ORACLE®

SPATIAL

April 2010

Oracle Spatial User Conference

ORACLE
SPATIAL

April 2010
Oracle Spatial
User Conference

Oracle Spatial

User Conference

April 29, 2010

Hyatt Regency Phoenix

Phoenix, Arizona USA

ORACLE SPATIAL

April 2010
Oracle Spatial
User Conference

Nick Padfield

Team Lead Cartographer, U.S. Census Bureau

ORACLE
SPATIAL

April 2010
Oracle Spatial
User Conference

Oracle Spatial Cartographic Database

Optimizing the U.S. Census Bureau's TIGER data for Mapping

ORACLE SPATIAL

April 2010

Oracle Spatial

User Conference

Presentation Topics

- Statistics, Databases, and Environment
- Cartographic Database
  - Topology Overview
  - Chaining Process
  - Dissolving Process
- Lessons Learned, Future Direction

ORACLE
SPATIAL

April 2010

Oracle Spatial

User Conference

Statistics, Databases, and Environment

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

ORACLE
SPATIAL

April 2010

Oracle Spatial

User Conference

Lots and lots of maps

- About <u>15 million</u> maps produced to date

- Maps Support 2010 Census
  - Field Maps for Census Field Workers
  - Publication Maps for Public Consumption

ORACLE
SPATIAL

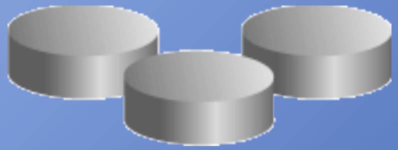
April 2010

Oracle Spatial

User Conference

Database Environment

- Oracle 10g, Release 2
- Real Application Cluster (RAC) Environment
- Have 46 distinct database instances
- Instances running on Linux Blade servers
- Have development, testing, and production environments



ORACLE SPATIAL

April 2010

Oracle Spatial

User Conference

Spatial Databases

- Live MAF/TIGER Database (2 TB)

  - Benchmark Databases (~1.6 TB)

  - Product Databases (~364 GB)

  - Cartographic Databases (~75 GB)

