SNIA COMPUTE + MEMORY + STORAGE SUMMIT

Architectures, Solutions, and Community
VIRTUAL EVENT, APRIL 11-12, 2023

Decomposing Compute to Grow Computational Storage

Presented by Aldrin Montana

Team

Aldrin Montana | Jeff LeFevre | Carlos Maltzahn | Peter Alvaro

Computer Science

Bianca Xue | Josh Stuart

Systems biology

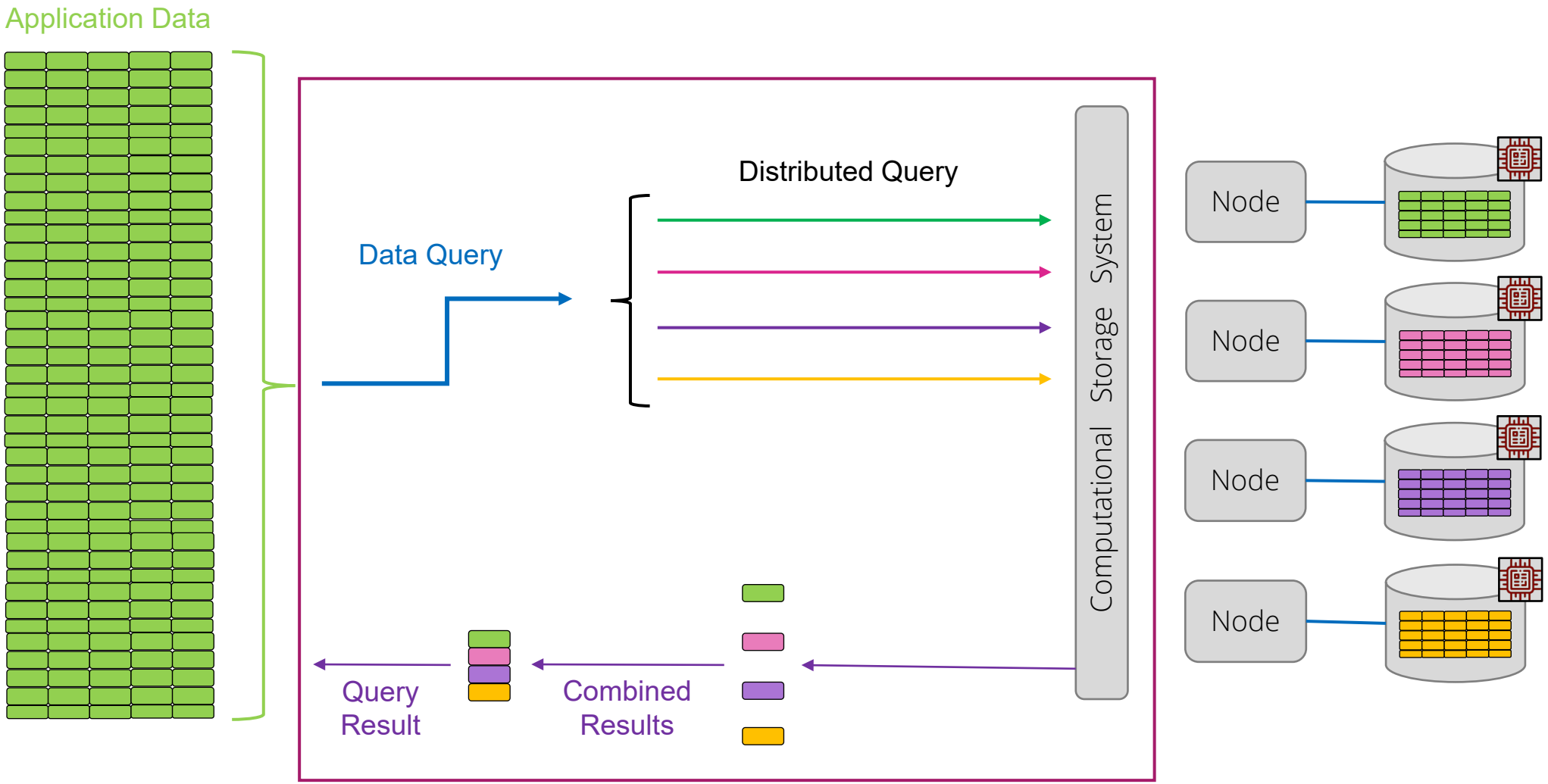
Philip Kufeldt

Storage Technologist

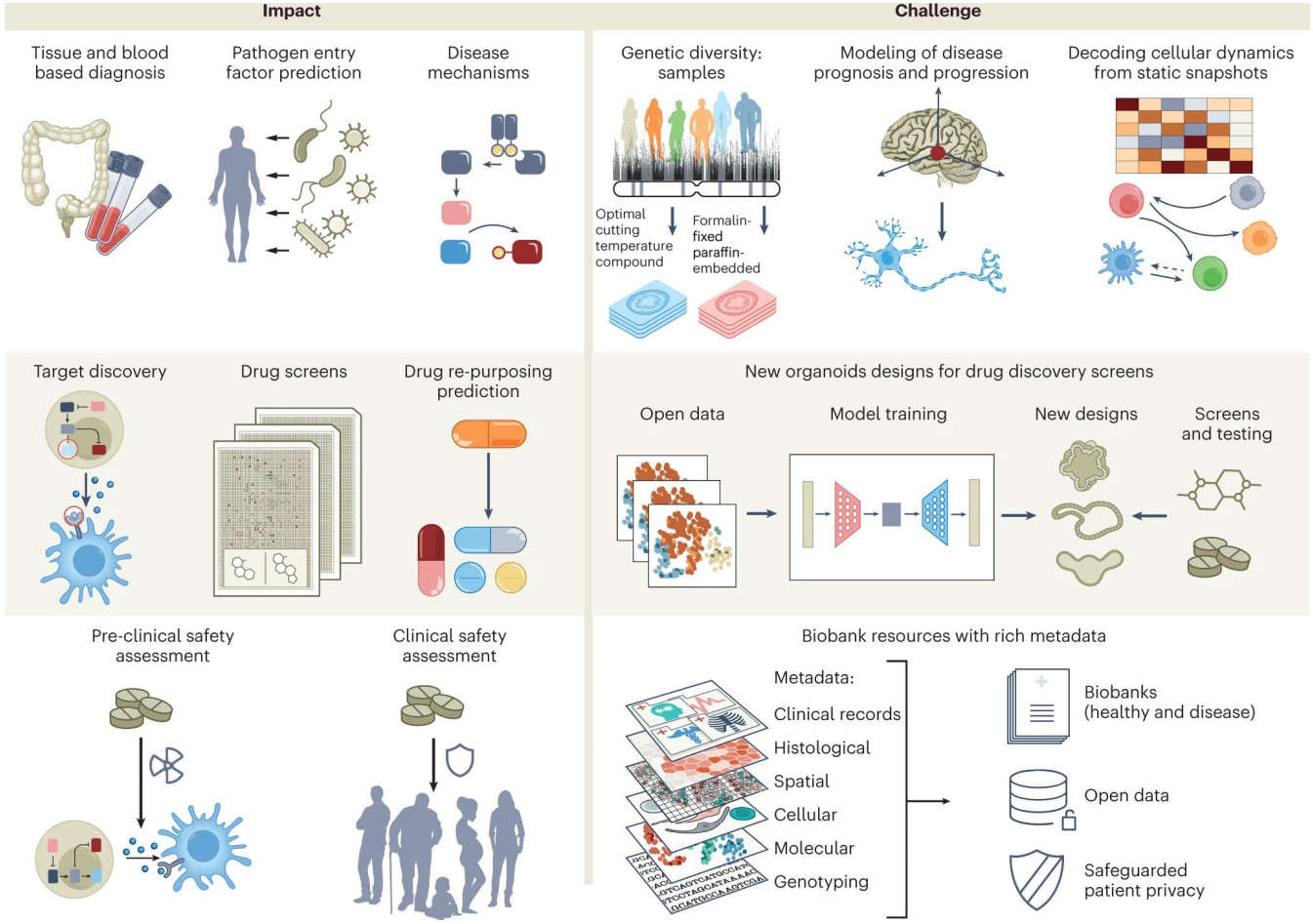
UC SANTA CRUZ

SEAGATE

Overview



Scientific Use Case



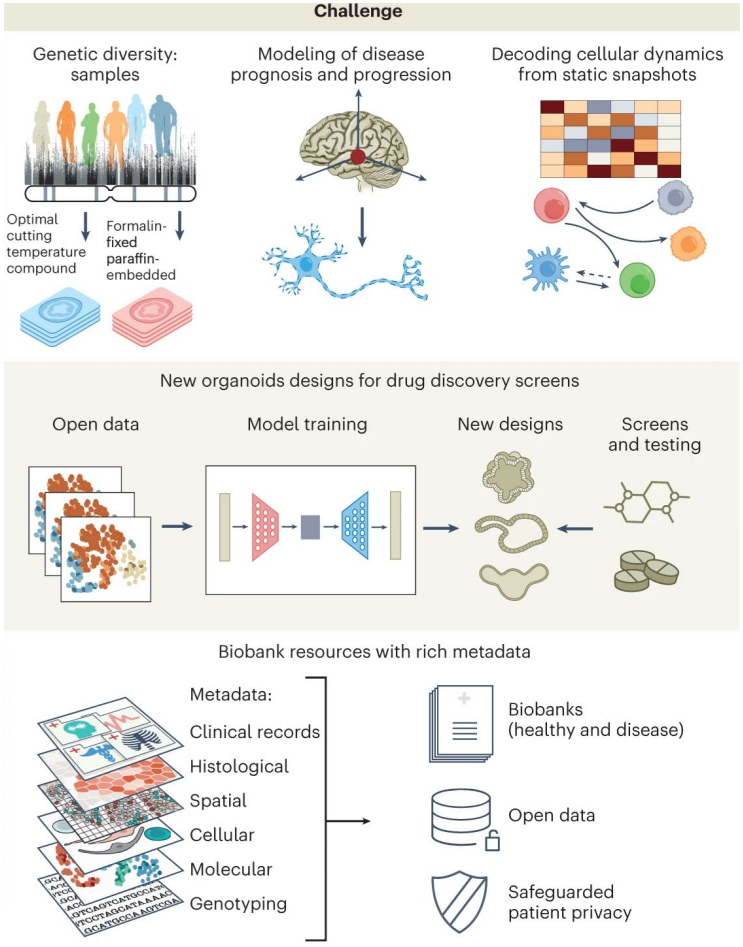
HUMAN CELL ATLAS

An international collaborative consortium

- charts the cell types in the healthy body
- across time from development
  to adulthood
  (eventually) to old age

Rood, J.E., Maartens, A., Hupalowska, A. *et al.* Impact of the Human Cell Atlas on medicine. *Nat Med* **28**, 2486–2496 (2022). https://doi.org/10.1038/s41591-022-02104-7

Scientific Use Case



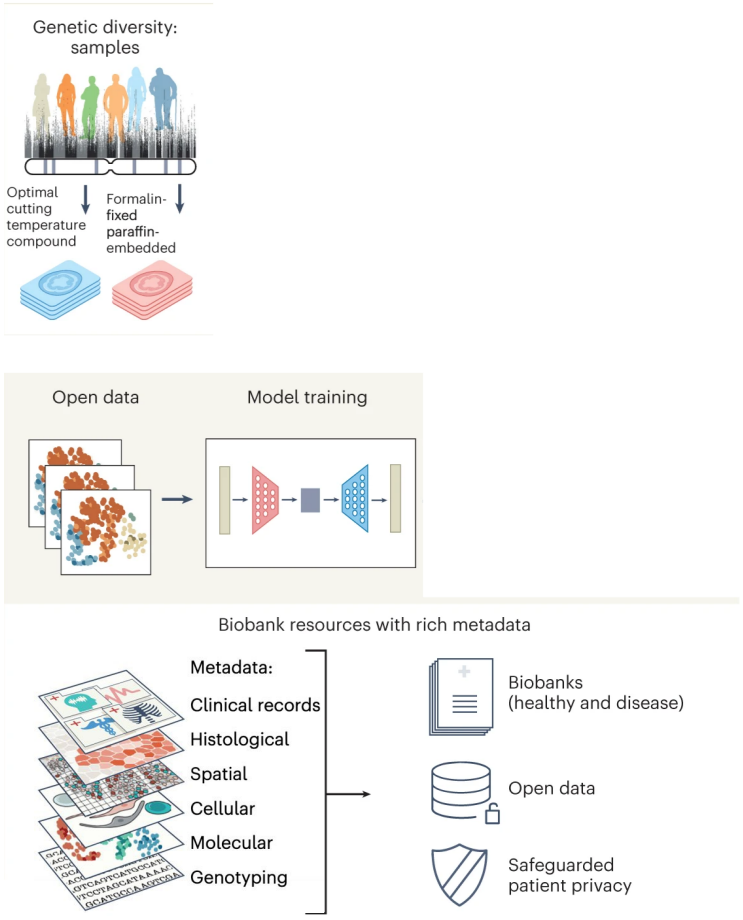
HUMAN CELL ATLAS

An international collaborative consortium

- charts the cell types in the healthy body
- across time from development
  to adulthood
  (eventually) to old age

Rood, J.E., Maartens, A., Hupalowska, A. *et al.* Impact of the Human Cell Atlas on medicine. *Nat Med* **28**, 2486–2496 (2022). https://doi.org/10.1038/s41591-022-02104-7

Scientific Use Case



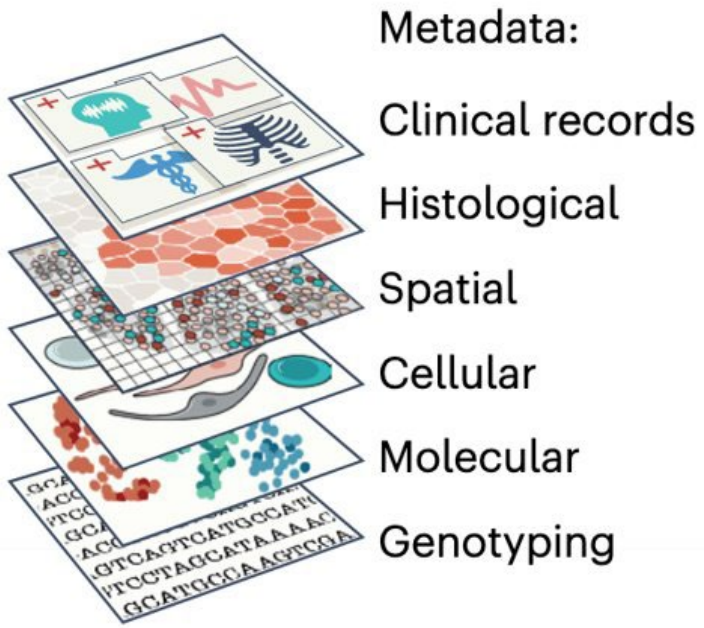
HUMAN CELL ATLAS

An international collaborative consortium

- charts the cell types in the healthy body
- across time from development
  to adulthood
  (eventually) to old age

Rood, J.E., Maartens, A., Hupalowska, A. *et al.* Impact of the Human Cell Atlas on medicine. *Nat Med* **28**, 2486–2496 (2022). https://doi.org/10.1038/s41591-022-02104-7

Scientific Use Case



HUMAN CELL ATLAS

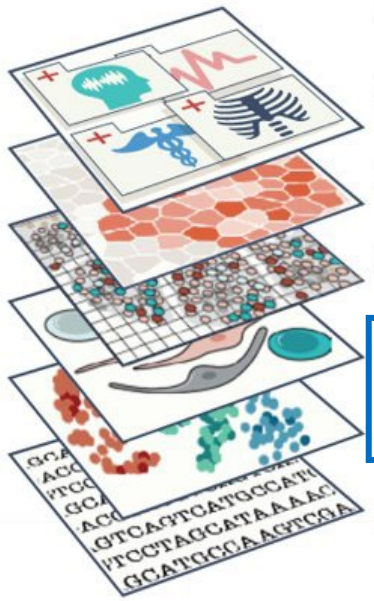
An international collaborative consortium

- charts the cell types in the healthy body
- across time from development
  to adulthood
  (eventually) to old age

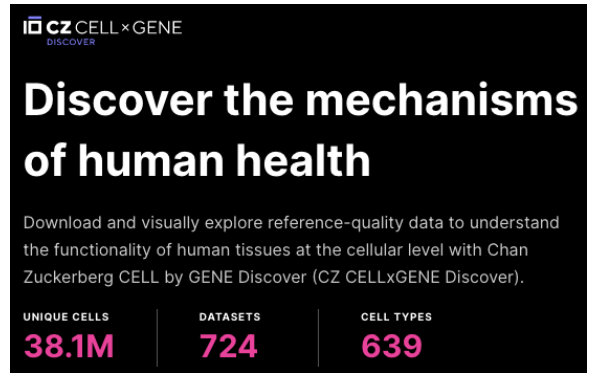
Rood, J.E., Maartens, A., Hupalowska, A. *et al.* Impact of the Human Cell Atlas on medicine. *Nat Med* **28**, 2486–2496 (2022). https://doi.org/10.1038/s41591-022-02104-7

Scientific Use Case

Chan Zuckerberg Initiative



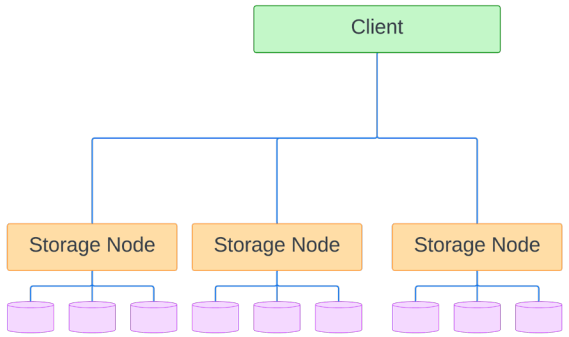
Metadata:	Magnitude
Clinical records	TiB
Histological	TiB
Spatial	TiB
Cellular	TiB
Molecular	TiB
Genotyping	GiB

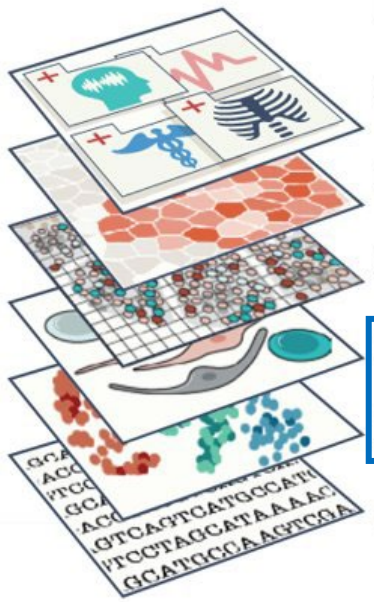


Cells	Unique Proteins	Size
1 M	25 K	0.1 TB
100 M	25 K	10 TB
1 B	25 K	100 TB
10 B	25 K	1000 TB
10 B	60 K	2400 TB

