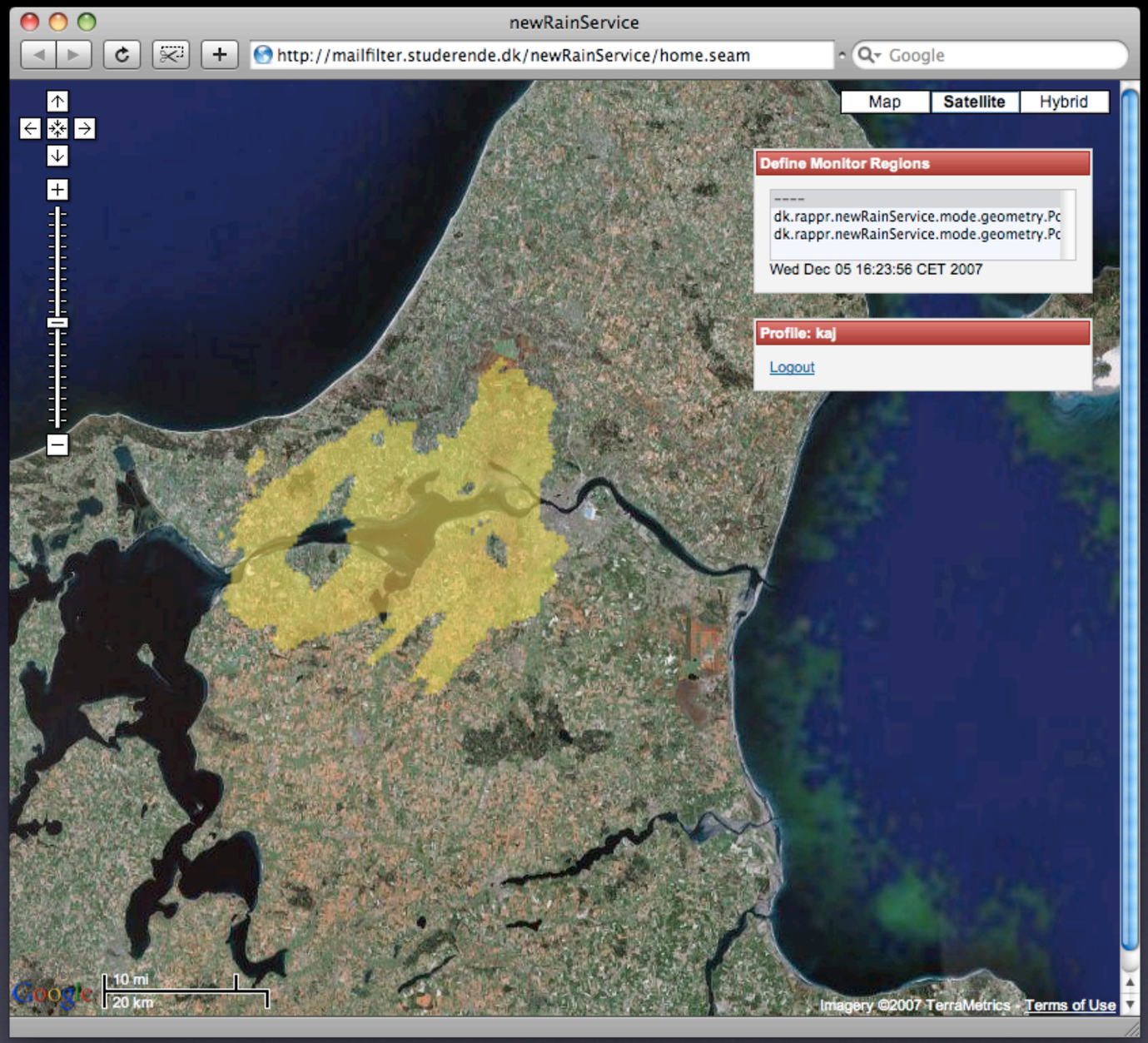
Weathr


newRainService
http://mailfilter.studerende.dk/newRainService/home.seam
Google
Map
Satellite
Hybrid
Define Monitor Regions
dk.rappr.newRainService.mode.geometry.Pc
dk.rappr.newRainService.mode.geometry.Pc
Wed Dec 05 16:23:56 CET 2007
Profile: kaj
Logout
10 mi
20 km
Imagery ©2007 TerraMetrics - Terms of Use

