



LUCENE / SOLR  
**REVOLUTION** / 2016

OCTOBER 11-14, 2016 • BOSTON, MA



# Near Real time Indexing

Building Real Time Search Index For E-Commerce

Umesh Prasad  
Tech Lead @ Flipkart

Thejus V M  
Data Architect @ Flipkart



# Agenda

- Search @ Flipkart
- Need for Real Time Search
- SolrCloud Solution
- Our approach
- Q & A



## Introducing Pixel Phone by Google

Up to ₹27,000 off on exchange

PRE-ORDER NOW

\*T&C Apply

## Deals of the Day

⌚ 23 : 24 : 10



Kids' Wear

Minimum 60% Off

Little Stars, AJ Dezines...



Under Rs.1,479

Sennheiser, Phillips..

Headphones



Acer Aspire E APU Quad Core ...

4GB RAM | 1TB HDD

Acer Aspire A10 Laptop



SanDisk SDDDC2-032G-G46 / S...

Just at Rs.

₹629 ₹925



Top-selling books

Min. 45% Off

Fiction, Self-help, Academ...



By appointment delivery  
from Samsung Plaza

[KNOW MORE](#)

**Samsung**  
**Delight** 

Introducing  
moto Z with motomods

COMING SOON



Launching  
honor 8<sup>°°</sup> at ₹29,999

+ EXCHANGE OFFER

No Cost EMI for 24 months



Launching  
Zenfone 3 Laser

AT JUST ₹18,999

+ Exchange Offer | No Cost EMI





Prachi Jain  
@MissPrachiD

Follow

It was my first shopping experience with @Flipkart. Got my order before time. Satisfied. Thank You! 😊  
#BigBillionDays



Vishalitiwari  
@vishalitiwari291

Follow

Mumbai rain destroyed my box but still the product was on time thank you @Flipkart  
#BigBillionDays #bigbilliondaysareback



Pranay  
@TheEngineerInMe

Follow

Pre-ordered stuff on @amazonIN on 30th Aug. Still not delivered.  
Ordered stuff on @Flipkart on 2nd Oct. Delivered today. In bubble wrap. #win

5:31 PM - 4 Oct 2016

4h 12 1 1



Prerna Srivastava  
@prerna\_maniac

Follow

LUCENE / SOLR  
REVOLUTION / 2016



Large BBD 2016  
Customer  
Delight Stories

Creating an Awesome  
★★★★★  
Customer Experience

Great experience. Awesome user demonstration and communication. Executive was so deliverable and clear for each interaction with me.

Alok Singh Negi



Kripashankar

Timely service. Thanks for such a fast and amazing service. I wish same promptness in installation too.

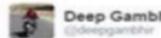
Ready had nice experience during shopping with flipkart.

Ankit Jain



Monu

WHAT IS YOUR CUSTOMER DELIGHT STORY?



Deep Gambhir  
@Deepgambhir

Follow

@\_sachinbansal @Flipkart ordered a microwave at 11 AM and its delivered at 5 PM..superfast delivery and service  
#BigBillionDay ZipZapZoom...



AVA  
@stunningmoon

Following

I ordered a hair brush on @Flipkart on Saturday evening..on Sunday morning it was delivered home. No need of prime when Flipkart is here. :P

RETWEET 1 LIKES 4



12:15 PM - 3 Oct 2016



Prerna Srivastava  
@prerna\_maniac

Follow

U guys never disappoint me @Flipkart, got my products delivered 2 days b4 the date given & in perfect condition.Thanq 4 making r lives easy

10:52 AM - 2 Oct 2016



It's amazing to see people happy with @Flipkart doing better numbers than its foreign competition :-)  
#Respect

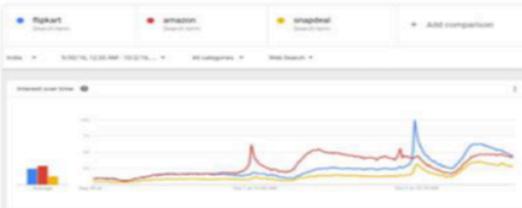
12:09 AM - 4 Oct 2016



Sameer Parwani  
@spanwani

Follow

According to search volume, Flipkart's Big Billion Days is still the king of Diwali sales



Tweet from the founder of CouponDunia



Maahir Panchal  
@maahirpanchal

Follow

That moment when @Flipkart offers exchange service in my location but @amazonIN fails to do so. 😂

1:04 AM - 3 Oct 2016



Mukund  
@mukund02

Follow

#LatestSmartphonesOnFlipkart offers r simply super.. @Flipkart is killing @amazonIN market share this festive season.

12:46 AM - 3 Oct 2016

Lucidworks

# Traffic @ Flipkart

- Peak Traffic
  - ~ 800K active users
  - ~ 160K requests per second
- Search Traffic
  - ~ 40K searches per second (Service)
  - ~ 10K searches per second (Solr )
- Latency
  - Median : 11 ms
  - 99th percentile : 1.1 second

# Search @ Flipkart

- Catalogue
  - ~ 50 main categories
  - ~ 5000 sub-categories
  - ~ 231 million documents
  - ~ 90 million SKUs
  - ~ 160 million listings
- E-commerce Marketplace
  - ~ 100K Sellers
  - Local Sellers
  - Regional Availability
  - Logistics Constraints

# E-commerce Search

- Heavy usage of drill down filters
- Heavy usage of faceting
- Only top results matter
- Results grouped/collapsed by products
- *Serviceability and delivery experience MATTERS*

# Agenda

- Search @ Flipkart
- Need for Real Time Search
- SolrCloud Solution
- Our approach
- Q & A

# Sorry, Stock Over !!?

SHOWING 234,138

Search within Footwear 

BROWSE

▼ Footwear

- Men's Footwear
- Women's Footwear
- Kids' & Infant Footwear
- Shoe Accessories

Looking for offers?  
All the Great Offers you wish for, now on the Flipkart App!  
[OPEN IN APP](#)