Rood, J.E., Maartens, A., Hupalowska, A. *et al.* Impact of the Human Cell Atlas on medicine. *Nat Med* **28**, 2486–2496 (2022). https://doi.org/10.1038/s41591-022-02104-7

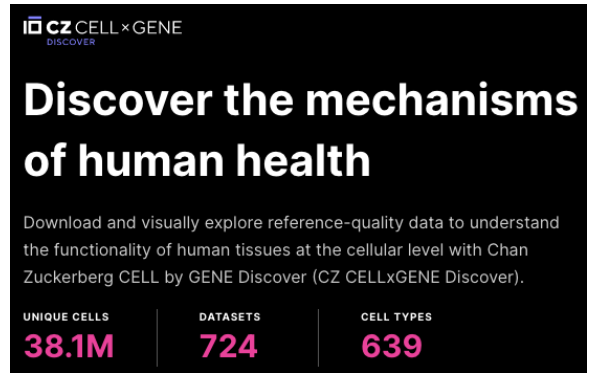
Scientific Use Case





Metadata:	Magnitude
Clinical records	TiB
Histological	TiB
Spatial	TiB
Cellular	TiB
Molecular	TiB
Genotyping	GiB

Chan Zuckerberg Initiative

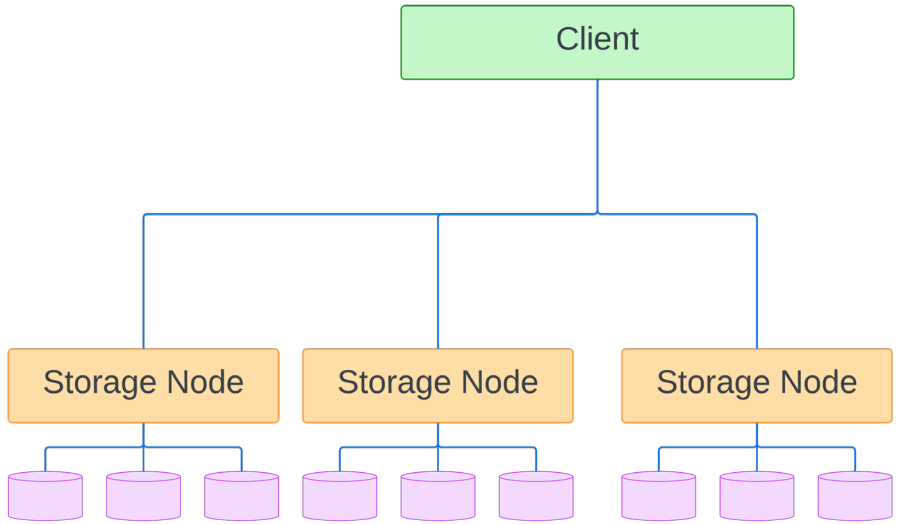


Cells	Unique Proteins	Size
1 M	25 K	0.1 TB
100 M	25 K	10 TB
1 B	25 K	100 TB
10 B	25 K	1000 TB
10 B	60 K	2400 TB

Rood, J.E., Maartens, A., Hupalowska, A. *et al.* Impact of the Human Cell Atlas on medicine. *Nat Med* **28**, 2486–2496 (2022). https://doi.org/10.1038/s41591-022-02104-7

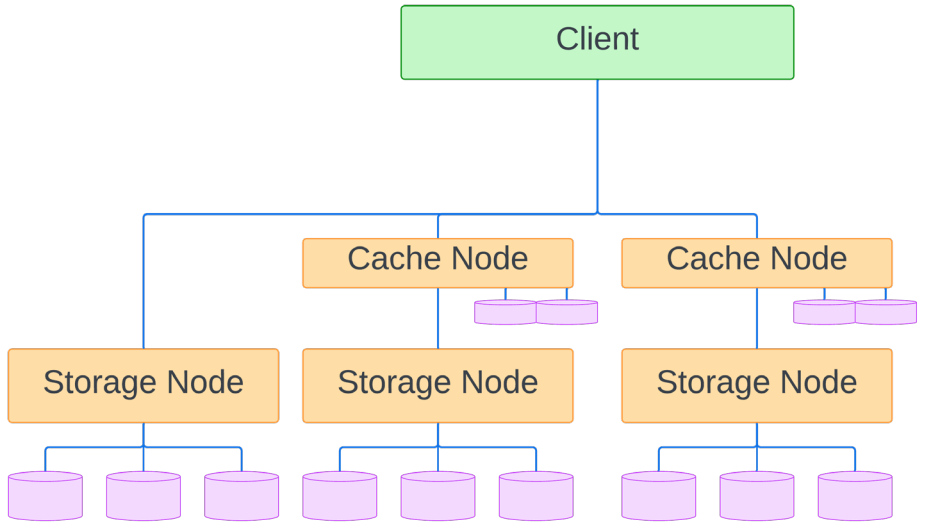
Computational Storage Environment

- Storage hierarchy



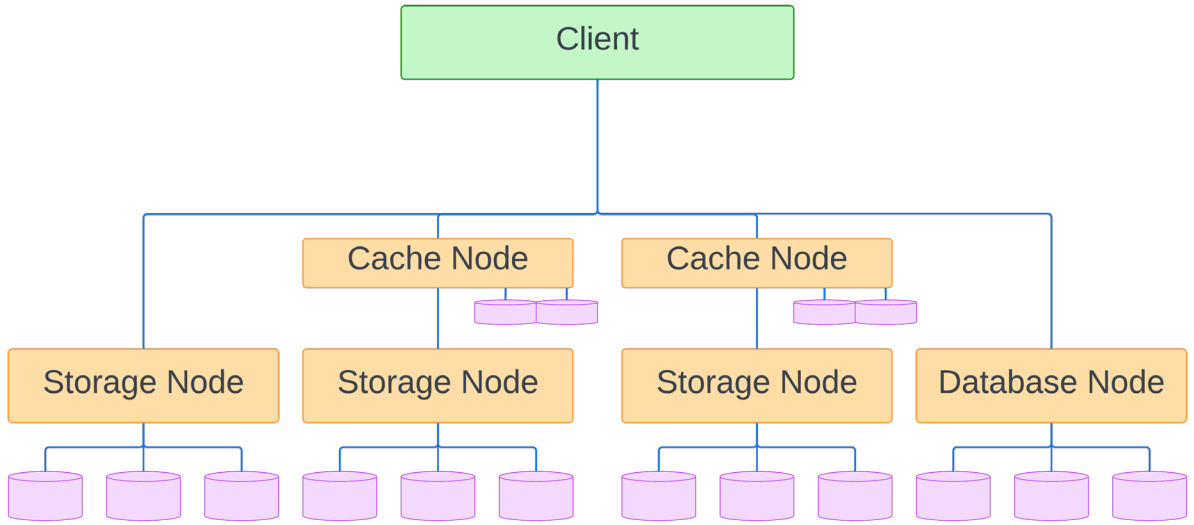
Computational Storage Environment

- Storage hierarchy
  - May become deeper



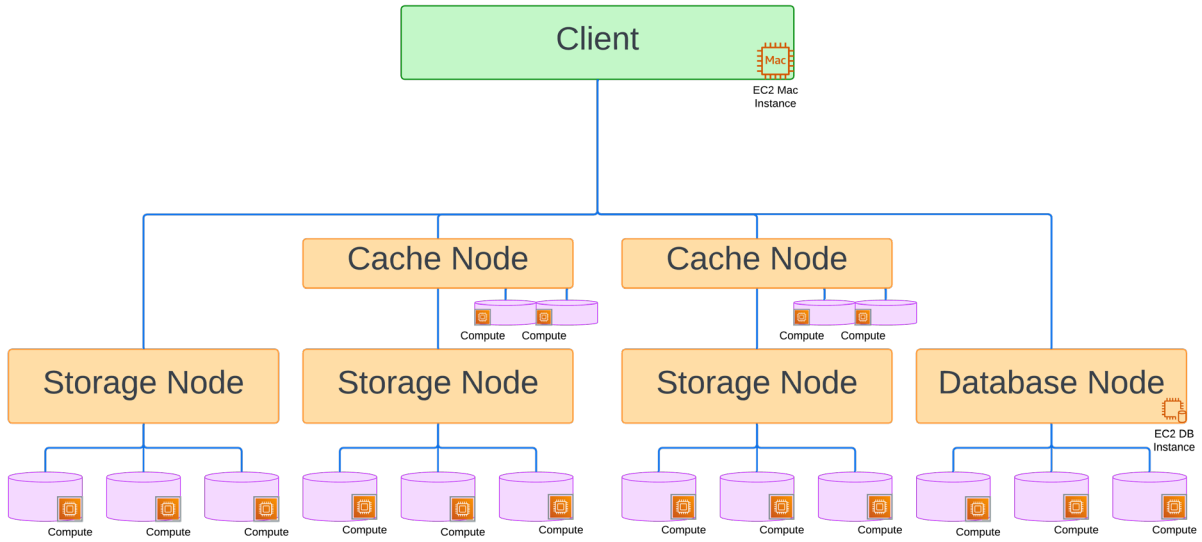
Computational Storage Environment

- Storage hierarchy
  - May become deeper
  - May become wider



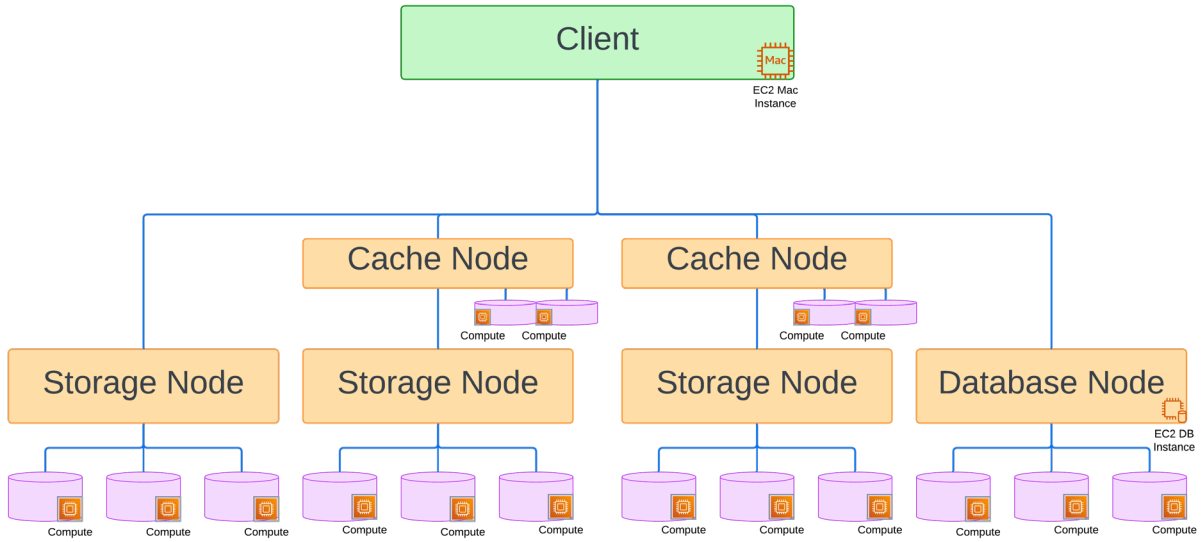
Computational Storage Environment

- Storage hierarchy
  - May become deeper
  - May become wider
  - Will gain more processors
    - And gain more heterogeneity



Computational Storage Environment

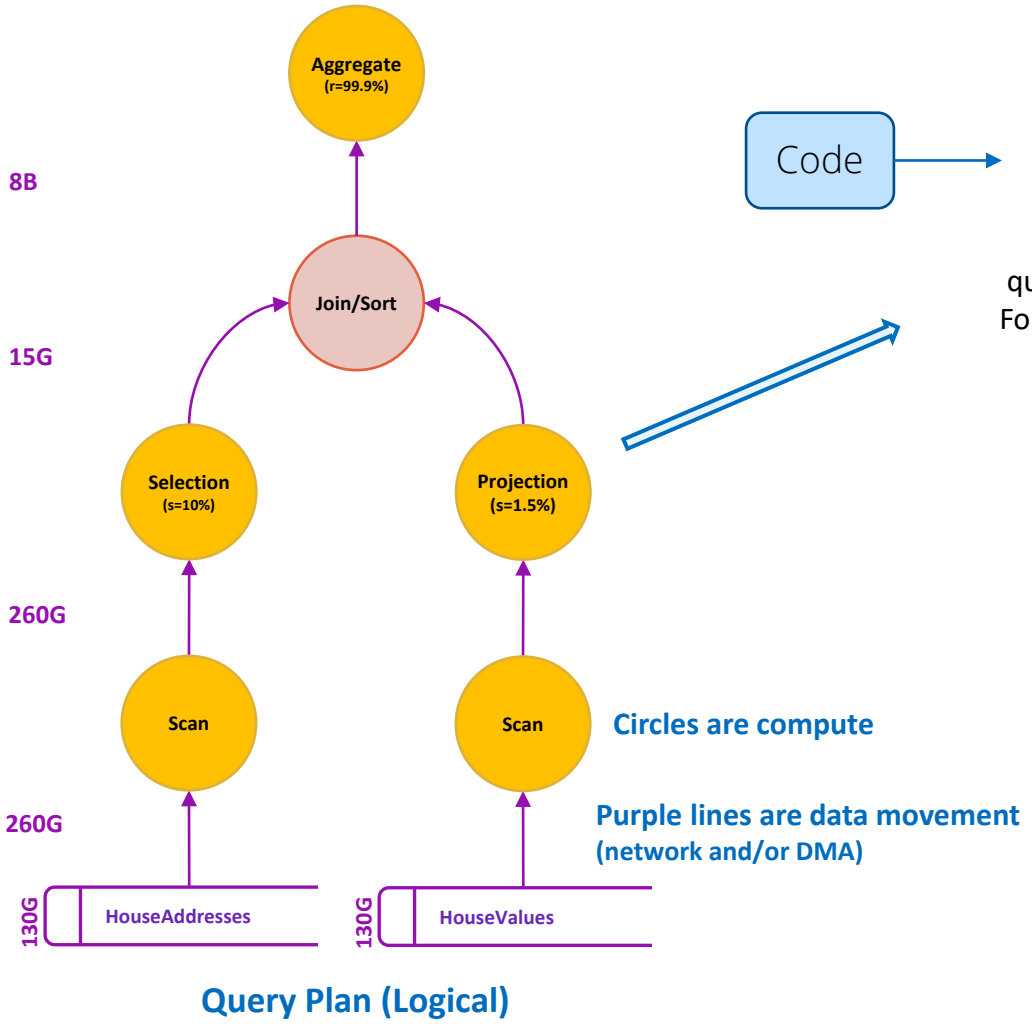
- Storage hierarchy
  - May become deeper
  - May become wider
  - Will gain more processors
    - And gain more heterogeneity

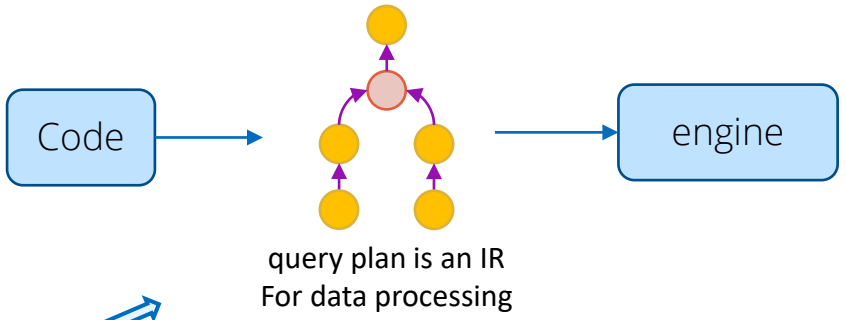


We are in need of a
Computational IO Stack

Decomposable Queries

Illustration – Query Plan



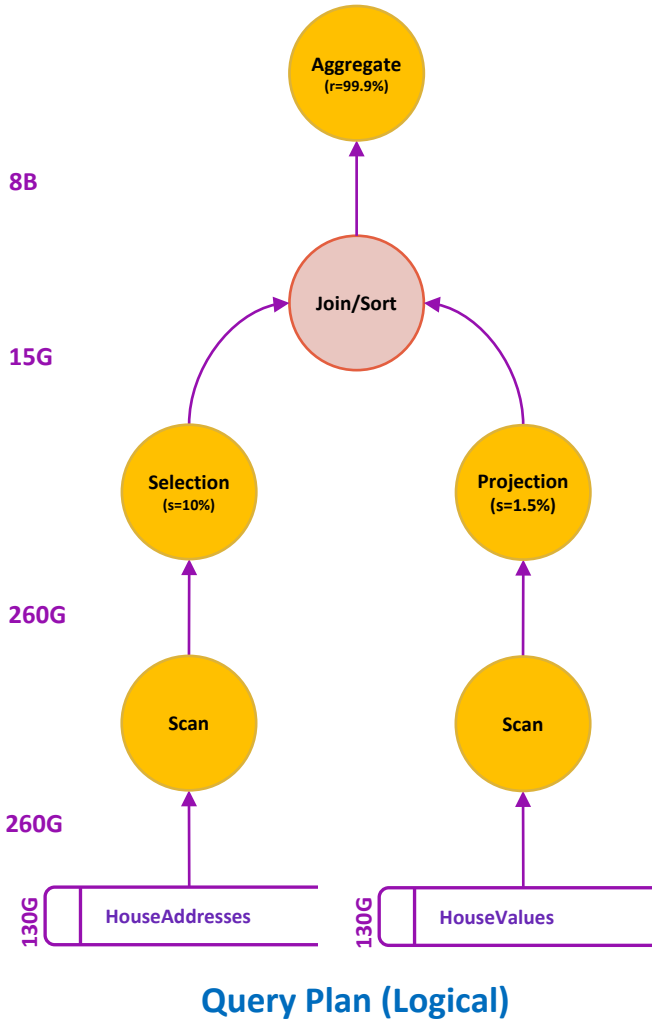


Question: What zip code in CA has the highest average value of houses?

