5 Steps to an Awesome Apache Cassandra™ Data Model

Patrick McFadin
VP Developer Relations, DataStax
@PatrickMcFadin

Relational Data Models

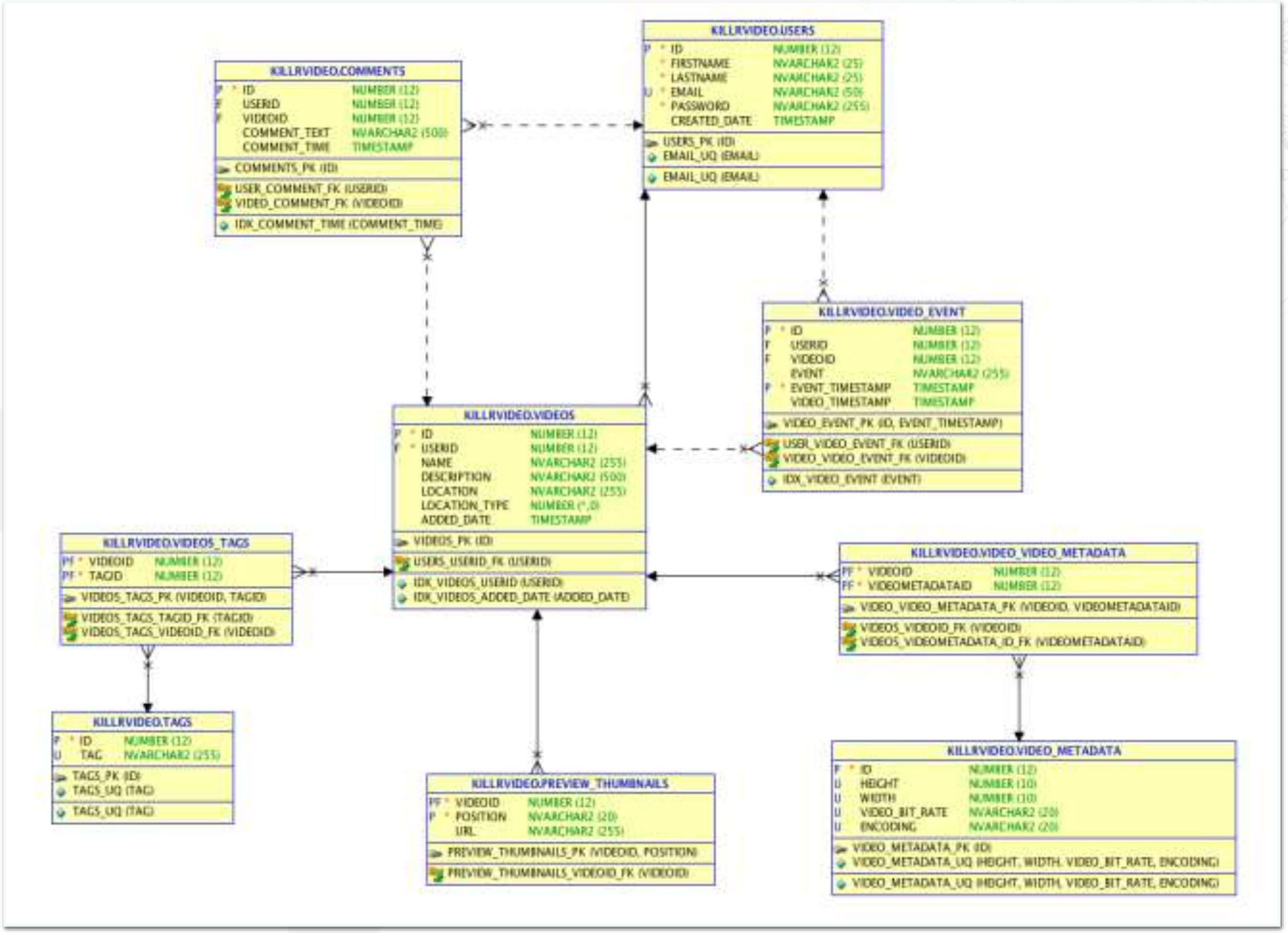
- 5 normal forms
- Foreign Keys
- Joins

Employees

deptId	First	Last
1	Edgar	Codd
2	Raymond	Boyce

Department

id	Dept
1	Engineering
2	Math


KILLRVIDEO.COMMENTS
P * ID NUMBER (12)
F USERID NUMBER (12)
F VIDEOID NUMBER (12)
COMMENT_TEXT NVARCHAR2 (500)
COMMENT_TIME TIMESTAMP
COMMENTS_PK (ID)
USER_COMMENT_FK (USERID)
VIDEO_COMMENT_FK (VIDEOID)
IDX_COMMENT_TIME (COMMENT_TIME)
KILLRVIDEO.USERS
P * ID NUMBER (12)
* FIRSTNAME NVARCHAR2 (25)
* LASTNAME NVARCHAR2 (25)
U * EMAIL NVARCHAR2 (50)
* PASSWORD NVARCHAR2 (255)
CREATED_DATE TIMESTAMP
USERS_PK (ID)
EMAIL_UQ (EMAIL)
EMAIL_UQ (EMAIL)
KILLRVIDEO.VIDEO_EVENT
P * ID NUMBER (12)
F USERID NUMBER (12)
F VIDEOID NUMBER (12)
EVENT NVARCHAR2 (255)
P * EVENT_TIMESTAMP TIMESTAMP
VIDEO_TIMESTAMP TIMESTAMP
VIDEO_EVENT_PK (ID, EVENT_TIMESTAMP)
USER_VIDEO_EVENT_FK (USERID)
VIDEO_VIDEO_EVENT_FK (VIDEOID)
IDX_VIDEO_EVENT (EVENT)
KILLRVIDEO.VIDEOS
P * ID NUMBER (12)
F * USERID NUMBER (12)
NAME NVARCHAR2 (255)
DESCRIPTION NVARCHAR2 (500)
LOCATION NVARCHAR2 (255)
LOCATION_TYPE NUMBER (*,0)
ADDED_DATE TIMESTAMP
VIDEOS_PK (ID)
USERS_USERID_FK (USERID)
IDX_VIDEOS_USERID (USERID)
IDX_VIDEOS_ADDED_DATE (ADDED_DATE)
KILLRVIDEO.VIDEOS_TAGS
PF * VIDEOID NUMBER (12)
PF * TAGID NUMBER (12)
VIDEOS_TAGS_PK (VIDEOID, TAGID)
VIDEOS_TAGS_TAGID_FK (TAGID)
VIDEOS_TAGS_VIDEOID_FK (VIDEOID)
KILLRVIDEO.VIDEO_VIDEO_METADATA
PF * VIDEOID NUMBER (12)
PF * VIDEOMETADATAID NUMBER (12)
VIDEO_VIDEO_METADATA_PK (VIDEOID, VIDEOMETADATAID)
VIDEOS_VIDEOID_FK (VIDEOID)
VIDEOS_VIDEOMETADATA_ID_FK (VIDEOMETADATAID)
KILLRVIDEO.TAGS
P * ID NUMBER (12)
U TAG NVARCHAR2 (255)
TAGS_PK (ID)
TAGS_UQ (TAG)
TAGS_UQ (TAG)
KILLRVIDEO.PREVIEW_THUMBNAILS
PF * VIDEOID NUMBER (12)
P * POSITION NVARCHAR2 (20)
URL NVARCHAR2 (255)
PREVIEW_THUMBNAILS_PK (VIDEOID, POSITION)
PREVIEW_THUMBNAILS_VIDEOID_FK (VIDEOID)
KILLRVIDEO.VIDEO_METADATA
P * ID NUMBER (12)
U HEIGHT NUMBER (10)
U WIDTH NUMBER (10)
U VIDEO_BIT_RATE NVARCHAR2 (20)
U ENCODING NVARCHAR2 (20)
VIDEO_METADATA_PK (ID)
VIDEO_METADATA_UQ (HEIGHT, WIDTH, VIDEO_BIT_RATE, ENCODING)
VIDEO_METADATA_UQ (HEIGHT, WIDTH, VIDEO_BIT_RATE, ENCODING)

