Meeting 12/20/17

Tuesday, December 19, 2017 10:22 PM

Agenda:

- Figure out how multi-byte tokens for token tables with more than 256 tokens work
- Figure out how Oracle distinguishes between token bytes and row data
- Learn more about block header data

Attendees:

Andy
Trevor, Dylan

Notes:

- Found that Oracle uses bytes fa - ff for special cases
  - Fa starts a multi-byte column count
  - Fb starts a multi-byte token number
  - Each of these is followed by the two bytes that represent the column count or token respectively
  - Ff represents a null column. No column count Is used with this case
  - Ff can be used between two non null columns to represent a null column as well
  - Fc, fd, and fe could possibly be for other special cases but no such cases were found
- For single byte values, Oracle knows any byte from 00 to c7 is a token value and any byte from c9 to f9 is a column count
- For single-byte column count values, the value is the column count plus 200
- The use of c8 can not yet be determined. It may never be used
- Fa - ff can be used in the second two bytes of a multi-byte token or column count as simply their number value
- Using this encoding, Oracle can always tell if something is a token byte or column count byte and how much is actual column data
- Also found examples of rows containing more than one token
- Found that all tokens in a row don't have to come before all actual row data. The row can have a token, then actual data, then another token.
- For header data, found that the itl's always start at exactly the same point in each block's hex dump. There are always exactly 44 bytes before the itl's start, which means the rest of the block header data is always in that 44 bytes.