[140M House in US; 14M in CA; 1000B/record]

Illustration – Storage Hierarchy

Question: What zip code in CA has the highest average value of houses?
[140M House in US; 14M in CA; 1000B/record]



Query Plan (Logical)

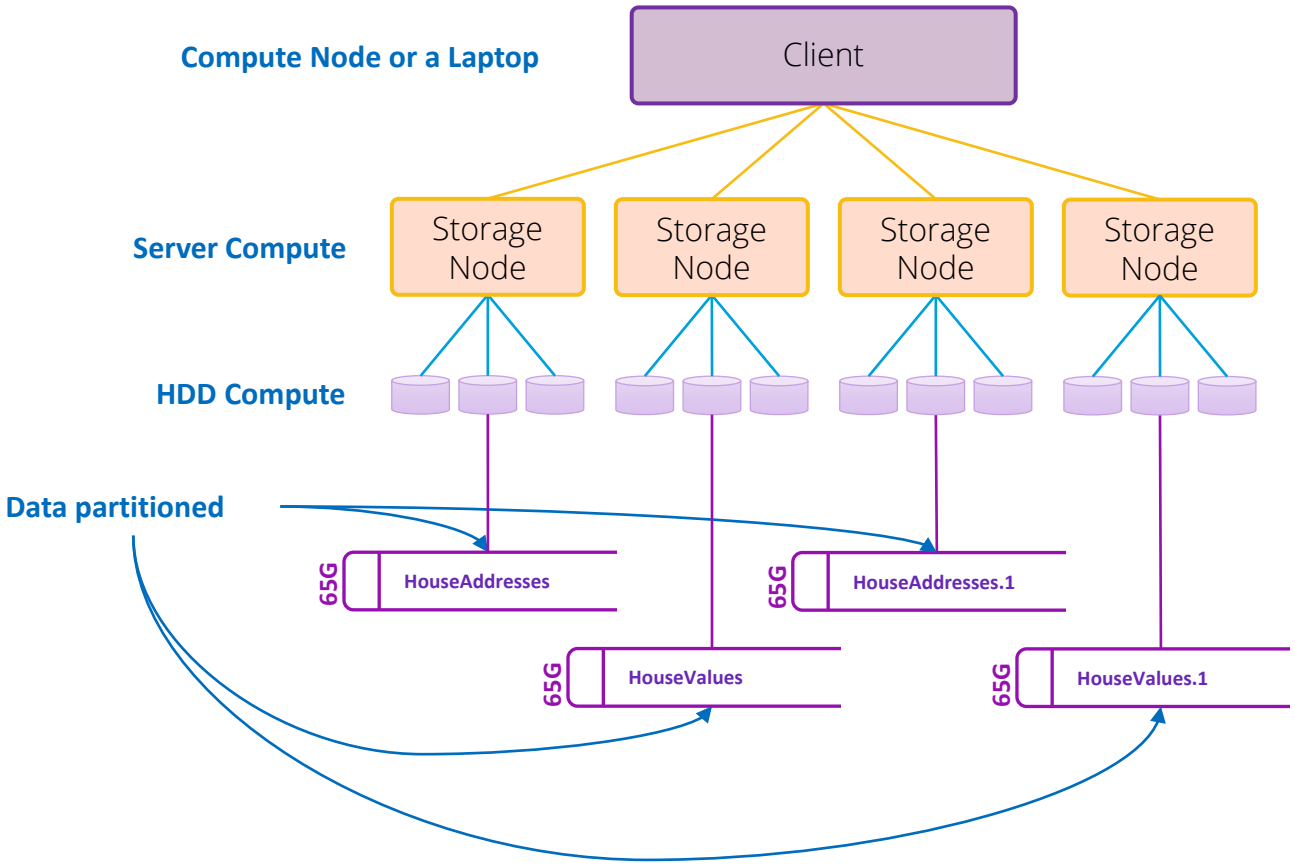
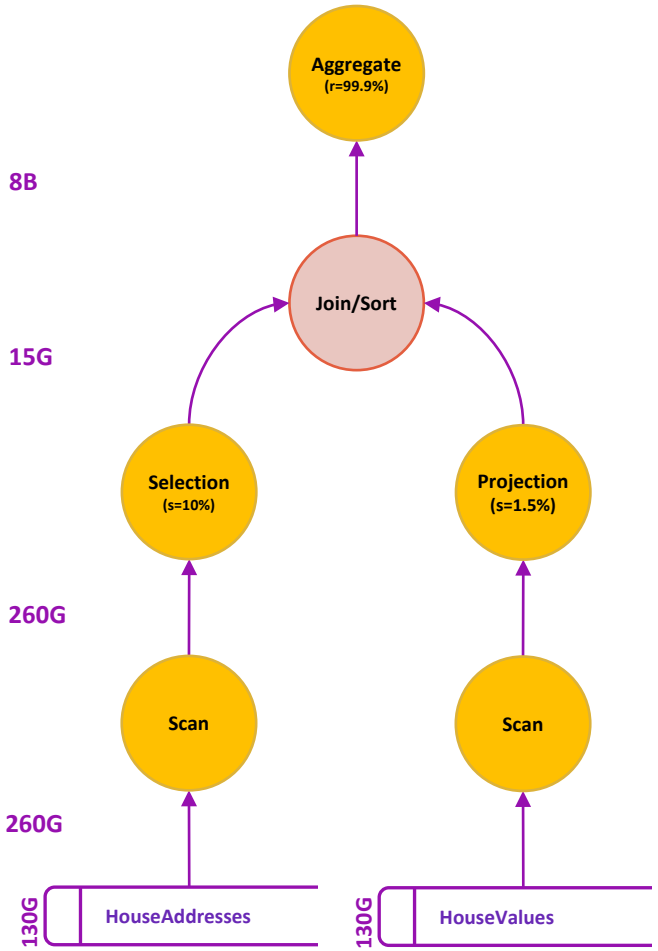


Illustration – Storage Hierarchy

Question: What zip code in CA has the highest average value of houses?
[140M House in US; 14M in CA; 1000B/record]



Query Plan (Logical)

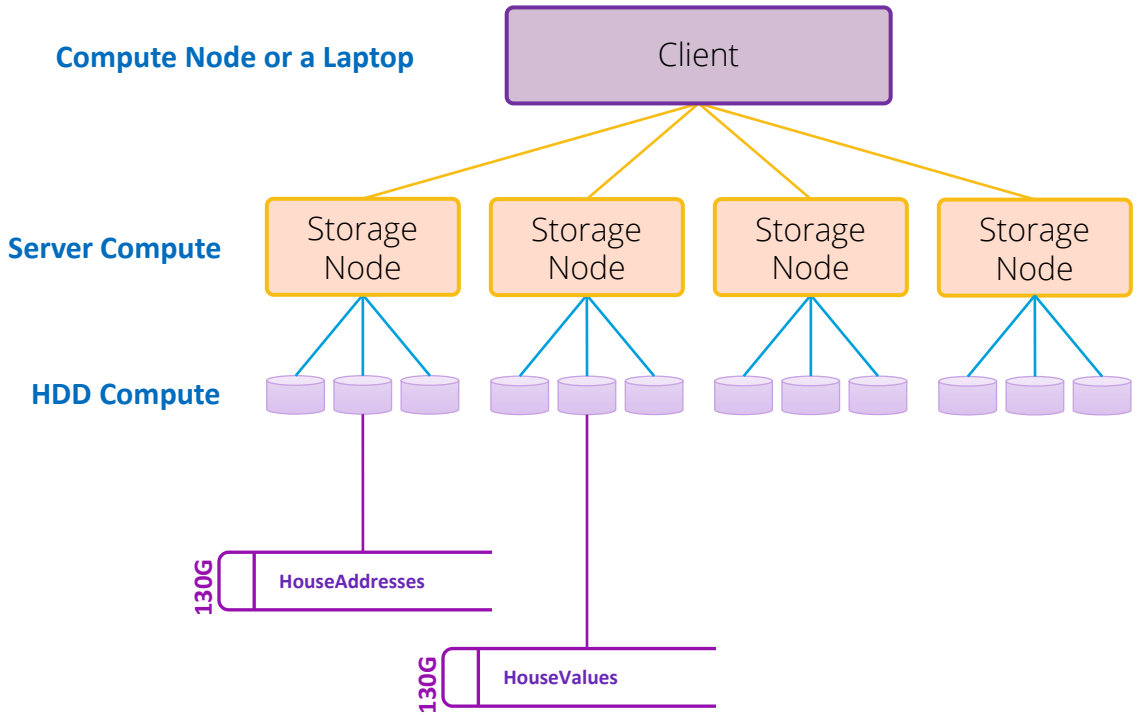
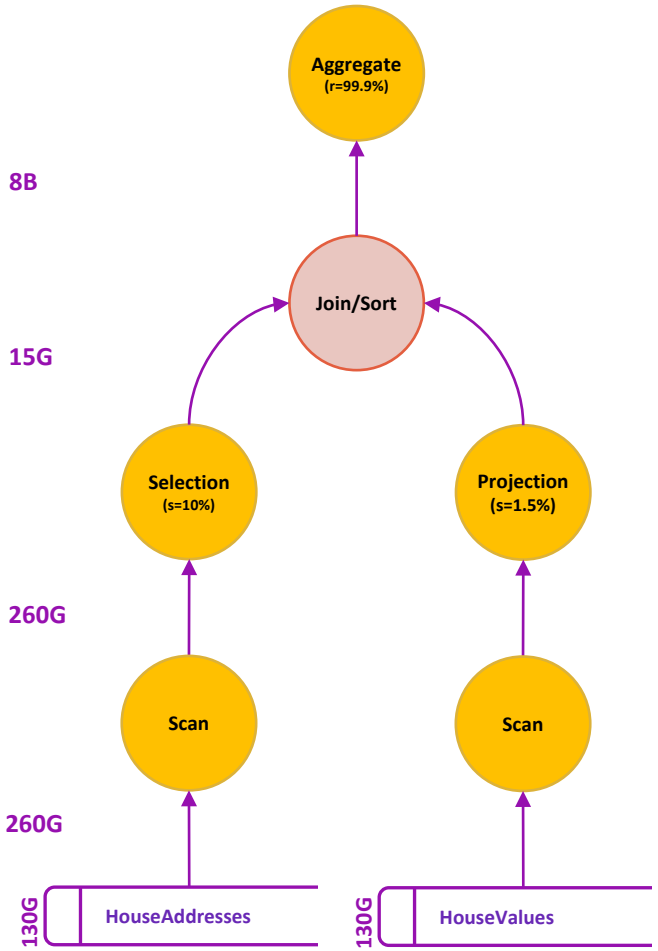
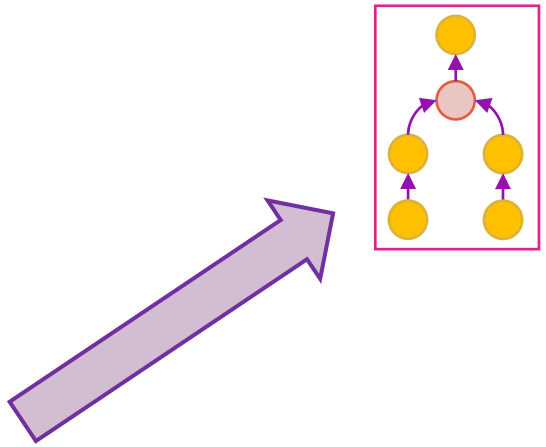


Illustration – Execution (Client)

Question: What zip code in CA has the highest average value of houses?
[140M House in US; 14M in CA; 1000B/record]



Query Plan (Logical)



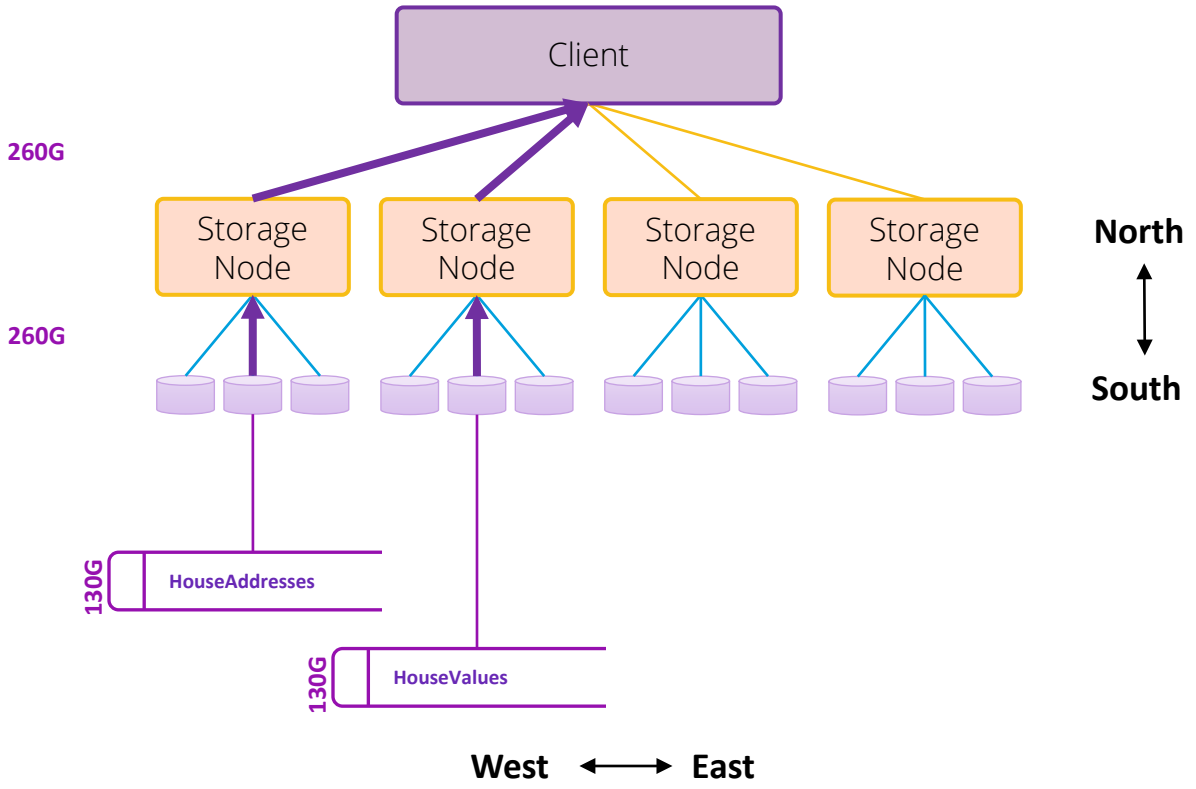
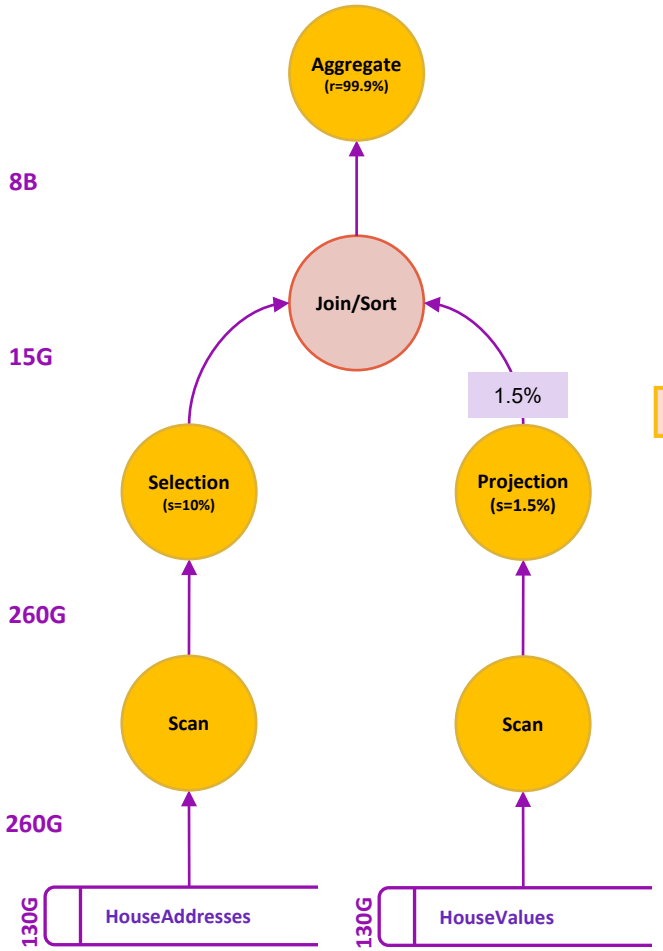
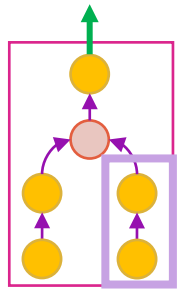


Illustration – Execution (Servers)

Question: What zip code in CA has the highest average value of houses?
[140M House in US; 14M in CA; 1000B/record]



Query Plan (Logical)



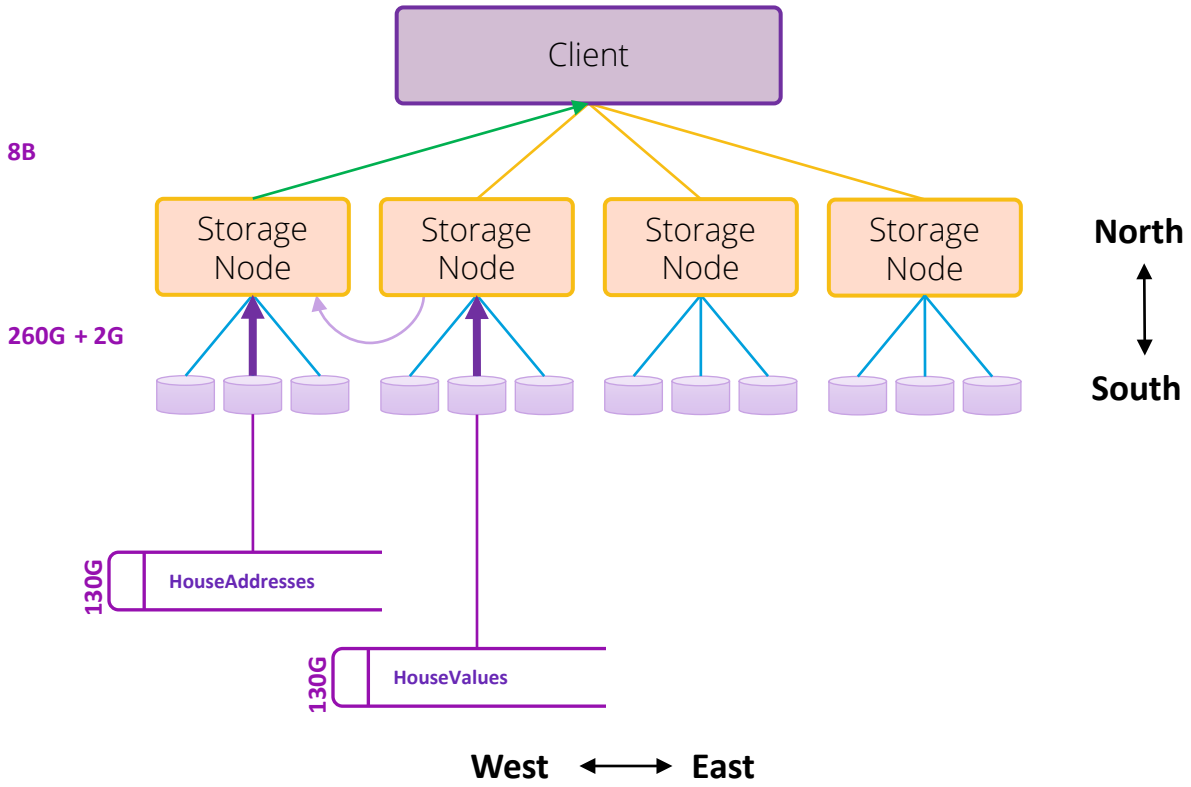
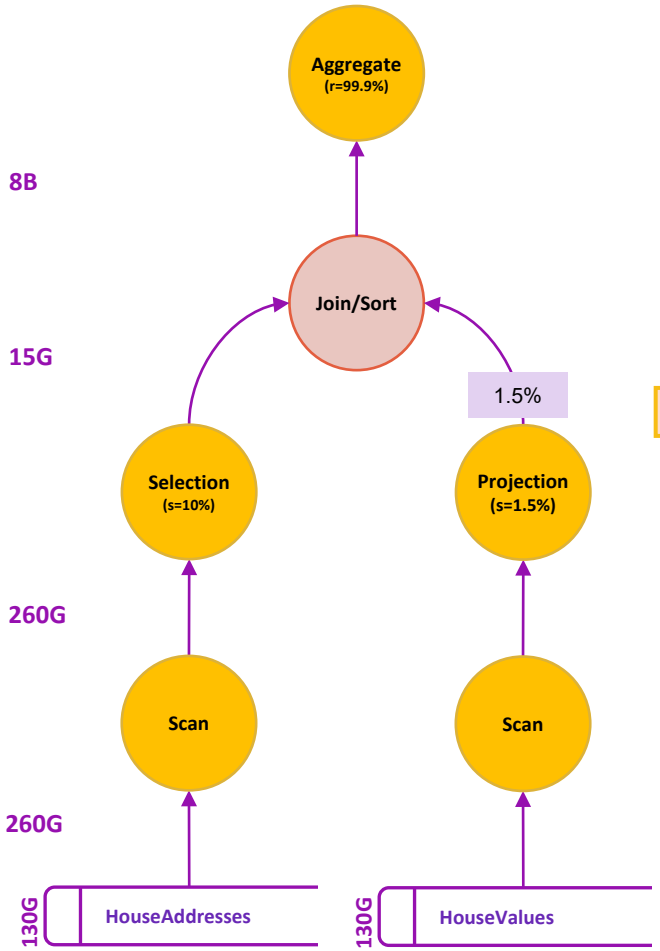
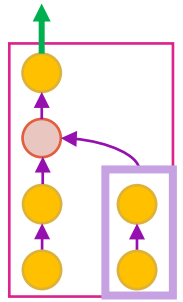