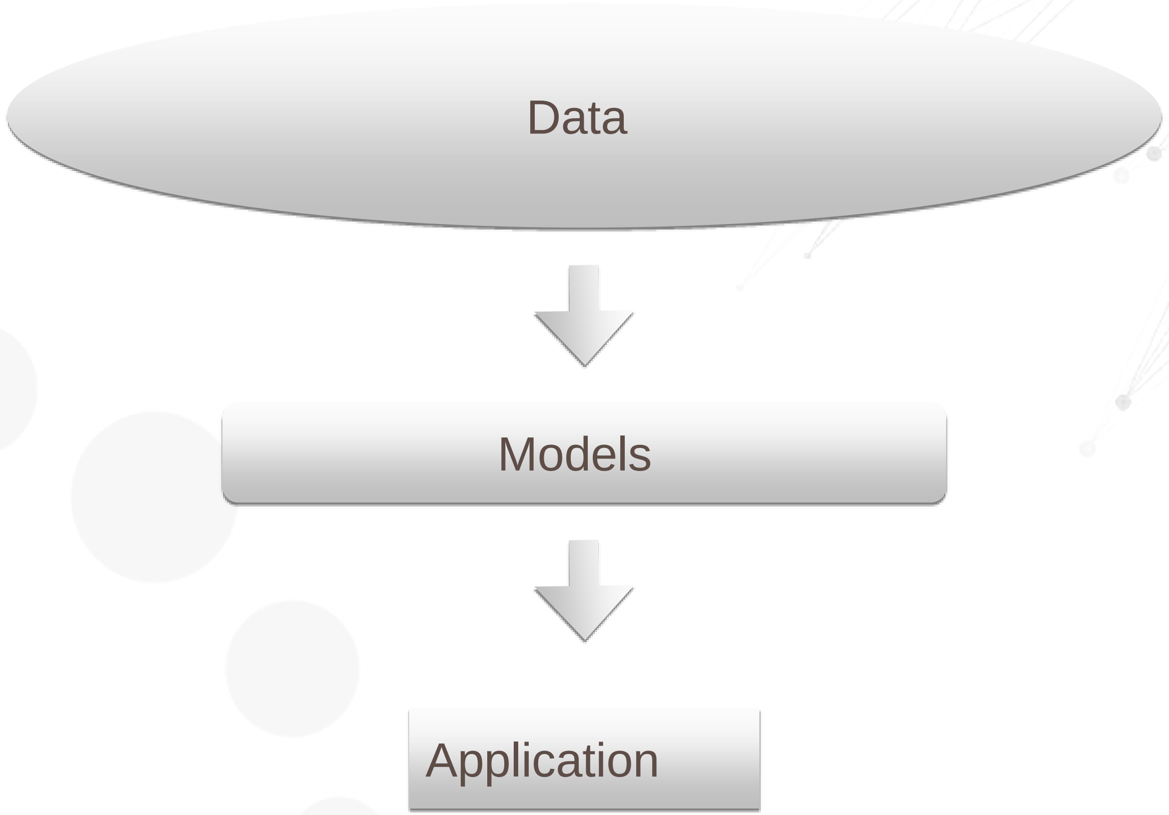
Relational Modeling

- Create entity table
- Add constraints
- Index fields
- Foreign Key relationships

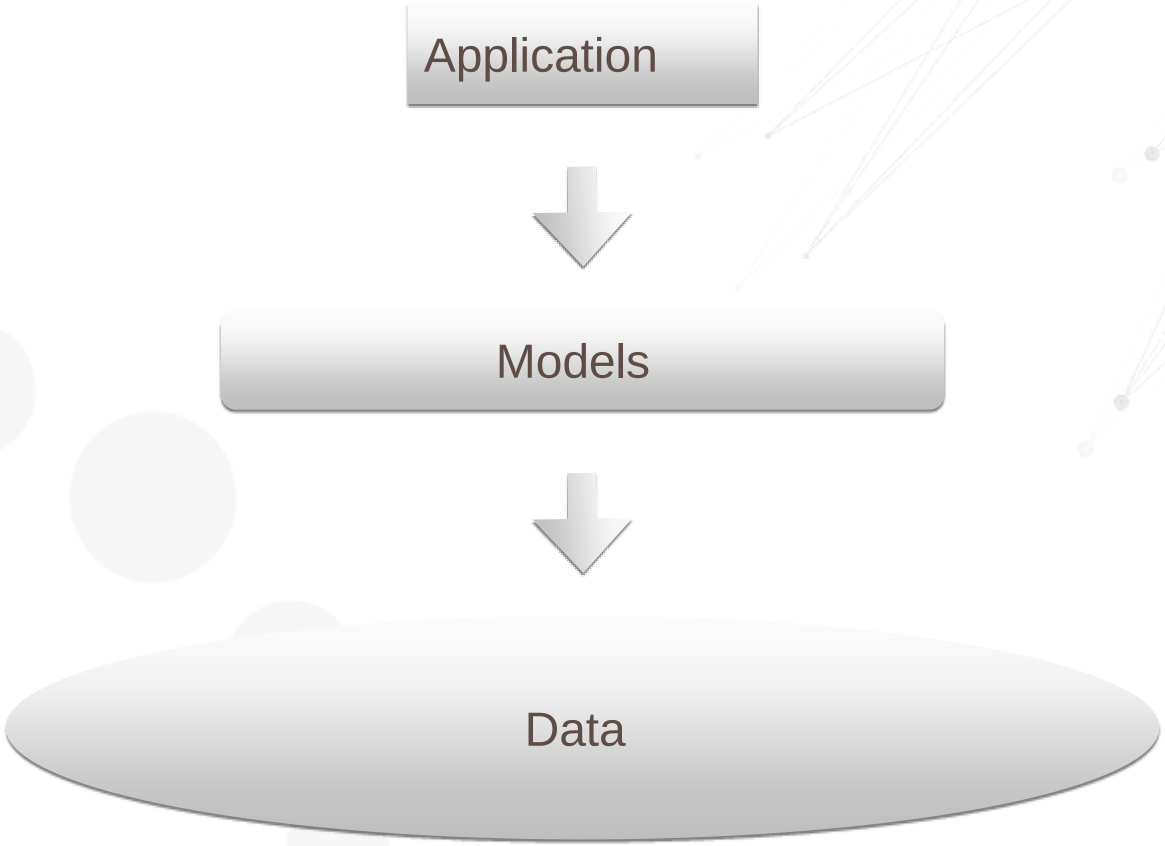
```
CREATE TABLE users (
  id          number(12) NOT NULL ,
  firstname    nvarchar2(25) NOT NULL ,
  lastname     nvarchar2(25) NOT NULL,
  email       nvarchar2(50) NOT NULL,
  password     nvarchar2(255) NOT NULL,
  created_date timestamp(6),
  PRIMARY KEY (id),
  CONSTRAINT email_uq UNIQUE (email)
);

-- Users by email address index
CREATE INDEX idx_users_email ON users (email);
```

```
CREATE TABLE videos (
  id number(12),
  userid number(12) NOT NULL,
  name nvarchar2(255),
  description nvarchar2(500),
  location nvarchar2(255),
  location_type int,
  added_date timestamp,
  CONSTRAINT users_userid_fk
    FOREIGN KEY (userid)
    REFERENCES users (Id) ON DELETE CASCADE,
  PRIMARY KEY (id)
);
```

