

Garmin International Segments and Leaderboards



OVERVIEW

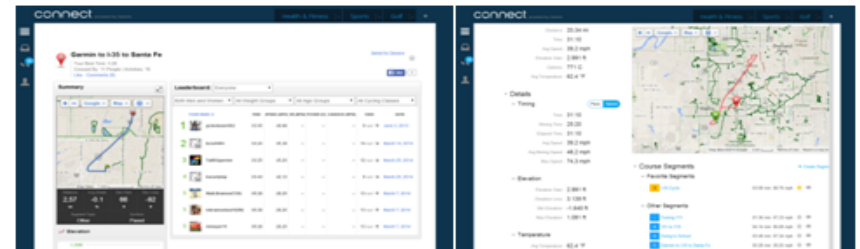
- Allow users to compete for fastest time climbing a hill or sprinting down a straightaway and rank the leaders.
- Users define the starting line, path, and finish line for each competitive “segment”
- Match a user’s fitness activities to segments and rank the activity

CHALLENGES / OPPORTUNITIES

- Be able to match activities to segments within seconds
- Initial loading for 300+ million activities
- Data quality

SOLUTIONS

- Oracle Exadata Machine (Half Rack)
- Oracle Database 11g Enterprise Edition
 - Spatial Option with Linear Referencing System
 - Partitioning
 - Parallel Pipelined Functions



RESULTS

- Stores and simplifies processing of more than 5 billion miles of user activities in a 40TB database
- Able to process and match activities to segments in seconds and able to match new segments to 5+ TB of activity data in minutes
- Allows for additional real time features on Garmin devices (Edge 1000)
- Enables additional reporting features about the use of Garmin Fitness and Wellness products

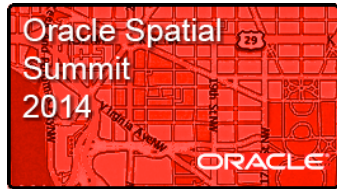
ORACLE®

May 2014
Oracle Spatial Summit

ORACLE®



May 21, 2014
Walter E. Washington Convention Center
Washington, DC USA

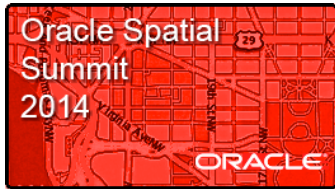


Steve Mitchell

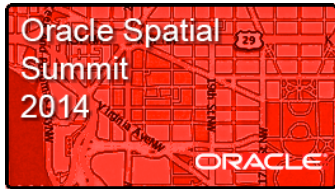
Garmin Software Engineer

Tim Gerber

Database Administrator



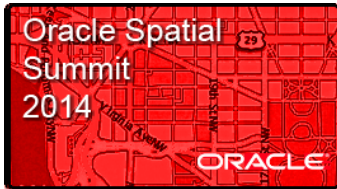
How Garmin Connect Manages and Analyzes 5 Billion Miles of Fitness GPS Data



Program Agenda

- Overview of Garmin's use of Oracle Spatial
- Segments and Leader Boards
- Challenges and their Solutions
- Benefits
- Q&A