REFINE

BRAND

Search by Brand 

- Puma
- Nike
- Reebok
- Fila
- Adidas
- Lee Cooper
- Woodland

Popular High Price Low Price New

  
OUT OF STOCK

Reebok REE TRACTION Running Shoes  
★★★★★ (17 ratings) ▶  
Rs. 1,499 Rs. 2,299 34%



Puma Limnos CAT 2 DP Sneakers  
★★★★★ (66 ratings) ▶  
Rs. 1,379 Rs. 2,299 40%



Puma Atom Fashion II DP Running Shoes  
★★★★★ (12 ratings) ▶  
Rs. 1,619 Rs. 2,699 40%

[See Similar Products ▾](#)







# Damn !! Is Offer Over ??



flipkart.com

Search for a product, category or brand

SEARCH

View all offers (3) »

MINIMUM 20% OFF on Nikon High-End Cameras

Rs. 10001 - Rs. 15000 (0)

Rs. 15001 - Rs. 20000 (0)

Rs. 20001 - Rs. 30000 (1)

Rs. 30001 - Rs. 50000 (1)

Rs. 50001 and Above (0)

Type

Search by Type

- Point & Shoot (8)
- DSLR (2)
- SLR (0)
- Sports & Action (0)
- Advanced Point & Shoot (0)
- Lens Style Camera (0)
- Camcorder (0)

Rs. 4,654

Nikon Coolpix L30 Point & Shoot Camera (Black) and 2 more variants  
Rs-6,950 22% OFF

OUT OF STOCK

1 OFFER

Rs. 5,414

Nikon Coolpix S3600 Point & Shoot Camera (Black) and 4 more variants  
Rs-6,950 22% OFF

OUT OF STOCK

1 OFFER

Rs. 4,690

Nikon Coolpix S2800 Point & Shoot Camera (Silver) and 5 more variants  
Rs-6,950 21% OFF

OUT OF STOCK

1 OFFER

Rs. 4,740

Nikon Coolpix S2800 Point & Shoot Camera (Red) and 5 more variants  
Rs-6,460 27% OFF

OUT OF STOCK

1 OFFER

Brand

Search by Brand

- Nikon (10)
- Canon (0)
- Panasonic (0)
- Fujifilm (0)
- Pentax (0)

OUT OF STOCK

1 OFFER

# What !! All Steal Deals Gone ??



Before Rs. 6,999  Offer: Sold Out	Now Rs. 99  Zync Z99 Tablet <a href="#">SHOP NOW</a>	Before Rs. 799  Offer: Sold Out	Now Rs. 1  Transcend JetFlash 360 16 GB Pen ... <a href="#">SHOP NOW</a>	Before Rs. 14,790  Offer: Sold Out	Now Rs. 1,390  Offer: Sold Out
Flat 90% Off		Flat 90% Off			
 Offer: Sold Out	 Offer: Sold Out				

# Product / Listing: Important Attributes

Home > Mobiles & Accessories > Mobiles > Apple Mobiles > Apple iPhone 6 (Gold, 16 GB)



## Apple iPhone 6 (Gold, 16 GB)

★★★★★ 855 242 REVIEWS

- 6.9 mm Thickness
- 4.7 inch Touchscreen
- iOS 8
- 8 MP Primary Camera

Color



Storage

128 GB    16 GB    64 GB

[View Compatible Accessories](#)

Available with 32 Sellers at 560030 [Change](#)

Rs. 45,493

SOLD BY

iDreams

4.1 / 5

Advantage

Product aka SKU

Listings

catalogue service

Availability Service

Seller Rating Service

Promise Service

Offer Service

Pricing Service

SELLERS

RATING ↓

DELIVERED BY ↑

OFFERS

PRICE ↑

Konnect

○ 10 day Replacement Guarantee. ?

4.2 / 5

\* Tue, 15th Dec: Rs. 50

₹ Cash on Delivery is Available ?

Rs. 48,568

EMI: Rs. 2,355 ?

STIPL Advantage

○ 30 day Replacement Guarantee. ?

4.1 / 5

\* Mon, 14th Dec: FREE

₹ Cash on Delivery is Available ?

Thu, 10th Dec : Rs. 60

APP ONLY OFFER +

Rs. 46,999

EMI: Rs. 2,279 ?

iDreams Advantage

○ 30 day Replacement Guarantee. ?

4.1 / 5

\* Fri, 11th Dec: FREE

₹ Cash on Delivery is Available ?

APP ONLY OFFER +

Rs. 45,493

EMI: Rs. 2,206 ?

KPJENTERPRISE

○ 10 day Replacement Guarantee. ?

3.9 / 5

\* Mon, 14th Dec: FREE

Rs. 51,000

EMI: Rs. 2,473 ?

# Summary : Lucene Document

- Product/SKU (Parent Document)
  - Listing (Child Document)
- Query : Mostly SKU Attributes (Free Text)
- Filters : SKU + Listing Attributes (Drill Down)
- Ranking : SKU + Listing Attributes (Explicit/  
Relevance)
- *Index Time Join aka Block Join  
Performance* (Best Performance)

# Out Of Stock, but Why Show?

BROWSE  
✓ Footwear  
Men's Footwear  
Women's Footwear  
Kids' & Infant Footwear  
Shoe Accessories

## Looking for offers?

All the Great Offers you wish for, now on the Flipkart App!

