

**Binary Chief**

**Softwarový projekt**

**(Vedúci projektu: RNDr. David Bednárek)**

Vladimír Bedecs

15. 9. 2008

# Obsah

|  |          |
|--|----------|
| <b>MOTIVÁCIA</b> .....                                 | <b>3</b> |
| • Problém A: .....                                     | 3        |
| • Problém B: .....                                     | 3        |
| • Problém C: .....                                     | 3        |
| <b>CIEĽ PROJEKTU</b> .....                             | <b>3</b> |
| 2. Požiadavky na riešenie problému A.....              | 3        |
| 3. Požiadavky na riešenie problému B.....              | 4        |
| 4. Požiadavky na riešenie problému C.....              | 4        |
| <b>APLIKÁCIA PROJEKTU</b> .....                        | <b>5</b> |
| <b>VŠEOBECNÉ POŽIADAVKY NA RIEŠENIE PROJEKTU</b> ..... | <b>5</b> |

## Motivácia

Binárne súbory sú výhodným úložiskom dát. Na rozdiel od iných známych spôsobov uchovávania dát, nemajú takmer žiaden priestorový, či výpočtový overhead. To znamená, že dátové jednotky binárneho súboru nie je nutné nijako tagovať, ako v XML, či v iných textových formátoch a prístup k nim nevyžaduje väčšinou žiadnu softwarovú medzivrstvu ako v prípade dát uložených v relačných, či v iných databázach. Inou výhodou binárneho súboru je jeho kompaktnosť. Pre tieto výhody a mnohé ďalšie je binárny súbor bežnou formou úložiska dát napríklad na embedded systémoch, kde sú priestorová aj výpočtová kapacita do značnej miery obmedzené. Táto technológia má však aj niekoľko problémov:

- **Problém A:** Na rozdiel od XML, je však skoro nemožné rozumne zobrazovať obsah binárneho súboru, či dáta „editovať ručne“ bez pomoci nejakej softwarovej medzivrstvy, pretože samotné dáta v binárnom súbore nevypovedajú nič o jeho štruktúre a väčšinou ani nie sú znakového charakteru.
- **Problém B:** Ak je súbor rozsiahly (niekoľko gigabajtový), môže byť prístup k dátam nielen výpočtovo, ale aj pamäťovo veľmi náročný.
- **Problém C:** Štruktúra binárneho súboru môže byť závislá na samotných dátach. To znamená, že o skutočnej štruktúre binárneho súboru sa niekedy musí rozhodnúť až po prečítaní jeho relevantných častí.

## Cieľ projektu

Cieľom projektu je vytvoriť software riešiaci hore uvedené problémy binárnych súborov na dodanom binárnom súbore popisujúcom mapu Európy. Riešenie zahŕňa návrh jazyka pre popis štruktúry binárneho súboru. Vytvorený software potom umožní vďaka vopred definovanej štruktúre binárneho súboru, zobraziť v ňom obsiahnuté dátové jednotky na vyššej úrovni abstrakcie ako obyčajný hex editor. Pre rýchly a pamäťovo nenáročný prístup k dátam vygeneruje kód read-only rozhrania(ní) pre aplikácie pracujúce s nimi. Tiež umožní, aby sa presná definícia binárneho súboru zostavila až v čase parsovania binárneho súboru, v závislosti na dátach.

### 1. Požiadavky na riešenie problému A

Po návrhu vhodného jazyka vo formáte XML na popis štruktúry binárneho súboru, bude cieľom projektu implementovať schopnosť „pozerať sa na binárny súbor cez jeho definíciu“. Definícia teda priradí elementárnym dátovým jednotkám (byte, word, dword,...) názvy a umožní ich združovať so komplexnejších pomenovaných jednotiek (štruktúra, pole). Interface tohto modulu využije jednoduché a intuitívne GUI, ktoré definíciu zobrazí a umožní aj jej editáciu. Takisto umožní

zobrazovať a editovať aj dáta v binárnom súbore vzhľadom na vytvorenú definíciu.

