Efficient Data Dissemination through a Storageless Web Database

Thomas Hammel

Bernadette Yetso

prepared for DARPA SensIT Workshop

7 November 2002

Fantastic Data

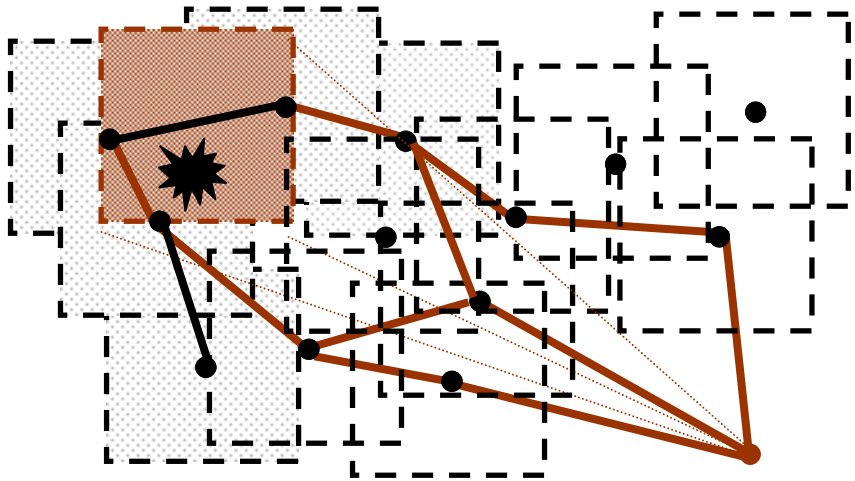
207 Prospect Avenue, San Francisco, California 94110, USA • (415) 643-9555 • (415) 643-9554 • www.fandat.com

What cache does for you

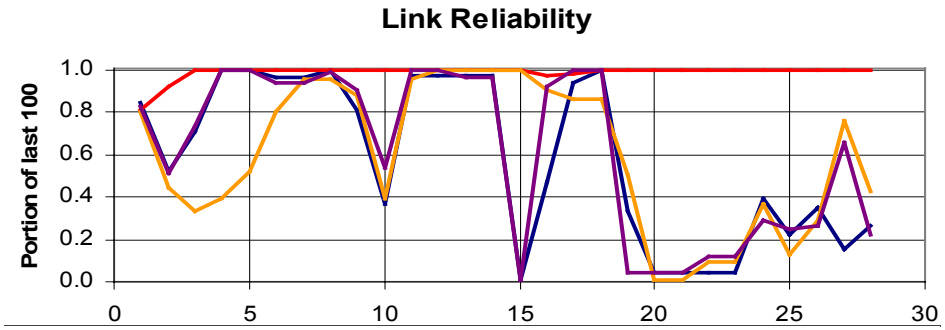
- **Data storage**
  - **producer and consumer do not need to directly communicate**
  - **easy to arrange data replay for test and debug**
- **Data dissemination**
  - **all data is available for remote access**
  - **filters automatically determined to satisfy application queries**
- **Primary key for data naming**
  - **automatic consistency enforcement**
  - **data stream merging**
- **Multiple access methods**
  - **real time notification of changes**
  - **search and extract (by query with where clause)**
- **Partial access to structures**
  - **subset of fields**
- **Implemented safely as a server for multiple clients**

