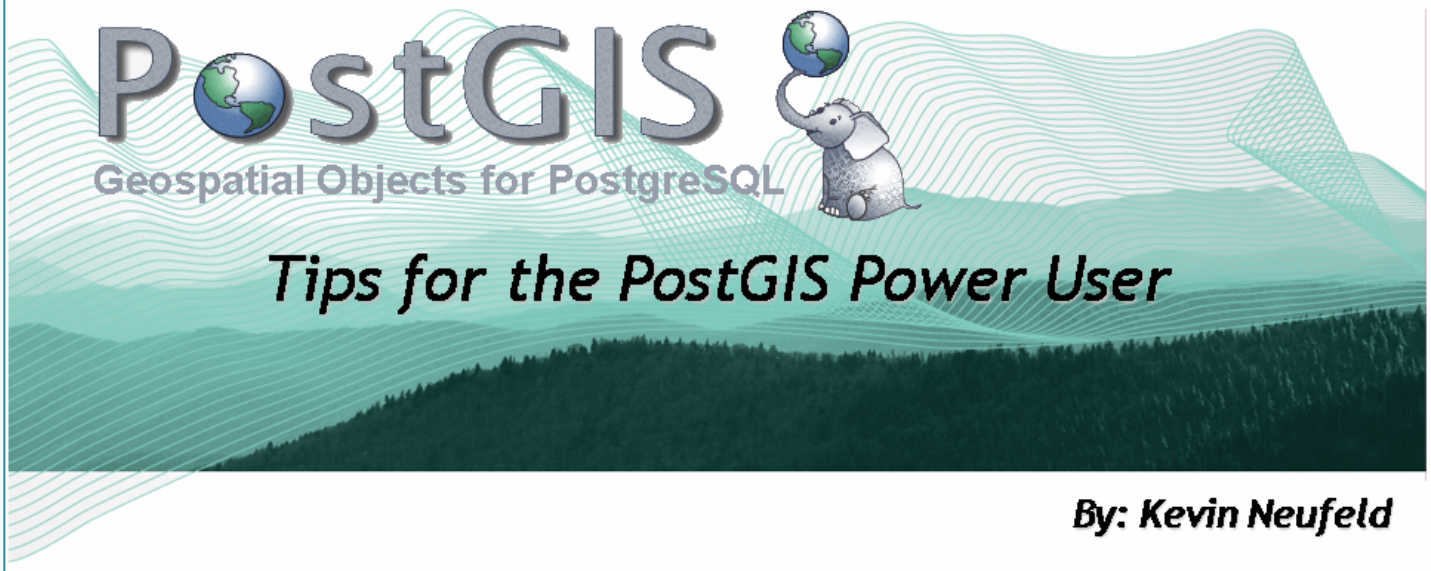
WWW.REFRACTIONS.NET

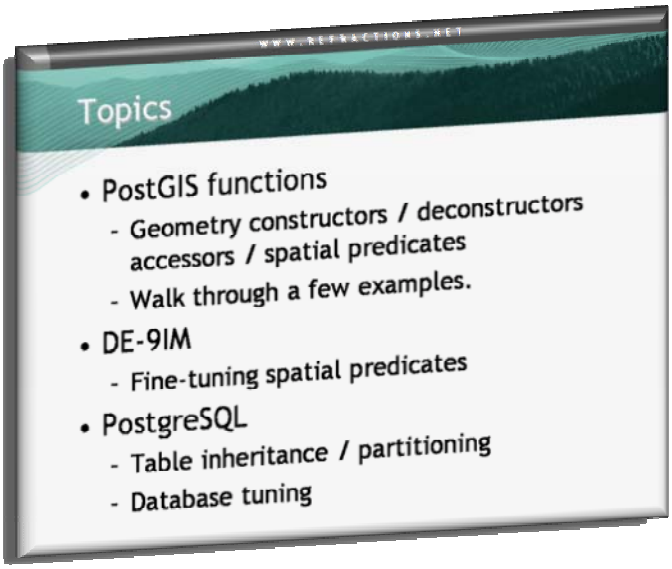


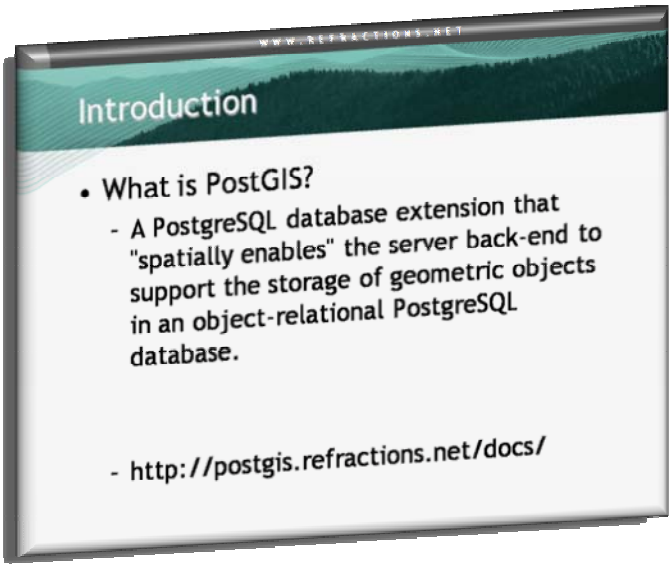
Tips for the PostGIS Power User

By: Kevin Neufeld

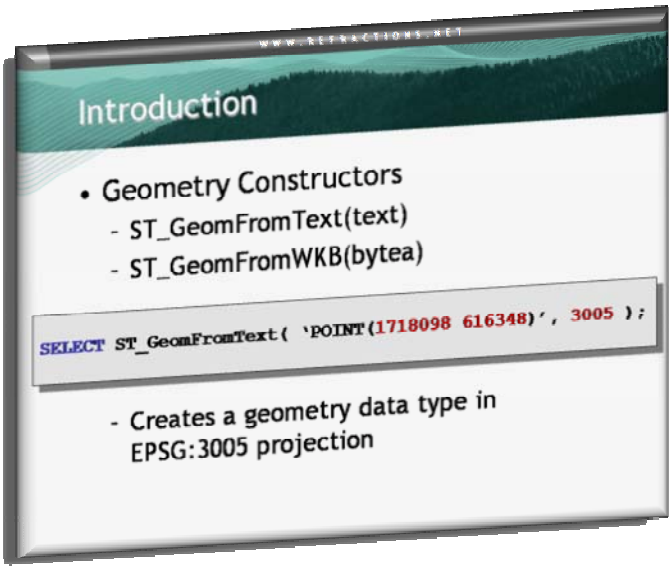
Refractions RESEARCH

THE GEOSPATIAL EXPERTS



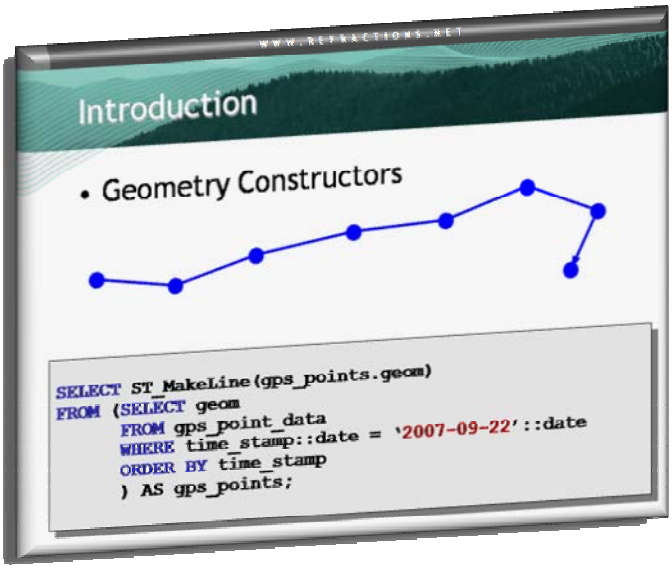


PostGIS also provides mechanisms to efficiently access these geometry types and perform spatial operations on them.

