Google
Developers

All the Ships in the World

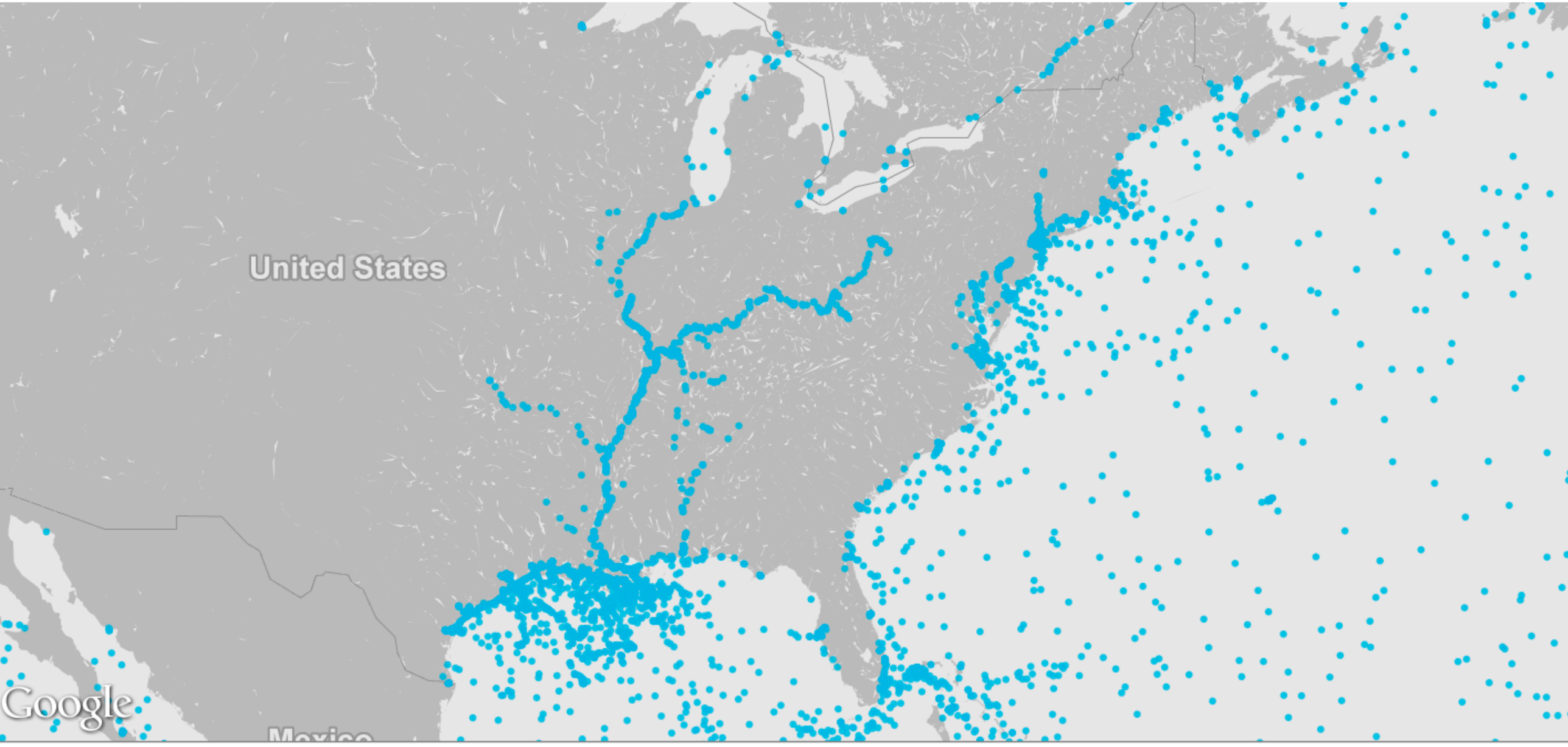
Francesc Campoy, Kurt Schwehr and
Mano Marks

Google I/O

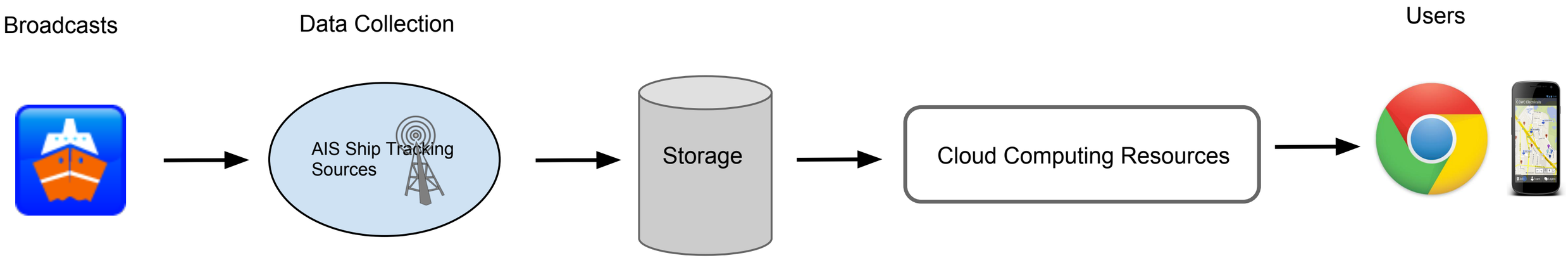
Tons of ships!

75K Ships

400M Points


United States
Google

What we're doing




use ALL the Clouds!!!

Gathering Data

What is AIS?