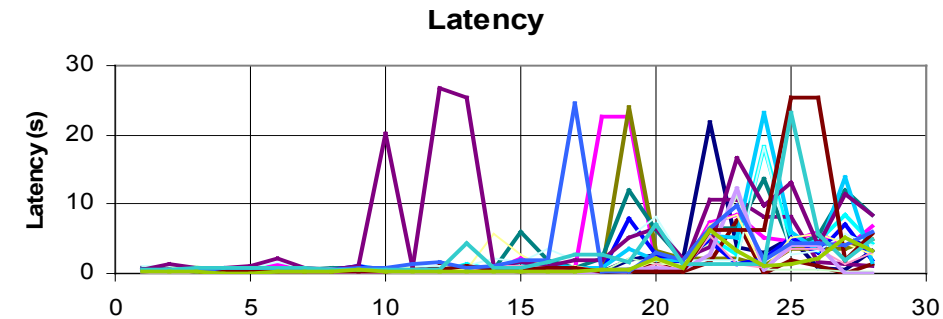
Data Dissemination

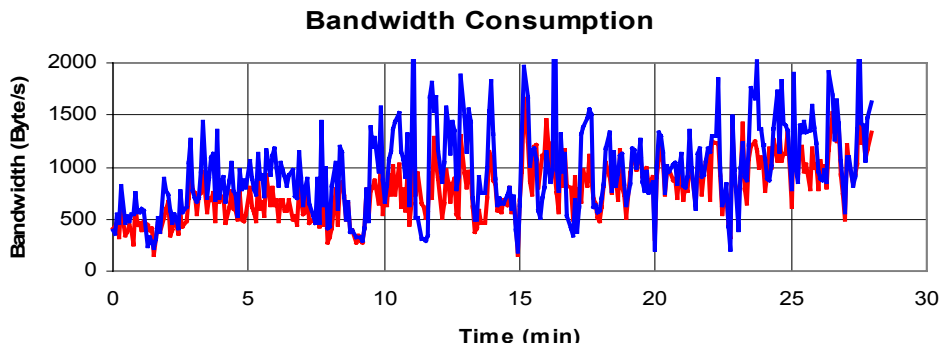
- **Dissemination triggered by interest**
  - **tracker**
  - **UMd gateway and display**
- **Data attribute routing**
  - **specified by query where clauses**
- **Guaranteed delivery for data as long as it remains valid**
  - **uses redundant paths through network**
  - **may send multiple times if link reliability is low**
  - **not all data will be delivered, some will become obsolete or expire before delivery**
  - **suppress on overhearing**
- **Using both Sensoria radios**



Performance Metrics







- Data cache measures performance continuously
  - delivery latency
  - bandwidth consumption
  - memory consumption
  - processor consumption
  - number of gaps
  - number of retransmissions
  - link quality
- Some metrics affect future performance, e.g. link quality
- Want to accurately measure these metrics on a stable network

Sample data from node 11
27 October 2002

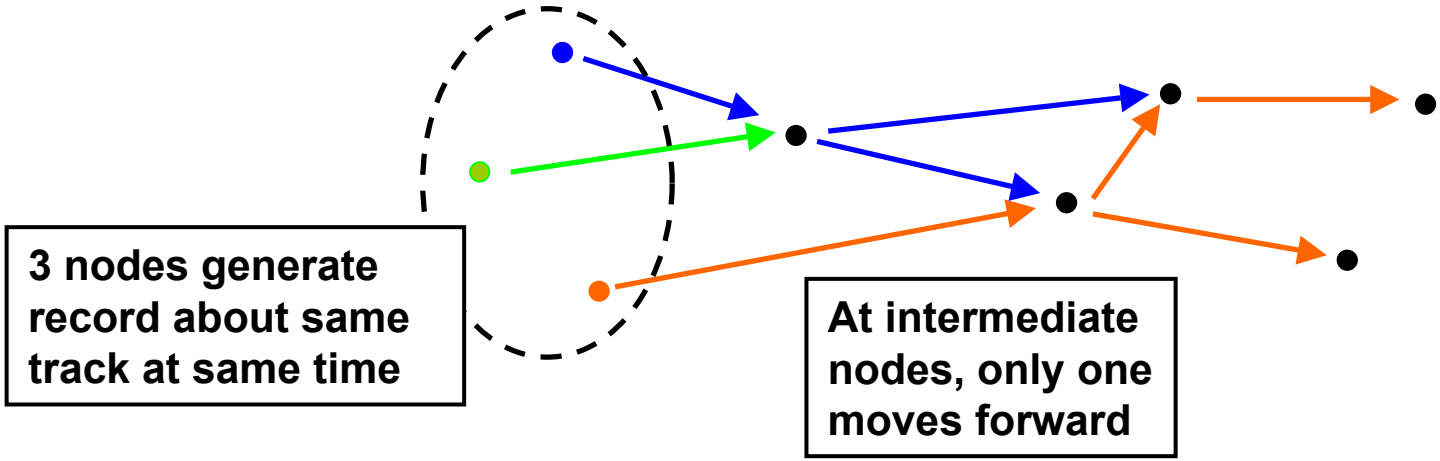
Fantastic Data

207 Prospect Avenue, San Francisco, California 94110, USA • (415) 643-9555 • (415) 643-9554 • www.fandat.com

Data naming

- **primary key in application's name space**
  - **correct naming allows merging of data streams enroute**

