HADOOP AT NOKIA

JOSH DEVINS, NOKIA

HADOOP MEETUP, JANUARY 2011
BERLIN

Two parts:
 * technical setup
 * applications

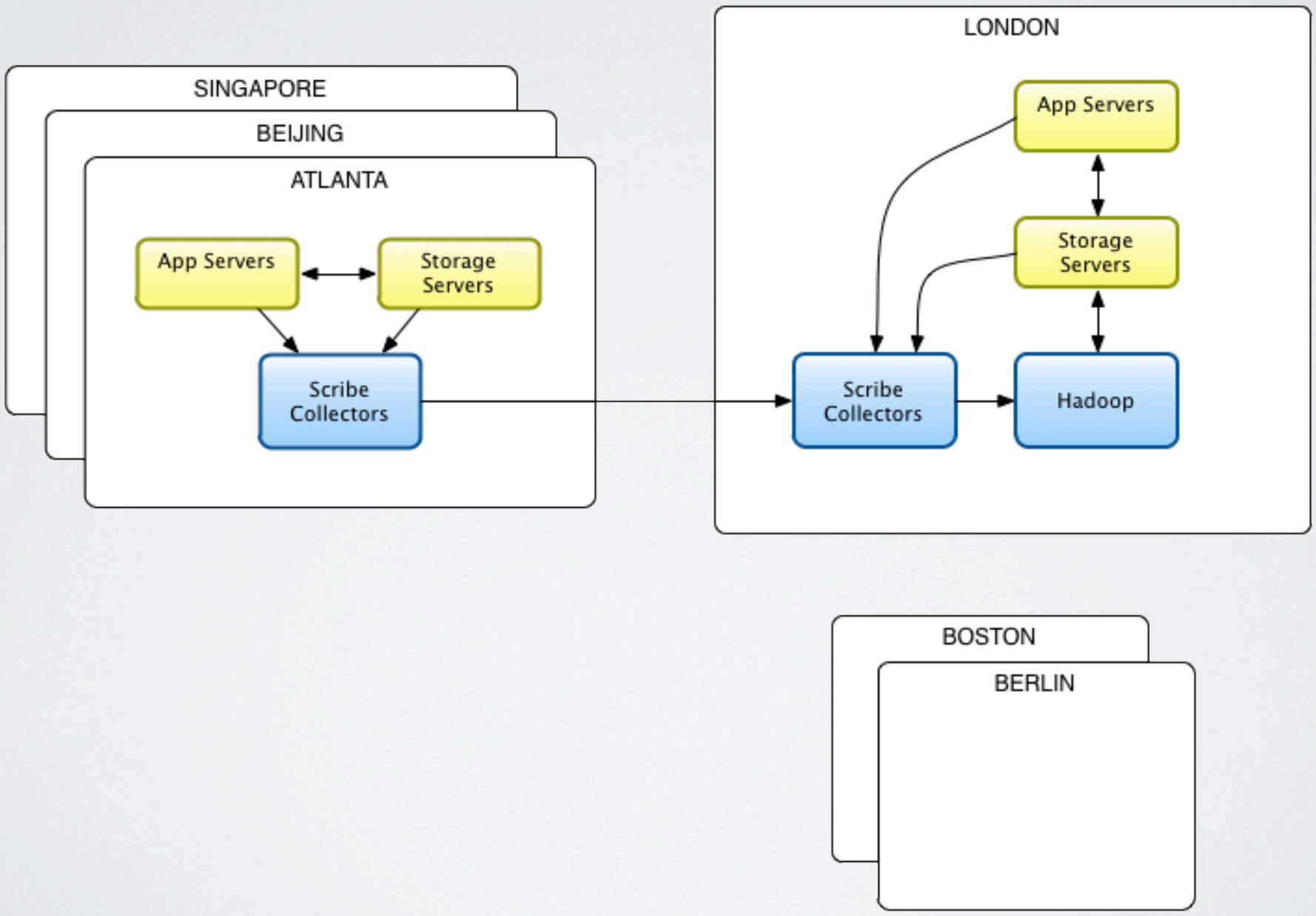
before starting Question: Hadoop experience levels from none to some to lots, and what about cluster mgmt?

TECHNICAL SETUP

Thursday, January 27, 2011

http://www.flickr.com/photos/josecamoessilva/2873298422/sizes/o/in/photostream/

GLOBAL ARCHITECTURE



Thursday, January 27, 2011

Scribe for logging
agents on local machines forward to downstream collector nodes
collectors forward on to more downstream nodes or to final destination(s) like HDFS
buffering at each stage to deal with network outages
must consider the storage available on ALL nodes where Scribe is running to determine your risk of potential data loss since Scribe buffers to local disk
global Scribe not deployed yet, but London DC is done
do it all over again? probably use Flume, but Scribe was being researched before Flume existed
 * much more flexible and easily extensible
 * more reliability guarantees and tunable (data loss acceptable or not)
 * can also do syslog wire protocol which is nice for compatibility's sake

DATA NODE HARDWARE

DC	**LONDON**	**BERLIN**
cores	12x (w/ HT)	4x 2.00 GHz (w/ HT)
RAM	48GB	16GB
disks	12x 2TB	4x 1TB
storage	24TB	4TB
LAN	1Gb	2x 1Gb (bonded)

Thursday, January 27, 2011

http://www.flickr.com/photos/torkildr/3462607995/in/photostream/

BERLIN
- HP DL160 G6
- 1x Quad-core Intel Xeon E5504 @ 2.00 GHz (4-cores total)
- 16GB DDR3 RAM
- 4x 1TB 7200 RPM SATA
- 2x 1Gb LAN
- iLO Lights-Out 100 Advanced

MASTER NODE HARDWARE

DC	**LONDON**	**BERLIN**
cores	12x (w/ HT)	8x 2.00 GHz (w/ HT)
RAM	48GB	32GB
disks	12x 2TB	4x 1TB
storage	24TB	4TB
LAN	10Gb	4x 1Gb (bonded, DRBD/Heartbeat)

Thursday, January 27, 2011

BERLIN

- HP DL160 G6
- 2x Quad-core Intel Xeon E5504 @ 2.00 GHz (8-cores total)
- 32GB DDR3 RAM
- 4x 1TB 7200 RPM SATA
- 4x 1Gb Ethernet (2x LAN, 2x DRBD/Heartbeat)
- iLO Lights-Out 100 Advanced (hadoop-master[01-02]-ilo.devbln)

MEANING?

- Size
  - Berlin: 2 master nodes, 13 data nodes, ~17TB HDFS
  - London: "large enough to handle a year's worth of activity log data, with plans for rapid expansion"
- Scribe
  - 250,000 1KB msg/sec
  - 244MB/sec, 14.3GB/hr, 343GB/day

Thursday, January 27, 2011

Berlin: 1 rack, 2 switches for...
London: it's secret!

PHYSICAL OR CLOUD?

http://www.flickr.com/photos/dumbledad/4745475799/

Question: How many run own clusters of physical hardware vs AWS or virtualized?
actual decision is completely dependent on may factors including maybe existing DC, data set sizes, etc.

PHYSICAL OR CLOUD?

- Physical
  - Capital cost
    - 1 rack w/ 2x switches
    - 15x HP DL160 servers
    - ~€20,000
- Annual operating costs
  - power and cooling: €5,265 @ €0.24 kWh
  - rent: €3,600
  - hardware support contract: €2,000 (disks replaced on warranty)
  - €10,865

Thursday, January 27, 2011

http://www.flickr.com/photos/dumbledad/4745475799/

Question: How many run own clusters of physical hardware vs AWS or virtualized?
actual decision is completely dependent on may factors including maybe existing DC, data set sizes, etc.

PHYSICAL OR CLOUD?

- Physical
  - Capital cost
    - 1 rack w/ 2x switches
    - 15x HP DL160 servers
    - ~€20,000
  - Annual operating costs
    - power and cooling: €5,265 @ €0.24 kWh
    - rent: €3,600
    - hardware support contract: €2,000 (disks replaced on warranty)
    - €10,865
- Cloud (AWS)
  - Elastic MR, 15 extra large nodes, 10% utilized: $1,560
  - S3, 5TB: $7,800
  - $9,360 or €6,835

Thursday, January 27, 2011

http://www.flickr.com/photos/dumbledad/4745475799/

Question: How many run own clusters of physical hardware vs AWS or virtualized?
actual decision is completely dependent on may factors including maybe existing DC, data set sizes, etc.

PHYSICAL OR CLOUD?

- Physical
  - Capital cost
    - 1 rack w/ 2x switches
    - 15x HP DL160 servers
    - ~€20,000
- Annual operating costs
  - power and cooling: €5,265 @ €0.24 kWh
  - rent: €3,600
  - hardware support contract: €2,000 (disks replaced on warranty)
  - €10,865

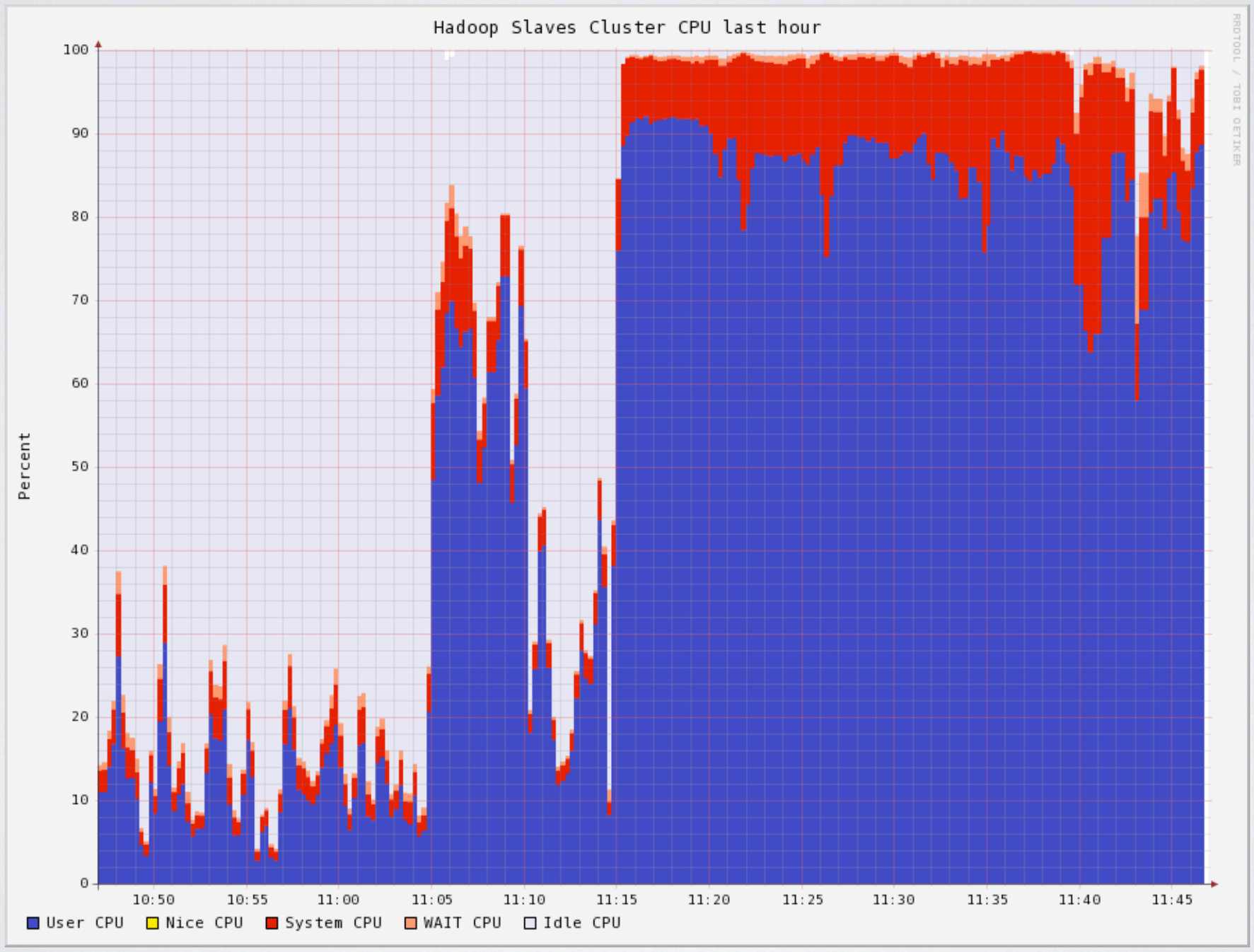
- Cloud (AWS)
  - Elastic MR, 15 extra large nodes, 10% utilized: $1,560
  - S3, 5TB: $7,800
  - $9,360 or €6,835