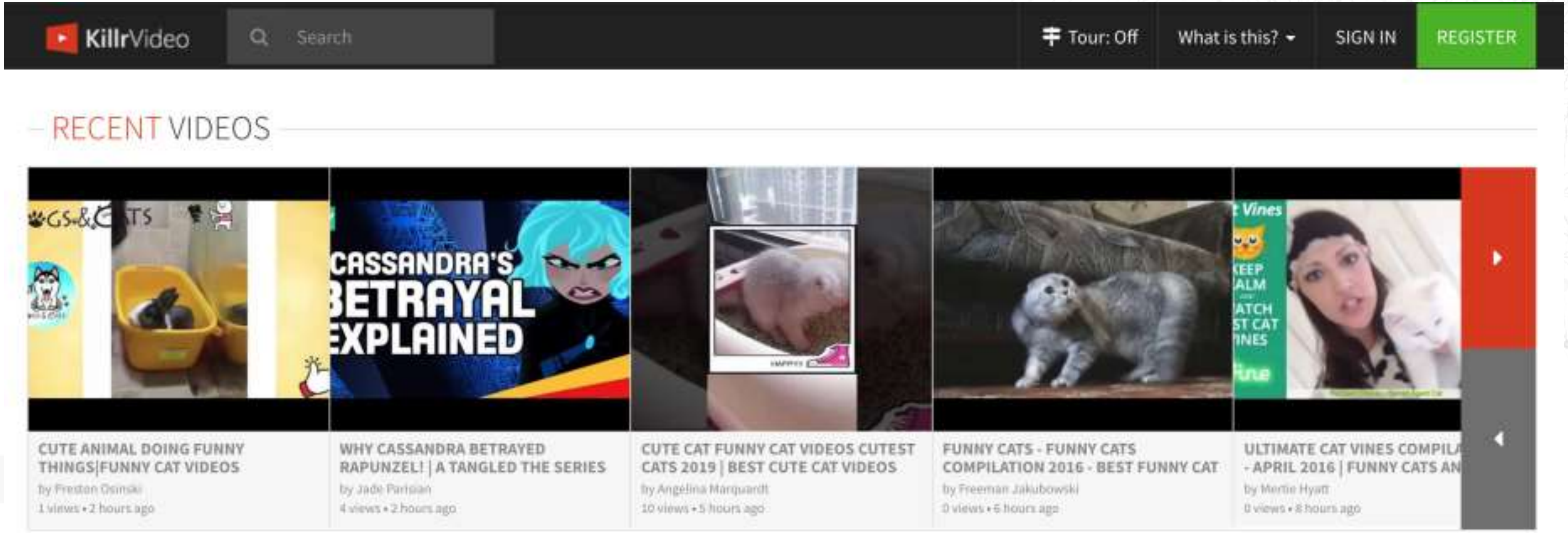
Relational Modeling



Cassandra Modeling



killrvideo.com



RECENT VIDEOS

CUTE ANIMAL DOING FUNNY THINGS|FUNNY CAT VIDEOS
by Preston Osinski
1 views • 2 hours ago

WHY CASSANDRA BETRAYED RAPUNZEL! | A TANGLED THE SERIES
by Jade Parisian
4 views • 2 hours ago