Illustration – Execution (Servers)

Question: What zip code in CA has the highest average value of houses?
[140M House in US; 14M in CA; 1000B/record]



Query Plan (Logical)



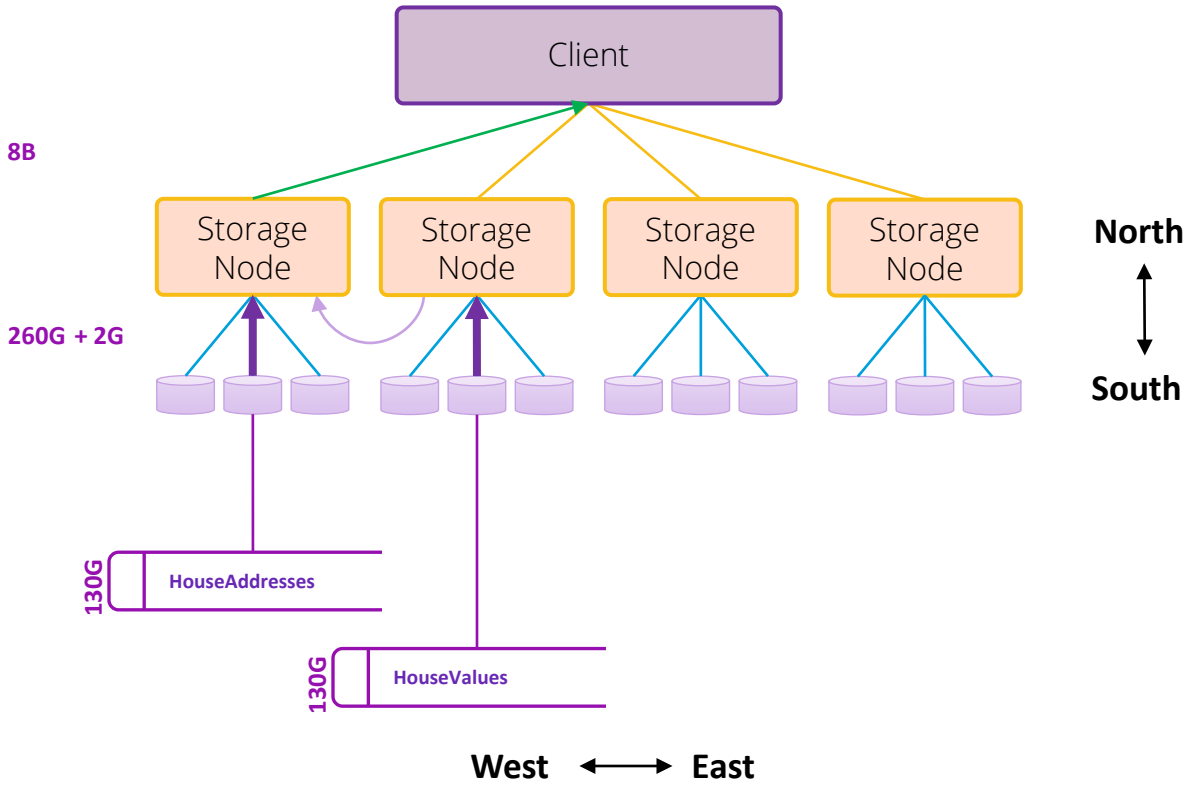
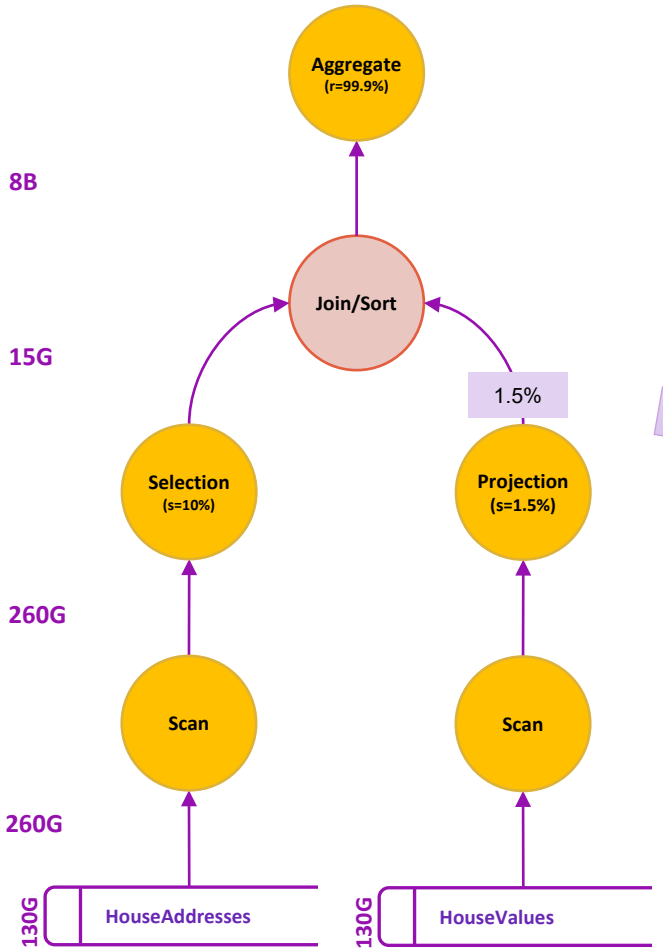
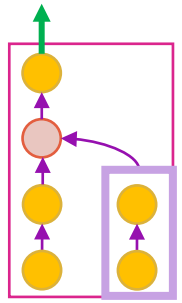


Illustration – Execution (Drives)

Question: What zip code in CA has the highest average value of houses?
[140M House in US; 14M in CA; 1000B/record]



Query Plan (Logical)



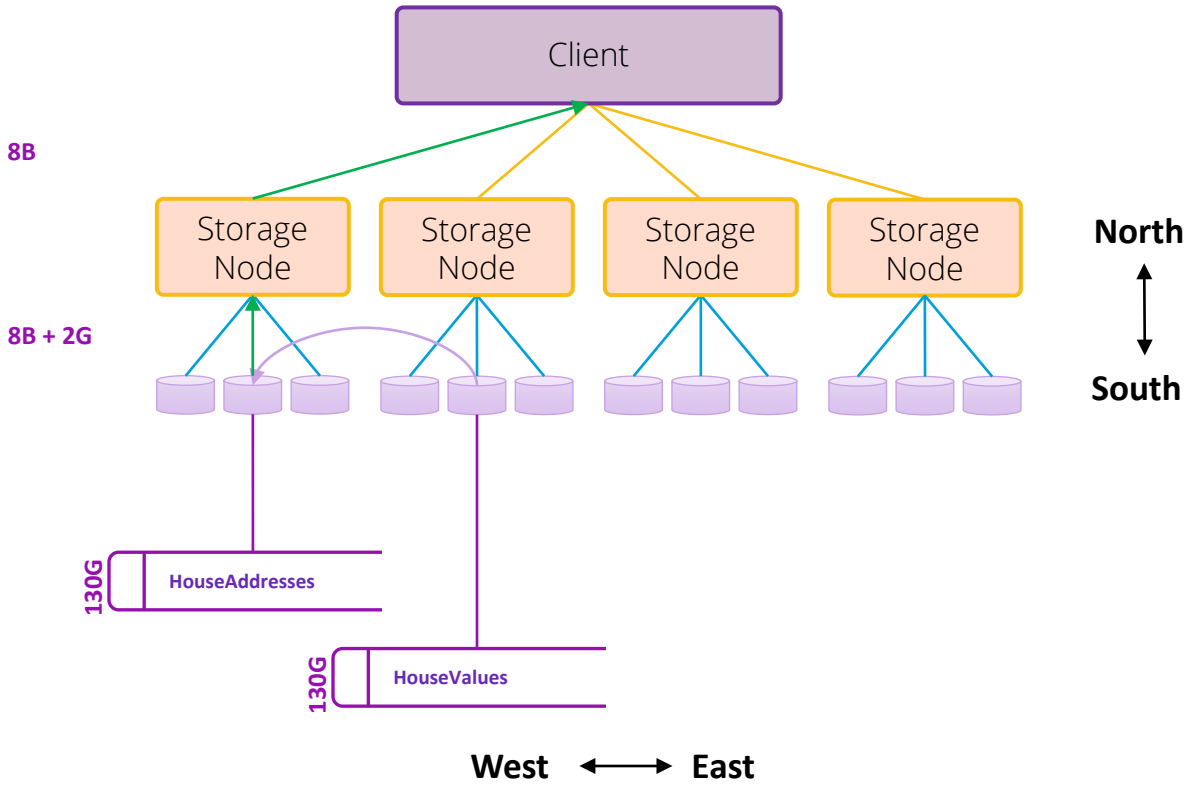
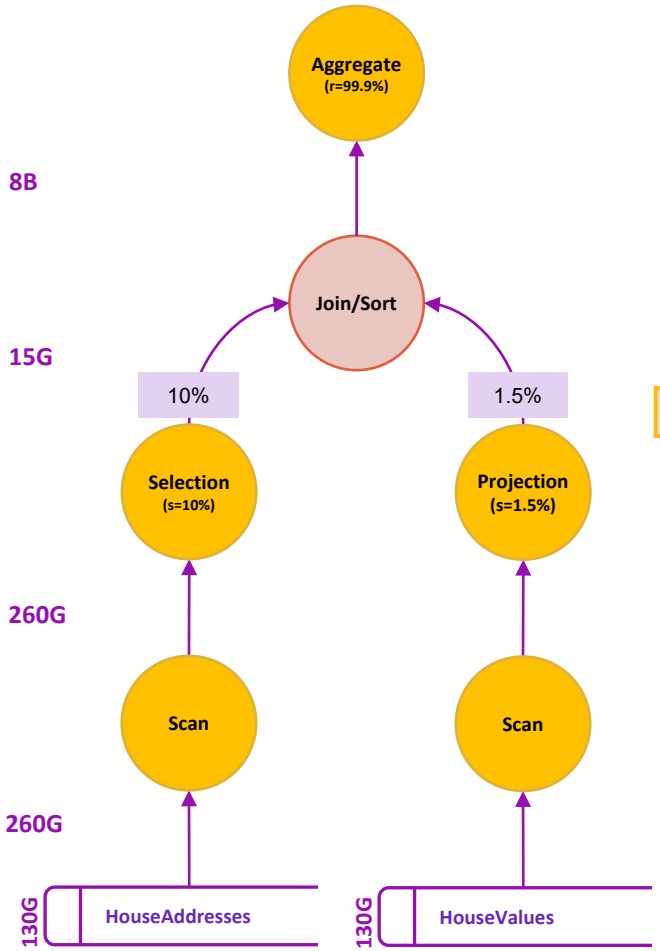
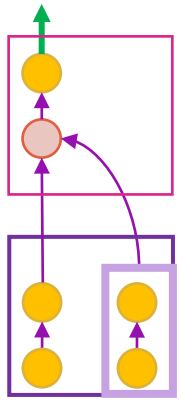


Illustration – Execution (Mixed)

Question: What zip code in CA has the highest average value of houses?
[140M House in US; 14M in CA; 1000B/record]



Query Plan (Logical)



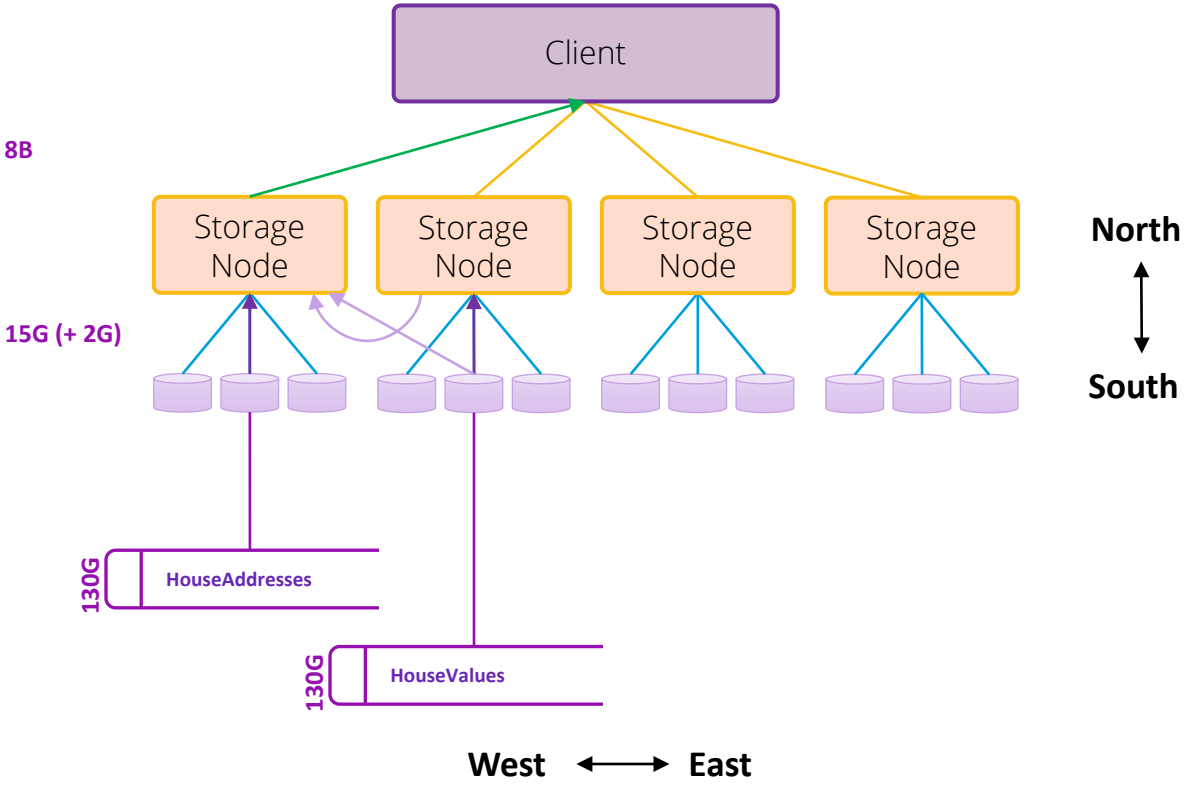
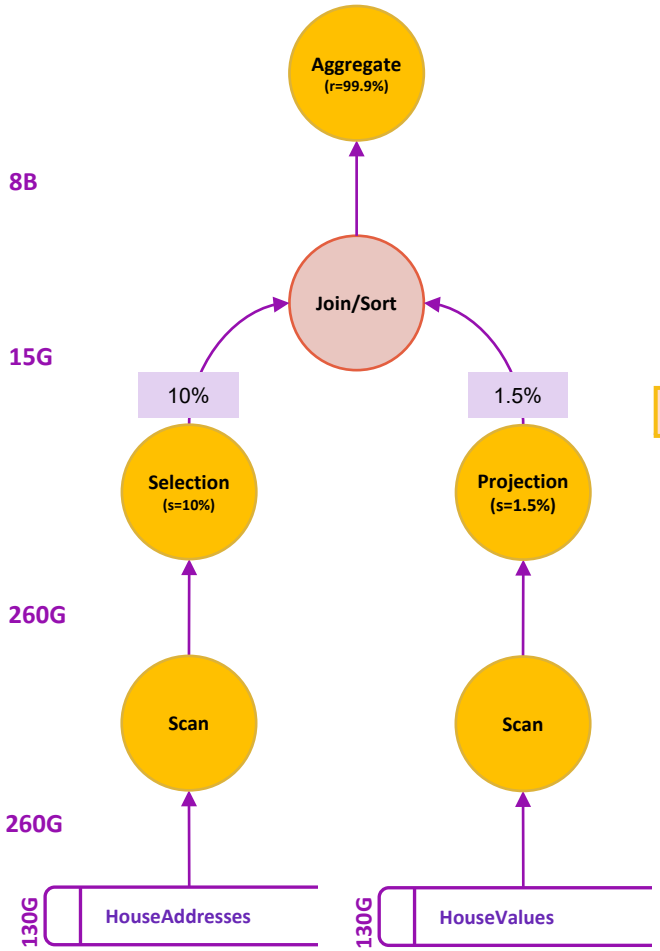
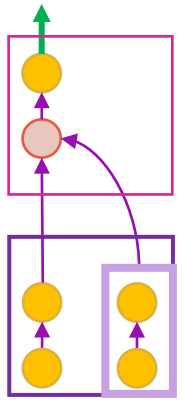