Thursday, January 27, 2011

http://www.flickr.com/photos/dumbledad/4745475799/

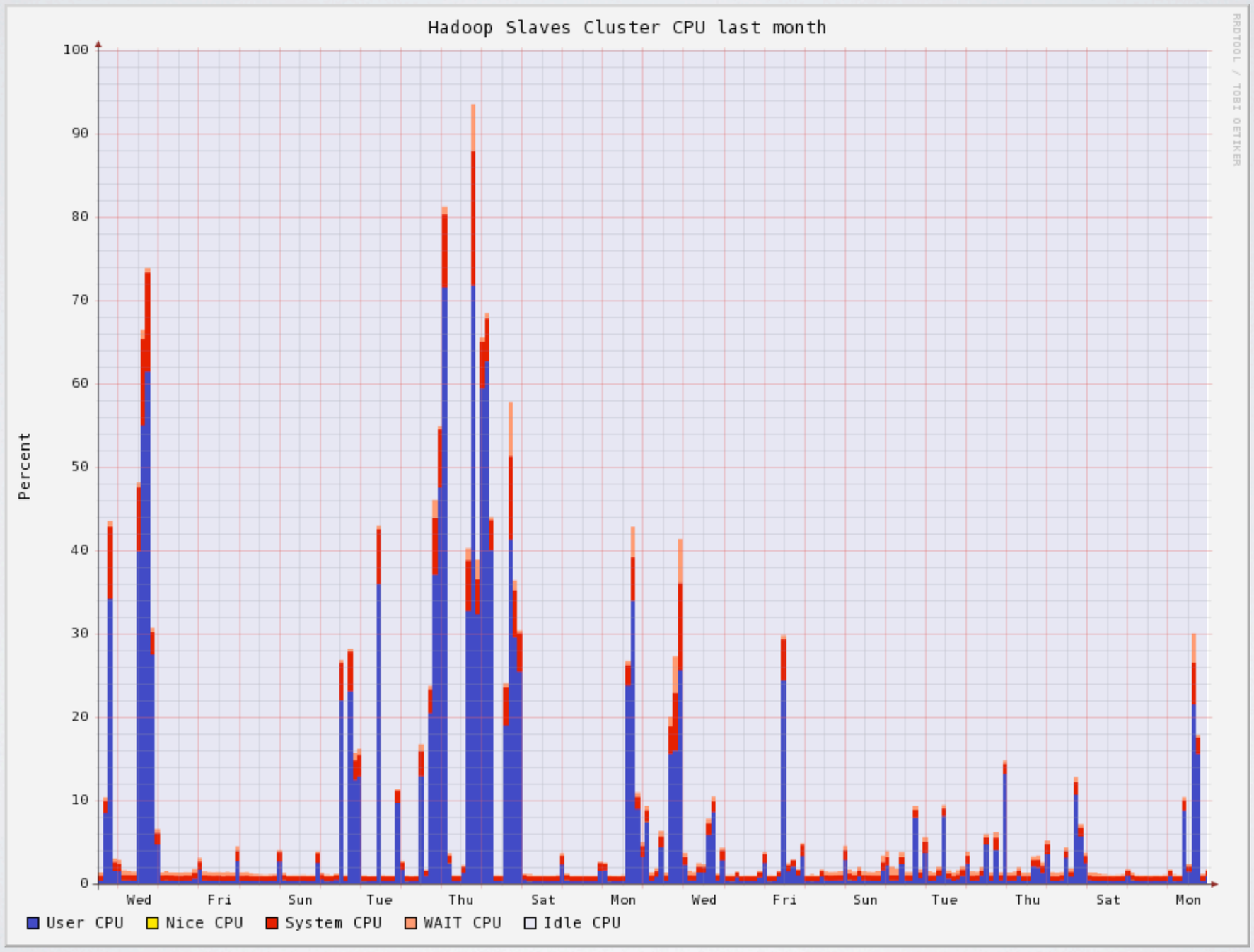
Question: How many run own clusters of physical hardware vs AWS or virtualized?
actual decision is completely dependent on may factors including maybe existing DC, data set sizes, etc.

UTILIZATION



here’s what to show your boss if you want hardware

UTILIZATION



here’s what to show your boss if you want the cloud

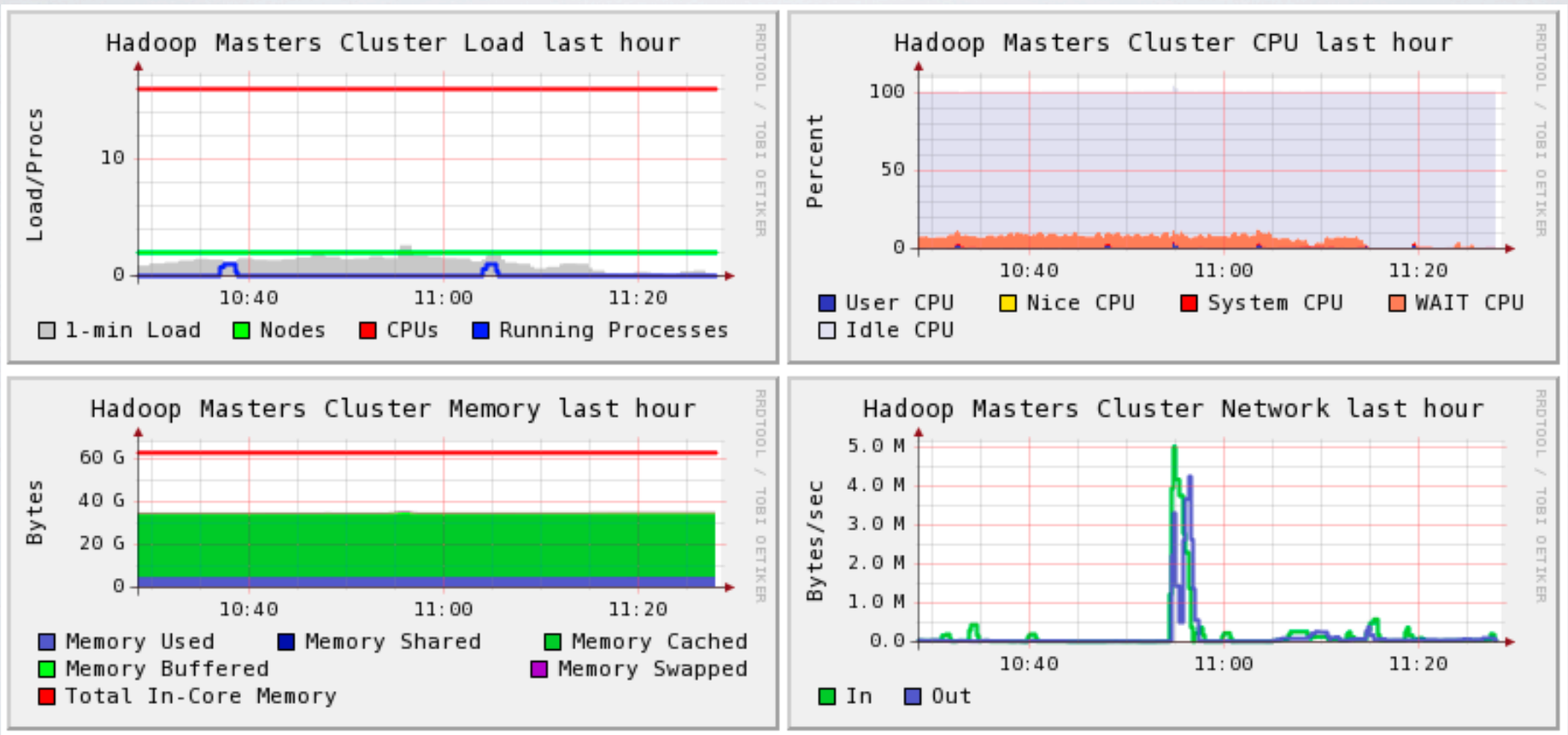
GRAPHING AND MONITORING

- Ganglia for graphing/trending
  - "native" support in Hadoop to push metrics to Ganglia
    - map or reduce tasks running, slots open, HDFS I/O, etc.
  - excellent for system graphing like CPU, memory, etc.
  - scales out horizontally
  - no configuration - just push metrics from nodes to collectors and they will graph it
- Nagios for monitoring
  - built into our Puppet infrastructure
  - machines go up, automatically into Nagios with basic system checks
  - scriptable to easily check other things like JMX