Garmin Segments and Leaderboards

Leveraging Oracle Spatial for Challenges

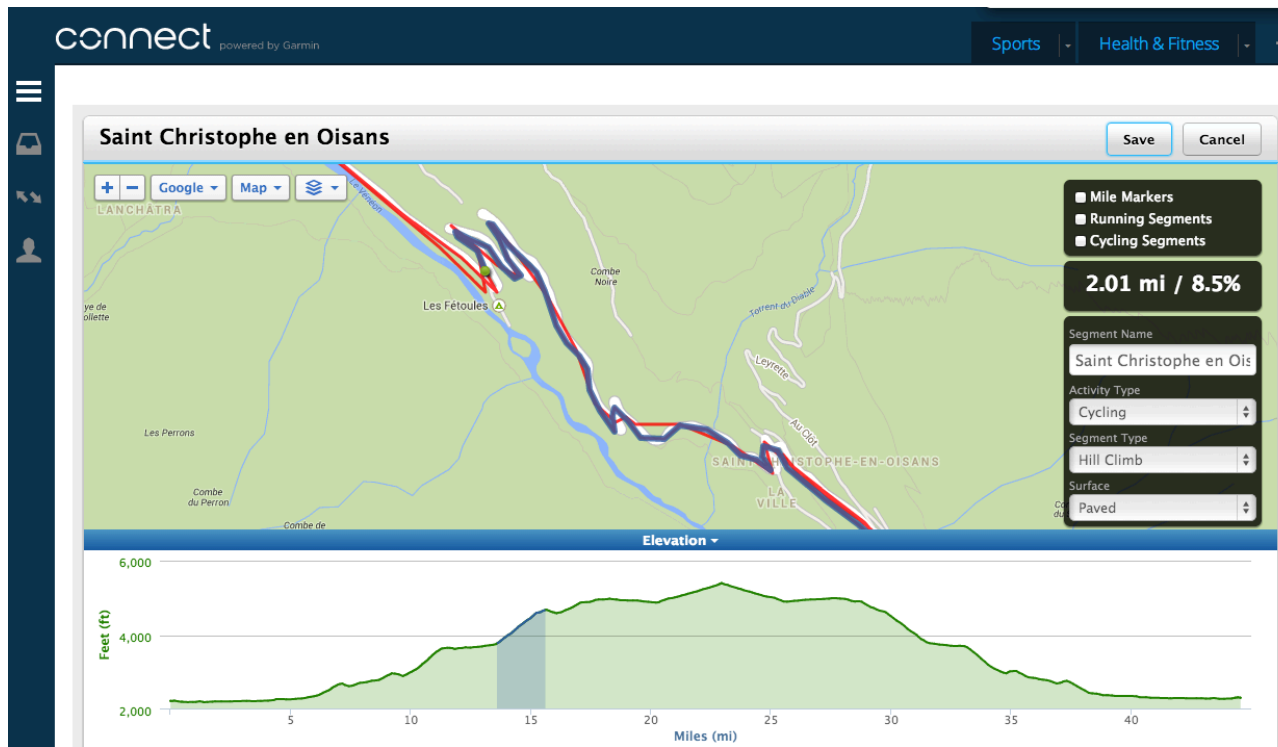
- Customers wanted a way to challenge one another.
- Segments provide a way to compete along a stretch of road or trail.
- Leaderboards rank user activities on a given challenge segment.
- Oracle Spatial and Graph plus Linear Referencing System are used to match user activities to segments and extract the elapsed time for ranking.

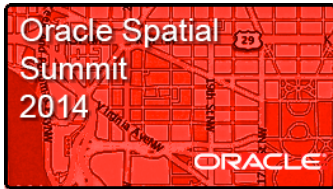




Garmin Connect Segments and Leaderboards

Defining Segments from User Activities





Garmin Segments and Leaderboards

Ranking Users Activities

connect powered by Garmin

Sports Health & Fitness

Saint Christophe en Oisans
Crossed By: 3 People | Activities: 4

Like 0

Summary

Distance	Avg Grade	Elev Gain	Elev Loss
2.01 mi	8.5 %	902 ft	0 ft

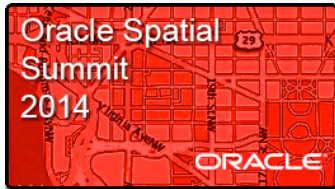
Segment Type: Hill Climb | Surface: Paved

Elevation

Leaderboard: Everyone

Both Men and Women | All Weight Groups | All Age Groups | All Cycling Classes

	TIME	SPEED (MPH)	HR (BPM)	POWER (W)	CADENCE (RPM)	WIND	DATE
1	gctest1	19:05	6.32	144	--	--	17 mph ↓ July 1, 2013
1	gctest2	19:05	6.32	144	--	--	17 mph ↓ July 1, 2013
3	gctest3	20:57	5.75	165	--	65	17 mph ↓ July 14, 2013




GARMIN

V I D E O

Meet the Edge 1000



Garmin Segments and Leaderboards

Distilling Data for the Edge 1000



The file sent to an Edge 1000 is like one drop from a distillery.

There is a lot of energy spent behind the scenes to produce it.

Oracle Spatial drives the process.