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

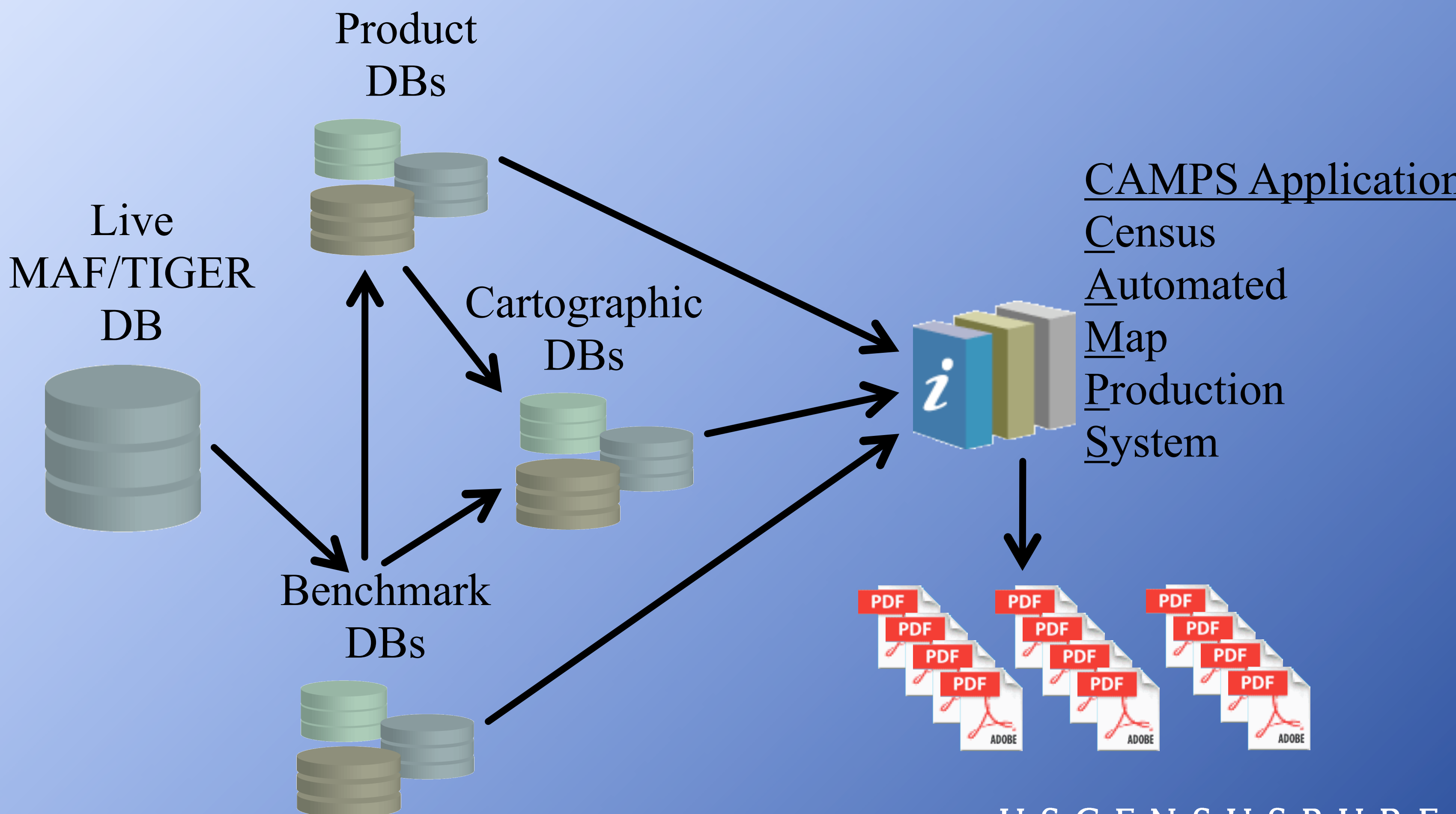
ORACLE SPATIAL

April 2010

Oracle Spatial

User Conference

Workflow Involved with Creating a Map



U S C E N S U S B U R E A U

Helping You Make Informed Decisions

ORACLE
SPATIAL

April 2010
Oracle Spatial
User Conference

Live MAF/TIGER Database Statistics

- 254 million primitives
  - 159 million nodes
  - 73 million edges
  - 22 million faces
- 16 million legal and statistical geographies
- 55 different types of geographic areas
- 100 million housing unit locations
- 23 million road features
- 9 million other physical features