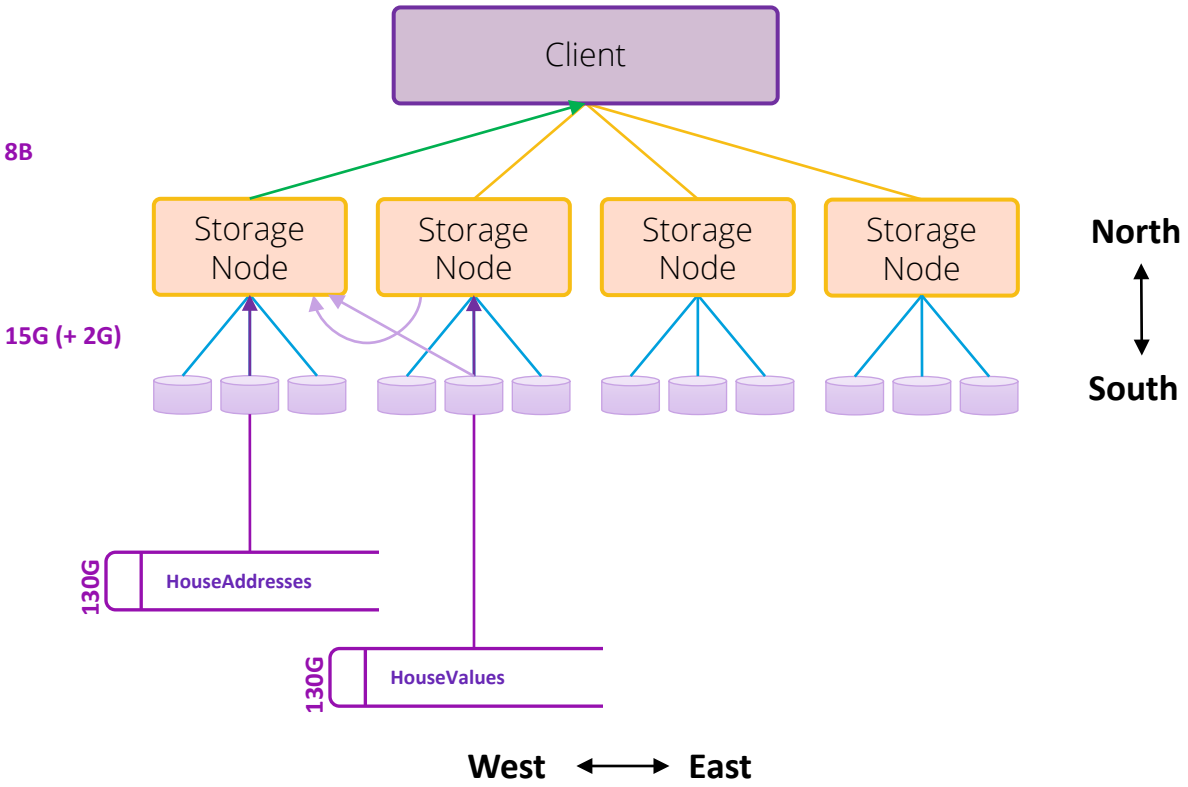
Illustration – Execution (Mixed)

Question: What zip code in CA has the highest average value of houses?
[140M House in US; 14M in CA; 1000B/record]



Query Plan (Logical)

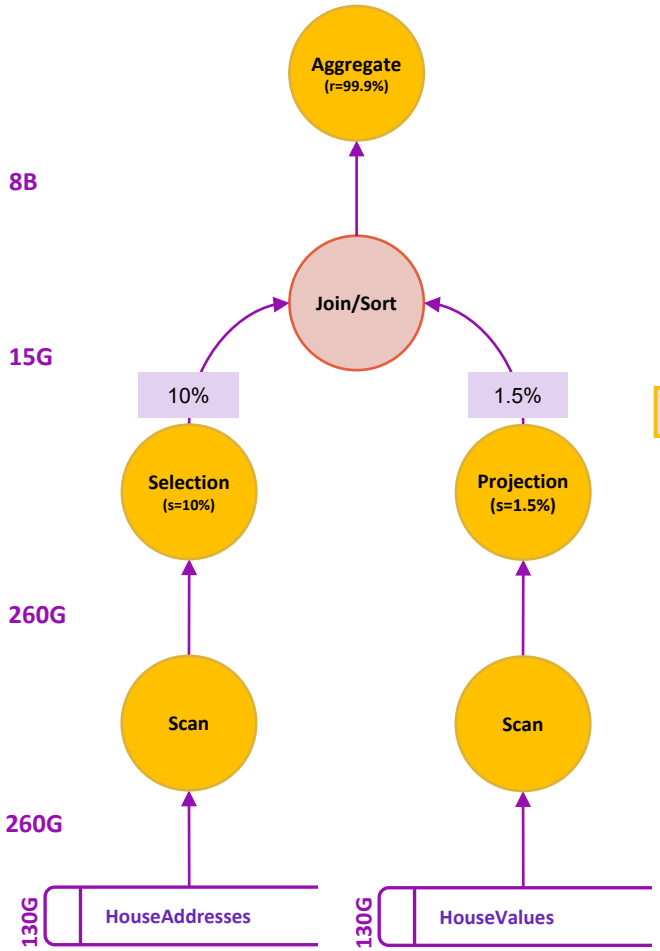




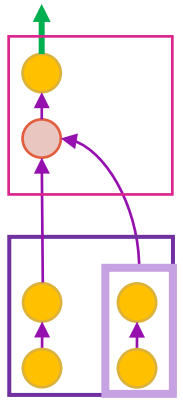
All examples yield the same answer with different efficiencies

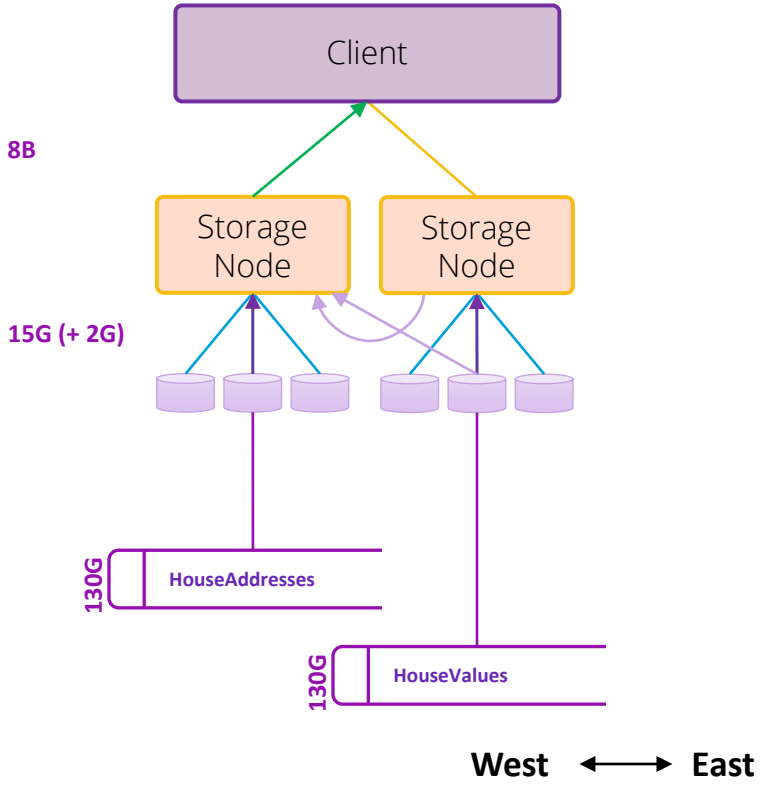
Illustration – Execution (Mixed)

Question: What zip code in CA has the highest average value of houses?
[140M House in US; 14M in CA; 1000B/record]



Query Plan (Logical)



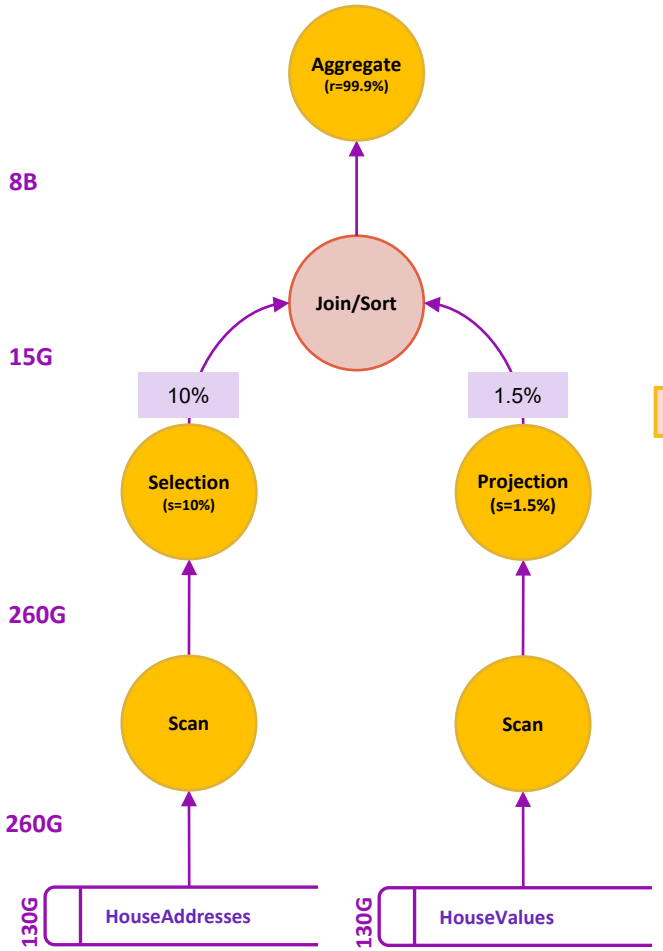


North

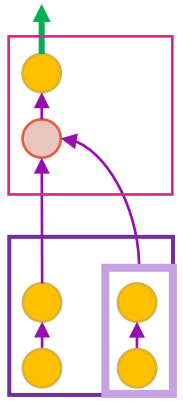
South

Illustration –Decomposition

Question: What zip code in CA has the highest average value of houses?
[140M House in US; 14M in CA; 1000B/record]



Query Plan (Logical)



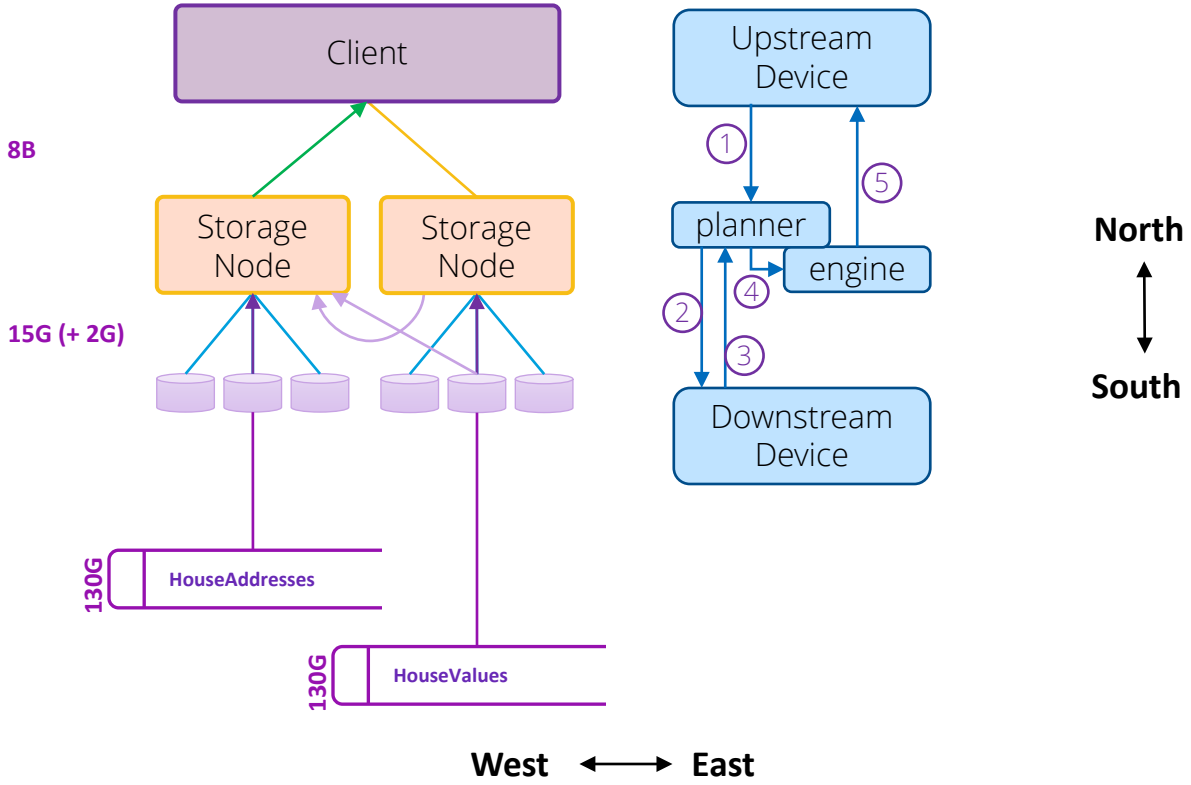
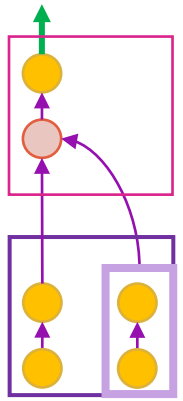


Illustration – Tooling

Question: What zip code in CA has the highest average value of houses?
[140M House in US; 14M in CA; 1000B/record]



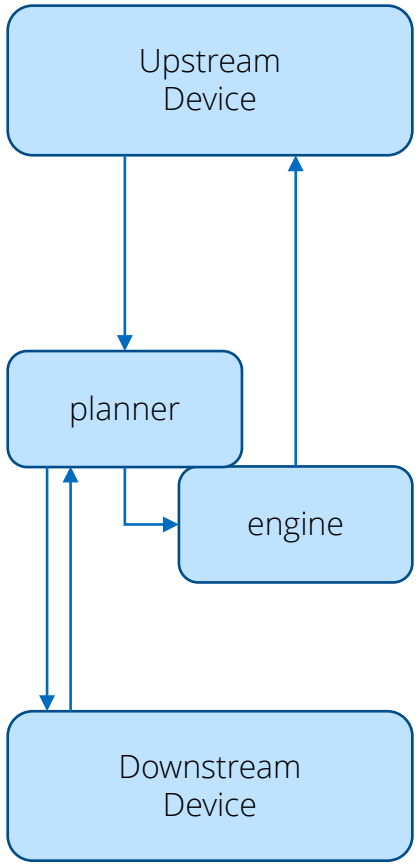
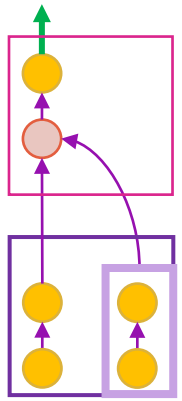
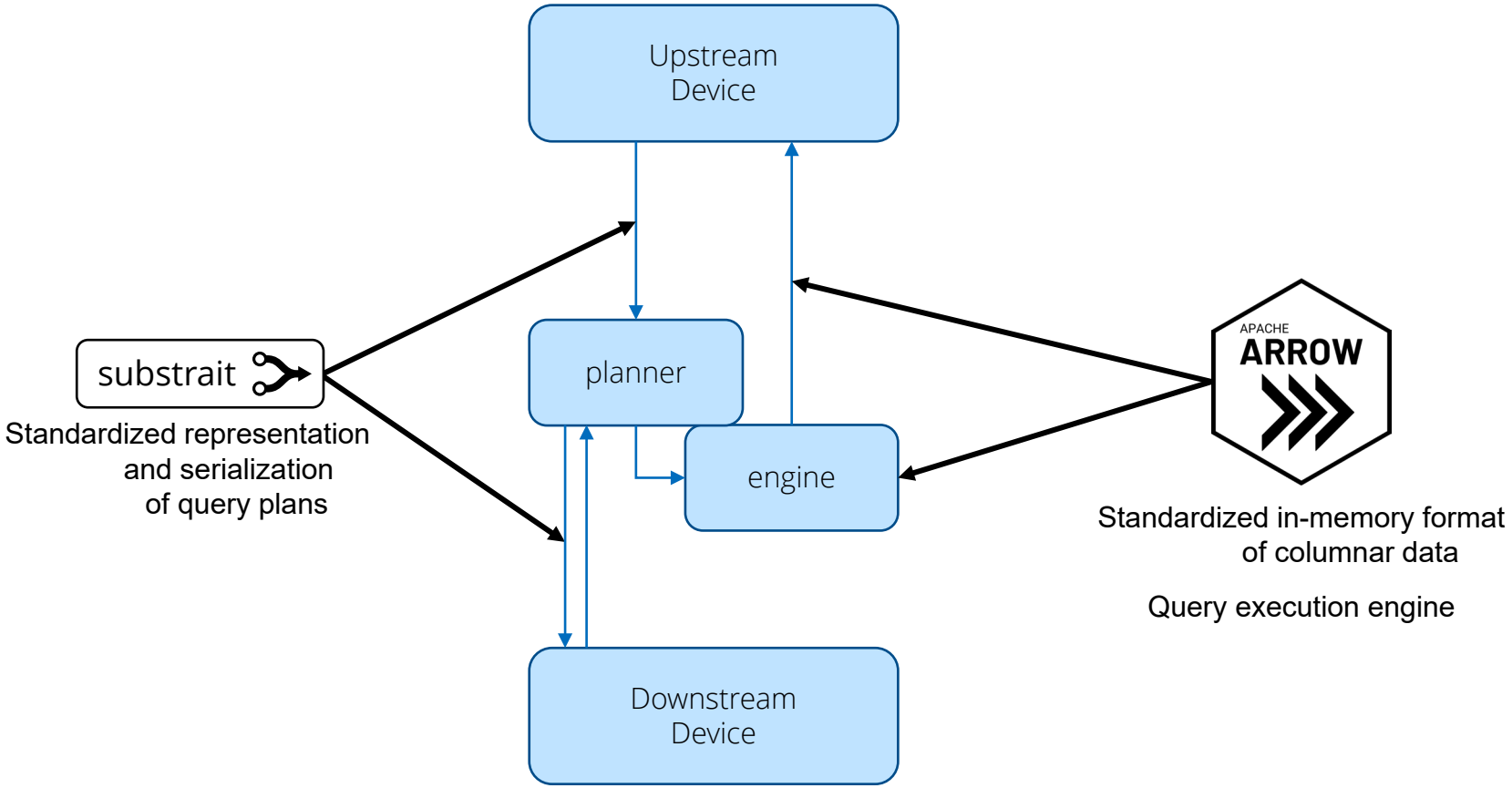


Illustration – Tooling

Question: What zip code in CA has the highest average value of houses?
[140M House in US; 14M in CA; 1000B/record]