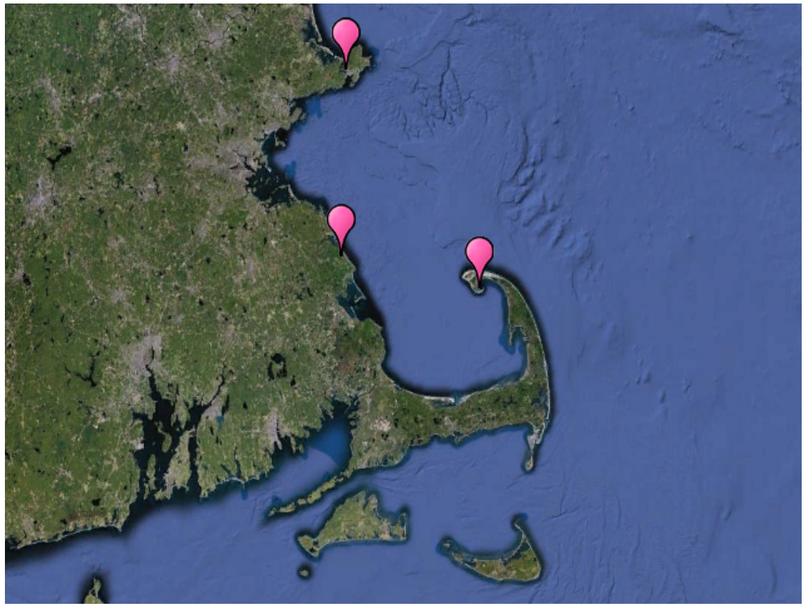
- VHF transceivers on ships
- Designed in the 1990's
- Help ships navigate
- Line of sight

Stats	Ship Positions	Unique Vessels	Date Range for demo
SpaceQuest Satellite Global Coverage	40M	76K	May 2012-present

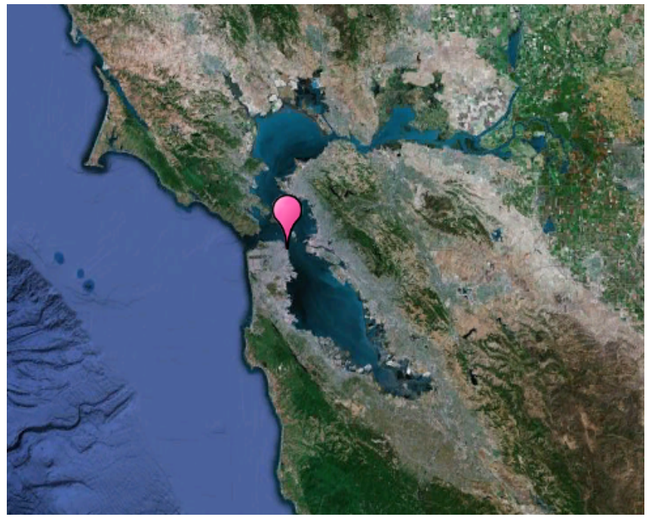
SpaceQuest

Analyze

Stats	Ship Positions	Unique Vessels	Date Range for demo
SpaceQuest Satellite Global Coverage	40M	76K	May 2012-present
NOAA, Boston, MA	364M	9K	2005-Present

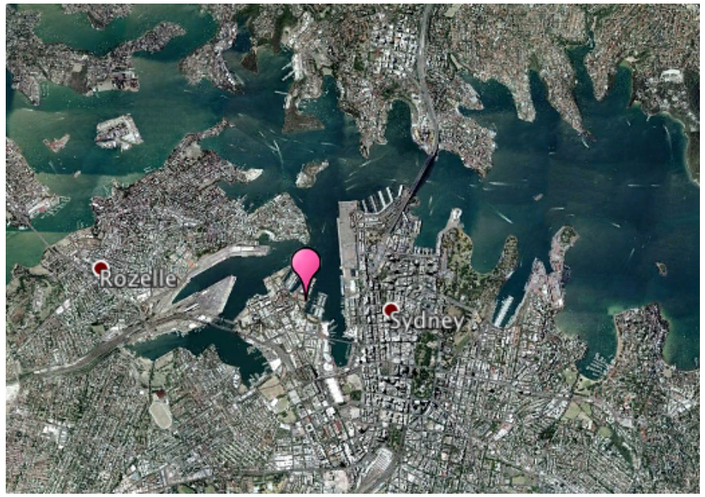


Stats	Ship Positions	Unique Vessels	Date Range for demo
SpaceQuest Satellite Global Coverage	40M	76K	May 2012-present
NOAA, Boston, MA	364M	9K	2005-Present
Google, San Francisco, CA	240K	171	Last week



Google

Stats	Ship Positions	Unique Vessels	Date Range for demo
SpaceQuest Satellite Global Coverage	40M	76K	May 2012-present
NOAA, Boston, MA	364M	9K	2005-Present
Google, San Francisco, CA	240K	171	Last week
Google, Sydney, Australia	134K	48	Last week



