GROUPBY 2020 | MAY 12-13

Let’s Dive into SQL Server I/O To Improve T-SQL Performance

Andy Yun

Principal Solutions Engineer - SentryOne

Andy Yun

Principal Solutions Engineer

- SQL Server DBA & DB Developer
- Chicago Suburban User Group Chapter Leader
- Chicago SQL Association – Director-at-Large
- Working with SQL Server since 2001
- Speaking since Early 2014
- Microsoft MVP (2017-2018)

@SQLBek - sqlbek@gmail.com

https://blogs.sentryone.com/andyyun/

https://www.github.com/sqlbek/

SentryOne™

Setting the Stage

Pack It Up!

Are We There Yet?

Microsoft® SQL Server®

Have You Ever…

- Had a query only returns 5 records, why is it slow?
- Had a query typically runs in 5 seconds, but is now running slow?
- Wondered what your query is really doing behind the scenes?

Today's Agenda

1. Internals: Query Lifecycle
2. Internals: Data Structures
3. Internals: Navigating Heaps & Indexes
4. T-SQL & I/O Scenarios

Goal:

To explore SQL Server I/O, to better understand query performance issues

Session Level:

- 200 - 300
- Splash of 400

SentryOne

Internals: Query Lifecycle

SentryOne™

I Like Analogies

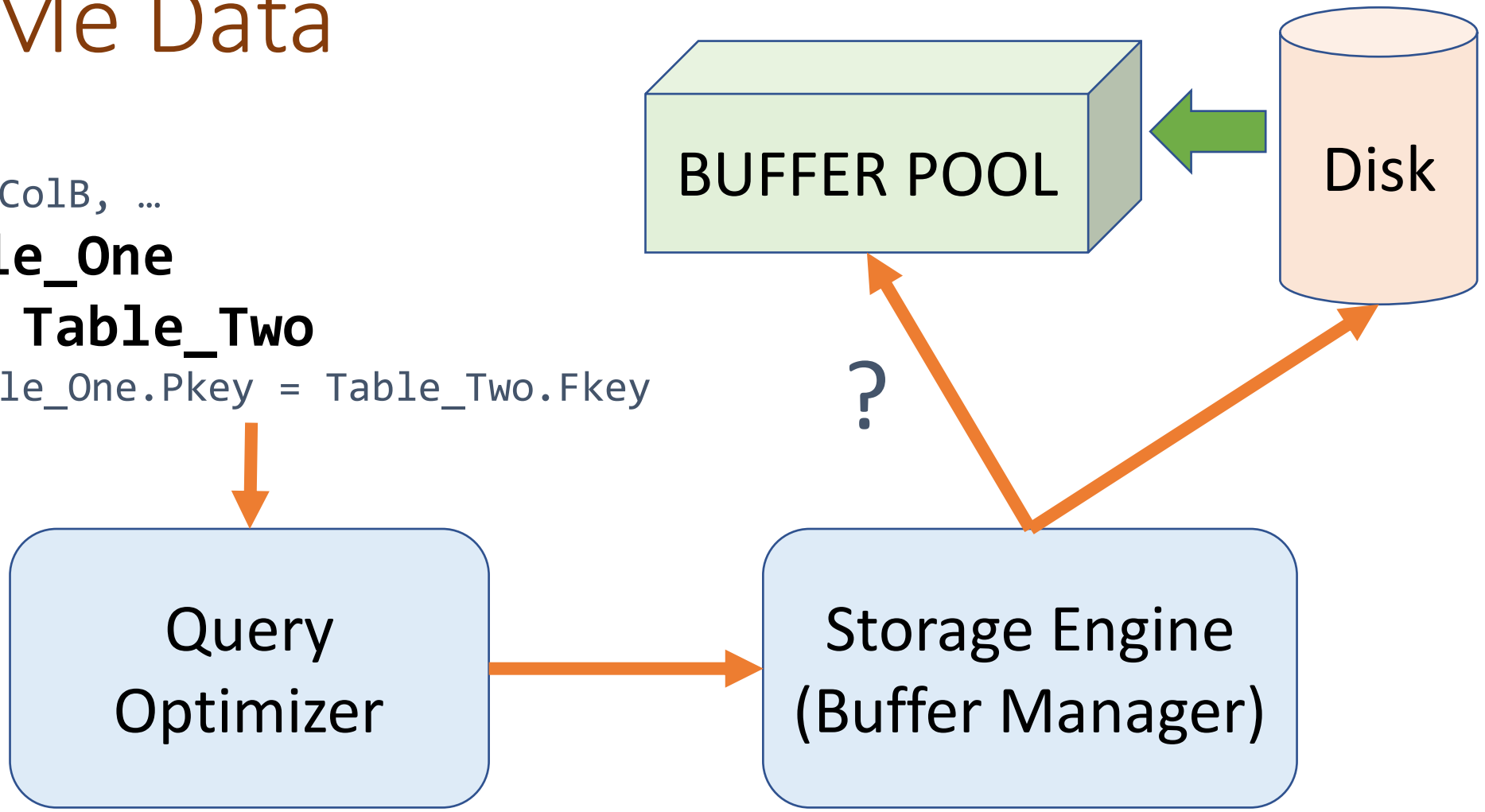
T-SQL Query

Query Optimizer

Storage Engine

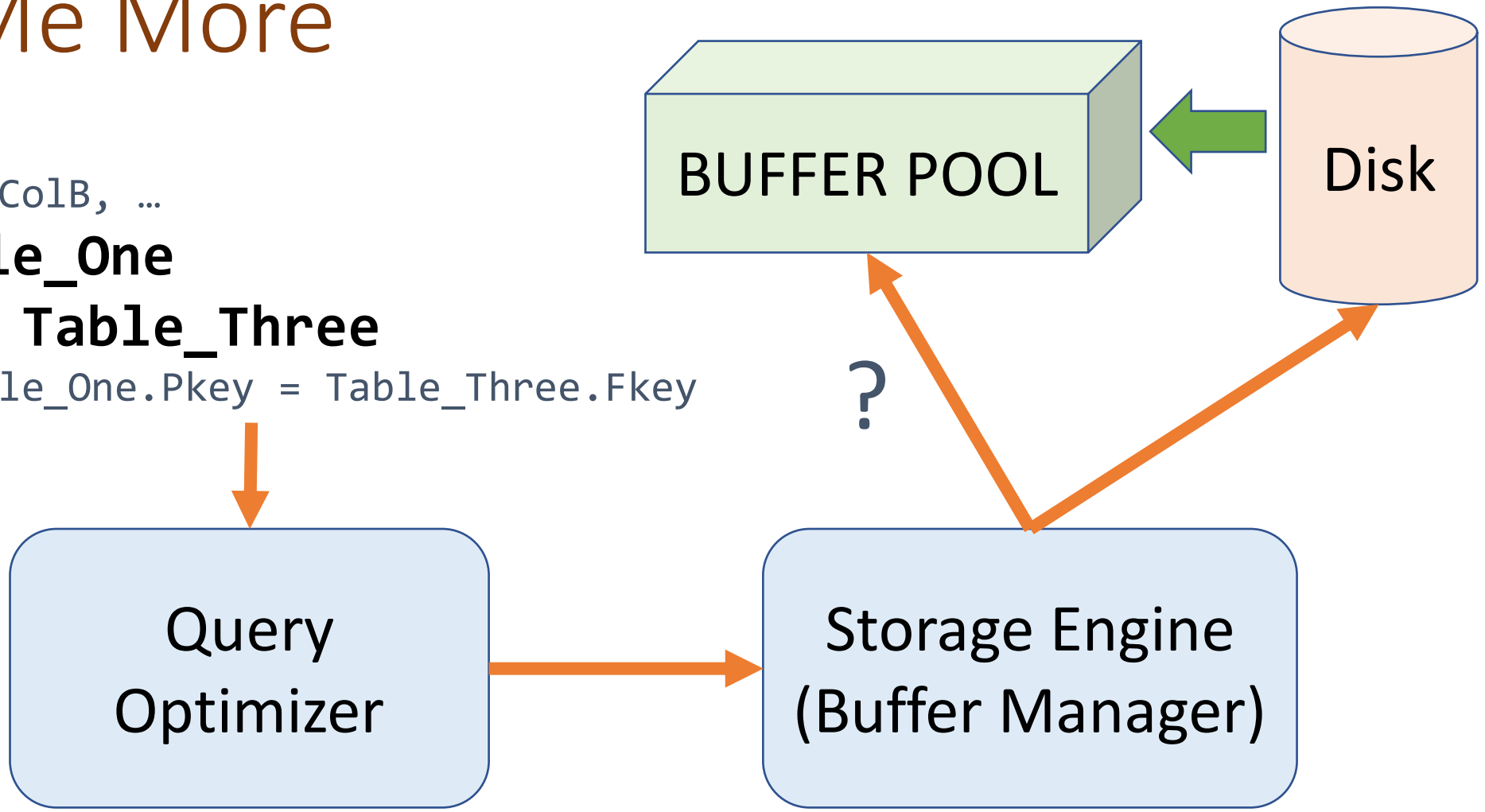
Give Me Data

```
SELECT
    ColA, ColB, …
FROM Table_One
INNER JOIN Table_Two
    ON Table_One.Pkey = Table_Two.Fkey
```