[OPEN IN APP](#)

## REFINE

### BRAND

Search by Brand



- Puma
- Nike
- Reebok
- Fila
- Adidas
- Lee Cooper
- Woodland

Index has Stale Availability Data



Reebok REE TRACTION Running Shoes

★★★★★ (17 ratings) ▶  
Rs. 1,499 Rs. 2,299 34%

SHOWING 234,138

Popular

High Price

Low Price

New

Search with



Puma Limnos CAT 2 DP Sneakers

★★★★★ (66 ratings) ▶  
Rs. 1,379 Rs. 2,299 40%



Puma Atom Fashion II DP Running Shoes

★★★★★ (12 ratings) ▶  
Rs. 1,619 Rs. 2,699 40%

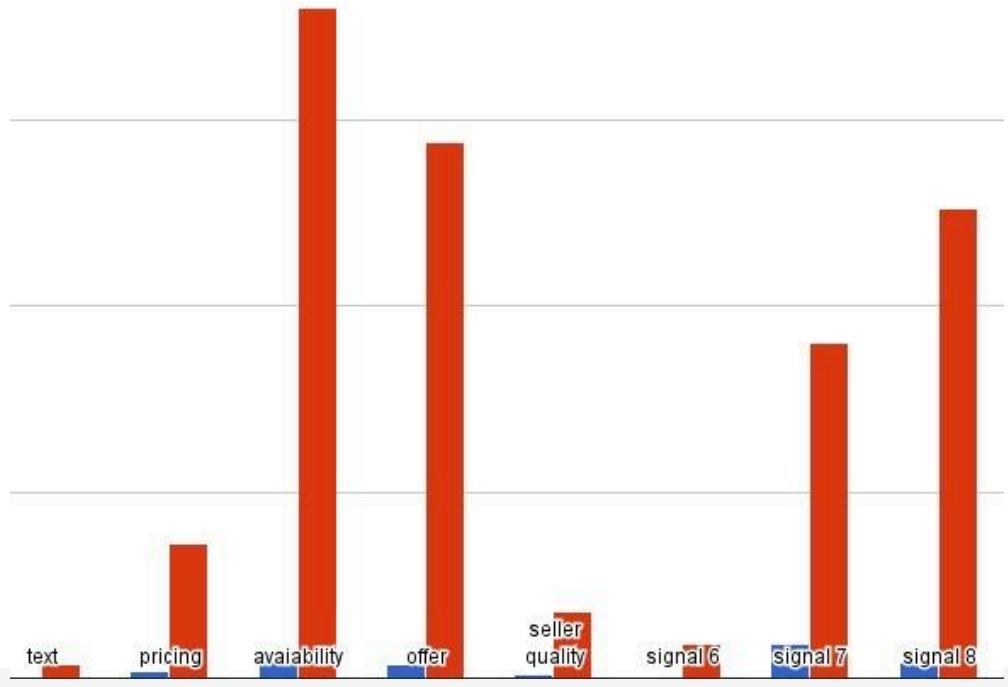
[See Similar Products ▾](#)

[See Similar Products ▾](#)



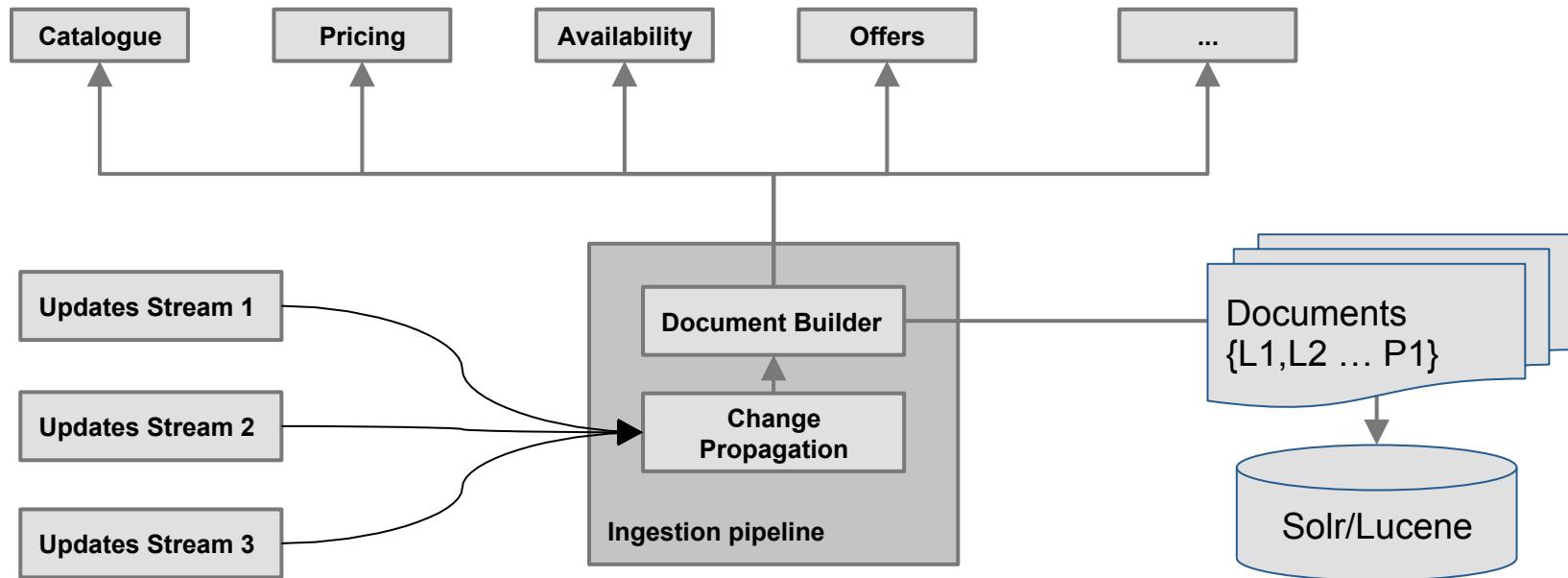
# Challenge 1 : High Update Rates

	updates / sec		updates / hr	
	normal	Peak		
text / catalogue	~10	~100	~100K	~100K
pricing	~100	~1K	~10 million	~10 million
availability	~100	~10K	~10 million	~10 million
offer	~100	~10K	~10 million	~10 million
seller rating	~10	~1K	~1 million	~1 million
signal 6	~10	~100	~1 million	~1 million
signal 7	~100	~10K	~10 million	~10 million
signal 8	~100	~10K	~10 million	~10 million