Google

~~All~~ Most of the Ships in the World

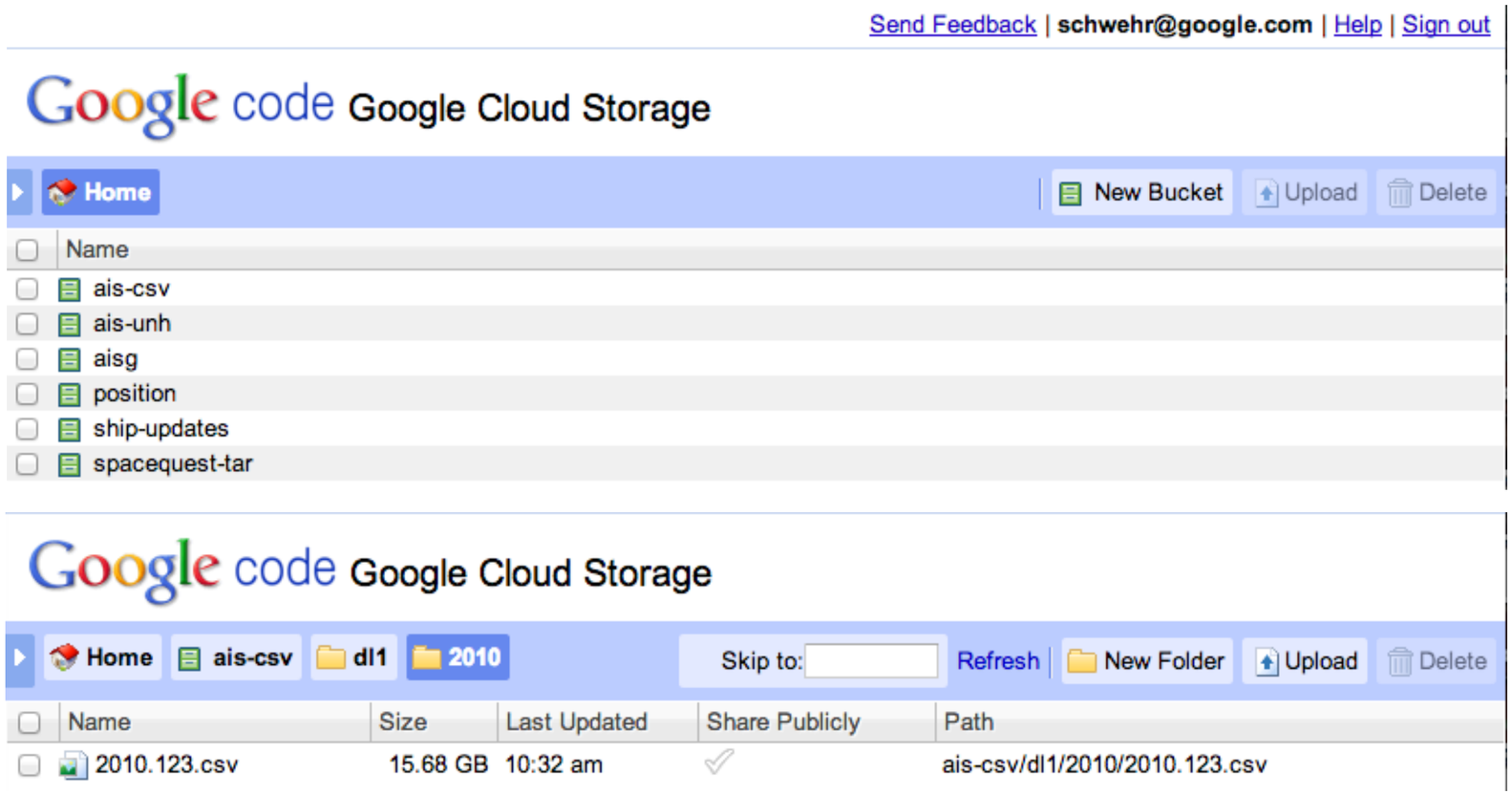
Challenges

- Data size
- Data quality
- Overlapping messages

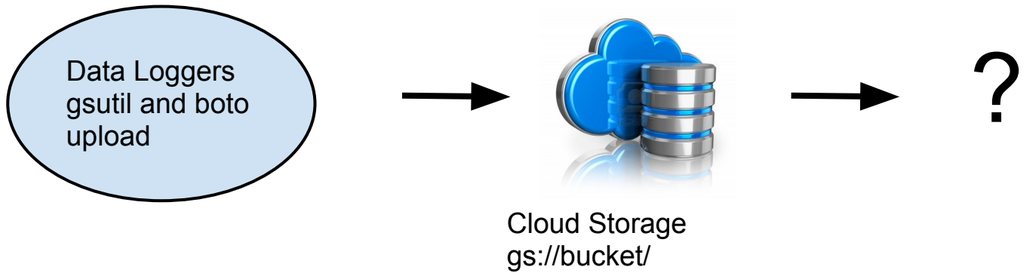
Storing Data

Cloud Storage

- Access Control Lists (ACLs)
- Up-to and beyond 4 PB
- Fast access to data



Architecture: storing data

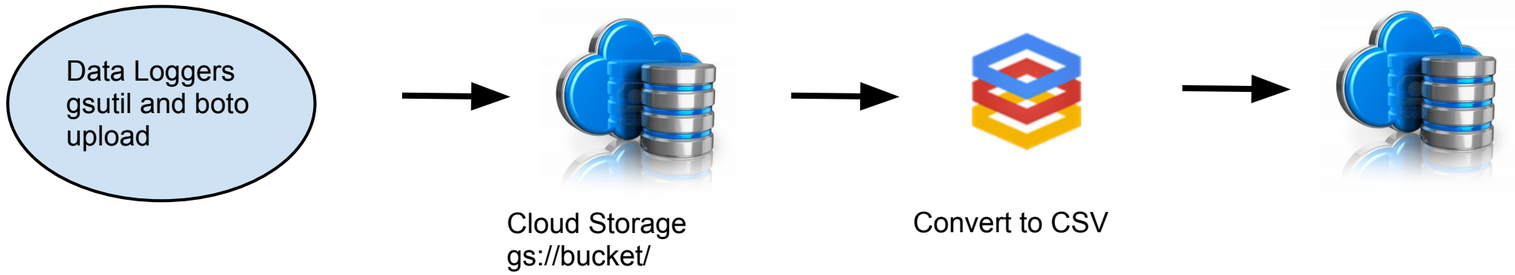


!AIVDM,1,1,,A,34hShb8siVo?dN`E`gL52T5N00r@,0*6A,rgooglesfospe2m6b,1368132768.04

Compute Engine

- Generic Linux instances
- Run code in any language
- Service accounts make security easy

