# Moving Viadeo to AWS

Julien Simon, *Chief Technology Officer*

jsimon@viadeoteam.com
@julsimon

# Viadeo

## France

- Founded in 2005
- Market leader, with 9m members
- Very strong ties with local recruiters
- Strong corporate offering, with 21% of top 1,000 companies recruiting on Viadeo

## Russia

- JV 50/50 with local partner,
- 1m members
- Position in high quality professional profiles

## China

- Acquired in 2008
- #1 with 25m members
- Trusted source of professional data in China
- Bringing trust into the professional world
- Strong local management team

## French-speaking countries in Africa

- Ideally positioned organically with strong momentum
- #1 with 3m members
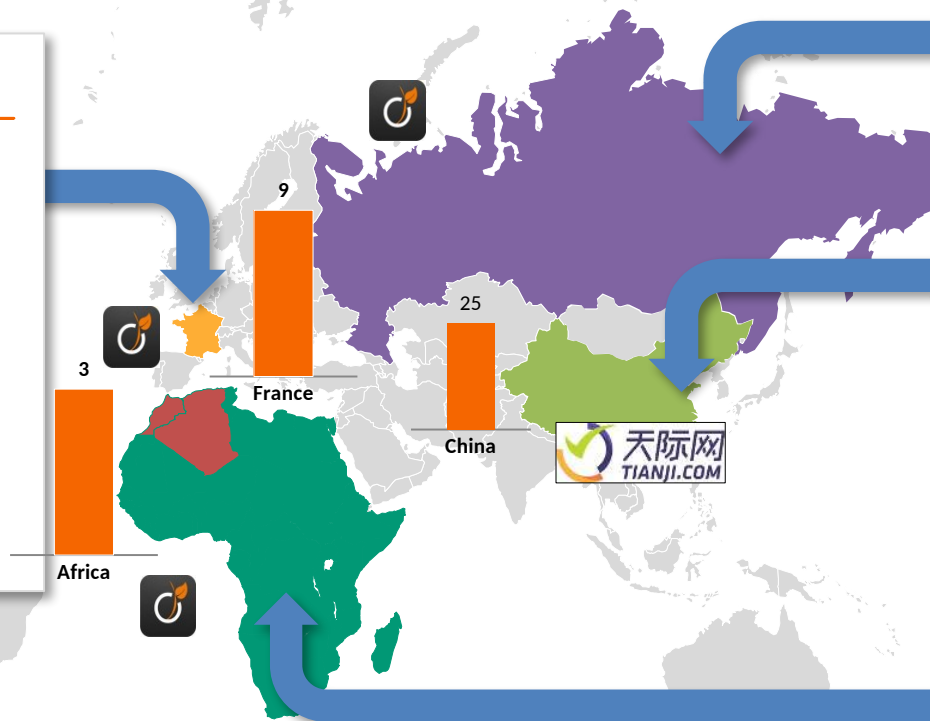
9
France

3
Africa

25
China

天际网
TIANJI.COM

viadeo

Rechercher    Recherche avancée    18    3

Profil    Contacts    Jobs & Formations    Entreprises    Participer    Accès recruteur    Premium

Julien Simon ★ ∞
Recrute des spécialistes Java/Data/Sécurité

∞ « Rencontrons-nous » ×

Chief Technology Officer, Viadeo / Modifier

Aperçu de mon profil

📷 Modifier ma photo

Paris, France / Modifier    130 contacts    Partager mon profil  f  g+  🐦  ♂
/ Modifier mes cartes de visite

En résumé    Mes contacts    Ma bibliothèque    Mon activité sur Viadeo    + Nouveau !

Ma présentation / Modifier

***
*** VIADEO RECRUTE
*** http://www.viadeo.com/fr/company/viadeo/job
***

Directeur Technique / Chief Technology Officer

- Définition et pilotage de projets stratégiques complexes (design, build, run)
- Capacité à innover et à transformer de manière continue (organisation, technologie, produit)
- Relations et négociations internationales (clients, partenaires, fournisseurs)
- Encadrement d'équipes jusqu'à 150 personnes en France et à l'étranger

Ma bibliothèque / Modifier

Java 8 in Action: Lambdas, streams, and function...
Raoul-Gabriel Urma, Mario Fusco, Alan Mycroft
★ ★ ★ ★ ☆

The Pragmatic Programmer
★ ★ ★ ★ ★

Systems Performance:

Face à Face
Rencontres autour de votre métier

Ils ont un objectif    Ils recrutent

Rejoignez nos face-à-face métier

Carrière    Réinitialiser

Intérêts des projets / missions
Responsabilité / Autonomie
Reconnaissance du manager
Formation et développement des compétences
Evolution de carrière / Mobilité

Projets

Evolution    Responsabilité

Formation    Reconnaissance

👤 4.2 /5    👥 3.8 /5
mon évaluation    la moyenne

---

viadeo

Rechercher    Recherche avancée    17

Profil    Contacts    Jobs & Formations    Entreprises    Participer    Accès recruteur    Premium

Quelle banque en ligne peut vous faire économiser jusqu'à 191€/an sur vos frais bancaires ?

Comme Laura, retrouvez vos connaissances et devenez 3 fois plus visible !    ×

Email    Importer

Viadeo protège votre vie privée ! Votre mot de passe ne sera pas mémorisé. Aucun message ne sera envoyé à vos contacts sans votre autorisation.

Publier un lien, un texte...

Nicolas PARRON suit une entreprise    ⏱ 1m

DCNS
DCNS    + Suivre
+10000 employés

J'aime

Emmanuel DETROY a de nouveaux contacts    ⏱ 41m

Sylvia Mund    + Ajouter
Comédienne voix off allemande et chant

Marine Tumelaire    + Ajouter
Artist illustrator, comics book author and colorist, Art teacher

LES CONNAISSEZ-VOUS ?
plus de 100 contacts potentiels    ↻

Marie Jeanne Vogel    ×
Manager - Ressources...

Frédéric Cesbron    ×
Chef de projet études,...

Christophe GACEM    ×
Directeur des système...

> Voir mes contacts potentiels

AVEC LE SERVICE
> VOTRE POTENTIEL MARCHÉ
UN CONSULTANT APEC VOUS AIDE
À IDENTIFIER LES SECTEURS
QUI VOUS CORRESPONDENT.

**Rencontres professionnelles**

écran à la réalité, il n'y a plus qu'un pas !
x membres Viadeo de vous rencontrer dans la vraie vie
∞

fév
d'hui

1468 visites depuis votre inscription

Qui a consulté mon profil ?

> Voir tous mes visiteurs

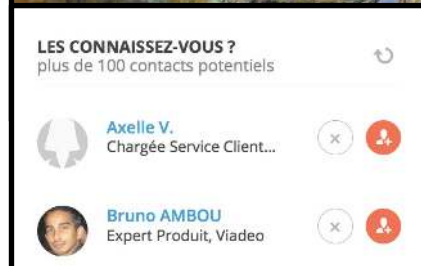RISTRAM et Michel Bellone aiment cela

# Our current platform

- Infrastructure : 250 Linux servers hosted in San Francisco, CA

- Backends : MySQL, ElasticSearch, Hbase, Spark, Hadoop

- Service-oriented platform : « Kasper », Java
  - CQRS : *Command-Query Responsibility Segregation*
  - ES : *Event Sourcing*
  - DDD : *Domain Driven Design*

- Web application : « Limbo », Node.js

- Mobile applications : iOS and Android
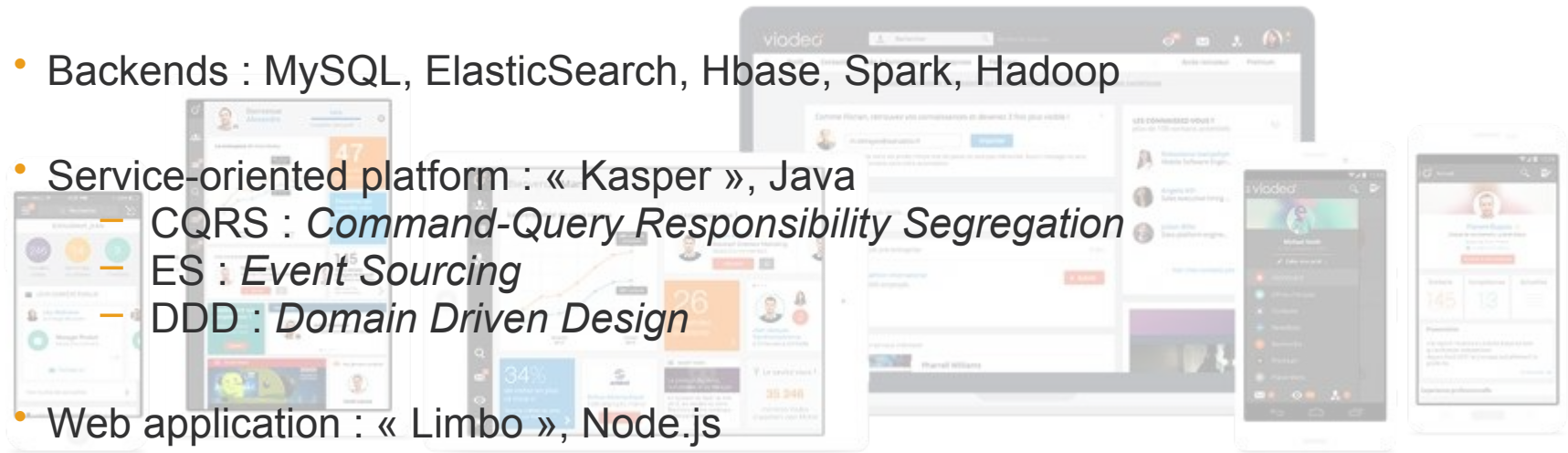
# AWS @ Viadeo : the story so far