# Challenge 2 : Micro Services

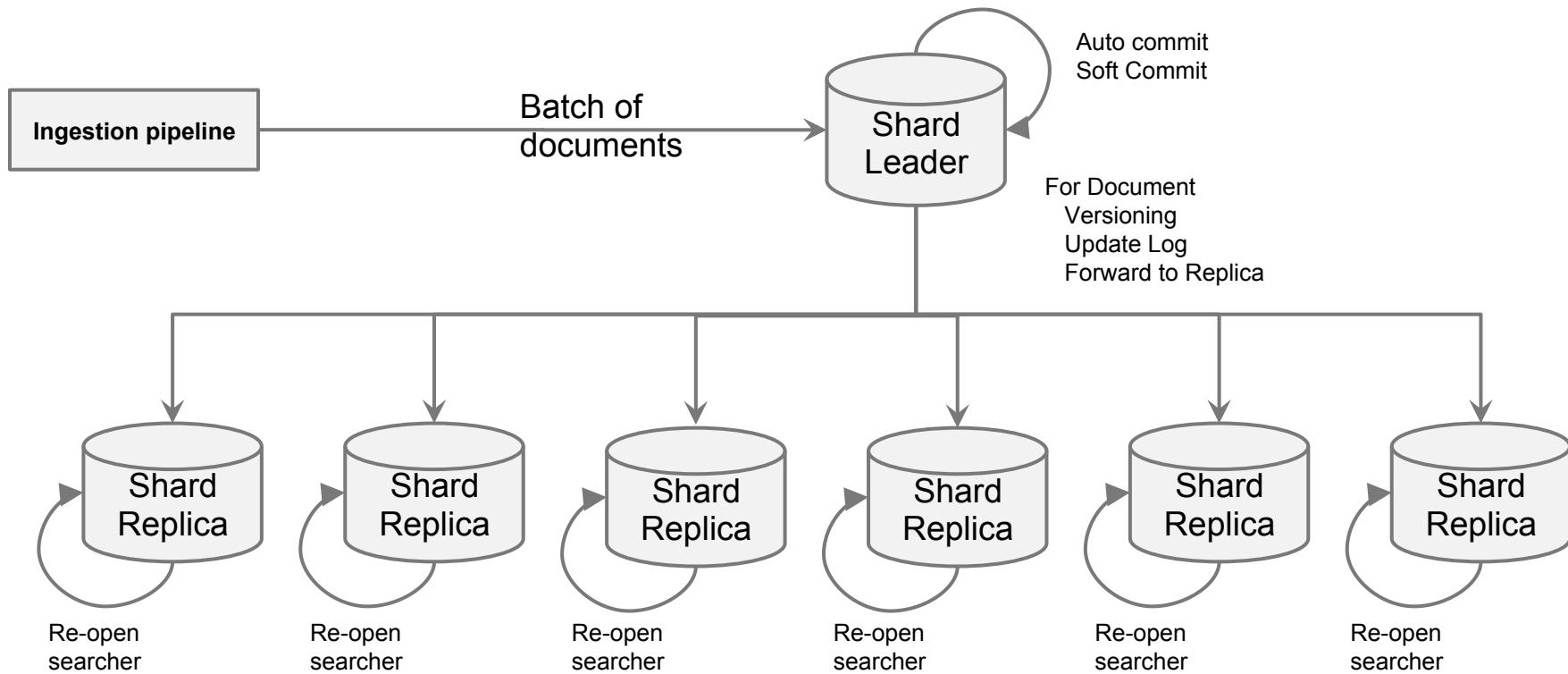
- Lucene doesn't support Partial Updates
- Update = Delete + Add



# Agenda

- Search @ Flipkart
- Need for Real Time Search
- SolrCloud Solution
- Our approach
- Q & A

# SolrCloud for NRT



# SolrCloud Evaluation

- Update = Delete + Add
  - Block Join Index ⇒ Update Whole Block (Product + Listings)
- Updated Document gets streamed to all replicas in sync
  - Reduces indexing throughput
- Soft commit is Not Free
  - Soft commit ⇒ In Memory Segment
  - Lots of Merges
  - Huge document churn / deletes
  - All caches still need to be re-generated
  - Filter Cache miss specially hurts performance

# Agenda

- Search @ Flipkart
- Need for Real Time Index
- SolrCloud Solution
  - Our approach
- Q & A