Get Me More

```
SELECT
    ColA, ColB, …
FROM Table_One
INNER JOIN Table_Three
    ON Table_One.Pkey = Table_Three.Fkey
```



Logical Reads vs Physical Reads

How do the two relate?

Which is more important?

Disregard logical reads & only focus on physical?

Where Did We Pack That Thing?!

1. Room = empty
2. Where did you pack X?

3. Retrieve 5 boxes to examine
   = 5 physical reads

4. Open & examine 5 boxes
   = 5 logical reads

1. Room = 5 boxes
2. Where did you pack Y?

3. Retrieve 3 boxes to examine
   = 3 physical reads

4. Open & examine 8 boxes
   = 8 logical reads

SentryOne

Key Takeaways

Physical Reads (Disk I/O) are slower than Logical Reads (Buffer Pool I/O)

Remember that ALL work is executed against data pages in the Buffer Pool

Internals: Data Structures

SentryOne™

Just Get Me One Record

```
SELECT TOP 1
     ColA, ColB
FROM Table_One
```

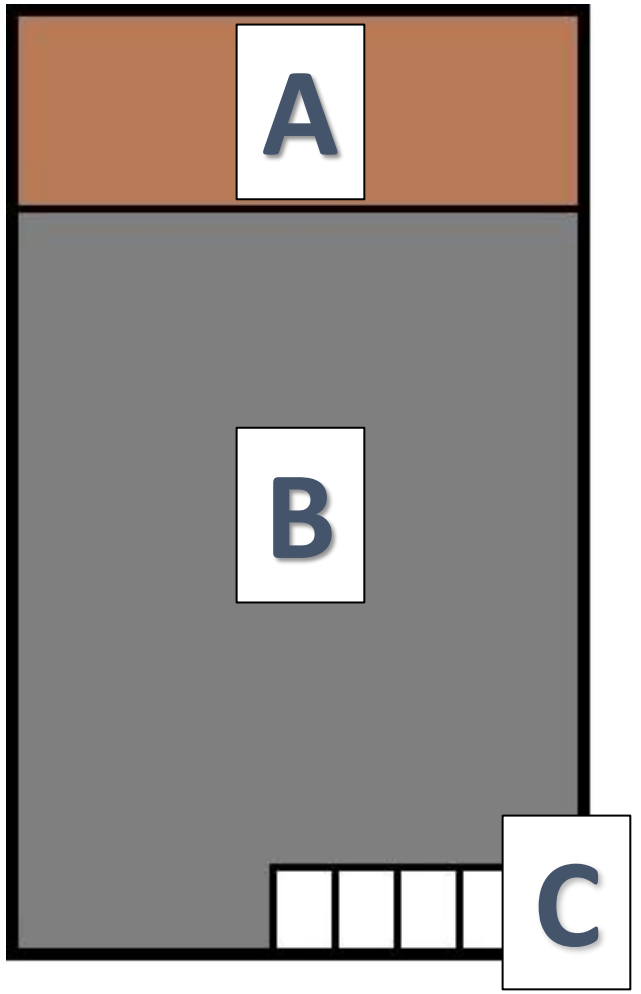
?

Data Page: Overview

Comprised of 3 components:

A. Page Header: 96 bytes
B. Data Records
C. Record or Slot Array:
   2 bytes per record

All Pages have a Fixed Size of 8KB or 8,192 bytes

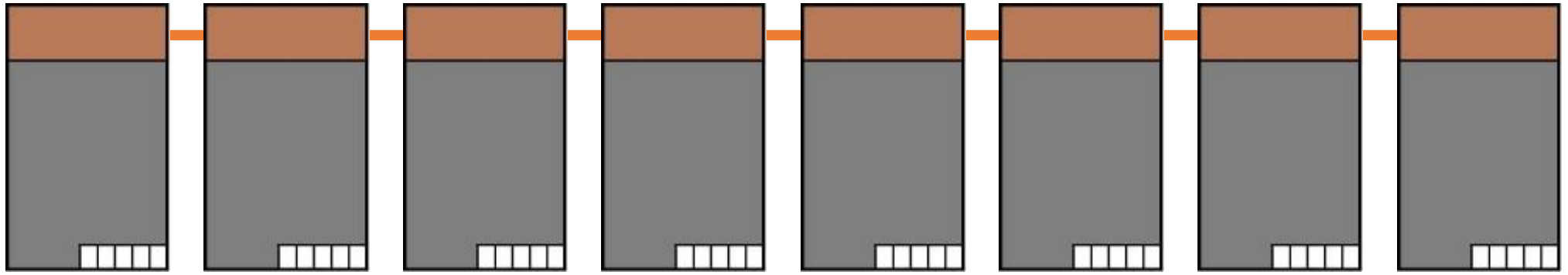


A collection of data pages is…

```
CREATE TABLE MyTable (
    MyTableID INT,
    ColumnOne VARCHAR(15),
    DateCreated DATE
);
```

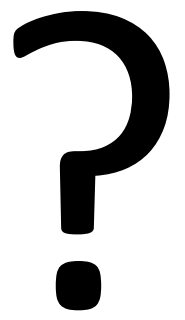
?