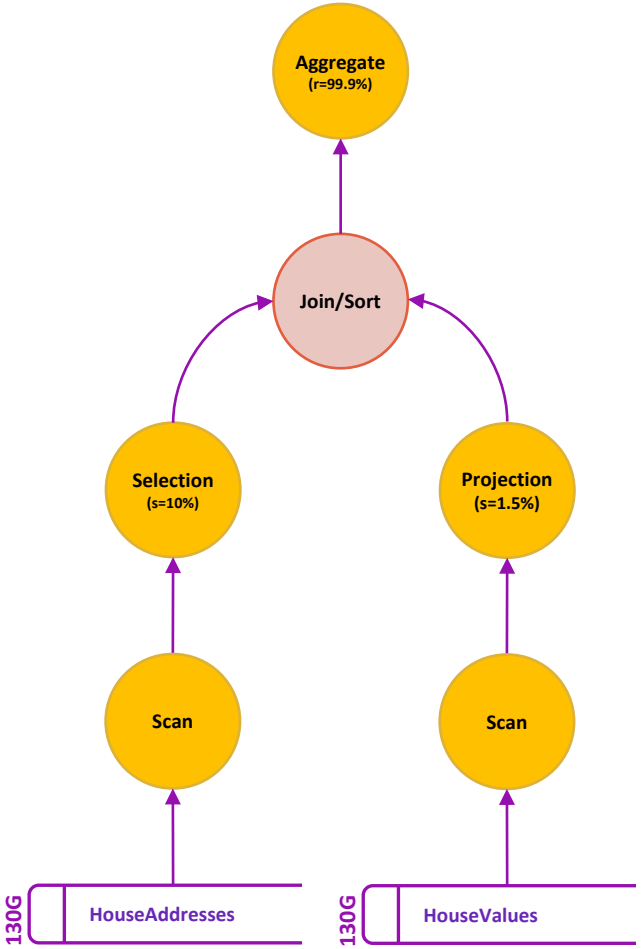


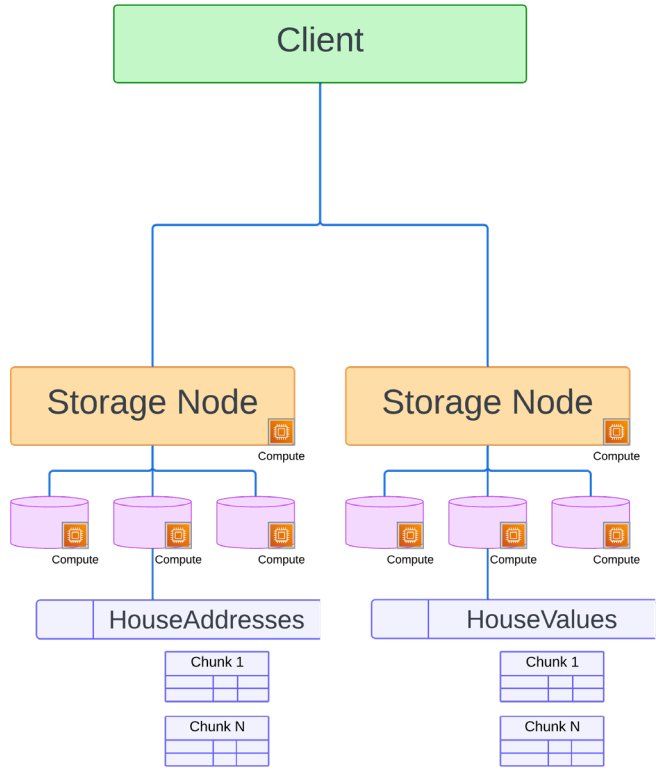


Using Substrait

Decomposable Queries – Using Substrait

- Send query plan to each object
- Determine “relevant” sub-plan
  - Lookup catalog
  - Maximize sub-plan
    - Intersect with catalog objects
    - Start with current object name
- Pushdown sub-plan

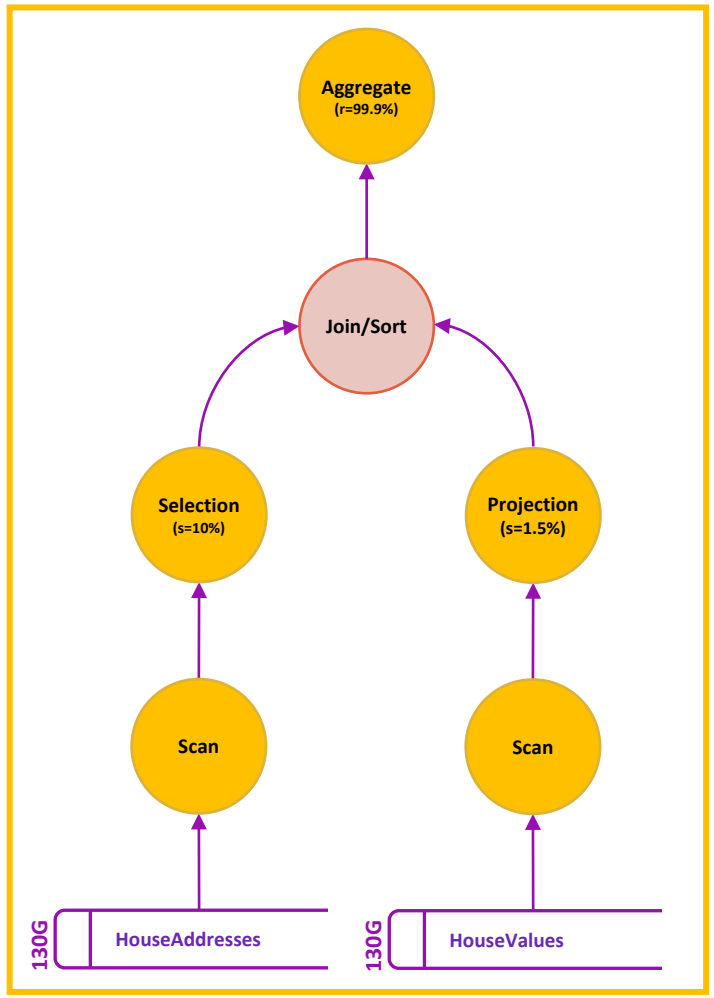


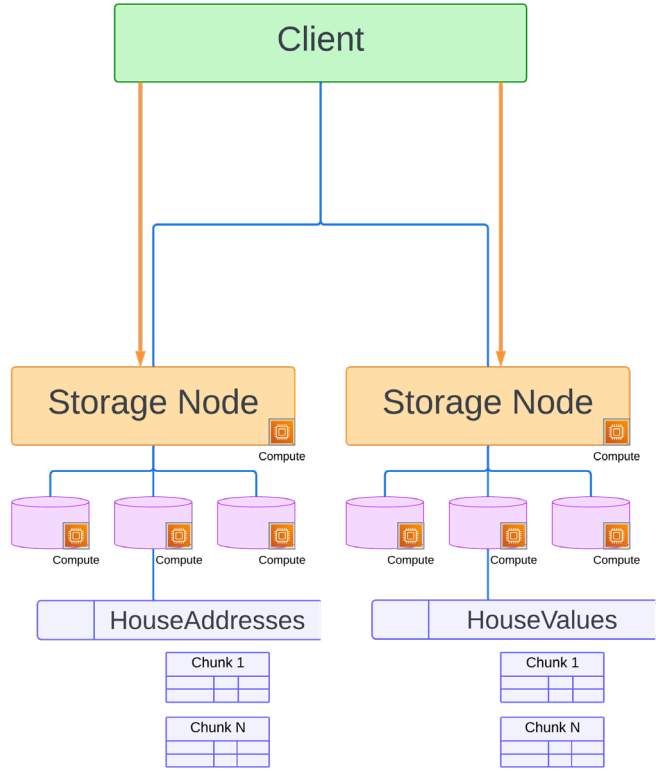


Decomposable Queries – Using Substrait

- Send query plan to each object

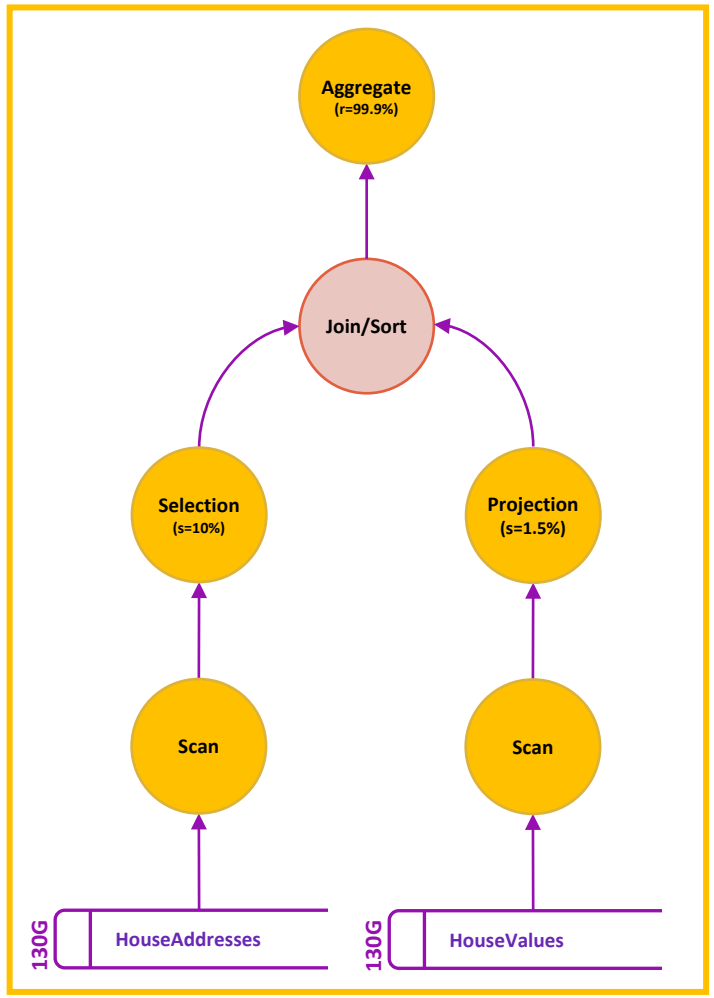
- Determine “relevant” sub-plan
  - Lookup catalog
  - Maximize sub-plan
    - Intersect with catalog objects
    - Start with current object name
- Pushdown sub-plan

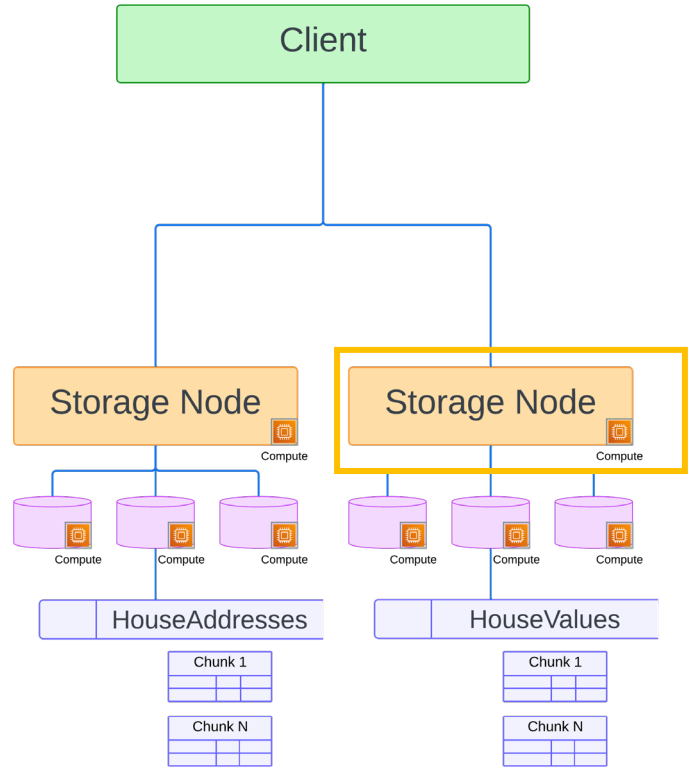




Decomposable Queries – Using Substrait

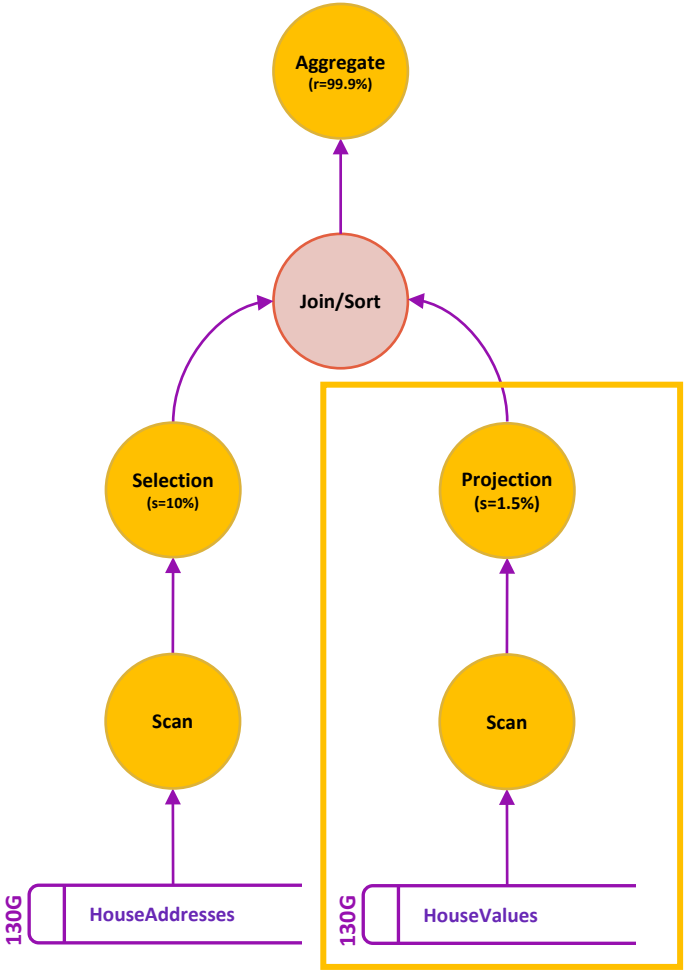
- Send query plan to each object
- Determine “relevant” sub-plan
  - Lookup catalog
  - Maximize sub-plan
    - Intersect with catalog objects
    - Start with current object name
- Pushdown sub-plan

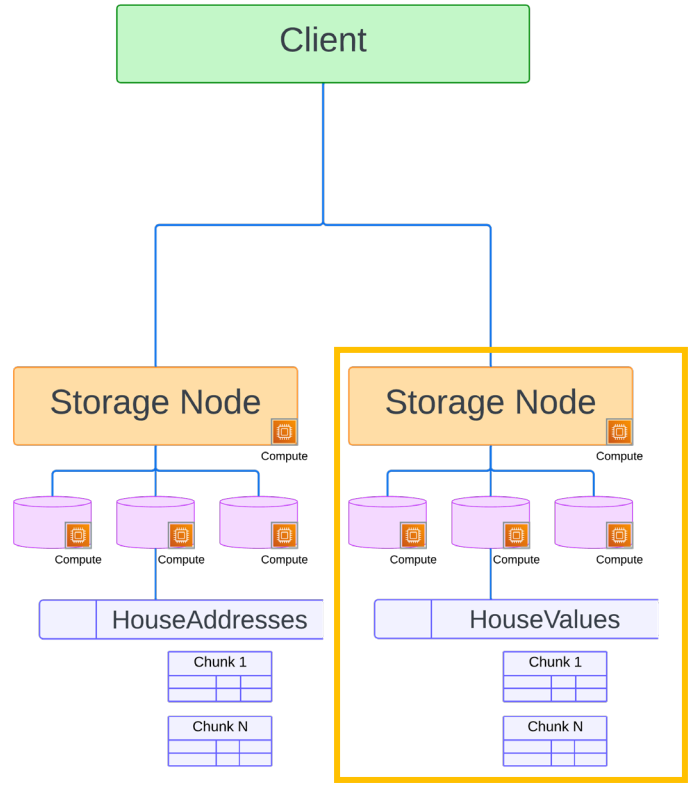




Decomposable Queries – Using Substrait

- Send query plan to each object
- Determine “relevant” sub-plan
  - Lookup catalog
  - Maximize sub-plan
    - Intersect with catalog objects
    - Start with current object name
- Pushdown sub-plan



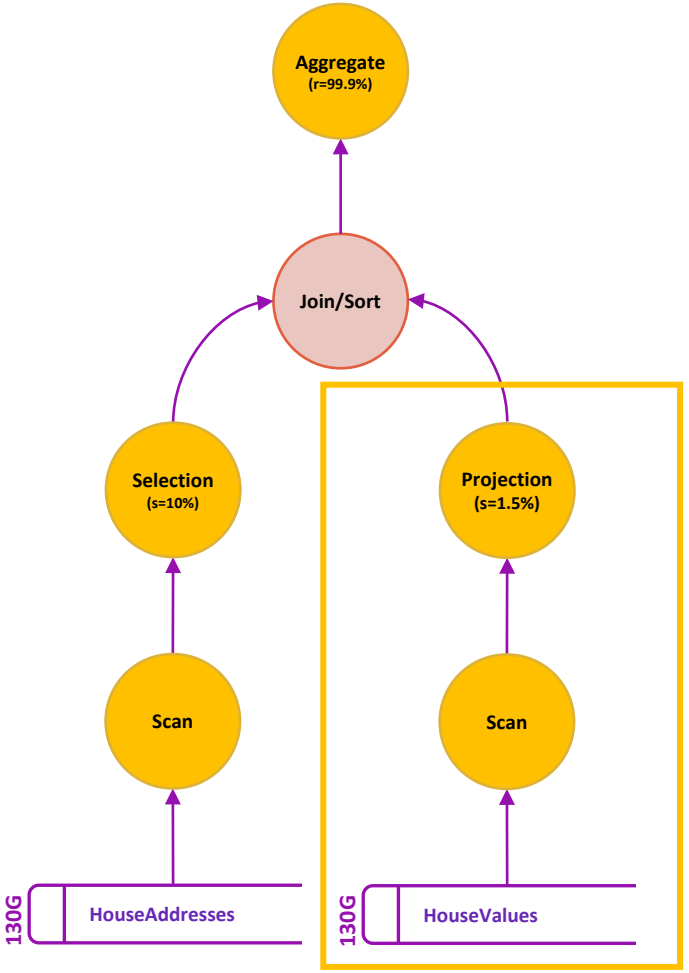


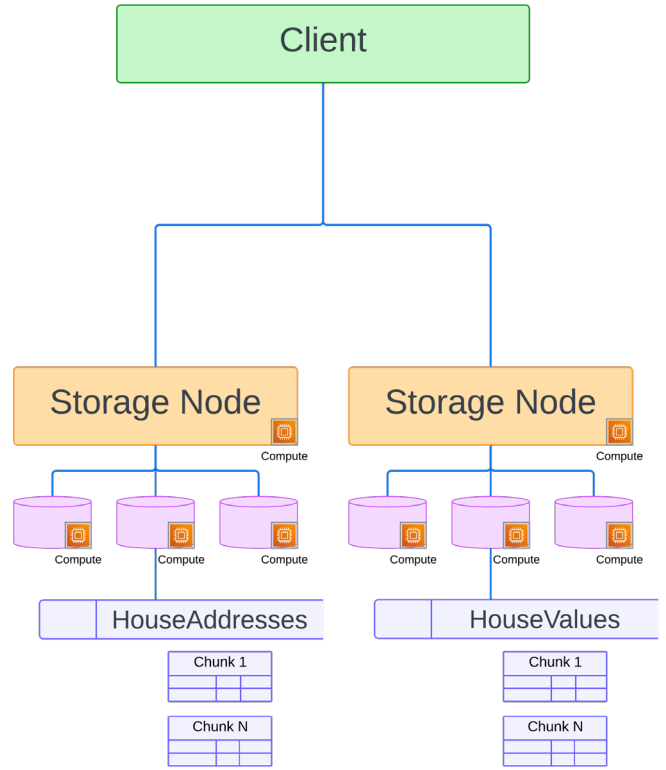
Decomposable Queries – Using Substrait

- Send query plan

```
input {
  project {
    common {
      emit {
        output_mapping: 9
        output_mapping: 10
        output_mapping: 11
        output_mapping: 12
      }
    }
    input {
      read {
        base_schema { names: "region" ...  }
        named_table { names: "HouseValues" }
      }
    }
    expressions { selection { ... } }
  }
}
```