Thursday, January 27, 2011

basically always have up: jobtracker, Oozie, Ganglia

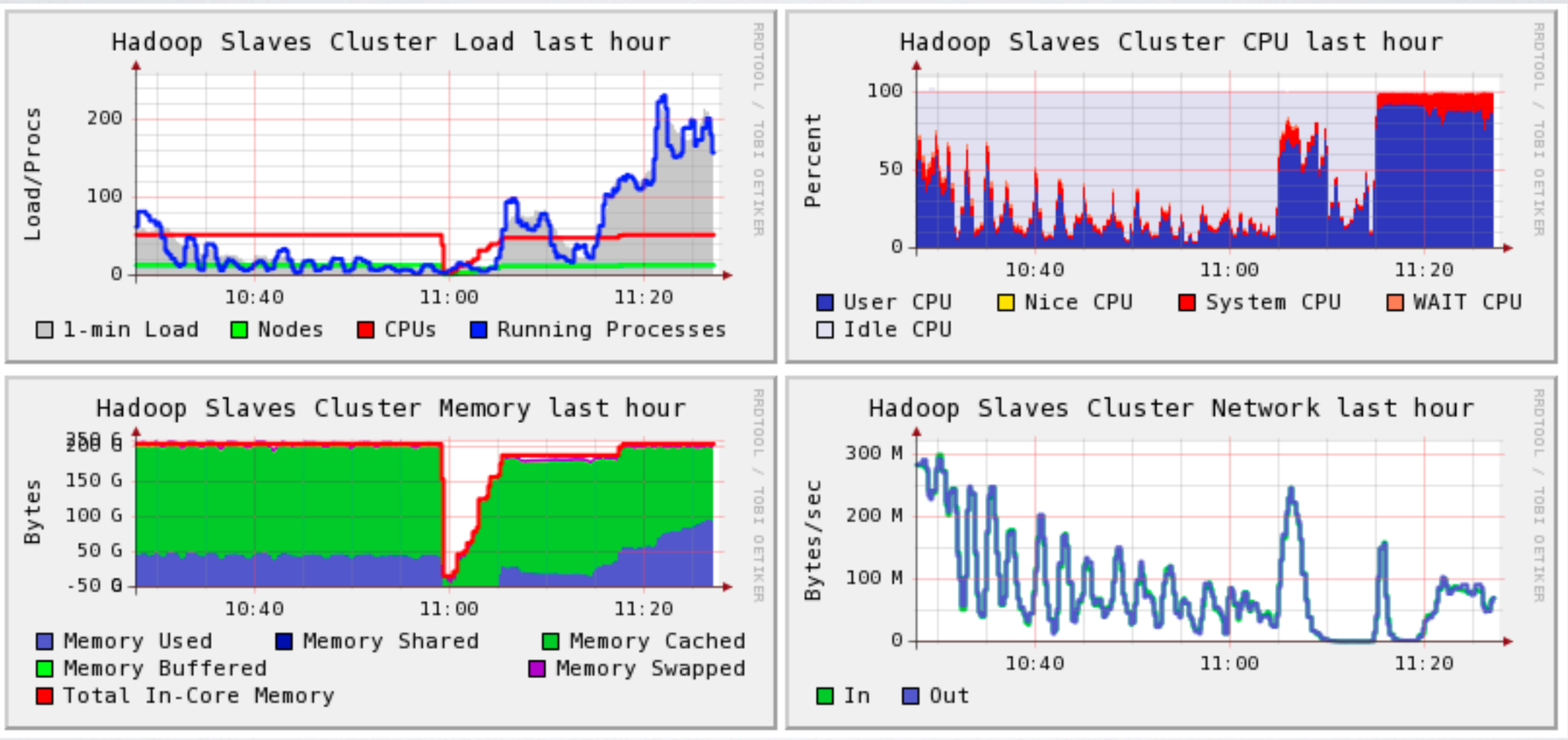
GANGLIA



Thursday, January 27, 2011

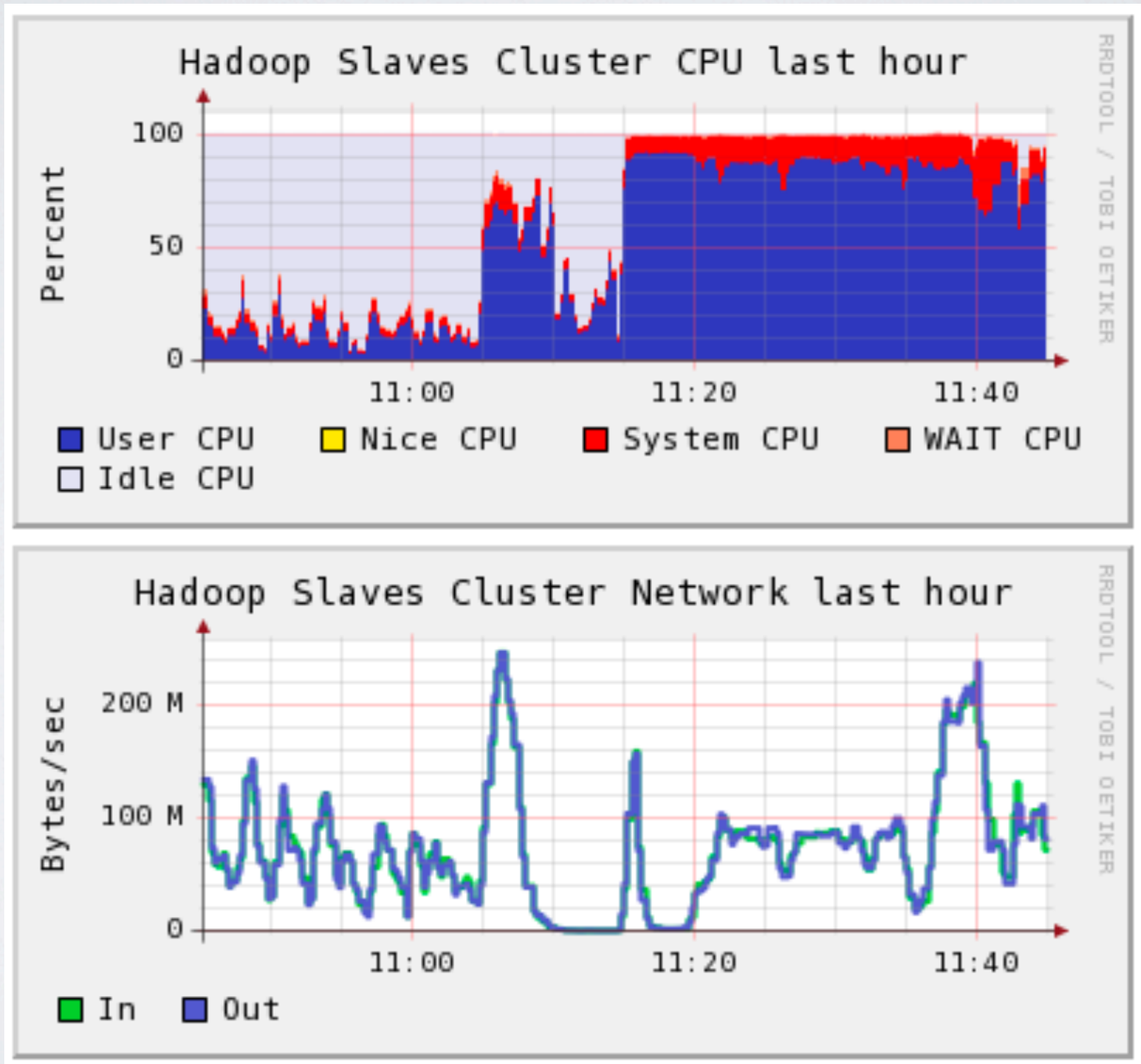
master nodes are mostly idle

GANGLIA



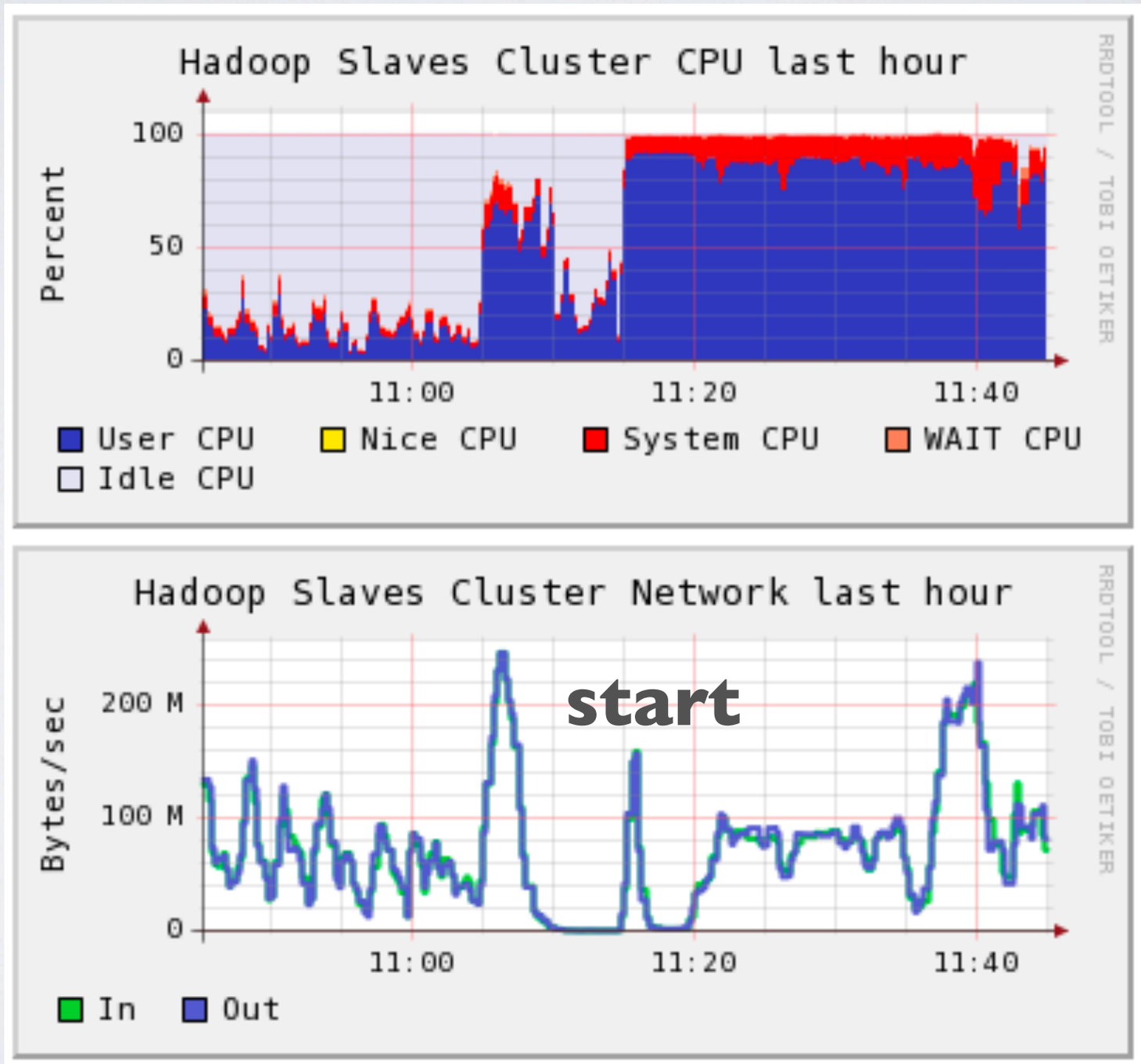
data nodes overview

GANGLIA



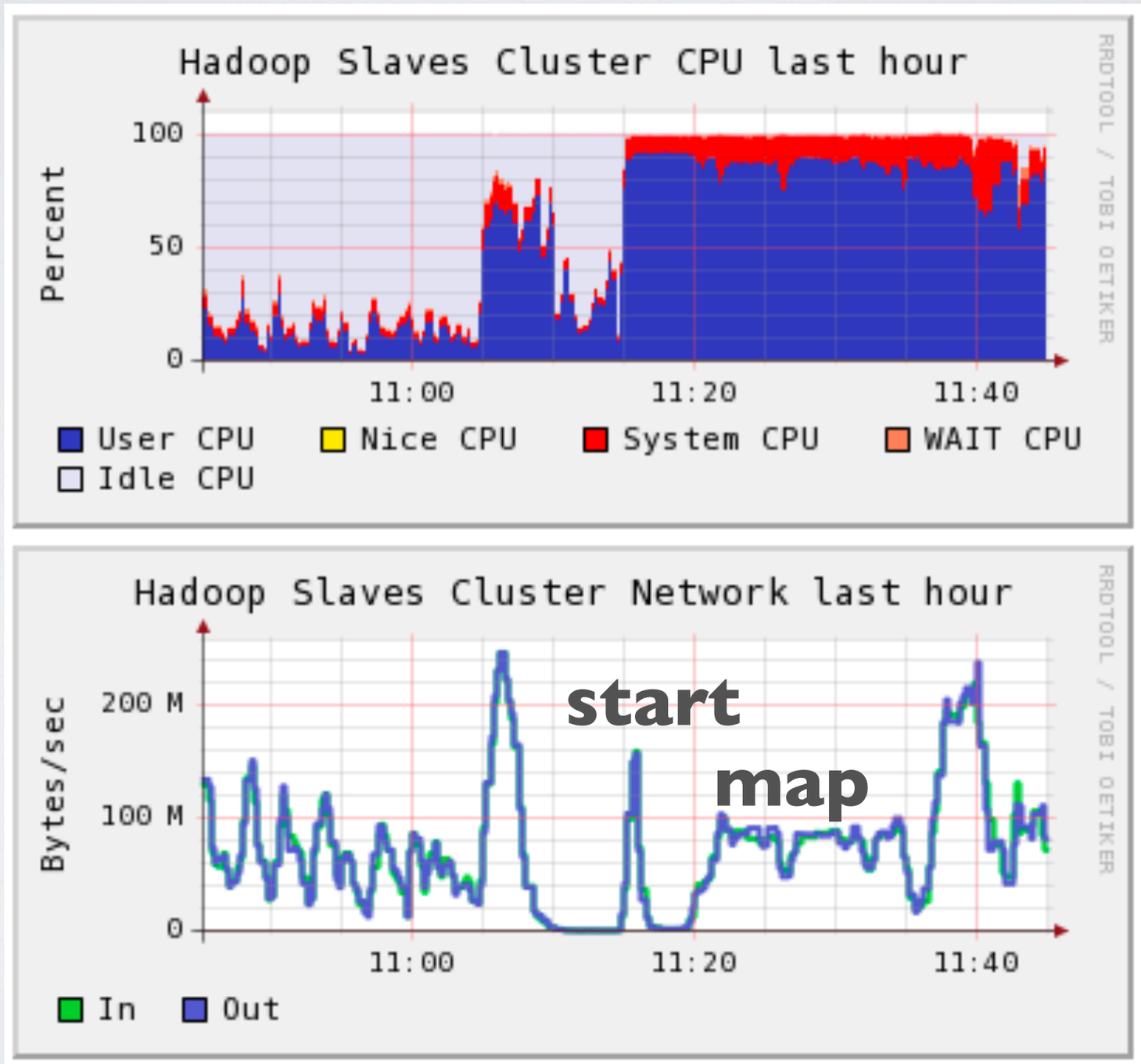
detail view can see actually the phases of a map reduce job
not totally accurate here since there were multiple jobs running at the same time

GANGLIA



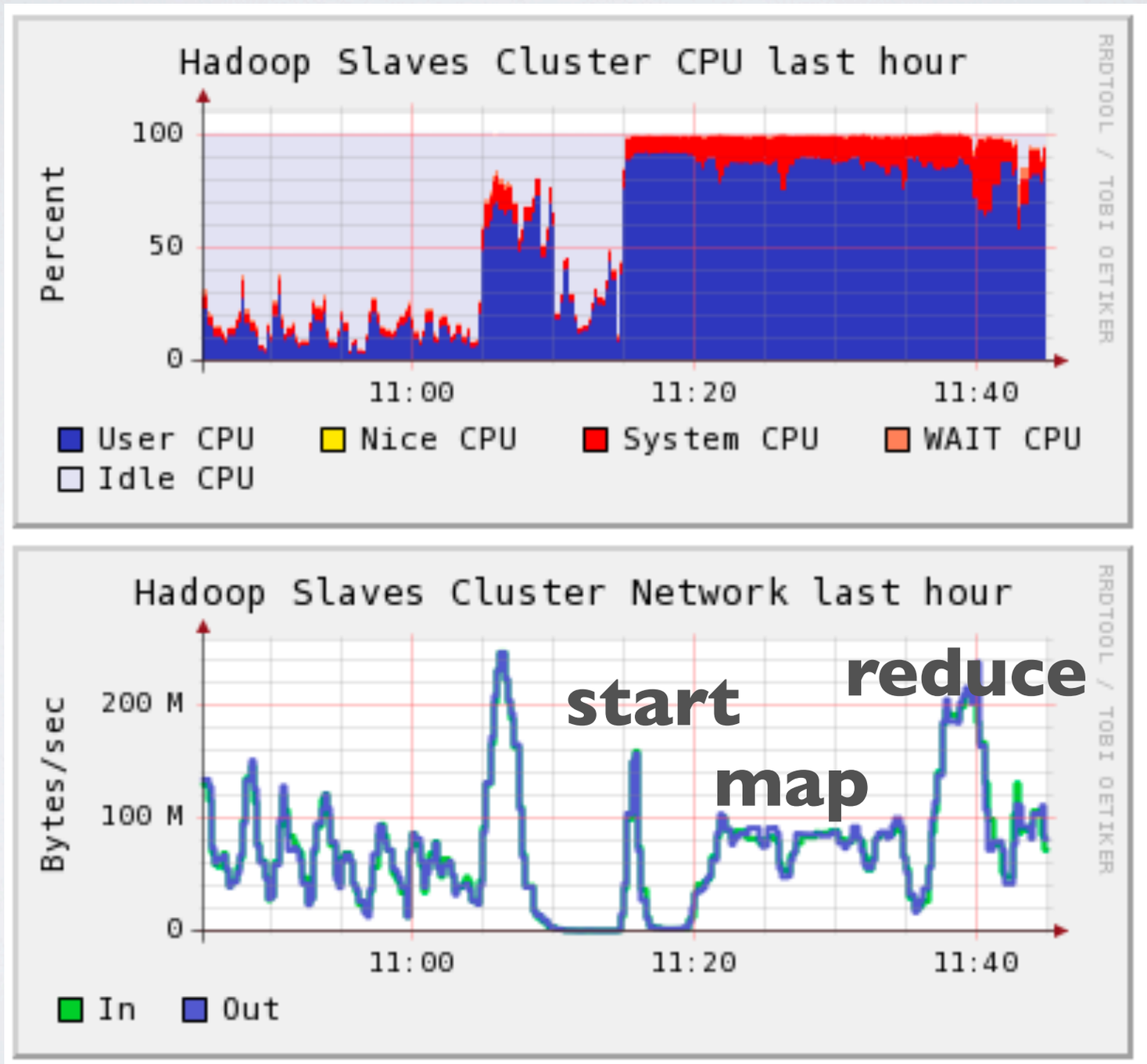
detail view can see actually the phases of a map reduce job
not totally accurate here since there were multiple jobs running at the same time

GANGLIA



detail view can see actually the phases of a map reduce job
not totally accurate here since there were multiple jobs running at the same time

GANGLIA



detail view can see actually the phases of a map reduce job
not totally accurate here since there were multiple jobs running at the same time

SCHEDULER

Running Jobs

Jobid	Priority	User	Name	Map % Complete	Map Total	Maps Completed	Reduce % Complete	Reduce Total	Reduces Completed	Job Scheduling Information
job_201011182036_6991	NORMAL	waldauka	oozie:launcher:T=pig:W=ab_evaluation:A=group_nearby_clicks_by_uuid:ID=0000101-101227224010260-oozie-W	0.00%	1	0	0.00%	0	0	NA
job_201011182036_6992	NORMAL	waldauka	oozie:launcher:T=pig:W=ab_evaluation:A=group_nearby_responses_by_uuid:ID=0000101-101227224010260-oozie-W	0.00%	1	0	0.00%	0	0	NA
job_201011182036_6993	NORMAL	waldauka	PigLatin:groupNearbyClicksByUuid.pig	0.00%	39	0	0.00%	1	0	NA
job_201011182036_6994	NORMAL	waldauka	oozie:launcher:T=pig:W=ab_evaluation:A=group_action_clicks_by_uuid:ID=0000101-101227224010260-oozie-W	0.00%	1	0	0.00%	0	0	NA
job_201011182036_6995	NORMAL	waldauka	PigLatin:groupActionClicksByUuid.pig	0.00%	39	0	0.00%	1	0	NA
job_201011182036_6996	NORMAL	waldauka	PigLatin:groupNearbyResponsesByUuid.pig	0.00%	521	0	0.00%	1	0	NA
job_201011182036_7003	NORMAL	waldauka	oozie:launcher:T=pig:W=ab_evaluation:A=group_nearby_clicks_by_uuid:ID=0000103-101227224010260-oozie-W	0.00%	1	0	0.00%	0	0	NA
job_201011182036_7004	NORMAL	waldauka	oozie:launcher:T=pig:W=ab_evaluation:A=group_nearby_responses_by_uuid:ID=0000103-101227224010260-oozie-W	0.00%	1	0	0.00%	0	0	NA
job_201011182036_7005	NORMAL	waldauka	oozie:launcher:T=pig:W=ab_evaluation:A=group_action_clicks_by_uuid:ID=0000103-101227224010260-oozie-W	0.00%	1	0	0.00%	0	0	NA
job_201011182036_7007	NORMAL	g2harris	PigLatin:lookahead_queries.pig	8.82%	12832	1133	0.69%	1	0	NA
job_201011182036_7008	NORMAL	devins	PigLatin:Search : Where : No Results	57.73%	1352	775	18.95%	26	0	NA
job_201011182036_7009	NORMAL	biddulph	PigLatin:starbucks.pig	44.96%	661	292	9.17%	80	0	NA
job_201011182036_7012	LOW	hadoop	Dashboard Parsing for 2011-01-17	6.66%	15	1	0.00%	0	0	NA

