Thursday, 10 December 2020 7:19 PM
(i) Gasher Requirements
1 Estimate scale
(3) Schema + APIS
Design Goals
(5) Remove bottlenecks and scale according to the requirement
X × ×
Video ingestion System!
To Requirements.
1) Requirements.  To Any signed in user can upload a rideo
o filter by language, category
o filter by language, caterjory o generate thumbnail, preview.  The video must adhere to copyright
claims.  Support greo-restrictions  Supported formats: mp4, mov
© size: 100 MB
verity video quality (720p)
Escale down the video to 320p,
$480p$ $2 \times 10^{5}$
2) Scale 2 Scale 2 Billion users
1 % users uploods vide de every day
TPS $\frac{2 \times 10^9}{100} = \frac{20 \text{ M}}{2 \times 10^7} = \frac{2 \times 10^7}{2 \times 3600} \approx \frac{250}{500} = \frac{20 \times 3600}{24 \times 3600} \approx \frac{1000}{200} = \frac{20 \times 10^9}{200} $
side o
$\odot$ 50%. Users consume at $\sim$ 10 k/sec
Storage: 200 M uploade
X 100 x 18 / agree
$Z = 2 \times 10^{10} \text{ MI}$ $= 20 \text{ PB}$
3) Schema
Juser Video
- type - user-id
- vidéo-id - languages
- categories.
formai
- lang - thum brail
video-eategon - preview
- cat-id - video-id
I a a fe a ord
category - id name
name
a Design Goals
Partition -
<ul> <li>Partition</li> <li>Consistent = ? Yes</li> </ul>
<ul> <li>Partition</li> <li>Consistent = ? Yes</li> <li>Availability = ? Yes</li> <li>lafency: can be high</li> </ul>
<ul> <li>Partition</li> <li>Consistent = ? Yes</li> </ul>
<ul> <li>Partition</li> <li>Consistent = ? Yes</li> <li>Availability = ? Yes</li> <li>lafency: can be high</li> </ul>
<ul> <li>Partition</li> <li>Consistent =? Yes</li> <li>Availibility =? Yes</li> <li>latency: can be high</li> <li>scale: horizontally.</li> <li>Initial Design</li> </ul>
<ul> <li>Partition</li> <li>Consistent =? Yes</li> <li>Availability =? Yes</li> <li>Lafency: can be high</li> <li>scale: horizontally.</li> <li>Initial Design</li> </ul>
<ul> <li>Partition</li> <li>Consistent =? Yes</li> <li>Availability =? Yes</li> <li>Iafency: can be high</li> <li>scale: horizontally.</li> <li>Initial Design</li> <li>Copyright (hear)</li> <li>check</li> </ul>
<ul> <li>Portition</li> <li>Consistent =? Yes</li> <li>Availability =? Yes</li> <li>latency: can be high</li> <li>scale: horizontally.</li> <li>Initial Design</li> <li>( hash)</li> <li>check</li> <li>apply country lows rate</li> <li>cooperate</li> </ul>
Partition  Consistent =? Yes  Availability =? Yes  Initial Design  Initial Design  Company to the country lows appropriate (saperate)  2. apply country lows (saperate)  3. format check (metadota)
Partition  Consistent =? Yes  Nailibility =? Yes  Instial Design  Copyright (nesh)  Check  2. apply country lows says sen  3. for mat check (netadoda)  4. check all (sep)  video frames (sys)
Partition  Consistent =? Yes  Nailibility =? Yes  Instial Design  Copyright (nesh)  Check  2. apply country lows says sen  3. for mat check (netadoda)  4. check all (sep)  video frames (sys)
O Consistent =? Yes  O Availibility =? Yes  O latency: can be high  O scale: horizontally.  (b) Initial Design  C API S ROST  1. copyright (hear)  check  2. apply country lows all  check all (sep yide of frames (sys)  Video frames (sys)  5. convert in (sel)  multiple formats.
