24 HOURS OF PASS

The Ambiguous Case of Off-Row Storage in In-Memory OLTP

Dmitri Korotkevitch, aboutsqlserver.com
Moderated By: **Sander Stad**

PREVIEW EDITION

Microsoft

IDERA

ATTUNITY

Hewlett Packard Enterprise

SQLSentry

Technical Assistance

Sample Webinar

1. How would you rate this session overall?
- ○ Excellent
- ○ Good
- ○ Average
- ○ Fair
- ○ Poor

2. How would you rate the speakers' presentation skills?
- ○ Excellent
- ○ Good
- ○ Average
- ○ Fair
- ○ Poor

Submit

Type question here.

Webcams   Zoom: 53%

If you **require assistance** during the session, type your inquiry into the question pane on the right side.

Maximize your screen with the zoom button on the top of the presentation window.

Please fill in the short evaluation following the session. It will appear in your web browser

Thank You

Microsoft

microsoft.com

Empower users with new insights through familiar tools while balancing the need for IT to monitor and manage user created content. Deliver access to all data types across structured and unstructured sources.

IDERA

idera.com

IDERA's award-winning SQL Server database solutions and multi-platform database, application and cloud monitoring tools ensure your business never slows down.

ATTUNITY

attunity.com

Attunity, a leader in data integration and management software, helps move, transform and analyze data efficiently in SQL Server/Azure environments.

Hewlett Packard Enterprise

SQLSentry®

JOIN PASS

PASS is a not-for-profit organization which offers year-round learning opportunities to data professionals

Membership is free, join today at **www.sqlpass.org**

Access to online training and content

Enjoy discounted event rates

Join Local Chapters and Virtual Chapters

Get advance notice of member exclusives

Save on PASS Summit 2016 Registration!

- The world's largest gathering of SQL Server & BI professionals
- Learn from the world's top data experts, in over 190 technical sessions
- More than 4000 attendees from all over the world
- Meet the Microsoft engineering team!

Save $200 right now using discount code **24HOP200!**

$2,195
until September 18, 2016

www.passsummit.com

BIO

20+ Years in IT

15+ Years working with SQL Server

Microsoft Data Platform MVP

Microsoft Certified Master

Author:

- Pro SQL Server Internals
- Expert SQL Server In-Memory OLTP

http://aboutsqlserver.com

dk@aboutsqlserver.com

@aboutsqlserver

The Ambiguous Case of Off-Row Storage in In-Memory OLTP

Dmitri Korotkevitch, aboutsqlserver.com

PREVIEW EDITION

In-Memory OLTP...

(C) Rasmin https://www.flickr.com/photos/rasmin/3484200308/

(C) Neal Gillis https://www.flickr.com/photos/ngillis/8520807103/

In-Memory OLTP...