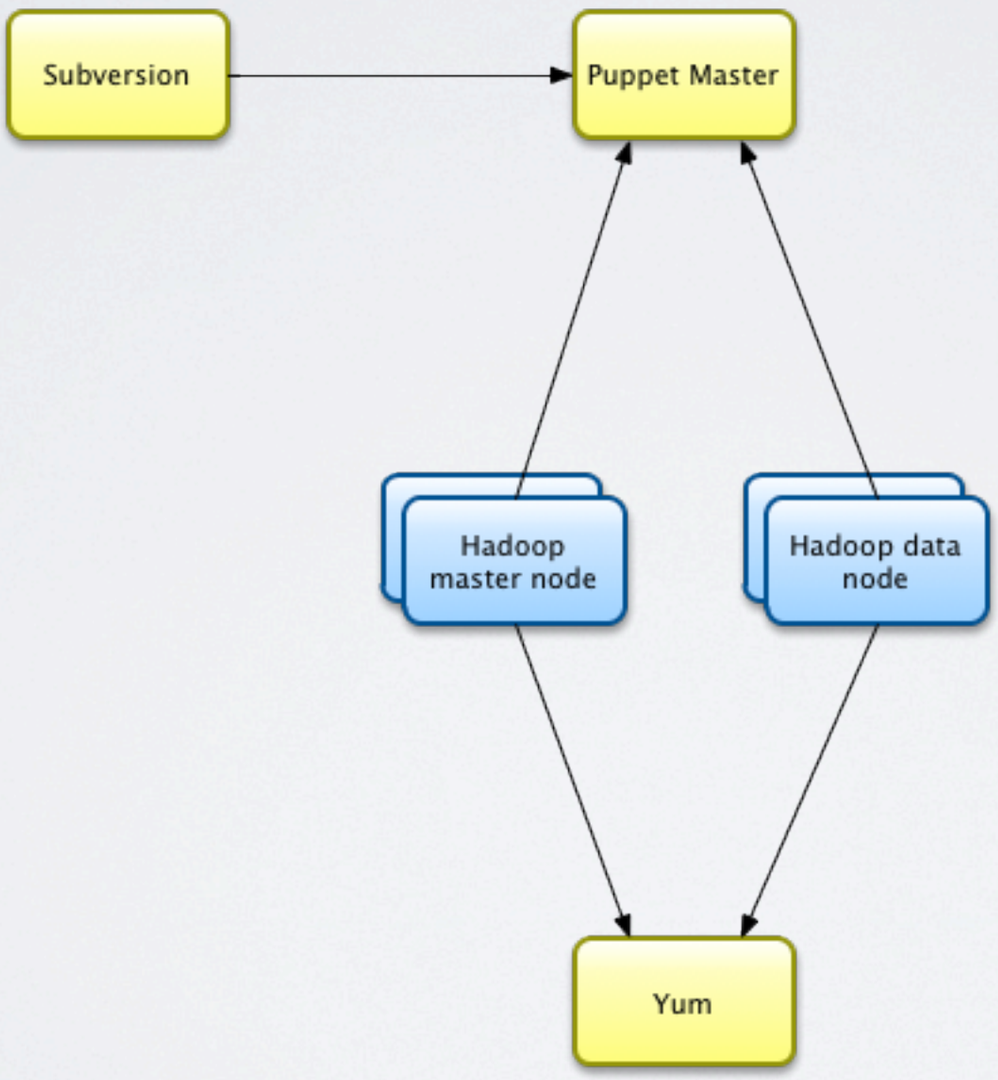
(side note, we use the fairshare scheduler which works pretty well)

INFRASTRUCTURE MANAGEMENT

we use Puppet
Question: Who here has used or knows of Puppet/Chef/etc?

pros
 * used throughout the rest of our infrastructure
 * all configuration of Hadoop and machines is in source control/Subversion
cons
 * no push from central
 * can only pull from each node (pssh is your friend, poke Puppet on all nodes)
 * that's it, Puppet rules