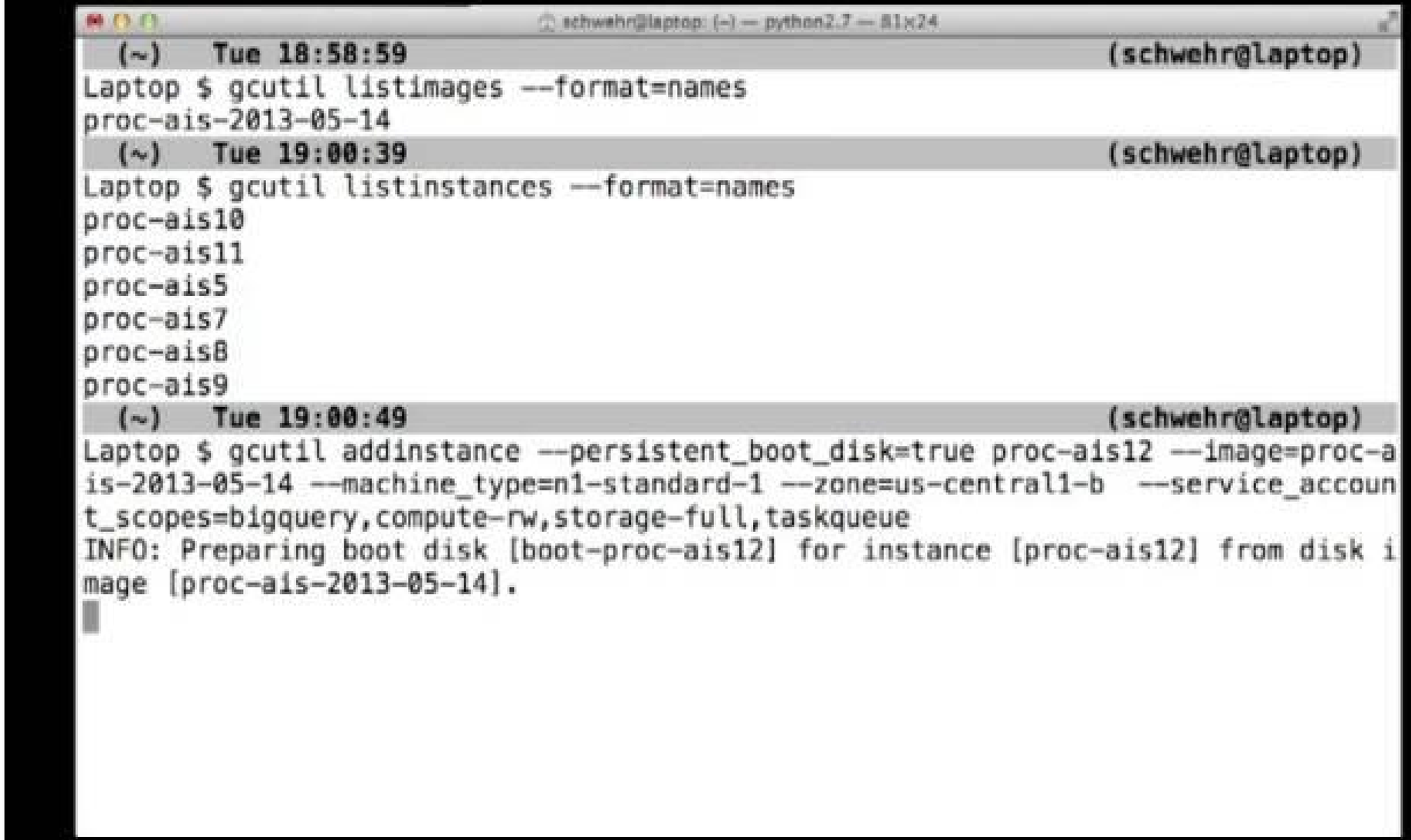
Architecture: storing data



!AIVDM,1,1,,A,34hShb8siVo?dN`E`gL52T5N00r@,0*6A,rgooglesfospe2m6b,1368132768.04

MMSI	Navigation Status	ROT	SOG	Longitude	Latitude	COG	True Heading
319353000	8	-17	10.2	-122.4005	37.81267	129	130

Open Source C++ decoder: libais


schwehr@laptop: (~) — python2.7 — 81×24
(~) Tue 18:58:59 (schwehr@laptop)
Laptop $ gcutil listimages --format=names
proc-ais-2013-05-14
(~) Tue 19:00:39 (schwehr@laptop)
Laptop $ gcutil listinstances --format=names
proc-ais10
proc-ais11
proc-ais5
proc-ais7
proc-ais8
proc-ais9
(~) Tue 19:00:49 (schwehr@laptop)
Laptop $ gcutil addinstance --persistent_boot_disk=true proc-ais12 --image=proc-a
is-2013-05-14 --machine_type=n1-standard-1 --zone=us-central1-b --service_accoun
t_scopes=bigquery,compute-rw,storage-full,taskqueue
INFO: Preparing boot disk [boot-proc-ais12] for instance [proc-ais12] from disk i
mage [proc-ais-2013-05-14].

```
  (~)   Tue 10:46:38                                          (schwehr@laptop)
Laptop $ gcutil addinstance --persistent_boot_disk=true proc-ais10 --image=proc-
ais-2013-05-14 --machine_type=n1-standard-1 --zone=us-central1-b  --service_acco
unt_scopes=bigquery,cloudsql,compute-rw,storage-full,taskqueue
INFO: Preparing boot disk [boot-proc-ais10] for instance [proc-ais10] from disk
image [proc-ais-2013-05-14].
INFO: Waiting for insert of disk boot-proc-ais10. Sleeping for 3s.

  (~)   Tue 10:48:21                                          (schwehr@laptop)
Laptop $ gcutil ssh proc-ais10
INFO: Zone for 'proc-ais10' detected as u'us-central1-b'.

schwehr@proc-ais10:~$
```