PASS

Agenda

Hash and nonclustered (range) indexes overview

LOB and row-overflow storage in In-Memory OLTP

Columnstore indexes on memory-optimized tables

In-Memory Data Storage

```sql
create table dbo.Community
(
    Name nvarchar(64) not null
        constraint IDX_NAME primary key nonclustered hash
        with (bucket_count=1024),
    City nvarchar(32) not null,
    FirstName nvarchar(64) not null,

    index IDX_CITY nonclustered hash(City)
    with (bucket_count=128),
)
with (memory_optimized=on, durability=schema_and_data);
```
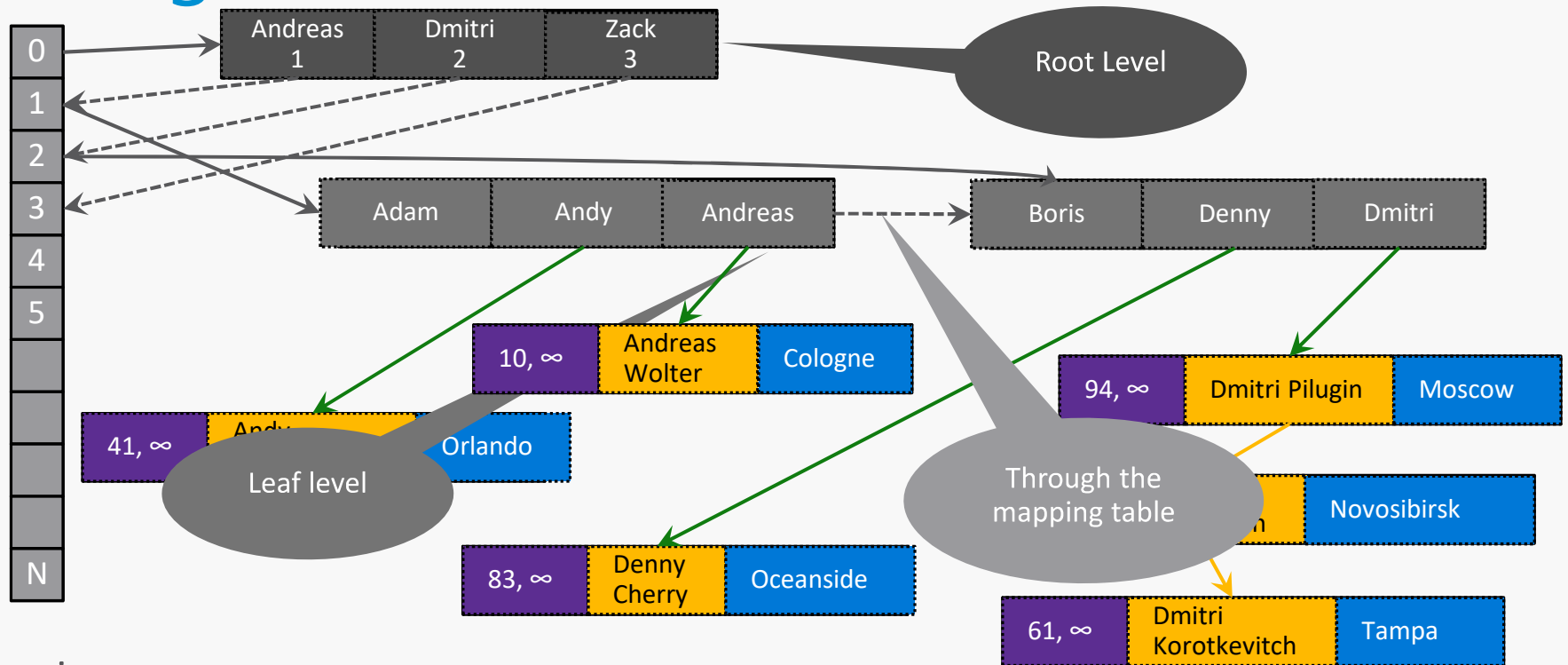
Hash Indexes
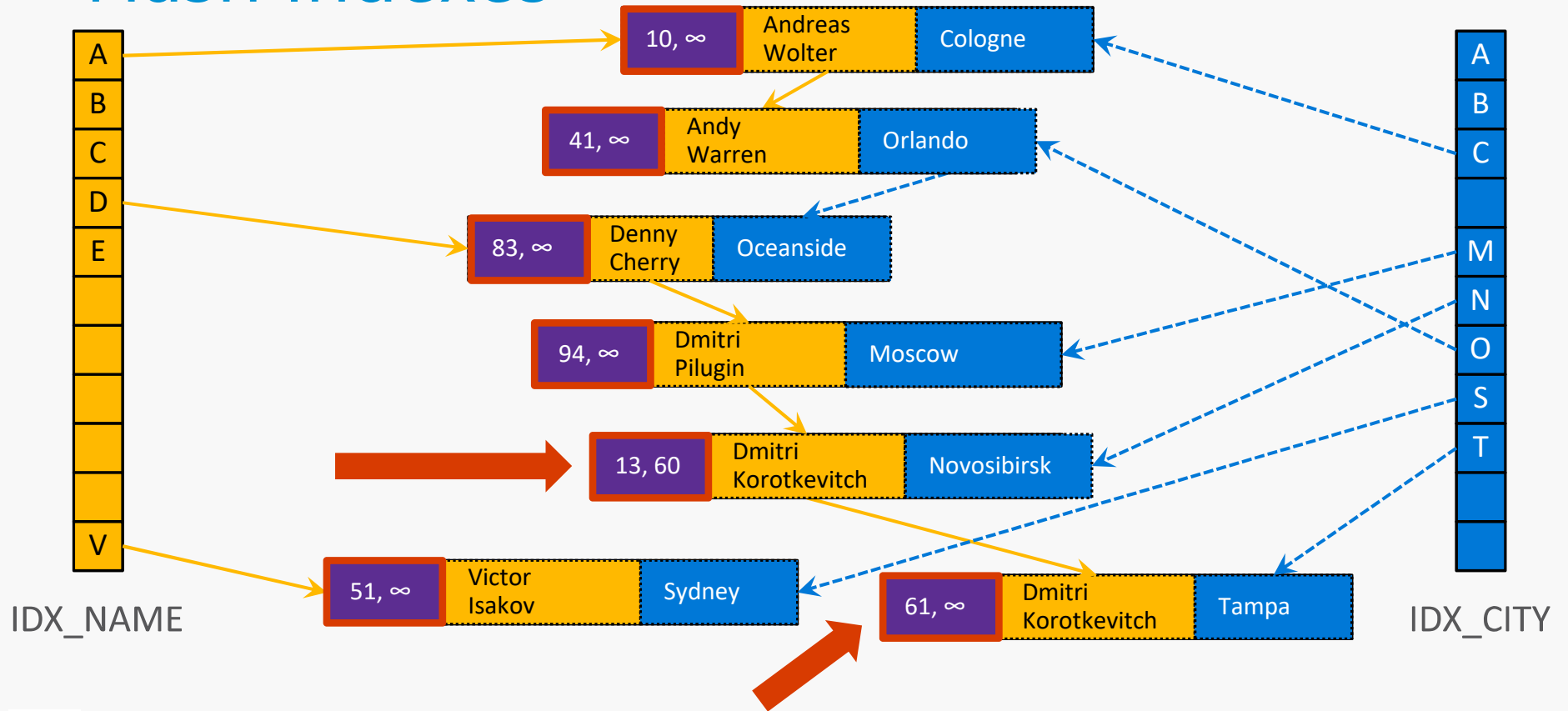
Range Index

```
alter table dbo.Community
    add index IDX_FIRSTNAME
    nonclustered(FirstName);
```