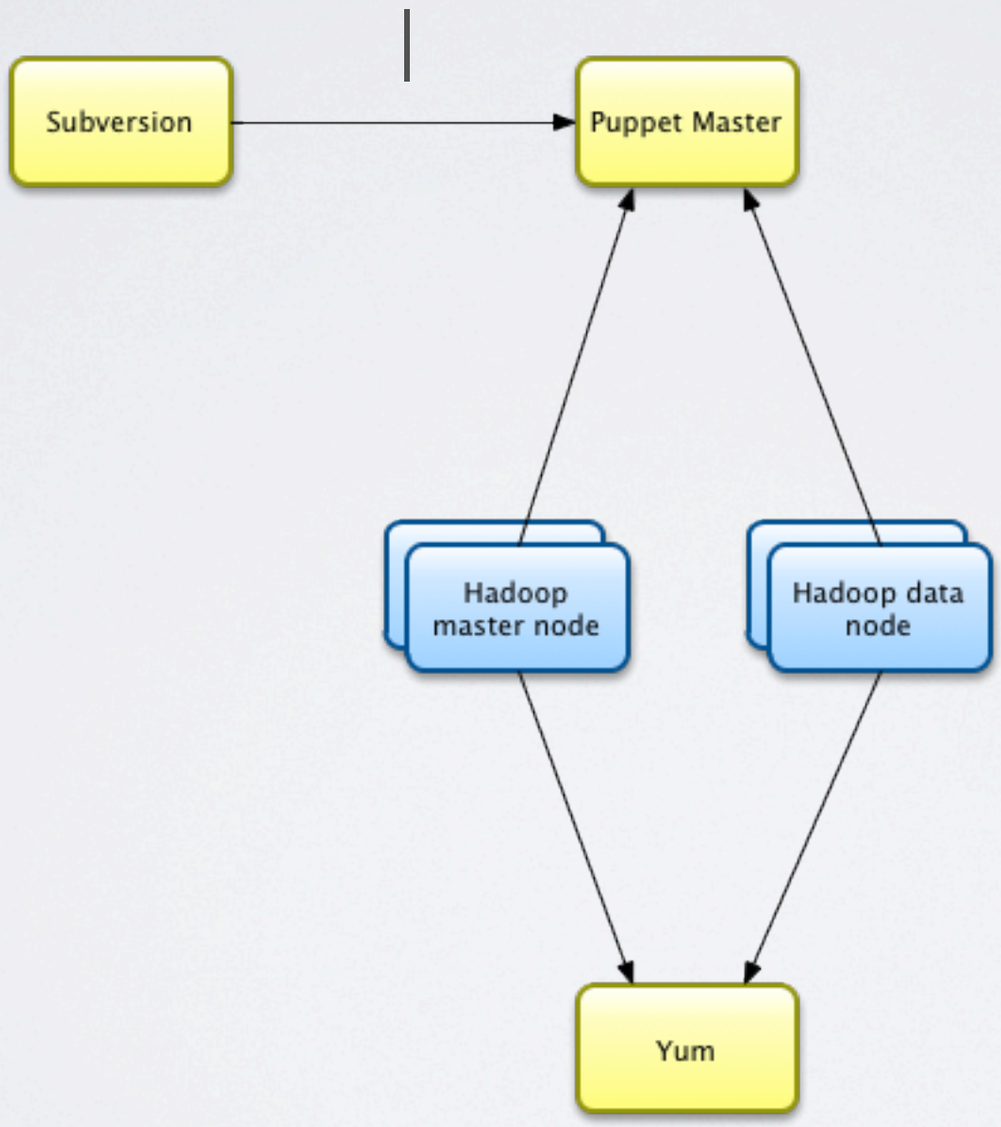
PUPPET FOR HADOOP



Thursday, January 27, 2011

more or less there are 3 steps in the Puppet chain

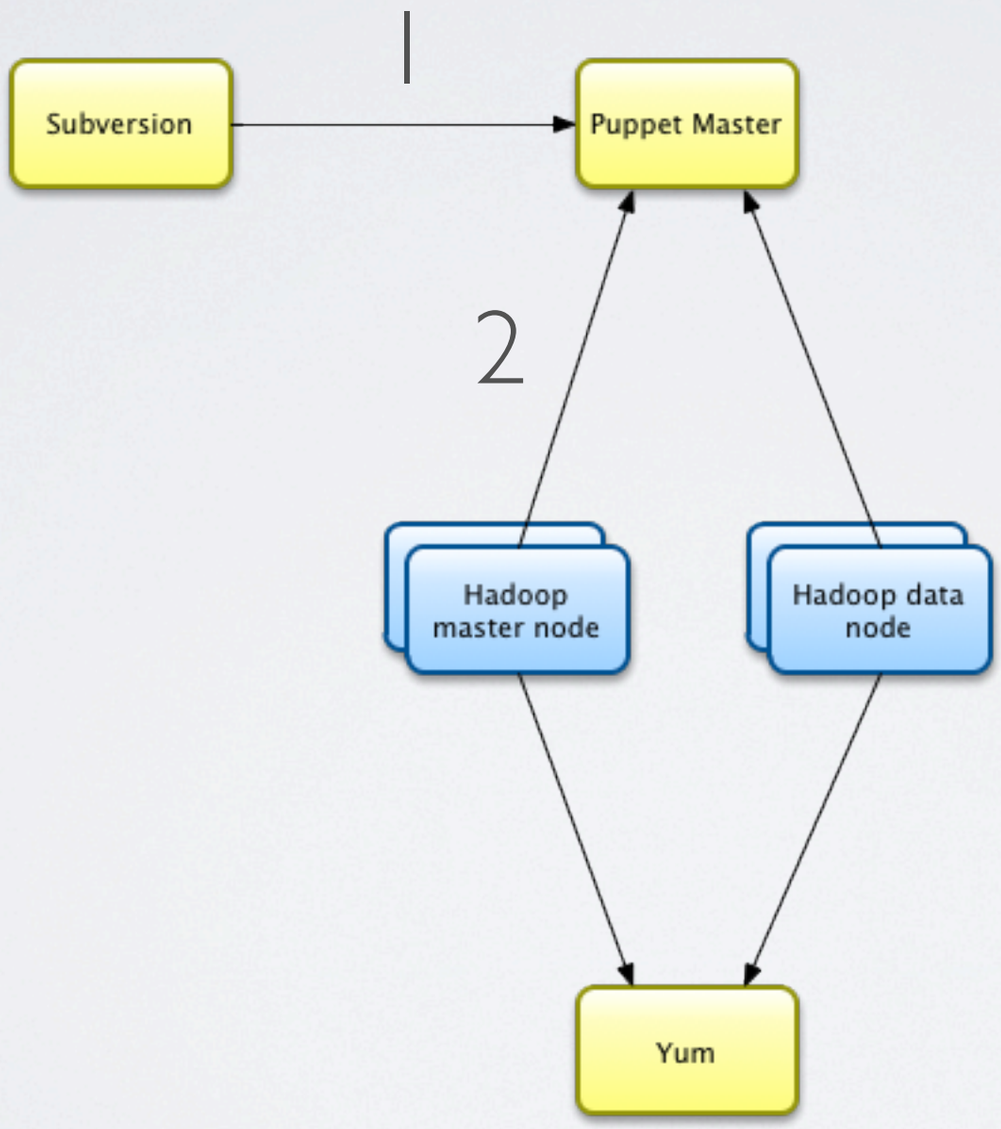
PUPPET FOR HADOOP



Thursday, January 27, 2011

more or less there are 3 steps in the Puppet chain

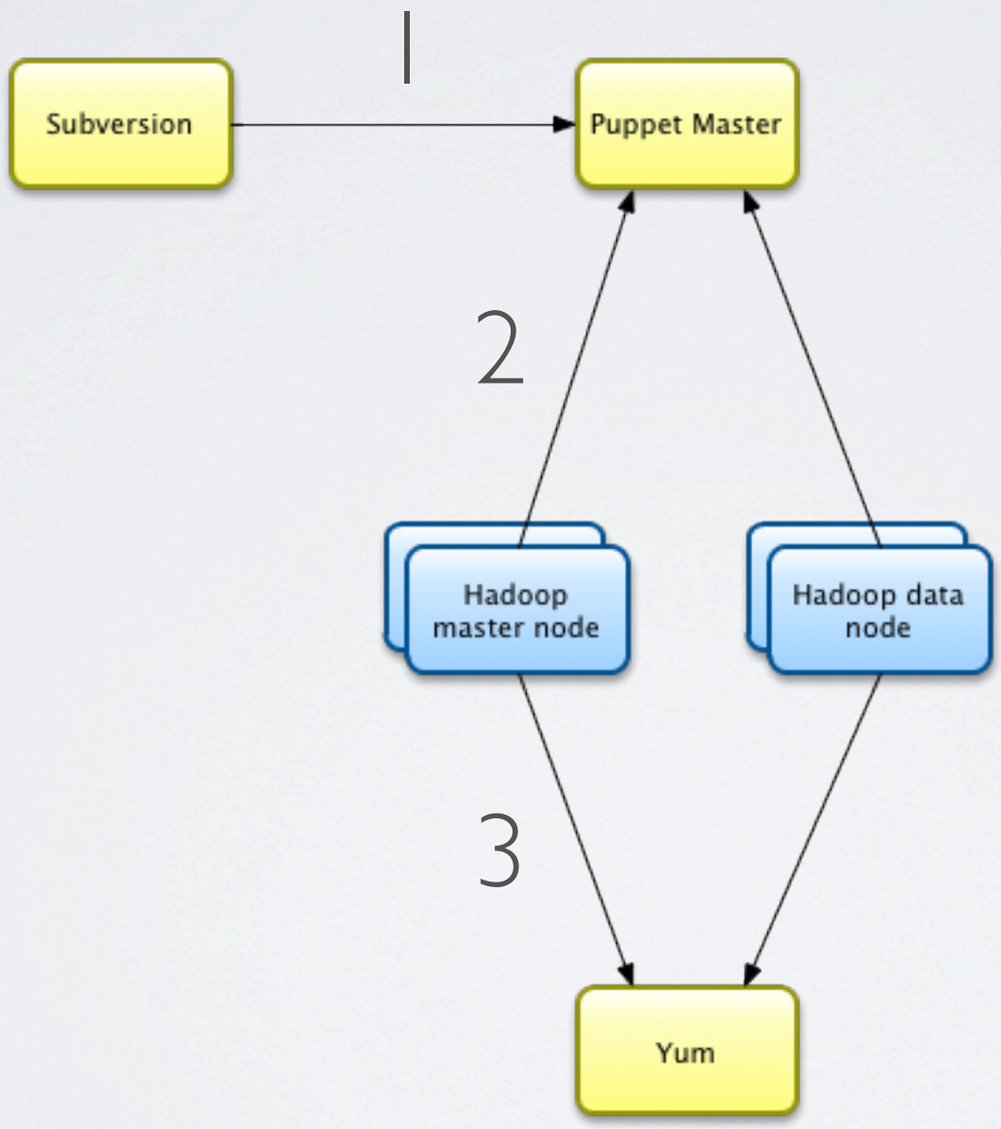
PUPPET FOR HADOOP



Thursday, January 27, 2011

more or less there are 3 steps in the Puppet chain

PUPPET FOR HADOOP



Thursday, January 27, 2011

more or less there are 3 steps in the Puppet chain

PUPPET FOR HADOOP

```
package {
    hadoop: ensure => '0.20.2+320-14';
    rsync: ensure => installed;
    lzo: ensure => installed;
    lzo-devel: ensure => installed;
}

service {
    iptables:
        ensure => stopped,
        enable => false;
}

# Hadoop account
include account::users::la::hadoop

file {
    '/home/hadoop/.ssh/id_rsa':
        mode => 600,
        source => 'puppet:///modules/hadoop/home/hadoop/.ssh/id_rsa';
}
```

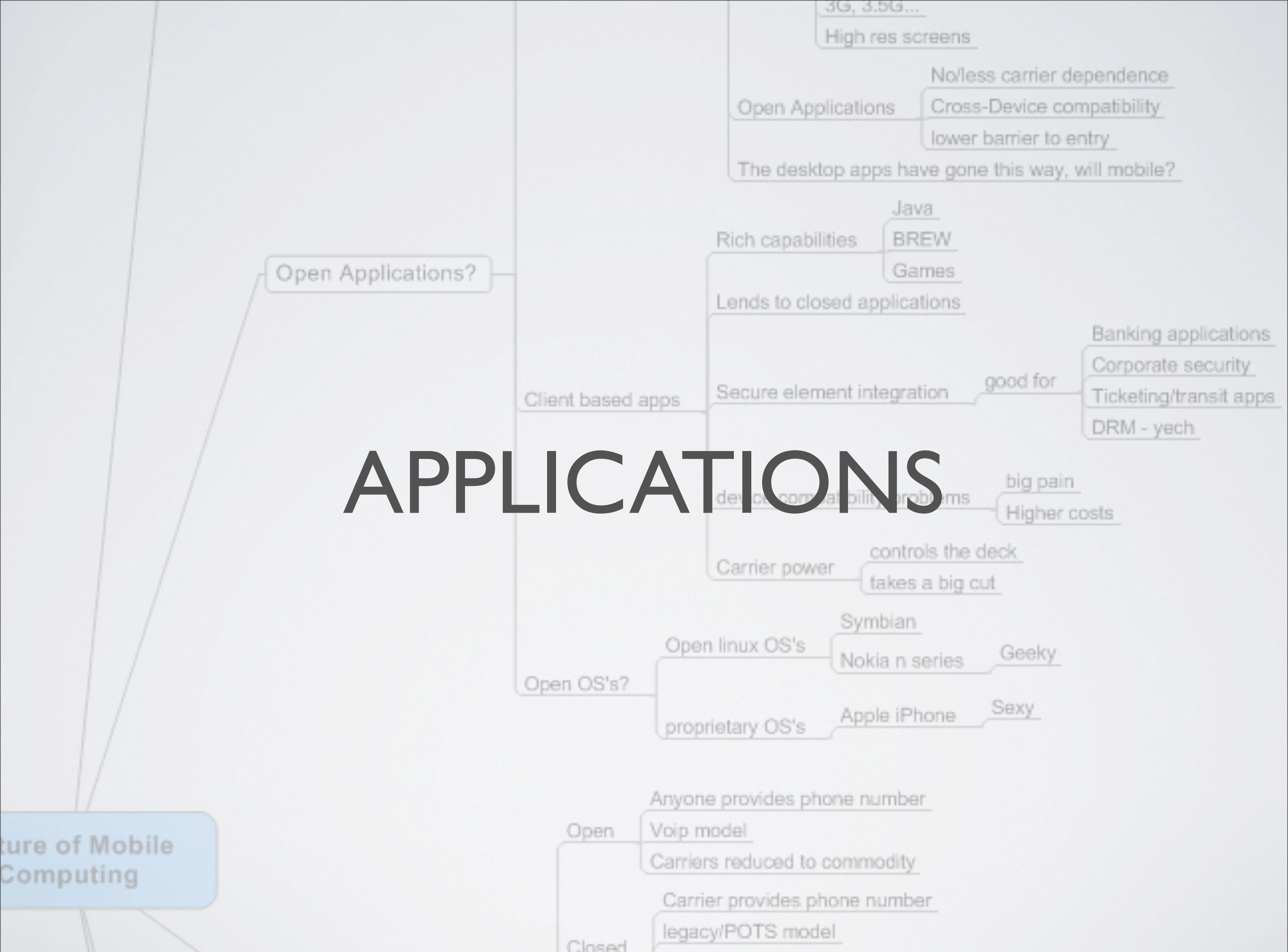

example Puppet manifest
note: we rolled our own RPMs from the Cloudera packages since we didn't like where Cloudera put stuff on the servers and wanted a bit more control

PUPPET FOR HADOOP

```
file {
    # raw configuration files
    '/etc/hadoop/core-site.xml':
        source => "$src_dir/core-site.xml";
    '/etc/hadoop/hdfs-site.xml':
        source => "$src_dir/hdfs-site.xml";
    '/etc/hadoop/mapred-site.xml':
        source => "$src_dir/mapred-site.xml";
    '/etc/hadoop/fair-scheduler.xml':
        source => "${src_dir}/fair-scheduler.xml";
    '/etc/hadoop/masters':
        source => "$src_dir/masters";
    '/etc/hadoop/slaves':
        source => "$src_dir/slaves";

    # templated configuration files
    '/etc/hadoop/hadoop-env.sh':
        content => template ('hadoop/conf/hadoop-env.sh.erb'),
        mode => 555;
    '/etc/hadoop/log4j.properties':
        content => template ('hadoop/conf/log4j.properties.erb');
}
```


Hadoop config files



http://www.flickr.com/photos/thomaspurves/1039363039/sizes/o/in/photostream/

that wrap's up the setup stuff, any questions on that?

REPORTING

operational - access logs, throughput, general usage, dashboards
business reporting - what are all of the products doing, how do they compare to other months
ad-hoc - random business queries

almost all of this goes through Pig
there are several pipelines that use Oozie tie together parts
lots of parsing and decoding in Java MR job, then Pig for the heavy lifting
mostly goes into a RDBMS using Sqoop for display and querying in other tools
currently using Tableau to do live dashboards

other than reporting, we also occasionally do some data exploration, which can be quite fun

any guesses what this is a plot of?

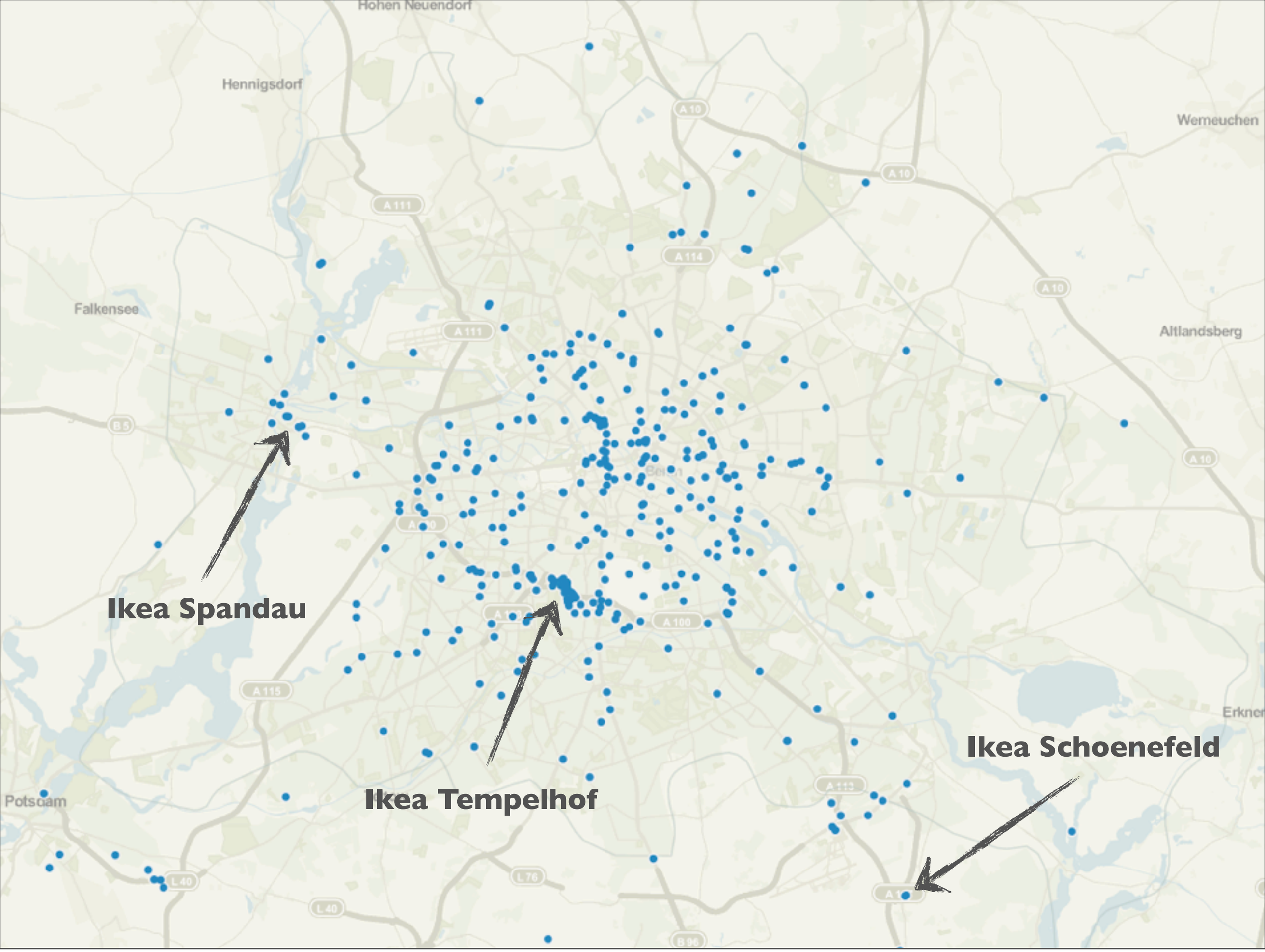
geo-searches for Ikea!

other than reporting, we also occasionally do some data exploration, which can be quite fun
any guesses what this is a plot of?
geo-searches for Ikea!