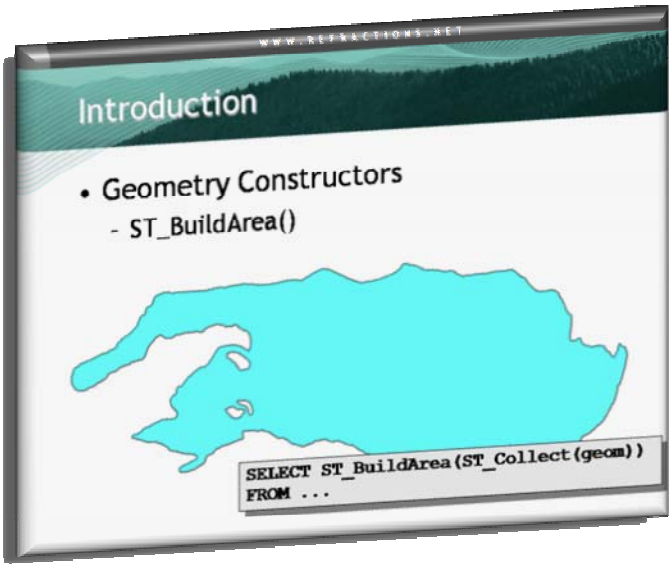


Of course, one would very rarely enter data manually this way, but rather import data directly into PostGIS using the shape dumper (shp2pgsql) or programmatically using your preferred language.

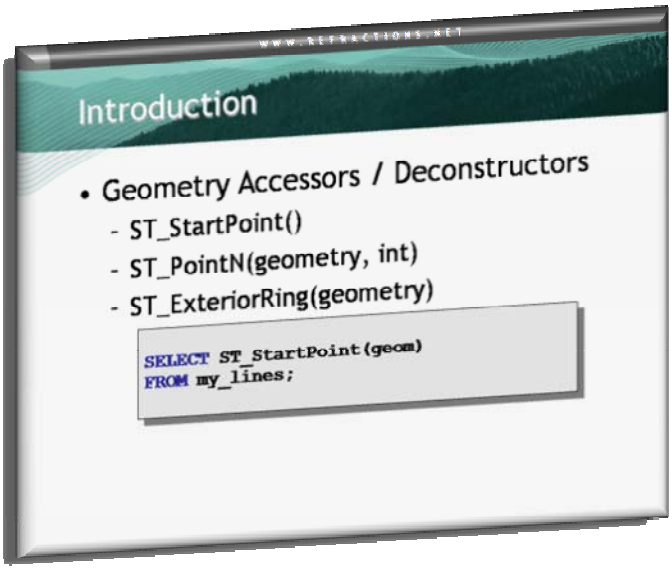
See Appendix A for example JAVA code demonstrating reading from and writing to a PostGIS enabled PostgreSQL database.