Presentation overview

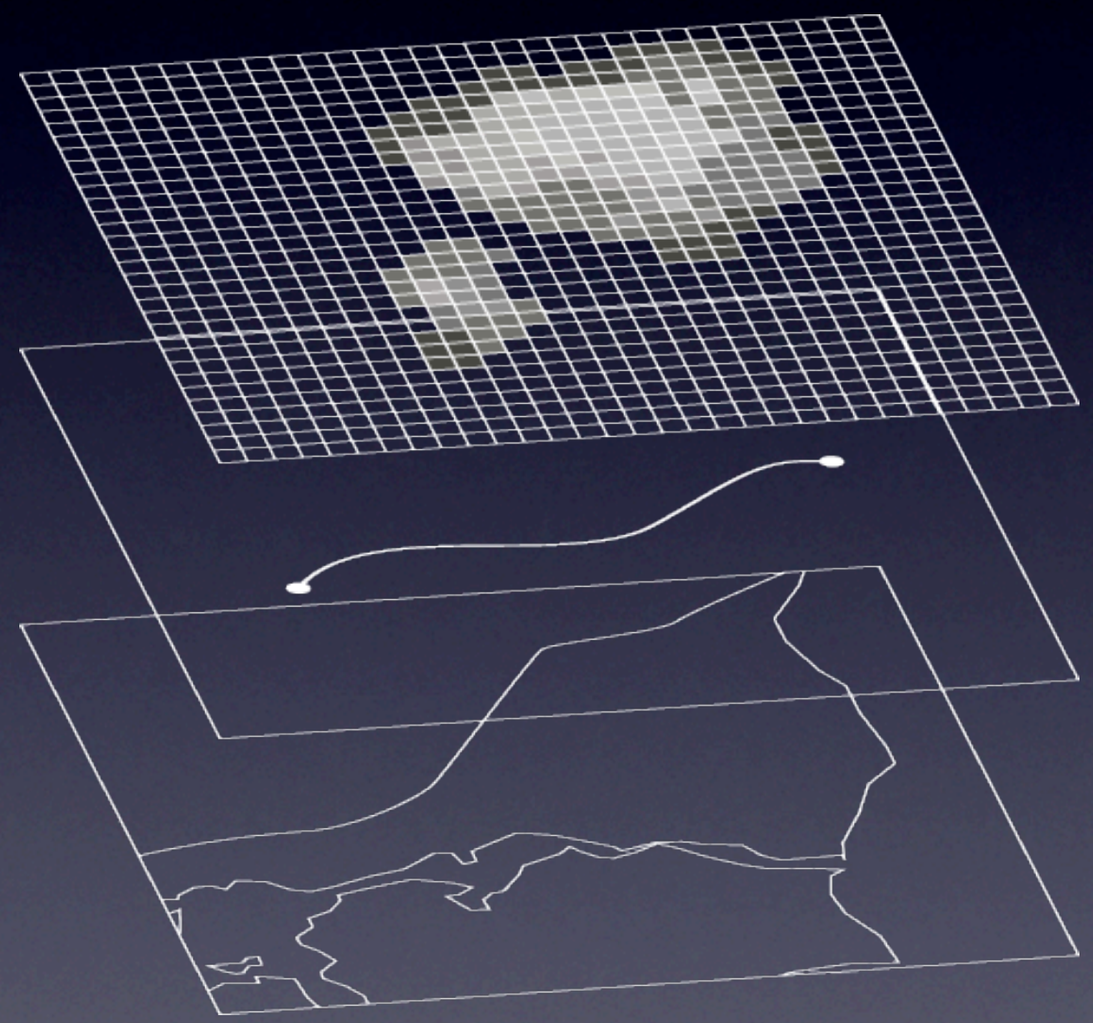
- Introduction
- Project contributions
- Software architecture
- Abstract data types (ADTs)
- Performance studies
- Related work
- Status
- Future work

Introduction

- Data comes from a Local Area Weather Radar (LAWR)
- We warn users of upcoming rainfall

Introduction

- 240×240 matrix
- Each cell is 500m×500m and has a value [0;255]
- We query:
  - Single cells
  - Routes
  - Polygons



Contributions

- Grand scale vision:
  - Warn StreamSpin users of rainfall
  - Route recognition
  - Warn web users of rainfall
  - Scalable up to *n* radar stations and *n* users with personal profiles

Contributions

- More specifically, we contribute with:
  - A first-step prototype
  - Modularized framework
  - Efficient ADTs (storage & query)

Software Architecture

Initial Considerations

- GeoTools (Java Library)
  - Highly advanced Java API implementing the GeoAPI.
  - Being advanced is also its downside.
- Oracle Spatial and GeoRaster.
  - Not interesting at this point, as we expect to gain more from processing queries in memory for current data.