Range Index

Hash vs. Range Indexes

Hash Indexes

- Useful only with point-lookups and equality joins
  - SQL Server needs to calculate the hash of the key
  - All columns from the composite keys must be present
- Faster than nonclustered (range) indexes <u>when</u> bucket_count is correct
  - Ideal bucket_count is 1.5-2 times of the # of unique index keys.
- Good choice for catalog entities

Range Indexes

- Similar to on-disk B-Tree indexes
  - Unidirectional
- Good choice for the range scans
- Should be used when correct bucket_count estimation is impossible

Demo: bucket_count and performance

On-Disk Allocation Units

IN-ROW: Main data row structure and all fixed-length data types

LOB: Variable-length data
- (n)varchar(max); varbinary(max) when size > 8,000 bytes

ROW-OVERFLOW: Variable-length data that does not fit IN-ROW
- (n)varchar(N); varbinary(N)
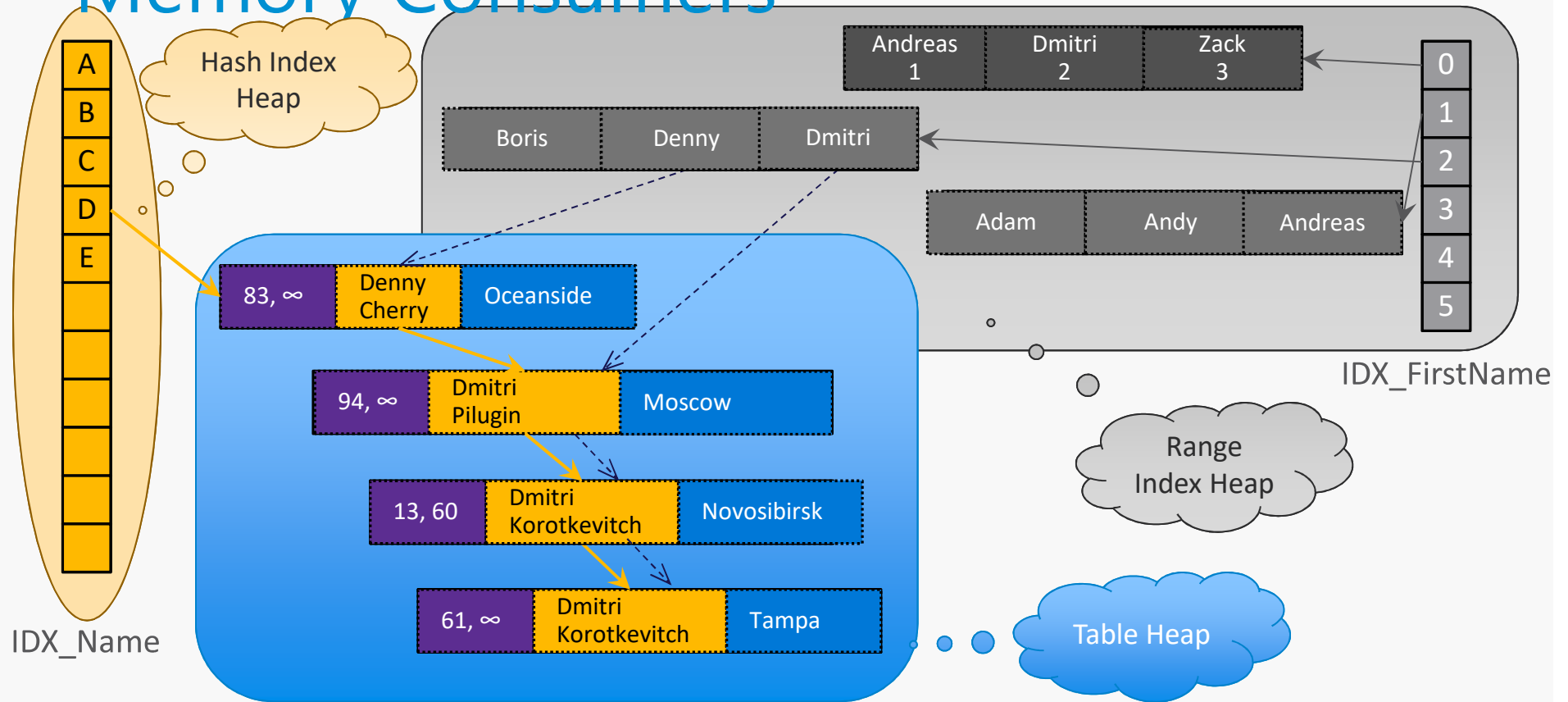- (n)varchar(max); varbinary(max) when size <= 8,000 bytes