U S C E N S U S B U R E A U
Helping You Make Informed Decisions

ORACLE SPATIAL

April 2010

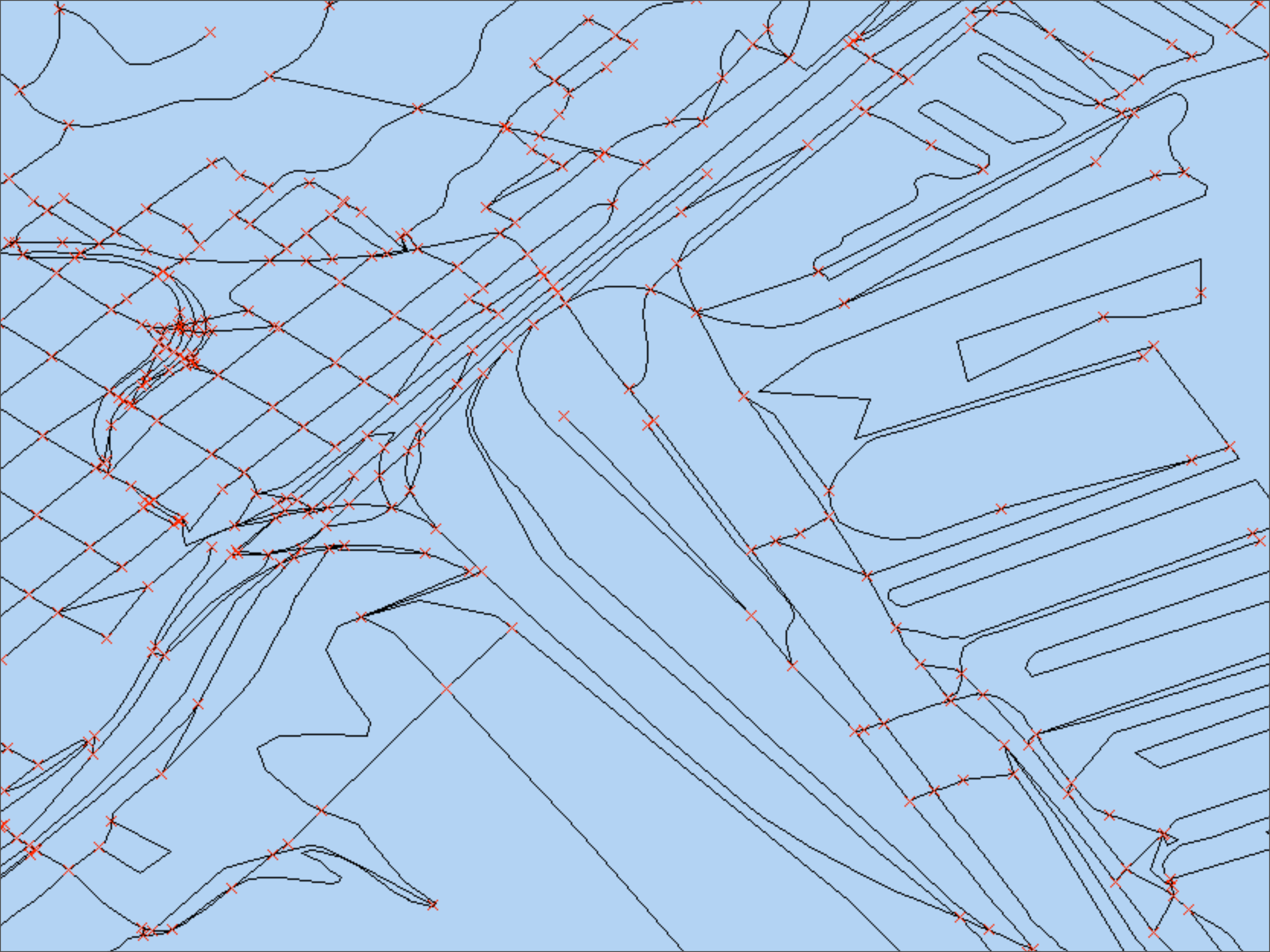
Oracle Spatial

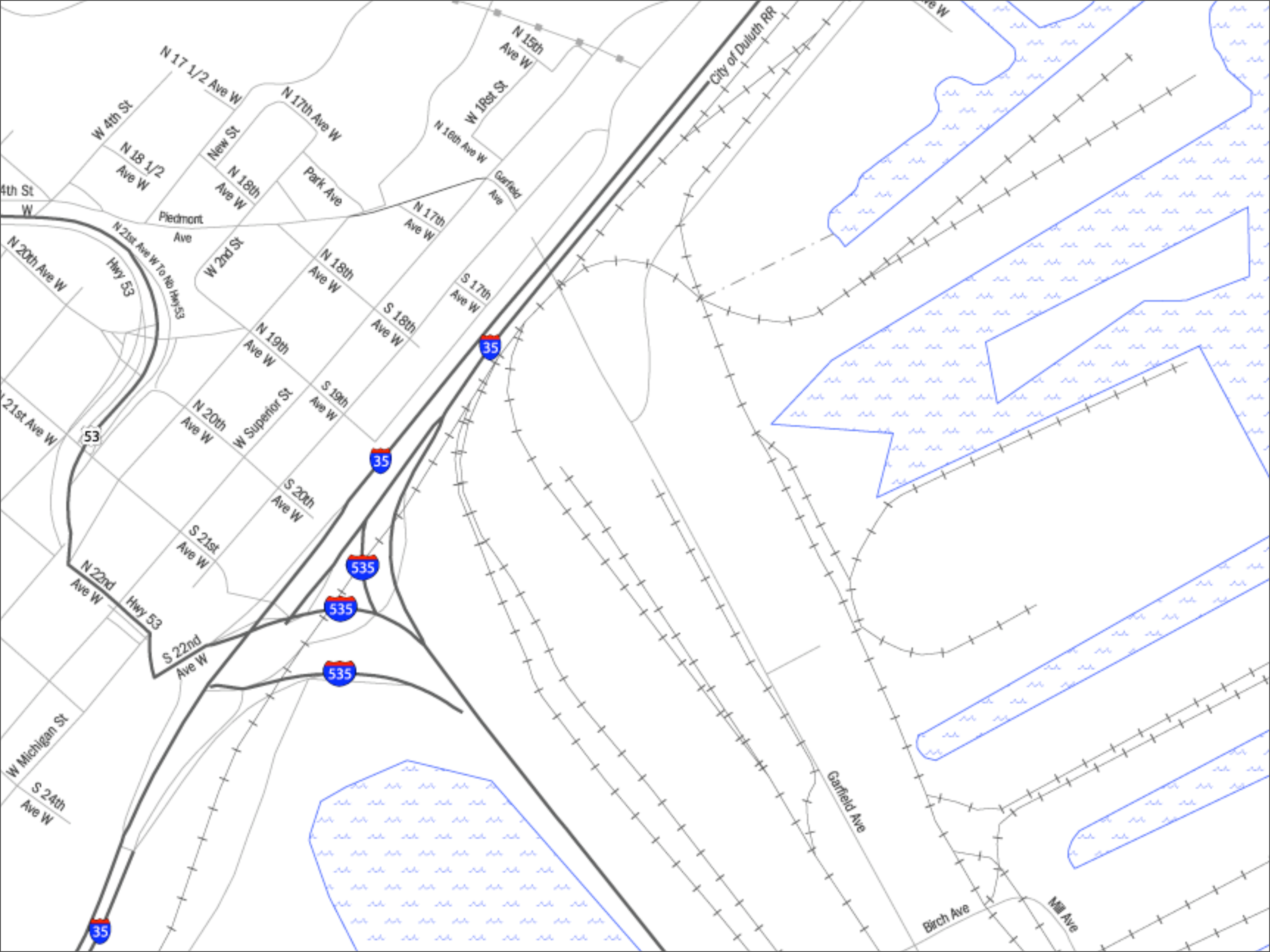
User Conference

Cartographic Database

U S C E N S U S B U R E A U

Helping You Make Informed Decisions





N 17 1/2 Ave W
N 17th Ave W
N 15th Ave W
W 1st St
N 16th Ave W
City of Duluth RR
W 4th St
New St
N 18 1/2 Ave W
N 18th Ave W
Park Ave
Garfield Ave
4th St
W
Piedmont Ave
N 17th Ave W
N 21st Ave W To Nb Hwy53
N 20th Ave W
Hwy 53
W 2nd St
N 18th Ave W
S 17th Ave W
S 18th Ave W
N 19th Ave W
35
S 19th Ave W
N 20th Ave W
W Superior St
N 21st Ave W
53
35
S 20th Ave W
S 21st Ave W
535
N 22nd Ave W
535
Hwy 53
S 22nd Ave W
535
W Michigan St
S 24th Ave W
35
Garfield Ave
Birch Ave
Mill Ave

ORACLE SPATIAL

April 2010

Oracle Spatial

User Conference

Cartographic Database Overview

- Predominantly stores features
- Denormalized database
- Data aggregation is key
  - Chaining
  - Dissolving
- Optimized for mapping throughput

April 2010

Oracle Spatial

User Conference

Cartographic Database Creation Process

- Korn shell master script
- Parallel processing with heavy reliance on DBMS_SCHEDULER
- Use of triggers to kick off subsequent Cartographic Database Units-of-Work

