



Giornate sugli sbocchi professionali
del corso di laurea in
Biotecnologie Industriali
(BIOTIN)

Oristano, 23/24 aprile 2013



<http://biocloud.unica.it>

Gabriele Milia
milia.ga@unica.it



Cloud Computing

 **BIOCLOUD**
SEARCH ENGINE





The NIST definition of Cloud Computing

*Cloud computing is a model for enabling ubiquitous, convenient, **on-demand network access** to a shared pool of configurable **computing resources** (e.g., networks, servers, storage, applications, and services) that can be **rapidly provisioned** and released with **minimal management effort** or service provider interaction.*

September 2011





Essential Characteristics

1. On-demand self-service



2. Broad network access



3. Resource pooling (storage, processing, memory, network bandwidth).



4. Rapid elasticity



5. Measured service



Everything as a Service (XaaS)



"IaaS"

Infrastructure-as-a-Service

host



"PaaS"

Platform-as-a-Service

build



"SaaS"

Software-as-a-Service

consume

TARGET

Administrator

Developer

User



Google
App Engine



Google App Engine

- ✓ Piattaforma di hosting e sviluppo di applicazioni di Google
- ✓ Applicazioni web a traffico elevato sul cloud di Google
- ✓ Applicazioni costruite su GAE utilizzano la **stessa tecnologia che fornisce velocità e affidabilità ai siti web di Google**



Google
App Engine



python

Java



Vantaggi GAE

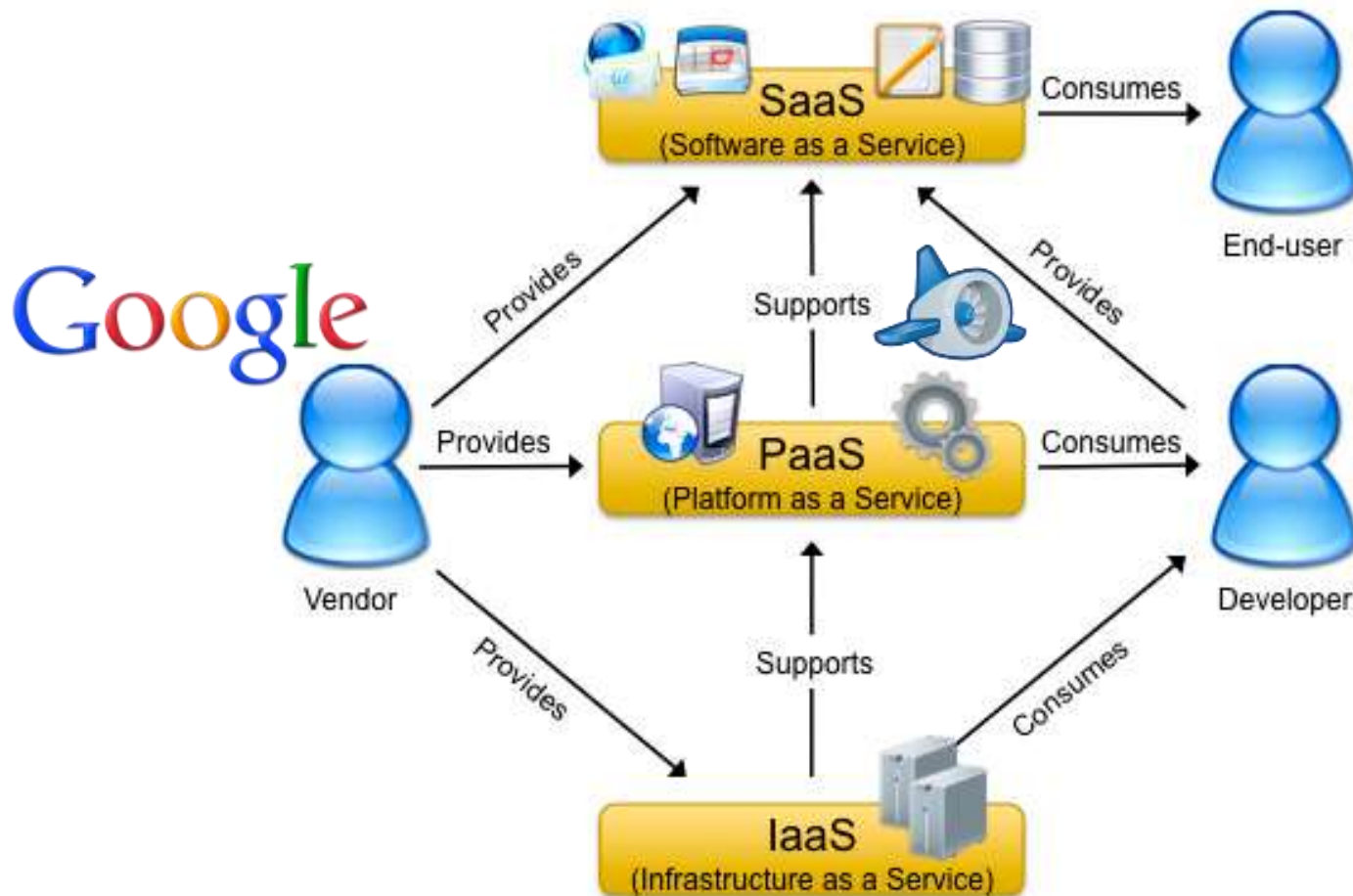
Easy to build, Easy to scale, Easy to maintain