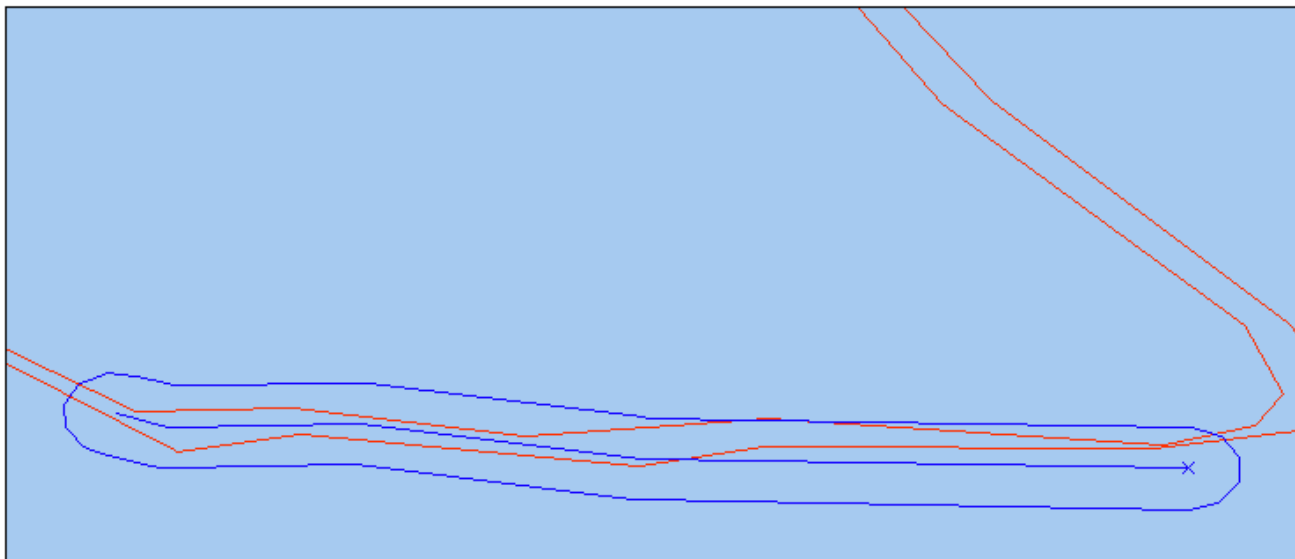


Garmin Segments and Leaderboards

Matching with Oracle Spatial and Graph



Simple Spatial Query Visualizer [source file](#)



Criteria:

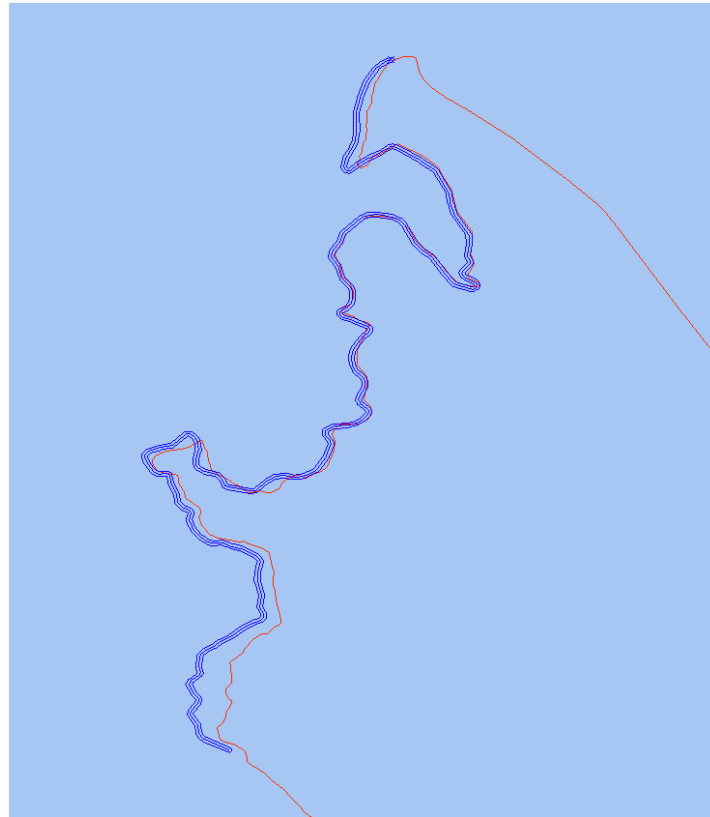
- Start
- End
- Direction
- Path





Garmin Segments and Leaderboards

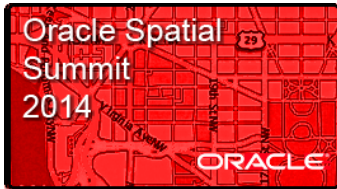
Segment Matching Challenges



GPS Variance:

- Buildings
- Trees
- Canyons
- Satellite drift
- Equipment



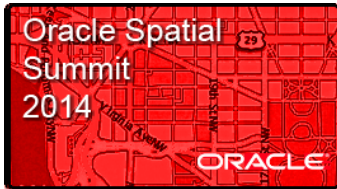


Garmin Segments and Leaderboards

Oracle Spatial and Graph to the Rescue

- `SDO_ANYINERACT` casts a broad net for possible match candidates.
- `SDO_GEOM.SDO_BUFFER` defines the distance from the start & end within which a candidate must pass.
- PL/SQL loops through sections of path between the start and end to make sure the paths roughly match.



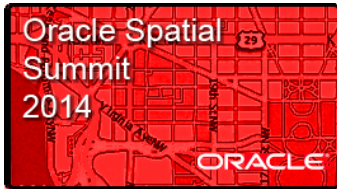


Garmin Segments and Leaderboards

Oracle Linear Reference System to the Rescue

- Needed a way to measure elapsed time.
- LRS time measure gives time entered & exited.
- Time comparison tests directionality.
- `SDO_LRS.REDEFINE_GEOM_SEGMENT` function allows switching to distance measure when necessary for other uses.