Heap

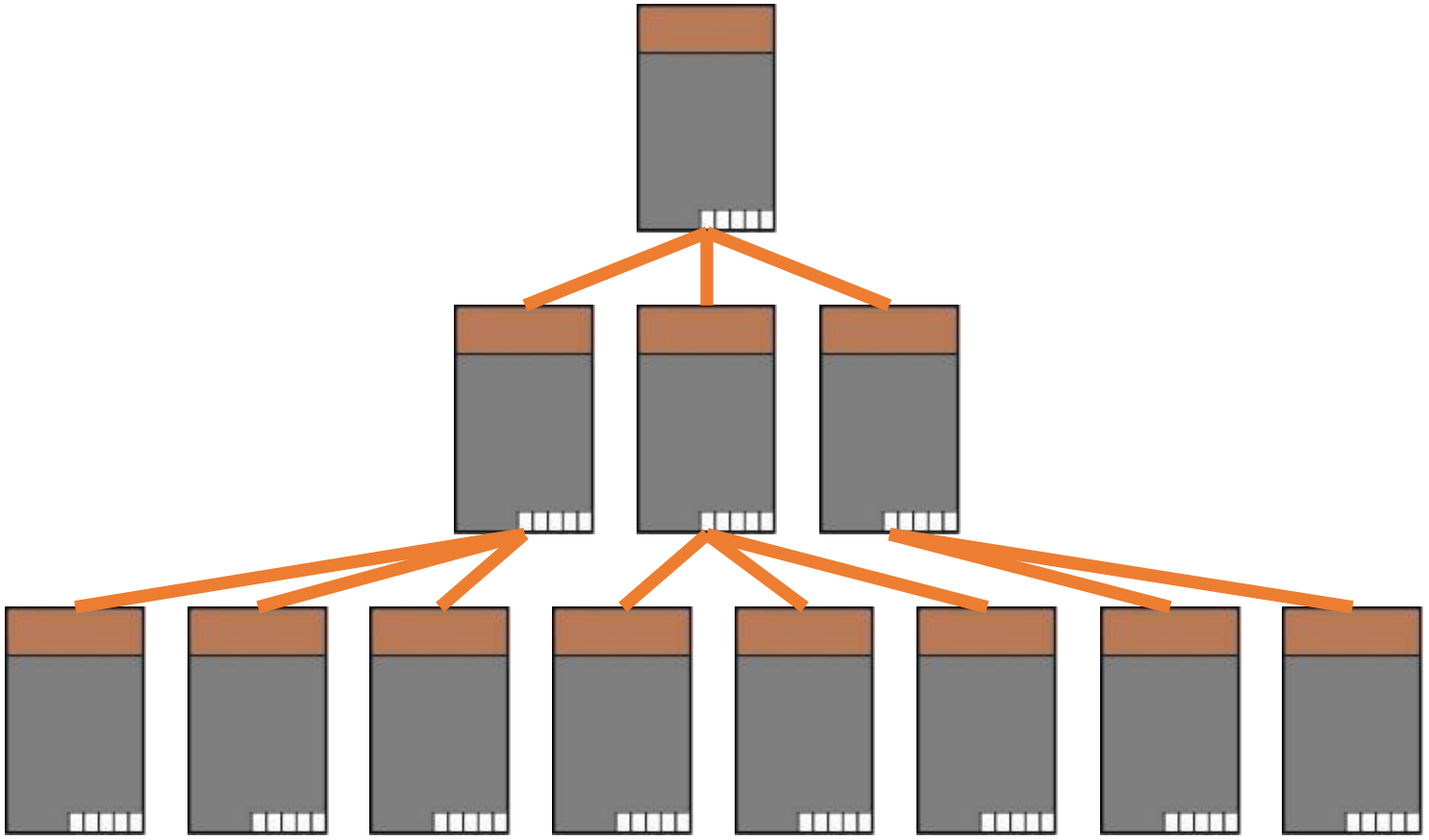


I Want Order!

```
CREATE CLUSTERED INDEX
     CK_MyTable
ON MyTable (
     MyTableID INT
);
```



