

PhoenixGraph

Yaswanth Biruduraju

Mihir Agarwal

Smeet Dedhia

Garvit Gopalani

Prajna Prabhu

Manipal Institute of Technology, Manipal



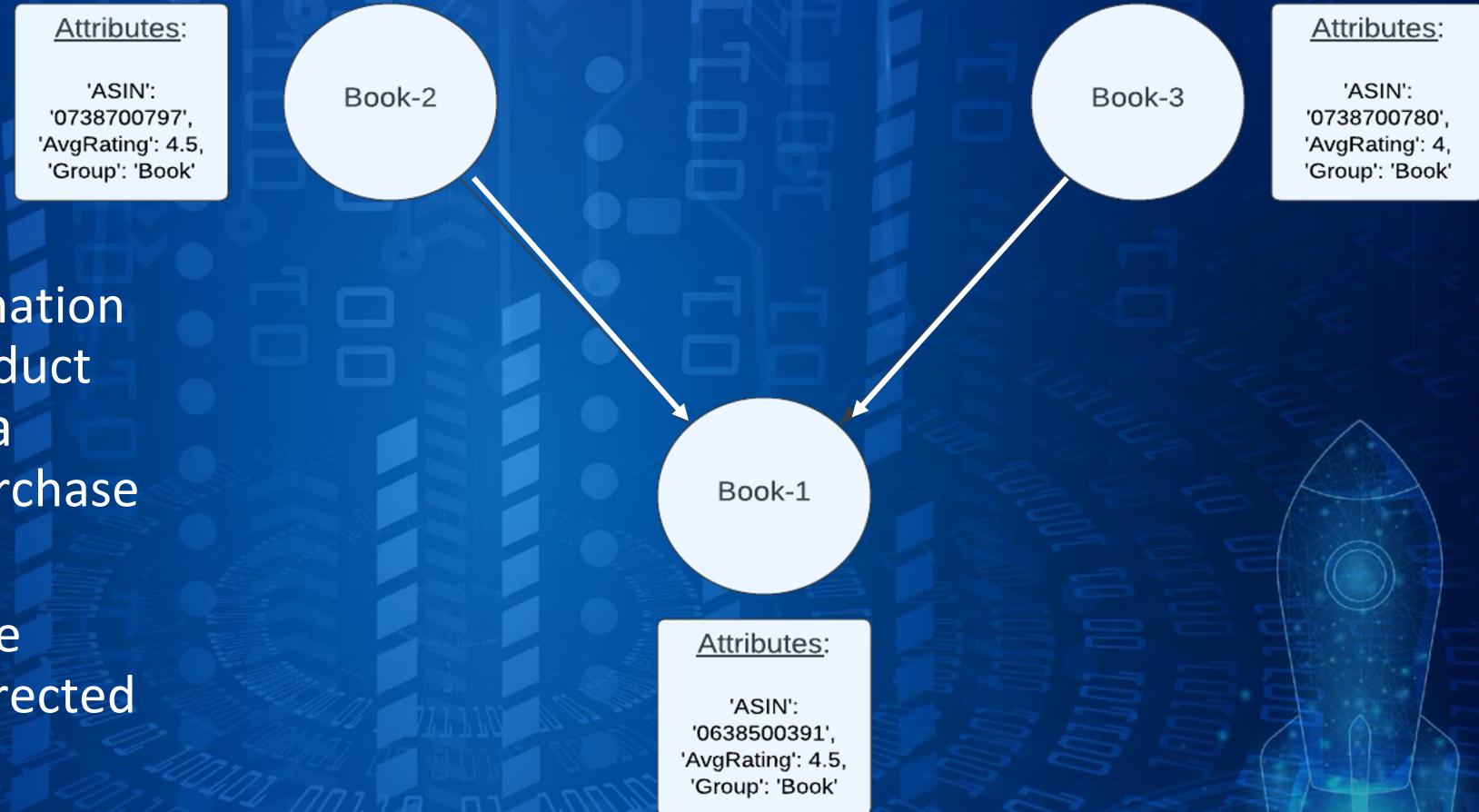
Data Set

- Metadata
 - Contains information about each product
- Transactional data
 - Contains Co-purchase information

Both these data sets are combined to make a directed graph.

