How to migrate 75Gb database, with 564 tables, 5000+ stored procedures, 813 triggers, which is working 24x7, with ~400 users in less than 4 months
About IBSurgeon

• Tools and consulting
• Platinum Sponsor of Firebird
• Founded in 2002: 9 years of Firebird recoveries and consulting
• Based in Moscow, Russia

www.ib-aid.com  www.ibsurgeon.com
About Profitmed

- [http://www.profitmed.net/](http://www.profitmed.net/)
- Pharmaceutical distributor Russia
- 24x7 work mode for 2 warehouses (16000 sqm + 7000 sqm)
- 12x7 work mode for office
- **400** peak users
- **75 Gb** database size
- 64 Gb RAM, RAID 10 (Dell Storage), Xeon 4x6 cores
Why to migrate?

• From Firebird 1.5
  • Old (2004), no professional support
  • Bad memory management for huge queries
  • Slow garbage collection (1.5 hour+)
  • Slow backup (2 hours)
  • Firebird Classic 1.5 [actually] does not use more than 6 CPU cores

All together = SLOW.
Simple approach is impossible

The only available timeframe for switch to the new version – Christmas & New Year Eve
Special approach needed

- Fork database
- Prepare patch
- Apply patch

- Metadata changes
- Business logic changes
- Performance optimization
Infrastructure for migration

- 2 servers (similar)
  - Production is 64Gb RAM, Dell Storage
  - Test is 32Gb RAM, RAID1
- Tools to verify the migration
  - SQL queries compatibility
  - SQL queries plans - performance
Steps

1. Prepare metadata
2. Test convert data to 2.5
3. Application migration
   1. Check SQL queries in applications for compatibility,
   2. Change SQLs if necessary
   3. Check execution plans of SQL queries
   4. Change plans if necessary
4. Test run
5. Final run
Step 1: Prepare metadata

1. Extract metadata from Firebird 1.5 to script
   `Isql -x` (or using GUI tools)
2. Run metadata script in Firebird 2.5
3. Get output as “errors.txt” and analyze it
4. Patch 1.5 database (should be compatible with 1.5 and 2.5), and external script
Errors in pure metadata – part 1

1. **Ambiguous field name** between table X and table Y *(need to use aliases!)*
   ~40%

2. Data type unknown. **Blob sub_types bigger than 1** *(text)* are for internal use only. *(wrong sub_type in BLOB definitions)*
   ~10%

3. Attempt to update read-only column *(changes in AFTER UPDATE/AFTER DELETE triggers)*
   ~7%
Errors in pure metadata – part 2

4. In 1.5 – function TRIM was in UDF, now it’s embedded function with different parameters  ~5%
5. Expression evaluation not supported. **Strings cannot be multiplied** in dialect 3 (explicit casting needed).  ~1%
6. New Keywords (GLOBAL)
7. Other errors  ~1%
Summary of Step 1

• 800+ errors
• 2 weeks to fix
• Scripts:
  • Fixes in existing Firebird 1.5 database
  • Patch25.sql - External script to patch 2.5 database

Next: we are ready to convert data to Firebird 2.5
Step 2: Test convert data to 2.5

1. Backup patched 1.5 database
2. Restore with `--fix_fss_metadata` and `--fix_fss_data` options
3. Apply Patch25.sql script to restored database
4. Backup/restore under 2.5

Backup – 2 hours, restore - 6 hours.
Summary of step 2

• 2 days
• We have database in 100% Firebird 2.5 compatible format

• Next: need to check and change SQL queries in all our applications
Step 3: Application migration

1. Check SQL queries in applications for compatibility,
2. Change SQLs if necessary
3. Check execution plans of SQL queries
4. Change plans if necessary

This is the most complex step!
How to check SQL queries for compatibility

- We needed to log all SQL queries from **all applications**
  - Own-written applications
  - Closed-source application
Tip: logging is complicated

- **MON$** tables does not help – they are snapshots (and in 2.1+ only) and make the heavy load
- **FBTrace API** is 2.5+ only

- **FBScanner** is IBSurgeon’s commercial tool which analyze the network traffic and stores the full log, it’s the only true logging solution
Several workstations were run through FBScanner one by one to reduce performance impact and make log more “linear”.

**FBScanner logged SQL queries**

**Diagram:**
- **Computer1**:
  - FBScanner
  - 3050

- **Server**:
  - Firebird
  - 3050

- **Firebird clients (applications)**, connected through TCP
All SQL queries were stored to the log