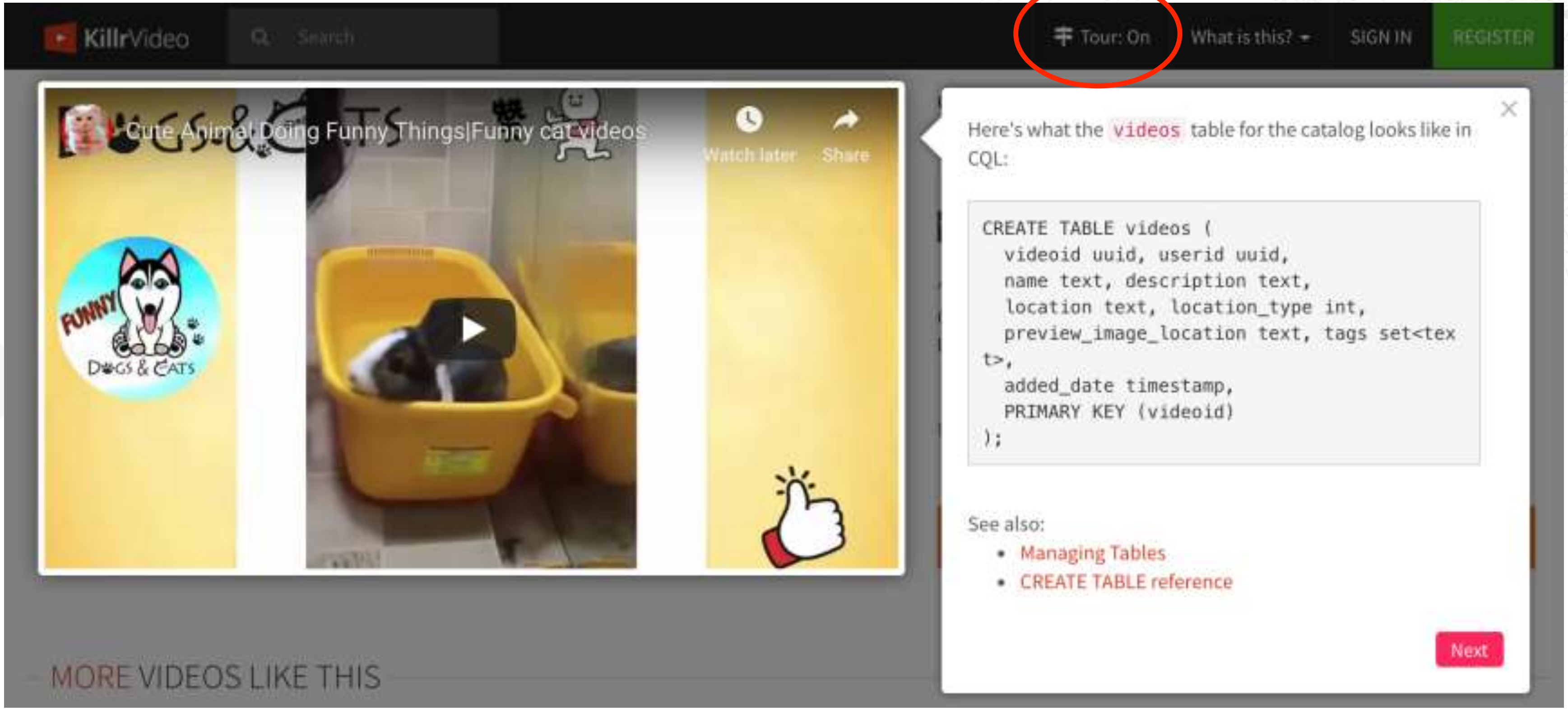
CUTE CAT FUNNY CAT VIDEOS CUTEST CATS 2019 | BEST CUTE CAT VIDEOS
by Angelina Marquardt
10 views • 5 hours ago

FUNNY CATS - FUNNY CATS COMPILATION 2016 - BEST FUNNY CAT
by Freeman Jakubowski
0 views • 6 hours ago

ULTIMATE CAT VINES COMPILA - APRIL 2016 | FUNNY CATS AN
by Mertie Hyatt
0 views • 8 hours ago

- Think a YouTube competitor
  - Users add videos, rate them, comment on them, etc.
  - Can search for videos by tag

killrvideo.com



Here's what the `videos` table for the catalog looks like in CQL:

```
CREATE TABLE videos (
  videoid uuid, userid uuid,
  name text, description text,
  location text, location_type int,
  preview_image_location text, tags set<tex
t>,
  added_date timestamp,
  PRIMARY KEY (videoid)
);
```

See also:

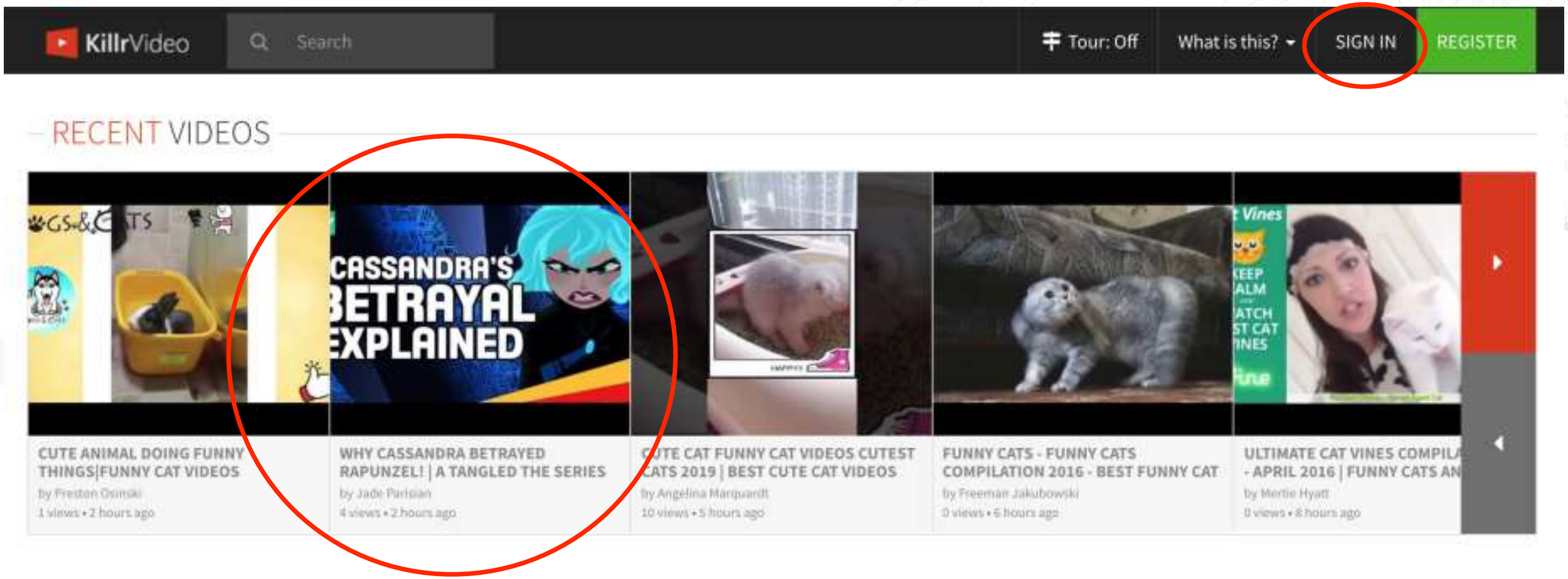
- Managing Tables
- CREATE TABLE reference

Next

1. Build Application Workflow

Workflow?

