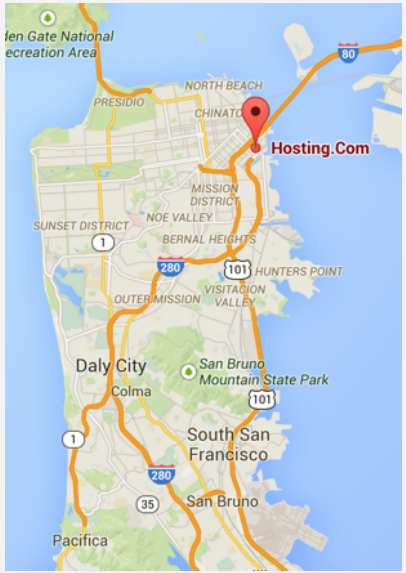
Infrastructure

GEX - 19/12/2014

viadeo

Current status

- Web platform hosted by Hosting.com in their San Francisco datacenter (Downtown @ 630 3rd Street)
- Datacenter is from an older generation, but that's not a problem in itself



- 325 hardware servers, 15 racks
- Our infrastructure is well-cared for and under control
- Life is good, then? Well, yes :) However...

Issue #1 : hardware aging

- No new server purchase since August'13
- End of '14: 67 servers over 4 years old (20% of total platform)
- End of '15: count rises to 143 (40%)

- Failure rate significantly rises after 3.5 years
- Total Cost of Ownership (TCO) degrades compared to newer server generations: lower density, higher power consumption

→ Best practices advise to replace servers every 3 to 3.5 years

IDC study (2012): "*The Cost of Retaining Aging Infrastructure*"

Issue #2 : expired support

- New servers are usually supported by a 3-year contract, which Viadeo did purchase
- End of ‘14: 132 servers no longer supported (40% of total platform)
- End of ‘15: count rises to 277 (85%)
- Very little spare hardware, none for some configurations

- Issues #1 & #2

  → growing risk of unsupported hardware issues in 2015

  → production issues that may not be easy to fix quickly
- Renewing 3-year support is costly

  → buying new servers is a better deal

Issue #3 : earthquakes

- Downtown San Francisco is a highly seismic area
- MMI 8 risk (max is 10)
  - *Nearly everything thrown down from shelves, cabinets, and walls*
  - *Poorly constructed buildings suffer partial or full collapse. Some well constructed buildings are damaged. Unreinforced walls fall, etc.*
- Even if the datacenter is intact, what about roads (for fuel deliveries), water (for AC), network connectivity, staff, etc. Sandy in NYC was a wake-up call!
  → Assume that severe downtime will occur and ask yourself: "do we have a plan?"

IBM datacenter (Dec'11, MMI 6 in Tokyo)

Napa Valley (Aug'14, MMI 7)

http://resilience.abag.ca.gov/earthquakes/sanfrancisco/
http://www.datacenterknowledge.com/archives/2012/11/01/ny-data-centers-battle-back-from-storm-damage/

Issue #4 : no Disaster Recovery plan

- Disaster Recovery: “how do we restore some level of service ASAP?”

  ≠ Business Continuity Plan: “how do we keep running the company no matter what (even if we’re 100% down)?”

- Today:
  - All data (mySQL, HBase, NFS) is replicated to the Amazon cloud at least once a day

    → no more than 24 hours of data would be lost
- And that’s it :(
  - No secondary datacenter to switch to
  - No procedures to rebuild our infrastructure
  - No idea how long it would even take (1 week? 2 weeks?)

OK, now what?

- Option #1: buy new servers and/or renew support
  - Solves issues #1 & #2, at the expense of CAPEX/OPEX upfront
  - No progress on #3 & #4.
- Option #2: build infrastructure in another datacenter
  - Solves all issues, but means years of inefficient CAPEX/OPEX: our business doesn't NEED a 2nd DC
- Option #3 : the so-called "Virtualization Project"
  - Consolidate all our infrastructure in the same DC (blades/VMWare)
  - Solves #1 & #2 elegantly, reducing TCO in the long run
  - 500K€ (60% CAPEX)
  - No (or marginal) progress on issues #3 & #4
  - Would have been great 2 years ago, but now less and less relevant as Limbo/Kasper migration eases pressure on our infrastructure

Option #4: jump into the Amazon cloud

- No more hardware issues. Ever (~~#1 & #2~~)
- Safe from natural disasters (~~#3~~)
- Ability to quickly (re)deploy “anywhere”. For “free” (~~#4~~)
- Zero CAPEX. OPEX can be aligned to business health

- We’re already using AWS for Analytics
- Excellent relationship, PR opportunities :)
- IT team super motivated by the project!

Proposal for next steps

- Finance / Legal
  - Estimate AWS costs at cruising speed (factoring in “free” Disaster Recovery)
  - Estimate cost overlap between DC ramp-down and AWS ramp-up
  - Cancel automatic renewal on all S.F. services, renegotiate dates for end of service
  - Check any potential issue with privacy and fiscal rules: Safe Harbor applies as long as data is stored in the US, but AWS has multiple locations in different states
- Tech
  - Start to decommission servers & racks ASAP (clean-up & cost reduction)
  - Build a demo platform in AWS, which could scale “instantly” to a production platform (*mid-2015 reasonable?*). Then (and only then), start moving everything else to AWS
  - In parallel: build and test basic Disaster Recovery procedures to switch from S.F. to AWS
  - Excluding contractual engagements, be ready to close the S.F. DC *at the end of 2015?*
- HR

Thanks for listening!