Upgrading from Firebird 1.5 to 2.5 and beyond

Pavel Cisar, IBPhoenix

Firebird Conference 2016, Prague

You certainly know the sugar, but what is the price?

Agenda

- Problems you may face
- Planning your upgrade
- Upgrading database
- Upgrading applications
- Changes in configuration

Problems you may face

Do not be intimidated by their number!

Problem categories

- Data (intl, types, conversions etc.)
- SQL (rules, syntax and semantics)
- PSQL (rules, syntax and semantics)
- Optimizer
- Security (changes)
- Engine (configuration, behaviour)
- UDF's
- Loopholes closed
- New "undocumented features" (a.k.a. bugs)

Problems arising in 2.0

- new reserved words and changes (namely CLOSE, OPEN, LEADING, TRAILING, ROWS, LOWER, TRIM, BOTH, ROW_COUNT, HOUR, DAY, MONTH, YEAR)
- DSQL parsing of table aliases is stricter (no ambiguties)
- Multiple hits to same column now illegal in UPDATE
- views made updatable via triggers no longer perform direct table operations
- CHECK constraint now allows NULL value
- change in concatenation overflow check (from compile to run-time)

- stricter validation of user-specified plans + they must refer to all tables in query

  enhancements to NULL Logic

  - New [NOT] DISTINCT test treats two NULL operands as equal

  - NULL comparison rule relaxed

-
  - NULLs ordering changed to comply with standard (now lowest)

- OLD context variables are now read-only. NEW context variables are now read-only in AFTER-triggers as well.

- ORDER BY <ordinal-number> now causes SELECT * expansion

- improved type coercion in UNIONs

- new string functions (that may collide with your UDF's, namely LOWER and TRIM)

- changes in NULL processing in UDF's

  UDFs added and changed:

  - IB_UDF_rand() and IB_UDF_srand()

-
  - IB_UDF_lower() (collision with LOWER)
- new INTL interface for Non-ASCII character sets (initial)

- restrictions on certain PDB parameters for regular users

- optimizer changes (namely predicates NOT IN and ALL may be slow)

- changed GC in SS (cooperative, background, combined)

  changes in configuration

  - ExternalFileAccess

  - LegacyHash

-
  - TCPNoNagle has changed

  - Security in Firebird 2

  - Direct connections to the security database are no longer allowed

  - Non-SYSDBA users no longer can see other users' accounts in the security database

  - Remote attachments to the server without a login and password are now prohibited

  - new security database (security2.fdb)

Problems arising in 2.1

- Metadata text conversion (Firebird versions 2.0.x had two problems related to character sets and metadata extraction)

- new reserved words and changes (namely START)

- New built-in functions (a lot of them!)

- Enhancements to functions (EXTRACT(), DATEADD(), DATEDIFF(), TRUNC())

- Type flag for stored procedures (selectable, non-selectable, legacy)

- RDB$DB_KEY returns NULL in outer joins
  To conform with standards, the character length of the result of applying SUBSTRING() to a VARCHAR or CHAR is now a VARCHAR of the same character length declared for or deduced from the value in the first argument.

- Data type of an aggregation, CASE or UNION result of mixed data.

- Changes in optimizer

- New INTL interface for non-ASCII character sets (ICU all the way)

- MaxFileSystemCache, DatabaseGrowthIncrement
- IPCName behaviour has changed

- SortMemBlockSize changed to TempCacheBlockSize

- SortMemUpperLimit changed to TempCacheUpperLimit

- forced writes on linux now works!

- Enlarged disk allocation chunks

- Bypass Filesystem Caching on Superserver

Problems arising in 2.5

- reserved words

- statements now return an SQL-2003 standard 5-alphanumeric SQLSTATE completion code

- changed logic in (UPDATE) SET clause (see OldSetClauseSemantics config. option)

- UDF Safeguard (memory check)

- rejection of inconsistent TPB options

- The parameter MaxFileSystemCache has been renamed to FileSystemCacheThreshold.
  Default Database Location. The configuration parameter DatabaseAccess now has more "meaning" attached to it. In the absence of any other indication, the first location defined in the "Restrict" list for DatabaseAccess is taken by the engine as the default location for creating a new database and for locating a database where the
- connection parameters do not specify either an alias or the full path specification.

- Parameters LockSemCount and LockSignal are deprecated.

Problems arising in 3.0

- Reserved words (namely BOOLEAN, OFFSET, SCROLL, OVER, TRUE, FALSE, UNKNOWN, ROW, RETURN)

- "Per-Database" Configuration

- New security (incl. multiple security databases, SQL to manage users, authentication etc.)

- an authenticated login is no longer required for embedded connection on POSIX platforms

- System tables are now Read-only
- Improperly mixed explicit and implicit joins are no longer supported.