No. of Nodes: 4,03,394

No. of Edges: 33,87,388



Solution overview - PhoenixGraph

1. Mobile Application Integration
2. Business Dashboard
3. Janusgraph Performance Analysis
4. Proposed Architecture



Application Demo

API endpoints

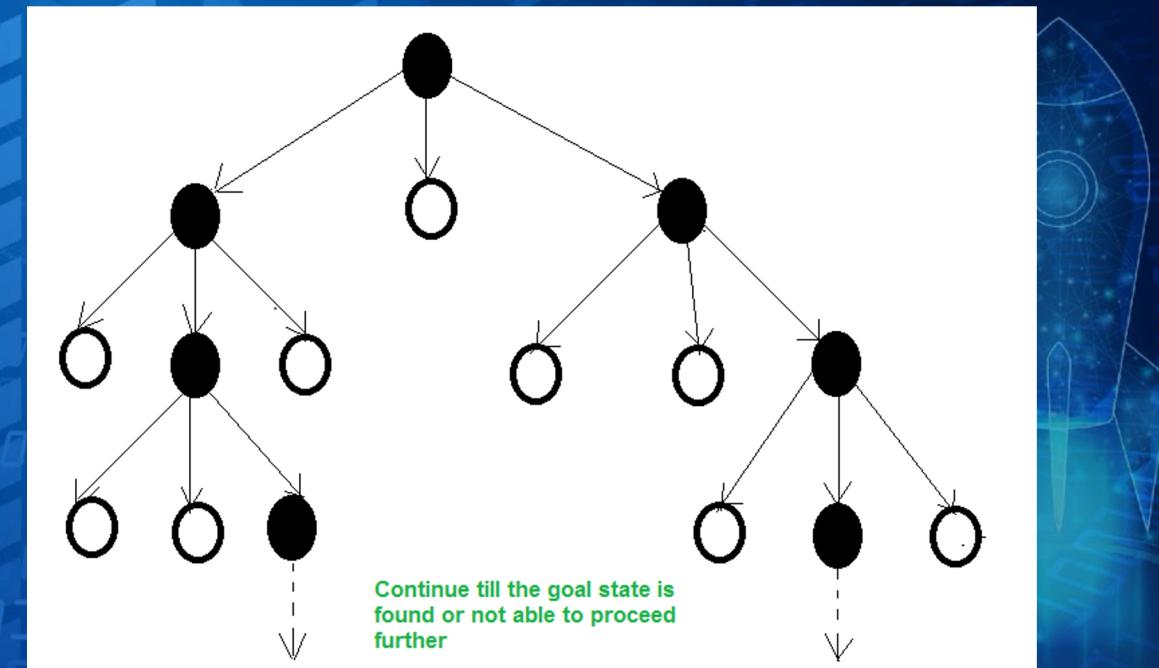
- a. Similar Products
- b. Recommended products
- c. Search



Business Dashboard Demo

API endpoints

- a. Popular Bundles (no. of neighbours, salesrank, average rating)
- b. Category wise Bundles
- c. Product Analysis



JanusGraph Performance Analytics

Monitoring of Janusgraph resource utilization.