B-Tree – Clustered & Non-Clustered Idx

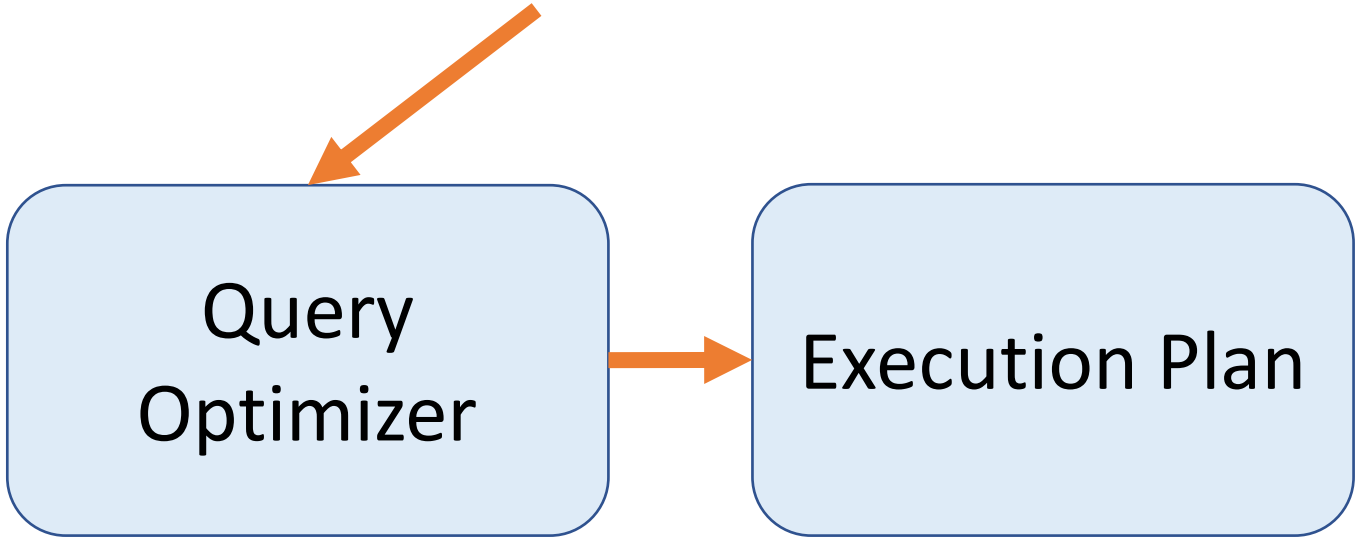


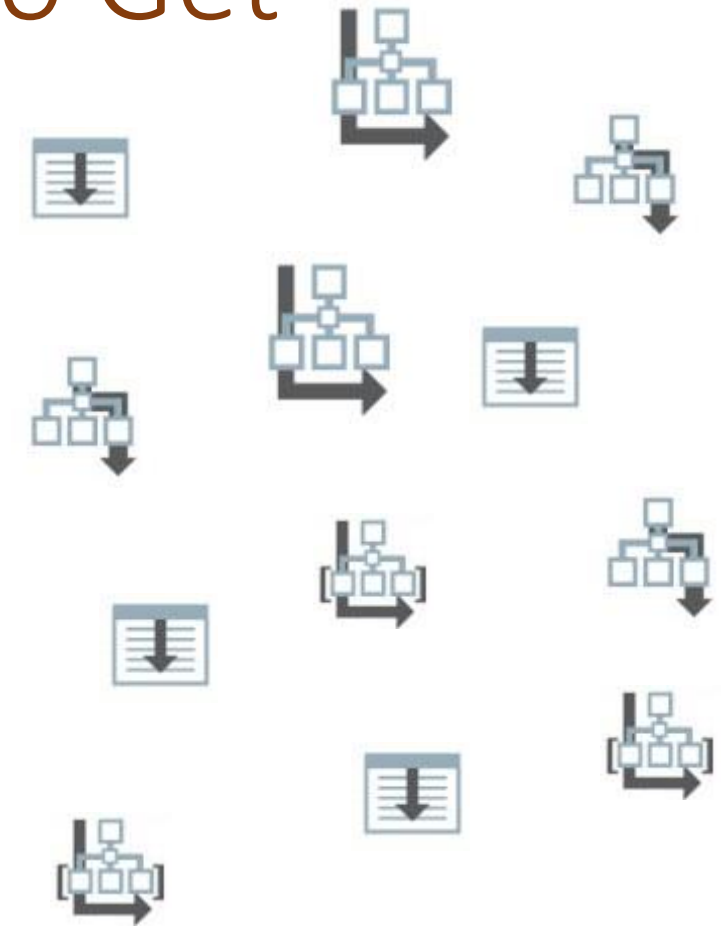
Internals: Navigating Heaps & Indexes

SentryOne™

T-SQL: How To Get vs What To Get

```
SELECT
    ColA, ColB, …
FROM Table_One
INNER JOIN Table_Two
    ON Table_One.Pkey = Table_Two.Fkey
```

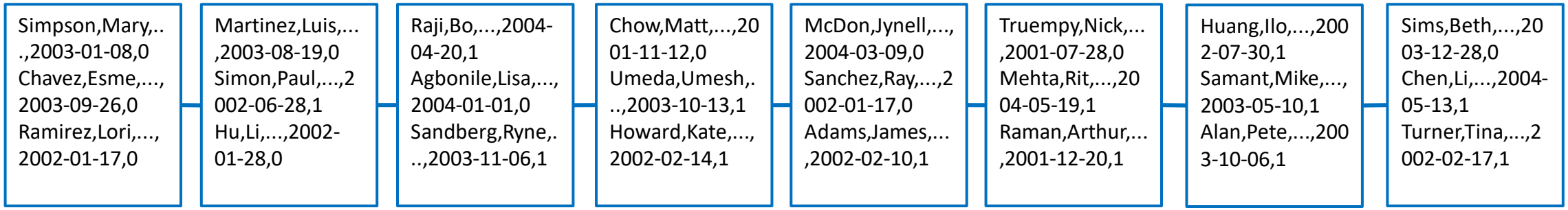




Customer Table

```
CREATE TABLE dbo.Customer (
        FirstName           VARCHAR(50),
        LastName            VARCHAR(50),
        Address             VARCHAR(50),
        City                VARCHAR(50),
        State               CHAR(2),
        ZipCode             CHAR(5),
        Email               VARCHAR(50),
        FirstVisit          DATE,
        RepeatCustomer      BIT
);
```

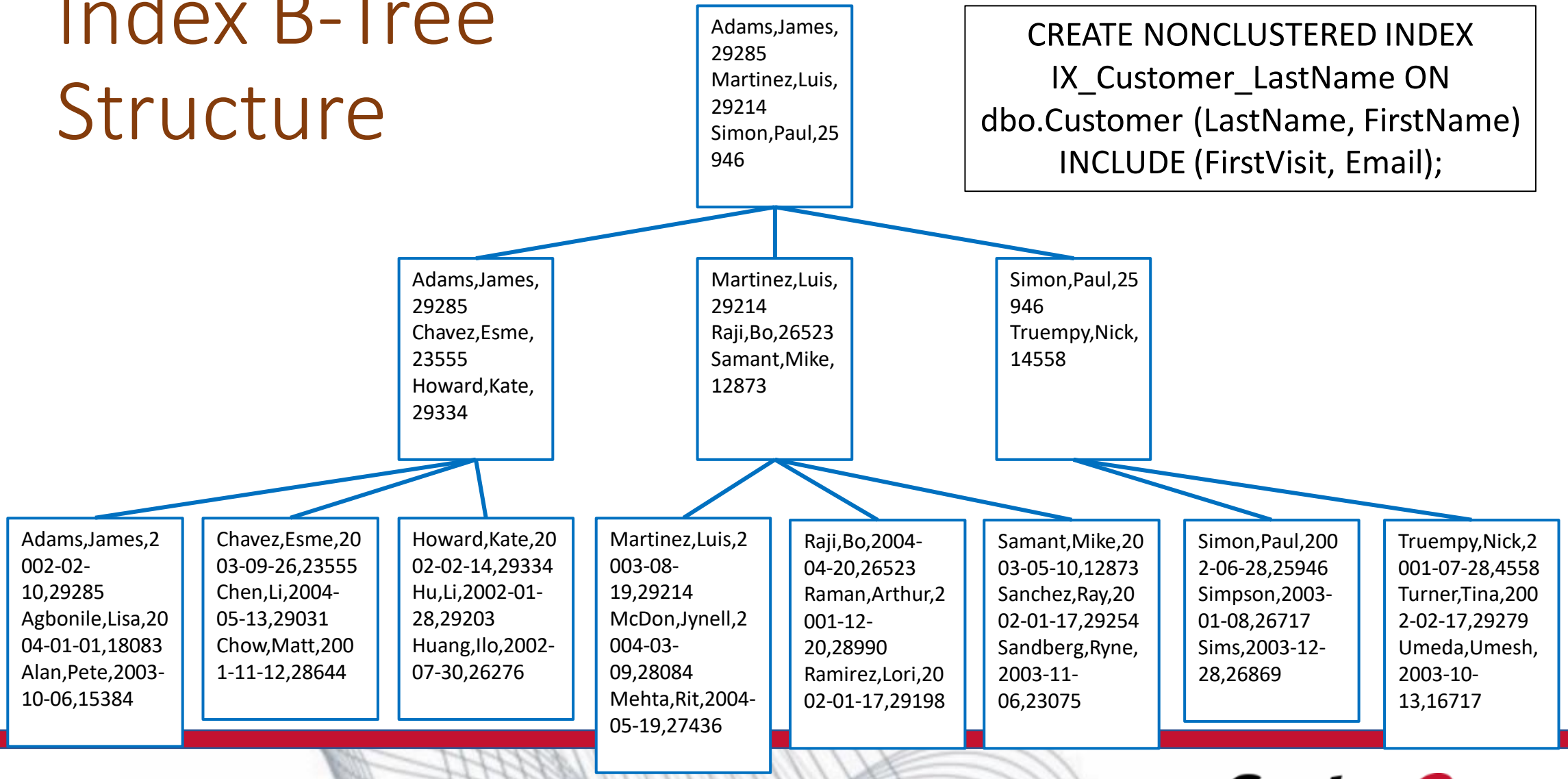
Heap



SentryOne

```
CREATE TABLE dbo.Customer (
        CustomerID          INT IDENTITY(1, 1) PRIMARY KEY CLUSTERED,
        FirstName           VARCHAR(50),
        LastName            VARCHAR(50),
        Address             VARCHAR(50),
        City                VARCHAR(50),
        State               CHAR(2),
        ZipCode             CHAR(5),
        Email               VARCHAR(50),
        FirstVisit          DATE,
        RepeatCustomer      BIT
);
        CREATE NONCLUSTERED INDEX IX_Customer_LastName
            ON dbo.Customer (LastName, FirstName)
                INCLUDE (FirstVisit, Email);
```

Index B-Tree Structure




```
CREATE NONCLUSTERED INDEX
IX_Customer_LastName ON
dbo.Customer (LastName, FirstName)
INCLUDE (FirstVisit, Email);
```