substrait



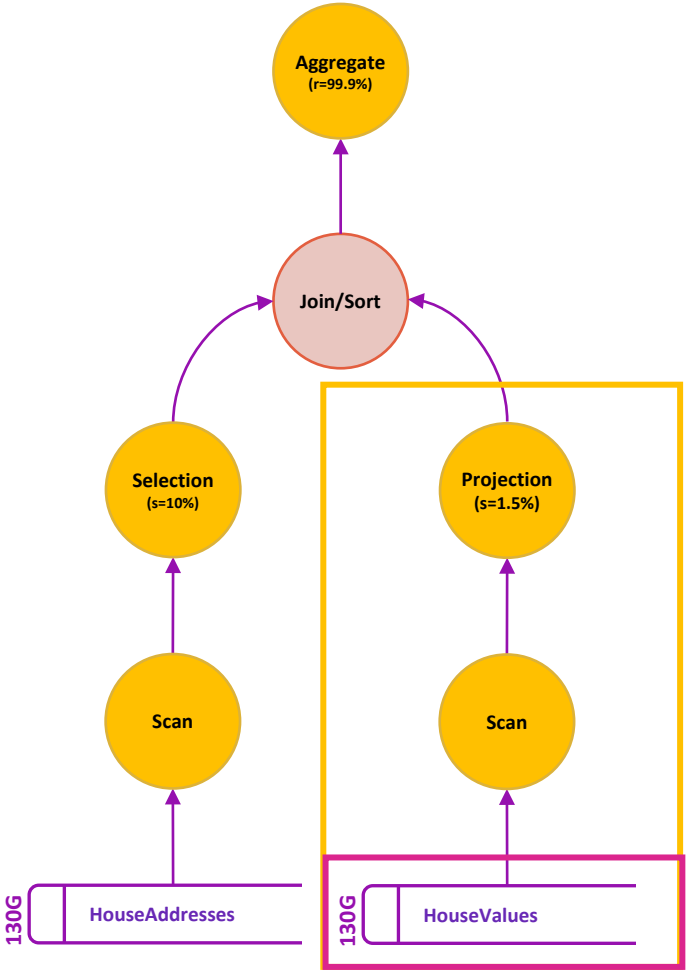


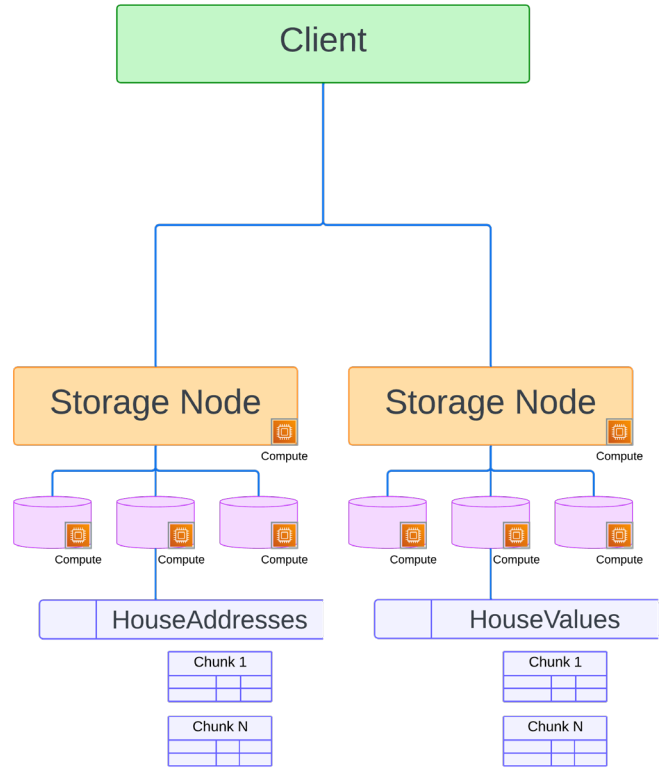
Decomposable Queries – Using Substrait

- Send query plan

```
input {
  project {
    common {
      emit {
        output_mapping: 9
        output_mapping: 10
        output_mapping: 11
        output_mapping: 12
      }
    }
    input {
      read {
        base_schema { names: "region" ... }
        named_table { names: "HouseValues" }
      }
    }
    expressions { selection { ... } }
  }
}
```

substrait



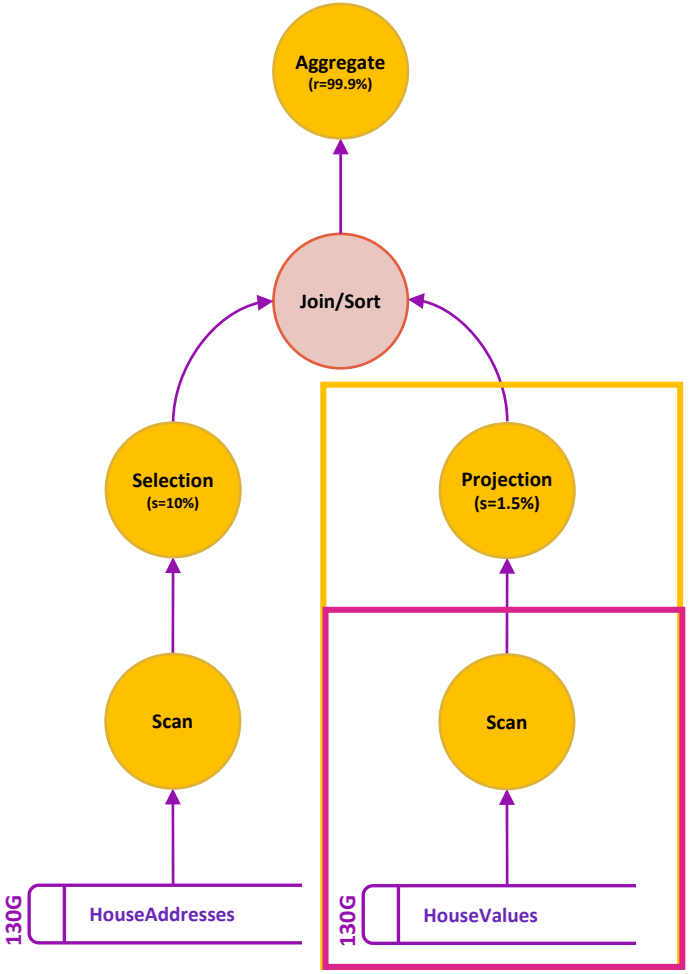


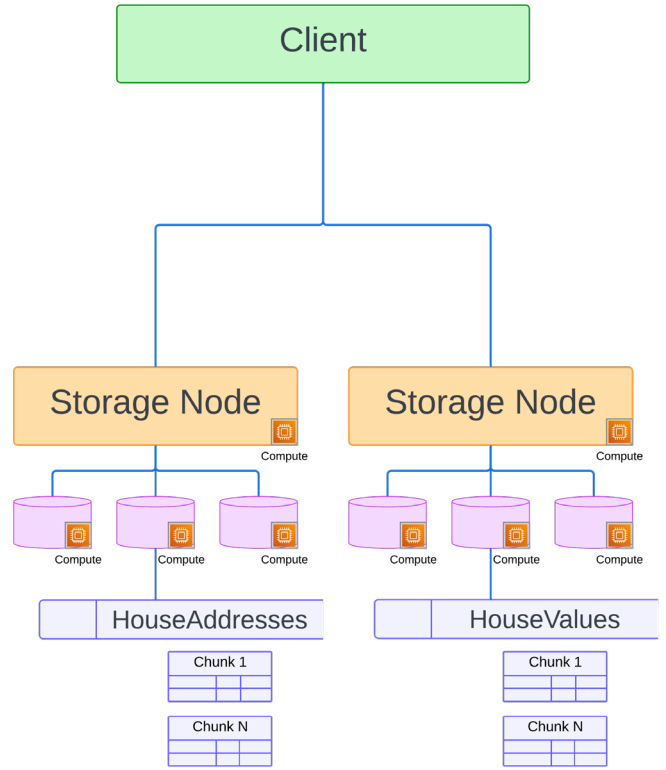
Decomposable Queries – Using Substrait

- Send query plan

```
input {
  project {
    common {
      emit {
        output_mapping: 9
        output_mapping: 10
        output_mapping: 11
        output_mapping: 12
      }
    }
    input {
      read {
        base_schema { names: "region" ... }
        named_table { names: "HouseValues" }
      }
    }
    expressions { selection { ... } }
  }
}
```

substrait



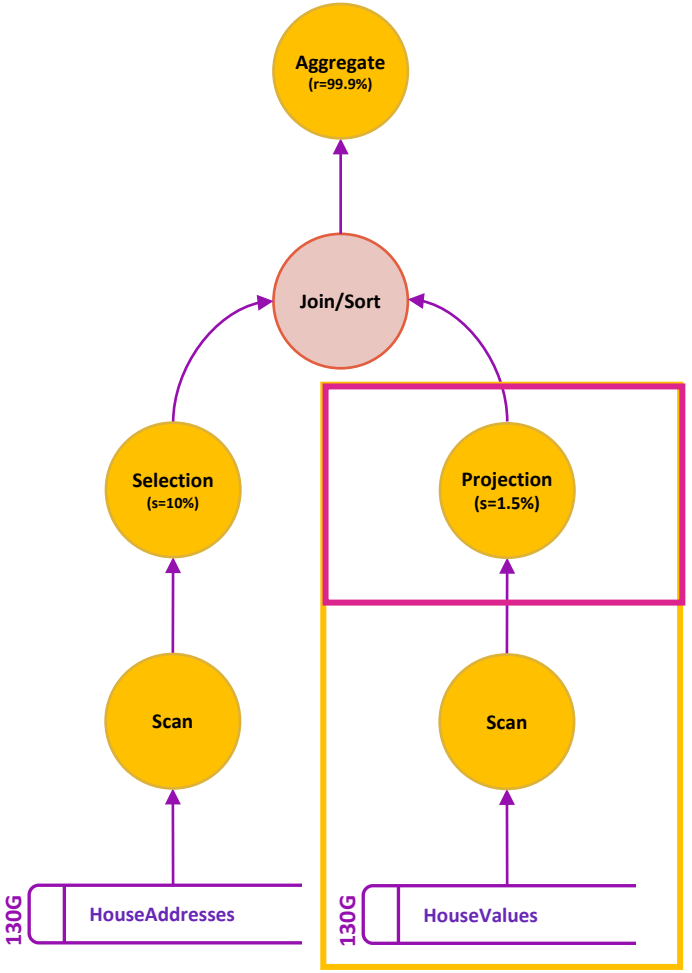


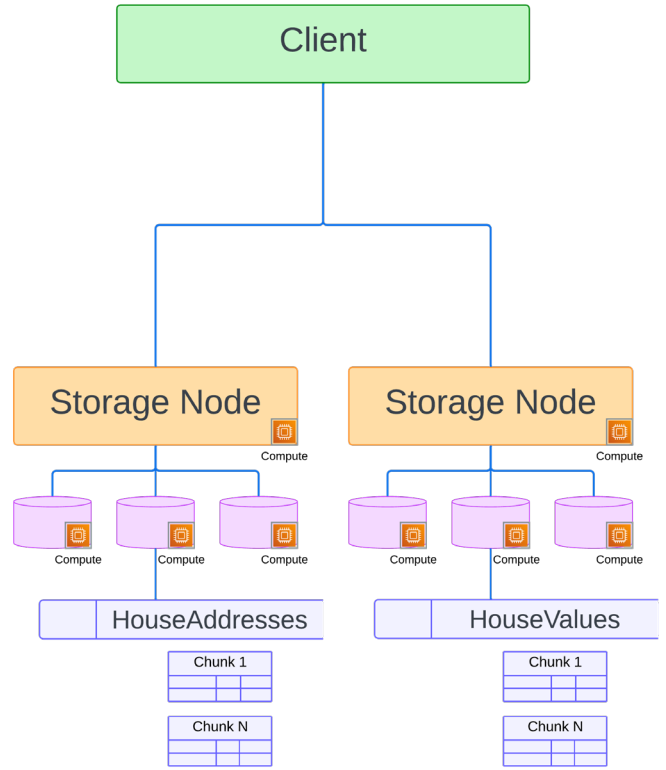
Decomposable Queries – Using Substrait

- Send query plan

```
input {
  project {
    common {
      emit {
        output_mapping: 9
        output_mapping: 10
        output_mapping: 11
        output_mapping: 12
      }
    }
    input {
      read {
        base_schema { names: "region" ...  }
        named_table { names: "HouseValues" }
      }
    }
    expressions { selection { ... } }
  }
}
```

substrait





Decomposable Queries – Using Substrait

- Send query plan

```
input {
  project {
    common {
      emit {
        output_mapping: 9
        output_mapping: 10
        output_mapping: 11
        output_mapping: 12
      }
    }
    input {
      read {
        base_schema { names: "region" ...  }
        named_table { names: "HouseValues" }
      }
    }
    expressions { selection { ... } }
  }
}
```

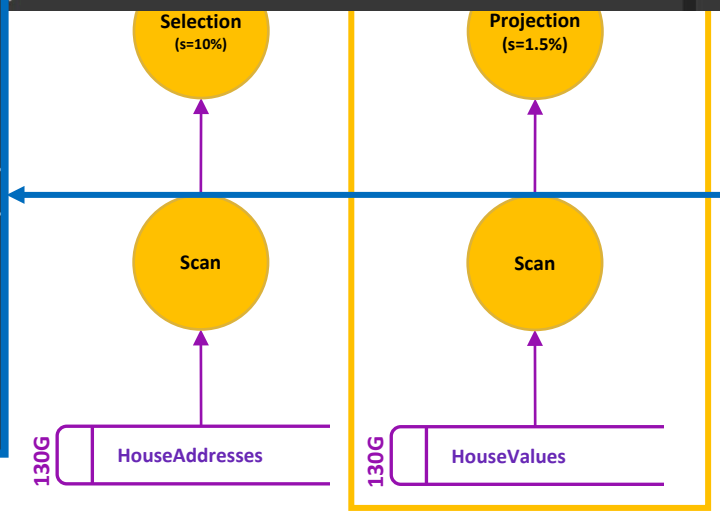
substrait

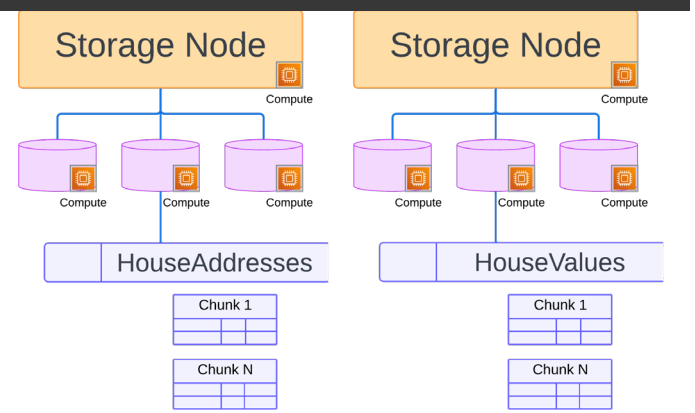
```
# Schema for median house sales
HouseValues = ibis.table(
    [   # Column names and types
        ('region'             , 'string')
       ,('end_month'          , 'string')
       ,('end_year'           , 'int   ')
       ,('median_price'       , 'int   ')
       ,('homes_sold'         , 'int   ')
       ,('new_listings'       , 'int   ')
       ,('inventory'          , 'int   ')
       ,('days_on_market'     , 'int   ')
       ,('average_sale_to_list', 'float ')
    ]
   ,name='HouseValues' # Table name
)
```

```
# Define the query as an ibis expression
query_expr = HouseValues[ 'region'
                         ,'end_month'
                         ,'end_year'
                         ,'median_price']

# SubstraitCompiler translates: ibis expr -> substrait plan
substrait_compiler = SubstraitCompiler()
proto_msg = substrait_compiler.compile(query_expr)
```

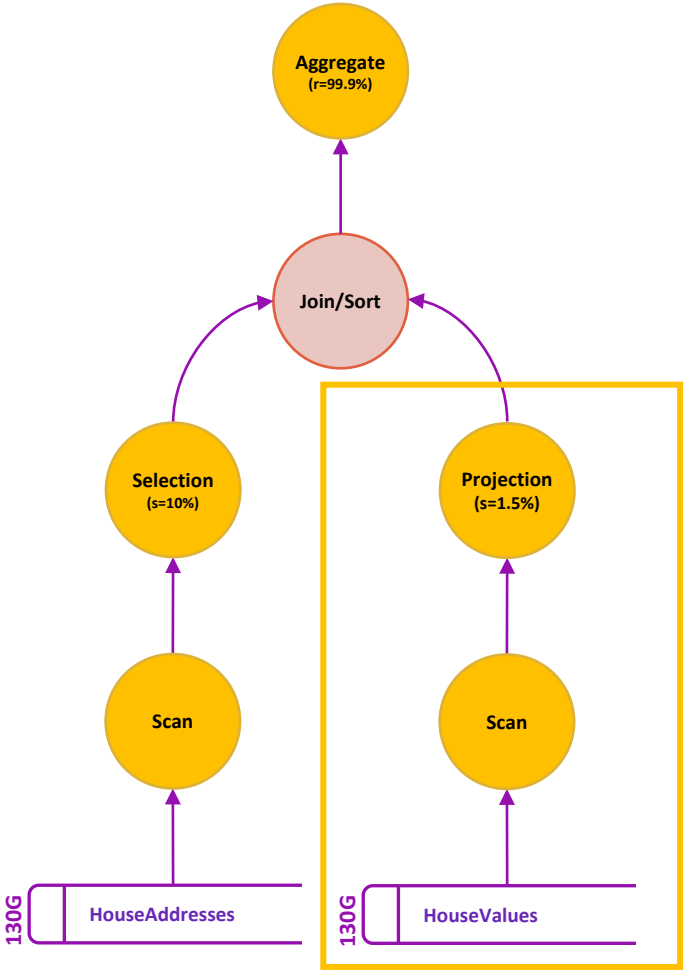
Ibis project

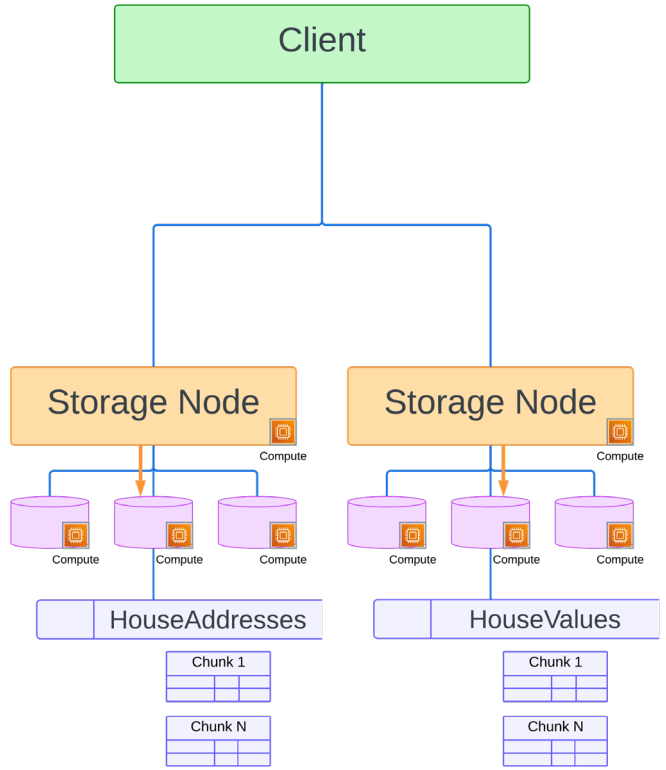




Decomposable Queries – Using Substrait

- Send query plan to each object
- Determine “relevant” sub-plan
  - Lookup catalog
  - Maximize sub-plan
    - Intersect with catalog objects
    - Start with current object name
- Pushdown sub-plan