Goals

- Obtain a flexible architecture facilitating:
  - support for different types of storage devices
  - support for storing grids in various formats (binary, ascii, compressed etc.)
  - support for dynamic ADT loading
    - ADT plugins
  - support for both in-application querying and delegating logic to DBMS (stored procedures)
  - support for both “stream-based” query and “request-based” query (inspiration: XML => SAX vs DOM)
- Not all features will be leveraged in this coming release, but will serve as bricks for coming thesis work

Implementation Details

- Implementation using the Seam Framework
  - Component based application framework making a “seam”-less integration of EJB3 standard, JSF (Facelets) and several other open source products
- Development by configuration
  - Employes several well-known patterns without the usual tedious boiler-plate code (intensive use of annotations).
  - IoC - Inversion of Control (Dependency Injection and Outjection) provides loose coupling
- Currently deployed for public on a Pentium 3 with 512MB of memory and SCSI drives running a JBoss AS at:

http://mailfilter.studerende.dk/newRainService/

Components

- Main Controller
- Data Monitor Component
- Data Access Layer
- StreamSpin Component
- Versatile parse mechanism with callback events
- Spatio Temporal Data Access Layer

Data Monitor Component

- Initialized automatically when framework itself has concluded initialization leveraging the EJB Timer mechanism
- Polls remote store via SFTP for new incoming data with a given interval (10 secs)
- On new available data, file is downloaded and `datareceived` notification is emitted
- Executed in separate thread
- Application Scoped

Weathr Controller Component

- Responsibility of the general application flow dealing with monitoring and actions on incoming data
- Observes the monitor for new data and initiates initial parsing registering relevant callback handlers:
  - Image generation (saves png image to disk for the UI)
  - User profile querying (querying all saved profile geometries on parse time)
  - ADT generation and registration
    (user provides a query to the system)

Data Access Layer Component

- For storage and retrieval of historic data
- Employs a DAO style architecture for data-access implementing a common data access generic typed interface.
  - Date range selects, returning the single nearest in time
  - Singe point in time with a given +/- threshold, returning single nearest in time
  - Select specific dataitem given some ID (eg. filename or integer value). Must be defined through generic type argument
  - Insert (save)
- Currently only a FileSystemWeatherGridDAO is fully implemented, and a generic DBMS is partially present.