SentryOne

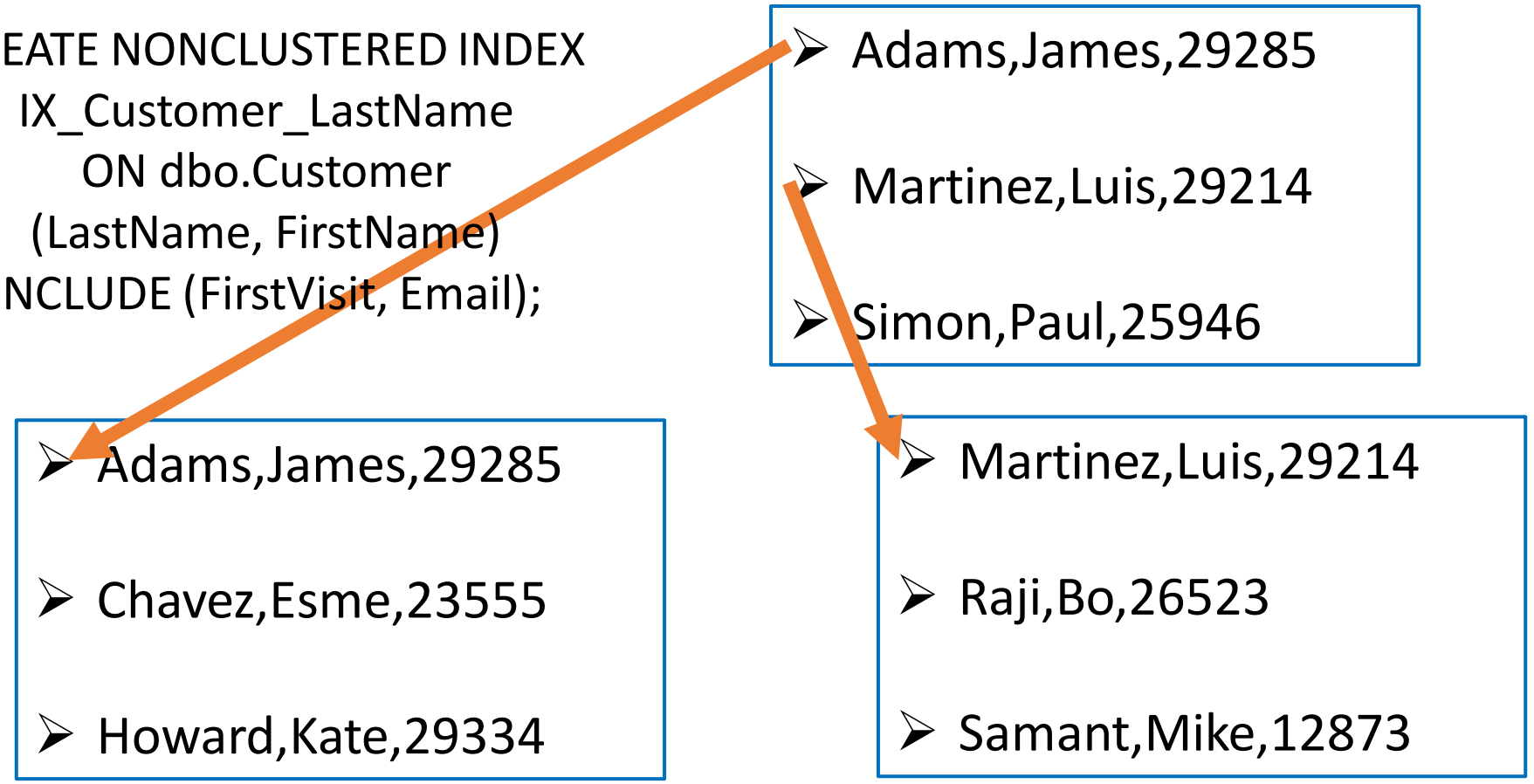
Root Level

```
CREATE NONCLUSTERED INDEX
IX_Customer_LastName
ON dbo.Customer
(LastName, FirstName)
INCLUDE (FirstVisit, Email);
```

- Adams,James,29285
- Martinez,Luis,29214
- Simon,Paul,25946

Intermediate Level

```
CREATE NONCLUSTERED INDEX
IX_Customer_LastName
ON dbo.Customer
(LastName, FirstName)
INCLUDE (FirstVisit, Email);
```



Leaf Level

```
CREATE NONCLUSTERED INDEX
IX_Customer_LastName
ON dbo.Customer
(LastName, FirstName)
INCLUDE (FirstVisit, Email);
```

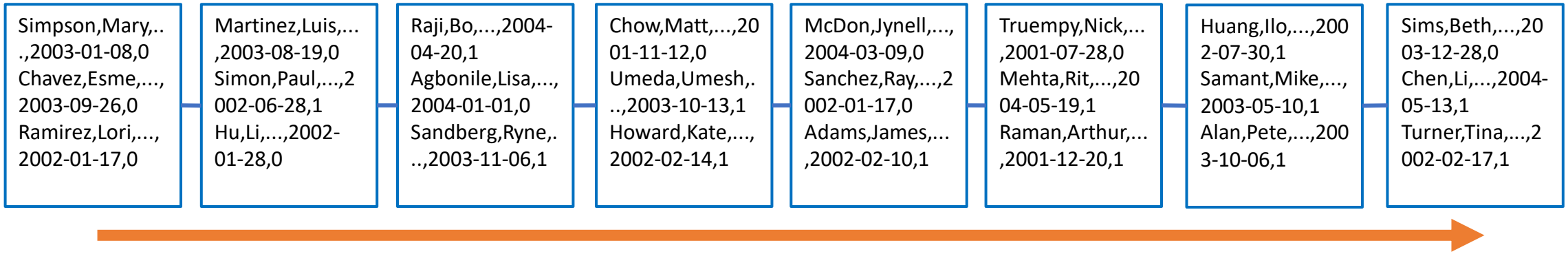
- Adams,James,29285
- Chavez,Esme,23555
- Howard,Kate,29334

- Adams,James,2002-02-10,jada@x.com,29285
- Agbonile,Lisa,2004-01-01,lagb@x.com,18083
- Alan,Pete,2003-10-06,pala@x.com15384

- Chavez,Esme,2003-09-26,echa@x.com,23555
- Chen,Li,2004-05-13,lche@x.com29031
- Chow,Matt,2001-11-12,mcho@x.com28644

Querying a Heap

```
SELECT
  FirstName, LastName, FirstVisit
FROM Customer
WHERE LastName = 'Dean'
```