Thursday, January 27, 2011

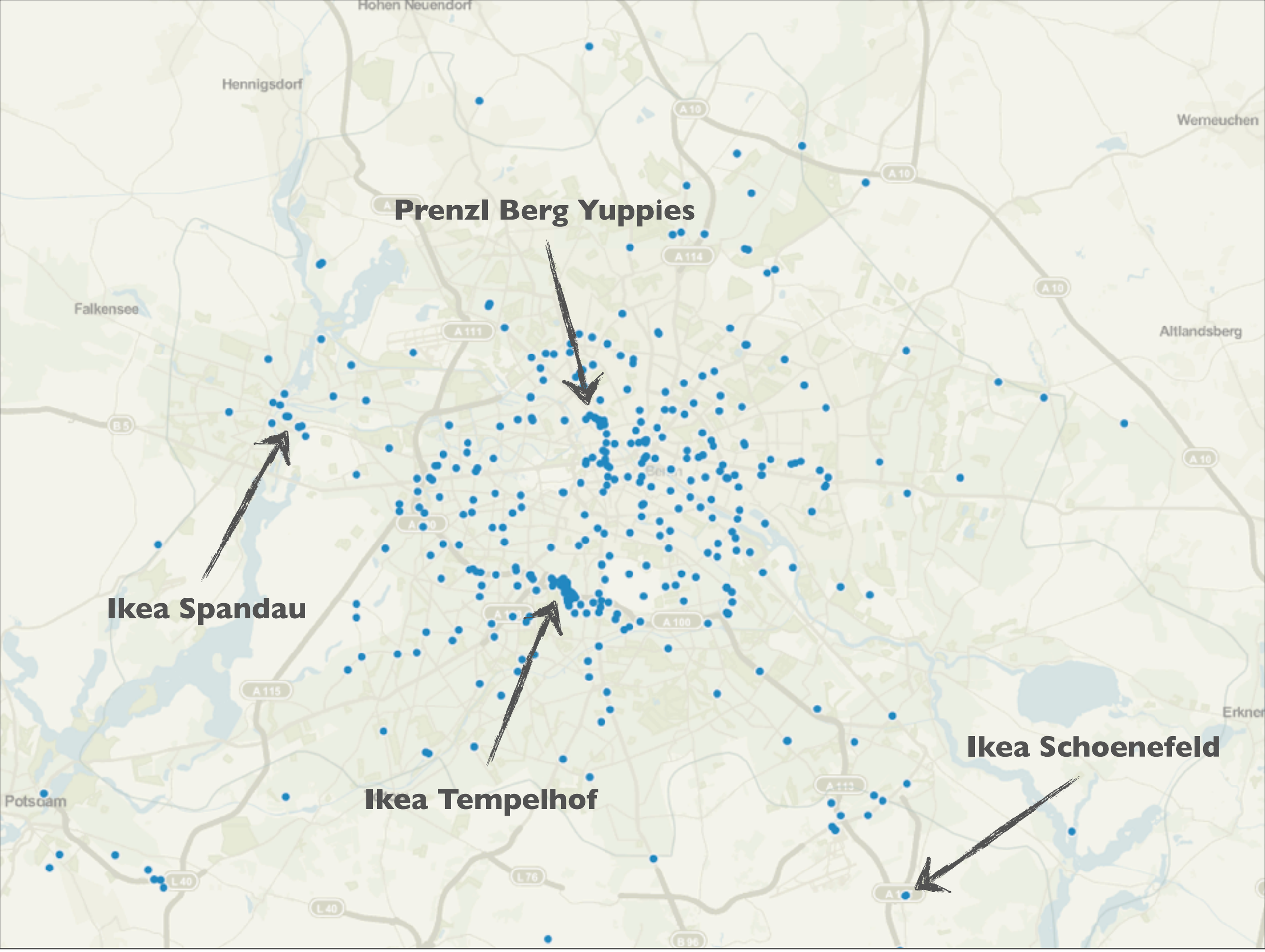
Ikea geo-searches bounded to Berlin
can we make any assumptions about what the actual locations are?
kind of, but not much data here
clearly there is a Tempelhof cluster but the others are not very evident
certainly shows the relative popularity of all the locations
Ikea Lichtenberg was not open yet during this time frame

Thursday, January 27, 2011

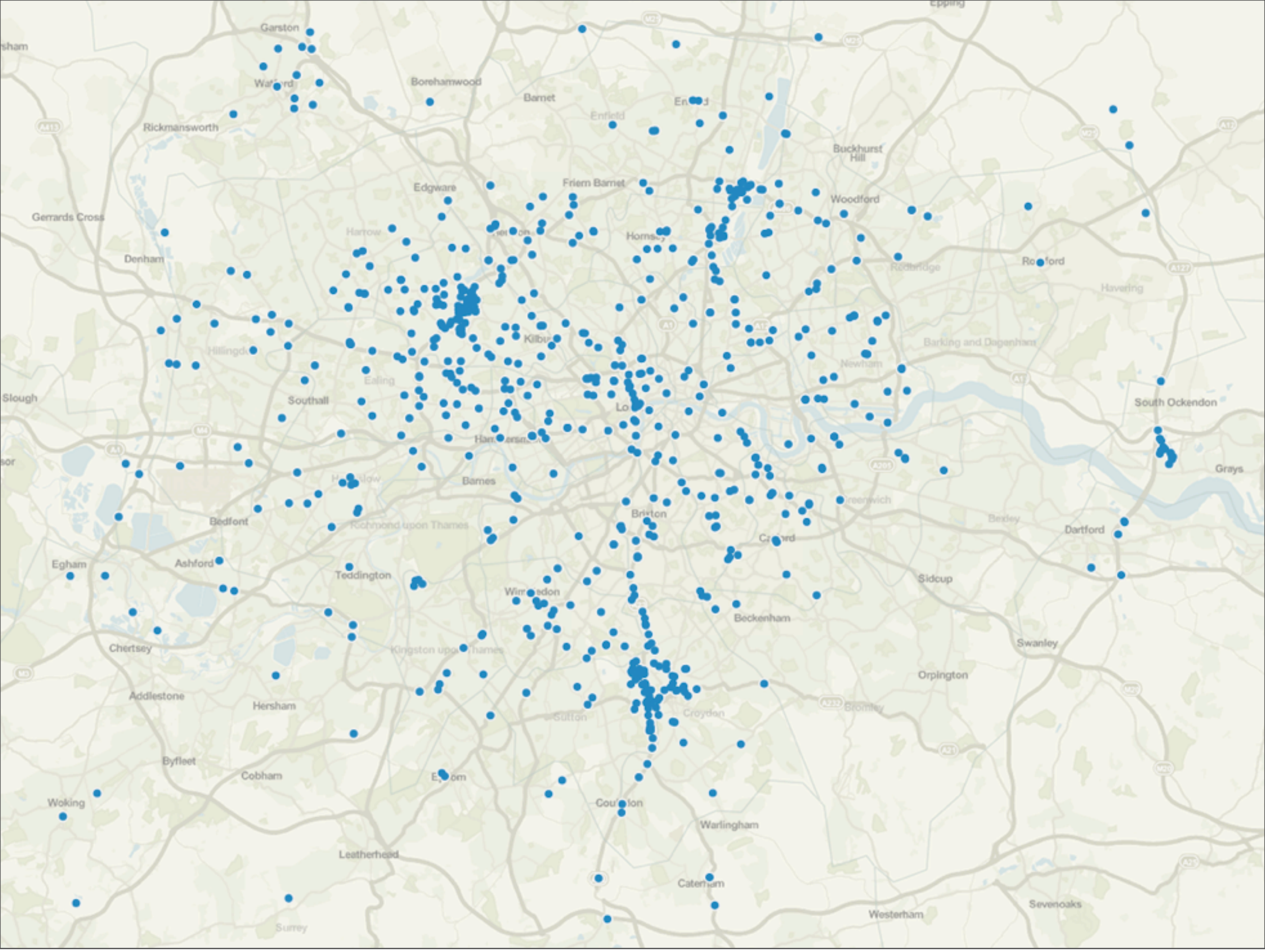


Ikea geo-searches bounded to Berlin
can we make any assumptions about what the actual locations are?
kind of, but not much data here
clearly there is a Tempelhof cluster but the others are not very evident
certainly shows the relative popularity of all the locations
Ikea Lichtenberg was not open yet during this time frame

Thursday, January 27, 2011

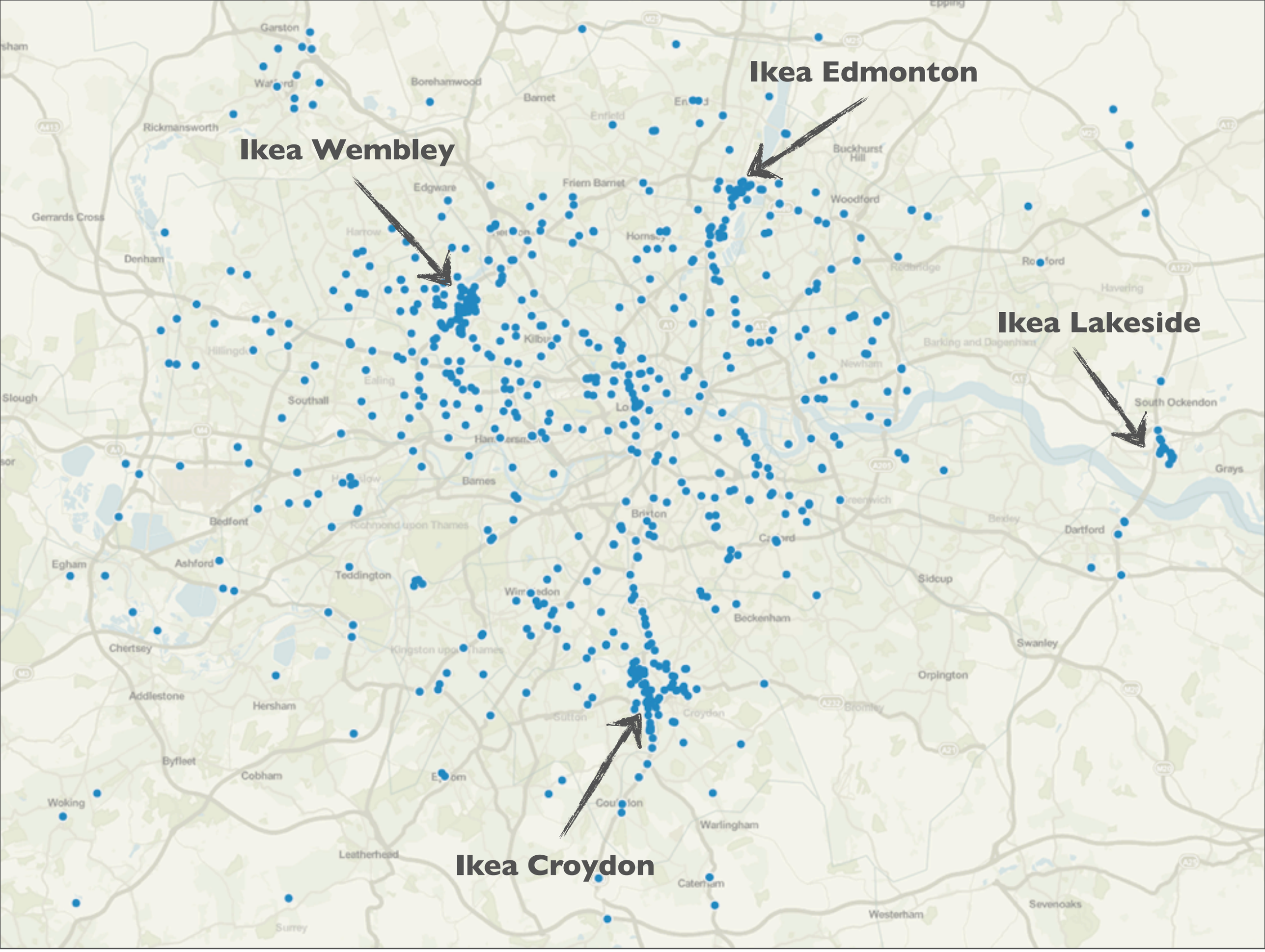


Ikea geo-searches bounded to Berlin
can we make any assumptions about what the actual locations are?
kind of, but not much data here
clearly there is a Tempelhof cluster but the others are not very evident
certainly shows the relative popularity of all the locations
Ikea Lichtenberg was not open yet during this time frame



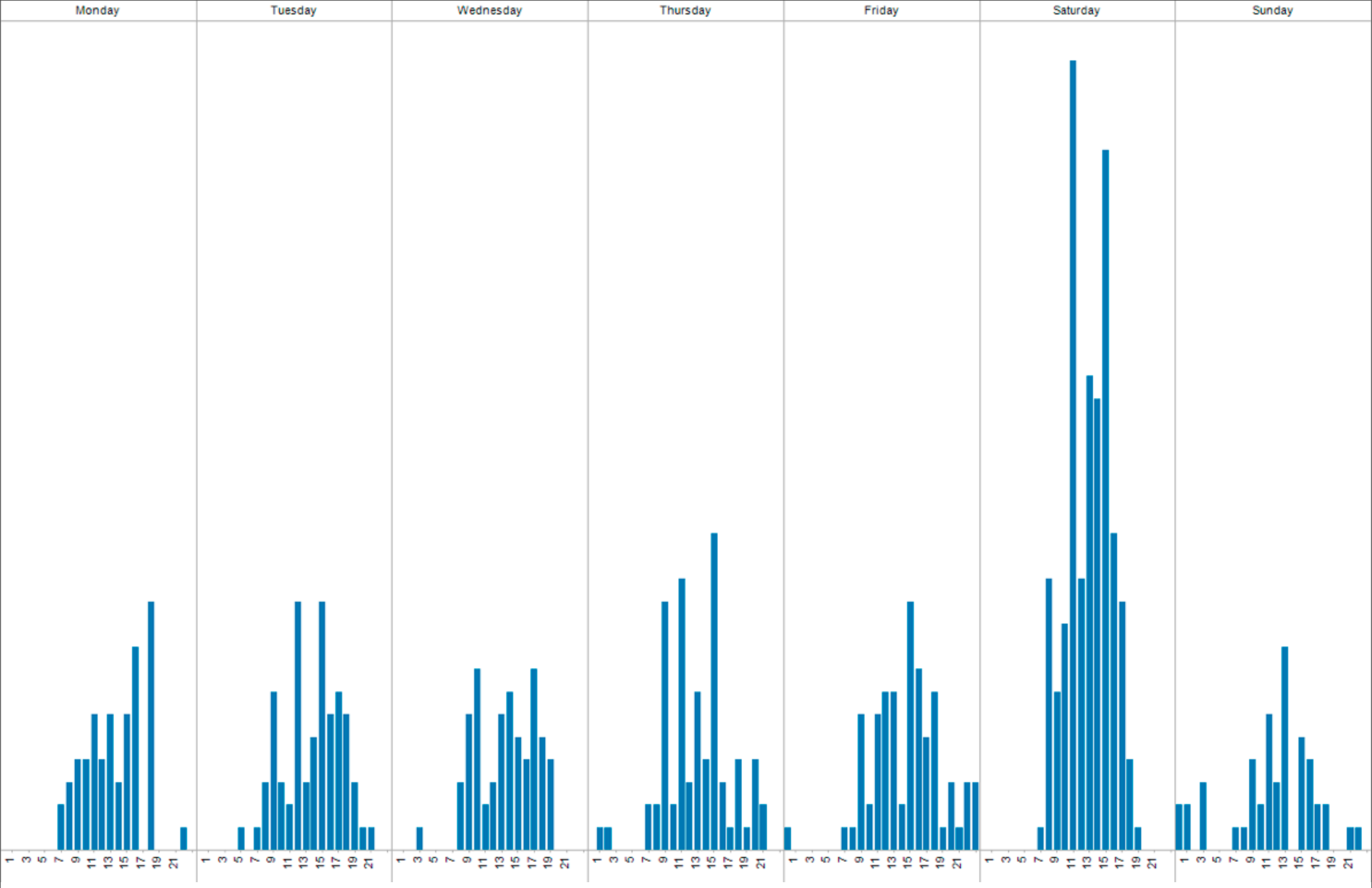
Ikea geo-searches bounded to London
can we make any assumptions about what the actual locations are?
turns out we can!
using a clustering algorithm like K-Means (maybe from Mahout) we probably could guess

> this is considering search location, what about time?



