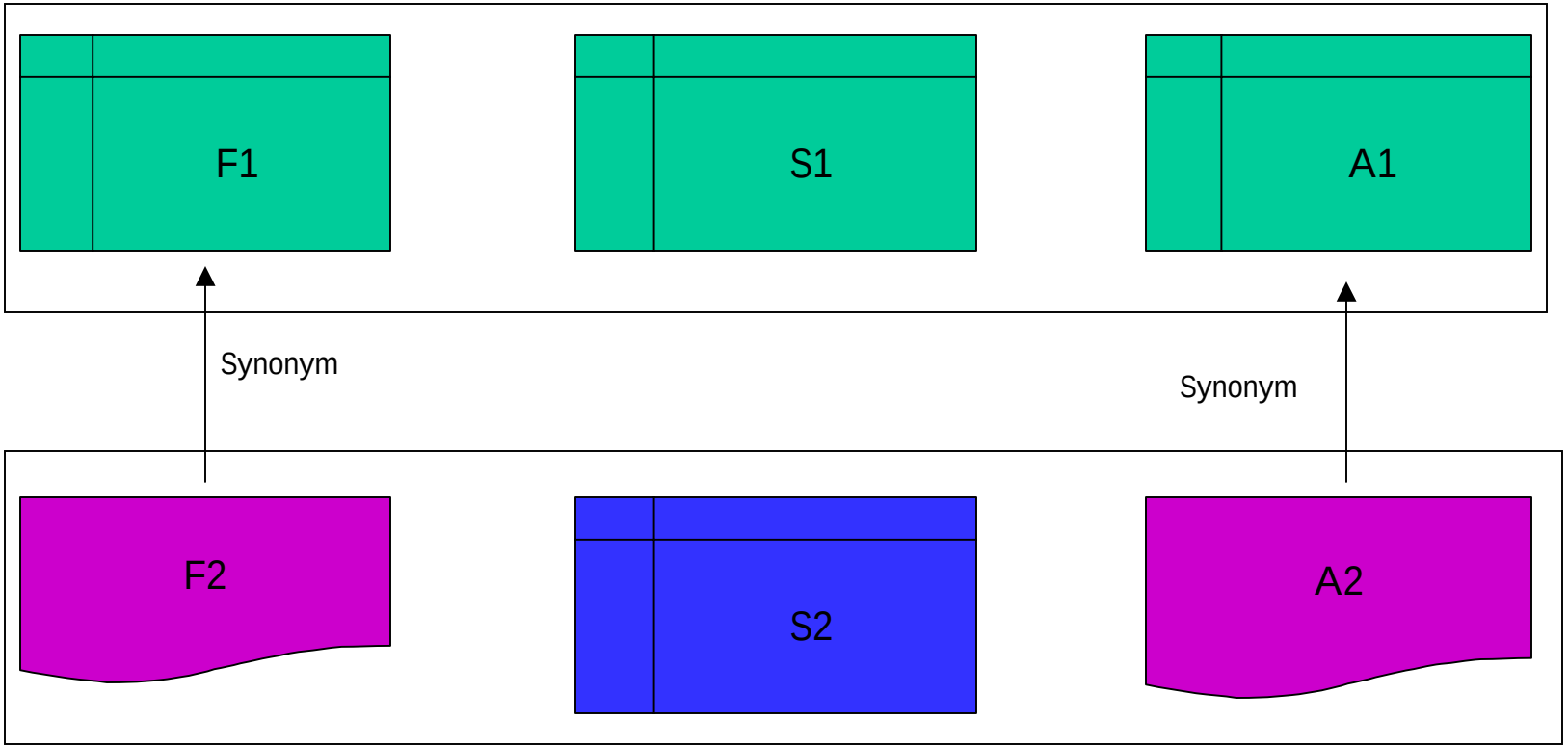
Virtual Layers



How Do I Do It ?

- Create “dummy” layer using sdelayer.
- Put layer in load IO mode.
- Using “sqlplus” drop new Ftable and Atable
- Create a synonym for the original tables as the new Ftable and the new Atable.
- Data and Geometry are the same.
- Spatially indexed differently.
- User accesses layer based on view extent.