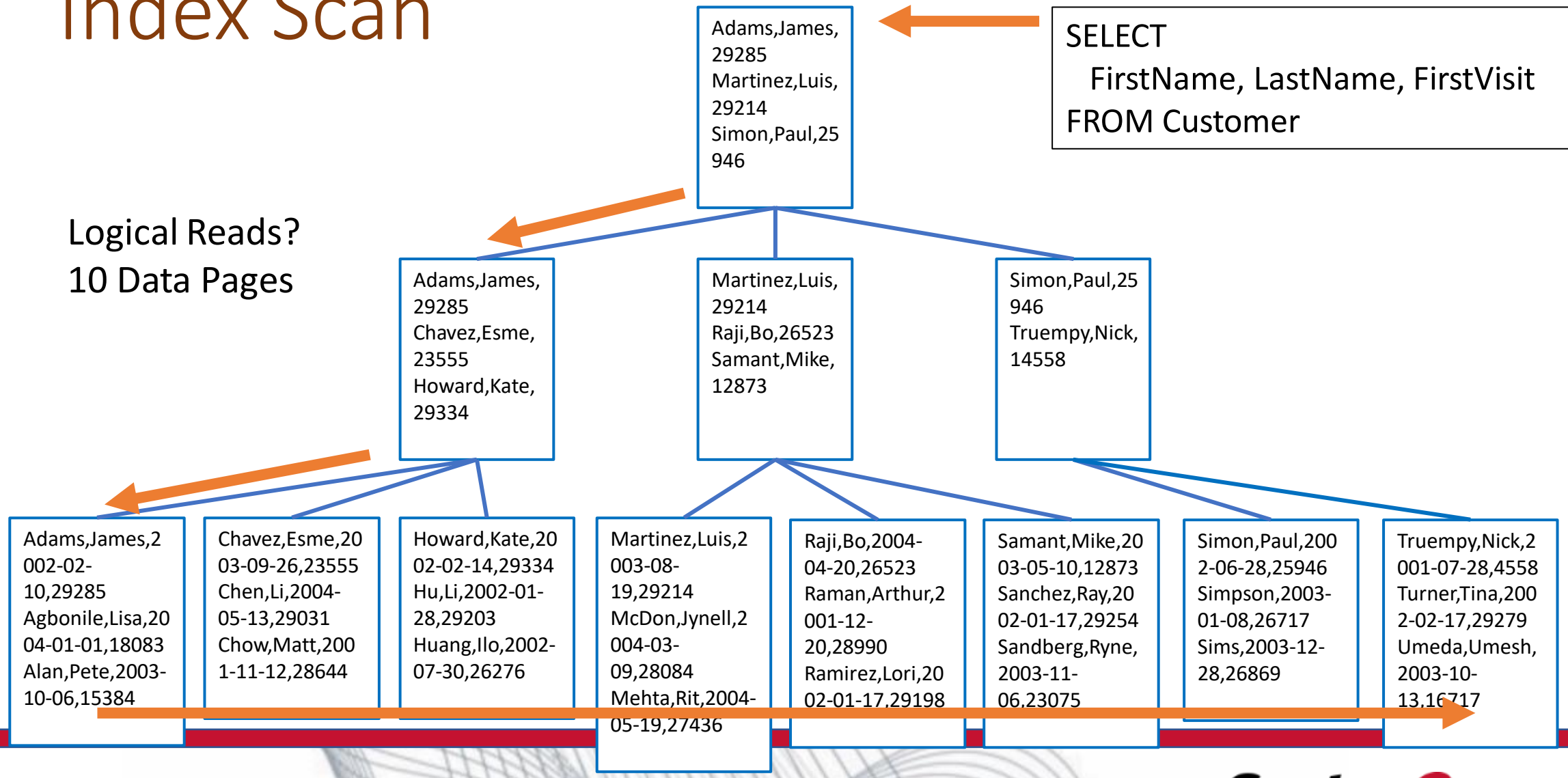


Logical Reads?