<T extends AbstractWeatherGrid

Data Access Component

Registered Grid File ContentHandlers

Grid Data Stream Parser

Database Connector

FileSystem Connector

..... Connector

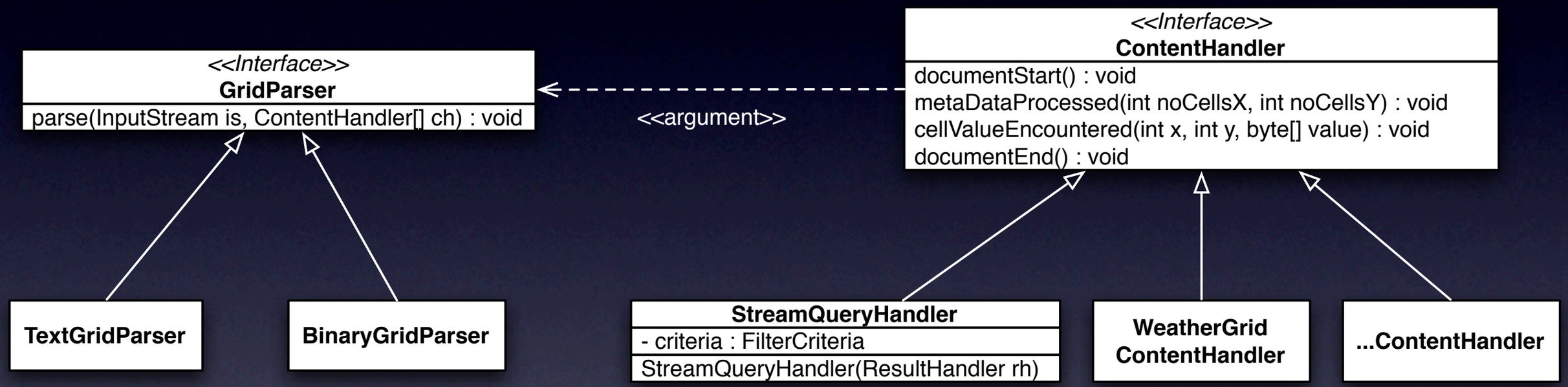
DBMS

Filesystem

...I.5

Parser

- Versatile Grid stream parsing mechanism using callback handlers
- Heavily inspired by SAX for stream based XML parsing.
- Parser Interface describes only one method
  `parse(InputStream is, ContentHandler ch)`
- Each implemented parser must call 4 methods on the ContentHandler at given stages of the parsing process
  - `startDocument(), endDocument(), metaValuesProcessed(GridMetaData gmd), cellValueEncountered(int x, int y, byte value)`



Parser w/ callback event-mechanism

Customizing behaviour at parse time

StreamSpin Component

- Handles all communication with the Streamspin service. Incoming and outgoing.
- Dispatches incoming user locations
- Submission of user messages
  - Creates new service and signs up user for submission, submits message and destroys service

Needs to be done

- Constrained by unknown factor regarding provision of forecasts in realizing a true service.