O Consistent = ? Yes  O Availibility = ? Yes  O latency: can be high  O scale: horizontally.  (B) Initial Design  (C) API S ROT  1. copyright (hest)  Check  2. apply country lows  check (metadoda)  4. check all (sep)  video frames (sys)  5. convert in  multiple formats.  6. generate Thumbrails (sep)
O Consistent =? Yes  O Availibility =? Yes  O latency: can be high  O scale: horizontally.  (b) Initial Design  (c) API   S   RDFI  1. copyright (hest)  Check  2. apply country lows rate  2. apply country lows rate  3. format check (metadoda)  4. check all (sep)  video frames (aps)  5. convert in (sup)  multiple formats.  6. generate Thumbnails (sep)  7. generate preview (ss)
O Consistent = ? Yes  O Availibility = ? Yes  O latency: can be high  O scale: horizontally.  (B) Initial Design  (C) API S ROT  1. copyright (hest)  Check  2. apply country lows  check (metadoda)  4. check all (sep)  video frames (sys)  5. convert in  multiple formats.  6. generate Thumbrails (sep)
Pertition  Consistent = ? Yes  Availibility = ? Yes  Intel Design  Linital Design  Linital Design  Concert Service  2. apply country lows 2. apply country lows 4. check all (sep)  Indeed frames (syl)  S. convert in (sep)  multiple formats.  Generate Trumbneils (sp)  7. generate preview (ss)  8. convert to (ower (ss)  9. save  bully code
Partition  Consistent =? Yes  Nailibility =? Yes  Initial Design  Copyright (nest)  check  apply country lows role  role of frames (sys)  sconvert in (self)  multiple formats.  generate Thumbnails (sy)  res  econvert to (ower (ss)  res  sources
Partition  Consistent = ? Yes  Nacilibility = ? Yes  Interpretation  Scale: horizontally.  Sinitial Design  Consistent  1. copyright (nest)  Check  2. apply county lows rate  check all (sep)  video frames (sys)  S. convert in (sep)  mattiple formats.  6. generate Thumbhails (sp)  7. generate preview (ss)  8. convert to (ower (ss)  9. save  bully code  can fail without alberty
O Consistent = ? Yes  O Availability = ? Yes  O latency: can be high  O scale: horizontally.  (appropriate (not)  1. copyright (not)  2. apply country (nos rate)  4. check all (sep video frames (sys)  5. convert in (sur)  multiple formats.  6. generate Thumbrails (sep)  7. generate preview (ss)  8. convert to lower (ss)  8. convert to lower (ss)  9. save  O bully code  O each stp can take of ferent  o can feil without allerty  g start the pipeline grain.  Traffic will be uneven (Queue)
O Consistent = ? Yes  O Availability = ? Yes  O latency: can be high  O scale: horizontally.  (appropriate (not)  1. copyright (not)  2. apply country (nos rate)  4. check all (sep video frames (sys)  5. convert in (sur)  multiple formats.  6. generate Thumbrails (sep)  7. generate preview (ss)  8. convert to lower (ss)  8. convert to lower (ss)  9. save  O bully code  O each stp can take of ferent  o can feil without allerty  g start the pipeline grain.  Traffic will be uneven (Queue)
O Consistent = ? Yes  O Availibility = ? Yes  O latency: can be high  O scale: horizontially.  (appropriate (nest)  1. coppyright (nest)  2. apply county (superally)  3. format check (netabola)  4. check all (sep video frames (sp)  5. convert in (sep)  multiple formats.  6. generate Thumbnails (sep)  7. generate preview (ss)  8. convert to lower (ss)  8. convert to lower (ss)  9. save  O bully code  o each step can take of flerent (sine.  o an fail without allerthy  a start the pipeline again.  Traffic will be uneven (Queue)  odon't need to respond inwedicably
O Consistent = ? Yes  O Availability = ? Yes  O latency: can be high  O scale: horizontally.  (appropriate (not)  1. copyright (not)  2. apply country (nos rate)  4. check all (sep video frames (sys)  5. convert in (sur)  multiple formats.  6. generate Thumbrails (sep)  7. generate preview (ss)  8. convert to lower (ss)  8. convert to lower (ss)  9. save  O bully code  O each stp can take of ferent  o can feil without allerty  g start the pipeline grain.  Traffic will be uneven (Queue)
