# P2P NETS - SEARCHING FOR NEW VIRUSES

Jiří SCHÄFER, Master Degree Programme (5)
Dept. of Intelligent Systems, FIT, BUT
E-mail: xschaf03@stud.fit.vutbr.cz

Supervised by: Dr. Daniel Cvrček

**ABSTRACT**

The application developed as a part of the project is able to connect to a network as a users interested in particular files. These files are downloaded, analysed, and information is stored in a database or resend to the analyzation software.

## 1 ÚVOD

Cílem této práce je navrhnout modulární systém, který by vyhledával v peer-to-peer sítích nové viry a analyzoval je. Tento projekt je vyvíjen ve spolupráci s firmou Grisoft, přičemž firma Grisoft dodá blackbox pro samotnou analýzu souborů. Tato část projektu je zaměřená na vyhledávání souborů a jejich následné stahování.

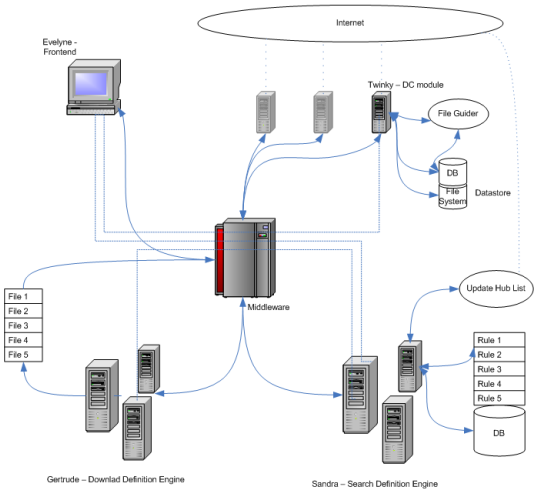
## 2 NÁVRH PROJEKTU

Systém je rozdělen do několika na sobě nezávislých modulů, které spolu komunikují přes střední vrstvu - middleware, ten zároveň funguje jako směrovač zpráv a buffer pro jednotlivé moduly. Moduly jsou rozdělené do skupin podle funkčních vlastností. Daly by se rozdělit do několika skupin: moduly definiční, moduly pro stahování, frontend a backend.

Tato architektura dovoluje značnou autonomnost modulů, systém je navržen jako distribuovaný, to znamená že jednotlivé moduly mohou být spuštěny na různých počítačích, v různých lokalitách a komunikují spolu přes střední vrstvu - middleware, který musí naopak běžet na pevně daném počítači, aby se k němu ostatní moduly mohly připojovat.

### 2.1 CELKOVÉ SCHÉMA ZAPOJENÍ

Obr. 1 Fyzické propojení mezi moduly je realizováno pomocí BSD socketů. Middleware naslouchá na portu a při připojení bloku se ustanoví spojení, které je udržováno až do zániku bloku.




Obrázek 1: Celkové schéma zapojení

## 2.2 SANDRA - SEARCH DEFINITION ENGINE

Tento modul patří do skupiny definičních modulů. V první řadě určuje kam se má modul pro stahování připojit. Po připojení dodává výrazy pro hledání v dané síti, tyto výrazy se mění s každým místem pro připojení. Pravidla pro hledání v sítích se také mění postupně při analýze stažených souborů. Pokud je při analýze podezření na infikaci souboru neznámým virem, změní se priorita vyhledávání tak abychom nalezli další souboru se stejnými(podobnými) parametry infikaci a potvrdili, nebo vyvrátili infikaci novým druhem viru.

## 2.3 GERTRUDE - DOWNLOAD DEFINITION ENGINE

V tomto modulu se jednak provádí analýza stažených souborů a určuje se které soubory se budou stahovat.

**Analýza výsledků hledání** Z modulu pro stahování se zašlou výsledky hledání přímo do tohoto modulu a zde se určí zdali se soubor bude stahovat, nebo ne. Toto rozhodování se provádí na základě výsledků předešlé analýzy souborů a na implicitních pravidlech, která jsou uložena v databázi. Vybrané soubory se poté předají zpět modulu pro stahování ke stažení.

**Analýza souborů** Analýzu souborů provádí Blackbox dodaný firmou Grisoft. Jedná se o DLL knihovnu které na vstup dáme soubor a výstupem je informace zda je daný soubor infikován či nikoliv a jestli je třeba dostahovat další část souboru pro další analýzu.

## 2.4 EVELYNE - FRONTEND

Frontend má dvě hlavní úlohy, slouží jednak jako GUI pro celý systém a jednak pro vnější řízení systému.

**Zjišťování stavových informací systému** Frontend může běžet jako klasické WWW rozhraní ve kterém se zobrazují jednotlivé moduly, jejich stavové informace, právě prováděné akce jako stahování souborů (včetně informací o tom jaký soubor se stahuje, odkud se stahuje ...) atp.

**Modifikace stavu systému** Vzhledem k tomu že se jedná o distribuovaný systém je třeba mít nějaké zařízení pomocí kterého bude možné měnit nastavení každého připojeného modulu. Vezmeme-li si například definiční moduly, vidíme že vzorce pro vyhledávání se mění podle implicitních pravidel a pomocí výsledků analýzy souborů. Pomocí Frontendu se dají měnit tyto implicitní pravidla. Dají se dají měnit priority pro stahování, místa por připojování atp.

### 2.5 TWINKY - MODUL PRO PŘIPOJENÍ DO SÍTĚ DIRECTCONNECT

Tento modul zajišťuje připojení do sítě DirectConnect, komunikaci v této síti a stahování souborů z této sítě. Na základě informací z definičních modulů provede připojení na hub (jeden ze základních uzlů sítě DirectConnect) a zasílá požadavky na hledání souborů. Výsledky hledání přeposílá přes middleware modulům pro stahování, které určí jestli se bude soubor stahovat, které jeho části a s jakou prioritou. V modulu se takovýto požadavek na soubor vloží do fronty a začne se stahovat hned jak na něj dojde řada.

## 3 ZÁVĚR

Tento systém je schopen vyhledávat a analyzovat nové viry, získávat poznatky o nových typech virů a pomáhá vyvíjet obrany proti nim.

Vzhledem k tomu že je použita architektura middleware/publisher-subscriber je možné místo jednoho velkého modulu pro připojení do mnoha peer-to-peer sítí vytvořit několik modulů menších a specializovanějších. V případ odstavení dané p2p sítě (nebo modulu určeného pro komunikaci s touto sítí) se nemusí celý projekt zastavovat, jednoduše se odpojí daný modul a systém poběží dále.

## REFERENCE

[1] http://www.dcpp.net/