

Onyx

విజయ్ కీర్త్న (@vijaykiran)



The Problem

Convert data into information



**Capture visitor data for a website
Store into a structured data store**



Simple Program running on single computer



Capture visitor data for a **high-traffic** website
Store into a structured data store



Vertical Scaling

More Memory

More Processors

More Disks



Capture visitor data for a **high-traffic** website
Process and store into a structured data store



Vertical Scaling Limit

Can't add more Memory

Can't add more Processors

Can't add more Disks



Horizontal Scaling

More Nodes!



*Hadoop Map Reduce, Storm, Spark, Flink, Sqoop,
Cascading, Cascatalog*



What's wrong with these ?

Hadoop Map Reduce, Storm, Spark, Flink, Sqoop,
Cascading, Cascalog*

They are not written in clojure



**Complex Information Model
Non-Standard Code
Differences between Batch and Streaming**



Enter Onyx!



A masterless, cloud scale, fault tolerant, high performance, distributed computation system



You build directed acyclic graph of computational paths in the program



Workflow

:in

|

:increment

`[[:in :increment] [:increment :out]]`

|

:out



Workflow

:in

|

:increment

`[[:in :increment] [:increment :out]]`

|

:out



Workflow

```
;;;  
;;      input  
;;      ^  
;; processing-1 processing-2  
;; |       |  
;; output-1   output-2
```

```
[[:input :processing-1]  
[:input :processing-2]  
[:processing-1 :output-1]  
[:processing-2 :output-2]]
```



Catalog



```
[{:onyx/name :in
  :onyx/plugin :onyx.plugin.core-async/input
  :onyx/type :input
  :onyx/medium :core.async}
{:onyx/name :inc
  :onyx/fn :onyx.peer.min-peers-test/my-inc
  :onyx/type :function}
{:onyx/name :out
  :onyx/plugin :onyx.plugin.core-async/output
  :onyx/type :output
  :onyx/medium :core.async}]
```

Segment



Clojure Map - Data flowing through the
Data flow graph

Task



Unit of work - input, processing, output

Flow Conditions



Predicates to control data flow direction

```
[{:flow/from :input-stream
  :flow/to [:inc]
  :flow/predicate :my.ns/even-or-zero?
  :flow/doc "Emits segment if the number even or zero."}]
```

Peer



A node or physical machine in the
cluster

A lots of other good stuff in Documentation!





Podcast Episode

<https://defn.audio/2016/09/06/episode-9-onyx-with-mike-and-lucas/>

Workshop!!!



<https://github.com/onyx-platform/learn-onyx>



Learn More about Onyx

- <http://www.onyxplatform.org>
- <https://gitter.im/onyx-platform/onyx>