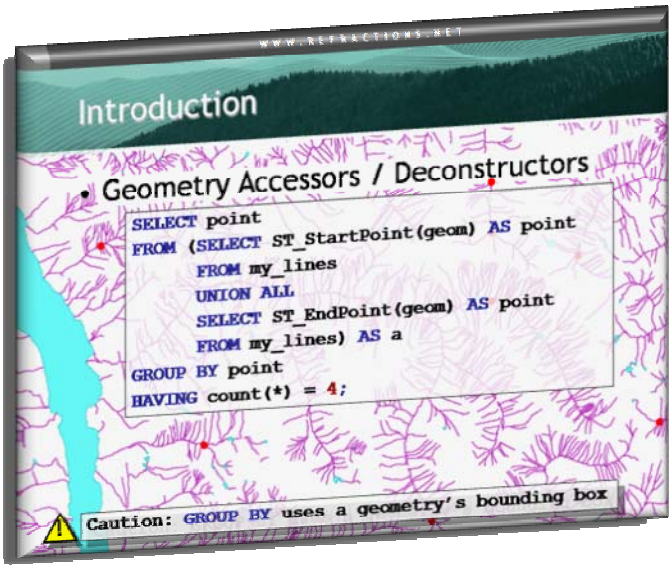


A typical example of creating a LineString from an ordered set of Points.

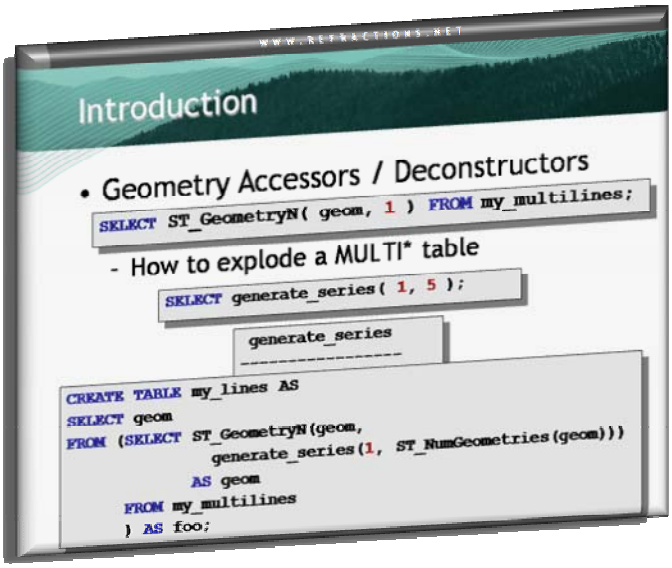


ST_BuildArea() works well when supplied a single geometry collection of LineStrings. Note though, that ST_BuildArea() works a little like ST_Union(), in that shared lines in adjacent polygons will get removed, merging polygons. Conversely, ST_Polygonize() will return "all" polygons formed by linear input. In this example, this includes a polygon with 3 holes in addition to the three inner rings as simple polygons.

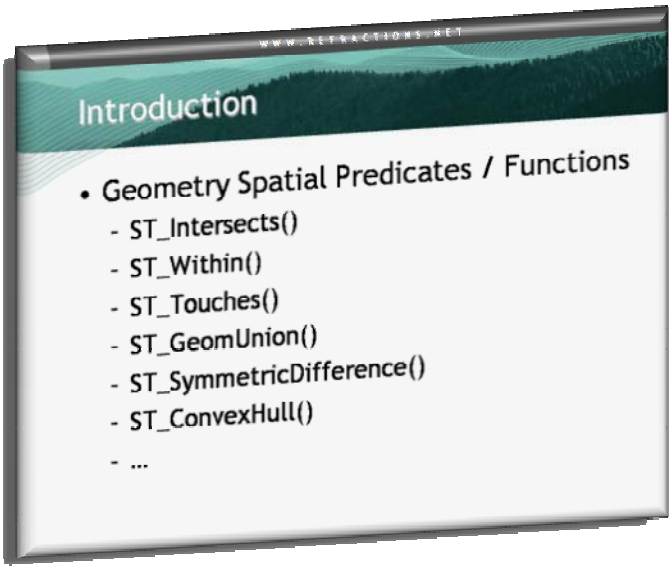


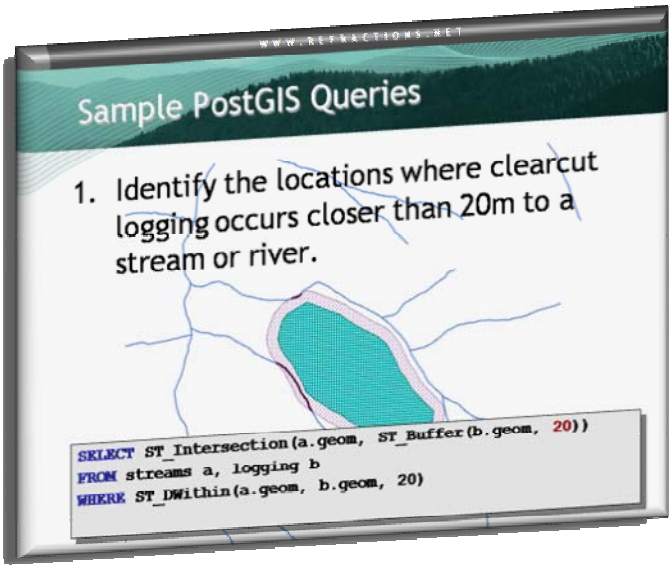


This query will find coincident start and end points, identifying locations in a linear network where a confluence of degree 4 occurs.

