infochimps platform

Your complete, turn-key Big Data stack

What is the Infochimps Platform?

The Infochimps Platform provides the Big Data infrastructure and expertise necessary to put data at your fingertips. In order to provide our market-leading data catalog of thousands of data sets and APIs, our team of data scientists and engineers had to create a uniquely powerful data system -- able to ingest and deliver massive amounts of data that could run on-demand, resource-intensive analytics.

Our unique technology suite and best practices are now available to you, unlocking the potential of the Big Data you already have and the even bigger data available online.

Infochimps provides the maximum value in fully-featured hosting, speed of implementation, responsiveness, expertise, and peace of mind. Your system will be fully integrated into your technology environment, yet able to tap into the infinite scalability and elasticity of the Cloud. Your time can be spent training your staff how to use Big Data, not focused on building and managing the systems to access it.

The Big Picture

The Big Data Problem
With a plethora of data sources and companies, the problem isn't data availability. The many types of data and the speed at which data changes, along with sheer volume, are today's most daunting challenges as businesses struggle to make sense of it all. Volume, velocity, and variety -- they're the hallmarks of the Big Data era we live in now.

Variety comes in the form of web logs, unstructured textual information from social networks, CRM, Point-of-Sale and a myriad of other data types. Velocity breeds velocity. Fast-changing data drives demand for deep analytic insights delivered in hours, minutes, or in extreme cases, seconds, instead of weekly or monthly reports that once sufficed.

The Power of _Ironfan_
The Infochimps Platform is your Cloud-ready Big Data infrastructure. This means having unprecedented scalability and elasticity when needed -- using the resources required when you need them. That's why we created _Ironfan_, an open-source offering authored by Infochimps, an elegant tool for rapidly declaring and utilizing resources that serves as the configuration-provisioning-deployment foundation for your stack. Also, the flexibility enabled by _Ironfan_ means your architecture can change as your business evolves.

Staying Focused on Your Business
Getting visibility of your data doesn't require you hire a full time staff or buy tons of new infrastructure. Your organization needs to focus on the analysis and data science that makes your business tick! Infochimps gives you focus and peace of mind by ensuring your Big Data system is running efficiently and reliably with on-demand scaling and minimal excess capacity.

at a glance

Get Data At Scale

We'll help you import any number of data sources. Internal or external. We'll make sure you reliably scale to the volume, velocity, and variety of data you need.

Augment & Transform Data

We give you the full toolset to run the analytics and processes you need. You can even augment and process data as it flows in.

Delivery & Managed Hosting

We make sure data is at your fingertips, whether that's an Infochimps-managed Cloud database, enterprise API, your existing stack, or bulk files.