Garmin Segments and Leaderboards

Building an Oracle Spatial and LRS Solution

- Define data type (Oracle LRS w/time measure)
- Spatially index data
- Load test data.
- Validate data.





Garmin Segments and Leaderboards

Define USER_SDO_GEOM_METADATA

```
INSERT INTO USER_SDO_GEOM_METADATA (
    table_name, column_name, diminfo, srid
) VALUES (
    'SEGMENT', 'POLYLINE',
    SDO_DIM_ARRAY (
    SDO_DIM_ELEMENT('Longitude', -180, 180, 0.05),
    SDO_DIM_ELEMENT('Latitude', -90, 90, 0.05),
    SDO_DIM_ELEMENT('Time', 0, 0, 0.05)), 8307);
```





Garmin Segments and Leaderboards

Loading Spatially Indexed LRS Data

```
new JGeometry(  
  3302, // Line string; 3D (x, y, m), 3rd is dimension  
  8307, // SRID for latitude/longitude (8307)  
  {1,2,1}; // offset, line, straight,  
  SDO.ordinates(geom)  
);
```





Garmin Segments and Leaderboards

Validating Data

```
SELECT COUNT(*) FROM  
GEO_ACTIVITY a  
WHERE  
SDO_GEOM.VALIDATE_GEOMETRY  
( a.polyline, .005) = 'FALSE';
```





Garmin Segments and Leaderboards

Fixing Problems

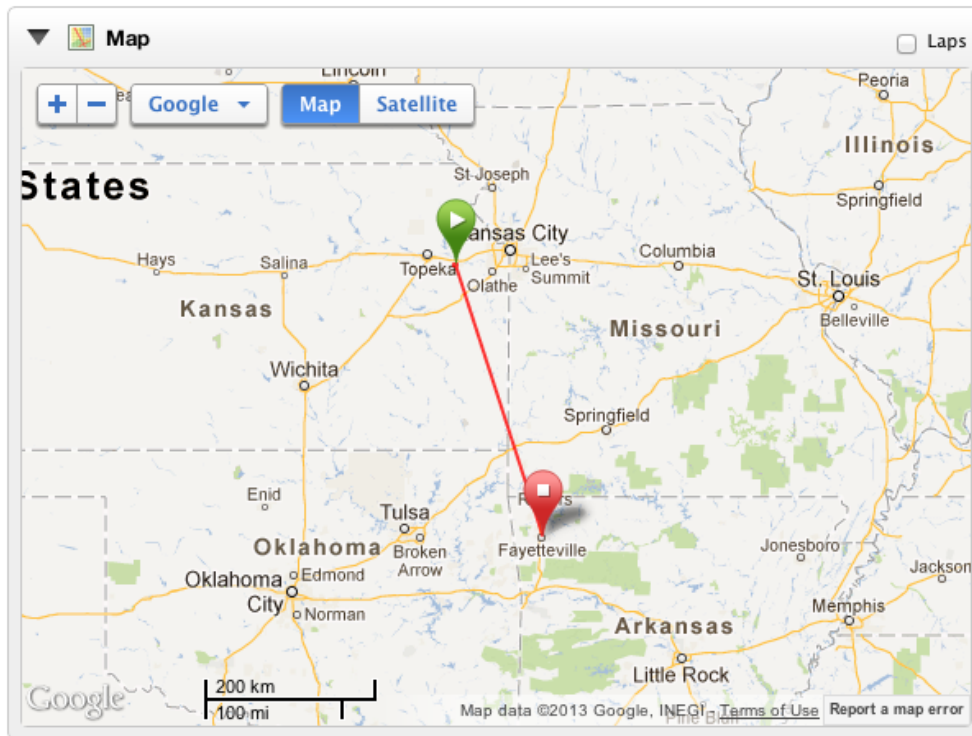
```
UPDATE GEO_ACTIVITY a  
SET a.polyline =  
SDO_UTIL.REMOVE_DUPLICATE_VERTICES  
( a.polyline, .005)
```





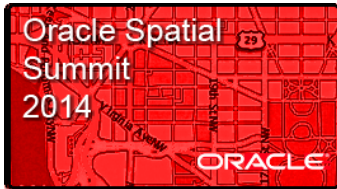
Garmin Segments and Leaderboards

Identifying User Data Problems



User paused multisport device at a triathlon in one city and then resumed it at home.