Architecture: Bucket Notifications & Task Queues

Serving data

Big Query

- Query terabyte data on a scale of seconds, not minutes (or hours)
- Write queries fast with a SQL-like syntax
- RESTful API means easy integration

Loading data to BigQuery

Bucket notifications, App Engine, and BigQuery

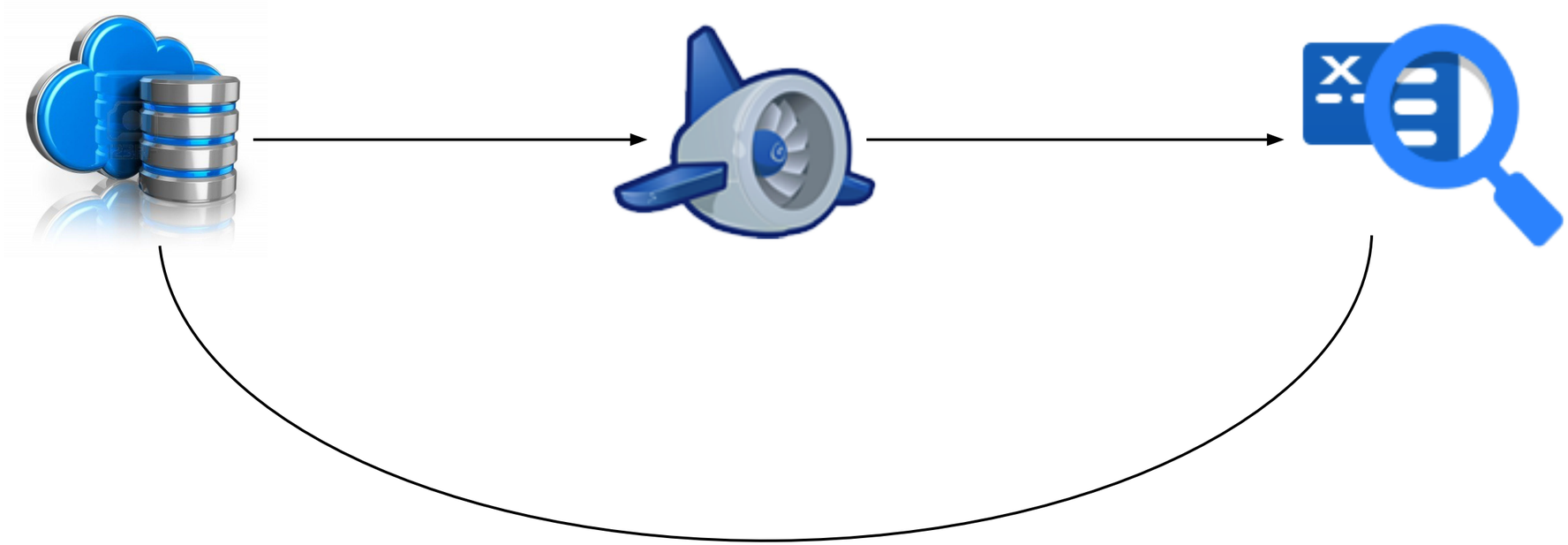


Table Details: pos123

Schema | Details | Query Table

Table Info

Table ID	ais-demo-v1:io.pos123
Table Size	69.2 GB
Number of Rows	374,601,235
Creation Time	7:43pm, 22 Apr 2013
Last Modified	4:11pm, 25 Apr 2013

Preview

Row	id	repeat_indicator	mmsi	nav_status	rot_over_range	rot	sog	position_accuracy	x	
1	1	0	366880370	0	false	0.0	0.10000000149	0	-71.0250930786	42.3
2	1	0	367010390	0	true	-731.386474609	7.80000019073	0	-70.9610595703	42.3
3	1	0	367010390	0	true	-731.386474609	7.80000019073	0	-70.9610595703	42.3
4	3	0	636090339	5	false	0.0	0.10000000149	0	-70.8772583008	42.5
5	1	0	867029470	0	true	-731.386474609	0.20000000298	0	-71.0401000977	42.3

New Query

```
SELECT id, lon, lat, t
FROM (
  SELECT
    mmsi AS id,
    HOUR(rcvr_time) as hour,
    AVG(x) as lon,
    AVG(y) as lat,
    INTEGER(AVG(TIMESTAMP_TO_SEC(rcvr_time))) as t
  FROM [io.pos123]
  WHERE DATEDIFF(CURRENT_TIMESTAMP(), rcvr_time) < 90
  AND x <= 180 AND y <= 90
  AND mmsi >= 100000000 AND mmsi <= 999999999
  GROUP BY id, hour
  ORDER BY id, hour
)
```

How fast is fast?