infochimps platform

in focus

At Infochimps, we created our own special social influence score called **Trstrank**, taking advantage of our vast historical record of tweets and Twitter users. However, generating this regularly calculated user ranking was a tremendous task for a young startup.
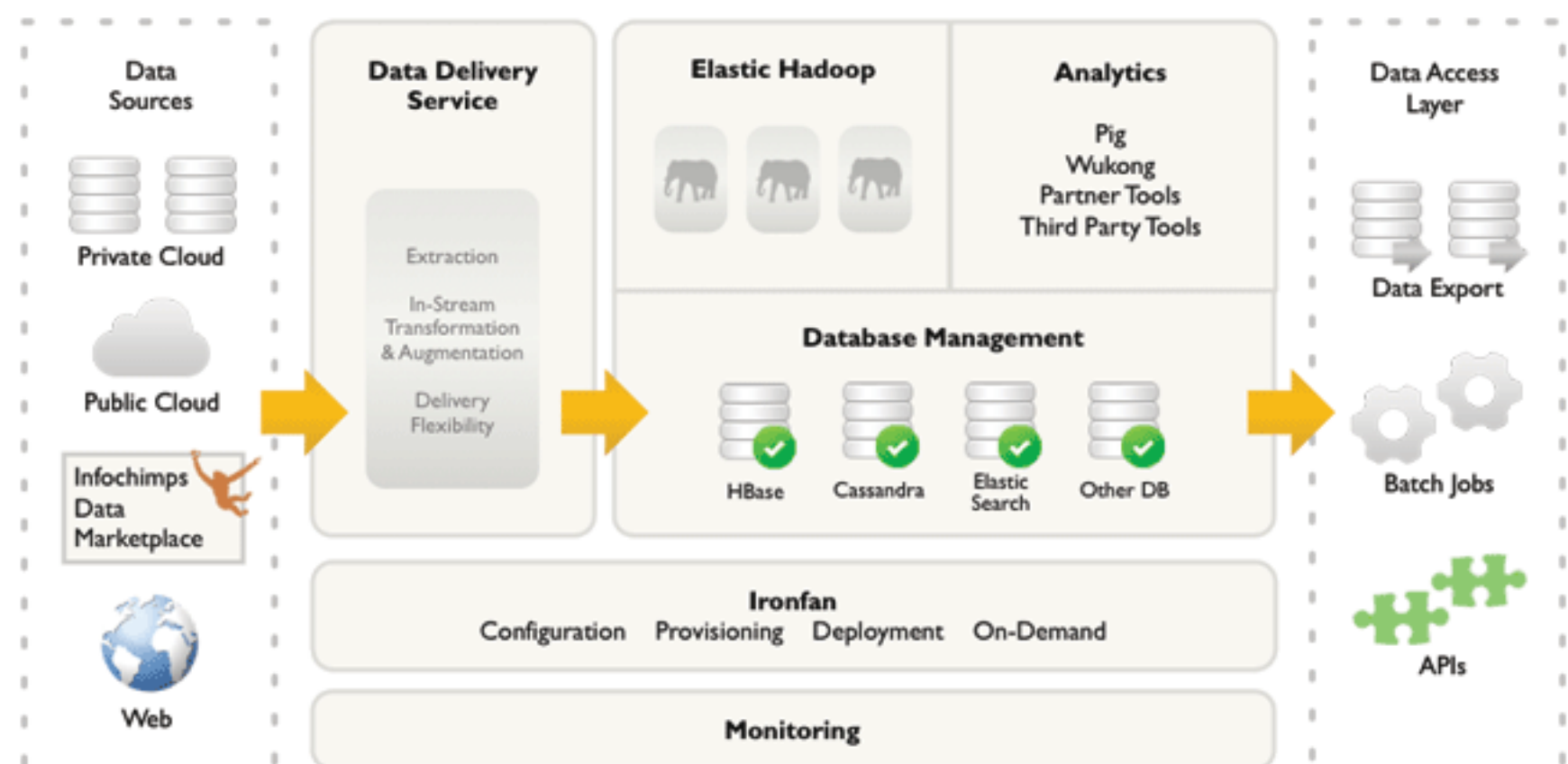
Amazon Web Services' (AWS) Cloud provided the flexibility and scalability to use and pay for only what we needed. However, projects like Trstrank require intensive ramp up time on server setup and machine configuration. Manual provisioning could take 2 weeks. With *Chef*, it would take up to 4 days to provision nodes individually.

To make spinning machines up and down even faster, Infochimps authored open-source *Ironfan*. Using *Ironfan*, it only takes 2 hours and 35 lines of code to provision and spin up the 30 nodes that run Trstrank. Coupled with our *Infochimps Data Delivery Service* ingesting 20 million tweets per day, Trstrank isn't just feasible. It's easy.

Infochimps Platform Solution Architecture

Data Delivery Service

Ensure sure your data gets from A to B with scalable data transport. In addition to standard bulk data import, we utilize technologies like Apache Flume to extract, split, transform, and recombine multiple data streams at terabyte scale in real-time. Combined with our best-in-class proprietary software, Data Delivery Service makes it easy to extend the power of Flume, while seamlessly integrating into your existing data environment with flexible delivery options.

Database Management

Take advantage of the right database technology for the job, whether it's HBase, Cassandra, Elastic Search, or others. We ensure your data storage environment can handle the scalability demands of your business. Whether we provide an on-premise or hosted solution, you will receive reliable monitoring, administration, and maintenance of your data storage.

Elastic Hadoop

Get Hadoop resources when you need them, whether scheduled, ad-hoc, or dedicated. This breadth in capability enables nightly batch jobs, compliance or testing clusters, science systems, and

production systems. With Ironfan, Infochimps' automated systems provisioning tool, as its foundation, Elastic Hadoop lets you tune the resources specially for the job at hand. Now it's easy to create map or reduce specialist machines, high compute machines, high memory machines, etc. as needed.

Analytics

Transform, analyze, and visualize your data with the best tools available. Use Pig, Infochimps' Wukong, and other platforms with Hadoop to execute the data calculations and algorithms your business requires. Take advantage of our partner offerings like Datameer and Karmasphere or other third-party systems to extract additional value from your data.

contact us

Tell us about your Big Data challenges.

PHONE
1-855-DATA-FUN
(1-855-328-2386)

EMAIL
sales@infochimps.com