User logs in



RECENT VIDEOS

CUTE ANIMAL DOING FUNNY THINGS|FUNNY CAT VIDEOS
by Preston Osinski
1 views • 2 hours ago

WHY CASSANDRA BETRAYED RAPUNZEL! | A TANGLED THE SERIES
by Jade Parisian
4 views • 2 hours ago

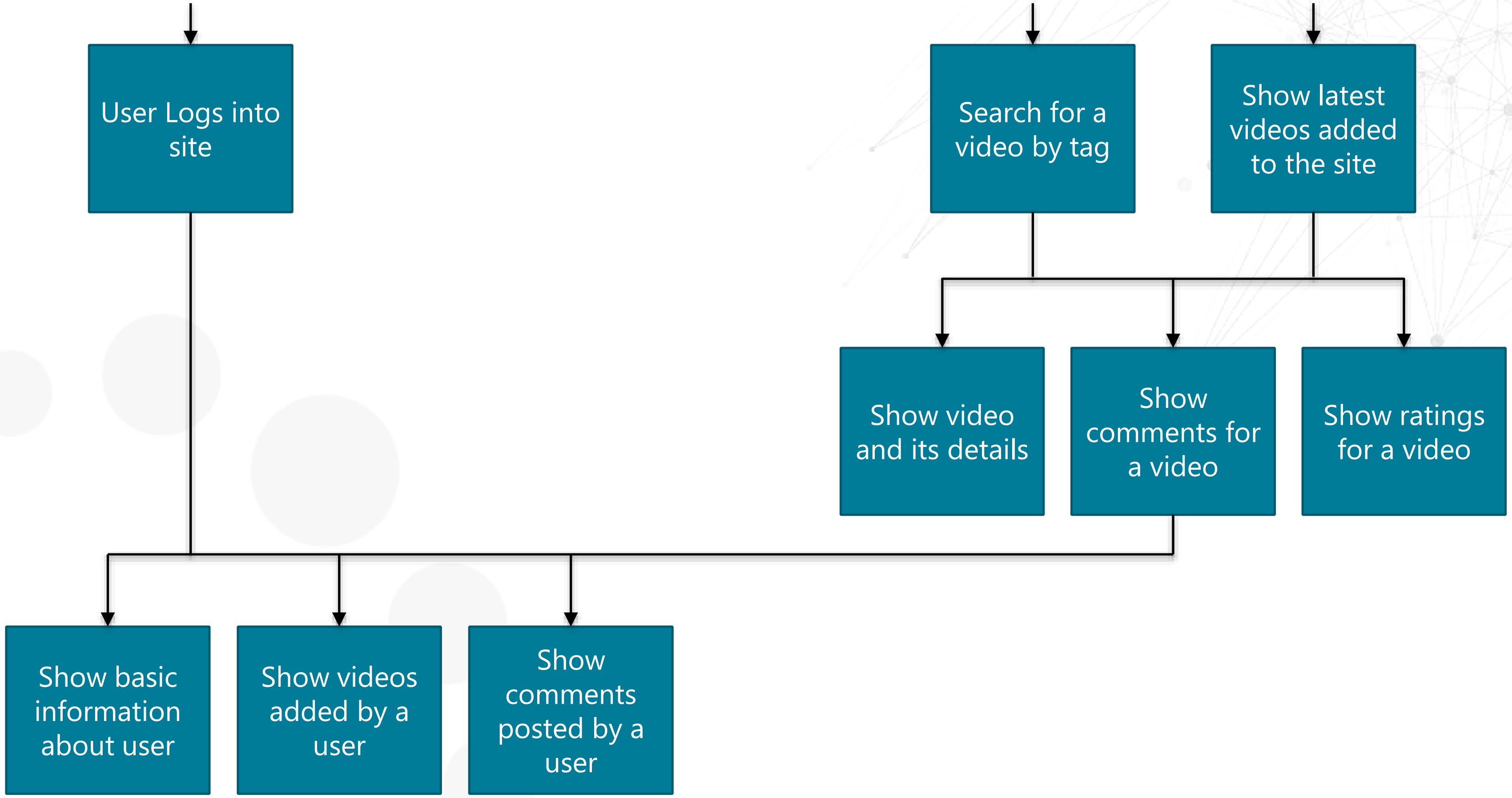
CUTE CAT FUNNY CAT VIDEOS CUTEST CATS 2019 | BEST CUTE CAT VIDEOS
by Angelina Marquardt
10 views • 5 hours ago

FUNNY CATS - FUNNY CATS COMPILATION 2016 - BEST FUNNY CAT
by Freeman Jakubowski
0 views • 6 hours ago

ULTIMATE CAT VINES COMPILA - APRIL 2016 | FUNNY CATS AN
by Mertie Hyatt
0 views • 8 hours ago