- More specialized Data Access Layer which suits on top of our own Query engine and the possibility of delegating queries to a DBMS leveraging fx. Oracle Spatial and GeoRaster or any other through eg. Stored Procs.
- Treats our query engine as a Source of Data.
  - Interface is partially defined:
    - `getGridsWithRain(Geometry geom, [int minutesInAdvance])`
    - `isRaining(Geometry geom, [int minutesInAdvance]);`

Future version

- Statistical queries on historic data. Wrapping data access layer must implement methods such as:
  - `getRainOnGeometry(Geometry geom, Date from, Date to)`
  - `getAvgRainOnGeometry(Geometry geom, Date from, Date to)`
  - `...`

ADTs

Weather Radar

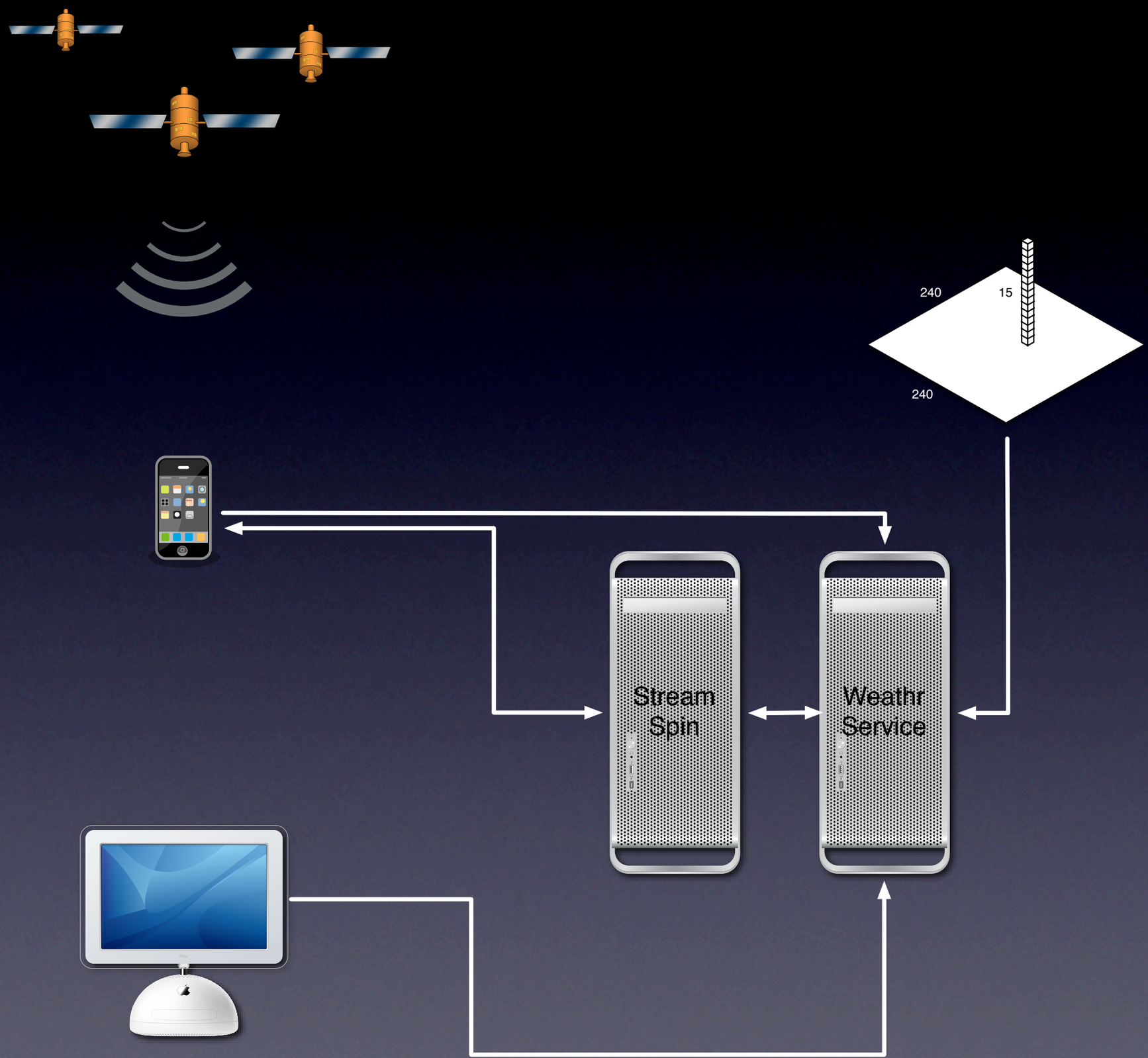
Frequency [MHz]	9410 ± 30 MHz
Output Power [kW]	25
Maximum range [km]	60 km
Grid resolution (pixel size) [meter]	500x500(60 km range) 250x250 (30 km range) 300x300 (15 km range) 100x100 (15 km range)
Image Frequency [minutes]	1 or 5
Antenna- Slotted wave guide [m]	2.44
Horizontal beam width (BWH)	0.95°
Vertical beam width (BWV) (measured from horizontal)	±10°
Rotation Speed [rpm]	24
Power consumption [kWh/ year]	5700