Ikea geo-searches bounded to London
can we make any assumptions about what the actual locations are?
turns out we can!
using a clustering algorithm like K-Means (maybe from Mahout) we probably could guess

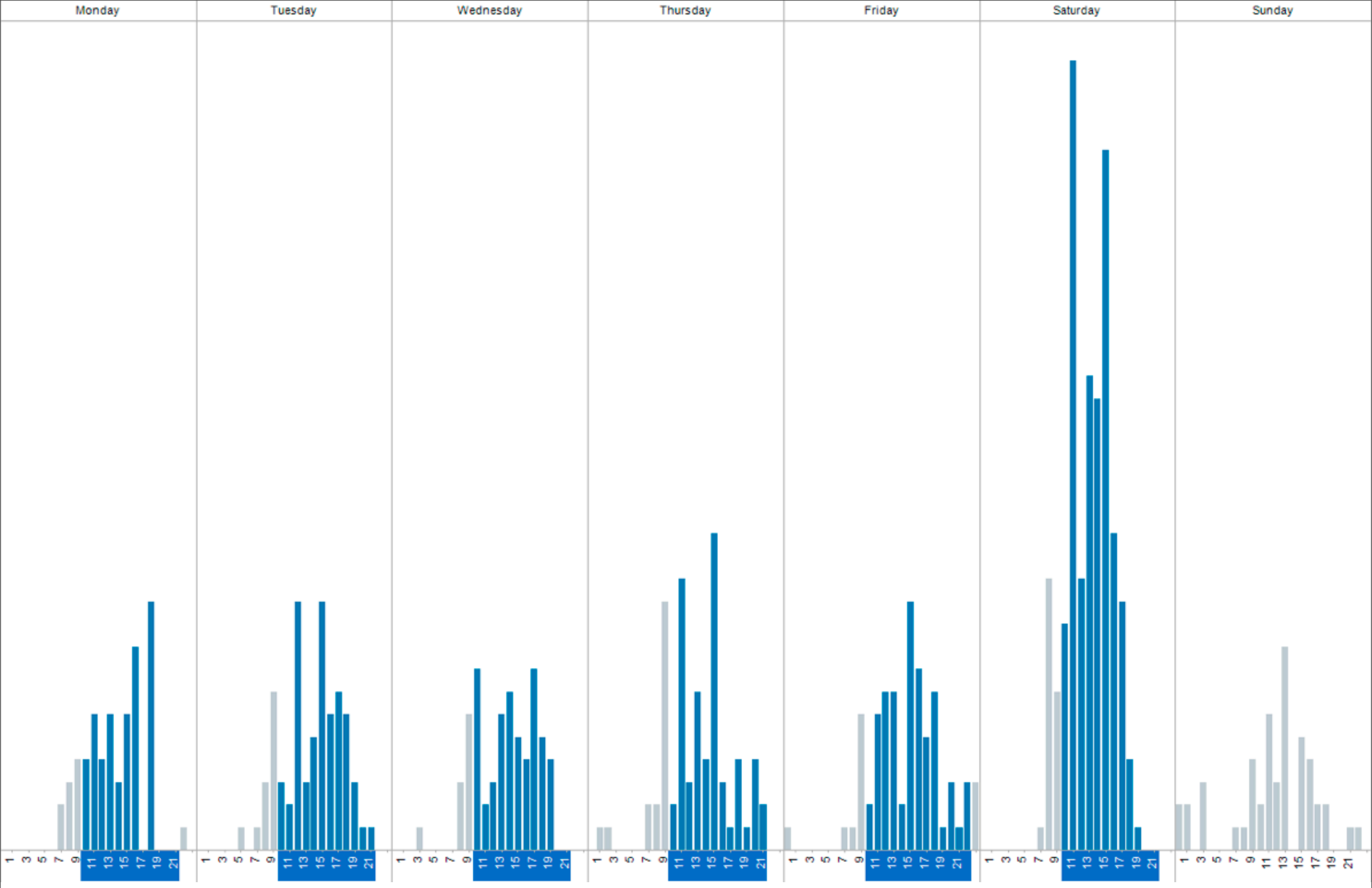
> this is considering search location, what about time?



Berlin

Thursday, January 27, 2011

distribution of searches over days of the week and hours of the day
certainly can make some comments about the hours that Berliners are awake
can we make assumptions about average opening hours?

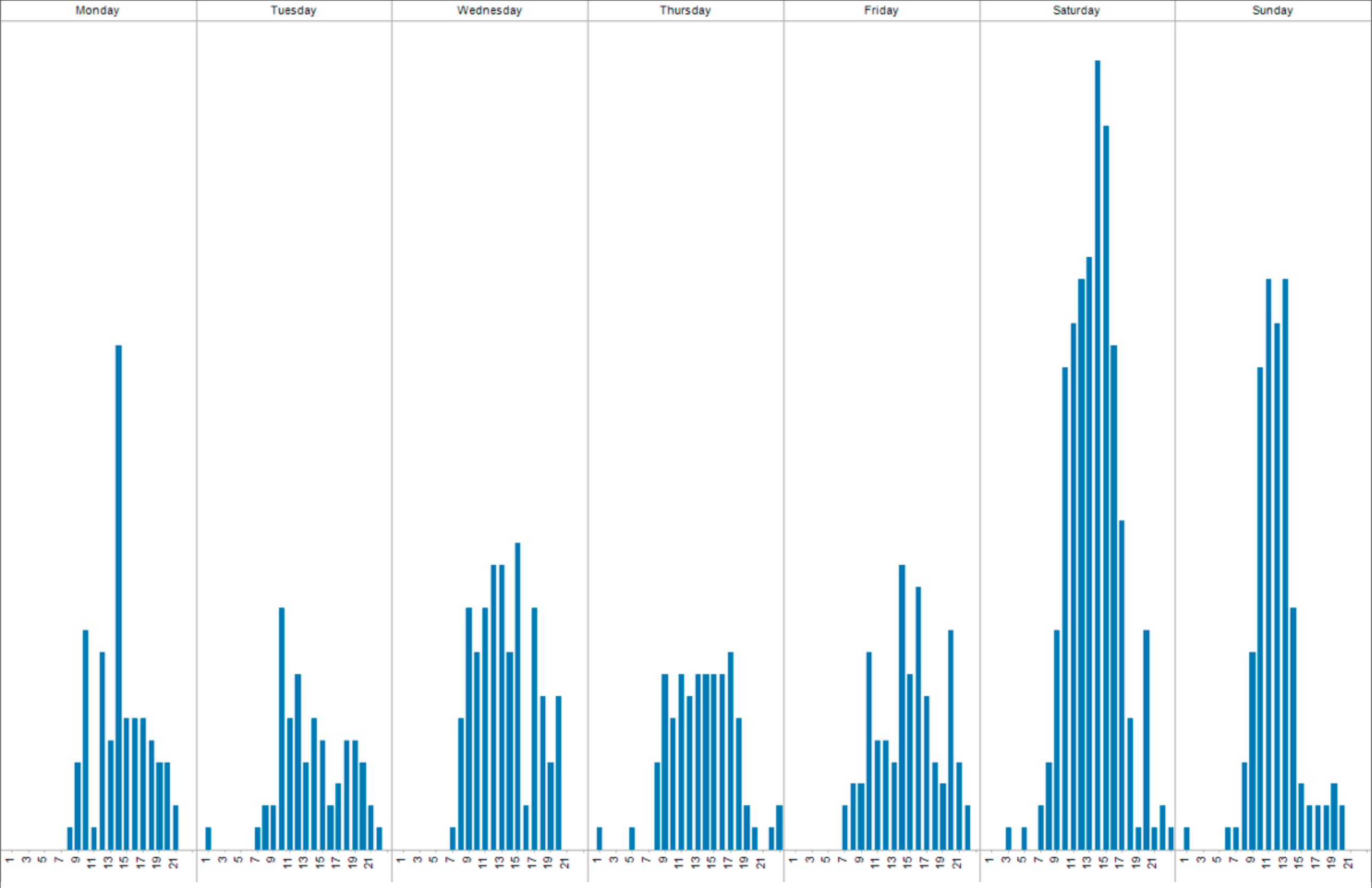


Berlin

Thursday, January 27, 2011

upwards trend a couple hours before opening
can also clearly make some statements about the best time to visit Ikea in Berlin - Sat night!

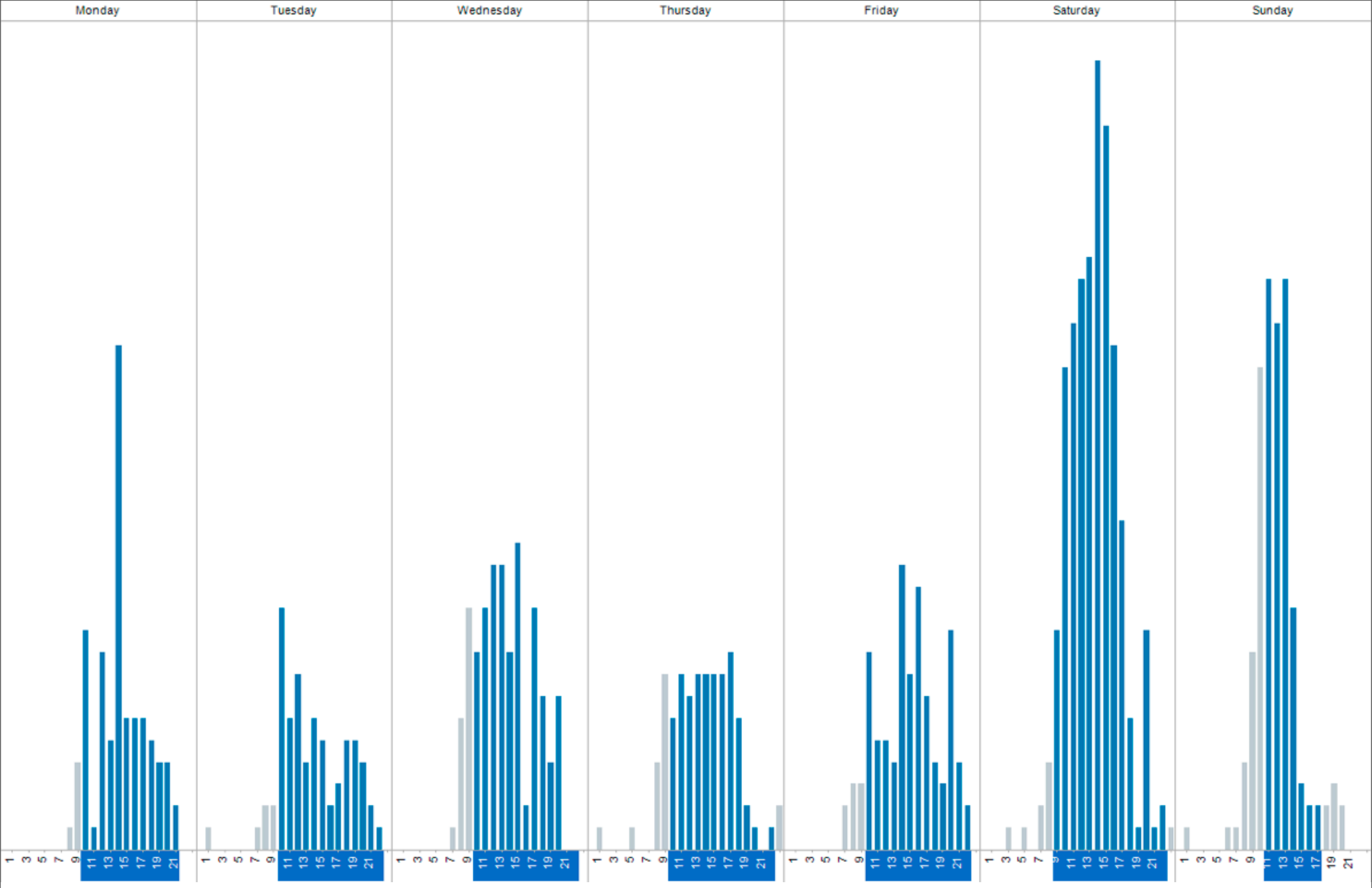
BERLIN
 * Mon-Fri 10am-9pm
 * Saturday 10am-10pm