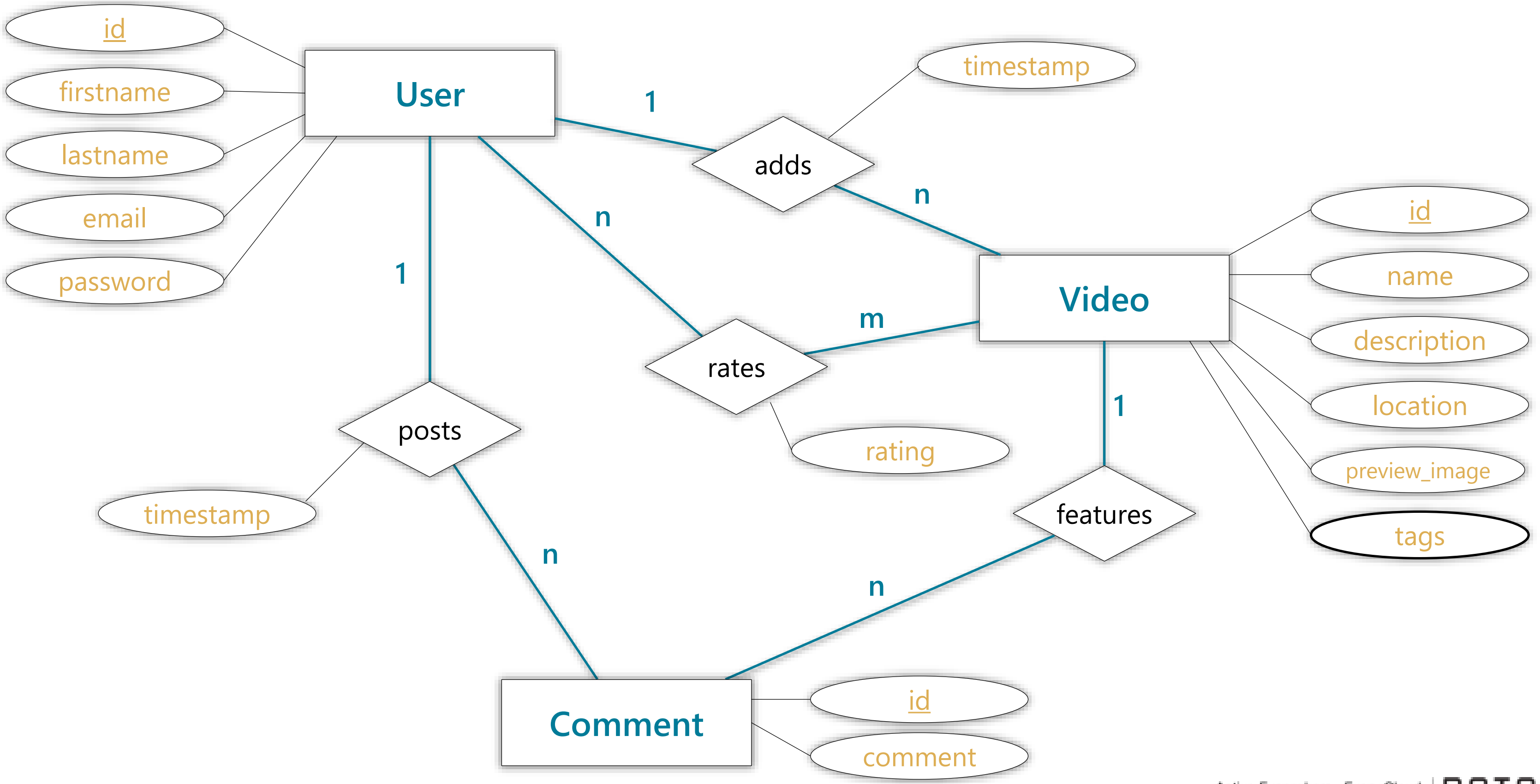
User selects video

Some Application Workflows in KillrVideo


User Logs into site
Search for a video by tag
Show latest videos added to the site
Show video and its details
Show comments for a video
Show ratings for a video
Show basic information about user
Show videos added by a user
Show comments posted by a user

2. Model Your Queries

Some of the Entities and Relationships in KillrVideo



Some Queries in KillrVideo to Support Workflows

Users

User Logs into site	Find user by email address
Show basic information about user	Find user by id

Comments

Show comments for a video	Find comments by video (latest first)
Show comments posted by a user	Find comments by user (latest first)

Ratings

Show ratings for a video	Find ratings by video

Some Queries in KillrVideo to Support Workflows

Videos

Search for a video by tag	Find video by tag
Show latest videos added to the site	Find videos by date (latest first)
Show video and its details	Find video by id
Show videos added by a user	Find videos by user (latest first)

3. Make Your Tables

Moving From Workflows

Entities
Single Name

User

Comment

Video

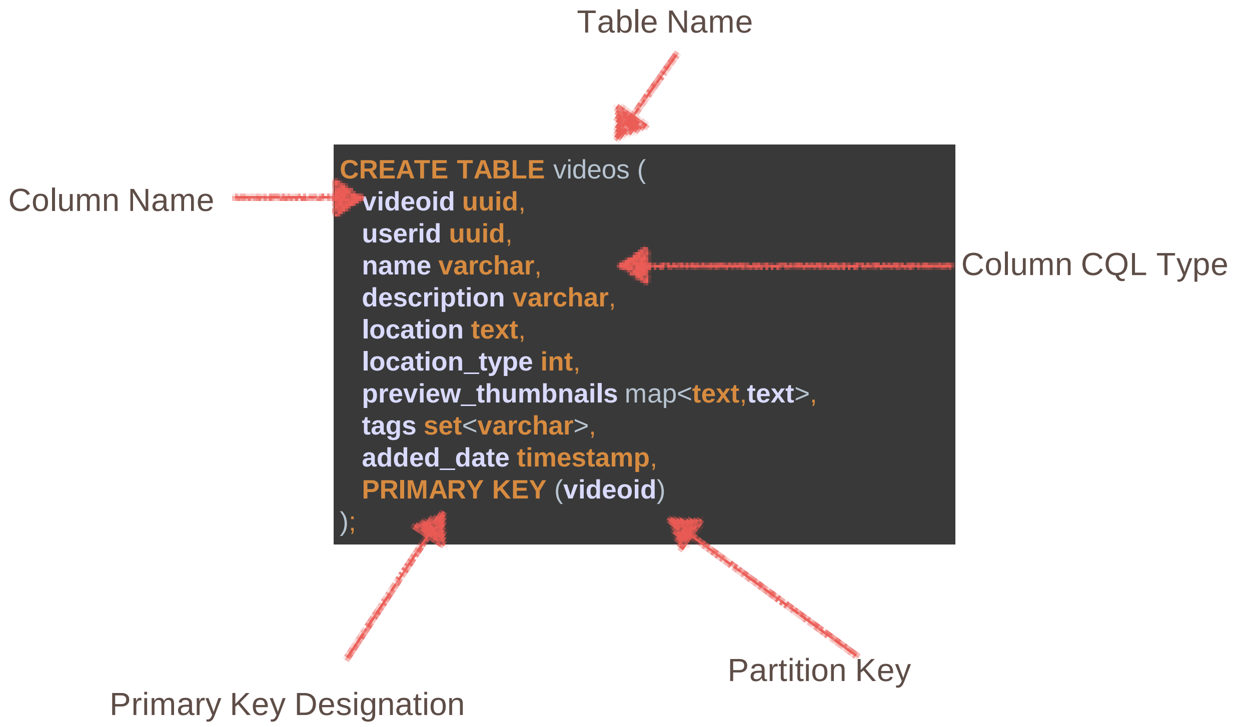
Relationships or Look Up
Descriptive Name

Show comments posted by a user	Find comments by user (latest first)
Search for a video by tag	Find video by tag

“Static” Table

Show video and its details

Find video by id



```
CREATE TABLE videos (
  videoid uuid,
  userid uuid,
  name varchar,
  description varchar,
  location text,
  location_type int,
  preview_thumbnails map<text,text>,
  tags set<varchar>,
  added_date timestamp,
  PRIMARY KEY (videoid)
);
```

“Dynamic” Table

Search for a video by tag

Find video by tag

```
CREATE TABLE videos_by_tag (
  tag text,
  videoid uuid,
  added_date timestamp,
  name text,
  preview_image_location text,
  tagged_date timestamp,
  PRIMARY KEY (tag, videoid)
);
```

Partition Key

Clustering Column

Users – The Cassandra Way

User Logs into site

Find user by email address

Show basic information about user

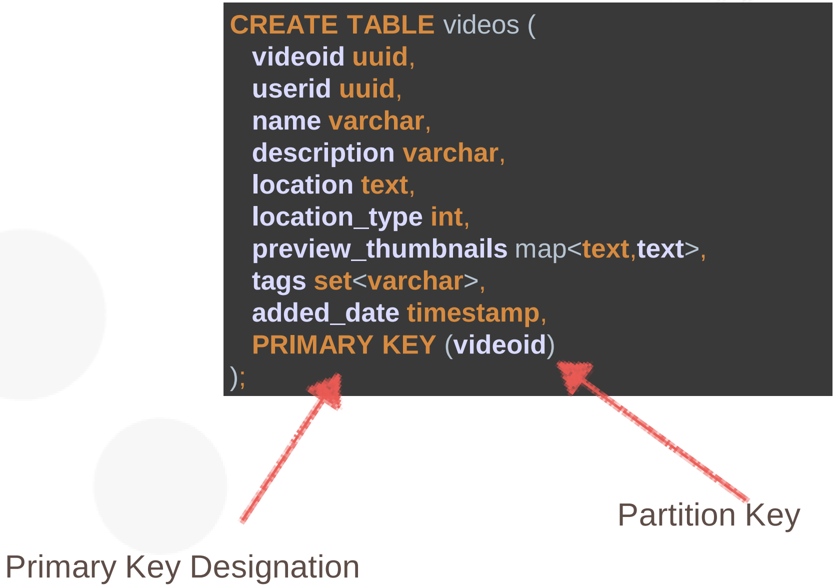
Find user by id

```
CREATE TABLE user_credentials (
   email text,
   password text,
   userid uuid,
   PRIMARY KEY (email)
);
```

```
CREATE TABLE users (
   userid uuid,
   firstname text,
   lastname text,
   email text,
   created_date timestamp,
   PRIMARY KEY (userid)
);
```