New Query

```
SELECT id, lon, lat, t
FROM (
  SELECT
    mmsi AS id,
    HOUR(rcvr_time) as hour,
    AVG(x) as lon,
    AVG(y) as lat,
    INTEGER(AVG(TIMESTAMP_TO_SEC(rcvr_time))) as t
  FROM [io.pos123]
  WHERE DATEDIFF(CURRENT_TIMESTAMP(), rcvr_time) < 90
  AND x <= 180 AND y <= 90
  AND mmsi >= 100000000 AND mmsi <= 999999999
  GROUP BY id, hour
  ORDER BY id, hour
)
```

RUN QUERY Query complete (5.8s elapsed, 14.0 GB processed)

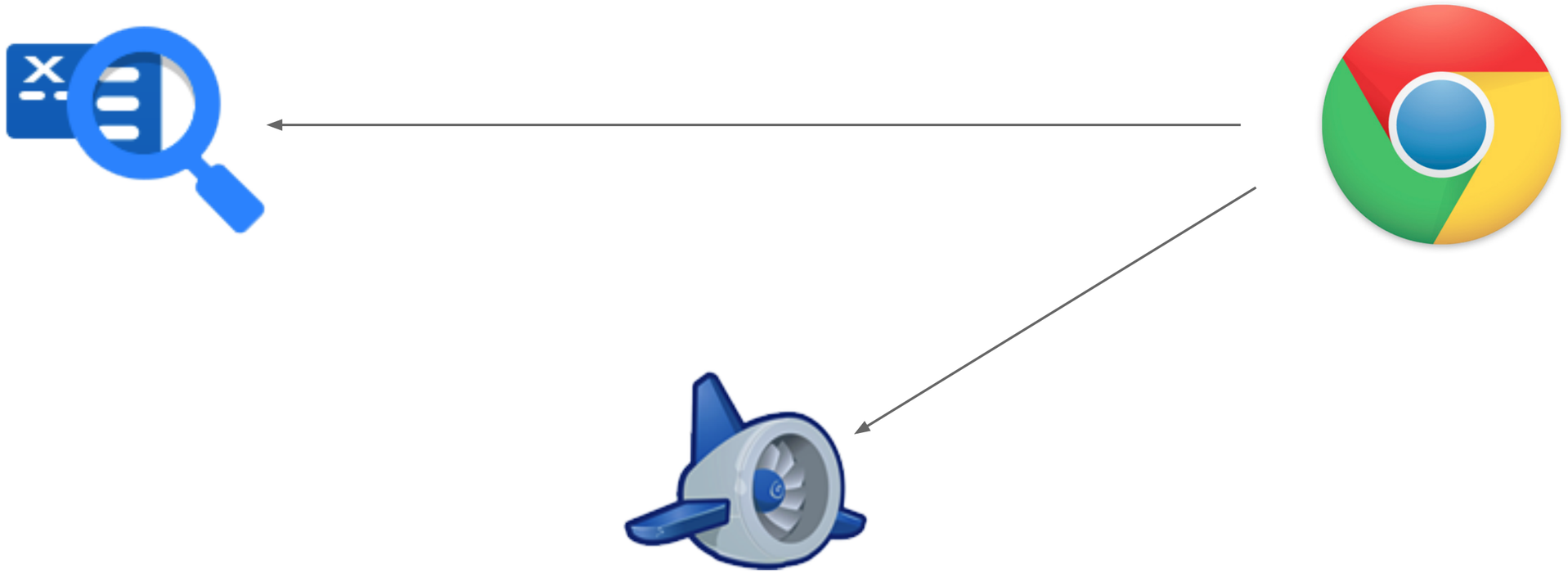
Query Results 9:55pm, 8 May 2013

Download as CSV | Save as Table

Row	id	lon	lat	t
1	210348000	-70.99343719482718	42.38785250854559	1361897990
2	210348000	-70.97638637388977	42.34629394121074	1361913635
3	210348000	-70.86484985352	42.38179375784	1361833531
4	210348000	-70.54536020755938	42.61340510844843	1361800768
5	210348000	-70.91886259104795	42.37115990625895	1361909501

