Meeting the needs of Science in the Cloud

Presented by CloudSigma freedom through technology

Strategic Goal

Helix Nebula, the Science Cloud

is a partnership that has been created

to support the **massive IT requirements**

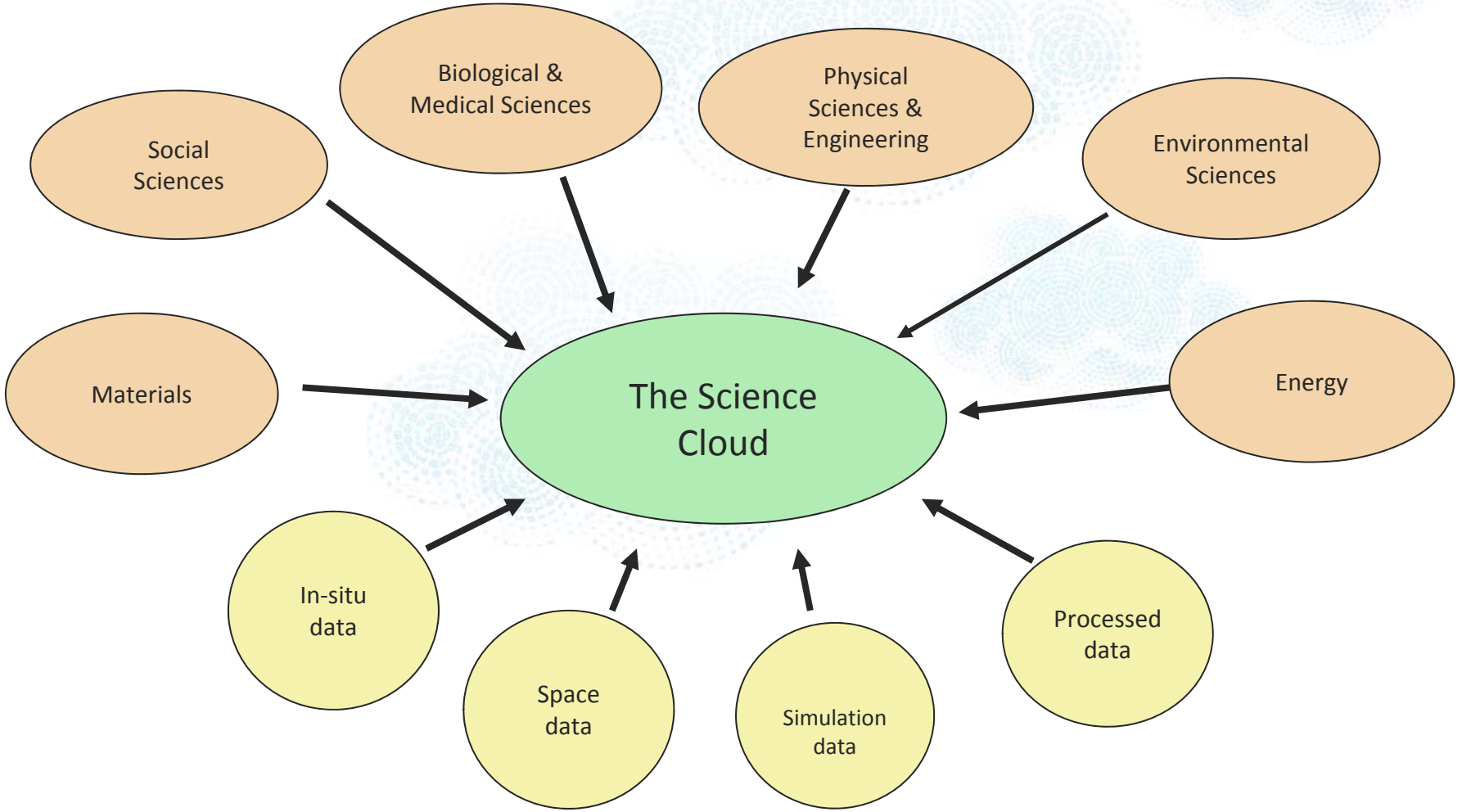
of European scientists

and **create a Cloud computing market**

for the public sector in **Europe**.