4. Get The Primary Key Right

Partition Key



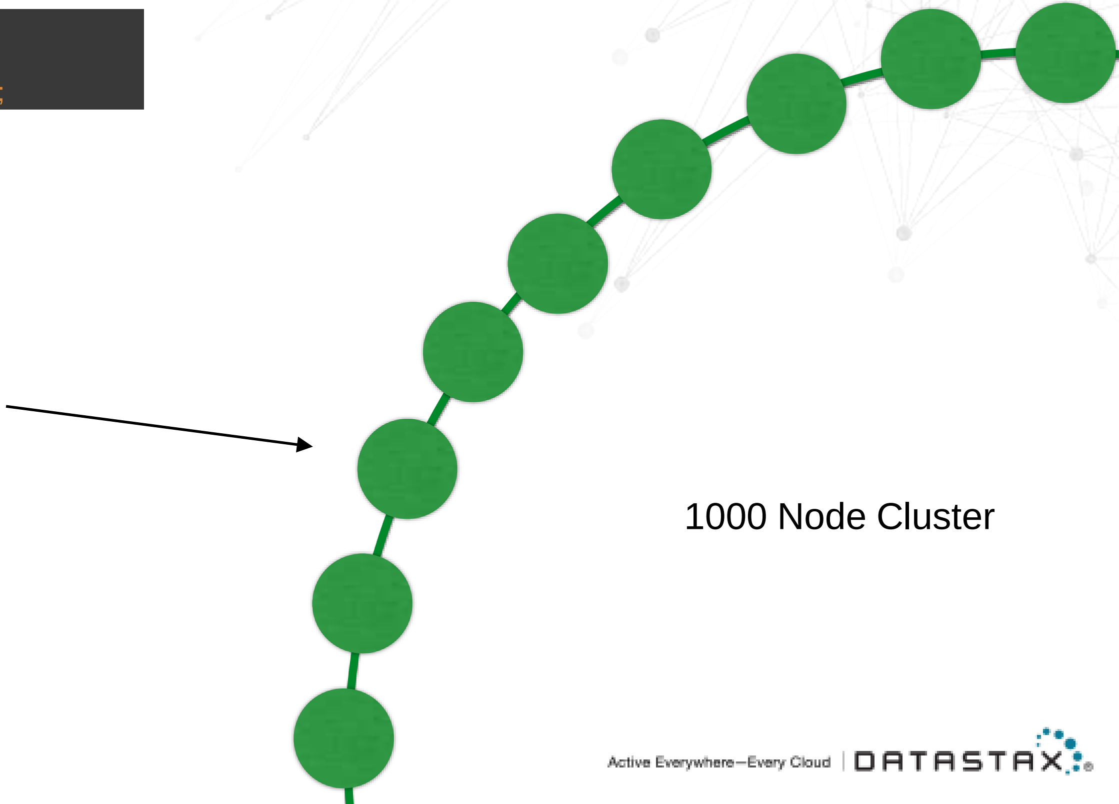
```
CREATE TABLE videos (
  videoid uuid,
  userid uuid,
  name varchar,
  description varchar,
  location text,
  location_type int,
  preview_thumbnails map<text,text>,
  tags set<varchar>,
  added_date timestamp,
  PRIMARY KEY (videoid)
);
```

Locality

```
SELECT name, description, added_date
FROM videos
WHERE videoid = 06049cbb-dfed-421f-b889-5f649a0de1ed;
```

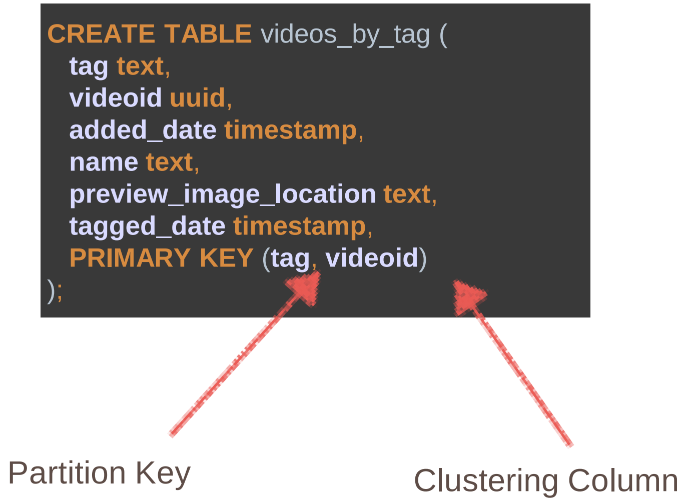
Partition Key: **REQUIRED**

videoid = 06049cbb-dfed-421f-b889-5f649a0de1ed



1000 Node Cluster

Why Dynamic?



```
CREATE TABLE videos_by_tag (
   tag text,
   videoid uuid,
   added_date timestamp,
   name text,
   preview_image_location text,
   tagged_date timestamp,
   PRIMARY KEY (tag, videoid)
);
```

Primary key relationship

PRIMARY KEY (tag,videoid)

Primary key relationship

PRIMARY KEY (tag,videoid)

Partition Key

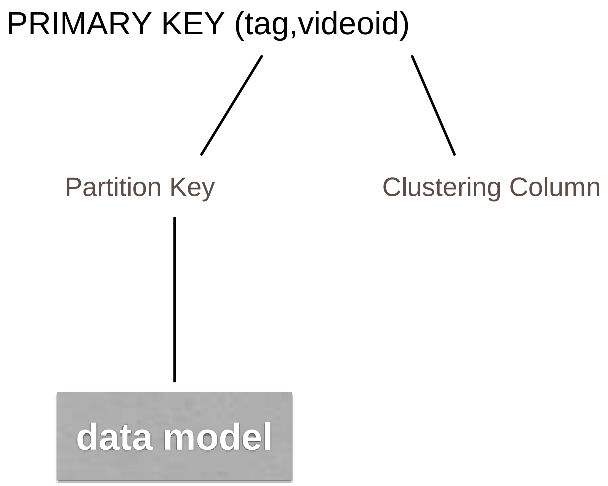
Primary key relationship

PRIMARY KEY (tag,videoid)