London

Thursday, January 27, 2011

more data points again so we get smoother results



London

Thursday, January 27, 2011

LONDON

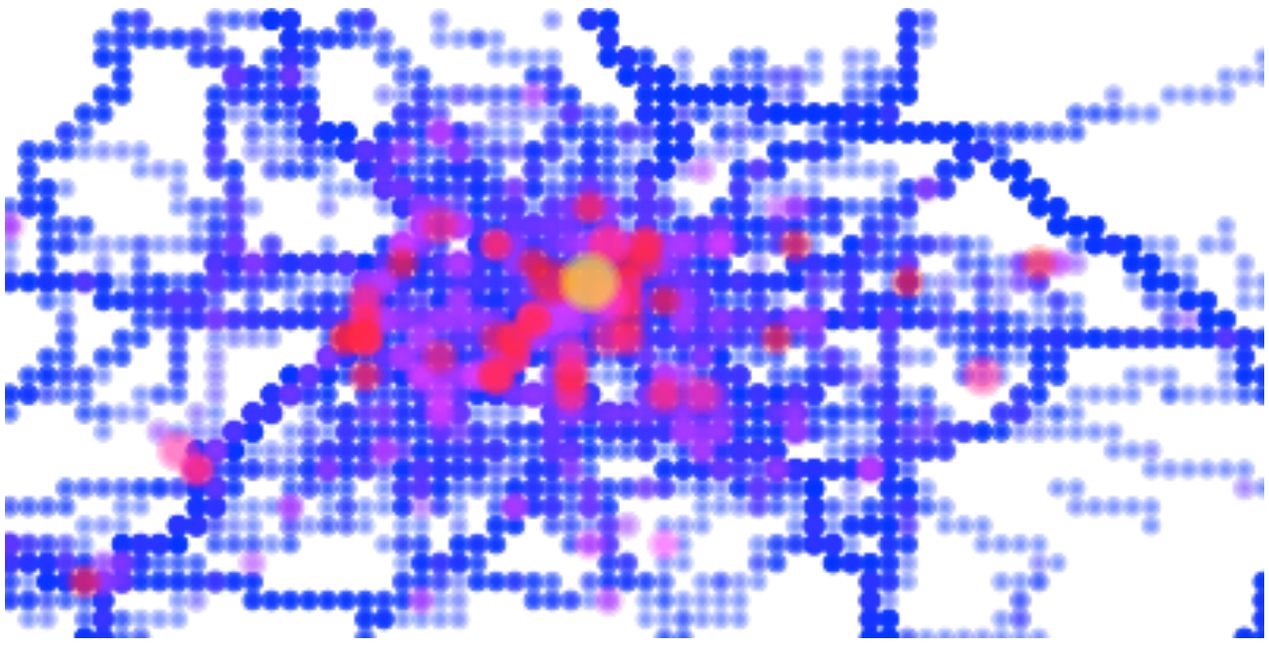
* Mon-Fri 10am-10pm
* Saturday 9am-10pm
* Sunday 11am-5pm

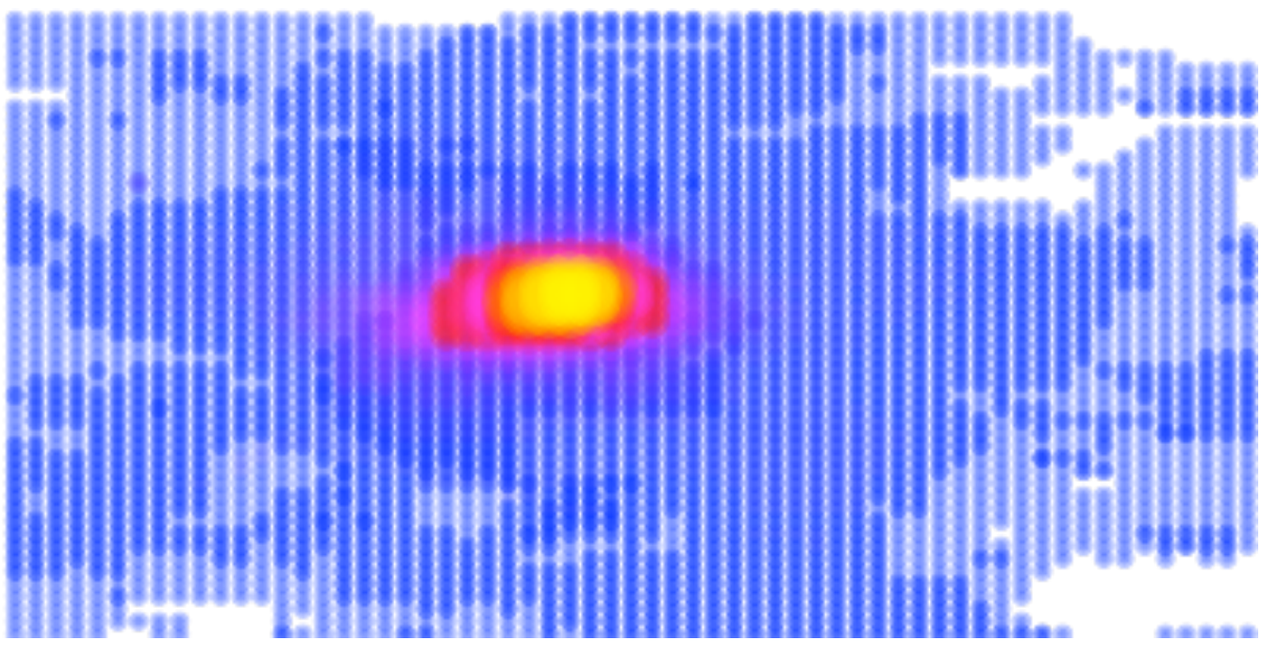
> potential revenue stream?
> what to do with this data or data like this?

PRODUCTIZING

taking data and ideas and turning this into something useful, features that mean something
often the next step after data mining and exploration
either static data shipped to devices or web products, or live data that is constantly fed back to web products/web services

BERLIN

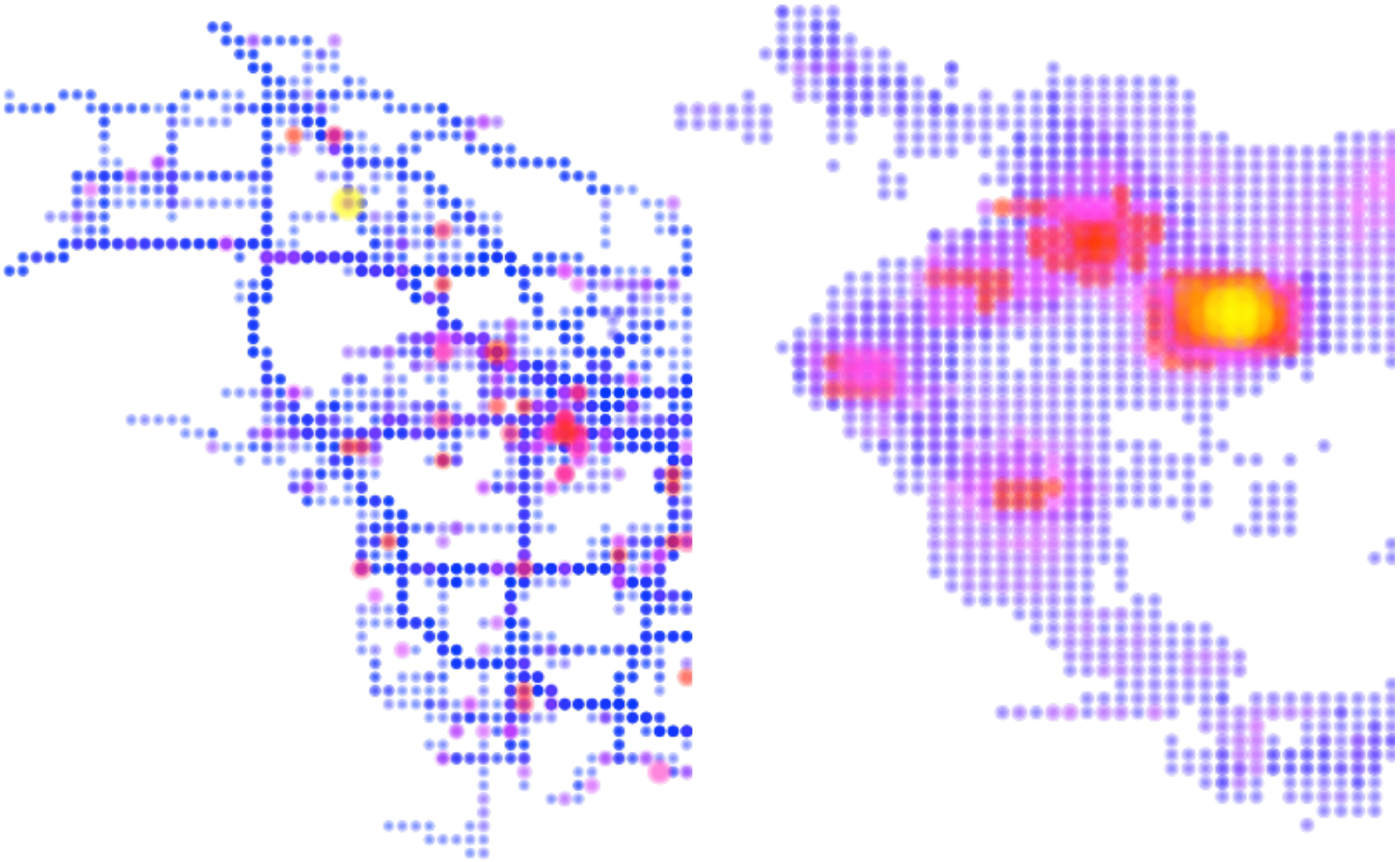




another example of something that can be productized

Berlin
 * traffic sensors
 * map tiles

LOS ANGELES



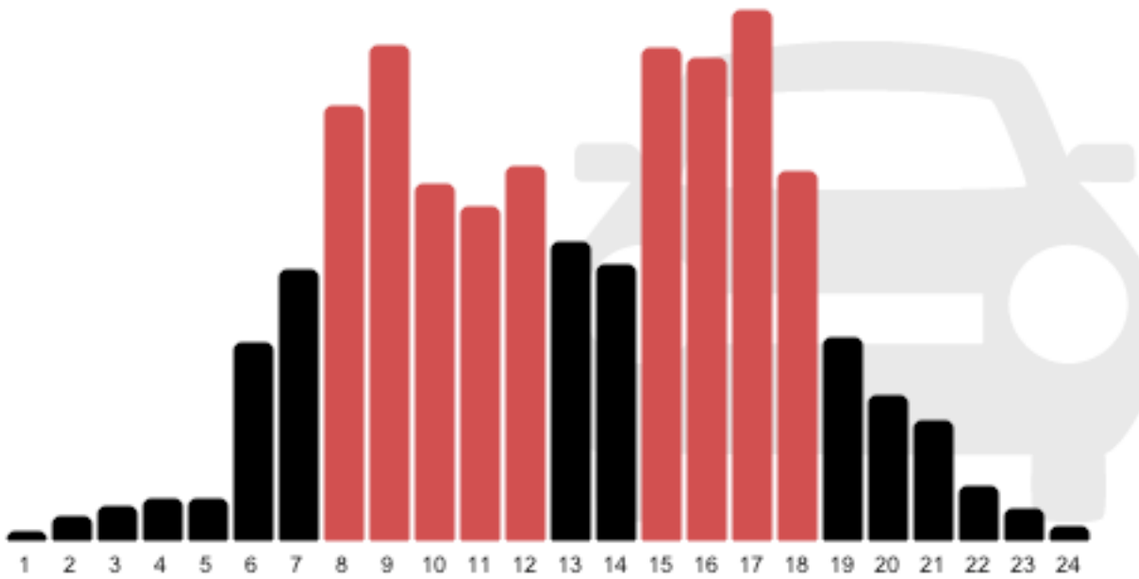
Thursday, January 27, 2011

LA

* traffic sensors
* map tiles

BERLIN

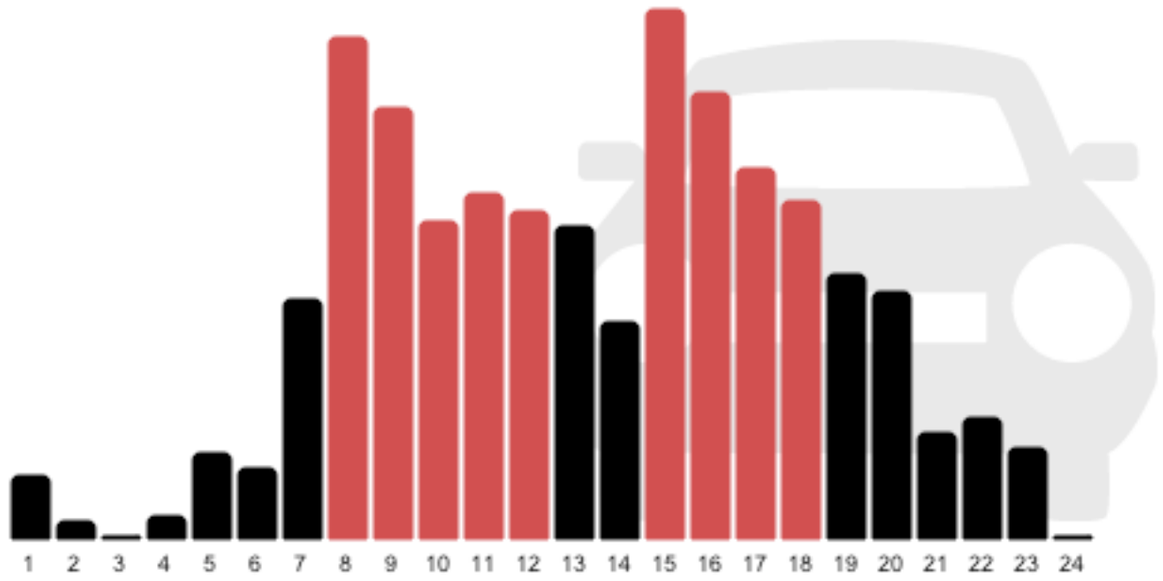
TRAFFIC
VOLUME PER HOUR



STARBUCKS INDEX
MEASURES STARBUCKS PENETRATION

LA

TRAFFIC
VOLUME PER HOUR



STARBUCKS INDEX
MEASURES STARBUCKS PENETRATION

Starbucks index comes from POI data set, not from the heatmaps you just saw

JOIN US!

- Nokia is hiring in Berlin!
  - analytics engineers, smart data folks
  - software engineers
  - operations
  - josh.devins@nokia.com
  - www.nokia.com/careers

THANKS!

JOSH DEVINS www.joshdevins.net @joshdevins