The Science Cloud : INPUT

The Science Cloud : a unique mine of scientific data



EC is strongly supporting Helix Nebula

Europe's Information Society
Thematic Portal

European Commission >

Policies | Activities | Culture & Society | Economy & Work | Education & Training | Quality of Life | IS Industry | Regions / World | Research & Innovation

Newsroom
Calendar
Funding Opportunities
Calls for Proposals
Calls for Tender
Press Packs
Press Releases
Library
Audiovisual
Speeches
Search Newsroom
Register & Subscribe
Login
Services for the Press

News ::

Helix Nebula makes Europe become Cloud Active in Science

(25/06/2012) The European Commission just took another step towards a Cloud Active Europe with the signature of a grant agreement for Helix Nebula. The Commission is giving €1.8 million in funding to this initiative which will develop cloud services in the scientific domain by partnering main research infrastructures in Europe (CERN, EMBL and ESA) with main stakeholders in the industry (e.g. Atos, T-systems, etc). The Commission's grant will support moving flagship applications (Atlas HEP, Genomic Assembly in the Cloud and Supersites Exploitation Platform) to the cloud, while identifying business opportunities for cloud services for the scientific community.

The signature of this agreement is one step in the implementation of the Digital Agenda for Europe, which calls for the further development of e-Infrastructures and the establishment of an EU strategy for cloud computing for government and science.

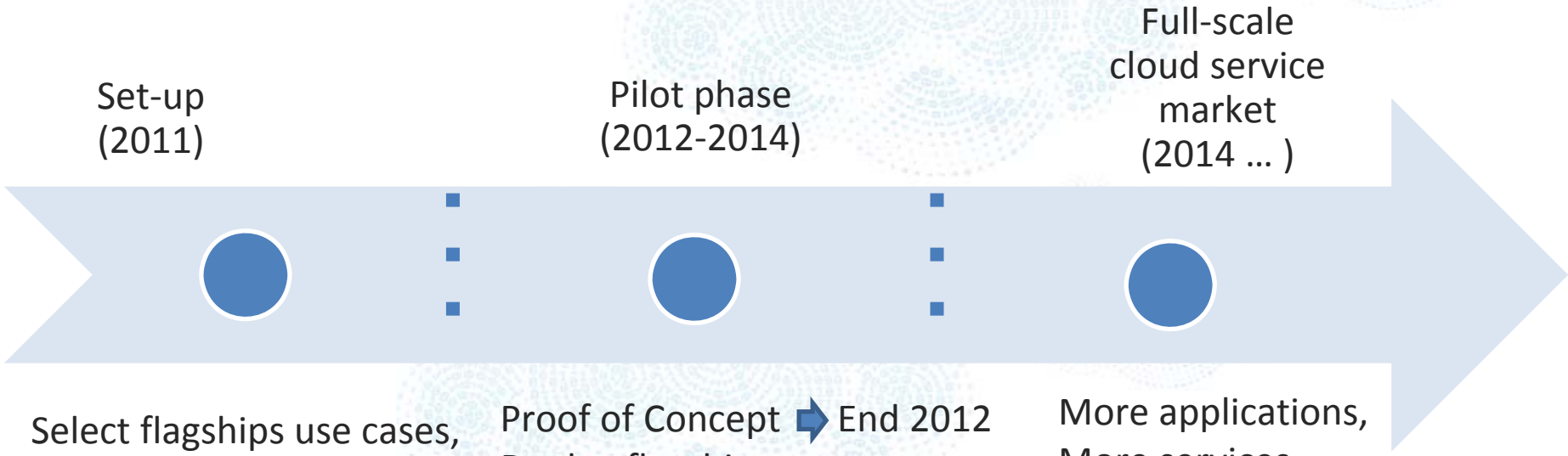
Helix Nebula as a cloud partnership for science launched under the e-Infrastructures activity of FP7, is a forerunner of the forthcoming Integrated Cloud Computing Strategy for the European Union expected to be adopted by mid-July and in particular is complementary to the European Cloud Partnership to promote public sector take-up of cloud services that was announced by Neelie Kroes in Davos.

More information

Digital Agenda for Europe

Digital Agenda Cloud Computing home page

Timeline



Select flagships use cases, identify service providers, define governance model

Proof of Concept ➡ End 2012
Deploy flagships,
Analysis of functionality, performance & financial model

More applications,
More services,
More users,
More service providers

Become a new member !

as:

Users

Service Providers

Adopters

Interested Parties

Become a new member

Over the next two years the Helix Nebula Consortium is expected to involve an increasing number of members.

The potential members are likely to be stemming from the categories below:

Users: An organisation that applies to become a user member should be a science or space organisation and commit to provide at least one flagship use case for cloud computing that can be verified and validated through a Proof of Concept with multiple service providers. Users will name a representative to participate in regular meetings of the Helix Nebula Users Board.

Service Providers: An organisation that applies to become a service provider member should commit to support a minimum set of cloud computing services and perform at least one Proof of Concept of a flagship use case with a user. Service Providers can apply to one or more of the following categories: Connectivity Provider, Infrastructure-as-a-Service Provider (IaaS) Platform-as-a-Service Provider (PaaS), Software-as-a-Service Provider (SaaS), Integrator, Consultant or Broker. Service Providers will name a representative to participate in regular meetings of the Helix Nebula Service Providers Board.