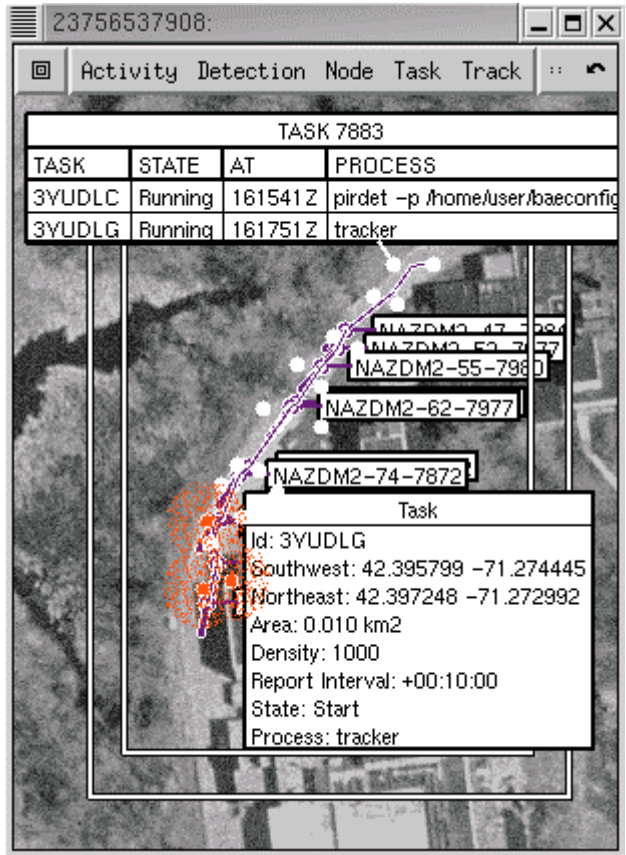
create table track (id c, time u32, … , primary(id,time));



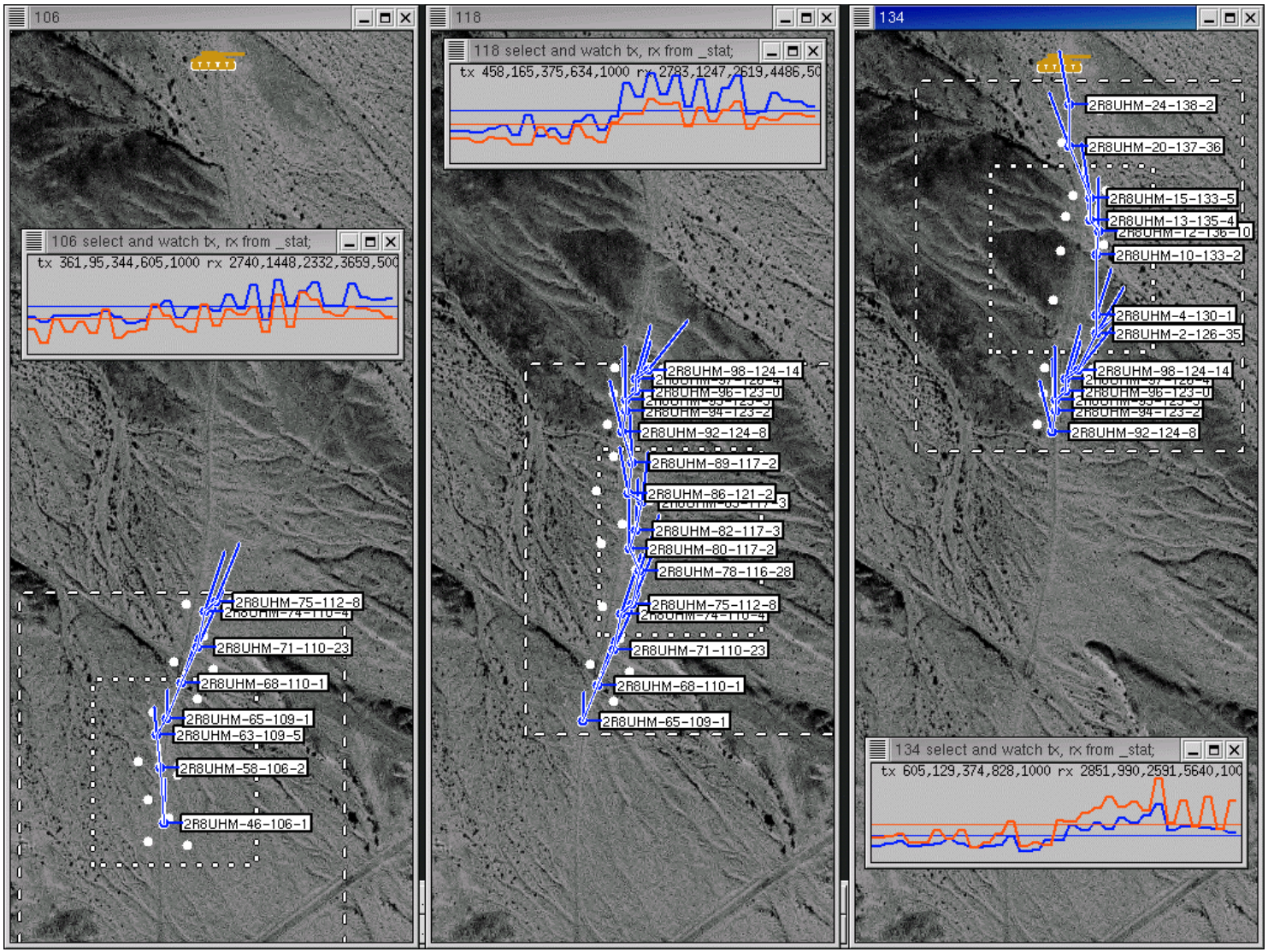
- **sequence number in database's name space**
  - **used for bookkeeping**
  - **consistency enforcement**
  - **retransmission and repair**