Virtual Layers



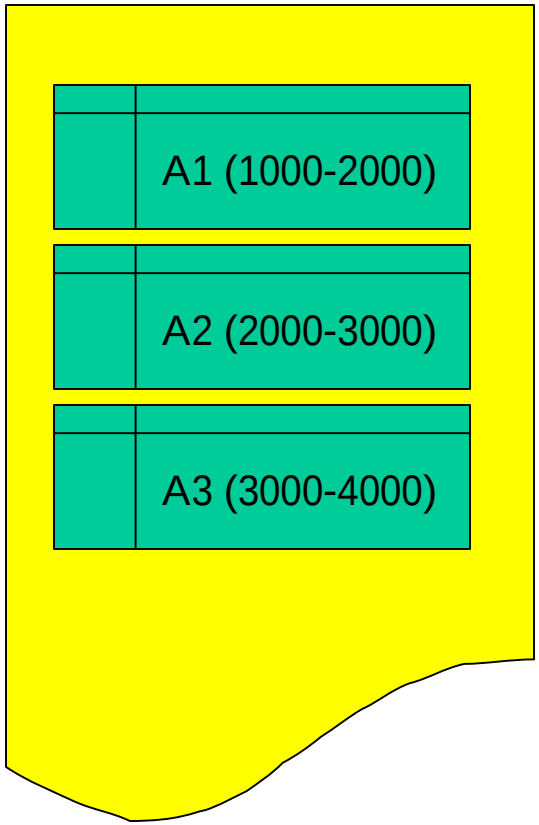
Partitioning Very Large Data

- TRUE, SDE enables seamless dataset.
- For sanity and maintenance sake, partition the layer into smaller layers.
- Create a “master” layer that “points” to the partitioned layer.
- Oracle 8.x partition scheme (no blobs)

Oracle 8 Partition

- Instead of a table of a 10 million rows.
- Create a view as “select union all” of 10 one million row tables.
- Create range of fids for each layer in the view.

SDE Partition




```
CREATE VIEW A100 AS
SELECT * FROM A1 UNION ALL
SEELCT * FROM A2 UNION ALL
SELECT * FROM A3…..
```

SDE Shape-Table

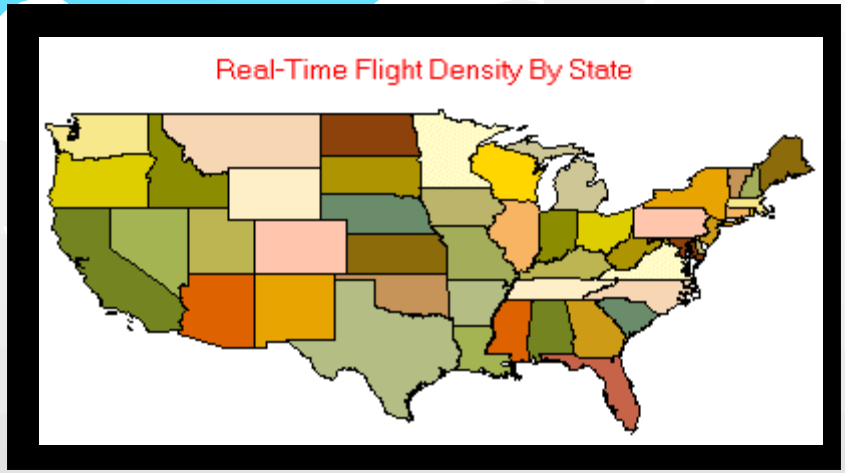
- "C" API Functions.
- CLIENT side operation.
- In memory collection of spatially enabled features.
- Each feature can be associated with a User-Defined Data.
- All the Spatial Search Methods are Applicable.
- Very Fast.