Adopters: An organisation may apply to become an adopter, if they initially do not want to be directly involved in the flagship use cases but wish to make use of the Helix Nebula's products and services on a pay-per-use basis and be able to provide feedback. Adopters may be invited to Boards (either Service Providers or Users) meetings for information purposes. Adopters will name a representative to be informed or invited by the Helix Nebula Consortium on a regular basis.

Interested Parties: An organisation may apply to become an Interested Party, if they initially do not want to be directly involved in the flagship use cases but wish to be kept informed of the work of or use information provided by the Consortium.

For all types of members a formal acceptance procedure will apply (specified hereafter), to be implemented by the Helix Nebula Management Team following the receipt of the membership request. **Membership application implies acceptance of the vision of Helix Nebula** as outlined in the strategic plan Strategic Plan for a Scientific Cloud Computing infrastructure for Europe, CERN-OPEN-2011-036, August 2011, **and willingness to collaborate with the other partners** in order to achieve the vision. All decisions must be adopted by consensus of both the user organisations and service provider companies. In case of lack of consensus, decision will be taken by a qualified majority of all members of the two Boards, which must include the positive vote of the public user organisations.

All users and service providers applying to become a member of the Helix Nebula Consortium (except as an Interested Party) agree to sign a multi-lateral NDA prior to becoming an active member.

For more details and updates about how to join, write to us at contact@helix-nebula.eu

Events

Helix Nebula @ DCI Workshop, 18 September 2012, Prague

Bob Jones (CERN) and Michael Symonds (Atos) will participate to the DCI Workshop *"Distributed Computing Infrastructures for e-Science: Future Perspectives"*.

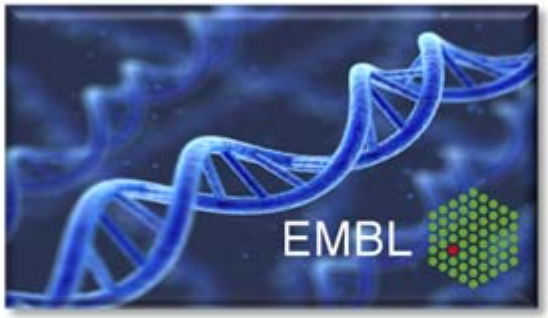
More

Participants

Helix Nebula Pilot Phase

Flagship use cases

Initial Flagship Use Cases

ATLAS High Energy Physics Cloud Use	Genomic Assembly in the Cloud	SuperSites Exploitation Platform
		
To support the computing capacity needs for the ATLAS experiment	A new service to simplify large scale genome analysis; for a deeper insight into evolution and biodiversity	To create an Earth Observation platform, focusing on earthquake and volcano research

- Template agreed by demand and supply side
- Eligibility review and analysis with cloud service suppliers

Flagship deployments
First results - 1

- Proof of Concept stage within the Pilot Phase started January 2012
- Each flagship has been deployed with a series of providers independently

CERN, EMBL and ESA succeeded in deploying scientific applications each involving tens of thousands of jobs running at data centres operated by Atos, CloudSigma and T-Systems

HELIX NEBULA
THESCIENCECLOUD

Flagship deployments
First results - 2

- **CERN** was able to run simulations previously executed on the Worldwide LHC Computing Grid by quickly deploying ATLAS experiment flagship application on the Cloud.
- **EMBL** successfully deployed and tested their novel software pipeline for large-scale genomic analysis using real world large genomic data sets.
- **ESA** successfully tested large-scale data processing and dissemination for its radar satellites using different cloud provider infrastructure.

Flagship deployments
First results - 3

- The PoC extensively evaluated **scalability**, **performance** and **on-demand provisioning of resources** for high performance computing and **fast data storage** in the cloud computing resources provided by **Atos**, **CloudSigma** and **T-Systems**

- In addition to the infrastructure providers, SME's such as **SixSq**, **Terradue** and **The Server Labs** were vital to get the flagship applications up and running.