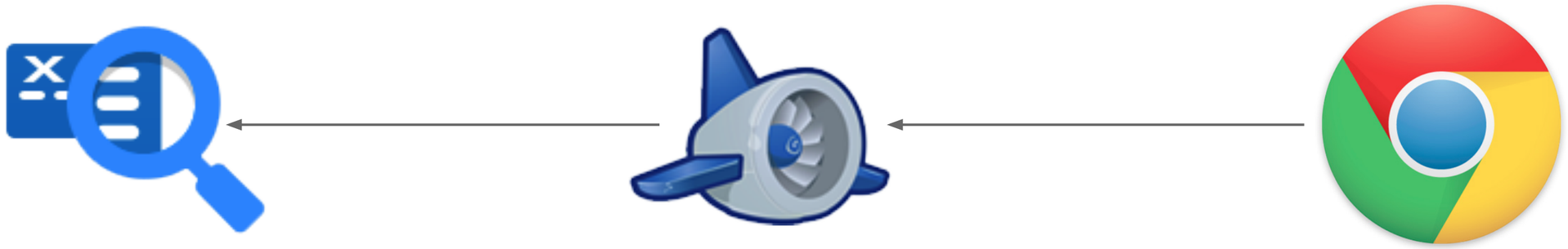
Serving architectures

Using BigQuery REST API from JavaScript



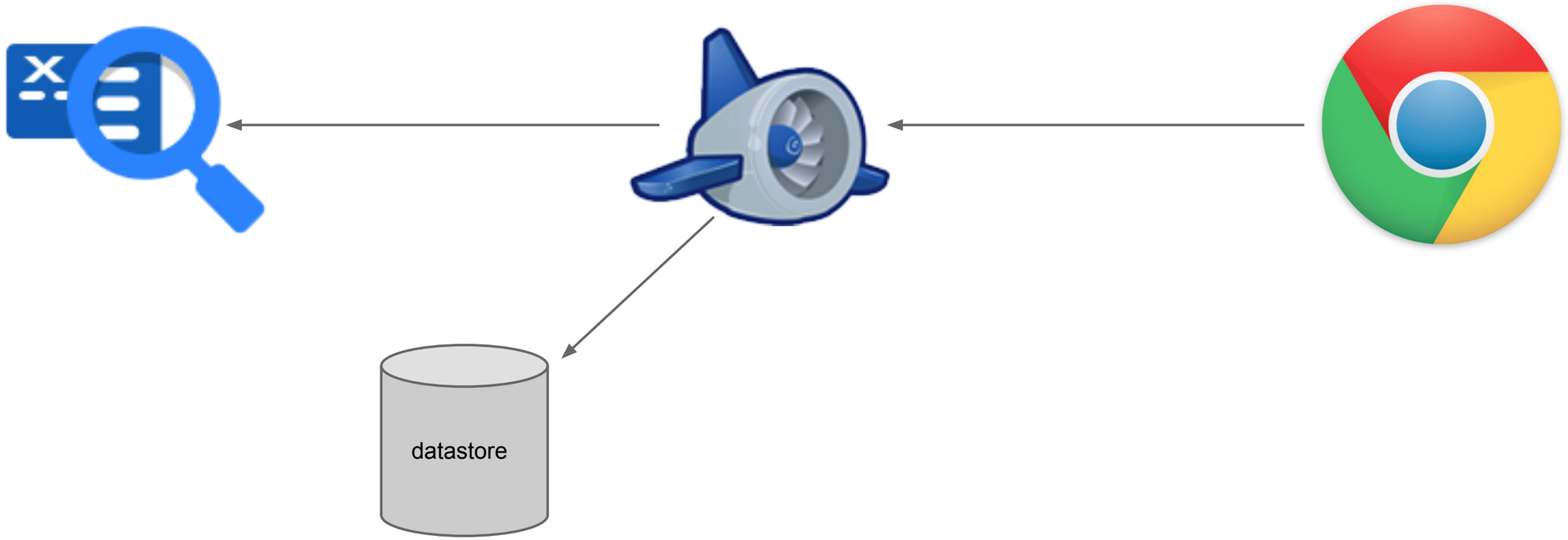
Serving architectures

Using BigQuery REST API from App Engine with Go



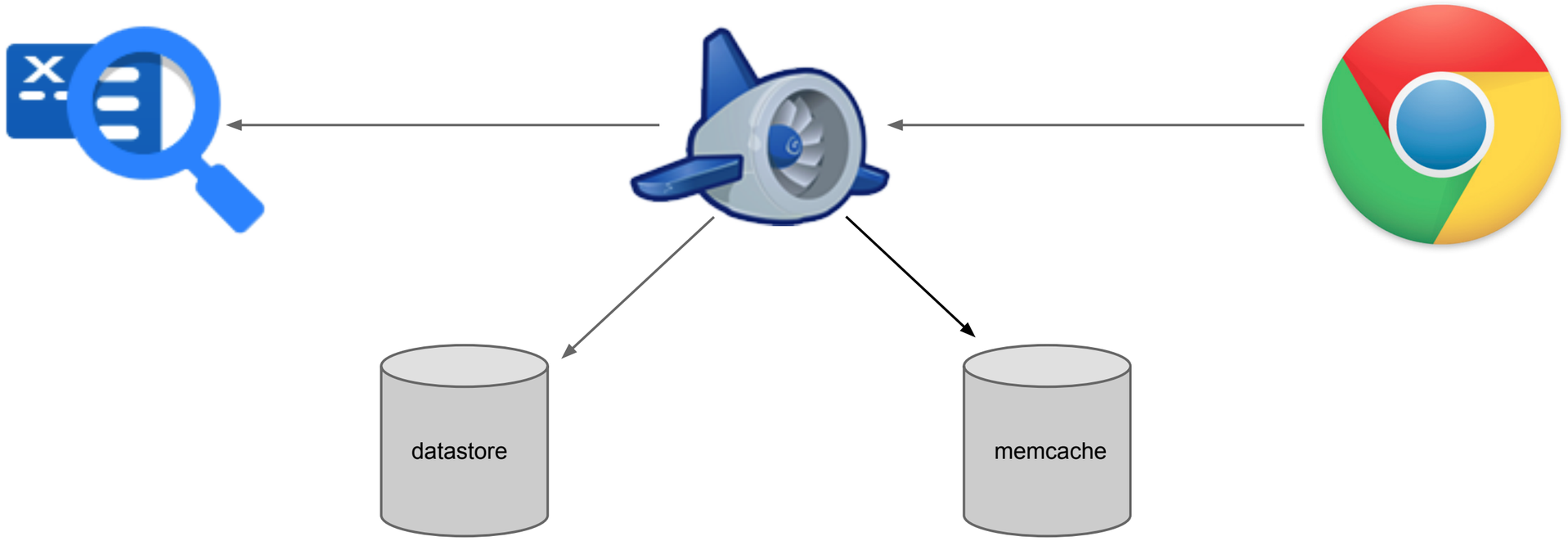
Serving architectures

Materializing query results on datastore



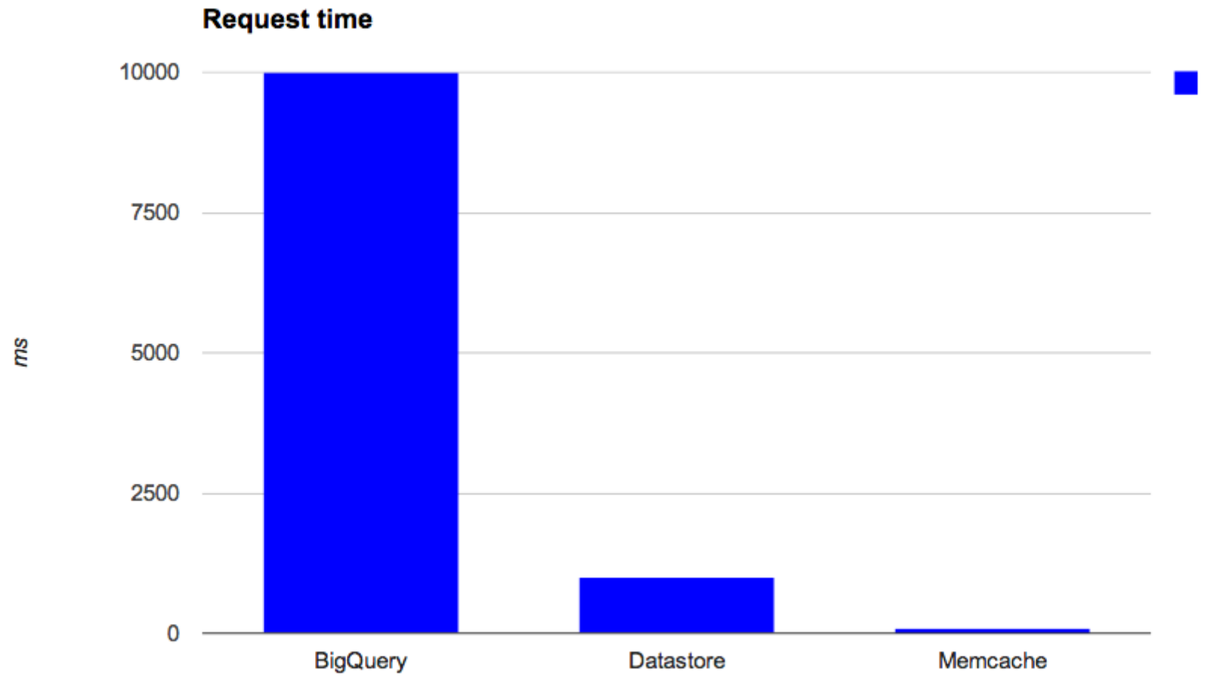
Serving architectures

Caching the formatted results on memcache

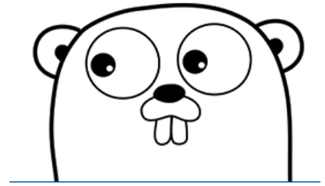


Serving architectures

Speed! Speed! Speed!



Go on App Engine

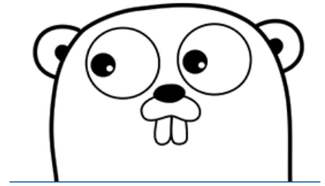


Go as the language for the cloud

- Speed!
  - Fastest instance startup time
  - Compiled to machine code
  - Less CPU, fewer instances
- Memory efficiency
  - Fewer, more affordable instances
- Concurrency
  - Concurrent I/O and network ops
  - Easy performance increase

Go

Concurrency made easy

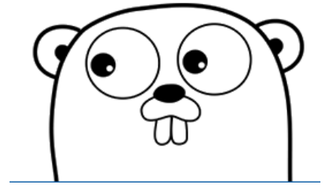


Go

```
3          i := &memcache.Item{
4              Key:   key,
5              Value: value,
6          }
7          memcache.Set(c, i)
```

Go

Concurrency made easy

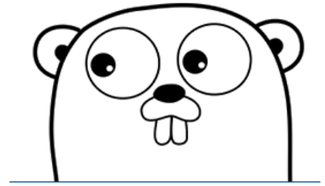


Go

```
2     go func() {
3         i := &memcache.Item{
4             Key:   key,
5             Value: value,
6         }
7         memcache.Set(c, i)
8     }()
```

Go

Concurrency made easy