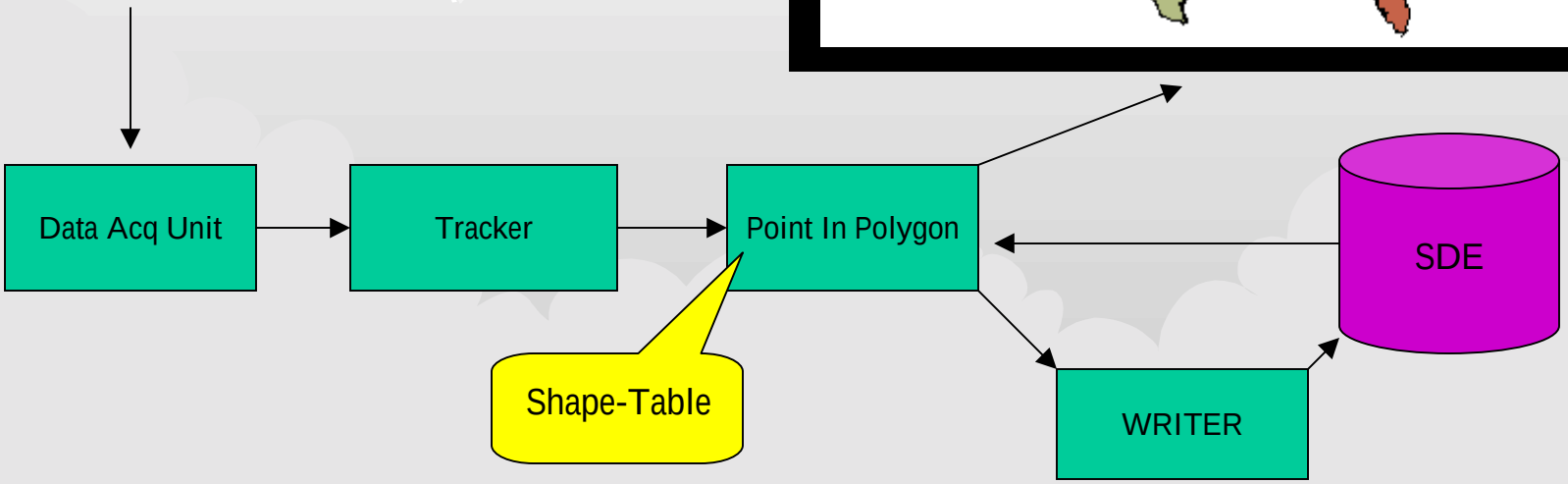
SDE Shape-Table Usage

- High volume of spatial requests on a layer.
- Reduction of Disk I/O
- Reduction of Network Traffic.
- Construction of user-defined spatial objects that need to spatially enabled.
- Real-Time Tracking with Spatial Properties.