HELIX NEBULA
THESCIENCECLOUD

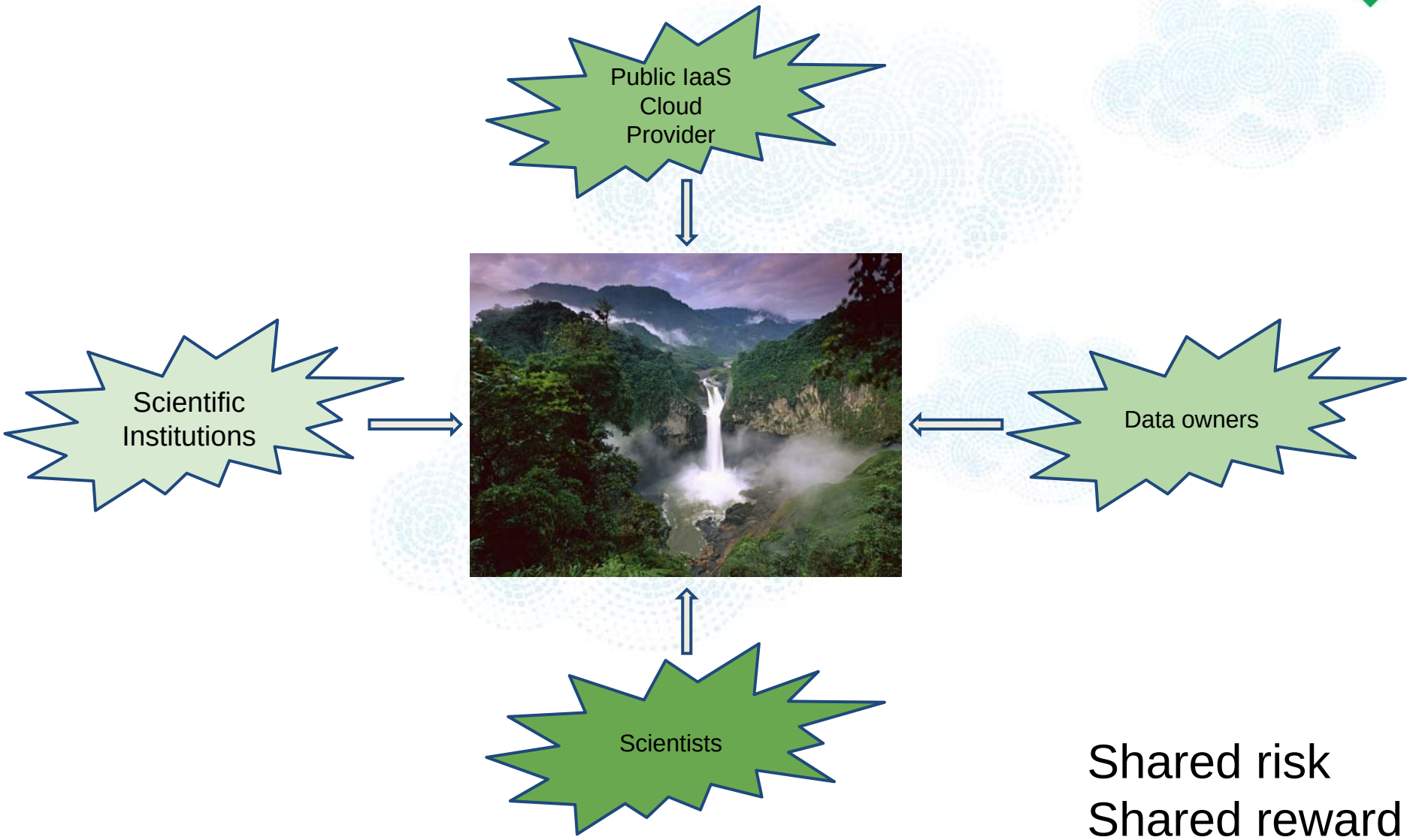
What's up now ?

The Helix Nebula consortium is now focussing on identifying a **common set of interfaces** for suppliers and users before the next wave of deployments, building on the lessons learned from the PoC.

The Ecosystem Model

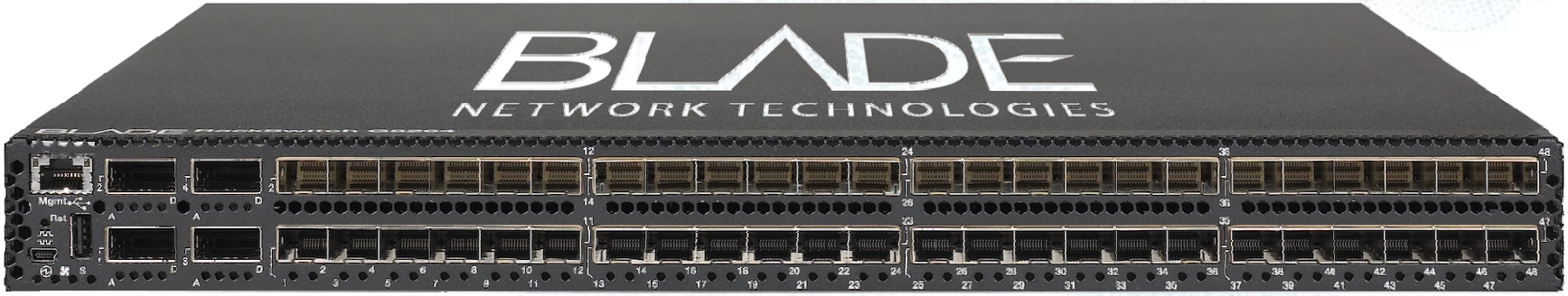
An **ecosystem** is a community of living organisms (plants, animals and microbes) in conjunction with the nonliving components of their environment (things like air, water and mineral soil), interacting as a system.