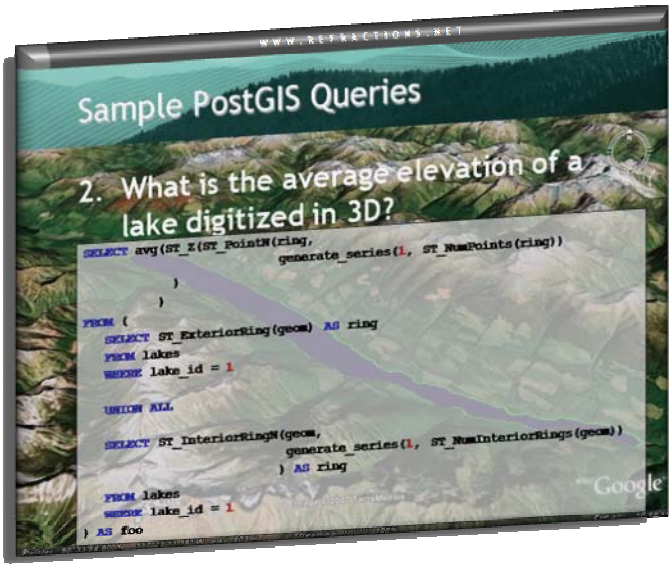


generate_series() is a set returning function with a return type of "int". Caveat: This function does not scale well, i.e. extracting points from large linestring with excess of 50,000 coordinates. Alternatively, one could use ST_Dump(), however, this function is not currently an aggregate, nor will it extract points from other lines or polygons.

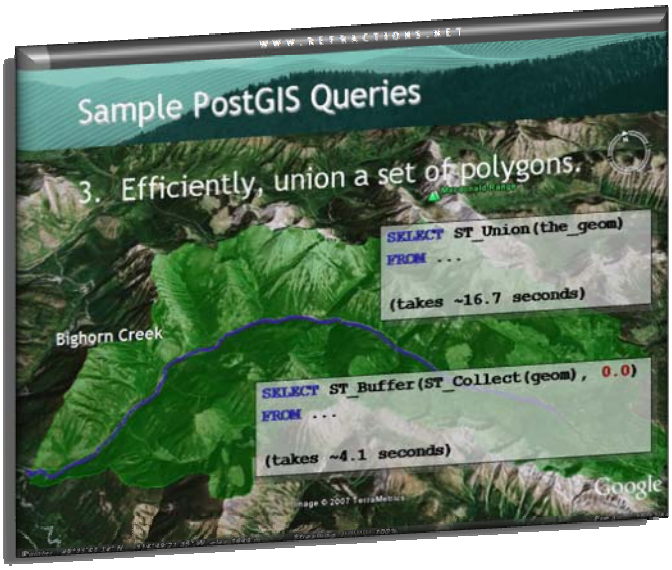




ST_Buffer() is an expensive operation. Don't use it to determine distance relationships.



This query will extract all the linework from a polygon (the exterior ring followed by all the interior rings) and further extract all the points from these rings.



Instead of using ST_Union(), adding one geometry at a time to a new union-ed geometry, ST_Buffer(0) will perform this operation quicker since it is operating on a single geometry object.

Sample PostGIS Queries

4. Find all docks that are contained completely within a lake, not touching a lake bank.

What PostGIS functions would you use?
~~ST_Within~~? ~~ST_Contains~~? ~~ST_Touches~~?

DE-9IM The Dimensionally Extended - Nine Intersection Model

- Approach
  - make pair-wise tests of the intersections between the Interiors, Boundaries, and Exteriors of two geometries and to represent these relationships in an "intersection" matrix

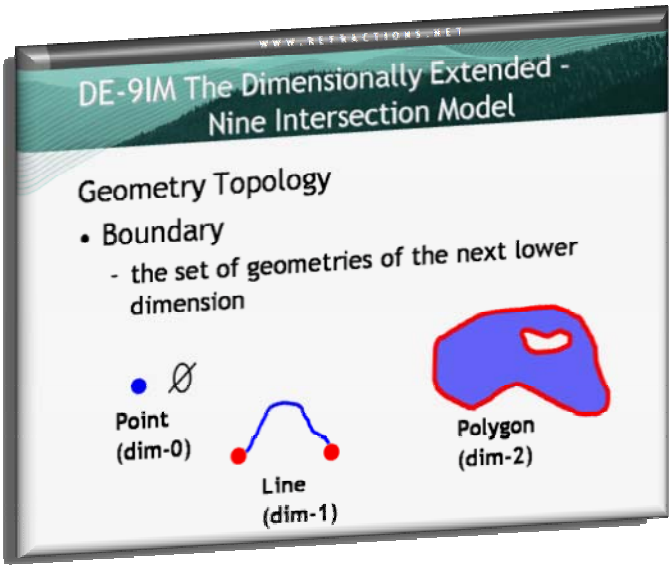
DE-9IM The Dimensionally Extended - Nine Intersection Model