Data Set

- Data set of 240×240 fields (57600)
- Precipitation is specified as a numeric value ranging from 1 to 255
- No precipitation is expressed by the numeric value zero (0)
- A new data set is received each five minutes through an FTP connection

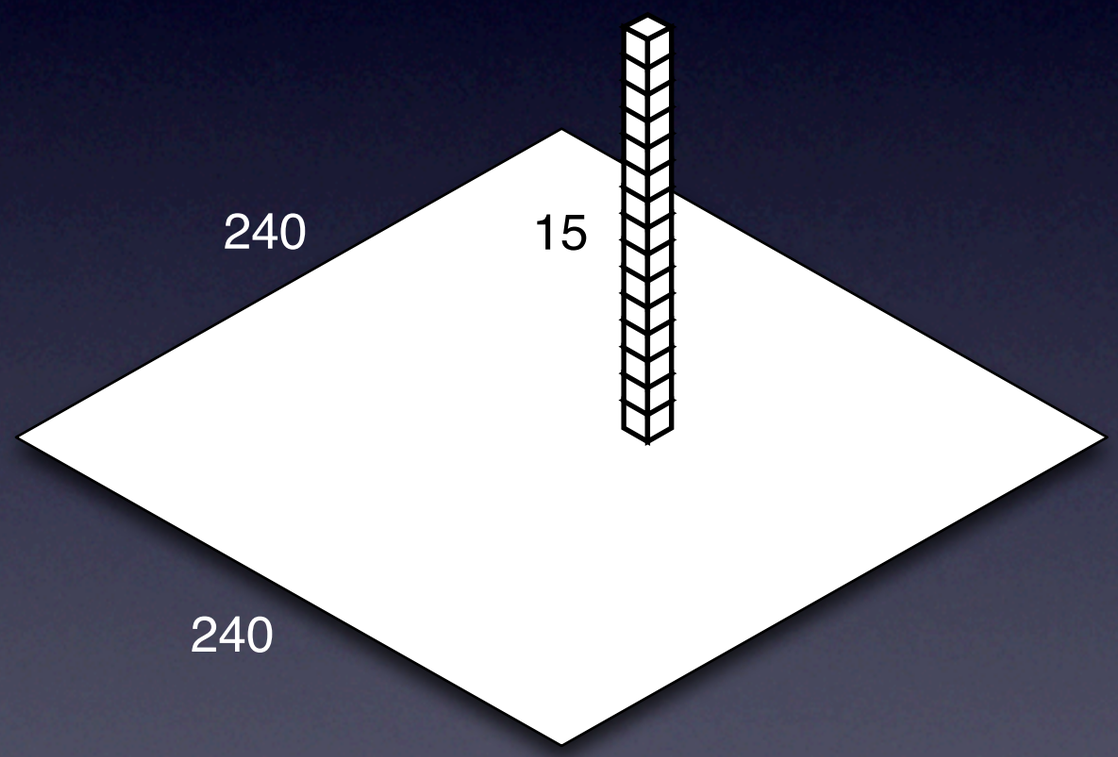
Data Set Statistical Analysis

- Close to three weeks of data sets have been analyzed:
  - 72,4% of the rows consist of all zeros
  - 97,2% of the fields in a data set contains a zero
  - 1,8% of the data sets is a complete set of zeros
    - Some noise values occur


240
15
240
Stream
Spin
Weathr
Service

ADT: Grid

- `byte[][][]`
- x-axis (North-South)
  — currently 240
- y-axis (East-West)
  — currently 240
- Time intervals
  (Forecasts)
  — currently 15



ADT: Hash Map

- HashMap <Key, Value>
- <Key> consists of a coordinate pair *(x,y)*, specifying location in the weather grid
- <Value> consists of an array of bytes. The array describes the precipitation in a single cell in time
- Number of forecasts [15, 70 minutes into the future]
- Forecast interval [5 minutes]

Minutes	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Precipitation	-	2	23	58	93	72	41	11	-	-	-	-	-	-	12

Performance Studies

- Space vs. precipitation concentration of ADTs
- Performance
- Single-cell queries.
- Route (multiple cells) queries
- Polygon queries for the various ADTs
- Polygon, cell vs. MBR
- Caching queries
- Caching precipitation data

Related Work

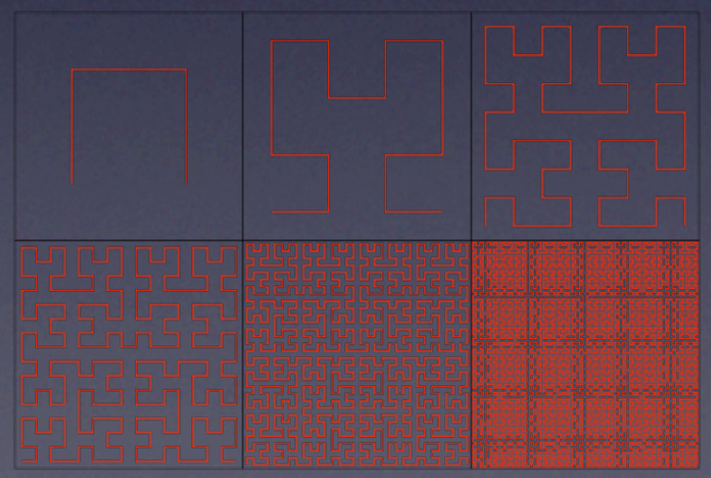
- Different spatial query types:
  - Stationary queries on moving objects
  - Moving queries on stationary objects
  - Moving queries on moving objects
- We differ since our data is always complete
  - No prediction is necessary
  - There are no moving objects

Related Work

- Main memory: Implementation of data structures in main memory
- Sparse data: Sparse data is by nature easily compressed
  - 97,5% zero cells
  - 72,4% zero-rows (October 2007)

Related Work

- Space filling curves:
  - Curve that visits every point/cell in a space
  - Maps multidimensional data to one dimension.
- Map e.g. *(x,y)* to a derived key: *(110,100) ⇒ 101110*



Related Work

- Other indices to index spatial data:
  - B-trees
  - B^+-trees
  - R-trees (MBR)

Related Work

- Other Similar projects:
  - IBM Thomas J. Watson Research Center: Visualization of data from a weather simulation
  - Departement de Linguistique Universite de Montreal has researched in synthesizes marine weather forecasts

Status

- The application: Nearing "completion"
- Article outline:
  - Introduction, related work, software architecture and implementation, StreamSpin, ADTs, performance study, conclusion, and future work

Future Work

- Statistics on historical data using a DBMS back-end
- In-memory solution is infeasible
- E.g. a farmer wants to monitor the amount of rainfall his land has received
- Oracle provides GeoRaster for raster (grid) data

Future Work

- Support for multiple radars
- Handle overlapping data grids
- Scalability
- Variable grid cell size

Feedback

- We would appreciate feedback on:
  - Abstract data types
  - Related work (anything comes to mind?)