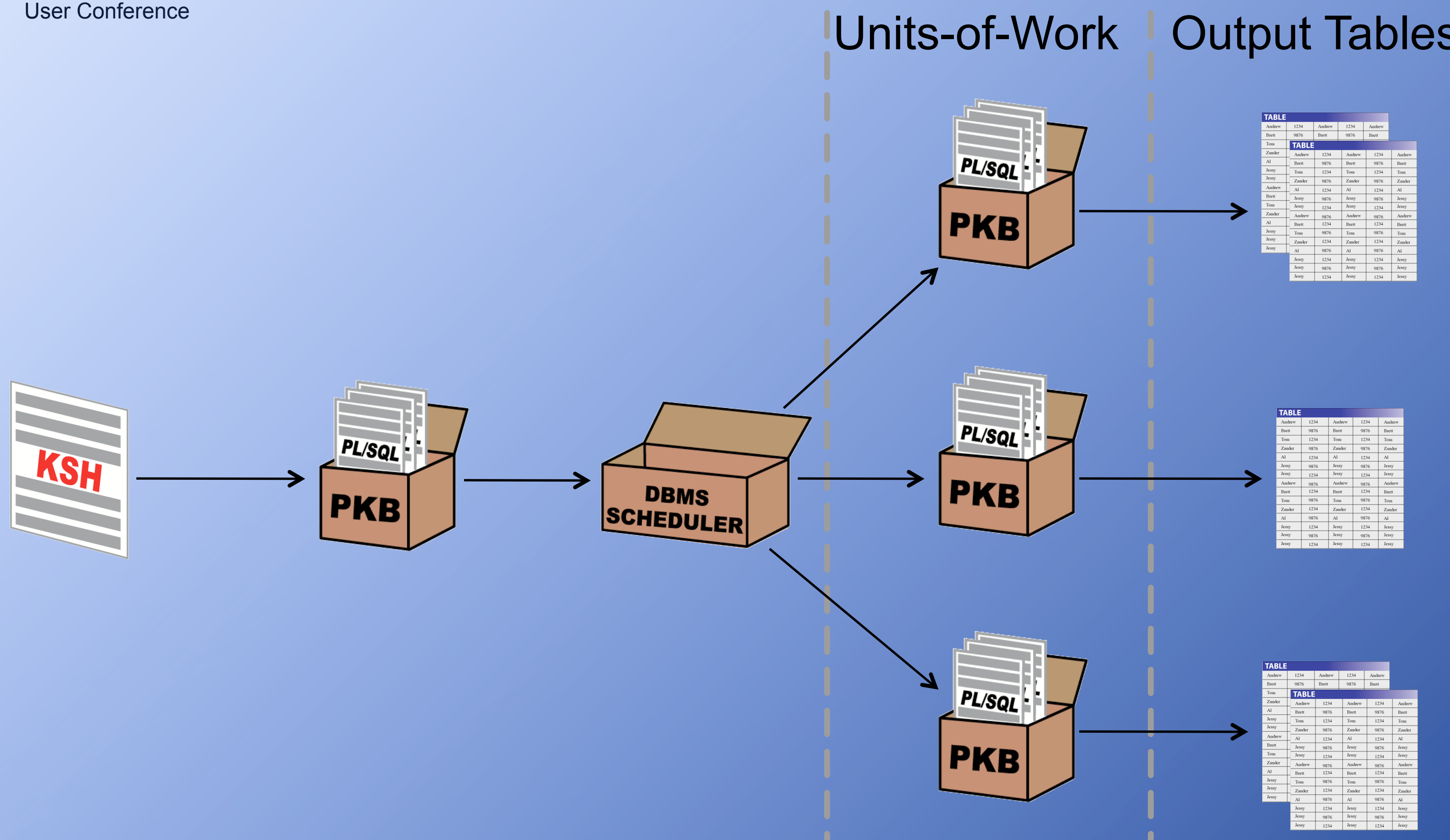
ORACLE
SPATIAL

April 2010

Oracle Spatial

User Conference

Cartographic Database Creation Process



U S C E N S U S B U R E A U

Helping You Make Informed Decisions

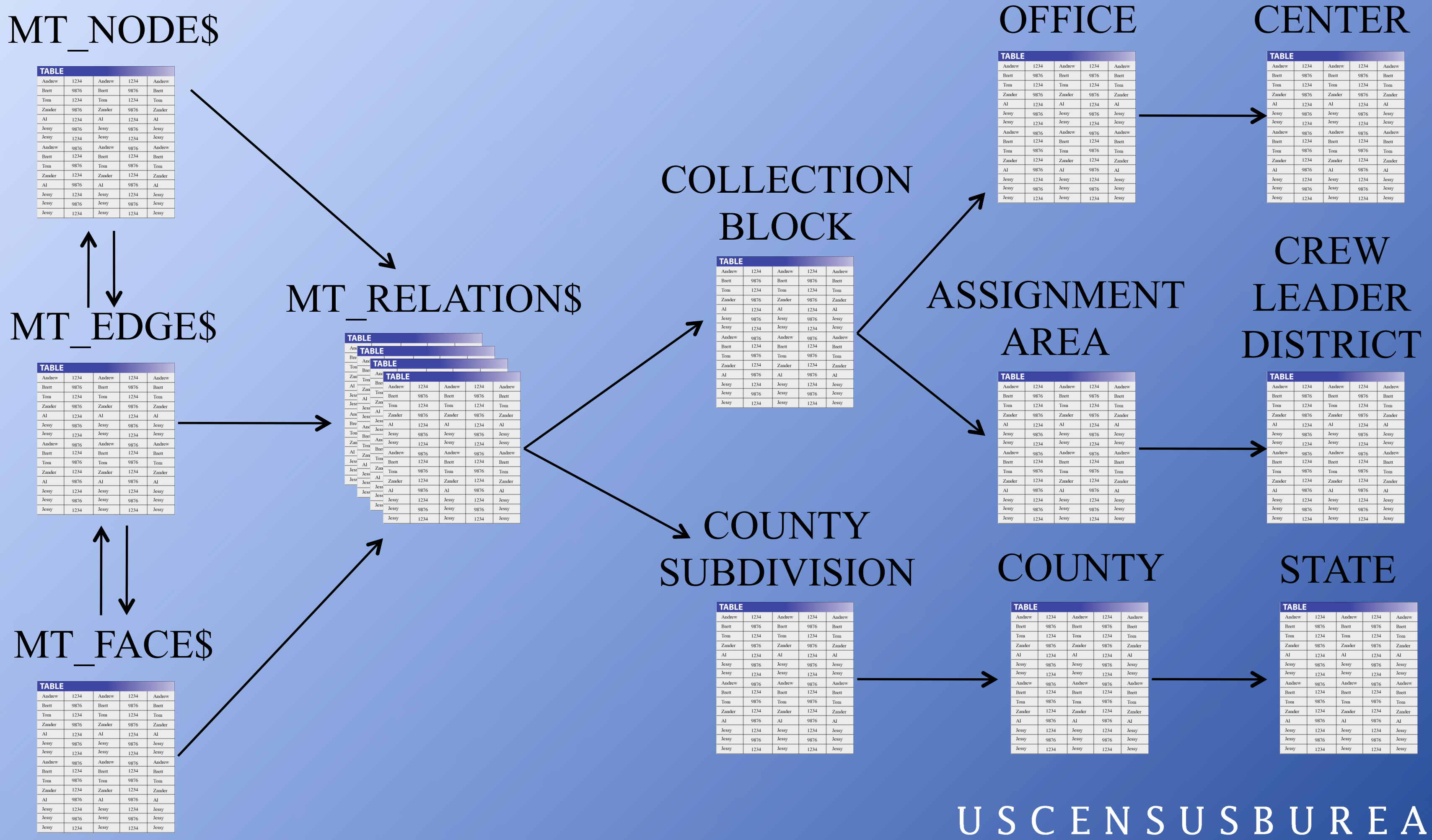
ORACLE SPATIAL

Topology Model Overview

April 2010

Oracle Spatial

User Conference



U S C E N S U S B U R E A U

Helping You Make Informed Decisions

ORACLE
SPATIAL

April 2010

Oracle Spatial

User Conference

Cartographic Database Chaining Requirements

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

ORACLE SPATIAL

April 2010

Oracle Spatial

User Conference

Cartographic Database Chaining Using Topology


```
SELECT topo_id AS edge_id
FROM mt_relation$ WHERE tg_layer_id = 1101
AND topo_type = 2 AND tg_id = 1105;
```

MT_RELATION$

TG_ID	TOPO_ID	TOPO_TYPE	TG_LAYER_ID
1105	8	2	1101
1105	9	2	1101
1105	10	2	1101


```
SELECT start_node_id, end_node_id
FROM mt_edge$ WHERE edge_id IN (8,9);
```

MT_EDGE$

EDGE_ID	START_NODE_ID	END_NODE_ID
8	3	4
9	4	5

SDO_UTIL.CONCAT_LINES()

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

Oracle Spatial

April 2010

Oracle Spatial

User Conference

Cartographic Database Chaining Optimization

- Data Filtering and classification
  - Filter and classify on tabular attributes
  - Filter and classify on Spatial Connectivity