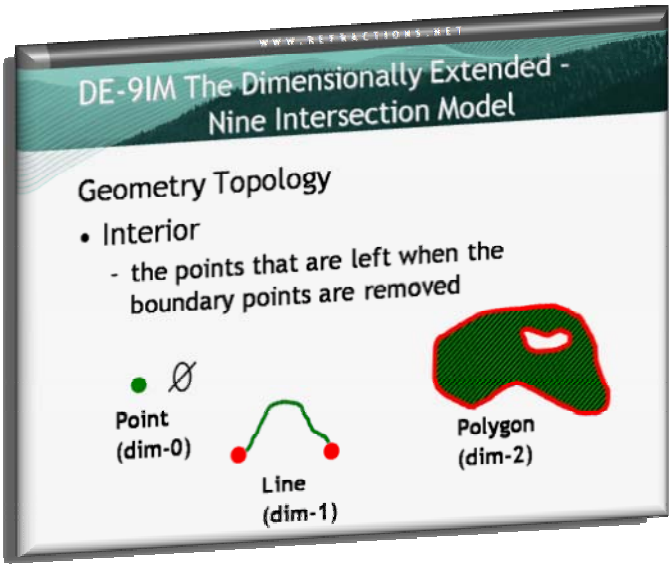
	Interior	Boundary	Exterior
Interior	*dim(I(a) ∩ I(b))*	*dim(I(a) ∩ B(b))*	*dim(I(a) ∩ E(b))*
Boundary	*dim(B(a) ∩ I(b))*	*dim(B(a) ∩ B(b))*	*dim(B(a) ∩ E(b))*
Exterior	*dim(E(a) ∩ I(b))*	*dim(E(a) ∩ B(b))*	*dim(E(a) ∩ E(b))*

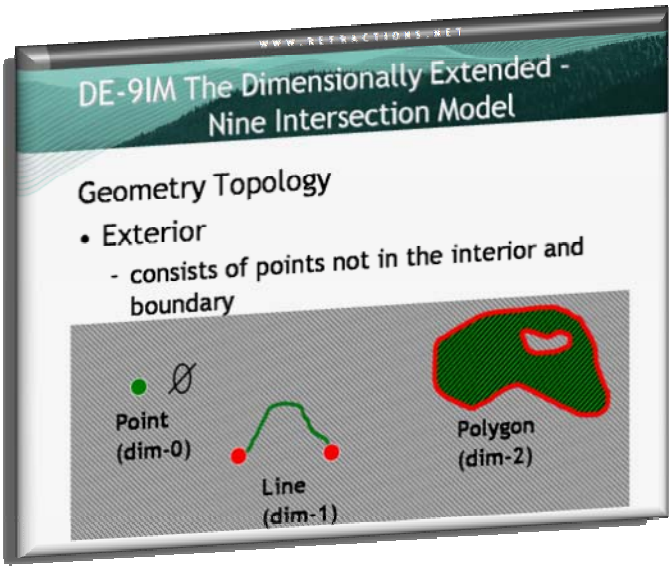
Possible values:
{T, F, *, 0, 1, 2}

Where:
T == {0,1,2}
F == empty set
* == don't care
0 == dimensional 0 - point
1 == dimensional 1 - line
2 == dimensional 2 - area

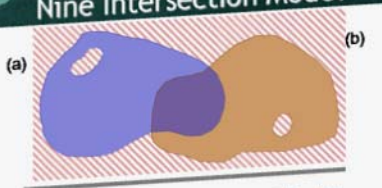
- where ***dim()*** is the maximum dimension of the intersections involved.


DE-9IM The Dimensionally Extended - Nine Intersection Model
Geometry Topology
• Boundary
- the set of geometries of the next lower dimension
Point (dim-0)
Line (dim-1)
Polygon (dim-2)


DE-9IM The Dimensionally Extended - Nine Intersection Model
Geometry Topology
• Interior
- the points that are left when the boundary points are removed
Point (dim-0)
Line (dim-1)
Polygon (dim-2)


DE-9IM The Dimensionally Extended - Nine Intersection Model
Geometry Topology
• Exterior
- consists of points not in the interior and boundary
Point (dim-0)
Line (dim-1)
Polygon (dim-2)

DE-9IM The Dimensionally Extended - Nine Intersection Model



	Interior	Boundary	Exterior
Interior	2	1	2
Boundary	1	0	1
Exterior	2	1	2

ST_Relate(a, b) = '212101212'

Sample PostGIS Queries

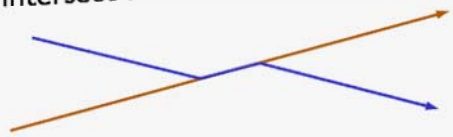
4. Find all docks that are contained completely within a lake, not touching a lake bank.

```
SELECT a.id
FROM docks a, lakes b
WHERE a.geom && b.geom
AND ST_Relate(a.geom, b.geom, 'TFFTFF212');
```

'TFFTFF212' means we are interested in cases where the interior and boundary of the dock intersects the interior (and not the boundary) of the lake.