# Lucene Index

ProductA

brand : Apple

availability : T

price : 45000

ProductB

brand : Samsung

availability : T

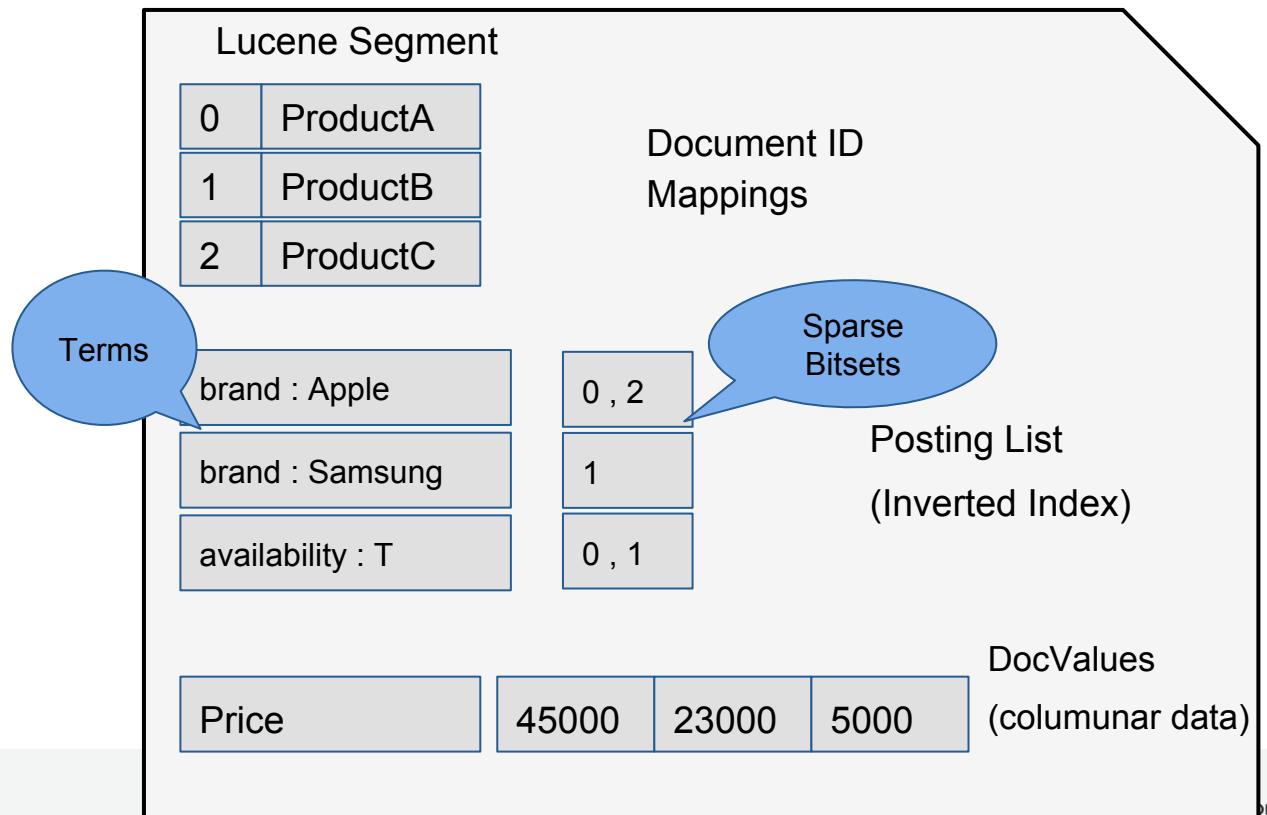
price : 23000

ProductC

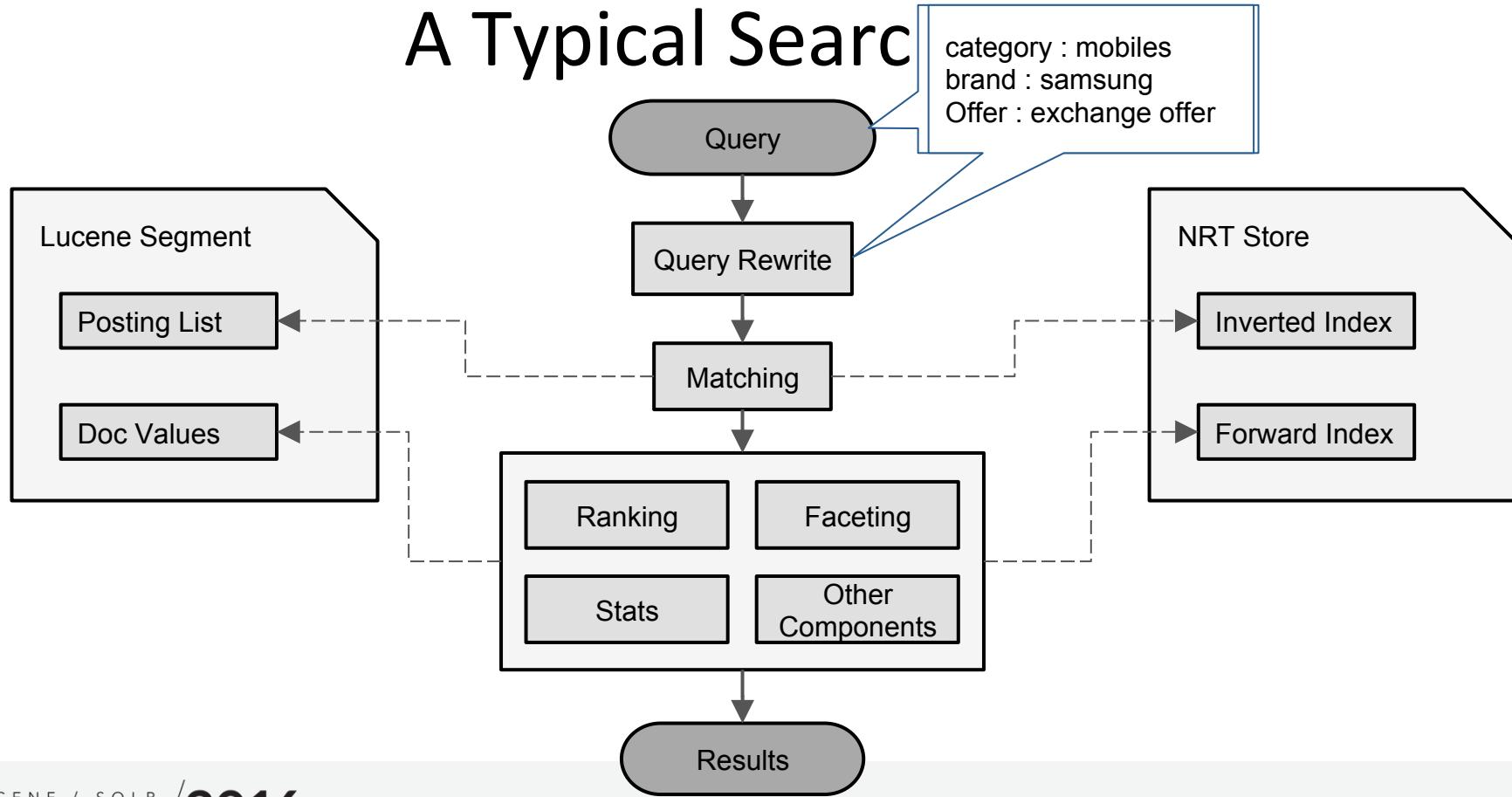
brand : Apple

availability : F

price : 5000