- Use of continual optimization of temporary tables

ORACLE
SPATIAL

April 2010

Oracle Spatial

User Conference

Cartographic Database Dissolving Requirements

BEFORE
DISSOLVING

AFTER
DISSOLVING

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

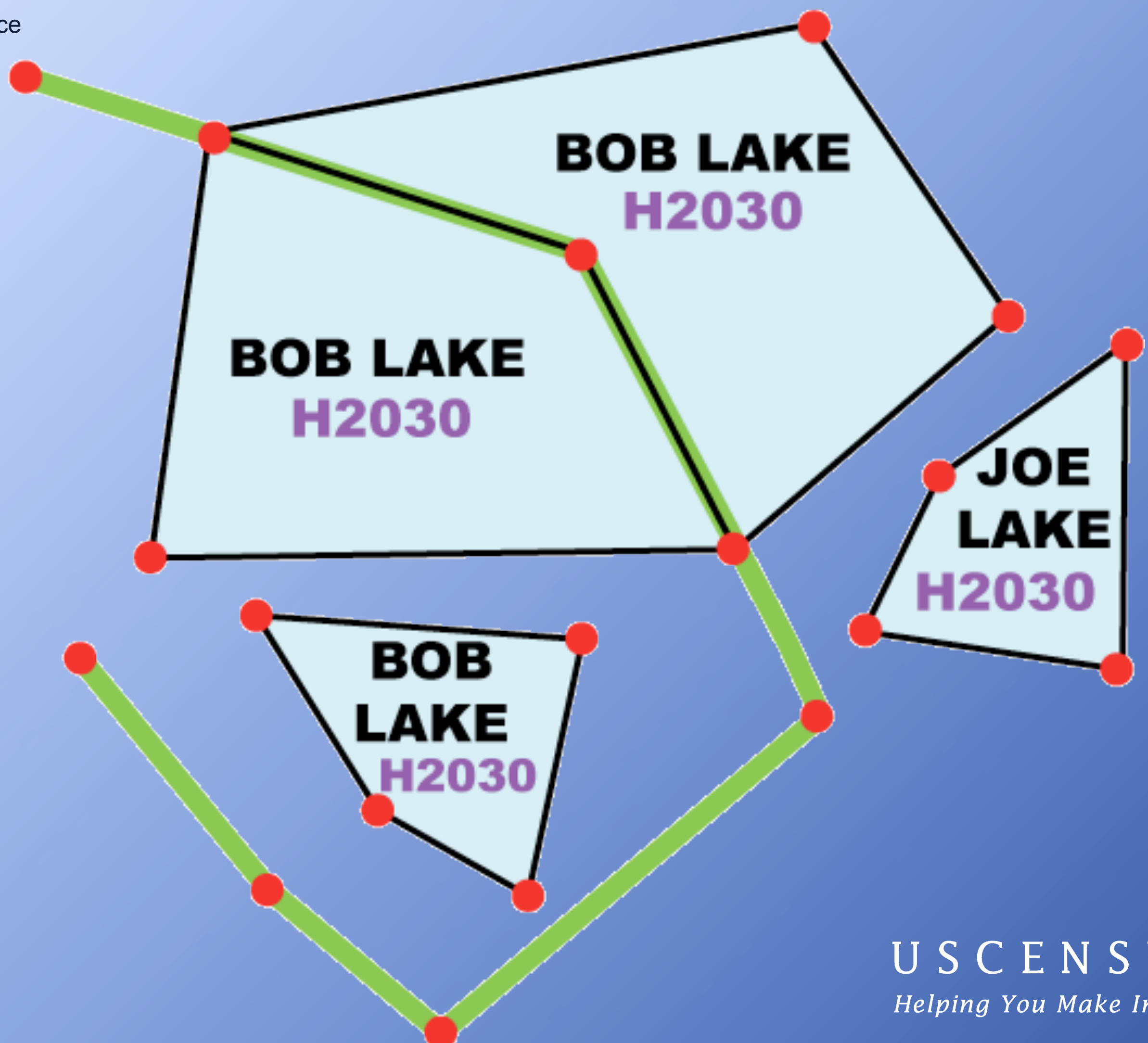
Cartographic Database Dissolving Using Topology

ORACLE SPATIAL

April 2010

Oracle Spatial

User Conference



US CENSUS BUREAU

Helping You Make Informed Decisions

ORACLE
SPATIAL

April 2010

Oracle Spatial

User Conference

Cartographic Database Dissolving Optimization

Optimizing Tricks

- Filtering data is important
- Classifying data is important
- Aggregating clustered polygons vs. disperse polygons
- Aggregating large numbers of polygons (max 100 at a time)