Garmin Segments and Leaderboards

Hardware Data Problems

▼ **Summary**

Distance:	13,343.84 mi
Time:	:00
Avg Speed:	-159,217.5 mph
Elevation Gain:	425 ft
Avg Temperature:	-3,471,668.6 °F

▼ **Details**

▼ ⌚ **Timing** Pace Speed

Time:	:00
Moving Time:	596:31:24
Elapsed Time:	:00
Avg Speed:	-159,217.5 mph
Avg Moving Speed:	2.5 mph
Max Speed:	4.3 mph

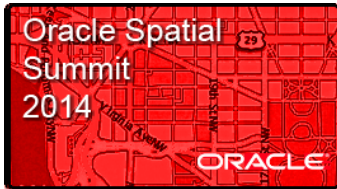
▼ 📈 **Elevation**

Elevation Gain:	425 ft
-----------------	--------

▼ 🗺 **Map**

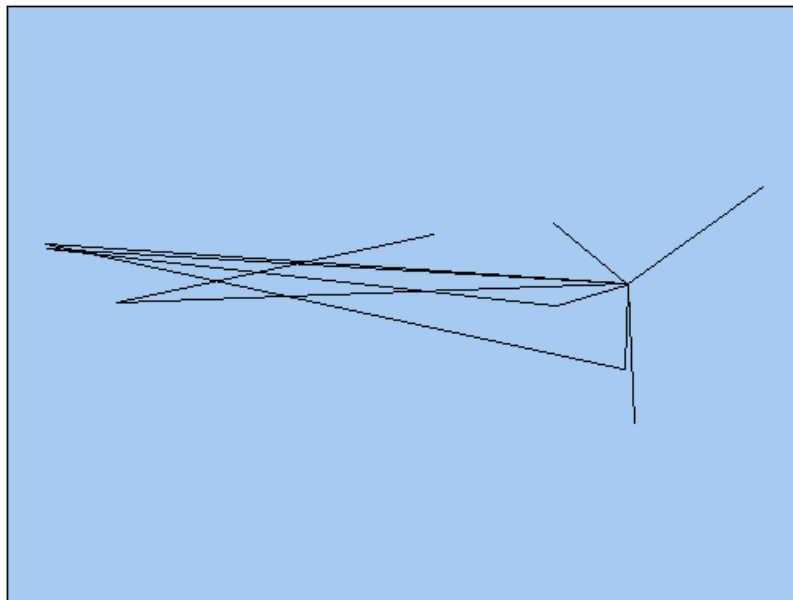
+ - Google Map Satellite

A Google Map showing the Pacific Ocean region. A red line is drawn from Thailand in Southeast Asia to the United States in North America. The map includes labels for Russia, Mongolia, China, South Korea, Japan, Thailand, Indonesia, Papua New Guinea, Australia, Canada, United States, Mexico, Colombia, Peru, and Chile. The North Pacific Ocean and South Pacific Ocean are also labeled. A scale bar at the bottom indicates 2000 km and 2000 mi. Map data is attributed to ©2013 MapLink, Tele Atlas - Terms of Use.



Garmin Segments and Leaderboards

Export/Import Data Problems



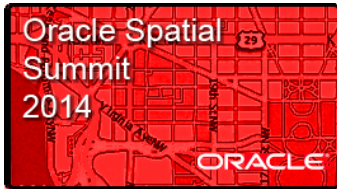
20,696 mile polyline made up of 134 activities exported from Training Center desktop application.

Click on the map to: A set of small, light blue icons for map navigation, including a magnifying glass, a location pin, a compass, and a refresh symbol.

Datasource: map width: height: AA

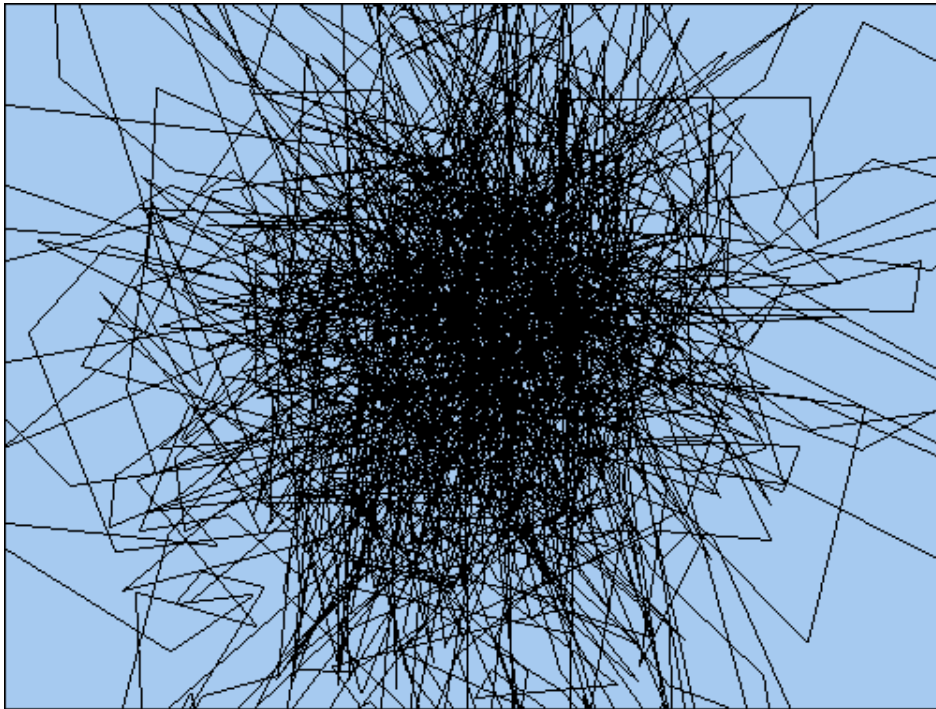
```
query 1: select POLYLINE_ORIGINAL from GEO_ACTIVITY where ACTIVITY_PK = 194987205
```