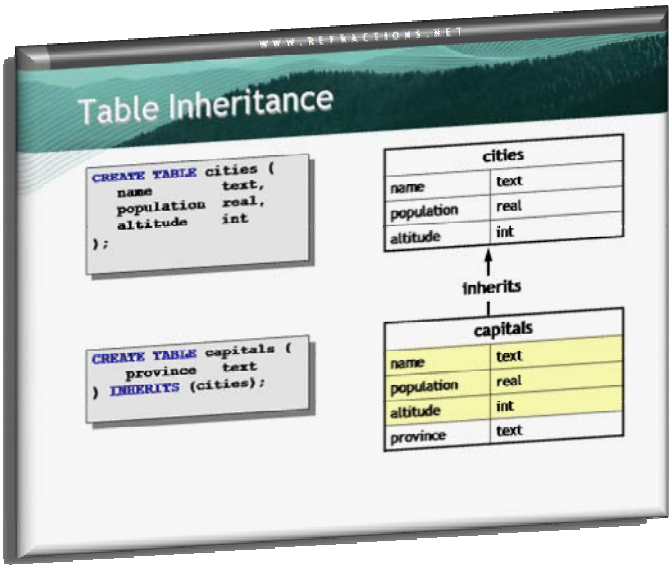
Sample PostGIS Queries

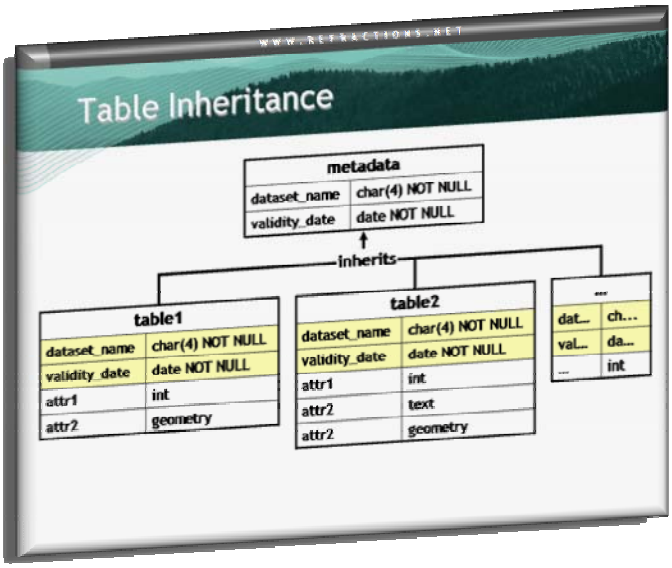
5. Identify linear spatial features that intersect on a line and not at a point.



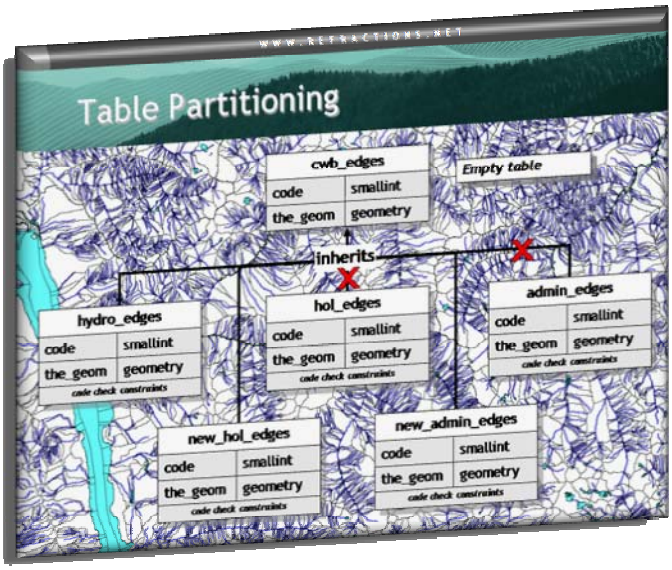
```
SELECT a.id, intersection(a.geom, b.geom)
FROM mylines a, mylines b
WHERE a.id != b.id
AND a.geom && b.geom
AND ST_Relate(a.geom, b.geom, '1*1***1**');
```

'1*1***1**' means we are looking for cases where the interiors intersect on a line, and the interior of both lines intersect the exterior of the other.