Task management

- **remote task control**
  - **area of operation (may overlap)**
  - **density**
  - **status: start, stop, don't care**
  - **process specification: executable, arguments**
- **task specification inserted into network (reverse of query)**
- **coordinates with local neighborhood to start and maintain desired task density**
  - **random task start**
  - **continuous task supervision**
  - **scheduled task status reporting to local neighborhood**
  - **task and node failure detection**
  - **restart on another node to maintain density**
  - **may stop local tasks if too dense**

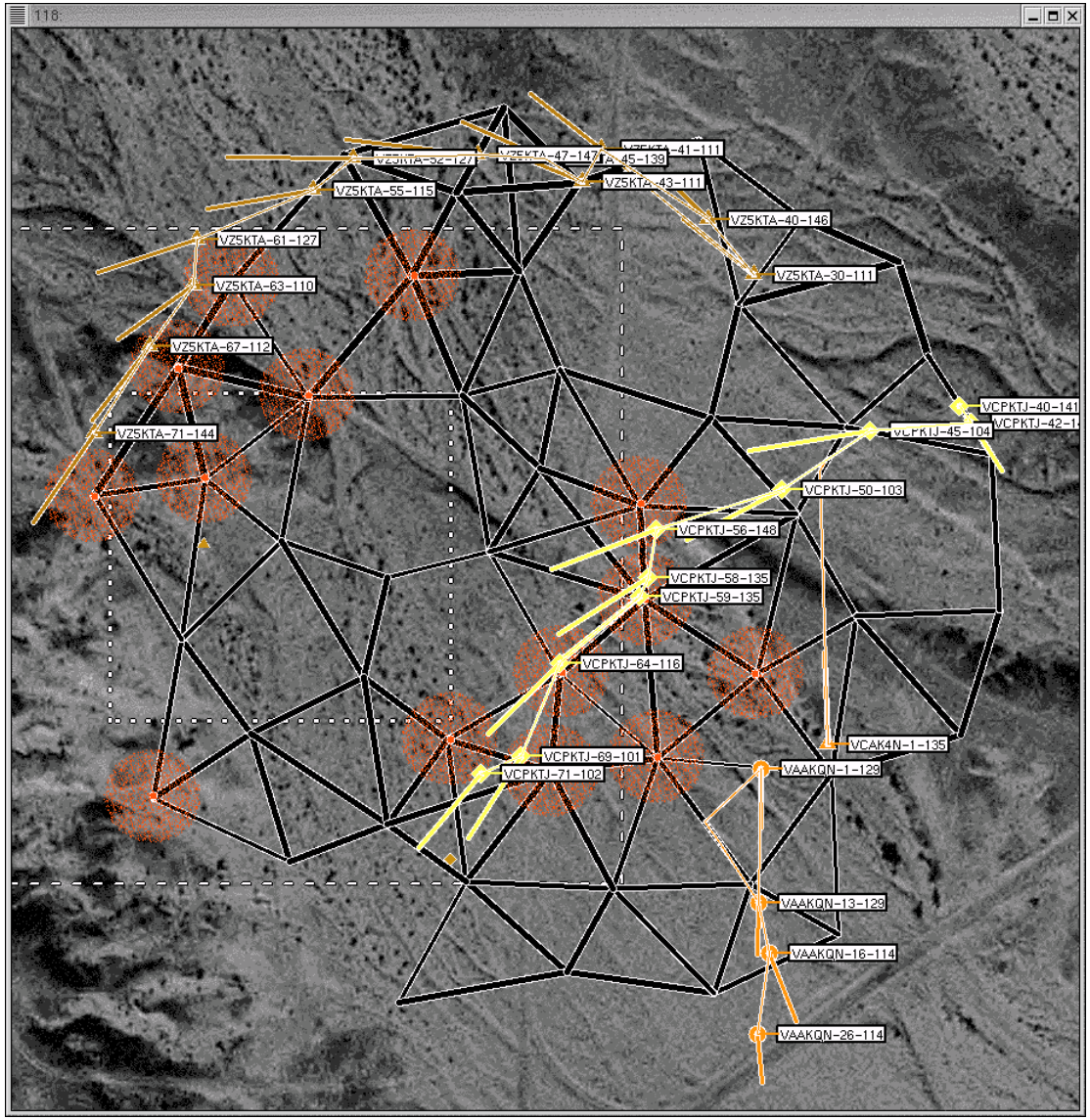


Tracking Snapshot 3 (Jan 02)