8 Data Pages

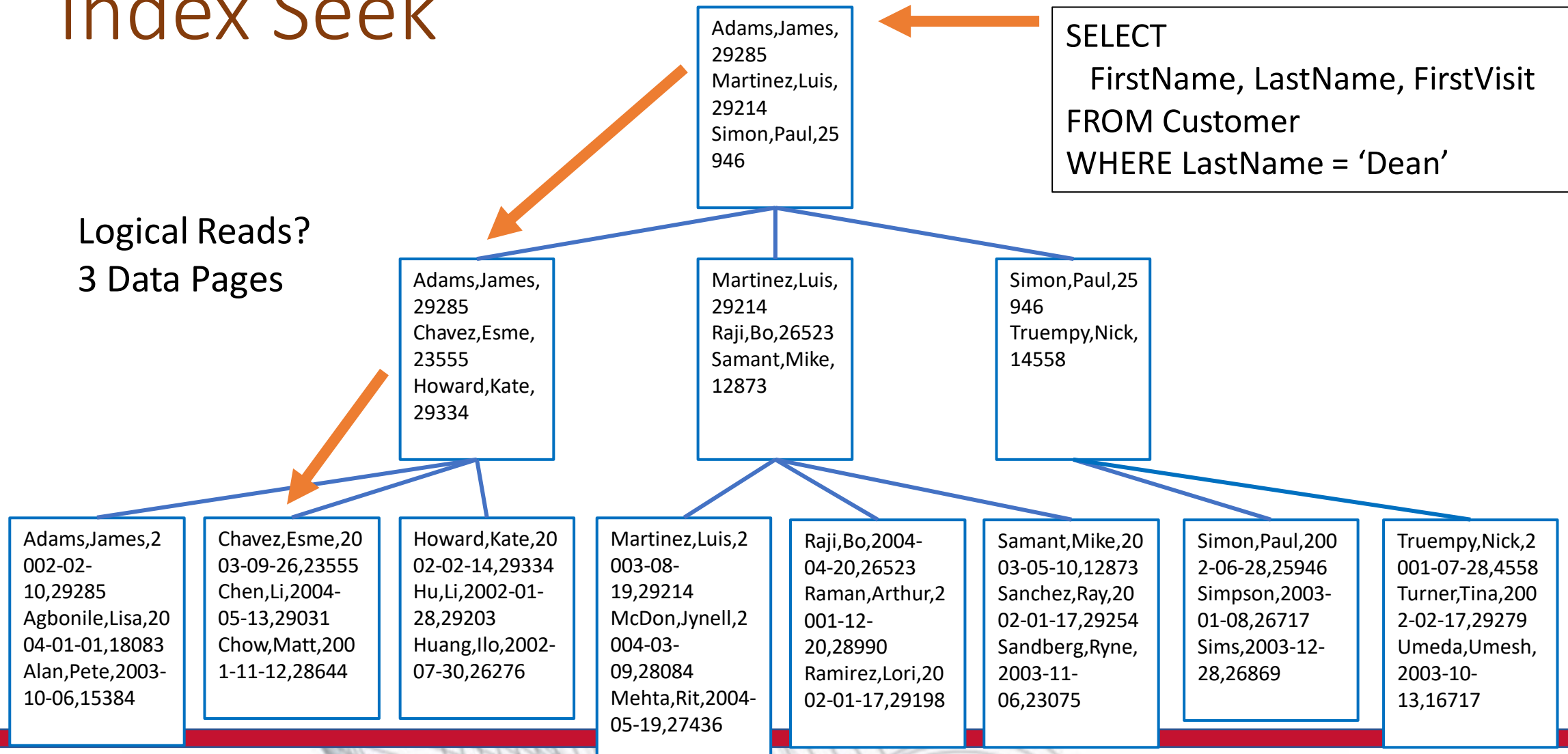
SentryOne

Index Scan



SentryOne

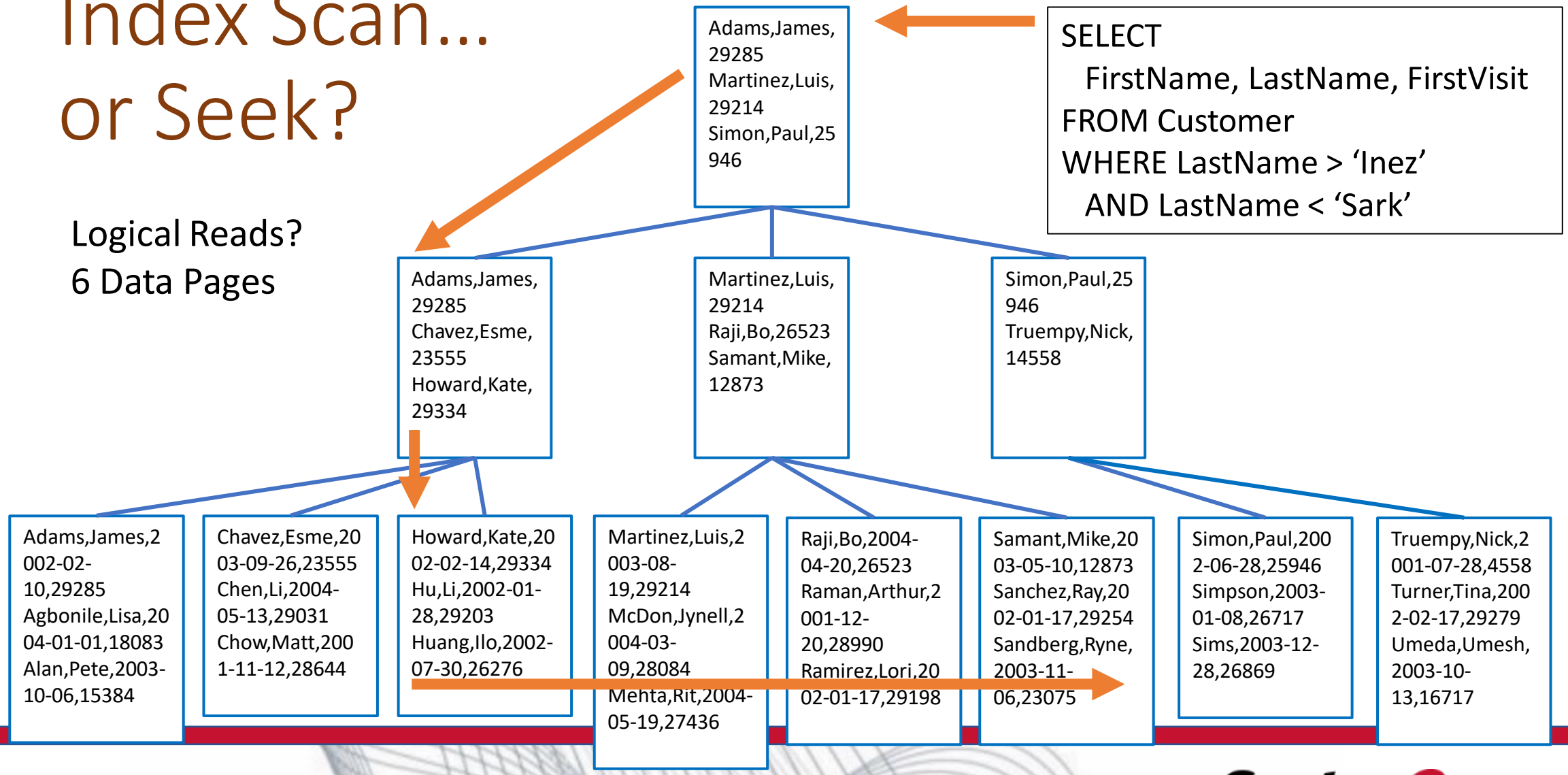
Index Seek



SentryOne

Index Scan… or Seek?