- 2011 : file storage (S3)

- 2012 : Viadeo office flooded! Internal servers destroyed
  → All services rebuilt in AWS in **48h** (VPC, EC2)

- 2013 : data processing (Hadoop, EMR)

- 2014 : more data !
  - New analytics infrastructure :
    Snowplow → S3 → Redshift
    (≈ 20 million events / day for starters)
  - Content personalization : EMR, Spark



LES CONNAISSEZ-VOUS ?
plus de 100 contacts potentiels

Axelle V.
Chargée Service Client...

Bruno AMBOU
Expert Produit, Viadeo

Cet article pourrait vous intéresser

205 jours... avant que les entreprises ne découvrent qu'elles sont cyber-attaquées

En usurpant l'identité des administrateurs informatiques, les hackers sont capables de passer inaperçus de longs mois dans le réseau informatique de leurs victimes. Combien faut-il de temps à une entr...

Source : usine-digitale.fr

J'aime    5

# AWS @ Viadeo : the story so far

- Infrastructure live in 3 regions (us-east-1, us-west-2, eu-west-1)
- 4 VPC
- ≈ 100 EC2 instances (half production, half development)
- ≈ 15 TB in S3
- 2 Redshift clusters (5 instances each)

- AWS services currently in production     at Viadeo :
  - Core infra : VPC, Route 53, IAM, S3, CloudFront,
  - Instances : Elastic Load Balancing, Elastic Beanstalk, CloudFormation, EC2
  - Data storage & processing : RDS, EMR, Redshift

# Enter physical infrastructure challenges…

- Improving our **agility**
  - How can we deploy infrastructure as often as we deploy code (i.e. every day)?
  - How can we experiment quickly and at (almost) zero cost?

- Optimizing our **spend**
  - How can we avoid CAPEX peaks caused by hardware refresh?
  - How do we best adapt spend to traffic and business conditions?

- Implementing a **stronger Disaster Recovery** plan
  - …without building a 2nd DC which we don't need anyway

- Scaling storage & CPU for **data processing**
  - Do we really need big servers, lots of rack space and lots of power?
  - How do we efficiently handle unpredictable workloads?

# Why we decided to move everything to AWS

- We want to **focus on our real job** : building a great service
  - No more issues with hardware vendors, no more licensing hell. Ever!
  - Experimenting, (re)deploying & scaling are just a few clicks away
  - Great ecosystem of SaaS partners running in AWS

- We've been using it for a while now and **it just works**
  - AWS CloudFormation is infrastructure as code
  - « Time to server » is in minutes, not days

- We're **ready** and **eager** for it : Agile and DevOps are strong Viadeo values