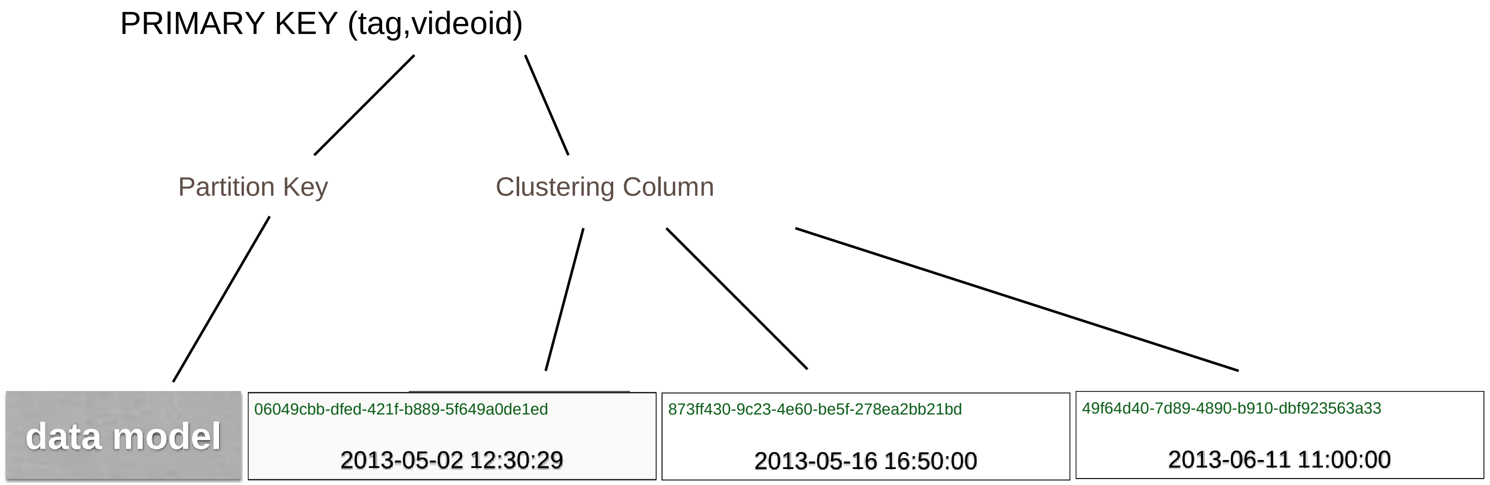
Partition Key

Clustering Column

Primary key relationship



Primary key relationship


PRIMARY KEY (tag,videoid)
Partition Key
Clustering Column
data model
06049cbb-dfed-421f-b889-5f649a0de1ed
2013-05-02 12:30:29
873ff430-9c23-4e60-be5f-278ea2bb21bd
2013-05-16 16:50:00
49f64d40-7d89-4890-b910-dbf923563a33
2013-06-11 11:00:00

5. Use Data Types Effectively

Data Types

1 - Data Marshalling

2 - Controlling Order

```
CREATE TABLE videos (
   videoid uuid,
   userid uuid,
   name varchar,
   description varchar,
   location text,
   location_type int,
   preview_thumbnails map<text,text>,
   tags set<varchar>,
   added_date timestamp,
   PRIMARY KEY (videoid)
);
```

Full Schema!

Controlling Order

- Controls row ordering when used as clustering column
- Default is ASC and can be overridden

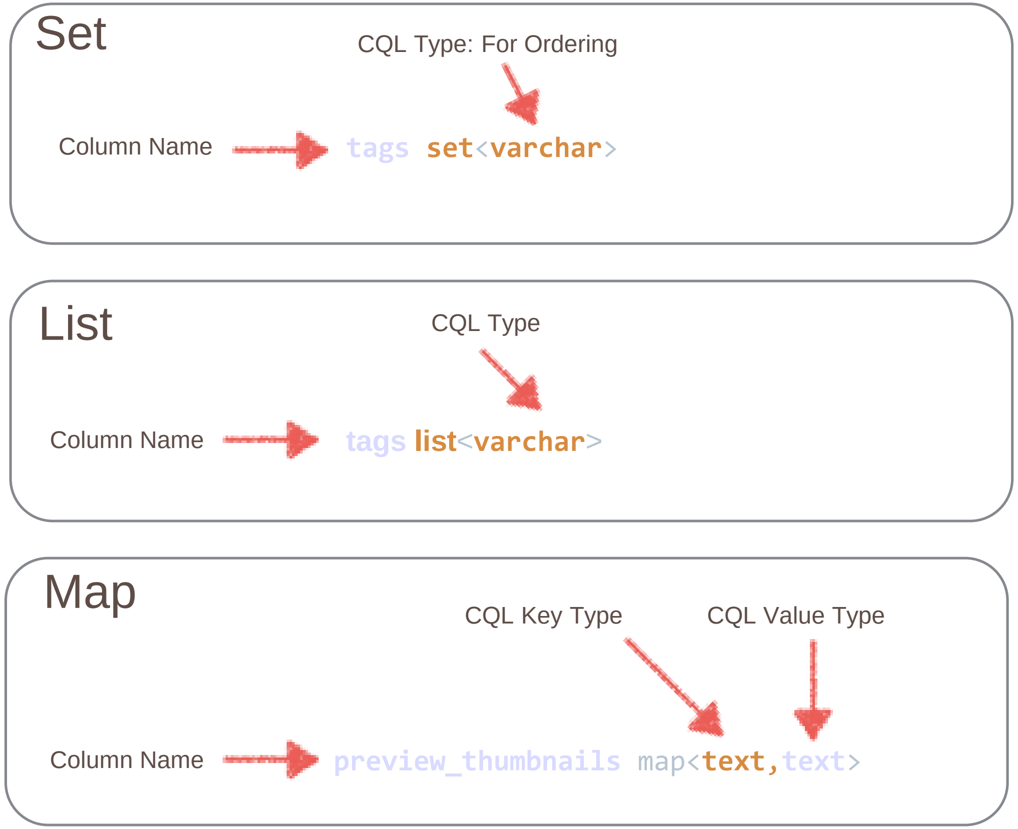
INT
VARCHAR
DATE
TIMESTAMP
TIMEUUID

Special Java Type Matches

Most types are obvious to Java, but….

CQL type	Java type
decimal	java.math.BigDecimal
float	java.lang.Float
double	java.lang.Double
varint	java.math.BigInteger

Collections



```
CREATE TABLE videos (
   videoid uuid,
   userid uuid,
   name varchar,
   description varchar,
   location text,
   location_type int,
   preview_thumbnails map<text,text>,
   tags set<varchar>,
   added_date timestamp,
   PRIMARY KEY (videoid)
);
```

What now?

Go do it!

Open Source	Open Source with Support	Full Enterprise Edition
Apache Cassandra™	DataStax Distribution of Apache Cassandra™	DataStax Enterprise

Thank you