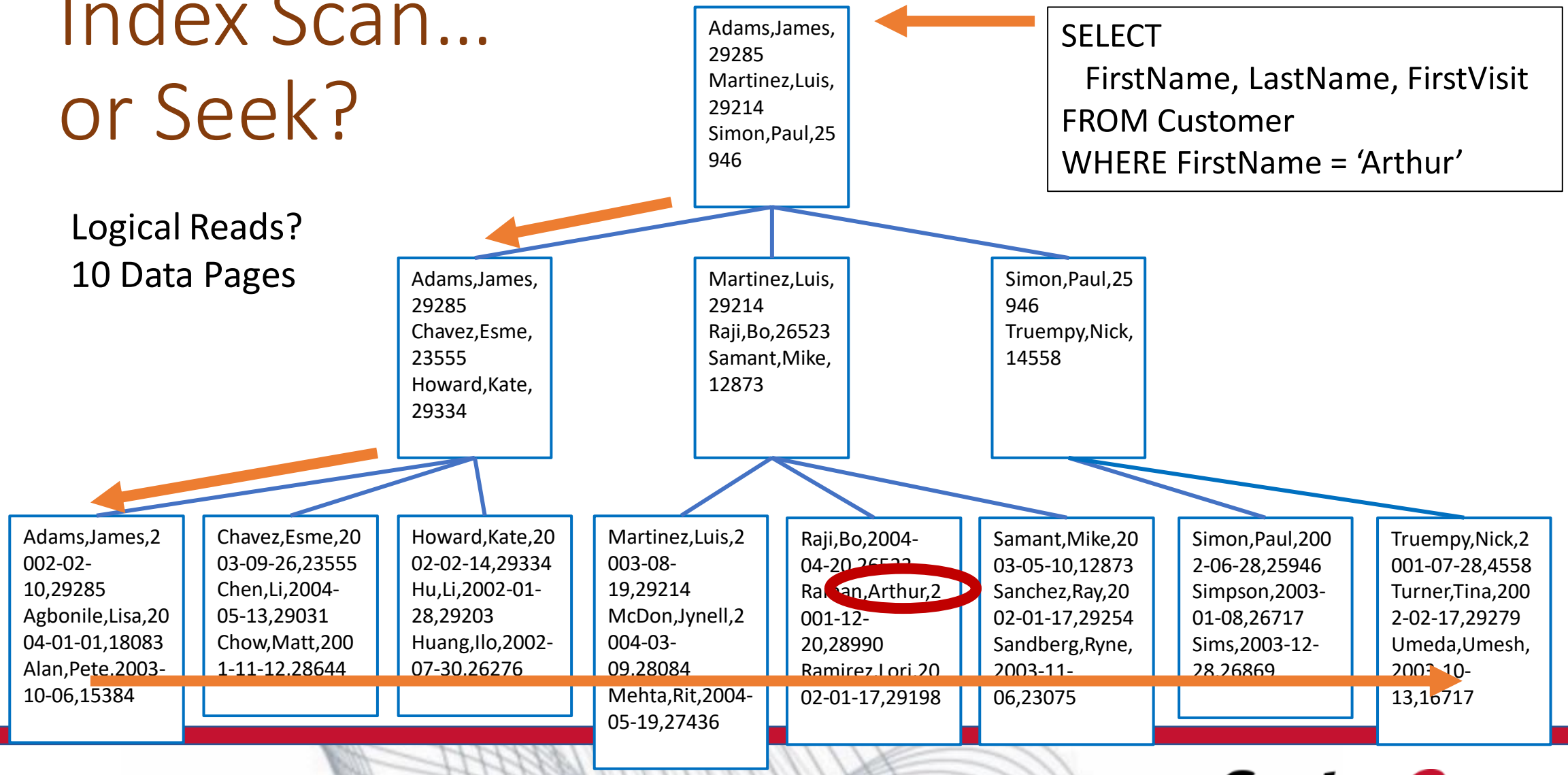
Logical Reads?
6 Data Pages



SentryOne

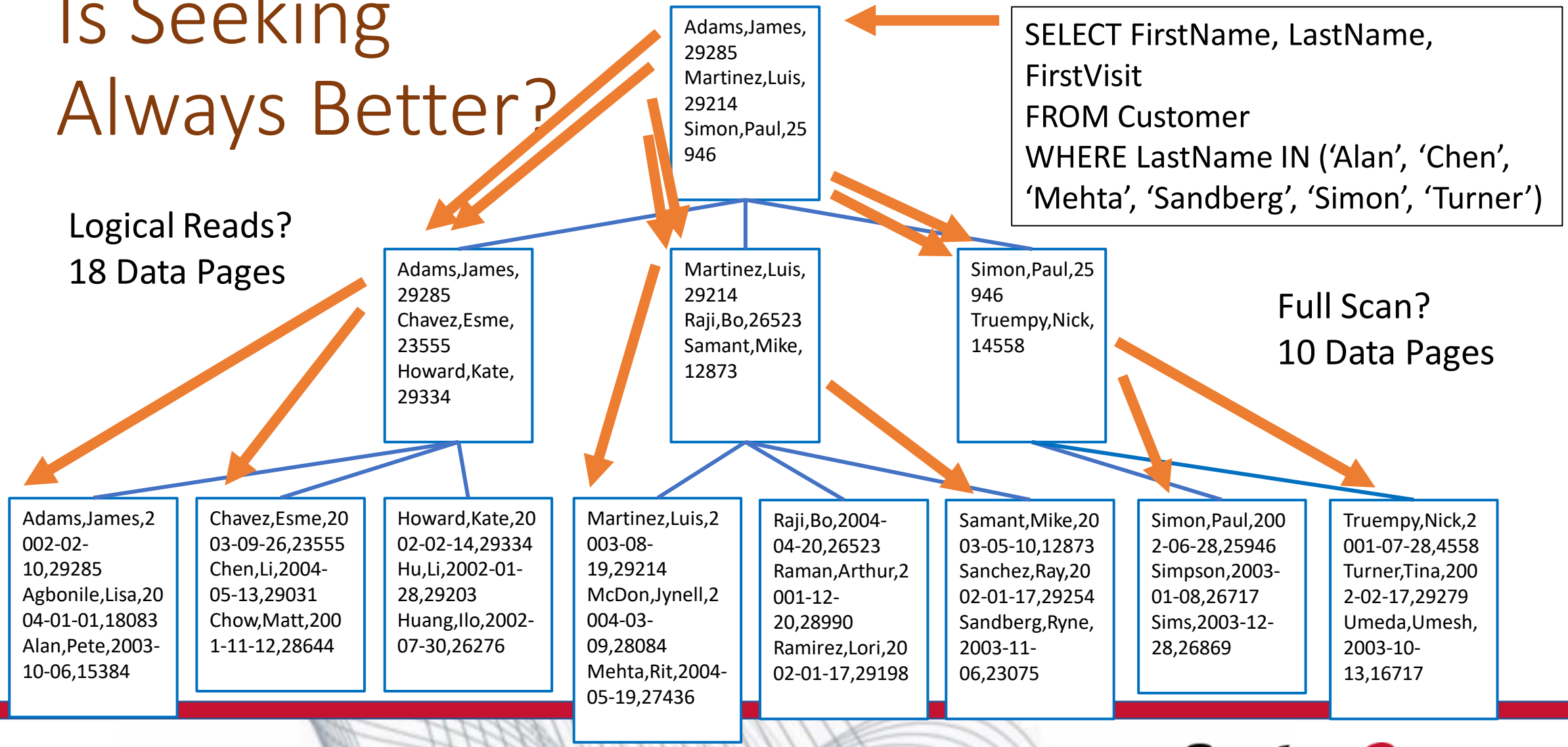
Index Scan… or Seek?

Logical Reads?
10 Data Pages



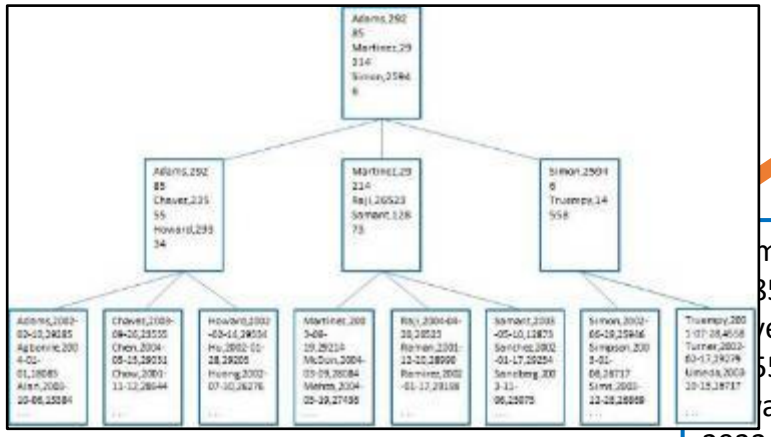
SentryOne

Is Seeking Always Better?



SentryOne

Key Lookup



Adams,James,
29285
Martinez,Luis,
29214
Simon,Paul,25
946

```
SELECT
   FirstName, LastName,
   FirstVisit, City, State
FROM Customer
WHERE LastName = 'Huang'
```

ms,James,
85
vez,Esme,
55
ard,Kate,
29334

Martinez,Luis,
29214
Raji,Bo,26523
Samant,Mike,
12873

Simon,Paul,25
946
Truempy,Nick,
14558

Adams,James,2
002-02-
10,29285
Agbonile,Lisa,20
04-01-01,18083
Alan,Pete,2003-
10-06,15384

Chavez,E me,20
03-09-26,2 555
Chen,Li,2004-
05-13,29031
Chow,Matt,200
1-11-12,28644

Howard,Kate,20
02-02-14,29334
Hu,Li,2002-01-
28,29203
Huang,Ilo,2002-
07-30,26276

Martinez,Luis,2
003-08-
19,29214
McDon,Jynell,2
004-03-
09,28084
Mehta,Rit,2004-
05-19,27436

Raji,Bo,2004-
04-20,26523
Raman,Arthur,2
001-12-
20,28990
Ramirez,Lori,20
02-01-17,29198

Samant,Mike,20
03-05-10,12873
Sanchez,Ray,20
02-01-17,29254
Sandberg,Ryne,
2003-11-
06,23075

Simon,Paul,200
2-06-28,25946
Simpson,2003-
01-08,26717
Sims,2003-12-
28,26869

Truempy,Nick,2
001-07-28,4558
Turner,Tina,200
2-02-17,29279
Umeda,Umesh,
2003-10-
13,16717

SentryOne

Index Operation Recap

- Index Key Order Matters. Can only seek on leading key column
- Add Appropriate Columns as Included Columns to "Cover a Query"
- Seeks are not necessarily better than Scans – Check Properties in Execution Plan

DEMO:
T-SQL & I/O Scenarios

Parting Thoughts

Widen your perspective

Server resources are among your entire workload

Be mindful of “unseen” work being done to fulfil a query

Thank you!

https://github.com/SQLBek

Andy Yun | @SQLBek
ayun@sentryone.com | SQLBek@gmail.com
http://blogs.sentryone.com/andyyun/
http://sqlbek.wordpress.com

Want To Learn More About SentryOne?
Book a demo with me!
http://www.sentryone.com/BookAndy/

SentryOne