Who feeds the eco-system?



Shared risk
Shared reward

Advantage #1: Reduced Data Churn



The cloud becomes a hub, a virtual switch between ecosystem participants.

Advantage #2: Elasticity

Infrastructure should be organic, shaping itself to customer requirements over time.

Requirements define the infrastructure; services should not be limited by inflexible infrastructure.

Advantage #3: Increased Service Choice

HELIX NEBULA THESCIENCECLOUD

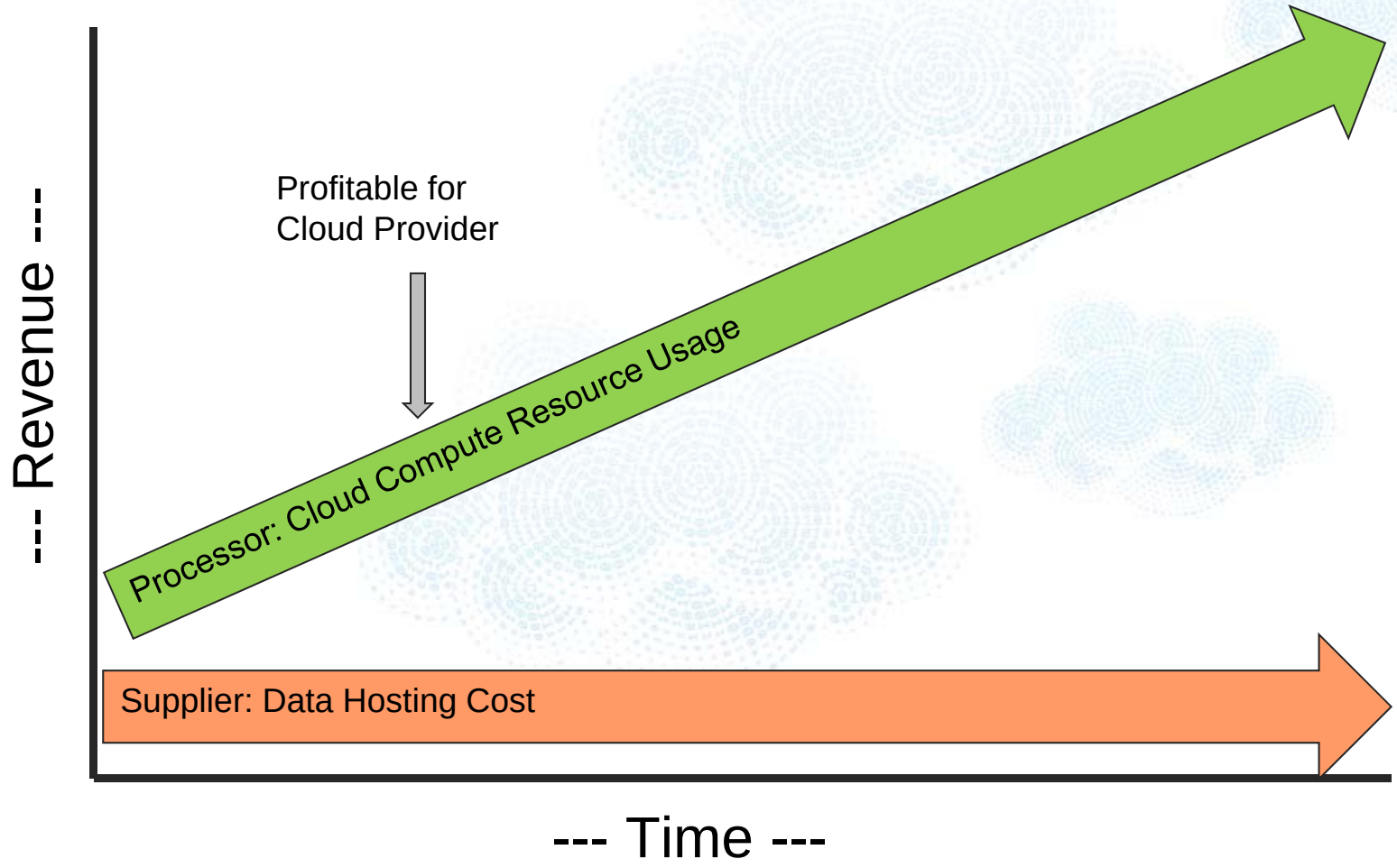
Cost Model

- In Public Clouds
  - Storage is the lowest cost and least profitable resource
  - CPU and RAM (Compute) generate higher margins
  - Burst compute has a low TCO compared to own/co-lo
- Science Users are cost constrained
  - Server and Data Center costs are high
  - Infrastructure (HVAC, power, etc.) costs continue to rise
  - There is less and less funding available
  - Much science work is “burst” in nature
  - There is a cost to data replication, who pays ?