- Memory Usage (Runtime)
- CPU Load

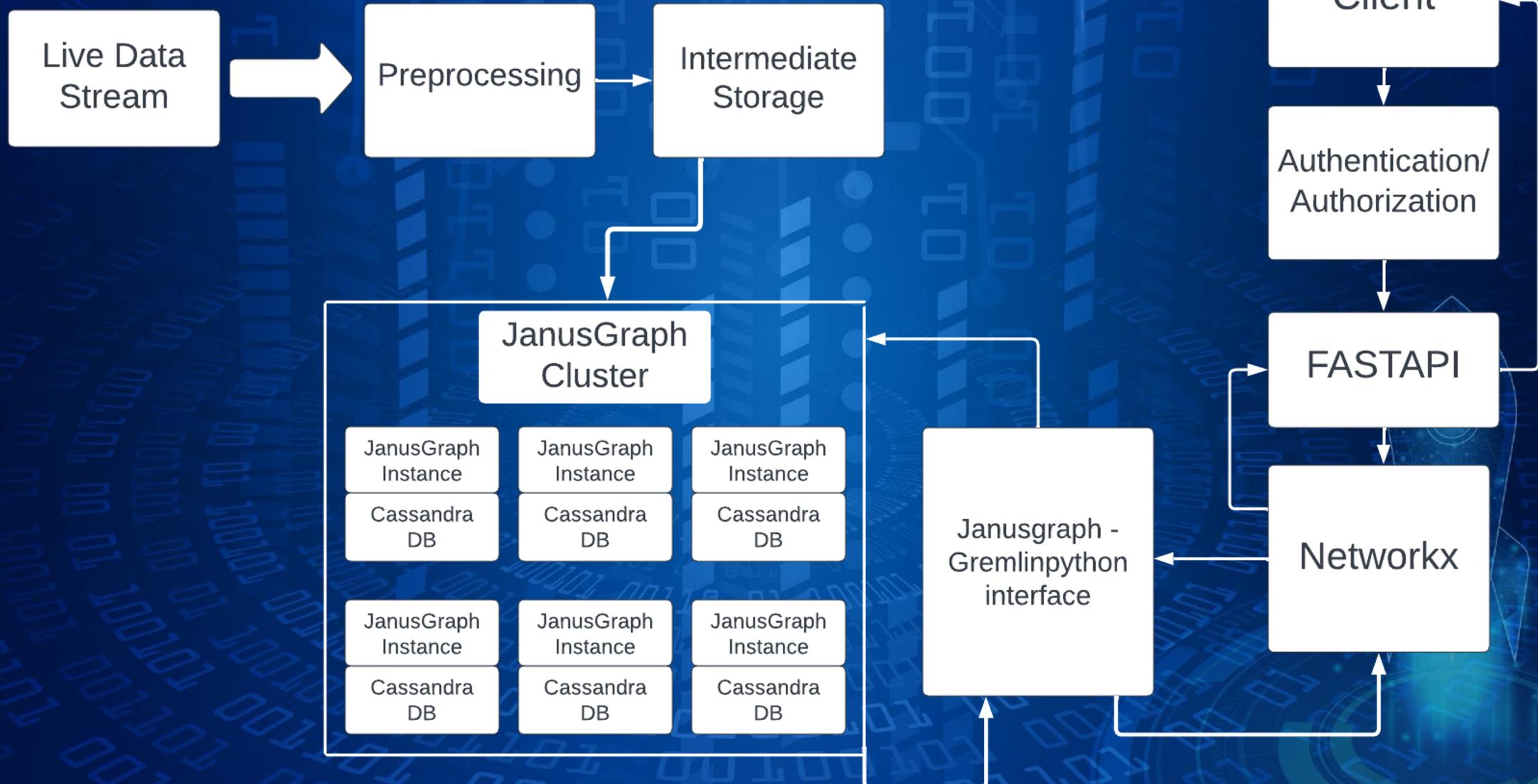
Usage statistics:

- Memory Footprint (Storage)
- Time to Query

Graph Visualization



Proposed architecture



Conclusion

- Unique Selling Point
 - Report generation (Automation)
 - Business dashboard (custom bundling)
 - End to End distributed architecture
 - Cost (Neo4j Enterprise)
- Future scope:
 - Role Based Access Control
 - Dataset aggregation
 - Image based recommendations



Challenges

- Exclusion of good quality products due to lesser number of reviews and co-purchases
- Defining an optimal metric to sort the edges for creation of bundles
- Memory limitations on small scale systems with Janusgraph.