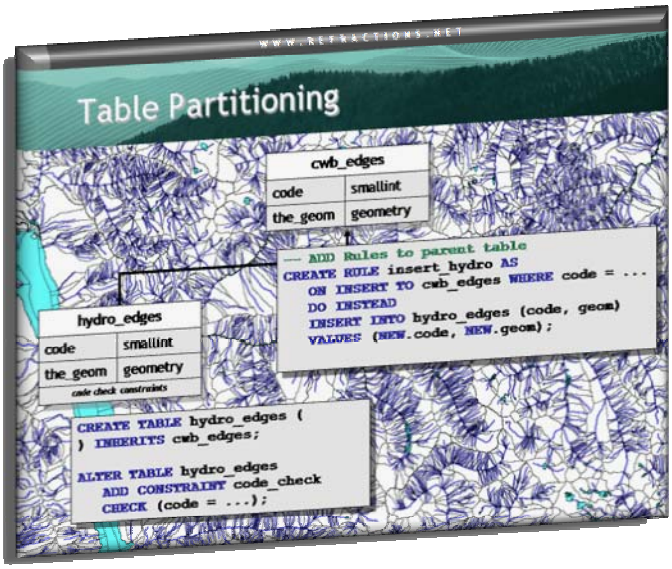




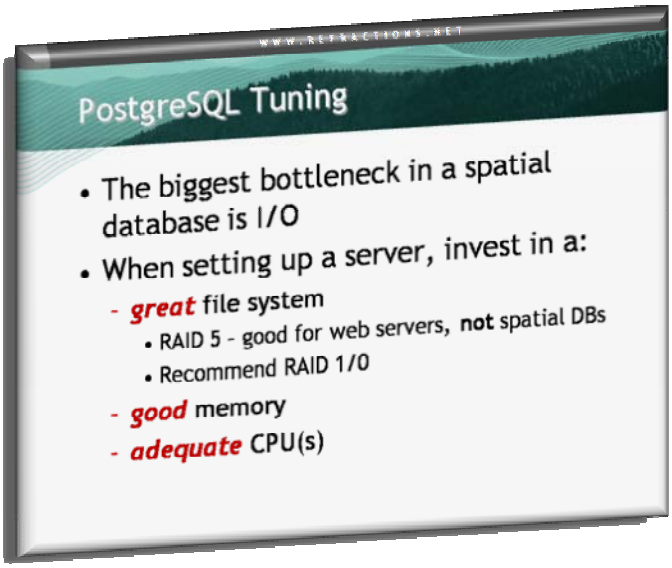
Since certain table constraints are inherited, a DBA could use inheritance to enforce spatial constraints on tables. I.E. an empty table could be created so that all child tables are known to have a certain SRID, dimension, or geometry type.



It is through the concept of table inheritance that table partitioning can be achieved. Due to the table's check constraints, the query planner explicitly knows in which child table(s) your data resides.

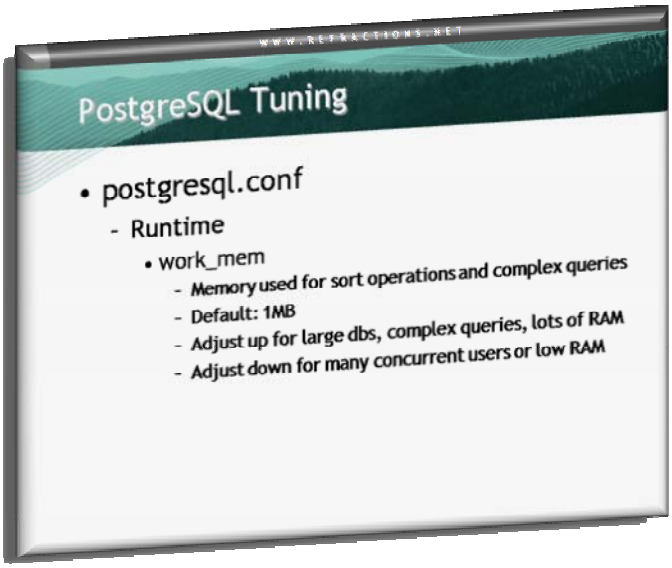


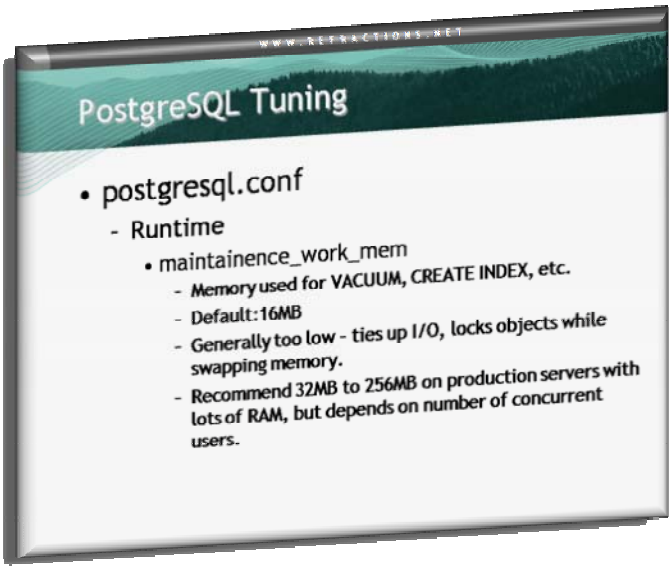
RULEs are used to divert INSERTs into the appropriate child table and to control UPDATEs among all children tables.

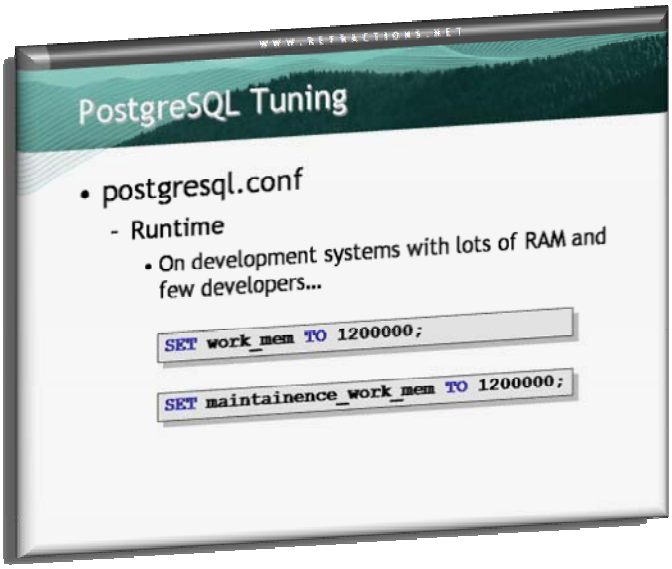


shared_buffers is very dependant on the db use case and the number of concurrent connections. This can be high on a web server with many connections where work_mem is set low. On a development box with few users running complex queries (thus needing a high work_mem setting), this setting will be lower since the available memory will be lower.

