

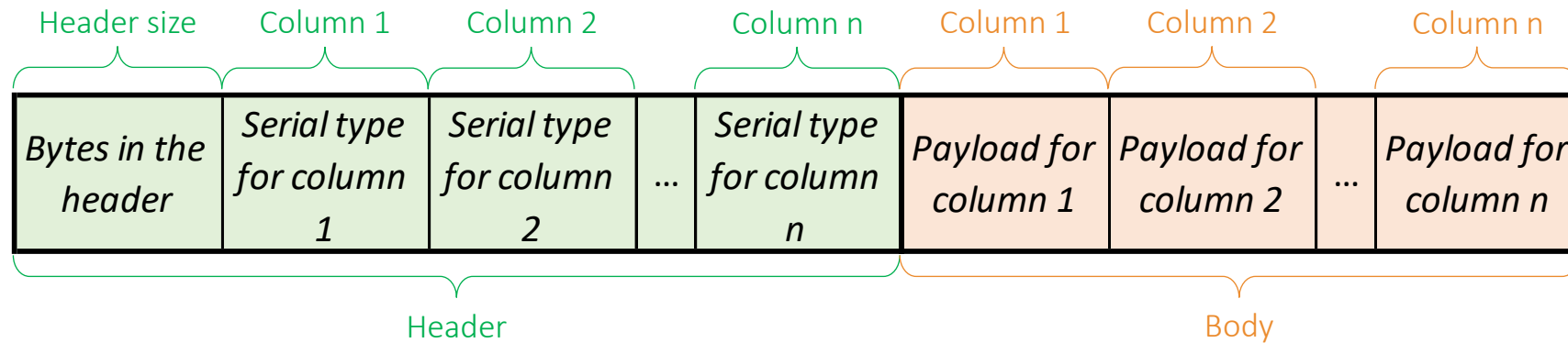


Architektur und Implementierung von Datenbanksystemen

Task 2

Team 3 - Gründlinger Diana, Targa Aaron, Klotz Thomas, Huber Marcel, Thalmann Matthias

SQLite Record Format



- Header
 - *Header size (varint)*: Length of the header in bytes (including header size field)
 - *One varint per column*: Serial type, defining the type and length of the column
- Body
 - *Values for the columns* → not each column must have a value

Serial Type Codes (1)

Serial Type	Column size	Meaning
0	0	<i>NULL</i>
1	1	8-bit integer
2	2	16-bit integer
...		
7	8	64-bit float
8	0	Integer with value 0
9	0	Integer with value 1
$N \geq 12$ and even	$(N-12)/2$	Blob
$N \geq 13$ and odd	$(N-13)/2$	String

Serial Type Codes (2)

- Type defines the size of the column in bytes
- Multiple types of integer (8/16/24/32/48/64 bits)
 - Big-endian integers → Highest value byte on smallest memory address
- 10/11 are reserved for internal use
- Strings are stored without the *nul* terminator
- Columns with a serial type of 0, 8, 9, 12 and 13 have a length of 0 in the body

Variable Length Integer (varint)

- 64-bit integer
- Big endian
- Huffman encoded → less bits for small positive values
- Between 1-9 bytes in length

Example Record (1)

PersonID	Name	Birthday	PhoneNumber
1	Max Mustermann	2000-01-01	NULL
2	Bob Kaufmann	1987-04-12	+43 0664 123 45 67
3	Emma Resch	1991-12-02	NULL

Header					Body			
Header size	Column 1	Column 2	Column 3	Column 4	Column 2		Column 3	
5	9	41	33	0	77	... 110	50	... 49

Column 1 has a value of 1
(integer)

Character 77 in UTF-8 = M

Column 4 is *NULL*

Column 3 has a string with 10 bytes

Column 2 has a string (since ≥ 13 and odd) with 14 bytes;
Assuming that it is encoded in UTF-8, each of the "basic" characters takes up 1 byte

$$\frac{33 - 13}{2} = 10$$

Example Record (2)

PersonID	Name	Birthday	PhoneNumber
1	Max Mustermann	2000-01-01	<i>NULL</i>
2	Bob Kaufmann	1987-04-12	+43 0664 123 45 67
3	Emma Resch	1991-12-02	<i>NULL</i>

Header					Body				
Header size	Column 1	Column 2	Column 3	Column 4	Column 2			Column 3	
5	9	41	33	0	77	...	110	50	... 49

Header size of 5 bytes

Column 1 is empty, since it is 1
(defined in header)

Column 4 is empty, since it is *NULL*
(defined in header)

Reference to our project

- **int** → one of the integer types (serial type 1-6)
 - if the value is 0 or 1, serial codes 8/9 are used
- **float** → serial type 7
- **varchar** → string (serial type ≥ 13 and odd)