Search...**Fixed Deposit**

Get high returns of up to 8.60% p.a
on your investment

[Open FD](#)

• • •

Recent Purchases:

Dell Laptop
Rs.45000



ABC Recliner
Rs.25000



LG Refrigerator
Rs.80000



Oppo 11 pro
Rs.19000



Smart Watch X
Rs.12000



Ibanez Guitar
Rs.80000



Products you may like



Espresso Beans

Rs.8000



Matress

Rs.2000



Wireless Speaker

Rs.30000



GO Pro Hero09

Rs.3600



Ergonomic Chair

Rs.10000



Earphone BT

Rs.25000



Dashboard

Search...

John Doe
Admin

Bajaj

Dashboard

Order

Products

Settings

Popular Bundles



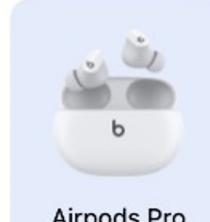
Phone Case



iPhone 11



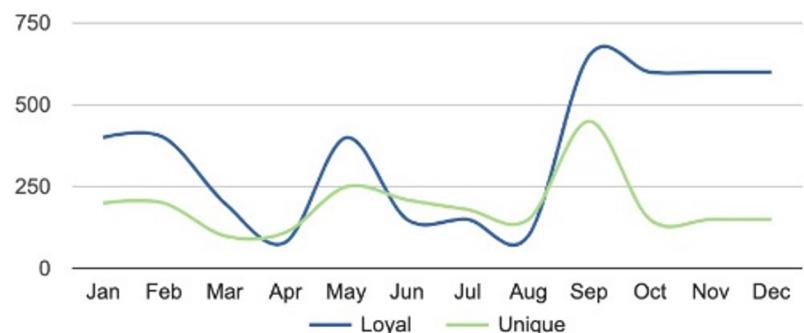
Charger



Airpods Pro

Electronics

Activities

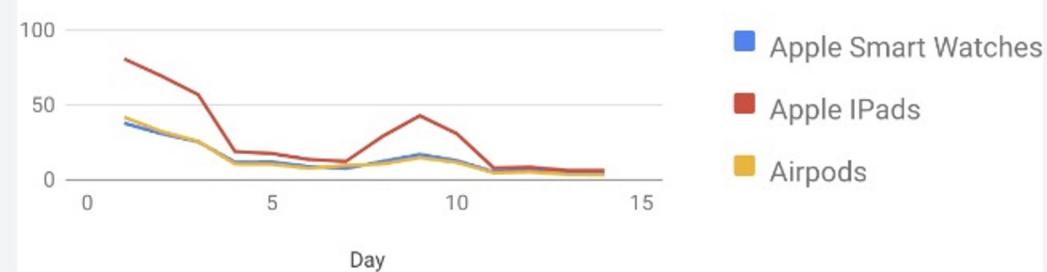


Top Products

PRODUCT NAME	POPULARITY	CATEGORY	PRICE
Apple Smart Watches	45%	Laptop	Rs. 150000
Apple iPads	35%	Laptop PC	Rs. 75000
Airpods	15%	Accessories	Rs. 25000

Sales

in millions of dollars (USD)



Logout

JanusGraph

Load Data:

Files from SNAP Amazon Copurchase dataset can be directly imported

Node Attributes (CSV File)



Drag and drop file here

Limit 200MB per file • CSV

Browse files



finalpreprocesseddata.csv 37.5MB



EdgeList (TSV)