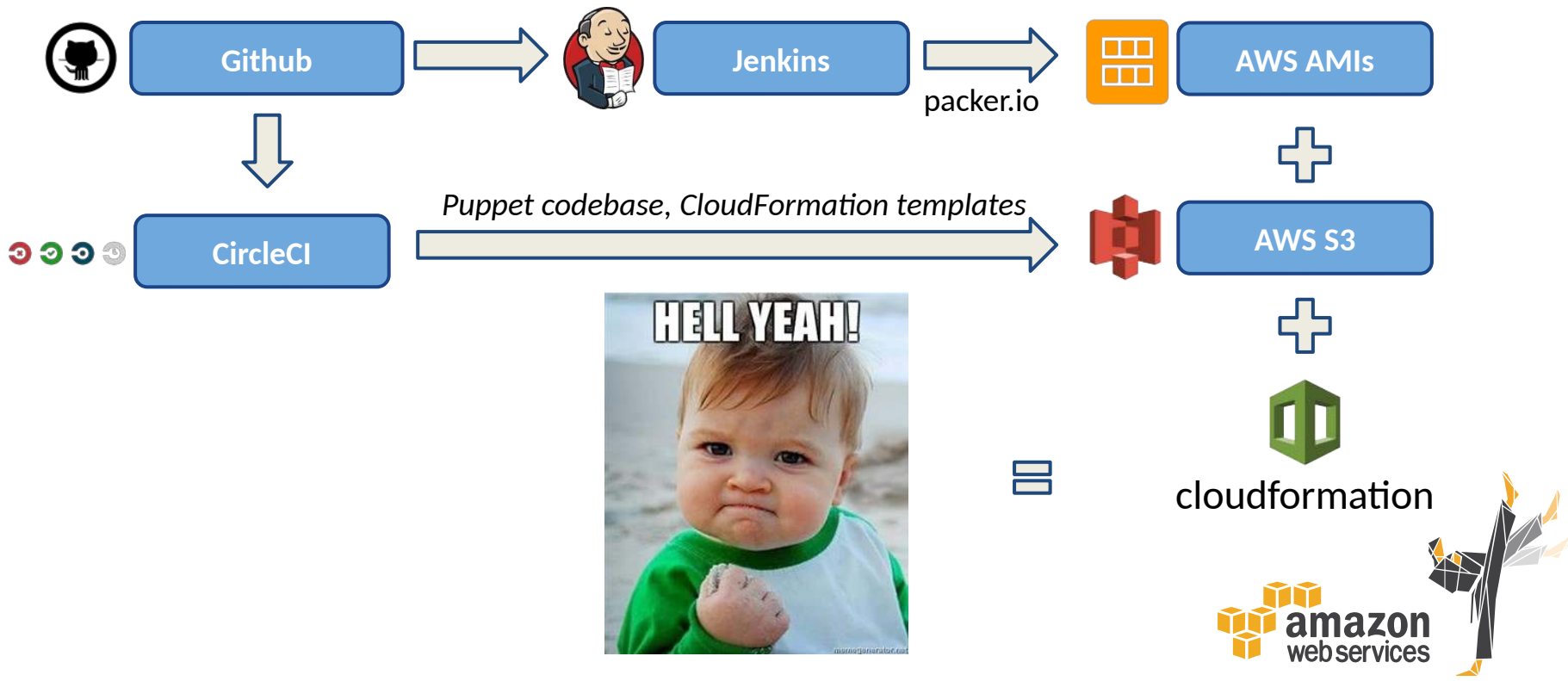- We want to **measure ROI** and **optimize cost**

# Our AWS manifesto

- Key objectives : **automation**, **scalability**, **safety**

- Continuous integration & delivery on all layers
  (infrastructure, instances, applications)

- Infrastructure : deep integration with AWS services for maximum leverage

- Applications : use our current stack for now, unless the benefit/cost ratio of
  adopting an equivalent AWS service is too good to pass (Redshift)

- Move everything, but no big bang : the move will be gradual
  → plan & build with parallel run in mind (DC + AWS)

amazon
web services

# Infrastructure as code : done!



Github → Jenkins → AWS AMIs

packer.io

Github → CircleCI

Puppet codebase, CloudFormation templates → AWS S3

AWS AMIs + AWS S3 + cloudformation = HELL YEAH!

amazon web services

# New challenges ;)

- Some performance & cost trade-off in the short term
- Technical debt (aka «  guess what's under the rug? »)
    - Layer 7 load balancing rules (*evil*), legacy filers, etc.
    - Servers running really old versions of whatever: keep or reinstall?
    - Any cruft ignored over the years. Time to clean up!


- Parallel run requires good connectivity →AWS DirectConnect
- CloudFront performance not great across regions → need a 2nd CDN
- Negotiating early termination of the legacy hosting contract !

# Conclusion



- 5 years ago : «Cloud computing? Why?»
- Now : «NO cloud computing? Why?»
- Some good reasons, but mostly really bad ones

- Cloud computing a fashion? I don't think so.
- Cloud = infrastructure in digital form (like photos, music, movies, money, etc.)

- Using AWS helps Viadeo change every day,

  by speeding up innovation and delivery

- See you there!