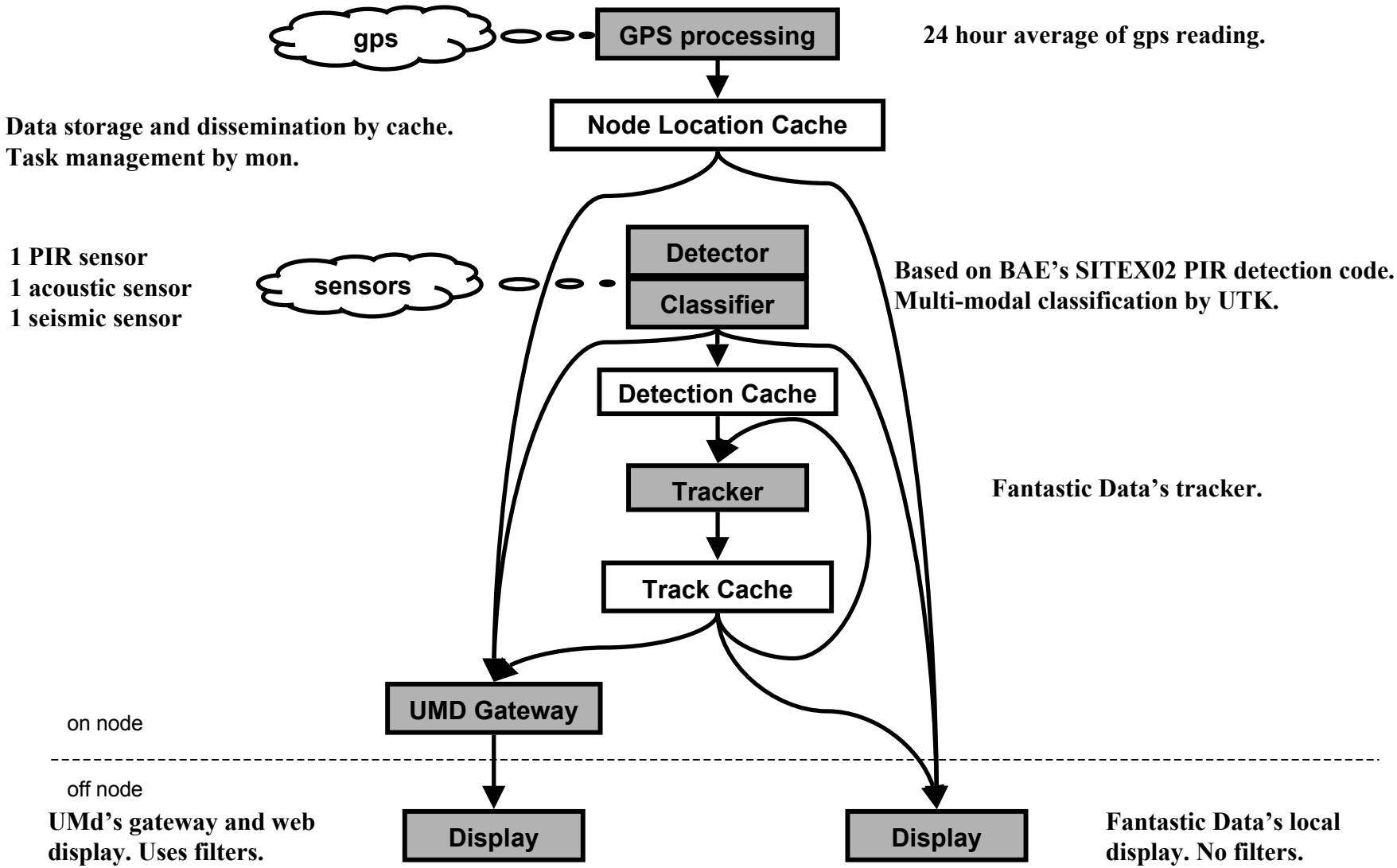
Fantastic Data

50 node simulation

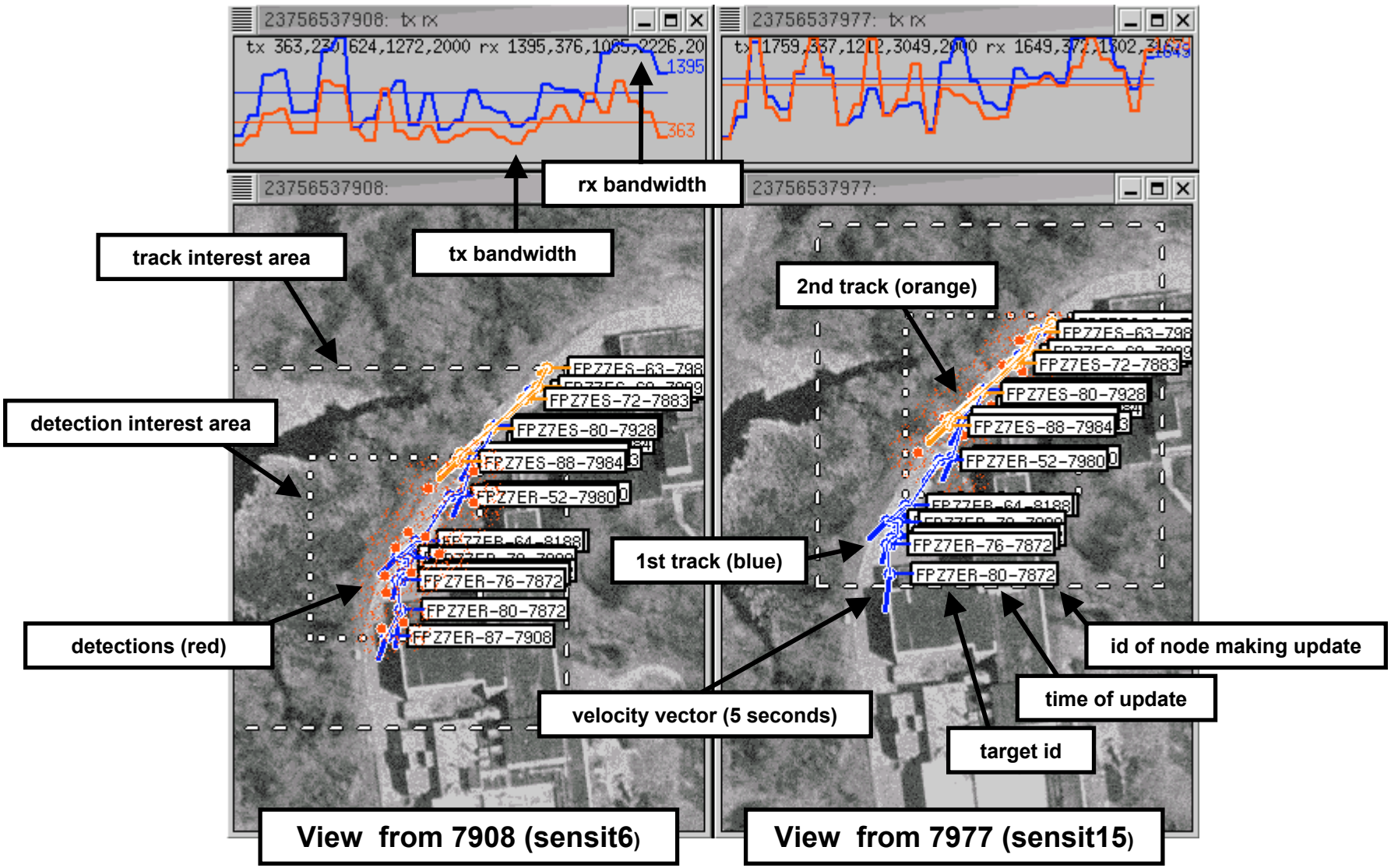


Fantastic Data

System Block Diagram



Tracking Snapshot



Fantastic Data