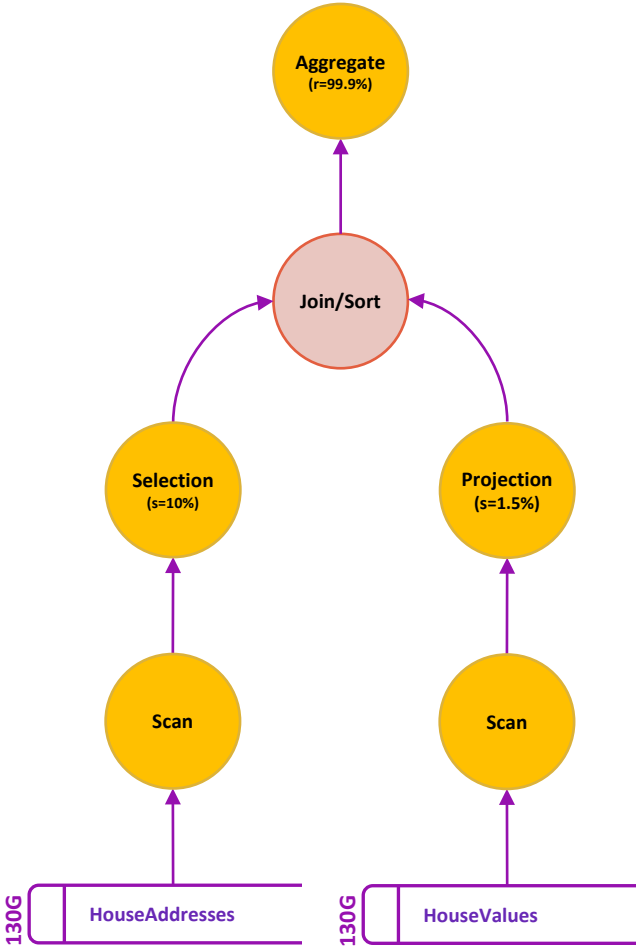


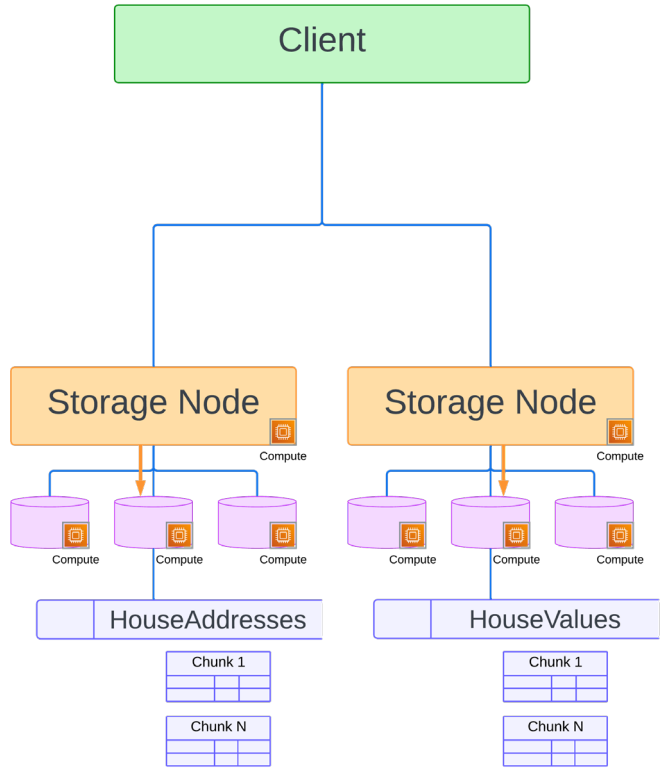


Using Apache Arrow

Decomposable Queries – Using Apache Arrow

- Execute Plan
  - Use first K chunks as canary
  - Return pushback plan
  - Return results





Decomposable Queries – Using Apache Arrow

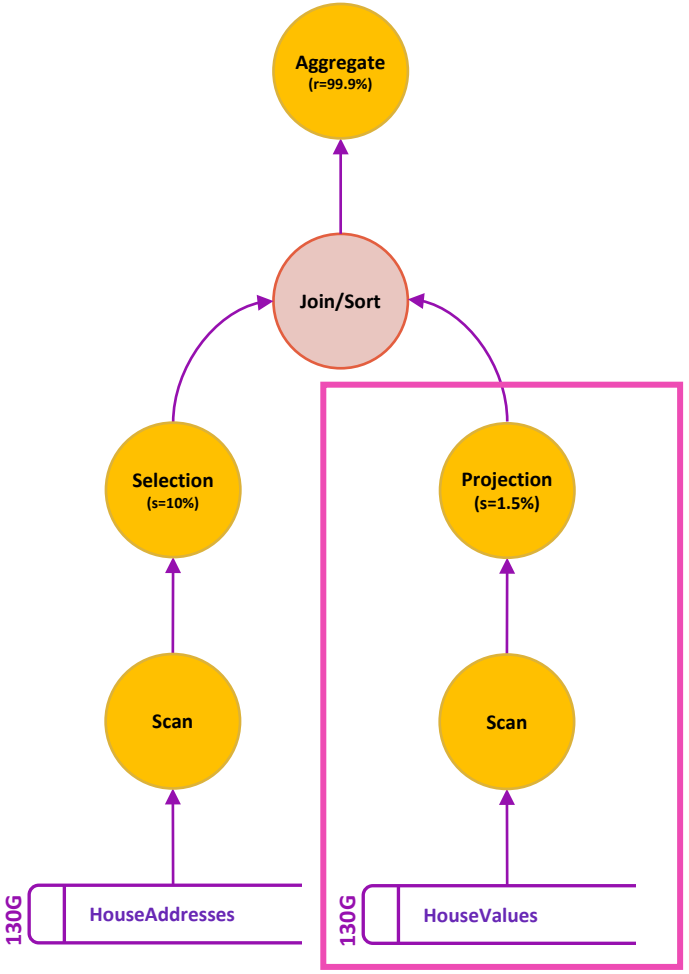
- Execute Plan
  - Use first K chunks as canary
  - Return pushback plan
  - Return results

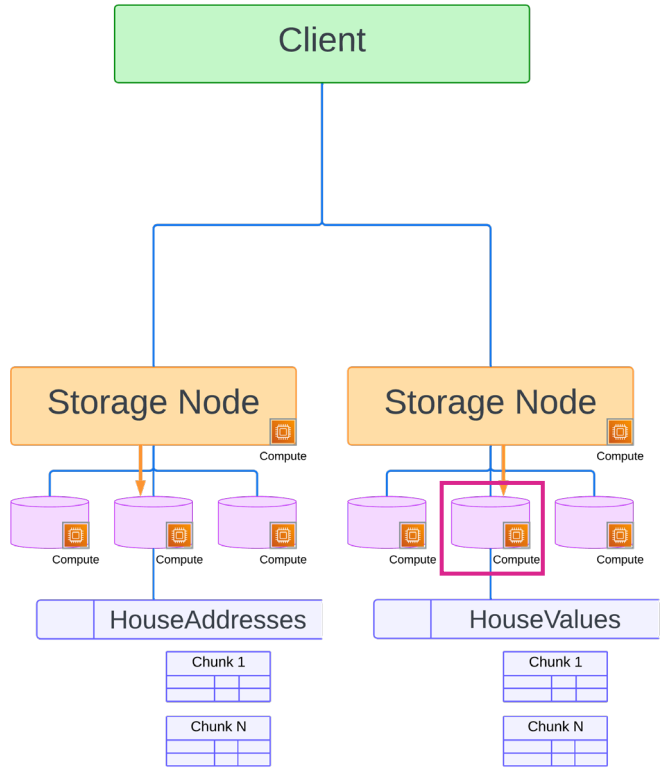
```
def ExampleTableProvider(table_names, expected_schema=None):
    # table_names is a list of strings representing a single table
    tname        = '.'.join(table_names)
    source_table = TableFromCSV(FilePathsByTable[tname])
    print(source_table.to_pandas())

    return source_table

def ExecuteSubstrait(substrait_plan: bytes) -> pyarrow.Table:
    print('Executing substrait...')
    result_reader = substrait.run_query(
         substrait_plan
        ,table_provider=ExampleTableProvider
    )

    print('Query plan executed')
    return result_reader.read_all()
```





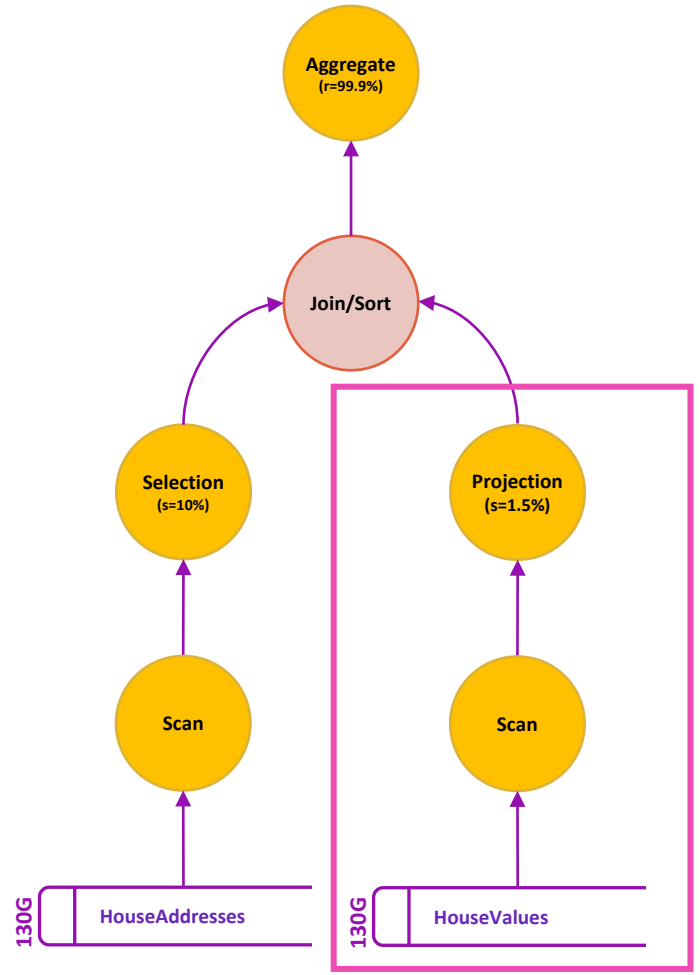
Decomposable Queries – Using Apache Arrow

```
def ExampleTableProvider(table_names, expected_schema=None):
    # table_names is a list of strings representing a single table
    tname        = '.'.join(table_names)
    source_table = TableFromCSV(FilePathsByTable[tname])
    print(source_table.to_pandas())

    return source_table

def ExecuteSubstrait(substrait_plan: bytes) -> pyarrow.Table:
    print('Executing substrait...')
    result_reader = substrait.run_query(
         substrait_plan
        ,table_provider=ExampleTableProvider
    )

    print('Query plan executed')
    return result_reader.read_all()
```



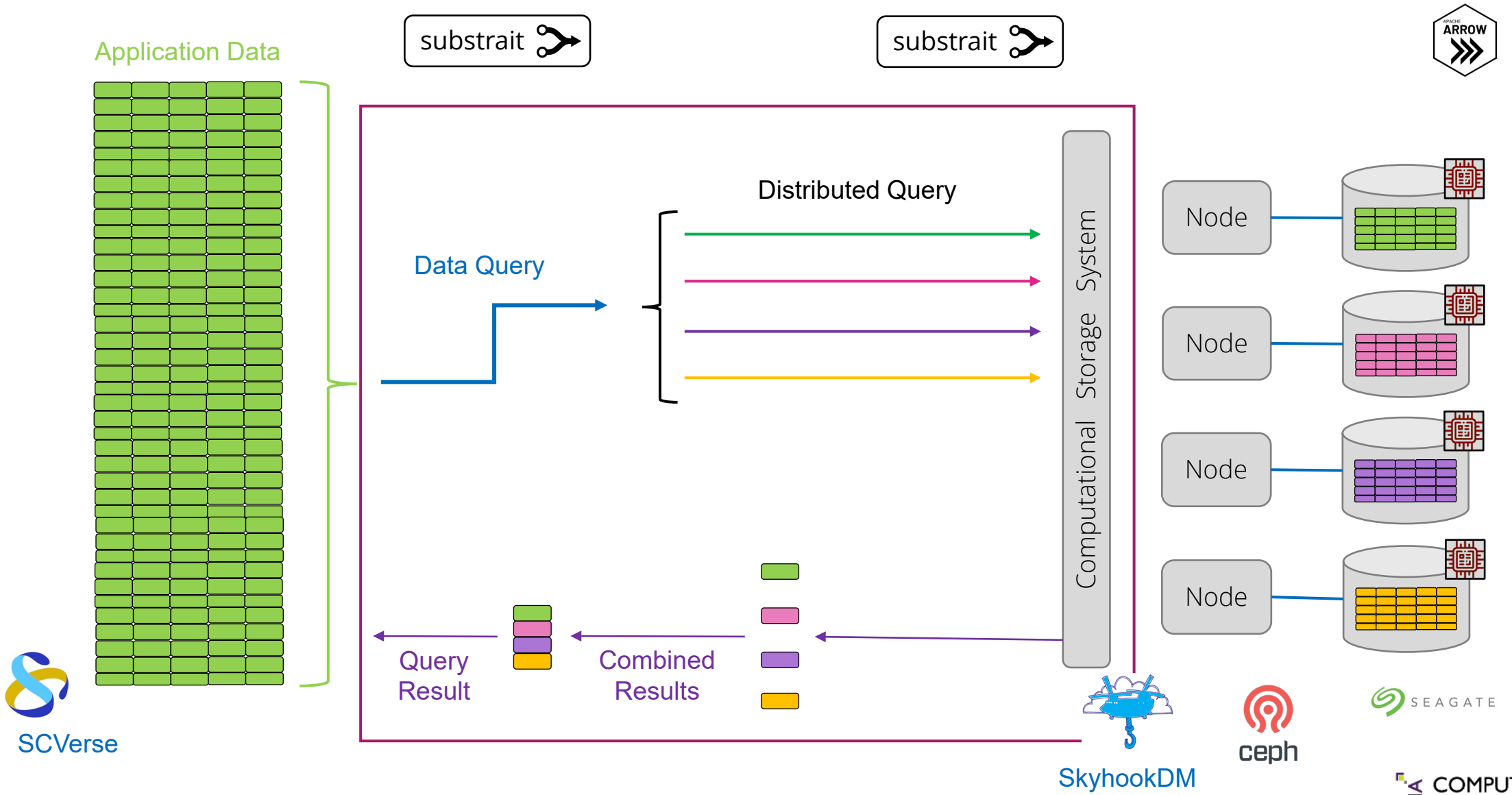
Decomposable Queries – Using Apache Arrow

```
(demo-cms-py3.11) 00:26 octalene@octalene-air ~/c/d/demo-cms  (mainline) python demo_cms/execute_substrait.py
Executing substrait...
Before Execution
                region end_month  end_year  median_price  homes_sold  new_listings  inventory  days_on_market  average_sale_to_list
0       Acalanes Ridge, CA  February      2012        590000           1             1          1              18                 0.967
1       Acalanes Ridge, CA     March      2012        743000           2             1          1              20                 0.974
2       Acalanes Ridge, CA       May      2012        961000           1             2          4              31                 0.916
3       Acalanes Ridge, CA      June      2012        600000           1             2          4               5                 1.002
4       Acalanes Ridge, CA      July      2012        671000           2             2          4              40                 0.991
...                    ...       ...       ...           ...         ...           ...        ...             ...                   ...
142264        Zayante, CA     April      2022        550000           1             0          1              57                 0.972
142265        Zayante, CA      June      2022        725000           1             1          4              23                 1.115
142266        Zayante, CA      July      2022        500000           1             0          2              16                 1.002
142267        Zayante, CA    August      2022        643000           2             0          1              73                 0.971
142268        Zayante, CA   October      2022        600000           1             0          0              28                 1.002

[142269 rows x 9 columns]
Query plan executed
After Execution
             FieldPath(0) FieldPath(1)  FieldPath(2)  FieldPath(3)
0       Acalanes Ridge, CA     February          2012        590000
1       Acalanes Ridge, CA        March          2012        743000
2       Acalanes Ridge, CA          May          2012        961000
3       Acalanes Ridge, CA         June          2012        600000
4       Acalanes Ridge, CA         July          2012        671000
...                    ...          ...           ...           ...
142264        Zayante, CA        April          2022        550000
142265        Zayante, CA         June          2022        725000
142266        Zayante, CA         July          2022        500000
142267        Zayante, CA       August          2022        643000
142268        Zayante, CA      October          2022        600000

[142269 rows x 4 columns]
```


Overview



References

[Diagram] Rood, J.E., Maartens, A., Hupalowska, A. *et al.* Impact of the Human Cell Atlas on medicine. *Nat Med* **28**, 2486–2496 (2022). https://doi.org/10.1038/s41591-022-02104-7

[Screenshot] Chan-Zuckerberg Institute. CELLxGENE Discover. https://cellxgene.cziscience.com/

[Repository] https://github.com/drin/demo-cms-skytether

[HCA website] https://www.humancellatlas.org/
[Example Dataset] https://www.redfin.com/news/data-center/

[Tool] Substrait homepage: https://substrait.io/
[Tool] Ibis homepage: https://ibis-project.org/
[Tool] Apache Arrow homepage: https://arrow.apache.org/

[Tool] Ibis-substrait module: https://github.com/ibis-project/ibis-substrait
[Tool] pyarrow substrait module: https://github.com/apache/arrow/blob/main/python/pyarrow/_substrait.pyx

SNIA COMPUTE + MEMORY + STORAGE SUMMIT

Architectures, Solutions, and Community
VIRTUAL EVENT, APRIL 11-12, 2023

Please take a moment to rate this session.

Your feedback is important to us.

Post-Summit, visit www.snia.org/cms-summit for additional content.