Decision is made based on the row/column data

In-Memory OLTP Memory Consumers

In-Memory OLTP objects allocate memory from separate VARHEAPs

- Data rows
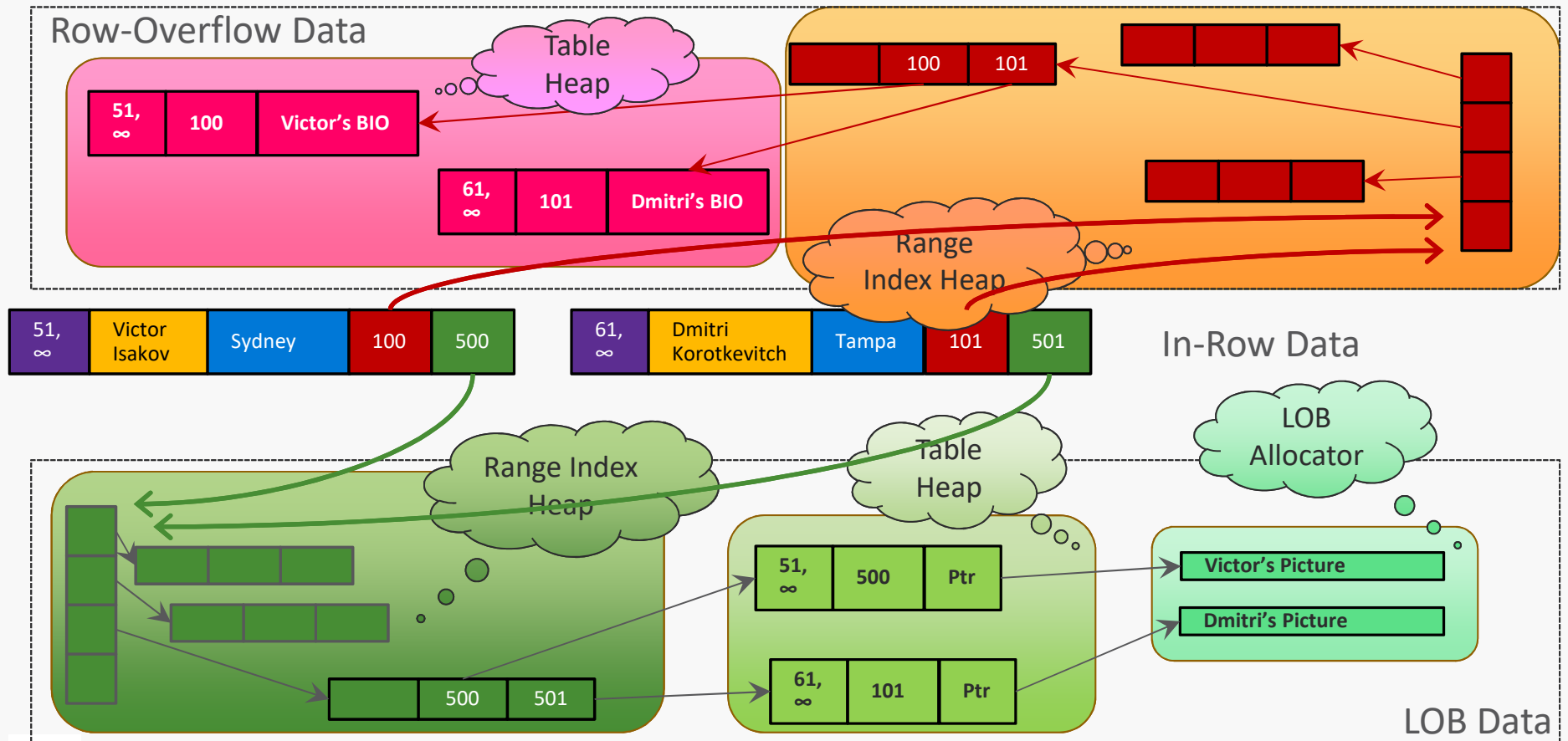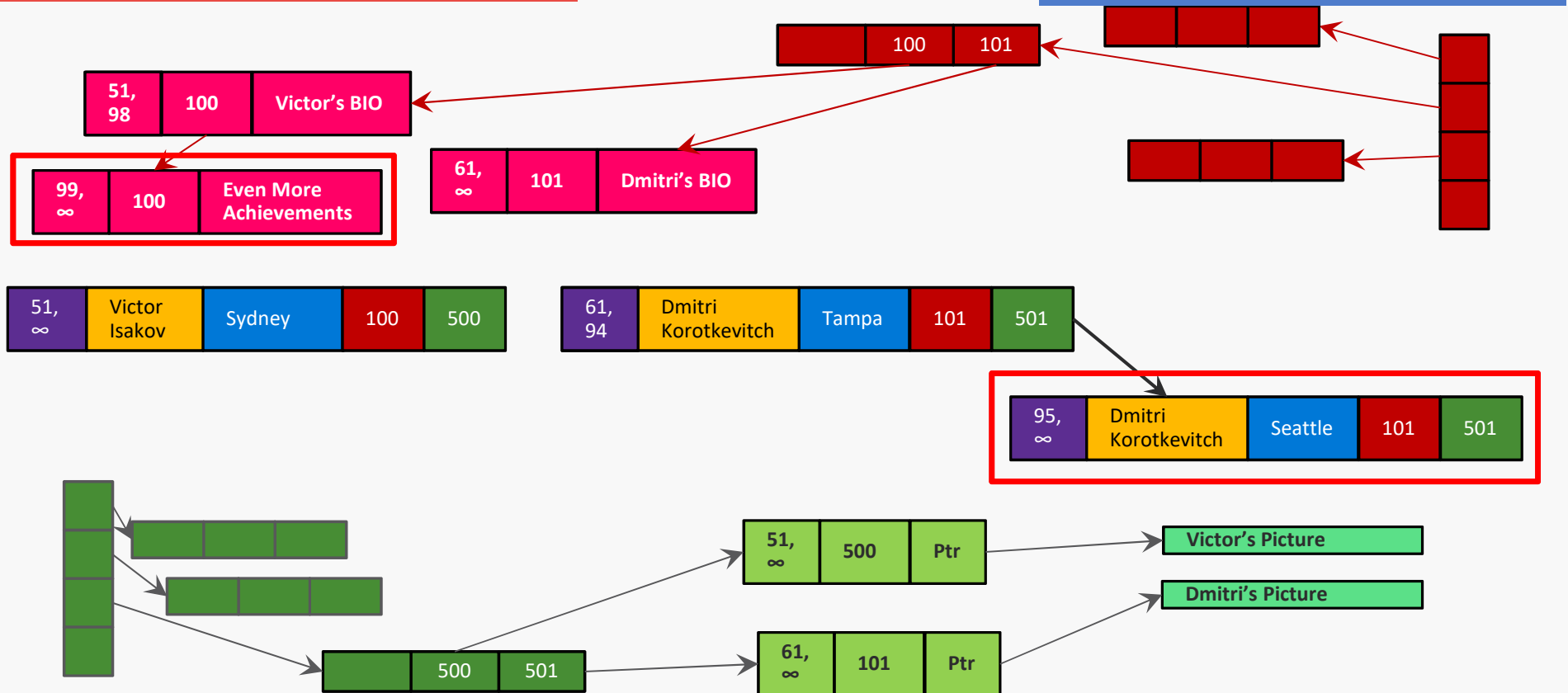- Hash tables in Hash indexes
- Range indexes