Cost Model

- Cloud Ecosystem Solution
  - Host the databases for close to zero cost
  - Grant the Owner full access and control of the Data / IP
  - Provide an easy to use compute platform in the cloud
    - Pre-loaded science software, cluster management, Virtual GPU's, S3 storage, SSD disks, simple & transparent billing model
  - Work with the Data Owner to attract Compute users
    - Pre-paid credits, discounts, capped duration projects, PhD's
  - Add more databases to the Ecosystem
    - EO data + WHO data = Better mosquito control planning
    - Ocean data, climate, genes, drug interactions, etc., etc., etc.

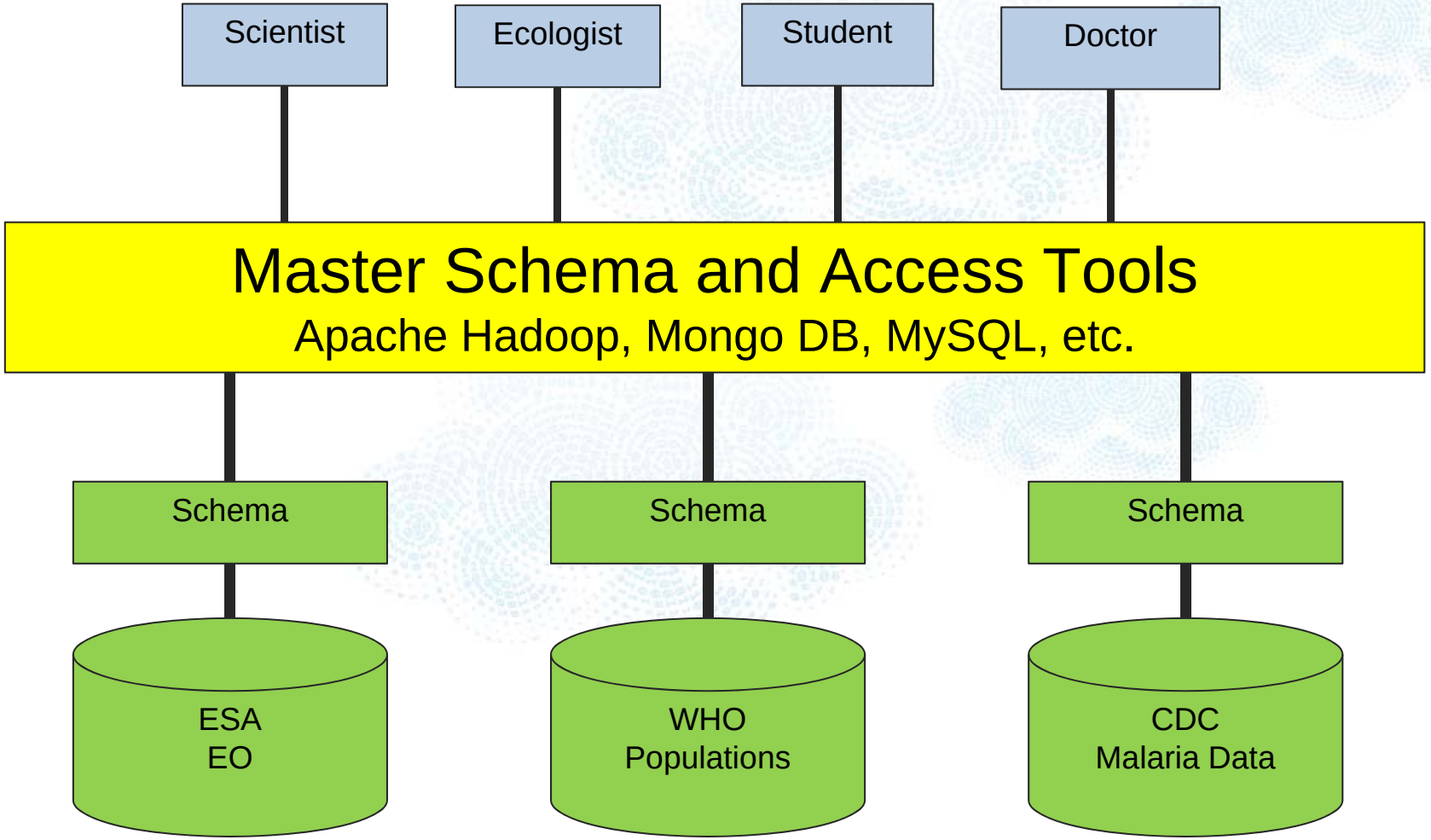
Cost Model


--- Revenue ---
Profitable for
Cloud Provider
Processor: Cloud Compute Resource Usage
Supplier: Data Hosting Cost
--- Time ---