Shape-Table Application

ACID,Latitude,Longitude,Altitude





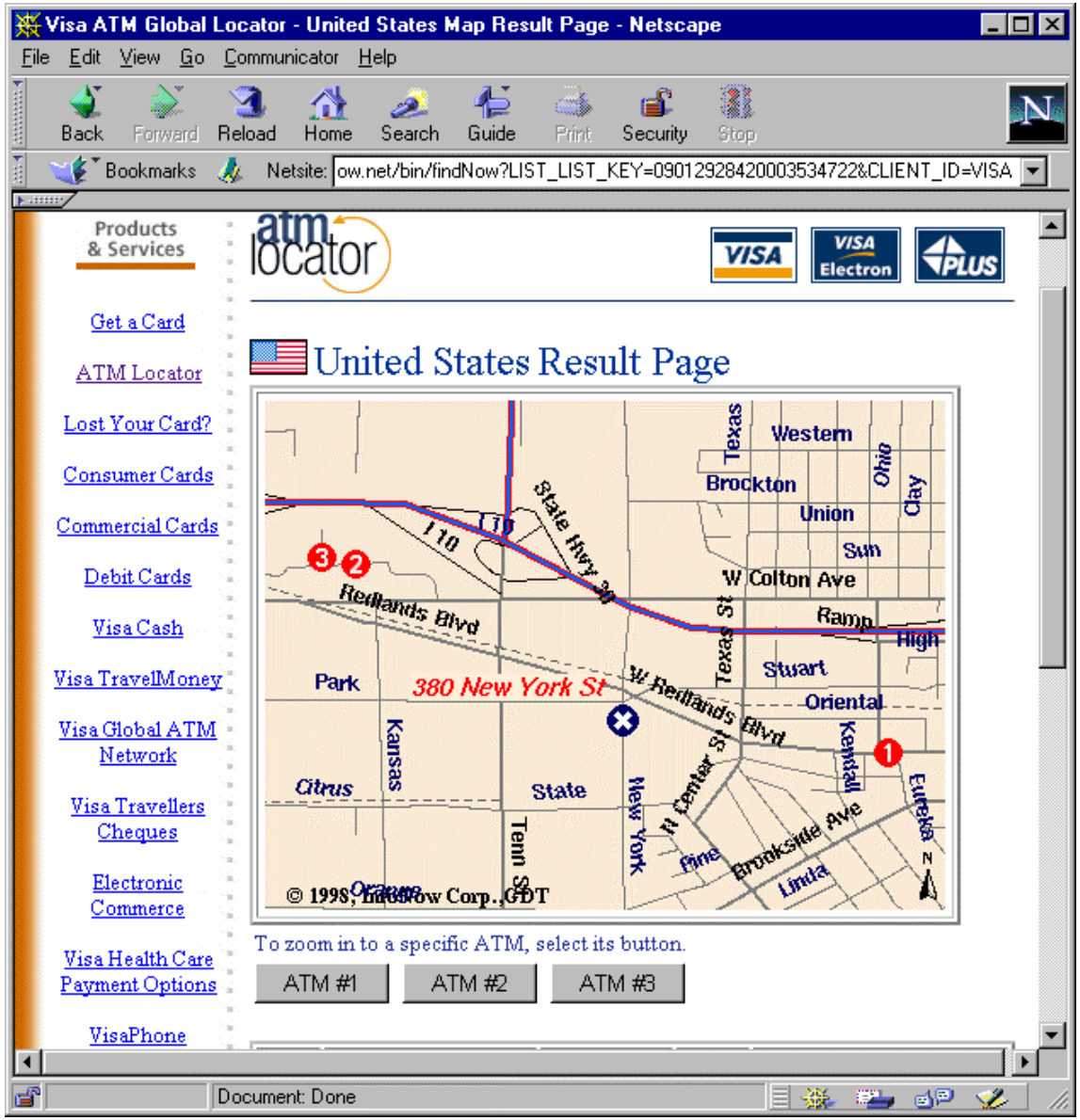
Other Shape-Table Applications

- Route tracker from GPS feed with Fuzzy logic analysis.
  - Closest Street.
  - Aligned with Street.
  - Most probable street continuation.
- Post processing geographical associations.

SDE on the Web

- Visa ATM application
- Realtor.com.
- ARC Data Online.
- MSP Flight Track Data.


Visa ATM Global Locator - United States - Netscape
File Edit View Go Communicator Help
Back Forward Reload Home Search Guide Print Security Stop
Bookmarks Netsite: ttp://visaatm.infonow.net/bin/findNow?LIST_LIST_KEY=09012928290003534647
VISA
Home
Products & Services
Offers & Promotions
Consumer Tips
For Businesses
Sponsorships & Events
New Technologies
About Visa
Products & Services
atm locator
VISA
VISA Electron
PLUS
Get a Card
ATM Locator
Lost Your Card?
Consumer Cards
Commercial Cards
Debit Cards
Visa Cash
Visa TravelMoney
Visa Global ATM Network
Visa Travellers Cheques
Electronic Commerce
United States
To find the three closest ATM's to your location, please type information and click "SUBMIT". For a map, you must enter a valid street address or cross-streets, as well as a city and state. A minimum of a city and state is required for a list-only response.
1 Street Address
380 new york st
1a Or Cross Street
2 City
redlands
3 State/Province
CALIFORNIA
4 Postal Code
92373
Map and List
List Only
SUBMIT
North America
Click to select
Go!
Document: Done


Visa ATM Global Locator - United States Map Result Page - Netscape
File Edit View Go Communicator Help
Back Forward Reload Home Search Guide Print Security Stop
Bookmarks Netsite: ow.net/bin/findNow?LIST_LIST_KEY=0901292842000353472&CLIENT_ID=VISA
Products & Services
Get a Card
ATM Locator
Lost Your Card?
Consumer Cards
Commercial Cards
Debit Cards
Visa Cash
Visa TravelMoney
Visa Global ATM Network
Visa Travellers Cheques
Electronic Commerce
Visa Health Care Payment Options
VisaPhone
atm locator
VISA
VISA Electron
PLUS
United States Result Page
380 New York St
Redlands Blvd
W Redlands Blvd
W Colton Ave
State Hwy 30
I 10
Texas St
Western
Brockton
Union
Sun
Ramp
High
Stuart
Oriental
Park
Kansas
Citrus
State
Tenn St
New York
N Center St
Kendall
Eureka
Brookside Ave
Linda
Pine
Orange
© 1998, Indivow Corp., GDT
To zoom in to a specific ATM, select its button.
ATM #1
ATM #2
ATM #3
Document: Done

SDE on the Web

- Integrate with a Web Server.
- Communicate via CGI or Servlet.
- Handle lots of users.
- Fast Reply.
- Load Balancing.
- Scalable.
- Flexible.