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

ORACLE SPATIAL

April 2010

Oracle Spatial

User Conference

Cartographic Database Dissolving via Pipeline

```
SELECT SDO_AGGR_UNION(SDOAGGRTYPE (ugeom, .05)) ugeom
     FROM
     (
         SELECT SDO_AGGR_UNION(SDOAGGRTYPE (ugeom, .05)) ugeom
             FROM
             (
                 SELECT SDO_AGGR_UNION(SDOAGGRTYPE(ugeom, .05)) ugeom
                     FROM
                     (
                         SELECT SDO_AGGR_UNION(SDOAGGRTYPE(sdogeometry, .05)) ugeom
                         FROM '||TempTable||' b
                         GROUP BY MOD(ROWNUM, 1000)
                     )
                     GROUP BY MOD(ROWNUM, 100)
             )
             GROUP BY MOD(ROWNUM, 10)
     );
```

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

ORACLE
SPATIAL

April 2010

Oracle Spatial

User Conference

Lessons Learned, Future Direction

US CENSUS BUREAU

Helping You Make Informed Decisions

ORACLE SPATIAL

April 2010

Oracle Spatial

User Conference

Lessons Learned

Topology Queries vs. Spatial Functions

- Leveraging topology in large enterprise level spatial databases far out performs traditional spatial operations.

- The topology model is excellent for both database update/maintenance and querying data from a database.

ORACLE SPATIAL

April 2010

Oracle Spatial

User Conference

Lessons Learned

Performance Tips

- Filtering data
- Classifying data
  - Tabular classification
  - Spatial classification
- Parallel processing
- Use of LIMITS
- Frequently COMMIT
- Bind Variables
- Get familiar with explain plans and resulting costs
- Use appropriate indexes.
  - Bitmap
  - BTREE

ORACLE SPATIAL

April 2010

Oracle Spatial

User Conference

Lessons Learned

BTREE vs. Bitmap Indexes

```
CREATE OR REPLACE FUNCTION bitmap_index_candidate(pColumnName VARCHAR2, pTableName VARCHAR2,
pSchemaName VARCHAR2 DEFAULT USER) RETURN BOOLEAN AS
    ColumnName              VARCHAR2(30)         := UPPER(pColumnName);
    TableName               VARCHAR2(30)         := UPPER(pTableName);
    SchemaName                         VARCHAR2(30)         := UPPER(pSchemaName);
    RecCount                           NUMBER;
    CardCnt                            NUMBER;
    MaxRepeatValueCount                NUMBER;
    sql_stmt                           VARCHAR2(4000);
    BitmapIndexNeeded       BOOLEAN              := FALSE;
BEGIN
    -- Figure Out Record Count (RecCount)-----------------------------------------------------------------------------------
    EXECUTE IMMEDIATE 'SELECT COUNT(*) FROM '||SchemaName||'.'||TableName INTO RecCount;
    -- Figure out Cardinality (CardinalityCount) -------------------------------------------------------------------------------
    EXECUTE IMMEDIATE 'SELECT COUNT(DISTINCT('||ColumnName||')) FROM ' ||SchemaName||'.'||TableName INTO CardCnt;
    -- Test to see if Bitmap Index Condition is met -------------------------------------------------------------------------------
    IF ((CardCnt/RecCount) < 0.01) THEN
        BitmapIndexNeeded := TRUE;
    ELSE
        -- Figure out Maximum Repeat Value Count for the column of interest ------------------------------------------------
        sql_stmt := 'SELECT MAX(COUNT(*)) FROM '||SchemaName||'.'||TableName||' GROUP BY '||ColumnName;
        EXECUTE IMMEDIATE sql_stmt INTO MaxRepeatValueCount;
        IF (MaxRepeatValueCount > 100) THEN
            BitmapIndexNeeded := TRUE;
        END IF;
    END IF;
    RETURN BitmapIndexNeeded;
END bitmap_index_candidate;
/
```

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

ORACLE
SPATIAL

April 2010
Oracle Spatial
User Conference

Future Direction with the Cartographic Database

- Used to support publication PDF map products
  - Integration of Oracle Spatial with COTS GIS software
  - Research integration of large national data sets
- Used to support web mapping endeavors
  - Optimize tables for quick data retrieval
    - Spatial sorting techniques
    - Use of partitioned tables
  - Research the best database design to achieve goals

ORACLE
SPATIAL

April 2010
Oracle Spatial
User Conference

Questions?

Nick Padfield
Nick.A.Padfield@census.gov

U.S. Census Bureau
Geography Division
Cartographic Products Branch

USCENSUSBUREAU
Helping You Make Informed Decisions

April 2010

Oracle Spatial

User Conference