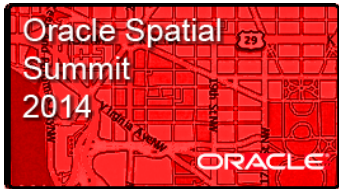
Garmin Segments and Leaderboards

Identifying User Data Problems



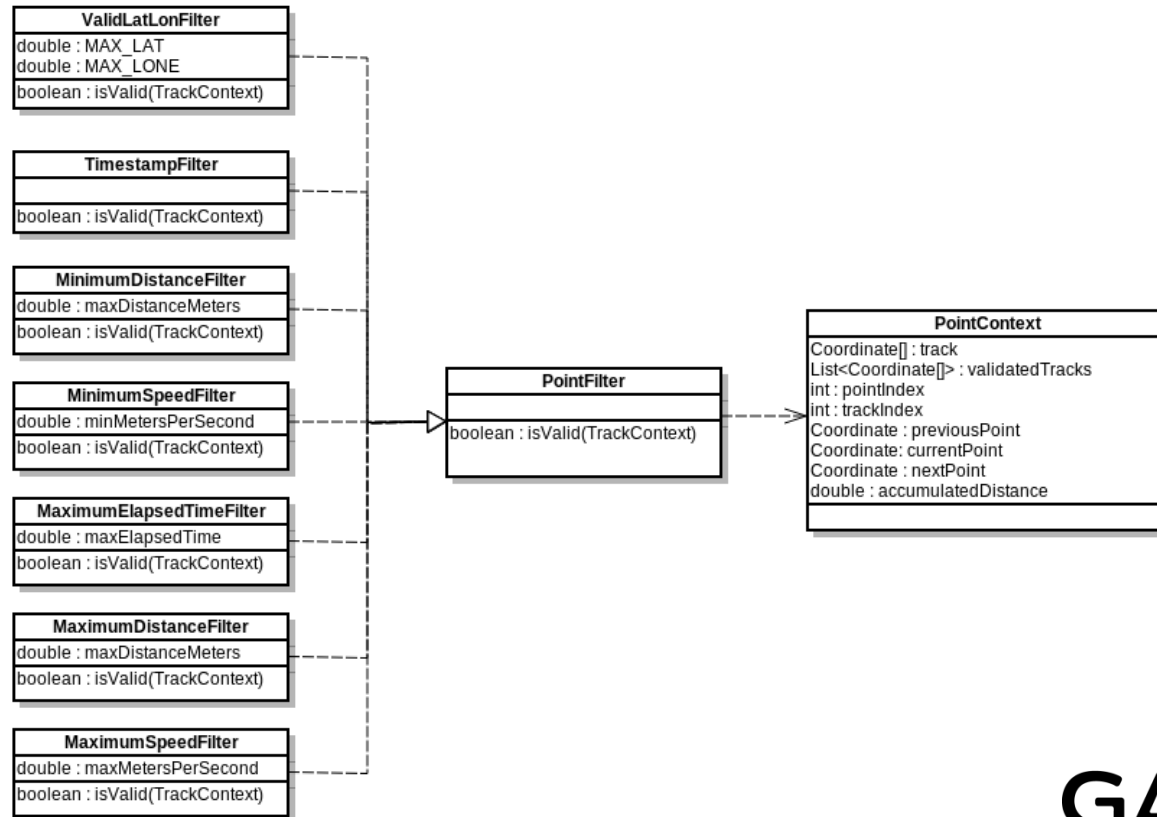
Indoor treadmill activity with GPS enabled.





Programming Solution

Point Filter Class Diagram





Garmin Segments and Leaderboards

Segment Summary

- Segments give users a way to compete with each other.
- Oracle Spatial and Graph allows us to match user activities to segments.
- Oracle LRS allows us to measure performance.
- The difficult part was provisioning clean data and overcoming differences in GPS data.