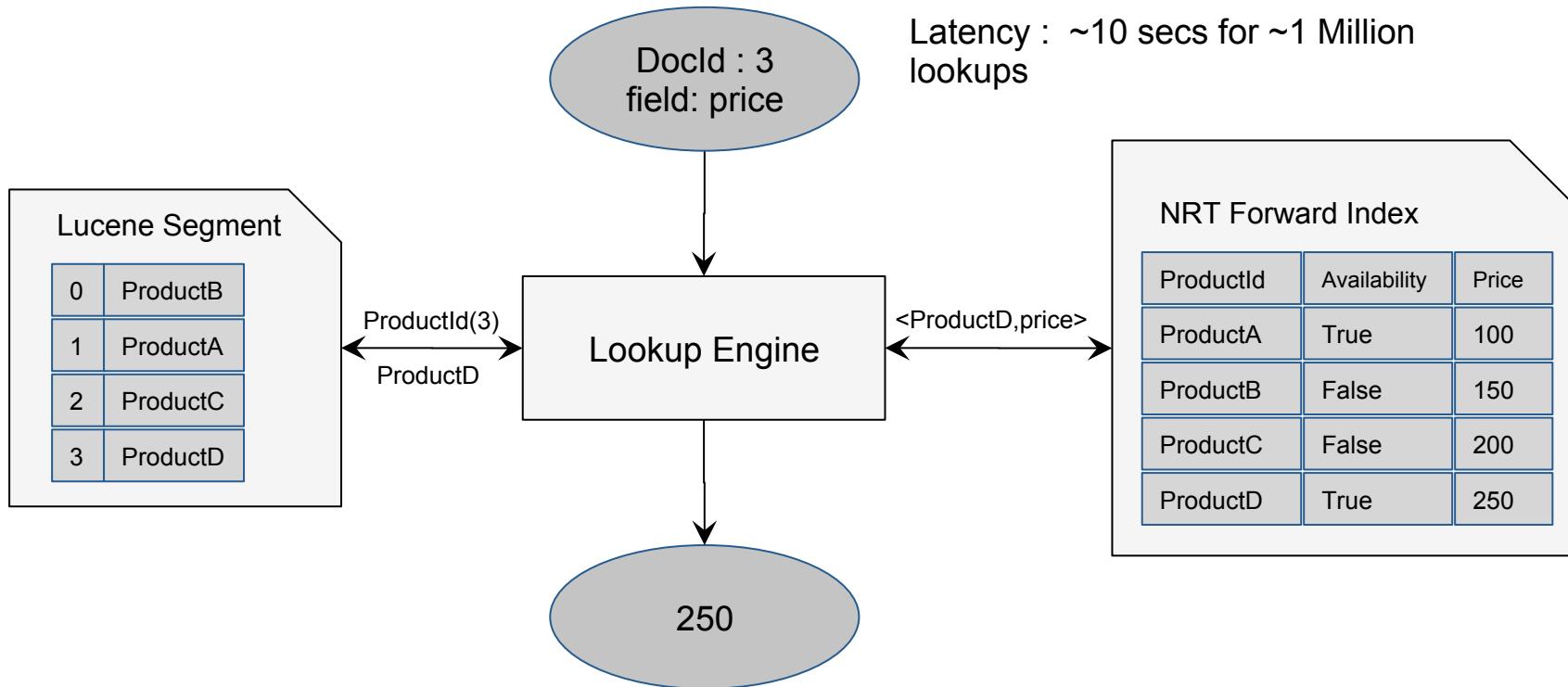
# A Typical Search



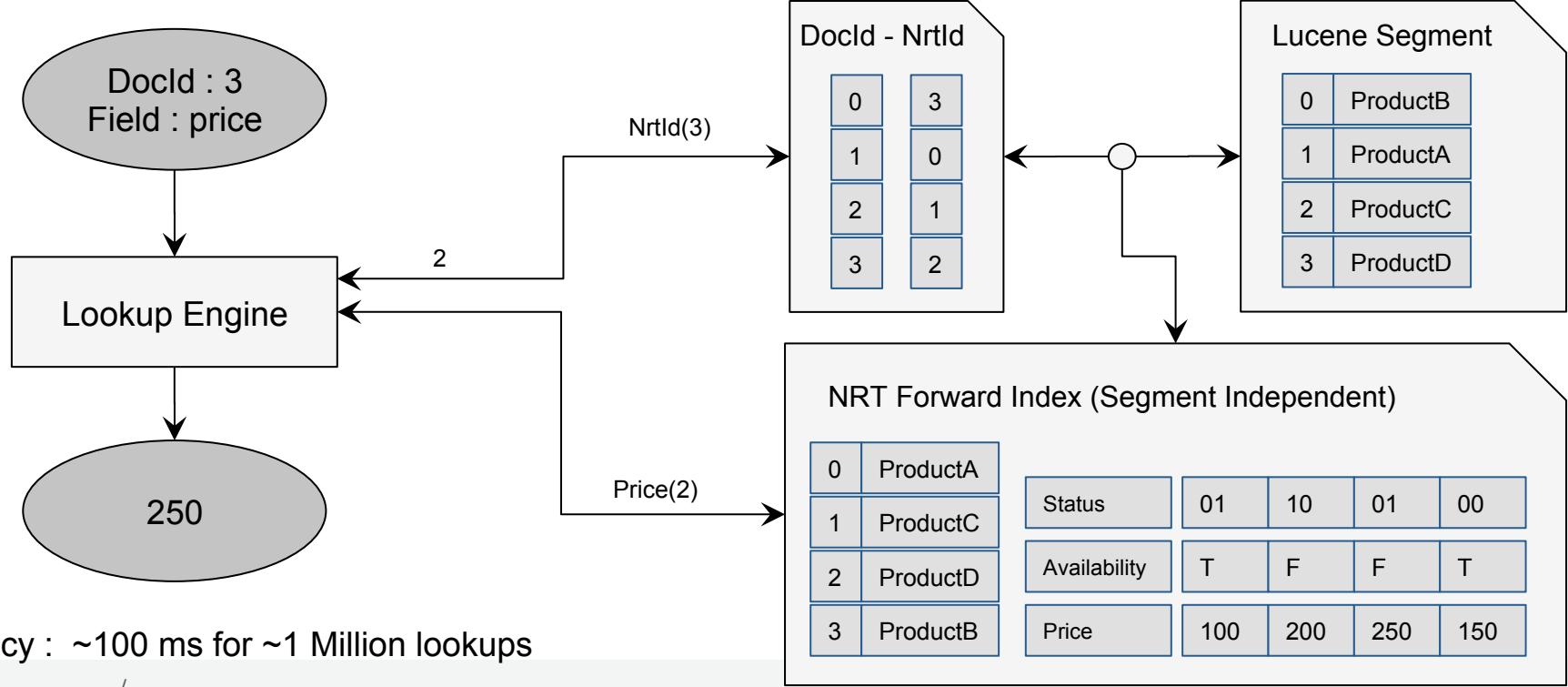
# NRT Forward Index - Considerations

- Lookup efficiency
  - 50th percentile : ~10K matches
  - 99th percentile : ~1 million matches
- Data on Java heap
  - Memory efficiency