Drag and drop file here

Limit 200MB per file • TXT

Browse files

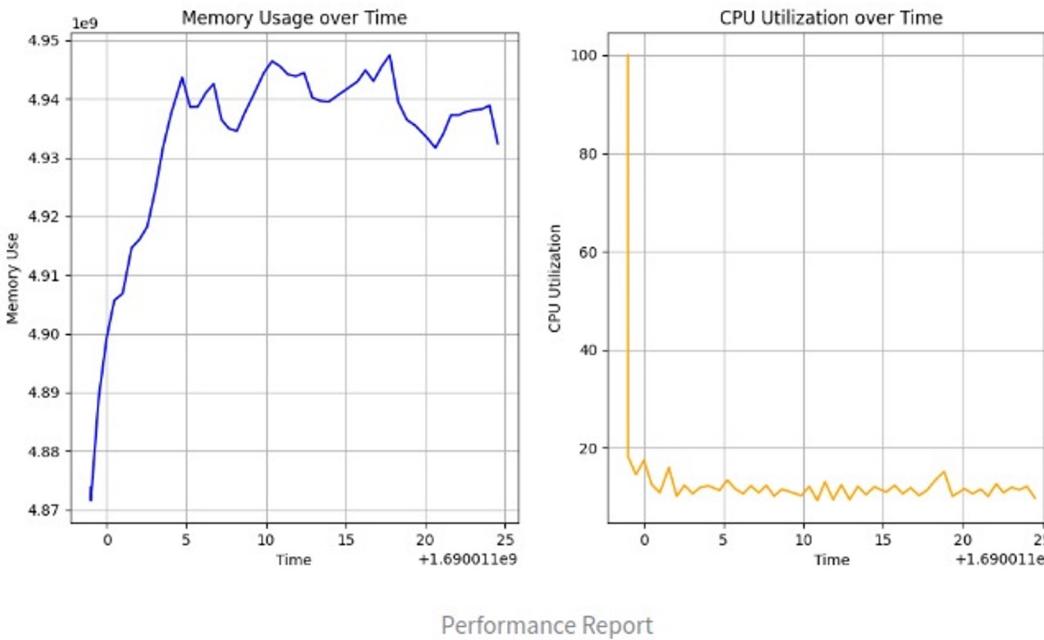


amazon0601.txt 47.9MB



Graph Creation:

Server Configuration: 1024MB Buffer size, 2GB memory size



Performance Report:

Total time elapsed: `25.545371532440186` sec

Total Nodes Inserted: `50` (out of 3.34 Lakh)

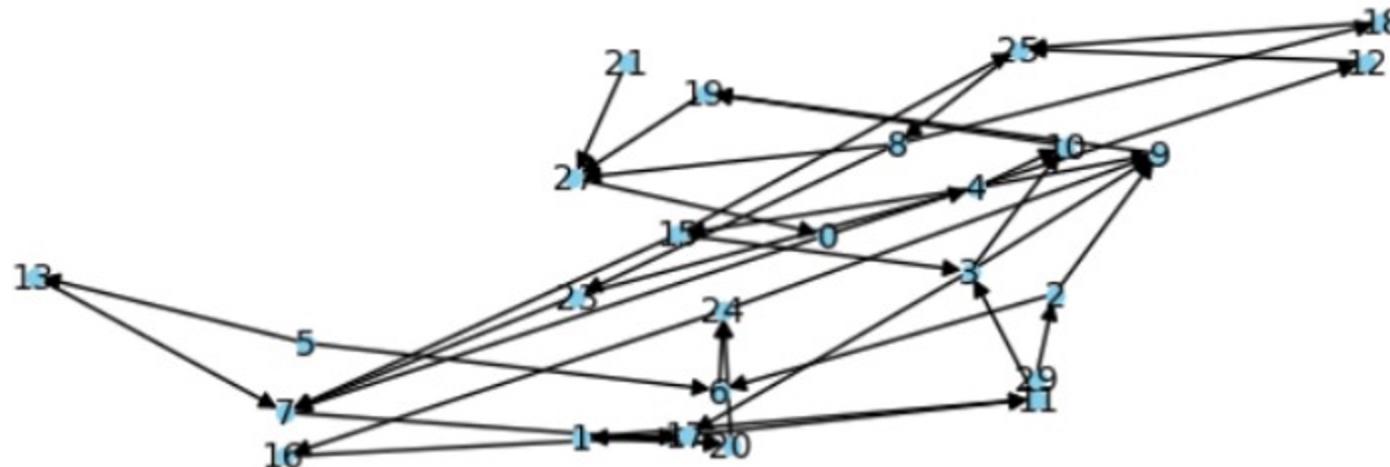
Total memory used: `58.564608` MB

Expected memory need for 3.34L nodes: `391.21158144` GB

Time to Query entire graph: `0.025698423385620117` sec



Visualization:



Graph Visualization