The World of Partitioning

Divide and Conquer the Data

- The Garmin Connect database is 40 TB and growing at a pace of 2 TB per month.
- Of this 40 TB, 5 TB is spatial data used for segments and leader boards.
- Putting this data into a single non-partitioned table is unrealistic.
- To get the best performance, we needed to partition the tables.





The World of Partitioning

Our Partitioning Strategy

GEO_ACTIVITY
(Partitioning Key: ACTIVITY_TYPE PK, START_TIME GMT)
Partition High Value: 1, 01/01/2009
Partition High Value: 1, 01/01/2010
Partition High Value: 1, 01/01/2011
Partition High Value: 1, 01/01/2011
.
Partition High Value: 1, 03/01/2014
Partition High Value: 1, 04/01/2014
Partition High Value: 2, 01/01/2009
Partition High Value: 2, 01/01/2010
Partition High Value: 2, 01/01/2011
Partition High Value: 2, 01/01/2011
.
Partition High Value: 2, 03/01/2014
Partition High Value: 2, 04/01/2014

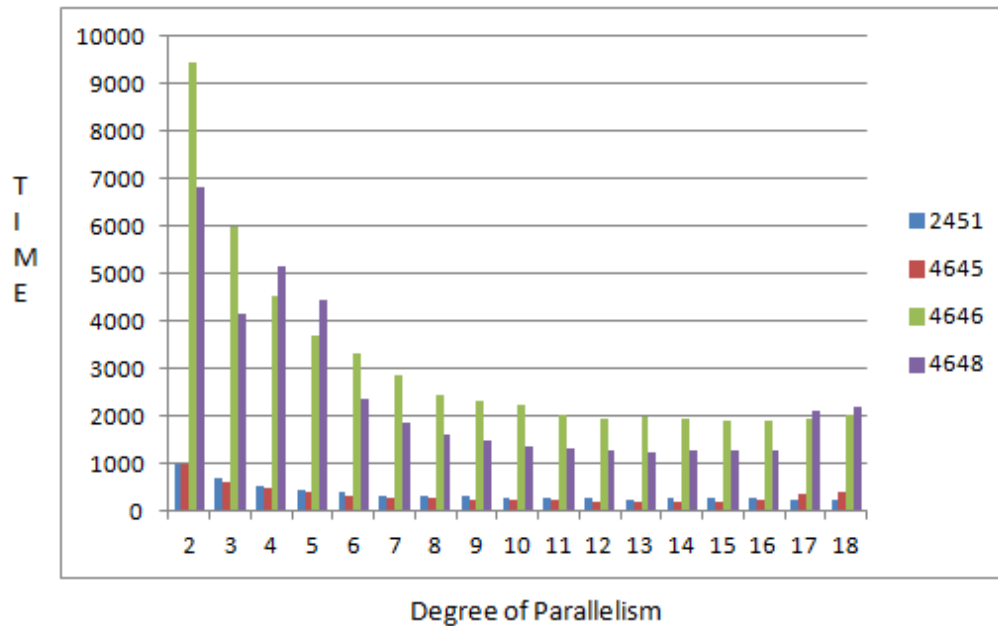
- All of our queries will be based around activity type (i.e. running and cycling) and date.
- Spatial indexing allows for partitioning but not sub-partitioning.
- Therefore we chose to use a composite key of activity_type, date with range partitioning.



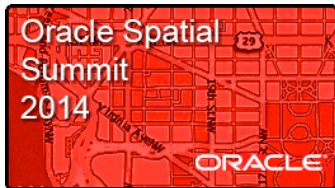


Power of Parallelism

Run Time of Segment Matching

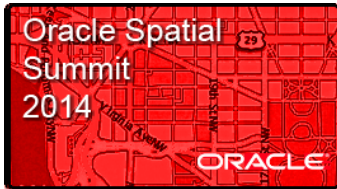


Running tests with different degrees of parallelism helped drive out the optimal number of parallel processes for the segment matching.