O Consistent = ? Yes  O Availibility = ? Yes  O latency: can be high  O scale: horizontially.  (appropriate (nest)  1. coppyright (nest)  2. apply county (superally)  3. format check (netabola)  4. check all (sep video frames (sp)  5. convert in (sep)  multiple formats.  6. generate Thumbnails (sep)  7. generate preview (ss)  8. convert to lower (ss)  8. convert to lower (ss)  9. save  O bully code  o each step can take of flerent (sine.  o an fail without allerthy  a start the pipeline again.  Traffic will be uneven (Queue)  odon't need to respond inwedicably
O Consistent = ? Yes  O Availibility = ? Yes  O latency: can be high  O scale: horizontially.  (appropriate (nest)  1. coppyright (nest)  2. apply county (superally)  3. format check (netabola)  4. check all (sep video frames (sp)  5. convert in (sep)  multiple formats.  6. generate Thumbnails (sep)  7. generate preview (ss)  8. convert to lower (ss)  8. convert to lower (ss)  9. save  O bully code  o each step can take of flerent (sine.  o an fail without allerthy  a start the pipeline again.  Traffic will be uneven (Queue)  odon't need to respond inwedicably
Consistent =? Yes  Consistent =? Yes  Availibility =? Yes  intency: can be high  scale: horizonially.  Listed Design  Check  apply country in a check (metadolo)  apply country in a check (metadolo)  acheck all (set)  sconert in (set)  convert in (set)  convert in (set)  generate preview (ss)  convert to lower (ss)  convert to lower (ss)  convert to lower (ss)  con fail without alkerty  distant the pipeline again.  traffic will be uneven (Queue)  and oban't need to respond inmediaty  Applications  Applications  Applications  Single
O consistent = ? Yes  O knowlibility = ? Yes  O latercy: can be high  O scale: horizonially.  (S) Initial Design  (I repry that (not)  2. apply country has be recovered.  3. format check (notedorn)  4. check all (seq)  video frames (set)  5. convert in (sel)  multiple formats.  6. generale numbrails (2°)  7. generale preview (ss)  8. convert to lower (ss)  8. convert to lower (ss)  1. save  (Soully code  (Source)  (Soully code  (Soully code  (Source)  (S
Consistent = ? Yes  Consistent = ? Yes  Nailibility = ? Yes  Intercy : car. be high  Scale : horizontally.  (Intercy)  Interce   Same    I
O consistent = ? Yes  O knowlibility = ? Yes  O latercy: can be high  O scale: horizonially.  (S) Initial Design  (I repry that (not)  2. apply country has be recovered.  3. format check (notedorn)  4. check all (seq)  video frames (set)  5. convert in (sel)  multiple formats.  6. generale numbrails (2°)  7. generale preview (ss)  8. convert to lower (ss)  8. convert to lower (ss)  1. save  (Soully code  (Source)  (Soully code  (Soully code  (Source)  (S
O Partition  O Consistent = ? Yes  Newlibrity = ? Yes  O lecterary: can be high  O scale: hard sometably.  Distal Design  Lichal Design  Licopyright (not)  2. apply country loss with  2. apply country loss with  3. format check (not state)  4. check all (of)  1. copyright (not)  2. apply country loss with  4. check all (of)  1. copyright (not)  4. check all (of)  5. convert in (of)  6. generale previous (ss)  7. generale previous (ss)  8. convert to lower (ss)  8. convert to lower (ss)  9. save  To bully so de  O can fail without alertry  2 shart the pipeline again.  To traffic will be uneven (Queue)  To den't need to respond innoclarly  Solarit need to respond innoclarly  To Apply Down D.  Sall
Densistent = ? Yes  Densistent = Per high  Descent = Per high  Descent = Per p