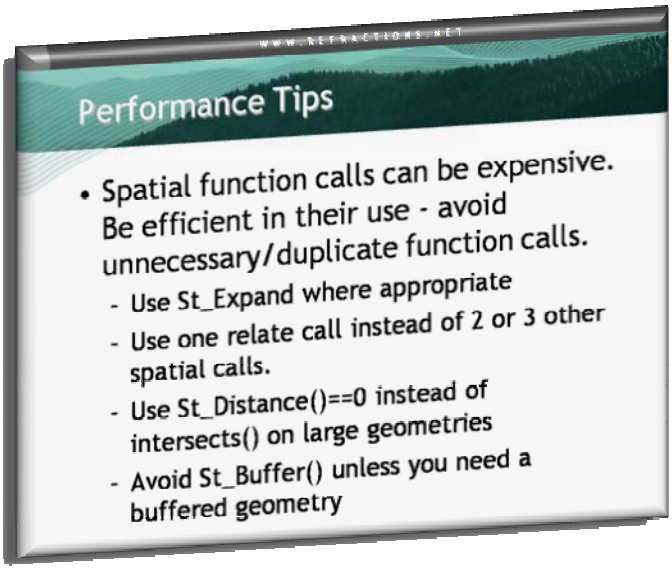
If you want to utilize the power of table partitioning, enable this setting, otherwise the query planner won't use constraint checking.

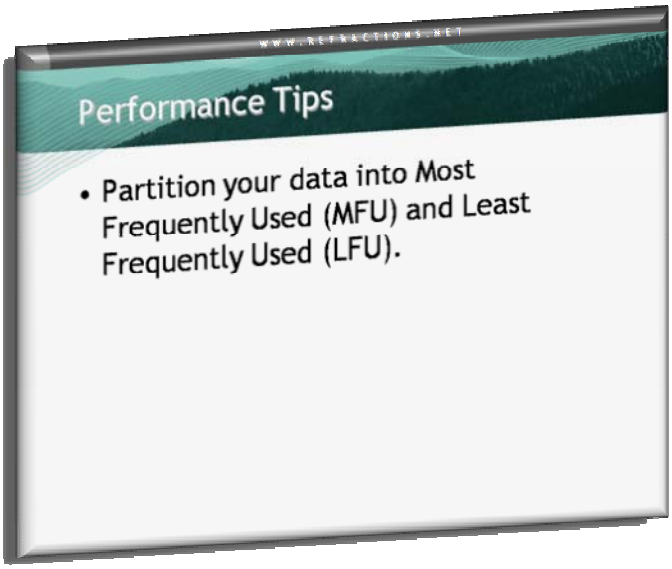


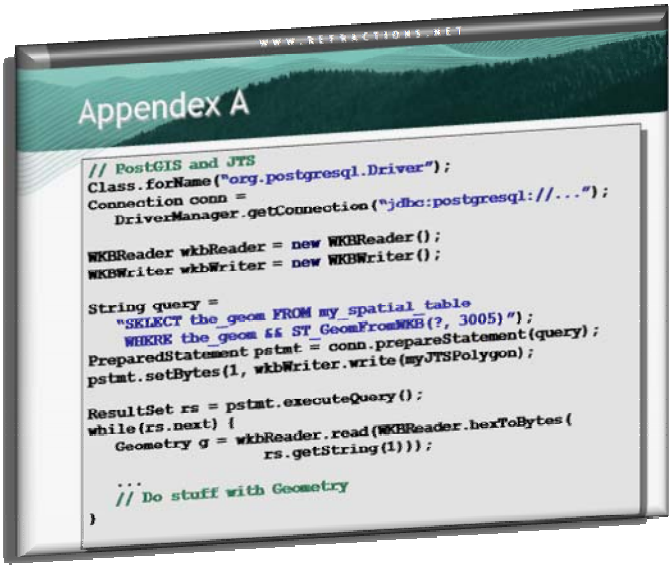



PostgreSQL Tuning
• postgresql.conf
- Runtime
• On development systems with lots of RAM and few developers...
SET work_mem TO 1200000;
SET maintainence_work_mem TO 1200000;


PostgreSQL Tuning
• postgresql.conf
- Runtime
• client_min_messages
SET client_min_messages to DEBUG;
- Useful when writing PL/Pgsql functions.
CREATE FUNCTION my_function () RETURNS TEXT AS
$BODY$
BEGIN
...
RAISE DEBUG 'myvar: %' var;
...


Performance Tips
• Spatial function calls can be expensive. Be efficient in their use - avoid unnecessary/duplicate function calls.
- Use St_Expand where appropriate
- Use one relate call instead of 2 or 3 other spatial calls.
- Use St_Distance()==0 instead of intersects() on large geometries
- Avoid St_Buffer() unless you need a buffered geometry





Some sample code that will read and write JTS Geometries from and to a PostGIS database.