Power of Parallelism

SEGMENT NAME	PARALLEL 2	PARALLEL 4	PARALLEL 8	PARALLEL 10	PARALLEL 12	PARALLEL 16
tour de longchamps	108123	44082	21977	17825	16254	17041
Garmin Velothon Berlin	963	539	315	273	256	260
Bugaksan Mountain	6809	5137	1624	1380	1270	1255



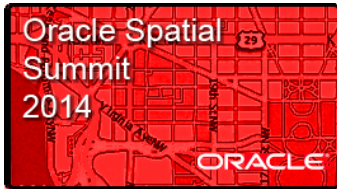
Power of Parallelism

Parallel Pipelined Functions

- All of our segment matching is using a PL/SQL package.
- To help performance within the package, we have taken advantage of parallel pipelined functions.

```
FUNCTION get_sections_for_segment (segment_pk_v      NUMBER,  
    buffer_dist      NUMBER,  
    interval_in_meters  NUMBER,  
    max_skip_percent  NUMBER,  
    source_table_cursor IN for_segment_cursor_type)  
RETURN activity_segment_table_type DETERMINISTIC  
PIPELINED PARALLEL_ENABLE  
(PARTITION source_table_cursor BY HASH (geo_activity_pk))
```





The Data Load Challenge

The Problem

How do you process, analyze and load 300 million fitness activities in a minimal time frame all while new activities are being created?





The Data Load Challenge

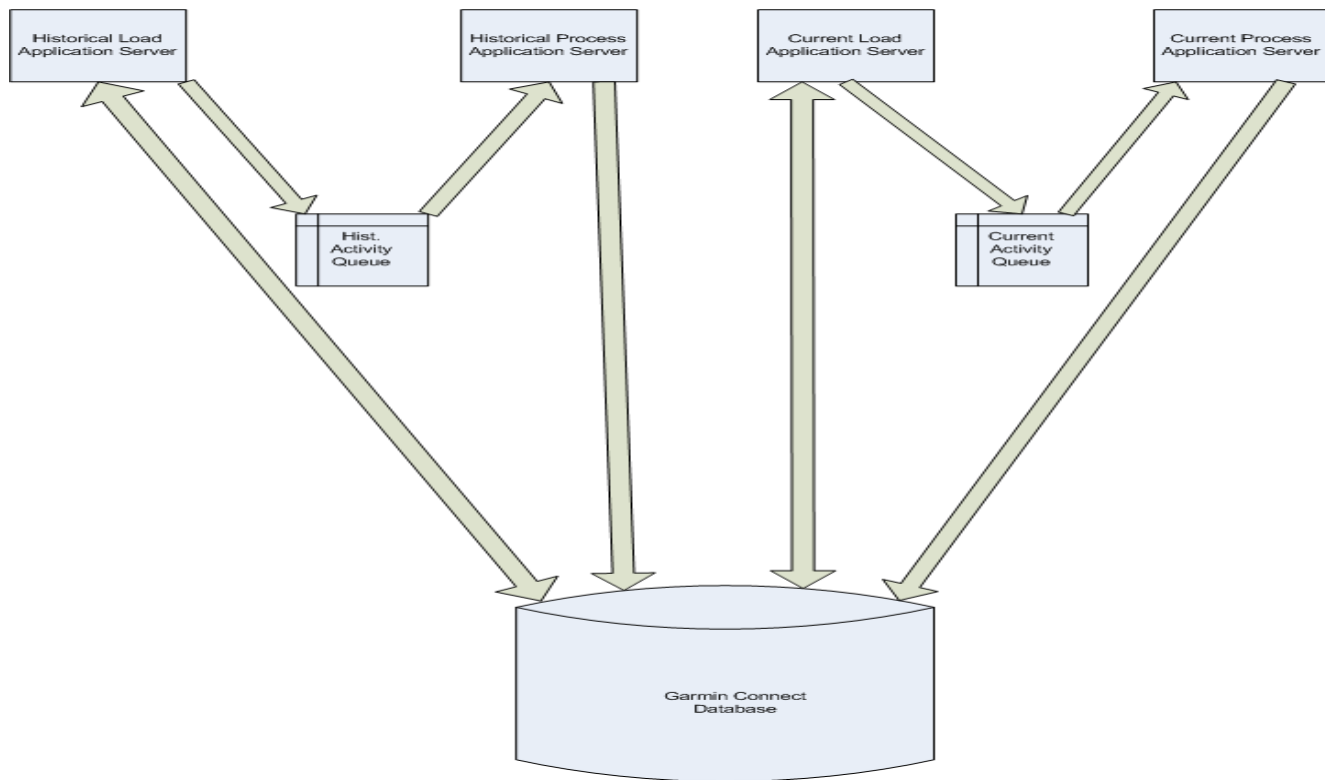
The Chosen Solution

- We were able to utilize several queues to place current activities being created in one queue and the historical activities in another queue.
- Processed both queues continuously until geo activities for all activities were created.
- We were able to process, analyze, load all 300 Million fitness activities in the database in under 20 days.





The Load Process





Benefits

The Results are In!

- Provides “clean” versions of our data.
- Able to provide users results within seconds.
- Significantly less time and resources consumed to get to production over writing custom code.
- Allows real time features for Garmin products.
- Enables additional reporting information about Garmin products.



Q&A

[View the Video of the Garmin Edge 1000 in action.](#)