Memory Consumers

Hash Index Heap

Range Index Heap

Table Heap

IDX_Name

IDX_FirstName

Andreas 1	Dmitri 2	Zack 3

Boris	Denny	Dmitri

Adam	Andy	Andreas

| 0 |
| 1 |
| 2 |
| 3 |
| 4 |
| 5 |

83, ∞	Denny Cherry	Oceanside

94, ∞	Dmitri Pilugin	Moscow

13, 60	Dmitri Korotkevitch	Novosibirsk

61, ∞	Dmitri Korotkevitch	Tampa

A
B
C
D
E

Demo: Memory Consumers

Off-Row Storage

Row-Overflow Data

Table Heap

| 51, ∞ | 100 | Victor's BIO |

| 61, ∞ | 101 | Dmitri's BIO |

| | 100 | 101 |

Range Index Heap

| 51, ∞ | Victor Isakov | Sydney | 100 | 500 |

| 61, ∞ | Dmitri Korotkevitch | Tampa | 101 | 501 |

In-Row Data

Range Index Heap

Table Heap

LOB Allocator

| 500 | 501 |

| 51, ∞ | 500 | Ptr |

| 61, ∞ | 101 | Ptr |

Victor's Picture

Dmitri's Picture

LOB Data

Demo: Off-Row Columns and Performance
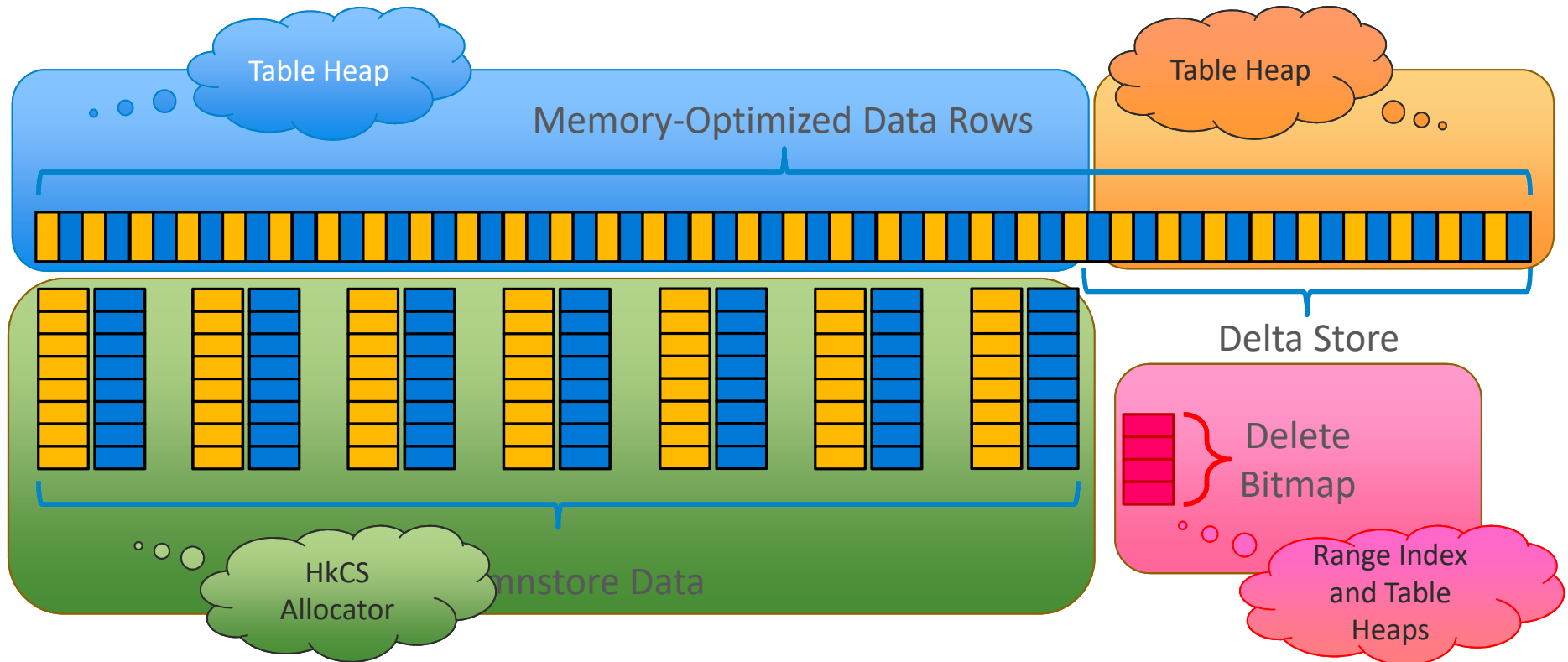
Off-Row Columns

Stored in the separate internal tables
- Referenced through artificial internal keys
- Have separate lifetime from in-row data
- Decided at CREATE TABLE phase based on data type/column size

Add significant overhead
- 50+ extra bytes per each value
- Slower INSERT/DELETE operations
- Slower SELECT: Not covered by In-Row indexes

Demo: Columnstore Indexes

Columnstore Indexes

Delta Store is another Table Heap with memory-optimized rows
- Tuple mover compresses the rows and moves them to the main Table Heap when it estimates ~>1M rows in delta store
- COMPRESSION_DELAY index option delays row compression

Delete Bitmap is the range index

Large Delta Stores and Delete Bitmaps negatively affect performance

Tables with columnstore indexes do not support off-row storage

Takeaways

Do not use HASH indexes unless you can correctly estimate and maintain the bucket_count

Avoid off-row storage in In-Memory OLTP
- Defining columns as varchar(max) *just because* is clearly the bad idea
- Performance critical cases can benefit from multi-column design

Control the size of columnstore index delta store and delete bitmap with COMPRESSION_DELAY option

We covered

Hash and nonclustered (range) indexes

Off-row data storage in In-Memory OLTP

Columnstore indexes in In-Memory OLTP

Email: dk@aboutsqlserver.com
Blog: http://aboutsqlserver.com

Questions?

Coming up next!

Understanding and Monitoring Tempdb
Vicky Harp

Microsoft
IDERA
ATTUNITY
Hewlett Packard Enterprise
SQLSentry

PREVIEW EDITION PASS SUMMIT

Follow **@pass24hop**

Share your thoughts with **#pass24hop** & **#sqlpass**

Thank You for Attending

PREVIEW EDITION **PASS** SUMMIT

PREVIEW EDITION

PASS
SUMMIT