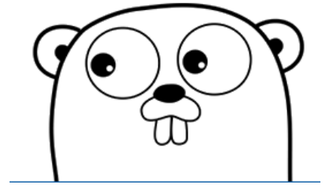


Go

```
1     errc := make(chan error)
2     go func() {
3         i := &memcache.Item{
4             Key:   key,
5             Value: value,
6         }
7         errc <- memcache.Set(c, i)
8     }()
9
10    <-errc
```

Go

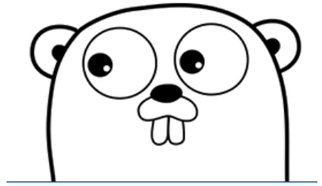
Concurrency made easy



Go

```
1     errc := make(chan error, 1)
2     go func() {
3         i := &memcache.Item{
4             Key:   key,
5             Value: value,
6         }
7         errc <- memcache.Set(c, i)
8     }()
9     select {
10    case <-errc:
11    case <-time.After(time.Second):
12    }
```

More on Go



Talks today

2:35pm - Room 7 - **High Performance Apps with Go on App Engine**

4:25pm - Room 7 - **Advanced Go Concurrency Patterns**

5:20pm - Room 2 - **Fireside Chat with the Go team**

Codelab tomorrow

9:00am - Room 3 - **Whispering Gophers: network programming in Go**

Office hours

3:00-3:45pm today | 1:45-2:30pm tomorrow

Displaying data

The Google Maps API

- Global and local context
- Multiplatform
- CanvasLayer syncs WebGL with Maps in JS

Not only ships!

Any kind of sensor data would do

- Flights
- Taxi/car service
- Trucks
- Anything else

developers.google.com/maps

developers.google.com/cloud

golang.org

Thank you!

+ManoMarks

+FrancescCampoyFlores

+KurtSchwehr

Google
Developers