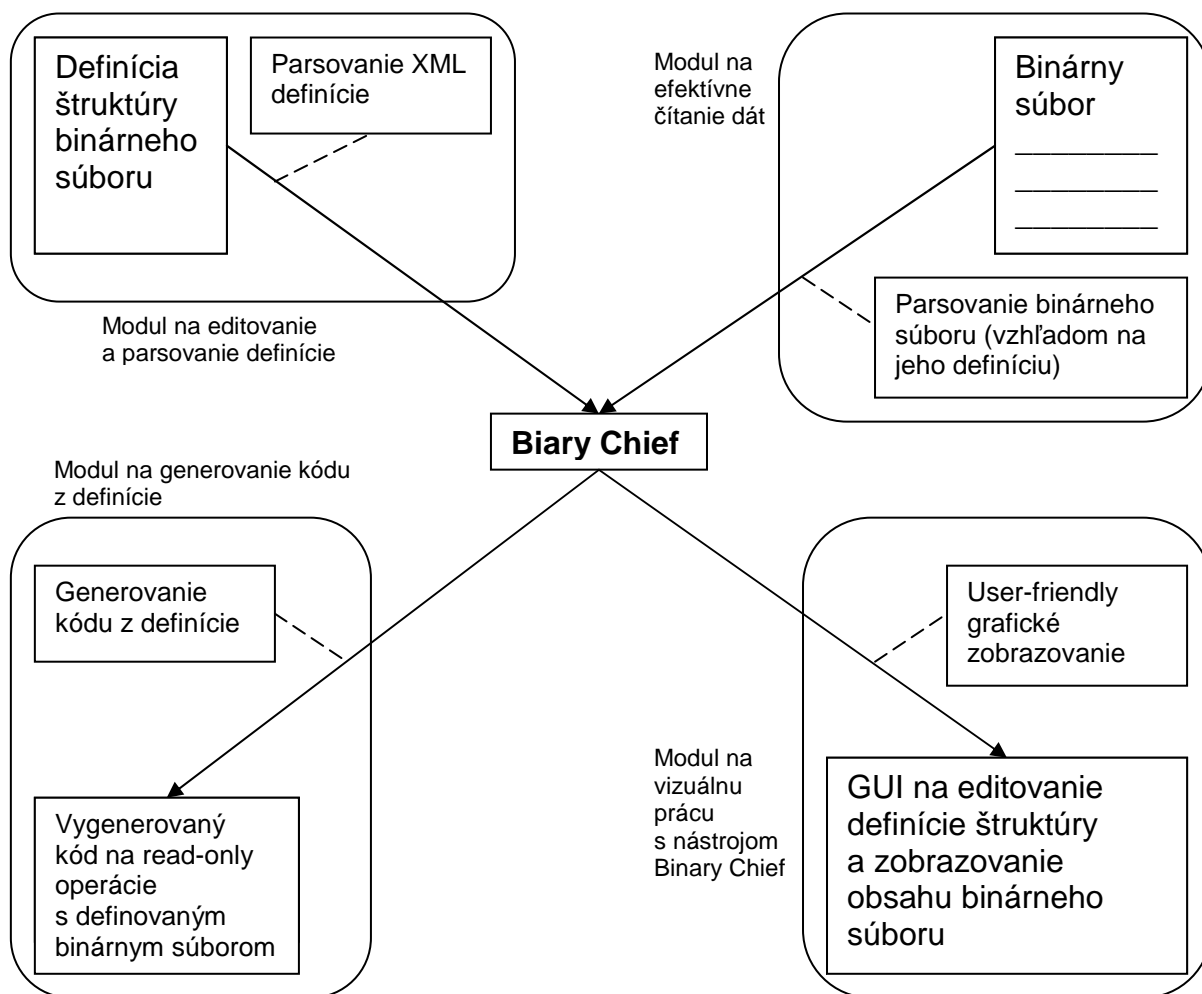
## **2. Požiadavky na riešenie problému B**

Z definície štruktúry binárneho súboru musí byť projekt schopný vygenerovať kód pre rýchly a pamäťovo nenáročný prístup k dátam obsiahnuté v danom binárnom súbore. To znamená, že sa súbor nebude načítavať celý do pamäte, ale iba nevyhnutná podmnožina dát potrebná pre prečítanie požadovaných dát. Vytvorí tak štandardnú softwarovú medzivrstvu pre read-only operácie, ktorú môžu aplikácie využívať na prístup k súboru. Podporované budú známe prístupy SAX a DOM z technológie XML, avšak implementované nad binárnym súborom. GUI opäť poskytne možnosť špecifikácie parametrov generovania a jeho samotné spúšťanie.

## **3. Požiadavky na riešenie problému C**

Štruktúra binárneho súboru môže byť závislá na samotných dátach. Do jazyka definície je teda nutné zaviesť systém, ktorým bude možné vyjadriť všetky alternatívne štruktúry binárneho súboru, ktoré môžu počas parsovania vzniknúť. Projekt pomocou tohto systému v definícii rozhodne na základe relevantných hodnôt v binárnom súbore, ktorú z alternatívnych definícií uplatní.

Obrázok 1: Rozdelenie projektu na moduly



## Aplikácia projektu

Projekt bude aplikovaný na 4GB binárny súbor, s existujúcou špecifikáciou štruktúry, popisujúci mapu Európy pre embedded navigačné systémy. Schopnosti a funkčnosť implementácie budú teda jednoducho demonštrovateľné napríklad kreslením samotnej mapy pomocou vygenerovanej prístupovej read-only vrstvy.

## Všeobecné požiadavky na riešenie projektu

Projekt bude napísaný v jazyku C++. Musí byť do čo najväčšej miery platformne nezávislý. Bude sa snažiť využívať dostupné existujúce knižnice (napr. Boost, TinyXml, ...) aby sa vývoj sústredil najmä na podstatu projektu. Kód musí byť na vysokej úrovni týkajúcej sa nielen prehľadnosti, ale aj návrhu. Preto by v ňom mali byť rozsiahlo využívané návrhové vzory či iné „dobré praktiky“.

Na projekte by malo pracovať 4 – 5 študentov s výbornou znalosťou C++ a objektového návrhu. Predpokladaná dĺžka trvania vypracovania projektu je 6 - 7 mesiacov.