- ✓ Affidabilità, Performance, Sicurezza
- ✓ Scalabilità automatica
- ✓ Focus sull'applicazione
- ✓ Approccio pay-as-you-go

- ✓ Free quota per app per day
- ✓ GAE pricing

<https://cloud.google.com/pricing/>





Google

BioCloud Team



Administration Console

Application: **biocloud-test [High Replication]**

[Community Support](#) [« My Applications](#)

Version: **All Versions**

Main

[Dashboard](#)

[Instances](#)

[Logs](#)

[Versions](#)

[Backends](#)

[Cron Jobs](#)

[Task Queues](#)

[Quota Details](#)

Data

[Datastore](#)

[Indexes](#)

[Datastore Viewer](#)

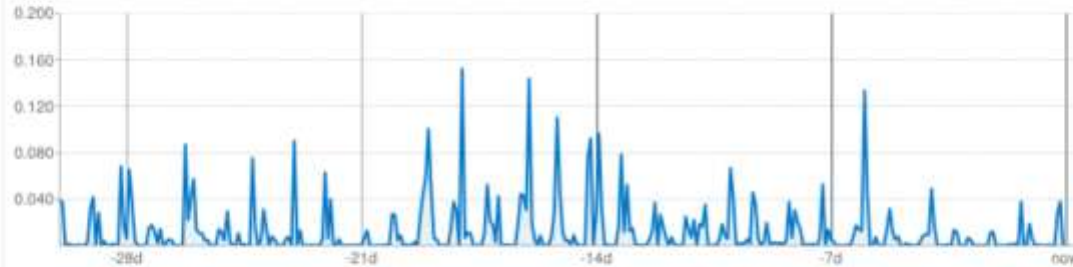
[Datastore](#)

[Statistics](#)

Charts

Requests/Second

30 mins 3 hrs 6 hrs 12 hrs 24 hrs 2 days 4 days 7 days 14 days 30 days



Instances

Number of Instances - [Details](#)

Average QPS

Average Latency

Average Memory

1 total

0.000

Unknown ms

58.8 MBytes

Billing Status: Free - [Settings](#)

Quotas reset every 24 hours. Next reset: 17 hrs



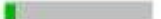









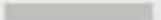




Quota details

- Data
 - [Datastore](#)
 - [Indexes](#)
 - [Datastore Viewer](#)
 - [Datastore Statistics](#)
 - [Blob Viewer](#)
 - [Prospective Search](#)
 - [Text Search](#)
 - [Datastore Admin](#)
 - [Memcache Viewer](#)
- Administration
 - [Application Settings](#)
 - [Permissions](#)
 - [Blacklist](#)
 - [Admin Logs](#)
- Billing
 - [Billing Settings](#)
 - [Billing History](#)

Instances ⓘ

Number of Instances - Details	Average QPS	Average Latency	Average Memory
1 total	0.000	Unknown ms	58.8 MBytes