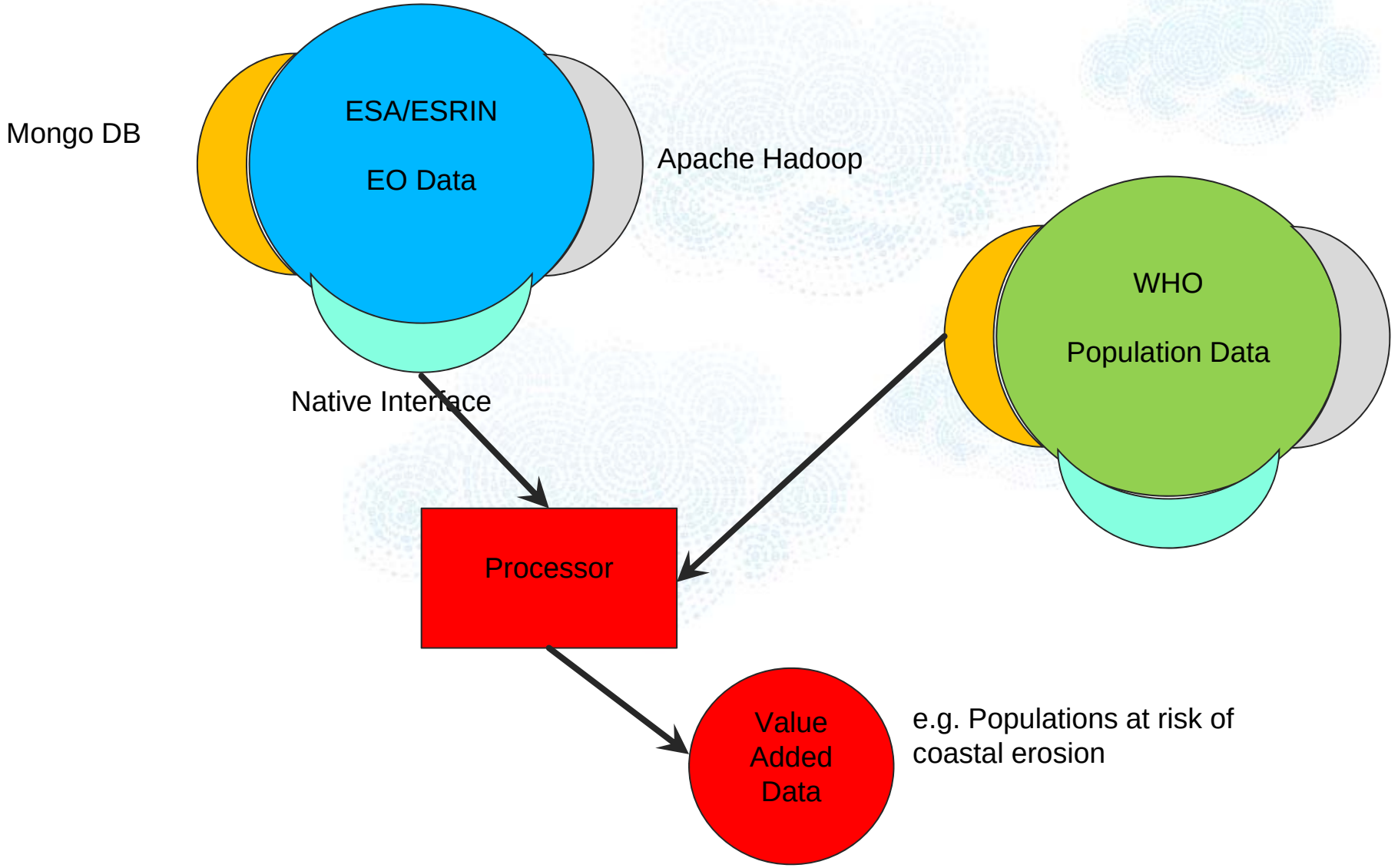
Meta-data / Catalog

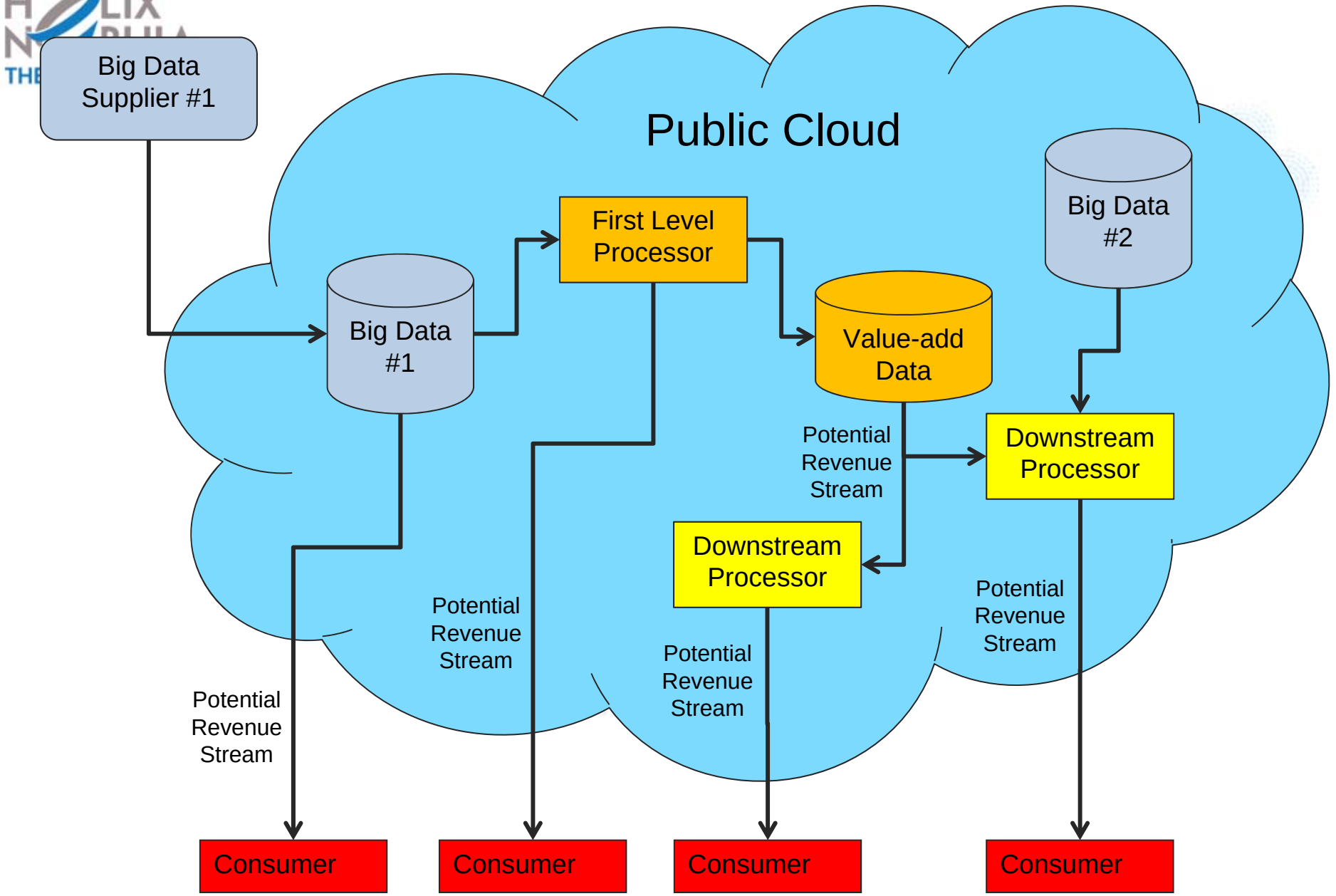
- Processors will require a Master Catalog of Data
  - What data is available, costs, latency, currency, etc.
  - Master Schemas – industry standards ?
  - Understanding granularity, range, scope, etc.
- Possible Meta-data Schemas to Expose
  - Native / Legacy – what you have today
  - Apache Hadoop
  - Mongo DB
  - SQL (Oracle, MySQL, etc.)
  - Others

Data Model


Scientist
Ecologist
Student
Doctor
Master Schema and Access Tools
Apache Hadoop, Mongo DB, MySQL, etc.
Schema
Schema
Schema
ESA
EO
WHO
Populations
CDC
Malaria Data

Ecosystem Meta-data Processing Model





The Big Picture

Contacts

Helix Nebula

Email:contact@helix-nebula.eu
Twitter: HelixNebulaSC
Website: http://www.helix-nebula.eu/

CloudSigma

Email: robert@cloudsigma.com
Twitter: @cloudsigma
Website: http://www.cloudsigma.com