SDE on the Web

- Adopted a Multi-Tier architecture
- Separate
  - Business Logic
  - Data Access
  - Data Presentation
- Platform independent
- Protocol independent

SDE on the Web

Data Presentation

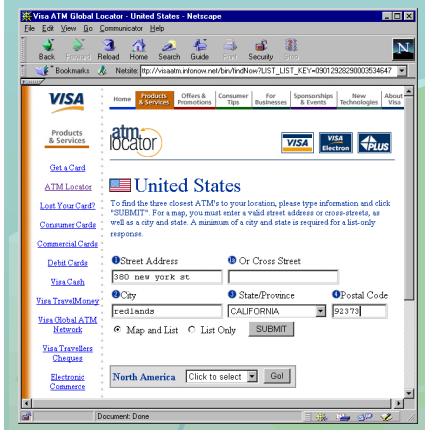
Graphics, Tables

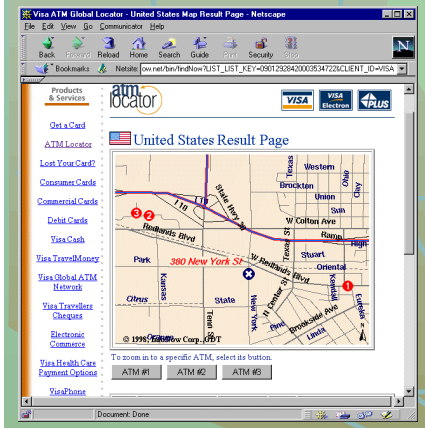
Business Logic That Calls Data Access Layer Objects

Math, If-then-else logic

Data Access Agent that does one thing very well

Divide and Conquer Approach.


Visa ATM Global Locator - United States - Netscape
United States
To find the three closest ATM's to your location, please type information and click "SUBMIT". For a map, you must enter a valid street address or cross-streets, as well as a city and state. A minimum of a city and state is required for a list-only response.
Street Address
Or Cross Street
380 new york st
City
State/Province
Postal Code
redlands
CALIFORNIA
92373
Map and List
List Only
SUBMIT
North America
Click to select
Go!
Document: Done


Visa ATM Global Locator - United States Map Result Page - Netscape
United States Result Page
380 New York St
To zoom in to a specific ATM, select its button.
ATM #1
ATM #2
ATM #3
Document: Done

WWW
WWW.VISA.COM
Web Server
CGI/SERVLET
Business Logic
IMS CORE
GEOCODER
PROXIMITY
MAPPER
Data Access
SDE
Data Repository

SDE on the Web

- Create SDE Agents with persistent connection.
- Agents are small and do one thing and one thing very well.
- Agent should be platform and protocol independent.
- Agents with Read-Only features should adopt caching strategy all the way to the middle tier.
- Design with reuse in mind.

SDE on the Web Stats

- SUN Enterprise 6000
  - 8 CPU
  - 4 Gigabyte of RAM (2 for Oracle SGA)
  - 12 Disk Controllers managing 350 Gigabyte of Storage (RAID 1)
- 22 Mappers, 2 Geocoders, 6 Query.
- 2 seconds / map (most time in gif file creation)
- 300,000 maps / day
- 20,000 requests / hour
- 340 SDE Layers

WEB SDE

- Enable Remote Access.
- Too Many Options To Remember.
- Need Friendly GUI.
- Platform Independent.
- Use existing SDE admin tools.
- The Web Tools Are Cool.

Thank You
Questions And Answers

Mraad@esri.com

ftp.esri.com


Sharing Geographic Knowledge
UC
99
ESRI
Nineteenth Annual User Conference