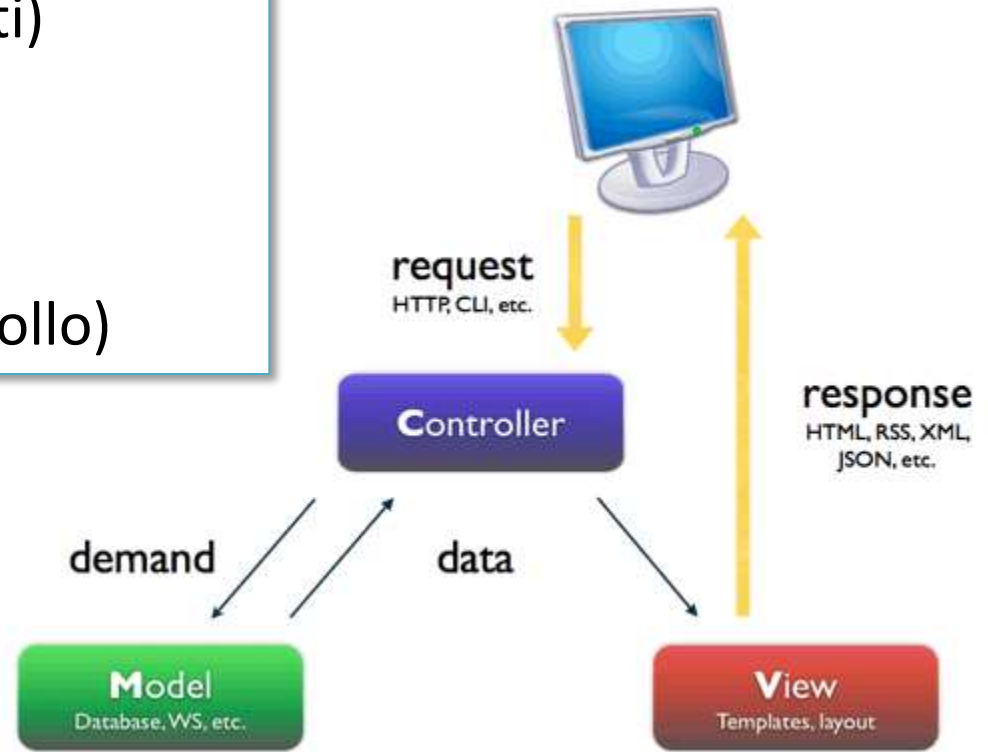
Billing Status: Free - [Settings](#) Quotas reset every 24 hours. Next reset: 17 hrs ⓘ

Resource	Usage
Frontend Instance Hours	 7% 2.09 of 28.00 Instance Hours
Backend Instance Hours	 0% 0.00 of 9.00 Instance Hours
Datastore Stored Data	 23% 0.23 of 1.00 GBytes
Logs Stored Data	 3% 0.03 of 1.00 GBytes
Task Queue Stored Task Bytes	 0% 0.00 of 0.49 GBytes
Blobstore Stored Data	 0% 0.00 of 5.00 GBytes
Code and Static File Storage	 1% 0.01 of 1.00 GBytes
Datastore Write Operations	 0% 0.00 of 0.05 Million Ops
Datastore Read Operations	 2% 0.00 of 0.05 Million Ops
Datastore Small Operations	 0% 0.00 of 0.05 Million Ops
Outgoing Bandwidth	 1% 0.01 of 1.00 GBytes
Recipients Emailed	 0% 0 of 100
Stanzas Sent	 0% 0 of 10,000
Channels Created	 0% 0 of 100
Logs Read Bandwidth	 0% 0.00 of 0.10 GBytes



Pattern MVC

- ✓ Model (organizzazione dati)
- ✓ View (presentazione dati)
- ✓ Controller (Logica di controllo)





Model

App Engine Datastore

- ✓ NoSQL schemaless object datastore
- ✓ Query engine
- ✓ Atomic transactions



High Replication Datastore (HRD)

- ✓ Replicazione dei dati su diversi datacenter
- ✓ Garantire
 - Elevata affidabilità
 - Elevata disponibilità



View

Responsive Web Design

- ✓ HTML + CSS
- ✓ JavaScript (jQuery + JQuery UI)
- ✓ Django web framework





Controller

- ✓ Python (linguaggio di programmazione ad alto livello)
- ✓ Librerie fornite da Google
- ✓ Biopython





Dipendenza dai servizi esterni

- ✓ NCBI Entrez Programming Utilities;
- ✓ UniChem RESTful Web Service;
- ✓ database identifier mapping service (Uniprot);
- ✓ STRING (<http://string-db.org/>);
- ✓ mygene.info REST web services (autocompletamento)



un1Chem





Servizi esterni

Vantaggi

- ✓ Dati aggiornati
- ✓ Web Application *“leggera”*

Svantaggi

- ✓ Dati limitati se i servizi esterni non funzionano.
- ✓ Policy (es. numero richieste/secondo NCBI)



Memcache

high-performance, distributed memory
object caching system



Conclusioni

- ✓ Struttura flessibile e modulare
- ✓ Downtime pari a zero

Future Work

- ✓ Migliorare aggiornamento dei dati memorizzati nel datastore
- ✓ Aggiungere nuove fonti dati





REGIONE AUTÒNOMA DE SARDIGNA
REGIONE AUTONOMA DELLA SARDEGNA

Progetto finanziato dalla Regione Autonoma della Sardegna

Legge regionale 7 agosto 2007, n. 7

“Promozione della ricerca scientifica e dell’innovazione tecnologica in Sardegna”

“DENIS: Dataspaces Enhancing the Next Internet in Sardinia”

Responsabile Scientifico

Prof.ssa Nicoletta Dessì

Università di Cagliari