Complete SQL texts with parameters and (optionally) plans!
<table>
<thead>
<tr>
<th>ID</th>
<th>IDATTACHMENT</th>
<th>IDTRANSID</th>
<th>NUM</th>
<th>START_TIME</th>
<th>PREPARE_TIME</th>
<th>EXECUTE_TIME</th>
<th>END_TIME</th>
<th>SQL_TEXT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1150545</td>
<td>1128961</td>
<td>1128964</td>
<td>13313</td>
<td>30.12.2010 15:56:22</td>
<td>0</td>
<td>0</td>
<td>30.12.2010 15:56:23</td>
<td>SELECT d.did, d.dcode, d.dname, d.dstate, d.ddate, d.dparent, d.dnum1, d.dnum2, d.dnum5, d.dnum9</td>
</tr>
<tr>
<td>1150546</td>
<td>1128961</td>
<td>1128964</td>
<td>13314</td>
<td>30.12.2010 15:56:23</td>
<td>0</td>
<td>0</td>
<td>30.12.2010 15:56:33</td>
<td>SELECT d.did, d.dcode, d.dname, d.dstate, d.ddate, d.dparent, d.dnum1, d.dnum2, d.dnum5, d.dnum9</td>
</tr>
<tr>
<td>1150547</td>
<td>1128961</td>
<td>1128964</td>
<td>13315</td>
<td>30.12.2010 15:56:46</td>
<td>0</td>
<td>0</td>
<td>30.12.2010 15:57:46</td>
<td>/* select iid, iid1 from link1 */</td>
</tr>
<tr>
<td>1150548</td>
<td>1128961</td>
<td>1128964</td>
<td>13316</td>
<td>30.12.2010 15:57:00</td>
<td>0</td>
<td>0</td>
<td>30.12.2010 15:58:09</td>
<td>select l.id, l.name, l.text1</td>
</tr>
<tr>
<td>1150549</td>
<td>1128961</td>
<td>1128964</td>
<td>13317</td>
<td>30.12.2010 15:58:09</td>
<td>0</td>
<td>0</td>
<td>30.12.2010 15:58:13</td>
<td>/* select iid, iid1 from link1 */</td>
</tr>
<tr>
<td>1150550</td>
<td>1128961</td>
<td>1128964</td>
<td>13319</td>
<td>30.12.2010 15:58:09</td>
<td>0</td>
<td>0</td>
<td>30.12.2010 15:58:13</td>
<td>select l.id, l.name, l.text1</td>
</tr>
<tr>
<td>1150551</td>
<td>1128961</td>
<td>1128964</td>
<td>13321</td>
<td>30.12.2010 15:58:13</td>
<td>0</td>
<td>0</td>
<td>30.12.2010 15:58:13</td>
<td>/* select iid, iid1 from link1 */</td>
</tr>
<tr>
<td>1150552</td>
<td>1128961</td>
<td>1128964</td>
<td>13322</td>
<td>30.12.2010 15:58:13</td>
<td>0</td>
<td>0</td>
<td>30.12.2010 15:58:13</td>
<td>select l.id, l.name, l.text1</td>
</tr>
<tr>
<td>1150553</td>
<td>1128961</td>
<td>1128964</td>
<td>13324</td>
<td>30.12.2010 15:58:13</td>
<td>0</td>
<td>0</td>
<td>30.12.2010 15:58:13</td>
<td>/* select iid, iid1 from link1 */</td>
</tr>
</tbody>
</table>

```sql
SELECT d.did, d.dcode, d.ddate1, d.ddate3, d.ddate4, d.did1, d.did6, d.dstate, d.ddate, d.dparent, d.dnum1, d.dnum2, d.dnum5, d.dnum9, l.name, l.text1,
  (SELECT FIRST 1 rf2_b_line(val,1) FROM xecblob WHERE id = l.id)
  trend,
  (SELECT kname FROM cfg WHERE kid = d.dstate) state_name,
  (SELECT lname FROM lib WHERE lid = d.did8) trend,
  (SELECT lname FROM lib WHERE lid = d.did9) manager_name,
  lib2800.lname as manager_name,
  lib5800$l1.name as dolqnost
FROM xecdata
```
To find incompatible SQLs

We need to “play” log to the copy of 2.5 database
• Make prepare
• Try to execute
• Catch exceptions/errors if any
• Log SQL execution plans
• Log SQL execution times

Find the differences in plans and delays
FBScanner played the log

FBScanner

Log database

Copy of database (in 2.5 format)
Differences in plans and execution times between 1.5 and 2.5 are highlighted.
Play log results

- We had ~55000 queries to analyze
- Only 280 has different plans
- ~400 has slower Execution time at 2.5 than 1.5
- ~50 queries raised exceptions

Only ~750 queries from 55000 required investigation.
Sorted log was exported to Excel
select agb.eid, agb.kollast, agb.eid5, agb.first_eid, agb.did3, agb.did4, agb.marker, agb.place, agb.mnp
from p101_ant_goods_balance agb
where agb.ddate <= :i$ddate and agb.l800 = :i$l800 and agb.did3 = rf2_abs(3409) and agb.marker = :i$marker and agb.did <> :i$did

• 2.5 PLAN (AGB INDEX (IDX$AGB_L800_DDDATE_ECON_EMP))
• 1.5 PLAN (AGB INDEX (IDX$AGB_L800_DDDATE_ECON_EMP, IX$AGB_DID3))
Another problems

- New Firebird keywords in closed-sourced software
  - Was hacked 😊
- UDF rFunc for 64-bit Windows and Linux
  - AUDFL [https://www.assembla.com/wiki/show/audfl_rfunc](https://www.assembla.com/wiki/show/audfl_rfunc)
Results of migration

• At Firebird 2.5
  • Better performance in queries
  • Fast garbage collection (20 minutes instead of 1.5 hours)
  • Backup time reduced (1.5 hour instead of 2 hours)
  • Better use of multi-CPU hardware
  • 64 bit version of Firebird available
  • EXECUTION STATEMENT and other SQL features
Optimization Pack “Rodizio” offer

• To optimize databases we offer 1 year subscription to install as many as you need FBScanner+IBAnalyst+IBTM instances
• This instances will run indefinitely (no restrictions)
• Should be activated
  • Silent bundle is also possible, requires signing VAR agreement

Today – EUR 1200 (reg EUR 1500)
Next presentation – tomorrow 14-00

- Nov 26, 14-00
- Supercharging Firebird production systems: transactions, garbage, maintenance
Thanks and Contacts

Blog
http://FirebirdSurgeon.blogspot.com

Twitter
http://Twitter.com/ibsurgeon

Web
http://www.ib-aid.com