TITANA HIGHLY SCALABLE, DISTRIBUTED GRAPH DATABASE

Amazon EC2 m1.large machines
- Titan/Cassandra cluster
  - 10 million transactions per day
  - (99.99% success rate)

1 of the 5 operations is randomly chosen using a biased coin toss.

TECHNIQUES FOR SCALE

**Graph Management**
- Data Management
- Edge Compression
- Vertex-Centric Indices

**OLAP**
-OLAP
- Map/Reduce
- OLAP

**OLAP STORY**
-TITAN
- FAUNUS
- FULGORA

A SOCIAL STRESS TEST

A massive-scale single-relational graph

80% of SQL queries support

A Massively Scalable, Distributed Graph Database

**SCHEMA**

**DATASTRUCTURE**

**Indexes**

**Traversal**

**Frames**

**Media?**, "WWW2010.

"What is Twitter 2009."

Kwak, H., Lee, C., Park, H., Moon, S., "What is Twitter, a Social Network or a News Medium?" WWW2010.

**Compression**

**Management**

**Architecture Cost**

6 Amazon EC2 m1.large machines

Titan/Cassandra cluster

$32.36 per hour x 6 machines x 24 hours

$46.08 per day

14 Amazon EC2 m1.small machines

Read/write servers

$0.06 per hour x 14 machines x 24 hours

$26.68 per day

Cost for running this architecture: $72.96 per day

100 million transactions per day

(99.99% success rate)

$72.96 / 110 = $0.66 per million transactions

A Graph API

A Graph Traversal Language

TINKERPOP INTEGRATION

TITAN's bim/gremlin

\( g \) = TitanFactory.open(...)

... Titan/Cassandra

\( g \).out('follows').name

===> spurius

===> rome

===> gremlin

\( g \).v(1).out('follows').name

===> horatius

===> lars

===> horatius

Astrid's siege on Rome will end soon.

"Astrid: unlikely as your soldiers are weak and few.

Spurius: we will assemble our remaining men on the bridge.

Lars: your siege on Rome will end now.

Rome history will look kindly on the great roman empire.

Horatius: together we will defeat the etruscan invaders, co @rome

Horatius: RT @rome history will look kindly on the great roman empire.

A SOCIAL STRESS TEST

The follows graph was generated from Twitter 2009.

41,700,000 user vertices

1,470,000,000 follows edges

123456789

+5

7 bytes

24 bytes

~10,500 concurrent users

~130ms

~140ms

~270ms

~500ms

TITAN

Read/write servers

~26.68 per day

Cost for running this architecture: $72.96 per day

100 million transactions per day

(99.99% success rate)

$72.96 / 110 = $0.66 per million transactions

ARCHITECTURE COST

6 Amazon EC2 m1.large machines

Titan/Cassandra cluster

$32.36 per hour x 6 machines x 24 hours

$46.08 per day

14 Amazon EC2 m1.small machines

Read/write servers

$0.06 per hour x 14 machines x 24 hours

$26.68 per day

Cost for running this architecture: $72.96 per day

100 million transactions per day

(99.99% success rate)

$72.96 / 110 = $0.66 per million transactions