- The names for column and table aliases and for local variables names in PSQL are now restricted to 31 bytes in length.

- User names are treated as SQL identifiers and are accordingly now limited to 31 bytes in length.

- Cursor Output Columns Must Be Named (aliased, important for expressions)

- aliases.conf is replaced by databases.conf
- The previously deprecated firebird.conf parameters CompleteBooleanEvaluation, OldColumnNaming and OldSetClauseSemantics are no longer supported anymore and have been removed.

- Some statements may now work differently due to the "cursor stability" improvement. Statements affected will be:

  - • those that modify the table that is being explicitly or implicitly selected from within the same statement
- • (as a side effect) some MERGE statements, that might work differently if multiple matches are possible.

Planning your upgrade

Basic steps

- Choose your target version (upgrade or rewrite?)

- Reserve your time and other resources

- Write down a check list

Choose your target version

- From 1.5
    - • Moving to 2.5 is tough upgrade
    - • Moving to 3.0 is better as full rewrite
- From 2.0
    - • Moving to 2.5 is moderately complex upgrade
    - • Moving to 3.0 is tough upgrade
- From 2.1
    - • Moving to 2.5 is straightforward upgrade
    - • Moving to 3.0 is complex upgrade

Making a detailed plan

- Start with UDF's
- Raw database upgrade
- Moving data
- Database verification
- Handling security issues
- Engine configuration
- Checking your application(s)
- Performance testing
- Revisiting (P)SQL

UDFs

- 64-bit version
- Can you replace some with new internal functions or other new feature?
- Separate testing under new FB version

Raw database upgrade

- Start with metadata script
- Execute it untill you encounter first problem. You may use isql with -b switch or write your own tool.
- Note the problem, fix it and repeat

Most typical problems

- UDF vs. internal function collision
- Collision with reserved words
- Count of column list and variable list do not match (because of *)
- Incomplete user PLAN (table <XXX> is not referenced in plan)
- Invalid expression in the ORDER BY clause (column name vs. aggregate)
- attempted update of read-only column (OLD. & .NEW)
- Count of read-write columns does not equal count of values (INSERT without column list)
- Selectable procedure doesn't contain SUSPEND statement

Moving data

Check potential problems with:

- - data types
- - domains
- - character sets
- Use a data pump (like fbcopy) or write your own tool (Python and FDB may come handy)

Database verification

Verify correct function of:

- - VIEWs
- - CHECK constraints
- - Triggers
- - Stored procedures
- Check error handling, it needs to use SQLSTATE

Handling security issues

- New security database
- Authentication (native, trusted, 3.0)
- Priviledges (new options)
- Newly imposed constraint (especially for non-SYSDBA users)
- Security significantly changed in FB 3.0

Engine configuration

- New, changed or deprecated options like ExternalFileAccess, LegacyHash, FileSystemCacheTreshold, FileSystemCacheSize, MaxUnflushedWites, MaxUnflushedWiteTime, OldSetClauseSemantics, RelaxedAliasChecking, LockMemSize, LockHashSlots, GCPolicy.
- Run Environment Optimization Guidelines changed.

Checking your application(s)

- Connectivity package(s)
- Attachment and Transaction options
- Data Input and Output (data types, character sets, parameters, performance characteristics, new possibilities)
- Handling SQL statements
- Error handling
- Security
- Maintenance routines
- Deployment

Performance testing

If you're upgrading from 1.5, be prepared to face significant performance problems because version 2.5 and later ones are very different

- - New memory management
- - New INTL support

- Significant changes in cache management, threading, IPC, error handling etc.
- Significantly changed optimizer
- A lot of new features

Isolating problem areas and bottlenecks:

- Compare execution plans
- Check your use of BLOBs and long VARCHARs
- Focus on most used parts (loops etc.)
- Test execution of all your code paths
- Load testing. Monitoring tables and trace may help you find problem areas.

Revisiting (P)SQL

- You may have to change how you do things to solve performance problems. Take advantage of new features!
  - EXECUTE BLOCK
  - Derrived tables
  - New internal functions
  - RETURNING clause
  - Expression indices
  - Global Temporary Tables
  - Context variables
- More features you may use:
  - Database triggers
  - MERGE and UPDATE OR INSERT statements
  - Regular expressions
  - Autonomous transactions and other extensions to EXECUTE STATEMENT
  - Common Table Expressions (CTE)
- If your target is 3.0, you may also use:
  - BOOLEAN data type
  - Window (Analytical) Functions
  - Statistical Functions (Aggregate, Linear Regression)
  - PSQL Stored Functions, Sub-routines and Packages

Thanks for your attention. Questions?

contact: pcisar@ibphoenix.cz www.ibphoenix.com