# NRT Forward Index - Naive Implementation

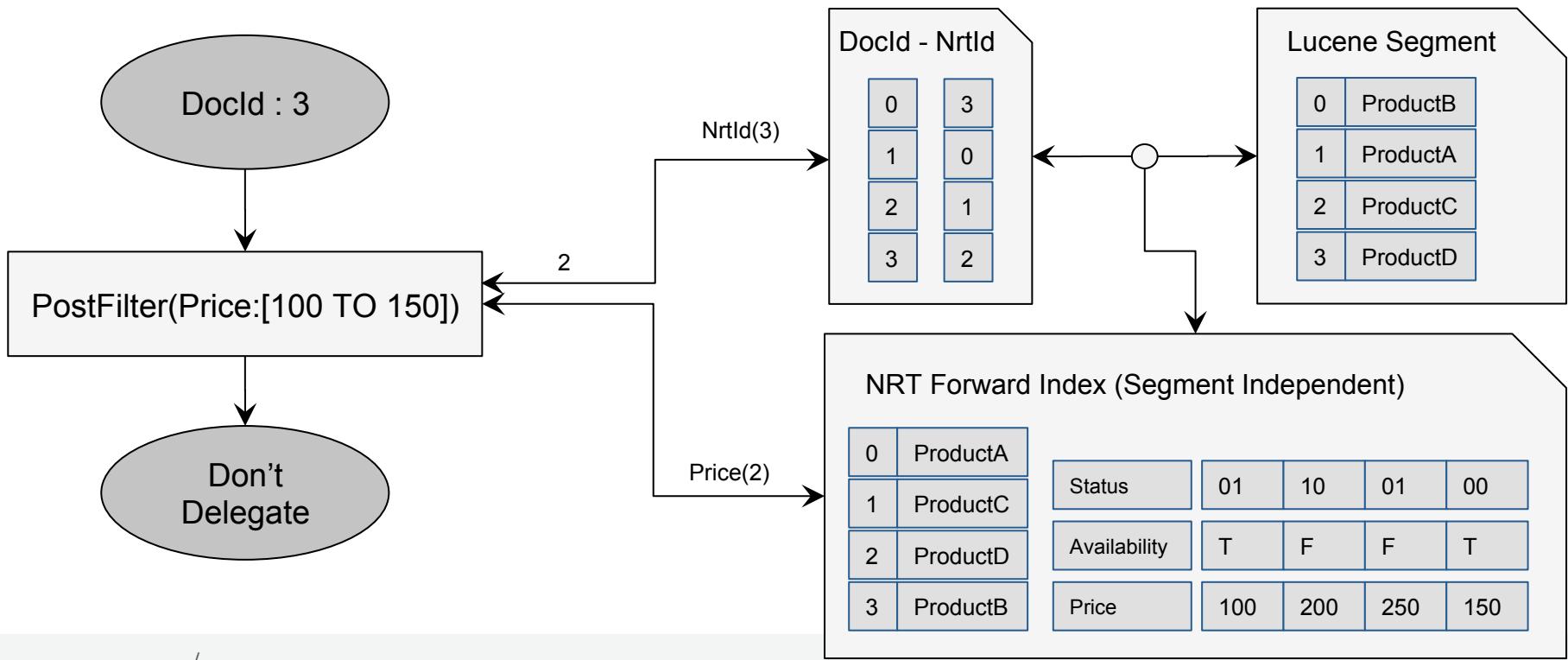


# NRT Store - Forward Index Optimized

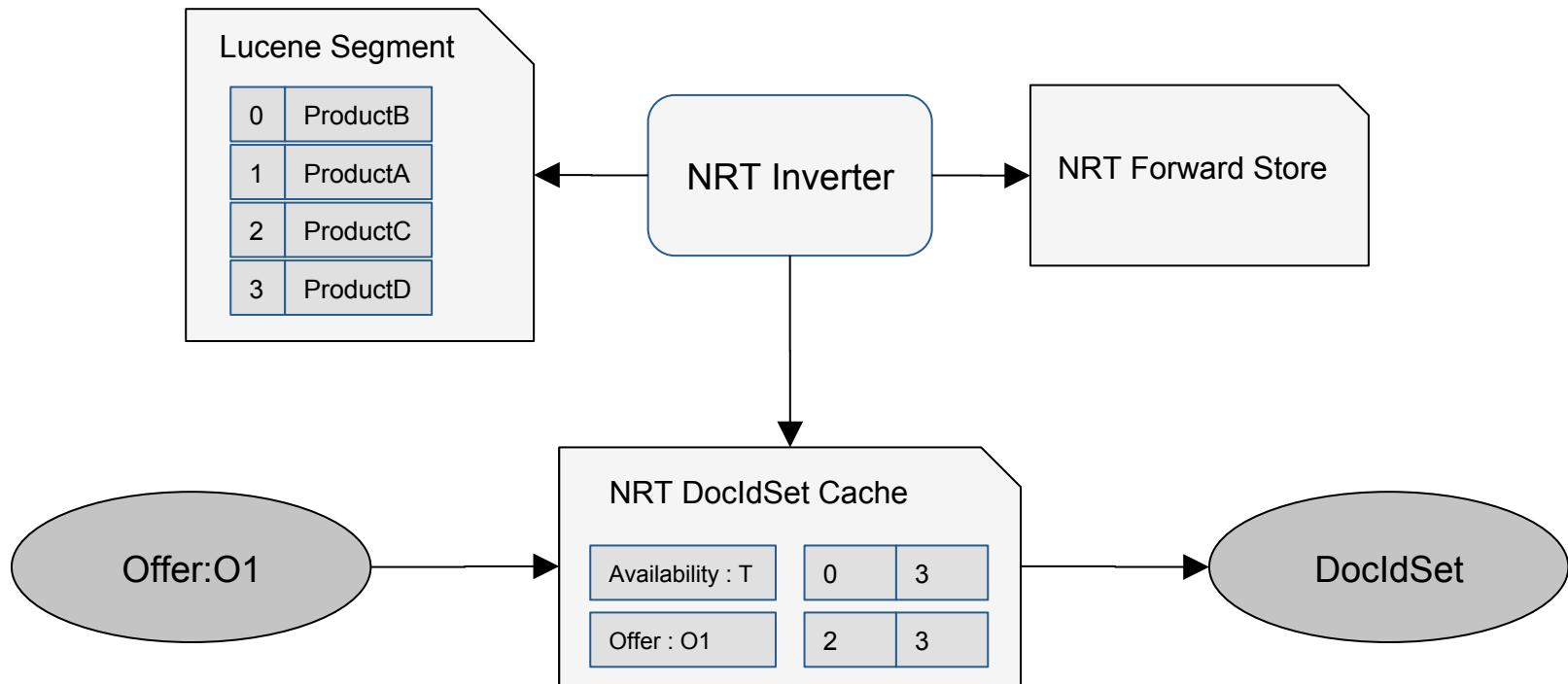


Latency : ~100 ms for ~1 Million lookups

# NRT Store Filter - PostFilter



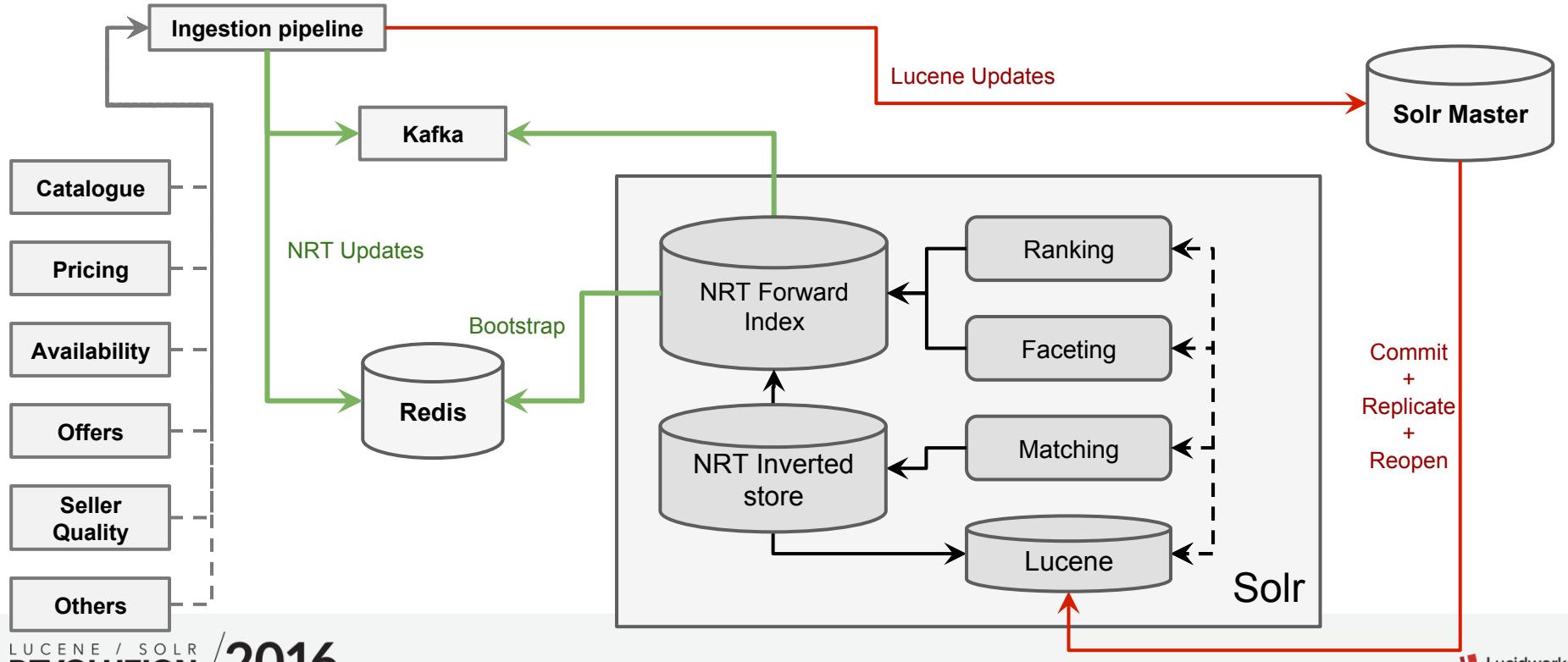
# NRT Store - Invert index



# Solr Integration Points

- ValueSources
- Filtering
  - Custom Filter Implementation for cached DocIdSet
  - Custom PostFilter
- Query
  - Wrapper over Filter
- Custom FacetComponent

# Near Real Time Solr Architecture



# Accomplishments

- Real time sorting
- Real time filtering : PostFilter
  - Higher latency
- Near real time filtering : cached DocIdSet
  - No consistency between lookup and filtering
- Independent of lucene commits
- Query latency comparable to DocValues
  - Consistent 99% performance

# Accomplishments @ Flipkart

- Real time consumption for ~150 Signals
- Reduction in shown out of stock products by 2X
- Production instances of ~50K updates/second real time

Thank you  
&  
Questions